International Journal of Modern Physics B

Causal potency of consciousness in the physical world

Danko D. Georgiev
Institute for Advanced Study, 30 Vasilaki Papadopulu Str.
Varna, 9010, Bulgaria
danko@q-bits.org

Received (11 March 2023)
Revised (24 April 2023)
Accepted (5 May 2023)

The evolution of the human mind through natural selection mandates that our conscious experiences are causally potent in order to leave a tangible impact upon the surrounding physical world. Any attempt to construct a functional theory of the conscious mind within the framework of classical physics, however, inevitably leads to causally impotent conscious experiences in direct contradiction to evolution theory. Here, we derive several rigorous theorems that identify the origin of the latter impasse in the mathematical properties of ordinary differential equations employed in combination with the alleged functional production of the mind by the brain. Then, we demonstrate that a mind–brain theory consistent with causally potent conscious experiences is provided by modern quantum physics, in which the unobservable conscious mind is reductively identified with the quantum state of the brain and the observable brain is constructed by the physical measurement of quantum brain observables. The resulting quantum stochastic dynamics obtained from sequential quantum measurements of the brain is governed by stochastic differential equations, which permit genuine free will exercised through sequential conscious choices of future courses of action. Thus, quantum reductionism provides a solid theoretical foundation for the causal potency of consciousness, free will and cultural transmission.

Keywords: ordinary differential equations; quantum indeterminism; stochastic differential equations.

PACS numbers: 03.67.-a, 03.65.Ud, 03.65.Ta, 87.19.La

Contents

1. Introduction

1.1. *The sense of agency*

We are *sentient beings* that possess an *inner psychological world* or a *stream of conscious experiences*, which we simply refer to as a *mind*.[1–3] In fact, we are what our conscious minds are. It is only through our conscious experiences that we are able to access the surrounding world, comprehend it and act upon it.[4] Our personal thoughts, aims, goals and desires motivate us to strive towards achieving a healthy, prosperous and happy life.[5] The subjective awareness of initiating, executing, and controlling our volitional actions in the physical world corroborates daily our *sense of agency*. In the absence of conscious experiences, however, such as during general anesthesia,[6–8] syncope[9] or coma,[10] we lose our *feeling of agency* within the world. Furthermore, exactly because we are conscious agents with causative potency, it is possible for civil law to establish blameworthiness for actions that are considered wrongdoings[11–13] and ethics can hold us morally responsible for what ensues from our behavior.[14,15]

1.2. *The problem of mental causation*

The necessity to describe our conscious minds within the framework of available physical theories, however, inevitably confronts us with the problem of *mental causation*, namely, how is it possible and what is the physical mechanism through which our consciousness is able to affect the physical world. Apparently, branding the conscious mind as "non-physical" is not going to be helpful because it will immediately put the mind outside of physics, and consequently if the mind is not subject to any physical laws it would be impossible to derive any mental causation upon the physical world.

Because physics is the most fundamental scientific discipline, it is expected to encompass *everything in existence* and study the *entirety of reality* using mathematical principles.[16] Our conscious minds do exist, therefore they are real and have to be defined as "physical"[17,18] thereby enjoying the privilege to be considered a valid subject for discussion by physical theories. Moreover, if a physical theory does not include consciousness or makes incorrect predictions with regard to consciousness, then such a physical theory should be deemed either *falsified* or *incomplete*. Therefore, conscious experiences should be considered *physical* and their causal effects upon the physical world should be described by a *physical theory*—either a physical theory that we already have or a physical theory that will be constructed in the future. The focus of the present work will be on comparison of available *classical* or *quantum* physical theories in view of demonstrating that quantum physics already has all the necessary mathematical ingredients for accommodating a causally potent consciousness.

1.3. *Evolutionary theory mandates causally potent consciousness*

The evolution of consciousness and development of culture in primates,[19–24] including chimpanzees,[25–29] bonobos,[30,31] gorillas,[32] orangutans,[33–36] snow monkeys[37–39] and early humans,[40–48] is an established empirical fact that entails the causal potency of consciousness in the physical world.[49–52] From this stronghold, we will investigate the implications of different physical approaches to modeling consciousness, including *functionalism* or *reductionism*, within the context of *classical* or *quantum* physics. Utilizing the mathematical properties of ordinary differential equations or stochastic differential equations, we will derive rigorous theorems, which constrain the available solutions to the problem of mental causation. In particular, we will show that within available physical theories, it is quantum physics that provides the only plausible framework for a physical theory of causally potent consciousness and free will.

1.4. *A synopsis of the presentation on the causal potency of consciousness*

The subsequent presentation is structured as follows: In Section 2, we formulate the problem of mental causation and briefly review the neuroanatomy and neurophysiology of the brain cortex, peripheral neural system, sensory organs and effector organs. Next, in Section 3 we explain why the problem of mental causation is unsolvable in classical physics. We also illustrate the mathematical theory of ordinary differential equations and prove theorems that eliminate classical functionalism and classical reductionism as plausible theories of consciousness. Then, in Section 4 we illustrate the mathematical theory of stochastic differential equations and prove theorems that recognize quantum reductionism as the most plausible theory of consciousness. In order to make the exposition self-contained, in Section 5 we elaborate on free will and its representation using stochastic processes, in Section 6 we illustrate how knowledge acquisition or learning generates dynamic biases and varying amounts of free will, in Section 7 we explain how quantum entanglement leads to mind binding and constrains free will, and in Section 8 we describe the physical mechanism underpinning the wave function collapse and disentanglement. Lastly, in Section 9 we discuss the significance of the presented results and how they provide a consistent account of the natural evolution of the human mind.

2. The causal potency problem

We react to sensory stimuli that are present in the surrounding world. The neurophysiological account of our reaction starts with transduction of the sensory stimulus into an electric signal by the sensory organs. For example, the eyes convert visible light into electric currents in the retina, the ears convert audible sound into electric currents in the cochlea, and the skin converts mechanical pressure from touch into electric currents in encapsulated nerve endings of Meissner corpuscles.[53,54] The sensory information carried by these electric signals then propagates

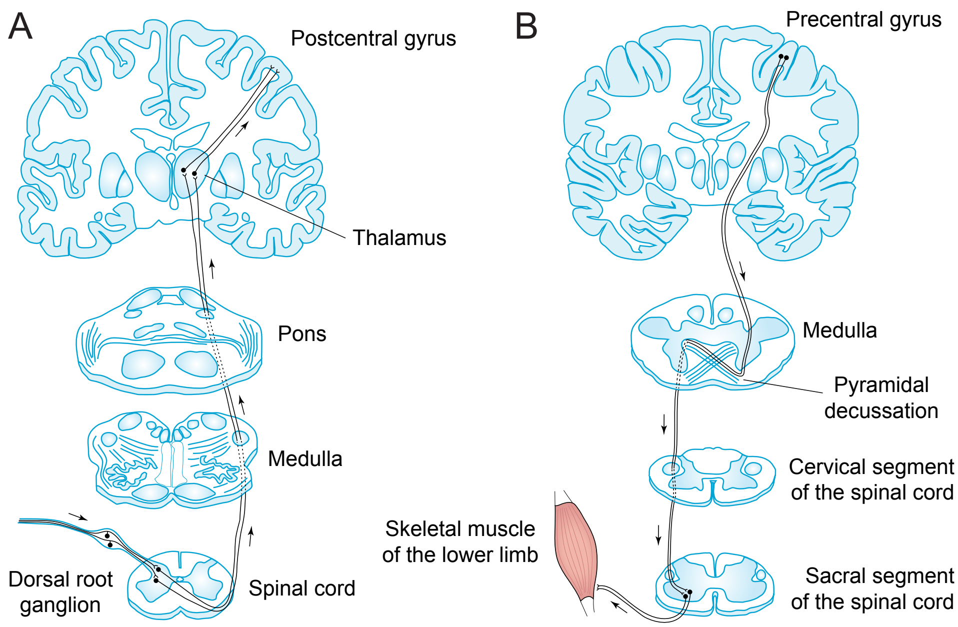

Fig. 1. The human brain interacts with its physical environment. Electric signals in the form of action potentials mediate the physiological communication between the human brain cortex and the body: the somatosensory pathway (A) delivers sensory information from the body to the somatosensory cortex in the postcentral gyrus, whereas the somatomotor pathway (B) delivers motor information from the motor cortex in the precentral gyrus to the body muscles. The spinal cord segments, medulla and pons are represented with their transversal sections, whereas thalamus and cortex are shown in frontal slice. Modified from Ref. 4.

along sensory pathways that reach corresponding sensory areas in the brain cortex (Fig. 1A). The voluntary decision for action is executed by the motor area of the brain cortex, which sends motor electric signals toward the effector organs such as the skeletal muscles that move the body (Fig. 1B). Therefore, the overall process of reacting to sensory stimuli is a sequential composition of inputting the sensory information from the body to the brain cortex and outputting the voluntary choice of action from the brain cortex to the body.[4]

Clinical observations from injuries of the human nervous system have accumulated an overwhelming amount of evidence that the brain cortex is the *seat of consciousness*.[55] This is consistent with physiological delays of at least 50 ms temporal duration between the application of the sensory stimulus at the sensory organ and the conscious perception of the sensation in the brain cortex or between the conscious decision to elicit a voluntary movement in the brain cortex and the actual contraction of the skeletal muscle.[56–58]

The somatosensory pathways from the body towards the brain cortex and the somatomotor pathways from the brain cortex towards the body could be either bypassed or replaced with brain–machine interfaces, which deliver sensory infor-

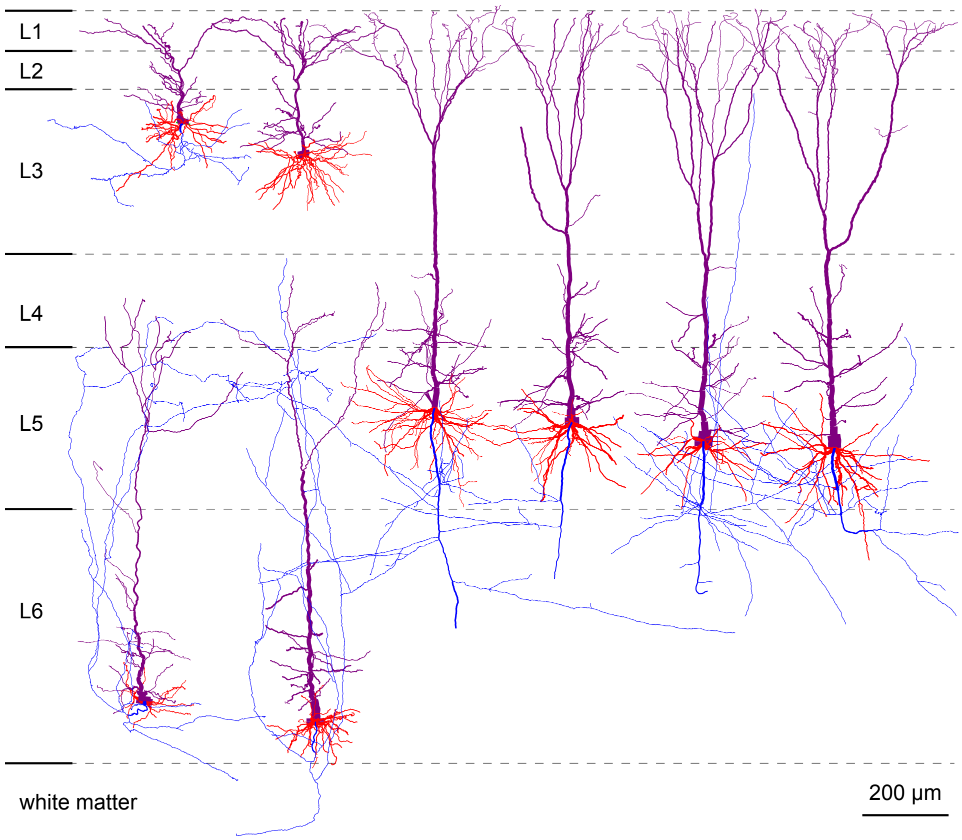

Fig. 2. Layered structure (L1–L6) of the gray matter of rat neocortex based on digital reconstructions from NeuroMorpho.org of Layer 2-3 pyramidal neurons (NMO_49059, NMO_49054), Layer 5 pyramidal neurons (NMO_77908, NMO_77904, NMO_77905, NMO_77920) and Layer 6 pyramidal neurons (NMO_09382, NMO_64646). Basal dendrites are rendered in red, apical dendrites in purple, and axons in blue. Neuron identification numbers are listed from left to right in the order the rendered reconstructions are assembled. Neuron reconstructions can be retrieved by their identification numbers at `https://neuromorpho.org/KeywordSearch.jsp`.

mation directly to the brain cortex where they are consciously experienced[59] or receive motor information directly from the brain cortex for the control of robotic devices.[60] The engineering of brain–machine interfaces and their practical application in neurorehabilitation is assisted by the layered organization of pyramidal neurons inside the gray matter of the brain cortex.[61]

The microscopic neuroanatomy of a cross section of the brain cortex reveals columns of vertically stacked pyramidal neurons[62–68] organized into 6 layers (Fig. 2). The first layer L1 is closest to the cortical surface, whereas the sixth layer L6 lays deepest. The dendrites of pyramidal neurons, which receive electric information, extend towards the superficial cortical layers and form dense arborizations.[69] The axons of pyramidal neurons, which output electric information, extend towards the underlying white matter whose characteristic color is due to the myelin sheets that insulate the axons from leaking their ionic currents into nearby inactive axons.[70]

The collective excitation of columns of pyramidal neurons creates multiple individual electric potentials that summate spatiotemporally into a larger *local field potential* that can be recorded with micro-electrodes[71] implanted in the extracellular space of cortical tissue.[60,72] If the recorded electric activity from the brain cortex is forwarded to computer program that controls a robotic arm, the conscious mind is able to train itself after several months of practice to control the robotic arm without actually moving any of the body muscles. Thus, our conscious minds appear to be causally potent agents within the physical world, because if they were not, conscious control of brain–machine interfaces would not have been possible.[4,73] The problem of mental causation is to explain how the conscious mind is able to physically affect the electric activity of cortical pyramidal neurons. Whether such an explanation is possible depends critically on the physical approach chosen for addressing the mind–brain problem. We will elaborate on this in great detail next.

3. Causal potency of consciousness in classical physics

The world of classical physics is based on two fundamental postulates. First, it is assumed that all physical quantities are *observable*. This means that the physical states of classical systems can be measured with physical instruments. Second, it is assumed that the *time dynamics* of physical states is governed by a *system of ordinary differential equations* (ODEs). This means that given an initial state $S(0)$ of a classical system, its future time evolution $S(t)$ is deterministic and can be calculated with absolute certainty and arbitrarily high precision in principle.

Example 1. (Deterministic dynamics) The mathematical properties of classical deterministic dynamics can be illustrated with the following 3-dimensional chaotic jerky Lorenz-like system whose time evolution is governed by a single third-order ordinary differential equation (ODE) of a time-dependent variable $x(t)$[74]

$$\left(\frac{d}{dt}\right)^3 x = -\left(\frac{d}{dt}\right)^2 x - 4\frac{d}{dt}x + 5x - x^3 \tag{1}$$

With the use of the following substitutions

$$x_1 = x \tag{2}$$

$$x_2 = \frac{d}{dt}x = \frac{d}{dt}x_1 \tag{3}$$

$$x_3 = \left(\frac{d}{dt}\right)^2 x = \frac{d}{dt}x_2 \tag{4}$$

$$\frac{d}{dt}x_3 = \left(\frac{d}{dt}\right)^3 x \tag{5}$$

the single third-order ODE (1) can be re-written as a system of three first-order ODEs

$$\begin{cases} \frac{d}{dt}x_1 &= x_2 \\ \frac{d}{dt}x_2 &= x_3 \\ \frac{d}{dt}x_3 &= -x_3 - 4x_2 + 5x_1 - x_1^3 \end{cases} \tag{6}$$

The main point emphasized here is that any single higher-order ODE can always be re-written in mathematically equivalent form as a system of several first-order ODEs. Once we have the system of first-order ODEs, we can simulate the trajectory in time of the physical system for any initial state $S(0)$ given as a set of observables $\{x_1(0), x_2(0), x_3(0)\}$ at the initial time $t = 0$. The determinism of the computed dynamics is manifested in the fact that no matter how many simulation runs we perform (with the same initial state and a fixed level of precision), the resulting trajectory always remains the same (Fig. 3).

The mathematical properties of ordinary differential equations may look innocent, but create serious difficulties once we start bringing forward putative mind–brain models.

3.1. *Classical functionalism implies causally impotent consciousness*

Definition 1. (Functionalism) *Functionalism* is the philosophical stance that the observable brain *produces* the unobservable conscious experiences. The main feature is that the brain Φ produces the mind Ψ, namely $\Phi \to \Psi$, but the brain Φ and the mind Ψ are mathematically distinct entities, $\Phi \neq \Psi$. The symbol $\to$ is used to indicate *functional production*. Here, the word *produces* can be replaced with other synonyms without changing the definition of functionalism, e.g., the brain *gives rise* to the mind, the brain *generates* the mind or the brain *creates* the mind. The same meaning can also be expressed as: the mind *emerges from* the brain, the mind *originates from* the brain or the mind *is constructed by* the brain.

