

Consciousness as the collective excitation of a brainwide web – understanding consciousness from below quantum fields to above neuronal networks

Suzy Lidström and Roland E. Allen

Department of Physics and Astronomy, Texas A&M University, College Station, Texas
77843-4242

E-mail: suzy.lidstrom@gmail.com, allen@tamu.edu

Abstract. In the spirit of previous ideas in the neuroscience community, but with a more physics-oriented perspective, we propose that consciousness can be described as the collective excitation of a *brainwide web* of neurones. This picture is inspired by the fact that, in all major areas of physics, a collective excitation has just as much physical reality as a particle or other localized object. The brainwide web extends into those regions (neuronal networks) where processed information is received from the senses, memories, etc. (emerging out of unconscious processes in prior networks). It unifies those regions (plus motor control regions) via the vast complexity of the neuronal interactions that it spans. A crude analogy is the worldwide web, which extends into servers which have been prepared by external agents. Other crude analogies are the many collective excitations in physics. True understanding of consciousness must rely mostly on experiment, but since probes of living brains have limited precision, there is an important role for large-scale simulations in revealing details and complementing experiment. Petascale computational facilities should make it possible to perform simulations with realistic complexity.

In discussing consciousness, we need a vocabulary that prevents confusion of different concepts. Continuing with a notation introduced elsewhere [1], we use the subscripts O and P to refer respectively to the phenomena of nature accessible to an external *observer* and the inner reality of nature experienced by a *participant*, with direct conscious experiences being only one example. The “hard problem” of David Chalmers [2] is the problem of explaining consciousness_P: Why do we have the specific inner experiences of seeing red, feeling pain, etc.? As argued elsewhere [1], it is not clear that this problem has meaning for observers or participants embedded within nature. Like the question “Why is there something rather than nothing?” [3], it appears either to be meaningless as an answerable problem or else to require a deeper level of understanding than we can now imagine.

Here we focus on understanding consciousness_O, which can potentially be addressed at various levels:

1. Quantum fields, which are the ultimate reality of present-day physics. Presumably biological consciousness primarily involves the electron and photon fields, since the quark, gluon, ... fields are perturbed mostly at inappropriate distance and energy scales. However, nuclear motion is of central importance in processes like the release of neurotransmitters, and



nuclear spins are of central importance in experimental techniques like functional magnetic resonance imaging (fMRI).

2. Particles, which are excitations of quantum fields (analogous to the quantized excitations of a harmonic oscillator). In a quantum description, one can form collective excitations that involve only one species of particle (e.g. plasma oscillations) or several species, as in polaritons (electron-photon hybrid modes) or polarons (electron-phonon hybrid modes, with phonons being the quanta of nuclear motion). A central point is that a collective excitation has just as much physical reality as a particle or other localized object. If one reverts to level 1 above, excitations of quantum fields can be defined in different ways according to the physical situation. For example, an observer far from a black hole and one who is nearby will define (level 2) particles differently, using different vacua and particle detectors, even though the (level 1) quantum fields in curved spacetime are well-defined. This is the origin of the Hawking effect [4]. Similarly, one can choose different descriptions of quantum excitations in material systems – for example, by performing a canonical transformation from atomic to normal mode coordinates for the vibrations of a molecule or material. In this case, a phonon (quantum of vibration) has just as much physical reality as an atom.
3. Atoms, and then molecules (with a macromolecule often consisting of an enormous number of atoms). At this level one still needs quantum mechanics, but the most nuanced features of quantum theory, such as entanglement, are very unlikely to be relevant, because a normal body temperature implies extremely short decoherence times [5]. Thermal agitation also implies that the states discussed below are to be interpreted as thermal averages over the microscopic states accessible at thermal energies.
4. The biochemistry and biophysics of synapses, ion channels, etc. At this point we are already into an area of biological science which is rich in potential for important discoveries. (Of course, this is also true for some aspects of the preceding level, such as protein structure and protein folding.)
5. The architecture of neuronal networks. In the brain there are many kinds of neurones, with axons of various lengths, thousands of dendrites per neurone, and differences in other attributes, comprising networks of astounding intricacy. In the formal model proposed here, there is one well-defined global network – the *brainwide web* discussed below – which can be identified as the physical system in which consciousness is located.
6. The time-dependent states of these networks, which are arrayed in a hierarchy, again of astounding complexity. These states must be viewed as collective modes of neurones. At the top of this hierarchy is the brainwide web proposed here, which includes those networks where final representations are formed of what is perceived, remembered, ... The networks which feed information into these components in the visual cortex etc. are not themselves part of the brainwide web (i.e. the conscious brain), so all their enormously sophisticated processing is unconscious.

