Journal of Statistical Physics (2020) 180:612–629
https://doi.org/10.1007/s10955-019-02483-1

Towards a Mathematical Model of the Brain

Lai-Sang Young[1,2]

Received: 5 November 2019 / Accepted: 31 December 2019 / Published online: 16 January 2020

Abstract
This article presents an idealized mathematical model of the cerebral cortex, focusing on the dynamical interaction of neurons. The author proposes a network architecture more consistent with neuroanatomy than in previous studies, borrows ideas from nonequilibrium statistical mechanics and calls attention to the fact that the brain is a large and complex dynamical system. The ideas proposed are illustrated with a realistic model of the visual cortex.

Keywords Cortical dynamics · Neuronal interaction · Nonequilibrium steady states · Visual cortex

The aim of this paper is to put into a mathematical framework some ideas related to the dynamics of the cerebral cortex, to be thought of as a large network of spiking neurons. It is a product of my attempt to conceptualize, to build an idealized picture of cortical dynamics that is as consistent as possible with neurobiology, and to make contact with mathematical physics at the same time.

Many before me have shared their insights and have made invaluable contributions to theoretical neuroscience, beginning with the ground-breaking work of Hodgkin–Huxley more than half a century ago modeling the ionic mechanisms underlying the initiation and propagation of action potentials in single neurons [23]. Very influential also are the mean-field approach to network dynamics proposed by Wilson and Cowan in the 1970s [38], and the much cited balanced state idea for neuronal populations [37]. Mathematical techniques and ideas introduced more recently include the approximation of neuronal dynamics in sparsely coupled networks by stochastic differential equations [7], weak-coupling limits [11], neural fields [5], statistical mechanics formalisms [9] and normalization [21], to mention just a small sample of the advances in the literature. Each author brings a different perspective, and theoretical neuroscience is made richer by these diverse views.

Communicated by Ivan Corwin.

This research is partially supported by NSF Grants 1734854 and 1901009.

✉ Lai-Sang Young
lsy@cims.nyu.edu

[1] Courant Institute of Mathematical Sciences, New York University, New York, NY 10012, USA

[2] School of Mathematics and School of Natural Sciences, Institute for Advanced Study, Princeton, NJ 08540, USA

As to how this paper might add to existing ideas, I would like to think that

(1) the network structure proposed here bears a closer resemblance to real neuroanatomy than in most previous theories;
(2) I will point to a natural analogy with nonequilibrium statistical mechanics that to my knowledge is novel, and
(3) dynamical interaction of neurons including the competition between Excitatory and Inhibitory populations is given the attention that it deserves.

I will propose to decompose the cortical network into a collection $\{\mathcal{L}_{i,j}\}$ of subnetworks corresponding to the *layers of individual cortical regions*, $\mathcal{L}_{i,j}$ being the jth layer of the ith cortical region. These are the basic units in our decomposition. Each $\mathcal{L}_{i,j}$ is a 2D network of neurons, comprised of local circuits possibly with longer-range connections. These networks have spatially homogeneous structures, a ready-made setup for statistical mechanics. As connectivity is densest within local circuits, this is where dynamical interactions of neurons are the most intense, making them natural playgrounds for studying neuronal dynamics. The collection $\{\mathcal{L}_{i,j}\}$ is interconnected by location-specific connections, i.e., neurons in a layer project to specific areas in other layers modeling the inter-laminar and inter-areal connections in real cortex.

Throughout the paper, I will use results from a previously built model of the visual cortex to illustrate the ideas proposed, so let me give some information on the nature of this model. For a number of years now I have been part of a team to build a large-scale computational model of the monkey visual cortex, which is very similar to human visual cortex. As our aim is to shed light on how real cortex works, biological realism is essential, and we have used dozens of sets of experimental data to guide our model construction and parameter selection. Ultimately we would like to have a mechanistic model with the capability to reproduce all of the basic phenomena of the primary visual cortex, and we are quite close to achieving that for the input layer of the magnocellular pathway; see [14,16]. This modeling work has shaped much of my thinking on the subject.

The account given in this paper is, I believe, a reasonably accurate depiction of the visual cortex given the level of abstraction, and the validity of this description extends, to varying degrees, beyond the visual cortex to other regions of cortex, which are known to have similarities and also differences with the sensory cortices. Because of the relative availability of data, the visual cortex serves as a window into the rest of cortex, about much of which the science is not sufficiently far along and I am not sufficiently knowledgeable. So when I write about "the brain" or "cortex" in the rest of this paper, the picture is an extrapolated one and should be understood as such.

This paper is organized as follows: Sect. 1 contains some neurobiological background, which we provide to justify the cortical structure proposed. This sets the stage for Sect. 2, which focuses on the dynamics within local circuits, and for Sect. 3, where analogies with nonequilibrium statistical mechanics are proposed. The main points are summarized in Sect. 4.

1 An Anatomy-Based Network Architecture

A number of theoretical neuroscience papers have studied network models of randomly connected Excitatory and Inhibitory neurons defined by, for example, Erdös–Renyi graphs. In these models, the entire population is statistically homogeneous, both in terms of network connections and input, and system size is often taken to infinity for mathematical tractability

[7,37]. Such models have contributed to the theoretical understanding of neuronal populations, but as all would agree, the cerebral cortex is highly inhomogeneous with different regions responsible for highly differentiated functions; it cannot be represented by networks with such homogeneous features.

The next level up is to model cortex as a network of networks, i.e., as a network the nodes of which are themselves homogeneous networks as described in the last paragraph; see e.g. [26]. More general still are networks with a hierarchical structure, with homogeneously connected networks appearing at the bottom level. Such models recognize the distinction between brain regions, as well as the further subdivision into subregions, but I will argue below that this is still not the most natural way to capture the characteristic anatomical structure within the cerebral cortex.

1.1 A Little Bit of Neuroanatomy

For the convenience of the reader, we include here a very rough account of the structure of the cerebral cortex. The reader is invited to consult standard neuroscience texts for more information.

The cerebral cortex is a thin sheet of neural tissue 2–3 mm in thickness, folded underneath the skull. For simplicity we think of it as a vast network of neurons; in reality, there is more to it than that. Neuroscientists have divided the area of this surface into many labeled regions, such as the sensory cortices corresponding to the different senses (e.g. visual, auditory, olfaction, touch), the motor cortex, which is involved in the planning, control and execution of voluntary movements, the prefrontal cortex, which is generally associated with planning complex cognitive behavior, decision making, and so on. These cortices can be thought of as subnetworks, interconnected in complicated ways; see [3].

Each of the cortices in the last paragraph is typically further subdivided into anatomically and functionally distinct cortical regions. For example, the visual cortex is divided into V1, V2, V3, V4 and V5 (also known as MT). Visual signals received through the eye enter the visual cortex via V1; the information is passed from there to V2 and MT. These five subregions of the visual cortex are interconnected with feedforward and feedback loops, with V4 (associated with geometric shapes) and MT (responsible for motion) having the most interareal connections, i.e. having the most connections with cortical regions outside of the visual cortex; see [35,36].

