



# Evolution of the Wilson–Cowan equations

Hugh R. Wilson<sup>1</sup> · Jack D. Cowan<sup>2</sup>

Published online: 19 November 2021

© The Author(s), under exclusive licence to Springer-Verlag GmbH Germany, part of Springer Nature 2021

## Abstract

The Wilson–Cowan equations were developed to provide a simplified yet powerful description of neural network dynamics. As such, they embraced nonlinear dynamics, but in an interpretable form. Most importantly, it was the first mathematical formulation to emphasize the significance of interactions between excitatory and inhibitory neural populations, thereby incorporating both cooperation and competition. Subsequent research by many has documented the Wilson–Cowan significance in such diverse fields as visual hallucinations, memory, binocular rivalry, and epilepsy. The fact that these equations are still being used to elucidate a wide range of phenomena attests to their validity as a dynamical approximation to more detailed descriptions of complex neural computations.

## 1 Introduction

Astronomy and physics were the first of the natural sciences to be enriched by an intricate link to mathematics. Biology, although developing exceptionally powerful theories such as Darwinian evolution, lagged far behind historically in incorporating mathematics as an integral, predictive approach in its own right. However, mathematics was beginning to encroach on biology, and particularly neurobiology, in deep ways. Hodgkin and Huxley (1952) solved the dynamics of the action potential by introducing nonlinear dynamics in four coupled equations. This was rapidly recognized to be brilliant, and they received the Nobel prize in 1962.

Upon finishing my PhD in theoretical chemistry at The University of Chicago in 1969, I (Wilson) was extremely fortunate to be offered a postdoctoral fellowship by Jack Cowan at the university. He informed me that nonlinear dynamics

were beginning to have a major impact on neuroscience, and he encouraged me to start working on this subject. That was the beginning of the work that produced the Wilson–Cowan equations.

The intellectual background must begin with the Hodgkin–Huxley equations (Hodgkin and Huxley 1952). These four nonlinear differential equations are acknowledged to explain action potentials, key to all neural computation. As emphasized previously (Wilson 1999), nonlinear dynamics was the essential ingredient in providing a convincing explanation. The experiments, upon which the Hodgkin–Huxley equations were based, required poisoning various ion channels selectively. Thus, Hodgkin and Huxley had at their disposal a measured action potential plus different ionic flows following poisoning of various channels. The glue that put the picture together convincingly was nonlinear dynamics, which definitively showed that combination of the independent ion channel results did indeed generate an exceptionally accurate squid action potential. It should be noted that the differential equation computations were performed on an adding machine and took eleven days of computation just to predict one spike.

Given this background, Beurle (1956) noted the mathematical complexity of the Hodgkin–Huxley equations and sought a more simple approximation. Specifically, he introduced the concept of neural populations, which could naturally be described by the fraction of active neurones at any given point and time. As a physicist, he developed equations that described the propagation of neural activity waves across a one-dimensional tissue. Simplified nonlinear

---

Communicated by Peter J. Thomas.

---

To highlight the scientific impact of our Journal over the last decades, we asked authors of highly influential papers to reflect on the history of their study, the long-term effect it had, and future perspectives of their research. We trust the reader will enjoy these first-person accounts of the history of big ideas in Biological Cybernetics.

---

✉ Hugh R. Wilson  
hrwilson@yorku.ca

<sup>1</sup> Centre for Vision Research, York University, Toronto, Canada

<sup>2</sup> Department of Mathematics, University of Chicago, Chicago, USA

dynamics were manifest in his formulation, and this led him to an analytical travelling wave solution for neural activity of the form:

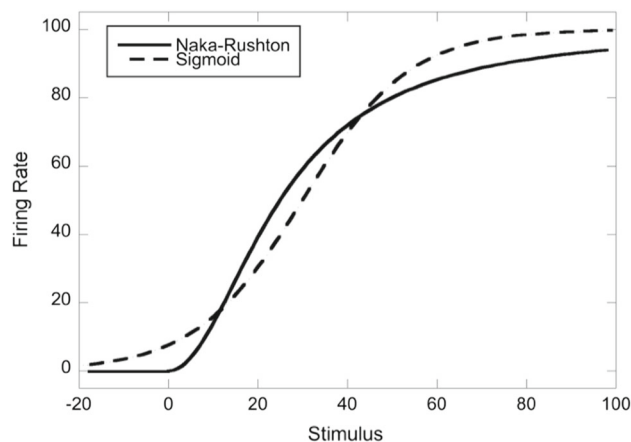
$$F(x - vt) = \frac{M}{2 \cosh^2(k(x - vt))} \quad (1)$$

where  $v$  is the velocity of the wave (to the right here) and  $F$  is the proportion of neurones active during passage of the wave.  $M$  and  $k$  are constants. Beurle acknowledged that this solution was unstable, as initial excitation to an amplitude slightly greater than that for Eq. (1) led to a transient increase in wave amplitude to saturation, whereas initial excitation to a slightly lower amplitude led to attenuation. In fact, the solution to Beurle's equation, termed a soliton, is also a solution to the Korteweg–De Vries equation that describes propagation of water waves in shallow channels (Korteweg and De Vries 1895). It is a soliton that conserves an infinite number of quantities.

Although providing several major insights, Beurle (1956) made a number of errors. First, he modelled the cortex as an unstable system that must be delicately balanced to remain plausible. Under these circumstances, he produced a system that was conservative, rather than dissipative, as the brain is known to be. Finally, and perhaps, the root cause of these problems was the omission of inhibition as a co-equal factor in brain function.

## 2 Wilson–Cowan equations

We designed the Wilson–Cowan equations to directly reflect the nonlinear dynamics inherent in excitatory–inhibitory interactions in cortical tissue. In addition, these equations were intended to be simpler than the Hodgkin–Huxley equations (1952) so that the dynamics of much larger populations of neurones could be explored. Given Beurle's (1956) insight, we chose to describe the activity of localized populations of neurones rather than the spiking of single neurones. However, it was clear that unstable soliton solutions could not effectively describe neural dynamics, so we developed equations that could produce travelling waves only under pathological conditions, such as epilepsy. Crucially, we argued that studies of neural activity must focus on the balance between excitatory and inhibitory activity in cooperating and competing neural populations. Our first results examined local temporal interactions of local neural populations (Wilson and Cowan 1972). Using phase plane techniques, it was shown that these equations could produce asymptotically stable excited states suggestive of short-term memory, limit cycle oscillations suggesting periodic motor control, and several more complex behaviours.



