Wilson–Cowan Equations for Neocortical Dynamics

Jack D. Cowan¹ · Jeremy Neuman² · Wim van Drongelen³

Received: 27 July 2015 / Accepted: 18 December 2015 / Published online: 04 January 2016
© 2016 Cowan et al. This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

Abstract In 1972–1973 Wilson and Cowan introduced a mathematical model of the population dynamics of synaptically coupled excitatory and inhibitory neurons in the neocortex. The model dealt only with the mean numbers of activated and quiescent excitatory and inhibitory neurons, and said nothing about fluctuations and correlations of such activity. However, in 1997 Ohira and Cowan, and then in 2007–2009 Buice and Cowan introduced Markov models of such activity that included fluctuation and correlation effects. Here we show how both models can be used to provide a quantitative account of the population dynamics of neocortical activity.

We first describe how the Markov models account for many recent measurements of the resting or spontaneous activity of the neocortex. In particular we show that the power spectrum of large-scale neocortical activity has a Brownian motion baseline, and that the statistical structure of the random bursts of spiking activity found near the resting state indicates that such a state can be represented as a percolation process on a random graph, called directed percolation.

Other data indicate that resting cortex exhibits pair correlations between neighboring populations of cells, the amplitudes of which decay slowly with distance, whereas

² J.D. Cowan
cowan@math.uchicago.edu
J. Neuman
jneuman@uchicago.edu
W. van Drongelen
wvandron@peds.bsd.uchicago.edu

¹ Department of Mathematics, University of Chicago, 5734 South University Avenue, Chicago, IL 60637, USA
² Department of Physics, University of Chicago, 5720 South Ellis Avenue, Chicago, IL 60637, USA
³ Department of Pediatrics, University of Chicago, KCBD 900 East 57th Street, Chicago, IL 60637, USA
stimulated cortex exhibits pair correlations which decay rapidly with distance. Here we show how the Markov model can account for the behavior of the pair correlations.

Finally we show how the 1972–1973 Wilson–Cowan equations can account for recent data which indicates that there are at least two distinct modes of cortical responses to stimuli. In mode 1 a low intensity stimulus triggers a wave that propagates at a velocity of about 0.3 m/s, with an amplitude that decays exponentially. In mode 2 a high intensity stimulus triggers a larger response that remains local and does not propagate to neighboring regions.

**Keywords** Wilson–Cowan equations · Bogdanov–Takens bifurcation · Propagating decaying LFP and VSD waves · Localized decaying LFP and VSD responses · Neural network master equation · Directed percolation phase transition · Pair-correlations

# 1 Introduction

The analysis of large-scale brain activity is a difficult problem. There are about 50 billion neurons in the cortex of the human brain: 80% are excitatory, whereas the remaining 20% are inhibitory. Each neuron has about seven thousand axon terminals from other neurons, but there is some redundancy in the connectivity so that it has effective connections from about 80 other neurons, mostly nearest neighbors. Each neuron is actually a complex switching device, but in this review, we introduce only the simplest cellular model, that neurons are binary switches, either quiescent or activated. It follows that there are approximately $10^{1.5 \times 10^{10}}$ configurations of activated or quiescent neurons. Such a large configuration space suggests the need to use statistical methods to analyze large-scale brain activity. In addition there is some degree of microscopic randomness in neural connectivity, and there are also random fluctuations of neural activity, both of which also support the need for a statistical treatment, as noted by Sholl in 1956 [1].

# 2 Experimental Data on Large-Scale Brain Activity

There is a large body of data on large-scale brain activity, including electroencephalographic (EEG) recordings with large electrodes from the surface of the scalp, functional magnetic resonance (fMRI) measurements of blood flow in different brain regions (also large-scale), local field potentials (LFP) recorded with smaller electrodes, microelectrode recordings from or near individual neurons, or (currently) microelectrode arrays which can record the simultaneous activity of many neighboring neurons. Currently there are also new techniques for forming optical images of local brain activity, using voltage sensitive dyes (VSD). All such recordings can be classified as either spontaneous or resting activity, or stimulus-driven evoked activity.
Fig. 1 The upper trace is the first recording of spontaneous electrical activity from the human scalp. The lower trace is a 10 Hz oscillation. [Reproduced from [3]]

Fig. 2 The power spectrum of the occipital EEG of a resting, awake human. [Reproduced from [4]]

2.1 Resting Activity

We first consider the resting brain activity of unanesthetized animals first observed in animals by Caton in 1875 [2], and in humans by Berger in 1924 [3]. Recordings from the human scalp are referred to as electroencephalographs (EEG) and are measured via electrodes on the unshaven scalp. The voltage differences measured between such electrode pairs are about 50 µV. Figure 1 shows a typical EEG recording. It will be seen that there are intermittent bursts of 10 Hz oscillations in the scalp activity. These oscillations comprise the alpha rhythm, seen in awake relaxed humans, mainly in the occipital region of the brain which processes visual signals from the eyes. Figure 2 shows the power spectrum of such activity. It will be seen that there is a pronounced peak in the power spectrum at around 10 Hz and a secondary peak around 20 Hz. This peak is said to be in the range of the beta rhythm of occipital EEG activity. Interestingly if the contributions of such peaks are eliminated, what is left can be fitted with the function $a/(b + f^2)$, where $a$ and $b$ are constants, and $f$ is the frequency in Hz. Figure 3 shows such a function and its fit to the EEG power spectrum. It is important to note that this power spectrum fit is that of Brownian motion, which suggests that resting brain activity is largely desynchronized and random.

Other measurements of resting brain activity have been carried out on lightly anesthetized animals using local field potential recordings of spiking neuron activity, or
Fig. 3 The left panel shows the function $75/(3 + f^2)$, the right panel the fit of such a function to the EEG power spectrum shown in Fig. 2.

Fig. 4 The left panel shows the power spectra of LFP recordings from a cat’s visual cortex in response to sine-wave modulated grating patterns. [Reproduced from [5].] The right panel shows fMRI recordings of both resting and stimulated human brain activity, and their associated power spectra. [Reproduced from [6]]

else via fMRI measurements of blood flows in the brain that accompany unanesthetized brain activity. Figure 4 shows examples. Note the fit of the Brownian motion power spectrum $125/(5 + f^2)$ to the resting LFP.