Connectivity within subregions like V1, . . ., V5 is still quite complicated, nor is there a natural way to further subdivide the area of e.g. V1 into a finite number of structurally simpler subregions. There is, however, a third dimension: Though cortex is largely 2D, this thin neural tissue in fact has 6 layers (in reality more because of further subdivisions within layers). Each layer has its own characteristic distribution of neurons the types and connection patterns of which vary from layer to layer. These subnetworks are not disjoint, but most of the connections to a neuron are lateral or intra-laminar, the axons of presynaptic neurons extending roughly horizontally in a 2D-like structure.

Zooming in on a single layer within a cortical region, one sees a continuum of local circuits: each neuron is embedded in a relatively densely connected local circuit of Excitatory and Inhibitory neurons. Depending on layer, there may or may not be longer-range connections, connections that can be several times longer than those in local circuits but still localized relative to the area of the entire layer. A salient feature of each cortical layer is the spatial homogeneity of its structures, meaning at two different locations of the layer, the surrounding local structures look similar.

In addition to connections from within the same layer, there are interlaminar connections within a cortical region, mostly between neurons located roughly "above" or "below" one another with some amount of convergence and divergence, i.e., each neuron synapses on groups of neurons in its target layers and receives signals from groups of neurons, often from multiple source layers. Some layers project outside of the cortical region; some do not. Interareal connections, like interlaminar ones, are primarily excitatory.

This completes our very brief review of neuroanatomy.

1.2 Proposed Idealization

We propose to organize the cortical network along the lines described in Sect. 1.1. That is, we follow the conventional subdivision by neuroscientists of the 2D area of cortex into regions, subregions, subsubregions and so on, according to function and anatomy, and let $\{\mathcal{R}_1, \mathcal{R}_2, \ldots, \mathcal{R}_n\}$ denote the resulting partition of the 2D cortical sheet. Each $\mathcal{R}_i$ is further subdivided into anatomically distinct layers $\{\mathcal{L}_{i,j}\}$, with $\mathcal{L}_{i,j}$ denoting the jth layer in cortical region $\mathcal{R}_i$. The collection of $\{\mathcal{L}_{i,j}\}$ will be the basic units of our decomposition.

Identifying a piece of cortex with the neuronal network it contains, we have described a coarse decomposition of the (global) cortical network into subnetworks $\{\mathcal{R}_i\}$, which are further decomposed into subsubnetworks $\{\mathcal{L}_{i,j}\}$. There are many connections among the $\{\mathcal{L}_{i,j}\}$. Connections between $\mathcal{L}_{i,j}$ and $\mathcal{L}_{i',j'}$ with $i = i'$ and $j \neq j'$, or with $i \neq i'$, correspond to interlaminar and interareal connections. These connections are mostly Excitatory, and they tend to be location-specific, i.e., specific groups of neurons in $\mathcal{L}_{i,j}$ projects to specific regions of $\mathcal{L}_{i',j'}$.

Each $\mathcal{L}_{i,j}$ is populated by a continuum of local circuits of Excitatory and Inhibitory neurons, with possibly longer-range connections. The connection structures within $\mathcal{L}_{i,j}$ are oriented in a sufficiently "horizontal" manner that it is not unreasonable to idealize $\mathcal{L}_{i,j}$ as a 2D network. These 2D networks are, needless to say, not infinite in extent, but they are much larger than their local structures, lending legitimacy to thinking of them as "infinitely large". Spatial structural homogeneity, referring to the similarity in spatial structures at different locations within each $\mathcal{L}_{i,j}$, corresponds in our mathematical idealization to *translational invariance*.

We remark that the $\mathcal{L}_{i,j}$ defined above are very different from the homogeneously connected and homogeneously driven networks studied in e.g. [7,37], and results from these earlier studies do not apply. Each $\mathcal{L}_{i,j}$ has a 2D geometry, a notion of distance, and most of the connections within the network are local or not much longer than local. Inputs received by each $\mathcal{L}_{i,j}$ are far from spatially homogeneous: connections among the different $\mathcal{L}_{i,j}$ are not spatially homogeneous as we have discussed, and external inputs are typically far from spatially homogeneous as we will see. The situation is very different, and we will argue in the pages to follow that this is a more realistic depiction of cortex.

The rest of this paper is focused on the 2D networks $\mathcal{L}_{i,j}$. We will explain cortical dynamics by identifying their similarities and differences with concepts and techniques from mathematical physics. Connections among the different $\mathcal{L}_{i,j}$ belong in the domain of neuroanatomy; we will not have more to say about that. We propose to analyze the dynamics within each $\mathcal{L}_{i,j}$ in two separate steps.

Dynamics of local populations Though local circuits in real cortical layers merge into one another and cannot be isolated, we will, in a further simplification, study them as isolated groups of homogeneously connected Excitatory and Inhibitory neurons. The is similar in spirit to previous studies except that we will avoid making assumptions for mathematical tractability

that are inconsistent with neuroanatomy, and we place greater emphasis on the dynamics of competition between Excitatory and Inhibitory neurons and on emergent behaviors. This is discussed in Sect. 2.

Dynamics across a cortical layer In Sect. 3, the focus is on the dynamics on a larger scale, across the 2D surfaces $\mathcal{L}_{i,j}$. Our main goals here are (i) to demonstrate the similarity in setup with nonequilibrium statistical mechanics using for illustration a realistic model of an input layer of the monkey primary visual cortex, and (ii) to explain how ideas from statistical mechanics can be leveraged to obtain a systematic understanding of the "macroscopic" observations in neuroscience.

2 Dynamics of Local Populations

The setting of this section is that of an isolated network of integrate-and-fire neurons, intended to model local circuits in the cerebral cortex. A description of the model is given in Sect. 2.1. The concept of dynamic equilibrium is discussed in Sect. 2.2, and two different aspects of the steady state dynamics of this system are illustrated in Sects. 2.3 and 2.4.

2.1 Model Description

We describe here an idealized model of a network of integrate-and-fire neurons. The material in Sects. 2.1–2.3 applies in principle to networks of arbitrary sizes and connectivities. For readers who prefer more definite contexts, we have in mind networks that resemble local cortical circuits, which typically consist of a few hundred to a couple thousand neurons. We assume that about 75% of the neurons are Excitatory (E) and 25% are Inhibitory (I). Following neurophysiology, E-to-E connectivity is relatively sparse, at $\sim$ 15%, while E-to-I, I-to-E, and I-to-I, are more densely connected, at $\sim$ 60% [24,31]. Connectivity between specific pairs of neurons is determined randomly, following the probabilities above. The network receives external input which are Excitatory only and which we assume for simplicity to be Poisson and independent from neuron to neuron.

The membrane potential dynamics of individual neurons are assumed to obey the leaky integrate-and-fire (LIF) equation, which we recall: Let v denote the membrane potential of a neuron. We put v in normalized units (following [29]), setting the rest potential and spiking threshold to be $v = 0$ and $v = 1$ respectively, i.e., the neuron spikes when v reaches 1, and is reset to $v = 0$ following a spike, where v rests for 2–3 ms before resuming its evolution described by the equation

$$\frac{dv}{dt} = -g_R v - g_E(t)(v - V_E) - g_I(t)(v - V_I)\,. \tag{1}$$

Here $g_R = 50\ \mathrm{s}^{-1}$ is the leakage conductance, $V_E = 14/3$ and $V_I = -2/3$ are the Excitatory and Inhibitory reversal potentials, following textbook values [27]. The time-dependent quantities $g_E(t)$ and $g_I(t)$ are the E and I-conductances of the neuron, the dynamic equations of which are given below. The quantities $g_E(t)(v - V_E)$ and $g_I(t)(v - V_I)$ are called the Excitatory and Inhibitory currents respectively.

