View metadata, citation and similar papers at core.ac.uk

brought to you by CORE

provided by Online Research Database In Technology

Technical University of Denmark

Dynamics of Lage Spiking Neural Networks

Lerchner, Alexander; Bothe, Hans-Heinrich; Hertz, John

Publication date:
2004

Document Version
Early version, also known as pre-print

Link back to DTU Orbit

Citation (APA):
Lerchner, A., Bothe, H-H., & Hertz, J. (2004). Dynamics of Lage Spiking Neural Networks.

DTU Library
Technical Information Center of Denmark

General rights
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Dynamics of Large Spiking Neural Networks

Modeling Circuits in the Central Nervous System

Alexander Lerchner

Submitted in partial fulfillment of the
requirements for obtaining a Ph.D. degree at

Ørsted•DTU
Technical University of Denmark (DTU)

June 17, 2004

Contents

Preface

The research project presented in this thesis was carried out at Ørsted-DTU, Technical University of Denmark (DTU) and at the Nordic Institute for Theoretical Physics (NORDITA), with Associate Professor Hans-Heinrich Bothe, Ørsted-DTU, as principal supervisor and Professor John Hertz, NORDITA, as project supervisor. Part of the research was carried out at the Center for Neural Science, New York University (NYU), in the laboratory of Professor John Rinzel. The study was funded by a DTU Grant Award (full PhD scholarship, September 2000 - December 2003) and by NORDITA (guest researcher, January 2004 - April 2004).

I am extremely grateful to Professor John Hertz for his guidance and advice. The success of this project in its present form would have been impossible without his generous support.

I am thankful to Professor John Rinzel for giving me the opportunity to work with him in one of the most exciting and stimulating environments for neuroscientists today (CNS at NYU). Thanks are due to Michael O'Donovan and Joël Tabak for scrutinizing my theoretical considerations on chick spinal cord dynamics in the light of experimental evidence.

I want to thank everyone at NORDITA for providing such a pleasant working environment, and the staff at Ørsted-DTU for their kind assistance.

Finally, I am thankful beyond description to the most important person in my life, my wife Mandana. This work has profited in many ways from her razor-sharp intellect in numerous discussions, from her support during hard work under time-pressure, and from her love at all times.

Chapter 1

Introduction

The number of neurons in the central nervous system, consisting of the brain and the spinal cord, is astronomically large. The cerebral cortex in humans contains between 20 and 30 billion neurons (Mountcastle, 1998), with about $20,000$–$40,000$ neurons per cubic millimeter. Even if we want to consider as small a subsystem as possible, we are facing network sizes on the order of many thousand neurons: Cortical neurons receive thousands of inputs from other neurons, at least half of them from local circuitry (not more than half a millimeter in diameter) in which any neuron is connected to any other neuron with a probability of about 10%, independent of distance. Thus, if we want to study circuits in the central nervous system, we are immediately confronted with large spiking neural networks.

A common approach to study the dynamics of such systems is to restrict the description of neuronal outputs to averaged quantities, i.e., firing rates, instead of describing precisely timed action potentials. In these models, the firing rate as a function of input current is given by hand. There are good arguments in favor of such firing-rate models as compared to (naive ways of using) spiking models (see, e.g., Dayan and Abbott, 2001, chapter 7).

However, there are several limitations to rate models, and we overcome some of them with the two different kinds of spiking models presented in this work.

While it is easier to solve certain aspects of network dynamics analytically using rate models, it is often difficult to provide physiological interpretations of introduced parameters. For example, there has been success in constructing a rate model with qualitative similar dynamics as the observed spontaneous activity in the chick spinal cord, but it remains unclear how these can

emerge from neuron properties. Here, we suggest a different solution based on a new synapse model, and show which physiological mechanisms could account for the dynamics by simulating and analyzing a full spiking neural network.

Some fundamental questions about cortical dynamics cannot be addressed by rate-models at all. Most notably, they cannot say anything about firing statistics. Cortical neurons fire highly irregularly in response to sensory stimuli, with firing statistics close to that of a completely random Poisson process. How is the irregularity in neuronal firing controlled by intrinsic network parameters? What are the mechanisms behind irregular firing? We present a complete mean-field theory that allows us to describe firing statistics, including firing correlations, for spiking cortical network models in a completely self-consistent manner. The theory is implemented by a numerical algorithm that can be used for any kind of neuron and synapse model. With these tools, one can for the first time provide both qualitative and quantitative explanations for observed firing statistics – and even for conductance- and membrane potential fluctuations – of cortical neurons *in vivo*, without the need to make assumptions about statistics of the neuronal input.

Organization of the thesis

The main part of this work is presented in a collection of papers reprinted at the end of the thesis.

In the first part, we summarize the theories and the main results in a closed form, using a consistent notation and pointing to interdependencies where applicable. In the sections corresponding to the two shorter papers, we provide important additional information about the models and present more detailed results. Generally, we discuss the results for all the models in greater detail and from a more global perspective. This way, we hope to make the implications of our findings more accessible to readers who are not actively involved in research in this field.

Chapter 2

General background

This chapter serves as an introduction to biological concepts and terminology, as far as they are relevant to both understanding and motivating the particular modeling approaches in this work.

2.1 Neurons and synapses

Neurons are the functional building blocks of nervous systems. They are connected via synapses to form networks. Figure 2.1 shows a cartoon representation of a typical neuron in the brain. Neurons receive inputs from other neurons in the network via postsynaptic terminals, which are located on specialized structures called dendrites, or directly on the cell body (soma). The inputs can either be excitatory or inhibitory and their effects are integrated in the soma. Provided that there is enough excitation, an *action potential* is created (one says the neuron "fires") that propagates in form a stereotyped voltage pulse (also called a "spike") through the axon to presynaptic terminals. Spikes that arrive at presynaptic terminals cause a release of neurotransmitter onto postsynaptic terminals of further neurons.

A few numbers serve to illustrate how densely connected neurons can be: The cell body of a typical cortical neuron ranges in diameter from about 10 to 50 micrometer and the dentritic tree receives, on average, 2 synaptic inputs per micrometer. The dendritic length density is about 400 meter per cubic millimeter, and the axonal length density is even higher, 3200 meter per cubic millimeter (Rolls and Deco, 2002).

Neurons are not in direct contact with each other (with the exception of

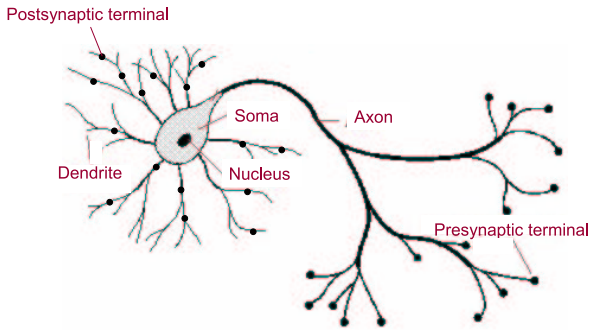

Figure 2.1: Structure of a neuron. The neuron receives inputs from other neurons via its postsynaptic terminals on the dendrite. The resulting input currents get integrated at the soma. If the soma gets sufficiently depolarized, an action potential ("spike", "firing") is initiated and propagated through the axon to the presynaptic terminals, which then provide inputs to postsynaptic terminals of other neurons. Figure adapted from http://vv.carleton.ca/ neil/neural/neuron-a.html

electrical synapses, which we do not consider here because they play a minor role in the networks that we consider in this work) – they influence each other only via spikes. For this reason, spikes are considered the carriers of information in neural networks, and individual neurons are characterized by their spike responses to, e.g., sensory stimuli.

Like all cells in the body, neurons are surrounded by a cell membrane consisting of a lipid-bilayer, interspersed with proteins. The membrane-spanning proteins of interest to us here are called ion-channels and receptors. We will come back to receptors later, when we have a closer look at the functionality of synapses. The membrane acts as an insulator, separating two liquid, electrically conducting media. Ion channels in the membrane allow specific ions to pass. There are concentration differences of various ion species between the inside and the outside of the cell, leading to a partial flow of ions to the other side until the force due to the concentration gradient is counterbalanced by the resulting voltage gradient in the other direction: an equilibrium potential is established that keeps the membrane polarized at the so-called resting potential, unless disturbed by input currents. The resting potential of cortical neurons is on the order of -70 mV.

Action potentials are initiated and propagated by voltage sensitive ion

channels that open (that is, they change their conductance) only when the membrane potential gets sufficiently depolarized, i.e., when the membrane potential reaches a threshold value ("the firing threshold").

Hodgkin and Huxley (1952) were the first to provide a detailed model of membrane-potential kinetics, including a precise description of how the action potential is caused by (active) changes in membrane conductance to Na^+ and K^+ ions.

However, membrane-potential kinetics in the subthreshold regime can be approximated well by a passive R-C circuit with a battery that maintains the resting potential u_0. The well-insulating lipid bilayer determines the capacitance C and the ion channels that are (passively) open below threshold determine the resistance, resp., conductance g. Thus, for membrane potentials $u < \theta$, with a threshold θ obey in this simplified description the equation of motion

$$C\frac{\mathrm{d}u}{\mathrm{d}t} + g(u - u_0) = I, \tag{2.1}$$

where I stands for an input current. The membrane acts thus as a leaky integrator of the current I. The shape of an action potential is a stereotyped, very brief (with a duration of about 1 millisecond) pulse that depolarizes the membrane to slightly positive values, followed by a re-polarization close to rest. This behavior can be simplified by "inserting" such a pulse when the neuron fires, followed by a reset of the membrane potential to rest anytime the membrane potential reaches threshold. This model is called a leaky *integrate-and-fire* (LIAF) neuron.

We will use such LIAF neurons in all our spiking neural network models described in this work, and we measure the membrane potential relative to the resting potential u_0 in units of the threshold θ (i.e., with our re-scaled u, we have $u_0 = 0$ and $\theta = 1$):

$$\tau\frac{\mathrm{d}u}{\mathrm{d}t} = -u + \hat{I}, \tag{2.2}$$

with $\hat{I} = I/g$ and $\tau = C/g$. The input-output relationship (firing rate as a function of input current), also called the "gain function" can be calculated for a constant current $\hat{I}_0 > \theta$ as

$$r = \left[\tau \ln\left(\frac{\hat{I}_0}{\hat{I}_0 - \theta}\right)\right]^{-1}, \tag{2.3}$$

or, with an absolute refractory period τ_r for which we keep the neuron at reset-level after spiking:

$$r = \left[\tau_r + \tau \ln\left(\frac{\hat{I}_0}{\hat{I}_0 - \theta}\right)\right]^{-1}. \tag{2.4}$$

We have mentioned above that neurons affect other neurons by emitting spikes, which cause a release of neurotransmitter onto postsynaptic terminals. The arrival of the neurotransmitter is sensed at the postsynaptic terminal via receptor molecules, which are the second kind of transmembrane proteins (beside ion channels) that are import for our purposes. These receptors then open transiently (a few milliseconds for the types we consider here) ion channels. In fact, the types we consider here are receptors and ion channels at the same time, so-called *ligand-gated ion channels*. Opening ion channels means increasing the conductance for the species of ions they are selective to, causing current to either flow in (excitation) or out (inhibition). In the cortex, neurons that release the neurotransmitter glutamate have an excitatory effect on postsynaptic neurons, because glutamate causes the opening of ion channels for cations that have a reversal (= equilibrium) potential around 0 mV. Thus, they have a depolarizing effect on the postsynaptic neuron, bringing its membrane potential closer to threshold. Neurons that release GABA (γ-aminobutyric acid) cause opening of Cl^- channels, making the postsynaptic neuron more depolarized because Cl^- has a reversal potential that lies below the resting value.

We can model the effect of synaptic input to the integrate-and-fire model (2.2) by writing the current $\hat{I}$ as $g_s(t)(u - V_s)$ where $g_s(t)$ describes the temporal synaptic conductance change due to opening of the ion channels, and where V_s is the reversal potential of that particular synapse:

$$\tau \frac{\mathrm{d}u}{\mathrm{d}t} = -u - g_s(t)(u - V_s) \tag{2.5}$$

A recent study by Meffin et al. (2004) shows that integrate-and-fire neurons with such conductance-based synapses have very similar dynamics to Hodgkin-Huxley model neurons regarding firing statistics and membrane potential statistics for neuronal inputs like those for cortical neurons *in vivo*. Thus, for large networks, this simple model offers a computationally very economic and yet physiologically faithful choice for simulating cortical dynamics.

2.2 Cerebral cortex

The cerebral cortex comprises about 60–65% of the volume of the human brain, of which it is the outer mantle. If spread flat, it would occupy about 2,600 cm^2 (Mountcastle, 1998). It is thus no surprise that it has to be highly convoluted in order to fit into our skulls. The connectivity within the cortex is remarkably similar across all areas and modalities. Vertically, it is organized into six distinct layers, usually referred to by roman numbers. Sensory input reaches the primary sensory areas mostly via layer IV. The local cortical networks (cortical columns) modeled later in this work may be more precisely referred to as layer IV of these cortical columns. When we speak of a local network or a cortical column, we mean circuits consisting of neurons that have similar firing characteristics in response to, e.g., sensory input. Such neurons are arranged close to each other, within a diameter not more than half a millimeter in horizontal directions (parallel to the surface) but which are arranged in vertical columns, thus the name. Within that length scale, neurons receive similar external input, have similar recurrent connectivity and fire in a similar manner. A correct description of neuronal input has therefore to be consistent with the output of the neuron itself. This is what we mean by the important concept of self-consistency, which will be central for the mean-field models developed in this work.

External input (all long-range input) to cortical neurons is always excitatory, whether from sensory sources, from lower brain centers or from more distant cortical areas. In the same vein, all long-range "output" of cortical processing is excitatory. Inhibitory neurons are only locally connected. Thus, only excitatory neurons carry signals that are received ("seen") by other brain centers.

We will frequently refer to the visual cortex, which is positioned on the occipital lobe on the back of our brain (when we consider the eyes as positioned in the front). Visual stimuli enter the eye in form of light waves, which are transformed to neuronal signals by the retina on the back of the eyeball. From there, the signals get propagated to neurons in the *lateral geniculate nucleus* (LGN), which in turn project to the primary visual cortex (V1).

Much research has been done to characterize primary cortical neurons by their firing responses to visual stimuli. Many neurons are orientation sensitive, that is, they respond strongest to visual stimuli of a specific orientation, and with gradually decreasing firing rates as the orientation becomes less optimal. Neurons with the same orientation preference are arranged vertically

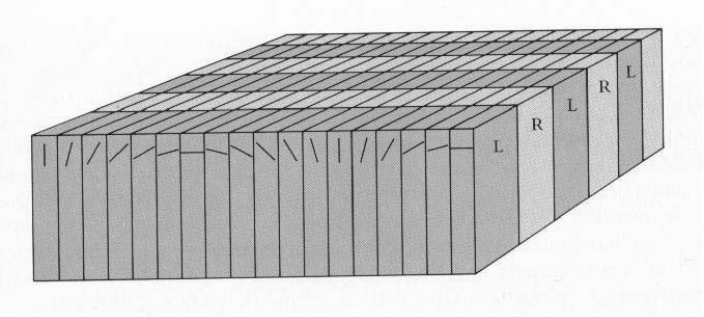

Figure 2.2: Regular arrangements of orientation columns in primary visual cortex. The indications L and R stand for left and right eye, respectively. Figure adopted from http://webvision.med.utah.edu/imageswv/columns.jpg

(into orientation columns) and neurons for different orientations are topographically organized in a regular manner. Figure 2.2 shows such a regular arrangement of varying orientations, which would be found in the cortex along the boundaries of ocular dominance stripes, i.e., where neurons on one side react strongest to the left eye and those on the other side to the right eye, or vice versa. A set of orientation columns of all possible orientations reacting to the same place in the visual field (i.e., they share the same receptive field) are referred to as an *orientation hypercolumn*. Since it is not possible to arrange orientation columns in a continuously regular manner (so that orientation preferences change continuously from one position to the next) on the two-dimensional surface of the cortex for topological reasons, there must be necessarily singularities in the arrangement. These are called pinwheel centers, and an entire pinwheel is a circular arrangement of orientation columns around a pinwheel center. We will refer to pinwheels later, when we discuss possible extensions to our simple orientation hypercolumn model.

Finally, we want to point out some important experimental findings from measurements of visual cortical neurons *in vivo* that have been topics of intense research in the field of neuroscience:

- Spiking is highly irregular in response to repeated visual stimuli. The spike count variance tends to be somewhat higher than that for a Poisson process.

- Cortical neurons *in vivo* exhibit a threshold-linear gain function, in contrast to what we have derived in (2.3) and (2.4); however, the same neurons in vitro show gain functions more similar to (2.4).

- Cortical neurons in vivo, especially during stimulation, but also "at rest" have much higher input conductances than the same neurons *in vitro*, making their characteristic membrane time constant shorter than expected from single-neuron investigations.

We can provide plausible explanations for underlying mechanisms of all these facts with help of the mean-field models that we present in this work.

Chapter 3

New synapse model – Spontaneous activity in the developing chick spinal cord

Spontaneous activity is a common phenomenon in most, if not all, developing networks of the central nervous system (CNS) (for reviews see O'Donovan, 1999; Wenner and O'Donovan, 2001). This activity is remarkably similar for such diverse network architectures as the hippocampus, the retina, and the spinal cord. It comprises recurring *episodes*, containing one or several *cycles* of high network activity, that alternate with *silent phases*. All of these developing networks have in common that they are hyperxcitable due to the fact that synaptic contacts that are inhibitory in the mature system (such as GABA-ergic synapses) are functionally excitatory during these stages of development.

There is widespread consensus that spontaneous activity plays important roles in the development of the nervous system itself, as well as of muscles, bones and other tissues that are affected by, e.g., the resulting spontaneous movements. Despite its universality and its importance, the mechanisms and dynamics underlying spontaneous activity are still poorly understood. While there has been some success in modeling episodic activity employing mutually independent fast and slow (activity-dependent) *network depression* (Tabak et al., 2000, 2001), it is not clear which physiological mechanisms should account for these dynamics, nor is it clear how these dynamics can emerge from neuron properties (we are not aware of successful attempts to implement these dynamics in full network models). In addition, these mod-

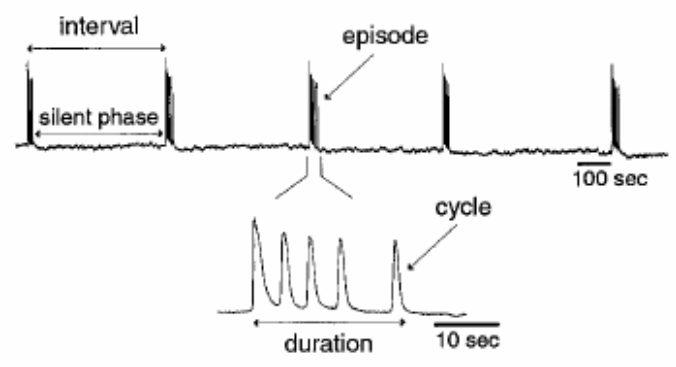

Figure 3.1: Spontaneous activity recorded from chick spinal cord. Episodes of cyclic activity alternate with silent phases. The cycling frequency decreases during the course of an episode. Figure adopted from Tabak et al. (2000).

els ascribe slow activity-dependent depression a causal role in terminating episodes, which is not reconcilable with experimental data, as we will discuss later.

Here, we propose a novel way of thinking about the problem by stressing the consequence of the fact that all these networks are hyperexcitable: once they get activated, positive feedback creates an avalanche of activity that drives them to their physiological limits. Then, the weakest or least resourceful parts of the system will necessarily set a boundary. Looking at the problem in this way, we can convince ourselves that the limited number of vesicles in the tiny presynaptic boutons of CNS nerve terminals (see, e.g., Harata et al., 2001) are likely to pose a major limitation on this run-away effect. For this reason, we introduce a new model for presynaptic dynamics that accounts for the limited number of readily available vesicles (Lerchner and Rinzel, 2004). The model is in agreement with recent findings on vesicle release mechanisms and structural organization of presynaptic CNS terminals. We show how the proposed synaptic dynamics can account for the generic qualitative features of spontaneous activity.

In this study, we concentrate on data obtained from the developing chick spinal cord, because it is probably the best studied developing system with respect to spontaneous network activity (see, e.g., Ho and O'Donovan, 1993; Chub and O'Donovan, 1998; Fedirchuk et al., 1999; Wenner and O'Donovan, 2001; Chub and O'Donovan, 2001). Reported episode durations are 5–90 sec, cycle rates 0.1–2 Hz, and the length of the silent phases 2–30 min. An

example of a recording, taken from Tabak et al. (2000), is shown in Figure 3.1.

We can summarize the properties that should be explained by a model that successfully reproduces the dynamics of spontaneous activity in the developing chick spinal cord as follows:

1. Recurring spontaneous initiation and termination of episodes
2. Cyclic activity within episodes
3. Cycle frequency decreases during episode
4. Shorter episodes start with lower cycle frequency; episode length can be predicted well by initial cycle frequency
5. Positive correlation between episode duration and length of preceding silent phase; episode termination is thus deterministic
6. Initiation of episodes is stochastic
7. Episodes have an absolute refractory period: not possible to evoke an episode prematurely for about 100 sec after termination of an episode
8. Slow synaptic depression sets in after the end of an episode and reaches its peak about 1–1.5 minutes later
9. Network architecture not important (same phenomenon in retina, hippocampus, and many other networks)

All of these features can be accounted for by our model, as we will discuss below.

3.1 Model and methods

The presynaptic terminal of a synapse contains vesicles of neurotransmitters. It is known that these vesicles do not wait there in an unstructured way for getting released. Instead, only vesicles that are in the *active zone*, already docked to the cell membrane or close to it, can be released in response to an incoming action potential. There is also a larger storage of vesicles that lie dormant behind the active zone, which have to be transported into the

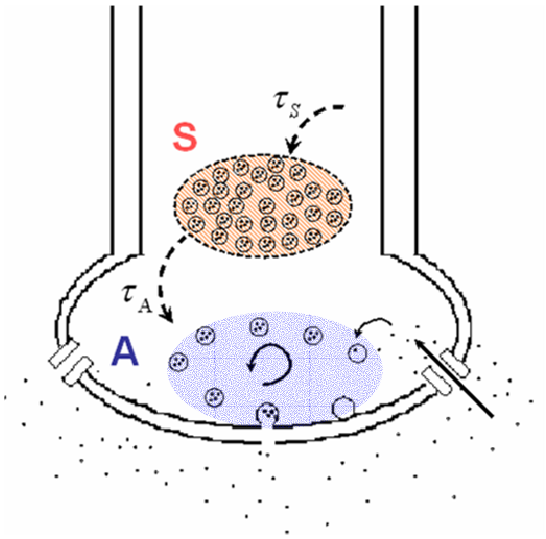

Figure 3.2: New synapse model. The presynaptic terminal contains two pools of neurotransmitter vesicles: an active pool A of readily releasable vesicles and a larger storage pool S. Vesicles in S can only be released after transition to A. The speed of transition, parameterized by τ_{A}, slows down with decreasing sizes of S. Unless S is filled up to its maximum, it gets replenished at a constant but slow rate, determined by the time constant τ_{S}. Also indicated is the partial recycling of released neurotransmitter. While recycling plays an important role in mature synapses, it is not included in our model for reasons outlined in the text.

active zone before they can get released. We will refer to these collectively as the *storage pool* S.

In our synapse model, we term the vesicles that can be readily released (in response to an action potential entering the presynaptic terminal) the *active pool* A. See Figure 3.2 for a cartoon representation of our synapse model.

Readily releasable vesicles (RRVs) in the real synapse are those that are currently in the active zone (less than 10 in the small CNS nerve terminals) plus those that can be quickly recruited to the active zone. The total number of vesicles in CNS nerve terminals is estimated between 200 and 520 (Harris and Sultan, 1995; Schikorski and Stevens, 1997). A fraction of about 15–20% of these are readily releasable (Harata et al., 2001). These RRVs undergo exo-

endocytotic recycling, that is, part of the neurotransmitter that gets released into the synaptic cleft is re-uptaken to allow quick refilling of emptied RRVs. In the mature synapse, vesicle recycling is a fast process, which explains why the small number of RRVs does not seem to pose a major limitation under "normal, mature" working conditions in the brain (Harata et al., 2001). We argue that vesicle recycling gets more efficient during development, for reasons that we will discuss later. Our model represents an early stage in the development of the chick spinal cord (embryonic day 7.5), so we do not model vesicle recycling explicitly here.

In our model, we assume that A has a maximum capacity $A_{\max}$. Provided that A is not empty, a presynaptic spike causes a vesicle release and the size of A gets reduced by one. The effect of the released vesicle in terms of postsynaptic input current is scaled by the momentary size of A relative to $A_{\max}$. This is a way of implementing the averaged effect of reduced release probability for smaller numbers of available vesicles (Harris and Sultan, 1995).

Unless A is filled up to its maximum, it gets constantly, but slowly, replenished by transition of vesicles from the storage pool S into A, with a rate that decreases with decreasing sizes of S. We formalize this concept by defining a vesicle transition variable V_{tr} that represents the fraction of a vesicle that is about to be transferred:

$$\tau_{\mathrm{A}} \frac{\mathrm{d}V_{\mathrm{tr}}(t)}{\mathrm{d}t} = \begin{cases} S + S_{\max} & \text{if } (A < A_{\max}) \wedge (S > 0) \\ 0 & \text{else,} \end{cases} \tag{3.1}$$

where $S_{\max}$ is the maximum capacity of the storage pool S. Actual transfer of a (whole) vesicle ($S \to S - 1$ and $A \to A + 1$) only occurs when V_{tr} exceeds 1, after which $V_{\mathrm{tr}} \to V_{\mathrm{tr}} - 1$. Thus, the pool sizes are confined to integer numbers. The role of $S_{\max}$ on the right-hand-side of (3.1) is to ensure that the transition rate remains finite for vanishing sizes of S, so that S can be completely depleted under excessive stimulation.

The storage pool gets replenished with "new" vesicles via a metabolic process that is much slower than the transition from S to A. Similar to the transition process, we keep track of a vesicle generation variable V_{gen} that obeys

$$\tau_{\mathrm{S}} \frac{\mathrm{d}V_{\mathrm{gen}}(t)}{\mathrm{d}t} = \begin{cases} \text{const.} & \text{if } S < S_{\max} \\ 0 & \text{if } S = S_{\max}, \end{cases} \tag{3.2}$$

with a time constant $\tau_{\mathrm{S}} \gg \tau_{\mathrm{A}}$. A new vesicle is "created" whenever V_{gen} exceeds 1, by setting $S \to S + 1$ and $V_{\mathrm{gen}} \to V_{\mathrm{gen}} - 1$.

In addition to the presynaptic modeling described above, we need to account for the experimentally observed *postsynaptic* depression. This kind of activity-dependent depression is a slow time-delayed process. Even though the postsynaptic depression is caused by the high activity during the episodes (as experiments confirm), there is almost no depression at the beginning of the following silent phase. The slow depression sets in steeply at the end of the episode, reaching its maximum only 1–1.5 min later, then relaxing slowly back over the entire silent phase.

We model the time-delayed postsynaptic depression by a mechanism that integrates over the firing activity of the postsynaptic neuron within a time window on the order of tens of seconds, and which takes effect after a further delay of the same order. Even though experimental studies indicate that somatic activity seems not to be required for the depression (Fedirchuk et al., 1999), postsynaptic activity certainly is. Since high postsynaptic activity and high somatic activity always coincide in both our model and during spontaneous activity in chick spinal cord, we do not introduce a qualitative error by making this simplification. The somatic activity within the time window is then compared with a "neutral range" of activity. If the activity exceeds the neutral range, then the synaptic strength is scaled down, if it is lower, it is scaled up. Within the neutral range, no adjustments are made to the postsynaptic sensitivity.

During the time course of an episode, the activity in the chick spinal cord remains at a low level, which we model by random firing at a low rate for several seconds. Preliminary investigations indicate that this activity need not be modeled explicitly if the network is big enough (thousands of neurons, which is a realistic number for the real network) with sufficient variation in the synapse parameters. Due to computational limitations, we include this detail in the current form of the model to allow smaller network sizes (hundreds of neurons).

Spontaneous activity (as the name suggests) does not require external stimulation to be initiated. Prolonged intervals without network – and thus synaptic – activity are accompanied by more and more frequent spontaneous synaptic events, peaking in frequency and amplitude before the spontaneous start of an episode (O'Donovan, 1999). We model these events as random spontaneous vesicle release (Poisson noise) that increases in frequency in the absence of action potential-triggered release.

The model network consists of randomly connected leaky integrate-and-fire neurons without external input. All synaptic contacts are excitatory,

governed by the pre- an postsynaptic dynamics outlined above.

3.2 Results

Our model network generated spontaneous episodic activity with the same qualitative features as observed in the developing chick spinal cord. We varied the size of the network (from a few hundred to a few thousands of neurons) and the connection probabilities (between 10% and 100%, i.e., full connectivity) and could find dynamics like those shown in Figure 3.3 for all cases investigated.

To compare our results with extracellular recordings such as that shown in Figure 3.1, we needed to calculate the time course of the overall network activity. We did this by averaging the number of spikes per millisecond (ms) over a sliding time window of 20 ms. The maximum average activity amounts then to 0.25 due to our choice of an absolute refractory period equal to 3 ms.

Figure 3.3 shows a typical time course of both a spontaneous and an evoked episode, aligned with corresponding synaptic states from a randomly chosen synapse in the network. In the top panel of the figure, the overall network activity is plotted, showing the episodic nature with cyclic activity within episodes. The first of the two episodes started spontaneously and the decrease in cycle frequency can be clearly seen. The second episode was evoked prematurely by transiently driving the membrane potential of all neurons above threshold. The duration of the early evoked episode is shorter. We generally found a positive correlation between episode length and length of the preceding silent phase. In addition, early evoked episodes always started with lower cycle frequencies than spontaneous episodes.

The states of the storage pool size and the active pool size in the second and third panel of Figure 3.3 show the intimate relationship between pool depletion (episodes terminate when the storage pool is emptied, cycles terminate when the active pool is emptied) and network activity in our model.

Finally, in the bottom panel of the figure, the corresponding time course of the delayed postsynaptic depression of a sample neuron is shown. This depression variable is a multiplicative factor of the effective synaptic strength. Thus, a value of 1 means no depression. We allowed this variable to assume slightly negative values, which reproduced qualitatively the observed hyperpolarization of membrane potentials shortly after an episode (data not shown).

3.3 Discussion

We showed how spontaneous activity in developing networks, which is remarkably similar across network architectures and neuron types involved, can emerge from a few general principles of synaptic dynamics. The key idea is exhaustion of presynaptic vesicle pools due to the extreme, self-sustaining activation in these recurrently connected hyperexcitable networks. We introduced a new synapse model that aims to capture recent findings about vesicle release properties in a quantitative manner.

In our chick spinal cord model, episodes are triggered by increasing amounts of spontaneously released synaptic vesicles in absence of stimulation. The length of an episode is determined by the overall amount of vesicles at the start of the episode, while repeated depletion of readily releasable vesicles produces the cyclic nature within episodes. In contrast to previous modeling approaches, slow post-synaptic depression does not terminate episodes; its importance with respect to the observed activity pattern lies in removing network hyperexcitability after episodes.

We deliberately kept the description of the dynamics of all processes involved as simple as possible, to illustrate the generality of the model: not the fine details of the processes involved are important but the overall functional organization together with the relative sizes of time constants and pool sizes. We showed how such a setup can naturally account for the observed dynamics.

We will now review the list of experimental findings for the developing chick spinal cord that need to be accounted for, and discuss for each point briefly the explanation provided by our model. To distinguish predictions from known data (according to the experimental studies cited above), we mark experimentally established data with (EE).

1. *Recurring spontaneous initiation and termination of episodes* Episodes are spontaneously initiated by the combination of increasing synaptic noise in absence of activity (EE), together with decreasing post-synaptic depression (EE). Episodes are spontaneously terminated by exhaustion of the vesicle reservoir in the presynaptic terminal.

2. *Cyclic activity within episodes* Cyclic activity within episodes is caused by repeated exhaustion of readily releasable vesicles. When the active pools of most synapses are exhausted, network activity drops to a low

level until sufficiently refilled active pools can support a new avalanche once again.

3. *Cycle frequency decreases during episode* Increasing transition times from the shrinking storage pool to the active pool during the course of an episode causes slower refilling of the active pool up to the point where a new avalanche of activity can be started.

4. *Shorter episodes start with lower cycle frequency; episode length can be predicted well by initial cycle frequency* Episodes are shorter when the storage pool is only partly refilled at the beginning of the episode. Since the transition rate from storage pool to active pool is determined by the size of the storage pool, we can predict the size of the storage pool (and thus the remaining length of the episode) from the cycle frequency.

5. *Positive correlation between episode duration and length of preceding silent phase; episode termination is thus deterministic* At the beginning of the silent phase, the storage pools are empty. Thus, the length of the preceding silent phase determines the size of the storage pool at the end of the silent phase. Using the above arguments, we can thus predict when the episode terminates from the length of the silent phase.

6. *Initiation of episodes is stochastic* The time for spontaneous initiation depends on the coincidence of sufficiently many spontaneous synaptic events, which in turn are stochastic.

7. *Episodes have an absolute refractory period: not possible to evoke an episode prematurely for about 100 sec after termination of an episode* Transient loss of hyperexcitability after the end of an episode (EE) prevents avalanches of activity, even under stimulation.

8. *Slow synaptic depression sets in after the end of an episode and reaches its peak about 1–1.5 minutes later* Slow synaptic depression is activity-dependent but time delayed. Models that assume a causal role for the termination of episodes cannot account for this fact.

9. *Network architecture not important (same phenomenon in retina, hippocampus, and many other networks)* Provided that avalanches of excessive network activity can spread quickly enough, the synaptic ex-

haustion processes will be sufficiently synchronized to determine the network dynamics.

Our model can indirectly account for the change in episode dynamics during development of the chick spinal cord. It is known that the length of episodes increases with age of the embryo, and the activity between cycles remains higher. We expect such an effect in our model, if we incorporate vesicle recycling that gets more efficient with the age of the synapse.

There is little doubt that the existence of spontaneous activity is important for the development of neural systems and the entire organism. However, it has not been clear why it has exactly the dynamics that are observed. These dynamics are surprisingly stable even when the network is disrupted: it gets re-established with almost the same time course after recovery (Chub and O'Donovan, 1998). The recovery to practically the same activity pattern led to speculations that the particular shape of the time course must be important – why else would the system take the efforts re-establish this particular pattern? Our answer to this question in light of our model is surprisingly simple: The exact time course may not be important, or at least not actively regulated; rather, it might be difficult to *avoid* such a pattern given the excessive self-stimulation in hyperexcitable networks in combination with the dynamics of limited synaptic resources.