2.1.1 Isolated Neocortex

But the most detailed studies, and the most information about the nature of spontaneous activity, has been obtained from studies of isolated neocortical slabs. The first detailed studies were carried out in the early 1950s by DeLisle Burns, on isolated slabs of parietal neocortex [7, 8]. The main relevant result was that very lightly anesthetized slabs spontaneously generated bursts of propagating activity from a number of randomly occurring sites. Any variation of the level of anesthesia, either up or down, abolished the activity.
However, it was not until 2003 that a systematic study of such burst activity was carried out by Beggs and Plenz [9] using isolated slabs of rat somatosensory cortex, either in mature tissue cultures, or else in slices. The tissue cultures exhibited spontaneous bursts of propagating activity in the form of local field potentials recorded at microelectrodes. The slices, however, were silent until stimulated with NMDA, a glutamate-receptor agonist, in combination with a dopamine D1-receptor agonist. In contrast to DeLisle Burns, Beggs and Plenz used an $8 \times 8$ microelectrode array to record local field potentials (LFPs) in the slab. The main result of their experiments is summarized in Figs. 5 and 6.

Beggs and Plenz’s conclusion is that such bursts of activity are avalanches defined as follows: the configuration of active electrodes in the array during one time bin of width $\Delta t$ is termed a frame, and a sequence of frames preceded and followed by blank frames is called an avalanche. However, successive frames are not highly correlated, so the activity is not wave-like: it is in fact self-similar, and in addition, the avalanche size distribution follows the power law $P[n] \propto n^\alpha$. In addition the exponent $\alpha$ is approximately $-1.5$. This is the mean-field exponent of a critical branching process [10]. This result was a step beyond that of Softky and Koch [11] who found Poisson-like spiking activity in individual cortical neurons, and introduced the possibility of criticality in brain dynamics. In fact this mean-field exponent turns up in several kinds of percolation processes on random graphs, including both isotropic and directed percolation. But branching and annihilating random walks are equivalent to directed
percolation, so it is possible that what Beggs and Plenz observed in cortical slices was a form of directed percolation. We will return to this topic later.

2.2 Driven or Stimulated Activity

In case there is an external stimulus, neocortical dynamics indicates a very different picture. It turns out that there is a big difference in the responses to weak stimuli, compared to those triggered by stronger stimuli. In addition correlations between pairs of neurons in driven neocortex have a shorter length scale than those found in spontaneous activity.

2.2.1 Weak Stimuli

The basic result for weak stimuli is that the cortical response is a propagating wave whose amplitude decays exponentially with distance. Figure 7 shows the cortical responses to low amplitude stimuli in the form of spikes, recorded by an implanted microelectrode array in three monkey visual cortices by Nauhaus et al. [13]. Each row shows data from the spike-triggered local field potentials (LFP) from a single location. The first column shows the dependence of time to peak of the LFP as a function of the cortical distance from the triggering electrode, and estimated propagation velocities. The second column shows the propagating wave, both as a pseudo-colored image, and as a plot of wave amplitude vs. distance from the triggering electrode, together with estimates of the space-constants of the decaying waves. The third column shows average LFP waveforms at three locations from the triggering spike.

It will be seen that the response is indeed a traveling LFP, whose velocity is about 25–30 cm/s. In addition the LFP amplitude decays exponentially, with a decay constant $\lambda$ of about 3 mm.
Fig. 7  Spikes of low amplitude initiate traveling waves of LFP in the cortex. See text for details. [Reproduced from [13]]

Fig. 8  Spikes of larger amplitude initiate standing waves of LFP in the cortex. See text for details. [Reproduced from [13]]

2.2.2 Strong Stimuli

In contrast the basic result for strong stimuli is that cortical responses to such stimuli are much more localized. Figure 8 shows a comparison of cortical responses to weak and strong stimuli [13]. It will be seen that responses to larger stimuli remain essentially localized. These observations immediately suggest a role for inhibition in localizing such responses.

2.2.3 Correlations

The basic result for correlations is that correlations between pairs of LFP fall off with separation distance, and such a falloff is much greater for strong stimuli than for
weaker ones; see Fig. 9. Thus strong stimuli weaken the intrinsic pair correlations that exist in spontaneous activity. See Lampl et al. and others [16–19]. These observations also suggest a role for inhibition.

To explain all these observations we need to understand the competing roles of neural excitation and inhibition in neural population dynamics. We therefore give a short account of the history and development of the Wilson–Cowan neural population equations.

3 Neural Population Equations

3.1 Introduction

Following early work by Shimbel and Rapaport [20], Beurle [21] focused, not on the activity of single neurons, but on the proportion of neurons activated per unit time in a given volume element of a slice or slab of neocortex, denoted by \( n(x, t) \). For all practical purposes this can be taken to be equivalent to the spike-triggered LFP and VSD described earlier.

Beurle introduced the update equation

\[
n(x, t + \tau) = q(x, t) f[n(x, t)],
\]

where \( q(x, t) \) is the density of quiescent neurons in the given volume element, and \( f[n(x, t)] \) the proportion of neurons receiving exactly threshold excitation. [There is an implicit assumption that individual neurons are of the integrate-and-fire variety.]

There are three points to note here.

1. By assuming that \( n(t + \tau) = q(t) f[n(t)] \) Beurle ignored the effects of fluctuations and correlations on the dynamics. It is not true that \( q \) and \( f[n] \) are statistically independent quantities, as was first pointed out in [22].
2. The update equation is incorrect. \( f[n] \) should be the proportion of neurons receiving at least threshold excitation, as was first noted by Uttley [23].

This proportion can be expressed [24] as:

\[
f[n] = \int_{-\infty}^{n} P(n_{\text{TH}}) dn_{\text{TH}} = \int_{-\infty}^{\infty} \vartheta[n - n_{\text{TH}}] P(n_{\text{TH}}) dn_{\text{TH}} = \langle \vartheta[n] \rangle, \tag{2}
\]
where \( \vartheta[n] \) is the Heaviside step function and \( \langle \vartheta[n] \rangle \) is the average of \( \vartheta[n] \) over the probability distribution of thresholds \( P(n_{TH}) \).

