Neural Mechanisms and Computational Principles of Adaptive Sensory Processing

vorgelegt von

Diplom-Ingenieur
Klaus Wimmer

aus Eferding, Österreich

von der Fakultät IV – Elektrotechnik und Informatik
der Technischen Universität Berlin
zur Erlangung des akademischen Grades

Doktor der Naturwissenschaften
Dr. rer. nat.

genehmigte Dissertation

Promotionsausschuss:

Vorsitzender: Prof. Dr. Manfred Opper
Berichter: Prof. Dr. Klaus Obermayer
Berichter: Prof. Valentin Dragoi, Ph.D.

Tag der wissenschaftlichen Aussprache: 8. Oktober 2009

Berlin 2009

D 83

What is now proven
was once only imagined.

(William Blake)

Acknowledgments

I am very grateful to Prof. Klaus Obermayer, the supervisor of this thesis, for giving me the opportunity to be part of his inspiring "Neural Information Processing Group" at the Technische Universität Berlin. Discussions with him were always focused, he was always critical, and I would like to thank him for directing me towards interesting problems and shaping my scientific thinking. This thesis would not have been possible without his encouragement of interdisciplinary work with experimentalists, and all the invited guests, workshops, and meetings with our collaborators.

I am indepted to Prof. Mriganka Sur for sharing the experimental data that inspired the modeling studies presented in this thesis. During several short stays in his research lab at MIT, I gained many insights into neurophysiology and experimental research in neuroscience. Collaborating with Prof. Mriganka Sur, James Schummers, and Jitendra Sharma was highly stimulating and inspiring.

I would also like to thank Matthias Hennig and Jannis Hildebrandt from the Humboldt Universität zu Berlin, who collaborated on the project about adaptation in the cricket auditory system. I very much enjoyed the fruitful collaboration.

Many thanks to Prof. Valentin Dragoi for being a member of my thesis committee. While I briefly visited his lab in Houston, I was fascinated not only by the scientific work, but also by the kindness of this lab members; I am especially indebted to Diego Gutnisky for his exceptional hospitality.

Thanks to all my colleagues at the Neural Information Processing Group and at the Bernstein Center for Computational Neuroscience. They all contributed in making our research group not only my intellectual home. In particular, I would like to thank Marcel Stimberg and Robert Martin – working with them on several projects was a pleasure and a lot of fun. Kamil Adiloglu was a great roommate. Many thanks to Lars Schwabe and Joshua Young who helped me getting started in the field of Computational Neuroscience and shared many exciting ideas with me. I also like to thank Thorsten Dietzsch, Gidi Farhi, Christian Rodloff – co-supervising their projects and master theses was a pleasure for me.

Thanks go to my family and to all my friends, and a very special thank goes to María Cañizares Rivas for being in my life.

Summary

Neural systems compute dynamic representations of the environment: sensory neurons respond selectively to sensory stimuli, and adapt their "code" according to recently received sensory input. Despite the ubiquity of adaptation throughout neural processing across different species and sensory modalities, the functional benefits of adaptation often remain unclear. Ultimately, a complete understanding of the underlying computational principles requires insight into the neural circuitry and the physiological mechanisms that carry out these computations. Here, we study the generation and adaptation of sensory representations in low-level sensory processing in two model systems of different complexity: the auditory pathway of crickets and the primary visual cortex (V1) of cats.

We first investigate the generation and dynamics of orientation selectivity in V1. Physiological and anatomical data indicate that the response properties of individual neurons can only be fully understood in the context of their local circuitry. Specifically, orientation tuned responses of a V1 neuron and its temporal response characteristics depend on the neuron's position in the orientation preference map and the recurrent inputs related to that position. We systematically explore a whole class of network models that incorporate the structure of the orientation preference map and conclude that only a network operating in a regime where excitatory and inhibitory recurrent inputs are balanced and dominate the feed-forward input is consistent with the experimental data. Furthermore, we argue that one functional benefit of this "balanced recurrent" operating regime is its enhanced sensitivity to modulations of the balance between excitation and inhibition.

We then focus on intracortical synaptic depression as a potential mechanism through which adaptation alters recurrent processing and thus V1's representation of sensory stimuli. We simulate orientation adaptation in network models with different strengths of synaptic depression and select those models that predict adaptation-induced changes in orientation tuning and perceptual read-out consistent with experimental data. The relative strength of synaptic depression of the selected models is consistent with *in vitro* data, making intracortical synaptic depression a plausible mechanism of orientation adaptation. The best fit of the experimental data is found for the balanced recurrent network, which also correctly predicts an enhanced capacity for adaptive changes close to pinwheel centers.

Finally, we study the optimality of adaptive systems in the simpler auditory pathway of the cricket. Unlike in primary visual cortex, where information is encoded by populations of neurons, just a single neuron conveys information about high frequency sounds to the cricket's brain. This enables us to accurately quantify the mutual information between a sensory stimulus and its neural representation. We design an experiment that allows us to distinguish whether this representation follows an infomax principle or a selective coding principle. We find that adaptation leads to a reduction rather than to an overall enhancement in information transmission, inconsistent with the infomax principle. However, adaptation also selectively decreases the amount of information that is transmitted about background signals, facilitating the detection of behaviorally relevant signals.

Contents

1. Introduction

1.1. Adaptive sensory processing

How do our brains process the sensory information coming from the eyes in order to shape our perception of the outside world? Understanding the computational principles and the neural mechanisms that underlie sensory processing is a major challenge in neuroscience research. Ultimately, it is some representation of the sensory information that the brain uses to perform higher-level processing, such as recognizing objects, making decisions, and navigating our body through a dynamically changing environment.

Indeed, the fact that the environment is dynamically changing is a crucial issue for understanding perception. For example, when in motion we constantly view individual objects from different perspectives, or while they are occluded by other objects, or lit from a different angle or intensity of light, and so on. Remarkably, our brains can deal with this complexity and easily recognize the same object in a variety of situations. How does the brain achieve this? We are still far from answering this question – our understanding of the higher-order visual brain areas that deal with object recognition is limited: the algorithms for performing these tasks remain largely undiscovered, and the underlying neural machinery that executes them not well understood.

However, much more is known about the main cortical visual area for "lower-level" processing, the primary visual cortex (V1). The computations performed in V1 – the extraction of elementary features such as contour orientation within an image – are simple enough to be measured precisely by presenting basic visual stimuli and recording neural responses. Yet, the responses of V1 neurons are complex enough to be interesting; this makes the visual cortex an ideal model system for studying neural computations. It has long been thought that V1 neurons respond faithfully and unalterably to physical properties of visual stimuli within their receptive fields. For instance, the preferred orientation of a cortical cell was thought to be shaped during development and to remain unchanged later in life. However, this view has changed as it became clear that the responses of V1 are remarkably flexible or "plastic" even during normal sensory processing; for example, they are strongly influenced by the history of the visual input. This suggests that one strategy the brain uses to deal with an ever changing environment is to adapt its sensory representations.

Physiological mechanisms

A physiologically well characterized form of adaptation in V1 is pattern adaptation. Pattern adaptation occurs across several stimulus dimensions, such as contrast (Maffei et al., 1973; Ohzawa et al., 1985; Carandini and Ferster, 1997; Sanchez-Vives et al., 2000b), orientation (Müller et al., 1999; Dragoi et al., 2000), or direction of motion (Hammond et al., 1985). An important characteristic of adaptation is that it is a rapid form of plasticity[1] that has a reversible effect on neuronal selectivity, i. e., the responses adapt on short time scales, and recover to their pre-adapted state when the source of adaptation is removed. How adaptation alters sensory representations, and in turn perception, can be demonstrated by simple psychophysical experiments (Figure 1.1). After prolonged viewing of the oblique lines (Figure 1.1A), the vertical lines appear briefly as if they were tilted in the opposite direction. This well-known phenomenon is called the tilt aftereffect (Gibson and Radner, 1937). Adaptation often leads to a reduction in sensitivity to stimulus attributes. For instance, in contrast adaptation (Figure 1.1C), the exposure to a high-contrast stimulus reduces the apparent contrast of a test stimulus.

What is the neural substrate of these aftereffects? Contrast adaptation shifts the contrast response curve of individual neurons towards higher contrast levels on a time scale of seconds (Figure 1.1D, Ohzawa et al., 1985; Carandini, 2000). A simple explanation is thus, that adaptation "fatigues" the neurons that respond most strongly; the fatigued neurons then respond less than they normally would, leading to changes in the perception. Contrast adaptation in V1 is mainly meditated by a sodium-gated potassium channel, as described by Sanchez-Vives et al. (2000a,b). This channel is triggered by the sodium influx that occurs with synaptic input and the generation of action potentials and it leads to a strong hyperpolarization of individual neurons (Carandini and Ferster, 1997). There is also a subcortical contribution to contrast adaptation, because contrast adaptation already occurs to some degree in earlier processing stages than V1 (Solomon et al., 2004); in fact, it can already be found in retinal circuits (Baccus and Meister, 2002).

Concerning orientation adaptation, it has been shown experimentally in monkey (Müller et al., 1999) and in cat (Dragoi et al., 2000), that adaptation to orientated gratings systematically changes neural tuning curves (Figure 1.1B): The response at the adapting orientation decreases, the tuning curve broadens for neurons with a preferred orientation

[1] Neural responses in V1 are also modified after visual discrimination learning (Schoups et al., 2001; Gilbert et al., 2001), during short-term visual memory (Supèr et al., 2001), or by top-down processes reflecting an internal state such as attention and expectation (Sharma et al., 2003). These forms of plasticity are beyond the scope of this thesis; we focus on bottom-up plasticity during ongoing sensory processing.

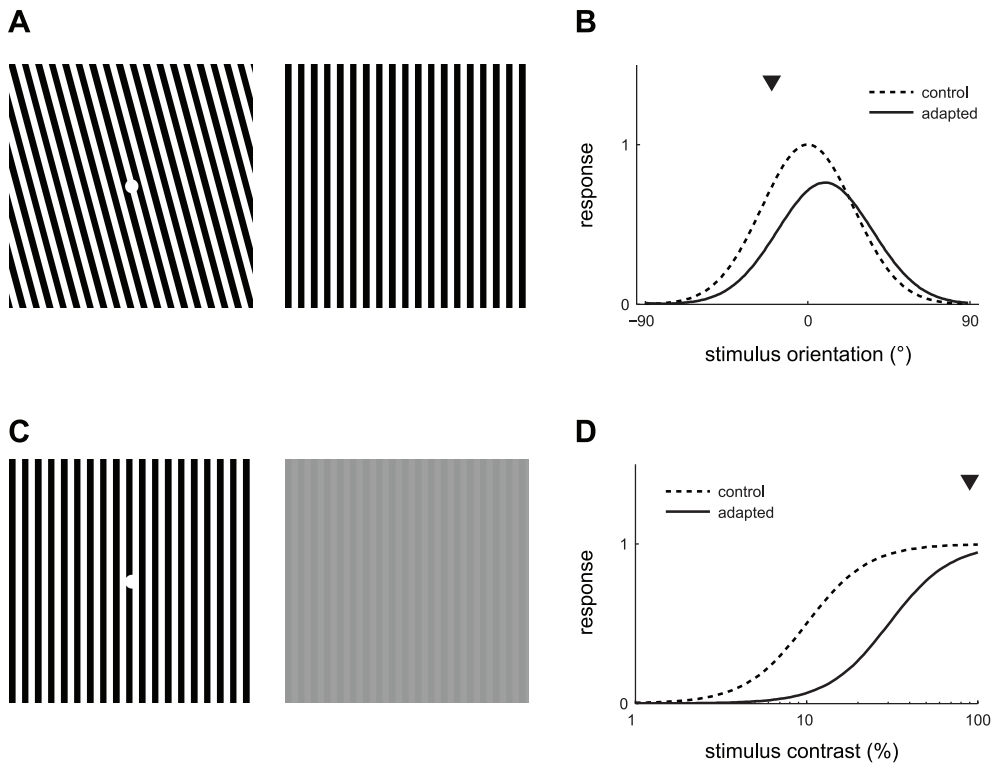


Figure 1.1 | Adaptation causes transient changes in visual perception (A, C) and in the responses of neurons in primary visual cortex (B, D). **A** | Tilt aftereffect. Fixate on the white dot inside the left image for 30 seconds, moving your eye slightly to avoid developing strong afterimages. Then quickly fixate the right image. The vertical lines should appear slightly tilted clockwise; this phenomenon is called the tilt aftereffect. **B** | Adaptation-induced plasticity in orientation tuning. Adaptation to an oriented grating (marked by the triangle) alters neuronal tuning curves. When the adapting orientation is close to a neuron's preferred orientation, the tuning curve broadens, its peak response decreases, and the curve shifts away from the adapting orientation. **C** | Perceptual reduction in apparent contrast. Fixate on the white dot inside the left image for 30 seconds, moving your eyes back and forth to prevent afterimages. Then quickly move your gaze to the right image. The low-contrast grating should briefly be invisible. **D** | Adaptation with a high contrast stimulus (triangle) reduces contrast sensitivity of individual neurons by shifting the contrast-response function to the right.

close to the adapting orientation (and sharpens for neurons with a preferred orientation far from the adapting orientation), and the preferred orientation of the neuron shifts repulsively away from the orientation of the adapting stimulus. These tuning curve changes show that adaptation is more than just "neuronal fatigue"; adaptation seems to reorganize tuning curves through a combination of suppression (on the flank close to the adapting stimulus) and enhancement (on the flank far away from the adapting stimulus), probably mediated through interactions of recurrently connected neurons. While the physiological mechanisms underlying contrast adaptation are relatively well understood, the exact site of orientation adaptation remains unknown.

One aim of this thesis is to study the physiological mechanisms that underlie orientation adaptation using an explicit model of orientation processing in V1. Specifically, we focus on a mechanistic model of the local cortical network and explore the hypothesis that depression of intracortical synapses (Abbott et al., 1997; Tsodyks and Markram, 1997; Varela et al., 1997) provides the substrate for orientation adaptation. Depression of intracortical synapses is consistent with a recent *in vivo* study in visual cortex (Boudreau and Ferster, 2005). Furthermore, both short-term synaptic depression and orientation adaptation occur at multiple time scales from tens of milliseconds to seconds (synaptic depression: Varela et al. 1997; Galarreta and Hestrin 1998; Zucker and Regehr 2002; orientation adaptation: Felsen et al. 2002; Müller et al. 1999; Dragoi et al. 2002, 2000). An intrinsic hyperpolarization of the presynaptic neurons might also contribute to orientation adaptation in V1 neurons (as in the case of contrast adaptation), but it occurs at a time scale of 30 seconds to minutes. We also want to bridge the gap between the physiologically observed tuning curve changes and the psychophysical tilt aftereffect. In order to link the neural responses to perception one needs to know how the rest of the brain "reads out" orientation from the responses of a population of V1 neurons (see below). It has been shown theoretically that, assuming an optimal decoder and independent neural responses, the reversible shifts observed in individual V1 neurons are in quantitative agreement with the tilt aftereffect (Jin et al., 2005). Here, we want to build on these results and explore the possibility that the adaptation effects emerge as a result of recurrent processing.

We use a bottom-up approach for modeling orientation tuning, and numerically simulate network models of primary visual cortex. An important aspect that we take into account is that response properties of individual neurons can only be understood in the context of the local circuits they are embedded in. However, the exact role that local circuits play in the computation of orientation selectivity is still debated. In a first step, we therefore develop a detailed two-dimensional network model that incorporates the structure of the orientation preference map and use recent anatomical and physiological data to constrain the recurrent interactions between a neuron and its local cortical

neighborhood. We then use this network to study how synaptic depression, caused by adaptation to orientation stimuli, influences recurrent processing.

Computational principles

While visual illusions like the tilt and the contrast aftereffects may seem like an anomaly in visual processing or even like a negative effect of adaptation, there is some evidence that they are actually the result of optimal processing. The benefits of adaptation are intuitive in the case of contrast coding, where adaptation shifts the contrast response curve towards higher contrast levels (Figure 1.1D). This can be understood as an adjustment of the limited dynamic range of the neuron to the range of contrast values observed in the recent past, in order to ensure an accurate representation for contrast values of contextual relevance. Theoretically, such an adjustment to changing contrast statistics is predicted by the efficient coding hypothesis (Adorján et al., 1999b; Schwabe and Obermayer, 2003). The efficient coding hypothesis (Barlow, 1961, 2001), grounded in information theory, states that sensory systems seek to provide an efficient representation of the natural environment. To achieve this goal, single sensory neurons should fully employ their information transmission capacity (this is called the information maximization principle or infomax, Linsker, 1988), and different neurons should operate as independent encoders by decorrelating their responses in order to reduce redundancy. In this way, also the tilt aftereffect (Figure 1.1A) is consistent with efficient coding, as has been shown in various modeling studies (Wainwright, 1999; Clifford et al., 2000, 2001; Schwartz et al., 2007). An abstract Bayesian inference model provides an alternative explanation; here the tilt aftereffect arises through the optimal combination of sensory observations with prior knowledge (Stocker and Simoncelli, 2006).

Orientation adaptation not only produces a perceptual bias (the tilt aftereffect), but also affects orientation discrimination. The discrimination performance is improved around the adapting orientation (Regan and Beverley, 1985) and the orthogonal orientation (Clifford et al., 2001), but is markedly impaired for angles of 10°-15° between adapter and test. These psychophysical effects need adaptation times between 5 sec and 1 min. Interestingly, the adaptive changes in orientation tuning occur at multiple time scales, from tens of milliseconds (Felsen et al., 2002) over hundreds of milliseconds (Müller et al., 1999; Dragoi et al., 2002) to several minutes (Dragoi et al., 2000). Adaptation might therefore continuously recalibrate neural responses even during natural viewing where monkeys (and humans) make saccadic eye movements several times per second. Indeed, Dragoi et al. (2002) showed that short-term adaptation to oriented gratings improves orientation discrimination for orthogonal orientations by sharpening neural tuning curves. Importantly, they also showed in studies of natural images (Dragoi et al.,

2002) and of eye movements (Dragoi and Sur, 2006) that saccades to distant image patches of largely dissimilar orientation structure are more likely than saccades to image patches with intermediate orientation differences.

Until now most experimental studies have investigated how adaptation can enhance the coding efficiency of an individual cortical neuron. However, perception relies on sensory information encoded in the activity of populations of neurons; thus, it is often difficult to link the results of these studies to a functional benefit for the organism. How individual neurons contribute to a population code, and how a population code can be interpreted, remains an active research topic (Deneve et al., 1999; Pouget et al., 2003; Clatworthy et al., 2003; Kang et al., 2004; Jazayeri and Movshon, 2006; Butts and Goldman, 2006). It is commonly accepted that correlations between neural responses can have a large impact on population coding (Averbeck et al., 2006), and in turn on the accuracy of a perceptual decision such as in a discrimination task (Cohen and Newsome, 2009). In a recent study, Gutnisky and Dragoi (2008) found that adaptation improves the accuracy of population coding by selectively reducing the mean and the variability of pairwise correlations between neurons. However, it might be necessary to take interactions between neurons beyond pairwise correlations into account to fully understand population coding (Roudi et al., 2009). Furthermore, perception involves processing distributed across multiple interconnected cortical areas (de Lafuente and Romo, 2006), where the responses of neurons are modulated by signals coming from other neural populations, making it difficult to discern stimulus-driven, adaptive effects from top-down, behavioral effects such as attention.

An alternative to studying adaptation in large-scale sensory systems that rely on a distributed population code is to use the simple sensory networks of invertebrates, which involve representations that are often encoded in the activity of single neurons and not heavily influenced by feedback signals (see for instance Brenner et al., 2000; Fairhall et al., 2001). Part of this thesis describes our use of this approach, where we study the computational principles of stimulus coding in the auditory pathway of crickets. Using a top-down approach, we explore how adaptation changes the response properties of the AN2 neuron and its information transmission capacity. The AN2 neuron is the only neuron of a local network that conveys information about high frequency sounds to the cricket's brain; it is thus a bottleneck for information transmission. Interestingly, the sound-intensity response curve of the AN2 neuron is reminiscent of the contrast response curve of V1 neurons (Figure 1.1D). However, in contrast to visual processing in V1, auditory processing at the stage of the AN2 neuron is purely feed-forward; adaptation is thus driven by the stimulus only, and not by task-dependent top-down processes. This, together with the fact that the AN2 neuron has a clear behavioral role in evasion of ultrasonic sounds coming from bats, makes the AN2 an ideal model for studying

the computational principles underlying sensory adaptation. Combining electrophysiological and modeling techniques, we ask which computational principle – infomax (Linsker, 1988) or selective coding (Sobel and Tank, 1994) – can predict the responses of this sensory neuron to a carefully designed stimulus.

Apart from the common topic – the generation and adaptation of sensory representations – the different parts of this thesis also share a common methodology. For modeling orientation tuning in V1, we use model-based data analysis in order to assess a continuum of network models, that encompasses the full range from feed-forward via inhibition-and excitation dominated models to models with excitation and inhibition in balance. Similarly, to shed light on the mechanisms underlying pattern adaptation in V1, we again explore the full range of synaptic depression parameters. Thus, in contrast to traditional modeling approaches[2], we determine the complete space of models able to account for the experimental data, rather than merely demonstrating one model to be compatible with the data set. In the same spirit, we use a Bayesian approach for quantifying how adaptation changes sound-intensity response curves in the cricket auditory system. Bayesian data analysis yields the joint posterior distribution of the model parameters which allows us to calculate precise confidence limits of adaptive changes as well as of derived quantities such as mutual information between stimulus and response. This is again in contrast to traditional modeling approaches, which preclude such an analysis because only one model is fit to the experimental data.

[2] Similar approaches have been used to explore the conductance space of detailed single cell models or small networks composed of a few a neurons (Marder et al., 2007), but applications to larger networks are rare.

1.2. Addressed questions and outline

A large part of the work presented in this thesis deals with orientation processing in the primary visual cortex. In order to provide the necessary background, we start with an overview of the early visual system in Chapter 2. Beyond a mere summary of the anatomy and physiology, which can be found in many good textbooks, this chapter also discusses previous functional and mechanistic models for V1 receptive fields and emphasizes experimental findings concerning the relationship between single neuron responses and the local cortical circuit.

The original research is presented in Chapter 3 to Chapter 6. Each chapter is self-contained and includes an introduction to the discussed subject, a review of previous work, and a concluding remark. Details of the developed network models can be found in the Appendix. In particular we address the following questions:

1. Can we infer the *mechanism* underlying the computation of orientation selectivity from experimentally-characterized relationships between the local cortical neighborhood and the input and response properties of single neurons?

2. Does short-term synaptic plasticity provide a *mechanism* that can account for the physiological (tuning curve changes) and perceptual (tilt aftereffect) consequences of adaptation in the visual system?

3. Which *computational principle* can explain adaptation in a first order auditory interneuron in the cricket?

The first question is considered in Chapter 3 and Chapter 4, where we investigate the mechanisms of orientation tuning in primary visual cortex (V1). V1 performs a remarkable transformation of the visual information it receives from the eyes. Cortical neurons are sensitive to orientation in a way that their afferent inputs are not, which means that they somehow compute this feature selectivity. The mechanism underlying this computation has been discussed extensively, giving rise to two different views. In feed-forward models, orientation selectivity arises primarily through the convergence of thalamocortical inputs, while in recurrent models it arises through a weak afferent bias and strong intracortical interactions (see Chapter 2 for a discussion of these models). The question how orientation tuning is computed is of broad interest because, given the similarity of neocortical circuits (Creutzfeldt, 1977), similar mechanisms might be employed in different cortical areas for different tasks.