For an E-neuron, the evolution of its E-conductance is given by

$$g_E(t) = S^{EE} \sum G_E(t - t^E_{\text{spike}}) \tag{2}$$

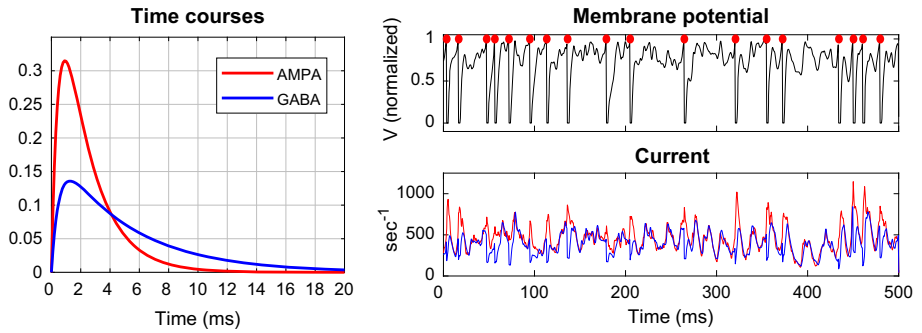

Fig. 1 Left: The functions $G_E(t)$ (red) and $G_I(t)$ (blue) in the definition of E and I-conductances. When an Excitatory spike is received, the conductance $g_E(t)$ of the postsynaptic neuron is elevated for several milliseconds. The same happens to $g_I(t)$ when an I-spike is received. The durations of elevation for AMPA and GABA follow known data. Right top: Time trace of membrane potential (in normalized units) of a sample E-neuron. Red dot indicates a spike is fired at that moment in time. Right bottom: E (red) and I (blue)-currents entering the same neuron. Often though not always, a spike is fired when the E-current exceeds the I-current by a sufficiently large amount (Color figure online)

where S^{EE} is the synaptic coupling weight from E-to-E, $G_E(s)$ is the normalized excitatory postsynaptic conductance (EPSC) function shown in Fig. 1, left; t^E_{spike} represents the points in time an E-spike is received by the neuron in question, and the sum is taken over the arrival times of all incoming E-spikes. These spikes can come from other E-neurons from within the network, or from the external Poisson drive. Every time an Excitatory spike is received, the E- conductance of the neuron is elevated for a few milliseconds (ms) following the curve in Fig. 1. Likewise, the I-conductance of this neuron $g_I(t)$ is described by

$$g_I(t) = S^{EI} \sum G_I(t - t^I_{\text{spike}}) \tag{3}$$

where S^{EI} is the synaptic coupling weight from I-to-E and $G_I(s)$ is the IPSC function in Fig. 1, left. Here t^I_{spike} represents the times at which Inhibitory spikes are received by the neuron in question, and I-spikes can come only from I-neurons in the local population.

If the neuron in question is of Inhibitory type, then the description above holds verbatim except that one must replace S^{EE} by S^{IE} and S^{EI} by S^{II}.

The 6 main parameters of the model, therefore, are the four synaptic coupling weights S^{EE}, S^{EI}, S^{IE}, S^{II}, together with dr_E and dr_I, the rates of external drive for E and I neurons in the population. Exact values of synaptic coupling weights are unimportant in the present discussion, but as was indicated earlier, we have in mind synaptic couplings that produce firing rates consistent with those from real cortex; see e.g. [14], Methods. To simulate external drive, we simply increase the values of dr_E and dr_I above background levels.

This completes a description of the basic version of our model. The setup is entirely standard; no novelty is claimed.

We finish with the following two remarks.

The first is that simple as Eqs. (1)–(3) are, there is no known way to express the firing rates of the system in terms of the parameters above. This is because in a normally functioning network, each neuron receives a steady stream of E and I-spikes, producing fluctuating E and I-currents the balance of which drives the spike firing of the neuron; see Fig. 1, right. Each spike fired by a neuron contributes, in turn, to modifying the conductances of its postsynaptic

neurons. The complexity of the dynamics comes not from the LIF equation governing the evolution of individual neurons but from the recurrent coupling relationship among neurons.

A second remark, directed mostly at the theoretical neuroscience community, is that our models are not the same as those studied in [6,7]. The theory described in [6,7] is predicated on a sparse-coupling assumption, requiring that if N denotes system size, then (i) $N \to \infty$, and (ii) (number of connections per neuron)/$N \to 0$ as $N \to \infty$. These assumptions imply in particular that the sets of presynaptic neurons of two randomly chosen neurons in the population are largely independent. We do not make assumptions of this kind. We do not exclude them, but they are not what we have in mind, as local circuits are generally not that large [2] and most cortical networks are not sparsely coupled; see [24,31].

2.2 E-I Interaction and Dynamic Equilibria

From the discussion above, we see that the spike firing of a neuron is governed by two competing forces: excitation coming from the spiking of E-cells in the network and from external drive, and inhibition from the spiking of I-cells. These two forces push the membrane potential of the neuron in opposite directions, causing the neuron to spike when excitation has a sufficiently large (though temporary) edge, and to be quiescent when inhibition rules; see Fig. 1, right.

From the point of view of a single neuron, the mechanisms are simple enough. The dynamical picture is less clear on the level of population activity, because when E-cells spike, they excite both the Excitatory and Inhibitory neurons postsynaptic to them; likewise, the spiking of I-cells suppresses both E and I-cells. Consider, as an example, the effects of increasing excitatory input to I-cells. This is known to cause I-firing rate to increase, and E-firing rate to decrease, but under certain circumstances it is also known to lower the firing rates of both E and I-cells, the decreased spiking of I-cells perhaps due to the loss of Excitatory input from the E-population. The dynamical interaction between the E and I-populations is complex and nuanced. It is a wide open area of research about which much remains to be understood.

Given a network of neurons driven by homogeneous Poisson processes, the system will, under "most" conditions, tend to a unique stationary state described by an invariant probability measure. There are exceptions to this, such as when the connectivity permits multiple steady states with no way to transition from one to another. Let us not consider these "pathological" situations.

In a stationary state, we think of the Excitatory and Inhibitory forces as having negotiated a *dynamic equilibrium*, a state of the system in which the opposing actions of the E and I-populations produce recurrent patterns of dynamics that remain statistically unchanged over time. One of the most fundamental theoretical questions, then, is to understand the properties of these dynamic equilibria.

2.3 Firing Rates and Currents in Dynamic Equilibria

Here we focus on two of the most important observables, firing rates and currents, and show how mean-field ideas can be used to gain insight into these quantities when the system is in a dynamic equilibrium.

