Cortical Population Dynamics and Psychophysics

Udo A. Ernst and Christian W. Eurich
Institute for Theoretical Neurophysics, University of Bremen
(uernst,eurich)@physik.uni-bremen.de
January 20, 2002

1 Introduction

Visual cortex is one of the most extensively studied regions in the mammalian brain. Over the past decade, numerous experimental studies have accumulated anatomical, physiological, and psychophysical knowledge about its properties. Much effort has been put into theoretical and computational work to reproduce basic phenomena and to explain their underlying mechanisms.

In this review, we discuss one specific computational approach that has been successfully applied to a variety of problems on different levels of cortical information processing. The approach describes the cortical population dynamics in the form of structurally simple differential equations for the neurons’ firing activities. The model class has been introduced by Wilson and Cowan (1972, 1973), and is still very popular for, in our opinion, two reasons: first, it is powerful enough to reproduce a variety of cortical phenomena and it captures the dynamics of neuronal populations seen in numerous experiments. Second, its degree of complexity is still low enough to allow for analytical treatment which yields an understanding of the mechanisms leading to the observed behaviour.

In the next section, we will introduce the model class and discuss some of its basic properties. The following sections will then show, how this model can be applied to explain dynamical properties of the primate visual system on different levels, reaching from single neuron properties like selectivity for the orientation of a stimulus, up to higher cognitive functions related to the binding and processing of stimulus features in psychophysical discrimination experiments.

The goal of our contribution is to show that a model which abstracts from biophysical details is often sufficient to identify possible neuronal mechanisms of cortical information processing. The diversity of the examples we mention demonstrates that even such a simplifying approach can put seemingly unrelated or even controversial findings into one coherent, unifying picture.

2 The Wilson-and-Cowan model class

For a basic introduction to differential equations in the context of neural systems and the class of models described here we refer the reader to the textbook by Wilson (1999).
Single units. The basic unit of the model is a neuronal population. The dynamics of an uncoupled population is described by an ordinary differential equation for its activity $A(t)$ – to be interpreted as the population firing rate – which consists of a decay term and the synaptic input $I(t)$,

$$\frac{dA(t)}{dt} = -A(t) + h(I(t)). \quad (1)$$

$\tau$ is a time constant, and $h$ is a monotonically increasing gain function describing the firing frequency of a population in dependence of its synaptic input $I$. Wilson and Cowan (1972) derived this equation from a more general integro-differential equation by applying a temporal filter (time course graining). Therefore, the resulting equation (1) is structurally simple, but not exact – one has to bear in mind that temporal variations on a small time scale have been averaged out.

Columns. Cortical nervous tissue contains both excitatory and inhibitory neurons in a dense network. A general model of this network necessarily has to include both cell types forming one excitatory and one inhibitory population, denoted by indices $e$ and $i$, respectively. Both populations, each of which typically represents some hundred single neurons, are mutually connected with weights $w_{ee}, w_{ie}, w_{ei},$ and $w_{ii}$. Index pairs like $ei$ are to be interpreted as a connection originating at the excitatory population and targeting the inhibitory population. We will identify the resulting dynamical system (Wilson and Cowan 1972) with the concept of a cortical column:

$$\tau_e \frac{dA_e(t)}{dt} = -A_e(t) + h_e(w_{ee}A_e(t) - w_{ie}A_i(t) + I_e(t)) \quad (2)$$

$$\tau_i \frac{dA_i(t)}{dt} = -A_i(t) + h_i(w_{ei}A_e(t) - w_{ii}A_i(t) + I_i(t)). \quad (3)$$

At this point we would like to note that two very similar model classes exist in the literature. Their dynamics differ in the sense that in the first class, $A$ describes the activation or the firing rate of a population (in this case the nonlinearity in $h$ is applied to the total synaptic input, $A = -A + h(wA + I)$), while in the second class, $A$ denotes the membrane potential (in that case $h$ is applied directly to $A$, $\dot{A} = -A + w h(A + I)$. The reader of the original publications should not be confused, because both variants lead to qualitatively similar results, and are often equally well suited to tackle a specific modeling problem.