In *Chapter 3*, which is based on Wimmer et al. (2009), we argue that experimental data concerning connectivity patterns and neural responses with respect to the orientation preference map provides constraints to distinguish between various theories of

orientation tuning. We investigate the full range of orientation selectivity models from feed-forward via inhibition- and excitation-dominated to balanced recurrent computation. Since neurons in primary visual cortex are embedded in an orientation preference map with a non-uniform layout, their short-range lateral inputs are determined by the position in the orientation map and by the strengths of these lateral inputs (which is dependent upon the relative locations of the neurons providing them). While data is available on the range of these lateral interactions, their strengths are unknown. Here, we use recent experimental data from intracellular recordings (Mariño et al., 2005) to constrain a Hodgkin-Huxley network model and to pin down the "operating regime" of V1, i. e., the relative strength of afferent excitatory, lateral excitatory and lateral inhibitory connections.

In *Chapter 4*, which is based on Schummers et al. (2007) and Stimberg et al. (2009), we extend the modeling study from Chapter 3 to the temporal domain. Using temporal filters to describe the dynamics of the afferent input to each model cell in the Hodgkin-Huxley network model, we simulated "reverse correlation" experiments. Single-unit recordings from cat primary visual cortex under this paradigm (Schummers et al., 2007) showed that neurons close to pinwheel centers and neurons in orientation domains exhibit a similar time course in their averaged responses, but differences in their inter-cell variability. These characteristics of the temporal dynamics of V1 can only be reproduced in the Hodgkin-Huxley network model operating in a balanced recurrent regime. One interesting consequence of this operating regime is that the network is susceptible to small perturbations of excitation or inhibition. Together, Chapter 3 and Chapter 4 highlight the importance of knowledge about the local network connectivity in understanding and modeling orientation selectivity in the visual cortex. This work, while not dealing with the topic of adaptation itself, provides a crucial prerequisite for studying the effects of adaptation on orientation tuning in Chapter 5.

In *Chapter 5*, we use a mean-field version of the Hodgkin-Huxley network model to investigate the mechanisms through which adaptation changes V1's representation of sensory stimuli. One motivation for this study was the physiological evidence supporting the idea that synaptic depression might underlie the experimentally-observed adaptive changes. We also wanted to test our prediction that even a small modification of the balance between excitation and inhibition in the highly recurrent network regime makes large variations in neuronal responses possible (see Chapter 3). Thus, we explore whether short-term synaptic plasticity evokes adaptive changes in orientation tuning at the neuronal level and whether these changes correctly predict the magnitude of the perceptual tilt aftereffect. This is in contrast to previous modeling studies that have ignored empirical constraints or have focused exclusively either on the perceptual (tilt aftereffect) or physiological (tuning curve changes) effects associated with orientation

adaptation. Our model exploration gives insights into how specific combinations of synaptic plasticity at different sites (excitatory and inhibitory synapses with excitatory and inhibitory target cells) contribute to the reshaping of orientation tuning curves. The set of models that is in quantitative agreement with the experimental data correctly predicts that adaptation-induced changes are more pronounced close to pinwheel centers (cf. Dragoi et al., 2001) than in orientation domains.

In *Chapter 6*, which is based on Wimmer et al. (2008), we focus on the computational principles that underlie adaptation in the cricket auditory system. Specifically, we designed a sound stimulus for which two competing hypotheses – infomax (a form of the efficient coding principle) and selective coding (a form of figure-background separation) – predict different AN2 neuron tuning curve changes. In other words, we sought to distinguish if the AN2 neuron is optimized for transmitting as much information as possible or for preprocessing sensory information in order to transmit only behaviorally relevant signals. Even if one of these theories is correct, architectural and physiological constraints might prevent a system from achieving the optimum predicted by the "correct" theory. However, because we compare the improvement in information transmission according to the two theories, we can discern which is more appropriate regardless of whether optimal performance is observed. Finally, we argue that that this approach is of broad interest because it is highly applicable to other sensory systems.

2. Background

2.1. The early visual system

The visual system can be loosely defined as all the parts of the brain that in some way analyze information about the visual world, usually conveyed through the eyes. This can, for example, be determined through the study of neuronal responses in model organisms – e. g., macaque monkey and cat – by testing for correlations between visual stimuli and neuronal activity. By studying these correlations, one obtains what is commonly called a receptive field: a description of the type of stimuli, including location, temporal and spatial structure, and possibly other aspects that elicits a neuronal response.

The functional approach to studying the visual system is supplemented by an anatomical approach. The connectivity pattern between different parts of the visual system can be studied in tracer studies, and based on morphological criteria, different areas of the visual cortex can be distinguished. Findings in both these approaches complement each other well; areas identified using anatomical techniques contain neurons with similar functional properties. Moreover, a widespread principle of organization of early areas in the visual system is that of retinotopy: The relative position of neurons to each other within an area can be mapped smoothly onto their relative receptive field regions in the visual field (or the corresponding regions of the retina). Generally, the functional and anatomical descriptions of visual areas tend to agree.

Exploration of the functional properties of neurons in different visual areas has led to the widespread view that processing in the visual system is hierarchically organized. The receptive fields of areas close (in terms of neuronal connections) to the eyes, i. e., at the "input" end of the system, are found to be small and simple in structure, providing information about a small proportion of the visual field. Visual areas that receive information from these early processing stages show larger, more complex receptive fields, areas receiving input from these show an even higher degree in complexity and abstraction (Lennie, 2003). This hierarchical view has been strengthened by findings of consistent connectivity patterns, and anatomical and functional studies have arrived at convergent descriptions of this processing hierarchy (Felleman and Essen, 1991).

The hierarchical view is useful for a structural characterization of the visual processing stream, whose subcortical pathways are depicted in Figure 2.1. It starts with light falling on the retina of the eyes. From the retina, information is transmitted by the retinal

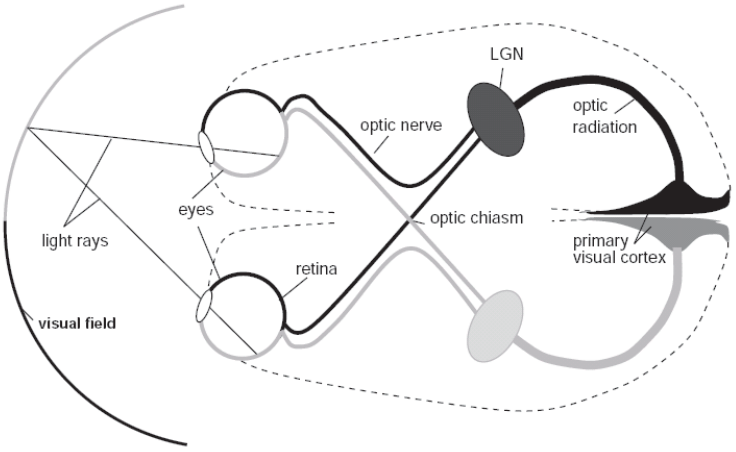


Figure 2.1 | Overview of the peripheral and subcortical visual pathways in higher mammals. Visual information originating from the left and right visual field is processed in the right and left cortical hemisphere.

ganglion cells, whose main projections innervate the lateral geniculate nucleus (LGN). Half of these projections of each eye cross at the optic chiasm to the other side of the LGN, such that for each hemifield, the representations of both eyes are mapped to the contralateral hemisphere. The main projections from LGN go to the primary visual cortex in the occipital lobe. It is also referred to as striate cortex, area 17, or V1 (the latter usually for primates; however, V1 will be used for the rest of the thesis, also for cat area 17). V1 is the first cortical visual area. While up to this point information processing is completely monocular, here for the first time, information of the two eyes is combined. The processing stream branches out at this stage. Two main streams are distinguished, the dorsal stream and the ventral stream. The dorsal stream is thought to be responsible for the visual control of action including, for example, the analysis of motion. Areas of the dorsal stream include MT and MST, where cells show motion sensitivity and egocentric coding of the visual field. The ventral stream is thought to be involved in feature processing and perception. Areas in this stream include V3, V4, and IT, where cells have larger receptive fields, are sensitive to color, shapes, possibly even objects and faces (IT).

The next sections will give an overview over the properties of the first three areas in this processing stream, the retina, LGN, and V1. This introduction is thought as a primer for the rest of this thesis; a detailed description of the visual pathway can be found in textbooks such as Kandel et al. (2000) or Chalupa and Werner (2003).

The retina

The retina is at the periphery of the visual system. The eye's optical system projects an image onto the retina, where luminance is translated into neuronal impulses. This conversion is performed by a densely packed array of photoreceptors. The human retina contains approximately 10^8 photoreceptors – rods and cones – and is responsible for a significant amount of processing in order to compress the signal into 10^6 fibers of the optic nerve (Meister and Berry, 1999). Rods are responsible for vision at low light levels, cones at greater light levels. In humans and some primates (including macaque), cones are subdivided into three classes, according to their wavelength sensitivity. The signals from the photoreceptors are processed in the plexiform layers of the retina and finally carried by the axons of the ganglion cells, in the form of action potentials, out of the retina.

A major product of processing in the retina is an adaptation of the signals coming from its afferents to the mean luminance in the visual world (Shapley and Enroth-Cugell, 1984). This is necessary in order to accommodate the working range of several orders of magnitude of luminance in the limited response range of visual neurons. As a result, retinal ganglion cells respond to contrast, that is, relative luminance differences or ratios, rather than to luminance itself.

The receptive field of retinal ganglion cells consists of a center and a larger surround that interact subtractively. An ON-center cell is one in which a bright light on the center of the receptive field evokes a positive response (an increase in firing rate) and a dark light evokes a negative response. An OFF-center cell does the opposite. Stimulation in different parts of these receptive fields has a linear effect on the firing rate of the cells, which can be modeled using the difference of two Gaussian functions. Combining this linear filter with a static nonlinearity and a Poisson generator, one obtains the so called linear–nonlinear–Poisson (LNP) model. Although the LNP model ignors all of the complexity of retinal circuitry it has been very successful in describing the responses of both the sustained (parvocellular pathway) and transient (magnocellular pathway) cells (Carandini et al., 2005).

The lateral geniculate nucleus (LGN)

The lateral geniculate nucleus is part of the thalamus. In primates, it consists of six major layers of retinotopically organized visually responsive neurons, two magnocellular layers (for motion processing) and four parvocellular layers (for color/form processing). In cat, it has three layers. The strongest retinal input to the LGN originates from ganglion cells of the X/parvocellular type and of the Y/magnocellular type. Additional input to LGN

relay cells originates from other geniculate neurons, from subcortical structures, and from cortex. The thalamus has traditionally been viewed as a simple relay station that passes on messages to the cortex largely as they come in. Evidence is now beginning to accumulate that it may actually play a much more dynamic role in the processing and transmission of information (see Sherman and Guillery, 2000).

The neurons in the LGN show a center-surround organization of their receptive fields, like in retinal ganglion cells. While the spatial properties of the LGN neurons resemble those of their retinal counterparts closely, the temporal properties have been shown to differ. In particular, two types of neurons have been found, lagged and non-lagged cells (Wolfe and Palmer, 1998; Cai et al., 1997). As the name suggests, the difference between the cells lies in their temporal response properties, lagged cells showing a temporally offset response to stimulation.

The primary visual cortex (V1)

V1 (or primary visual cortex, striate cortex, or area 17) is located in the occipital lobe of the brain. The majority of visual input to the cortex passes through V1, and it is mostly received from the LGN. Thus, in a hierarchical view of the visual system, V1 is two levels above the retina, one level above the LGN, and at the bottom of the hierarchy within cortex (Felleman and Essen, 1991). The outputs of V1 feed into both, the dorsal and the ventral processing stream, making V1 the main provider of visual information for motion processing as well as feature processing. It is the last stage where one may still find a representation of the full visual scene before processing becomes specialized, giving the neuronal transformations performed in V1 additional significance.

The overall organization is, like in the LGN, retinotopic, and it is the first stage at which information coming from the two eyes is combined. The monocular input into V1 is processed in separate ocular dominance columns, elongated stripes distributed across the surface of V1. This ordered mapping of eye preference across the surface of visual cortex gives rise to the so called ocular dominance map. A given V1 neuron not only has a preference for stimuli being shown to the right versus the left eye and at a particular location in visual space, but also responds preferentially to stimuli with a particular orientation, a particular spatial frequency (the spacing of the bars for a grating stimulus), a particular direction of movement, and other features (for a review see Lennie and Movshon, 2005). As we move across the cortical surface, we find interrelated maps of these neuronal response properties (Obermayer and Blasdel, 1993; Hübener et al., 1997; Yu et al., 2005). The underlying organizing principle is maximizing the coverage (i. e., the cortex' ability to respond to all possible stimulus properties at every point in visual space) while providing a smooth representation of the feature space.

The organization of V1 in depth also shows regularities. Like all neocortical areas, it consists of six main laminae. The main input receiving layer is layer 4, where most projections from LGN terminate. It is vital to remember that the LGN input into V1 neurons represents only a small proportion of their input connections. The consensus figure from the anatomy is that the thalamus provides around 5 % of the excitatory synapses in its main target layer (Douglas and Martin, 2007). Most of the connections in layer 4 are intrinsic connections originating within V1, such as layer 6 pyramidal cells and spiny stellate cells. The main route of information flow in the local processing is from layer 4 through layers 2 + 3 to layer 5 and then layer 6, which sends projections back to thalamus. The functional significance of these layers has still not been established conclusively. See Douglas and Martin (2004), Binzegger et al. (2004) for a detailed description of this cortical microcircuit and Hirsch and Martinez (2006) for a review of the laminar differences in sensory processing.

Mainly the neurons in the superficial layers 2/3 also communicate via long-range lateral connections over distances of up to several millimeters. Another source of inputs to V1 neurons comes from inter-areal feedback connections between V1 and extrastriate areas. In contrast to the local connections, that show an isotropic connection pattern, long-range and feedback connections predominantly contact neurons with a similar preferred orientation. A review of the anatomy of V1 can be found in Lund et al. (2003).

2.2. Orientation selectivity in V1

Functional models of V1 receptive fields

In a seminal study of the responses of V1 neurons, Hubel and Wiesel established, initially in cat (Hubel and Wiesel, 1962) and later in macaque monkey (Hubel and Wiesel, 1968) that neurons in V1 are selective for stimulus size and orientation. The orientation preferences of V1 neurons are usually mapped using gratings or bars of various orientations. The neuronal response varies smoothly as a function of orientation, giving rise to the so called orientation tuning curve (Figure 2.2D, E). Classically, two different types of neurons are distinguished: The receptive fields of *simple cells* are characterized by discrete parallel elongated regions which show either ON or OFF responses; *complex cells* do not have such separate excitatory and inhibitory regions, but rather respond to both ON and OFF stimulation. Due to this different structure, drifting gratings elicit different response patterns in the two idealized cell models. Simple cells will respond only when the grating approximately covers the ON and OFF regions with the correct polarity. They are thus sensitive to the spatial phase of the presented grating, and their response to the drifting grating is modulated with the temporal frequency of the drifting (Movshon

et al., 1978b). Complex cells, on the other hand, do not show the discrete organization of the activating regions. Thus, their responses will be more uniform throughout the drifting (Movshon et al., 1978a). This property of simple and complex cells is often used to identify them.

The receptive field of a simple cell can be modeled using a linear spatiotemporal filter (e. g., a Gabor filter) followed by an output-nonlinearity, i. e., an LNP model as described above for retinal ganglion cells. Complex cells are commonly modeled with two phase-shifted filters whose outputs are squared and summed (the energy model, Adelson and Bergen, 1985; Spitzer and Hochstein, 1985) before being passed through the threshold nonlinearity. This is what could be called the "old standard model" that is being replaced because it fails to account for important, experimentally observed receptive field properties, such as the change in gain with the average illumination of the receptive field or the suppression of responses by stimuli presented outside the receptive field.

These shortcomings led to the development of a new, more comprehensive, standard model for V1 cells (Rust and Movshon, 2005; Carandini et al., 2005) incorporating multiple additional elements. These include multiple linear input filters whose output is combined nonlinearly and subject to various gain control mechanisms (e. g., luminance gain, contrast gain) that change responses depending on the combination of stimuli being presented. Further components of the model capture nonlinearities caused by the stimulation history (e. g., contrast adaptation) and spatial context effects (e. g., suppression caused by stimuli outside the classical receptive field). Current versions of this model capture approximately 35 % of the explainable variance of V1 neurons in response to natural stimuli (Carandini et al., 2005).

Mechanistic models of orientation tuning

To this point, we have specified a receptive field in terms of computations performed on an image. Such a purely functional description of receptive fields can describe the mapping from the image space onto the activity of a single neuron; in other words, it can tell us *what* the visual system is computing. We now move on to a more mechanistic description of V1, one that serves for asking the question: *how* does the visual system do it? Mechanistic models incorporate findings from anatomy and biophysics, and often try to describe not only single cells, but local networks.

Orientation selectivity is a property that "emerges" in V1: cortical neurons are selective for stimulus orientation while the receptive fields of their thalamic input are almost circular and therefore relatively unselective. The question how orientation selectivity emerges has provoked much controversy over the years because orientation selectivity

can be seen as the prototypical example of a cortical computation, and certainly the one that is best characterized experimentally. Here we briefly describe the basic ideas behind the models and their weaknesses that have driven the debate. See Ferster and Miller (2000); Sompolinsky and Shapley (1997); Teich and Qian (2006); Priebe and Ferster (2008) for detailed reviews of various models and relevant experimental data.

Hubel and Wiesel (1962) proposed a feed-forward model in which orientation selectivity arises due to an appropriate wiring from the LGN to V1, ensuring that V1 cells receive inputs from aligned ON and OFF subregions. There is substantial experimental support for this assumption (Reid and Alonso, 1995; Ferster et al., 1996; Alonso et al., 2001). Incorporating the intrinsic nonlinear properties of cortical neurons, Finn et al. (2007) also reconciled the feed-forward idea with the experimental finding that orientation tuning width is independent of stimulus contrast. However, there are also features of orientation tuning that this model cannot explain, e. g., why a neuron's orientation selectivity can be reduced through pharmacological inactivation of a cortical site in the vicinity (Crook et al., 1997).

Recurrent models (sometimes also called feedback models) offer a completely different view of how orientation selectivity arises. They rely on intracortical excitation, inhibition, or a combination of both that refine selectivity relative to a bias provided by thalamocortical inputs. Inhibition models state that intracortical inhibition suppresses non-preferred responses and hereby sharpens weakly tuned feed-forward input (Wörgötter and Koch, 1991; Troyer et al., 1998; McLaughlin et al., 2000; Wielaard et al., 2001). It is a challenge for models of this type to explain why an intracellular blockade of inhibition has a negligible effect on the sharpness of orientation tuning (Nelson et al., 1994). Furthermore, measurements of inhibitory synaptic inputs to V1 neurons have shown that inhibition tends to be strongest at the preferred orientation rather than at the orthogonal orientation (Monier et al., 2003; Mariño et al., 2005). Another hypothesis is that strong recurrent excitation is necessary to generate orientation tuning (Douglas et al., 1995; Ben-Yishai et al., 1995; Somers et al., 1995; Adorján et al., 1999a; Mariño et al., 2005). Depending on the model, recurrent inhibition is unselective or sharply tuned, only necessary to generate stable network activity or exquisitely balanced with excitation. In the extreme case, the so called marginal phase (Ben-Yishai et al., 1995; Adorján et al., 1999a), orientation selectivity is determined exclusively by the pattern of recurrent connections and arises from almost untuned afferent input. Responses of the marginal phase model to multiple simultaneously presented stimuli are inconsistent with experimental data (Carandini and Ringach, 1997). Furthermore, silencing cortical responses through cooling does not abolish the selectivity of orientation tuning (Ferster et al., 1996), but rather leads to a reduction in the response strength (Girardin and Martin, 2009).

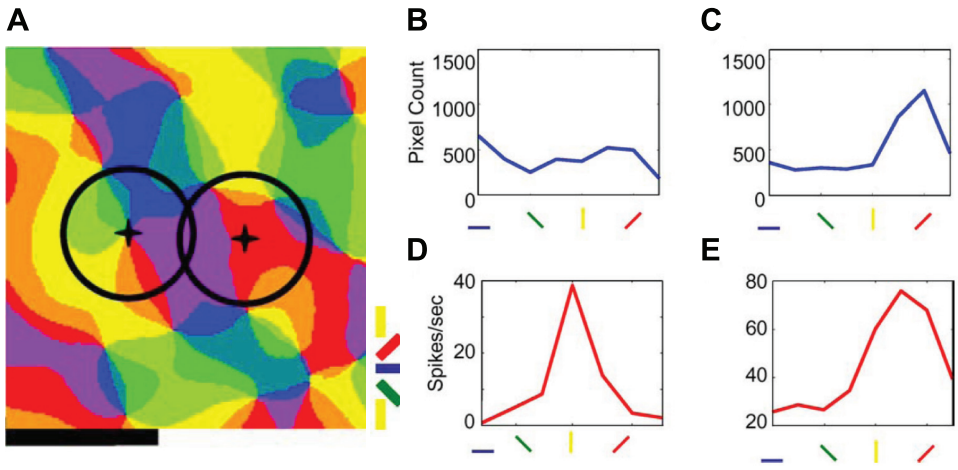


Figure 2.2 | Relationship between the orientation preference map and responses of V1 neurons. **A** | Part of an orientation map obtained from cat V1 with optical imaging of intrinsic signal responses to drifting gratings. The color of each pixel denotes the preferred orientation of cells located at that pixel (colored bars at the bottom right). The horizontal scale bar represents 1 mm. The two circles with radius 400 µm denote the regions of the local connections to the points at the center of each circle. **B, C** | Histograms of the number of pixels representing each orientation that fall within the two circles shown in (A), for the left circle that is centered on a pinwheel center (B), and for the right circle that is centered on an orientation domain (C). **D, E** | Orientation tuning curves of V1 cells from the two locations marked by crosses in (A). Figure adapted from Schummers et al. (2004).

Why has it been so difficult to distinguish between the feed-forward and the different variants of the recurrent hypothesis? Possibly the problem of generating orientation selectivity is underconstrained. Many models simplified the cortical architecture to a one-dimensional "hypercolumn" and exclusively focused on explaining the output, i. e., the spiking activity, of cortical cells. There is, however, a large variability in the synaptic inputs of single neurons (Monier et al., 2003), possibly due to laminar differences (Ringach et al., 2002; Hirsch and Martinez, 2006) and inhomogeneities of the local circuit (Schummers et al., 2004). To pin down the potential mechanisms underlying orientation tuning, it seems thus necessary to impose further constraints on the models by taking into account anatomical and physiological data.

The orientation map and local circuits

It has proven difficult to link functional properties of individual neurons to the underlying neuronal circuits. One way to bridge this gap are variations in the functional connectivity provided through the cortical map structure. For example, the position of a

neuron relative to the orientation map influences the lateral inputs that a neuron receives. We will show in Chapter 3 and Chapter 4 that this relation of local computations to the orientation map can be used to constrain the space of possible orientation tuning models. Here, we give a short overview of the local connectivity and its consequences for local computations.