This implies that the function \( f[n] \) should have the form of a probability distribution function, not a probability density. In Cowan [25] the logistic or sigmoid form,

\[
f[n] = \left[ 1 + \exp[-n] \right]^{-1} = \frac{1}{2} \left[ 1 + \tanh \left( \frac{n}{2} \right) \right]
\]

(3)

was introduced, as an analytic approximation to the Heaviside step function used in McCulloch–Pitts neurons [26]. This indicates that the required continuum equations should represent the dynamics of a population of integrate-and-fire neurons in which there is a random distribution of thresholds.

The corrected version of Beurle’s equation takes the form

\[
n(x, t + \tau) = q(x, t) f[n(x, t)]
\]

\[
= q(x, t) \left[ \int_{-\infty}^{t'} dt' \int_{-\infty}^{\infty} d\xi' \alpha(t - t')[\beta(x - x')n(x', t') + h(x, t')] \right],
\]

(4)

where

\[
q(x, t) = 1 - \int_{t-r}^{t} n(x, t);
\]

(5)

\( r = 1 \) ms is the (absolute) refractory period or width of the action potential, and

\[
\alpha(t - t') = \alpha_0 e^{-(t-t')/\tau}, \quad \beta(x - x') = be^{-|x-x'|/\sigma}
\]

(6)

are the impulse response function and spatially homogeneous weighting function of the continuum model, with membrane time constant \( \tau \sim 10 \) ms, and space constant \( \sigma \sim 100 \) \( \mu \text{m} \).

3. Beurle’s formulation does not explicitly incorporate a role for inhibitory neurons.

### 3.2 The Wilson–Cowan Equations

Wilson and Cowan corrected and extended Beurle’s work and introduced equations for the population dynamics of a spatially homogeneous population of coupled excitatory and inhibitory binary neurons [24], and its extension to spatially inhomogeneous populations [27]. These equations take the forms

\[
\tau \frac{dE}{dt} = -E(t) + (1 - rE(t)) f_E[w_{EE}E - w_{EI}I + h_E(t)];
\]

\[
\tau \frac{dI}{dt} = -I(t) + (1 - rI(t)) f_I[w_{IE}E - w_{II}I + h_I(t)].
\]

(7)
for the spatially homogeneous case, and
\[
\tau \frac{\partial E(x, t)}{\partial t} = -E(x, t) + (1 - rE(x, t)) \times f_E \left[ \int_{-\infty}^{\infty} \rho_E d x' \beta_{EE}(x - x')E(x', t) - \int_{-\infty}^{\infty} \rho_I d x' \beta_{EI}(x - x')I(x', t) + h_E(x, t) \right],
\]
\[
\tau \frac{\partial I(x, t)}{\partial t} = -I(x, t) + (1 - rI(x, t)) \times f_I \left[ \int_{-\infty}^{\infty} \rho_E d x' \beta_{IE}(x - x')E(x', t) - \int_{-\infty}^{\infty} \rho_I d x' \beta_{II}(x - x')I(x', t) + h_I(x, t) \right],
\]
for the continuum form of the spatial case, in which \( \rho_E \) and \( \rho_I \) are, respectively, the packing densities of excitatory and inhibitory cells in the cortical slab.

Note that \( f_E[n] \) and \( f_I[n] \) are modified versions of the firing rate function \( f[n] \) introduced in Eq. (3), such that \( f_E[0] = f_I[0] = 0 \).

Note also that the variables \( E(x, t) \) and \( I(x, t) \) are time coarse-grained, i.e.
\[
E(x, t) = \int_{-\infty}^{t} dt' \alpha(t - t')n_E(x, t') \quad \text{and} \quad I(x, t) = \int_{-\infty}^{t} dt' \alpha(t - t')n_I(x, t'),
\]
where \( n_E(x, t) \) and \( n_I(x, t) \) are the proportions of excitatory and inhibitory neurons activated per unit time. It follows from Eq. (4) that \( \alpha(t) \) acts as a low-pass filter, and therefore that \( E(x, t) \) and \( I(x, t) \) are low-pass filtered version of \( n_E(x, t') \) and \( n_I(x, t') \), respectively. The net effect of such a coarse-graining is to remove oscillatory components of neural population responses greater than 100 Hz.

### 3.3 Attractor Dynamics

A major feature of Eq. (7) is that it supports different kinds of asymptotically stable equilibria. Figure 10 shows two such equilibrium patterns: There is also another phase plane portrait in which the equilibrium is a damped oscillation, i.e., a stable focus. In fact by varying the synaptic weights \( w_{EH} \) and \( w_{IH} \) or \( a = w_{EE} w_{II} \) and \( b = w_{IE} w_{EI} \) we can move from one portrait to another. It turns out that there is a substantial literature dealing with the way in which such changes occur, The mathematical technique for analyzing these transformations is bifurcation theory, and it was first applied to neural problems 53 years ago by Fitzhugh [28], but first applied systematically by Ermentrout and Cowan [29–31] in a series of papers on the dynamics of the mean-field Wilson–Cowan equations. Subsequent studies by Borisyuk
Fig. 10 The left panel shows the $E$–$I$ phase plane and nullclines of Eq. (7). The intersections of the two nullclines are equilibrium or fixed points of the equations. Those labeled (+) are stable, those labeled (−) are unstable. Parameters: $w_{EE} = 12$, $w_{EI} = 4$, $w_{IE} = 13$, $w_{II} = 11$, $n_H = 0$. The stable fixed points are nodes. The right panel shows an equilibrium which is periodic in time. Parameters: $w_{EE} = 16$, $w_{EI} = 12$, $w_{IE} = 15$, $w_{II} = 3$, $n_H = 1.25$. In this case the equilibrium is a limit cycle. [Redrawn from [24]]

Fig. 11 The left panel shows bifurcations of Eq. (7) in the spatially homogeneous case, organized around the Bogdanov–Takens (BT) bifurcation. SN1 and SN2 are saddle-node bifurcations. AH is an Andronov–Hopf bifurcation, and SHO is a saddle homoclinic-orbit bifurcation. Note that $a$ and $b$ are the control parameters introduced earlier. The right panel shows the nullcline structure of a Bogdanov–Takens bifurcation. At the Bogdanov–Takens point, a stable node (open circle) coalesces with an unstable point. [Redrawn from [34]]

and Kirillov [32] and Hoppenstaedt and Izhikevich [33] have greatly extended this analysis.