The classical aspect of brain modeling comes from the requirement that all physical observables of the brain are governed by ordinary differential equations (ODEs). This means that given the initial state of the brain $\Phi(0)$ we can compute deterministically (i.e., solve numerically using a computer program) the future state of the brain $\Phi(t)$ for any time $t > 0$. Furthermore, because the mind Ψ is unobservable, it is not present as a variable in the set of ODEs that govern brain dynamics. The combination of classical physics and functionalism inevitably implies that the mind is causally impotent and cannot affect anything inside the physical world.[4]

Theorem 1. *Classical functionalism implies that the conscious mind lacks causal potency and is unable to affect the physical world.*

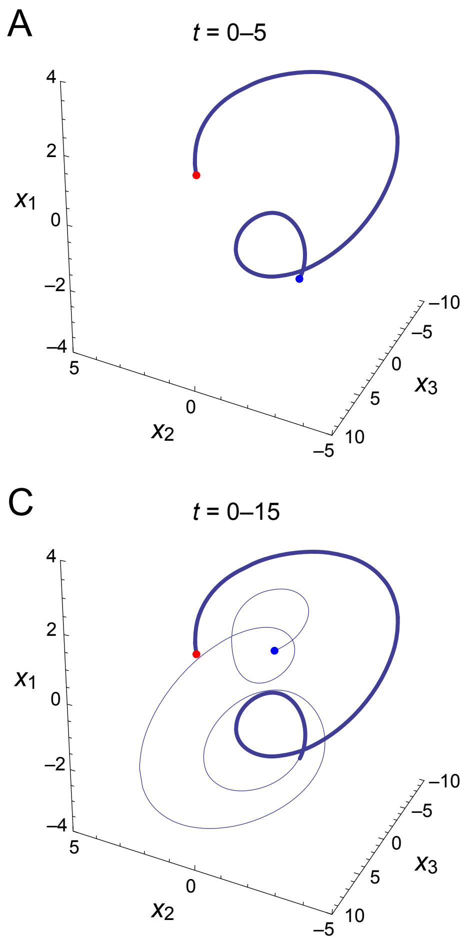

Fig. 3. Deterministic dynamics of a physical system whose time evolution is governed by the ordinary differential equation (1). Each of the four simulation runs starts from the same initial state $S(0)$ with $x_1(0) = 1$, $x_2(0) = 2$ and $x_3(0) = 3$, but lasts for different period of time t =0–5 in (A), t =0–10 in (B), t =0–15 in (C) or t =0–20 in (D). The characteristic feature of deterministic dynamics is that no matter how many runs are performed with the same initial state, the initial segment of the trajectory for the period t =0–5 (thick purple line) will always remain the same. We have performed four simulations with different durations because if we had simulated the same period of time t =0–5 consecutively four times, we would have ended with four identical panels in the composite figure without any evidence that these were four different simulations. The initial state is shown with a red point, whereas the final state is shown with a blue point. The units of t and x_1 are arbitrary, whereas the units of x_2 and x_3 are fixed as corresponding rates of change by (3) and (4).

Proof. The main premises of classical functionalism can be summarized as follows:

Premise 1. The act of functional production of the mind by the brain, $\Phi(t) \to \Psi(t)$, entails an infinite list of productions at each time point t: $\Phi(0) \to \Psi(0)$,

$\Phi(t_1) \to \Psi(t_1)$, $\Phi(t_2) \to \Psi(t_2)$, ..., $\Phi(t_n) \to \Psi(t_n)$.

Premise 2. The physical states of the brain $\Phi(t)$ and the surrounding world $W(t)$ are governed by an explicitly given system of ordinary differential equations (ODEs) in which only physical observables of the brain and the world are present. The interaction between the brain and the surrounding world is reflected in the fact that the system of ODEs may not be separable in terms of brain and world variables.

From the second premise, which gives explicitly the system of physical equations, we can use the initial states of the brain $\Phi(0)$ and the world $W(0)$ to deterministically compute the future states of the brain $\Phi(t)$ and the world $W(t)$ for any future time $t > 0$. These future states $\Phi(t)$ and $W(t)$ will remain exactly the same regardless of what conscious experiences $\Psi(0), \Psi(t_1), \Psi(t_2), \ldots$ are produced by the brain, due to the fact that the conscious experiences do not enter in the physical equations. Moreover, there will be no difference whatsoever between the two cases in which any conscious experiences $\Psi(0), \Psi(t_1), \Psi(t_2), \ldots$ are produced (case 1) or not produced (case 2), since the computation of future states $\Phi(t)$ and $W(t)$ does not require knowledge of $\Psi(0), \Psi(t_1), \Psi(t_2), \ldots$ The causal effect of the conscious experiences is determined by subtraction of the brain and world dynamics obtained in the absence of conscious experiences (case 2) from the dynamics obtained in the presence of conscious experiences (case 1). Because both dynamics are the same, after the subtraction we are left with zero causal effect of conscious experiences

$$\Phi_{\text{case 1}}(t) - \Phi_{\text{case 2}}(t) = 0 \tag{7}$$

$$W_{\text{case 1}}(t) - W_{\text{case 2}}(t) = 0 \tag{8}$$

Therefore, the conscious experiences $\Psi(0), \Psi(t_1), \Psi(t_2), \ldots$ that are produced by the classical brain cannot affect anything in the brain or the surrounding world.

The proof of the theorem remains unaffected even if we assume that the brain has to perform some "function" in order to generate conscious experiences. Because the brain "function" does not correspond to a single brain state at a single time point, but corresponds to a sequence of brain states that form a trajectory from $\Phi(t_0)$ to $\Phi(t_1)$ to generate the conscious experience $\Psi(0)$, we only need to discretize the first premise as follows: $[\Phi(t_0) \sim \Phi(t_1)] \to \Psi(0)$, $[\Phi(t_1) \sim \Phi(t_2)] \to \Psi(t_1)$, $[\Phi(t_2) \sim \Phi(t_3)] \to \Psi(t_2)$, ..., etc., where the symbol $\sim$ indicates the unique trajectory with the given initial and final state. Solving the system of ODEs uses only the second premise, which again reproduces the lack of causal potency of the functionally generated mind upon the brain or the surrounding world. $\square$

The implications of Theorem 1 are that neither human consciousness nor animal consciousness could have evolved through natural selection in classical functionalism.[4] Furthermore, if conscious experiences were unable to affect anything in the physical world, then their presence would have been utterly meaningless.[75] Since the human mind has evolved naturally and has left cultural artifacts such as tools,[76] musical instruments,[47] hand-carved statuettes[46] and wall paintings in prehistoric caves,[40,77–84] it follows that classical functionalism has to be false.

Some philosophers have attempted to express the mathematical theory behind ordinary differential equations (ODEs) as a list of independent verbal postulates and have misleadingly identified the contents of Theorem 1 as the "problem of mental causation".[85,86] The main philosophical goal was to isolate the verbal postulate that is "wrong" and fix it. Ultimately, it was concluded that the "causal closure" of the physical world is the main culprit leading to causally impotent consciousness, hence it is mandatory to resort to "reductionism" if we were to have a theory of causally potent consciousness.[87] Although we consider that the overall move towards reductionism is on the right track, there are several important inaccuracies that make this previous philosophical work wanting:

First, the mathematical theory of ordinary differential equations (ODEs) cannot be split into independent verbal postulates from which one is allowed to chose from. Instead, the theory of ODEs is build upon prerequisite mathematical concepts introduced in ordinary calculus such as mathematical *functions*, *series*, *limits*, *derivatives* and *integrals*.[88] The resulting mathematical properties of the solutions of ODEs, including *existence*, *determinism* and *uniqueness* given exact initial condition, are inherited all together and cannot be dropped selectively as one pleases.

Second, a major culprit leading to causally impotent consciousness is the fact that the conscious states *are not present inside* the system of ODEs. In fact, whether the mind is labeled as "physical" or "non-physical" is irrelevant for its "closure" from the physical world. Even if the mind is defined to be "physical", it will not be able to affect the brain or the surrounding world provided that the mind is absent from the ODEs that govern the time dynamics of the brain and the surrounding world.

Third, the *determinism* of ODEs is indispensable for the proof of Theorem 1. The subtraction procedure between solutions obtained with the same initial conditions in different simulation runs returns zero value only for ODEs, but not for stochastic differential equations (SDEs) as we shall see in Section 4.

Fourth, the contents of Theorem 1 is not itself a "problem", but rather it is an "indicator" that at least one or maybe both of the listed premises are false in the actual world in which we live in, where evolution of human consciousness is possible through natural selection.

The importance of Theorem 1 is that it provides concrete guidelines on how a putative physical theory of causally potent consciousness should look like. In particular, from the constructed proof of the theorem, it is clear that the conscious experiences $\Psi(0), \Psi(t_1), \Psi(t_2), \ldots, \Psi(t_n)$ *need to be present inside* the physical equations in order to affect the future dynamics of the brain and the physical world. This endorses a form of "reductionism", albeit not necessarily a classical one as we shall see next.

3.2. *Classical reductionism implies lack of free will*

Definition 2. (Reductionism) *Reductionism* is the philosophical stance that the brain (or a part of the brain) *is identical with* the mind, $\Phi = \Psi$. We have already discussed clinical evidence that the seat of human consciousness is located in the *brain cortex* (cf. Section 2). Pinpointing the exact brain area hosting consciousness, however, may not be relevant for the subsequent discussion because cortical pyramidal neurons are comprised from the same basic chemical elements as other cells in the body, hence the experiential aspect is immediately attributed to all living matter. Since the logical *identity* relation goes in both ways, it implies not only that human conscious experiences are built up from chemical atoms, but also that chemical atoms outside of the human brain *are* conscious experiences whose only distinction is that they are not "our" conscious experiences. In other words, reductionism endorses a form of "panpsychism" or "panexperientialism", according to which the physical world is comprised from conscious experiences. These experiences can be attributed to multiple minds, thereby avoiding the depressing thesis of "solipsism," according to which our mind is the only thing that exists in the universe.

Expressed as a percentage from the total mass, the elemental composition of the brain is 74.4% oxygen (O), 12.1% carbon (C), 10.6% hydrogen (H), 1.8% nitrogen (N), 0.4% phosphorus (P), 0.1% sulfur (S) and 0.6% other elements (Na, K, Mg, Cl, etc.). Our estimates for the brain elements are based on the chemical composition of the brain including 78.9% water, 11.3% lipids, 9.0% proteins, 0.2% nucleic acids and 0.6% electrolytes.[89–91] The chemical composition of the brain contains fewer chemical elements compared to the complete periodic table of chemical elements. This leaves the theoretical possibility to identify conscious experiences with only those chemical elements that appear in the brain, while leaving other elements that occur naturally in the surrounding world as lacking any experiences. Here, we have not entertained such a partial attribution of experiences to only part of the world because all chemical elements are ultimately composed from protons, neutrons and electrons. If experiences are attributed to all elementary particles, the result will be exactly *panexperientialism*.

The reductive claim expressed by the *mind–brain identity*, $\Phi = \Psi$, is that the two distinct labels "Φ" and "Ψ" refer to the same thing in physical reality. For example, when we refer to the real person Alice, we may use either the personal name "Alice" or words like "she" or "herself". Even though the three words, "Alice", "she" and "herself", are different, they all mean the same thing, which is that particular individual named Alice. The distinction between the label "Alice" and the real person Alice is the same as between the *map* and the *territory*.[92] In other words, the mind–brain identity stipulates that when we talk about "pyramidal neurons firing electric spikes in the brain cortex", we literally talk about conscious experiences that exist in the real world. Or to put it differently, the reductionism denies the philosophical stance of "naive realism" according to which the phrase "pyramidal neurons firing electric spikes in the brain cortex" refers to existing insentient pyramidal neurons firing electric spikes in the brain cortex. If what exists in the world as *territory* is only the conscious mind (experience), then the "brain" is just a label on the *map* of our scientific theory that refers to the existing mind. In contrast, the mind–brain problem is unavoidable in "naive realism" where the "brain" refers to an existing brain composed on insentient matter that is then expected to somehow generate the conscious experiences (sentience) that comprise the conscious mind. Having clarified the meaning of the mind–brain identity thesis, we are ready for its implications in classical physics:

Theorem 2. *Classical reductionism implies that the conscious mind is causally potent and affects the physical world. Such conscious mind, however, lacks free will because it cannot choose its dynamics.*

Proof. Provided with an initial state of the brain $\Phi(0)$ and the world $W(0)$, it is always possible to deterministically compute the future states of the brain $\Phi(t)$ and the world $W(t)$ for any future $t > 0$. However, since the brain and the mind are identical, $\Phi(t) = \Psi(t)$, for all $t \geq 0$, it follows trivially that the initial state of the mind $\Psi(0)$ affects the future state of the brain $\Phi(t)$. The brain is part of the world, which means that the mind affects the world. This is not surprising because from the mind–brain identity, the nature of the physical world becomes essentially mental. Because consciousness is present in the physical equations, the subtraction argument based on simulation of brain dynamics in the presence or absence of conscious experiences cannot be applied to reductionism due to the fact that the *identity of an entity with itself* cannot be turned off. By logical necessity, it is always true that $\Phi(t) = \Phi(t)$ and it is always false that $\Phi(t) \neq \Phi(t)$. Therefore, if $\Phi(t)$ is a mental state, then its mental nature cannot be turned off without negating $\Phi(t)$ itself. This concludes the proof of the first part of the theorem. The second part, which concerns the lack of free will, follows directly from the deterministic dynamics of ODEs. $\square$

The main advantage of reductionism is that the conscious mind is present in the physical equations that govern the future time dynamics of the physical world. The resulting panexperientialism then provides a mental substrate from which more complex and elaborate minds could evolve through natural selection. In the words of the psychologist William James, the natural evolution of the human mind requires the existence of "mind dust" in nature.[1] We agree with that conclusion, but we have formulated it in terms of physical reductionism and presence of conscious experiences in the system of ODEs that govern the behavior of the physical world.

Still, there is an unsettling inconsistency between classical physics and panexperientialism, because classical physics postulates that the brain is "observable", whereas the phenomenological, qualitative nature of conscious experiences is "unobservable".[17,18] To fix this serious problem, one would need to admit that not all aspects of physical reality are observable, hence classical physics needs to be repaired in some form or another. Fortunately, modern physicists have already found a better replacement for classical physics, which is provided by quantum physics.[93–95] Framing the mind–brain problem in the context of quantum physics, not only solves the "unobservability" and the "causal potency" of consciousness, but also introduces "free will" as we shall see next.

4. Causal potency of consciousness in quantum physics

Quantum physics conflicts conceptually with classical physics in a remarkable way. Because quantum systems possess the ability to accomplish physical tasks that are classically impossible, their "quantumness" or "quantum nature" is a valuable *physical resource* that is worth having.[96,97] Noteworthy, in the quantum world, there is a fundamental dichotomy between "existence" and "observability" because *what exists* is different from *what can be observed*.[4,17,18] Mathematically, this difference is expressed in the fact that *quantum states* are *vectors* $|\Psi\rangle$ in Hilbert space $\mathcal{H}$, whereas *quantum observables* $\hat{A}$ are *operators* on the Hilbert space $\mathcal{H}$.[93,94,98–100] For example, the general quantum state of a spin-1 particle can be written as

$$|\Psi\rangle = \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{pmatrix} = \alpha_1 |\uparrow_z\rangle + \alpha_2 |\bigcirc_z\rangle + \alpha_3 |\downarrow_z\rangle \tag{9}$$

where $|\uparrow_z\rangle$, $|\bigcirc_z\rangle$ and $|\downarrow_z\rangle$ are the eigenvectors of the z-component of the spin-1 observable

$$\hat{\sigma}_z = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix} = 1|\uparrow_z\rangle\langle\uparrow_z| + 0|\bigcirc_z\rangle\langle\bigcirc_z| - 1|\downarrow_z\rangle\langle\downarrow_z| \tag{10}$$

with corresponding eigenvalues 1, 0 and -1. The relationship between eigenvectors and eigenvalues of quantum observables and quantum measurements of quantum states is given by the Born rule[101,102] as follows: Suppose that we have the quantum state $|\Psi\rangle$ and we measure the quantum observable $\hat{\sigma}_z$. Then the quantum system chooses indeterministically the outcomes from the eigenvalues of $\hat{\sigma}_z$ and at the same performs a quantum jump into the corresponding eigenvector of $\hat{\sigma}_z$. The probabilities for the different choices are computed from the quantum probability amplitudes that define the quantum state $|\Psi\rangle$, namely, the quantum system produces the eigenvalue outcome 1 and jumps into the eigenstate $|\uparrow_z\rangle$ with probability $|\alpha_1|^2$, produces the eigenvalue outcome 0 and jumps into the eigenstate $|\bigcirc_z\rangle$ with probability $|\alpha_2|^2$, or produces the eigenvalue outcome -1 and jumps into the eigenstate $|\downarrow_z\rangle$ with probability $|\alpha_3|^2$. Symbolically, we write the probabilities for the alternative quantum jumps as

$$\text{prob}\,(|\Psi\rangle \hookrightarrow |\uparrow_z\rangle) = \langle\Psi|\uparrow_z\rangle\langle\uparrow_z|\Psi\rangle = |\alpha_1|^2 \tag{11}$$

$$\text{prob}\,(|\Psi\rangle \hookrightarrow |\bigcirc_z\rangle) = \langle\Psi|\bigcirc_z\rangle\langle\bigcirc_z|\Psi\rangle = |\alpha_2|^2 \tag{12}$$

$$\text{prob}\,(|\Psi\rangle \hookrightarrow |\downarrow_z\rangle) = \langle\Psi|\downarrow_z\rangle\langle\downarrow_z|\Psi\rangle = |\alpha_3|^2 \tag{13}$$

Quantum indeterminism is a characteristic feature of quantum systems because for the description of the outcomes of a sequence of quantum measurements performed on a dynamic quantum system, one needs to introduce a stochastic process $\Gamma(t)$ and solve *stochastic differential equations* (SDEs).[103] Because the *stochastic calculus* that is needed to solve SDEs was developed by the Japanese mathematician Kiyosi Itô,[104–107] it is often referred to as *Itô calculus*.[108]

Example 2. (Stochastic dynamics) To illustrate the mathematical properties of quantum stochastic dynamics, consider the jerky Lorenz-like system into which is injected the quantum stochastic process $\Gamma(t)$

$$\left(\frac{d}{dt}\right)^3 x = -\left(\frac{d}{dt}\right)^2 x - 4\frac{d}{dt}x + 5x - x^3 + \Gamma \tag{14}$$

where $\Gamma(t)$ is a continuous function obtained through linear interpolation of a sequence of unbiased choices from the set $\{1, 0, -1\}$ performed at unit time intervals. The remarkable feature of stochastic dynamics is that every simulation run produces almost surely a different result (Fig. 4). This is a direct consequence from the fact that the stochastic process $\Gamma(t)$ produces different sequences of chosen outcomes for different simulation runs. If the simulation is run for t units of time, the stochastic process $\Gamma(t)$ will involve $t+1$ unbiased choices. Therefore, the probability to produce two identical simulation runs is $\text{prob}(\text{run } 1 = \text{run } 2) = \left(\frac{1}{3}\right)^{t+1}$, which approaches zero in the limit $t \to \infty$. It is in this sense that different stochastic runs produce "almost surely" different results.