The local circuit provides the major source of inputs to a V1 neuron (Binzegger et al., 2004; Thomson and Lamy, 2007). Anatomical data shows that local excitatory and inhibitory connections are spatially isotropic with a radius of influence of approximately 250 μm (Yousef et al., 2001; Mariño et al., 2005). Thus, the excitatory and inhibitory inputs that a neuron receives will strongly depend on the local neighborhood. Depending on the neuron's position in the orientation preference map, this neighborhood can be very different (Figure 2.2A). In one extreme, there are regions in the map, close to the singularities (pinwheel centers), where neurons with most or all of the preferred orientations are represented in a small neighborhood. In the other extreme, there are regions, where one particular preferred orientation dominates and only varies slowly with location (orientation domains). Consequently, neurons near pinwheel centers have local connections with neurons of a wide range of orientation preferences; neurons far from pinwheel centers, on the other hand, mainly connect to neurons sharing similar orientation preference. This is quantified by the histograms of neighboring pixels in the orientation map in Figure 2.2B, C. These pixel counts are of course only an indirect measure of the inputs to a V1 neuron. However, Schummers et al. (2002) and Mariño et al. (2005) experimentally measured the relationship between orientation specificity of the surrounding local cortical network (calculated from the pixels in the orientation map) and the orientation specificity of the subthreshold inputs to cat V1 neurons, by combining optical imaging of orientation maps with intracellular recordings. They found that map position is strongly correlated with the tuning of the membrane potential, the excitatory and the inhibitory inputs. Interestingly, while the tuning of the total recurrent input is broader for pinwheel cells the spike responses are nevertheless highly selective, similar to orientation domain neurons (Figure 2.2D, E, see also Maldonado et al., 1997; Dragoi et al., 2001; Mariño et al., 2005; Ohki et al., 2006).

3. Operating regimes for computing orientation selectivity

This chapter is based on Wimmer et al. (2009). The work presented in this chapter was done in close collaboration with Marcel Stimberg.

Abstract

In this chapter, we investigate the computational role of the local recurrent network in primary visual cortex. To address this issue, we analyze intracellular recording data of cat V1, which combine measuring the tuning of a range of neuronal properties with a precise localization of the recording sites in the orientation preference map. For the analysis, we consider a network model of Hodgkin-Huxley type neurons arranged according to a biologically plausible two-dimensional topographic orientation preference map. We then systematically vary the strength of the recurrent excitation and inhibition relative to the strength of the afferent input. Each parametrization gives rise to a different model instance for which the tuning of model neurons at different locations of the orientation map is compared to the experimentally measured orientation tuning of membrane potential, spike output, excitatory, and inhibitory conductances. A quantitative analysis shows that the data provides strong evidence for a network model in which the afferent input is dominated by strong, balanced contributions of recurrent excitation and inhibition. This recurrent regime is close to a regime of "instability", where strong, self-sustained activity of the network occurs. The firing rate of neurons in the best-fitting network is particularly sensitive to modulation, which could be one of the functional benefits of a network operating in this particular regime.

3.1. Introduction

One of the major tasks of primary visual cortex (V1) is to compute a useful representation of the environment. Orientation tuning in V1 has long served as a paradigmatic example of a cortical computation because the thalamic cells providing afferent input to the visual cortex lack orientation selectivity while the cortical cells are orientation-selective. Early feed-forward models (Hubel and Wiesel, 1962), combining the center-surround receptive fields of lateral geniculate nucleus to give rise to orientation selectivity, have been shown to be over-simplistic. Nonetheless, a debate remains regarding the relative contribution of afferent and recurrent excitatory and inhibitory influences to the emergence of

orientation tuning (Reid and Alonso, 1996; Sompolinsky and Shapley, 1997; Ferster and Miller, 2000; Ringach et al., 2003; Finn et al., 2007). Modeling studies (for a review see e. g., Teich and Qian, 2006) have also shown that information processing in cortex changes dramatically with this "cortical operating regime", i. e., depending on the relative strengths of the afferent and the different recurrent inputs (Somers et al., 1995; Ben-Yishai et al., 1995; Hansel and Sompolinsky, 1996; Troyer et al., 1998; McLaughlin et al., 2000; Kang et al., 2003). The wide range of models operating in different regimes that are discussed in the literature are an indication that models of V1 orientation selectivity are under-constrained.

Recently, experimental studies have investigated how a cell's orientation tuning depends on its position in the orientation preference map (Maldonado et al., 1997; Schummers et al., 2002; Mariño et al., 2005; Ohki et al., 2006; Nauhaus et al., 2008). Experimental findings concerning the relationship between orientation selectivity of the spike output and the synaptic inputs are presented in Chapter 2; an experimentally measured orientation map is shown in Figure 2.2. Here, we assess whether the specific location dependence of the tuning of internal neuronal properties can provide sufficient constraints to determine the corresponding cortical operating regime. The data originates from intracellular recordings of cat V1 (Mariño et al., 2005), combined with optical imaging in order to determine the position of the recording sites within the orientation preference map. This allowed to measure, *in vivo*, the output (firing rate) of neurons, the input (excitatory and inhibitory conductances) and a state variable (membrane potential) as a function of the position in the orientation map. Figure 3.1 shows the experimentally observed tuning strength of each of these properties depending on the distribution of orientation selective cells in the neighborhood of each neuron. The x-axis in each of the subplots denotes the orientation selectivity index of the map (map OSI; see Appendix A) for each of the recorded cells, which is a measure for a cell's position in the orientation map varying from close to a pinwheel center (towards the left end of the axis), where the neighboring neurons show a range of orientation preferences, or within an iso-orientation domain (towards the right end of the axis), where the neighboring neurons have very similar orientation preferences. Each y-axis denotes the tuning of the individual properties measured in the experiment, quantified by the orientation selectivity index (OSI; see Appendix A), ranging from 0 (unselective) to 1 (perfectly selective). The tuning of the membrane potential (V_m) as well as the tuning of the total excitatory (g_e) and inhibitory (g_i) conductances vary strongly with map location. Specifically, the conductances and the membrane potential are sharper tuned for neurons within an iso-orientation domain, where the neighboring neurons have very similar orientation preferences, as compared to neurons close to a pinwheel center, where the neighboring neurons show a broad range of orientation preferences. However, the firing rate (f) of

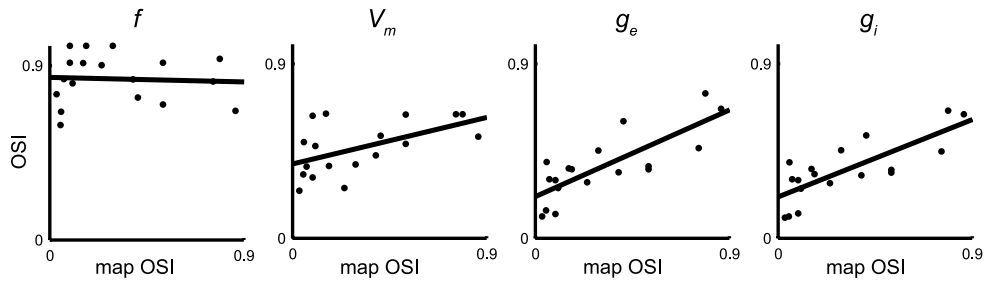


Figure 3.1 | Dependence of the synaptic inputs and the responses of V1 cells on the position in the orientation preference map. Variation of the orientation selectivity indices (OSI, cf. Appendix A) of the firing rate (f), the average membrane potential (V_m), and the excitatory (g_e) and inhibitory (g_i) input conductances of neurons in cat V1 with the map OSI (the orientation selectivity index of the orientation map at the location of the measured neuron). Dots indicate the experimentally measured values from 18 cells (Mariño et al., 2005). Solid lines show the result of a linear regression. The slopes (values ± 95 % confidence interval) are -0.02 ± 0.24 (f), 0.27 ± 0.22 (V_m), 0.49 ± 0.20 (g_e), 0.44 ± 0.19 (g_i).

V1 neurons is highly selective near pinwheel centers and in orientation domains (see also Maldonado et al., 1997; Ohki et al., 2006).

Here, we show that this specific map-location dependence of neuronal properties imposes strong constraints on the operating regime of a generic network composed of Hodgkin-Huxley type model neurons. The model takes into account that the lateral inputs a cell receives are determined (i) by the position in the orientation map and (ii) by the way that synaptic inputs are pooled across the map. The synaptic pooling radius has been shown experimentally to be independent of map location (Mariño et al., 2005), resulting in essentially different local recurrent networks depending on whether the neighborhood is made up of neurons with similar preferred orientation, such as in an iso-orientation domain, or is highly non-uniform, such as close to a pinwheel. The strength of lateral connections, on the other hand, is unknown. Mariño et al. (2005) have shown that their data is compatible with a model showing strong recurrent excitation and inhibition. However, this approach cannot rule out alternative explanations accounting for the emergence of orientation tuning in V1. Here, we systematically explore the model space, varying the strength of the recurrent excitation and inhibition. This allows us to test the full range of models, including feed-forward-, inhibition- and excitation-dominated models as well as balanced recurrent models, and to determine those that are compatible with the observed data.

3.2. Network model and quantitative evaluation

The Hodgkin-Huxley network model

The network model consists of Hodgkin-Huxley type point neurons (Destexhe and Paré, 1999; Destexhe et al., 2001). Synaptic conductances were modeled as originating from $GABA_A$-, AMPA-, and NMDA-like receptors (Destexhe et al., 1998); additional conductances represent background activity (Ornstein-Uhlenbeck conductance noise). Orientation preferences were assigned according to the location in an artificial orientation map (see Appendix B). This orientation map was calibrated such that the pinwheel distance and the spread of recurrent connections matches experimental data (Mariño et al., 2005). The network was composed of 50×50 excitatory and $1/3 \times (50 \times 50)$ inhibitory neurons, and corresponds to a patch of cortex $1.56 \times 1.56\,\mathrm{mm}^2$ in size. In order to avoid boundary effects, we used periodic boundary conditions. Intracortical synaptic connections were modeled as spatially isotropic with the same radial profile ($\sigma_E = \sigma_I = 125\,\mu\mathrm{m}$) for excitatory and inhibitory cells. Afferent inputs to excitatory and inhibitory cortical cells were moderately tuned (circular Gaussian tuning function with $\sigma_{\mathrm{Aff}} = 27.5°$) and modeled as Poisson spike trains. In order to calculate the orientation tuning of V_{m}, g_{e}, g_{i}, and f, an input spike train with a constant rate was applied and the network was simulated for 1.5 seconds with 0.25 ms resolution (usually, the network settled into a steady state after a few hundred milliseconds). For analyzing the tuning properties, we calculated the firing rate, the average membrane potential, and the average total conductances for every cell in an interval between 0.5 s and 1.5 s. A detailed description of the Hodgkin-Huxley network model can be found in Appendix C.

Quantitative evaluation of network models

In order to compare the model results to the experimental data, we analyze the orientation tuning properties of V_{m}, g_{e}, g_{i}, and f for each neuron in the simulated network model using the orientation selectivity index (see Appendix A). In addition, the map OSI was used to characterize the sharpness of the recurrent input a cell receives based on the artificial orientation preference map (see Appendix B). The dependence of each tuning property on the local map OSI was then described by a linear regression line using the least squares method (i. e., an OSI-OSI plot as shown in Figure 3.1 for the experimental data). These linear fits provided a good description of the relationship between map OSI and the tuning of the neuronal properties in the simulations (mean squared deviation around the regression lines was typically below 0.0025 and never above a value of 0.015) as well as in the experimental data (mean squared deviation was between 0.009 (g_{i}) and 0.015 (f)).

The slope of the linear regression of this dependence is a single number, which we used to characterize a network parameterization. In order to find the regions of parameter space where the linear relationship predicted by the models is compatible with the experimental data, the confidence interval for the slope of the linear fit to the data was used. Specifically, we regard a network model as being consistent with the experimental data if the OSI-OSI slope values for V_m, g_e, g_i, and f lie inside the 95 % confidence interval for the slope of the corresponding experimentally measured tuning property (see Figure 3.1).

3.3. Simulation results

The parameter space of the class of network models considered here is spanned by the peak conductance of synaptic excitatory connections to excitatory ($\overline{g}_{EE}$) and inhibitory ($\overline{g}_{IE}$) neurons. We first characterize the operating regimes found in this model space, before comparing the location dependence of tuning observed in the different models with that found experimentally.

Operating regimes of the network model

The operating regimes of a firing rate model can be defined in terms of the strength and shape of the effective recurrent input (Kang et al., 2003, see also Appendix D). The definitions of Kang et al. (2003), however, are based on the analytical solution of a linear firing rate model where all neurons are above threshold and cannot be applied directly to the non-linear Hodgkin-Huxley network model used here. Therefore, we characterize the parameter space explored here using a numerical definition of the operating regimes. This definition is based on the orientation tuning of the input currents to the excitatory model cells in the orientation domain ($0.6 <$ map OSI < 0.9). Specifically, if the sum of input currents is positive (negative) for all presented orientations, recurrent excitation (inhibition) is dominant, and the regime thus excitatory (EXC; respectively inhibitory, INH). If the sum of input currents has a positive maximum and a negative minimum (i. e., Mexican-hat like), a model receives significant excitation as well as inhibition and we refer to such a model as operating in the recurrent regime (REC). An example for the orientation tuning properties observed in the recurrent regime is shown in Figure 3.2B. Finally, if the sum of the absolute values of the currents through excitatory and inhibitory recurrent synapses of the model cells (at preferred orientation) is less than 30 % of the current through afferent synapses, the afferent drive is dominant and we call such regimes feed-forward (FF).

The regions of parameter space corresponding to these operating regimes are depicted

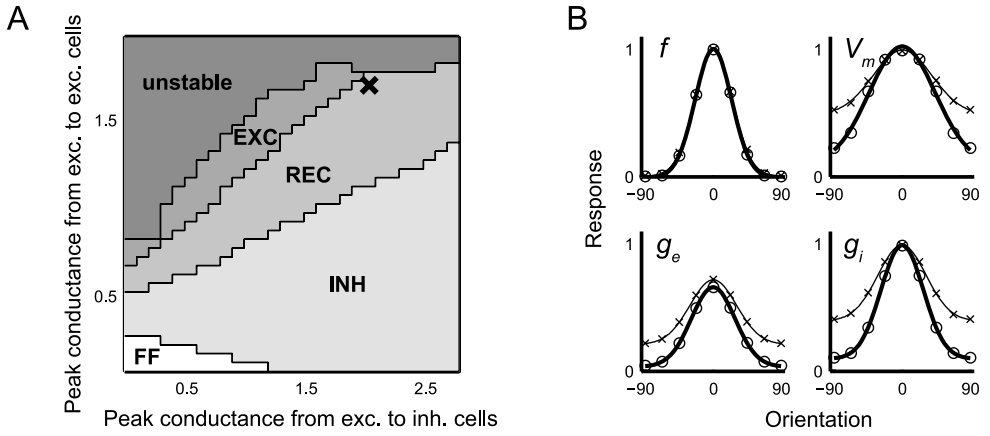


Figure 3.2 | Operating regimes of the Hodgkin-Huxley network model. **A** | Dependence of the operating regime of the network model on the peak conductance of synaptic excitatory connections to excitatory ($\overline{g}_{\mathrm{EE}}$) and inhibitory ($\overline{g}_{\mathrm{IE}}$) neurons: FF – feed-forward, EXC – recurrent excitatory dominated, INH – recurrent inhibitory dominated, REC – strong recurrent excitation and inhibition, and unstable. The conductances are given as multiples of the afferent peak conductance of excitatory neurons ($\overline{g}_{\mathrm{E}}^{\mathrm{Aff}}$). The figure summarizes simulation results for 38 × 28 different values of $\overline{g}_{\mathrm{EE}}$ and $\overline{g}_{\mathrm{IE}}$. **B** | Tuning curves for one example network in the REC regime (marked by a cross in A). Mean responses across cells are shown for the firing rate (f), the membrane potential (V_{m}), the total excitatory (g_{e}), and the total inhibitory conductance (g_{i}), separately for cells in iso-orientation domains (0.6 < map OSI < 0.9, thick lines) and cells close to pinwheel centers (map OSI < 0.3, thin lines). For each cell, responses were individually aligned to its preferred orientation and normalized to its maximum response; for the V_{m} tuning curve, the mean membrane potential without any stimulation ($V_m = -64.5\,\mathrm{mV}$) was subtracted beforehand. To allow comparison of the magnitude of g_{i} and g_{e} responses, both types of conductances were normalized to the maximum g_{i} response.

in Figure 3.2A as a function of the peak conductance of synaptic excitatory connections to excitatory ($\overline{g}_{\mathrm{EE}}$) and inhibitory ($\overline{g}_{\mathrm{IE}}$) neurons. We refer to the network as "unstable" if the model neurons show strong responses (average firing rate exceeds 100 Hz) and remain at high firing rates if the afferent input is turned off; i. e., the network shows self-sustained activity. In this regime, the model neurons lose their orientation tuning.

Orientation tuning properties in the different operating regimes

We analyzed the dependence of the orientation tuning properties on the operating regimes and compared them to the experimental data. For every combination of $\overline{g}_{\mathrm{EE}}$ and $\overline{g}_{\mathrm{IE}}$, we simulated the responses of neurons in the network model to oriented stimuli in order to measure the orientation tuning of V_{m}, f, g_{e} and g_{i} (see Section 3.2). The

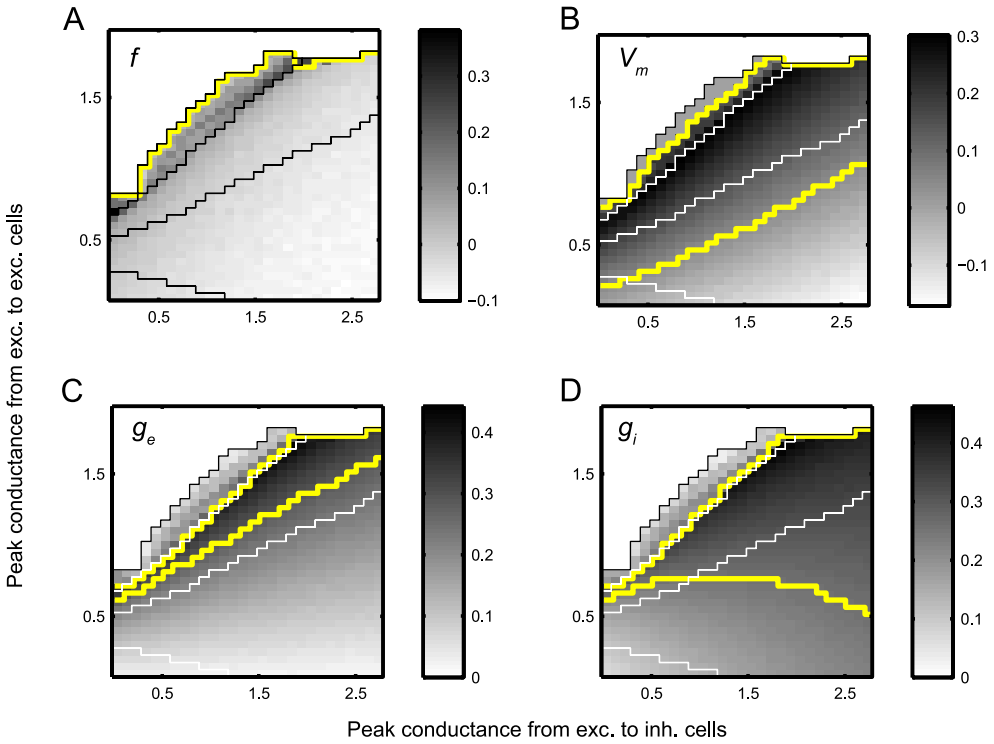


Figure 3.3 | Map-location dependence of orientation tuning of the conductances, the membrane potential, and the firing rate in the network model. The figure shows the slope values of the OSI-OSI regression lines (in gray values) as a function of the peak conductance of synaptic excitatory connections to excitatory ($\overline{g}_{\mathrm{EE}}$) and inhibitory ($\overline{g}_{\mathrm{IE}}$) neurons, separately for the spike rate (A), the membrane potential (B), the total synaptic excitatory (C), and inhibitory conductance (D). The conductances are given as multiples of the afferent peak conductance of excitatory neurons ($\overline{g}_{\mathrm{E}}^{\mathrm{Aff}}$). Thin lines denote the borders of the different operating regimes (cf. Figure 3.2). The region delimited by the thick yellow line corresponds to slope values within the 95 % confidence interval of the corresponding experimental data. Note that in (A) this region covers the whole range of operating regimes except the unstable regime. The figure summarizes simulation results for 38 × 28 different values of excitatory ($\overline{g}_{\mathrm{EE}}$) and inhibitory ($\overline{g}_{\mathrm{IE}}$) peak conductances.

OSI of each of the four quantities can then be plotted against the map OSI to reveal the dependence of the tuning on the map location (similar to the experimental data shown in Figure 3.1). The slope of the linear regression of this OSI-OSI dependence was used to characterize the different operating points of the network. Figure 3.3 shows these slopes for the tuning of f, V_m, g_e and g_i, as a function of $\overline{g}_{EE}$ and $\overline{g}_{IE}$ of the respective Hodgkin-Huxley network models (gray scale). Model networks with strong recurrent excitation (large values of $\overline{g}_{EE}$), as in the REC regime, predict steeper slopes than networks with less recurrent excitation. In other words, as the regime becomes increasingly more recurrently dominated, the recurrent contribution leads to sharper tuning of V_m, g_e and g_i in neurons within iso-orientation domains as compared to neurons near the pinwheel centers. However, yet closer to the line of instability the map-dependence of the tuning almost vanishes (slope approaching zero). This reflects the strong excitatory recurrent input in the EXC regime which leads to an overall increase in the network activity that is almost untuned and therefore provides very similar input to all neurons, regardless of map location. Also, the strongly inhibitory-dominated regimes (large values of $\overline{g}_{IE}$) at the bottom right corner of Figure 3.3 are of interest. Here, the slope of the location dependence becomes negative for the tuning of firing rate f and membrane potential V_m. Such a sharpening of the tuning close to pinwheels in an inhibition dominated regime has been observed elsewhere (McLaughlin et al., 2000).

Comparing the slope of the OSI-OSI regression lines to the 95 % confidence interval of the slopes estimated from the experimental data (Figure 3.1) allows us to determine those regions in parameter space that are compatible with the data (yellow contours in Figure 3.3). The observed location-independence of the firing rate tuning is compatible with all stable models in the parameter space (Figure 3.3A) and therefore does not constrain the model class. In contrast to this, the observed location-dependence of the membrane potential tuning (Figure 3.3B) and the inhibitory conductance tuning (Figure 3.3D) excludes most of the feed-forward and about half of the inhibitory-dominated regime. Most information, however, is gained from the observed location-dependence of the excitatory conductance tuning (Figure 3.3C). It constrains the network to operate in either a recurrent regime with strong excitation and inhibition or in a slightly excitatory-dominated regime.

Only the strongly recurrent regime satisfies all constraints

Combining the constraints imposed by the OSI-OSI relationship of the four measured quantities (yellow contour in both panels of Figure 3.4), we can conclude that the experimental data constrains the network to operate in a recurrent operating regime with strong recurrent excitation and inhibition.