The left panel of Fig. 11 shows the detailed structure around such bifurcations. Evidently the saddle-node and Andronov–Hopf bifurcations lie near the Bogdanov–Takens bifurcation. Thus all the bifurcations described in the spatially homogeneous Wilson–Cowan equations lie close to such a bifurcation in the $(a, b)$-plane. The Bogdanov–Takens bifurcation depends on two control parameters $a$ and $b$, and is therefore of codimension 2. In such a bifurcation an equilibrium point can simultaneously become a marginally stable saddle and an Andronov–Hopf point. So at the bifurcation point the eigenvalues of its stability matrix have zero real parts. In addition the right panel of Fig. 11 shows how the fast E-nullcline and the slow I-nullcline intersect. The first point of contact of the two nullclines is the Bogdanov–Takens bifurcation point. The two nullclines remain close together over a large part of the subsequent $E$–$I$ phase space before diverging. As we will later discuss, this property...
of the nullclines is closely connected with the existence of a balance between excitatory and inhibitory currents in the network described by the Wilson–Cowan equations, and therefore with the existence of avalanches in stochastic Wilson–Cowan equations [35].

4 Stochastic Neural Dynamics

4.1 Introduction

To develop such equations we need to reformulate neural population dynamics as a Markov process. We first consider the representation of the dynamics of a cortical sheet or slab comprising a single spatially homogeneous network of $N$ excitatory binary neurons. Such neurons transition from a quiescent state $q$ to an activated state $a$ at the rate $f$ and back again to the quiescent state $q$ at the rate $\alpha$, as shown in Fig. 12.

4.2 A Master Equation for a Network of Excitatory Neurons

The first step is to formulate a master equation describing the evolution of the probability distribution of neural activity $P_n(t)$ in such a network. Consider first $n$ activated neurons, each becoming quiescent at the rate $\alpha$. This produces a flow out of the state $n$ at rate $\alpha$, proportional to $p_n(t)$, hence a term in the master equation of the form $-\alpha n P_n(t)$. Similarly the flow into $n$ from the state $n+1$ produces a term $\alpha(n+1) P_{n+1}(t)$. The net effect is the term

$$\alpha \left[ (n+1) P_{n+1}(t) - n P_n(t) \right].$$

(10)

Now consider the $N-n$ quiescent neurons in state $n$, each prepared to spike at rate $f[s_E(n)]$, leading to the term $-(N-n) f[s_E(n)] P_n(t)$, in which the total input is $s_E(n) = I(n)/I_{TH} = (w_{EE} n + h_E)/I_{TH}$, and $f[s_E(n)]$ is the function shown in Fig. 13, a low-noise version of Eq. (3).

The flow into the state $n$ from the state $n-1$ is therefore $(N-n+1) \times f[s_E(n-1)] P_{n-1}(t)$, and the total contribution from excitatory spikes is then

$$(N-n+1) f[s_E(n-1)] P_{n-1}(t) - (N-n) f[s_E(n)] P_n(t).$$

(11)
It follows that the probability $P_n(t)$ evolves according to the master equation
\[
\frac{dP_n(t)}{dt} = \alpha \left[ (n + 1) P_{n+1}(t) - n P_n(t) \right] + (N - n + 1) f\left[ s_E(n-1) \right] P_{n-1}(t) - (N - n) f\left[ s_E(n) \right] P_n(t). \tag{12}
\]
It is easy to derive an evolution equation for $\langle n(t) \rangle$, the average number of active neurons in the network, using standard methods. The equation takes the form
\[
\frac{d\langle n(t) \rangle}{dt} = -\alpha \langle n(t) \rangle + (N - \langle n(t) \rangle) f\left[ \langle s_E(n) \rangle \right], \tag{13}
\]
where $\langle s_E(n) \rangle = w_{EE} \langle n \rangle + h_{EE}$, and is the simplest form of Eq. (7) for a single excitatory population. Such a mean-field equation can be obtained in a number of different ways, in particular by using the van Kampen “system-size expansion” of Eq. (12) about a locally stable equilibrium [36]. However, as is well known, this expansion breaks down at a marginally stable critical point, e.g. at a Bogdanov–Takens point, and a different method must be used to analyze such a situation.

Before proceeding we note that these equations can be extended to cover the situation introduced in Eq. (7) which incorporates spatial effects. The variable $n(t)/N$ is extended to $n(x, t)$ representing the density of active neurons at the cortical location $x$ at time $t$, and the total input current $I(n)$ becomes the current density
\[
I(n(x)) = \int d^d x' w_{EE}(x - x') n(x') + h_E(x). \tag{14}
\]

### 4.3 A Master Equation for a Network of Excitatory and Inhibitory Neurons

Since about 1/5th of all cortical neurons are inhibitory, it is important to include the effects of such inhibition. We therefore extend Eq. (10) to include inhibitory neurons. The result is the master equation:
\[
\frac{dP(n_E, n_I, t)}{dt} = \alpha_E \left[ (n_E + 1) P(n_E + 1, n_I, t) - n_E P(n_E, n_I, t) \right]
+ \left[ (N_E - n_E + 1) f_E\left[ s_E(n_E - 1, n_I) \right] P(n_E - 1, n_I, t) \right].
\]
\[
- (N_E - n_E) f_E(n_E, n_I) P(n_E, n_I, t) \\
+ \alpha_I [(n_I + 1) P(n_E, n_I + 1, t) - n_I P(n_E, n_I, t)] \\
+ \left[ (N_I - n_I + 1) f_I(n_E, n_I - 1) P(n_E, n_I - 1, t) \\
- (N_I - n_I) f_I(n_E, n_I) P(n_E, n_I, t) \right].
\]

(15)

See Benayoun et al. [35] for a derivation of this equation. It is easy to derive Eq. (7) from this master equation. However, there is much more information as regards stochastic neural dynamics contained in Eq. (15) than is contained in such an equation. We refer, of course, to the effects of intrinsic fluctuations and of correlations.

5 Analyzing Intrinsic Fluctuations

To analyze such effects we need to look more closely at the attractor dynamics of Eq. (7). There are two cases to consider. In case 1, the attractor is either an asymptotically stable node or focus, or else a limit cycle. In case 2, the attractor is only marginally stable. In nonlinear dynamics this is a bifurcation point, e.g. a Bogdanov–Takens point, or a saddle node or Andronov–Hopf point. In statistical mechanics this is the critical point of a phase transition.