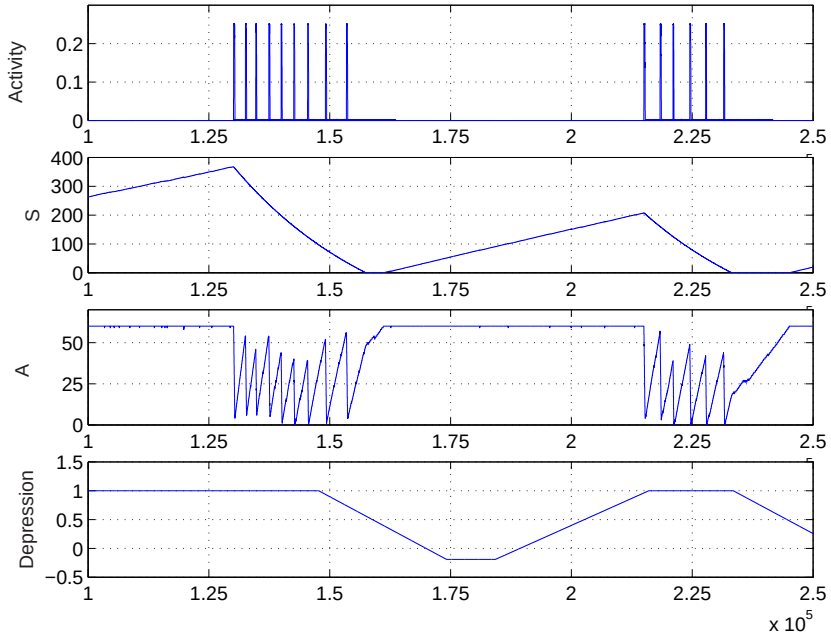

Figure 3.3: Network activity, pool sizes, and postsynaptic depression. *Top panel:* Network activity with two episodes. The first episode initiated spontaneously, while the second one was evoked prematurely, resulting in a shorter episode length. *Second panel:* Time course of the storage pool size S. Episode termination is caused by depletion of the storage pool. *Third panel:* Time course of the active pool A. Avalanches in network activity cause depletion of A, which in turn stops the avalanche and the cycle. *Bottom panel:* Slow, time-delayed postsynaptic depression. The postsynaptic depression removes the hyperexcitability from the network, causing an effective refractory period for initiating new avalanches.

Chapter 4

Mean field models – Cortical networks

Irregular firing of cortical neurons in response to sensory stimuli is a well studied phenomenon (for studies in visual cortex see, e.g., Heggelund and Albus, 1978; Dean, 1981; Vogels et al., 1989; Shadlen and Newsome, 1998; Gershon et al., 1998). Firing statistics tend to be close to that of a completely random Poisson process, raising the question where this apparent noisiness stems from. Considering known anatomical properties of the cortex (high connectivity) and physiological properties of cortical neurons (long passive time constant, absence of noisy behavior *in vitro*), the picture gets even more puzzling, and simple neuron models, like point integrate-and-fire neurons, seem to be inadequate to capture cortical dynamics (Softky and Koch, 1993). Shadlen and Newsome (1998) pointed out that a balance between excitation and inhibition, with an average membrane potential just below threshold and firing caused by fluctuations in the membrane potential, can account for such irregular firing even for integrate-and-fire neurons with long membrane time constants. That such a balance with stable irregular firing can naturally emerge from a simple feedback mechanism (without the need to fine-tune any parameters) between excitation and inhibition was first shown by van Vreeswijk and Sompolinsky (1996, 1998) for a mean-field model of a network of binary neurons. Amit and Brunel (1997a,b); Brunel (2000) investigated a similar model, developing a mean-field theory for a balanced network of integrate-and-fire neurons and showing the existence of stable asynchronous states with irregular firing. However, these mean-field theories cannot address the question of how the amount of firing irregularity

(or, more generally, the nature of firing correlations) is controlled by intrinsic network parameters: The van Vreeswijk-Sompolinsky treatment does not use spiking neurons and the Amit-Brunel treatment is not self-consistent regarding firing correlations because it assumes the neuronal input to be white noise. It was first shown in a preliminary study by Hertz et al. (2003) how a balanced all-inhibitory network with excitatory external input can be treated fully self-consistently, using the systematic formulation of mean-field theory due to Fulvi Mari (2000). Here, we show how to extend this basic model in several ways: In Section 4.1, we analyze a model of a generic cortical column with an excitatatory and an inhibitory population, stimulated by (excitatory) external input. Then, in Section 4.2, we show how to include conductance-based synapses in the self-consistent description of a cortical column. The analysis and results for this model allows us to elucidate how firing irregularity and high-conductance states of cortical neurons *in vivo* are related. Finally, in Section 4.3, we include functional architecture into the mean-field description and study a model of an orientation hypercolumn in primary visual cortex. We will see how several, seemingly independent experimental findings can be accounted for by the model. In addition, we make quantitative predictions on tuning properties of both neuronal firing and neuronal input currents.

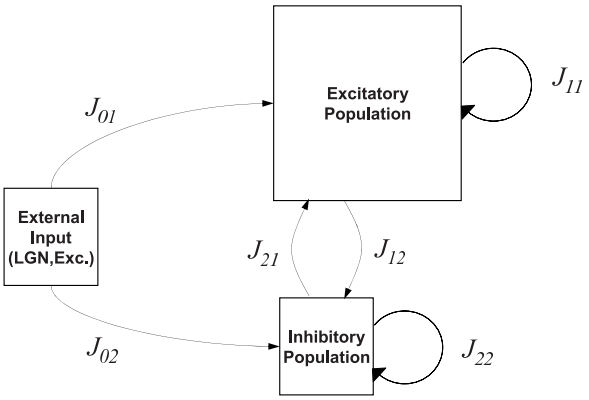

Figure 4.1: Structure of the single-column model.

4.1 Cortical column: current-based synapses

In our first cortical network model, we connect neurons with "current-based synapses", i.e., we model the effect of a presynaptic spike as a stereotyped postsynaptic input current. Since we consider synapses with dynamics that are fast compared to the membrane time constant of our integrate-and-fire neurons, we simplify the spike response to a jump in the membrane potential by an amount that is determined by the "strength" of the synapse.

4.1.1 Model and methods

The network model for the cortical column consists of an excitatory population of size N_1 and an inhibitory one of size N_2. The neurons are connected randomly so that any given neuron in population $a \in \{1, 2\}$ receives synaptic input from any other neuron of population $b \in \{1, 2\}$ with probability K_b/N_b. In our calculations, we take this probability to be 10%, independent of both the source population b and the target population a. (Throughout this chapter, we use the indices a and b to denote target and source population, respectively.) The strengths of the resulting synapse J_{ij}^{ab} between presynaptic neuron j and postsynaptic neuron i is taken to be $J_{ab}/\sqrt{K_b}$ with J_{ab} of order 1, respectively $J_{ij}^{ab} = 0$ if there is no connection. With this scaling, the mean input from the (on average) K_b presynaptic neurons of population b is of order $K_b/\sqrt{K_b} = \sqrt{K_b}$, but the fluctuations are of order 1 (cf. van Vreeswijk and Sompolinsky, 1996), the same order as the distance

between reset and threshold of our model integrate-and-fire neurons. While this scaling will turn out to be convenient for our balanced network model, it also translates to synaptic strengths in the right biological ballpark, given (the realistic range of) K_b on the order of a few thousand.

The cortical column receives excitatory input from an external population indexed by 0. Completely analogously as for the intracortical connections, we assume random connectivity and synapses of strength $J_{a0}/\sqrt{K_0}$.

We can then write the subthreshold dynamics of the membrane potentials as

$$\frac{\mathrm{d}u_i^a}{\mathrm{d}t} = -\frac{u_i^a}{\tau} + \sum_{b=0}^{2}\sum_{j=1}^{N_b} J_{ij}^{ab} S_j^b(t), \tag{4.1}$$

where $S_j^b(t) = \sum_s \delta(t - t_{jb}^s)$ is the spike train of neuron j in population b. For simplicity, we chose the membrane time constant τ to be the same for all neurons, but we sampled the firing thresholds from a narrow normal distribution with mean 1 and standard distribution 0.1. The reset is fixed at zero. We do not implement an absolute refractory period since we consider a low-rate regime only.

Because of the dilute random connectivity, each neuron receives a high number of uncorrelated inputs (van Vreeswijk and Sompolinsky, 1998), provided that the network operates in a low rate regime. We can thus employ the central limit theorem and describe the input as a Gaussian random process, for which we need to determine the mean and correlation function self-consistently (see Fulvi Mari, 2000, for a formal proof of this general mean field approach).

In this chapter, we will consider stationary firing rates only, to keep the notation and explanations more transparent – without loss of generality. There is no qualitative difference between stationary and non-stationary rates mainly for two reasons: First, the theory is general enough that no re-formulation is needed to include non-stationary rates. Replacing time averaged quantities (e.g., rates r_b) in the following treatment with their time dependent counterparts (e.g., $r_b(t)$) yields the more general time-dependent case. Second, the highly connected network converges to the balanced state very quickly, with stability eigenvalues proportional to $\sqrt{K_0}$ (van Vreeswijk and Sompolinsky, 1998). Thus, in the mean-field limit of $K_0 \to \infty$, the network follows changes in the mean input instantaneously, which keeps it in a quasi-stationary state. For biologically realistic values of finite, but large connectivity (as in our numerical calculations) the network is still fast enough

to follow smooth rate changes on the millisecond scale. However, more care needs to be taken for calculating the solutions numerically with the algorithm described below, as discussed in Lerchner et al. (2004c).

To separate the mean of the currents from their fluctuations, it is convenient to apply such separations to the description of both the synaptic weights J_{ij}^{ab} and the spike trains $S_j^b(t)$ in (4.1):

$$J_{ij}^{ab} = \overline{J_{ij}^{ab}} + \delta J_{ij}^{ab} \tag{4.2}$$

$$S_j^b(t) = r_b + \delta r_j^b + \delta S_j^b(t). \tag{4.3}$$

Throughout this chapter, we use the bar-notation for averaging over neuron populations, which will always apply to the running index j. Here, $r_b = \overline{r_j^b} = 1/N_b \sum_j r_j^b$ is the average rate of the neurons in population b. The difference between this average rate and the rate of the individual neuron j is denoted δr_j^b. These two components are both static, describing time-averaged quantities. The temporal fluctuations of the spike train and their possible correlations in time are captured by the third term on the right-hand side of (4.3), $\delta S_j^b(t)$. With help of this separation, we can calculate a description of the recurrent currents of the following form:

$$I_a^{\mathrm{rec}}(t) = \sum_{b=1}^{2} J_{ab} \left[\sqrt{K_b} r_b + \sqrt{1 - K_b/N_b} \left(\left(\overline{(r_j^b)^2} \right)^{\frac{1}{2}} x_b + \xi_b(t) \right) \right] \tag{4.4}$$

We have dropped the neuron index i because such a statistical description of the input current reduces the network problem to single neuron problems, one for each population. For the external currents, a similar formulation – albeit simpler, because of the pure feed-forward nature – can be made. Following we will provide a qualitative description of Equation (4.4). (For a more detailed account with a slightly different notation see Lerchner et al., 2004c). The square brackets on the right-hand side of (4.4) contain three different terms. The first one is the population mean, $\sqrt{K_b} r_b$. It is much larger (of order $\sqrt{K_b} \gg 1$) than the other two, which are both of order 1. In the second term, x_b is a unit-variance random number, causing a random offset of the mean. This offset is scaled by the two factors $\sqrt{1 - K_b/N_b}$ and $\overline{(r_j^b)^2}$, which reflect the facts that individual neurons vary in their number of inputs and that there is a distribution of rates among the input neurons. Thus, sampling values for x_b corresponds to sampling individual neurons in the population. We will come back to this concept when we discuss how the input currents can

be evaluated numerically. Finally, the third term in the square brackets of (4.4) describes the fluctuations in the recurrent input current, together with their temporal correlations. The expression $\xi_b(t)$ stands for a realization of a Gaussian random process obeying

$$\langle \xi_b(t)\xi_b(t')\rangle = C_b(t-t'), \tag{4.5}$$

where $C_b(t-t')$ denotes the average autocorrelation function of the fluctuations in the spike trains of neurons in population b, given by

$$C_b(t-t') = \frac{1}{N_b}\sum_{j=1}^{N_b}\langle \delta S_j^b(t)\delta S_j^b(t')\rangle. \tag{4.6}$$

With the operation $\langle\cdot\rangle$ we mean averaging over "trials", i.e., realizations of random processes, such as stochastic spike trains. Like the static fluctuations in the mean input, these dynamic fluctuations also contain the scaling factor $\sqrt{1-K_b/N_b}$ that reflects the random connectivity. For a fully connected network this factor is zero because then there can be no randomness in the connectivity. Therefore, the temporal fluctuations in the input currents (and thus ultimately in the neuronal firing) are due to the static fluctuations in the connectivity. Thus, this model offers an explanation of how the apparent noisy firing statistics of cortical neurons *in vivo* may originate in the quenched disorder of the cortical connectivity, without requiring any stochastic processes.

The balance condition

As mentioned earlier, we are interested in finding self-consistent solutions to this model for *low rates*. In analyzing the recurrent input currents (4.4), we have observed that the means of both the excitatory and the inhibitory inputs are large, of order $\sqrt{K_b} \gg 1$ for $b = 1, 2$, respectively. If either of the inputs dominates, the neuron would fire at extremely high rates (dominant excitation) or not at all (dominant inhibition). The only non-trivial case is a balance of the mean input currents so that the mean membrane potential stays just below threshold, with occasional threshold crossings of the membrane potential caused by the fluctuations in the input. Such a *balance condition* has to apply to the overall input, including the external one, and can be formalized using our mean field description as

$$\sum_{b=0}^{2} J_{ab}\sqrt{K_b}r_b = \mathcal{O}(1). \tag{4.7}$$

At this point, it may appear that such a balance requires a delicate fine tuning of the parameters, specifically of the synaptic strengths J_{ab}. This, however, turns out not to be the case. It is due to the negative feedback within the highly recurrently connected cortical network that such a balance is established dynamically for a large range of inputs without being sensitive on exact values of J_{ab}, as long as certain constraints on their relative sizes are met (such as that cortical inhibition has to be stronger than cortical excitation). This robustness with respect to values of the synaptic strengths is one of the reasons why balanced network models can offer such an attractive explanation for cortical dynamics.

Within our model, the balance condition is simply a consequence of our requirement to find self-consistent solutions for the rates and the correlation functions. We can use the formulation of the balance condition (4.7) to calculate analytically good estimates for the rates. This is in general not possible for the correlation functions, for which we will instead present a numerical procedure below. (The numerical procedure also provides *exact* solutions for the rates and the rate fluctuations). To estimate the rates, we can divide (4.7) by $\sqrt{K_0}$ and neglect the resulting term of $\mathcal{O}(1/\sqrt{K_0})$ on the right-hand side of the equation to get

$$\sum_{b=0}^{2} \hat{J}_{ab} r_b = 0, \tag{4.8}$$

with $\hat{J}_{ab} = J_{ab}\sqrt{K_b/K_0}$. These are two linear equations in the two unknowns r_a, $a = 1, 2$, with the solution

$$r_a = \sum_{b=1}^{2} [\hat{\mathsf{J}}^{-1}]_{ab} J_{b0} r_0, \tag{4.9}$$

where $\hat{\mathsf{J}}^{-1}$ is the inverse of the 2×2 matrix with elements $\hat{J}_{ab}$, $a, b = 1, 2$.

Numerical procedure

To find self-consistent solutions for the complete mean field model, we employ an iterative algorithm that may be summarized as follows:

1. Start with initial estimates for the firing statistics, i.e., the rates, the rate fluctuations, and the correlation functions. For the rate estimates,

we use the approximation derived above. We neglect the rate fluctuations in the first step and use a white noise approximation for the correlation functions.

2. Use the estimates of the firing statistics to compute many realizations (trials) of the Gaussian random processes describing the input currents. By drawing new values for x_b in (4.4) from trial to trial, we effectively sample over the population, which is important to finally obtain self-consistent solutions for the rate fluctuations.

3. Simulate single neurons driven by the input currents calculated in step 2 and collect their firing statistics.

4. Calculate improved estimates of the rates, rate fluctuations and correlations with help of the statistics derived in step 3.

5. Repeat steps 2–4 until convergence.

Such an algorithm was first used to calculate the remanent magnetization of a mean field model for spin glasses by Eisfeller and Opper (1992).

After convergence of the iterative procedure, we know the solutions for the population statistics. These can then be used to collect firing statistics from *individual* neurons by keeping the values for x_b fixed over trials.

4.1.2 Results

We are interested in the question of which model parameters, and ultimately which dynamic mechanisms, control the amount of irregularity in the neuronal firing. One major finding of the present study is that firing irregularity depends mainly on the strength of the synapses: stronger synapses generally lead to higher irregularity in spike counts, irrespective whether excitatory, inhibitory, or both kind of synapses are increased in strength. For this reason, we keep the relative sizes between excitation and inhibition fixed in this brief presentation of the main results and refer to Lerchner et al. (2004c) for a more detailed treatment. Here, we use the following values for the parameters J_{ab}:

$$\begin{pmatrix} J_{11} & J_{12} \\ J_{21} & J_{22} \end{pmatrix} = J_s \begin{pmatrix} 0.5 & -2 \\ 1 & -2 \end{pmatrix}, \tag{4.10}$$

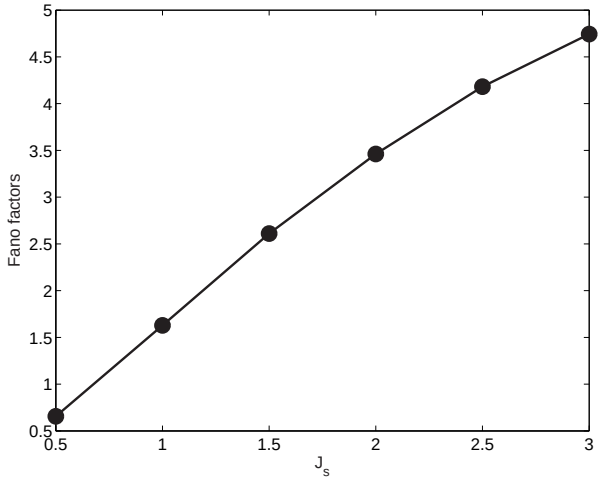

Figure 4.2: Firing irregularity depends on synaptic strengths. Fano factors are plotted as a function of the overall factor for synaptic strengths J_s. Stronger synapses generally lead to higher Fano factors. There is a smooth transition from subpoissonian ($F < 1$) over Poisson ($F = 1$) to superpossionian ($F > 1$) firing statistics.

for the relative sizes of the intracortical connections, and $J_{10} = J_s \cdot 1$, $J_{20} = J_s \cdot 0.5$ for the connections from the external population. We vary the scaling factor J_s to study the effect of synaptic overall-strength on the amount of firing irregularity. To characterize the irregularity in firing itself, we use the Fano factor F, defined as the variance in spike count divided by the mean spike count. For a Poisson process $F = 1$, while $F \neq 1$ always implies temporal correlations in the spike times: $F > 1$ ("superpoissonian" statistics) indicates a tendency towards "bursty" spiking behavior, and $F < 1$ ("subpoissonian") indicates a tendency towards well-separated spikes.

Figure 4.2 shows that stronger synapses, reflected by increasing values of the synaptic scaling factor J_s, lead to higher Fano factors. The transition from subpossionian to superpossonian behavior is smooth, without any special role for $F = 1$, the temporally uncorrelated case. Fano factors obtained from measurements on cortical neurons *in vivo* tend to lie in the same range.

In Figure 4.3 we compare data obtained from experiments with numerical results of our model. The experimental data shown in the left panel are from Gershon et al. (1998), where monkeys were presented with flashed, stationary visual patterns, and spike statistics from neurons in primary visual cortex

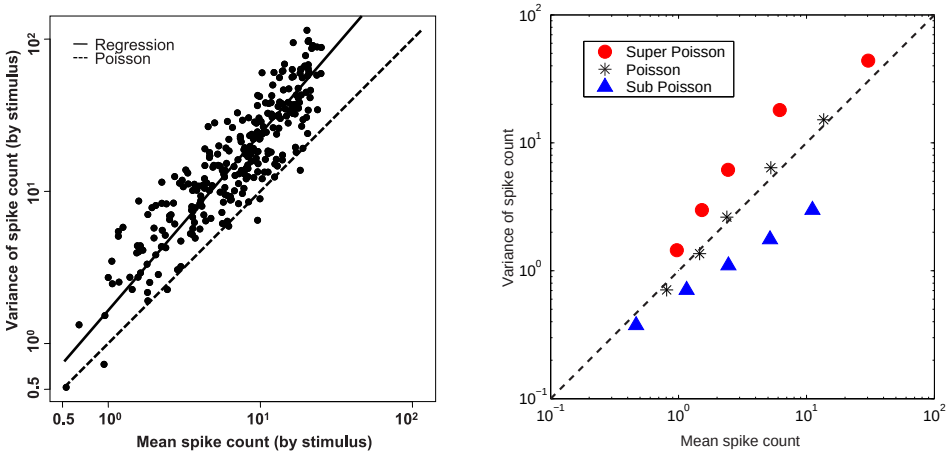

Figure 4.3: Approximately linear relationship between the logarithm of the spike count variance and logarithm of the mean spike count in both experiment and model. *Left panel:* Experimental data obtained from a single neuron in monkey V1 *in vivo*. Each point reflects the statistics obtained for a specific stimulus (a flashed stationary pattern). The sampling error due to the limited number of trials can account for all of the variation from the regression line. *Right panel:* Numerical results for three different values of overall synaptic strength J_s (red circles: 1.25, black stars: 0.75, and blue triangles: 0.375), varying the external input rate. For $J_s = 1.25$ the data look qualitatively like those from the experiments shown in the left panel.

(V1) were collected. The logarithm of the spike count variance is plotted against the logarithm of the spike count mean for a single ("typical") neuron. Each data point reflects the statistics obtained for a specific stimulus. The following general trends can be seen: First, the ratio between variance and mean (i.e., the Fano factor F) is systematically higher than one, which is typical for spike statistics obtained from neurons in V1. Second, the relationship between the log-variance and the log-mean is approximately linear, indicating a good fit of the variance by a power law function of the mean count. In the right panel of Figure 4.3 we show numerical results of our model for three different values of overall synaptic strength ($J_s = 1.25, 0.75,$ and 0.375), varying the input rate r_0. For $J_s = 1.25$ the data look qualitatively like those from the experiments. The approximate linear relationship is evident for all three cases with the exception of a deviation towards lower variance at high rates that is caused by an increasing role of refractoriness

in limiting the possible variability of interspike intervals.

4.1.3 Discussion

We showed how to do a complete mean-field theory for a network of integrate-and-fire neurons that allows one to calculate self-consistently not only the firing rates, but also the correlation functions of the firing. The solutions can in general not be calculated analytically, so we complemented the theory with a numerical procedure for solving the equations. This treatment makes it possible to quantify the amount of irregularity as a function of network parameters. For the specific model analyzed here, the firing irregularity is mainly controlled by the strength of the synapses.

Why do stronger synapses lead to higher irregularity in spike counts, as shown in Figure 4.2? The membrane potential performs a random walk in a quadratic potential, resulting in an approximately Gaussian distribution. There is an absorbing barrier at threshold $\theta = 1$ and re-injection at the reset level (equal to 0). Between any two spikes, the membrane potential has to "diffuse on top of a drift" from the reset to the threshold. While the distance between reset and threshold is fixed, the step size of the random walk is scaled by the synaptic strengths: each input spike from a neuron in population b to a neuron in population a causes a jump in the membrane potential proportional to J_{ab}. The drift is always in the direction of the bottom of the quadratic potential, where the mean of the membrane potential distribution lies. Since J_{ab} controls the step size, it also controls the width of the distribution. The dependence can be shown by integrating the fluctuating part $\delta I_i^{ab}(t)$ of the input current $I_i^{ab}(t)$ (see Equation (4.4)). Defining

$$\delta I_i^{ab}(t) = J_{ab}\sqrt{1 - K_b/N_b}\xi_i^{ab}(t), \tag{4.11}$$

we can calculate the variance of the membrane potential distribution as

$$\langle \delta V_{ab}(t)\delta V_{ab}(t')\rangle = J_{ab}^2(1 - K_b/N_b)\int \mathrm{d}t_1\mathrm{d}t_2 e^{-(t-t_1)/\tau}e^{-(t'-t_2)/\tau}C_b(t,t'), \tag{4.12}$$

where we have ignored the threshold effect, which makes no qualitative difference in the low-rate limit we are interested in. Note that the width is scaled by J_{ab} but not by the number of inputs – only the connection probability K_b/N_b enters. So we see that stronger synapses result in bigger step sizes of the random walk and thus in broader membrane potential distributions.

The mean of the distribution is always lower than threshold. For a narrow distribution, it lies close to threshold (above reset), while it can lie well below reset for a wide distribution.

Let us first consider the case of weak synapses – and why they lead to more regular spike trains (subpoissonian firing statistics): Between any two spikes, the membrane potential has to drift from the reset up to the mean of the distribution close to threshold before the small steps of the random walk can drive the neuron above threshold. Thus, a threshold crossing immediately after reset is extremely unlikely, which leads to an effective relative refractory period after each spike, reflected in lower Fano factors.

We now consider strong synapses, which lead to more irregular spike counts (superpoissonian statistics): In this case the step size of the random walk is so big that the membrane potential can bridge the distance between reset and threshold with only a few steps. This leads to a higher probability for very short intervals, i.e., a tendency towards bursting. Considering a fixed rate, the higher number of shorter interspike intervals is necessarily counterbalanced by a higher number of longer intervals. These longer intervals are due to a drift *away* from the threshold after reset, since the mean of the distribution lies below reset in this case, as mentioned above.

Our analytical estimation of the firing rates (4.9) tells us that there is a linear relationship between the strength of the external input (firing rate r_0) and the firing rate r_a of the cortical neurons for both excitatory ($a = 1$) and inhibitory ($a = 2$) neurons. Figure 4.4 shows simulation results confirming this kind of dependence. It has been known for a long time that cortical neu-

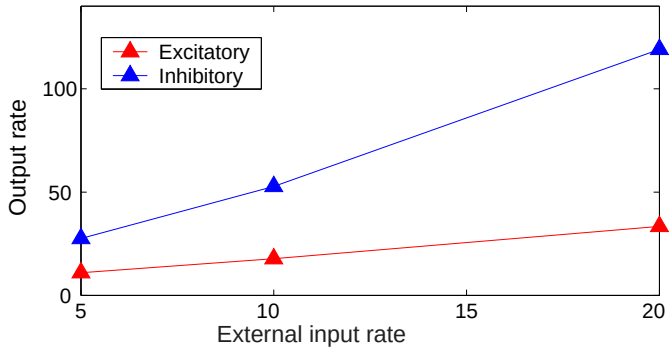

Figure 4.4: Linear relationship between firing rate of the external input and output rate for both excitatory and inhibitory neurons.

rons exhibit such a linear input-output relationship, but the mechanisms behind this somewhat surprising property (considering the many non-linearities in the system) have not been clear. Consequently, models of cortical networks – especially on the level of rate models – have frequently assumed a threshold-linear gain function for single cortical neurons, despite the fact that single neurons *in vitro* do not show such behavior. Clearly, we are observing a network effect here. At first sight, it may seem that there must be a very precise computational system at work that manages to compensate all the non-linearities involved in the system. However, our model provides a much simpler answer to this puzzle (the same answer was already obtained by van Vreeswijk and Sompolinsky, 1996, 1998, for their model using binary neurons). We have derived this property, reflected in Equation (4.9), from the balance condition, which in turn expresses the quasi-steady state achieved by the negative feedback between excitation and inhibition. (Note that we do not impose balance artificially – it is a consequence of the dynamics, not an assumption entering the model.) Negative feedback is thus a necessary ingredient, but there is a more fundamental reason for the linearity. It is rooted in the fact that the sum over a high number of (practically) uncorrelated random processes always results in a random process with very simple properties. This fact is captured mathematically in what is known as the Central Limit Theorem, which lies at the very heart of our mean-field description.

Following the argumentation above, our model not only provides possible explanations for several – seemingly independent – properties of cortical dynamics. It also provides a straightforward answer to a basic issue that has been discussed heatedly between neuroscientists for many years: *"Why is neuronal firing in the cerebral cortex noisy? Is it really noisy? Or do we just not understand the complexity of the neuronal code and the timing of single spikes is important?"* The maybe surprising answer in the light of our theory is:

Neocortical neurons fire noisily because it makes the computation much simpler and much more robust.

However, such an answer raises more questions. Where does the noise come from? Individual neurons behave perfectly deterministic when removed from the brain and put into a dish. Where are the noisy processes in the brain? The solution provided by our model is:

There is no need for noisy processes. The brain could be perfectly deterministic and still appear to be noisy.

What anatomists observe is "static noise" in the connectivity. What we

conclude is that the network converts it into apparent temporal noise in the neuronal firing. We say 'apparent' because in our network there is not really noise at all: what we are facing is a state of deterministic chaos, indistinguishable from "real" noise for all practical purposes (such as collecting firing statistics from neurons in the visual cortex in response to repeated stimuli).

We like to point out that the very concept of "static noise" (or "random connectivity", as we normally refer to it) is only meaningful in systems with many parts. The same is true for the validity of using the Central Limit Theorem as a base of description. This means that the study of *large* spiking neural networks can offer explanations for cortical dynamics that cannot be derived from models with only a limited number of neurons.

4.2 Cortical column: conductance-based synapses

In the following mean field model, we will use a conductance-based description for the effect of a presynaptic spike on the target neuron. This will add more realism because the amount of input current that enters the cell through ion channels in the postsynaptic membrane really depends on the momentary voltage across that membrane.

Even though we still consider synaptic dynamics that are fast compared to the (passive) membrane time constant of our integrate-and-fire neurons, we will now include a more detailed description of the synaptic dynamics. In contrast to the neuron model with current-based synapses, it will turn out that now the effective membrane time constant can become so short that even fast synaptic dynamics have strong influence on the firing statistics.

4.2.1 Model and methods

We use the same model for the generic cortical column as described above for the case of current-based synapses, i.e., the same random connectivity between the two populations and from the external population. The main difference is the effect of presynaptic spikes on the target neurons. With conductance-based synapses, the equation of motion for the subthreshold membrane potentials becomes

$$\frac{\mathrm{d}u_i^a(t)}{\mathrm{d}t} = -g_\mathrm{L} u_i^a(t) - \sum_{b=0}^{2} \sum_{j=1}^{N_b} g_{ij}^{ab}(t)(u_i^a(t) - V_b). \tag{4.13}$$

The conductance g_L pertains to the "leaky" part of the integrate-and-fire neuron. It therefore corresponds to the inverse of the membrane time constant τ in (4.1). The synaptic reversal potentials are denoted V_b, where $V_0 = V_1$ is higher than threshold (excitatory synapses) and V_2 is below the normalized resting potential of 0 (inhibitory synapses). The time-dependent conductances $g_{ij}^{ab}(t)$ describe the postsynaptic conductance changes of neuron i in population a in response to the presynaptic spike train $S_j^b(t)$ of neuron j in population b:

$$g_{ij}^{ab}(t) = \frac{g_{ab}^0}{\sqrt{K_b}} \int_{-\infty}^{t} \mathrm{d}t' K(t - t') S_j^b(t'), \tag{4.14}$$

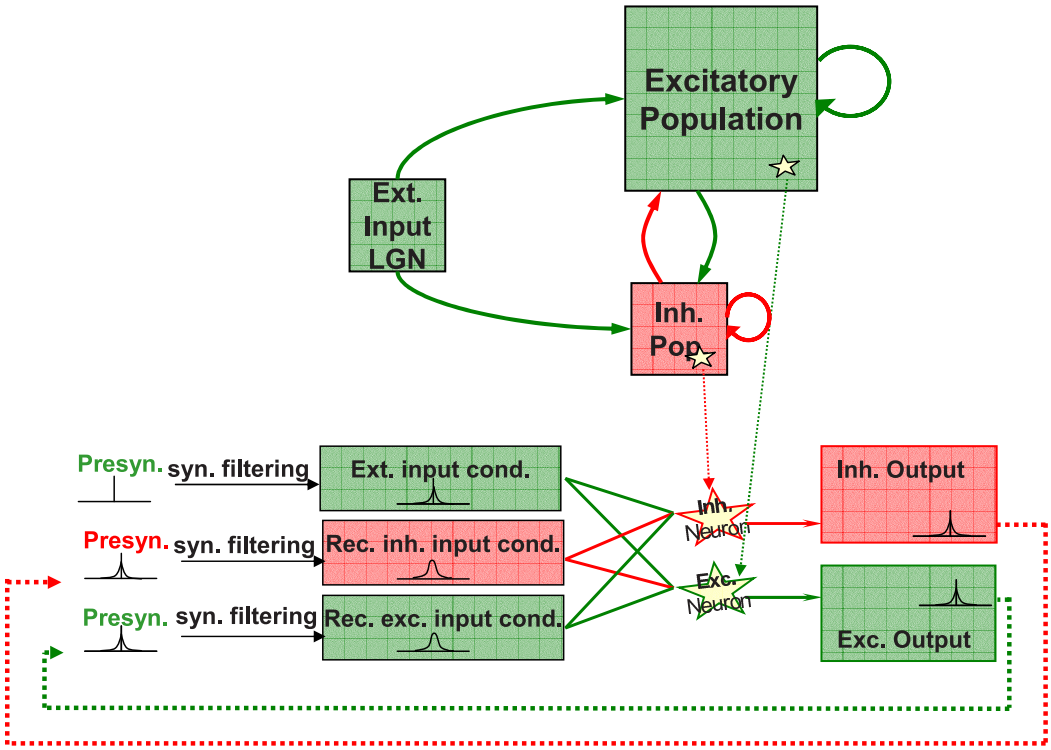

Figure 4.5: Diagram of the network structure and of the algorithm for the conductance-based model. The importance of coloring the noise correctly is indicated by the shapes of the autocorrelations in the lower part of the figure (Presyn., input cond., and Output).

if there is a connection between the two neurons, otherwise $g_{ij}^{ab}(t) = 0$. The parameters g_{ab}^0 are taken to be of order 1, which results in an analogous scaling of the synaptic strengths as in the current-based model. The filter kernel $K(t)$ describes the temporal shape of the conductance change in response to a single spike. We take it to be

$$K(t) = \frac{1}{\tau_2 - \tau_1} \left(\mathrm{e}^{-t/\tau_2} - \mathrm{e}^{-t/\tau_1} \right), \tag{4.15}$$

where τ_1 is the average opening time and $\tau_2 > \tau_1$ is the average closing time of the ion channels in the postsynaptic terminal. Figure 4.6 shows the shape of the kernel $K(t)$ given by (4.15) for $\tau_1 = 1$ ms and two different values of τ_2. Note that the integral over the kernel is normalized to 1.

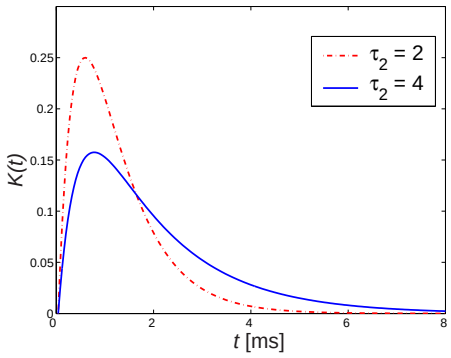

Figure 4.6: Shape of the conductance kernel $K(t)$ defined in (4.15) for a characteristic opening time of $\tau_1 = 1$ ms and two different values of the closing time τ_2. .