Since physics has come so far during the past 120 years, it is quite plausible that new physics may lie ahead, with yet another level:

0. A deeper understanding which explains quantum fields, and which may even render the “hard problem” meaningful.

However, level 0 physics will not invalidate quantum field theory, which has been demonstrated to describe the world with remarkable success and accuracy. Even gravity must ultimately be consistent with quantum theory, as can be demonstrated by a thought experiment

involving a “Schrödinger planet”: Imagine that the trajectory of a planet can be perturbed by a radioactive decay or other quantum event or measurement. The gravitational field must follow the planet, so one has two outcomes for the trajectory of the planet’s gravitational field in the same way that one has two outcomes for the state of Schrödinger’s cat. In the path integral formulation of quantum mechanics, the gravitational field, as well as the matter fields and other force fields, must follow all allowed trajectories.

Since consciousness is one aspect of physical reality, and physical reality is described with great success by quantum fields, we conclude that consciousness must be expressible in terms of quantum fields. In other words, we conclude that consciousness can be regarded as a state $|\Psi(t)\rangle$ of these fields. (Again, this state must properly be regarded as a nearly classical thermal average over specific quantum states, and interference phenomena for specific states are precluded by rapid decoherence.) This is the most fundamental physics perspective currently available.

Here we are using the Dirac bra-ket notation [6], with a general state represented by a ket like $|\Psi\rangle$, and the amplitude for this state being in a specific state like $|\psi\rangle$ given by its projection onto this state (the bra-ket), $\langle\psi|\Psi\rangle$. An example is the wavefunction $\phi(\vec{r}, t) = \langle\vec{r}|\phi(t)\rangle$ for a particle in a state $|\phi(t)\rangle$, where $|\vec{r}\rangle$ is the state of a particle located exactly at position $\vec{r}$.

There are two aspects of $|\Psi(t)\rangle$ that are immediately revealed just by introspection: It has boundaries, which means it must be localized in a *coherent domain of fields* within nature. And it spans a very large number of degrees of freedom x , with time-dependent amplitudes

$$\Psi(x, t) = \langle x | \Psi(t) \rangle . \quad (1)$$

The content of consciousness is constantly shifting from one centre of attention to another – from one object to another in the visual field, from one memory to another, from one sound to another, ..., or equally often from a visual object to a sound or a thought, ... This is the time dependence of the state $|\Psi(t)\rangle$. Let $|X\rangle$ represent some set of degrees of freedom – for example a perceived face or a remembered melody. The amplitude for $|X\rangle$ being in the consciousness at time t is $\langle X | \Psi(t) \rangle$. I.e., this is the degree to which one is conscious of $|X\rangle$.

From a biology perspective [7–15] – based on a vast number of brilliant experimental studies which it is impossible to cite here – it is clear that biological consciousness must involve *neuronal networks* in specific ways. The coherence of consciousness implies that it must be localized in a *coherent domain of neurones*, which we call the brainwide web. A crude analogy is the worldwide web, which extends into servers which have been prepared by external agents, in the same way that the brainwide web extends into regions that have been prepared by unconscious processes in other networks lower in the hierarchy. With the brainwide web regarded as a well-defined (but open) system, its time-dependent state $|\Psi(t)\rangle$ can be represented by whatever neuronal states $|\psi_n(t)\rangle$ are most appropriate for experimental study or simulation:

$$|\Psi(t)\rangle = \sum_n a_n(t) |\psi_n(t)\rangle . \quad (2)$$

If the $|\psi_n(t)\rangle$ are chosen to be orthonormal – i.e., $\langle\psi_n(t)|\psi_{n'}(t)\rangle = \delta_{nn'}$, where $\delta_{nn'} = 0$ if $n' \neq n$ and 1 if $n' = n$ – then the amplitudes satisfy