5.1 The System-Size Expansion

The system-size expansion was introduced by van Kampen [36] to analyze the effects of intrinsic fluctuations in case 1. The intuition behind this approach comes from the idea that if neurons are independently activated, then the total activity in an excitatory neural network in such a case is Gaussian distributed, with mean activity \( \langle n_E(t) \rangle \) proportional to \( N \), the total number of neurons in the network, and standard distribution proportional to \( \sqrt{N} \). So the number of neurons activated at a given time can be represented by the variable

\[
k = N n_E + \sqrt{N} \xi_E,
\]

(16)

where \( \xi_E \) is a Gaussian random perturbation.

The deterministic term satisfies Eq. (7), the random variable satisfies the linear Langevin equation

\[
\frac{d\xi_E}{dt} = A \xi_E + \sqrt{\alpha E n_E + (1 - n_E) f_E(n_E)} \eta_E
\]

(17)

to order \( N^{-1/2} \), where \( A \) is a constant and \( \eta_E \) is an independent white noise variable, whose amplitudes are calculated from Eq. (7).

An early version of this application of the system-size expansion can be found in Ohira and Cowan [37]. The extension to the excitatory and inhibitory neural network introduced in Eq. (7) is to be found in Benayoun et al. [35]. This paper is notable for its use of the Gillespie algorithm [38]. In this algorithm the simulation time is advanced only when the network’s state is updated, and the time intervals \( dt \) are
Fig. 14  Raster plot of the spiking patterns in a network of $N = 800$ excitatory neurons. Each black dot represents a neural spike. The mean activity $\langle n_E(t) \rangle$ is represented by the blue trace. Simulation using the Gillespie algorithm with parameter values $h_E = h_I = 0.001$, $w_0 = w_E - w_I = 0.2$, and $w_E + w_I = 0.8$. [Redrawn from [35]]

Fig. 15  Phase plane plots of the activity shown in Fig. 14 showing the vector field (blue) and nullclines $\dot{E} = 0$ (magenta) and $\dot{I} = 0$ (red), of Eq. (1) and plots of a deterministic (black) and a stochastic (green) trajectory starting from identical initial conditions. [Redrawn from [35]]

random variables dependent upon the network state. The simulation is carried out for a network in which certain symmetry conditions are introduced. These conditions are

$$w_{IE} = w_{EE} = w_E; \quad w_{EI} = w_{II} = w_I; \quad w_E - w_I = w_0,$$

where $w_0$ is kept constant. Figures 14, 15, 16, 17 show the results.

It should be evident from a study of these figures that the location of the fixed point of Eq. (7) remains unchanged as $w_E + w_I$ increases from 0.8 to 13.8, but the stochastic trajectory (green) becomes increasingly spread out as the nullclines become more parallel. Such a feature is also evident in the right panel of Fig. 11 in which the nullcline structure of the Bogdanov–Takens bifurcation is shown. It is also evident that a qualitative change has taken place in the nature of the activity: it has changed from random fluctuations to random bursts. Figures 18 and 19 make this clear.
Fig. 16  Raster plot of the spiking patterns in a network of $N = 800$ excitatory neurons. Each black dot represents a neural spike. The mean activity $\langle n_E(t) \rangle$ is represented by the blue trace. Simulation using the Gillespie algorithm with parameter values $h_E = h_I = 0.001$, $w_0 = w_E - w_I = 0.2$, and $w_E + w_I = 13.8$. [Redrawn from [35]]

Fig. 17  Phase plane plots of the activity shown in Fig. 16 showing the vector field (blue) and nullclines $\dot{E} = 0$ (magenta) and $\dot{I} = 0$ (red), of Eq. (1) and plots of a deterministic (black) and a stochastic (green) trajectory starting from identical initial conditions. [Redrawn from [35]]

5.2 Symmetries and Power Laws

It will be seen that the simulations described above, in which the network symmetry represented in Eq. (17) is present, have uncovered an important property, namely that a stochastic version of Eq. (7) incorporating such a symmetry can spontaneously generate random activity in the form of bursts, whose statistical distribution is a power law. The other important property concerns the basic network dynamics generating such bursts.

We first note the experimental data provided by DeLisle Burns [7] and Beggs and Plenz [9] described in the introduction, and then we discuss the underlying neurodynamics. The main result of the Beggs–Plenz observations is that isolated slices generate bursting behavior similar to that found in the simulations, with a power law burst distribution with slope exponent of $\beta = -1.5$. This should be compared with the simulation data shown in Fig. 18 in which $\beta = -1.62$. Note, however, that the geometry
5.2.1 Random Bursting

We turn now to the neuro-dynamics underlying random bursting. We first note that the fixed point of the dynamics remains unchanged as \( w_E + w_I \) increases from \( 0.8 \to 13.8 \), and \( n_E = n_I \). We also recall by Eq. (18) that \( w_E - w_I = w_0 = 0.2 \), of our network simulation is not comparable with that of a cortical slice. It remains to carry out simulations of the stochastic version of Eq. (7) on a 2-dimensional lattice. Work on this is currently ongoing. In any event, the Beggs–Plenz paper generated a great deal of interest in the possibility of critical behavior in the sense of statistical physics existing in stochastic neural dynamics, including the possibility that brain dynamics exhibits self-organized criticality. In the later parts of this paper, we briefly address this possibility.
so that as the network begins to fire in random bursts,

\[ w_0 \ll w_E + w_I. \]  

This inequality has a number of consequences [35, 39]. Most importantly, it allows a particular change of variables in Eq. (12) extended to include inhibition.

\[ \frac{d\langle n_E(t)\rangle}{dt} = -\alpha \langle n_E(t) \rangle + (1 - \langle n_E(t) \rangle) f[\langle s \rangle], \]

\[ \frac{d\langle n_I(t)\rangle}{dt} = -\alpha \langle n_I(t) \rangle + (1 - \langle n_I(t) \rangle) f[\langle s \rangle], \]

where \( \langle s \rangle = w_En_E - w_In_I + h \), and \( \langle n_E \rangle \) and \( \langle n_I \rangle \) are interpreted as the mean fractions of activated neurons in the network.