We can derive a mean field description of this model using the same principles and techniques as for the one with current-based synapses if we perform the separation of mean and fluctuations on the conductances g^0_{ab} instead of on the weights J_{ab} (see Equation (4.2)). We then obtain statistical descriptions of the overall "input conductances" $g_{ab}(t)$ and the network problem becomes once again one for single neurons that obey

$$\frac{\mathrm{d}u_a(t)}{\mathrm{d}t} = -g_{\mathrm{L}}u_a(t) - \sum_{b=0}^{2} g_{ab}(t)(u_a(t) - V_b). \tag{4.16}$$

Similar to the input currents (4.4) in the current-based model, we can calculate the following statistical description of the time dependent conductances:

$$g_{ab}(t) = g^0_{ab}\left[\sqrt{K_b}r_b + \sqrt{1 - K_b/N_b}\left(\left(\overline{(r^b_j)^2}\right)^{\frac{1}{2}} x_b + \xi_b(t)\right)\right]. \tag{4.17}$$

There is a difference between $\xi_b(t)$ in (4.4) and (4.17) due to the synaptic dynamics that we have included in the conductance-based model. Now we need the correlation functions of the *synaptically filtered* spike trains to obtain the Gaussian random processes describing the temporal fluctuations and correlations:

$$\langle \xi_{ab}(t)\xi_{ab}(t')\rangle = \tilde{C}_b(t - t') \tag{4.18}$$

with

$$\tilde{C}_b(t - t') = \int_{-\infty}^{t} \mathrm{d}t_1 K(t - t_1) \int_{-\infty}^{t} \mathrm{d}t_2 K(t - t_2) C_b(t_1 - t_2). \tag{4.19}$$

The balance condition

As for the model with current-based synapses, we are looking for low-rate solutions, for which the mean input currents must nearly cancel. We cannot read off these currents directly from the mean field description (4.17), as we could do for the analogous case for the current-based synapses (4.4). Now the input currents depend on the momentary levels of the membrane potential $u_a(t)$. However, in a low-rate steady state we can replace $u_a(t)$ by its time-averaged mean value $\overline{u_a}$ and use (4.16) to arrive at a balance condition for the conductance-based model.

$$\sum_{b=0}^{2} g_{ab}^0 \sqrt{K_b}(\overline{u_a} - V_b) = \mathcal{O}(1), \tag{4.20}$$

where we have absorbed the small contribution of the leakage term into the right-hand side of this equation. In order to solve it for getting estimates of the firing rates, we first need to know the mean membrane potential $\overline{u_a}$. We can approximate it by the firing threshold θ_a for reasons that will become clear shortly. Then the solution can be found exactly analogously to the previous model, and we get

$$r_a = \sum_{b=1}^{2} [\hat{\mathsf{J}}_{\text{eff}}^{-1}]_{ab} \tilde{J}_{b0}^{\text{eff}} r_0, \tag{4.21}$$

where $\tilde{J}_{ab}^{\text{eff}} = g_{ab}^0(V_b - \theta_a)$, $\hat{J}_{ab}^{\text{eff}} = \tilde{J}_{ab}^{\text{eff}} \sqrt{K_b/K_0}$, and $\hat{\mathsf{J}}_{\text{eff}}^{-1}$ is the inverse of the matrix with the elements $\hat{J}_{ab}^{\text{eff}}$, $a, b = 1, 2$ (cf. Equation (4.9)).

High conductance states

Leaky integrate-and-fire (LIAF) neurons with conductance-based synapses are qualitatively different from LIAF neurons with current-based synapses. Generally, the membrane time constant of the neuron is inversely proportional to the membrane conductance: the higher the conductance, the lower the membrane time constant and the faster the membrane potential dynamics of the neuron. The only membrane conductance entering the current-based description is the fixed (i.e., passive) leakage conductance determining the (passive) membrane time constant τ, as in the equation for the subthreshold dynamics of the current-based model (4.1). In contrast, the equation for the membrane potential dynamics in the conductance-based case (4.13) contains

in addition as many time-dependent conductance terms as there are synaptic contacts. By rearranging (4.13) into the same form as (4.1), with a leakage term and a current-source term, we get

$$\frac{\mathrm{d}u_i^a(t)}{\mathrm{d}t} = -\left[g_\mathrm{L} + \sum_{b=0}^{2}\sum_{j=1}^{N_b} g_{ij}^{ab}(t)\right] u_i^a(t) + \sum_{b=0}^{2}\sum_{j=1}^{N_b} g_{ij}^{ab}(t) V_b \tag{4.22}$$

$$= -[g_i^{a,\mathrm{tot}}(t)] u_i^a(t) + \sum_{b=0}^{2}\sum_{j=1}^{N_b} g_{ij}^{ab}(t) V_b, \tag{4.23}$$

with a time- and activity-dependent total leakage conductance $g_i^{a,\mathrm{tot}}(t)$. In the mean field limit $N_b \to \infty$ with $K_b/N_b = \text{const.}$, the sum $\sum_{j=1}^{N_b} g_{ij}^{ab}(t)$ diverges and the *effective membrane time constant* $\tau_i^{a,\mathrm{eff}}(t) := g_i^{a,\mathrm{tot}}(t)^{-1}$ goes to zero. In our model, we consider finite but large network sizes for which the mean of the total conductance is of order $\sqrt{K_b} \gg 1$ (see Equation (4.17)) resulting in total conductances that are much higher than the passive leakage conductance. Such *high-conductance states* accompanied by fast neuronal dynamics are regularly observed in cortical neurons *in vivo* (for review see Destexhe et al., 2003). In fact, the effective membrane time constant in both real cortex and our model can become shorter than fast synaptic dynamics that have time constants on the order of a few milliseconds. For this reason, it was essential that we included detailed synaptic dynamics in our conductance-based description.

High-conductance states not only introduce fast neuronal dynamics that enable membrane potential changes on a much faster time scale than for the model with current-based synapses. There is also another, less obvious qualitative difference regarding the postsynaptic potentials (PSPs) in response to presynaptic spikes: in states of high conductance, the PSPs reduce in amplitude – the slower the synaptic dynamics, the more pronounced the effect, as we will see shortly. Reduced amplitudes of the PSPs in response to stochastically arriving spikes translate to reduced step sizes in the random walk that the membrane potential performs. This causes a decrease in the spread of the membrane potential distribution (which is approximately Gaussian at low rates). Thus, high-conductance states lead to smaller fluctuations in the membrane potential.

What causes the reduction of PSP amplitude in states of high conductance compared to a resting neuron? It is immediately obvious from Equation (4.13) that the postsynaptic current amplitude is proportional to the

driving force $(u_i^a(t) - V_b)$. However, when the membrane potential dynamics are fast compared to the dynamics of the synapse, then there can be a substantial *leakage* of current during the time course of a single spike response. The increase in leakage conductance becomes obvious in our reformulated versions (4.22) and (4.23) where we have already alluded to this effect by calling $g_i^{a,\mathrm{tot}}(t)$ an activity dependent *leakage* conductance. From this perspective it becomes also clear why this effect is stronger for slower synaptic dynamics: the longer the time course of a single postsynaptic event, the more of the postsynaptic current will already leak out during its time course, leading to smaller PSPs. Clearly, this effect can only play a substantial role when the membrane dynamics are fast compared to the synaptic dynamics, i.e., when the neuron is in a high-conductance state. Our analysis explains simulation results reported by Kuhn et al. (2004), in which the effect of synaptic bombardment on PSPs was investigated.

Numerical procedure

The algorithm for finding self-consistent firing statistics for the conductance-based model is essentially the same as the one for the current-based model described in the previous section. However, there are two important differences: First, we cannot compute the input currents directly from the firing statistics. Rather, we have to compute the "input conductances" and let the input currents evolve simultaneously with the membrane potentials in the explicit simulations of the single neuron problems. Second, we are now considering synaptic dynamics, i.e., the spike statistics of presynaptic neurons affect the postsynaptic neuron only after a filtering stage at the synapse (see Equation (4.19)).

We have schematized the numerical procedure for the conductance-based model in the lower part of Figure 4.5. The cartoon images of autocorrelation functions at three stages ('presynaptic', 'input conductance', and 'output') are meant to stress the importance of doing the mean-field calculations self-consistently not only for the rates, but also for the correlations: synaptic filtering introduces correlations that influence the firing statistics (neuronal output) – provided that the neuronal dynamics are fast enough, as in high-conductance states. These correlations in the output feed then back to the neuronal input and get further enhanced by synaptic filtering. Thus, coloring the noise correctly is important to capture the full amount of temporal correlations in conductance-based models.

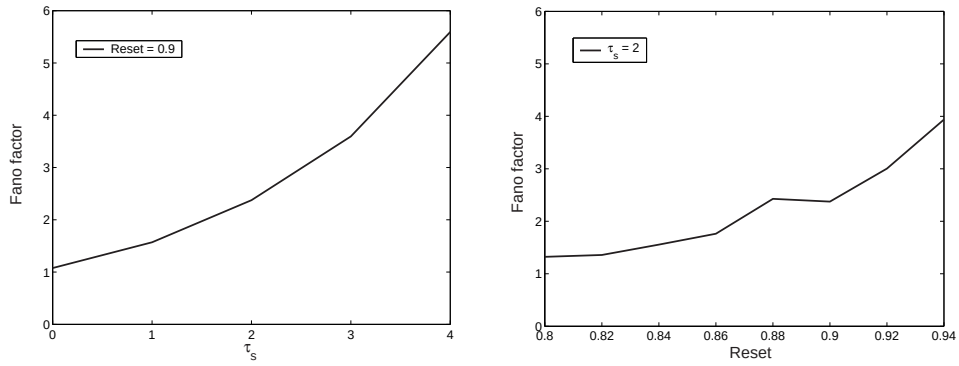

Figure 4.7: Dependence of the Fano factor on the synaptic time constant and the difference between reset and threshold. *Left panel:* Slower synaptic dynamics lead to higher Fano factors. In the case of instantaneous synapses ($\tau_s = 0$), temporal correlations in the spike trains disappear and Poisson-like firing statistics ($F \approx 1$) are observed. *Right panel:* Fano factor as a function of the reset level for a fixed firing threshold equal to 1, and $\tau_s = 2$ ms. Decreasing distances between reset and threshold lead to increasing Fano factors.

4.2.2 Results

As for the model with current-based synapses, we are interested in pinning down which model parameters control the irregularity of firing, as characterized by, e.g., the Fano factor. In the theoretical treatment above, we have already stressed the importance of even fast synaptic dynamics in states of high conductance. As expected, we observed such high-conductance states in all our simulations. In the left panel of Figure 4.7 we show the Fano factor as a function of the synaptic time constant $\tau_s = 0, 1, \cdots, 4$ ms (denoted τ_2 in (4.15)). Firing irregularity depends sensitively on these synaptic dynamics, with increasing Fano factors for increasing τ_s. As in all our simulations, the passive membrane time constant was fixed at $\tau = g_{\mathrm{L}}^{-1} = 10$ ms. For $\tau_s = 0$, we observed consistently Fano factors close to 1. In the right panel of Figure 4.7, Fano factors are plotted as a function of spike-reset level, where we kept the firing threshold fixed at 1. (See also Lerchner et al., 2004a).

Figure 4.8 shows how the width of the membrane potential distribution changes with the synaptic time constant. The results confirm predictions based on our theoretical analysis of high-conductance states above: longer synaptic time constants lead to narrower membrane potential distributions.

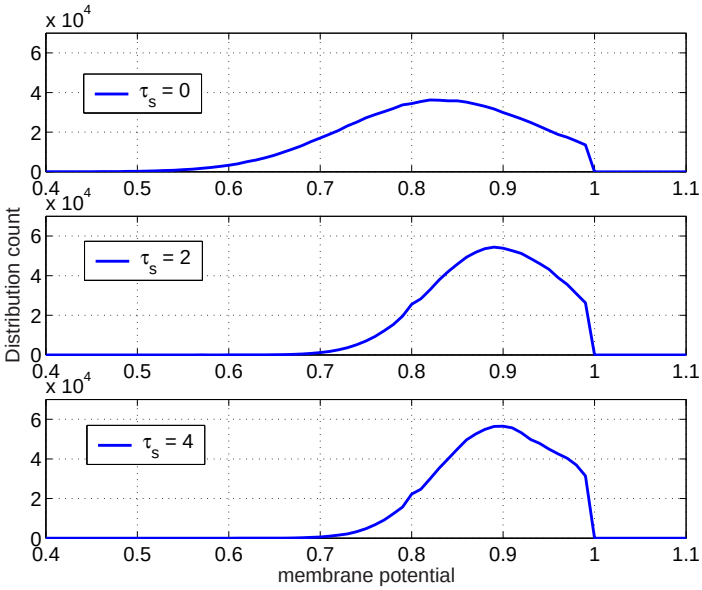

Figure 4.8: The width of the membrane potential distribution depends on the synaptic time constant τ_s. Longer synaptic time constants result in narrower distributions.

Note that longer synaptic time constants also lead to higher Fano factors, as seen in the left panel of Figure 4.7. This is at odds with what we learned from the current-based model, where wider membrane potential distributions were always accompanied by higher Fano factors. The reason lies in different neuronal dynamics for the two models, which lead to different mechanisms controlling the amount of irregularity in the neuronal firing. We will discuss these mechanisms in more detail below, where we will see that the degree of firing irregularity in the conductance-based model is mainly determined by the amount of temporal correlations in the input (which are controlled by the synaptic time constant).

In Figure 4.9 we show how the width of the membrane potential distribution depends on the density of the connectivity. To control the expected overall number of synaptic inputs from both populations, we introduced a general connectivity factor K_{fac} and took $K_1 = K_{\text{fac}} \cdot 100$, $K_2 = K_0 = K_{\text{fac}} \cdot 15$ with $K_a/N_a = 0.1$ fixed. The three membrane potential distributions shown in Figure 4.9 were obtained from calculations with $K_{\text{fac}} = 15, 30$, and 60, translating to on average 1500, 3000, and 6000 excitatory inputs, respectively.

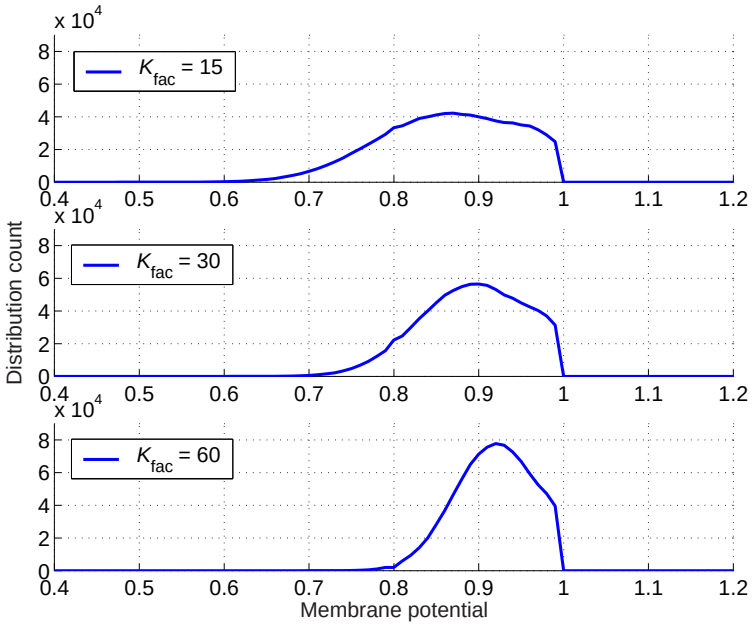

Figure 4.9: The width of the membrane potential distribution depends on the size of the connectivity. Membrane potential distributions are shown for three different values of the overall connectivity factor K_{fac} that controls the average number of synaptic inputs. Higher connectivity results in narrower distributions.

It can be seen that higher number of inputs result in narrower distributions. We can understand this result by using Equation (4.22), which tells us that the total leakage conductance $g_i^{a,\text{tot}}(t)$ increases with the number of inputs: the summation runs over more terms that are proportional to the (approximately same) firing rate. As explained in the analysis of high-conductance states above, a higher total conductance results in smaller postsynaptic potentials and thus in narrower membrane potential distributions.

4.2.3 Discussion

The basic mechanism for *producing* irregular firing in the conductance-based model is the same as for the current-based version discussed in Section 4.1: random but high connectivity and a negative feedback between excitatory and inhibitory populations lead to a random walk of the membrane potential

with a mean value below threshold and occasional excursions above threshold.

As for the current-based model, the degree of irregularity is controlled by the step size of the random walk relative to the distance between reset and threshold. For this reason, we observed higher Fano factors for reset levels that were closer to threshold, as shown in the right panel of Figure 4.7.

We also observed some important differences regarding the mechanisms that control the *degree* of irregularity in the two models. Probably the most straightforward difference is that the short effective membrane time constant observed in the conductance-based version makes synaptic filtering, even on a fast time scale, an important concept. The left panel in figure 4.7 shows the sensitive dependence of the Fano factor on the synaptic time constant. However, we saw in Figure 4.8 that there must be an additional, different mechanism at work. The increasing spike-count irregularity for longer synaptic time constants cannot be explained by random-walk dynamics of the membrane potential: instead of wider membrane potential distributions at higher Fano factors, we even observed narrower distributions. What then, makes the firing more irregular for longer synaptic time constants? The answer is that longer synaptic time constants lead to longer temporal correlations in the input currents which in turn introduce a higher tendency towards bursty firing. There is an elegant way to explain this phenomenon by rearranging the equation of motion for the membrane potential (4.13) via (4.23) yet one step further:

$$\begin{aligned} \frac{\mathrm{d}u_i^a(t)}{\mathrm{d}t} &= -g_i^{a,\mathrm{tot}}(t)\left[u_i^a(t) - \frac{\sum_{b=0}^{2}\sum_{j=1}^{N_b} g_{ij}^{ab}(t)V_b}{g_i^{a,\mathrm{tot}}(t)}\right] \\ &= -g_i^{a,\mathrm{tot}}(t)\left[u_i^a(t) - V_{\mathrm{S}}(t)\right], \end{aligned} \tag{4.24}$$

where we follow Shelley et al. (2002) in calling $V_{\mathrm{S}}(t)$ the *effective reversal potential*. We can see directly from (4.24) that the membrane potential tries to follow $V_{\mathrm{S}}(t)$ with a time constant of $[g_i^{a,\mathrm{tot}}(t)]^{-1}$. In states of high conductance, that time constant can be very short, while the effective reversal potential changes on the same time scale as the synaptic dynamics. Whenever $V_{\mathrm{S}}(t)$ makes an excursion above threshold, the membrane potential tries to follow it until it hits threshold and gets reset. For longer synaptic time constants, $V_{\mathrm{S}}(t)$ can stay longer above threshold, making the neuron fire repeatedly (emitting a burst of spikes) due to the fast neuronal dynamics in the high-conductance state. As mentioned previously, more bursty firing means higher Fano factors. Figure 4.10 shows that the membrane potential fol-

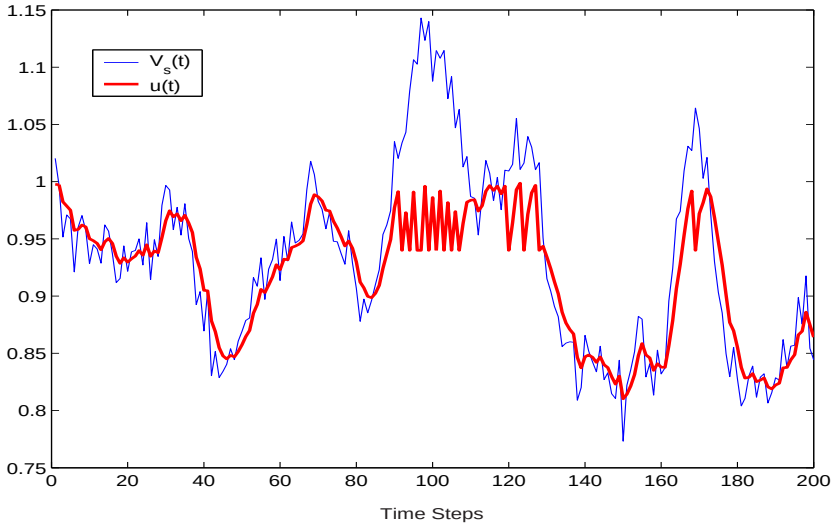

Figure 4.10: Emergence of bursty behavior in states of high conductance: The membrane potential $u(t)$ follows effective reversal potential $V_s(t)$ closely, except when $V_s(t)$ is above threshold. As long as $V_s(t)$ is super-threshold, the neuron fires repeatedly.

lows the effective reversal potential $V_{\mathrm{S}}(t)$ closely except when $V_{\mathrm{S}}(t)$ is above threshold. Synaptic filtering causes temporal correlations in $V_{\mathrm{S}}(t)$. Without synaptic dynamics, these correlations are removed and thus there is no tendency to bursty behavior, reflected by Fano factors close to 1 (as in the left panel of Figure 4.7 for $\tau_s = 0$).

We saw in Figure 4.9 that the membrane potential distribution becomes narrower as the connectivity increases. There is no limit to this effect, which leads to a pathological state for the mean-field limit of $K_a \to \infty$ with K_b/N_b fixed: in this case the total leakage conductance diverges and postsynaptic potentials vanish in amplitude. This is another qualitative difference between this and the current-based model.

Relevance of this model

Explaining the mechanisms underlying irregular firing of neocortical neurons *in vivo* has been a long-standing challenge. This work offers a biologically plausible solution to this problem, covering and explaining a host of independent experimental observations.

One may wonder how biologically plausible a model with integrate-and-fire (IAF) neurons can possibly be. This question was recently addressed in a work by Meffin et al. (2004), where current-based IAF neurons, conductance-based IAF neurons, and Hodgkin-Huxley model neurons were compared in the light of neocortical dynamics. It was found that the conductance-based synapse is the essential mechanism required to explain a range of data on cortical neurons *in vivo* and that the particular model of spike initiation is not crucial. While the current-based IAF neuron and the conductance-based one showed several qualitative differences (in agreement with our study), the conductance-based IAF and the Hodgkin-Huxley neurons showed almost identical dynamics regarding firing statistics and membrane potential statistics. We are thus optimistic that our theory can provide physiologically meaningful explanations.

Experimental evidence for the ubiquity of high-conductance states in neocortical neurons *in vivo* (see, e.g., Borg-Graham et al., 1998) has drawn much attention recently. Rudolph and Destexhe (2003) even propose high-conductance states to be essential for maintaining irregular firing activity, on the grounds that in such states there is no need for exact balance between excitation and inhibition. The authors are especially skeptical towards the consensus that has emerged for explaining irregular firing states: that the membrane potential stays close to firing threshold and that action potentials are triggered by fluctuations in the membrane potential. They argue that such a state would, e.g., require a fine-tuned balance between excitation and inhibition and that it were presently unclear how such a subthreshold regime might appear physiologically. In contrast to that, our model together with our theory provides a natural explanation for the emergence of such states, without requiring any fine-tuning and without imposing exact balance between excitation and inhibition. In fact, we do not assume balance at all; we just look for self-consistent asynchronous solutions. In our model, the dynamical balance emerges as a consequence of the network dynamics. It is true that we need to assume exact balance to derive analytically rate *estimates*, but the theory itself and the solutions found by the numerical procedure do not need such a stringent condition. In both real cortex and our model, the negative feedback between inhibitory and excitatory populations always produces sufficiently balanced states (notwithstanding pathological exceptions). Thus, we do not argue that balance plays a causal role in producing irregular firing – rather, we point out that it emerges naturally from the network dynamics. The key to solving the problem is to use a truly self-

consistent description. This brings us back to a fundamental problem in the proposition of Rudolph and Destexhe (2003) that high-conductance states are essential for maintaining an irregular firing state. In their investigations, they *assumed* irregular firing of presynaptic neurons to begin with and concluded that high-conductance states make it easier for the target neuron to fire irregularly in response as well. That approach, however, addresses only a part of the problem and circumvents the real question about the emergence of irregular firing states by simply assuming them in the input. With our theory it is for the first time possible not only to explain the mechanisms for irregular firing (the combination of quenched disorder in the connectivity, relatively strong synapses, and negative feedback between excitation and inhibition), but also to quantify the degree of irregularity – and more generally the firing correlations – as functions of intrinsic network parameters.

We have already mentioned the study by Meffin et al. (2004), in which the differences between a current-based IAF model, a conductance-based one, and a Hodgkin-Huxley model were carefully investigated. One of their conclusions was that there are different mechanisms underlying the irregular firing in the current-based and the conductance-based model. We do agree that there are different mechanisms controlling the *amount* of irregularity and the nature of firing correlations in the two models, as analyzed above. However, the mechanisms underlying the emergence of irregularity itself can be the very same, as we have demonstrated in this work.

In their review on high-conductance states of neocortical neurons *in vivo*, Destexhe et al. (2003) point out that most of the network-level modeling studies with IAF neurons did not use conductance-based synapses ("only one study has reported the genesis of self-sustained stochastic states with conductance-based synapses") and that it is consequently impossible to compare the states in these models with conductance measurements *in vivo*. Here, we provide a network-level study that elucidates how high-conductance states and firing correlations are intimately related in highly connected, random "balanced" networks – without a causal role for high-conductance states in the emergence of irregularity per se.

4.3 Orientation hypercolumn in V1

The primary visual cortex (V1) contains neurons that respond strongest to elongated visual stimuli of a specific orientation, the neuron's *preferred orientation* (PO). Such orientation selective neurons exhibit a tuned response to other orientations, with decreasing firing rates as the orientation angle becomes more dissimilar, until the firing vanishes for orientations outside the *tuning width* of the neuron. Neurons with the same PO are grouped together, forming cortical columns. A collection of columns that span the entire range of orientation angles (and that respond to the same area in the visual field) is called an *orientation hypercolumn*.

We can extend the mean field theory developed for a single cortical column to an orientation hypercolumn by introducing structural inhomogeneity into the network and solving the mean field equations simultaneously for multiple cortical columns.

We will see that several key features of orientation selective neurons in V1 can be reproduced and explained within the model. Among these are a tuning width that is invariant with respect to the stimulus contrast, irregular firing with statistics including Fano factors greater than 1, and an almost linear relationship between stimulus contrast and firing rate. Beyond that, we can make quantitative predictions about the tuning of the firing irregularity and that of the input currents.

4.3.1 Model and methods

We model the orientation hypercolumn as an ensemble of n orientation columns with preferred orientations θ equally spaced between $-\pi/2$ and $\pi/2$. Each column consists of an inhibitory and an excitatory subpopulation of sizes N_1/n and N_2/n, respectively. We introduce structure to the random connectivity of the network by grading the density of connections between orientation populations by the similarity of their respective angles. Specifically, any given neuron in sub-population a of column θ receives synaptic input from any other neuron of sub-population b of column θ' with probability

$$P_{ab}(\theta - \theta') = \frac{K_b}{N_b} \left(1 + \gamma \cos 2(\theta - \theta')\right), \tag{4.25}$$

where K_b is the expected overall number of inputs from neurons in population b. As for the single-column model, we take the ratio K_b/N_b to be independent

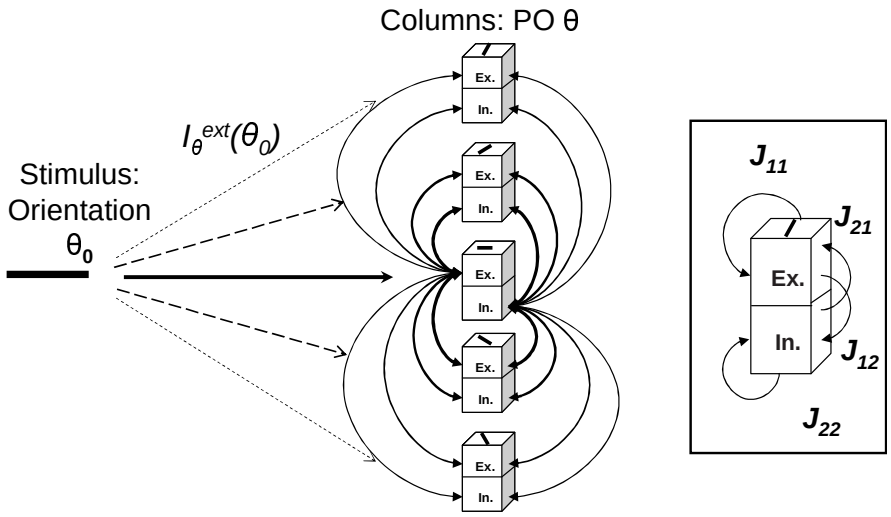

Figure 4.11: Structure of the orientation hypercolumn model. The network consists of multiple orientation columns, interconnected with densities that depend on the similarity of the preferred orientations (PO). Connections are only indicated between one column and the rest of the network. *Inset:* the local structure of each orientation column corresponds to that of the single-column model described in Section 4.1.

of b. The parameter $\gamma \in (0, 1)$ defines the degree of tuning, which we assume to be the same for the inhibitory and the excitatory population. The functional form of (4.25) implies that the connection probability between neurons decreases as their difference in PO increases. This is motivated by the fact that orientation columns with similar PO tend to lie closer together on the cortical surface than ones with dissimilar PO, and anatomical evidence that the connection probability between cortical neurons decreases with distance.

The following analysis is performed for current-based synapses and we use the same scaling as described in the single-column model in Section **??**: each nonzero synapse between a neuron in population b and a neuron in population a has strength $J_{ab}/\sqrt{K_b}$, independent of their preferred orientations.

The orientation hypercolumn receives excitatory input from the *lateral geniculate nucleus* (LGN) with a strength proportional to the stimulus contrast I_a^{ext}, and a weak tuning to the stimulus orientation θ_0. We model this external input and its tuning as originating from a feed-forward connectivity from the LGN, as for example in the classical model of Hubel and Wiesel

(1962). For simplicity, we take it to be constant in time and the same for all neurons within population column $a\theta$:

$$I_{a\theta}^{\text{ext}}(\theta_0) = I_a^{\text{ext}}(1 + \epsilon \cos 2(\theta - \theta_0)), \tag{4.26}$$

where the degree of tuning is determined by the parameter ϵ, which is positive and smaller than 1 to ensure that the the external input is always excitatory (as the input from LGN in real cortex). We followed Ben-Yishai et al. (1995) in choosing the simplest possible general form that is periodic with period π for both the intracortical connectivity (4.25) and the external input current (4.26).

We can now write the equation of motion for the subthreshold membrane potential of neuron i in orientation population $a\theta$ as

$$\frac{\mathrm{d}u_i^{a\theta}(t)}{\mathrm{d}t} = -\frac{u_i^{a\theta}(t)}{\tau} + I_{a\theta}^{\text{ext}}(\theta_0) + I_i^{a\theta,\text{rec}}(t), \tag{4.27}$$

where the recurrent input current $I_i^{a\theta,\text{rec}}(t)$ is given by

$$I_i^{a\theta,\text{rec}}(t) = \sum_{b=1}^{2} \sum_{\theta'=\theta_1}^{\theta_n} \sum_{j=1}^{N_b/n} J_{ij}^{a\theta,b\theta'} S_j^{b\theta'}(t). \tag{4.28}$$

To derive a mean field description of the recurrent current, we start by separating the synaptic strengths $J_{ij}^{a\theta,b\theta'}$ and the spike trains $S_j^{b\theta'}(t)$ into mean- and fluctuating parts, analogously to (4.2) and (4.3). We then obtain a set of equations, defining the Gaussian random processes describing the recurrent inputs to the $2n$ column populations $a\theta$:

$$I_{a\theta}^{\text{rec}}(t) = \sum_{b=1}^{2} J_{ab} \left(\sqrt{K_b} A_b + \sqrt{1 - K_b/N_b} B_b(t) \right), \tag{4.29}$$

with

$$A_b = \frac{1}{n} \sum_{\theta'=\theta_1}^{\theta_n} (1 + \gamma \cos 2(\theta - \theta')) r_b(\theta') \tag{4.30}$$

$$B_b(t) = \frac{1}{\sqrt{n}} \sum_{\theta'=\theta_1}^{\theta_n} \sqrt{1 + \gamma \cos 2(\theta - \theta')} \left(\left(\overline{(r_j^{b\theta'})^2} \right)^{\frac{1}{2}} x_{b\theta'} + \xi_{b\theta'}(t) \right) \tag{4.31}$$

The part $J_{ab}\sqrt{K_b} A_b$ of $I_{a\theta}^{\text{rec}}(t)$ is time independent and describes the population mean common for each neuron in column population $a\theta$ (here we

consider stationary rates only to highlight the main principles). It is proportional to A_b, which is an average over the column population rates $r_b(\theta')$ weighted by the density of connections between the source columns θ' and the target column θ.

The part $\sqrt{1 - K_b/N_b}\, B_b(t)$ in the statistical description of the recurrent inputs (4.29) contains the terms that are due to the randomness in the connectivity – thus it vanishes in case of full connectivity ($K_b = N_b$). The term $B_b(t)$ contains averages of the rate fluctuations $\overline{(r_j^{b\theta'})^2}$ and the correctly colored temporal fluctuations $\xi_{b\theta'}(t)$, weighted by the square root of the connection densities. The unit variance random number $x_{b\theta'}$ in (4.31) accounts for the fact that the specific number of connections from the source column population $b\theta'$ to the target neuron is random, while the rate fluctuations accounts for the distribution of firing rates within $b\theta'$. Analogous to the previous mean field models, picking a specific set $\{x_{b\theta'} : b = 1, 2;\, \theta' = \theta_1, \ldots, \theta_n\}$ amounts to sampling a specific neuron in the target population $a\theta$. Finally, the temporal fluctuations in the recurrent input current (4.29) are captured by the Gaussian random processes $\xi_{b\theta'}(t)$, which obey

$$\langle \xi_{b\theta'}(t)\xi_{b\theta'}(t')\rangle = C_{b\theta'}(t - t'). \tag{4.32}$$

Here, $C_{b\theta'}(t - t')$ denotes the average autocorrelation function of the fluctuations in the spike trains of neurons with PO θ' in population b, given by

$$C_{b\theta'}(t - t') = \frac{1}{N_b/n} \sum_{j=1}^{N_b/n} \langle \delta S_j^{b\theta'}(t)\delta S_j^{b\theta'}(t')\rangle. \tag{4.33}$$

The balance condition

We model the orientation hypercolumn as multiple interconnected (generic) columns. Thus, the reasoning that led us to the formulation of the balance condition for a single column in Section 4.1 can be directly applied to the hypercolumn, provided that we require it to hold for all of the orientation columns simultaneously: the input currents in Equation (4.27) have to nearly cancel for each column population $a\theta$, leading to the following set of $2n$ equations for $a = 1, 2$ and $\theta = \theta_1, \ldots, \theta_n$:

$$\sqrt{K_0}\hat{I}_a^{\text{ext}}(1 + \epsilon \cos 2(\theta - \theta_0)) + \sum_{b=1}^{2} J_{ab}\sqrt{K_b}A_b = \mathcal{O}(1), \tag{4.34}$$

where we have used the mean-field description (4.29) of the mean recurrent currents.

We will now outline how one can find analytical estimates of the orientation population rates $r_a(\theta)$ from (4.34). As mentioned previously, orientation selective neurons in V1 tend to have suppressed firing for stimulus orientations θ_0 that differ more than the tuning width θ_c ("critical angle") from the neuron's preferred orientation θ: $r_a(\theta) = 0$ for $|\theta - \theta_0| > \theta_c$. We will call this property *narrow tuning* in contrast to *broad tuning* with non-vanishing firing rates for all stimulus orientations. The solutions for the broadly tuned case are easier to find and they provide insight into the expected form of the analogous solutions to the narrowly tuned case; i.e, they provide us with an *ansatz* for solving the biologically more realistic narrow case.