$$a_n(t) = \langle\psi_n(t)|\Psi(t)\rangle . \quad (3)$$

The $a_n(t)$ might be amplitudes for experiences localized in particular regions – in the visual cortex, in memory, etc. They are clearly driven by many influences which are both outside and within consciousness, from the barking of a dog to a worrisome memory. We represent this by the purely formal equation

$$a_n(t) = F_n[\psi, a, \chi] \quad (4)$$

where ψ , a , and χ respectively represent the neuronal states $|\psi_n(t)\rangle$ accessible to consciousness, the amplitudes $a_n(t)$ themselves, and direct influences χ_i outside consciousness (as opposed to indirect influences mediated by the $|\psi_n(t)\rangle$). A trivial example of a χ_i would be a stimulant, depressant, or hallucinogen entering the brain. The more obvious influences are the networks

within the brainwide web which acquire information processed by unconscious networks. It is they that interact to focus attention on inputs that may be enticing, threatening, ...

If $|N\rangle$ is the state of a specific neural network N , the amplitude of consciousness at a given moment t in that network is $\langle N|\Psi(t)\rangle$, and this is a quantity that is in principle accessible to experiment.

The functionals $F_n[\psi, a, \chi]$ are currently even more unknown and mysterious than the one in density-functional theory [16], which is treated with semiempirical models. A corresponding model for consciousness and other aspects of neuroscience will have to be vastly more sophisticated and complex.

As an aside, we note that the word consciousness as used here applies to any sufficiently complex *coherent domain* which has the qualities implicit in the functional $F_n[\psi, a, \chi]$ written above – including richness of experience and a capacity for self-awareness. A bat [17], bee [18], or creature of the remote future with artificial intelligence will certainly have a different consciousness_P than human beings – with e.g. different inner experiences red_P for the same colour. And they will have varying levels of consciousness. But the word “consciousness” is not misapplied to any biological or nonbiological creature that satisfies the conditions implied here. On the other hand, it is certainly a misuse of language to say that an electron or a stone or a nebula is conscious. This misuse reflects confusion between consciousness and the mysterious inner reality of nature. There is a reality_O accessible to science and a reality_P which we cannot directly access (with any science that is currently imaginable), but very little of reality_P is conscious.

We also note that our picture is similar to those proposed by others – for example, in Refs. [7] and [13] and by Tononi [19] – but here we place more emphasis on the physics setting and the physical reality of conscious experience.

Understanding of consciousness will come primarily from increasingly sophisticated experiments, employing methods that largely are developed by the physics community. In Fig. 1, we show an example of a promising new technique [20].

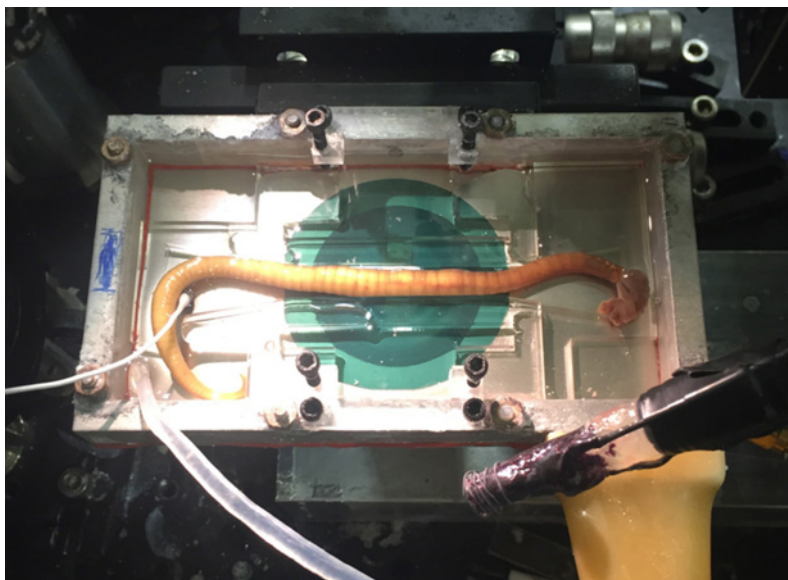


Figure 1. Noninvasive detection of action potentials with single-neurone sensitivity, using a sensor composed of quantum defects (nitrogen-vacancy colour centres in a diamond chip), with no observed adverse effect on the animal (marine fanworm *Myxicola infundibulum*). After Ref. [20]. Credit: John Barry, Matthew Turner, Jennifer Schloss, and Ronald Walsworth.