The theory of stochastic differential equations (SDEs) is very rich and includes the theory of ordinary differential equations (ODEs) as a special case. There are two ways that one can obtain deterministic dynamics. First option is to consider the expectation values of the quantum measurements performed on the quantum system.[96, 109–111] This means that one needs to collect a sufficiently large sample of individual stochastic trajectories and then compute a single average trajectory. The computation of such an average trajectory may or may not be useful and its physical interpretation may or may not be meaningful due to potential presence of outliers. For example, the average of 1 billionaire and 999 poor people will be a "millionaire". Learning that the "average" person from the group of 1000 people is a "millionaire", however, is useless and utterly misleading because the distribution of the sample consists of 99.9% of poor people. Thus, working with expectation values and interpreting those expectation values requires good understanding of statistics[112] and quantum foundations.[113] Second option to obtain deterministic dynamics is to consider highly biased probability distributions with zero variance. In Example 2, we have considered unbiased stochastic process $\Gamma(t)$, which makes the resulting simulated trajectories equally likely. However, a uniform probability distribution can be continuously transformed into a highly biased probability distribution that is narrowly peaked onto a single outcome with zero variance. Quantum measurement

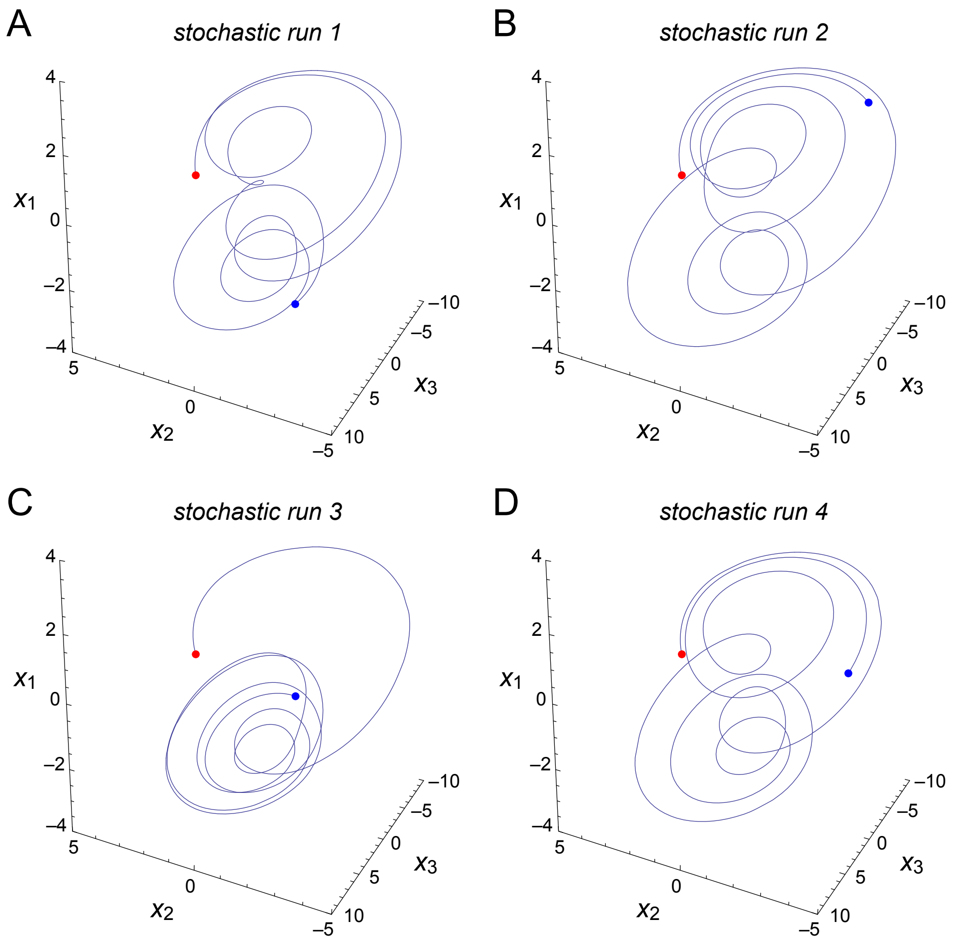

Fig. 4. Stochastic dynamics of a physical system whose time evolution is governed by the stochastic differential equation (14). Each of the four simulation runs (A–D) starts from the same initial state $S(0)$ with $x_1(0) = 1$, $x_2(0) = 2$ and $x_3(0) = 3$ and lasts for exactly the same period of time t =0–20. The characteristic feature of stochastic dynamics is that for each run the dynamic trajectory is almost surely going to be different as the outcomes of each stochastic choice are drawn from a certain probability distribution. The initial state is shown with a red point, whereas the final state is shown with a blue point. The units of t and x_1 are arbitrary, whereas the units of x_2 and x_3 are fixed as corresponding rates of change by (3) and (4).

theory[99,114] allows for physical realization of the full spectrum of quantum probability distributions from a completely uniform distribution to a highly nonuniform distribution consisting of a single narrow peak onto a single outcome. This means that the behavior of quantum systems depends critically on the measurement context. For some quantum measurement contexts, the resulting dynamic trajectory may appear to be random, whereas for other quantum measurement contexts the resulting dynamic trajectory may appear to be deterministic[115] (see also Section 5). That is why quantum physics is indispensable for the proper understanding of consciousness and free will. Furthermore, the causal potency of consciousness is no longer threatened in the quantum world, as we shall demonstrate next.

4.1. *Quantum functionalism does not guarantee causally potent consciousness*

Quantum functionalism understood as the *quantum brain state* $|\Phi\rangle$ producing the *conscious mind* Ψ is not in itself directly incompatible with causally potent consciousness. If the quantum stochastic dynamics of the brain state $|\Phi(t)\rangle$ is governed by a system of SDEs, it would follow that different simulation runs produce different results, hence the subtraction of two different stochastic runs will not be zero. This leaves room for the introduction of conscious action that is causally potent.

It should be noted that the brain is an open quantum system interacting with its physical environment, because the brain inputs sensory information and outputs motor information (Fig. 1). Nevertheless, we prefer writing the quantum brain state with the use of the brain quantum state vector $|\Phi\rangle$, rather than the brain density matrix $\hat{\rho}$, because we model "genuine" quantum stochastic dynamics that is due to the presence of objective wave function collapses[116–118] that lead to actualization of single measurement results with intermittent disentanglement of the brain and its physical environment[4] (see also Section 8). In fact, it can be shown that "no collapse" models of quantum mechanics, in which the whole universe as a closed system is described only by the Schrödinger equation, are no different than the classical models based on ordinary differential equations (ODEs). In particular, the Schrödinger equation is an ODE and the partial trace operation is perfectly deterministic procedure, which would imply that Theorems 1 and 2 apply to such "no collapse" quantum theory of consciousness. Extensive criticism of epiphenomenal consciousness in "no collapse" models of quantum mechanics has already been presented elsewhere[4] and will not be repeated here. Instead, we would like to focus on the resolution of the problem with mental causation provided by genuine stochastic "quantum jumps" that are mathematical representation of the objective physical wave function collapses in "dynamical collapse" models of quantum mechanics.

Theorem 3. *Quantum functionalism does not guarantee causally potent consciousness, but leaves room for upgrading the theory to a form of interactive mind–brain dualism.*

Proof. The main premises of quantum functionalism can be summarized as follows:

Premise 1. The act of functional production of the mind by the brain, $|\Phi(t)\rangle \to \Psi(t)$, entails an infinite list of productions at each time point t: $|\Phi(0)\rangle \to \Psi(0)$, $|\Phi(t_1)\rangle \to \Psi(t_1)$, $|\Phi(t_2)\rangle \to \Psi(t_2)$, $\ldots$, $|\Phi(t_n)\rangle \to \Psi(t_n)$.

Premise 2. The quantum states of the brain $|\Phi(t)\rangle$ and the surrounding world $|W(t)\rangle$ are governed by an explicitly given system of stochastic differential equations

(SDEs) in which only physical observables of the brain and the world are present.

From the second premise, we can infer that for each simulation run the quantum states of the brain $|\Phi(t)\rangle$ and the surrounding world $|W(t)\rangle$ undergo sequences of quantum jumps, which we will denote with the symbol $\hookrightarrow$ as follows:

$$|\Phi(t)\rangle = |\Phi(0)\rangle \hookrightarrow |\Phi(t_1)\rangle \hookrightarrow |\Phi(t_2)\rangle \hookrightarrow \ldots \hookrightarrow |\Phi(t_n)\rangle \hookrightarrow \ldots \tag{15}$$

$$|W(t)\rangle = |W(0)\rangle \hookrightarrow |W(1)\rangle \hookrightarrow |W(2)\rangle \hookrightarrow \ldots \hookrightarrow |W(t_n)\rangle \hookrightarrow \ldots \tag{16}$$

The unitary quantum interaction between the brain and its environment will lead to production of quantum entangled clusters of neurons in the brain cortex that will disentangle to the corresponding product states $|\Phi(t)\rangle \otimes |W(t)\rangle$ at the instances $t_0, t_1, \ldots, t_n$ when the definite measurement results are actualized through "dynamical collapses"[4] (see also Section 8). All the unitary quantum dynamics resulting from the Schrödinger equation, which is an ODE, will be dependent on the actual brain quantum Hamiltonian and is purposefully left implicit in the $\hookrightarrow$ symbols to prevent unnecessary distraction. The *conceptual highlight* in the above mathematical description is that, in general, performing a subtraction for two different stochastic runs 1 and 2 will not produce a zero result

$$|\Phi(t)\rangle_{\text{run 1}} - |\Phi(t)\rangle_{\text{run 2}} \neq 0 \tag{17}$$

$$|W(t)\rangle_{\text{run 1}} - |W(t)\rangle_{\text{run 2}} \neq 0 \tag{18}$$

This non-zero difference does not have to be attributed to the action of the conscious mind $\Psi(t)$, but if the quantum functionalism wants to have a theory of consciousness that is consistent with natural evolution, then it is possible to introduce as a postulate that the conscious mind is the agent that chooses the particular outcomes for the brain states at each quantum jump. The chosen brain states will then affect the state of the surrounding world, and the resulting theory will be a form of interactive mind–brain dualism. □

Despite that quantum physics is not incompatible with functionalism and causally potent consciousness, the very idea of "production" or "emergence" of consciousness is problematic for the following reasons.

First, the postulated emergence of consciousness is *ad hoc* and not different from the postulation of occurrence of "miracles". For example, one may as well postulate that a flutter of fairies appeared from the brain, then they performed a dance to welcome their fairy queen, and finally they all decided what the next brain state should be as an outcome of the current quantum jump. In other words, it is quite unsettling that the brain did not have the capacity to perform the choice of the quantum jump by itself without resort to any external agency.

Second, even if the emergence of a conscious mind Ψ_i is granted for each quantum brain state $|\Phi_i\rangle$, it is not clear what prevents the possibility of paranormal action? For example, how is it possible that the conscious mind Ψ_i recognizes that it can act upon the quantum brain state $|\Phi_i\rangle$ but not on another present quantum brain state $|\Phi_j\rangle$? Or to put it in more familiar terms, what is the physical mechanism that prevents Alice's consciousness to act upon Bob's brain and vice versa?

Third, even more severe objection to functionalism is the fact that it is easy for one to falsify it empirically. For example, to postulate that our conscious mind chooses what our brain state should be, implies that we are knowingly selecting which neuron in our brain will be firing and which neuron will remain silent, which ion channel is be open for the passage of electric current and which ion channel will remain closed. Introspectively, we can verify that we have no idea what brain state we are choosing. What is more, no human being on Earth has the slightest idea of what the actual chemical composition of their own brain is. Therefore, it seems that we are not knowingly choosing our brain states.

Fortunately, a safe way out from all these problems is provided by *quantum reductionism*.

4.2. *Quantum reductionism guarantees causally potent consciousness and free will*

Quantum reductionism identifies the quantum state of the brain $|\Phi\rangle$ with the conscious mind $|\Psi\rangle$ at all times, namely, $|\Phi(t)\rangle = |\Psi(t)\rangle$. Because the identity of a thing with itself cannot be logically turned off, it would then follow that all quantum states in the quantum world are comprised of conscious experiences. Thus, the resulting *quantum panexperientialism* provides a mental fabric for physical reality in which complex minds could evolve naturally from simpler minds.[4]

Theorem 4. *Quantum reductionism implies that the conscious mind is causally potent and affects the physical world. Such conscious mind possesses free will because it is able to choose among future courses of action.*

Proof. Since the quantum brain state and the conscious mind are identical, $|\Phi(t)\rangle = |\Psi(t)\rangle$, for all t, it follows trivially that the initial state of the mind $|\Psi(0)\rangle$ affects the future quantum state of the brain $|\Phi(t)\rangle$. The brain is part of the world, which means that the mind affects the world. The possession of free will follows from the stochastic dynamics. A quantitative measure for the amount of manifested free will is provided by the expected information gain from learning the actual sequence of choices made by the conscious mind (for details, see Sections 5 and 6). $\square$

One of the advantages of quantum reductionism is that because the quantum brain–mind identity relation goes in both ways, it would follow that the conscious mind has all of the properties of a quantum state and satisfies the axioms of a vector in Hilbert space. To highlight this fact, we no longer use the bare symbol "Ψ" for the conscious mind, but rather insert it inside a ket $|\Psi\rangle$ following Dirac's bra-ket notation.[93,119] This is highly informative from a theoretical perspective because one becomes equipped with quantum information theoretic *no-go theorems* that can be applied to consciousness in order to determine some of its physical properties.[4,120] For example, the quantum state vector of quantum physical systems is not observable due to a theorem by Busch,[121] which in turn explains why the conscious experiences are not observable.[4,17,18] Deriving the "unobservability" of conscious experiences as a prediction from the physical theory of consciousness is a remarkable achievement, especially when compared to classical reductionism, which is ripped by the internal inconsistency between the *unobservable mind* and the *observable brain*.