We start by rewriting A_b, the weighted average over the column population rates $r_a(\theta)$ defined in (4.30), in a continuum notation

$$A_b = \int_{-\pi/2}^{\pi/2} \frac{\mathrm{d}\theta'}{\pi} (1 + \gamma \cos 2(\theta - \theta')) r_b(\theta'), \tag{4.35}$$

so that (4.34) becomes a set of integral equations for $r_a(\theta)$. In the broadly tuned case, which we consider first for the reasons outlined above, these can be solved by using a Fourier expansion of the rates around the stimulus orientation, $r_a(\theta) = r_{a,0} + r_{a,2} \cos 2(\theta - \theta_0) + \cdots$, and with help of the identity $\cos(\alpha - \beta) = \cos\alpha \cos\beta + \sin\alpha \sin\beta$. For finding approximate solutions to the mean rates, we set the small term of order 1 on the right-hand side of (4.34) to zero. Due to our choice of the functional form for the connection probabilities (4.25) and for the external input current (4.26) in which we only retained the two first Fourier components of a more general description with period π, the solutions turn out to be particularly simple with the same low number of Fourier components:

$$\sqrt{K_0}\hat{I}_a^{\mathrm{ext}}(1 + \epsilon \cos 2(\theta - \theta_0)) + \sum_{b=1}^{2} \sqrt{K_b} J_{ab} \left[r_{b,0} + \frac{1}{2}\gamma r_{b,2} \cos 2(\theta - \theta_0) \right] = 0. \tag{4.36}$$

The resulting equations for each of the two Fourier components are directly analogous to the balance conditions (4.8) and (4.20) obtained for the previous models. We arrive thus at solutions for each of the Fourier components of the rates that look similar to (4.9) and (4.21):

$$r_{a,0} \;=\; -\sum_{b=1}^{2} (\hat{\mathsf{J}}^{-1})_{ab} \hat{I}_b^{\mathrm{ext}} \tag{4.37}$$

$$r_{a,2} = -\frac{2\epsilon}{\gamma}\sum_{b=1}^{2}(\hat{\mathsf{J}}^{-1})_{ab}\hat{I}_b^{\text{ext}} = \frac{2\epsilon}{\gamma}r_{a,0}, \tag{4.38}$$

where the matrix $\hat{\mathsf{J}}$ is composed of the elements $\hat{J}_{ab} = J_{ab}\sqrt{K_b/K_0}$. Equation (4.38) tells us that these solutions for the broadly tuned case correspond to parameter values of $\epsilon \in (0, \gamma/2]$, since the firing rates $r_a(\theta)$ must be non-negative for all θ and θ_0.

We can now turn to the narrowly tuned case, for which we try the *ansatz* that the solution for non-vanishing rates has the same form as for the broadly tuned case, and is equal to zero otherwise. This can be formally written as

$$r_a(\theta) = \begin{cases} r_{a,0} + r_{a,2}\cos 2(\theta - \theta_0) & \text{for } |\theta - \theta_0| < \theta_c \\ 0 & \text{for } |\theta - \theta_0| \geq \theta_c, \end{cases} \tag{4.39}$$

where $\theta_c = -1/2\cos^{-1}(r_{a,0}/r_{a,2})$. We do not need a population index a for the tuning width θ_c because it turns out to be equal for excitatory and inhibitory neurons due to our population-independent choices of ϵ and γ. It is therefore convenient to restrict the integration in (4.35) to the range for non-zero rates, defined by the (yet unknown) tuning width θ_c. To make the dependence on θ_c explicit, we rewrite the part of the *ansatz* for $|\theta' - \theta_0| < \theta_c$ in the form

$$r_a(\theta) = r_{a,2}(\cos 2(\theta - \theta_0) - \cos 2\theta_c) \tag{4.40}$$

and solve for the two unknowns θ_c and $r_{a,2}$ instead of $r_{a,0}$ and $r_{a,2}$. It turns out that this *ansatz* leads to valid solutions within the parameter regime $\epsilon \in (\gamma/2, \gamma]$, which obey, analogously to (4.36),

$$\begin{aligned} 0 &= \sqrt{K_0}\hat{I}_a^{\text{ext}}(1 + \epsilon\cos 2(\theta - \theta_0)) \\ &+ \sum_{b=1}^{2}\sqrt{K_b}J_{ab}r_{b,2}\left[f_0(\theta_c) + \gamma f_2(\theta_c)\cos 2(\theta - \theta_0)\right], \end{aligned} \tag{4.41}$$

where

$$f_0(\theta_c) = \int_{-\theta_c}^{\theta_c}\frac{\mathrm{d}\theta'}{\pi}(\cos 2\theta' - \cos 2\theta_c) \tag{4.42}$$

$$= \frac{1}{\pi}(\sin 2\theta_c - 2\theta_c\cos 2\theta_c) \tag{4.43}$$

and

$$f_2(\theta_c) = \int_{-\theta_c}^{\theta_c}\frac{\mathrm{d}\theta'}{\pi}\cos 2\theta'(\cos 2\theta' - \cos 2\theta_c) \tag{4.44}$$

$$= \frac{1}{\pi}(\theta_c - \frac{1}{4}\sin 4\theta_c). \tag{4.45}$$

Similar expressions appeared in a different model studied by Ben-Yishai et al. (1995) and Hansel and Sompolinsky (1998), from whom we have borrowed the notation.

As for the broadly tuned case, we can solve these equations for each of the two Fourier components separately, yielding a pair of equations for $r_{a,2}$ and θ_c for each angle θ:

$$\hat{I}_a^{\text{ext}} + \sum_{b=1}^{2} \hat{J}_{ab} r_{b,2} f_0(\theta_c) = 0 \tag{4.46}$$

$$\epsilon \hat{I}_a^{\text{ext}} + \gamma \sum_{b=1}^{2} \hat{J}_{ab} r_{b,2} f_2(\theta_c) = 0 \tag{4.47}$$

We can solve for θ_c via dividing (4.47) by (4.46), which eliminates the dependence on $r_{a,2}$, yielding a single equation in the unknown θ_c:

$$\frac{f_2(\theta_c)}{f_0(\theta_c)} = \frac{\epsilon}{\gamma} \tag{4.48}$$

Figure 4.12 shows the dependence of θ_c on the ratio ϵ/γ according to (4.48).

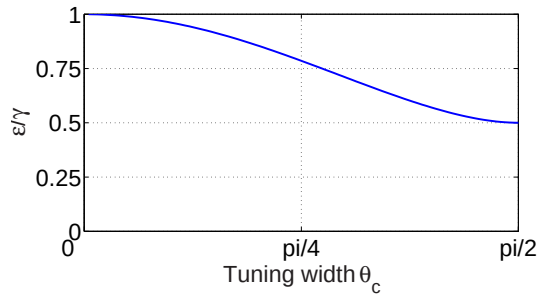

Figure 4.12: Shape of the function $f_2(\theta_c)/f_0(\theta_c) = \epsilon/\gamma$.

We can see directly from this figure that narrowly tuned solutions can only be found for $\epsilon \in (\gamma/2, \gamma)$. Positive ϵ below that range lead to broad tuning, as shown above. For $\epsilon \to \gamma$, the tuning width θ_c becomes vanishingly small, and for $\epsilon > \gamma$ there are no solutions to the balance condition (4.41) for θ_c in the meaningful range $(0, \pi/2]$.

Finally, with the knowledge of θ_c, we arrive at a familiar form for the rate solutions (cf. (4.9), (4.21), and (4.37)):

$$r_{a,2} = -\frac{1}{f_0(\theta_c)} \sum_{b=1}^{2} (\hat{\mathsf{J}}^{-1})_{ab} \hat{I}_b^{\text{ext}}. \tag{4.49}$$

The first Fourier component of the rate is then obtained via $r_{a,2}$ by using the equality $r_{a,0} = -r_{a,2} \cos 2\theta_c$.

The tuning width θ_c, which can be determined from (4.48), depends only on the two parameters ϵ and γ that specify the degree of tuning of the external input from the LGN and of the intracortical connectivity, respectively. In particular, we find that the tuning width is independent of the stimulus contrast in our model. As noted earlier, a tuning width that is invariant with regard to the stimulus contrast is one of the main properties observed in orientation selective cortical neurons that cannot be accounted for in pure feed-forward models. In our model, it is a consequence of the dynamical balance that is continually maintained by the negative feedback between the excitatory and the inhibitory neuron populations within the cortex.

Further inspection of (4.48) tells us that stronger modulation of the external input tuning (larger ϵ) results in narrower tuning of the neuronal firing. However, stronger modulation of the intracortical connectivity (larger γ for ϵ fixed) results in *less sharp tuning* of the neuronal firing. This may seem counterintuitive at first sight, but it is a natural consequence of the fact that the intracortical connectivity is predominantly *inhibitory*, which leads to competitive behavior between individual orientation columns. This competition culminates in a winner-takes-all behavior for "flat" cortical connectivity with $\gamma = 0$, which leads theoretically to a δ-function-like tuning centered at the stimulus orientation.

Tuning of the neuronal input noise

It is in general not possible to provide an analytical expression of the self-consistent input noise. As for the single-column models analyzed in Section 4.1, we have to resort to numerical evaluations on a case-by-case basis. However, we can obtain some insight into properties of the tuning of the input noise when we consider a general feature of the firing correlation functions $C_{a\theta}(t-t')$ defined in (4.33): for irregular firing states, they always have a piece proportional to $r_a(\theta)\delta(t-t')$. This delta function at $t = t'$ provides

a flat contribution to the noise spectrum that can be calculated analytically. Specifically, we find that the high-frequency limit of the neuronal input noise has the same orientation tuning as the external input to the neuron. To obtain this result, we use the mean field description of the recurrent current (4.29) together with a continuum formulation for the weighted average $B_b(t)$ defined in (4.31) to derive the following expression of the dynamic noise in the input current:

$$\langle \delta I_{a\theta}^{\mathrm{rec}}(t) \delta I_{a\theta}^{\mathrm{rec}}(t') \rangle = \sum_{b=1}^{2} J_{ab}^2 (1 - K_b/N_b) \int_{-\pi/2}^{\pi/2} \frac{\mathrm{d}\theta'}{\pi} 1 + \gamma \cos 2(\theta - \theta')) C_{b\theta'}(t - t'). \quad (4.50)$$

For $t \to t'$, or, equivalently for $\omega \to \infty$ in the frequency domain, we get then

$$\lim_{\omega \to \infty} \langle |\delta I_{a\theta}^{\mathrm{rec}}(\omega)|^2 \rangle = -\hat{I}_a^{\mathrm{ext}} [1 + \epsilon \cos 2(\theta - \theta_0)] \sum_{b=1}^{2} J_{ab}^2 (1 - K_b/N_b) \sum_{c=1}^{2} (\hat{\mathbf{J}}^{-1})_{bc} \hat{I}_c^{\mathrm{ext}}. \quad (4.51)$$

Details of the derivation can be found in Lerchner et al. (2004b).

Numerical procedure

For the orientation hypercolumn model, we have to solve the mean field equations simultaneously for multiple cortical columns. The procedure to do so is directly analogous to the one described for the single column model. Now we need to determine and update the statistics for $2n$ orientation populations $a\theta$ (instead of 2 populations) at each iteration. However, much computing time can be saved by making use of the inherent symmetry of our model: to determine the self-consistent population statistics for the n orientation columns, we only need to explicitly simulate (many trials of) single neurons corresponding to half of the columns and mirror the results for the other half.

Once the algorithm has converged and we have found the population statistics for the rates, the rate fluctuations and the correlations, we can collect statistics of individual neurons. Individual neurons have specific connection patterns (randomly drawn but held fixed) from the $2n$ source orientation populations, which brakes the symmetry for the calculations. A specific connection pattern (and therefore a specific neuron) is determined by drawing a set of $2n$ random numbers $x_{b\theta'}$ (see Equation (4.31))

$$\{x_{b\theta'} : b = 1, 2;\ \theta' = \theta_1, \ldots, \theta_n\}, \quad (4.52)$$

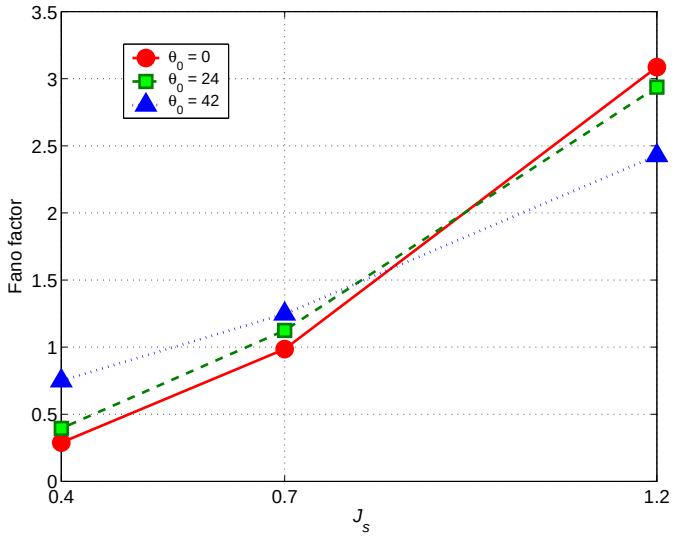

Figure 4.13: Dependence of the irregularity in firing on synaptic strengths. The Fano factor is plotted as a function of the overall synaptic strength J_s for three different stimulus orientations: $\theta_0 = 0$ (the neuron's preferred orientation), $\theta_0 = 24$, and $\theta_0 = 42 \approx \theta_c$ (the tuning width). For all stimulus orientations, stronger synapses generally lead to higher spike count variances.

which needs to be kept fixed over all trials for collecting the statistics for that neuron.

4.3.2 Results

In agreement with the results obtained from the current-based single-column model, we find that the firing irregularity of neurons within the orientation hypercolumn is mainly controlled by the strength of the synapses: Stronger synapses lead to higher Fano factors. Figure 4.13 shows the Fano factor as a function of the overall synaptic strength J_s for three different stimulus orientations. The effect is most pronounced at the neuron's preferred orientation, but it is not much diminished even close to the tuning width of the neuron (where firing rates are very low).

In Figure 4.14 we compare experimental data obtained from an orientation selective neuron in cat primary visual cortex (Sclar and Freeman, 1962) with numerical results for an individual neuron of our hypercolumn model.

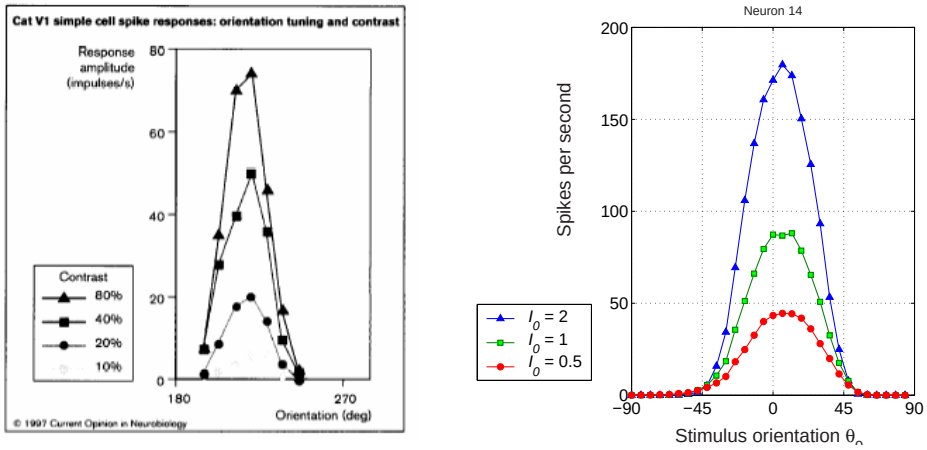

Figure 4.14: The tuning width is invariant with respect to contrast. *Left panel:* Experimental data from a simple cell in cat V1. Orientation tuning was measured at different contrasts. Data from Sclar and Freeman (1962); figure adopted from Sompolinsky and Shapley (1997) *Right panel:* Numerical results for an individual neuron within our model. The same qualitative features as in the experiment can be seen, including an almost linear input-output relationship between stimulus contrast and firing rate.

The data from the model look qualitatively like those from experiments. Most importantly, the tuning width is invariant with respect to the stimulus contrast, as predicted by the theory (see Equation (4.48)). In addition, the data from both experiment and model show an almost linear relationship between stimulus contrast and firing rate. We anticipated this result from our analytically derived rate estimate (4.49).

In our analytical treatment, we have calculated that the high-frequency neuronal input noise power has the same tuning as the external input (4.51). Figure 4.15 shows simulation results that confirm this prediction. The tuning of the noise variance in the neuronal input current (triangles) is overlaid with the tuning of the external input (solid line). We adjusted the scaling of the two y-axes in Figure 4.15 and aligned them at 0 (see the lower cut-off) so that the modulation strengths can be compared directly.

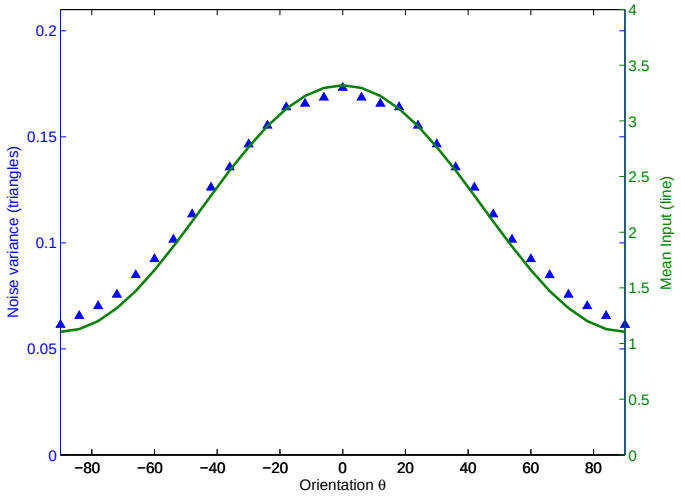

Figure 4.15: The neuronal input noise has the same orientation tuning as the external input. Data collected from a neuron with preferred orientation $\theta = 0$. The noise variance of the input current is plotted as a function of orientation (blue triangles), overlaid with the tuning of the external input (green solid line).

4.3.3 Discussion

In this section, we showed how the complete mean-field theory for randomly connected spiking neurons can be extended to networks with structural inhomogeneity, making it possible to incorporate functional architecture.

We applied the theory to a simple model of an orientation hypercolumn in primary visual cortex and obtained a range of results that are in agreement with data obtained from experiment.

With help of the numerical procedure, introduced in the previous sections and adapted here for this model, we are in a position to make quantitative statements about the orientation tuning of both firing statistics and input current statistics. Some of the tuning properties of the firing statistics are experimentally established and can be reproduced and explained by our model. Other properties, like the orientation dependence of the Fano factors (for more details see Lerchner et al., 2004b) are straightforward to obtain experimentally with methods that have been employed previously, while the predictions concerning the tuning of the input currents and their fluctuations in orientation selective neurons are more difficult to test *in vivo*.

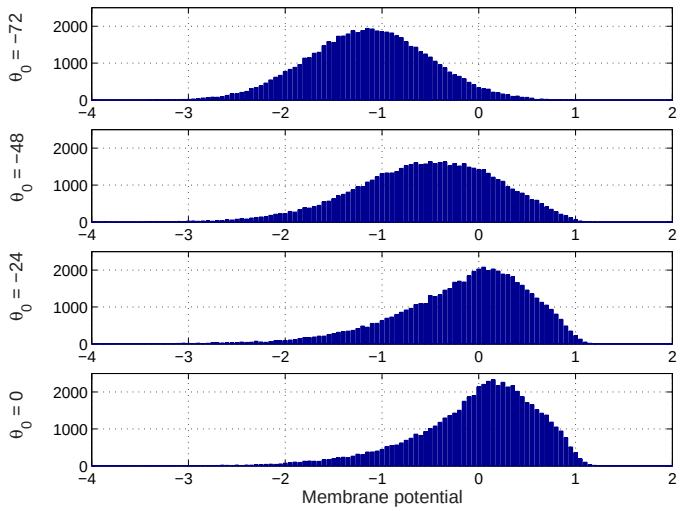

Figure 4.16: Tuning of the membrane potential distribution. Histogram counts collected from neurons with PO $\theta = 0$ and tuning width $\theta_c \approx 43$ (as in the right panel of Figure 4.14) for four different stimulus orientations θ_0. The mean of the distributions move away from threshold as the stimulus orientation becomes less optimal (panels from bottom to top).

Here, we performed the analysis and the numerical investigations with current-based integrate-and-fire neurons, which limits the realism of the model in a few respects, such as membrane potential statistics. Figure 4.16 shows how the membrane potential distribution depends on the stimulus orientation. Histogram counts were collected from neurons with PO $\theta = 0$ and tuning width $\theta_c \approx 43$ (the corresponding tuning of the firing rate is shown in the right panel of Figure 4.14). We plotted the histograms for four different stimulus orientations $\theta_0 = 0, 24, 48$ an 72. As a general trend, we see that the mean of the distributions shifts away from threshold as the stimulus becomes more dissimilar from the neurons preferred orientation. Clearly, the width of the distributions and their centering (especially for the cases outside the tuning width, shown in the upper two panels) are physiologically unrealistic: With our scaling, where the resting potential is set to 0 and threshold to 1, the reversal potential for inhibitory synapses would correspond to a value of about $-2/3$. Thus, the membrane potentials should be confined between $-2/3$ and 1. However, preliminary investigations on the hypercolumn

model with conductance-based synapses show that the general trend (the mean of the (narrower) distributions shifts away from threshold as θ and θ_0 become more dissimilar) is the same. From this, and first of all from what we have learned about the similarities and differences between current-based and conductance-based model for a single column, we have reason to expect that many underlying mechanisms can already be explained by the simpler model.

Conductance-based synapses can be incorporated into this mean-field approach by exactly the same means as shown for the single-column model in the previous section. Due to the generality of the approach and due to the explicit simulation of the synapse and the neuron in the numerical procedure, it is possible to incorporate more of the known physiology. In fact, any synapse model, possibly including short-term facilitation and depression, and even long-term potentiation and depression (i.e., learning) can be used for the numerical procedure. The same is true for the neuron model, which may be Hodgkin-Huxley-like or even a morphologically detailed multi-compartment model. While the use of such detailed models is computationally prohibitive for large-scale simulations of full networks, our mean-field approach does make it possible to study the dynamics of large spiking neural networks with computationally more expensive neurons and synapses. In our approach, the size of the network enters only as a parameter in the computations. The computationally demanding aspects of our approach are the number of trials needed to collect the firing statistics, the number of iterations needed for the algorithm to converge to the self-consistent solutions, and finally, especially important for more complex network architectures, the number of populations for which the firing statistics have to be determined.

Using the techniques presented in this section, it is conceptually straightforward to include much more of the known functional architecture, such as an entire orientation pinwheel in the primary visual cortex. In the same manner, further coding features such as spatial phase or spatial frequency can be incorporated. Seen from a computational perspective, one should be careful when extending the model in these directions, because the number of "populations" for which the firing statistics have to be calculated increases quickly with the level of anatomical detail.

Bibliography

Amit, D. and Brunel, N. (1997a). Dynamics of a recurrent network of spiking neurons before and following learning. *Network*, 8:373–404.

Amit, D. and Brunel, N. (1997b). Model of spontaneous activity and local structured activity during delay periods in the cerebral cortex. *Cereb Cortex*, 7:237–252.

Ben-Yishai, R., Bar-Or, R. L., and Sompolinsky, H. (1995). Theory of orientation tuning in visual cortex. *Proc. Natl. Acad. Sci.*, 92:3844–3848.

Borg-Graham, L. J., Monier, C., and Fregnac, Y. (1998). Visual input evokes transient and strong shunting inhibition in visual cortical neurons. *Nature*, 393:369–374.

Brunel, N. (2000). Dynamics of sparsely connected networks of excitatory and inhibitory spiking neurons. *J Comput Neurosci*, 8:183–208.

Chub, N. and O'Donovan, M. J. (1998). Blockade and recovery of spontaneous rhythmic activity after application of neurotransmitter antagonists to spinal networks of the chick embryo. *J Neurosci*, 18(1):294–306.

Chub, N. and O'Donovan, M. J. (2001). Post-episode depression of GABAergic transmission in spinal neurons of the chick embryo. *J Neurophysiol*, 85:2166–2176.

Dayan, P. and Abbott, L. F. (2001). *Theoretical neuroscience: computational and mathematical modeling of neural systems*. MIT Press.

Dean, A. F. (1981). The variability of discharge of simple cells in the cat striate cortex. *Exp Brain Res*, 44:437–440.

Destexhe, A., Rudolph, M., and Par, D. (2003). The high-conductance state of neocortical neurons *in vivo*. *Nature Reviews Neurosci.*, 4:739–751.

Eisfeller, H. and Opper, M. (1992). New method for studying the dynamics of disordered spin systems without finite-size effects. *Phys Rev Lett*, 68:2094–2097.

Fedirchuk, B., Wenner, P., Whelan, P. J., Ho, S., Tabak, J., and O'Donovan, M. J. (1999). Spontaneous network activity transiently depresses synaptic transmission in the embryonic chick spinal cord. *J Neurosci*, 19(6):2102–2112.

Fulvi Mari, C. (2000). Random networks of spiking neurons: instability in the *xenopus* tadpole moto-neuron pattern. *Phys Rev Lett*, 85:210–213.

Gershon, E., Wiener, M. C., Latham, P. E., and Richmond, B. J. (1998). Coding strategies in monkey V1 and inferior temporal cortex. *J Neurophysiol*, 79:1135–1144.

Hansel, D. and Sompolinsky, H. (1998). *Methods in Neuronal Modeling: from Synapse to Networks*, chapter Modeling feature selectivity in local cortical circuits. MIT Press.

Harata, N., Pyle, J. L., Aravanis, A. M., Mozhayeva, M., Kavalali, E. T., and Tsien, R. W. (2001). Limited numbers of recycling vesicles in small CNS nerve terminals: implications for neural signaling and vesicular cycling. *Trends Neurosci*, 24:637–643.

Harris, K. and Sultan, P. (1995). Variation in the number, location and size of synaptic vesicles provides an anatomical basis for the nonuniform probability of release at hippocampal CA1 synapses. *Neuropharmacology*, 34:1387–1395.

Heggelund, P. and Albus, K. (1978). Response variability and orientation discrimination of single cells in striate cortex of cat. *Exp Brain Res*, 32:197–211.

Hertz, J., Richmond, B., and Nilsen, K. (2003). Anomalous response variability in a balanced cortical network model. *Neurocomputing*, 52–54:787–792.

Ho, S. and O'Donovan, M. J. (1993). Regionalization and intersegmental coordination of rhythm-generating networks in the spinal cord of the chick embryo. *J Neurosci*, 13(b):1354–1371.

Hodgkin, A. L. and Huxley, A. F. (1952). A quantitative description of ion currents and its applications to conduction and excitation in nerve membranes. *J. Physiol. (Lond.)*, 117:500–544.

Hubel, D. H. and Wiesel, D. N. (1962). Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. *J. Physiol. Lond.*, 160:106–154.

Kuhn, A., Aertsen, A., and Rotter, S. (2004). Neuronal integration of synaptic input in the fluctuation-driven regime. *J Neurosci*, 24(10):2345–2356.

Lerchner, A., Ahmadi, M., and Hertz, J. (2004a). High conductance states in a mean field cortical network model. *in press: Neurocomputing*.

Lerchner, A. and Rinzel, J. (2004). Synaptic model for spontaneous activity in developing networks. *to be published in Neurocomputing*.

Lerchner, A., Sterner, G., Hertz, J., and Ahmadi, M. (2004b). Mean field theory for a balanced hypercolumn model of orientation selectivity in primary visual cortex. *submitted to Networks: Comp in Neural Syst*.

Lerchner, A., Ursta, C., Hertz, J., Ahmadi, M., and Ruffiot, P. (2004c). Response variability in balanced cortical networks. *submitted to Neural Compuation*.

Meffin, H., Burkitt, A. N., and Grayden, D. B. (2004). An analytical model for the 'large, fluctuating synaptic conductance state' of neocortical neurons *in vivo*. *J Comp Neurosci*, 16:159–175.

Mountcastle, V. B. (1998). *Perceptual Neuroscience: The Cerebral Cortex*. Harvard Univ Pr.

O'Donovan, M. J. (1999). The origin of spontaneous activity in developing networks of the vertebrate nervous system. *Curr Opin Neurobiol*, 9:94–104.

Rolls, E. and Deco, G. (2002). *Computational Neuroscience of Vision*. Oxford University Press.

Rudolph, M. and Destexhe, A. (2003). The discharche variability of neocortical neurons during high conductance states. *Neuroscience*, 119:855–873.

Schikorski, T. and Stevens, C. (1997). Quantitative ultrastructural analysis of hippocampal excitatory synapses. *J Neurosci*, 17:5858–5867.

Sclar, G. and Freeman, R. (1962). Orientation selectivity in cats striate cortex is invariant with stimulus contrast. *Exp Brain Res*, 46:457–461.

Shadlen, M. N. and Newsome, W. T. (1998). The variable discharge of cortical neurons: implications for connectivity, computation, and information coding. *J Neurosci*, 18:3870–3896.

Shelley, M., McLaughlin, D., Shapley, R., and Wielaard, J. (2002). States of high conductance in a large-scale model of the visual cortex. *J Comput Neurosci*, 13:93–109.

Softky, W. R. and Koch, C. (1993). The highly irregular firing of cortical cells is inconsistent with temporal integration of random EPSPs. *J Neurosci*, 13(1):334–350.

Sompolinsky, H. and Shapley, R. (1997). New perspectives on the mechanisms for orientation selectivity. *Curr Opinion Neurobiol*, 7:514–522.

Tabak, J., Rinzel, J., and O'Donovan, M. J. (2001). The role of activity-dependent network depression in the expression and self-regulation of spontaneous activity in the developing spinal cord. *J Neurosci*, 21(22):8966–8978.

Tabak, J., Senn, W., O'Donovan, M. J., and Rinzel, J. (2000). Modeling of spontaneous activity in developing spinal cord using activity-dependent depression in an excitatory network. *J Neurosci*, 20(8):3041–3056.

van Vreeswijk, C. and Sompolinsky, H. (1996). Chaos in neuronal networks with balanced excitatory and inhibitory activity. *Science*, 274:1724–1726.

van Vreeswijk, C. and Sompolinsky, H. (1998). Chaotic balanced state in a model of cortical circuits. *Neural Comp*, 10:1321–1371.

Vogels, R., Spileers, W., and Orban, G. A. (1989). The response variability of striate cortical neurons in the behaving monkey. *Exp Brain Res*, 77:432–436.

Wenner, P. and O'Donovan, M. J. (2001). Mechanisms that initiate spontaneous network activity in the developing chick spinal cord. *J Neurophysiol*, 86:1481–1498.

List of papers

1. Lerchner, A. and Rinzel, J. (2004). Synaptic model for spontaneous activity in developing networks. *to be published in Neurocomputing.*

2. Lerchner, A., Ursta, C., Hertz, J., Ahmadi, M., and Ruffiot, P. (2004). Response variability in balanced cortical networks. *submitted to Neural Compuation.*

3. Lerchner, A., Ahmadi, M., and Hertz, J. (2004). High conductance states in a mean field cortical network model. *in press: Neurocomputing.*

4. Lerchner, A., Sterner, G., Hertz, J., and Ahmadi, M. (2004). Mean field theory for a balanced hypercolumn model of orientation selectivity in primary visual cortex. *submitted to Networks: Comp in Neural Syst.*

Synaptic Model for Spontaneous Activity in Developing Networks

Alexander Lerchner

Oersted-DTU, Technical University of Denmark, 2800 Lyngby, Denmark

John Rinzel

Center for Neural Science and Courant Institute of Mathematical Sciences
New York University, New York, New York, 10003

Abstract

Spontaneous rhythmic activity is an almost universal phenomenon in developing neural networks. The activity in these hyperexcitable networks is comprised of recurring "episodes" consisting of "cycles" of high activity that alternate with "silent phases" with little or no activity. We introduce a new model of synaptic dynamics that takes into account that only a fraction of the vesicles stored in a synapse is readily available for release. We show that our model can reproduce spontaneous rhythmic activity with the same general features as observed in experiments, including a positive correlation between episode length and length of the preceding silent phase.

Key words: synapse model, hyperexcitable network, synaptic vesicle pools

1 Introduction

Network-driven spontaneous activity in developing neural networks is observed in many parts of the nervous system, including the hippocampus, the retina, and the spinal cord (see O'Donovan [3] for a review). Despite the different network architectures and the different neuron types involved, the general dynamics of this activity is always similar: recurring *episodes* of synchronous discharge, separated by *silent phases*. In the chick spinal cord, the episode

Email addresses: al@oersted.dtu.dk (Alexander Lerchner), rinzel@cns.nyu.edu (John Rinzel).

duration is 30–90 sec with a rhythmic *cycle* rate of 0.1–2 Hz that decreases during the course of each episode.

All of these networks share a common feature during this stage of development: they are hyperexcitable. For example, connections that are inhibitory in the mature chick spinal cord are functionally excitatatory in the embryo. This hyperexcitability in combination with limited resources within synapses are the key elements of our model. Especially the small terminals of synapses in the central nervous system (as compared to, e.g., neuromuscular junctions) have only a very limited number of readily available vesicles. Studies on hippocampal neurons indicate that only 15% of all vesicles in a synaptic terminal can be readily released [1]. The total number of vesicles in synaptic boutons of CNS neurons is estimated between 200 and 520 (see [2] and [6]).

In our model, we term the vesicles that can be readily released the *active pool* A, and the vesicles that need more time to be recruited (by transition into the active pool) the *storage pool* S. We show how the repeated depletion of A can produce cycles within an episode, and how the overall length of an episode is determined by the size of S at the onset of the episode. The start of an episode is triggered by spontaneous vesicle release at the synapses ("synaptic noise"), whereas slow postsynaptic depression reduces the excitability of the network immediately after an episode.

The spontaneous and evoked episodes in our model network share several of the qualitative features found in measurements of the developing chick spinal cord [7]: a decreasing cycling frequency during each episode, a positive correlation between episode duration and the length of the preceding silent phase, and a positive correlation between the length of the episode and the initial cycling frequency.

2 The Model

We introduce a new model for presynaptic dynamics that accounts for the fact that CNS nerve terminals are very tiny and contain relatively few vesicles (usually about 200 overall). Only a fraction of about 15–20% of these vesicles can engage readily in exo-endocytotic recycling [1]. Recycling of released vesicles is a fast process, making the apparent number of readily releasable vesicles (size of the active pool A) higher than their actual number within the synaptic terminal. For CNS neurons, recycling (and the compensation for loss of neurotransmitter) is fast enough that this small number does not seem to pose a major limitation for activity levels within the normal physiological range. Within a hyperexcitable network, however, synapses can be driven to the limit of their capacity due to the positive feedback in the network. Effectively,

exhausting the active pool may be a kind of activity-dependent depression acting on a short time scale: reduced sizes of A will likely be accompanied by reduced probability of vesicle release [2]. In our model, we assume that A gets constantly, but slowly, replenished by transition of vesicles from the storage pool S into A, and that the rate of this replenishment decreases with decreasing sizes of S. Finally, an even slower metabolic process is responsible to replenish S by generating "new" vesicles.

Experimentally, a slow time-delayed postsynaptic depression is observed after each episode. This slow depression reaches its maximum 1–1.5 min *after* the end of the episode and decays slowly during the entire silent phase [4]. We model this depression by a mechanism that integrates over the activity of the postsynaptic neuron within a time window of 10 sec, and which takes effect with a further delay of 15 sec. The activity within the window is compared with a range of activity in which the neuron does not adjust its sensitivity to inputs. If the activity is too high (relative to the neutral range), the synaptic strength is scaled down, if it is too low, it is scaled up. The synaptic strength is allowed to get slightly negative in response to very high activity, which can lead to post-episode hyperpolarizations as observed in intracellular recordings [3].