Estimation of firing rates While there is no explicit way to express firing rates in terms of model parameters, it does not mean they cannot be estimated. Let f_E and f_I denote the firing

rates of the network in dynamic equilibrium. Fixing $T > 0$ and integrating the left side of Eq (1) from time 0 to time T without resetting v to 0 after each spike, we obtain the number of spikes fired on the time interval $[0, T]$. Integrating both sides of the equation and neglecting the leak term, we obtain

$$\#\text{ spikes fired } \approx \text{ total E-current } - \text{ total I-current.}$$

Total E-current here refers to the total E-current received by the neuron in question on the time interval $[0, T]$, counting both current from within the network and from external drive. I-current comes only from within the network. Here as in the discussion to follow it is convenient to speak of I-current as though it was a positive quantity; we are really referring to the magnitude of the current, hence the negative sign in the formula above. From this formula, it follows that f_E and f_I can be estimated as follows:

$$f_E \approx (N^{EE} f_E + \mathrm{dr}_E) S^{EE} \times 4 - N^{EI} f_I S^{EI} \times 1.33 \tag{4}$$

where N^{XY} is the mean number of presynaptic Y-cells for neurons of X-type. In the first term on the right side, $N^{EE} f_E + \mathrm{dr}_E$ is the average number of E-spikes received on the time interval in question, $N^{EE} f_E$ of them coming from the network and the rest from external drive. Assuming (as a rough approximation) that mean $v = 2/3$, we have replaced $V_E - v$ by its mean of $14/3 - 2/3 = 4$ and $v - V_I$ by $2/3 - (-2/3) = 4/3$. The corresponding equation for f_I is

$$f_I \approx (N^{IE} f_E + \mathrm{dr}_I) S^{IE} \times 4 - N^{II} f_I S^{II} \times 1.33\,. \tag{5}$$

Solving Equations 4 and 5 give estimates of f_E and f_I.

Similar ideas were used in [28], which verified that values of f_E and f_I found by solving these linear equations give, for a good range of examples, reasonable approximations of the mean firing rates of the system compared to simulated values. Mean-field ideas along these lines were introduced by [38] much earlier.

A source of error in the estimation above is that it does not take into consideration spiking patterns, which can have an effect on firing rate. For example, E-spikes that arrive in a concentrated fashion are more likely to produce spikes in the postsynaptic neuron than the same number of spikes spread out evenly in time, because the leak term (which was neglected in the estimation above) causes a fraction of the current to dissipate before it produces a spike. We will see a concrete example of this in Sect. 3. More generally, the estimation above assumes that the system is mean-driven. Fluctuation-driven systems can produce nonlinear effects not captured by the linear equations above.

Balances and imbalances of E and I-currents Using the ideas above, we now make a connection to the balanced-state theory of [37], which asserts that in homogeneously connected networks, E and I-currents are roughly equal as system size becomes infinitely large. That is, if C_E and C_I denote the E and I-currents entering a neuron, then the idea of balanced states suggests that C_E and C_I should be comparable in magnitude.

Broadly interpreted, this idea was shown to be valid in realistic models of V1 (e.g. [14]), where local circuits are not infinitely large: it was found that

$$C_E - C_I \ll C_E,\ C_I \tag{6}$$

where "$\ll$" means the quantity on the left is $10 - 30\%$ those on the right depending on circumstance.

Important as (6) is, we point out that it is not just the balancing of E and I-currents that is of interest; the degree of their *imbalance*, i.e., $C_E - C_I$, is equally if not more important, as that

controls the firing rate of the postsynaptic neuron as explained in the firing rates discussion above. Indeed the discussion in the first half of Sect. 2.3 offers the following clarification with regard to balanced state ideas:

1. Keeping firing rates, hence $C_E - C_I$, roughly constant, the larger the E and I-currents, the closer C_E/C_I is to 1. Larger currents require larger numbers of presynaptic neurons, hence larger system size. It is however not system size *per se* but the numbers of presynaptic neurons that count.
2. Fixing the network (including the numbers of presynaptic neurons), the larger the firing rate, the greater the imbalance, i.e., the larger the gap $C_E - C_I$ between the two currents.

These statements are valid for systems of all sizes.

We finish with the following observation in relation to condition (6). When $C_E - C_I \ll C_E, C_I$, an increase in external drive to E-cells (or an increase in drive biased toward E-cells) that constitutes only a small percentage change in E-current can cause $C_E - C_I$ to increase by a large percentage, hence a significant increase in firing rate. In this sense condition (6) is related to the concept of *high gain*.

2.4 An Emergent Phenomenon: Gamma-Band Rhythms

Dynamic equilibria are characterized not just by firing rates but by spiking patterns and subthreshold activity. In this section we will describe an emergent phenomenon that has been observed ubiquitously in the real brain, namely gamma-band rhythms. A network phenomenon is called *emergent* if it occurs as a result of the dynamical interaction among neurons and is not apparent or not easily deduced from properties of individual neurons.

The spiking of individual neurons is well known to be irregular. Interspike intervals (ISI) are known to have long exponential [30] or heavier tails [1]. See Fig. 2, left. One way to understand that heuristically is to approximate the dynamics of the membrane potential v by a random walk with a drift, so that the first passage time, the time it takes such a walk to go from 0 to 1, is described by an inverse Gaussian [19]. This fact about the spiking behavior of individual neurons is well documented both in real cortex and in models.

Collectively, neuronal dynamics present a quite different picture: Neuronal populations in real cortex are known to produce irregular oscillations with wandering frequencies in the gamma band (30–90 Hz). This phenomenon has been captured and analyzed using spectral power densities by experimentalists [20,22]. The rhythm is broad-band, even episodic [10, 12], yet it is clearly detectable, and it occurs in many parts of the brain, prompting one to ask if it is there to serve any specific purpose or if it is simply a by-product of neuronal processes.

Gamma-band activity has been a challenge to theorists. The first model, due to Kopell and known as PING, explains this rhythm in terms of an E-population-spike (the entire E-population firing synchronously) followed by an I-population-spike which suppresses all for a certain period of time, with the E-population-spike to be repeated when the inhibition wears off. This mechanism produces highly regular oscillations [4]. Others proposed adding noise terms to oscillatory motion e.g. [8]. These early models successfully captured the oscillatory aspect of the phenomenon but failed to capture the broad-band, episodic nature of gamma-band activity.

The mechanism proposed by Rangan and Young in [32], and subsequently used in [13,15, 33], produced more realistic results. The mechanism, called *Recurrent Excitation-Inhibition* (REI) can be described as follows: Gamma-band activity occurs when the system is under drive. From the crossing of threshold by a few E-cells, recurrent excitation leads to the

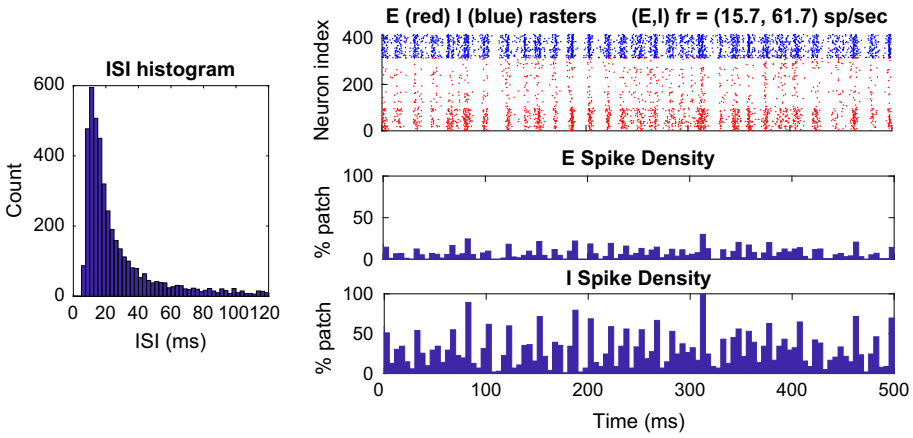

Fig. 2 Left: Interspike intervals for a typical neuron. Right top: Rasters of spike firing in a local population under drive: a dot is plotted every time a spike is fired. The x-axis is time, and the y-axis is neuron number. E-neurons (red), I-neurons (blue). Note the presence of an irregular but clearly detectable rhythm. This rhythm represents emergent collective behavior, to be contrasted with the times between spikes for a single neuron shown on the left. Right bottom: Percentage of E and I neurons spiking in 5 ms bins. The fraction of E-cells spiking together in a 5 ms bin is usually no more than 10% (Color figure online)