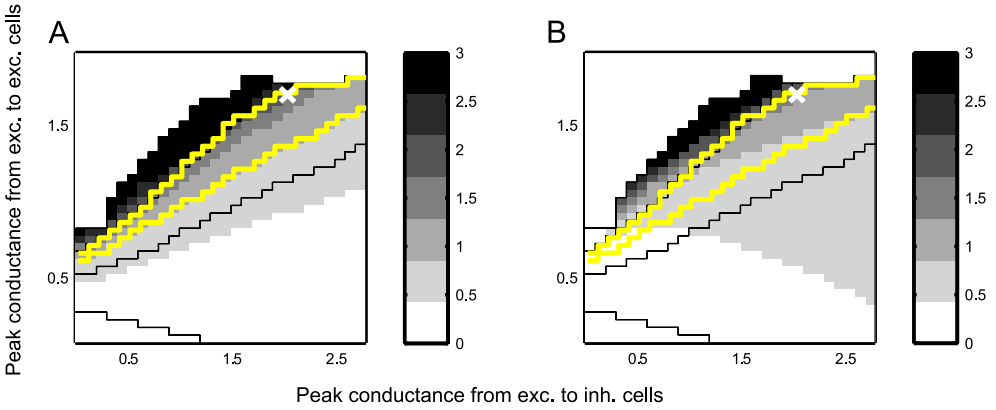


Figure 3.4 | Analysis of the results of the Hodgkin-Huxley network model. Ratio between (A) the excitatory current through the recurrent synapses and the current through afferent synapses of excitatory model cells and between (B) the inhibitory recurrent and the excitatory afferent current (in gray values). Currents were calculated for stimuli at the cells' preferred orientations, and averaged over all model cells within orientation domains ($0.6 <$ map OSI < 0.9). The region delimited by the thick yellow line corresponds to slope values that are in the 95 % confidence interval for each experimentally measured quantity (spike rate, membrane potential, the total synaptic excitatory, and inhibitory conductance). The white cross at $(2.0, 1.7)$ denotes the combination of model parameters that yields the best fit to the experimental data (see text). Thin lines denote the borders of the different operating regimes (cf. Figure 3.2).

In addition, we calculated the sum of squared differences between the data points (Figure 3.1) and the OSI-OSI relationship predicted by the model, for each operating regime. The "best fitting" operating regime, which had the lowest squared difference, is marked with a cross in Figure 3.4. The corresponding simulated orientation tuning curves for orientation domain and pinwheel cells are shown in Figure 3.2B.

In line with the definition of the operating regimes, the excitatory current through the recurrent synapses (gray values in Figure 3.4A) plays a negligible role in the feed-forward and in most of the inhibitory-dominated regimes. Only in the recurrent and the excitatory-dominated regime is the recurrent current stronger than the afferent current. A similar observation holds for the inhibitory current (Figure 3.4B). The strong recurrent currents in the excitatory-dominated regime reflect the strong overall activity that reduce the map-location dependence of the total excitatory and inhibitory conductances (cf. Figure 3.3C and Figure 3.3D).

We can thus conclude that the operating regime is in a region of parameter space that is characterized by significant excitatory and inhibitory recurrent contributions, which are approximately balanced and dominate the afferent input.

3.4. Discussion

Although much is known about the anatomy of lateral connections in the primary visual cortex of cat, the strengths of synapses formed by short-range connections are largely unknown. In our study, we use intracellular physiological measurements to constrain the strengths of these connections. Extensively exploring the parameter space of a spiking neural network model, we find that neither feed-forward dominated, nor recurrent excitatory- or inhibitory-dominated networks are consistent with the tuning properties observed *in vivo*. We therefore conclude that the cortical network in cat V1 operates in a regime with a dominant recurrent influence that is approximately balanced between inhibition and excitation, and dominates the afferent input.

This main finding – that tuning properties of cat V1 are best explained by a network operating in a regime with strong recurrent excitation and inhibition – is robust against variation of the values chosen for other parameters not varied here, such as the strengths of the inhibitory synapses ($\overline{g}_{\mathrm{II}}$ and $\overline{g}_{\mathrm{EI}}$) or the spatial range for lateral excitatory and inhibitory connections (data not shown; a detailed discussion on the robustness of the simulation results can be found in Stimberg et al., 2009). Nevertheless, the network architecture is based on a range of basic assumptions: e. g., all neurons in the network receive equally sharply tuned input. The explicit inclusion of location dependence of the input tuning might well lead to tuning properties compatible with the experimental data in different operating regimes. However, there is no evidence supporting such a location dependence of the afferent input and therefore assuming location-independent input seemed the most prudent basis for this analysis. Another assumption is the absence of untuned inhibition, since the inhibitory neurons in the network presented here receive tuned afferent input, too. The existence of an untuned inhibitory subpopulation is still a matter of debate (compare e. g., Hirsch et al., 2003; Cardin et al., 2007; Nowak et al., 2008). Naturally, such an untuned component would considerably reduce the location dependence of the inhibitory conductance g_{i}. Given that in our exploration only a small region of parameter space exists where the slope of g_{i} is steeper than in the experiment, a major contribution of such an untuned inhibition seems incompatible with the data.

Our analysis demonstrates that the network model is compatible with the data only if it operates in a regime that – due to the strong recurrent connections – is close to the line of instability across which the network exhibits strong self-sustained activity. Such a network is very sensitive to changes in its governing parameters, e. g., small changes in connection strengths lead to large changes in the overall firing rate: In the regimes close to the line of instability, increasing $\overline{g}_{\mathrm{EE}}$ by just 5 % typically leads to increases in firing rate of around 40 % (EXC), respectively 20 % (REC). In the other regimes (FF and INH) firing rate only changes by around 2 % – 3 %. In the "best fitting" operating

regime, a 10 % change in firing rate, which is of similar magnitude as observed firing rate changes under attention in macaque V1 (McAdams and Maunsell, 1999), is easily achieved by increasing $\overline{g}_{\mathrm{EE}}$ by 2 %. It therefore seems plausible that one benefit of being in such a regime is the possibility of significantly changing the "operating point" of the network through only small adjustments of the underlying parameters or feedback signals. Candidates for such an adjustment could be contextual modulations, adaptation or attentional effects.

Recapitulation

The cortical mechanisms that underlie the computation of orientation selectivity in primary visual cortex have been discussed throughout the last four decades. In this chapter, we showed how physiological measurements of the dependence of a cell's orientation tuning properties on its position in the orientation preference map can be used to constrain the operating regime of a biologically realistic Hodgkin-Huxley network model. We found that strong and approximately balanced recurrent excitation and inhibition are necessary to account for the experimental data. Orientation selectivity arises from the moderately tuned afferent input that is slightly sharpened and amplified by the recurrent interactions.

An interesting prediction of our model is that the intracortical interactions establish an operating point close to a line of instability, across which the network activity increases dramatically. Operating in such a regime would make the cortical responses particularly sensitive to modulatory effects or feedback signals. An exploration of this hypothesis will be presented in Chapter 5, where we take adaptation-induced short-term plasticity as an example of a modulatory, temporal shift of the operating point.

To this point, we abstracted from the dynamics of the V1 responses and based our analysis on the steady state the network reaches when presented with one non-changing orientation, i. e., a grating with a time-invariant orientation. In the next chapter, we investigate if a network model operating in the recurrent regime is also compatible with the dynamic properties of orientation tuning in V1.

4. Dynamics of orientation tuning in primary visual cortex

This chapter is based on Schummers et al. (2007) and Stimberg et al. (2009). The work presented in this chapter was done in collaboration with Marcel Stimberg and Robert Martin.

Abstract

Analysis of the time-course of the orientation tuning of responses in primary visual cortex can provide insight into the circuitry underlying orientation tuning. Extending the analysis of Chapter 3 to the temporal domain, we use recent experimental data from reverse correlation experiments to further constrain a Hodgkin-Huxley network model. Single-unit recordings from cat primary visual cortex showed that neurons close to pinwheel centers and neurons in orientation domains exhibit a similar time course in their averaged responses, but differences in their inter-cell variability. The mean responses of orientation domain cells are more similar to one another than those of pinwheel cells. We investigated how the temporal characteristics in pinwheel and orientation domain neurons vary with different parameterization of the Hodgkin-Huxley network. We find that in an excitation dominated regime the responses of orientation domain cells are markedly longer than those of cells close to pinwheel centers. The response curves of pinwheel and orientation domain cells are relatively similar for a wide range of feed-forward, recurrent, and moderately inhibition dominated regimes. The difference in the variance of the temporal responses between pinwheel and orientation domain cells observed *in vivo* can only be observed in the excitatory-dominated and the balanced recurrent regime of the network model. We show that the differential variability can be attributed to the variability present in the afferent input provided by neurons with different temporal characteristics. The recurrent connectivity removes some of this variability, but the degree of "smoothing" differs between orientation domains and pinwheel centers. In sum, we find that the in-vivo responses are not compatible with either excitatory or inhibitory dominated regimes, but can only be reproduced in a balanced recurrent regime.

4.1. Introduction

The dynamics of V1 cell responses can provide insight into the circuit and the mechanisms underlying orientation tuning. Several experimental studies have analyzed the orientation tuning dynamics of V1 neurons, for example using the reverse correlation paradigm (e. g., Ringach et al., 1997; Schummers et al., 2007). In particular, this type of analysis has been used to distinguish between thalamocortical inputs and intracortical excitatory or inhibitory inputs, and thus, to estimate their respective roles in the generation of orientation selectivity. The idea behind this approach is that synaptic inputs that a cell receives from different sources may be separable in time, and the afferent and recurrent drive may therefore be relatively more prominent at different periods of the response. For instance, if orientation selectivity is generated by a purely feed-forward mechanism – the convergence of inputs from the lateral geniculate nucleus (Hubel and Wiesel, 1962) – neurons should be equally selective during the initial and late periods of the response. Alternatively, if orientation selectivity arises from specific intracortical recurrent interactions, neurons should be less selective during the initial part of the response compared to the later period. Therefore, measurements of the time-course of response enhancement and suppression may be able to distinguish between different models of orientation tuning.

Some studies have demonstrated that the tuning curves derived from early portions of the visual response are quite different from those derived from later in the response (Chen et al., 2005; Ringach et al., 2003; Sharon and Grinvald, 2002; Shevelev et al., 1993; Volgushev et al., 1995), whereas others have found that orientation selectivity is relatively constant throughout the duration of the visual response (Celebrini et al., 1993; Gillespie et al., 2001; Mazer et al., 2002; Nishimoto et al., 2005). Recent results of Schummers et al. (2007) showed that at least part of this discrepancies can be attributed to different response dynamics resulting from differences in local cortical inputs. The response dynamics of cortical neurons reflect circuitry based on both vertical (i. e., the laminar position; see also Martinez et al., 2002) and horizontal location (the location in the orientation preference map; see also Schummers et al., 2002; Monier et al., 2003; Mariño et al., 2005) within cortical networks. Other factors influencing the diversity of results may include cell class, receptive field type, and differences in intracortical connectivity, probably reflecting functional differences (e. g., directional selective vs. non-selective cells).

Here, we focus on the characteristic signature of the location in the orientation preference map that is reflected in the response dynamics. We will first summarize the experimental findings of Schummers et al. (2007) concerning this issue and then use a Hodgkin-Huxley network model of V1, as in Chapter 3, to investigate which constraints

the experimental data impose on the strength of afferent, excitatory recurrent, and inhibitory recurrent connections.

4.2. Experimental findings

The experiments were performed by James Schummers in the laboratory of Mriganka Sur at MIT. Here we briefly summarize the experimental findings that are necessary as a background for this chapter. For the details we refer to Schummers et al. (2007).

Orientation preference maps were obtained using optical imaging of intrinsic signals from cat V1. Following optical imaging, single unit recordings were performed using a "reverse correlation" stimulus paradigm. The stimulus is a time series consisting of randomly chosen 20 ms frames of one of 16 high contrast gratings with 22.5° orientation spacing. The spike times were then reverse correlated with the stimulus sequence to estimate the linear relationship between stimulus orientation and firing probability. Thus, V1 responses were described by the probability of different stimulus orientations a certain time lag τ before a spike is elicited. The reverse correlation procedure is described in detail in Ringach et al. (1997); Schummers et al. (2007).

Neurons near pinwheel centers receive intracortical inputs from a wider range of orientations (Schummers et al., 2002; Yousef et al., 2001) than neurons in the center of orientation domains (cf. Section 2.2). Do the response dynamics reveal the different inputs that pinwheel and orientation domain neurons receive? Figure 4.1 shows the average tuning curves of the two populations (red: pinwheel cells, blue: orientation domain cells) over the time-course of the response. The tuning curves demonstrate, as with previous measurements of steady state responses (Dragoi et al., 2001; Maldonado et al., 1997; Schummers et al., 2002; Mariño et al., 2005), that cells have sharp tuning at pinwheel as well as orientation domain locations. The average responses of pinwheel and orientation domain neurons are very similar. Furthermore, there is no evidence of instability, shifts, multiple peaks, or any other gross differences in tuning in the cells near pinwheel centers.

To explore the relationship between map location and response dynamics in more detail, we examined the time-course of responses for the preferred orientation of each cell. Figure 4.2A shows again that the average response for the preferred orientation is similar for pinwheel and orientation domain cells. However, the variability of the response time-course is larger in pinwheel neurons (Figure 4.2B). For the preferred orientation, higher variance in the pinwheel population was most prominently following the peak response, during the decay phase.

Is this inter-cell variability significantly higher in the population of pinwheel cells or is it caused by a few outliers? To test this, we employed a point-wise bootstrap method:

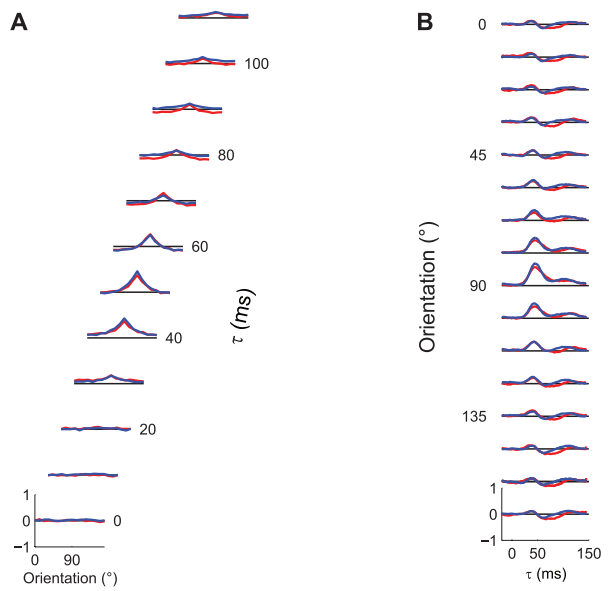


Figure 4.1 | Orientation dynamics for pinwheel center and orientation domain neurons. **A** | For each τ, the average tuning curve for all the pinwheel neurons (red; $n = 31$) and orientation domain neurons (blue; $n = 55$) is plotted. The vertical scale is the same for all values of τ. The tuning curves represent the average of normalized individual tuning curves, with zero representing the blank response (black line), and 1 representing the maximal response at all τ's and orientations. Thus, points in the tuning curve above the black line indicate enhancement of firing, whereas points below the black line represent suppression of firing. **B** | Average time-course of responses for pinwheel and orientation domain cells.

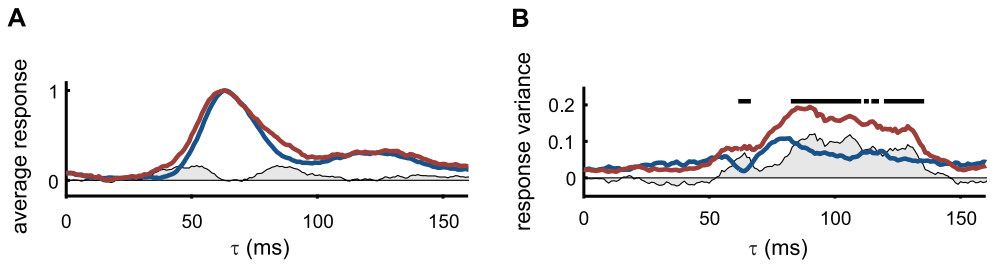


Figure 4.2 | Time-course of pinwheel and orientation domain cells at their preferred orientation. **A** | Average of the normalized responses of cells close to pinwheel centers (red line) and within orientation domains (blue line). Both curves are normalized to a peak of one; the shaded area visualizes the difference between pinwheel center and orientation domain responses. **B** | Variance of the normalized temporal responses at the preferred orientation and for each point in time, when averaging cells close to pinwheel centers (red line) and cells within orientation domains (blue line). The difference in variance is denoted by the shaded area. Black dots mark points in time, at which the variance of the response of pinwheel neurons is significantly higher than the variance of the response of orientation domain neurons, assessed using a bootstrap method.

15 pinwheel and 15 orientation domain cells were drawn at random from the set of all recorded cells. For every draw, the mean and the variance of the temporal response were calculated separately for pinwheel and orientation domain cells. For each time lag τ it was then counted in which proportion of draws the variance of the pinwheel cell responses was larger than that for the orientation domain cells. If this proportion was above 95 %, we regarded the difference as significant. Using this procedure, we found that the variance of pinwheel cells is indeed significantly higher after the peak of the response (Figure 4.2B).

Figure 4.3A plots the time-course of variance for pinwheel (red) and domain (blue) cells for each stimulus orientation. The figure shows that, for all orientations, the variance was higher in pinwheel cells for the entire response duration, particularly after the peak of the response. This suggests that the timing in orientation domain cells is much more uniform, while the timing of pinwheel center cells is much more heterogeneous.

To examine the possibility that the differences in the timing of pinwheel and domain responses are simply due to differences in the distributions of laminar position in the two groups, we plotted the differences in timing for each layer independently. Figure 4.3B shows the time-course of the response to the preferred orientation as a function of map location (columns), and laminar position (rows). There is more variability in pinwheel neurons in all three layers. These analyses suggest that despite having similar mean population tuning curves, there is substantially more individual variability in the timing of enhancement and suppression that shape the tuning curves of pinwheel cells. One potential explanation for this effect is the local cortical network surrounding the sites classified as pinwheel center, which are more heterogeneous than those in orientation domains (Mariño et al., 2005).

4.3. Modeling the dependence of the orientation tuning dynamics on map location in a Hodgkin-Huxley network model

To elucidate the mechanisms behind the observation that pinwheel cells show similar average response dynamics but have much higher variability than their orientation domain counterparts, we simulated the reverse correlation experiments in a large-scale neural network of a patch of V1 containing four pinwheel centers. Exactly the same model was used in the last chapter to determine which steady-state operating regime generates orientation tuning properties consistent with experimental findings. While up to this point we abstracted from the dynamics of the V1 responses, using a time-invariant input, we now include temporal filters, which describe the dynamics of the afferent input to each model cell. To realistically model the time-course of visual responses, the inputs to the model were filtered using temporal kernels matched to the impulse response

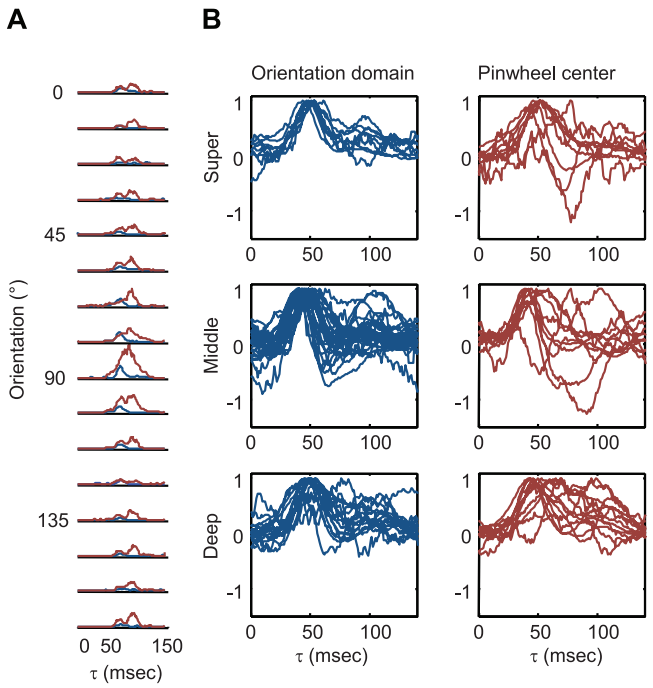


Figure 4.3 | Response timing is more variable near pinwheel centers. **A** | Plots of population variance in response amplitude as a function of time for each stimulus orientation. Pinwheel neurons ($n = 31$) are plotted in red and orientation domain neurons ($n = 55$) in blue. **B** | Response time-course of individual neurons, grouped as a function of laminar position and orientation map position. Orientation domain cells are plotted in the left column, and pinwheel center neurons in the right column. Superficial layer cells are plotted in the top row, middle layer cells are plotted in the middle row and deep layer cells are plotted in the bottom row.

functions of LGN cells (see Section C.5). A detailed description of the Hodgkin-Huxley network model can be found in Appendix C.