**Fig. 1** Sigmoid function used in the original Wilson–Cowan equations (dashed line) compared with a sigmoid (Naka–Rushton function, solid curve) designed as a more accurate approximation of cortical dynamics

Within a year, this local model was extended to include interactions among excitatory ( $E$ ) and inhibitory ( $I$ ) neural populations across space (Wilson and Cowan 1973), and this was published in *Kybernetik*, the parent of *Biological Cybernetics*. The original equations for the  $E(x,t)$  and  $I(x,t)$  populations are:

$$\begin{aligned} \tau_E \frac{\partial E}{\partial t} &= -E + (1 - rE)S_E(\beta_{EE}(x) \otimes E - \beta_{IE}(x) \otimes I + P(x, t)) \\ \tau_I \frac{\partial I}{\partial t} &= -I + (1 - rI)S_I(\beta_{EI}(x) \otimes E - \beta_{II}(x) \otimes I + Q(x, t)) \end{aligned} \quad (2)$$

where  $\tau_E$  and  $\tau_I$  are the respective time constants on the order of 10–15 ms,  $r$  is the refractory period,  $S$  is a sigmoid function increasing monotonically from its minimum at  $-\infty$  to its maximum value at  $+\infty$ , and  $P$  and  $Q$  describe external inputs to the respective populations. A logistic function was originally used for  $S$ , but as the exact mathematical form of sigmoid does not change the qualitative dynamics, other forms have more recently been used based on the experimental literature. For example, responses of cortical neurones have been described by a Naka–Rushton equation (Naka and Rushton 1966) with an exponent of approximately 2.4 (Sclar et al. 1990). A Naka–Rushton function with an exponent of 2 has also been used to facilitate mathematical solutions for equilibrium states (Wilson 1999). It is described by Eq. (3) and is plotted in Fig. 1.

$$\begin{aligned} S(x) &= \frac{M \cdot x^2}{\sigma^2 + x^2}, \quad x \geq 0 \\ S(x) &= 0, \quad x \leq 0 \end{aligned} \quad (3)$$

In this equation,  $M$  is the maximum firing rate, while  $\sigma$  defines the semi-saturation level, as when  $x = \sigma$ ,  $S = M/2$ . Values of  $M = 100$ ,  $\sigma = 25$  are shown in the figure.

The network inputs to each population are defined by spatial convolutions, denoted by  $\otimes$  in Eq. (2). In the original formulation the kernels were all functions of distance, and data available then suggested that they should be decaying exponentials of distance (Sholl 1956). Thus, the first convolution in Eq. (2) took the form:

$$\beta_{EE}(x) \otimes E(x) = \int_{-\infty}^{\infty} e^{-|x-x'|/\omega_{EE}} E(x') dx' \quad (4)$$

which describes recurrent  $E$  to  $E$  connections. The remaining three convolutions similarly describe all possible connections among the  $E$  and  $I$  populations. For different neural populations, different space constants can produce different ranges of interaction, thus permitting recurrent, long-range inhibition. More recent data have suggested the use of Gaussian kernels rather than decaying exponentials of distance, but this does not change the dynamics qualitatively.

Finally, it rapidly became apparent that the refractory period  $r$  mainly reduced the maximum firing rate but did not affect the dynamics substantially. Thus, most subsequent studies have set  $r = 0$ , thereby eliminating the term before the sigmoid input function. We should also emphasize here that values of parameters have largely been omitted below to emphasize major conceptual developments in the evolution of these equations. Details are available in the original references.

In this form the Wilson–Cowan equations exhibit a substantial range of dynamical modes, depending on the parameters chosen, that suggested explanations of a variety of cortical phenomena. The model could produce spatially stable inhomogeneous steady states that stored information dynamically and suggested a basis for short term memory. This is illustrated in Fig. 2, where the asymptotic  $E$  (red) and  $I$  (blue) activity is shown following brief activation at the two loci marked with arrows. Recurrent  $E$ – $E$  connections cooperatively maintain neural activity, while longer range  $I$ – $E$  connections maintain the localization. Another parameter regime produced spatially localized limit cycle oscillations with likely relevance to motor control. In addition, yet another set of parameters resulted in the generation of travelling waves, which were suggestive of epilepsy. Details of both the local computations, in which the convolution in Eq. 4 is replaced with simple multiplication by a weight constant, plus the one-dimensional spatial equations incorporating convolution, can be found in the original articles (Wilson and Cowan 1972, 1973). MATLAB scripts for many simulations along with parameter values are available elsewhere (Wilson 1999).

Although all of the applications of Wilson–Cowan described below have resulted from simulations, it is important to note that closed-form analytic solutions have very

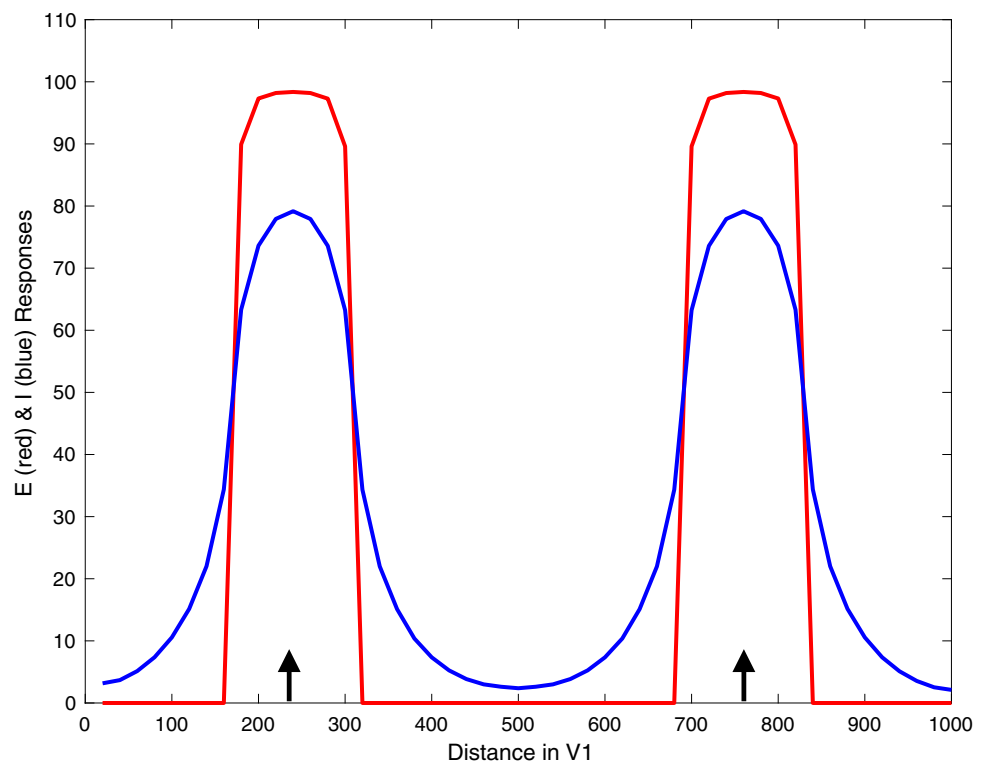
recently been obtained using specific forms of the sigmoidal nonlinearity and particular parameter values (Cowan et al. 2021). For these particular cases, travelling soliton waves of the basic form of Eq. (1) have been obtained. This is true even with the extension of Beurle’s formulation to include inhibition. These, of course, are solutions under particular functional conditions and do not encompass the much broader range of Wilson–Cowan dynamics as applied to particular areas of cortex.