Another achievement of quantum reductionism is the ability to explain the physical difference between the unobservable mind and the observable brain (Fig. 5). In a quantum world, what can be observed is different from what exists in the form of a quantum state $|\Psi\rangle$. This physical difference is reflected in the mathematical formalism of quantum theory where quantum observables are described by operators $\hat{A}$ on Hilbert space $\mathcal{H}$. The observed outcomes in quantum measurements are the eigenvalues of the measured quantum brain observables. Because the eigenvalues are just numbers, these can be represented by bits of classical information and stored on a digital file as a string of 0s and 1s. This classical information is the "observable brain" and can be communicated to multiple external observers. For example, a microscopic picture of the anatomical organ (brain) that is inside your skull shows a neural network of neurons.[122] This neuronal picture is different from the conscious experiences that exists in reality and is compatible with the fact that we do not introspectively perceive ourselves as a collection of neurons. In other words, the picture of neurons is not the conscious mind itself, but what the conscious mind looks like from a third-person point of view. The same picture of the observable brain can be copied, multiplied, communicated to and simultaneously studied by multiple neuroscientists. The amount of classical bits of information that can be obtained while observing the quantum state of the brain $|\Psi\rangle$ is bound by Holevo's theorem in quantum information theory.[17,18,123]

Clear conceptual distinction between the "unobservable" mind and the "observable" brain is preserved by avoiding the usage of a brain density matrix $\hat{\rho}$, which is a *quantum observable*. For a pure brain state, i.e., $\mathrm{Tr}(\hat{\rho}^2) = 1$ it follows that $\hat{\rho} = |\Psi\rangle\langle\Psi|$ hence the state can be equivalently written as a ket vector $|\Psi\rangle$. Component parts of a composite quantum entangled system, however, do not have their own ket vectors, hence do not have their own minds, whereas they always have a reduced density matrix $\hat{\rho}$ that is not pure, i.e., $\mathrm{Tr}(\hat{\rho}^2) < 1$.[4] In other words, possessing a density matrix is not something special that is useful for *demarcation of mind boundaries* because any collection of quantum particles is guaranteed to have a reduced density matrix. On the other hand, it is indeed something special for a collection of quantum particles to have their collective quantum state vector $|\Psi\rangle$ because in general not every collection of quantum particles is guaranteed to have a quantum state vector. What is more, the factorizability of $|\Psi\rangle$ is informative with regard to mind boundaries, namely, a nonfactorizable $|\Psi\rangle$ corresponds to a single mind, whereas a factorizable $|\Psi\rangle = |\psi_1\rangle \otimes |\psi_2\rangle \otimes \ldots \otimes |\psi_k\rangle$ corresponds to a

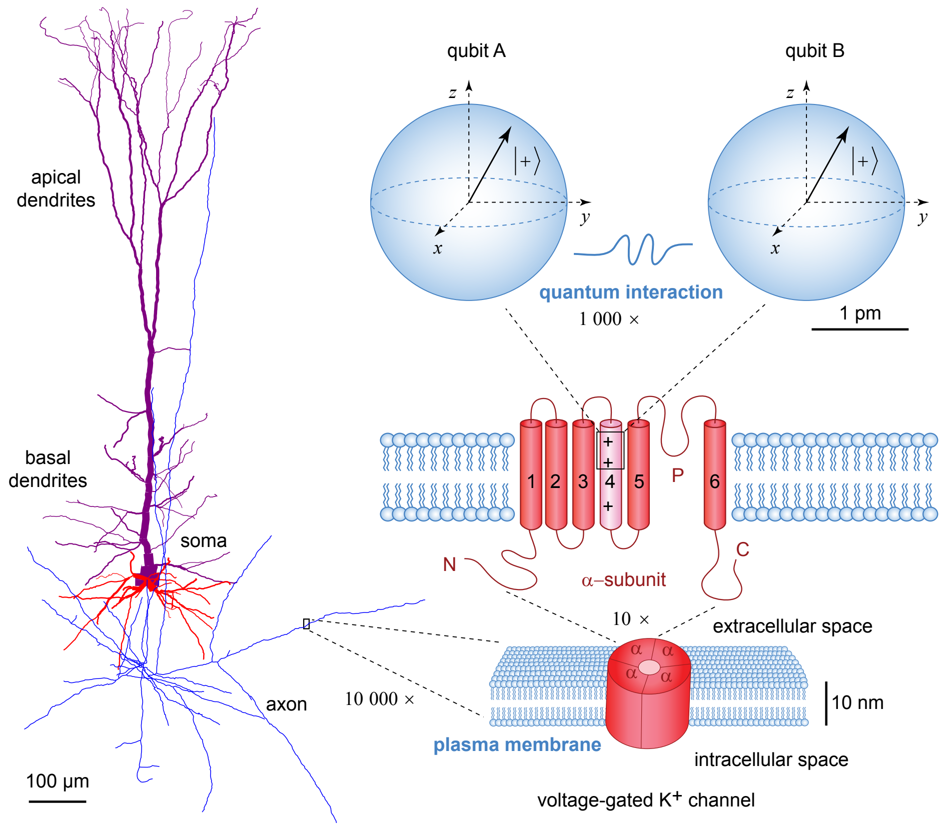

Fig. 5. Different levels of organization of physical processes within the central nervous system. At the microscopic scale, the brain cortex is composed of neurons, which form neural networks. The morphology of the rendered pyramidal neuron (NMO_77905) from layer 5 of rat somatosensory cortex (http://NeuroMorpho.Org) reflects the functional specialization of dendrites and axon for the input and output of electric signals, respectively. At the nanoscale, the electric activity of neurons is generated by voltage-gated ion channels, which are inserted in the neuronal plasma membrane. As an example of ion channel is shown a single voltage-gated K^+ channel composed of four protein α-subunits. Each subunit has six α-helices traversing the plasma membrane. The 4th α-helix is positively charged and acts as voltage sensor. At the picoscale, individual elementary electric charges within the protein voltage sensor could be modeled as qubits represented by Bloch spheres. For the diameter of each qubit is used the Compton wavelength of electron. Consecutive magnifications from micrometer (μm) to picometer (pm) scale are indicated by $\times$ symbol. Modified from Ref. 124.

collection of k separate minds[4,124] (see also Section 7).

Example 3. (Quantum reductionism forbids mind uploading) Within the quantum reductive approach, profound insights into the nature of consciousness arise from the characterization of every quantum state $|\Psi\rangle$ as a quantum coherent superposition of quantum probability amplitudes for potential future quantum events to occur. The *actualization* of one of these potential events occurs during quantum measurements with their associated quantum jumps. For example, consider the chemical composition of the voltage sensor of voltage-gated K^+ channel, which is built up from carbon (C), nitrogen (N), oxygen (O) and hydrogen (H) atoms[125–127] (Fig. 5). The positively charged hydrogen nuclei (protons) in the positively charged Lysine or Arginine amino acid residues can be located in different locations in space, the conformation of which determines whether the voltage-gated K^+ channel is in open or closed state.[128–130] Different conformations of Lysine or Arginine are realized by different distributions of C, N, O and H atoms. Each distribution could be realized with quantum probability given by the absolute square of the corresponding quantum probability amplitudes. Importantly, these quantum probability amplitudes are non-zero only for the mentioned C, N, O and H atoms, but are zero for atoms of other chemical elements. This leads to impossibility of the quantum probability amplitudes to be separated from their physical substrate. In modern science fiction scenarios, it is often imagined that human consciousness could be uploaded onto silicon-based computer chips.[131–134] Quantum reductionism, however, forbids physically such possibility. Indeed, let us imagine for the sake of argument that the quantum state of the voltage-gated K^+ channel could be cloned onto the silicon chip. Chemically, the silicon chip is comprised of silicon (Si) atoms. If the quantum state of the silicon chip has perfectly become the quantum state of the voltage-gated K^+ channel, then it would contain zero quantum probability amplitude for measurement of silicon (Si) atoms. Consequently, after we interact the silicon chip to measure it, we will observe the actualization of a voltage-gated K^+ channel. Similarly, if we were able to clone the complete quantum state of the brain onto a silicon chip, the silicon chip would turn into an organic brain tissue upon observation. Logically, the quantum probability amplitude for actualization of a brain does actualize a brain, but does not actualize a silicon chip. That is why the quantum information contained in the quantum probability amplitudes of a quantum state $|\Psi\rangle$ is fundamentally inseparable from its physical substrate. The quantum probability amplitudes are for the actualization of something, and this "something" is what we call the physical substrate of the quantum state.

Quantum stochastic dynamics of the quantum brain state $|\Psi(t)\rangle$, which is comprised of conscious experiences, solves at once both the causal potency problem and the free will problem.[115] If one performs a series of quantum measurements upon the quantum brain state $|\Psi(t)\rangle$, the result will be a sequence of actualized outcomes

$$|\Psi(t)\rangle = |\Psi(0)\rangle \hookrightarrow |\Psi(t_1)\rangle \hookrightarrow |\Psi(t_2)\rangle \hookrightarrow \ldots \hookrightarrow |\Psi(t_n)\rangle \hookrightarrow \ldots \tag{19}$$

The mathematics of stochastic differential equations does not have a placeholder for indicating the agent that makes the choices. This makes the mathematical theory generally applicable to many different contexts where the agent could be either *internal* or *external* to the system.

Example 4. (Genuinely stochastic dynamics) For the particular context when the agent is the modeled system itself, we say that the dynamics is *genuinely* stochastic and the modeled system possesses free will because it is able to make choices from a set of physically possible future courses of action.

Example 5. (Effectively stochastic dynamics) In the context when the agent is external to the modeled system, we say that the dynamics is *effectively* stochastic but the modeled system does not posses free will because it is unable to make choices. For example, consider two human players playing a chess game. When only the chess pieces are described, they move stochastically on the chess board because the two human players make choices and use the muscles of their hands to move the chess pieces. Since each chess piece is not an active agent itself, it exhibits effective stochastic dynamics where the word "effective" means that it only "looks like" stochastic dynamics due to the fact that the external cause is left out from the description. Indeed, if the two human players are included in the description and their choices are described by stochastic processes, then the motion of the chess pieces is completely deterministic as they "copy" exactly the outcome of the external human choice. The determinism of the copying action is due to the presence of conditional probabilities that are only selected from the set $\{0\%, 100\%\}$.

Example 6. (Simulated stochastic dynamics) In the special context when we are the external agent to a studied genuinely stochastic system, we say that we perform a simulation. In the process of solving a system of stochastic differential equations (SDEs), we make weighted choices ourselves and attribute the obtained results to the simulated system because it *could have produced* genuinely the same outcomes with the same probability. Of course, when we use computer programs to perform the simulations we do not even make weighted choices ourselves, but rather relegate the task to *pseudorandom number generator*, which is perfectly deterministic process that produces outcomes with the required statistics. Because we do not know how the pseudorandom number generator works and do not control directly its initialization state, we pretend that our "ignorance" of the outcome of the deterministic pseudorandom number generation is a good enough substitution for the genuine stochastic process that generates truly random numbers.

In the physical world, only quantum systems could exhibit genuinely stochastic behavior. The Brownian motion of a classical particle in a fluid could be effectively modeled with a stochastic differential equation (SDE), but the resulting trajectory is deterministic if the exact positions and velocities of the rest of the particles in the fluid are taken into account.[135] In other words, all the information that is needed to predict exactly the trajectory of the Brownian particle is available somewhere in the environment. If one wants to transmit securely a secret message along a cryptographic channel, it would be unwise to rely on effective stochasticity because someone may find a way to extract the hidden information from the environment. Instead, one could use a genuine quantum system, which is guaranteed to produce truly random numbers.[136,137] Thus, quantum systems are the only physical systems that could exhibit genuinely stochastic dynamics and manifest their free will.[115]

The action of the conscious mind on its own brain is no longer mysterious in quantum reductionism. In quantum functionalism, it was assumed that it is the brain that produces the mind and then the mind should somehow find a way to affect its own brain. In quantum reductionism, the roles are reversed and in a fundamentally mental physical world it is the mind that produces the "observable brain". The conscious mind is the set of all physical potentialities that could be actualized during quantum measurement, whereas the observable brain is the classical information that characterizes the actual outcome that has been chosen by the mind. Thus, it is logically impossible for the actualized mind's choice not to be that mind's choice. For example, if Alice's observable brain is the actualized choice of Alice's mind, it is impossible for Alice to choose Bob's observable brain, because Alice's mind is not Bob's mind.

The most challenging problem met by quantum functionalism (cf. Section 4.1) was the empirical fact that we do not knowingly choose our observable brain state. When we choose to move our hand, we experience a desire to voluntary move our hand without knowing which neuron in our motor brain cortex is firing electrically. In quantum reductionism, there is no mystery to such introspective testimony because the quantum jumps given in Eq. (19) describe a stochastic dynamics from one conscious experience to another conscious experience. For example, we have a desire to consciously move our hand and then we evolve into a conscious state in which we have triggered the hand motion. The conscious state with triggered hand motion does not also have extra self-referential information of how our brain looks like from a third-person perspective. Quantum measurement theory[114,138] further makes it clear that a quantum state does not observe itself because a quantum state cannot measure itself. The quantum measurement generates communicable information, i.e. classical bits of information.[4] When the conscious mind experiences certain qualia, e.g. the redness of a red rose, no communicable information is generated with respect to the phenomenological nature of those qualia, e.g. what is it like to experience the redness of a red rose. Therefore, the conscious mind *experiences* itself, but *does not observe* itself in the technical sense of the word "observation" understood as "physical measurement".[17,18] Furthermore, the conscious mind should not have an extra knowledge of what the observable brain is, because the brain picture is dependent on the measuring instrument. For example, microscopic image of pyramidal neurons in the brain cortex could be taken at different magnification and with different resolution, the ongoing electrophysiological processes in dendrites, axons and synaptic junctions could be described in molecular language,[120,139] and so on. Holevo's theorem in quantum information theory puts a strict upper bound on the amount of classical information that can be obtained by an external observer from a given quantum state $|\Psi\rangle$. Thus, the "observable brain" is nothing but the physically admissible upper limit of classical information that can be communicated to an external observer with regard to each and every actualized choice of the dynamically evolving conscious mind $|\Psi(t)\rangle$.

5. Free will and stochastic processes

Definition 3. (Free will) *Free will* is the inherent capacity of a physical system to make genuine choices among at least two physical outcomes.[4,115]

The *act of choosing* always selects *one actualized outcome* from *several available possible outcomes*. Consequently, the information obtained from a *single choice* is insufficient for an external observer to differentiate between a deterministic system and a system endowed with free will. This is because the single choice produces a single outcome without any accompanying evidence of the physical existence of the other potential alternative outcomes. Instead, the external observer needs to collect information from *repeated choices* performed by the target system and accumulate a statistical probability distribution for different actualized outcomes. If the probability distribution is 100% peaked onto a single outcome, then the system is guaranteed to be deterministic. However, if the probability distribution is spread over several outcomes, then this could be a manifestation of the target system making genuine choices among those several outcomes. Whether or not the target system actually makes choices can only be answered by the physical laws that describe the properties of physical reality. If the physical laws allow/disallow the capacity for making genuine choices, then real physical systems can/cannot make genuine choices. The mathematical representation of the *act of choosing* necessitates the adoption of a *generalized type of processes* known as *stochastic processes*.

Definition 4. (Stochastic process) A *stochastic process* Γ is a generalized type of process that implements the actualization of a single outcome x_a selected from a sample space $X = \{x_1, x_2, \ldots, x_n\}$, where the probability weights for the alternative outcomes are given by corresponding probability distribution $P = \{p_1, p_2, \ldots, p_n\}$ with normalized sum $\sum_n p_n = 1$.

In modern mathematical software, the function that generates an instance of a stochastic process is often called *weighted random choice*. Typically, the latter terminology does not create any confusion because the majority of mathematical models of stochastic phenomena focus on the presence of external noise due to uncontrollable factors in the environment. The attachment of the word *random* to the phrase *weighted choice*, however, is misleading and may obscure the fact that a stochastic process Γ can generate a completely deterministic trajectory. In fact, the most important ingredient in the definition of the stochastic process Γ is the probability distribution $P = \{p_1, p_2, \ldots, p_n\}$ and it can continuously vary from a completely biased distribution peaked onto a single outcome to completely unbiased distribution equally spread over all possible outcomes. Therefore, without having any information about the probability distribution $P = \{p_1, p_2, \ldots, p_n\}$, the *stochastic process* Γ *can be any process* including a completely deterministic one.

Example 7. (Stochastic processes with different bias) The simplest stochastic process Γ can be realized with a sample space consisting of only two outcomes $X = \{0, 1\}$, each of which can be realized with corresponding probability weight given by $P = \{p_0, p_1\}$. The bias in the two-outcome probability distribution can be quantified as

$$\mathscr{B} = p_1 - p_0 \tag{20}$$

with range $\mathscr{B} \in [-1, 1]$, where $\mathscr{B} = 1$ indicates a completely biased (deterministic) distribution such that $x = 1$ always occurs, $\mathscr{B} = -1$ indicates a completely biased (deterministic) distribution such that $x = 1$ never occurs, and $\mathscr{B} = 0$ indicates a completely unbiased (indeterministic) distribution such that $x = 0$ and $x = 1$ are equally likely to occur. Individual stochastic runs of $n = 100$ repetitions of the weighted choice with different values of the bias $\mathscr{B}$ are shown in Fig. 6. Comparative analysis of the obtained stochastic plots for different values of the bias $\mathscr{B}$ clearly demonstrates that the number of transitions between $x = 0$ and $x = 1$ decreases with the increase of the bias towards unity, $\mathscr{B} \to 1$, and in the limit of complete bias, $\mathscr{B} = 1$, the trajectory generated by the stochastic process is completely deterministic. The Kolmogorov complexity and algorithmic randomness (incompressibility) of the realized sequences of 0s and 1s also decrease as $\mathscr{B} \to 1$ due to the appearance of longer and longer strings of consecutive 1s.

Now, after we have shown that the term *stochastic* does not by itself imply anything concrete with regard to *randomness* or *determinism* without taking into account the actual values of the *bias* $\mathscr{B}$, we are ready to discuss how the free will is constrained by the presence of non-zero bias $|\mathscr{B}| \neq 0$. Also, we will show that randomness is the external manifestation of free will as perceived by observers outside the target agent.

Example 8. (External manifestation of free will) Suppose that Alice and Bob are agents endowed with free will and each one performs a series of $n = 100$ completely unbiased choices of $X = \{0, 1\}$ with $P = \{p_0 = \frac{1}{2}, p_1 = \frac{1}{2}\}$, hence $\mathscr{B} = 0$. From the perspective of Bob, the sequence of outcomes chosen by Alice is going to be statistically random due to approximately equal occurrence of 0s and 1s. From the point of view of Alice, there is nothing random but only a manifestation of her own free will because she is the agent who made each of the choices and for each choice she was able to choose otherwise. When Bob performs his series of $n = 100$ completely unbiased choices, the roles are reversed, namely, it was Bob who made the choices and manifested his free will, while Alice perceives Bob's choices to be random. The sequences of chosen outcomes produced by Alice or Bob are statistically indistinguishable in terms of *algorithmic randomness* or *incompressibility*, which is maximal. Thus, *randomness* by itself as a statistical property of the generated outcomes is not something that is incompatible with *free will.* What is important for the attribution of free will is the location of the physical source of stochasticity, whether it is inside Alice or inside Bob.