elevation of subthreshold activity and the spiking of more E and I-cells. We stress that when an E-cell spikes, it excites both E and I-cells postsynaptic to it, so both populations are activated simultaneously. Depending on how the membrane potentials of neurons postsynaptic to the spiking cells are positioned, I-cells can step in to stop this developing event quickly, or the recruitment of more cells to join the event continues for a millisecond or two longer before the suppressive effects of I-spikes are felt. Since I-neurons are fairly densely connected to the rest of the local population, this pushback tends to hyperpolarize a good majority of the cells in the system. The decay of the suppressive effect, and the depolarization of E-cells, helped along by the external drive, allows the scenario to be repeated once again. This explains the production of a rhythm. Time constants in the LIF equations cause the cycles to have frequencies in the gamma band.

In the real brain, the phenomenon described above is produced in local neuronal populations. For models of the type described in Sect. 2.1, a similar rhythm emerges provided the E and I-forces are sufficiently balanced so neither one dominates. We stress that in REI, synchronization is very partial, usually involving only a small fraction of the population (see Fig. 2, right). The size of the spiking event, i.e., the number of neurons participating, is variable from beat to beat depending on the membrane potentials of neurons postsynaptic to the ones crossing threshold, as are the gaps between spiking events: larger events, for example, tend to recruit more I-cells resulting in a stronger hyperpolarization of most neurons in the network and a longer time gap before the next event. These characteristics emulate the episodic nature of data on gamma-band activity better than previous models.

This brings us to the end of Sect. 2. Models of isolated neuronal populations can be as simple or as sophisticated as one wishes. We have considered a very minimal version, in part to convey the fact that even for models of such simplicity—describable by a handful of parameters—the dynamic equilibria "negotiated" by the Excitatory and Inhibitory populations have rich and varied properties including emergent dynamics. It has been our experience that on the local circuit level, mean-field computations offer useful baseline

information provided "corrections" are made for phenomena not captured by these estimates.

3 Dynamics Across a Cortical Layer

Having focused on local circuits in the last section, we now turn our attention to entire layers of cortical regions, the $\mathcal{L}_{i,j}$-networks introduced in Sect. 1.2. When stimulated, the dynamics on these extended 2D networks are highly inhomogeneous. We will propose a way to conceptualize them borrowing ideas from nonequilibrium statistical mechanics. A realistic model of the visual cortex is used both to illustrate and to justify the ideas proposed; relevant facts about this model are recalled in Sect. 3.1. Section 3.2 contains the main proposal of this paper.

3.1 Monkey V1 Cortex: Example of a Cortical Network

We describe here an example of a cortical network corresponding to a single layer of a specific region of the cerebral cortex. The animal in question is the macaque monkey, the cortical region is the primary visual cortex V1, and the layer is 4Cα, the input layer in the magno-pathway (one of the two major pathways) to V1. The material below is mostly taken from [14], which modeled a small patch of this layer located at about 5° eccentricity. The same structure (with varying cortical magnification factor) can be assumed to hold in a much larger surface of this layer.

Below are a few key features of this network relevant for the discussion to follow.

1. One of the most important features of V1 is that it has a retinotopic map. That is, there is a natural mapping from the monkey's 2D visual field to its retinal surface, and an almost point-to-point correspondence between retinal and cortical locations. Peripheral vision excepted, the feedforward input received by a local population in Layer 4Cα comes from a very small region of the monkey's visual field, and as cortical location varies, the corresponding region of visual field varies continuously with it.
2. Another important function of V1 is orientation selectivity (OS). This means most neurons in V1 have a preferred orientation, in the sense that when an "edge" or contour in the neuron's receptive field is aligned with its preferred orientation, the neuron is excited and its firing rate is elevated, and when the edge is orthogonal to its preferred orientation, its firing rate barely rises above the spontaneous level. Now neurons with similar receptive fields and similar orientation preferences tend to be located close to one another on the cortical surface, so that the cortical network can be thought of as divided into orientation domains.

Figure 3 (left) depicts the structures above in the model in [14]: Each of the 9 square represents a "hypercolumn" consisting of $\sim$ 4000 neurons with overlapping receptive fields. Orientation domains are organized around a pinwheel center within each hypercolumn; their intended orientation preferences are indicated by the triple bars. These structures have been mapped experimentally using voltage-sensitive dyes or optimal imaging; see e.g. [34].

3. Each neuron in Layer 4Cα receives feedforward input from the lateral geniculate nucleus (LGN), lateral input from within the same layer, and input from Layer 6 of V1 which we think of as feedback. In our model, 60–70% of the excitatory input to a neuron in Layer 4Cα comes from lateral interaction, about 20% comes from feedback, LGN comprises

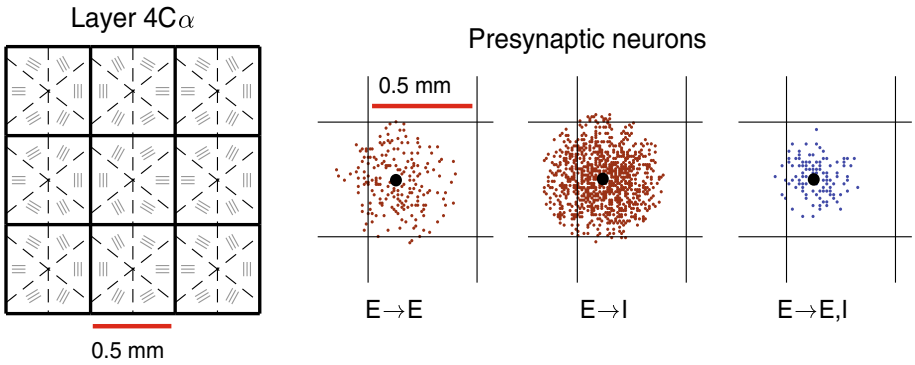

Fig. 3 Left: The 9-hypercolumn model of Layer 4Cα of the monkey V1 in [14]. Domains with intended orientation preferences are organized around pinwheel centers, "intended" in the sense that neurons in these domains receive input from LGN cells with certain spatial alignments. Right: Sets of neurons presynaptic to E or I-cells (black, at center of cloud); squares are shown to indicate the sizes of hypercolumns

a little over 10%, and the rest comes from all other modulatory influences combined. These numbers are believed to be reasonably realistic.