Within episodes, the network sustains a state of low activity between the cycle peaks. In the current form of the model, this activity is simulated by random neuronal firing at a low rate for a limited time after each cycle.

Finally, we account for observed spontaneous synaptic events that become more frequent and stronger throughout the silent phase, peaking before the spontaneous start of an episode [3]. We model these events as spontaneous vesicle release (Poisson noise) that increases in frequency during the course of the silent phase.

The synaptic model outlined above is used to connect leaky integrate-and-fire neurons randomly, to create an all-excitatory recurrent network without external input.

3 Results

The following results were obtained with a network size of $N = 300$ neurons, randomly connected with a probability of 0.4. We found the same qualitative behavior with bigger network sizes (we tried up to thousands of neurons) and different connection probabilities.

The top panel of Fig. 1 shows the average network activity plotted as function

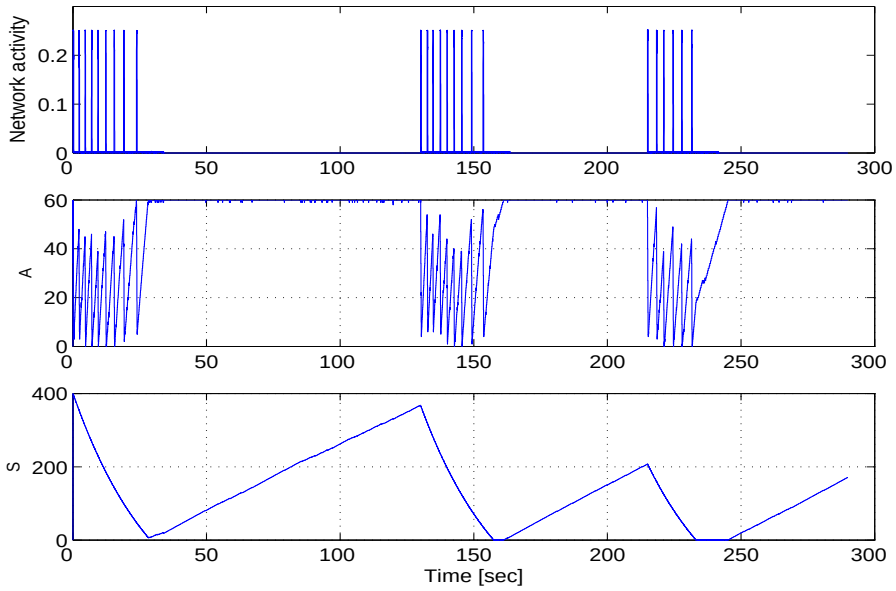

Fig. 1. Network activity, and pool sizes of a sample synapse. Top panel: Network activity showing three episodes; the second episode started spontaneously, while the third one was evoked early at time 215 sec, resulting in an episode of shorter length. The middle and the lower panel show time courses of pool sizes from a randomly chosen synapse. Middle panel: active pool A. Lower panel: storage pool S.

of time. The network activity was obtained by averaging the total number of spikes per millisecond (ms) over a sliding time window of 20 ms. With the absolute refractory period of 3 ms chosen in our simulations, the maximum average activity is 0.25. The figure shows the episodic nature of the activity, where three episodes of cycling activity can be seen. The first episode was evoked at the beginning of the simulation. The second episode started spontaneously, while the third one was evoked by driving the membrane potential of all neurons above threshold for a few milliseconds at time $t = 215$ sec. It can be seen that the cycling frequency decreases during the course of each episode. The prematurely evoked episode is shorter in length and starts with a lower cycling frequency.

Generally, we found a positive correlation between length of the silent phase and length of the following episode. We also observed that shorter episodes always started with a lower cycle frequency. These correlations are qualitatively similar to experimental findings [7].

The time course of a the active pool A and the storage pool S of a randomly chosen synapse in the network are plotted in Fig. 1 in the middle and lower panel, respectively. The size of A fluctuates synchronously to the cycles of activity within the episodes, while the size of S decreases monotonically during each episode until it reaches zero at the end of the episode. During the silent phase, S is refilled slowly.

4 Discussion

The nature of the correlations between episode length and length of the preceding silent phase in our model can be understood by comparing the time courses of the synaptic pool sizes with the network activity (Fig. 1). The activity cycles are reflected as an alternation between depletion and replenishment of the active pool A. Cycles are terminated because of exhaustion of the active pool, which reduces the network activity drastically and gives A time to be refilled, until a new avalanche of activity exhausts A once again. During the course of the episode, the storage pool S reduces in size, making the replenishment of A slower and therefore the interval between cycles increases. At some point, the storage pool S is depleted and the episode terminates. The length of the silent phase determines the size of S at the start of the following episode, which in turn determines the episode length.

This model also attempts to address an aspect of the the exceptional data that was ignored or unaccounted for in previous modeling studies. It is considered a puzzling observation that there is almost no postsynaptic depression at the end of the episode [4]. Rather, it sets in steeply afterwards to reach its maximum 1–1.5 min later, then relaxing slowly back over the entire silent phase to reach once again the high level it had at the end of the episode (see [7], Fig. 4C). In agreement with these observations, the slow postsynaptic depression is not causing the termination of episodes in our model. Instead, it removes temporarily the hyperexcitability in the network, preventing episodes to be started spontaneously before the storage pool is sufficiently replenished. The model network displays an absolute refractory period regarding episode initiation, in which episodes cannot even be triggered by external stimulation. The same phenomenon is observed experimentally in the developing chick spinal cord (see [7], Fig. 4B). Moreover, the slowly decaying postsynaptic depression modulates the amplitude of the spontaneous synaptic activities, which increase during the silent phase until they reach a maximum just before the initiation of a new episode. This description applies to both the developing chick spinal cord and the dynamics of our model.

There has been success in modeling spontaneous activity on the network-dynamics level. This was achieved by employing both slow and fast activity-dependent *network depression* that act independently from each other [5]. However, it has proven challenging to show how such network dynamics can emerge from neuron properties. This work is an attempt to provide such an explanation, albeit with different dynamics than have been considered so far. It suggests that the limited resources of synapses and their proposed internal dynamics may lead to the typical activity patterns observed in hyperexcitable networks of diverse architectures.

References

[1] Harata, N. *et al.* (2001) Limited numbers of recycling vesicles in small CNS nerve terminals: implications for neural signaling and vesicular cycling. Trends Neurosci 24:637–643

[2] Harris, K.M. and Sultan, P. (1995) Variation in the number, location and size of synaptic vesicles provides an anatomical basis for the nonuniform probability of release at hippocampal CA1 synapses. Neuropharmacology 34:1387-1395

[3] O'Donovan, M.J. (1999) The origin of spontaneous activity in developing networks of the vertebrate nervous system. Curr Opin Neurobiol 9:94–104

[4] O'Donovan, M.J., Chub, N. and Wenner, P. (1998) Mechanisms of spontaneous activity in developing spinal networks. J Neurobiol 37: 131-145

[5] Tabak, J., Senn, W., O'Donovan, M.J. and Rinzel J. (2000) Modeling of spontaneous activity in developing spinal cord using activity-dependent depression in an excitatory network. J Neurosci 20:3041–3056

[6] Schikorski, T. and Stevens, C.F. (1997) Quantitative ultrastructural analysis of hippocampal excitatory synapses. J Neurosci 17:5858–5867

[7] Tabak, J., Rinzel, J. and O'Donovan, M.J. (2001) The role of activity-dependent network depression in the expression and self-regulation of spontaneous activity in the developing spinal cord. J Neurosci 21:8966-8978

arXiv:q-bio.NC/0402022 v1 10 Feb 2004

Response Variability in Balanced Cortical Networks

Alexander Lerchner
Technical University of Denmark, 2800 Lyngby, Denmark

Cristina Ursta
Niels Bohr Institut, Blegdamsvej 17, 2100 Copenhagen Ø, Denmark

John Hertz
Nordita, Blegdamsvej 17, 2100 Copenhagen Ø, Denmark

Mandana Ahmadi
Nordita, Blegdamsvej 17, 2100 Copenhagen Ø, Denmark

Pauline Ruffiot
Université Joseph Fourier, Grenoble, France

February 10, 2004

Abstract

We study the spike statistics of neurons in a network with dynamically balanced excitation and inhibition. Our model, intended to represent a generic cortical column, comprises randomly connected excitatory and inhibitory leaky integrate–and–fire neurons, driven by excitatory input from an external population. The high connectivity permits a mean-field description in which synaptic currents can be treated as Gaussian noise, the mean and autocorrelation function of which are calculated self-consistently from the firing statistics of single model neurons. Within this description, we find that the irregularity of spike trains is controlled mainly by the strength of the synapses relative to the difference between the firing threshold and the post-firing reset level of the membrane potential. For moderately strong synapses we find spike statistics very similar to those observed in primary visual cortex.

1 Introduction

The observed irregularity and relatively low rates of the firing of neocortical neurons suggest strongly that excitatory and inhibitory input are nearly balanced. Such a balance, in turn, finds an attractive explanation in the mean-field descriptions of Amit and Brunel [1, 2, 3] and Van Vreeswijk and Sompolinsky

[4, 5]. In their theories, the balance does not have to be put in "by hand"; rather, it emerges self-consistently from the network dynamics. This success encourages us to study firing correlations and irregularity in models like theirs in greater detail. In particular, we would like to quantify the irregularity and identify the parameters of the network that control it. This is important because one can not extract the signal in neuronal spike trains correctly without a good characterization of the noise. Indeed, an incorrect noise model can lead to spurious conclusions about the nature of the signal, as demonstrated by Oram *et al* [6].

Response variability has been studied for a long time in primary visual cortex [9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19] and elsewhere [20, 17, 18, 21]. Most, though not all, of these studies found rather strong irregularity. As an example, we consider the findings of Gershon *et al* [17]. In their experiments, monkeys were presented with flashed, stationary visual patterns for several hundred ms. Repeated presentations of a given stimulus evoked varying numbers of spikes in different trials, though the mean number (as well as the PSTH) varied systematically from stimulus to stimulus. The statistical objects of interest to us here are the distributions of single-trial spike counts, for given fixed stimuli. Often one compares the data with a Poisson model of the spike trains, for which the count distribution $P(n) = m^n \mathrm{e}^{-m}/n!$. This distribution has the property that its mean $\langle n \rangle = m$ is equal to its variance $\langle \delta n^2 \rangle = \langle (n - \langle n \rangle)^2 \rangle$. However, the experimental finding was that the measured distributions were quite generally wider than this: $\langle \delta n^2 \rangle > m$. Furthermore, collecting data for many stimuli, the variance of the spike count was fit well by a power law function of the mean count: $\langle \delta n^2 \rangle \propto m^y$, with y typically in the range $1.2 - 1.4$, broadly consistent with the results of many of the other studies cited above.

Some of this observed variance could have a simple explanation: The condition of the animal might have changed between trials, so the intrinsic rate at which the neuron fires might differ from trial to trial, as suggested by Tolhurst *et al* [11]. But it is far from clear whether all the variance can be accounted for in this way. Moreover, there is no special reason to take a Poisson process as the null hypothesis, so we don't even really know how much variance we are trying to explain.

In this paper, we try to address the question of how much variability, or more generally, what firing correlations can be expected as consequence of the intrinsic dynamics of cortical neuronal networks. The theories of Amit and Brunel and of van Vreeswijk and Sompolinsky do not permit a consistent study of firing correlations. The Amit-Brunel treatment assumes that the input to neurons is uncorrelated in time (white noise). Thus, although one can calculate the variability of the firing [3], it is not self-consistent. Van Vreeswijk and Sompolinsky use a binary-neuron model with stochastic dynamics which makes it difficult, if not impossible, to study temporal correlations that might occur in networks of spiking neurons. Therefore, in this paper we do a complete mean-field theory for a network of leaky integrate-and-fire neurons, including, as self-consistently-determined order parameters, both firing rates and autocorrelation functions. A general formalism for doing this was introduced by Fulvi Mari [7]

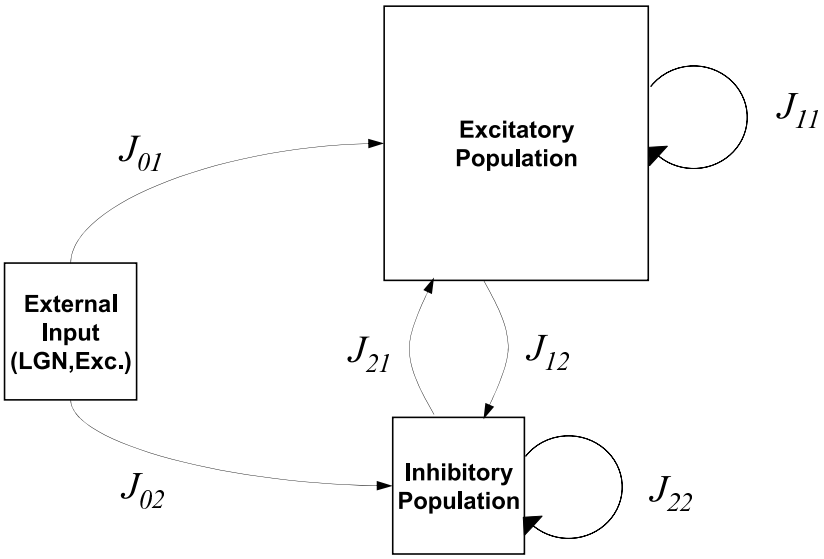

Figure 1: Structure of the model network.

and used for an all-excitatory network; here we employ it for a network with both excitatory and inhibitory neurons. A preliminary study of this approach for an all-inhibitory network was presented previously [22].

2 Model and Methods

The model network, indicated schematically in Fig. 1, consists of N_1 excitatory neurons and N_2 inhibitory ones. In this work we use leaky integrate-and-fire neurons, though the methods could be carried over directly to networks of other kinds of model neurons, such as conductance-based ones. They are randomly interconnected by synapses, both within and between populations, with the mean number of connections from population b to population a equal to K_b, independent of a. In specific calculations, we have used K_1 from 400 to 6400, and we take $K_2 = K_1/4$. The population sizes N_a do not enter directly in the mean field theory, only their ratios (the connection probabilities) K_a/N_a. We have used $K_a/N_a = 0.1$ for both excitatory and inhibitory connections, implying $N_1 = 4N_2$.

We scale the synaptic strengths in the way van Vreeswijk and Sompolinsky did[4, 5], with each nonzero synapse from population b to population a having the value $J_{ab}/\sqrt{K_b}$. The parameters J_{ab} are taken to be of order 1, so the net input current to a neuron from the K_b neurons in population b connected to it is of order $\sqrt{K_b}$. With this scaling, the fluctuations in this current are of order 1.

Similarly, we assume that the external input to any neuron is the sum of $K_0 \gg 1$ contributions from individual neurons (in the LGN, if we are thinking about modeling V1), each of order $1/\sqrt{K_0}$, so the net input is of order $\sqrt{K_0}$. In our calculations, we have used $K_0 = K_1$.

We point out that this scaling is just for convenience in thinking about the problem. In the balanced asynchronous firing state, the large excitatory and inhibitory input currents nearly cancel, leaving a net input current of order 1. Thus, for this choice, both the net mean current and its typical fluctuations are of order 1, which is convenient for analysis. The physiologically relevant assumptions are only that excitatory and inhibitory inputs are separately much larger than their sum and that the latter is of the same order as its fluctuations.

Our synapses are not modeled as conductances. Our synaptic strength simply defines the amplitude of the postsynaptic current pulse produced by a single presynaptic spike.

The model is formally specified by the sub-threshold equations of motion for the membrane potentials u_i^a ($a = 1, 2$, $i = 1, \ldots N_a$):

$$\frac{\mathrm{d}u_i^a}{\mathrm{d}t} = -\frac{u_i^a}{\tau} + \sum_{b=0}^{2} \sum_{j=1}^{N_b} J_{ij}^{ab} S_j^b(t), \tag{1}$$

together with the condition that when u_i^a reaches the threshold θ_a, the neuron spikes and the membrane potential is reset to a value u_r^a. The indices a or $b = 1$ or 2 label populations: $b = 0$ refers to the (excitatory) population providing the external input, $a = 1$ refers to the excitatory population and $a = 2$ to the inhibitory population. In (1), τ is the membrane time constant (taken the same for all neurons, for convenience), the strength of the synapse from neuron j in population b to neuron i in population a is denoted by J_{ij}^{ab}, and $S_i^b(t) = \sum_s \delta(t - t_{jb}^s)$ is the spike train of neuron j in population b. We have ignored transmission delays, and we take the thresholds $\theta_a = 1$ and the reset levels u_r^a equal to the rest value of the membrane potential, 0. In our calculations, the thresholds are given a Gaussian distribution with a standard deviation equal to 10% of the mean. Analogous variability in other single-cell parameters (such as membrane time constants) could also be included in the model, but for simplicity we do not do so here.

We assume that the neurons in the external input population ($b = 0$) fire as independent Poisson processes. However, the neurons in the network ($b = 1, 2$) are not in general Poissonian; it is their correlations that we want to find in this investigation.

Mean Field Theory: Stationary States

We describe the mean field theory and its computational implementation first for the case of stationary rates. When the connectivity is large and random, as we will assume here, each of the three terms in the sum on b on the right-hand side of (1) can be treated as a Gaussian random function with time-independent mean. The simplest case is $b = 0$, the external input. For simplicity, we assume that all N_0 neurons in the external population fire at the same rate, r_0. But because of the random connectivity, the net time-averaged input current they provide to a neuron in cortical population a can vary from neuron to neuron.

Assuming large, dilute connectivity ($K_0 \gg 1$ and $K_0 \ll N_0$), the central limit theorem then implies

$$\langle I_i^{a0}(t)\rangle = \sum_j J_{ij}^{a0} r_0 = \sum_j (\overline{J_{ij}^{a0}} + \delta J_{ij}^{a0}) r_0 = J_{a0} r_0 (\sqrt{K_0} + x_i^{a0}), \tag{2}$$

where x_i^{a0} is a Gaussian-distributed random number of unit variance. By $\langle \cdots \rangle$ we mean a time average or, equivalently, an average over "trials" (independent repetitions of the Poisson processes defining the input population neurons). We will generally use a bar over a quantity to indicate an average over the neuronal population or over the distribution of the J_{ij}^{ab}. (Note that these two kinds of averages are very different things.)

Writing the spike train $S_j^0(t)$ for neuron j in the input population as

$$S_j^0(t) = r_0 + \delta S_j^0(t), \tag{3}$$

with $\langle \delta S_j^0(t)\rangle = 0$, we can write the fluctuations around $\langle I_i^{a0}\rangle$ as

$$\delta I_i^{a0}(t) = \sum_j J_{ij}^{a0} \delta S_j^0(t) = J_{a0} \xi_i^{a0}(t) \tag{4}$$

where $\xi_i^{a0}(t)$ is white noise of power r_0:

$$\langle \xi_i^{a0}(t) \xi_i^{a0}(t')\rangle = r_0 \delta(t - t') \tag{5}$$

Thus, quite generally, the input has a large mean value, of order $\sqrt{K_0}$, plus Gaussian fluctuations of order 1. The fluctuations are of two kinds. One is constant for a given neuron, independent of time and trial and arises from the fact that the connectivity is random and the neurons in the input population have a distribution of rates. The other fluctuation is a dynamical one, with correlations (independent of i) reflecting the Poisson dynamics of the input population neurons.

The recurrent input terms $I_i^{ab}(t)$ also have large means and fluctuations, static and dynamic, of order 1, but certain features of their statistics are slightly different, as a systematic formal derivation [7, 8] proves. Here we do not give the derivation, but just describe the result, which is that $I_i^{ab}(t)$ can be written

$$I_i^{ab}(t) = J_{ab}[\sqrt{K_b} r_b + B_b x_i^{ab} + \sqrt{1 - K_b/N_b}\, \xi_i^{ab}(t)], \tag{6}$$

with

$$r_b = \overline{r_j^b} = \frac{1}{N_b} \sum_j r_j^b. \tag{7}$$

the average rate in population b, x_i^{ab} a unit-variance Gaussian random number,

$$B_b = \sqrt{\left(1 - \frac{K_b}{N_b}\right) \overline{(r_j^b)^2}} \tag{8}$$

and

$$\langle \xi_i^{ab}(t)\xi_i^{ab}(t')\rangle = C_b(t-t'). \tag{9}$$

Here $C_b(t-t')$ is the average autocorrelation function of the firing of neurons in population b,

$$C_b(t-t') = \frac{1}{N_b}\sum_j \langle \delta S_j^b(t)\delta S_j^b(t')\rangle, \tag{10}$$

with

$$\delta S_j^b(t) = S_j^b(t) - r_j^b. \tag{11}$$

Again, x_i^{ab} is time- and trial-independent, while the noise $\xi_i^{ab}(t)$ varies both in time within a trial and randomly from trial to trial. Note that for this model a correct and complete mean field theory has to include rate fluctuations, through $\overline{(r_j^b)^2}$, and the firing correlations, given by $C_b(t-t')$, as well as the mean rates.

The means of the recurrent input currents I_i^{ab} are completely analogous to the mean term in I_i^{a0}, but the effective noise is different in three ways

1. The amplitude B_b of the static noise component (the second term) contains a factor of the rms rate $\sqrt{\overline{(r_j^b)^2}}$, not r_b as in (2). The same would be true for the static input noise ($b = 0$) if we allowed a distribution of rates in the input population. So this difference is not an essential one. It occurs only because we made a simplifying assumption about the input population. However, we are not allowed to assume that about the neurons in the cortical network, which will always have a distribution of rates because of the random connectivity.

2. The neurons providing the source of these currents are not generally Poissonian, so their correlations appear in the statistics of the noise term.

3. The noise terms, both static and dynamic, have a factor $\sqrt{1-K_b/N_b}$ in front of them. This can be understood in the following way: It is the randomness in the synaptic connections in the network that generates these noise terms in the effective single-neuron problem; in general, they are proportional to $\overline{(\delta J_{ij}^{ab})^2}$, which is equal to $J_{ab}^2(1-K_b/N_b)/N_b$ in our model. In the limit of full connectivity, $K_b = N_b$, all J_{ij}^{ab} are equal and there is no randomness. Therefore there is no noise, as guaranteed here by this factor.

The self-consistency equations of mean field theory are simply the conditions that the average output statistics of the neurons, r_a, $\overline{(r_j^a)^2}$ and $C_a(t-t')$ are the same as those used to generate the inputs for single neurons using integrate-and-fire neurons with synaptic input currents given by (2), (4) and (6).

In an equivalent formulation, the second term in (6) can be omitted if the noise terms $\xi_i^{ab}(t)$ have correlations equal to the unsubtracted correlation function

$$C_b^{\rm tot}(t-t') = \frac{1}{N_b}\sum_j \langle S_j^b(t)S_j^b(t')\rangle \tag{12}$$

instead of (10). For $|t-t'| \to \infty$, $C_b^{\rm tot}(t-t') \to \overline{(r_j^2)}$, so $\xi_i^{ab}(t)$ acquires a random static component of mean square value $\overline{(r_j^b)^2}$.

In still another way to do it, one can use the square of the average rate, r_b^2 in place of $\overline{(r_j^b)^2}$ in Eq.(8) for B_b and employ noise with correlation function

$$\tilde{C}_b(t-t') = \frac{1}{N_b}\sum_j \langle (S_j^b(t) - r_b)(S_j^b(t') - r_b)\rangle. \tag{13}$$

For $|t-t'| \to \infty$,

$$\tilde{C}_b(t-t') \to \overline{(r_j^b - r_b)^2} \equiv \overline{(\delta r_j^b)^2}. \tag{14}$$

There are now two static random parts of $I_i^{ab}(t)$, one from the B_b term and one from the static component of the noise. Their sum is a Gaussian random number with standard deviation equal to B_b as given in (6). Thus these three ways of generating the input currents are all equivalent.

The balance condition

In a stationary, low-rate state, the mean membrane potential described by (1) has to be stationary. If excitation dominates, we have $\mathrm{d}u_i^a/\mathrm{d}t \propto \sqrt{K_0}$, implying a firing rate of order $\sqrt{K_0}$ (or one limited only by the refractory period of the neuron). If inhibition dominates, the neuron will never fire. The only way to have a stationary state at a low rate (less than one spike per membrane time constant) is to have the excitation and inhibition nearly cancel. Then the mean membrane potential can lie a little below threshold, and the neuron can fire occasionally due to the input current fluctuations. Thus, using (2) and (6), we have

$$\sum_{b=0}^{2} J_{ab}\sqrt{K_b} r_b = \mathcal{O}(1) \tag{15}$$

or, up to corrections of $\mathrm{O}(1/\sqrt{K_0})$,

$$\sum_{b=0}^{2} \hat{J}_{ab} r_b = 0 \tag{16}$$

with $\hat{J}_{ab} = J_{ab}\sqrt{K_b/K_0}$. These are two linear equations in the two unknowns r_a, $a = 1, 2$, with the solution

$$r_a = \sum_{b=1}^{2} [\hat{\mathsf{J}}^{-1}]_{ab} J_{b0} r_0, \tag{17}$$

where $\hat{\mathsf{J}}^{-1}$ is the inverse of the 2×2 matrix with elements $\hat{J}_{ab}$, $a, b = 1, 2$. If there is a stationary balanced state, the average rates of the excitatory and inhibitory populations are given by (17) (in the large-N limit).

This argument depends only on the rates, not on the correlations, and is exactly the same as that given by Amit and Brunel and by Sompolinsky and van Vreeswijk.

This calculation does not say whether this state is stable, however. To determine this, one can expand around this solution and examine the linear stability of the fluctuations, as done for their model by van Vreeswijk and Sompolinsky [5]. Here, we do not do this analytically, but rather check the stability of our states numerically within our algorithm.

Numerical procedure

For integrate-and-fire neurons in a stationary state, the mean field theory can be carried out analytically if a white-noise (Poisson firing) approximation is made [1, 2, 3]. But if firing correlations are to be taken into account, it is necessary to resort to numerical methods. Thus we simulate single neurons driven by Gaussian synaptic currents, collect their firing statistics to compute the rates r_a, rate fluctuations $\overline{(\delta r_j^a)^2}$ and correlations $C_a(t-t')$, and then use these to generate improved input current statistics. The cycle is repeated until the input and output statistics are consistent. This algorithm was first used by Eisfeller and Opper [23] to calculate the remanent magnetization of a mean field model for spin glasses.

Explicitly, we proceed as follows. We simulate single excitatory and inhibitory neurons over "trials" 100 integration timesteps long. (We will call each timestep a "millisecond". We have explored using smaller timesteps and verified that there are no qualitative changes in the results.) We start from estimates of the rates given by the balance condition, which makes the net mean input current vanish. Then the sum of the $\mathcal{O}(\sqrt{K_b})$ terms in (2) and (6) vanishes, leaving only the rate fluctuation and noise terms. We then run 10000 trials of single excitatory and inhibitory neurons, selecting on each trial random values of x_i^{ab} and $\xi_i^{ab}(t)$. Since at this point we do not have any estimates of either the rate fluctuations $\overline{(\delta r_j^b)^2}$ or the correlations $C_b(t-t')$, we use r_b^2 in place of $\overline{(r_j^b)^2}$ in Eq.(8)for B_b and use white noise for $\xi_i^{ab}(t)$: $C_b(t-t') \to r_b\delta(t-t')$.

The random choice of x_i from trial to trial effectively samples across the neuronal populations, so we can then collect the statistics r_a, $\overline{(r_j^a)^2}$ (or, equivalently, $\overline{(\delta r_j^a)^2}$), and $C_a(t-t')$ from these trials. These can be used to generate an improved estimate of the input noise statistics to be used in (6) in a second set of trials, which yields new spike statistics again. This procedure is iterated until the input and output statistics agree. This may take up to several hundred iterations, depending on network parameters and how the computation is organized.

If one tries this procedure in its naive form, i.e., using the output statistics directly to generate the input noise at the next step, it will lead to big oscillations and not converge. It is necessary to make small corrections (of relative order $1/K_0$) to the previous input noise statistics to guarantee convergence.

When one computes statistics from the trials in any iteration, the simplest

procedure involves calculating not (10), but rather $\tilde{C}_b(t-t')$ (Eq.(13)). From it, we can proceed in two ways. In the first, from its $|t-t'| \to \infty$ limit we can obtain $\overline{(\delta r_j^b)^2}$, and thereby $\overline{(r_j^b)^2} = r_b^2 + \overline{(\delta r_j^b)^2}$ for use in calculating B_b in (8). Subtracting this limiting value from $\tilde{C}_b(t-t')$ give us $C_b(t-t')$ (which vanishes for large $|t-t'|$) for use in generating the noise $\xi_i^{ab}(t)$. This is the first of the three methods described above.

Alternatively, we can use the third method: At each step of our iterative procedure we can generate noise directly with the correlations $\tilde{C}_b(t-t')$ (which are long-ranged in time) and use r_b^2 in place of $\overline{(r_j^b)^2}$ in calculating B_b (8). We have verified that the two methods give the same results when carried out numerically, though the second procedure converges more slowly.

While the true rates in the stationary case are time-independent and $C_a(t,t')$ is a function only of $t-t'$, the statistics collected over a finite set of noise-driven trials will not exactly have these stationarity properties. Therefore we improve the statistics and impose time-translational invariance by averaging the measured $r_a(t)$ and $\overline{(\delta r_j^a(t))^2}$ over t and averaging over the measured values $C_a(t,t')$ with a fixed $t-t'$.

After the iterative procedure converges, so that we have a good estimate of the statistics of the input, we want to run many trials *on a single neuron* and compute its firing statistics. This means that *the numbers x_i^{ab} ($b = 0, 1, 2$) should be held constant over these trials.* In this case it is necessary to subtract out the large $t-t'$ limit of $\tilde{C}_a(t-t')$ and use fixed x_i^{ab} (constant in time and across trials) to generate the input noise. (If we did it the other way, without the subtraction, we would effectively be assuming that x_i^{ab} changed randomly from trial to trial, which is not correct.)

In our calculations we have used 10000 trials to calculate these single-neuron firing statistics. We perform the subtraction of the long-time limit of $\tilde{C}_a(t-t')$ at $|t-t'| = 50$, and we have checked that (13) is flat beyond this point in all the cases we have done.

If we perform this kind of measurement separately for many values of the x_i^{ab}, we will be able to see how the firing statistics vary across the population. Here, however, we will confine most of our attention to what we call the "average neuron": the one with the average value (0) of all three x_i^{ab}.

In particular, we calculate the mean spike count in the 100-ms trials and its variance across trials. From this we can get the Fano factor F (the variance/mean ratio). We also compute the autocorrelation function, which offers a consistency check, since the Fano factor can also be obtained from

$$F = \frac{1}{r} \int_{-\infty}^{\infty} C(\tau) \mathrm{d}\tau. \tag{18}$$

(This formula is valid when the measurement period is much larger than the time over which $C(\tau)$ falls to zero.)

We will study below how these firing statistics vary as we change various parameters of the model: the input rates r_0, parameters that control the balance of excitation and inhibition, and the overall strength of the synapses. This will

give us some generic understanding of what controls the degree of irregularity of the neuronal firing.

Nonstationary Case

When the input population is not firing at a constant rate, almost the same calculational procedure can be followed, except that one does not average measured rates, their fluctuations or correlation function over time. To start out, we get initial instantaneous rate estimates from the balance condition, assuming that the time-dependent average input currents do not vary too quickly. (This condition is not very stringent; van Vreeswijk and Sompolinsky showed that the stability eigenvalues are proportional to $\sqrt{K_0}$, so if they have the right sign the convergence to the balanced state is very rapid.)

To do the iterative procedure to satisfy the self-consistency conditions of the theory, it is simplest to use the second of the two ways described above (not doing any subtraction until the final calculations with single neurons). In this case the expression (8) does not include rate fluctuations, and we get equations for the noise input currents just like (2), (4) and (6) except that the r_b are t-dependent and the correlation functions C_b and D_b depend on both t and t', not just their difference.

The only tricky part is the subtraction of the long-time limit of the correlation function, which is not simply defined.

We treat this problem in the following way. We examine the rate-normalized quantity

$$\hat{D}_a(t,t') = \frac{D_a(t,t')}{r_a(t)r_a(t')}. \tag{19}$$

We find that this quantity is time-translation invariant (i.e., a function only of $t - t'$) to a very good approximation, so we perform the subtraction of the long-time limit on it. Then multiplying the subtracted $\hat{D}$ by $r_a(t)r_a(t')$ gives a good approximation to the true correlation function $C_a(t,t')$. The meaning of this finding is, loosely speaking, that when the rates vary (slowly enough) in time, the correlation functions just inherit these rates as overall factors without changing anything else about the problem.

We will use the this time-dependent formulation below to simulate experiments like those of Gershon *et al* [17], where the LGN input $r_0(t)$ to visual cortical cells is time-dependent because of the flashing-on and off of the stimulus.

3 Results

The results presented in this chapter were obtained from simulations with parameters corresponding to population sizes of $N_1 = 40{,}000$ excitatory neurons and $N_2 = 10{,}000$ inhibitory neurons. With the above mentioned connection probabilities of $K_a/N_a = 0.1$, this translates to an average number of $K_1 = 4000$ excitatory inputs and $K_2 = 1000$ inhibitory inputs to each neuron. The average

number of external (excitatory) inputs K_0 was chosen to be equal to K_2. All neurons have the same membrane time constant τ of 10 ms.

To study the effect of various combinations in synaptic strength, we use the following generic form to define the intra-cortical weights J_{ab}:

$$\begin{pmatrix} J_{11} & J_{12} \\ J_{21} & J_{22} \end{pmatrix} = \begin{pmatrix} \epsilon & -2g \\ 1 & -2g \end{pmatrix} \tag{20}$$

For the synaptic strengths from the external population we use $J_{10} = 1$ and $J_{20} = \epsilon$. With this notation, g determines the strength of inhibition relative to excitation within the network, and ϵ the strength of intracortical excitation. Additionally, we scale the overall strength of the synapses with a multiplicative scaling factor denoted J_s so that each synapse has an actual weight of $J_s \cdot J_{ab}$, regardless of a and b.

Figure 2 summarizes how the firing firing statistics depend on all of the parameters g, ϵ, and J_s. The irregularity of spiking, as measured by the Fano factor, depends most sensitively on the overall scaling of the synaptic strength, J_s. The Fano factor increases systematically as J_s increases, and higher values of intracortical excitation ϵ also result in higher values of F. The same pattern holds for stronger intracortical inhibition, parameterized by g. For all of these cases the mean firing rate remains virtually unchanged due to the dynamic balance of excitation and inhibition in the network, whereas the fluctuations increase with the increase of any of the synaptic weights.

Interspike interval (ISI) distributions are shown in Figure 3 for three different values of J_s, keeping ϵ and g fixed at 0.5 and 1, respectively. For a Poisson spike train, the Fano factor $F = 1$, while $F > 1$ (which we term "superpoissonian") indicates a tendency of spikes occurring in clusters separated by accordingly longer empty intervals, and $F < 1$ ("subpoissonian") indicates more regularity, reflected by a narrower distribution. We have adjusted the input rate r_0 so that the output rate is the same in all three cases.

The top panel of Figure 3 shows the ISI distribution of a superpoissonian spike train, obtained for $J_s = 1.5$. Overlayed on the histogram of ISI counts is an exponential curve indicating a Poisson distribution with the same mean ISI length. Compared with the Poisson distribution, the superpoissonian spike train contains more short intervals, as seen by the peak at short lengths, and also more long intervals, causing a long tail. Necessarily, the interval count around the average ISI length is lower than that for a Poisson spike train.