However, there are two points which demonstrate that very large-scale *simulations* can have an important role in complementing experiment:

- Unless a dramatically more effective technique is invented for observing deep inside a living (and normal) human brain with atomic-scale precision, neuroscience experiments are limited to rather gross observations from the outside. This means that realistic (and therefore extremely large-scale) simulations are needed to study what neurones do at the microscopic scale, in organisms whose neural complexities transcend that of Fig. 1 by many orders of magnitude. An analogy is a supernova, which also can only be observed from the outside, with no possibility of getting at the detailed processes where the explosion actually occurs. In both cases, the exterior observations serve as boundary conditions on simulations of the interior processes. If the underlying theory is good enough (e.g. nuclear physics and fluid dynamics for the supernova, biochemistry and biophysics for the brain) well-designed simulations can provide information that complements the experiments or observations.
- We have entered the era of petascale computing. The connections in the human brain are also petascale. This means that we can finally perform realistic simulations of the neuronal processes in the brain, within models that simplify the dynamics to include only the neural connections. But well-chosen models of these processes, guided by experimental discoveries of how neurones work, should be sufficient to capture the essential processes behind consciousness. And, of course, revealing how the brain works will certainly have a strong impact on medical and other biological applications.

Understanding consciousness is just as exciting a goal for our time as understanding the origin of the universe and its laws.

References

- [1] Alexander G et al 2019 The sounds of science – a symphony for many instruments and voices, submitted; see the section by Allen R E and Lidström S: What is consciousness, and is there really a “hard problem”?
- [2] Chalmers D 1995 Facing up to the problem of consciousness *Journal of Consciousness Studies* **2** 200
- [3] Wipfel J F 2007 Thomas Aquinas on the ultimate why question: why is there anything at all rather than nothing whatsoever? *The Review of Metaphysics* **60** 731
- [4] Hawking S W 1976 Black holes and thermodynamics *Phys. Rev. D* **13** 191
- [5] Tegmark M 2000 Importance of quantum decoherence in brain processes *Phys. Rev. E* **61** 4194
- [6] Dirac P A M 1939 A new notation for quantum mechanics *Mathematical Proceedings of the Cambridge Philosophical Society* **35** 416
- [7] Baars B J 1989 *A cognitive theory of consciousness* (Cambridge: Cambridge University Press)
- [8] Crick F 1994 *The astonishing hypothesis – The scientific search for the soul* (USA: Touchstone)
- [9] Kandel E R, Schwartz J H, and Jessell T M 2000 *Principles of neural science* (USA: McGraw-Hill Companies Inc)
- [10] Koch C 2004 *The Quest for Consciousness – A neurobiological approach* (USA: Roberts and Company)
- [11] Edelman G M 2006 *Second nature – brain science and human knowledge* (USA: Yale University Press)
- [12] Northoff G, Heinzel A, de Greck M, Bermpohl F, Dobrowolny H, and Panksepp J 2006 Self-referential processing in our brain – A meta-analysis of imaging studies on the self *NeuroImage* **31** 440 and references therein
- [13] Dehaene S 2014 *Consciousness and the brain: Deciphering how the brain codes our thoughts* (USA: Penguin)
- [14] Koch C 2017 *Consciousness – Confessions of a romantic reductionist* (USA: MIT Press)
- [15] Demirtas M et al. 2019 Hierarchical Heterogeneity across Human Cortex Shapes Large-Scale Neural Dynamics *Neuron* **101** 1-14 and references therein
- [16] Kohn W 1964 Inhomogeneous Electron Gas *Phys. Rev.* **136** B864
- [17] Nagel T 1974 What is it like to be a bat? *The Philosophical Review* **83** 435
- [18] Feinberg T E and Mallatt J M 2018 *Consciousness Demystified* (USA: MIT Press)
- [19] Tononi G 2004 An information integration theory of consciousness *BMC Neurosci.* **5** 42
- [20] Barry J F, Turner M J, Schloss J M, Glenn D R, Song Y, Lukin M D, Park H, and Walsworth R L 2016 Optical magnetic detection of single-neuron action potential using quantum defects in diamond *P.N.A.S.* **113** 14133