4. In the model, connections within the layer can be summarized as follows: Each E-neuron in Layer 4Cα synapses on neurons within a region roughly the size of a hypercolumn; I-neurons have a shorter range, a little more than half of that of E-neurons; see Fig. 3, right. Neurons postsynaptic to a cell are roughly isotropic; they do not respect boundaries of orientation domains or hypercolumns. The probability of connection between two neurons falls off with distance, its density having roughly the shape of a truncated Gaussian. E-to-E connectivity is more sparse, assumed to be $\sim$ 15% at the center, while E-to-I, I-to-E and I-to-I connections are denser with peak connectivity 50–60%. These numbers are taken from known neuroanatomy; we refer the reader to [14] for details and references. These are the numbers we tried to emulate in the local population models in Sect. 2.
5. As to how OS (see Item 2 above) comes about, when a visual stimulus is presented to the eye, the signal is passed from the retina to LGN to V1. At least in primates, retinal and LGN cells have no orientation preferences. Hubel and Wiesel [25] proposed half a century ago that OS of V1 cells is derived from the spatial alignment of its LGN afferents. According to this theory, all cortical cells affected by a grating receive the same number of LGN spikes per unit time, but the arrival times of the spikes are more modulated for cells that are optimally driven, going up and down following the light and dark bands in the drifting grating, and that produces a much stronger response. (A cell is said to be *optimally driven* if its preferred orientation is aligned with that of the grating, orthogonally driven if it is orthogonal to it.)

The model in [14] demonstrated the effiacacy of Hubel and Wiesel's idea; firing rates rose to several times that in background when optimally driven. This is an example of what we were alluding in Sect. 2.3 when we wrote that currents do not tell the whole story: in the model cells in optimally and orthogonally driven domains received the same amount of current from LGN; it is the patterns of the spike trains exaggerated by recurrent excitation that made the difference.

This is most of what is needed for purposes of the discussion to follow.

3.2 Nonequilibrium Steady States for Cortical Networks

We wish to be clear from the start that we are not proposing any deep connection between the physics of heat conduction and the biology of signal processing. We appeal to statistical mechanics ideas because there is some superficial resemblance between the two setups, and we believe computational neuroscience can benefit from some of the ideas used to describe thermal systems driven out of equilibrium.

Relevant ideas from nonequilibrium statistical mechanics

Consider a homogeneous piece of matter either in isolation or in contact with a single heat reservoir. Its microscopic dynamics are described by a Hamiltonian system on a domain $\Omega \subset \mathbb{R}^d$, and large time statistics of particle motions are given by a distribution with density proportional to $e^{-\beta H}$ where H is the Hamiltonian and $\beta = T^{-1}$ is the inverse temperature.

Of interest to us here are systems that are driven out of equilibrium. Consider the same system placed in contact with multiple heat baths at unequal temperatures. In preparation for the discussion to follow, imagine a sheet of metal heated unevenly from underneath. The resulting dynamical system will, after a transient, settle down to a nonequilibrium steady state (NESS), the statistics of which are described by an invariant probability measure on the system's phase space (or space of all possible configurations). This probability distribution is generally a complicated object not representable in a simple form.

One way to understand an NESS is via the idea of local thermal equilibrium (LTE), which roughly speaking, asserts that in an NESS, there is a function $\beta(x)$, $x \in \Omega$, with the property that near each x in the physical domain, the invariant distribution is roughly describable by the equilibrium distribution with $\beta = \beta(x)$. More precisely, we say a system approaches LTE in the infinite volume limit if at each point x in the domain, the NESS restricted to a neighborhood of x of mesoscopic scale tends to a distribution whose density is increasingly proportional to $e^{-\beta(x)H}$. It is when this condition holds that one can make sense of the idea of local temperature for systems out of equilibrium, to be interpreted as $T(x) = 1/\beta(x)$ at the point x [17].

The setting can be more general than that above, e.g. it can involve mass as well as energy transport.

Viewing a nonequilibrium steady state as a collection of local equilibria simplifies the picture considerably: First, if each local distribution is describable by a finite number of parameters (such as energy, particle density), then the system as a whole is described by finitely many functions of x, and solving for finitely many unknown functions on a domain $\Omega \subset \mathbb{R}^d$ is orders of magnitude simpler than working with the space of invariant distributions of dynamical systems the dimensions of which grow with particle numbers. Second, assuming the local distributions are known from equilibrium theory, we have knowledge of the (energy, particle density) fluxes they produce. For certain classes of systems, solving for the functions above is reduced to balancing these fluxes among neighboring regions and with input values.

For an example of how these ideas were used to obtain information on the NESS of a class of mechanical systems, see [18].

Cortex is out of equilibrium

We now consider the driven dynamics on cortical layers, using as illustrative example the V1 model in [14]—though the ideas are not limited to this example.

We first draw some parallel between the setup in this V1 model, more generally in the $\mathcal{L}_{i,j}$-networks defined in Sect. 1.2, and that of statistical mechanics: Cortical circuits restricted to $\mathcal{L}_{i,j}$ are roughly 2D with translationally invariant structures. With neuronal dynamics governed by the same equations everywhere, the rules of local dynamics internal to the

network are also spatially homogeneous. These are the underlying structures in statistical mechanical systems.

Inputs to a cortical layer are analogous to external heat sources in statistical mechanics, and in statistical mechanics, the difference between a system being in or out of equilibrium is determined by the homogeneity of the heat source with which it is in contact. An important point here is that *when stimulated, inputs to cortex are rarely spatially homogeneous.*

As an example consider the network discussed in Sect. 3.1. When presented with a drifting grating of finite size, a region of Layer 4Cα corresponding roughly to the grating via the retinotopic map is stimulated. Within this region, domains with different orientation preferences are driven differently as explained in Item 5 of Sect. 3.1, resulting in a much more vigorous response in domains whose orientation preferences are aligned with the grating. In general, more complex visual stimuli (e.g. uneven contrast in different spatial locations, objects moving at different speeds) will result in different LGN responses which are passed along to cortex. For a system under drive, it is not just the inputs that are different; cortico-cortical interactions in response to the stimulus produces a dynamic that differs from location to location, and feedback is location dependent as well for the same reasons.

To summarize, cortical dynamics are characterized by spatially inhomogeneous inputs driving a system whose underlying structures and local dynamical rules are spatially homogeneous. This setup is not unique to V1: cortical layers that are farther from the stimulus source receive inputs from other layers, both interlaminar and interareal, and these inputs are almost always inhomogeneous. The setting is remarkably similar to that of nonequilibrium statistical mechanics.

Nonequilibrium steady states in cortex

Consider a situation where the visual stimulus is periodic (e.g. a drifting grating) or stationary. We assume as in Sect. 2 that the system has a unique invariant probability distribution to which it converges starting from most initial conditions. Since the system is under (inhomogeneous) drive, this invariant distribution is a *nonequilibrium steady state*.

Inspired by the idea of local thermal equilibrium (LTE), we propose that the NESS dynamics here can also be represented as a continuum of the local dynamic equilibria (LDE) in Sect. 2. Recall that the local population models in Sect. 2 are described by a finite set of parameters (synaptic coupling weights and inputs to E and I-cells), and mean firing rates can be estimated in terms of these parameters. Given a stationary stimulus, we propose that the NESS on a cortical layer can be estimated following the recipe used in statistical mechanics:

If we know the E and I-inputs to a local population $\mathcal{P}$, then in principle we should be able to estimate its mean E and I-firing rates (using e.g. mean-field computations with correction). The total Excitatory input to $\mathcal{P}$ is the sum of all Excitatory inputs it receives from within the same layer, local and long-range, from other layers, or directly from external stimuli if it is an input layer; and the same is true for the Inhibitory input it receives, though most of that will be local. The firing rates produced by $\mathcal{P}$ will contribute to the inputs to local populations to which $\mathcal{P}$ projects. To estimate an NESS, then, requires that we balance Excitation and Inhibition in much the same way that one balances heat fluxes to obtain macroscopic energy profiles in statistical mechanics.