The ISI distribution in the middle panel of Figure 3 belongs to a spike train with a Fano factor close to one, obtained for $J_s = 0.75$. The overlayed exponential reveals a deviation from the ISI count: while intervals of diminishing length are the most likely ones for a real Poisson process, our neuronal spike trains always show some refractoriness reflected by a dip at the shortest intervals. (We have not used an explicit refractory period in our model. The dip seen here simply reflects the fact that it takes a little time for the membrane potential distribution to return to its steady-state form after reset.) Apart from this deviation, however, there is a close resemblance between the observed distribution and the "predicted" one.

Finally, the lower panel of Figure 3 depicts a case with $F < 1$, with weaker synapses, leading to a stronger refractory effect and (since the rate is fixed) an accordingly narrower distribution around the average ISI length, as compared to the overlayed Poisson distribution. This distribution was obtained with weak synapses produced by a small scaling factor of $J_s = 0.375$.

As mentioned in the previous section, the Fano factor can also be obtained by integrating over the spike train autocorrelation divided by the spike rate (18). For a Poisson process the autocorrelation vanishes for all lags different from zero. In contrast, $F > 1$ (superpoissonian case) implies a positive integral over non-zero lags, whereas in the subpoissonian case there must be a negative area under the curve. Figure 4 shows examples of autocorrelations for all of the three cases. For the superpoissonian case (red dash-dot line), there is a "hill" of positive correlations for short intervals, reflecting the tendency toward spike clustering. The subpoissonian autocorrelation (blue dotted line) shows a valley of negative correlations for short intervals, indicating well separated spikes in a more regular spike train. The curve labeled as Poisson (black solid line) does have a small valley around zero lag, which reflects once more the refractoriness of neurons to fire at extremely short intervals, unlike a completely random Poisson process. (Actually, the measured F in this case is slightly greater than 1, implying that in this case the integral of the very small positive tail for $t > 2$ ms is slightly larger than that of the (more obvious) negative short-time dip.)

Measurements on V1 neurons in awake monkeys (see for example Gershon *et al.* [17]) suggest a linear relationship between the log variance and the log mean of stimulus-elicited spike counts. We find a similar dependence for neurons within our model network. Figure 5 shows results for three different values of J_s. In each case, five different values of the external input rate r_0 were tried, causing various mean spike counts and variances. The logarithm of the spike count variance is plotted as a function of the logarithm of the spike count mean, and a solid diagonal line indicates the identity, i.e, a Fano factor of exactly 1. We see that for the largest value of J_s used here, the data look qualitatively like those from experiments, with Fano factors in the range around 1.5 to 2.

Nonstationary Case

The results presented in the previous section were obtained with stationary inputs, while experimental data like those from [17] were collected from visual neurons subject to time-dependent inputs. Therefore, we performed calculations of the spike statistics in which the input population rate r_0 was time-dependent. The modeled temporal shape of $r_0(t)$ is depicted in Figure 6. It is the sum of three terms:

$$r_0(t) = R_0 + A(t) + B(t) \tag{21}$$

The first, R_0, is just a constant, as in the preceding section. The second term, $A(t)$, rises to a maximum over a 25-ms interval, remains constant for 50 ms,

A_0	0.375	0.375	0.500	0.500	0.750	0.750	1.000	1.000
B_0	0.125	0.375	0.125	0.375	0.250	0.750	0.250	0.750
F	1.14	1.2	1.22	1.23	1.29	1.36	1.37	1.4
A_0	1.500	1.500	2.000	2.000	3.000	3.000	4.000	4.000
B_0	0.500	1.500	0.500	1.500	1.000	3.000	1.000	3.000
F	1.48	1.5	1.55	1.53	1.57	1.41	1.43	1.34

Table 1: Stimulus parameters A_0 and B_0 for the results depicted in Figure 7, and the resulting Fano factors F.

and then falls off to zero over the final 25 ms.

$$A(t) = \begin{cases} 0.5A_0(1 - \cos(4t\pi/T)) & \text{for } 0 < t \leq T/4 \\ A_0 & \text{for } T/4 < t \leq 3T/4 \\ 0.5A_0(1 - \cos(4(T-t)\pi/T)) & \text{for } 3T/4 < t \leq T, \end{cases} \tag{22}$$

where T is the total simulation interval of 100 ms. The third term, B_0, rises to a maximum in the first 25 ms and then falls back to zero in the next 25 ms, remaining zero thereafter.

$$B(t) = \begin{cases} 0.5B_0(1 - \cos(4t\pi/T)) & \text{for } 0 < t \leq T/4 \\ 0.5B_0(1 - \cos(4(T/2-t)\pi/T)) & \text{for } T/4 < t \leq T/2 \\ 0 & \text{for } T/2 < t \leq T. \end{cases} \tag{23}$$

Figure 7 shows the logarithm of the spike count variance plotted against the logarithm of the spike count mean for various non-stationary inputs characterized by different values of A_0 and B_0. The graph shows results for $J_s = 1$, $\epsilon = 0.5$, $g = 1$, and a background rate of $R_0 = 0.1$. Table 1 shows the choice of the sixteen combinations of the stimulus parameters A_0 and B_0, together with the resulting Fano factors F for the simulated neuron.

The data look qualitatively like those obtained from *in-vivo* experiments [17] and are similar to the superpoissonian case in Figure 5. The neuron fires consistently in a superpoissonian regime with Fano factors slightly higher than 1 and an almost linear relationship between the log variance and the log mean for low spike counts. For higher spike counts, the curve bends towards values of lower Fano factors, just as for stationary inputs (Figure 5). In both cases, this bend reflects the the decrease in irregularity of firing caused by an increasingly prominent role of refractoriness for shorter interspike intervals.

4 Discussion

Cortical neurons receive thousands of both excitatory and inhibitory inputs, and despite the high number of inputs from nearby neurons with similar firing statistics and similar connectivity, their observed firing is very irregular [9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21]. Dynamically balanced excitation and inhibition through a simple feedback mechanism provides an explanation that

naturally accounts for this phenomenon without requiring fine tuning of the parameters [1, 2, 3, 4, 5]. Moreover, neurons in such model networks show an almost linear input-output relationship (input current versus firing frequency), as do neurons in the neocortex.

Here, we have extended the mean-field description of the dynamically balanced asynchronous firing state to analyze firing correlations. We found that the relationship between the observed irregularity of firing (spike count variance) and the firing rate (spike count mean) of the neurons resemble closely data collected from *in-vivo* experiments (see Figures 5 and 6). To do this, we developed a complete mean-field theory for a network of leaky integrate-and-fire neurons, in which both firing rates and correlation functions are determined self-consistently. Using an algorithm that allows us to find the solutions to the mean-field equations numerically, we could elucidate how the strength of synapses within the network influences the expected firing statistics of cortical neurons in a systematic manner (see Figure 2).

We have shown that the irregularity of firing, as measured by the Fano factor, increases with increasing synaptic strengths (Figure 2). Nearly Poisson statistics (with $F \approx 1$) are observed for moderately strong strengths, but the transition from subpoissonian to superpoissonian statistics is smooth, without a special role for $F = 1$.

The higher irregularity in the spike counts is always accompanied by a tendency toward more "bursty" firing. (These bursts are a network effect; the model contains only leaky integrate-and-fire neurons, which do not burst on their own.) This burstiness can best be seen in the spike train autocorrelation function (Figure 4), which acquires a hill of growing size and width around zero lag for increasing Fano factors. The interdependence between firing irregularity and bursting can be understood with help of the ISI distributions depicted in Figure 3: when the rate, and thus the average ISI, is kept constant, then any higher count for shorter-than-average ISIs must be accompanied by an accordingly higher count for longer ISIs (indicating bursts), and vice versa. Thus higher irregularity always goes hand in hand with a higher tendency toward temporal clustering of spikes.

Why do stronger synapses lead to higher irregularity in firing? The size of the input current fluctuations in (6) are controlled by the J_{ab}, and so, therefore, are the corresponding membrane potential fluctuations. Thus, for example, the width of the steady-state membrane potential distribution is proportional to J_s. We next have to consider where this distribution is centered. Remembering that, according to the balance condition, the firing rate is independent of J_s, the center of the distribution has to move farther away from threshold as J_s is increased in order to keep the rate fixed. Therefore, for very small J_s almost the entire equilibrium membrane potential distribution will lie well above the post-spike reset value, while for large J_s it will be mostly below reset.

Immediately after a spike, the membrane potential distribution is a delta-function centered at the reset (here 0). It then spreads and its mean moves up or down toward its equilibrium value. This equilibration will take about a membrane time constant. If the equilibrium value is well above zero (the

small-J_s case), the probability of reaching threshold will be suppressed during this time, implying a refractory dip in the ISI distribution and the correlation function and a tendency toward a Fano factor less than 1.

In the large-J_s case, on the other hand, where the membrane potential is reset much closer to the threshold than to its eventual equilibrium value, the initial rapid spread (with the width growing proportional to $J_s\sqrt{t}$) leads to an enhanced probability of early spikes. At short times this diffusive spread dominates the downward drift of the mean (which is only linear in t). Thus there is extra weight in the ISI distribution and a positive correlation function at these short times, leading to a Fano factor greater than 1.

Empirically, an approximate power-law relationship between the mean and variance of the spike count has frequently been observed for cortical neurons (see, e.g., [11, 13, 17, 20]). Our model shows the same qualitative feature (Figures 5 and 6), though we have no argument that the relation should be an exact power law. However, this agreement suggests that the model captures at least part of physics underlying the firing statistics.

As already observed, not all of the variability in measured neuron responses has to be explained in the manner outlined above. Changing conditions during the run of a single experiment may introduce extra irregularity, caused by collecting statistics over trials with different mean firing rates. The present analysis shows why – and how much – irregularity can be expected due to intrinsic cortical dynamics.

Our formulation of the mean-field theory is general enough to allow straightforward extensions to greater biological realism and to more complicated network architectures. We have introduced a generalization of this model with conductance-based synapses in another paper [24]. We have also extended the model to include systematic structure in the connections, modeling an orientation hypercolumn in the primary visual cortex [25]. Moreover, our algorithm for finding the mean-field solutions is not restricted to networks of integrate-and-fire neurons. It can be applied to any kind of neuronal model. Furthermore, any kind of synaptic dynamics can be incorporated by using synaptically filtered spike trains to compute the self-consistent solutions.

References

[1] D Amit and N Brunel, Model of spontaneous activity and local structured activity during delay periods in the cerebral cortex. Cereb Cortex 7, 237-252 (1997)

[2] D Amit and N Brunel, Dynamics of a recurrent network of spiking neurons before and following learning. Network 8, 373-404 (1997)

[3] N Brunel, Dynamics of sparsely connected networks of excitatory and inhibitory spiking neurons. J Comput Neurosci 8, 183-208 (2000)

[4] C van Vreeswijk and H Sompolinsky, Chaos in neuronal networks with balanced excitatory and inhibitory activity. Science 274, 1724-1726 (1996)

[5] C van Vreeswijk and H Sompolinsky, Chaotic balanced state in a model of cortical circuits. Neural Comp 10, 1321-1371 (1998)

[6] M W Oram, M C Wiener, R Lestienne and B J Richmond, Stochastic nature of precisely-timed spike patterns in visual system neural responses. J Neurophysiol 81, 3021-3033 (1999)

[7] C Fulvi Mari, Random networks of spiking neurons: instability in the *xenopus* tadpole moto-neuron pattern. Phys Rev Lett 85, 210-213 (2000)

[8] R Kree and A Zippelius, Continuous-time dynamics of asymmetrically diluted neural networks. Phys Rev A 36, 4421-4427 (1987)

[9] P Heggelund and K Albus, Response variability and orientation discrimination of single cells in in striate cortex of cat. Exp Brain Res 32, 197-211 (1978)

[10] A F Dean, The variability of discharge of simple cells in the cat striate cortex. Exp Brain Res 44, 437-440 (1981)

[11] D J Tolhurst, J A Movshon and I D Thompson, The dependence of response amplitude and variance of cat visual cortical neurones on stimulus contrast. Exp Brain Res 41, 414-419 (1981)

[12] D J Tolhurst, J A Movshon and A F Dean, The statistical reliability of signals in single neurons in cat and monkey visual cortex. Vision Res 23, 775-785 (1983)

[13] R Vogels, W Spileers and G A Orban, The response variability of striate cortical neurons in the behaving monkey. Exp Brain Res 77, 432-436 (1989)

[14] R J Snowden, S Treue and R A Andersen, The response of neurons in areas V1 and MT of the alert rhesus monkey to moving random dot patterns. Exp Brain Res 88, 389-400 (1992)

[15] M Gur, A Beylin and D M Snodderly, Response variability of neurons in primary visual cortex (V1) of alert monkeys. J Neurosci 17, 2914-2920 (1997)

[16] M N Shadlen and W T Newsome, The variable discharge of cortical neurons: implications for connectivity, computation, and information coding. J Neurosci 18, 3870-3896 (1998)

[17] E Gershon, M C Wiener, P E Latham and B J Richmond, Coding strategies in monkey V1 and inferior temporal cortex. J Neurophysiol 79, 1135-1144 (1998)

[18] P Kara, P Reinagel and R C Reid, Low response variability in simultaneously recorded retinal, thalamic, and cortical neurons. Neuron 27,635-646 (2000)

[19] G T Buracas, A M Zador, M R DeWeese and T D Albright, Efficient discrimination of temporal patterns by motion-sensitive neurons in primate visual cortex. Neuron 20, 959-969 (1998)

[20] D Lee, N L Port, W Kruse and A P Georgopoulos, Variability and correlated noise in the discharge of neurons in motor and parietal areas of primate cortex. J Neurosci 18, 1161-1170 (1998)

[21] M R DeWeese, M Wehr and A M Zador, Binary spiking in auditory cortex. J Neurosci 23, 7940-7949 (2003)

[22] J Hertz, B Richmond and K Nilsen, Anomalous response variability in a balanced cortical network model. Neurocomputing 52–54, 787–792 (2003)

[23] H Eisfeller and M Opper, New method for studying the dynamics of disordered spin systems without finite-size effects. Phys Rev Lett 68, 2094-2097 (1992)

[24] A Lerchner, M Ahmadi and J Hertz, High conductance states in a mean field cortical network model. CNS 2003, to be published in Neurocomputing (2004)

[25] J Hertz and G Sterner, Mean field model of an orientation hypercolumn. Program No. 911.19 Abstract Society for Neuroscience (2003)

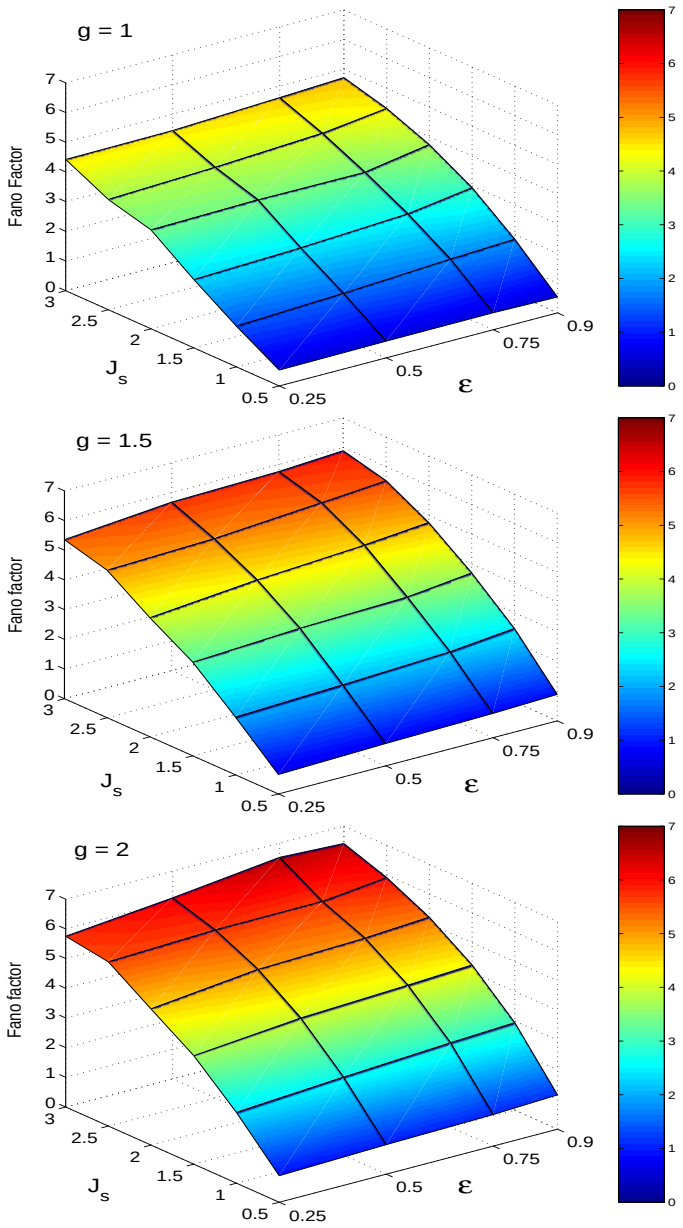

Figure 2: Fano factors as a function of overall synaptic strength J_s and intracortical excitation strength ϵ for three different inhibition factors: $g = 1$, 1.5, and 2, respectively. The increase of any of these parameters results in more irregular firing statistics as measured by the Fano factor.

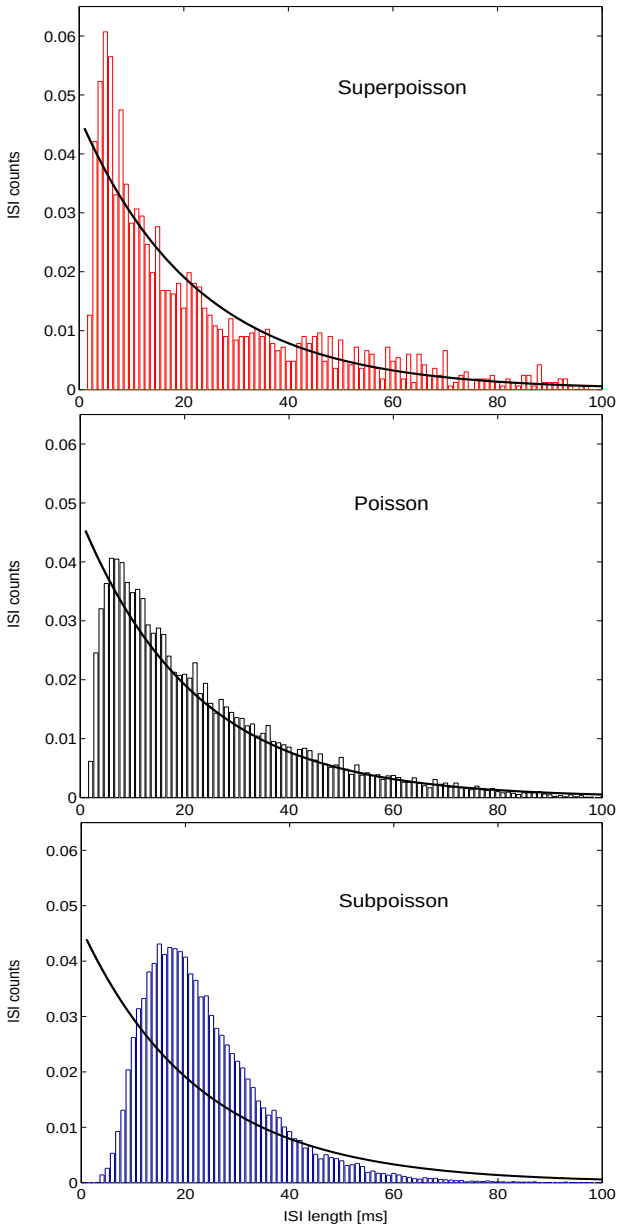

Figure 3: Interspike interval distributions for fixed $\epsilon = 0.5$ and $g = 1$, and three different values of overall synaptic strength J_s: 1.5 (superpoissonian), 0.75 (Poissonian), and 0.375 (subpoissonian). Overlayed on each figure is the exponential fall-off of a true Poisson distribution with the same average rate as in all of the three cases.

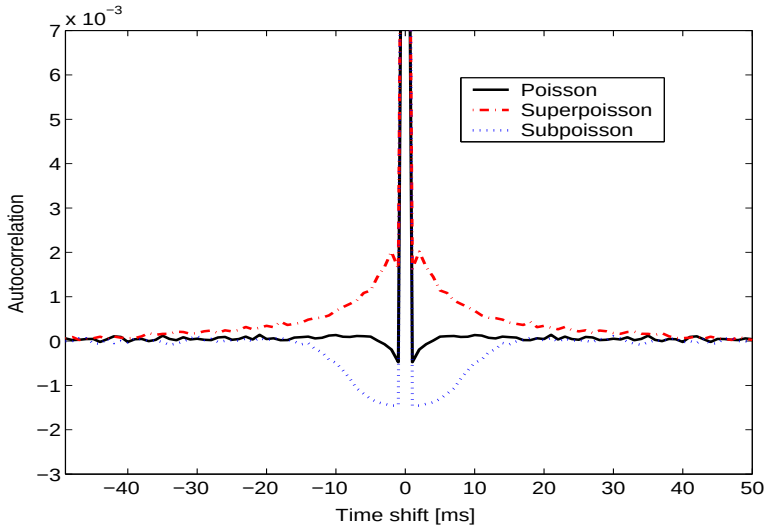

Figure 4: Three different spike train autocorrelations illustrating the relationship between the Fano factor F and the area under the curve. For $F = 1$ (Poissonian, black solid line), the autocorrelation is an almost perfect delta function. $F > 1$ (superpoissonian, red dash-dot line) is reflected by a hill generating a positive area, and $F < 1$ (subpoissonian, blue dotted line) is accompanied by a valley of negative correlations. (See the text for more details.)

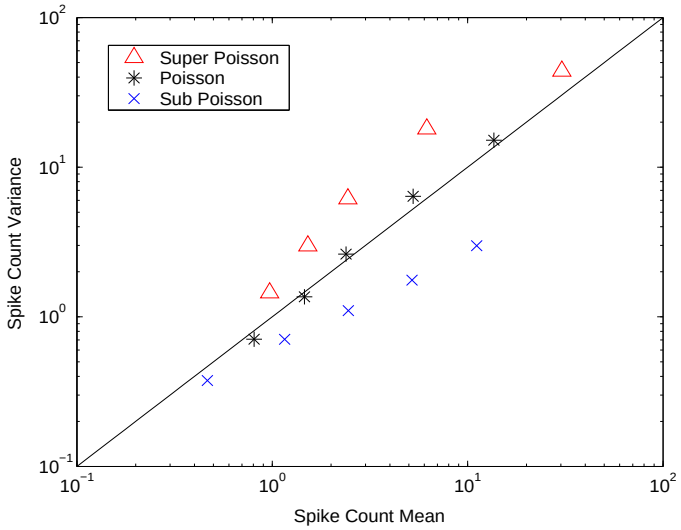

Figure 5: Spike count log(variance) vs. log(mean) for three different values of overall synaptic strength J_s, varying the external input rate r_0. For $J_s = 1.25$ (superpoissonian, red triangles) the data look qualitatively like those from experiments. The other values for J_s are 0.75 (Poisson, black stars) and 0.375 (subpoissonian, blue crosses).

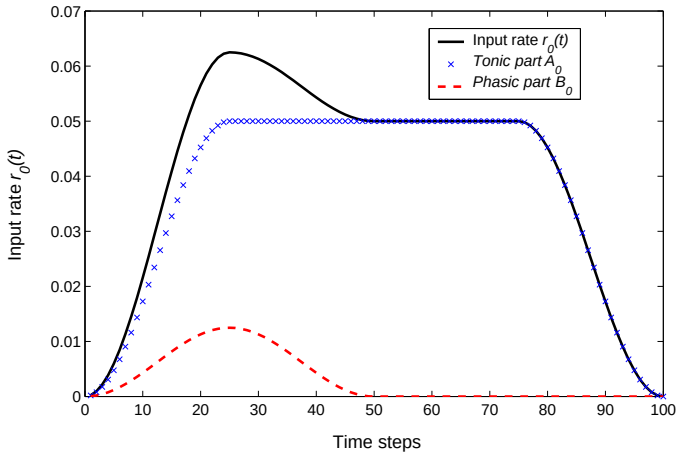

Figure 6: Parametrization of the time-dependent input rate $r_0(t)$. The input is modeled as the sum of three functions: (1) a stationary background rate (which is zero in this case); (2) a tonic part, which rises within the first 20 ms to a constant level of A_0 where it stays for 60 ms, falling back to zero within the last 20 ms; and (3) an initial phasic part, which is nonzero only in the first 50 ms, rising to a maximum value of B_0.

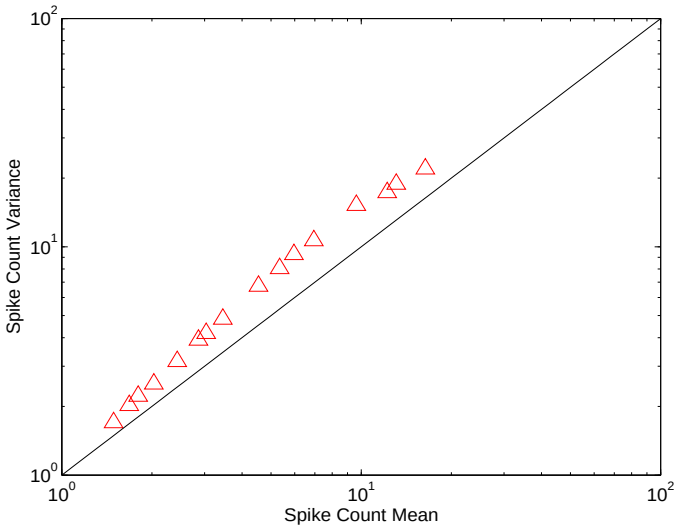

Figure 7: Spike count log(variance) vs. log(mean) for time-varying external inputs with varying overall strength. The neuron in the simulated network (red triangles) fires in a superpoissonian regime, with an almost linear relationship for low spike rates between the log variance and the log mean, resembling closely data obtained from in-vivo experiments. The diagonal solid line indicates the identity of variance and mean (Fano factor $F = 1$).

arXiv:q-bio.NC/0402026 v1 11 Feb 2004

High Conductance States in a Mean Field Cortical Network Model

Alexander Lerchner Mandana Ahmadi John Hertz

Nordita, Blegdamsvej 17, 2100 Copenhagen, Denmark

Abstract

Measured responses from visual cortical neurons show that spike times tend to be correlated rather than exactly Poisson distributed. Fano factors vary and are usually greater than 1 due to the tendency of spikes being clustered into bursts. We show that this behavior emerges naturally in a balanced cortical network model with random connectivity and conductance-based synapses. We employ mean field theory with correctly colored noise to describe temporal correlations in the neuronal activity. Our results illuminate the connection between two independent experimental findings: high conductance states of cortical neurons in their natural environment, and variable non-Poissonian spike statistics with Fano factors greater than 1.

Key words: synaptic conductances, response variability, cortical dynamics

1 Introduction

Neurons in primary visual cortex show a large increase in input conductance during visual activation: *in vivo* recordings (see, e.g., [1]) show that the conductance can rise to more than three times that of the resting state. Such *high conductance states* lead to faster neuronal dynamics than would be expected from the value of the passive membrane time constant, as pointed out by Shelley et al. [2]. We use mean field theory to study the firing statistics of a model network with balanced excitation and inhibition and observe consistently such high conductance states during stimulation.

In our study, we classify the irregularity of firing with the Fano factor F, defined as the ratio of the variance of the spike count to its mean. For tem-

Email addresses: lerchner@nordita.dk (Alexander Lerchner), ahmadi@nordita.dk (Mandana Ahmadi), hertz@nordita.dk (John Hertz).

porally uncorrelated spike trains (i.e., Poisson processes) $F = 1$, while $F > 1$ indicates a tendency for spike clustering (bursts), and $F < 1$ points to more regular firing with well separated spikes. Observed Fano factors for spike trains of primary cortical neurons during stimulation are usually greater than 1 and vary within an entire order of magnitude (see, e.g., [3]). We find the same dynamics in our model and are able to pin-point the relevant mechanisms: synaptic filtering leads to spike clustering in states of high conductance (thus $F > 1$), and Fano factors depend sensitively on variations in both threshold and synaptic time constants.

2 The Model

We investigate a cortical network model that exhibits self-consistently balanced excitation and inhibition. The model consists of two populations of neurons, an excitatory and an inhibitory one, with dilute random connectivity. The model neurons are governed by leaky integrate-and-fire subthreshold dynamics with conductance-based synapses. The membrane potential of neuron i in population a ($a = 1, 2$ for excitatory and inhibitory, respectively) obeys

$$\frac{du_a^i(t)}{dt} = -g_L u_a^i(t) - \sum_{b=0}^{2} \sum_{j=1}^{N_b} g_{ab}^{ij}(t)(u_a^i(t) - V_b). \tag{1}$$

The first sum runs over all populations b, including the excitatory input population representing input from the LGN and indexed by 0. The second sum runs over all neurons j in population b of size N_b. The reversal potential V_b for the excitatory inputs ($b = 0, 1$) is higher than the firing threshold, the one for the inhibitory inputs (V_2) is below the reset value. The constant leakage conductance g_L is the inverse of the membrane time constant τ_m.

The time dependent conductance $g_{ab}^{ij}(t)$ from neuron j in population b to neuron i in population a is taken as

$$g_{ab}^{ij}(t) = \frac{g_{ab}^0}{\sqrt{K_b}} \sum_s \exp(-(t - t_s^j)/\tau_b)\Theta(t - t_s^j) \tag{2}$$

if there is a connection between those two neurons, otherwise zero. The sum runs over all spikes s emitted by neuron j, τ_b is the synaptic time constant for the synapse of type b (excitatory or inhibitory), and Θ is the Heavyside step function. K_b denotes the average number of presynaptic neurons in population b. We followed van Vreeswijk and Sompolinsky [4] by scaling the conductances

with $1/\sqrt{K_b}$ so that their fluctuations are of order one, independent of network size.

3 Mean Field Theory

We use mean field theory to reduce the full network problem to two neurons: one for each population. This method is exact in the limit of large populations with homogeneous connection probabilities [5]. The neurons receive self-consistent inputs from their cortical environment, exploiting the fact that all neurons within a population exhibit the same firing statistics due to homogeneity. The time dependent conductance described in (2) can then be replaced by a realization of a Gaussian distributed random variable g_{ab} with mean

$$\langle g_{ab} \rangle = g_{ab}^0 \sqrt{K_b}\, r_b, \tag{3}$$

and covariance

$$\langle \delta g_{ab}(t)\, \delta g_{ab}(t') \rangle = (g_{ab}^0)^2 (1 - K_b/N_b)\, C_b(t - t'), \tag{4}$$

Here, r_b is the firing rate of the presynaptic neuron b, and $C_b(t - t')$ is the autocorrelation function of its spike train. A simple approximation of the autocorrelation, like the one used by [6] and [7], is to assume $g_{ab}(t)$ to be temporally uncorrelated (i.e., white noise), in which case it simplifies to $C_b(t - t') = r_b\, \delta(t - t')$. The term $(1 - K_b/N_b)$ is a correction for the finite connection concentration K_b/N_b and can be derived using the methods of [8].

The self-consistent balance condition is obtained by setting the net current in (1) to zero when the membrane potential is at threshold θ_a and the conductances have their mean values (3). In the large K_b-limit, it reads

$$\sum_{b=0}^{2} g_{ab}^0 \sqrt{K_b}\, r_b\, (\theta_a - V_b) = 0. \tag{5}$$

The distribution of the variables g_{ab} can be calculated numerically using an iterative approach [9]. One starts with a guess based on the balance equation (5) for the means and covariances and generates a large sample of specific realizations of $g_{ab}(t)$, which are used to integrate (1) to generate a large sample of spike trains. The latter can then be used to calculate new estimates of the means and covariances by applying (3) and (4) and correction of the initial guess towards the new values. These steps are repeated until convergence.

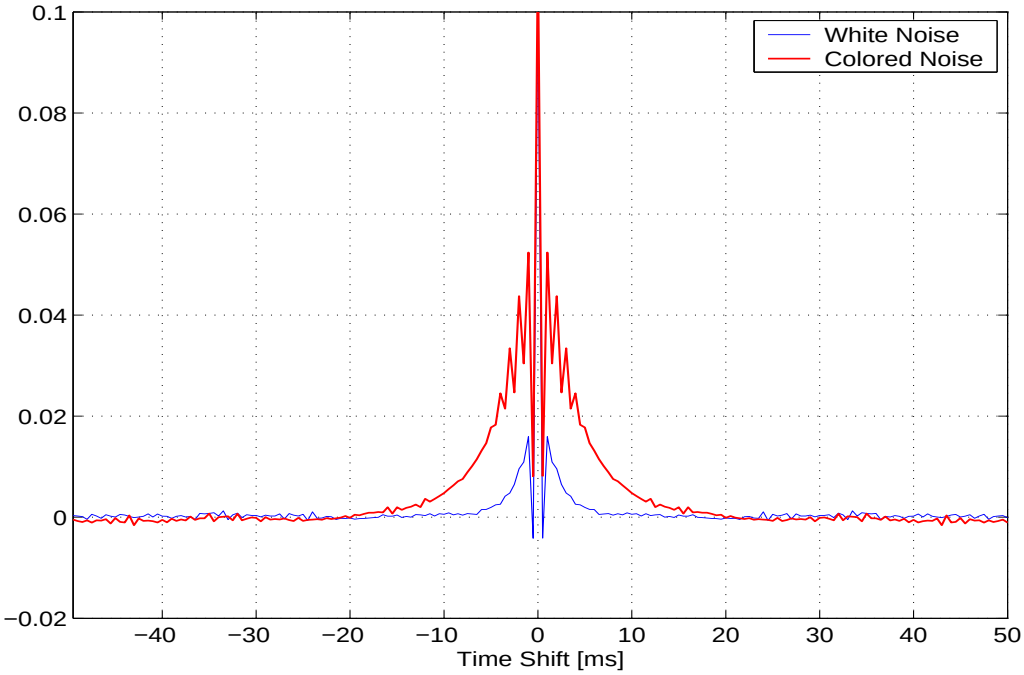

Fig. 1. Autocorrelation functions for white noise (blue) and colored noise (red). The white noise approximation underestimates the amount of temporal correlation in the neuron's firing.

4 Results

For the above described model, we chose parameters corresponding to population sizes of 16,000 excitatory neurons and 4,000 inhibitory neurons, representing a small patch of layer IV cat visual cortex. The neurons were connected randomly, with 10% connection probability between any two neurons. The firing threshold was fixed to 1, excitatory and inhibitory reversal potentials were set to $+14/3$ and $-2/3$, respectively, and the membrane time constant $\tau_m = g_L^{-1}$ was 10 ms. For the results presented here, the integration time step was 0.5 ms.

Figure 1 illustrates the importance of coloring the noise produced by intracortical activity. The white noise approximation underestimates both the correlation times and the strength of the correlations in the neuron's firing: its autocorrelation (blue) is both narrower and weaker than the one for colored noise (red).