Time-course of averaged responses

We first investigated how the temporal characteristics in pinwheel and orientation domain neurons vary with different parameterization of the Hodgkin-Huxley network. These parameterizations correspond to the different operating regimes of the network, as defined in Chapter 3. The temporal responses of four different models (one for each operating regime; cf. dots in Figure 4.4A) to reverse correlation stimulation are shown in Figure 4.4B, using the same format as for the experimental results in Figure 4.2A. Note that the timescale is shifted for the simulated data: here, zero corresponds to the

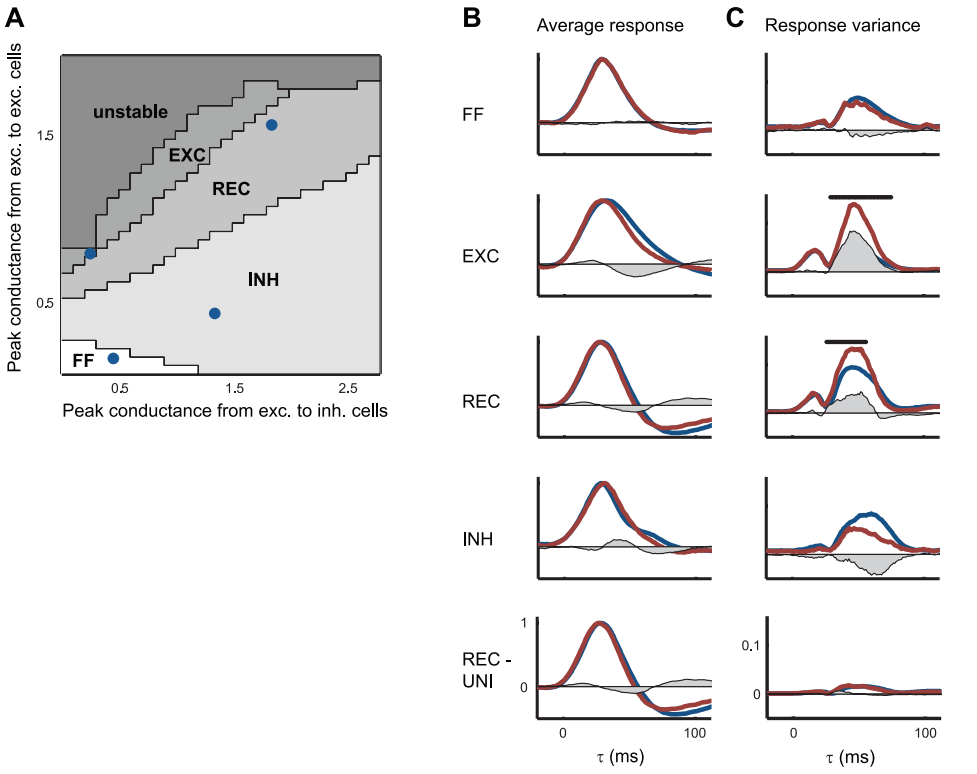


Figure 4.4 | Reverse correlation results for the Hodgkin-Huxley network model. **A** | Operating regimes of the network model (same as Figure 3.2A) as a function of the peak conductance of excitatory synaptic connections to excitatory ($\overline{g}_{\mathrm{EE}}$) and inhibitory ($\overline{g}_{\mathrm{IE}}$) neurons: FF – feed-forward; EXC – recurrent excitatory dominated; INH – recurrent inhibitory dominated; REC – strong recurrent excitation and inhibition. The conductances are given as multiples of the afferent peak conductance of excitatory neurons ($\overline{g}_{\mathrm{E}}^{\mathrm{Aff}}$). Dots indicate the parameterization of the example models in B and C, one for each parameter regime. **B** | Time-course of the responses, determined using the reverse correlation paradigm, at the cells' preferred orientation averaged over all pinwheel (map OSI < 0.3; red lines) and orientation domain (0.6 < map OSI < 0.9; blue lines) cells and normalized to a peak of one. Different plots show the results for the four example models marked in A. The model REC-UNI corresponds to the same operating point as REC; the difference is that for REC-UNI individual neurons all received as input a similar mix of the different afferent cell types (i. e., temporal input kernels). The gray shaded area denotes the difference between the averaged responses of pinwheel and orientation domain cells. **C** | Variance of the temporal responses at preferred orientation for cells close to pinwheels (red lines) and in orientation domain (blue lines) regions. The shaded area denotes the difference between the response variance of pinwheel and orientation domain cells. Black dots mark points in time at which the variance of the response of pinwheel neurons is significantly higher than the variance of the response of orientation domain neurons, assessed using a bootstrap method (cf. Section 4.2).

input onset for V1, in the experiment zero refers to stimulus onset. Because the network model does not include influences of long-range connectivity (omitted for simplicity) it cannot account for the late part of the response. Therefore, all analysis presented here is restricted to the period from 0 – 100 ms.

We find that in an excitation dominated regime (Figure 4.4B, "EXC"), the responses of orientation domain cells – for their preferred stimulus and averaged over all cells – are longer than those of cells close to pinwheel centers. This is intuitive: for an orientation domain cell, many neighboring cells have similar preferred orientations and, therefore, the excitation dominated recurrent input will amplify the response to the cell's preferred orientation most strongly. A pinwheel center cell, on the other hand, receives recurrent input from cells with a broader range of preferred orientations and, therefore, its recurrent input in response to the preferred orientation will not affect the cells response as much. It is thus the stronger recurrent drive in orientation domains that prolongs the response to a phasic input compared to pinwheel centers, where recurrent input is weak. However, such a prolonged response for cells in the orientation domain is inconsistent with the observations *in vivo* (Figure 4.2A). For the other example regimes, on the other hand, the phasic responses of orientation domain cells are similar (Figure 4.4B, "FF" and "REC") or even slightly shorter than those of the pinwheel cells ("INH"). In the recurrent regimes ("REC"), the time-course of the model neurons captures the onset and the phasic part of the neuronal responses well. The decay phase is less well described: the real neurons show a small second peak at approximately 120 ms (Figure 4.2A) and a plateau of sustained activity following their initial decline; the model neurons, on the other hand, decline further to below baseline. Nevertheless, as in the real neurons, throughout the whole time-course, the average responses of pinwheel and orientation domain neurons are similar. This is despite the fact that cells close to a pinwheel center receive recurrent input from cells with a much broader range of preferred orientations.

The difference between the response curves of pinwheel and orientation domain cells is quantified in Figure 4.5A. The figure shows the mean difference between the averaged response curves of pinwheel and orientation domain cells in the interval 0 – 100 ms (i. e., a measure for the size of the shaded areas in Figure 4.4B) as a function of $\overline{g}_{\mathrm{IE}}$ and $\overline{g}_{\mathrm{EE}}$. Averaged responses are relatively similar, i. e., differences between response curves are small to moderate (darker colors) for a relatively wide range of feed-forward, recurrent, and moderately inhibition dominated regimes. The values are compatible with what has been observed *in vivo* (0.06 for the interval 50 ms – 150 ms). For the example models shown in Figure 4.4A (dots) we obtain values of 0.02 (FF) and 0.06 ("REC" and "INH"), respectively. Strongly inhibitory as well as excitatory regimes show markedly larger differences (up to 0.15 for the more extreme cases, 0.08 for the "EXC" regime shown).

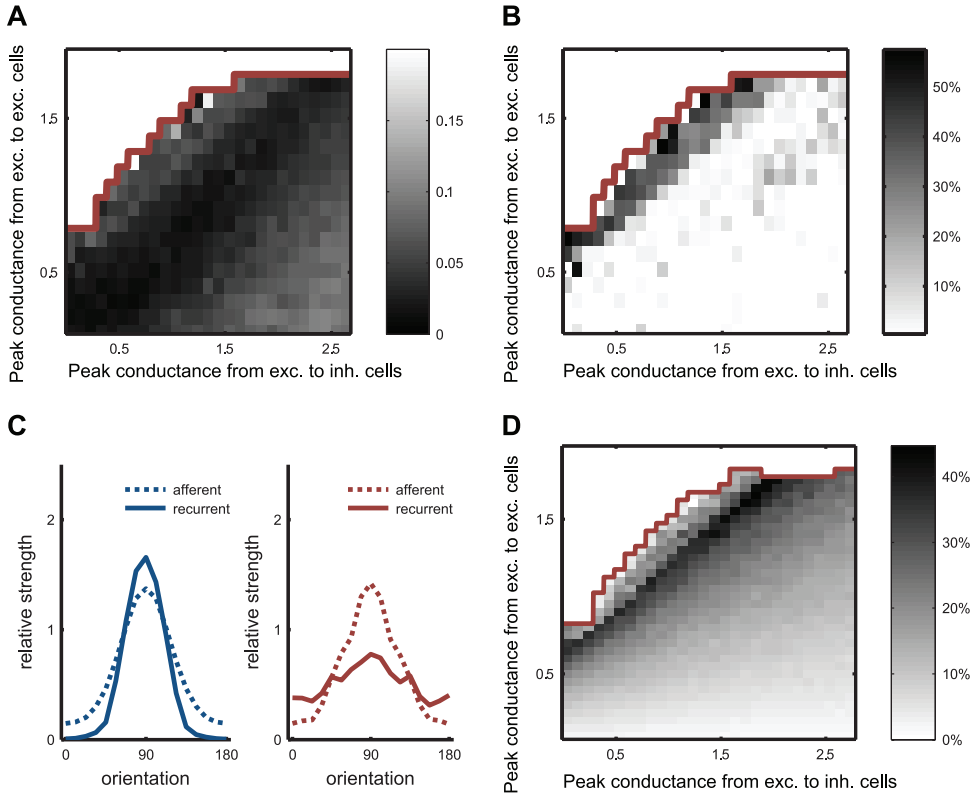


Figure 4.5 | Dependence of the temporal responses on the operating regime for the Hodgkin-Huxley network model. **A** | Mean difference in the time-course of the responses of pinwheel and orientation domain cells (this corresponds to the size of the shaded area in Figure 4.4B) as a function of $\overline{g}_{\mathrm{IE}}$ and $\overline{g}_{\mathrm{EE}}$. The mean difference between 0 ms (input onset) and 100 ms is coded by gray values for each of the 27×18 parameter combinations tested (see scale bar). The area above the solid line corresponds to parameter values for which the model network is "unstable". **B** | Fraction of time points in the first 100 ms for which the variance of response in pinwheel regions is significantly higher (cf. dots in Figure 4.4C; significance level > 0.95) than the variance of the response in orientation domains as a function of $\overline{g}_{\mathrm{IE}}$ and $\overline{g}_{\mathrm{EE}}$. The fraction of time (given in %) is coded by gray values (cf. scale bar) for each of the 27×18 combinations. To assess significance, we employed a point-wise bootstrap method (cf. Section 4.2). **C** | Tuning of the excitatory conductances of orientation domain neurons (left panel, blue) and pinwheel neurons (right panel, red) in the REC – regime. The afferent input (dotted lines) has the same strength for both, pinwheel and orientation domain cells. For the preferred orientation, the strength of the total recurrent excitatory conductance (red lines) is higher in orientation domain neurons. **D** | Differences in the contributions of the afferent drive between pinwheel areas and orientation domains. The contribution of the afferent drive was calculated as the ratio of the current through afferent synapses and the total excitatory current (all currents are taken at the cells' preferred orientations). The relative differences of this contribution between pinwheel and orientation domain cells are coded by gray values (in % , see scale bar). Thus, values greater than 0 indicate a higher contribution of the afferent drive in pinwheel neurons.

Inter-cell variability of the responses

The difference in the variance of the temporal responses between pinwheel and orientation domain cells observed *in vivo* provides a second important constraint. Figure 4.4C shows the response variance as a function of time for the selected example models. Only the excitatory-dominated (Figure 4.4C, "EXC") and the recurrent regime (Figure 4.4C, "REC") show the observed higher response variance in cells close to pinwheel centers. The variance of the pinwheel cell responses in the model is higher for all orientations, like for the experimental data (data not shown). To quantify the response variance as a function of $\overline{g}_{\mathrm{IE}}$ and $\overline{g}_{\mathrm{EE}}$, we compute, for every combination of parameters, the proportion of the interval 0 – 100 ms, for which the variance between pinwheel cells is significantly higher than the variance between orientation domain cells (using the bootstrap method described in Section 4.2). Figure 4.5B shows, that this fraction is highest along the "line of instability" ("EXC" and "REC" regimes), where values similar to the experimental data are achieved (Figure 4.2B). Experimental data show significantly increased variance for 49 % of the interval from 50 ms and 150 ms, though the variance difference is also partly in the late response, which the model cannot reproduce. Nonetheless the value compares well with the 30 % and more for the excitatory dominated and recurrent regimes (47 % and 30 %, respectively for the two examples shown), and is incompatible with the 0 % found for both, the feed-forward and inhibitory dominated regimes.

The more variable responses of the pinwheel neurons in highly recurrent regimes may appear trivial at first glance, since it might just reflect the difference in local circuitry: the recurrent inputs to the pinwheel cells are less uniform because they originate from cells with a broader range of preferred orientations. However, on closer inspection we find that rather than having the non-uniform recurrent connections of the pinwheel cells *introducing* variability into their responses, what actually happens is that the more uniform responses of orientation domain cells provide smoothing of the temporal variability already present in the input to the V1 cells. This can be seen in simulations where – without changing the operating point of the network – all network neurons received similar input: all cells then did in fact show smaller variance in their responses, and the difference between pinwheel and domain disappears (Figure 4.4C, "REC–UNI" simulation).

The cause of the differential smoothing effect in pinwheel and orientation domain (in the recurrent regime) can be observed directly by assessing the mean excitatory input conductances received from the different sources by pinwheel and orientation domain neurons as a function of stimulated orientation (Figure 4.5C). The relative contribution of feed-forward connections to a cell's inputs of the preferred orientation is far greater in the pinwheel than in the orientation domain. This means that the afferent input drives

pinwheel neurons more effectively than orientation domain neurons, which receive strong recurrent excitation. Thus, any variance present in this feed-forward input will have a more dominant effect on the cells response.

Also the dependence of the variance differences on $\overline{g}_{\mathrm{IE}}$ and $\overline{g}_{\mathrm{EE}}$ can be understood by considering the smoothing influence of the recurrent input. Only in highly recurrent regimes close to the line of instability the recurrent excitatory input into orientation domain cells can dominate their afferent input, whereas for pinwheel center neurons the relative contribution of the recurrent input is much weaker, leading to a particularly large difference of variability in this regime. This is confirmed by Figure 4.5D, which compares this relative strength of the afferent drive around pinwheel centers with that of orientation domain cells. As can be seen, only in regimes with a sufficiently strong excitatory drive, i. e., the regimes close to the line of instability, is the relative influence of the afferent drive stronger around pinwheel centers.

In sum, the temporal dynamics of V1 responses and their dependence on location in the orientation map also constrain the likely operating regime: the response shape as well as the similarity in the mean response is not compatible with either excitatory dominated or strongly inhibitory regimes; the response variability found *in vivo* can on the other hand only be reproduced in regimes close to the line of instability, i. e., excitatory dominated or balanced recurrent regimes. Taken together, only the strongly recurrent regime with balanced excitation and inhibition can account for both the observed similar time-course and the different variability of the responses of cells near and far from pinwheel centers. The results indicate variability already present in the afferent input as a likely cause of the variability in pinwheel neurons.

4.4. Discussion

We have investigated the influence of local circuit constitution on the dynamics of orientation tuning in cat V1. Experimental data shows that while the average time-course of responses is similar in orientation domains and pinwheel centers, there is much more variability in the timing at pinwheel centers. Using a realistic Hodgkin-Huxley network model incorporating orientation map topology, we demonstrated that this previously unreported phenomenon is the natural result of a regime with balanced excitation and inhibition receiving afferent inputs with diverse temporal kernels.

Robustness of the simulation results and model limitations

An important parameter that is not well characterized experimentally is the ratio of AMPA vs. NMDA receptors of excitatory synapses in visual cortices (Myme et al., 2003).

In the Hodgkin-Huxley network model, 70 % of the synapses were fast (AMPA-like) and 30 % were slow (NMDA-like). If the fraction of slow synapses is increased, the network remains stable for higher values of recurrent excitation. While under this manipulation the results stay qualitatively the same for the steady-state, the number of fast vs. slow synapses is crucial for obtaining a higher response variance in pinwheel regions compared to orientation domains. We tested several ratios of fast vs. slow synapses and found that more than about 50 % of the synapses have to be fast in order to obtain the experimentally observed higher variance close to pinwheel centers. This is intuitive: NMDA synapses have long time constants, slowing down the responsiveness of the network. Thus, fast afferent inputs are effectively prolonged and made more similar to one another, independent of the position in the orientation map.

The model responses in the strongly recurrent regime captures the main features of the responses of the real neurons with the exception of the late part of the response. This discrepancy between model and real neurons during the late phase of the response is likely due to the model's simplicity: Because of its restriction to the local recurrent network it does not incorporate any long-range horizontal and feedback connectivity, which may cause sustained activation of the neurons. Alternative explanations, such as global oscillatory behavior of cortex in response to the flashed grating stimulation are also difficult to reproduce in a simple one-layered network. Despite the simplicity of the model, the responses of model and real neurons are in good qualitative agreement during the first 100 ms.

The network's input is matched to the output of LGN neurons. Thus, strictly speaking the model only represents something akin to the input layers of V1. However, the basic result can easily be extended to other layers of V1. These, too, will receive strong visually driven input, likely originating in V1 simple cells. Such quasi-afferent inputs mediated through layer IV simple cells – while in time-course shifted compared to LGN – nonetheless are very similar in their temporal characteristics to those of LGN cells (Alonso et al., 2001; Wolfe and Palmer, 1998). Thus, the results regarding similar time-courses in pinwheel and orientation domain as well as regarding differential variance likely generalize.

Mechanism underlying the higher variability of pinwheel cells

We argue that the higher variability in response timing near pinwheel centers reflects the difference in the inputs from the local cortical circuits at those sites compared to orientation domains. There is evidence that the local circuit connectivity is isotropic in V1 (e. g., Mariño et al., 2005), and thus, in orientation domains, local inputs are integrated from a patch of cortex that contains a representation of only a narrow range

of orientations. All cells classified as orientation domain cells have a similar pattern of local inputs arising from neurons with a narrow range of preferred orientations. By contrast, the circuits near pinwheel centers are considerably more varied (Mariño et al., 2005; Schummers et al., 2002, 2004; Yousef et al., 2001). The network model implements such a pattern of local recurrent connectivity, and in our simulations, we were able to confirm that this higher variability in recurrent inputs leads to the observed larger variability of responses near pinwheel centers. It is worth noting that the local circuit variability alone is not sufficient to account for the differential output variability. Rather than actually introducing this variability through variability in the connectivity of pinwheel cells, the uniform recurrent connectivity in the orientation domain removes variability by integrating over the different inputs of their neighbors. This ability is reflected in the fact that the relative contribution of feed-forward connections to a cell's inputs of the preferred orientation is larger for pinwheel cells than for orientation domain cells.

One may hypothesize that other parameter manipulations may also result in the variance differences between pinwheel and orientation domain neurons, not requiring the variability to be present in the afferent input already. However, parameter changes, which may appear well-suited for introducing more variance into pinwheel neurons, such as varying the afferent input strength for different model cells or using a less symmetric orientation map for assigning the preferred orientation of the afferent inputs, did not lead to larger variance in the pinwheel neurons (data not shown). Thus, the temporally variable afferent input appears essential for reproducing the behavior of the real neurons in the model. Furthermore, there is experimental evidence that variability in the temporal response properties is already present in the responses of LGN neurons and remains present in the input receiving (simple) cells of V1 (Alonso et al., 2001; Wolfe and Palmer, 1998).

Operating regime of local cortical computations

Interestingly, only the strongly recurrent regime with balanced excitation and inhibition can account for all aspects of the *in vivo* data. This is fully consistent with the model-based analysis of stationary responses, i. e., the previously considered orientation tuning properties of the membrane potential and the input conductances of V1 neurons and their dependence on location in the orientation map (see Chapter 3). Both results constitute converging evidence for a cortical network operating in a regime close to the "border of instability", beyond which the network settles into a state of high, self-sustained activity.

Recapitulation

Beyond considering the steady state, the Hodgkin-Huxley network also allows investigating the temporal dynamics of orientation tuning. Here, we used experimentally measured tuning dynamics as an alternative constraint for pinning down the operating regime of the network. We added realistic temporal dynamics to the afferent input of the Hodgkin-Huxley network model, and obtained – for every parameterization of the model – temporal response kernels using the reverse correlation technique.

Our findings support the following conclusions: (i) The higher inter-cell variability close to pinwheel centers can be attributed to the variability present in the afferent input provided by neurons with different temporal characteristics. The recurrent connectivity in V1 removes some of this variability, but the degree of "smoothing" differs between orientation domains and pinwheel centers, causing the observed differences. (ii) The map-location dependence of the average temporal response properties of cat V1 neurons together with their variability can only be accounted for in a regime with significant recurrent contributions. A similar conclusion was reached analyzing the independent data set describing the steady-state behavior of V1 in Chapter 3, demonstrating the robustness of the results.

What are the benefits of an highly recurrent, approximately balanced operating regime for the processing of sensory signals? Networks operating close to the line of instability are very sensitive to modulations. Small changes in the network parameters, for instance, lead to relatively large changes in the neuronal activity. This supports the hypothesis that the balance between excitation and inhibition may allow for rapid transitions between relatively stable network states, permitting the modulation of neuronal responsiveness in a behaviorally relevant manner (Haider et al., 2006). Furthermore, it is conceivable that spatial and temporal context interactions (Schwartz et al., 2007) or attentional effects (McAdams and Maunsell, 1999) temporarily shift the operating point. It would be interesting to investigate if those modulatory influences can be explained in network models operating in a strongly recurrent, balanced regime.

In the next chapter, we take temporal context effects as an example and explore the possibility that the strongly recurrent regime facilitates adaptive processing of sensory stimuli. Specifically, we test if our model is consistent with the pronounced changes in orientation tuning observed in adaptation experiments.

5. Orientation adaptation and tilt aftereffect in a network model of V1

Abstract

The temporal context of a sensory stimulus influences the way the stimulus is represented by neural populations and, in turn, how the stimulus is perceived. A well studied adaptation phenomenon is the psychophysical tilt aftereffect, and the accompanying repulsive shifts and response changes of orientation tuning curves of individual neurons in primary visual cortex. Here, we use a firing rate model of orientation tuning to link the physiological and psychophysical results, and to investigate whether synaptic plasticity is a plausible mechanism underlying orientation adaptation. Since the strength of intracortical synaptic depression *in vivo* is unknown, we systematically explore its influence on recurrent processing in the model. Depending on the relative strength of depression for the different types of synapses, tuning curves of individual cells as well as the population responses show either attractive or repulsive shifts, and adaptation leads either to response suppression or facilitation. Interestingly, models consistent with experimentally observed physiological and psychophysically adaptation effects have underlying synaptic depression parameters that are consistent with *in vitro* studies. Moreover, we investigate how the tuning curve changes of model neurons depend on the local cortical neighborhood. The network model provides a mechanistic explanation for the experimental finding of larger adaptation-induced shifts close to pinwheel centers: adaptation affects the broadly tuned recurrent inputs near pinwheels more strongly, which in turn leads to larger tuning curve shifts. We do not observe strong adaptive tuning curve shifts and differences between pinwheel and orientation domain cells in models with weaker recurrent excitation or stronger inhibition. This suggests that one benefit of balanced recurrent processing is an enhanced sensitivity to adaptive changes and presumably also to other bottom-up and top-down modulations.