Now introduce the change of variables

\[ \Sigma = \frac{1}{2}(n_E + n_I), \quad \Delta = \frac{1}{2}(n_E - n_I), \]

so that Eq. (20) transforms into the equation

\[ \frac{d\langle \Sigma(t)\rangle}{dt} = -\alpha \langle \Sigma(t) \rangle + (1 - \langle \Sigma(t) \rangle) f[\langle s \rangle], \]

\[ \frac{d\langle \Delta(t)\rangle}{dt} = -\langle \Delta(t) \rangle(\alpha + f[\langle s \rangle]). \]

Such a transformation was introduced into neural dynamics by Murphy and Millar [39], and used by Benayoun et al. [35]. But it was introduced much earlier by Janssen [40] in a study of the statistical mechanics of stochastic Lotka–Volterra population equations on lattices, which are known to be closely related to stochastic neural population equations on lattices [41].

The important point about the transformed equations is that they are decoupled, with the unique stable solution \((\Sigma_0, 0)\), which is equivalent to \(n_E = n_I\) in the original variables. This is precisely the stable fixed point used in the simulations. Note also that, in the new variables \(\Sigma\) and \(\Delta\), the fixed point current is

\[ s = w_0 \Sigma + (w_E + w_I) \Delta + h. \]

So at the stable fixed point \((\Sigma_0, 0)\), \(s = w_0 \Sigma_0 + h\). Near such a fixed point, \(\Delta\) is only weakly sensitive to changes in \(\Sigma\), and \(\Sigma_0\) is unchanged when varying \(w_E + w_I\) for constant \(w_0\). Murphy and Miller called Eq. (20) an effective feed-forward system exhibiting a balance between excitatory and inhibitory currents, and a balanced amplification of a stimulus \(h\).

We can now perform a system-size expansion of the associated master equations [35], to obtain a two component linear Langevin equation for small Gaussian fluctuations about the stable fixed point \((\Sigma_0, 0)\). This takes the form

\[ \frac{d}{dt} \begin{pmatrix} \xi_{\Sigma} \\ \xi_{\Delta} \end{pmatrix} = \begin{pmatrix} -\lambda_1 & w_0 \\ 0 & -\lambda_2 \end{pmatrix} \begin{pmatrix} \xi_{\Sigma} \\ \xi_{\Delta} \end{pmatrix} + \sqrt{\alpha \Sigma_0} \begin{pmatrix} \eta_{\Sigma} \\ \eta_{\Delta} \end{pmatrix}, \]  

\[ \tag{24} \]
where the eigenvalues are \( \lambda_1 = (\alpha + f[s_0]) + (1 - \Sigma_0)w_0 f'[s_0] \) and \( \lambda_2 = (\alpha + f[s_0]) \), and \( w_{ff} = (1 - \Sigma_0)(w_E + w_I)f'[s_0] \).

The Jacobian matrix

\[
A = \begin{pmatrix}
-\lambda_1 & w_{ff} \\
0 & -\lambda_2
\end{pmatrix}
\]

is upper triangular and has eigenvalues \(-\lambda_1\) and \(-\lambda_2\). It follows that when \( w_0 \) is small and positive, then so are the eigenvalue magnitudes \( \lambda_1 \) and \( \lambda_2 \). So the eigenvalues are small and negative and the fixed point \((\Sigma_0, 0)\) is weakly stable. Evidently \( A \) lies close to the matrix

\[
B = \begin{pmatrix}
0 & w_{ff} \\
0 & 0
\end{pmatrix} = \begin{pmatrix}
0 & 1 \\
0 & 0
\end{pmatrix} w_{ff} = \bar{B} w_{ff}.
\]

But the matrix \( \bar{B} \) is the signature of the normal form of the Bogdanov–Takens bifurcation [33]. Thus the weakly stable node lies close to a Bogdanov–Takens bifurcation, as we have suggested.

5.3 Intrinsic Fluctuations at a Marginally Stable Fixed Point

We now turn to case 2, in which the network dynamics is at a marginally stable fixed point. As we showed earlier, such a fixed point is a Bogdanov–Takens point. We cannot use the system-size expansion at such a point, but we can use the methodology and formalism of statistical field theory [42–45]. However, for the neuro-dynamics considered in this article, case 1 applies: the resting and driven activities are all at or near a weakly stable fixed point. Despite this, the fact that the fixed point is only weakly stable indicates that the resting and weakly driven states lie in what has been called the fluctuation-driven region near the marginally stable fixed point [46]. Thus we need to outline some of the results of the analysis of case 2. The reader is referred to the details in the article by Cowan et al. [45].

The basic result is that the stochastic equivalent of the Bogdanov–Takens bifurcation is the critical point of a Directed Percolation phase transition, or DP [47]. In DP there are two stable states, separated by a marginally stable critical point. One of these is an absorbing state, corresponding to the neural population state in which all neurons are quiescent, so that the mean number of activated states or order parameter \( \langle n \rangle = 0 \). The other is one in which many neurons are activated, so that \( \langle n \rangle \neq 0 \) in the activated state. At a critical point the quiescent state becomes marginally stable and is driven by fluctuations into the activated state.

What is important for the present study is that in the neighborhood of such a critical point, i.e. in the fluctuation-driven regime, there are two significant features of the activity which relate to the experimental data we have described: (a) the resting behavior shows random burst behavior whose statistical signature is consistent with DP, i.e., the distribution of bursts follows a power law with slope exponent \(-1.5\), which is the slope of several forms of random percolation, including what is called mean-field DP [9, 10]; (b) intrinsic correlations are large, and pair correlations extend over significant cortical distances [18].
Fig. 20 The left panel shows the pair-correlation function for resting and driven activity, for additive Gaussian noise, the right panel that for resting and driven activity, for intrinsic noise, averaged over many simulations using the Gillespie algorithm. [Reproduced from [38]]

6 Modeling the Experimental Data

6.1 Resting Activity

6.1.1 Random Burst Activity

Assuming that the resting state occurs in the neighborhood of a weakly stable node or focus, to start with we can use the results of the system-size expansion of the $E$–$I$ master equation described earlier. The conclusion we reach is that in the case that there is a balance between excitation and inhibition, so that the network is at weakly stable node, or possible a focus, then random burst behavior with a power law slope exponent close to $-1.5$ is seen [35]. This is the result shown in Figs. 14–19, and of course the result is also completely consistent with the Beggs–Plenz data plotted in Figs. 5 and 6. We also note that these results are completely consistent with our recent analysis, Cowan [45], and with recent experimental data that demonstrates the sub-criticality of the resting state by Priesemann et al. [48].