Fano factors vary systematically with both the distance between reset and threshold and the synaptic time constant τ_s. Non-zero synaptic time constants produced consistently Fano factors greater than one. We varied the reset between 0.8 and 0.94 and τ_s between 0 and 6 ms, which resulted in values for F that span an entire order of magnitude, from slightly above 1 to approximately

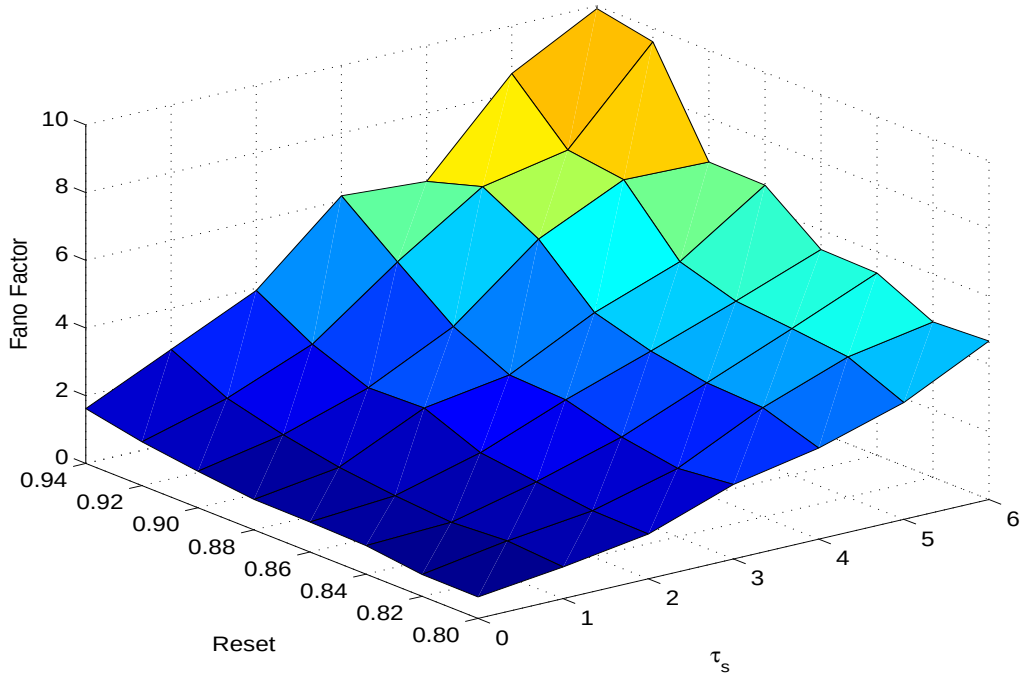

Fig. 2. Fano factors for a range of reset values and synaptic time constants τ_s. Longer synaptic time constants lead to increased clustering (bursts) of spikes, which is reflected in higher Fano factors.

10 for $\tau_s \geq 2$ ms (see Figure 2).

5 Discussion

In all our simulations, we observed that the membrane potential changed on a considerably faster time scale than the membrane time constant $\tau_m = 10$ ms. This behavior is only observed if conductance-based synapses are included in the integrate-and-fire neuron model. To understand this phenomenon, it is convenient to follow the notation of Shelley et al. [2] to rewrite the equation for the membrane potential dynamics (1) in the following form:

$$\frac{du_a(t)}{dt} = -g_T(t)\,(u_a(t) - V_S(t))\,, \tag{6}$$

with the *total conductance* $g_T(t) = g_L + \sum_b g_{ab}(t)$, and the *effective reversal potential* $V_S(t) = g_T(t)^{-1} \sum_b g_{ab}(t) V_b$. The membrane potential $u_a(t)$ follows the effective reversal potential with the input dependent *effective membrane time constant* $g_T(t)^{-1}$. The effective reversal potential changes on the time scale of the synaptic time constants, which are up to five times shorter than τ_m in our simulations. However, if the effective membrane time constant is shorter

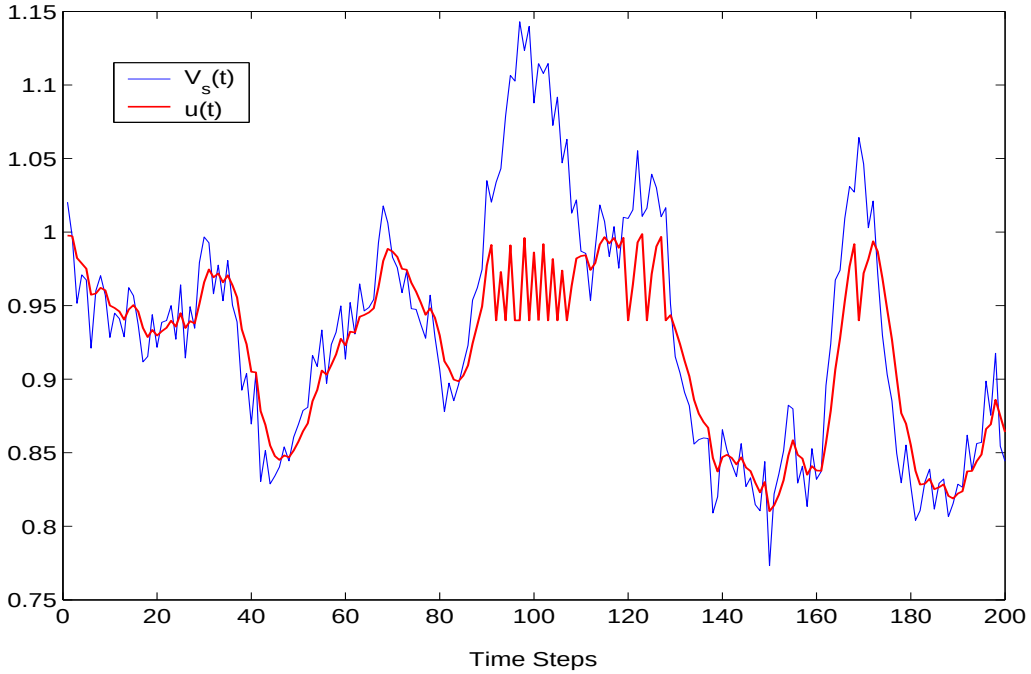

Fig. 3. The membrane potential $u(t)$ (red) follows the effective reversal potential $V_S(t)$ (blue) closely, except for detours when the neuron is reset due to firing. The membrane potential recovers fast enough to spike several times while $V_S(t)$ stays above threshold, thus producing bursts of spikes. Here, the threshold is set to 1 and the reset to 0.94.

than the synaptic time constant due to a large enough total conductance, then $u_a(t)$ can follow $V_S(t)$ closely, as observed in our simulations (see Figure 3).

In high conductance states, the firing statistics are strongly influenced by synaptic dynamics (see Figure 2). This is in contrast with strictly current based models, where the neuron reacts too slow to reflect fast synaptic dynamics in its firing. The 'synaptic filtering' of arriving spikes leads to temporal correlations in $V_S(t)$ and thus to temporal correlations (by way of spike clustering) in firing. Therefore, the model neurons receive temporally correlated input rather than white noise. For this reason, in mean field models dealing with conductance based dynamics, coloring the noise is important to arrive at the full amount of temporal correlation in firing statistics (see Figure 1). We confirmed these considerations by running simulations without synaptic filtering ($\tau_s = 0$). As expected, intra-cortical activity became uncorrelated and the white noise approximation produced the same result as coloring the noise correctly. In that case, Fano factors stayed close to 1 (see Figure 2), i.e, no tendency of spike clustering was observed.

Previous investigations showed that varying the distance between threshold and reset in balanced integrate-and-fire networks has a strong effect on the irregularity of the firing [10]. By including a conductance-based description of

synapses, we were now able to show the importance of synaptic time constants on firing statistics, even if they are several times smaller than the passive membrane time constant: Synaptic filtering facilitates clustering of spikes in states of high conductance.

References

[1] L J Borg-Graham, C Monier, and Y Fregnac, *Nature* **393**:369-374 (1998).

[2] M Shelley, D McLaughlin, R Shapley, and J Wielaard, *J. Comput. Neurosci*, **13**:93-109 (2002).

[3] E D Gershon, M C Wiener, P E Latham and B J Richmond, *J Neurophysiol* **79**:1135-1144 (1998).

[4] C van Vreeswijk and H Sompolinsky, *Science* **274**:1724-1726 (1996), *Neural Comp* **10**:1321-1371 (1998).

[5] C Fulvi Mari, *Phys Rev Lett* **85**:210-213 (2000).

[6] D J Amit and N Brunel, *Cerebral Cortex* **7**:237-252 (1997).

[7] N Brunel, *J Comput Neurosci* **8**:183-208 (2000).

[8] R Kree and A Zippelius, *Phys Rev A* **36**:4421-4427 (1987).

[9] H Eisfeller and M Opper, *Phys Rev Lett* **68**:2094-2097 (1992)

[10] J Hertz, B J Richmond, and K Nilsen, *CNS 2002*, to be published in *Neurocomputing* (2003).

arXiv:q-bio.NC/0403037 v1 25 Mar 2004

Mean field theory for a balanced hypercolumn model of orientation selectivity in primary visual cortex

A Lerchner†, G Sterner‡, J Hertz§ and M Ahmadi§

† Ørsted-DTU, Technical University of Denmark, 2800 Kgs. Lyngby, Denmark
‡ Department of Physics, University of Rochester, Rochester NY 14627
§ Nordita, Blegdamsvej 17, 2100 Copenhagen Ø, Denmark

E-mail: al@oersted.dtu.dk

Abstract. We present a complete mean field theory for a balanced state of a simple model of an orientation hypercolumn. The theory is complemented by a description of a numerical procedure for solving the mean-field equations quantitatively. With our treatment, we can determine self-consistently both the firing rates and the firing correlations, without being restricted to specific neuron models. Here, we solve the analytically derived mean-field equations numerically for integrate-and-fire neurons. Several known key properties of orientation selective cortical neurons emerge naturally from the description: Irregular firing with statistics close to – but not restricted to – Poisson statistics; an almost linear gain function (firing frequency as a function of stimulus contrast) of the neurons within the network; and a contrast-invariant tuning width of the neuronal firing. We find that the irregularity in firing depends sensitively on synaptic strengths. If Fano factors are bigger than 1, then they are so for all stimulus orientations that elicit firing. We also find that the tuning of the noise in the input current is the same as the tuning of the external input, while that for the mean input current depends on both the external input and the intracortical connectivity.

Submitted to: *Network: Computation in Neural Systems*

1. Introduction

Neurons in primary visual cortex (V1) fire highly irregularly in response to visual stimuli, but with reproducible firing rates. They do so despite the fact that they receive synaptic input from thousands of other cortical neurons, which would lead to fluctuations in the input that were small compared to the mean if excitatory and inhibitory inputs were not balanced [1]. There has been some success in describing how such a balance can emerge self-consistently from dynamics that are plausible for cortical networks. This was accomplished by mean field-descriptions by van Vreeswijk and Sompolinsky [2, 3] and Amit and Brunel [4, 5, 6]. However, their treatments do not permit a self-consistent calculation of firing correlations. How to do this correctly was first shown

for an all-inhibitory network by Hertz *et al.* [7] using the systematic formulation of mean field theory due to Fulvi Mari [8]. In a recent paper [9] we presented a mean-field theory for a balanced network model that allowed us to quantify how the irregularity in firing and, more generally, the firing correlations depend on intrinsic network properties such as synaptic strengths. The analysis was applied to a statistically homogeneous network, representing a cortical column composed of neurons with similar response characteristics. Here, we show how to extend this treatment to networks with systematic structure, consisting of multiple cortical columns. In particular, we model an orientation hypercolumn, composed of a set of orientation columns.

An orientation column contains neurons that respond strongest to elongated visual stimuli of a specific orientation, the *preferred orientation* (PO). Orientation selective neurons exhibit a tuned response to other orientations, with sharply decreasing firing rates as the similarity between PO and stimulus orientation decreases, until the firing is completely suppressed for orientations outside the *tuning width* of the neuron in question. An important feature of orientation tuning is that the tuning width is independent of the stimulus contrast [10]. It is not possible to capture this feature in a single-neuron description using a Hubel and Wiesel feed-forward connectivity [11] from the lateral geniculate nucleus (LGN); rather, cortical interactions are needed to achieve contrast-invariant tuning (for review see [12]). Ben-Yishai *et al* [13] proposed a model for which the tuning width is independent of the contrast, but a threshold-linear relationship between input current and firing rate was an assumption of the model, and the problem of the firing statistics was not addressed.

Here, we show how a contrast-invariant tuning width, an almost linear input-output relationship, and irregular firing can all be explained by a balanced hypercolumn model. With our mean-field treatment, we can quantify how certain network properties like synaptic strengths, tuning of the LGN input and of the intracortical connectivity influence the statistics and tuning of the neuronal firing. Using the Fano factor F (the ratio of spike count variance and mean spike count) to quantify the irregularity in firing, we find, e.g., that if F is significantly greater than 1 the orientation tuning of F reaches a maximum at the PO (Fano factors greater than 1 are normally observed for neurons in V1 [14]). We also make quantitative predictions about the tuning of the input currents and their fluctuations.

2. Model and Methods

We model a single orientation hypercolumn in primary visual cortex, with a simplified network architecture as indicated in Figure 1. The network comprises an excitatory population and an inhibitory one, of sizes N_1 and N_2, respectively. Each population is divided into n sub-populations (orientation columns), parameterized by an angle θ. The angles, spaced equally between $-\pi/2$ and $\pi/2$, indicate the preferred orientation (PO), to which the neurons in the corresponding column respond strongest.

We use leaky integrate-and-fire neurons and interconnect them randomly with a

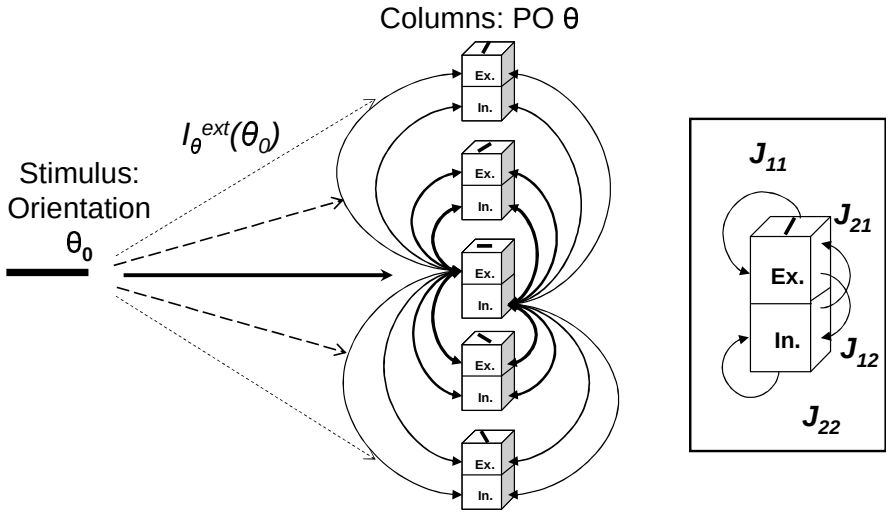

Figure 1. Structure of the model network. The hypercolumn consists of multiple orientation columns, each of which has an excitatory and an inhibitory subpopulation and is assigned a preferred orientation (PO) θ. Columns with more similar POs share on average more connections than more dissimilar ones (the density of connections is indicated only between one column and the rest, for clarity). The network receives excitatory external input, weakly tuned to the stimulus orientation θ_0. The inset shows a sketch for the connectivity and connection strengths J_{ab} within an orientation column.

connection probability $P_{ab}(\theta - \theta')$ that depends on the similarity of the POs. The probability that a neuron with PO θ (in population a) receives afferent input from a neuron with PO θ' in population b is taken as

$$P_{ab}(\theta - \theta') = \frac{K_b}{N_b}\left(1 + \gamma \cos 2(\theta - \theta')\right), \tag{1}$$

where K_b is the expected overall number of inputs from neurons in population b. We take the ratio K_b/N_b independent of b, i.e., excitatory and inhibitory neurons interconnect with the same probability in our model. The functional form of (1) is motivated by anatomical evidence that the connection probability between cortical neurons decreases as their distance increases, and by the fact that orientation columns with similar PO tend to lie closer together on the cortical surface than ones with dissimilar PO. We followed Ben-Yishai *et al.* [13] in choosing the simplest possible form that is periodic with period π. We assume that the degree of tuning, as measured by the parameter $\gamma \in (0, 1)$, is the same for both the inhibitory and the excitatory population.

Each nonzero synapse from a neuron in population b to one in population a is taken to have strength

$$J_{ij}^{a\theta,b\theta'} = \frac{J_{ab}}{\sqrt{K_b}} \tag{2}$$

where the parameters J_{ab} are of order 1. With this scaling, the fluctuations in the input current are also of order 1, the same order as the distance between reset and threshold of our model neurons (cf. van Vreeswijk and Sompolinsky [2, 3]).

The subthreshold dynamics of the membrane potentials are given by

$$\frac{\mathrm{d}u_i^{a\theta}(t)}{\mathrm{d}t} = -\frac{u_i^{a\theta}(t)}{\tau} + I_{a\theta}^{\mathrm{ext}}(\theta_0) + I_i^{a\theta,\mathrm{rec}}(t), \tag{3}$$

where the membrane time constant τ is chosen to be the same for all neurons. The excitatory external input from the LGN, $I_{a\theta}^{\mathrm{ext}}(\theta_0)$, is assumed to be (weakly) tuned to the orientation θ_0 of the stimulus due to a feed-forward connectivity from the LGN as in the classical model by Hubel and Wiesel [11]. For simplicity, we take it to be constant in time and the same for all neurons i within a column. The functional form we use is, similar to the tuning (1) of the intracortical connectivity,

$$I_{a\theta}^{\mathrm{ext}}(\theta_0) = I_a^{\mathrm{ext}}(1 + \epsilon\cos 2(\theta - \theta_0)), \tag{4}$$

where $\epsilon \in (0, 1)$ is the degree of tuning, which is assumed to be the same for both populations. (The condition $\epsilon < 1$ assures $I_{a\theta}^{\mathrm{ext}}(\theta_0)$ to be non-negative, i.e. excitatory, for all orientations). A more detailed model for this external input current, including temporal fluctuations and random connectivity, was briefly described in an overview article by Hertz *et al.* [15].

The recurrent input $I_i^{a\theta,\mathrm{rec}}(t)$ from within the model cortex is given by

$$I_i^{a\theta,\mathrm{rec}}(t) = \sum_{b=1}^{2}\sum_{\theta'=\theta_1}^{\theta_n}\sum_{j=1}^{N_b/n} J_{ij}^{a\theta,b\theta'} S_j^{b\theta'}(t), \tag{5}$$

where $S_j^{b\theta'}(t) = \sum_s \delta(t - t_{j\theta' b}^s)$ is the spike train of neuron j with PO θ' in population b.

Mean Field Theory

In the following mean-field description of the orientation hypercolumn model, we consider stationary firing only, for simplicity. However, the formulation is general enough to allow for non-stationary rates. We presented such a time-dependent treatment for a balanced single-column model elsewhere [9].

Because of the dilute random connectivity, each neuron receives a high number of uncorrelated inputs (we assume K_b to be large, but smaller than N_b). According to the central limit theorem, the recurrent input currents given by (5) can therefore be described as Gaussian random processes. For stationary rates, the mean input current is constant in time for any given neuron, although the level of the mean does vary from neuron to neuron due to the random connectivity. In a general mean-field theory, one must consider temporal correlations in these currents, i.e., not restrict the description of the random processes to white noise.

To separate the mean of the currents from their fluctuations ("noise"), it is convenient to apply such separations to the description of both the synaptic weights $J_{ij}^{a\theta,b\theta'}$ and the spike trains $S_j^{b\theta'}(t)$ in (5). For the weights we can write

$$J_{ij}^{a\theta,b\theta'} = \overline{J_{ij}^{a\theta,b\theta'}} + \delta J_{ij}^{a\theta,b\theta'}, \tag{6}$$

where the bar means averaging over the index j, i.e., the neurons in the source population:

$$\overline{J_{ij}^{a\theta,b\theta'}} = \frac{1}{N_b/n} \sum_{j=1}^{N_b/n} J_{ij}^{a\theta,b\theta'} \tag{7}$$

Generally, we use the bar-notation for averaging over neuron populations, which will always apply to the running index j in this work. To separate the spike trains into static and dynamic components, we write

$$S_j^{b\theta'}(t) = r_b(\theta') + \delta r_j^{b\theta'} + \delta S_j^{b\theta'}(t), \tag{8}$$

where $r_b(\theta') = \overline{r_j^{b\theta'}} = 1/(N_b/n)\sum_j r_j^{b\theta'}$ is the average rate of the neurons in subpopulation θ' of population b. The difference between this average rate and the actual rate of neuron j is denoted $\delta r_j^{b\theta'}$. These two components are both static, describing time-averaged quantities. The temporal fluctuations of the spike train and their possible correlations in time are captured by the third term on the right-hand side of (8), $\delta S_j^{b\theta'}(t)$. Using the central limit theorem and methods like those in [8] and [16] we can then derive the following mean-field formulation of the recurrent current:

$$I_{a\theta}^{\mathrm{rec}}(t) = \sum_{b=1}^{2} J_{ab} \left(\sqrt{K_b} A_b + \sqrt{1 - K_b/N_b} B_b(t) \right), \tag{9}$$

with

$$A_b = \frac{1}{n} \sum_{\theta'=\theta_1}^{\theta_n} (1 + \gamma \cos 2(\theta - \theta')) r_b(\theta') \tag{10}$$

$$B_b(t) = \frac{1}{n} \sum_{\theta'=\theta_1}^{\theta_n} \sqrt{1 + \gamma \cos 2(\theta - \theta')} \left(\left(\overline{(r_j^{b\theta'})^2} \right)^{\frac{1}{2}} x_{b\theta'} + \xi_{b\theta'}(t) \right) \tag{11}$$

where the values $x_{b\theta'}$ are drawn from a unit-variance normal distribution. Selecting specific values $x_{b\theta'}$ effectively samples different neurons within the column population. We have dropped the neuron index i because this statistical description of the input current reduces the network problem to single neuron problems – one for each column population, indexed by $a\theta$. The terms $\xi_{b\theta'}(t)$ stand for realizations of Gaussian random processes obeying

$$\langle \xi_{b\theta'}(t) \xi_{b\theta'}(t') \rangle = C_{b\theta'}(t - t'). \tag{12}$$

Here, $C_{b\theta'}(t - t')$ denotes the average autocorrelation function of the fluctuations in the spike trains of neurons with PO θ' in population b, given by

$$C_{b\theta'}(t - t') = \frac{1}{N_b/n} \sum_{j=1}^{N_b/n} \langle \delta S_j^{b\theta'}(t) \delta S_j^{b\theta'}(t') \rangle. \tag{13}$$

With the operation $\langle \cdot \rangle$ we mean averaging over realizations of random processes, such as stochastic spike trains. We will refer to such realizations as "trials" since they represent (responses to) repeated presentations of the same stimulus in experimental settings.

The balance condition

The input currents from the excitatory population and the inhibitory population have mean values of order $\sqrt{K_1} \gg 1$ and $\sqrt{K_2} \gg 1$, respectively (see Equation (9)). In addition, for the external input current (4) we take $I_a^{\rm ext} = \sqrt{K_0}\hat{I}_a^{\rm ext}$ with $\sqrt{K_0} \gg 1$. If the neurons are to exhibit irregular firing at a low rate, as cortical neurons do, these currents must nearly cancel and threshold crossings have to be caused by the fluctuations in the currents, which are of order 1. For our orientation hypercolumn model, this balance condition implies that the average input currents in (3) have to add up to zero for each orientation column θ:

$$\sqrt{K_0}\hat{I}_a^{\rm ext}(1 + \epsilon\cos 2(\theta - \theta_0)) + \sum_{b=1}^{2} J_{ab}\sqrt{K_b}A_b = \mathcal{O}(1), \tag{14}$$

where A_b is defined in (10). Here, we have ignored the contribution of the leakage current (the first term on the right-hand side of (3)), because it is small compared to the input currents, and because the balance condition (14) holds only up to corrections of $\mathcal{O}(1)$.

To solve these equations, we consider a continuum formulation for the weighted average over all angles instead of the discrete formulation in (10) and write

$$A_b = \int_{-\pi/2}^{\pi/2} \frac{{\rm d}\theta'}{\pi}(1 + \gamma\cos 2(\theta - \theta'))r_b(\theta'). \tag{15}$$

Then (14) becomes a pair of integral equations for $r_a(\theta)$.

In the *broadly tuned case* (all orientation columns respond with non-vanishing mean rates to every stimulus orientation), these integral equations can be solved directly. To do so, we perform a Fourier expansion centered at θ_0 of the mean rate within orientation column θ' and write $r_b(\theta') = r_{b,0} + r_{b,2}\cos 2(\theta' - \theta_0) + \cdots$. For both the input current and the connection probabilities, we have already used such Fourier notations with the fewest possible terms to retain a periodic function with period π. Due to that choice, all higher Fourier components for the mean currents vanish as well, and we get

$$\sqrt{K_0}\hat{I}_a^{\rm ext}(1 + \epsilon\cos 2(\theta - \theta_0)) + \sum_{b=1}^{2}\sqrt{K_b}J_{ab}[r_{b,0} + \frac{1}{2}\gamma r_{b,2}\cos 2(\theta - \theta_0)] = 0. \tag{16}$$

By solving for each of the two Fourier components of the mean rates separately, we obtain

$$r_{a,0} = -\sum_{b=1}^{2}(\hat{\mathsf{J}}^{-1})_{ab}\hat{I}_b^{\rm ext} \tag{17}$$

$$r_{a,2} = -\frac{2\epsilon}{\gamma}\sum_{b=1}^{2}(\hat{\mathsf{J}}^{-1})_{ab}\hat{I}_b^{\rm ext} = \frac{2\epsilon}{\gamma}r_{a,0}, \tag{18}$$

where the matrix $\hat{\mathsf{J}}$ is composed of the elements $\hat{J}_{ab} = J_{ab}\sqrt{K_b/K_0}$. Firing rates have to be non-negative, so this solution can only be valid for $\epsilon \in (0, \gamma/2]$. However, such a broad tuning is not normally observed for cortical neurons. Rather, orientation sensitive neurons tend to be more "narrowly tuned", with firing suppressed for stimulus

orientations θ_0 that differ too much from the neuron's preferred orientation θ: $r_a = 0$ for $|\theta - \theta_0| \geq \theta_c$ for some tuning width θ_c. Within the parameter regime $\epsilon \in (\gamma/2, \gamma]$ we find such narrowly tuned solutions to our model. The tuning width θ_c turns out to be the same for both excitatory and inhibitory neurons, which is a consequence of the population-independence of the tuning parameters ϵ and γ.

To find the solutions for the *narrowly tuned case*, we use our insight from the broadly tuned case and make the *ansatz*

$$r_b(\theta') = \begin{cases} r_{b,0} + r_{b,2}\cos 2(\theta' - \theta_0) & \text{for } |\theta' - \theta_0| < \theta_c^b \\ 0 & \text{for } |\theta' - \theta_0| \geq \theta_c^b, \end{cases} \tag{19}$$

where $\theta_c^b = -1/2\cos^{-1}(r_{b,0}/r_{b,2})$. As mentioned above, since we have assumed equal tuning in (1), θ_c^b is the same for both b. Thus, in (15) the integration is restricted to $|\theta' - \theta_0| < \theta_c$. Because $r_b(\theta') = 0$ at $\theta' - \theta_0 = \theta_c$, we can rewrite the part of the *ansatz* for $|\theta' - \theta_0| < \theta_c$ in the form

$$r_b(\theta') = r_{b,2}(\cos 2(\theta' - \theta_0) - \cos 2\theta_c). \tag{20}$$

With this approach, we can indeed find solutions for the tuning width and the rates from the balance condition (14). Analogous to the solution for the broadly tuned case (16), now the total mean-input current can be expressed as

$$\begin{aligned} \langle I^{a\theta,\mathrm{tot}} \rangle = &\sqrt{K_0}\hat{I}_a^{\mathrm{ext}}(1 + \epsilon\cos 2(\theta - \theta_0)) \\ &+ \sum_{b=1}^{2}\sqrt{K_b}J_{ab}[r_{b,2}f_0(\theta_c) + \gamma r_{b,2}f_2(\theta_c)\cos 2(\theta - \theta_0)], \end{aligned} \tag{21}$$

where

$$f_0(\theta_c) = \int_{-\theta_c}^{\theta_c}\frac{\mathrm{d}\theta'}{\pi}(\cos 2\theta' - \cos 2\theta_c) = \frac{1}{\pi}(\sin 2\theta_c - 2\theta_c\cos 2\theta_c) \tag{22}$$

$$f_2(\theta_c) = \int_{-\theta_c}^{\theta_c}\frac{\mathrm{d}\theta'}{\pi}\cos 2\theta'(\cos 2\theta' - \cos 2\theta_c) = \frac{1}{\pi}(\theta_c - \frac{1}{4}\sin 4\theta_c). \tag{23}$$

(We have borrowed the notation from Ben-Yishai *et al.* [13] who studied a different kind of model that contains similar expressions; see also [17]). Again, the total current (21) has to vanish for all orientation columns θ, so both the constant and the $\cos 2(\theta - \theta_0)$ terms vanish separately:

$$\hat{I}_a^{\mathrm{ext}} + \sum_{b=1}^{2}\hat{J}_{ab}r_{b,2}f_0(\theta_c) \quad = 0 \tag{24}$$

$$\epsilon\hat{I}_a^{\mathrm{ext}} + \gamma\sum_{b=1}^{2}\hat{J}_{ab}r_{b,2}f_2(\theta_c) = 0 \tag{25}$$

Dividing (24) by (25) yields

$$\frac{f_2(\theta_c)}{f_0(\theta_c)} = \frac{\epsilon}{\gamma}, \tag{26}$$

which can be solved for θ_c. Note that (26), and thus the tuning width of the mean rates, does not depend on the overall strength of the input, I_a^{ext} (i.e., the "contrast" of the stimulus). We find therefore contrast-invariant tuning of the mean rates as a result of

cortical interactions, in agreement with experimental findings [10]. Having calculated θ_c, we can find the mean rates with help of (24), via

$$r_{a,2} = -\frac{1}{f_0(\theta_c)} \sum_{b=1}^{2} (\hat{\mathbf{J}}^{-1})_{ab} \hat{I}_b^{\text{ext}}, \tag{27}$$

and by using the equality $r_{a,0} = -r_{a,2} \cos 2\theta_c$.

The above calculations show how cortical interactions are responsible for a narrowing of the tuning of the population firing rates, relative to the tuning of the input to the network. We can proceed one step further in our analytical treatment of the mean-field model and consider the *tuning of the neuronal input noise spectrum*. We can write the dynamic noise in the input current as

$$\langle \delta I_{a\theta}^{\text{rec}}(t) \delta I_{a\theta}^{\text{rec}}(t') \rangle = \sum_{b=1}^{2} J_{ab}^2 \int_{-\pi/2}^{\pi/2} \frac{\mathrm{d}\theta'}{\pi} (1 + \gamma \cos 2(\theta - \theta')) C_{b\theta'}(t - t'), \tag{28}$$

where we have used the continuum notation for the weighted averages. The correlation function $C_{b\theta'}(t - t')$ has a piece proportional to $r_b(\theta)\delta(t - t')$, which gives

$$\lim_{\omega \to \infty} \langle |\delta I_{a\theta}^{\text{rec}}(\omega)|^2 \rangle = \sum_{b=1}^{2} J_{ab}^2 \int_{-\pi/2}^{\pi/2} \frac{\mathrm{d}\theta'}{\pi} (1 + \gamma \cos 2(\theta - \theta')) r_b(\theta') \tag{29}$$

$$= \sum_{b=1}^{2} J_{ab}^2 [r_{b,2} f_0(\theta_c) + \gamma r_{b,2} f_2(\theta_c) \cos 2(\theta - \theta_0)]. \tag{30}$$

To obtain (30), we performed calculations analogous to the ones for solving the integrals for the rate equations. Using (26) and (27), we can then write the flat contribution to the noise spectrum as

$$\lim_{\omega \to \infty} \langle |\delta I_{a\theta}^{\text{rec}}(\omega)|^2 \rangle = -\hat{I}_a^{\text{ext}} [1 + \epsilon \cos 2(\theta - \theta_0)] \sum_{b=1}^{2} J_{ab}^2 \sum_{c=1}^{2} (\hat{\mathbf{J}}^{-1})_{bc} \hat{I}_c^{\text{ext}}, \tag{31}$$

This result states that the high-frequency limit of the neuronal input noise has the same orientation tuning as the external input to the neuron.

For $t \neq t'$, it is not possible to calculate analytically solutions to (28) because the correlation function $C_{b\theta'}(t - t')$ needs to be evaluated numerically. Similarly, the tuning of the irregularity in the neuronal firing (as described by, e.g., the Fano factor) can only be determined by solving the full mean-field model numerically.

3. Numerical procedure

In our simulations, we modeled the orientation hypercolumn as an assembly of 30 orientation columns, with their preferred orientations θ equally spaced between $-\pi/2$ and $\pi/2$ (or between -90 and 90 degrees, as in the figure captions). We used parameter values corresponding to $N_1 = 8000$ excitatory and $N_2 = 2000$ inhibitory neurons, and a membrane time constant of $\tau = 10$ ms for all neurons. The generic intra-cortical connection strengths J_{ab} in (2) were taken as

$$\begin{pmatrix} J_{11} & J_{12} \\ J_{21} & J_{22} \end{pmatrix} = \begin{pmatrix} 0.5 & -2 \\ 1 & -2 \end{pmatrix}. \tag{32}$$

The synaptic strengths of the afferent inputs from the LGN are taken to be stronger for the excitatory neurons; specifically, in (4), we chose $I_2^{\rm ext} = \frac{2}{3} I_1^{\rm ext}$. To study the role of the overall strength of synapses, we multiply the generic synaptic weights (including the strength of the external input) by a common scaling factor J_s.

We use an iterative approach that was originally developed for spin glass models [18] to find self-consistent solutions of the firing statistics given by the rates $r_a(\theta)$, the rate fluctuations $\overline{(r_j^{a\theta})^2}$, and the correlations $C_{a\theta}(t-t')$. We start with initial estimates of these quantities, which we obtain by using a white-noise approximation in the analytical treatment described above. We then generate many realizations of Gaussian synaptic currents using (4) and (9), which we use to drive single integrate-and-fire neurons. By collecting their firing statistics, we obtain improved estimates of the rates, rate fluctuations, and correlations. These are then used to repeat the cycle until the input and output statistics are consistent.

For the hypercolumn model, we need to determine these firing statistics for each population a (excitatory and inhibitory) within each orientation column θ. However, because of the inherent symmetry in the network topology, at each iteration step we only need to run simulations for half of the columns and mirror the results to obtain improved statistics for the entire network. To collect the firing statistics from the column population $a\theta$, we run many trials of single neurons that are sampled from the entire column population. This is achieved by generating Gaussian input currents that fluctuate not only in time (by generating realizations of the dynamic and appropriately colored input noise $\xi_{b\theta'}(t)$ in (11)), but also differ in their overall size due to the random numbers $x_{b\theta'}$ in (11), which reflects the fact that different neurons have in general different connectivity patterns. (Note that we have used here – as throughout the text – the indices $a\theta$ for referring to the "target column", whereas $b\theta'$ runs over all "source columns"). For a more detailed account on handling some of the subtleties in obtaining the correct statistics, see [9].

Once the procedure converges, which takes tens to hundreds of iterations, depending on the set of parameters and the specific approach taken, one has obtained a set of self-consistent firing statistics, describing the population responses for a specific network input (stimulus contrast and stimulus orientation). Equipped with these population statistics we can then calculate input and firing statistics for individual neurons. To specify such a neuron, we select a set

$$\{x_{b\theta'} : b = 1, 2;\ \theta' = \theta_1, \ldots, \theta_n\} \tag{33}$$

and keep it fixed over all trials to collect the statistics for that neuron. The $x_{b\theta'}$ represent the intrinsic variability across the population in the strength of synaptic input due to the randomness in the connectivity of the network.