Layers. A neuronal layer may be described as a multitude of columns arranged in a topographically ordered space. This space may have a varying number of dimensions, for example some authors have used a one-dimensional chain representing the orientation preference axis, others identify a two-dimensional layer with the surface of the cortical tissue. With $\vec{x}, \vec{x'} \in C$ denoting positions within such a layer, the columns are coupled by appropriately chosen functions $W_{(e,i,e,e,i)}(\vec{x}, \vec{x'})$ (the so-called lateral couplings). Mathematically, the neuronal layer is described as a pair of coupled partial differential equations (Wilson and Cowan 1973)

$$\tau_e \frac{\partial A_e(\vec{x}, t)}{\partial t} = -A_e(\vec{x}, t) + h_e \left( w_{ee} \int_C A_e(\vec{x'}, t) W_{ee}(\vec{x}, \vec{x'}) d\vec{x'} - w_{ie} \int_C A_i(\vec{x'}, t) W_{ei}(\vec{x}, \vec{x'}) d\vec{x'} + I_e(\vec{x}, t) \right) \quad (4)$$
\[
\tau_i \frac{\partial A_i(\vec{x}, t)}{\partial t} = -A_i(\vec{x}, t) + h_i \left( w_{ei} \int_C A_e(\vec{x}', t) W_{ei}(\vec{x}, \vec{x}') d\vec{x}' \right) + \left( w_{ui} \int_C A_i(\vec{x}', t) W_{ui}(\vec{x}, \vec{x}') d\vec{x}' + I_i(\vec{x}, t) \right) .
\]

The inputs \( I_{(e,i)} \) are typically calculated by integrating a stimulus \( S(\vec{x}, t) \) over an afferent coupling function \( V_{(e,i)}(\vec{x}, \vec{x}') \)

\[
I_{(e,i)}(\vec{x}, t) = \int_R S(\vec{x}', t) V_{(e,i)}(\vec{x}, \vec{x}') d\vec{x}',
\]

where \( \vec{x}, \vec{x}' \) are elements of an input space \( R \). A convenient choice for the lateral as well as the afferent couplings are functions decaying with the distance between two populations, as has been revealed in anatomical and physiological studies. Choosing \( W \) satisfying \( W(\vec{x}, \vec{x}') = W(|\vec{x} - \vec{x}'|) \), the model becomes translationally and rotationally invariant. A commonly used prototype for these kernels is an \( n \)-dimensional Gaussian function defined as

\[
W_{(ee,ei,ie,ui)}(|\vec{x} - \vec{x}'|) = \frac{1}{(2\pi \sigma^2_{(e,i)})^n} \exp \left( -\frac{(\vec{x} - \vec{x}')^2}{2\sigma^2_{(e,i)}} \right).
\]

The computational advantage of these kernels is that the integration reduces to a multiplication in Fourier space, which speeds up computation time considerably.

For a large part of this article, we will assume that connections originating from inhibitory populations will be longer than those originating from excitatory populations. Following this scheme, the effective coupling between two columns will have the shape of a Mexican hat (difference of Gaussians). This assumption, which is often made in modeling studies, is questionable in so far as long-ranging patchy excitatory connections exist, at least in the mature primary visual cortex. This may not be a problem, because the layout of primary visual cortex revealed by the structure of the orientation preference map suggests that inhibitory couplings dominate at least over intermediate distances. Nevertheless, we will also discuss which different or additional phenomena are observed in the presence of long-range axons in chapter 5. For a more detailed introduction to the notion of neural layers see Layered computation in neural networks by Mallot.