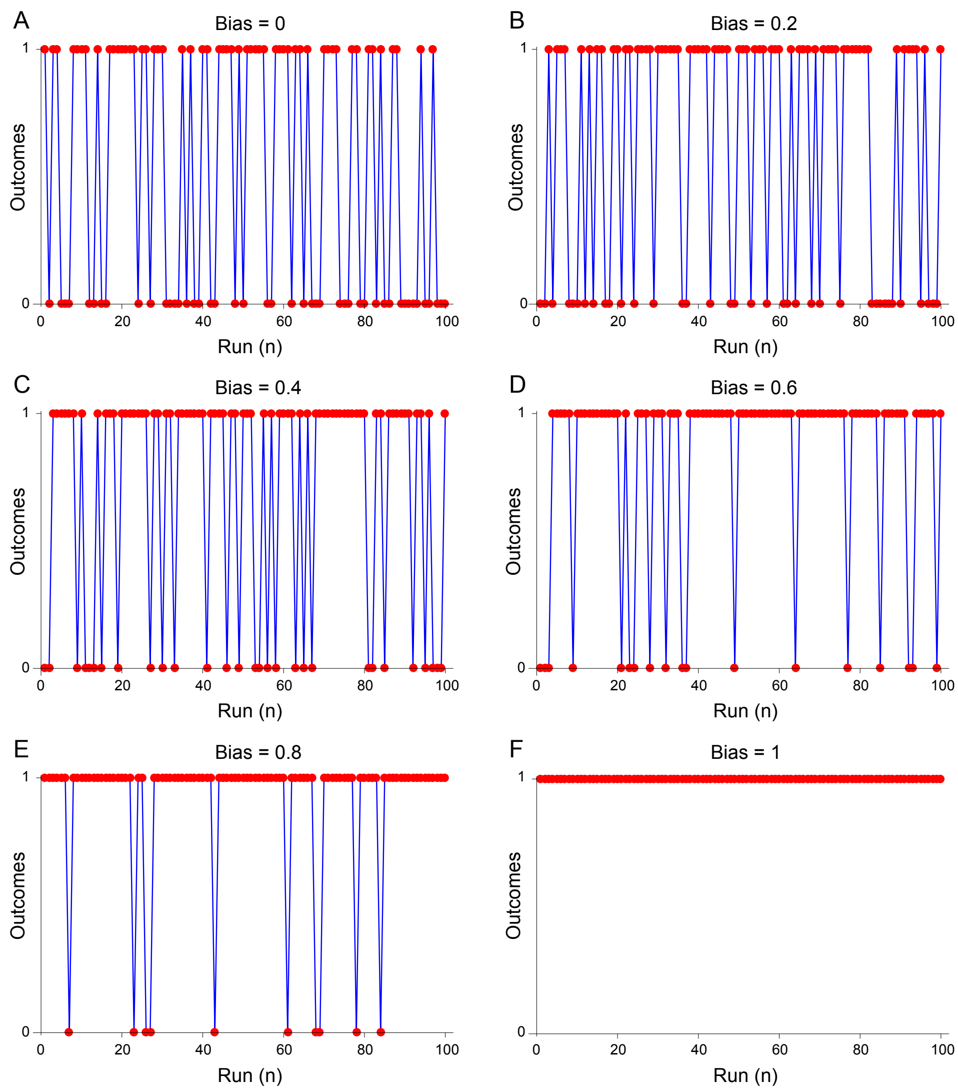

Fig. 6. Two-outcome stochastic processes Γ with different levels of the bias $\mathscr{B}$ in favor of the outcome $x = 1$ over the outcome $x = 0$. Each weighted choice was repeated for $n = 100$ times. The individual outcomes at each time point are shown as red points, whereas consecutive outcomes are connected with a blue line that serves as a visual guide for transitions between the two outcome values. For different levels of the bias $\mathscr{B}$ the stochastic process Γ can produce any trajectory from completely random to completely deterministic.

Definition 5. (Amount of free will) The *amount of free will* $\mathscr{F}$ of a target agent can be quantified by an external observer as the expected information gain in bits from learning the actualized outcome chosen by the target agent[115]

$$\mathscr{F} = -\sum_i p_i \log_2 p_i \tag{21}$$

The amount of free will $\mathscr{F}$ is strongly affected by the presence of an inherent bias $\mathscr{B}$, which leads to greater preference by the agent for one outcome over alternative outcomes.[115] For the simplest stochastic process Γ with a sample space consisting of only two outcomes $X = \{0, 1\}$, the individual probabilities can be expressed in terms of the bias (20) as follows

$$p_0 = \frac{1-\mathscr{B}}{2} \tag{22}$$

$$p_1 = \frac{1+\mathscr{B}}{2} \tag{23}$$

For $\mathscr{B} = 0$, the amount of free will $\mathscr{F}$ is maximal

$$\mathscr{F}(\mathscr{B} = 0) = -2 \times \frac{1}{2} \log_2 \frac{1}{2} = 1 \tag{24}$$

whereas for $|\mathscr{B}| = 1$, the amount of free will $\mathscr{F}$ is zero

$$\mathscr{F}(|\mathscr{B}| = 1) = -0 \log_2 0 - 1 \log_2 1 = 0 \tag{25}$$

Theorem 5. *The amount of free will for n repeated sequential choices of a stochastic process Γ with k possible outcomes $X = \{x_1, x_2, \ldots, x_k\}$ with probability weights $P = \{p_1, p_2, \ldots, p_k\}$ at each single time is cumulative and adds up to $n \times \mathscr{F}_0$, where $\mathscr{F}_0 = -\sum_{i=1}^{k} p_i \log_2 p_i$ is the amount of free will for each individual choice.*

Proof. Any stochastic processes with k possible outcomes $X = \{x_1, x_2, \ldots, x_k\}$ at a single time with probability weights $P = \{p_1, p_2, \ldots, p_k\}$, repeated at n times can be viewed as a single choice of the entire history of n outcomes $p_{i_1} p_{i_2} \ldots p_{i_n}$. Thus, the amount of free will is

$$\begin{aligned}
\mathscr{F} &= -\sum_{i_1=1}^{k} \sum_{i_2=1}^{k} \cdots \sum_{i_n=1}^{k} p_{i_1} p_{i_2} \ldots p_{i_n} \log_2 \left(p_{i_1} p_{i_2} \ldots p_{i_n}\right) \\
&= -n \left(p_1 + p_2 + \ldots + p_k\right)^{n-1} \sum_{i=1}^{k} p_i \log_2 p_i \\
&= -n \times 1^{n-1} \sum_{i=1}^{k} p_i \log_2 p_i = n \times \mathscr{F}_0
\end{aligned} \tag{26}$$

where we have used the fact that the probability weights are normalized to unity, $\sum_{i=1}^{k} p_i = 1$, together with the logarithmic product property converting log of a product into a sum of logs, namely, $\log(p_1 p_2) = \log p_1 + \log p_2$. □

An important consequence of Theorem 5 is that as long as the process is not deterministic, even strongly biased stochastic processes with very small but non-zero amount of free will per individual choice can produce arbitrary large changes in the resulting dynamics if the choices are repeated sufficient number of times. In the sample stochastic processes illustrated in Fig. 6, the total amount of free will exercised by the agent is $\mathscr{F} = 100$ bits for $\mathscr{B} = 0$, $\mathscr{F} \approx 97$ bits for $\mathscr{B} = 0.2$, $\mathscr{F} \approx 88$ bits for $\mathscr{B} = 0.4$, $\mathscr{F} \approx 72$ bits for $\mathscr{B} = 0.6$, $\mathscr{F} \approx 47$ bits for $\mathscr{B} = 0.8$, and $\mathscr{F} = 0$ bits for $\mathscr{B} = 1$.

6. Learning, biases and free will

The capacity for making choices is granted by physical laws. Because the evolutionary processes are governed by the physical laws of the universe, it is impossible for an organism to evolve a capacity that miraculously breaks the physical laws. What can be evolutionary achieved, however, is to acquire strategies for optimal utilization of the physical laws.[4,115]

In quantum mechanics, the measurement of quantum observables can generate the whole range of quantum probability distributions varying from completely biased to completely unbiased, where each probability distribution grants a different amount of free will $\mathscr{F}$ according to (21). From the Born rule, $P = \langle\Psi|\hat{A}|\Psi\rangle$ it is evident that the quantum probabilities change either by changing the quantum state of the measured system $|\Psi\rangle$ or changing the basis in which quantum measurement is performed, where the measurement basis is fixed by the eigenvectors of the measured quantum observable $\hat{A}$. Because typically the measured observable $\hat{A}$ and the measurement basis will be fixed by the environment, the organisms can acquire molecular mechanisms that modify the quantum state $|\Psi\rangle$ of their nervous system that undergoes repeated measurement by the environment. The preparation of different quantum state $|\Psi(t)\rangle$ for subsequent measurement could be viewed as a form of *knowledge acquisition* or *learning* as the time t progresses. Next, we will present the simplest quantum toy example involving a single qubit that is measured to produce a two-outcome sample space $X = \{0, 1\}$.

Example 9. (Dynamic biases and free will) Let the target system be a qubit (spin-$\frac{1}{2}$ particle) whose quantum state $|\Psi\rangle$ resides in a two-dimensional complex Hilbert space $\mathcal{H}$. Let the measurement by the environment determine the orientation of the spin along a fixed axis, which we can call the z-axis, in the real 3-dimensional space. The eigenbasis of the measurement is given by $\{|\downarrow_z\rangle, |\uparrow_z\rangle\}$ and the two-outcome sample space $X = \{0, 1\}$ is given by the eigenvalues corresponding to each eigenvector of the measured observable $\hat{A} = 1|\uparrow_z\rangle\langle\uparrow_z| + 0|\downarrow_z\rangle\langle\downarrow_z|$. If the spin of the qubit points away at an angle θ from the z_+-axis in the real 3D space (Fig. 7), the quantum state of the qubit in Hilbert space can be expressed as

$$|\Psi\rangle = \cos\left(\frac{\theta}{2}\right)|\uparrow_z\rangle + \sin\left(\frac{\theta}{2}\right)|\downarrow_z\rangle \tag{27}$$

where we have used the freedom to choose which direction in 3-dimensional space will be called x_+ thereby removing an inessential pure phase factor from one of the

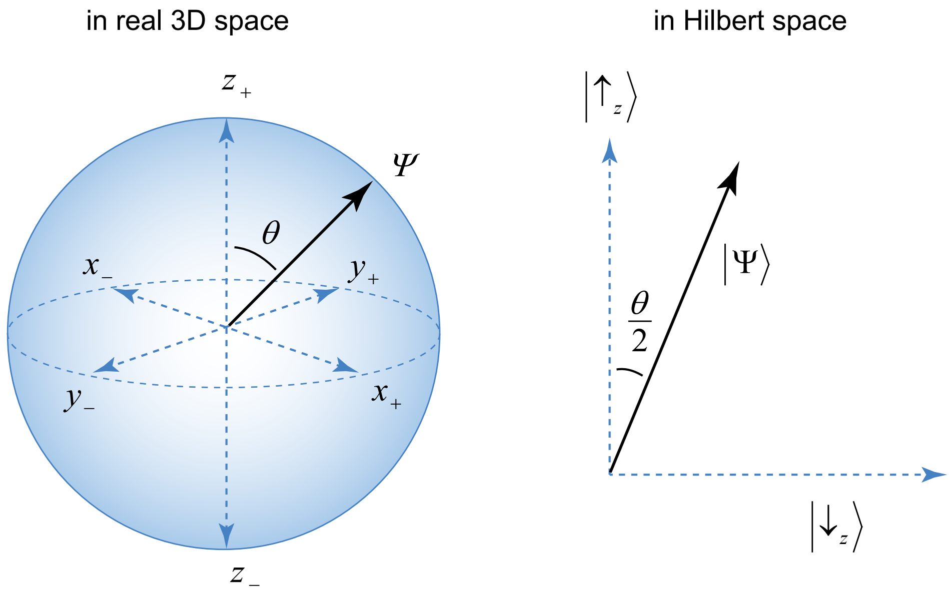

Fig. 7. Quantum toy example illustrating the physical realization of a stochastic process Γ with two-outcome sample space $X = \{0, 1\}$, eigenbasis of the quantum measurement given by $\{|\downarrow_z\rangle, |\uparrow_z\rangle\}$ and quantum probability distribution $P = \{p_0, p_1\}$ given by the Born rule.

superposed states. The Born rule determines the two probability weights for the actualization of each of the two possible outcomes

$$p_0 = \langle\Psi|\downarrow_z\rangle\langle\downarrow_z|\Psi\rangle = \sin^2\left(\frac{\theta}{2}\right) \tag{28}$$

$$p_1 = \langle\Psi|\uparrow_z\rangle\langle\uparrow_z|\Psi\rangle = \cos^2\left(\frac{\theta}{2}\right) \tag{29}$$

The weighted choice performed by the target qubit upon measurement is described by a stochastic process Γ with two-outcome sample space $X = \{0, 1\}$ and probability weights $P = \{p_0, p_1\}$ given by the Born rule. If the target qubit is not supported with some form of memory and it is *repeatedly prepared* in the same state $|\Psi\rangle$ and then *measured*[140,141] in the $\{|\downarrow_z\rangle, |\uparrow_z\rangle\}$ basis, the stochastic process will exhibit constant bias $\mathscr{B}$ in time

$$\mathscr{B} = \cos^2\left(\frac{\theta}{2}\right) - \sin^2\left(\frac{\theta}{2}\right) = \cos\theta \tag{30}$$

with constant free will in time

$$\mathscr{F} = -\frac{1-\mathscr{B}}{2}\log_2\left(\frac{1-\mathscr{B}}{2}\right) - \frac{1+\mathscr{B}}{2}\log_2\left(\frac{1+\mathscr{B}}{2}\right) \tag{31}$$

Inside the living brain endowed with memory and dopamine reward system, however, the qubit may participate in knowledge acquisition or learning in time t and exhibit a dynamic bias $\mathscr{B}(t)$ and dynamic amount of free will $\mathscr{F}(t)$. For example, if the outcome $|\uparrow_z\rangle$ is followed by reward, the qubit state $|\Psi(\theta,t)\rangle$ can be prepared with angle $\theta(t)\to 0$, increasing the bias $\mathscr{B}(t)\to 1$ and limiting the amount of free will $\mathscr{F}(t)\to 0$ with the expectation that future measurement of the spin along the z-axis will have an increased probability of choosing the outcome $|\uparrow_z\rangle$, hence again receiving reward. In contrast, if the outcome $|\uparrow_z\rangle$ is followed by punishment, the qubit state $|\Psi(\theta,t)\rangle$ can be prepared with angle $\theta(t)\to\pi$, increasing the bias in the opposite direction $\mathscr{B}(t)\to -1$ and limiting the amount of free will $\mathscr{F}(t)\to 0$ with the expectation that future measurement of the spin along the z-axis will have an increased probability of choosing the outcome $|\downarrow_z\rangle$, hence avoiding receiving another punishment. In the absence of previous experience or in the presence of conflicting information about the consequences of the two possible outcomes, the qubit state $|\Psi(\theta,t)\rangle$ can be prepared with angle $\theta(t)\to\frac{\pi}{2}$, decreasing the bias $\mathscr{B}(t)\to 0$ and increasing the amount of free will $\mathscr{F}(t)\to 1$ so that either outcome can be chosen and the consequences of each choice can be investigated through accumulation of new knowledge. In conclusion, the greater certainty in accumulated knowledge is associated with greater bias $\mathscr{B}(t)$ and lower amount of manifested free will $\mathscr{F}(t)$.

The single qubit in isolation is necessarily too simple and does not have the mechanisms of memory or reward. In the living brain, however, the elementary particles assemble into biomolecules that can store memories for prolonged periods of time and there are biochemical cascades that can be triggered by rewards or punishments. For example, voltage-gated ion channels can be subject to phosphorylation or dephosphorylation triggered by dopamine release, which modifies the sensitivity of their voltage-sensors to the transmembrane electric field and changes the probabilities for the ion channels to be in open or closed state (Fig. 5). The construction of precise quantum models of neuronal function is beyond the scope of this study and will be the subject of future research.

7. Quantum entanglement, mind binding and free will

The composite quantum state $|\Psi\rangle$ of k components resides in a tensor product Hilbert space $\mathcal{H}=\mathcal{H}_1\otimes\mathcal{H}_2\otimes\ldots\otimes\mathcal{H}_k$ formed by the individual Hilbert spaces $\mathcal{H}_1,\mathcal{H}_2,\ldots,\mathcal{H}_k$ of the corresponding components. Inside the composite Hilbert space $\mathcal{H}$, there are two kinds of states: those that are factorizable and those that are non-factorizable.

Definition 6. (Non-entangled quantum states) Factorizable composite quantum states in the form

$$|\Psi\rangle = |\psi_1\rangle\otimes|\psi_2\rangle\otimes\ldots\otimes|\psi_k\rangle \tag{32}$$

are called separable or non-entangled states.[142,143]

Definition 7. (Entangled quantum states) Non-factorizable composite quantum states, which cannot be expressed as a tensor product of component states

$$|\Psi\rangle \neq |\psi_1\rangle \otimes |\psi_2\rangle \otimes \ldots \otimes |\psi_k\rangle \tag{33}$$

are called non-separable or entangled states.[96,97]

In the quantum reductive approach, minds are attributed only to non-factorizable pure quantum states.[4] This means that quantum entanglement binds conscious experiences into a single unitary mind. The components of a quantum entangled state do not have definite individual state vectors, therefore it is not surprising that they are not endowed with individual minds. On the other hand, separable quantum states of the form (32) are such that both the composite system and the component systems have definite state vectors, namely, the composite state vector is $|\Psi\rangle$, whereas the individual component state vectors are $|\psi_1\rangle, |\psi_2\rangle, \ldots, |\psi_k\rangle$. In the latter case, the existence of a state vector such as $|\Psi\rangle$ is not sufficient to guarantee the attribution of a single mind, instead the separability of $|\Psi\rangle$ makes it a *collection of minds*. This is more easily understandable by taking into consideration that separability actually implies *statistical independence*.[124] In other words, two separable minds have no direct access to each other's conscious experiences and possess their own individual free will. The situation can be clarified with the following simplified toy example.