Striking as the similarities may be, it must be pointed out that this picture can only be interpreted as approximate, as an aid to conceptualization. For one thing, we are dealing with much smaller numbers ($\mathcal{O}(10^{11})$ neurons in the human cerebral cortex, $\mathcal{O}(10^7)$ on each layer of V1, $\mathcal{O}(10^3)$ in a local circuit); there are no infinite volume limits to speak of.

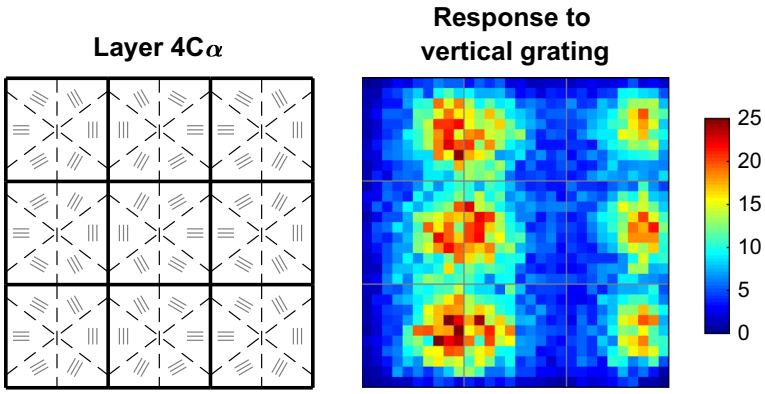

Fig. 4 Left: Model of Layer 4Cα of the V1 cortex in [14] divided into intended orientation domains (reproduced from Fig. 3 for the convenience of the reader). Right: Activity maps showing NESS responses to a vertical preferring grating. Each pixel represents a group of about 30 E-neurons; the colors represent mean firing rates averaged over 1 s. From the panel on the left, one sees that the vertical grating did indeed produce higher firing rates in the orientation domains identified as vertical preferring according to their LGN alignments. The range of colors in the right panel shows that in this 2D network of $\sim 40,000$ neurons, local populations have equilibrated to quite different dynamical regimes producing a wide range of firing rates (see color bar)

Nevertheless the idea of LDE seems both clear and natural in simulations, and the way these local equilibria fit together on cortical layers is consistent with what we have seen in model simulations of the macaque V1.

Figure 4 depicts the spiking activity during visual stimulation by a drifting grating in the 9 hypercolumn model cortex in [14]. The grating in question has vertical stripes. From the intended orientation domains shown in the left panel, one finds the regions that receive the most effective LGN drive, and they correspond roughly to the most brightly colored regions, i.e., the regions with the highest firing rates, in the right panel. The right panel can be seen as a snapshot of the NESS. Each pixel represents the activity of about 30 Excitatory neurons; its color is indication of the mean firing rates of these neurons averaged over a one second period, a snapshot of the LDE. The analogy with the metal sheet heated unevenly from below resulting in a nonconstant temperature profile is quite evident.

3.3 A Walk from NESS to NESS with Each Saccade

Much of ergodic theory including the notion of NESS is predicated on the concept of stationarity. Most of the sensory inputs received by cortex, on the other hand, are time-varying and not governed by fixed sets of statistical laws or patterns. How, then, is the notion of NESS relevant to cortical dynamics? Indeed the problem is not limited to neuroscience. The theory of dynamical systems has focused on unforced, periodically forced, or randomly forced systems. Mathematical techniques are lacking for dealing with time-varying, nonstationary forcing.

We do not have a general solution to offer, but want to end with an example to illustrate how NESS ideas can be taken a step beyond their usual boundaries, again using visual neuroscience as our guide.

Our reasoning is based on the differential in timescales of the various processes: Neurons operate in milliseconds (ms). In the visual cortex, transmission times between layers and

even cortical regions range from a few to 10–20 ms. In simulations using actual cell densities and biophysical time constants as was done in the modeling work in [14], response to the onset of a new stimulus starts quickly, and some degree of convergence occurs within 100+ ms. Compare this to the fact that we make no more than three saccades a second, and most of these saccades are very small, i.e., they do not involve drastic changes in visual stimuli. The arithmetic above suggests that for a nontrivial fraction of time, the dynamics within the layers of our visual cortex are either in, or moving towards, a nonequilibrium steady state.

The following dynamical picture is speculative, but it follows logically from the reasoning above: Suppose we are presented with a large and complex scene. As our eyes sample (quickly) the various aspects of it, the dynamics within the visual cortex can be described as drifting towards and transitioning between nonequilibrium steady states. To be sure, the set of NESS of any cortical layer is infinitely large. As we change our gaze, the dynamics within our visual cortex zigzag across an infinite landscape visiting one NESS after another, lingering longer at some (that catch our attention), never quite reaching others before moving on, and possibly revisiting some over and over. In short, NESS are the *limiting regimes* to which cortical dynamics are driven in response to time-varying, nonstationary visual stimuli, with convergence in the mathematical sense occurring to varying degrees.

4 Summary and Conclusion

Here again are the main points, summarized more succinctly in the hope of giving a clearer perspective:

1. *Relevant anatomical structures* We proposed to divide the 2D area of cortex into cortical regions the way it is conventionally done, then further subdivide into layers (e.g. the layers of V1). The resulting objects are biologically natural as they are anatomically and functionally distinct, and they are mathematically idealizable as infinitely large 2D networks with translationally invariant structures—based on the fact that local spatial structures are similar from point to point. This collection of 2D networks is then viewed as nodes of a larger network that is the cerebral cortex.
2. *Analogy with nonequilibrium statistical mechanics* Structural homogeneity does not imply dynamical homogeneity, and this is because external inputs to cortex are rarely spatially homogeneous. For the visual cortex, for example, features of the visual stimuli are carried by distinct patterns of drive at different cortical locations resulting in different activity profiles across the cortical surface. The setup is similar to that in nonequilibrium statistical mechanics, where a homogeneous piece of material is placed in contact with multiple unequal heat reservoirs. The analogy is superficial, but there are similarities in how fluxes are balanced and how excitation spreads.
3. *E–I interaction: a challenge for dynamicists* There are differences as well, the most prominent being that in cortex, local dynamics are dominated by the presence of two groups of agents, namely Excitatory and Inhibitory neurons, that exert opposing forces on the local population. Models of isolated local populations are interesting in their own right: competition of antagonistic and sympathetic forces, balancing acts, negotiated dynamic equilibria, and the potential for emergent network phenomena pose exciting new challenges for the dynamical systems community.
4. *From local dynamic equilibria to nonequilibrium steady states* It is proposed that following local thermal equilibrium ideas in statistical mechanics, one may be able to estimate cortical NESS by balancing Excitation and Inhibition produced in local populations.