5.1. Introduction

The perception of a sensory stimulus depends on what has been observed in the recent past, i. e., on the stimulus' temporal context. Typically, prolonged exposure to a stimulus leads to "repulsive" aftereffects, i. e., stimuli similar to the adapting stimulus appear to

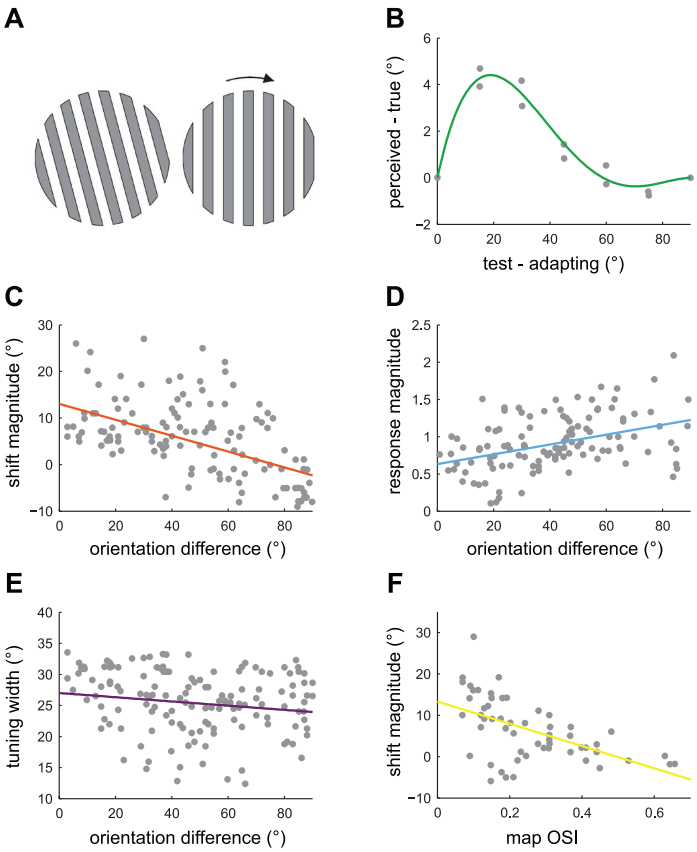


Figure 5.1 | Summary of experimental data on the psychophysical (A, B) and physiological effects (C–F) of adaptation to oriented stimuli. **A, B** | Tilt aftereffect. **A** | Adaptation to a tilted grating causes the vertical test grating to appear repulsed away from the adapting orientation. **B** | Repulsive shift of the perceived orientation as a function of the difference between the testing orientation and the adapting orientation. Psychophysical data (gray dots) replotted from Clifford et al. (2000). Solid line is the least square fit of a polynomial $\theta(90-\theta)a(1+b\theta)(1+c\theta)$ with parameters $a = 0.0061$, $b = -0.011$, $c = -0.017$ (Jin et al., 2005). **C–F** | Adaptation-induced plasticity of orientation tuning in V1 cells. **C** | Change in the preferred orientation after adaptation (positive numbers indicate repulsive shifts) is plotted against the difference between the adapting orientation and the control-preferred orientation of each neuron. Experimental data (gray dots) are taken from Dragoi et al. (2000) and fit by a regression line $a + b\theta$ with $a = 13$ and $b = -0.17$. **D** | Change in maximum firing rate relative to the control condition. Values greater (smaller) than 1 indicate facilitation (suppression). Data points are replotted from Jin et al. (2005) and fit by a regression line $a + b\theta$ with $a = 0.63$ and $b = 0.0066$. **E** | Width of tuning curves (standard deviation of a circular Gaussian tuning function) measured after adaptation. Data (gray dots) replotted from Jin et al. (2005) and fit by a regression line $a + b\theta$ with $a = 27$ and $b = -0.034$. **F** | Dependence of the adaptive change in the preferred orientation on the map OSI (the OSI of the orientation map at the location of the measured neuron; cf. Appendix A). Positive numbers indicate repulsive shifts. Data (gray dots) replotted from Dragoi et al. (2001) and fit by a regression line $a + b\theta$ with $a = 13.3$ and $b = -26.8$.

be more different from the adapting stimulus than they actually are. Psychophysical aftereffects (Clifford and Rhodes, 2005) have been found for stimulus properties including direction of motion (Levinson and Sekuler, 1976; Clifford, 2002), spatial frequency (Blakemore et al., 1970), stimulus orientation (Gibson and Radner, 1937), brightness (Eagleman et al., 2004), blur (Webster et al., 2002), and even for higher-level stimulus properties such as emotion, gender and attractiveness of faces (Rhodes et al., 2003; Webster et al., 2004). Underlying these phenomena is in many cases pattern-specific adaptation of visual neurons, leading to reversible changes in their response properties. How adaptation changes the stimulus encoding in single neurons and in neural populations is an active research area (for recent reviews see Clifford et al., 2007; Kohn, 2007; Schwartz et al., 2007). It has, however, proved difficult to link changes in neural response properties to both the cellular mechanisms they arise from and the perception they give rise to.

Here, we focus on the tilt aftereffect (Gibson and Radner, 1937; Mitchell and Muir, 1976; Clifford et al., 2000), a psychophysical phenomenon in which adaptation to a context stimulus causes an orientated test stimulus to appear repulsed away from the context orientation (Figure 5.1A–B; see also Figure 2.1A in Chapter 1). Neurophysiological studies (e. g., Müller et al., 1999; Dragoi et al., 2000, 2001; Felsen et al., 2002; Dragoi et al., 2002; Felsen et al., 2005) have identified the accompanying changes in orientation tuning of individual neurons in primary visual cortex: Through adaptation, tuning curves shift repulsively away from the adapting stimulus (Figure 5.1C; see also Figure 2.1B in Chapter 1). The shift magnitude depends on the difference between the preferred orientation of the neuron and the orientation of the adapting stimulus. The largest shifts occur for neurons with a small orientation difference (i. e., for neurons whose preferred orientation is close to the adapting orientation); neurons with large orientation differences either remain unaffected or show a slight attractive shift. Adaptation also leads to suppression or enhancement of firing rates (Figure 5.1D), and broadening or sharpening of tuning widths (Figure 5.1E), all depending on the difference between the preferred orientation of the neuron and the adapting stimulus. These data show that adaptation to a specific stimulus reorganizes orientation selective responses across the whole population of V1 neurons. This process can therefore only be understood considering the cortical network V1 neurons are embedded in. Indeed, Dragoi et al. (2001) found that adaptive tuning curve changes are not distributed uniformly across the cortical surface but depend on the position of a neuron relative to the orientation preference map (cf. Section 2.2). As shown in Figure 5.1F, neurons close to pinwheel centers show more pronounced repulsive shifts of their preferred orientation than neurons in orientation domains.

We have shown in Chapter 3 and Chapter 4 that the experimentally measured response

properties of V1 neurons are only consistent with the predictions of a Hodgkin-Huxley network model dominated by recurrent interactions and with balanced contributions from excitation and inhibition. Our analysis demonstrated that due to the strong recurrent contributions, the most likely operating point is close to the line of instability across which the network exhibits self-sustained activity. In this regime, small changes in network parameters disturbed the balance between excitation and inhibition and led to large changes in the neuronal activity (for instance, increasing the peak conductance of all excitatory-to-excitatory synapses by 5 % led to an increase in firing rate of more than 20 %). It is conceivable that modulatory processes such as adaptation influence the balance between excitation and inhibition. Since adaptation is stimulus-specific, it could lead to a local alteration of the cortical operating point, giving rise to the observed tuning curve changes. Furthermore, due to the structure of the orientation map, the balance is more crucial close to pinwheel centers, where neurons receive recurrent inputs at all orientations, whereas orientation domain neurons only receive recurrent inputs near their preferred orientation. This might make pinwheel cells more sensitive to changes in the intracortical inputs and orientation tuning more susceptible to adaptive modulation.

Here, we want to test the hypothesis that adaptation leads to a local change in the strength of recurrent excitatory and inhibitory connections, which in turn causes the changes in orientation selectivity. We use a mean-field (firing rate) model that is matched closely to the Hodkgin-Huxley network model used in the previous chapters (cf. Stimberg et al., 2009, Appendix D) and focus on depression of intracortical synapses as the underlying mechanism. Short-term synaptic depression has been found at multiple time scales from tens of milliseconds to seconds (Abbott et al., 1997; Varela et al., 1997; Tsodyks and Markram, 1997; Galarreta and Hestrin, 1998; Varela et al., 1999; Zucker and Regehr, 2002). This parallels the modulations of orientation tuning occurring from tens of milliseconds (Felsen et al., 2002) over hundreds of milliseconds (Müller et al., 1999; Dragoi et al., 2002) to several minutes (Dragoi et al., 2000). In principle, the multiple time scales may originate from diverse intrinsic or synaptic mechanisms. However, the similarity of the observed effects suggests a common underlying plasticity mechanism (Felsen et al., 2002; Schwabe and Obermayer, 2005).

Most of our knowledge about synaptic plasticity stems from *in vitro* studies; the strength of synaptic depression *in vivo* is in dispute (Chung et al., 2002; Boudreau and Ferster, 2005; Reig et al., 2006). We therefore explore, in a first set of analyses, how adaptive changes in orientation tuning of cells in the network model depend on the strength of synaptic depression of excitatory and inhibitory connections. For every parameterization of our model, the changes in orientation tuning were calculated in order to quantitatively compare the model responses to the physiological data (Figure 5.1C–F).

Moreover, we estimate the "perceived" stimulus orientation from the population response of the network and compare the bias caused by adaptation to the psychophysical data (Figure 5.1B). This is in contrast to previous modeling studies that have focused either exclusively on the perceptual tilt aftereffect (Wainwright, 1999; Clifford et al., 2000, 2001; Bednar and Miikkulainen, 2000) or on the neurophysiological tuning curve changes (Chelaru and Dragoi, 2008). Another previous study (Teich and Qian, 2003) concluded that there is a discrepancy between the psychophysical and the physiological data. However, Jin et al. (2005) showed that tuning curve shifts are in quantitative agreement with the tilt aftereffect, but in an abstract population coding model that does not provide a mechanistic explanation of the tuning curve shifts. Here we investigate how the tuning curve changes emerge as a result of recurrent processing. Together, the physiological and the psychophysical data provide strong constraints on the strength of the synaptic depression parameters of the model. An analysis of the set of models that are consistent with the experimental findings shows how different combinations of synaptic depression parameters can give rise to tuning curve modulations.

In a second set of analyses, we then use one of the consistent models to investigate how the observed changes in orientation tuning are generated through local interaction in the network. This allows us to pin down the mechanism that generates larger tuning curve shifts of neurons close to pinwheel centers.

5.2. Firing rate model and perceptual read-out

The firing rate network model

The firing rate model (cf. Kang et al., 2003; Stimberg et al., 2009) consists of two populations of threshold-linear neurons, each arranged in a 2-dimensional grid of 64×64 cells. The model architecture resembles the Hodgkin-Huxley network model used in Chapter 3 and Chapter 4. Both excitatory and inhibitory cells receive identically tuned feed-forward input, determined by an artificial orientation preference map consisting of four pinwheels (see Figure B.1 in Appendix B). Lateral connection strengths between the neurons are spatially isotropic and weighted according to a Gaussian distribution with the same spatial extent for excitation and inhibition. The main reason for using the firing rate model instead of the Hodgkin-Huxley model is the reduced simulation time, allowing us to explore a large range of model parameters. We have shown in Stimberg et al. (2009) that the firing rate network leads to very similar results as the Hodgkin-Huxley network model regarding the likely operating regime. In general, conductance-based models of large cortical neuronal networks can be described by simplified rate models, unless the underlying network state possess a high degree of

synchrony (Shriki et al., 2003).

The network model is similar to the model of Stimberg et al. (2009), but was extended to include dynamic intracortical synaptic connections. Synaptic depression was described by the mean-field equation derived by Tsodyks et al. (1998); see Equation D.5 in Appendix D. It is controlled by two parameters: the recovery time constant τ_{rec} which we set to 300 ms for all synapses, and U_{ij}, the utilization of synaptic efficacy (where $i, j \in \{\text{E}, \text{I}\}$ denote the excitatory and inhibitory populations; j denotes the presynaptic and i the postsynaptic population). In the detailed model of synaptic depression, involving the arrival times of individual spikes (Abbott et al., 1997; Tsodyks and Markram, 1997), each presynaptic spike activates a fraction U_{ij} of resources, which then quickly inactivate (with a time constant of few milliseconds) and recover with the time constant of τ_{rec}. Thus, besides the absolute connection strengths S_{ij}, the major parameters of the model are the U_{ij}'s, which determine the dynamics of the synaptic response. Experimentally, U values were found to cover almost the entire range from 0 to 1 (Tsodyks and Markram, 1997). In agreement with experimental data (Galarreta and Hestrin, 1998; Varela et al., 1999), we assume that both excitatory and inhibitory recurrent connections undergo depression. Thus, in our exploration of the phase space, we systematically vary U_{EE}, U_{IE}, U_{II} and U_{EI}. A detailed description of the firing rate model and the exact simulation protocol can be found in Appendix D.

Calculating the perceived stimulus orientation

In order to estimate the hypothetical perceived orientation from the population response of the firing rate network, we have to define how the rest of the brain "reads out" orientation from the responses of a population of V1 neurons. While various ideas exist how such population codes (Vogels, 1990; Deneve et al., 1999; Pouget et al., 2000, 2003) can be interpreted, this is still an open issue. Here, we compared the orientation estimates of two widely-used read-out methods: the population vector and the maximum-likelihood method.

In the population vector method (Vogels, 1990), the perceived orientation of a stimulus is constructed from the responses of all neurons using the vector average. Each neuron contributes a vector with orientation equal to twice its preferred orientation (as measured in the control condition) and length equal to its firing rate. Summation of these vectors gives a population vector, whose orientation is twice that of the perceived orientation. The formula for calculating the perceived orientation from the population response is the same as the one for calculating the preferred orientation of a single neuron (Appendix A, Equation A.3), with $R(\phi_i)$ denoting the response of a neuron with preferred orientation ϕ_i. For the calculation of the population vector the preferred

orientations ϕ_i have to be equally spaced. Because the preferred orientations are not exactly equally distributed in the artificial orientation map (cf. Appendix B) of the network, we employed bootstrap method: 60 cells with preferred orientations approximately (with a tolerance of 5°) equidistantly spaced between −90° to +90° were drawn at random from the set of all model cells 1000 times. For every draw, the perceived orientation was calculated. For reporting the results we use the mean and the standard deviation across the 1000 estimates of the perceived orientation.

The maximum-likelihood method (Pouget et al., 2000) estimates the perceived orientation as the orientation θ that maximizes the likelihood function $P(\vec{R}|\theta)$, where $\vec{R}$ is the vector of the firing rate responses of all neurons. Equivalently, the maximum-likelihood method can be described as fitting predetermined response templates to the population response and choosing the template that matches best (Deneve et al., 1999). The template corresponding to a perceived orientation θ is the population response to the stimulus with orientation θ before adaptation. Each of these templates is then fitted to the adapted population response by minimizing the squared error (i. e., assuming Gaussian distributed noise) through scaling the maximum of the template. The template that best matches, i. e., the one that achieves the lowest squared error, determines the perceived orientation.

5.3. Exploring the model space

We set up a firing rate model of a two-dimensional cortical orientation map in order to investigate how adaptation changes the orientation tuning properties of V1 cells depending on the strength of synaptic depression of excitatory and inhibitory connections. All excitatory and inhibitory neurons in the model receive orientation-tuned afferent inputs with similar tuning widths and with their preferred orientations being assigned according to the orientation map. The strength of recurrent excitatory synaptic connections (S_{EE} for excitatory postsynaptic cells, S_{IE} for inhibitory postsynaptic cells) and inhibitory connections (S_{EI} to excitatory cells, S_{II} for inhibitory) were chosen so that the model cells receive strong recurrent excitation and inhibition, both dominating the feed-forward input. We have shown in Chapter 3 (see also Wimmer et al., 2009; Stimberg et al., 2009) that only such an operating regime is compatible with experimental data on orientation tuning in primary visual cortex.

We then systematically varied the strength of the synaptic depression parameters U_{EE}, U_{IE}, U_{II} and U_{EI} for the four types of recurrent connections. Each parameterization gives rise to a different model instance, which is then used to simulate adaptation experiments. First, orientation tuning curves of model cells are measured in the control condition, then the network model is adapted for 1 s to a stimulus with 0° orientation,

and tuning curves are measured again (test condition). A detailed description of the simulation protocol can be found in Appendix D.

Visualizing the parameter space of the network

We use a technique called *dimensional stacking* (LeBlanc et al., 1990; Taylor et al., 2006) to visualize how adaptation influences the response properties of the network as a function of the four synaptic depression parameters. Dimensional stacking is a linear projection from a multidimensional parameter space to only two dimensions. For concreteness, we introduce dimensional stacking by applying it to the firing rate of the network models, measured after adaptation. Figure 5.2A shows the maximum firing rate of excitatory neurons (color-coded) in the adapted network models as a function of the parameters U_{II} and U_{IE} (depression strength of inhibitory and excitatory connections to inhibitory neurons) while the other two parameters (U_{EE} and U_{EI}) were kept fixed. The firing rate of excitatory neurons decreases with increasing depression of the connections from inhibitory to inhibitory neurons (increasing values of U_{II}, i. e., going from left to right along the x-axis) because increasing U_{II} leads to an increase in firing rate of inhibitory neurons. For increasing depression of excitatory connections to inhibitory neurons (increasing values of U_{IE}, i. e., going from bottom to top along the y-axis), firing rates increase until the network crosses the border to a parameter regime where the network is "unstable" (white squares), that is, where the firing rates exceed 100 Hz (see Appendix D). The firing rates increase because higher values of U_{IE} lead to less excitation of inhibitory neurons which in turn leads to less inhibition of excitatory neurons (via the connections from inhibitory to excitatory neurons).

Dimensional stacking enables us to visualize the results for the full exploration, where we independently vary all four depression parameters. This is achieved by making a montage of plots like the one shown in Figure 5.2A with each individual plot in the montage scanning U_{II} and U_{IE} and with the other two parameters (U_{EE} and U_{EI}) varying over the montage. The resulting dimensional stacking image of the maximum firing rate (Figure 5.2B) is easy to interpret. Synaptic depression is weakest for all connections on the bottom left of the figure. Firing rates decrease for higher values of U_{EE}, and increase for higher values of U_{EI}. The assignment of parameters to the inner "low order" axis and the outer "high order" axis influences the usefulness of the visualization (cf. Taylor et al., 2006). We chose to put U_{EE} and U_{EI} on the outer and U_{II}, U_{IE} on the inner axis, because this order leads to a compact representation of the unstable region (shown in white). Additionally, this order is particularly useful, because both depression parameter that lead to more inhibition (and consequently to lower firing rates of excitatory cells) vary along the x-axis, and both parameters that lead to more excitation vary along the y-axis.

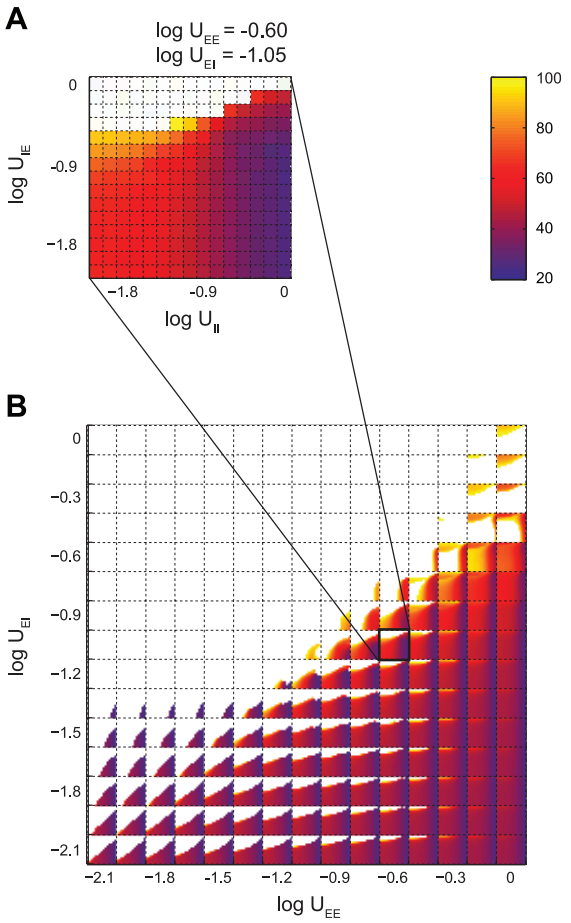


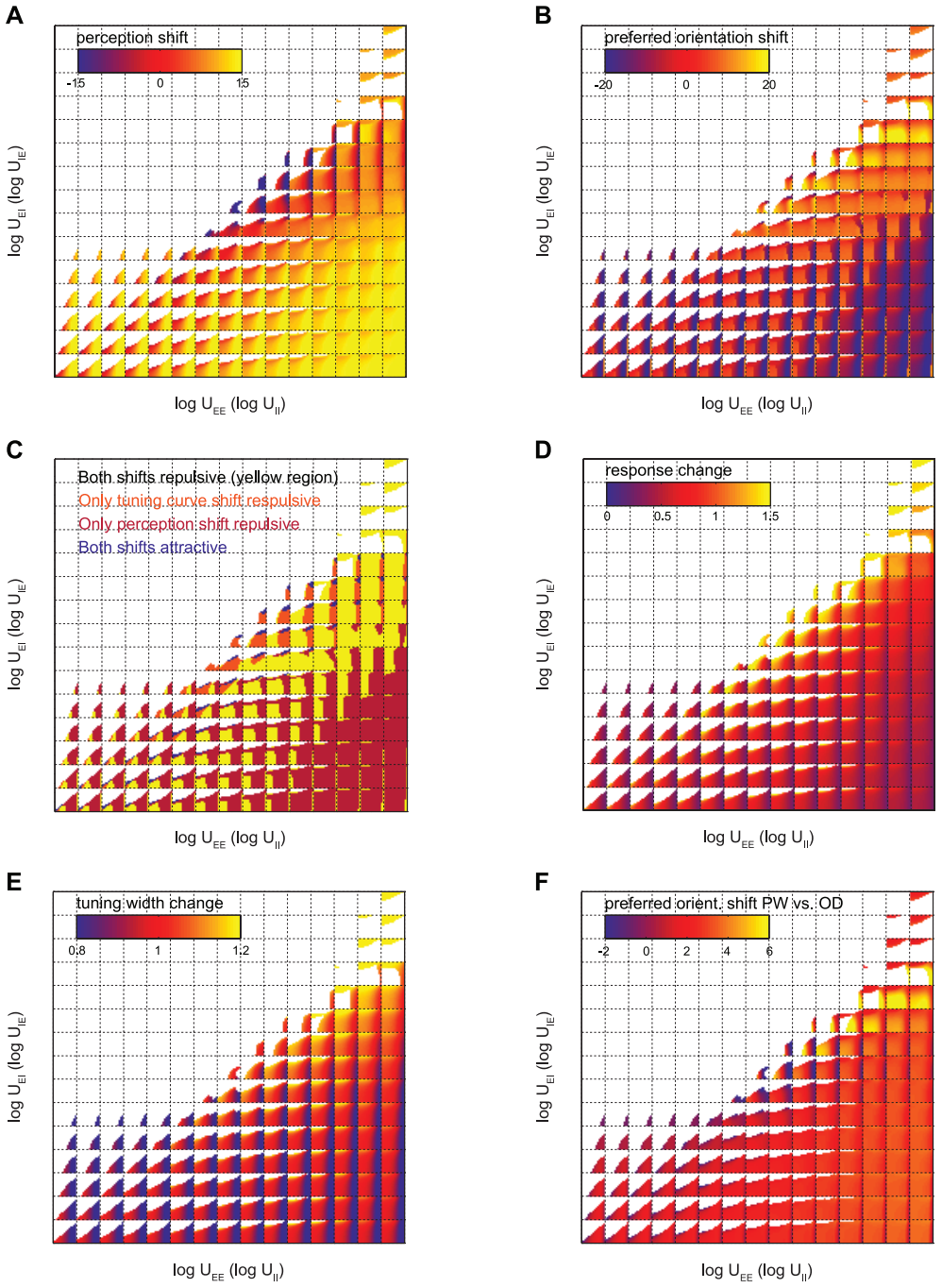
Figure 5.2 | Maximum firing rate of excitatory cells after adaptation as a function of the synaptic depression parameters U_{EE}, U_{IE}, U_{II} and U_{EI}. **A** | Firing rate as a function of the two independently varied parameters U_{II} (controlling depression of recurrent connections from inhibitory to inhibitory neurons) and U_{IE} (controlling depression of recurrent connections from excitatory to inhibitory neurons). The parameters were varied on a logarithmic scale between −2.1 and 0, corresponding to values from 0.0079 (almost no synaptic depression) and 1 (maximum synaptic depression). The two other parameters were kept fixed at $\log U_{EE} = -0.60$ and $\log U_{EI} = -1.05$. Colors (scale bar on the right) denote the value of the maximum firing rate across all excitatory neurons and all test stimulus orientations, after adaptation to a 0° stimulus. White squares denote the unstable region that was determined numerically. In particular, we consider networks firing at rates higher than 100 Hz as unstable (see Appendix D for the exact definition of the unstable region). The figure summarizes the simulation results from 15 × 15 different values of U_{II} and U_{IE}. **B** | To visualize the firing rate for the entire parameter space, the grid from (A) is embedded in a larger grid scanning the additional two depression parameters U_{EE} and U_{EI}. Within each square of the grid shown in (B) is a 15 × 15 grid scanning U_{II} and U_{IE}. Overall, a 225 × 225 grid is formed, scanning the 50 625 (15^4) possible combinations of depression parameters we tested.