Example 10. (Separate minds have independent free will) Suppose that Alice and Bob are two separate minds, each of which is modeled by a definite two-level state

$$|\Psi_A\rangle = \alpha_0 |\downarrow_z\rangle + \alpha_1 |\uparrow_z\rangle \tag{34}$$

$$|\Psi_B\rangle = \beta_0 |\downarrow_z\rangle + \beta_1 |\uparrow_z\rangle \tag{35}$$

The composite state is

$$|\Psi_{AB}\rangle = |\Psi_A\rangle \otimes |\Psi_B\rangle \tag{36}$$

which makes the resulting quantum probabilities $P = \{p_{00}, p_{01}, p_{10}, p_{11}\}$ for each of the four possible outcomes $\{|\downarrow_z\downarrow_z\rangle, |\downarrow_z\uparrow_z\rangle, |\uparrow_z\downarrow_z\rangle, |\uparrow_z\uparrow_z\rangle\}$ to be factorizable

$$p_{00} = |\alpha_0|^2 |\beta_0|^2, \quad p_{01} = |\alpha_0|^2 |\beta_1|^2, \quad p_{10} = |\alpha_1|^2 |\beta_0|^2, \quad p_{11} = |\alpha_1|^2 |\beta_1|^2 \tag{37}$$

This means that the amount of free will $\mathscr{F}_{AB}$ manifested by the composite system comprised of Alice and Bob is exactly the sum of the amount of free will $\mathscr{F}_A$ manifested by Alice and the amount of free will $\mathscr{F}_B$ manifested by Bob

$$\begin{aligned}
\mathscr{F}_{AB} &= -\sum_{i=0}^{1}\sum_{j=0}^{1} |\alpha_i|^2 |\beta_j|^2 \log_2 \left(|\alpha_i|^2 |\beta_j|^2\right) \\
&= -\left(|\beta_0|^2 + |\beta_1|^2\right) \sum_{i=0}^{1} |\alpha_i|^2 \log_2 |\alpha_i|^2 - \left(|\alpha_0|^2 + |\alpha_1|^2\right) \sum_{j=0}^{1} |\beta_j|^2 \log_2 |\beta_j|^2 \\
&= -\sum_{i=0}^{1} |\alpha_i|^2 \log_2 |\alpha_i|^2 - \sum_{j=0}^{1} |\beta_j|^2 \log_2 |\beta_j|^2 = \mathscr{F}_A + \mathscr{F}_B
\end{aligned} \tag{38}$$

where we have used the logarithmic product property and the normalization of the component states, namely, $|\alpha_0|^2 + |\alpha_1|^2 = 1$ and $|\beta_0|^2 + |\beta_1|^2 = 1$.

Example 11. (Entanglement binds components into a single mind) Consider now what would happen if the composite state of two components A and B was maximally quantum entangled

$$|\Psi_{AB}\rangle = \frac{1}{\sqrt{2}} \left(|\downarrow_z\uparrow_z\rangle + |\uparrow_z\downarrow_z\rangle\right) \tag{39}$$

In the quantum entangled state, neither component A nor component B can have their own state vector and the resulting quantum probabilities $P = \{p_{00}, p_{01}, p_{10}, p_{11}\}$ for each of the four possible outcomes $\{|\downarrow_z\downarrow_z\rangle, |\downarrow_z\uparrow_z\rangle, |\uparrow_z\downarrow_z\rangle, |\uparrow_z\uparrow_z\rangle\}$ are not factorizable

$$p_{00} = 0, \quad p_{01} = \frac{1}{2}, \quad p_{10} = \frac{1}{2}, \quad p_{11} = 0 \tag{40}$$

The reduced density matrices for A or B are completely mixed

$$\hat{\rho}_A = \hat{\rho}_B = \begin{pmatrix} \frac{1}{2} & 0 \\ 0 & \frac{1}{2} \end{pmatrix} \tag{41}$$

which means that when viewed locally the probability weights for both A and B are $P = \{\frac{1}{2}, \frac{1}{2}\}$ for the outcomes $\{|\downarrow_z\rangle, |\uparrow_z\rangle\}$. Without taking into consideration the quantum entanglement, from the local statistics one may incorrectly conclude that component A manifested $\mathscr{F}_A = 1$ bit of free will and component B also manifested $\mathscr{F}_B = 1$ bit of free will. When the quantum entanglement is taken into account, however, it can be correctly concluded that the composite system exhibited only $\mathscr{F}_{AB} = 1$ bit of free will because the outcomes produced by A and B were perfectly anti-correlated, when A chose $|\downarrow_z\rangle$ it was always the case that B chose $|\uparrow_z\rangle$ and when A chose $|\uparrow_z\rangle$ it was always the case that B chose $|\downarrow_z\rangle$. This apparent *subadditivity* of the manifested amount of free will, namely, $\mathscr{F}_{AB} = 1 < 1+1 = 2 = \mathscr{F}_A + \mathscr{F}_B$ is due to the *quantum correlations* resulting from quantum entanglement. In other words, the component subsystems cannot manifest their own free will independently from the rest of the composite quantum entangled state. Instead, it is only the composite quantum entangled system as a whole that manifests its own independent free will and imposes quantum correlations on the components. The attribution of free will to the composite entangled system, which has a non-factorizable state vector and a mind, but not to the component subsystems which have neither their own state vectors nor independent minds, provides a one-to-one correspondence between non-factorizable pure quantum state vectors, single minds and their independent free will.[4] To summarize, pure quantum entangled systems have a single mind and it is the composite quantum entangled system as a whole that possesses the free will.

8. Wave function collapse and disentanglement

The capacity of quantum systems to make genuine choices requires a physical *actualization* process that converts one of the *possible* outcomes into an *actual* outcome. It is exactly the act of actualization that makes the stochastic process Γ suitable for describing the concept of choosing. Without actualizations there can be only possibilities without anything ever happening in reality. In quantum physics, this means that the unitary quantum dynamics prescribed by the Schrödinger equation can only produce possibilities, but it needs a stochastic process governed by the Born rule in order to create actualities. The stochastic actualization process is referred to as *wave function collapse* and is expected to occur when the composite quantum system reaches a certain energy threshold $\mathscr{E}$. The energy threshold $\mathscr{E}$ for wave function collapse is a parameter to be determined experimentally,[118, 144, 145] but in the context of quantum reductive theories of consciousness it is expected to be way above the energies of individual elementary particles and slightly below the total energy consumed by the metabolically active human brain.[4] The existence of an energy threshold $\mathscr{E}$ for wave function collapse solves the measurement problem in quantum mechanics and prevents the whole universe into getting entangled into a single universal cosmic mind.[4, 18]

Example 12. (Schrödinger's cat) The measurement problem results from the unitary quantum dynamics prescribed by the Schrödinger equation

$$\imath\hbar\frac{d}{dt}|\Psi(t)\rangle = \hat{H}|\Psi(t)\rangle \tag{42}$$

where $\hat{H}$ is the Hamiltonian and $|\Psi(t)\rangle$ is the time-dependent quantum state vector of the system. The Schrödinger equation is a linear ordinary differential equation (ODE) with formal solution

$$|\Psi(t)\rangle = e^{-\frac{\imath}{\hbar}\hat{H}t}|\Psi(0)\rangle = \hat{U}|\Psi(0)\rangle \tag{43}$$

The linearity implies that if $|\Psi_1\rangle$ and $|\Psi_2\rangle$ are two solutions of the Schrödinger equation, then any linear combination of those two solutions is also a valid solution

$$|\Psi\rangle = \alpha_1|\Psi_1\rangle + \alpha_2|\Psi_2\rangle \tag{44}$$

The unitary time evolution operator $\hat{U} = e^{-\frac{\imath}{\hbar}\hat{H}t}$, which governs the quantum dynamics, is also linear

$$\hat{U}\left(\alpha_1|\Psi_1\rangle + \alpha_2|\Psi_2\rangle\right) = \alpha_1\hat{U}|\Psi_1\rangle + \alpha_2\hat{U}|\Psi_2\rangle \tag{45}$$

In order to see how the unitary quantum dynamics leads to quantum entangled superpositions of macroscopic devices, consider a single photon $|\gamma\rangle$, which if prepared in a state with horizontal polarization $|\gamma_H\rangle$ is detected by a macroscopic measuring device initially prepared in state $|M\rangle$ whose pointer moves horizontally to final state $|M_H\rangle$ and if prepared in a state with vertical polarization $|\gamma_V\rangle$ is detected by the same macroscopic measuring device $|M\rangle$ but the pointer moves vertically to final state $|M_V\rangle$. For each of the two alternative cases, the action of the unitary operator governing the quantum dynamics of the composite system composed of the photon and the measuring device can be written as

$$\hat{U}|\gamma_H\rangle|M\rangle = |\gamma_H\rangle|M_H\rangle \tag{46}$$

$$\hat{U}|\gamma_V\rangle|M\rangle = |\gamma_V\rangle|M_V\rangle \tag{47}$$

Now, consider what will happen if the photon is polarized diagonally $|\gamma_+\rangle = \frac{1}{\sqrt{2}}(|\gamma_H\rangle + |\gamma_V\rangle)$ before it is sent to the measuring device. The linearity of the unitary time evolution operator creates an entangled quantum state

$$\begin{aligned} \hat{U}|\gamma_+\rangle|M\rangle &= \hat{U}\,\frac{1}{\sqrt{2}}(|\gamma_H\rangle + |\gamma_V\rangle)\,|M\rangle \\ &= \frac{1}{\sqrt{2}}(\hat{U}|\gamma_H\rangle|M\rangle + \hat{U}|\gamma_V\rangle|M\rangle) \\ &= \frac{1}{\sqrt{2}}(|\gamma_H\rangle|M_H\rangle + |\gamma_V\rangle|M_V\rangle) \end{aligned} \tag{48}$$

Experimentally, we always observe only one of the two possible alternative readings of the measurement device, but never the entangled superposition of both. To highlight the conflict between experimental observations and entangled macroscopic superpositions, Erwin Schrödinger proposed that the measuring device be replaced by a cat placed inside a box equipped with a photosensitive mechanism that releases poisonous gas only if the photon is detected in one of the states, e.g., $|\gamma_H\rangle$. In this case, the resulting entangled macroscopic quantum superposition will include a cat that is simultaneously dead and alive.[146,147]

The conflict between quantum mechanics and the lack of observable macroscopic superpositions is solved by the energy threshold $\mathscr{E}$ for wave function collapse. If the measuring device is sufficiently large to pass the energy threshold $\mathscr{E}$, the quantum entangled state (48) undergoes stochastic disentanglement Γ to one of the two separable outcomes $\{|\gamma_H\rangle|M_H\rangle, |\gamma_V\rangle|M_V\rangle\}$ with probability weights given by the Born rule $P = \{\frac{1}{2}, \frac{1}{2}\}$. Here, we recall that the concise product notation of two kets actually implies the tensor product, namely, $|\gamma_H\rangle|M_H\rangle \equiv |\gamma_H\rangle \otimes |M_H\rangle$ and $|\gamma_V\rangle|M_V\rangle \equiv |\gamma_V\rangle \otimes |M_V\rangle$.

Using all of the technical concepts defined above, we are now able to briefly outline how the quantum reductive model of consciousness is supposed to operate.

Example 13. (Quantum reductive model of consciousness) The model of how our consciousness inputs sensory information from the environment, makes choices and outputs those choices to the environment implements a repeated cycle consisting of several steps.

Step 1. The starting point of the cycle can be defined to be a part of the anatomical brain cortex that is in a disentangled tensor product state of component biomolecules, ions or elementary particles, $|\Psi(t_1)\rangle = |\psi_1(t_1)\rangle \otimes |\psi_2(t_1)\rangle \otimes$

$\ldots \otimes |\psi_k(t_1)\rangle$. At this stage, the composite system $|\Psi(t_1)\rangle$ is a *collection of elementary minds*. Among the components $\{|\psi_1(t_1)\rangle, |\psi_2(t_1)\rangle, \ldots, |\psi_k(t_1)\rangle\}$ are included energy quanta coming from the environment that input sensory information from the surrounding world.

Step 2. Then, the components interact with each other and undergo unitary quantum dynamics prescribed by the Schrödinger equation. Nearby components entangle with each other to form small entangled clusters, then as time goes on these small entangled clusters entangle with each other to form larger entangled clusters, and so on, until a sufficiently large entangled cluster $|\Psi(t_2)\rangle$ reaches the energy threshold $\mathscr{E}$ for wave function collapse. At this stage, the quantum entanglement has bound the conscious experiences of all the component elementary minds into a single integrated conscious experience $|\Psi(t_2)\rangle$ that could be recognized as the *conscious self*. Furthermore, because the energy threshold $\mathscr{E}$ for wave function collapse is reached the *conscious self* has to make a choice and selects one of the available disentangled outcomes using its free will.

Step 3. After the choice is made, the chosen outcome by the conscious self is a disentangled state $|\Psi(t_3)\rangle = |\psi_1(t_3)\rangle \otimes |\psi_2(t_3)\rangle \otimes \ldots \otimes |\psi_k(t_3)\rangle$ such that some small portion of the disentangled components leaves the brain in order to output motor information towards the environment and gets replaced by incoming sensory energy quanta, for example, we can say that $|\psi_k(t_3)\rangle$ was an electric signal that left the brain and it was replaced by another sensory signal $|\psi_{k'}(t_3)\rangle$.

Step 4. The cycle can be considered complete when the growth of entangled clusters proceeds based on unitary quantum dynamics and another sufficiently large entangled cluster $|\Psi(t_4)\rangle$ reaches the energy threshold $\mathscr{E}$ for wave function collapse. This future version of the conscious self is composed by the same component subsystems indexed from 1 to $k-1$ together with the new sensory component k' and lacking the motor output component k sent to the environment. Thus, the single integrated conscious experience $|\Psi(t_4)\rangle$ is another instance of the conscious self that is aware of the new sensory information and bears the consequences of its past motor choice sent to the environment.

This quantum reductive cycle of consciousness has been constructed in such a way, so that it addresses several philosophical problems of consciousness at once (the mind–brain problem, the mind physical boundary problem, the mind binding problem, the mind causal potency problem, the free will problem, the mind inner privacy problem and the hard problem) and explains how the conscious self changes in time by making choices and how it can acquire knowledge based on the outcomes of past choices (for a detailed exposition and more extensive discussion, see Chapter 6 in Ref. 4).

9. Discussion

Our prehistoric ancestors have painted cave walls with exquisite art images of horses, bisons, mammoths, cave bears, lions, panthers, rhinoceroses, owls and other animals.[44,45,78–82] The anatomical description of these animals, some of which are long extinct, is precisely captured by the prehistoric artists.[148,149] Because the primary purpose of art is to elicit conscious experiences in the viewer, it is highly implausible to assume that the prehistoric artists have produced those paintings without having any experiences of the animals that were painted. Furthermore, due to the precise anatomical correspondence between the animals that were seen and the animals that were painted, it is fair to conclude that the conscious experiences of the prehistoric painters have been causally potent in producing the physical artifacts in the form of cave paintings.[83,84] The causal potency of consciousness in the physical world is a prerequisite for the evolution of human consciousness through natural selection of our animal ancestors. Yet, if one sticks to the premises of classical functionalism, it would appear that the conscious mind is causally impotent in the physical world and could not have evolved naturally. Therefore, one needs to reject classical functionalism and search for better physical alternatives.

Here, we have presented a thorough analysis of the ramifications of two contrasting approaches to consciousness, namely, functionalism or reductionism, within the two main physical frameworks provided by classical physics or quantum physics. To derive rigorous theoretical results about classical or quantum systems, we have comprehensively reviewed the mathematical theory behind ordinary differential equations (ODEs) or stochastic differential equations (SDEs) and have illustrated the characteristic properties of the resulting classical or quantum dynamics using minimal toy systems. Further consideration of the mind–brain relationship within each of the four theoretical schemes, namely, classical functionalism, classical reductionism, quantum functionalism or quantum reductionism, revealed that only the latter scheme supports both the causal potency of the conscious mind and free will.

Quantum reductionism endorses a form of panexperientialism, according to which elementary feelings are attributed to all quantum systems, including simple living organisms that do not have a nervous system. Such elementary experiences or elementary feelings would be also attributed to inanimate quantum materials with the explicit caveat that these are memoryless. In fact, our adult human consciousness is a combination of ongoing sensory experiences and a recalled memory of who we are. In medical practice, we have the striking observations of a newborn baby or an aged person with Alzheimer's disease, who definitely have conscious experiences, but either do not yet know who they are or have forgotten who they are. Without the ability to memorize past experiences, a quantum physical system will not be able to communicate to the surrounding world that it had those experiences. Therefore, the presence of transient memoryless experiences in inanimate quantum materials is not something to be bothered by, instead these are the primordial substrates for natural evolution from which the conscious experiences in living systems should have evolved, given the fact that living organisms use free energy and are capable of storing memories. During sleep, the energy consumption by the brain is rapidly suppressed, which is reflected by the fact that the conscious experiences during dreaming are often illogical and hard to recall. This further highlights the need of energy source for storing memories and recalling memories of past experiences, as opposed to merely having any form of experience.