In the years since the original publication of the Wilson–Cowan equations in *Kybernetik*, computer power has increased phenomenally. Relative to the Digital Equipment Corporation PDP8 on which the original simulations were done, a desktop iMac Pro runs more than  $10^7$  times faster (Wilson 2019)! This has engendered two major developments in neural modelling. First, much larger two-dimensional and multi-layer network simulations have become possible. Furthermore, these have incorporated more than just the two original  $E$  and  $I$  populations, thereby reflecting greater accuracy in describing cortical networks. In parallel, vastly more detailed simulations of individual neurones have been developed, some incorporating more than 1000 differential equations (Mainen et al. 1995). This reflects a trade-off between network complexity and single unit complexity constrained by available computer power. As the Wilson–Cowan equations emphasize the network approach, we shall focus on salient applications of this approach below. Details of parameter values are seldom given, as they are available in the original references.

### 3 Visual hallucinations

Drug-induced visual hallucinations frequently display one of a small number of geometric spatial patterns: concentric circles, radial spokes or arms spiralling outward from the centre (Siegel 1977). The natural question this raised was: what neuronal activity in the visual cortex might generate this percept in the absence of appropriate stimulation? The critical insight of Ermentrout and Cowan (1979) was that hallucinations could be explained by the spatially inhomogeneous steady states of the Wilson–Cowan equations if two additional factors were incorporated. First, the equations must be extended to two dimensions to represent the surface of visual cortex. The second key insight for explaining hallucinations was that the gradient of ganglion cells in the retina, ranging from densest in the fovea to very sparse in the periphery, meant that there must be a nonlinear mapping from the retina to the cortex. This mapping was shown to be well approximated by a complex logarithm in polar coordinates for the left and right half visual fields (Schwartz 1977, Schwartz 1980). If the retinal location of a stimulus point is described in polar coordinates as radius  $r$  and orientation  $\varphi$ , then the

**Fig. 2** Dynamical response of the Wilson–Cowan equations to brief stimulation by sufficiently strong pulses at the points marked by the vertical arrows. Due to recurrent excitation, network activity moves to two saturated peaks of  $E$  population activity (red curves), but these activity peaks are prevented from spreading by the spatially broader inhibition (blue curve). As multiple peaks can be thus stabilized, this was interpreted to provide a basis for short term memory



corresponding point of cortical projection in  $x, y$  coordinates is:

$$\begin{aligned} x &= \ln(1 + r) \\ y &= \phi \end{aligned} \quad (5)$$

This mapping to one hemisphere is illustrated in Fig. 3, where the bounding blue contour represents the vertical meridian, and the remaining three roughly horizontal lines, converging at the origin, represent orientations of  $45^\circ$ ,  $0^\circ$  (horizontal), and  $-45^\circ$  respectively. Had two concentric circles with radii of  $3.6^\circ$  and  $7.2^\circ$  been imaged on the retina, the cortical activation would be represented by the two almost parallel red horizontal bands. Conversely, a radial pattern starting a few degrees away from the fovea would have generated almost horizontal bands of activation due to this mapping. The critical insight was that these could be approximated as parallel neural activation patterns in V1 (Ermentrout and Cowan 1979).

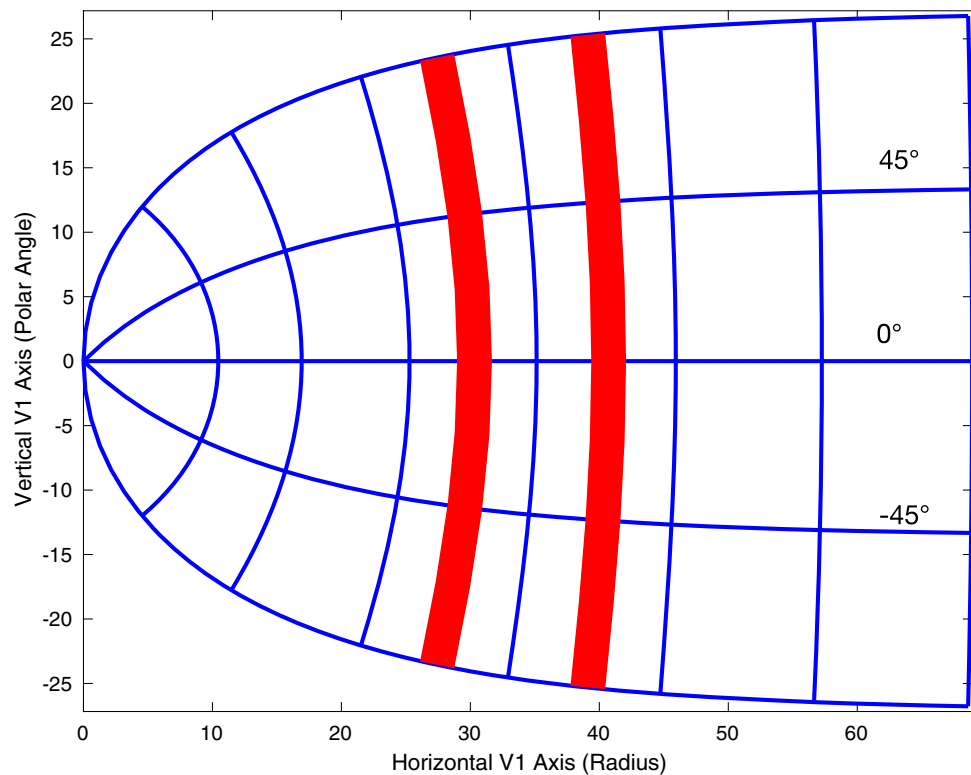
With the complex logarithmic mapping in 2D, plus an analysis of steady states for the Wilson–Cowan equations, Ermentrout and Cowan showed convincingly that visual hallucinations could be explained by asymptotically stable patterns of activated parallel lines of  $E$  neurons in V1. When projected back to the retina, these would have been concentric circles. Given drug activation of V1, higher levels of the visual system would have received the same stimulation from V1 as would have resulted from concentric

circles in the visual field. In addition, drug activation of horizontal contours of V1 activity would have resulted in the illusory percept of radial spokes, and activation of diagonal contours would have produced a hallucination of spirals. Complex checkerboards with checks increasing in size with distance from the fovea were also generated. Recall the minimal properties required: 2D generalization, asymptotically stable firing states of the network, and the empirically determined complex logarithmic mapping in Eq. (5).

This elegant explanation of visual hallucinations (Ermentrout and Cowan 1979) was very powerful, but it simplified by ignoring a very important organizing principle of V1: orientation columns (Hubel and Wiesel 1977). A more recent study by Cowan and colleagues has reexamined this by introducing multiple  $E$  populations, each tuned to a different peak orientation (Bressloff et al. 2001). In addition to incorporating multiple arrays of orientation-tuned  $E$  neurons, the  $E$  to  $E$  connectivity functions were altered to incorporate collinear facilitation. Physiological data from V1 had demonstrated the presence of long range  $E$ – $E$  interconnections among neurones with similar orientation preferences located at a distance from each other but aligned roughly collinearly (Ts'o and Gilbert 1988; Gilbert and Wiesel 1989). The addition of multiple orientations plus collinear facilitation to the Wilson–Cowan equations permitted a much wider range of visual hallucinations to be explained (Bressloff et al. 2001).