Adaptive changes in the firing rate model

In order to investigate how adaptation influences the orientation tuning properties of cells in our network model and thus also the perceived stimulus orientation, we determine – for every parameter combination – the changes in the population response, in the preferred orientation, in the maximum response, and in the tuning width. Figure 5.3 summarizes these results, using the dimensional stacking technique (cf. Figure 5.2B). Interestingly, depending on the relative strength of the depression parameters, tuning curves of individual cells as well as the population responses can either show attractive or repulsive shifts, i. e., towards or away from the adapting stimulus (Figure 5.3A,B). Within a large region of the parameter space (towards the bottom right of Figure 5.3A), perception shifts towards the adapting stimulus while the preferred orientation of individual cells shifts repulsively away from the adapting stimulus (Figure 5.3B). This finding is consistent with Jin et al. (2005), who showed that shifts in preferred orientation lead to opposite shifts in perception, unless they are balanced by suppression.

However, in the experimental data (Figure 5.1B,C), both perceived orientation and tuning curves generally shift repulsively away from the adapting stimulus (indicated by

Figure 5.3 *(facing page)* | Adaptive changes of perceived orientation (A) and neuronal tuning curves (B–F) as a function of the synaptic depression parameters of the firing rate model. Parameters were varied on a logarithmic scale, and the axis in the figures are assigned as in Figure 5.2B: excitatory (U_{EE}) and inhibitory connections (U_{EI}) targeting excitatory neurons vary with grid positions (outer axis), inhibitory (U_{II}) and excitatory (U_{IE}) connections targeting inhibitory neurons are varied within each square of the grid (inner axis). **A** | Maximum repulsive (positive values) or attractive (negative values) shift in perception, calculated from the population responses of the models using the maximum-likelihood method. For parameter combinations that led – depending on the difference between adapting and test stimulus – to repulsive or attractive shifts, the shift with the higher absolute value was assigned. Colors denote the shift value in degrees (see scale bar). **B** | Maximum preferred orientation change (in degrees, positive values: repulsive shifts) of any of the neurons in the model network, calculated by vector averaging. If neurons in the model network showed both shift directions, the shift with higher absolute value was assigned. **C** | Directions of perception and tuning curve shift (different colors). Perception and neuronal tuning curves shift repulsively away from the adapting stimulus – as found experimentally – in the yellow region. **D** | Average change in maximum firing rate relative to the control condition across all neurons. Values greater (smaller) than 1 indicate dominating facilitation (suppression). **E** | Ratio between the width of tuning curves after adaptation and in the control condition. Shown are the average values of this ratio across all neurons in the network. Values greater (smaller) than 1 indicate an overall broadening (sharpening) of tuning curves. **F** | Difference between the preferred orientation change of cells in pinwheel areas (map OSI < 0.3) and cells in orientation domains ($0.6 <$ map OSI < 0.9). The difference is calculated between the average tuning curve shift across all pinwheel and all orientation domain cells. White color corresponds to parameter values for which the model network becomes unstable (cf. Appendix D). The figure summarizes the simulation results for 50 625 (15^4) different values of the synaptic depression parameters.


A
perception shift
-15
0
15
log U_{EI} (log U_{IE})
log U_{EE} (log U_{II})
B
preferred orientation shift
-20
0
20
log U_{EI} (log U_{IE})
log U_{EE} (log U_{II})
C
Both shifts repulsive (yellow region)
Only tuning curve shift repulsive
Only perception shift repulsive
Both shifts attractive
log U_{EI} (log U_{IE})
log U_{EE} (log U_{II})
D
response change
0
0.5
1
1.5
log U_{EI} (log U_{IE})
log U_{EE} (log U_{II})
E
tuning width change
0.8
1
1.2
log U_{EI} (log U_{IE})
log U_{EE} (log U_{II})
F
preferred orient. shift PW vs. OD
-2
0
2
4
6
log U_{EI} (log U_{IE})
log U_{EE} (log U_{II})

positive shift values). The experimentally observed shifts can be used as a constraint on our network models. Thus, we defined four different classes of models, depending on the shift directions of perception (population response) and tuning curves: (i) both shifts repulsive, (ii, iii) either perception or tuning curve shift repulsive, and (iv) both shifts attractive. Figure 5.3C shows this four classes, indicating that a large region in parameter space (colored in yellow) is consistent with the experimental constraint that both shifts are repulsive.

In addition to the perception and tuning curves shifts, we also analyzed the changes in maximum response (Figure 5.3D), tuning widths (Figure 5.3E), and the difference in preferred orientation shifts for pinwheel cells and orientation domain cells (Figure 5.3F). Depending on the strength of synaptic depression, adaptation can lead to either response suppression or facilitation, and the width of tuning curves can either increase (broaden) or decrease (sharpen). We next sought to determine which of these models are plausible, given the available experimental evidence.

Model selection

In order to find network models that are consistent with the experimental data, we used the constraints that the experimental findings impose on the perception and tuning curve changes. We examined our population of 50 625 network models and selected those with properties that agreed with the experimental data of Figure 5.1 (cf. Clifford et al., 2000; Dragoi et al., 2000, 2001; Jin et al., 2005). To be classified as an *admissible* model, the population response of the network as well as the tuning properties of each neuron in the network had to lie within bounds chosen to contain the central 85 % of the experimental data points. The lower and upper bounds for the different properties were (i) perception shift: −3° to 6° (estimated from Mitchell and Muir, 1976; Clifford et al., 2000, see Figure 5.1B); (ii) preferred orientation shift: -6.06 to 17.99° (calculated from the data points in Figure 5.1C); (iii) maximum firing rate relative to the control condition: 0.40 to 1.50 (calculated from the data points in Figure 5.1D); (iv) tuning width relative to the control condition: 0.69 to 1.25 (calculated from the data points in Figure 5.1E). Additionally, the average preferred orientation shift had to be larger for pinwheel cells than for orientation domain cells (Figure 5.1F).

Selecting models using these criteria yielded 1014 admissible models corresponding to approximately 2 % of the parameter space. Figure 5.4A shows where these models are located in parameter space (blue region). All admissible models lie below the diagonal where U_{EI} equals U_{EE}. In other words, to be consistent with the experimental data, excitatory connections (targeting excitatory neurons) have to undergo stronger depression than inhibitory connections ($U_{\mathrm{EE}} > U_{\mathrm{EI}}$). To asses how much each of the experimental

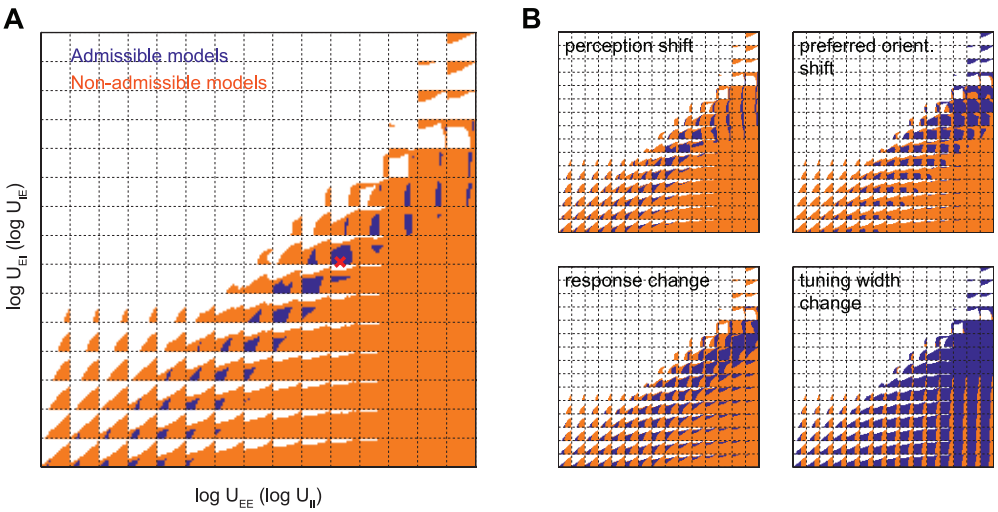


Figure 5.4 | Distribution of admissible models in the space of the synaptic depression parameters U_{EE}, U_{IE}, U_{II} and U_{EI}. **A** | Admissible models (shown in blue) were selected using consistency with the experimental data as a criterion. Perception shift, tuning curve shifts, firing rate changes, tuning width changes, and differences in shift magnitude between pinwheel and orientation domain cells must all lie within the interval containing 85 % of the corresponding experimental data points. Non-admissible models are shown in orange, and white corresponds to parameter values for which the model network becomes unstable. The red cross denotes an admissible model whose responses are analyzed in detail in Figure 5.8 – Figure 5.11. **B** | Classification of models as admissible (blue) or non-admissible (orange) when considering each of the constraints given by the experimental data separately. Admissible models shown in (A) fulfill all of the constraints, i. e., the blue region in (A) is the intersection of the all the blue regions shown in (B). Parameters were varied on a logarithmic scale and the axis in the figures are assigned as in Figure 5.2B.

measurements contributes to the selection of admissible models, we consider each of the constraints given by the experimental data separately (Figure 5.4B). As can be seen in the figure, the blue regions for the perception shift is smallest, suggesting that this property particularly restricts the parameter space of admissible models. The changes in tuning width are consistent with the experimental data for a wide range of model parameters, reflected by the large blue region for this property. However, this means that the changes in tuning width are much less informative than the other properties. Indeed, we found that no model can be excluded from the class of admissible models only because the tuning width change exceeds the bounds given by the experimental data, i. e., this property did not provide additional information.

The mean values of perceptive shifts and tuning curve changes of the admissible models are generally similar to the mean values computed from the experimental data (Figure 5.5A–E). The shifts in preferred orientation, however, are somewhat smaller than

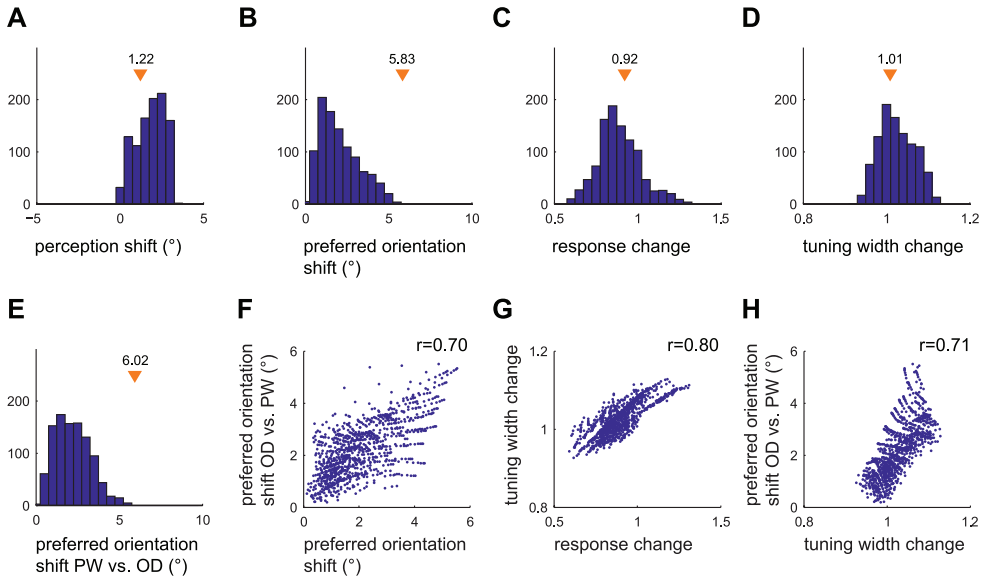


Figure 5.5 | Comparison of the properties of admissible models and experimental data. **A** | Histogram of the average shifts in perception across the whole range of test orientations. Perceptual shifts were calculated as the difference between the maximum-likelihood estimate of the stimulus orientation and the actual stimulus orientation. The triangle denotes the average perceptual shift calculated from the experimental data of Clifford et al. (2000) shown in Figure 5.1B. **B** | Histograms of the shift in preferred orientation compared to the experimental data of Dragoi et al. 2000 (triangle, cf. Figure 5.1B). Shifts were calculated as the difference between the preferred orientations (vector average of responses) of the neurons after adaptation and in the control condition. **C–E** | Histograms of the other properties of adapted tuning curves compared to the experimental data of Dragoi et al. 2000, 2001; Jin et al. 2005 (experimentally observed values are denoted by triangles; cf. Figure 5.1D–F for the experimental data). The y-axis in panels A–E is the number of admissible models. **F–H** | Correlations between pairs of model responses in the population of 1014 admissible models. Shown are the scatter plots of pairs of model responses that exhibited large ($r > 0.5$) and significant ($p < 0.001$) correlations.

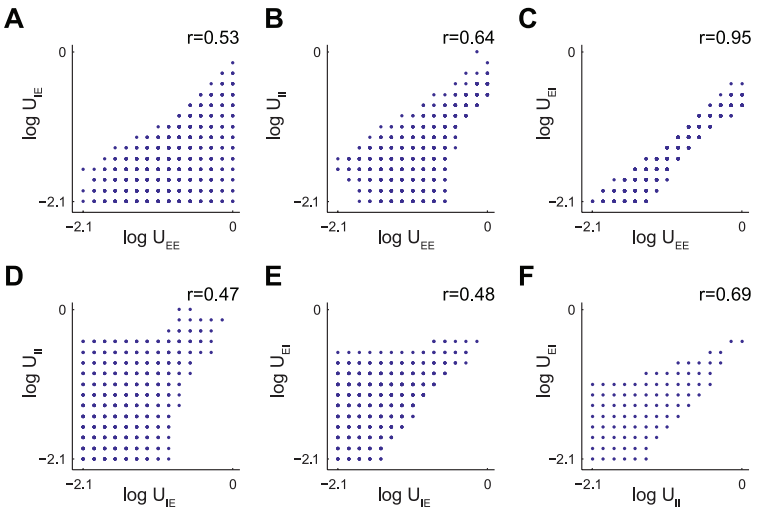


Figure 5.6 | Correlations between pairs of model parameters in the population of 1014 admissible models. **A–F** | Scatter plots of pairs of model parameters (logarithmic scale). Because of the simulation grid there are $15 \times 15 = 225$ possible combinations for each two model parameters. The linear correlation coefficient r is given on top of each figure ($p < 0.001$ in all cases).

in the experimental data. Figure 5.5F–H shows that some of the response properties of the admissible models are correlated. In particular, the shift in preferred orientation and the difference of the shift for pinwheel versus orientation domain cells is positively correlated, i. e., models with large shifts in preferred orientation also show large differences for pinwheel and domain cells. Thus, some models show tuning curve shifts with a similar magnitude as found experimentally.

Properties of the parameter space of admissible models

The properties of the model subspace spanned by the set of admissible models can provide insight into the mechanisms underlying the experimentally observed tuning curve changes. What attracts attention in Figure 5.4A is that it seems that all the admissible models (blue region) are part of a single connected "island" in parameter space. This conjecture was tested using connected component analysis (Shapiro and Stockman, 2001) on the sampling grid. The connected components algorithm determines whether there is a path (on the grid) from any admissible model to any other admissible model such that each model on the path is also part of the set of admissible models. The result was, that the set of admissible models indeed form a single connected component, suggesting that the model class is robust against small variations of the synaptic depression

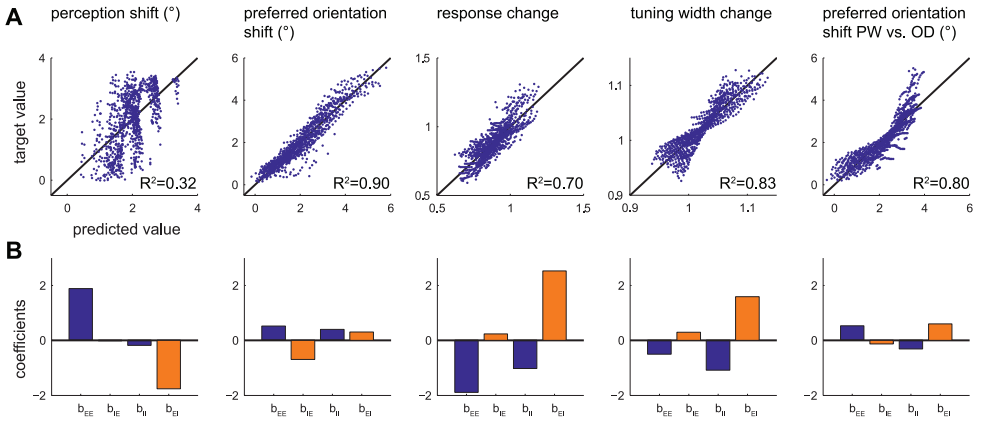


Figure 5.7 | Results of linear regression on model responses, using the model parameters U_{EE}, U_{IE}, U_{II} and U_{EI} as predictors. **A** | Quality of the linear fits for the different response properties. Response properties of the admissible models are plotted against the values predicted by multiple linear regression. A perfect fit would have all points on the identity line, shown in black. The R^2 values of the fits (given below the points) range between 0.32 for the perception shift to 0.90 for the preferred orientation shift. **B** | Linear regression coefficients b_{EE}, b_{IE}, b_{II}, b_{EI} for each of the fits shown in panel (A).

parameters. Although this is not a mathematical proof we believe that it is a reasonable conjecture given that the sampling of the parameter space is dense enough.

We next examined how the constraints imposed on the network models restrict the parameter space of admissible models. As can be seen in Figure 5.6, the depression parameters of admissible models span a large range. However, not all parameter combinations are possible, e. g., for larger values of U_{IE} in Figure 5.6A the parameter U_{EE} must be increased accordingly to stay in the region of admissible models. Furthermore, almost all admissible models lie below the "diagonal". This means that depression must be stronger for the excitatory–excitatory connections than for the excitatory–inhibitory connections. Generally, the parameters of admissible models are pairwise correlated. (Figure 5.6A-F). The parameters U_{EE} and U_{EI} – that were varied on the outer axis in Figure 5.4 – exhibit the strongest correlation. These results imply that while each single depression parameter can vary in a large range, the parameters cannot be varied independently. Rather, a change in one model parameter must be compensated by a corresponding change in other parameters in order to stay in the region of admissible models.

Linear regression analysis

We were curious to know how each synaptic depression parameter of the network model influences each of the response properties, such as the change in preferred orientation. The model equations provide of course a quantitative description of how the model properties vary as a function of the model parameters, but only implicitly. In order to find a simpler description that can be interpreted directly, we performed multiple linear regression on the model responses, using the model parameters U_{EE}, U_{IE}, U_{II} and U_{EI} as predictor variables. If linear regression provides a good fit to the actual property, the regression coefficients directly reflect the contribution of each model parameter.

Figure 5.7A shows that it is indeed possible to predict the magnitude of the adaptive changes with a simple linear model. In the figure, the actual model properties are plotted against the predicted values from linear regression, i. e., a perfect fit would have all points on the identity line. The quality of the fits for the changes in tuning curve properties is indicated by high R^2 values (in the range of 0.70 for the response change to 0.90 for the shift in preferred orientation). On the other hand, linear prediction failed to predict the shifts in perception ($R^2 = 0.32$, Figure 5.7A, left panel). The worse prediction performance might reflect the fact that the perception shift is a more indirect quantity. Whereas the tuning curve changes are a property of individual neurons, the perception shift is related to the population responses. Response suppression and shifts in preferred orientation generally have an opposite effect on the population response (Jin et al., 2005), possibly making it more difficult to predict the perception shift directly from the network parameters.

The way in which the model parameters determine each model property can be understood by examining the regression coefficients b_{EE}, b_{IE}, b_{II} and b_{EI} (corresponding to the model parameters U_{EE}, U_{IE}, U_{II} and U_{EI}). These coefficients are shown in Figure 5.7B. The sign of each coefficient reveals the direction of the effect that a model parameter has on a model property. The value of the coefficient indicates how much the predicted property is expected to increase (decrease) when the corresponding model parameter increases (decreases), holding all the other model parameters constant. Consider the response change (Figure 5.7B, third panel) as an example: b_{EE} is negative, indicating that a higher value of U_{EE} leads to smaller responses (more suppression). This is expected because larger values of U_{EE} lead to more depression of excitatory–excitatory connections, resulting in response suppression. If, on the other hand, the depression of connections from inhibitory neurons to excitatory neurons increases (higher values of U_{EI}), recurrent inhibition of excitatory neurons decreases and the responses of excitatory neurons increase. This is reflected in a large positive value for the coefficient b_{EI}. Other coefficients are less intuitive: the depression parameter for excitatory to inhibitory

connections (U_{IE}), for example, has a negative contribution to the preferred orientation shift while the depression parameter for inhibitory to excitatory connections (U_{EI}) has a positive contribution (Figure 5.7B, second panel).

In summary, linear regression analysis provides a quantitative description of how each model parameter influences each tuning property of the model network. This description is useful because it reveals the effects that synaptic depression exerts on recurrent processing.

5.4. Mechanisms underlying orientation tuning changes

To this point, we explored the entire space of possible network models by varying the synaptic depression parameters and analyzed the average changes that adaptation induces in the network models. Now we take closer look at one typical admissible model (the one marked by a red cross in Figure 5.4A) and examine its response properties in more detail. We first analyze the adaptive changes of the population response and of the tuning curves of individual cells. Finally, we focus on the mechanism underlying the stronger adaptive shifts close to pinwheel centers.

Population responses of the network

Figure 5.8 shows the population response of the selected admissible network model to two different test stimuli, after adapting to a stimulus with 0° orientation. If the test stimulus has the same orientation as the adapting stimulus (Figure 5.8A,B), the adaptive changes in the population response are symmetric around this orientation. The dominant effect is a suppression of the firing rate responses of neurons with a preferred orientation close to the adapting orientation, resulting in a suppression of the population response. This is true for orientation domain cells (0.6 < map OSI < 0.9) as well as for pinwheel center cells (map OSI < 0.3). If, on the other hand, the orientation of the test stimulus differs from the adapting stimulus (Figure 5.8C,D), the population response becomes asymmetric. Furthermore, the cell that fires strongest in response to the test stimulus has a preferred orientation (as determined in the control condition) that is different from the test stimulus orientation of 33.75°. The reason is that neurons with a preferred orientation close to the adapting orientation (0°) fire strongly in the adaptation phase. Hence, synapses originating from these neurons depress. Since neurons with a preferred orientation close to the adapting orientation receive many inputs from these depressed synapses they show more changes in their responses. That is why neurons with preferred orientation close to the adapting orientation of 0° show more changes in their responses than neurons close to the orientations −90° and +90°.