3 Dynamical regimes and orientation preference

**Linear and marginally stable regimes.** Resulting from different choices of the system parameters in (4,5), almost all model variants exhibit two different dynamical behaviours: if the strength of the afferent input dominates over the lateral feedback, a homogeneous and constant input will lead to an activation pattern which is also spatially and temporally constant. This steady state is stable against noise. The parameter regime where this behaviour occurs is called the \textit{linear regime} (upper region in Fig. 1). As soon as the inhibitory feedback gets weaker, or the excitatory feedback gets stronger, the system enters a second regime called the \textit{marginally stable regime} (Ben-Yishai et al., 1995). Now the steady state is unstable and even the smallest perturbation leads to the emergence of a pattern of activation clusters commonly called \textit{blobs} (inset in central region of Fig. 1). The mechanism for this type of pattern formation is easy to understand; if the input at one position is slightly increased, this perturbation of the steady state will be amplified by
the dominating excitatory feedback, while the longer-ranging inhibition will suppress the activity in the surround of the emerging blob. For related reading, cf. Winner-take-all mechanisms by Yuille and Geiger.

**Lateral feedback and orientation selectivity.** The existence of a marginally stable regime could have consequences for the emergence of orientation selectivity in primary visual cortical neurons. Ben-Yishai et al. (1995) observed that the shape of the blobs remains invariant against different input levels. In a one-dimensional model of a cortical hypercolumn, where \( x \in C \) is identified with the orientation preference \( \Phi \), they demonstrated that the response behaviour of neurons to oriented gratings is accurately reproduced: the orientation tuning width remains largely invariant under changes of the stimulus amplitude (contrast). This finding indicates that the cortical dynamics may be dominated by lateral feedback rather than by feedforward excitation. A weak afferent orientation bias as emerging from a Hubel-and-Wiesel arrangement of LGN receptive fields would then suffice to induce a sharply tuned orientation tuning curve. This idea with it pros and cons, and also the experimental evidence for the origin of orientation tuning, is discussed in detail in the contribution of Shapley and McLaughlin (see Orientations selectivity: models and neural mechanisms).

4 Inhomogeneities and cortical maps

**Localisation of activation clusters.** In the marginally stable regime, each perturbation lays the seed for the emergence of an activation cluster. This perturbation could be induced by the afferent input but also by structural inhomogeneities in the model. For example, the lateral coupling function may not be perfectly translationally and rotationally invariant but could be subject to small random jitter. Then even a homogeneous and constant input will lead to the emergence of activation clusters. Preferentially, these clusters will be located at positions, where by chance the lateral excitatory feedback is slightly stronger than at other positions nearby. If this jitter comes together with inhomogeneous afferent inputs, both effects will add up and blobs will choose positions where the afferent input plus the lateral feedback will be largest. Note that in this model, the noise breaks the symmetry of the coupling kernels and the model will be no longer rotationally and translationally invariant.

**Instantaneous emergence of cortical maps.** Ernst et al. (2001) simulated a two-dimensional Wilson-and-Cowan model in which they put a small amount of static noise on the lateral coupling matrix, as can be expected in a biological system with all its irregularities. They presented moving gratings or bars as stimuli, generating an inhomogeneous afferent input. By recording the model’s response to differently oriented gratings (Fig. 2A), orientation and direction preference maps naturally emerge when the blobs localise at the spatial inhomogeneities in the model cortex (Fig. 2B).

This model has several advantages over other approaches of map development, because it reproduces seemingly controversial findings from experimental studies: first, the structure of the maps shows up within milliseconds and does not require any learning. Second, due to the intracortical origin of the map structure seeded by the random jitter of the lateral connections, the feature maps are identical for stimulation of either of the two eyes. Third,
the gratings induce an oscillatory movement of the blobs around their preferred positions which is different for opposite directions of movement. This suggests a new mechanism for directional selectivity of the neuronal response (for a detailed discussion on directional selectivity see the article Directional Selectivity by Grzywacz and Merwine). And finally, the model reproduces the known relationships between different kinds of feature maps. Taken together, these properties qualify this approach as a model for the initial phase in cortical development where the coarse layout of the maps is determined, which then could get subsequently refined and rearranged by self-organising mechanisms (see Swindale, 1996, for an extensive review and discussion).

5 Long-range connections and contour integration

Up to now, the couplings have been chosen as if there were no long-ranging excitatory connections in the brain. However, those connections exist and they link preferentially neurons having similar orientation preferences. What dynamical phenomena may one expect if these connections are included?