5. *NESS as limiting regimes for nonstationary inputs* NESS are associated with stationary inputs, but natural inputs driving cortical dynamics are typically far from stationary. Using vision as an example, we hypothesize that NESS may represent limiting regimes of cortical responses to nonstationary inputs, in the sense that each saccade drives the visual cortex to a new NESS with partial convergence.

References

1. Baddeley, R., Abbott, L.F., Booth, M.C.A., Sengpiel, F., Freeman, T., Wakeman, E.A., Rolls, E.T.: Responses of neurons in primary and inferior temporal visual cortices to natural scenes. Proc. R Soc. Lond. B **264**, 1775–1783 (1997)
2. Beaulieu, C., Kisvarday, Z., Somogyi, P., Cynader, M., Cowey, A.: Quantitative distribution of GABA-immunopositive and -immunonegative neurons and synapses in the monkey striate cortex (area 17). Cereb. Cortex **2**, 295–309 (1992)
3. Binzegger, T., Douglas, R., Martin, K.: Topology and dynamics of the canonical circuit of cat V1. Neural Netw. **22**, 1071–78 (2009)
4. Börgers, C., Kopell, N.: Synchronization in networks of excitatory and inhibitory neurons with sparse, random connectivity. Neural Comput. **15**(3), 509–538 (2003)
5. Bressloff, P.: Waves in Neural Media: From single neurons to neural fields. Lecture Notes in Math. Modeling in the Life Sciences, Springer (2014)
6. Brunel, N., Hakim, V.: Fast global oscillations in networks of integrate-and-fire neurons with low firing rates. Neural Comput. **11**(7), 1621–1671 (1999)
7. Brunel, N.: Dynamics of sparsely connected networks of excitatory and inhibitory spiking neurons. J. Comput. Neurosci. **3**, 183–208 (2000)
8. Brunel, N., Wang, X.J.: What determines the frequency of fast network oscillations with irregular neural discharges? I. Synaptic dynamics and excitation-inhibition balance. J. Neurophysiol. **90**(1), 415–430 (2003)
9. Buice, M.A., Cowan, J.: Statistical mechanics of the neocortex. Prog. Biophys. Mol. Biol. **99**(2–3), 53–86 (2009)
10. Burns, S.P., Xing, D., Shapley, R.M.: Gamma-band activity in the local field potential of V1 cortex: a "clock or filtered noise? J. Neurosci. **31**, 9658–64 (2011)
11. Cai, D., Tao, L., Rangan, A.V., McLaughlin, D.W., et al.: Kinetic theory for neuronal network dynamics. Commun. Math. Sci. **4**(1), 97–127 (2006)
12. Cardin, J.A.: Snapshots of the brain in action: local circuit operations through the lens of oscillations. J. Neurosci. **36**, 10496–10504 (2006)
13. Chariker, L., Young, L.-S.: Emergent spike patterns in neuronal populations. J. Comput. Neurosci. **38**(1), 203–220 (2015)
14. Chariker, L., Shapley, R., Young, L.-S.: Orientation selectivity from very sparse LGN inputs in a comprehensive model of macaque V1 cortex. J. Neurosci. **36**, 12368–12384 (2016)
15. Chariker, L., Shapley, R., Young, L.-S.: Rhythm and synchrony in a cortical network model. J. Neurosci. **38**(40), 8621–8634 (2018)
16. Chariker, L., Shapley, R., Young, L.-S.: Contrast response in a comprehensive network model of V1. Under Review
17. De Groot, S., Mazur, P.: Nonequilibrium Thermodynamics. North Holland, Amsterdam (1962)
18. Eckmann, J.P., Young, L.-S.: Nonequilibrium energy profiles for a class of 1D models. Commun. Math. Phys. **262**(1), 237–267 (2006)
19. Gerstein, G.L., Mandelbrot, B.: Random walk models for the spike activity of a single neuron. Biophys. J. **4**, 41 (1964)
20. Gray, C.M., Singer, W.: Stimulus-specific neuronal oscillations in orientation columns of cat visual cortex. Proc. Natl. Acad. Sci. USA **86**, 1698–1702 (1989)
21. Heeger, D.: Theory of cortical function. Proc. Natl. Acad. Sci. USA **114**, 1773–1782 (2018)
22. Henrie, J.A., Shapley, R.: Lfp power spectra in v1 cortex: the graded effect of stimulus contrast. J. Neurophysiol. **94**(1), 479–490 (2005)
23. Hodgkin, A.L., Huxley, A.F.: A quantitative description of membrane current and its application to conduction and excitation in nerve. J. Physiol. **117**(4), 500–44 (1952)
24. Holmgren, C., Harkany, T., Svennenfors, B., Zilberter, Y.: Pyramidal cell communication within local networks in layer 2/3 of rat neocortex. J. Physiol. **551**, 139–53 (2003)

25. Hubel, D., Wiesel, T.: Receptive fields, binocular interaction and functional architecture in the cats visual cortex. J. Physiol. **160**, 106–154 (1962)
26. Joglekar, M., Mejias, J., Yang, G.R., Wang, X.-J.: Inter-areal balanced amplification enhances signal propagation in a large-scale circuit model of the primate cortex. Neuron **98**(1), 222–234.e8 (2018)
27. Koch, C.: Biophysics of Computation. Oxford Univ Press, Oxford (1999)
28. Li, Y., Chariker, L., Young, L.-S.: How well do reduced models capture the dynamics in models of interacting neurons? J. Math. Biol. **78**, 83 (2018). https://doi.org/10.1007/s00285-018-1268-0
29. McLaughlin, D., Shapley, R., Shelley, M., Wielaard, D.J.: A Neuronal Network Model of sharpening and dynamics of orientation tuning in an input layer of macaque primary visual cortex. Proc. Natl. Acad. Sci. USA **97**, 8087–8092 (2000)
30. Ostojic, S.: Interspike interval distributions of spiking neurons driven by fluctuating inputs. J. Neurophysiol. **106**, 361–373 (2011)
31. Oswald, A.M., Reyes, A.: Maturation of intrinsic and synaptic properties of layer 2/3 pyramidal neurons in mouse auditory cortex. J. Neurophysiol. **99**, 2998–3008 (2008)
32. Rangan, A.V., Young, L.-S.: Emergent dynamics in a model of visual cortex. J. Comput. Neurosci. **35**, 155–167 (2013)
33. Rangan, A.V., Young, L.-S.: Dynamics of spiking neurons: between homogeneity and synchrony. J. Comput. Neurosci. **34**(3), 433–460 (2013)
34. Tso, D.Y., Frostig, R.D., Lieke, E.E., Grinvald, A.: Functional organization of primate visual cortex revealed by high resolution optimal imaging. Science **249**(4967), 417–420 (1990)
35. Ungerleider, L.G., Desimone, R.: Cortical connections of visual area MT in the macaque. J. Comput. Neurol. **248**(2), 190–222 (1986)
36. Ungerleider, L.G., Galkin, T.W., Desimone, R., Gattass, R.: Cortical connections of area V4 in the macaque. Cereb. Cortex **3**, 477–99 (2008)
37. van Vreeswijk, C., Sompolinsky, H.: Chaotic balanced state in a model of cortical circuits. Neural Comput. **10**(6), 1321–1371 (1998)
38. Wilson, H., Cowan, J.D.: Excitatory and inhibitory interactions in localized populations of model neurons. Biophys. J. **12**(1), 1–24 (1972)