6.1.2 Pair Correlations

As to pair correlations associated with resting or spontaneous activity, we refer to Fig. 9 in which the measured resting pair correlation falls off with pair separation, in both cats and monkeys. This finding can be replicated within the theoretical framework we have established in two differing ways.

(a) We first make use of Eq. (7), the mean-field Wilson–Cowan equations for the 1D-spatial case, and simply add $\delta$-correlated Gaussian noise to the equations. The resulting pair-correlation function for resting activity is shown in the left panel of Fig. 20.  
(b) We then use the stochastic Wilson–Cowan master equation introduced in Eq. (14), extended to the spatial case. In such a case the noise is multiplicative and intrinsic, and we used the Gillespie algorithm [38] to simulate the process.
Fig. 21  A Variation in the LFP amplitude of decaying waves. The largest amplitude is the initial response to a brief weak current pulse. B The exponential decay of the LFP amplitude, as a function of distance traveled. C Time–distance plot of the peak amplitude indicating that the velocity of wave propagation is constant at about 0.3 m s$^{-1}$. D Localized LFP in response to a strong current pulse. E Rapid decay of the amplitude in a linear fashion. F Very slow propagation of the LFP.

Such simulations of the behavior of Wilson–Cowan equations replicate very accurately, the pair-correlation behavior shown in Fig. 9, reported in [13], both for resting activity and for driven activity.

6.2 Driven Activity

6.2.1 Weak Stimuli

We now consider the results reported by Carandini et al. [12–14], of traveling, decaying waves seen in LFP, shown in Figs. 7 and 8; and by Muller and Destexhe [15], in VSD recordings, in response to brief weak current pulses. These results can be replicated quite precisely in simulations of Eq. (8), in which the network dynamics is near the balanced state in which $E \approx I$. The top row of Fig. 21 shows a simulation of these simulations. These results should be compared with those plotted in Fig. 7. It should be clear that the simulations replicate very accurately, such data.

6.2.2 Strong Stimuli

The other result reported by Carandini et al. is that for strong stimuli the resulting LFP does not propagate very far and remains localized. This property was actually reported in Wilson and Cowan’s 1973 paper [27]! The bottom row of Fig. 21 shows a current simulation of this property, again in which the network state is approximately balanced.
6.3 Explaining the Differing Effects of Weak and Strong Stimuli

It is evident that there are big differences between the effects produced by weak and strong stimuli. What is the cause of such differences? Given that the only parameter in the Wilson–Cowan equations that is varied in the two cases is the stimulus intensity, this suggests that the property which causes the different responses is the level of inhibition. It must therefore be the case that the threshold for inhibitory activity is set high enough that weak stimuli do not trigger inhibitory effects, whereas strong enough stimuli do trigger such effects. Indeed this is one of the possibilities suggested by Carandini et al. in their papers. Thus inhibition blocks LFP (and VSD) propagation.

This possibility is also consistent with the effects of stimuli on pair correlations. We predict that the pair-correlation function should falloff more slowly in the case of resting or weakly driven activity, than in the case of stronger stimuli. Such a result would be consistent with the suggestions of Churchland et al. that one effect of stimuli is to lower noise levels.

7 Discussion

7.1 Early Work

The main results described in this article concern the use of the Wilson–Cowan equations to analyze the dynamics of large populations of interconnected neurons. Early workers, including Shimbel and Rapaport [20] and Beurle [21], appreciated the need to use a statistical formulation of such dynamics, but lacked the techniques to go beyond mean-field theory. The Wilson–Cowan equations [24, 27] were the first major attempt at a statistical theory, but still lacked a treatment of second and higher moments. However, what the equations did describe was mathematical conditions for attractor dynamics. Further work by Ermentrout and Cowan [29–31] and by Borisyuk and Kirillov [32], and Hoppenstaedt and Izhikevich [33, 34] used the mathematical techniques of bifurcation theory to more fully analyze such dynamics. The main result was that neural population dynamics is organized around a Bogdanov–Takens bifurcation point, in the neighborhood of which (in a phase space of two control parameters) are saddle-node and Andronov–Hopf bifurcations. Thus neural network dynamics contains locally stable equilibria in the form of stationary and oscillatory attractors.

7.2 The System-Size Expansion

The problem of going beyond the mean-field regime proved to be very difficult. Some progress was made by Ohira and Cowan [37] formulating stochastic neural dynamics in the neighborhood of a stable stationary equilibrium as a random Markov process and using the Van Kampen system-size expansion [36]. Further progress along these lines was made by Benayoun et al. [35] who formulated Eq. (7) as a random Markov process. But Benayoun et al. went further, by incorporating some symmetries into
Eq. (7) discovered by Murphy and Miller [39] which, in retrospect, located the stationary equilibrium of the equations near a Bogdanov–Takens point. The result was that the stochastic version of Eq. (7) generates the random bursts of activity we now refer to as avalanches. In addition the avalanche distribution was that of a power law, with a slope exponent $\beta = 1.6$. This value is close to that observed by Beggs and Plenz [9] in their observations of neural activity in an isolated cortical slab, of avalanche distributions with a slope exponent of $\beta = 1.5$.

### 7.3 A Statistical Theory of Neural Fluctuations

There remained the problem of developing a statistical theory for the fluctuations about a marginally stable critical point, such as a Bogdanov–Takens point. This problem was formulated by Cowan [42] and solved by Buice and Cowan [43, 44]. This is a major result since it connects the theory of stochastic neural populations at a critical point, with many well studied examples of other populations of interconnected units. Examples include percolation in random graphs, branching and annihilating random walks, catalytic reactions, interacting particles, contact processes, nuclear physics, and bacterial colonies. Many of these processes are subject to a phase transition, known as a directed percolation phase transition (DP), and all these processes have the same statistical properties, including the appearance of random bursts or avalanches.