Long-range connections. Several authors employ a connection scheme which has locally the shape of a Mexican-hat, but extends over that region sending out additional excitatory connections targeting inhibitory and excitatory populations with a similar orientation preference. The columns in these models have a position \((x_1, x_2)\) within the nervous tissue and an orientation preference \(\Phi\), thus \(C \ni \vec{x} = (x_1, x_2, \Phi)\). While the response of the classical model without long-range interactions would follow the dynamics described in the previous sections, the addition of long-range connections opens the possibility that spatially extented stimuli modulate this response. The modulation will depend on the strength or contrast, and on the orientation of the stimuli presented. One important aspect of this excitatory modulation is that the net effect on the column’s firing rate depends on the activation of the target column, especially in cases where the populations have different thresholds or gains. The reason for this is that long-range input converges onto inhibitory and excitatory target populations, thus the excitatory target population receives direct excitation and indirect inhibition. The balance between those two sources determines if the total input inhibits or excites the target population (differential interaction).

Non-classical receptive fields. Models with long-range connections have been examined to find an explanation for the so-called non-classical receptive fields of neurons. Most visual cortical cells have shown dramatic changes in their response to a stimulus within their normal receptive field, when an additional stimulus has been presented outside that region (this additional stimulus alone would elicit no response). Typical phenomena include an increase in the response, if the two stimuli have orthogonal orientations, and a decrease, if the stimuli are parallel in orientation (Sillito et al. (1995)). The latter modulatory effect may change its sign when the stimuli are presented at a lower luminance level. These findings can largely be explained by population models with long-range interactions; in particular it is easy to explain the sign change with dynamical properties relying on the differential interaction scheme (see Stetter et al. (2000) and references
Association fields. Another type of cortical coupling function is motivated by association fields measured in psychophysical experiments. Association fields quantify how the presentation of a bar at position \((x_1, x_2)\) with orientation \(\Phi\) will increase or decrease the threshold to detect a bar at position \((x'_1, x'_2)\) with orientation \(\Phi'\). The coupling matrix and model dimension are similar to the models employed in the last paragraph, except for one important difference: the coupling function \(W\) is chosen according to the association field and therefore not only orientation-selective, but also directionally biased. In other words, two columns best responding to oriented bars being aligned in succession will be connected with a positive weight, while two columns best responding to oriented bars being aligned in parallel will be connected with a negative weight (Fig. 3A) or remain unconnected.

Contour integration. An aspect of cortical information processing, which can be examined and understood in this type of model, is the dynamics of contour integration. Contours can be interpreted as a succession of aligned bars, thus a coupling matrix based on the association field is especially suited to enhance the activity of columns stimulated by elements of the contour, whereas the activity of columns stimulated by distractors becomes suppressed. Li (1999, 2001) has accumulated plenty of evidence that contour integration may be explained by this kind of cortical models (Figs. 3B,C). A close relation of modeling work and psychophysical experiment shows that the structural simplicity of the Wilson-Cowan model class allows to make specific predictions about certain experiments while opening the door to an understanding of the mechanisms working behind the scene.

6 Transient dynamics and feature bindung

The modelling approaches discussed so far have focused on the long-term behavior of solutions of Wilson-and-Cowan-type equations. In particular, steady states of the system and their stability have been associated with phenomena of cortical physiology and psychophysics. In this section, we study the transient dynamics of coupled neural populations and link it to perceptual phenomena in the context of feature binding.

Feature inheritance and shine-through. The spatiotemporal behavior of the visual system can be assessed psychophysically through experiments where stimuli are presented successively for short time intervals. The visual system is thus forced to work at its spatial and temporal limit resulting in illusions which elucidate cortical mechanisms of signal processing.

Two such illusions have recently been described (Herzog and Koch, 2001). In the so-called feature inheritance effect (Fig. 4A), a single vernier — two bars which are slightly displaced — is presented for a brief time (i.e., 10–30 ms, depending on the individual performance of the subject). The vernier is followed by a double grating of five nondisplaced bars which is presented for 300 ms. Psychophysically, subjects are not aware of the vernier but perceive a displaced grating. That is, the vernier is masked by the grating which inherits
the vernier’s displacement. The inheritance effect has also been demonstrated for other features such as orientation and apparent motion (Herzog and Koch, 2001).