So far, nothing has been said about spatiotemporal hallucinations. It is known that uniformly flickering light, par-

**Fig. 3** Complex log retina to cortex mapping as defined by Eq. 5. The bounding blue contour represents the vertical meridian, while the horizontal meridian ( $0^\circ$ ) and the two diagonals are as indicated. Axis units are in mm along the cortical surface. A reflected map represents the other half of the visual field in the opposite V1 cortex. The two almost parallel horizontal bands are the projections of two half concentric circles from the retina onto the cortex



ticularly near 10.0 Hz, can induce spatiotemporal illusions, including auras in migraine sufferers (Crotogino et al. 2001). Ermentrout and colleagues have shown that the Wilson–Cowan equations with appropriate parameters can accurately predict this behaviour (Rule et al. 2011). A mathematical analysis of the equations plus simulations demonstrated the existence of nonlinear spatiotemporal oscillations. In response to a flickering stimulus that was uniform across space, the initial network oscillation was unstable, but after a transient period, broke into a spatially alternating pattern of synchronous oscillations. An example is illustrated in Fig. 4, where a network with appropriate parameters for an active transient mode (Wilson 1999) generate such a pattern in response to uniform sinusoidal stimulation. This oscillation consists of three spatial loci becoming active, then decaying, and the interdigitated three competing active foci beginning to fire. Boundary conditions were periodic, and the number of active populations per half cycle is determined by the spatial extent of the network. Ermentrout and colleagues went on to conduct experiments using a uniform annulus within which uniform field, counterphase flicker was employed (Pearson et al. 2016). Subjects perceived a ring of equally spaced illusory grey blobs that alternated between clockwise and counterclockwise rotation. As the annulus effectively reduced the stimulus to one-dimension (Wilson et al. 2001), the 1D Wilson–Cowan equations provided a neural model that effectively explained the illusion. Further

research has also invoked Wilson–Cowan in explaining additional spatiotemporal oscillations (Bertalmio et al. 2020).

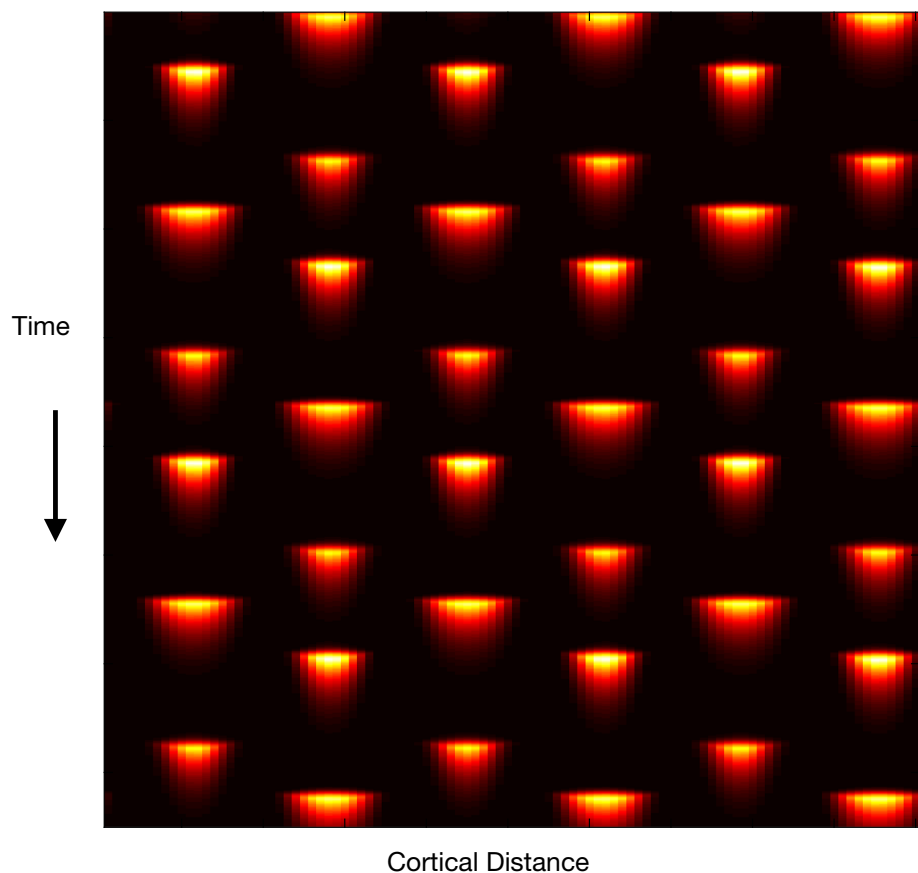
#### 4 Long-term memory

The activity depicted in Fig. 2 reflects short-range recurrent excitation localized by longer range inhibition. As such, the activity pattern is dependent on the neural connectivity, which is the same throughout this network. This suggested, however, that network connectivity could be learned from training examples and could therefore be used to encode long term memories. This was explored, first in networks with step function neural responses and subsequently with sigmoid nonlinearities by Hopfield (1982, 1984). It has been pointed out that Hopfield (1984) utilized a special case of the Wilson–Cowan equations in which the learned connection matrix was symmetric (i.e.  $\beta_{ij} = \beta_{ji}$ ), but no recurrent connections of a population to itself were permitted, so  $\beta_{ii} = 0$  (Destexhe and Sejnowski 2009). In addition, Hopfield networks permitted neurons to have both excitatory and inhibitory connections, so no explicitly inhibitory population was incorporated. However, a more realistic model with explicit populations of excitatory and inhibitory neurons can easily be developed (Wilson 1999).

In Hopfield networks, learning a pattern comprising  $N$  distinct neurons uses a Hebbian rule (Hebb 1949) to calculate the average cross-correlation between the responses of



**Fig. 4** Example of a spatiotemporal illusion resulting from uniform field flicker. The plot shows one spatial dimension on the abscissa and time increasing downward on the ordinate. E neuron activity levels are pseudocoloured as very low (black), intermediate (shades of red), and high (yellow). Although the stimulus is uniform flicker, the neural activity pattern bifurcates to a spatiotemporal alternation of competing activity foci



each  $ij$  pair ( $i \neq j$ ) active in the pattern. This cross-correlation (without a time lag that would encode causality) guarantees the symmetry of the connection matrix. Under these conditions, Hopfield constructed an energy function and proved that when stimulated with a sufficient percentage of a pattern, the network would asymptotically approach activity representing the full learned pattern.

Neural learning models have progressed far beyond Hopfield networks in the intervening years. Of particular importance have been deep learning networks. These networks incorporate multiple hierarchical levels of model neurons that learn their connection weights from the previous layers. In particular, errors between the desired output and the current output are calculated, and the relative error is assigned to the weights in the various layers based on the chain rule from calculus (Rumelhart et al. 1986). The current generic deep learning networks consist of hierarchical layers in which there is first a neighbourhood convolution with input from the previous layer, followed by a nonlinear transformation, such as the maximum within a neighbourhood, and then a spatial subsampling to a smaller upstream area (Reisenhuber and Poggio 1999; LeCun et al. 2015). Recently, alternative networks using both lateral interactions and feedback from higher areas have been shown to provide greater accuracy and enhanced biological plausibility (Spoerer et al. 2017). These

networks are consonant with multi-layer Wilson–Cowan networks, as they incorporate convolution of inputs followed by a sigmoid nonlinearity, and this approach has evolved to generate an enormous range of very powerful applications.