4. Results

We concentrate first on results describing response characteristics of neurons obtained from their firing statistics. It is possible to compare these results directly with known

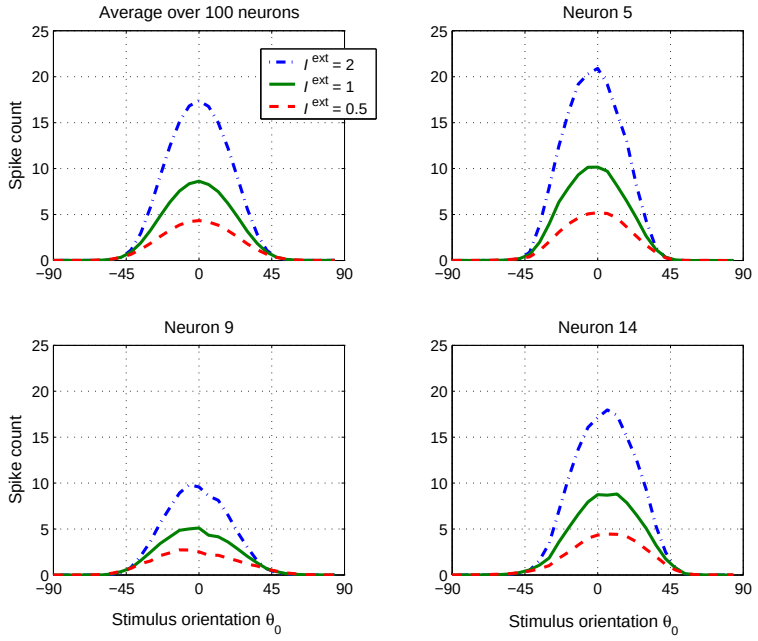

Figure 2. Contrast-invariant tuning width. Average over 100 neurons (upper-left panel) and three randomly chosen neurons. The parameter values for the stimulus tuning and the connectivity tuning were $\epsilon = 0.5$ and $\gamma = 0.625$, respectively, resulting in a tuning width of 43.2 degrees according to the calculations. Contrast-invariant tuning is observed for both averaged and single-neuron tuning, despite the small distortions and asymmetries for single neurons. (See the text for further details)

properties like contrast-invariant tuning, or with the variability in spike counts. We then describe results pertaining to properties of the neuronal input currents (and their orientation tunings) for the hypercolumn model.

4.1. Tuning of the neuronal firing

For the present model, we have shown analytically above that the tuning width of the column population rates is invariant with respect to the contrast of the stimulus (see Equation (26)). We investigated whether such contrast-invariant tuning is also observed for single, randomly chosen neurons. The number of afferent connections that a given neuron receives from neurons with another preferred orientation is a random number drawn from a probability distribution given by (1). This will in general distort the shape of the neuron's tuning curve. Figure 2 shows the tuning curves of three randomly chosen neurons from the column with $\theta = 0$ for three different contrasts $\hat{I}_a^{\text{ext}} = 0.5, 1$, and 2. For our network with 30 orientation columns and 2 populations, the resulting realization of the random connectivity to a single neuron is therefore determined by a set of 60 random numbers (see Equation (33)). To record the neuronal responses, these sets were held fixed, while the network was presented successively with stimuli of all orientations θ_0. Also shown in Figure 2 is the result of averaging over the tuning curves of $n = 100$ randomly chosen neurons. While the averaged tuning is both smooth and symmetric, the

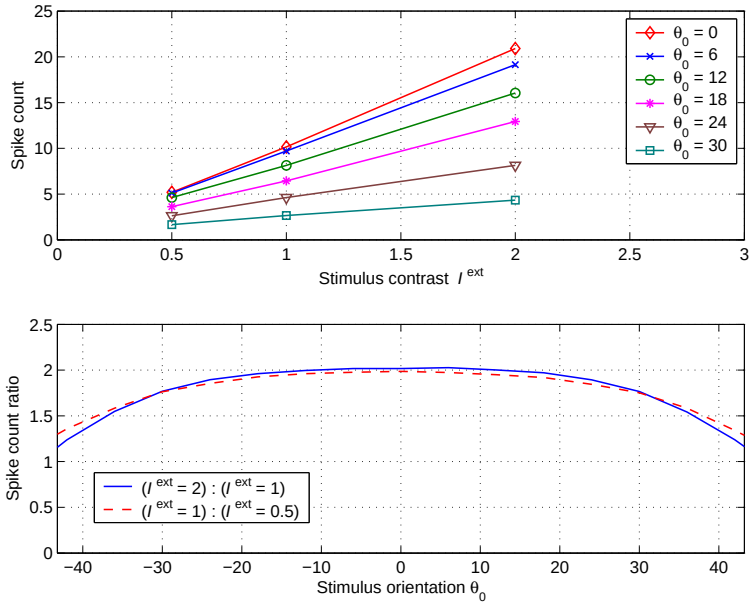

Figure 3. Tuning of the gain function. *Upper panel:* Spike count as a function of stimulus contrast, parameterized by the stimulus orientation. The input-output relationship is linear, and the slope decreases as the stimulus orientations θ_0 becomes more dissimilar to the neuron's preferred orientation $\theta = 0$ (results shown for neuron 5 in Figure 2). *Lower panel:* Spike count ratios for two pairs of spike counts resulting from doubling the contrast. At the preferred orientation (PO) and for orientations not too far from the PO, doubling the contrast doubles the spike count. For more dissimilar stimulus orientations, the ratios decrease systematically.

tuning curves of single neurons show small distortions and asymmetries. Additionally, the overall strength of the response varies from neuron to neuron. However, despite the somewhat irregular shapes, the contrast-invariance of the tuning width is preserved for single, randomly chosen neurons. The analytical treatment predicts a threshold-cosine shape of the tuning, while the curves shown here, including the averaged ones, show a rounded fall-off to zero with non-zero rates for angles just outside the tuning width. This "rounding artifact" appears to be due to a slow convergence of the numerical procedure at extremely low firing rates; the artifact is reduced when the algorithm is run for more iterations.

In all our simulations, we observe an almost linear input-output relationship between stimulus contrast and firing rate, in agreement with experiments (see, e.g., Figure 1 in [12]). Figure 3 shows how the input-output relationship depends on the stimulus orientation. In the upper panel of Figure 3, the spike count is plotted as a function of the external input strength $\hat{I}_1^{\text{ext}}$, i.e. the contrast of the stimulus, for a single neuron (neuron 5 of Figure 2). The slope changes systematically with stimulus orientation θ_0, getting smaller as the difference between the stimulus orientation and the neuron's preferred orientation increases. The lower panel of Figure 3 shows the spike count ratios of two pairs of spike counts that resulted from doubling the stimulus contrast. In contrast to the upper panel of Figure 3, these curves show results of

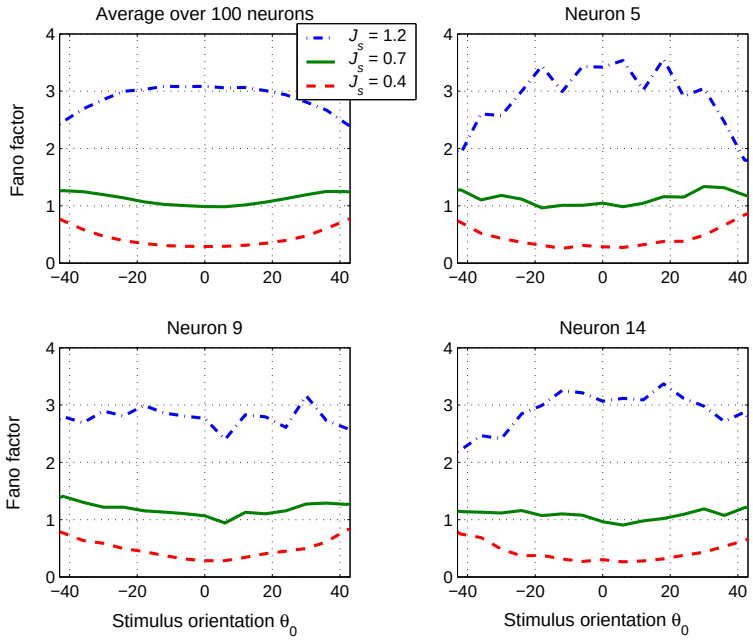

Figure 4. Tuning of the Fano factors. Tuning curves, parameterized by relative synaptic strengths J_s, are shown for the same three neurons as in Figure 2 and for an average over 100 neurons (upper left panel). The Fano factors F depend systematically on J_s: stronger synapses lead to higher Fano factors. On average, F stays either above 1 for all orientations or below 1 for all orientations. For $F \approx 1$, the tuning is almost flat, while it reaches a maximum (resp. minimum) at the preferred orientation for $F > 1$ (resp. $F < 1$).

averaging spike counts over 100 neurons, in order to make the general tendency clearer. It can be seen that for the preferred orientation, doubling the stimulus almost perfectly doubles the spike count (this is also true for single neurons, as can be read off from Figure 2). This relationship also holds for stimulus orientations away from the PO, until about 20 degrees difference, which is about half the tuning width of these neurons. For larger orientation differences, the ratio decreases. It seems likely that at large orientation differences (near the tuning width) this reduction is due to the rounding artifact for very low spike rates discussed above. For intermediate orientation differences, say 20–35 degrees, the reason for the reduction is not evident to us.

We characterize the irregularity in the neuronal firing by the Fano factor F. For a Poisson process $F = 1$, while $F \neq 1$ implies temporal correlations in the spike times: $F > 1$ indicates a tendency towards "bursty" spiking behavior, and $F < 1$ indicates more regular spike trains with narrower interspike interval (ISI) distributions. Figure 4 shows the tuning of the Fano factor for three different overall connection strengths $J_s = 0.4, 0.7$, and 1.2. As in Figure 2, the results for (the same) three individual neurons are shown, as well as an averaged tuning curve. It can be seen that the Fano factor depends systematically on the overall strength of connectivity: stronger synapses lead to more irregular spike counts. The averaged tuning curves reveal two further properties, which we observed consistently in all our simulations, performed with many

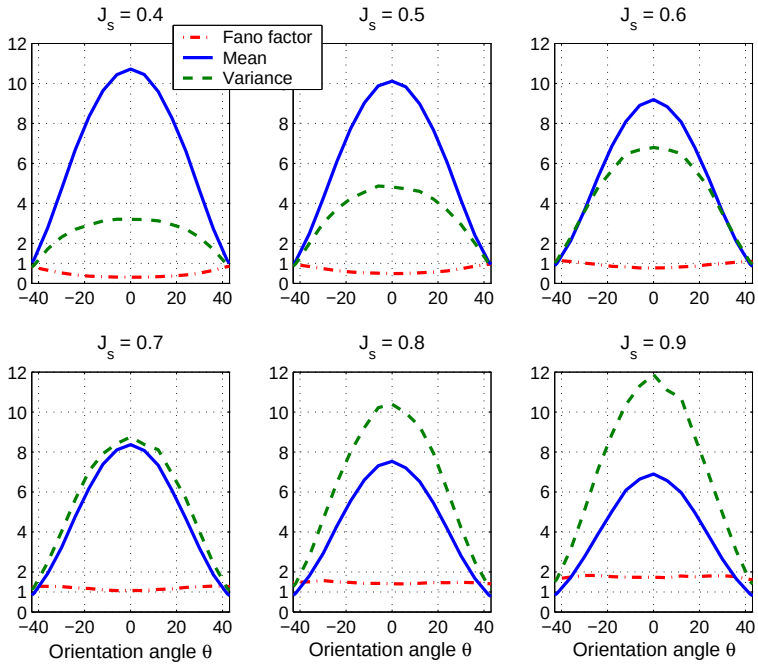

Figure 5. Analysis of Fano factor tuning: tuning of mean spike count and spike count variance for relative synaptic strengths $J_s = 0.4, 0.5, \ldots, 0.9$. For each J_s, the variance stays either below the mean or above the mean for all orientations (upper and lower panels, respectively), resulting in ratios $F < 1$ and $F > 1$ for all orientations. The variance increases with J_s – most sensitively at the preferred orientation (PO). For $F \approx 1$, the variance and mean tuning curves are almost identical, resulting in an almost flat tuning of their ratio F, while for $F \neq 1$ the ratios reach a minimum/maximum at the PO.

different sets of parameters: First, Fano factors are either less than 1 at all angles or greater than 1 at all angles. Second, if they are considerably greater than 1, they peak at the preferred orientation, falling off as the difference between stimulus orientation and PO increases; in the case where F stays below 1, the opposite tuning is observed, i.e., the Fano factor reaches a minimum at the preferred orientation. We can shed some light on the emergence of these two properties by looking at pairs of tuning curves for the spike count variance and the mean spike count and then systematically changing the connection strengths. We show these tuning curves for 6 different values of J_s in Figure 5. It can be seen that both the mean and the variance peak at the PO, falling off towards increasing angle differences. Furthermore, for $F \approx 1$ at $J_s = 0.7$, the tuning curves are nearly identical resulting in almost untuned Fano factors close to 1. For lower J_s values, the variance curve stays entirely below the mean curve, while the opposite is true for J_s values bigger than 0.7. Therefore, the ratio of the curves, which is the tuning curve of the Fano factor, stays either always below 1 or always above 1. The size of the spike count variance depends sensitively on the overall connection strengths J_s. Apparently, this sensitivity is strongest at the PO, decreasing towards greater angle differences. Therefore, the Fano factor reaches its minimum for the cases with $F < 1$

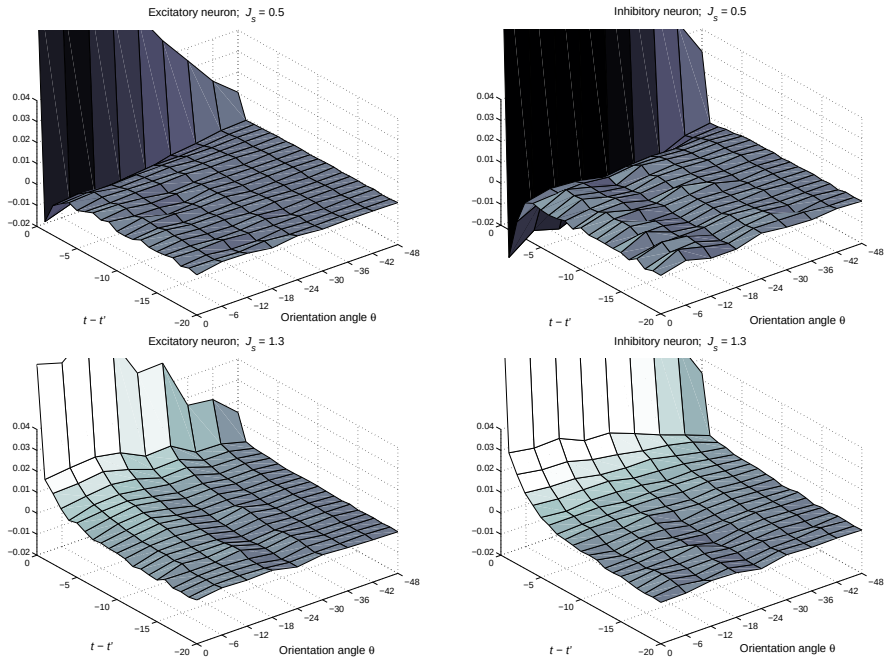

Figure 6. Autocorrelation tuning. *Upper panels:* Weak synapses with $J_s = 0.5$. There is a dip to negative values for small time differences. It decreases in strength at greater time differences. The dip indicates a relative refractoriness to emitting a spike immediately after a previous one, resulting in Fano factors $F < 1$. *Lower panels:* Strong synapses with $J_s = 1.3$. There is a hill of positive correlations for short intervals, falling off to zero for increasing time differences. The hill indicates a tendency toward clustered spikes, resulting in $F > 1$. The autocorrelations for excitatory neurons (left panels) and inhibitory neurons (right panels) show the same qualitative features, differing only in overall size.

(respectively its maximum for $F > 1$) when the stimulus is at the preferred orientation.

As already mentioned, Fano factors that deviate from 1 indicate temporal correlations in the spike trains. The nature of these correlations and their orientation dependence is summarized in Figure 6 for a case with $F < 1$ ($J_s = 0.5$; upper panels) and a case with $F > 1$ ($J_s = 1.3$; lower panels) for both excitatory neurons (left panels) and inhibitory ones (right panels). For $J_s = 0.5$, there is a negative dip for small time differences, indicating a relative refractoriness to emitting a spike immediately after a previous one. For stronger synapses ($J_s = 1.3$) there is no such refractoriness. On the contrary, for strong synapses, we observe positive correlations for small time differences. For both strong and weak synapses, the correlations are strongest at the preferred orientation and decrease monotonically for less optimal stimulus orientations. The autocorrelations for excitatory and inhibitory neurons show the same qualitative features, differing only in their overall size.

In Figure 7 we illustrate how the firing statistics depend on ϵ and γ, which determine how strongly the input current and the intracortical connectivity are tuned (see equations (4) and (1), respectively). Fano factor tuning curves (left panels) and firing rate tuning curves (right panels) for three different combinations of ϵ and γ are

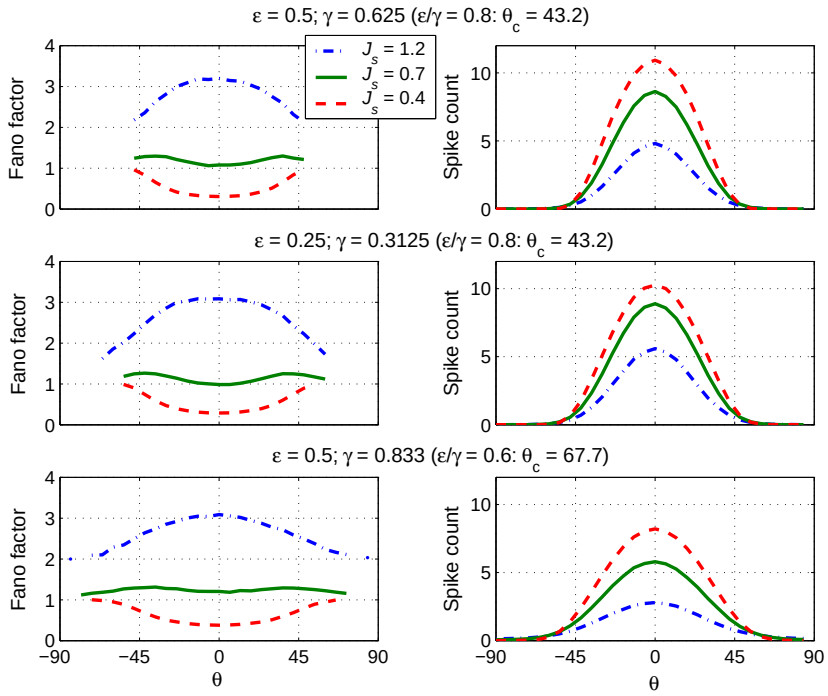

Figure 7. Dependence of the Fano factors on tuning parameters ϵ and γ at three different values of relative synaptic strengths J_s. Fano factors and mean spike counts are shown for three different combinations of ϵ (external input tuning) and γ (connectivity tuning). The tuning of both the Fano factors and the mean counts are controlled by the ratio ϵ/γ.

shown, parameterized by J_s, the scaling factor for the synaptic strengths. As shown analytically above, the ratio ϵ/γ determines the tuning width of the neuronal firing (see Equation (26)). This is reflected by the identical firing tuning widths in the first and second row of Figure 7, for both of which $\epsilon/\gamma = 0.8$, resulting in a tuning width of $\theta_c = 43.2$ degrees. The third row of Figure 7 shows results for the same external input tuning $\epsilon = 0.5$ as in the first row, but for a different ratio $\epsilon/\gamma = 0.6$. This results in $\theta_c = 67.7$ degrees and an accordingly broader tuning curve of the firing, plotted in the right panel of the third row. The curves for the Fano factor tuning in the left panels of Figure 7 suggest that the tuning of the firing irregularity is – just as the tuning of the firing itself – only dependent on the ratio ϵ/γ. (We consistently found this dependence in all our simulations.)

4.2. Tuning of the neuronal input current

Our analytical treatment of the balanced hypercolumn model reveals that the high-frequency neuronal input noise power has the same tuning as the external input. In Figure 8 we show simulation results of the noise tuning for the same three combinations of ϵ and γ as in Figure 7. For the panels in the first and the second row of Figure 8, $\epsilon/\gamma = 0.8$, but $\epsilon = 0.5$ and $\epsilon = 0.25$ in the upper and middle rows, respectively.

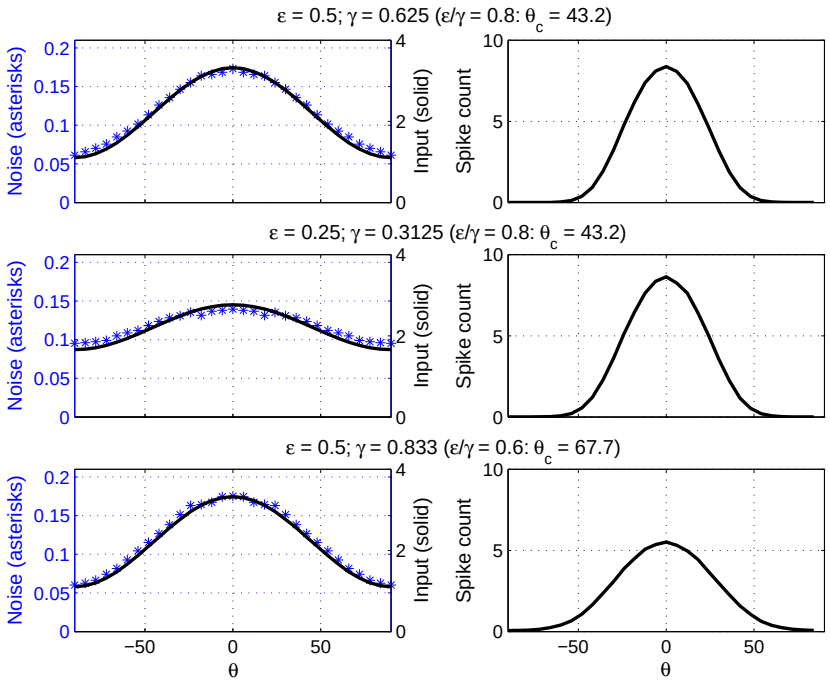

Figure 8. Dependence of the noise on tuning factors ϵ and γ. External input and dynamic input noise versus tuning of the neuronal firing for the same three combinations of ϵ and γ as in Figure 7. It can be seen that the tuning of the noise is determined by ϵ, while the tuning of the firing rate is determined by the ratio ϵ/γ.

While the tuning of the neuronal firing is identical for these two cases, the noise tuning is weaker in the middle row, reflecting the weaker tuning of the external input (left panels). The results presented in the third row of Figure 8 show a case with a broader tuning of the response, resulting from a different ratio between ϵ and γ, but with the same $\epsilon = 0.5$ as in the first row. For these two cases, the tunings on the input side – concerning external input and dynamic noise – are practically indistinguishable, while the tunings of the firing differ. Thus, the noise tuning is determined by ϵ, unlike the response tuning, which depends on the ratio ϵ/γ.

The balanced state for the orientation hypercolumn implies that the mean input currents (external and recurrent currents), which are each of $\mathcal{O}(\sqrt{K_a})$ with $K_a \gg 1$, cancel up to corrections of $\mathcal{O}(1)$. It is not straightforward to calculate the tuning of the resulting net mean current, since the balance condition (14) does not allow inferences about its size. However, the solutions obtained by the numerical algorithm provide direct access to the net mean currents, which we depict in Figure 9 for the same combinations of ϵ and γ as for the noise tuning in Figure 8. It is clear from Figure 9 that the tuning of the mean input, unlike the dynamic input noise tuning, is not determined by the tuning of of the external input. Rather, it seems to be the ratio ϵ/γ that primarily determines it, as suggested by the almost identical tunings for the two cases with identical ϵ/γ. Since the tuning of the external input and that of the noise variance are the same, the left panels of Figure 9 also show how the tuning of the noise compares to that of the

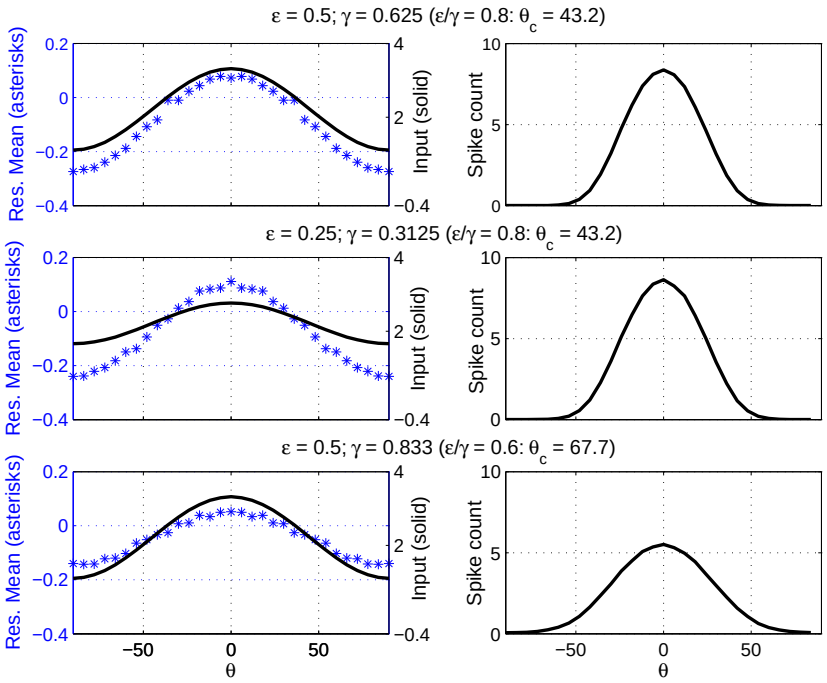

Figure 9. Dependence of the mean input current on tuning factors ϵ and γ. External input tuning and mean-input tuning versus tuning of the response for the same three combinations of ϵ and γ as in Figure 7 and Figure 8. The tuning of the mean input is not determined by ϵ; rather, as for the spike count tuning shown in the right panels, the ratio ϵ/γ plays an important role.

mean input current for the three combinations of ϵ and γ.

5. Discussion

In this work, we presented a complete mean field theory for a balanced network with structural inhomogeneity, together with an algorithm that allows one to find the self-consistent solutions for the mean rates, their cell-to-cell fluctuations, and the correlation functions. We applied the theory to a simple model of an orientation hypercolumn in primary visual cortex, comprised of integrate-and-fire neurons. Despite the relative simplicity of the model, the resulting dynamics capture several key properties known about responses of orientation selective cortical neurons *in vivo*. Within this description, we can pinpoint how the resulting neuronal dynamics are controlled by parameters of the model, and quantify their influence.

Specifically, we find contrast-invariant tuning of the neuronal firing not only for the population rates, as derived from the analytical treatment, but also for single, randomly chosen neurons. Moreover, the firing rate increases linearly with the strength of the input current (i.e., the contrast of the stimulus). Note that these are network effects originating in the dynamical balance between excitation and inhibition, not properties of isolated neurons. This is in agreement with experimental results, where such a linear

input-output relationship can only be found for cortical neurons *in vivo*, but not for single neurons *in vitro*.

Another network effect that emerges naturally from the self-consistent dynamic balance, in combination with the static randomness in the connectivity, is the irregularity in the neuronal firing. We are able to describe it quantitatively through the correlation functions, which are determined self-consistently in the theory. Such firing-statistical issues cannot be addressed in "rate models", which simply assume a particular relation between average input current or membrane potential and firing rate. While it is possible to calculate the firing variability in the mean-field treatment of Brunel [6], it cannot be done in a self-consistent manner because of the assumption that the neuronal input is uncorrelated in time (white noise). Here we color the noise self-consistently. Poisson-like statistics (Fano factor $F = 1$) are only one possibility within a continuum of firing statistics that depend sensitively on the strengths of the synapses: stronger synapses generally lead to higher Fano factors. The underlying mechanism can be summarized as follows: Stronger synapses increase the probability of a spike shortly after reset, which leads to a higher tendency of spikes occurring in "clusters", thereby increasing the spike count variance. A detailed account of this mechanism, involving the dependence of the membrane potential distribution on the synaptic strength can be found in [9], where the analysis was carried out for a single cortical column.

The mean field theory applied to the present model allows us to study tuning properties of both the neuronal firing and the neuronal input and their dependence on network parameters. Concerning the irregularity of firing, our results suggest that F stays either above 1 or below 1 for all orientations. Moreover, the modulation strength of F over angles increases, relative to the almost untuned case of $F \approx 1$, with increasing (resp. decreasing) overall values of F, reaching a maximum (resp. a minimum) at the preferred orientation.

Concerning the tuning of the input currents, we find analytically that the high-frequency input noise power has the same tuning as the external input to the neuron (which in turn is determined by a Hubel-Wiesel feed-forward connectivity from the LGN). In our numerical calculations we observe a close fit between the tuning of the overall input noise and the one of the external input. This suggests that the tuning of the external input may be a good predictor for the noise tuning, and vice versa. In contrast, we find that the tuning of the mean input current does not reflect the one of the external input, but is predominantly determined by the ratio ϵ/γ of the modulation strengths of the external input and the cortical interactions.

Some of our results (the existence of a stable, asynchronous low-rate state, contrast-invariant orientation tuning, and the inverse relation between the sharpness of orientation tuning and intracortical tuning strength γ) were obtained previously by Wolf *et al.* [19] in an extension of van Vreeswijk and Sompolinsky's stochastic binary model [2, 3] to a hypercolumn, but the treatment of a spiking neuron model and all the results for correlations of both input and output are new here. Also new is that we go beyond population statistics and make quantitative predictions about input and output

characteristics of *individual* neurons, which can be tested directly.

Firing irregularity of neurons in primary visual cortex has been investigated experimentally for a long time (see, e.g., [20, 21, 22, 23, 14]). Well studied is also the dependence of firing rate on the stimulus orientation [10, 24], but we are not aware of studies investigating the dependence of firing irregularity on the orientation. Our predictions concerning the tuning of the input currents (for both mean and noise) can be tested experimentally by systematically changing ϵ (the external input tuning strength) via changing the spatial modulation of the stimulus and then observing how the the mean and noise tunings are affected separately.

The mean field theory presented here, in combination with the numerical procedure for finding the self-consistent solutions, can be applied to models that capture more of the known neuronal and cortical physiology. For example, it is straightforward to incorporate conductance-based synapses into the hypercolumn model, as has already been done for a single-column model (see [25] and [15]). It is also straightforward to use different, possibly more realistic neuron models – even several kinds of neuron models within one given network model, since the neuronal dynamics are explicitly simulated within the numerical procedure for collecting the firing statistics. Here, we have shown how the theory can be applied to networks with non-homogenous architecture, using a simple one-dimensional model for a cortical hypercolumn. This model can be thought of as describing an annulus around a pinwheel center. Using the same general techniques as introduced here, the model can be extended to incorporate a two-dimensional geometry to describe an entire pinwheel. Similarly, as we have shown for orientation selectivity, it is possible to include other coding features, such as spatial phase, for example. Thus, the power of this mean-field approach lies in its generality, which makes it possible to quantify dynamics of balanced, highly connected networks.

References

[1] Softky W R and Koch C 1993 The highly irregular firing of cortical cells is inconsistent with temporal integration of random EPSPs. *J. Neurosc.* **13** 334–350

[2] van Vreeswijk C and Sompolinsky H 1996 Chaos in neuronal networks with balanced excitatory and inhibitory activity. *Science* **274** 1724–1726

[3] van Vreeswijk C and Sompolinsky H 1998 Chaotic balanced state in a model of cortical circuits. *Neural Comp.* **10** 1321–1371

[4] Amit D and Brunel N 1997 Model of spontaneous activity and local structured activity during delay periods in the cerebral cortex. *Cereb. Cortex* **7** 237–252

[5] Amit D and Brunel N 1997 Dynamics of a recurrent network of spiking neurons before and following learning. *Network* **8** 373–404

[6] Brunel N 2000 Dynamics of sparsely connected networks of excitatory and inhibitory spiking neurons. *J. Comput. Neurosci.* **8** 183-208

[7] Hertz J, Richmond B and Nilsen K 2003 Anomalous response variability in a balanced cortical network model. *Neurocomputing* **52–54** 787–792

[8] Fulvi Mari C 2000 Random networks of spiking neurons: instability in the *xenopus* tadpole moto-neuron pattern. *Phys. Rev. Lett.* **85** 210–213

[9] Lerchner A, Ursta C, Hertz J and Ahmadi M 2004 Response variability in balanced cortical networks. *submitted to Neural Computation*

[10] Sclar G and Freeman R 1982 Orientation selectivity in cat's striate cortex is invariant with stimulus contrast. *Exp. Brain Res.* **46** 457–461

[11] Hubel D H and Wiesel D N 1962 Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. *J. Physiol. Lond.* **160** 106–154

[12] Sompolinsky H and Shapley R 1997 New perspectives on the mechanisms for orientation selectivity. *Curr. Opinion Neurobiol.* **7** 514–522

[13] Ben-Yishai R, Lev Bar-Or R and Sompolinsky H 1995 Theory of orientation tuning in visual cortex. *Proc. Natl. Acad. Sci.* (USA) **92** 3844–3848

[14] Gershon E, Wiener M C, Latham P E and Richmond B J 1998 Coding strategies in monkey V1 and inferior temporal cortex. *J. Neurophysiol.* **79** 1135–1144

[15] Hertz J, Lerchner A and Ahmadi M 2004 Mean field methods for cortical network dynamics, *to be published in Springer Lect. Notes in Comp. Sci.*

[16] Kree R and Zippelius A 1987 Continuous-time dynamics of asymmetrically diluted neural networks. *Phys. Rev. A* **36** 4421–4427

[17] Hansel D and Sompolinsky H 1998 Modeling feature selectivity in local cortical circuits. in *Methods in Neuronal Modeling: from Synapse to Networks*, Koch C and Segev I, eds (MIT Press)

[18] Eisfeller H and Opper M 1992 New method for studying the dynamics of disordered spin systems without finite size effects. *Phys. Rev. Lett.* **68** 2094-2097

[19] Wolf F, van Vreeswijk C and Sompolinsky H 2001 Chaotic activity induces contrast invariant orientation tuning *Soc. Neurosci. Abstr.* 12.7

[20] Heggelund P and Albus K 1978 Response variability and orientation discrimination of single cells in in striate cortex of cat. *Exp. Brain Res.* **32** 197–211

[21] Dean A F 1981 The variability of discharge of simple cells in the cat striate cortex. *Exp. Brain Res.* **44** 437–440

[22] Tolhurst D J, Movshon J A and Thompson I D 1981 The dependence of response amplitude and variance of cat visual cortical neurones on stimulus contrast. *Exp. Brain Res.* **41** 414–419

[23] Snowden R J, Treue S and Andersen R A 1992 The response of neurons in areas V1 and MT of the alert rhesus monkey to moving random dot patterns. *Exp. Brain Res.* **88** 389–400

[24] Skottun B, Bradley A, Sclar G, Ohzawa I and Freeman R 1987 The effects of contrast on visual orientation and spatial frequency discrimination: a comarison of single cells and behavior *J. Neurophysiol.* **57** 773–786

[25] Lerchner A, Ahmadi M and Hertz J 2004 High conductance states in a mean field cortical network model. *in press: Neurocomputing*