Alternative quantum functional models of consciousness rely on the problematic idea that conscious experiences are somehow generated by an insentient brain substrate. For example, the Orch OR model proposed by Penrose and Hameroff postulates that the quantum brain dynamics is unconscious or subconscious, whereas it is the event of objective reduction that creates intermittently the "flashes" of conscious experiences with a frequency of 40 Hz. The problem with the latter claim is the reversed causal order, namely, if the event of objective reduction is causally potent in creating the phenomenological content of the conscious experience, then the generated conscious experience would be causally impotent to choose the outcome of the objective reduction. In contrast, the quantum reductionism advocated in this work holds that the quantum dynamics is continuously conscious and it is the ongoing consciousness that is causally potent in choosing the actual outcome of the objective reduction.

Philosophers often present together the problems of mental causation and free will.[150] In fact, the two problems are distinct but hierarchically organized as follows: (1) causally impotent mind cannot cause any course of action in the physical world, (2) causally potent mind that lacks free will can cause exactly one course of action in the physical world, and (3) causally potent mind that possesses free will can choose between two or more possible courses of action in the physical world. This hierarchy of the two problems implies that one could solve the problem of mental causation without solving the free will problem, whereas the converse is impossible, namely, solving the free will problem cannot be done without also solving the problem of mental causation.

In ethics, the hierarchical organization of the problems of mental causation and free will leads to different levels of blame attribution. If your conscious mind is causally potent, then you can be blamed for what you have done. If in addition to causal potency your conscious mind possesses free will through which to execute choices, then you can also be blamed for what you have not done. The logic of blame attribution is straightforward: (1) if your conscious mind is causally potent but lacks free will, then there is only a single course of action that you are able to cause in the physical world. Since you could not have done otherwise, you can only be blamed for what you have done, but not for what you have not done. Alternatively, (2) if your conscious mind possesses free will, then you are able to make a genuine choice among at least two possible courses of action. This immediately makes you morally responsible not only for what you have chosen to do, but also for what you have chosen not to do. In other words, the existence of free will comes with the burden of having to contemplate the consequences of your actions so that you not only avoid causing harm, but also avoid missing potential benefits had you acted differently.

In evolution theory, the lack of causal potency is much more harmful proposition compared to the lack of free will because a causally impotent human consciousness could not have evolved through natural selection. Consequently, the main goal of the present study was to derive rigorously the causal potency or impotency of consciousness within the four theoretical schemes including functionalism or reductionism in classical or quantum physics. Noteworthy, utilizing the mathematical properties of ordinary differential equations, we have demonstrated that classical functionalism is not a plausible physical theory of consciousness and classical physics is a disadvantageous framework for approaching the mind–brain problem. Then, we have established that quantum physics is not just an exotic extravagance, but an indispensable general theoretical framework supporting stochastic differential equations that are ideally tailored to describe dynamic trajectories produced by sequential choices. Further building upon the quantum information theoretic properties of quantum states and quantum observables, we have shown that quantum reductionism predicts unobservable conscious mind that is causally potent in choosing the future course of action of the observable brain. This explains how the human consciousness could have evolved through natural selection in our animal ancestors.

References

1. W. James. The Principles of Psychology. Volume 1. Henry Holt and Company, New York, 1890.
2. E. B. McGilvary. The stream of consciousness. *The Journal of Philosophy, Psychology and Scientific Methods* 1907; **4**(9):225–235. doi:10.2307/2010813.
3. T. Nagel. What is it like to be a bat?. *The Philosophical Review* 1974; **83**(4):435–450. JSTOR:2183914. doi:10.2307/2183914.
4. D. D. Georgiev. Quantum Information and Consciousness: A Gentle Introduction. CRC Press, Boca Raton, 2017. doi:10.1201/9780203732519.
5. C. Peterson, N. Park, M. E. P. Seligman. Orientations to happiness and life satisfaction: the full life versus the empty life. *Journal of Happiness Studies* 2005; **6**(1):25–41. doi:10.1007/s10902-004-1278-z.
6. U. Rudolph, B. Antkowiak. Molecular and neuronal substrates for general anaesthetics. *Nature Reviews Neuroscience* 2004; **5**(9):709–720. doi:10.1038/nrn1496.
7. N. P. Franks. General anaesthesia: from molecular targets to neuronal pathways of sleep and arousal. *Nature Reviews Neuroscience* 2008; **9**(5):370–386. doi:10.1038/nrn2372.
8. E. N. Brown, P. L. Purdon, C. J. Van Dort. General anesthesia and altered states of arousal: a systems neuroscience analysis. *Annual Review of Neuroscience* 2011; **34**(1):601–628. doi:10.1146/annurev-neuro-060909-153200.
9. W. N. Kapoor. Syncope. *New England Journal of Medicine* 2000; **343**(25):1856–1862. doi:10.1056/nejm200012213432507.
10. E. N. Brown, R. Lydic, N. D. Schiff. General anesthesia, sleep, and coma. *New England Journal of Medicine* 2010; **363**(27):2638–2650. doi:10.1056/nejmra0808281.
11. I. Fagelson. The last bastion of fault? Contributory negligence in actions for employers' liability. *The Modern Law Review* 1979; **42**(6):646–666. doi:10.1111/j.1468-2230.1979.tb01558.x.
12. A. J. E. Jaffey. Volenti non fit injuria. *The Cambridge Law Journal* 1985; **44**(1):87–110. doi:10.1017/s000819730011445x.
13. C. A. Hopkins. Causation and contributory negligence. *The Cambridge Law Journal* 2005; **64**(3):546–548. doi:10.1017/s0008197305296952.

14. J.-P. Sartre. Existentialism Is a Humanism. Yale University Press, New Haven, Connecticut, 2007.
15. R. Shafer-Landau. The Fundamentals of Ethics. 5th Edition. Oxford University Press, Oxford, 2020.
16. R. P. Feynman, R. B. Leighton, M. L. Sands. The Feynman Lectures on Physics. California Institute of Technology, Pasadena, California, 2013. `http://www.feynmanlectures.caltech.edu/`.
17. D. D. Georgiev. Inner privacy of conscious experiences and quantum information. *Biosystems* 2020; **187**:104051. `doi:10.1016/j.biosystems.2019.104051`.
18. D. D. Georgiev. Quantum information theoretic approach to the mind–brain problem. *Progress in Biophysics and Molecular Biology* 2020; **158**:16–32. `doi:10.1016/j.pbiomolbio.2020.08.002`.
19. K. Imanishi. Prehuman societies. Iwanami Shoten, Tokyo, 1951.
20. K. Imanishi. Human. Mainichi Press, Tokyo, 1952.
21. K. Imanishi. Learned behavior of Japanese monkeys. *Japanese Journal of Ethnology* 1957; **21**(3):185–189. `doi:10.14890/minkennewseries.21.3_185`.
22. K. Imanishi. Formation of human society. Vol. 40 of *NHK Books*. Japan Broadcast Publishing Association, Tokyo, 1966.
23. T. Matsuzawa. Primate Origins of Human Cognition and Behavior. Springer, Hong Kong, 2001. `doi:10.1007/978-4-431-09423-4`.
24. T. Gruber, K. Zuberbühler. Evolutionary origins of the human cultural mind. *The Psychologist* 2012; **25**(5):364–369.
25. J. Goodall. The Chimpanzees of Gombe: Patterns of Behavior. Bellknap Press, Boston, 1986.
26. A. Whiten, J. Goodall, W. C. McGrew, T. Nishida, V. Reynolds, Y. Sugiyama, C. E. G. Tutin, R. W. Wrangham, C. Boesch. Cultures in chimpanzees. *Nature* 1999; **399**(6737):682–685. `doi:10.1038/21415`.
27. T. Gruber, M. N. Muller, P. Strimling, R. Wrangham, K. Zuberbühler. Wild chimpanzees rely on cultural knowledge to solve an experimental honey acquisition task. *Current Biology* 2009; **19**(21):1806–1810. `doi:10.1016/j.cub.2009.08.060`.
28. J. Lind, P. Lindenfors. The number of cultural traits is correlated with female group size but not with male group size in chimpanzee communities. *PLOS ONE* 2010; **5**(3):e9241. `doi:10.1371/journal.pone.0009241`.
29. C. Boesch, A. K. Kalan, R. Mundry, M. Arandjelovic, S. Pika, P. Dieguez, E. A. Ayimisin, A. Barciela, C. Coupland, V. E. Egbe, M. Eno-Nku, J. Michael Fay, D. Fine, R. Adriana Hernandez-Aguilar, V. Hermans, P. Kadam, M. Kambi, M. Llana, G. Maretti, D. Morgan, M. Murai, E. Neil, S. Nicholl, L. J. Ormsby, R. Orume, L. Pacheco, A. Piel, C. Sanz, L. Sciaky, F. A. Stewart, N. Tagg, E. G. Wessling, J. Willie, H. S. Kühl. Chimpanzee ethnography reveals unexpected cultural diversity. *Nature Human Behaviour* 2020; **4**(9):910–916. `doi:10.1038/s41562-020-0890-1`.
30. L. Samuni, F. Wegdell, M. Surbeck. Behavioural diversity of bonobo prey preference as a potential cultural trait. *eLife* 2020; **9**:e59191. `doi:10.7554/eLife.59191`.
31. E. J. C. van Leeuwen, N. Staes, J. Verspeek, W. J. E. Hoppitt, J. M. G. Stevens. Social culture in bonobos. *Current Biology* 2020; **30**(6):R261–R262. `doi:10.1016/j.cub.2020.02.038`.
32. M. M. Robbins, C. Ando, K. A. Fawcett, C. C. Grueter, D. Hedwig, Y. Iwata, J. L. Lodwick, S. Masi, R. Salmi, T. S. Stoinski, A. Todd, V. Vercellio, J. Yamagiwa. Behavioral variation in gorillas: evidence of potential cultural traits. *PLOS ONE* 2016; **11**(9):e0160483. `doi:10.1371/journal.pone.0160483`.

33. M. Krützen, E. P. Willems, C. P. van Schaik. Culture and geographic variation in orangutan behavior. *Current Biology* 2011; **21**(21):1808–1812. doi:10.1016/j.cub.2011.09.017.
34. T. Gruber, I. Singleton, C. van Schaik. Sumatran orangutans differ in their cultural knowledge but not in their cognitive abilities. *Current Biology* 2012; **22**(23):2231–2235. doi:10.1016/j.cub.2012.09.041.
35. K. Koops, E. Visalberghi, C. P. van Schaik. The ecology of primate material culture. *Biology Letters* 2014; **10**(11):20140508. doi:10.1098/rsbl.2014.0508.
36. A. R. Lameira, G. Santamaría-Bonfil, D. Galeone, M. Gamba, M. E. Hardus, C. D. Knott, H. Morrogh-Bernard, M. G. Nowak, G. Campbell-Smith, S. A. Wich. Sociality predicts orangutan vocal phenotype. *Nature Ecology & Evolution* 2022; **6**(5):644–652. doi:10.1038/s41559-022-01689-z.
37. S. Kawamura. The process of sub-culture propagation among Japanese macaques. *Primates* 1959; **2**(1):43–60. doi:10.1007/BF01666110.
38. M. Kawai. Newly-acquired pre-cultural behavior of the natural troop of Japanese monkeys on Koshima islet. *Primates* 1965; **6**(1):1–30. doi:10.1007/bf01794457.
39. J.-B. Leca, M. A. Huffman, P. L. Vasey. The Monkeys of Stormy Mountain: 60 Years of Primatological Research on the Japanese Macaques of Arashiyama. Vol. 61 of *Cambridge Studies in Biological and Evolutionary Anthropology*. Cambridge University Press, Cambridge, 2012.
40. M. Gross. Cave art reveals human nature. *Current Biology* 2020; **30**(3):R95–R98. doi:10.1016/j.cub.2020.01.042.
41. B. Ehrenreich. The humanoid stain: Art lessons from our cave-dwelling ancestors. *The Baffler* 2019; **48**:6–14. https://thebaffler.com/salvos/the-humanoid-stain-ehrenreich. JSTOR:26842273.
42. N. J. Wade, D. Melcher. Cave art interpretation I. *Perception* 2006; **35**(5):577–580. doi:10.1068/p3505ed.
43. D. Melcher, N. J. Wade. Cave art interpretation II. *Perception* 2006; **35**(6):719–722. doi:10.1068/p3506ed.
44. J. C. Harris. Chauvet cave: the panel of horses. *Archives of General Psychiatry* 2011; **68**(9):869–870. doi:10.1001/archgenpsychiatry.2011.114.
45. W. Herzog. Cave of Forgotten Dreams. Creative Differences Productions, Inc., Los Angeles, 2010. https://www.imdb.com/title/tt1664894/.
46. N. J. Conard. A female figurine from the basal Aurignacian of Hohle Fels Cave in southwestern Germany. *Nature* 2009; **459**(7244):248–252. doi:10.1038/nature07995.
47. N. J. Conard, M. Malina, S. C. Münzel. New flutes document the earliest musical tradition in southwestern Germany. *Nature* 2009; **460**(7256):737–740. doi:10.1038/nature08169.
48. H. Valladas, J. Clottes, J. M. Geneste, M. A. Garcia, M. Arnold, H. Cachier, N. Tisnérat-Laborde. Palaeolithic paintings: evolution of prehistoric cave art. *Nature* 2001; **413**(6855):479–479. doi:10.1038/35097160.
49. K. R. Popper, J. C. Eccles. The Self and Its Brain: An Argument for Interactionism. Routledge & Kegan Paul, London, 1983. doi:10.4324/9780203537480.
50. J. C. Eccles. Evolution of the Brain: Creation of the Self. Routledge, London, 1991. doi:10.4324/9780203976661.
51. J. C. Eccles. Evolution of consciousness. *Proceedings of the National Academy of Sciences* 1992; **89**(16):7320–7324. doi:10.1073/pnas.89.16.7320.
52. J. C. Eccles. How the Self Controls Its Brain. Springer, Berlin, 1994. doi:10.1007/978-3-642-49224-2.

53. J. E. Hall, M. E. Hall. Guyton and Hall Textbook of Medical Physiology. 14th Edition. Elsevier, Amsterdam, 2021.
54. E. R. Kandel, J. D. Koester, S. H. Mack, S. A. Siegelbaum. Principles of Neural Science. 6th Edition. McGraw Hill, New York, 2021.
55. A. H. Ropper, M. A. Samuels, J. Klein, S. Prasad. Adams and Victor's Principles of Neurology. 11th Edition. McGraw-Hill, New York, 2019.
56. J. M. Cattell. The time taken up by cerebral operations. Part I & II. *Mind* 1886; **11**(42):220–242. doi:10.1093/mind/os-XI.42.220.
57. J. M. Cattell. The time taken up by cerebral operations. Part III. *Mind* 1886; **11**(43):377–392. doi:10.1093/mind/os-XI.43.377.
58. J. M. Cattell. The time taken up by cerebral operations. Part IV. *Mind* 1887; **11**(44):524–538. doi:10.1093/mind/os-XI.44.524.
59. W. H. Dobelle. Artificial vision for the blind by connecting a television camera to the visual cortex. *ASAIO Journal* 2000; **46**(1):3–9. doi:10.1097/00002480-200001000-00002.
60. L. R. Hochberg, M. D. Serruya, G. M. Friehs, J. A. Mukand, M. Saleh, A. H. Caplan, A. Branner, D. Chen, R. D. Penn, J. P. Donoghue. Neuronal ensemble control of prosthetic devices by a human with tetraplegia. *Nature* 2006; **442**(7099):164–171. doi:10.1038/nature04970.
61. D. D. Georgiev, S. K. Kolev, E. Cohen, J. F. Glazebrook. Computational capacity of pyramidal neurons in the cerebral cortex. *Brain Research* 2020; **1748**:147069. doi:10.1016/j.brainres.2020.147069.
62. R. Nieuwenhuys. The neocortex: an overview of its evolutionary development, structural organization and synaptology. *Anatomy and Embryology* 1994; **190**(4):307–337. doi:10.1007/bf00187291.
63. J. H. O. Hoffmann, H. S. Meyer, A. C. Schmitt, J. Straehle, T. Weitbrecht, B. Sakmann, M. Helmstaedter. Synaptic conductance estimates of the connection between local inhibitor interneurons and pyramidal neurons in layer 2/3 of a cortical column. *Cerebral Cortex* 2015; **25**(11):4415–4429. doi:10.1093/cercor/bhv039.
64. C. A. Anastassiou, R. Perin, G. Buzsáki, H. Markram, C. Koch. Cell type- and activity-dependent extracellular correlates of intracellular spiking. *Journal of Neurophysiology* 2015; **114**(1):608–623. doi:10.1152/jn.00628.2014.
65. M. Marx, D. Feldmeyer. Morphology and physiology of excitatory neurons in layer 6b of the somatosensory rat barrel cortex. *Cerebral Cortex* 2012; **23**(12):2803–2817. doi:10.1093/cercor/bhs254.
66. M. Marx, G. Qi, I. L. Hanganu-Opatz, W. Kilb, H. J. Luhmann, D. Feldmeyer. Neocortical layer 6b as a remnant of the subplate - a morphological comparison. *Cerebral Cortex* 2015; **27**(2):1011–1026. doi:10.1093/cercor/bhv279.
67. J. M. Bekkers. Pyramidal neurons. *Current Biology* 2011; **21**(24):R975. doi:10.1016/j.cub.2011.10.037.
68. G. A. Ascoli. Mobilizing the base of neuroscience data: the case of neuronal morphologies. *Nature Reviews Neuroscience* 2006; **7**(4):318–324. doi:10.1038/nrn1885.
69. A. Karimi, J. Odenthal, F. Drawitsch, K. M. Boergens, M. Helmstaedter. Cell-type specific innervation of cortical pyramidal cells at their apical dendrites. *eLife* 2020; **9**:e46876. doi:10.7554/eLife.46876.
70. C. Stadelmann, S. Timmler, A. Barrantes-Freer, M. Simons. Myelin in the central nervous system: structure, function, and pathology. *Physiological Reviews* 2019; **99**(3):1381–1431. doi:10.1152/physrev.00031.2018.