## 5 Binocular rivalry and travelling waves

The Wilson–Cowan equations have been used to explain a range of nonlinear visual phenomena, one of the most dramatic ones being binocular rivalry. Under normal stimulation, the eyes have evolved to sample the visual world from two slightly different visual perspectives, which the brain then combines to generate a percept of the third dimension, namely depth. However, when two radically different images (e.g. orthogonal gratings) are viewed independently by the two eyes, they cannot be interpreted in depth, and rivalry ensues. The brain then defaults to a stochastic oscillation in which first one monocular image and then the other is perceived, with the transitions between monocular images occurring approximately once every 2 s.

Before describing an explanation for binocular rivalry, it is necessary to review some more recent work on single neurons in the mammalian cortex. Since the Hodgkin–Huxley equations were developed to describe action potential generation

in the squid giant axon, studies of mammalian neocortical neurons have shown that many additional ion currents are present. In particular, excitatory neocortical neurons self-adapt as the result of a,  $Ca^{++}$  mediated  $K^+$  current that slowly hyperpolarizes the cell (McCormick and Williamson 1989; Sanchez-Vives et al. 2000). This current has an exponential time constant of a second or more, almost two orders of magnitude longer than typical synaptic currents. This fits well with the rate of alternations in binocular rivalry. Crucially, it is primarily excitatory rather than inhibitory neurons that possess this slow adapting current. The Wilson–Cowan equations have been extended to include this slow current by introducing an equation for a hyperpolarizing variable  $H$  (Wilson 1999, 2007; Wilson et al. 2000) Thus, the equation for  $E(x,t)$  in Eq. (2) is replaced by the two equations:

$$\begin{aligned} \tau_E \frac{\partial E}{\partial t} &= -E + S_E \left( \beta_{EE}(x) \otimes E \right. \\ &\quad \left. - \beta_{IE}(x) \otimes I + P(x, t) - gH \right) \\ \tau_H \frac{dH}{dt} &= -H + E \end{aligned} \tag{6}$$

with  $\tau_H$  approximately 1.0 s, almost two orders of magnitude greater than  $\tau_E$ . (Alternatively,  $H$  can be added to the semi-saturation constant  $\sigma$  in Eq. (3) with no significant qualitative differences.) The parameter  $g$  was assigned a value such that adaptation ultimately reduced the  $E$  firing rate to about 1/3 of its maximum, in agreement with electrophysiology. Finally, the stochastic component of rivalry can be simulated, if desired, by adding a Gaussian noise term to the  $H$  equation in Eq. (6), which produces a dominance time distribution that is well fit by either a log-normal or gamma distribution in accord with data (Fox and Herrmann 1967).

Given this embellishment of the Wilson–Cowan equations to include adaptation, binocular rivalry can now be explained. Separate equations describe excitatory neurons driven by the left and right eyes,  $E_L$  and  $E_R$  respectively. Competition between them is driven by inhibition,  $I_L$ , and  $I_R$ , from the opposite eyes. Thus, limit cycle competition emerges in which one eye first suppresses the other eye, but it then gradually adapts via its  $H$  current so the second eye can escape from the suppression and itself become dominant. This is illustrated in Fig. 5. The slow oscillation on a several second time scale is a result of the very long time constant  $\tau_H$  in Eq. (6). Note that the alternation is far too slow to be explained by ordinary synaptic inhibition.

Thus far, binocular rivalry had been treated as though it were a unitary phenomenon in which one monocular image uniformly replaced the other, but this is inaccurate. Rather, the suppressed image will first begin to replace the visible image at one point, and it will then transform into a travelling wave moving across the image from that point. To measure the travelling wave properties rivalry was restricted

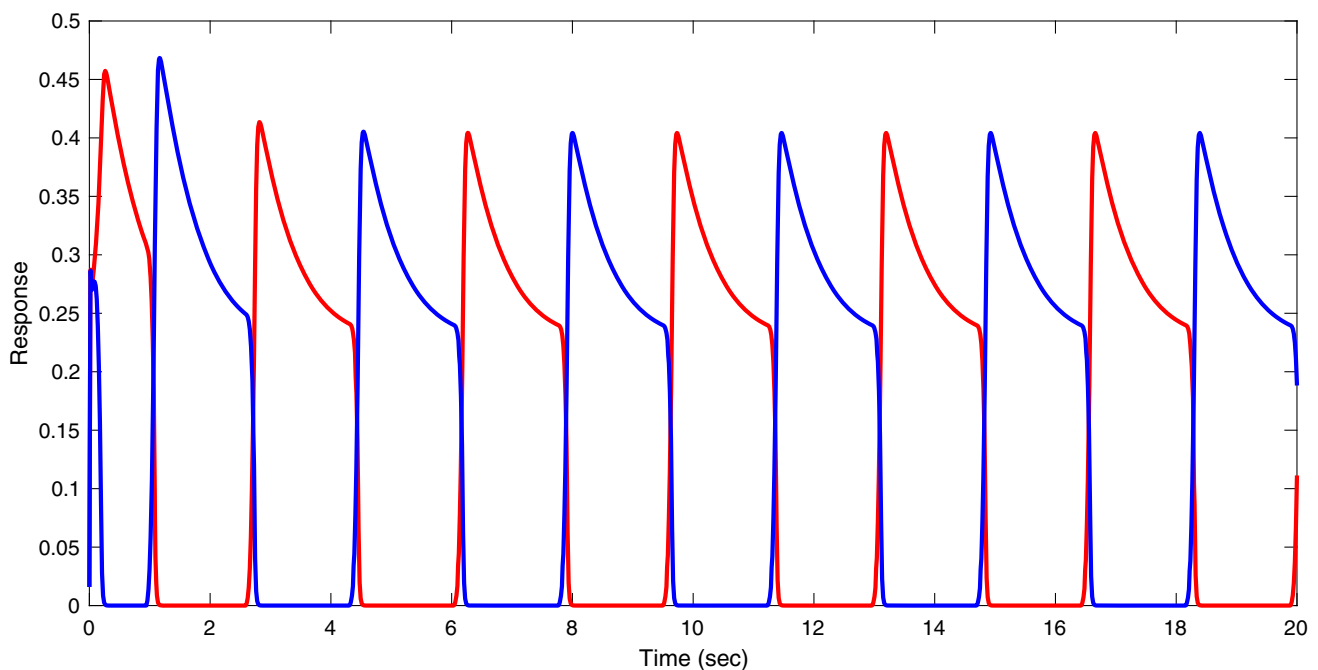
to a circular, effectively one-dimensional annulus or “race track” (Wilson et al. 2001). A wave of the suppressed pattern could then be triggered at any point around the circle, with the subject indicating when it reached the finish line. Using this psychophysical technique, it was shown that the wave travelled at a roughly constant speed across the cortex, which was estimated using the cortical mapping in Eq. (5) to be about 2.24 cm/s (Wilson et al. 2001). In an elegant subsequent fMRI experiment, wave speed was directly measured on the human cortex and found to be in good agreement with the psychophysical estimate (Lee et al. 2007).