Changes in the geometrical arrangement of the grating can modify or even abolish the feature inheritance effect. An example is shown in Fig. 4B where an extended grating of 25 bars follows the presentation of the vernier. In this case, subjects are aware of the vernier which appears superimposed on the grating (shine-through effect).

Vernier visibility as a transient effect. The spatiotemporal version of the Wilson-Cowan equations (Eqs. 4,5) can be used to account for the vernier visibility in the different masking conditions. In order to elucidate the underlying neural mechanisms a simple, one-dimensional version without the property of orientation tuning is employed.

Consider an excitatory and an inhibitory population of cortical neurons arranged along a one-dimensional axis, \( C = \mathbb{R} \). The input space is also taken to be one-dimensional, \( R = \mathbb{R} \). For reasons of simplicity, we assume symmetry in the weights (\( w_{ee} = w_{ci} \); \( w_{ic} = w_{ii} \)) and in the interaction kernels (\( W_{ee} = W_{ci} = W_{ez} \); \( W_{ic} = W_{zi} = W_{zi} \)). The latter are modelled as Gaussians, cf. eq. (7). The external input current is identical for both populations, \( I_e(x,t) = I_i(x,t) \equiv I(x,t) \), and is given by a convolution of the presented spatiotemporal stimulus intensity \( S(x,t) \) with a Mexican-hat type of filter \( V(|x-x'|) \) whose integral vanishes,

\[
V(|x-x'|) = \frac{1}{\sqrt{2\pi \sigma_{e}^{2}}} \exp \left( -\frac{(x-x')^2}{2\sigma_{e}^2} \right) - \frac{1}{\sqrt{2\pi \sigma_{i}^{2}}} \exp \left( -\frac{(x-x')^2}{2\sigma_{i}^2} \right),
\]

resembling on-off receptive field properties of LGN neurons. The stimulus \( S(x,t) \) takes the value 1 if it is part of the vernier or a bar element, and 0 otherwise.

The system parameters – kernel widths, synaptic weights, population time constants, and gain functions – are adjusted considering symmetries and relations in cortical anatomy and physiology.

Numerical results for the feature-inheritance and shine-through conditions as described above are given in Fig. 4. The grayscale-coded activities of the excitatory populations show peaks at the position of the vernier and at the edges of the gratings, whereas almost no activity emerges in the bulk of the gratings. A comparison of the central peaks reveals that in the feature-inheritance condition (Fig. 4A) the vernier activity decays earlier than in the shine-through condition (Fig. 4B). This is due to a strong inhibition by the active neurons representing the nearby edges of the grating. However, if the extended grating comprised of 25 bars, the edges are too far away to exert an influence on the center. The fast suppression of the vernier activity by the small grating shown in Fig. 4A leads to a complete masking of the vernier element and a subsequent erroneous binding of its feature, the displacement, to the grating. On the other hand, conditions which allow a longer persistence of the vernier activity like the one in Fig. 4B result in a conscious perception of the vernier and its displacement. Thus, the occurrence of feature inheritance or shine-through is explained with the transient dynamics of a Wilson-and-Cowan type model.

The model can be applied to a number of further stimulus conditions and provides quantitative predictions for the visibility of the vernier element with a single set of model parameters. The model is robust with respect to parameter changes, and the overall results are the same no matter if the dynamical equations are formulated for the population firing rates or the average membrane potentials. In fact, Li was also able to see
the described transient behavior in her cortex model (Li, personal communication). The reduced one-dimensional model presented here also yields an analytical access and allows the identification of neural mechanisms responsible for the observed psychophysical effects (Herzog, Ernst & Eurich, submitted).