71. Y. Long, J. Li, F. Yang, J. Wang, X. Wang. Wearable and implantable electroceuticals for therapeutic electrostimulations. *Advanced Science* 2021; **8**(8):2004023. `doi:10.1002/advs.202004023`.
72. E. J. Hwang, R. A. Andersen. Brain control of movement execution onset using local field potentials in posterior parietal cortex. *Journal of Neuroscience* 2009; **29**(45):14363. `doi:10.1523/jneurosci.2081-09.2009`.
73. D. D. Georgiev, I. Georgieva, Z. Gong, V. Nanjappan, G. V. Georgiev. Virtual reality for neurorehabilitation and cognitive enhancement. *Brain Sciences* 2021; **11**(2):221. `doi:10.3390/brainsci11020221`.
74. J. C. Sprott. Simplifications of the Lorenz attractor. *Nonlinear Dynamics, Psychology, and Life Sciences* 2009; **13**(3):271–278.
75. W. James. Are we automata?. *Mind* 1879; **4**(13):1–22. `doi:10.1093/mind/os-4.13.1`.
76. A. Zupancich, S. Nunziante-Cesaro, R. Blasco, J. Rosell, E. Cristiani, F. Venditti, C. Lemorini, R. Barkai, A. Gopher. Early evidence of stone tool use in bone working activities at Qesem Cave, Israel. *Scientific Reports* 2016; **6**(1):37686. `doi:10.1038/srep37686`.
77. M. C. Burkitt. Prehistoric art in caves and rock-shelters. *Nature* 1921; **107**(2693):460–464. `doi:10.1038/107460a0`.
78. J.-M. Chauvet, E. B. Deschamps, C. Hillaire. Dawn of Art: The Chauvet Cave. Harry N. Abrams, New York, 1996.
79. J.-M. Chauvet, E. B. Deschamps. Chauvet Cave: The Discovery of the World's Oldest Paintings. Thames & Hudson, London, 2001.
80. M. Aubert, A. Brumm, M. Ramli, T. Sutikna, E. W. Saptomo, B. Hakim, M. J. Morwood, G. D. van den Bergh, L. Kinsley, A. Dosseto. Pleistocene cave art from Sulawesi, Indonesia. *Nature* 2014; **514**(7521):223–227. `doi:10.1038/nature13422`.
81. D. L. Hoffmann, C. D. Standish, M. García-Diez, P. B. Pettitt, J. A. Milton, J. Zilhão, J. J. Alcolea-González, P. Cantalejo-Duarte, H. Collado, R. de Balbín, M. Lorblanchet, J. Ramos-Muñoz, G. C. Weniger, A. W. G. Pike. U-Th dating of carbonate crusts reveals Neandertal origin of Iberian cave art. *Science* 2018; **359**(6378):912–915. `doi:10.1126/science.aap7778`.
82. A. W. G. Pike, D. L. Hoffmann, M. García-Diez, P. B. Pettitt, J. Alcolea, R. De Balbín, C. González-Sainz, C. de las Heras, J. A. Lasheras, R. Montes, J. Zilhão. U-series dating of paleolithic art in 11 caves in Spain. *Science* 2012; **336**(6087):1409–1413. `doi:10.1126/science.1219957`.
83. J. D. Lewis-Williams. The Mind in the Cave: Consciousness and the Origins of Art. Thames & Hudson, London, 2002.
84. J. D. Lewis-Williams. Chauvet: the cave that changed expectations. *South African Journal of Science* 2003; **99**(5):191–194. `doi:10520/ejc97651`.
85. J. Kim. What is the problem of mental causation?. in: M. L. Dalla Chiara, K. Doets, D. Mundici, J. Van Benthem (Eds.), Structures and Norms in Science: Volume Two of the Tenth International Congress of Logic, Methodology and Philosophy of Science, Florence, August 1995. Vol. 260 of *Synthese Library*. Kluwer, Dordrecht, 1997. pp. 319–329. `doi:10.1007/978-94-017-0538-7_19`.
86. J. Kim. Mind in a Physical World: An Essay on the Mind-Body Problem and Mental Causation. MIT Press, Cambridge, Massachusetts, 1998. `doi:10.7551/mitpress/4629.001.0001`.
87. J. Kim. The nonreductivist's troubles with mental causation. in: J. Heil, A. Mele (Eds.), Mental Causation. Oxford University Press, Oxford, 2003. pp. 189–210.
88. G. Strang. Calculus. Wellesley Cambridge Press, Wellesley, Massachusetts, 1991.

89. J. S. O'Brien, E. L. Sampson. Lipid composition of the normal human brain: gray matter, white matter, and myelin. *Journal of Lipid Research* 1965; **6**(4):537–544. doi:10.1016/S0022-2275(20)39619-X.
90. M. Winick. Changes in nucleic acid and protein content of the human brain during growth. *Pediatric Research* 1968; **2**(5):352–355. doi:10.1203/00006450-196809000-00003.
91. R. L. Friede, K. H. Hu. A new approach for determining the volume of cerebral extracellular fluid and demonstration of its communication with c.s.f.. *Journal of Physiology* 1971; **218**(2):477–493. doi:10.1113/jphysiol.1971.sp009629.
92. A. Korzybski. Science and Sanity: An Introduction to Non-Aristotelian Systems and General Semantics. 5th Edition. Institute of General Semantics, Brooklyn, New York, 1994.
93. P. A. M. Dirac. The Principles of Quantum Mechanics. 4th Edition. Oxford University Press, Oxford, 1967.
94. L. Susskind, A. Friedman. Quantum Mechanics: The Theoretical Minimum. What You Need to Know to Start Doing Physics. Basic Books, New York, 2014.
95. M. L. Bhaumik. How Dirac's seminal contributions pave the way for comprehending nature's deeper designs. *Quanta* 2019; **8**:88–100. doi:10.12743/quanta.v8i1.96.
96. D. D. Georgiev, S. P. Gudder. Sensitivity of entanglement measures in bipartite pure quantum states. *Modern Physics Letters B* 2022; **36**(22):2250101. doi:10.1142/S0217984922501019.
97. D. Georgiev, E. Cohen. Entanglement measures for two-particle quantum histories. *Physical Review A* 2022; **106**(6):062437. doi:10.1103/PhysRevA.106.062437.
98. J. von Neumann. Mathematische Grundlagen der Quantenmechanik. Springer, Berlin, 1932. doi:10.1007/978-3-642-61409-5.
99. J. von Neumann. Mathematical Foundations of Quantum Mechanics. Vol. 53 of *Princeton Landmarks in Mathematics and Physics*. Princeton University Press, Princeton, New Jersey, 1955. doi:10.1515/9781400889921.
100. M. L. Bhaumik. Is the quantum state real in the Hilbert space formulation?. *Quanta* 2020; **9**:37–46. doi:10.12743/quanta.v9i1.142.
101. M. Born. Statistical interpretation of quantum mechanics. *Science* 1955; **122**(3172):675–679. doi:10.1126/science.122.3172.675.
102. L. Vaidman. Derivations of the Born rule. in: M. Hemmo, O. Shenker (Eds.), Quantum, Probability, Logic: The Work and Influence of Itamar Pitowsky. Springer, Cham, Switzerland, 2020. pp. 567–584. doi:10.1007/978-3-030-34316-3_26.
103. S. P. Gudder. Stochastic Methods in Quantum Mechanics. North Holland, New York, 1979.
104. K. Itô. Stochastic integral. *Proceedings of the Imperial Academy* 1944; **20**(8):519–524. doi:10.3792/pia/1195572786.
105. K. Itô. Stochastic differential equations in a differentiable manifold. *Nagoya Mathematical Journal* 1950; **1**:35–47. doi:10.1017/S0027763000022819.
106. K. Itô. Stochastic differentials. *Applied Mathematics and Optimization* 1975; **1**(4):374–381. doi:10.1007/bf01447959.
107. K. Itô. Foundations of Stochastic Differential Equations in Infinite Dimensional Spaces. Vol. 47 of *CBMS-NSF Regional Conference Series in Applied Mathematics*. Society for Industrial and Applied Mathematics, Philadelphia, 1984. doi:10.1137/1.9781611970234.
108. N. Ikeda, S. Watanabe, M. Fukushima, H. Kunita. Itô's Stochastic Calculus and Probability Theory. Springer, Tokyo, 1996. doi:10.1007/978-4-431-68532-6.

109. D. D. Georgiev, J. F. Glazebrook. On the quantum dynamics of Davydov solitons in protein α-helices. *Physica A: Statistical Mechanics and its Applications* 2019; **517**:257–269. doi:10.1016/j.physa.2018.11.026.
110. D. D. Georgiev, J. F. Glazebrook. Quantum transport and utilization of free energy in protein α-helices. *Advances in Quantum Chemistry* 2020; **82**:253–300. doi:10.1016/bs.aiq.2020.02.001.
111. D. D. Georgiev, J. F. Glazebrook. Thermal stability of solitons in protein α-helices. *Chaos, Solitons and Fractals* 2022; **155**:111644. doi:10.1016/j.chaos.2021.111644.
112. D. Huff. How to Lie with Statistics. W. W. Norton & Company, New York, 1993.
113. S. P. Gudder. Quantum Probability. Academic Press, San Diego, California, 1988.
114. P. Busch, P. J. Lahti, P. Mittelstaedt. The Quantum Theory of Measurement. Vol. 2 of *Lecture Notes in Physics Monographs*. Springer, Berlin, 1996. doi:10.1007/978-3-540-37205-9.
115. D. D. Georgiev. Quantum propensities in the brain cortex and free will. *BioSystems* 2021; **208**:104474. doi:10.1016/j.biosystems.2021.104474.
116. G. Ghirardi, A. Rimini, T. Weber. Unified dynamics for microscopic and macroscopic systems. *Physical Review D* 1986; **34**(2):470–491. doi:10.1103/PhysRevD.34.470.
117. G. Ghirardi, R. Grassi, P. Pearle. Relativistic dynamical reduction models: General framework and examples. *Foundations of Physics* 1990; **20**(11):1271–1316. doi:10.1007/bf01883487.
118. A. Bassi, G. Ghirardi. Dynamical reduction models. *Physics Reports* 2003; **379**(5):257–426. doi:10.1016/S0370-1573(03)00103-0.
119. P. A. M. Dirac. A new notation for quantum mechanics. *Mathematical Proceedings of the Cambridge Philosophical Society* 1939; **35**(3):416–418. doi:10.1017/S0305004100021162.
120. A. V. Melkikh. Thinking as a quantum phenomenon. *Biosystems* 2019; **176**:32–40. doi:10.1016/j.biosystems.2018.12.007.
121. P. Busch. Is the quantum state (an) observable?. in: R. S. Cohen, M. Horne, J. Stachel (Eds.), Potentiality, Entanglement and Passion-at-a-Distance: Quantum Mechanical Studies for Abner Shimony, Volume Two. Vol. 194 of *Boston Studies in the Philosophy of Science*. Kluwer, Dordrecht, 1997. pp. 61–70. doi:10.1007/978-94-017-2732-7_5.
122. J. DeFelipe, L. Alonso-Nanclares, J. I. Arellano. Microstructure of the neocortex: comparative aspects. *Journal of Neurocytology* 2002; **31**(3):299–316. doi:10.1023/a:1024130211265.
123. A. S. Holevo. Bounds for the quantity of information transmitted by a quantum communication channel. *Problems of Information Transmission* 1973; **9**(3):177–183.
124. D. D. Georgiev. Quantum information in neural systems. *Symmetry* 2021; **13**(5):773. doi:10.3390/sym13050773.
125. A. M. Kariev, P. Njau, M. E. Green. The open gate of the Kv1.2 channel: quantum calculations show the key role of hydration. *Biophysical Journal* 2014; **106**(3):548–555. doi:10.1016/j.bpj.2013.11.4495.
126. A. M. Kariev, M. E. Green. The role of proton transport in gating current in a voltage gated ion channel, as shown by quantum calculations. *Sensors* 2018; **18**(9):3143. doi:10.3390/s18093143.
127. A. M. Kariev, M. E. Green. Quantum calculations on ion channels: why are they more useful than classical calculations, and for which processes are they essential?. *Symmetry* 2021; **13**(4):655. doi:10.3390/sym13040655.

128. M. S. P. Sansom. Potassium channels: Watching a voltage-sensor tilt and twist. *Current Biology* 2000; **10**(5):R206–R209. doi:10.1016/S0960-9822(00)00354-7.
129. W. A. Catterall. Ion channel voltage sensors: structure, function, and pathophysiology. *Neuron* 2010; **67**(6):915–928. doi:10.1016/j.neuron.2010.08.021.
130. L. D. Islas. Functional diversity of potassium channel voltage-sensing domains. *Channels* 2016; **10**(3):202–213. doi:10.1080/19336950.2016.1141842.
131. S. Bamford. A framework for approaches to transfer of a mind's substrate. *International Journal of Machine Consciousness* 2012; **4**(1):23–34. doi:10.1142/s1793843012400021.
132. Y. Choe, J. Kwon, J. R. Chung. Time, consciousness, and mind uploading. *International Journal of Machine Consciousness* 2012; **4**(1):257–274. doi:10.1142/S179384301240015X.
133. W. Pfister, J. Paglen, J. Depp, M. Freeman, R. Hall, K. Mara, C. Murphy, C. Hauser, P. Bettany. Transcendence. Warner Bros. Pictures, Burbank, California, 2014. https://www.imdb.com/title/tt2209764/.
134. M. L. Cappuccio. Mind-upload. The ultimate challenge to the embodied mind theory. *Phenomenology and the Cognitive Sciences* 2017; **16**(3):425–448. doi:10.1007/s11097-016-9464-0.
135. P. Langevin. Sur la théorie du mouvement brownien. *Comptes Rendus de l'Académie des Sciences* 1908; **146**:530–533. https://gallica.bnf.fr/ark:/12148/bpt6k3100t/f3.item.
136. C. Gabriel, C. Wittmann, D. Sych, R. Dong, W. Mauerer, U. L. Andersen, C. Marquardt, G. Leuchs. A generator for unique quantum random numbers based on vacuum states. *Nature Photonics* 2010; **4**(10):711–715. doi:10.1038/nphoton.2010.197.
137. Y. Zhang, H.-P. Lo, A. Mink, T. Ikuta, T. Honjo, H. Takesue, W. J. Munro. A simple low-latency real-time certifiable quantum random number generator. *Nature Communications* 2021; **12**(1):1056. doi:10.1038/s41467-021-21069-8.
138. V. B. Braginsky, F. Y. Khalili. Quantum Measurement. Cambridge University Press, Cambridge, 1992. doi:10.1017/cbo9780511622748.
139. D. Johnston, S. M.-S. Wu. Foundations of Cellular Neurophysiology. Bradford Books. MIT Press, Cambridge, Massachusetts, 1995.
140. S. P. Gudder. Dual instruments and sequential products of observables. *Quanta* 2022; **11**:15–27. doi:10.12743/quanta.v11i1.197.
141. P. Busch, G. Cassinelli, P. J. Lahti. On the quantum theory of sequential measurements. *Foundations of Physics* 1990; **20**(7):757–775. doi:10.1007/BF01889690.
142. S. P. Gudder. Quantum entanglement: spooky action at a distance. *Quanta* 2020; **9**:1–6. doi:10.12743/quanta.v9i1.113.
143. S. P. Gudder. A theory of entanglement. *Quanta* 2020; **9**:7–15. doi:10.12743/quanta.v9i1.115.
144. A. Vinante, M. Carlesso, A. Bassi, A. Chiasera, S. Varas, P. Falferi, B. Margesin, R. Mezzena, H. Ulbricht. Narrowing the parameter space of collapse models with ultracold layered force sensors. *Physical Review Letters* 2020; **125**(10):100404. doi:10.1103/PhysRevLett.125.100404.
145. M. Carlesso, S. Donadi, L. Ferialdi, M. Paternostro, H. Ulbricht, A. Bassi. Present status and future challenges of non-interferometric tests of collapse models. *Nature Physics* 2022; **18**(3):243–250. doi:10.1038/s41567-021-01489-5.
146. E. Schrödinger. Die gegenwärtige Situation in der Quantenmechanik. I. *Naturwissenschaften* 1935; **23**(48):807–812. doi:10.1007/bf01491891.

147. E. Schrödinger. The present situation in quantum mechanics. in: J. A. Wheeler, W. H. Zurek (Eds.), Quantum Theory and Measurement. Princeton Series in Physics. Princeton University Press, Princeton, New Jersey, 1983. pp. 152–167.
148. B. Kurtén. The cave bear. *Scientific American* 1972; **226**(3):60–73. doi:10.1038/scientificamerican0372-60.
149. C. Bon, V. Berthonaud, P. Fosse, B. Gély, F. Maksud, R. Vitalis, M. Philippe, J. van der Plicht, J.-M. Elalouf. Low regional diversity of late cave bears mitochondrial DNA at the time of Chauvet Aurignacian paintings. *Journal of Archaeological Science* 2011; **38**(8):1886–1895. doi:10.1016/j.jas.2011.03.033.
150. C. Sartorio. Causation and Free Will. Oxford University Press, Oxford, 2016. doi:10.1093/acprof:oso/9780198746799.001.0001.