The observed rivalry wave propagation was shown to be predictable by a model based on the Wilson–Cowan equations with adaptation (Wilson et al. 2001). The left and right eye patterns were represented by independent groups of  $E_L$  and  $E_R$  neurons that were mutually inhibitory. As the suppressed neurons become dominant at the moving front of the wave, they inhibit previously dominant neurons in front of the wave, thereby generating a release from inhibition. A detailed mathematical analysis of this wave propagation was later developed based on a variant of the Wilson–Cowan equations (Bressloff and Webber 2012).

Rivalry only occurs when the two monocular images are so different that they prevent fusion and the extraction of depth. If the two images are oriented cosine gratings; for example, an interocular orientation difference up to about  $\pm 6^\circ$  will result in fusion and the depth percept of a grating tilted forward or backward in depth (Blake and Wilson 2011). For an interocular orientation difference greater than that, however, fusion is impossible and rivalry ensues. Importantly, when the interocular orientation difference is continuously varied it has been shown that the switch from fusion to rivalry involves hysteresis (Buckthought et al. 2008). Fusion, rivalry, and hysteresis have all been captured using Wilson–Cowan equations with  $H$  current adaptation as shown in Eq. (6) (Wilson 2017). To particularize the model for V1, the model comprised 12 different excitatory populations (difference of  $15^\circ$  in preferred orientation between them) for each eye plus four separate inhibitory populations at each orientation (two for each eye): one for small orientation contrast normalization and one for long range orientation rivalry respectively. By incorporating additional neural populations and particularizing the connectivity among cell populations, hysteresis between fusion and rivalry can be explained by the Wilson–Cowan approach.

## 6 Epilepsy

The Wilson–Cowan equations have also been applied to epilepsy. This was based upon the original observations that spatially localized limit cycles could exist and that travelling waves could occur should the inhibition be too weak (Wilson and Cowan 1973). Shusterman and Troy (2008) later



**Fig. 5** Example of binocular rivalry generated by the Wilson–Cowan model with adaptation. Following a brief transient, a limit cycle results with left monocular activity (red) alternating with right monocular activity (blue) about once every two seconds. Model responses are in units

of relative contrast. The stochastic component has been omitted here to emphasize the limit cycle produced by H current adaptation and reciprocal inhibition

showed that, with appropriate parameters, local oscillations would lead to travelling waves and then to synchronous sustained activity. The results were comparable to data recorded from cortical surface electrodes during passage of epileptic seizures.

The explanations of cortical travelling waves in focal epilepsy have also led to a suggested modification of the Wilson–Cowan equations. To describe an epileptiform cortex, an effect was incorporated that does not occur under healthy physiological conditions. The Hodgkin–Huxley equations show that if extracellular  $K^+$  builds up in the extracellular space too much, there is a bifurcation in which all spiking vanishes. This was simulated in Wilson–Cowan by replacing the sigmoid function in Eq. (3) by a Gaussian so that excessive activity could actually drive the firing rate down to zero. Introduction of this physiologically important clinical observation produced an accurate simulation of focal epilepsy and its spread (Meijer et al. 2015).

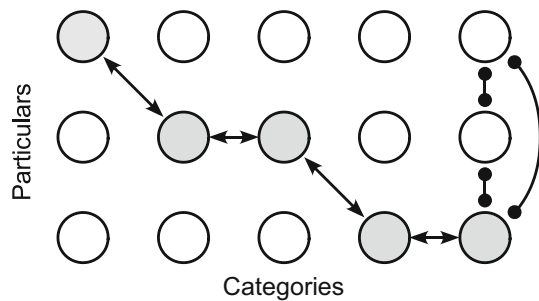
## 7 Decisions

Contemporary philosophy of mind has been strongly influenced by neuroscience. For example, Dennett (1996) has proposed that conscious decisions are the result of competition among a range of possibilities. Similarly, Dehaene, a

neuroscientist who has studied brain function in domains such a mathematics (Dehaene 1997), argued in a recent book on consciousness (Dehaene 2014): “Rivalry is, indeed, an apt metaphor for the constant fight for conscious access.” These opinions suggest that decision making by the brain might be usefully interpreted as a form of generalized rivalry among competing neural representations reflecting ideas based on past memories.

A candidate network for decisions, presumably mimicking areas in the prefrontal lobe, has been proposed related to the Wilson–Cowan network for rivalry (Wilson 2009; Wilson 2013). It can be argued that decisions among alternative interpretations or courses of action require reflection and the expenditure of neural energy in the cases where there is almost equal evidence in favour of several alternative courses of action. In this instance one typically considers each alternative in turn, seeking new evidence for or against it, and ultimately deciding on one alternative. The small network in Fig. 6 can be used to illustrate this. Imagine that each column of neural populations represents a category, perhaps the first as subject, second as verb, and so forth. The neural populations in each row are the particular possibilities for each category, such as I, you, he/she for subject; came, went, gave for verb, etc. Then a particular idea or possibility would be represented by a learned pattern association including one member of each column. With correlation learning (Hopfield





**Fig. 6** Model for learning and recalling a series of simple patterns. Each pattern is represented by one active population from each Category (e.g. five grey circles). During Hebbian learning all connections among all five units are symmetrically strengthened (only nearest neighbour connections are shown by double arrows to simplify diagram). Finally, all of the particular instances are mutually exclusive in any pattern and so are coupled via mutual inhibition. This is shown by the lines terminating in solid circles on the far right. Other vertical population arrays of particulars are also mutually inhibitory, although connections are not shown for clarity. Even this small network can store up to about five partially overlapping patterns. When a subset of these patterns are activated with nearly equivalent stimulus levels, they will become dominant one after another in a generalized rivalry oscillations. This is suggested to be a model for considering several alternatives for a course of action in deliberation. See text for other details

1982), the five populations in each pattern (grey for example) would have all their interconnection synapses strengthened as shown by arrows. Within a column all the possibilities are mutually exclusive in any one thought, hence mutual inhibition shown on the right by interconnections with solid circles. Under these conditions, the network will exhibit generalized rivalry in which one thought pattern will alternate or compete with several others in sequence. Furthermore, if a few patterns receive fairly weak evidence or input relative to others, they are automatically excluded by the dynamics from competition within the network. For a trivially small network of only 15 neurons, about five patterns can be learned (with partial pattern overlap), and from 2 to 5 of them will compete when receiving roughly comparable input. If this network is extended to a more realistic  $500 \times 1000$  neural populations or more, the link to decisions becomes quite plausible.