7 Discussion

The Wilson-and-Cowan model class yields a description of the behavior of coupled neural populations on a coarse timescale. Versions of the model include purely temporal behavior, spatio-temporal behavior in one and two spatial dimensions, and may incorporate further stimulus features such as the orientation of edges as additional model dimensions. The relatively simple structure of the equations allows for an analytical and thorough numerical access of the system dynamics. In the last years, the model class has been employed to successfully account for various physiological and psychophysical phenomena of the visual system such as orientation selectivity, cortical map formation, figure-ground segregation, feature binding, and masking effects.

Phenomena outside the visual system are beyond the scope of this review. The same holds for several dynamical aspects of the populations equations which have not been addressed; among these are hysteresis phenomena and limit cycle activity (Wilson and Cowan, 1972). For example, Tsodyks et al. (1997) have modelled oscillatory neural activity in rat hippocampus.

An important extension of the Wilson-and-Cowan model class is obtained if the simplification of time coarse graining is dropped. The search for appropriate equations describing the behavior of neural populations also on fast timescales and under the consideration of noise is a topic of much current interest. A suggestion which has been put forward in this context is the use of a Fokker-Planck equation; see the introduction by Gerstner (INTEGRATE-AND-FIRE NETWORKS) and Knight (2000) and references therein for a framework of a variety of such approaches.

Related reading

Directional selectivity; Integrate-and-fire networks; Layered computation in neural networks; Orientations selectivity: models and neural mechanisms; Visual cortex: anatomical structure and models of function; Winner-take-all mechanisms.
8 References


Figure captions

Figure 1: Phase diagram of the spatiotemporal Wilson-Cowan equations (4,5) for a two-dimensional sheet of neurons with threshold-linear $h$. Three dynamical regimes can be distinguished depending on the weights $w_e$ and $w_i$: the linear and the marginally stable regimes as described in the text, and a biologically implausible regime where neural firing rates diverge. Insets in the linear and marginally stable regimes show typical steady state activity pattern of the excitatory layer for a constant input $I$ plus a small amount of noise.

Figure 2: Orientation preference arises from the interaction of jitter in the neuronal connections with an oriented moving stimulus in a two-dimensional neural layer (i.e, $\vec{x} = (x_1, x_2)$). (A) Blob pattern emerging upon presentation of moving gratings, covering the whole input space $R$, having different orientations as shown by the bars. (B) Vectorial sum of the single condition blob patterns in (A) for different orientations coded in scales of grey. The picture strongly resembles orientation maps obtained experimentally with voltage-sensitive dyes (adapted from Ernst et al. (2001)).

Figure 3: (A) Coupling scheme $W$ connecting the center column at $\vec{x}$ with the surround columns at $\vec{x}'$ ($\vec{x} = (x_1, x_2, \Phi)$). Excitatory connections are marked with thin bars, inhibitory connections with broken bars. The orientation of the bars denotes the difference in orientation preference of the connected columns. (B) and (C) show sample stimuli $S(\vec{y})$, $\vec{y} \in R$, and the corresponding activation pattern $A(\vec{x})$, $\vec{x} \in C$. The activation level and the orientation preference of the corresponding column are coded by the thickness and the orientation of the bar, respectively. The connection scheme together with the model’s dynamics lead (B) to the enhancement of (orientation) discontinuities (edge detection), and (C) to an increase in the activity of detectors stimulated by a closed contour (contour integration). (Figs. adopted from Li (1999)).

Figure 4: Stimuli, visual percept and simulation result for the feature inheritance and shine-through effects. In the feature inheritance condition (A), a single vernier is followed by a grating of five bars (right panels). The percept is a displaced grating (top). The simulation shows the activity of the excitatory population in greyscale-code. The central peak resulting from the vernier is rapidly suppressed by the edge activity of the grating. In the shine-through condition (B), the vernier is followed by an elongated grating of 25 bars (right). Perceptually, the vernier looks superimposed on the grating (top). The simulation of the excitatory population reveals that in this case, the central vernier activity persists for a longer time leading to the conscious perception of the shine-through element.
Figures

Figure 1: Ernst and Eurich, Figure 1

Figure 2: Ernst and Eurich, Figure 2
Figure 3: Ernst and Eriuch, Figure 3

Figure 4: Ernst and Eriuch, Figure 4