### 7.4 Relation to Experimental Data

However, although the statistical theory is relevant to the pair-correlation problem, it is the mean-field Wilson–Cowan equations that proved to be necessary and sufficient to analyze neocortical responses to brief stimuli, both weak and strong. In our opinion the close fit between the data and the simulations of the Wilson–Cowan equations with fixed parameters is quite remarkable, especially given the fact that these equations were formulated some 45 to 50 years ago! More detailed papers dealing with these and other results on neocortical responses to stimuli are in preparation.

### Competing Interests

The authors declare that there are no competing interests.

### Authors’ Contributions

All authors contributed equally to the writing of this paper. All authors read and approved the final manuscript.

**Acknowledgements**  
We thank Dr. Mark Hereld, Argonne National Laboratory, for many helpful discussions and comments. JN was supported in part by the Dr. Ralph and Marian Falk Medical Research Trust and R01 NS084142-01 to Prof. van Drongelen.
References

1. Sholl DA. The organization of the cerebral cortex. London: Methuen; 1956.
2. Caton R. Br Med J. 1875;2:278.
3. Berger H. Arch Psychiatr Nervenkrankh. 1929;87:527–70.
4. Rowe DL, Robinson PA, Rennie CJ. J Theor Biol. 2004;231:413–33.
5. Henrie JA, Shapley R. J Neurophysiol. 2005;94:479–90.
6. Thurner S, Windischberger C, Moser E, Walla P, Barth M. Physica A. 2003;326:511–21.
7. Burns BD. J Physiol. 1951;112:156–75.
8. Burns BD. J Physiol. 1955;127:168–88.
9. Beggs JM, Plenz D. J Neurosci. 2003;23(35):11167–77.
10. Alstrom P. Phys Rev A. 1988;38(9):4905–6.
11. Softky WR, Koch C. J Neurosci. 1993;13(1):334–50.
12. Benucci A, Frazor RA, Carandini M. Neuron. 2007;55:103–17.
13. Nauhaus I, Busse L, Carandini M. Ringach DL. Nat Neurosci. 2009;12(1):70–6.
14. Nauhaus I, Busse L, Ringach DL, Carandini MN. J Neurosci. 2012;32(9):3088–94.
15. Muller L, Reynaud A, Chavane F, Destexhe A. Nat Commun. 2014;5:3675. doi:10.1038/ncomms4675.
16. Lampl I, Reichova I, Ferster D. Neuron. 1999;22:361–74.
17. Kohn A, Zandvakili A, Smith MA. Curr Opin Neurobiol. 2009;19:434–8.
18. Schulz DPA, Carandini M. Biol Rep. 2010;2:43. doi:10.3410/B2-43.
19. Churchland MM, Yu BM, Cunningham JP, Sugrue LP, Cohen MR, Corrado GS, Newsome WT, Clark AM, Hosseini P, Scott BB, Bradley DC, Smith MA, Kohn A, Movshon JA, Armstrong KM, Moore T, Chang SW, Snyder LH, Lisberger SG, Friebe NJ, Finn IM, Ferster D, Ryu SI, Santhanan G, Shenoy KV. Nat Neurosci. 2010;13(3):369–78.
20. Shimbel A, Rapoport A. Bull Math Biophys. 1948;10:41–55.
21. Beurle RL. Philos Trans R Soc Lond B. 1956;240(669):55–94.
22. Smith DR, Davidson CH. J ACM. 1962;9(2):268–79.
23. Uttley AM. Proc R Soc Lond B. 1955;144(915):229–40.
24. Wilson HR, Cowan JD. Biophys J. 1972;12:1–22.
25. Cowan JD. In: Caianiello ER, editor. Neural networks. Berlin: Springer; 1968. p. 181–8.
26. McCulloch WS, Pitts WH. Bull Math Biophys. 1943;5:115–33.
27. Wilson HR, Cowan JD. Kybernetik. 1973;13:55–80.
28. Fitzhugh R. Biophys J. 1961;1(6):445–66.
29. Ermentrout GB, Cowan JD. J Math Biol. 1979;7:265–80.
30. Ermentrout GB, Cowan JD. Biol Cybern. 1979;34:137–50.
31. Ermentrout GB, Cowan JD. SIAM J Appl Math. 1980;38:1–21.
32. Borisyuk RM, Kirillov AB. Biol Cybern. 1992;66(4):319–25.
33. Hoppenstaedt FC, Izhikevich EM. Weakly connected neural networks. Cambridge: MIT Press; 1997.
34. Izhikevich EM. Dynamical systems in neuroscience. Cambridge: MIT Press; 2007.
35. Benayoun M, Cowan JD, van Drongelen W, Wallace E. PLoS Comput Biol. 2010;6(7):e1000846.
36. Van Kampen NG. Stochastic processes in physics and chemistry. Amsterdam: North-Holland; 1981.
37. Ohira T, Cowan JD. In: Ellacott SW, Mason JC, Anderson I, editors. Mathematics of neural networks: models, algorithms and applications. Berlin: Springer; 1997. p. 290–4.
38. Gillespie D. J Phys Chem. 1977;81:2340–61.
39. Murphy BK, Miller KD. Neuron. 2009;61(4):635–48.
40. Janssen HK. J Stat Phys. 2001;103(5–6):801–39.
41. Cowan JD. In: Coombes S, biem Graben P, Pothast R, Wright JJ, editors. Neural field theory. New York: Springer; 2014. p. 47–99.
42. Cowan JD. In: Touretzsky DS, Lippman RP, Moody JE, editors. Advances in neural information processing systems. vol. 3. San Mateo: Morgan-Kaufmann; 1991. p. 62–8.
43. Buice MA, Cowan JD. Phys Rev E. 2007;75:051919.
44. Buice MA, Cowan JD. Prog Biophys Mol Biol. 2009;99(2, 3):53–86.
45. Cowan JD, Neuman J, Kiewiet B, van Drongelen W. J Stat Mech. 2013;3:04030.
46. Cai D, Tao L, Shelley M, McLaughlin DW. Proc Natl Acad Sci USA. 2004;101(20):7757–62.
47. Hinrichsen H. Adv Phys. 2000;49(7):815–958.
48. Priesemann V, Wirbal M, Valderama M, Pröpper R, Le Van Quyen M, Geisel T, Triesch J, Nikolić D, Munk MHJ. Front Syst Neurosci. 2014;8:108.