## 8 Discussion

The Wilson–Cowan equations have produced useful models and insightful explanations in many cortical areas and brain functions. The original spatial model (Wilson and Cowan 1973) was applied to several phenomena in V1, and this has since been extended to visual hallucinations (Ermentrout and Cowan 1979), multiple orientations defining E neuron groups (Bressloff et al. 2001), and multiple inhibitory groups in fusion and rivalry (Wilson 2017). In higher cortical areas Wilson–Cowan has served as a basis

for connectionist learning and memory first introduced in the Hopfield (1984) network. In addition, Wilson–Cowan has provided interpretation for travelling waves, both in binocular rivalry (Wilson et al. 2001; Lee et al. 2007) and in epilepsy (Shusterman and Troy 2008, Meijer et al. 2015). Finally, a generalization of rivalry has generated a possible explanation for decisions among several plausible alternative possibilities (Wilson 2009; Wilson 2013).

This range of applications of the Wilson–Cowan model depends on a number of extensions to the original model. Most obviously it is possible to generalize to multiple *E* and *I* populations reflective of particular cortical areas and functions. Among multiple populations there must be multiple population-to-population connectivity functions, which further individuate models. Furthermore, it has been shown to be important for connectivity functions to be learned (Hopfield 1984) in many cortical areas, and others have taken this much further in the development of multi-layer, deep learning convolutional networks (LeCun et al. 2015).

To this panoply of extensions must be added the ability to introduce ion currents that generate self-adaptation, particularly in *E* populations. Since the first development of Wilson–Cowan, it was discovered that many cortical pyramidal neurons, typically excitatory, incorporate slow  $\text{Ca}^{++}$  mediated  $\text{K}^{+}$  currents that serve to hyperpolarize active cells and thereby reduce their firing rate (McCormick and Williamson 1989; McCormick 1998). The key here is that this is a slow current, with a time constant much longer than the excitatory and inhibitory postsynaptic currents on which the Wilson–Cowan equations were based. Other suitably slow potentials can obviously be introduced as they are discovered. This, however, highlights one clear limitation of the Wilson–Cowan formalism: it cannot deal with extremely rapid neural variation, where simulation of individual spikes would be required. To cite but one example, humans can accurately discriminate the direction from which a sound emanates when the arrival time difference at the two ears is as small as 10 microseconds, or about 1/100 the width of an excitatory action potential. This exquisite sensitivity involves axonal delay lines and coincidence detection, and is clearly too fast for the Wilson–Cowan approach.

As the Wilson–Cowan approach has been very successful at incorporating additional populations, particularizing multiple sets of interconnections, and adding slow adaptive currents, it is appropriate to ask why this approach has succeeded in capturing key features of neural networks in numerous areas of the cortex. When we began this work a half century ago, it was frequently claimed that nonlinear dynamics was such a vast area, effectively infinite, that only linear approximations plus perhaps a few exact nonlinear solutions of particular equations were possible. Despite the clearly vast range of possible nonlinear systems, our argument then was that the nervous system depended on strong,

but understandable nonlinearities. Four key ideas incorporated in the original formulation epitomize this.

First, the simplification to neural populations and firing rates as espoused by Beurle (1956) simplified the description of neural networks relative to a detailed description using multiple equations to describe each individual spike (Hodgkin and Huxley 1952). As suggested previously (Wilson 1999), this is analogous to reporting data in terms of post stimulus time histograms (PSTH) rather than describing individual spike trains. Bins in PSTH are typically about 10–20 ms wide, which is reflected in the Wilson–Cowan time constants. Regarding human understanding, it is clear that the overwhelming wealth of our knowledge lies on the several second time scale, with memory vastly extending this back into the past. Thus, the time scale engendered by population dynamics in Wilson–Cowan fits naturally with human understanding as well.

Implicit in the paragraph above is the notion of an exponential time scale. The Wilson–Cowan equations were derived on the assumption that changes in firing rates of populations followed the time constants of typical neural post-synaptic potentials, EPSPs and IPSPs. Given multiple populations, multiple time constants have been incorporated. But by constraining our equations to time constants representing postsynaptic potentials, which implied ignoring individual spikes, a major simplification was accomplished: only population firing rates mattered. Compared to simulating individual spikes, this reduced computational requirements by almost two orders of magnitude. This meant that Wilson–Cowan could simulate networks about  $100 \times$  larger than spiking networks, given equivalent computing power.

Third, Wilson–Cowan was established on the extremely important physiological observation that excitatory and inhibitory neurons formed distinct, interacting populations (Wilson and Cowan 1972). This is now a commonplace, but it was ignored by all earlier attempts at network modelling. One can think of neural excitation as active and cooperative (“yang”), while inhibition functions to shut down excitatory activity by introducing competition (“yin”). This generates a cooperative-competitive dynamic that forms the basis of neural computation. Both components have been integral to Wilson–Cowan from the beginning.

This cooperative–competitive theme has been independently developed into a canonical model for neocortex (Douglas et al. 1989; Douglas and Martin 1991). These authors developed a model with two excitatory neuron populations, one comprising neurons in the supragranular cortical layers 2 and 3, and one comprising neurons in infragranular layers 5 and 6. Recurrent excitation both within and between the excitatory populations is incorporated into the model. The final group contains inhibitory neurons that are connected to themselves and to both excitatory groups to generate negative feedback. All neural populations are described by spike

rate dynamics. Thus, this candidate for a canonical circuit for neocortex may be regarded as an embellishment of Wilson–Cowan dynamics. Related  $E-I$  dynamics have also been used by Grossberg to develop theories of a large number of cortical functions (Grossberg 2021).

Finally, the sigmoid function is a key to the power of the Wilson–Cowan equations. The sigmoid captures the importance of neural thresholds, followed by the roughly linear activity increase with increasing stimulation, and finally by a compressive nonlinearity and saturation at high input levels. Unlike many nonlinear dynamical systems, his nonlinearity has facilitated state space analysis, as neural activity is constrained by the threshold to be  $\geq 0$ , and sigmoid saturation guaranteed that activity must be  $\leq M$ , the maximum value. This restriction to a hyper-cube in state space has proved valuable in many mathematical analyses, so the importance of the sigmoid to physiologically relevant neural modelling cannot be overstated.

The Wilson–Cowan equations have clearly proved their value at explaining a wide variety of neural functions in diverse cortical areas. Keys to this success are cooperative–competitive dynamics, population dynamics on a postsynaptic potential time scale, and sigmoid nonlinear boundedness by thresholds and saturation. Within this framework, there has been a rich evolution, and we suspect that they will continue to enhance our understanding of brain function.

## References

- Bertalmio M, Calatroni L, Franceschi V, Franceschiello B, Gomez-Villa A, Prandi D (2020) Visual illusions via neural dynamics: Wilson–Cowan-type models and the efficient representation principle. *J Neurophys* 123:1606–1618
- Beurle RL (1956a) Properties of a mass of cells capable of regenerating impulses. *Phil Trans R Soc B* 240:55–94
- Blake R, Wilson HR (2011) Binocular vision. *Vision Res* 51:754–770
- Bressloff PC, Cowan JD, Golubitsky M, Thomas PJ, Wiener MC (2001) Geometric visual hallucinations, Euclidean symmetry, and the functional architecture of striate cortex. *Phil Trans R Soc B* 356:299–330
- Bressloff PC, Webber MA (2012) Neural field model of binocular rivalry waves. *J Comput Neurosci* 32:233–252
- Buckthought A, Kim J, Wilson HR (2008) Hysteresis effects in stereopsis and binocular rivalry. *Vision Res* 48:819–830
- Cowan JD, Neuman J, Smith G, van Drongelen W (2021) Mathematical models for large-scale neuronal activity with exact closed form solutions. Submitted for publication
- Crotogino J, Feindel A, Wilkinson F (2001) Perceived scintillation rate of migraine aura. *Headache* 41:40–48
- Dehaene S (1997) *The number sense*. Oxford University Press, Oxford
- Dehaene S (2014) *Consciousness and the brain*. Viking Penguin, New York
- Dennett DC (1996) *Kinds of minds: toward an understanding of consciousness*. Basic Books, New York
- Destexhe A, Sejnowski TJ (2009) The Wilson–Cowan model, 36 years later. *Biol Cybern* 101:1–2

- Douglas RJ, Martin AC, Whitteridge D (1989) A canonical circuit for neocortex. *Neural Comput* 1:480–488
- Douglas RJ, Martin KAC (1991) A functional microcircuit for cat visual cortex. *J Physiol* 440:735–769
- Ermentrout GB, Cowan JD (1979) A mathematical theory of visual hallucination patterns. *Biol Cybern* 34:137–150
- Fox R, Herrmann J (1967) Stochastic properties of binocular rivalry alternations. *Percept Psychophys* 2:432–436
- Gilbert CD, Wiesel TN (1989) Columnar specificity of intrinsic horizontal and corticocortical connections in cat visual cortex. *J Neurosci* 9:2432–2442
- Grossberg S (2021) *Conscious mind resonant brain*. Oxford University Press, New York
- Hebb DO (1949) *The organization of behavior*. Wiley, New York
- Hodgkin AL, Huxley AF (1952) A quantitative description of membrane current and its application to conduction and excitation in nerve. *J Physiol* 117:500–544
- Hopfield JJ (1982) Neural networks and physical systems with emergent collective computational abilities. *Proc Natl Acad Sci USA* 79:2554–2558
- Hopfield JJ (1984) Neurons with graded response have collective computational properties like those of two-state neurons. *Proc Natl Acad Sci* 81:3088–3092
- Hubel DH, Wiesel TN (1977) Functional architecture of macaque monkey visual cortex. *Proc R Soc Lond B* 198:1–59
- Korteweg DJ, De Vries G (1895) On the change of form of long waves advancing in a rectangular canal, and on a new type of long stationary waves. *Phil Mag* 39:422–433
- LeCun Y, Bengio Y, Hinton G (2015) Deep learning. *Nature* 521:436–444
- Lee SH, Blake R, Heeger DJ (2007) Hierarchy of cortical responses underlying binocular rivalry. *Nat Neurosci* 10:1048–1054
- Mainen ZF, Jeoerges J, Huguenard JR, Sejnowski TJ (1995) A model for spike initiation in neocortical pyramidal neurons. *Neuron* 15:1427–1439
- McCormick DA (1998) Membrane properties and neurotransmitter actions. In: Shepherd GM (ed) *The synaptic organization of the brain*. Oxford University Press, New York, pp 37–75
- McCormick DA, Williamson A (1989) Convergence and divergence of neurotransmitter action in human cerebral cortex. *Proc Natl Acad Sci* 86:8098–8102
- Meijer HGE, Eissa TL et al (2015) Modeling focal epileptic activity in the Wilson–Cowan model with depolarization block. *J. Math. Neurosci.* 5:1
- Naka KI, Rushton WA (1966) S-potentials from colour units in the retina of fish. *J Physiol* 185:584–599
- Pearson J, Chiou R, Rogers S, Wicken M, Heitmann S, Ermentrout B (2016) Sensory dynamics of visual hallucinations in the normal population. *eLife* 5:e17072
- Reisenhuber M, Poggio T (1999) Hierarchical models of object recognition in cortex. *Nature Neurosci* 2:1019–1025
- Rule M, Stoffregen M, Ermentrout B (2011) A model of the origin and properties of flicker-induced geometric phosphenes. *PLoS Comput Biol* 7:e1002158
- Rumelhart DE, Hinton GE, Williams RJ (1986) Learning representations by back-propagating errors. *Nature* 323:533–536
- Sanchez-Vives MV, Nowak LG, McCormick DA (2000) Cellular mechanisms of long-lasting adaptation in visual cortical neurons in vitro. *J Neurosci* 20:4286–4299
- Schwartz EL (1980) Computational anatomy and functional architecture of striate cortex: a spatial mapping approach to perceptual coding. *Vision Res* 20:645–669
- Schwartz EL (1977) Spatial mapping in the primate sensory projection: analytic structure and relevance to perception. *Biol Cybern* 25:181–194
- Sclar G, Maunsell JHR, Lennie P (1990) Coding of image contrast in central visual pathways of the macaque monkey. *Vision Res* 30:1–10
- Sholl DA (1956) *The organization of the cerebral cortex*. Methuen, London
- Shusterman V, Troy WC (2008) From baseline to epileptiform activity: a path to synchronized rhythmicity in large-scale neural networks. *Phys Rev E* 77:061911
- Siegel RK (1977) Hallucinations. *Sci Am* 237(October):132–137
- Spoerer CJ, McClure P, Kriegeskorte N (2017) Recurrent convolutional neural networks: a better model of biological object recognition. *Front Psychol* 12:1551
- Ts'o DY, Gilbert CD (1988) The organization of chromatic and spatial interactions in the primate striate cortex. *J Neurosci* 8:1712–1727
- Wilson H (2019) Hyperchaos in Wilson–Cowan oscillator circuits. *J Neurophys* 122:2449–2457
- Wilson HR (1999) *Spikes, decisions, and actions: dynamical foundations of neuroscience*. Oxford University Press, Oxford
- Wilson HR (2007) Minimal physiological conditions for binocular rivalry and rivalry memory. *Vis Res* 47:2741–2750
- Wilson HR (2009) Requirements for conscious visual processing. In: Jenkin M, Harris LR (eds) *Cortical mechanisms of vision*. Cambridge University Press, Cambridge, pp 399–417
- Wilson HR (2013) Binocular rivalry: cooperation, competition, and decisions. In: Miller SM (ed) *The constitution of visual consciousness: lessons from binocular rivalry*. John Benjamin Publishing, Philadelphia, pp 281–303
- Wilson HR (2017) Binocular contrast, stereopsis, and rivalry: toward a dynamical synthesis. *Vision Res* 140:89–95
- Wilson HR, Blake R, Lee SH (2001) Dynamics of travelling waves in visual perception. *Nature* 412:907–910
- Wilson HR, Cowan JD (1972) Excitatory and inhibitory interactions in localized populations of model neurons. *Biophys J* 12:1–24
- Wilson HR, Cowan JD (1973) A mathematical theory of the functional dynamics of cortical and thalamic nervous tissue. *Kybernetik* 13:55–80
- Wilson HR, Krupa B, Wilkinson F (2000) Dynamics of perceptual oscillations in form vision. *Nature Neurosci* 3:170–176

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.