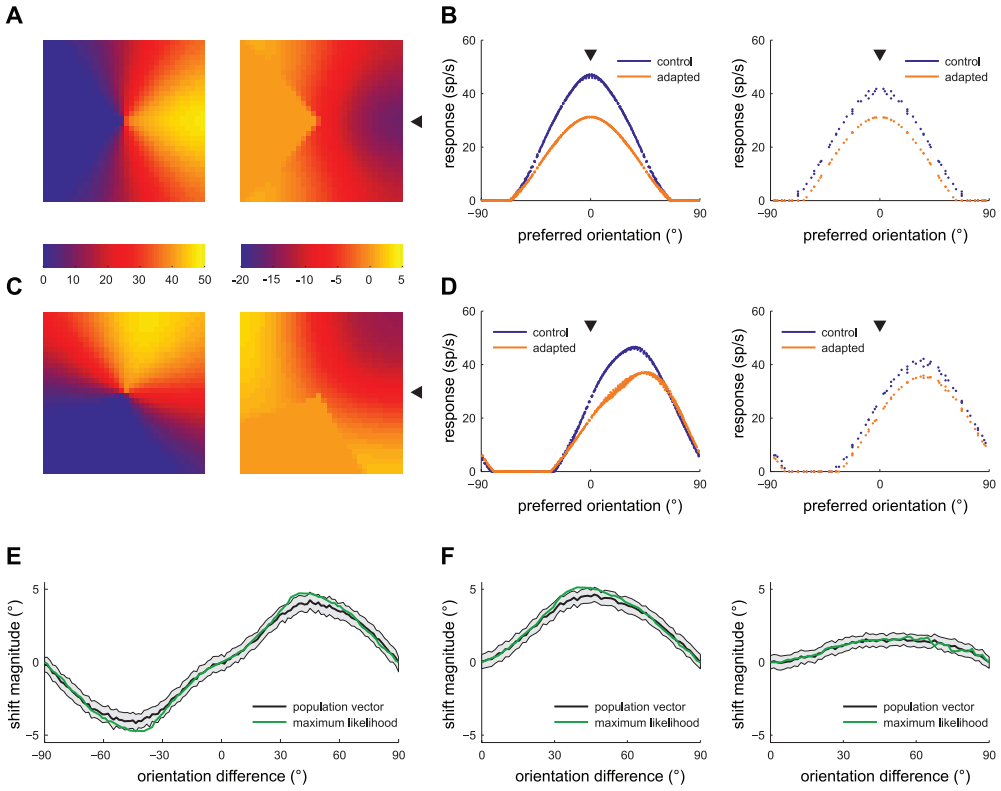


Figure 5.8 | Population responses and perception shifts in a selected admissible network model. **A** | Firing rate response of the network to a stimulus of 0° orientation in the control condition, i. e., before adaptation (on the left), and change of the network response through adaptation to a 0° stimulus (on the right). The triangle denotes the region in the map where the neurons' preferred orientation is equal to the orientation of the adapting stimulus (0°). The response change (on the right) is the difference in firing rate of each cell after and before adaptation, i. e., negative (positive) values correspond to response suppression (enhancement). Scale bars at the bottom indicate the firing rate (left) and the difference in firing rate (right). For clarity only the responses of neurons in one out of the four pinwheels are shown. **B** | Population responses to the 0° stimulus as a function of the preferred orientation of neurons, for neurons in orientation domains (left) and for neurons close to pinwheel centers (right). Each point is the response of one neuron. **C, D** | Responses of the network to a stimulus of 33.75° orientation, before and after adaptation to a 0° stimulus. **E** | Difference between perceived and true stimulus orientation as a function of the difference between test stimulus and adapting stimulus. The perceived orientation was estimated from the population responses using the maximum-likelihood method (green line) and the population-vector method (black line). The shaded area depicts the SD around the estimate from the population-vector method. **F** | Shift in perception evaluated only for cells in orientation domains (left) and close to pinwheel centers (right). The depression parameters of the selected network model were $\log U_{EE} = -0.60$, $\log U_{IE} = -2.10$, $\log U_{II} = -1.35$ and $\log U_{EI} = -1.05$.

As can be seen in Figure 5.8D, the population response of orientation domain cells shifts repulsively away from the adapting stimulus while the population response of pinwheel cells shifts only marginally. This is surprising: since pinwheel cells show more pronounced shifts in preferred orientation than orientation domain cells, one would expect that the pinwheel population response also shows a stronger shift. However, this intuition is wrong. Repulsive shifts of tuning curves cause attractive rather than repulsive shifts of the population response (see e. g., Jin et al., 2005). Suppression of tuning curves near the adapting orientation, on the other hand, causes repulsive shifts of the population response. In other words, suppression and repulsive tuning curve shifts have opposite effects on the population response. If suppression and repulsive tuning curve shifts coexist, as is the case here, a repulsive population response occurs when suppression dominates. Thus, taking the subset of cells that undergo larger repulsive tuning curve shifts necessarily results in a weaker population shift.

To quantify the impact of the adaptive reorganization of the population response on "perception", we estimated the hypothetical perceived orientation from the population responses using the maximum-likelihood and the population-vector method (Figure 5.8E). Both methods led to similar results that are in qualitative agreement with the experimentally observed perception shifts (cf. Figure 5.1B). Rather than considering the whole population of cells, the "perceived" orientation can also be calculated for subpopulations of orientation domain and pinwheel cells. As can be seen in Figure 5.8F, the perceptual shift in the population of orientation domain cells is very similar to the shift observed in the entire population while the perceptual shift in the population of pinwheel cells is much weaker. The weaker "perception" change calculated from the pinwheel cells is the consequence of the more pronounced repulsive tuning curve shifts of pinwheel cells as compared to orientation domain cells, as has been discussed above.

Orientation tuning properties of the network

The adaptive changes of the tuning curves of individual cells are shown in Figure 5.9, for all cells of the selected admissible network model. Tuning curves shift repulsively away from the adapting stimulus (Figure 5.9A). The largest shifts are observed for cells around the pinwheel center. This can be seen more clearly when plotting the shifts as a function of map OSI (Figure 5.9A, middle panel; see Figure 5.1F for the experimental data). We also plotted the shifts in preferred orientation against the difference between adapting and control–preferred orientation (Figure 5.9A, right panel; see Figure 5.1C for the experimental data). Interestingly, pinwheel cells exhibit larger shifts for a broader range of orientation differences, as has been found experimentally (Dragoi et al., 2001). This experimental finding was not used in the selection of admissible network models.

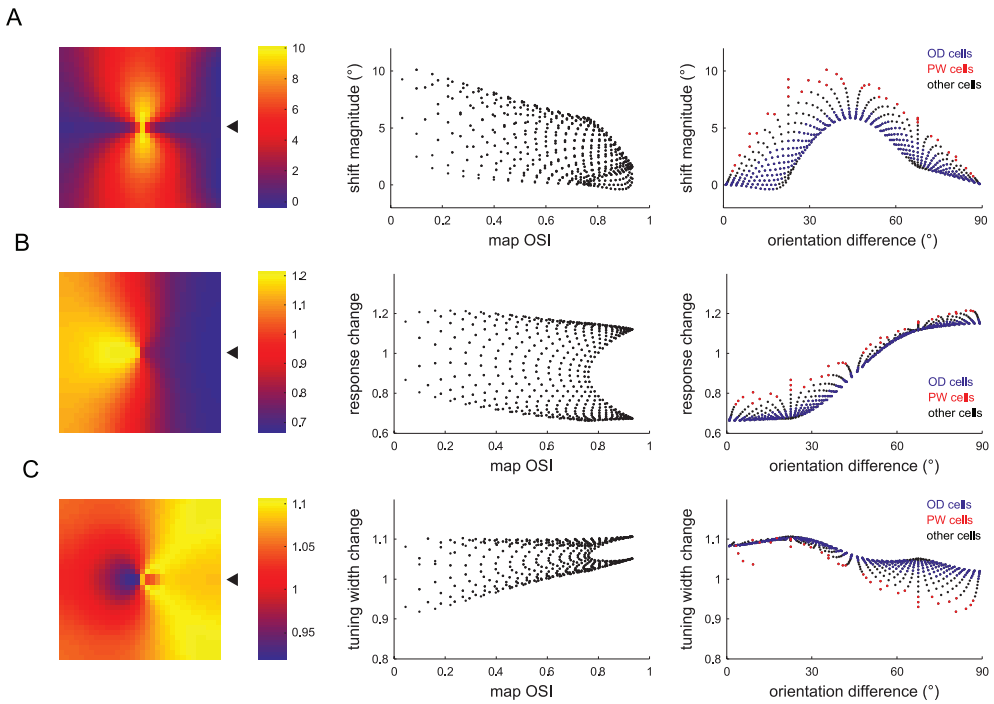


Figure 5.9 | Adaptive changes of orientation tuning curves for cells in a selected admissible network model. **A** | Change in the preferred orientation after adaptation is plotted for each cell at the corresponding location in the orientation map (left), as a function of map OSI (middle) and as a function of the difference between testing and adapting stimulus (right). Positive numbers indicate repulsive shifts away from the adapting stimulus. The triangle indicates the position in the map where the neuron's preferred orientation is equal to the adapting stimulus (0°). For clarity only the responses of neurons in one out of the four pinwheels are shown. Each point in the middle and the right panel corresponds to the tuning curve shift of one neuron. **B** | Adaptation-induced changes in maximum firing rate relative to the control condition. Changes were calculated for each neuron as the ratio of the maximum firing rate in response to any of the presented stimuli after adaptation and the maximum response in the control condition (before adaptation). Values greater (smaller) than 1 indicate facilitation (suppression). **C** | Adaptive changes in the width of tuning curves, measured as the half-width at half-height. The changes were calculated for each neuron as the ratio of the tuning width after and before adaptation. Values greater (smaller) than 1 indicate broadening (sharpening) of a cell's tuning curve through adaptation. The depression parameters of the selected network model were $\log U_{EE} = -0.60$, $\log U_{IE} = -2.10$, $\log U_{II} = -1.35$ and $\log U_{EI} = -1.05$.

While the tuning curve shifts depend on the map OSI as well as on the orientation difference, the magnitude of the response changes (Figure 5.9B) is mainly determined by the difference between adapting and control–preferred orientation and depends only weakly on the map OSI. The dependency on the orientation difference is very similar between the model and the experimental data (see Figure 5.1D).

Finally, the selectivity of tuning curves (Figure 5.9C) also changes as in the experimental data (Figure 5.1E): tuning curves of neurons with preferred orientation close to the adapting orientation show broadening while tuning curves of neurons far away of the adapting orientation show sharpening. The adaptive sharpening is more pronounced in cells close to pinwheel centers. It would be interesting to test if this characteristic difference between orientation domain and pinwheel center can also be found in the experimental data. However, the tuning width changes are rather small so that the 22.5° spacing of the stimuli used in Dragoi et al. (2001) could make it difficult to detect differences between pinwheel and orientation domain cells.

In summary, we find a good agreement of experimental data and model results, for the adaptive changes in the perceived orientation and the tuning curve properties. This agreement goes beyond the criteria used for the selection of admissible models, that were based on the comparison of average values rather than the detailed comparison of the dependencies on orientation difference and map location.

Dependence of adaptation-induced changes in orientation tuning on map location

We have shown that the network models – depending on the synaptic depression parameters – can show more pronounced adaptation-induced changes close to pinwheel centers, as has been found experimentally. In oder to shed light on the mechanisms that underlie the stronger effect of adaptation on pinwheel cells, we now examine how the tuning curve shifts of model neurons depend on the local cortical interactions.

To illustrate the map-dependence of adaptation, we picked out two neurons: one "pinwheel neuron" (PW) close to the pinwheel center and one "orientation domain neuron" (OD) far from the pinwheel center (Figure 5.10A). While both neurons have the same preferred orientation (22.5°), the map OSI of the PW neuron is much lower than the map OSI of the OD neuron (0.13 vs. 0.83; Figure 5.10B). Despite this difference in the local neighborhood, the firing rates of both neurons are sharply tuned (Figure 5.10C, D; cf. Mariño et al., 2005). The firing rate tuning curve of the PW cell shows a pronounced shift away from the adapting orientation, whereas the shift of the OD cell is only modest (shift magnitudes 8.5° vs. 2.0°). In the control condition, the recurrent as well as the afferent synaptic inputs of both the OD and the PW neuron are aligned to the preferred orientation of the firing rate tuning curve (Figure 5.10E, F). The synaptic inputs of the

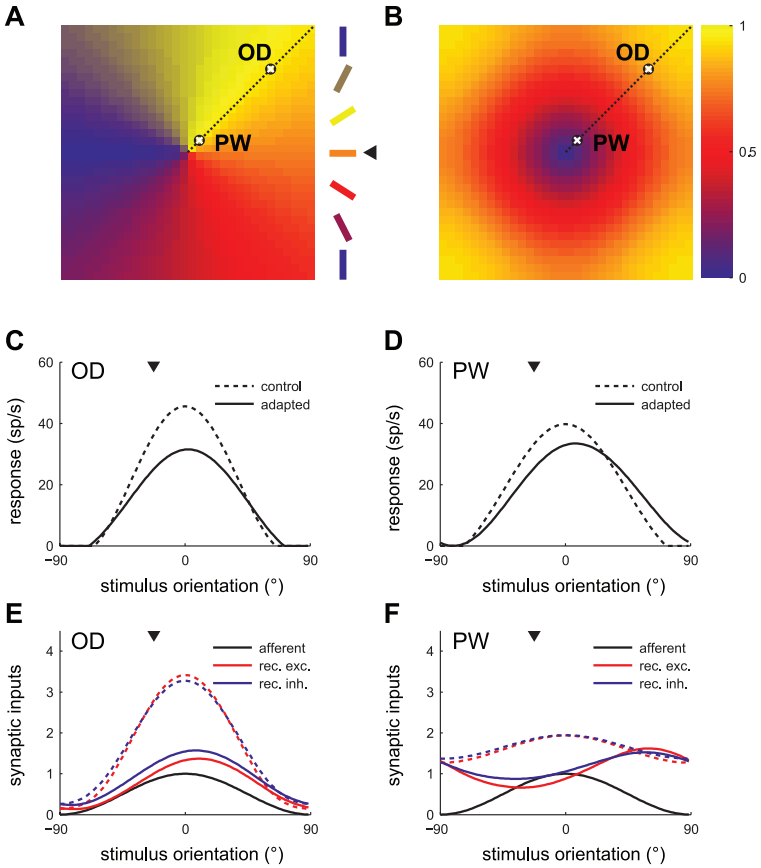


Figure 5.10 | Adaptive changes in orientation tuning of individual cells in the selected network model. **A** | Part of the artificial orientation preference map showing one out of the four pinwheels. The color of each pixel denotes the preferred orientation (before adaptation) of cells located at that pixel (colored bars at the right). OD (PW) denotes a representative cell from an orientation domain (close to a pinwheel center). Both cells have a preferred orientation of 22.5°, indicated by the dotted line. **B** | Map OSI of the artificial orientation map. The map OSI of the cell at location PW is much lower than the map OSI of the cell at location OD (0.13 vs. 0.83). **C, D** | Tuning curve of the firing rate of the OD cell (C) and the PW cell (D) before (dotted line) and after adaptation (solid line). The tuning curves and the triangle denoting the orientation of the adapting stimulus (0°) were shifted such that the peak response before adaptation is at 0°. **E, F** | Tuning curve of the recurrent excitatory (red lines) and the recurrent inhibitory (blue lines) inputs of the OD cell (E) and the PW cell (F). Dotted (solid) lines depict the tuning curves before (after) adaptation. All input tuning curves were normalized to the maximum of the afferent input. The afferent input (black lines) has the same strength for both pinwheel and orientation domain cells. Tuning curves were shifted as in (C) and (D). The depression parameters of the selected network model were $\log U_{EE} = -0.60$, $\log U_{IE} = -2.10$, $\log U_{II} = -1.35$ and $\log U_{EI} = -1.05$.

OD neuron are well-tuned, reflecting the fact that the OD neuron mainly integrates synaptic inputs from neurons with similar preferred orientations. Through adaptation, recurrent synapses depressed and the excitatory and inhibitory inputs to the OD neuron decreased (Figure 5.10E). Since the OD neuron and its neighboring neurons prefer similar orientations, adaptation causes only little inhomogeneity in the synaptic inputs. The PW neuron, however, integrates synaptic inputs from neurons with a broad range of preferred orientations. After adaptation, only synapses originating from neurons with preferred orientations close to the adapting orientation show strong depression. Hence, the synaptic inputs of the PW neuron are completely modified after adaptation (Figure 5.9F): synaptic inputs decrease for test orientations close to the adapting orientation and increase for orientations far away from the adapting orientation. The preferred orientation of the input tuning curves after adaptation is different from the preferred orientations of the input and the firing rate in the control condition. Furthermore, it seems that the recurrent inputs became slightly more orientation-selective.

Figure 5.11 shows the map-dependences of the adaptation-induced changes in orientation tuning for neurons with preferred orientation 22.5° and map OSIs between 0.04 and 0.93 (cf. the dotted line in Figure 5.10A). The figure confirms that the shift in preferred orientation of the firing rate and the synaptic inputs decreases with the distance to the pinwheel center (Figure 5.11A, B), and that the tuning of synaptic inputs indeed got more selective close to the pinwheel center (Figure 5.11C). It has been shown experimentally (Dragoi et al., 2000, 2001), that the shifts of the firing rate tuning curve are characterized by suppression of the responses on the flank of the tuning curve towards the adapting stimulus (near flank) and by facilitation of the responses on the flank of the tuning curve away from the adapting stimulus (far flank). Figure 5.11D–E shows that in the model adaptation also led to a suppression of firing rates for the preferred orientation (Figure 5.11D, top panel) and the near flank (Figure 5.11E, top panel) of the tuning curve. For the far flank of the tuning curve (Figure 5.11F, top panel), the magnitude of the changes in firing rate is related to the map OSI: responses of neurons close to the pinwheel center are facilitated, whereas responses of neurons in the orientation domain are slightly suppressed. Through adaptation, the magnitude of recurrent excitatory and inhibitory inputs (Figure 5.11D–E, bottom panel) decreased at the center as well as on the near and far flanks of the tuning curves. At the center and on the near flank recurrent inhibition dominates recurrent excitation after adaptation, leading to the observed decrease in firing rate. At the far flank, however, recurrent excitation became slightly stronger than recurrent inhibition, for small map OSIs, leading to response facilitation. In summary, we find that the broader orientation distribution of intracortical inputs to neurons close to pinwheel centers results in larger adaptive changes in orientation tuning of synaptic inputs and firing rates of these neurons.

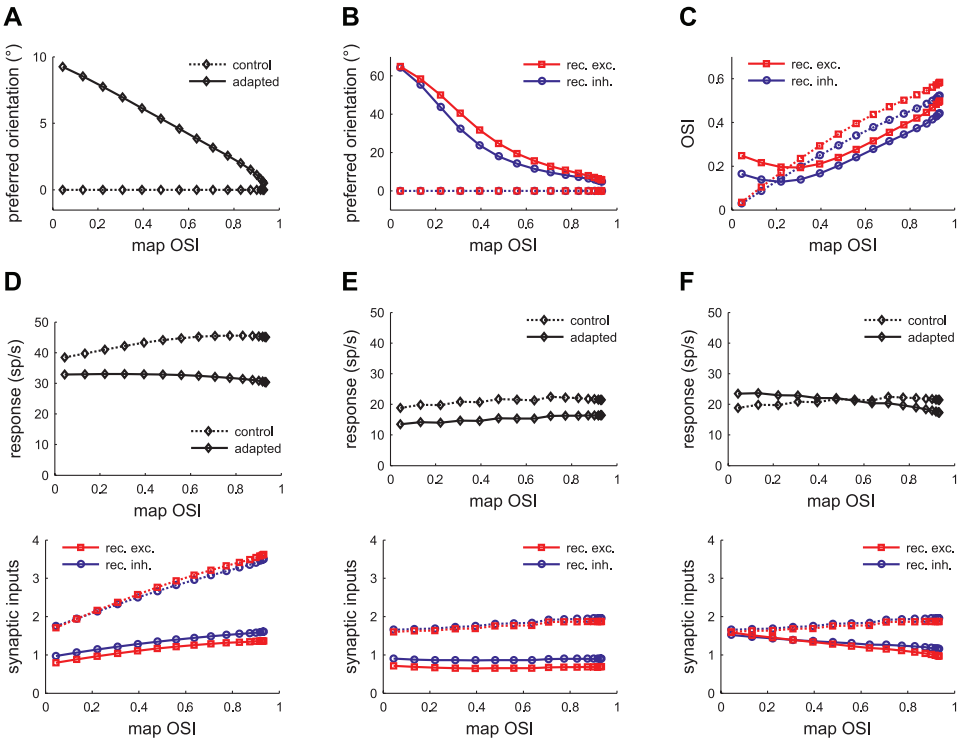


Figure 5.11 | Adaptive changes in orientation tuning properties for cells with 22.5° preferred orientation as a function of map OSI. The position in the orientation map of these cells in given by the dotted lines in Figure 5.10A, B. **A** | Preferred orientation of the cells in the adapted (solid line) and the control (dotted line) condition. The value of the preferred orientation in the control condition (22.5°) was subtracted before plotting. Preferred orientation was calculated from the firing rates using vector averaging. **B** | Difference between the preferred orientation of the recurrent excitatory and inhibitory inputs and the preferred orientation of the firing rate in the control condition. In the control condition, the preferred orientation of the recurrent excitatory and inhibitory inputs are both aligned to the firing rate (dotted line at 0°). In the adapted condition, the preferred orientation of recurrent inputs is shifted away from the preferred orientation (solid lines). **C** | Selectivity of the recurrent inputs before (dotted lines) and after (solid lines) adaptation, measured using the OSI. **D** | Firing rate (top) and synaptic inputs (bottom) in response to the preferred stimulus in the control condition (22.5°), before (dotted lines) and after (solid lines) adaptation. **E** | Firing rate (top) and synaptic inputs (bottom) on the "near" flank of the tuning curve, i. e., the flank of the tuning curve where the adapting stimulus is located. Shown are the responses to a −16.5° stimulus which evokes a firing rate before adaptation that is approximately 50 % of the maximum response. **F** | Firing rate (top) and synaptic inputs (bottom) on the "far" flank of the tuning curve, i. e., the opposite flank with respect to the adapting stimulus. Again, as in (E), the presented stimulus orientation was chosen so that the firing rate before adaptation is approximately 50 % of the maximum response (stimulus orientation 61.5°). The adapting orientation was 0°. The depression parameters of the selected network model were $\log U_{EE} = -0.60$, $\log U_{IE} = -2.10$, $\log U_{II} = -1.35$ and $\log U_{EI} = -1.05$.

5.5. Discussion

We used mechanistic models of orientation tuning to investigate the neural substrates that underlie temporal context effects (orientation adaptation) in primary visual cortex. As a plausible underlying mechanism, we focused on intracortical synaptic depression and used experimental data concerning both adaptive changes in orientation tuning (Dragoi et al., 2000, 2001; Jin et al., 2005) and changes in perception (Mitchell and Muir, 1976; Clifford et al., 2000) to constrain the strengths of the synaptic depression parameters. Consistent with *in vitro* data, our results provide strong evidence for those network models where excitatory-to-excitatory synapses depress more than excitatory-to-inhibitory synapses and inhibitory-to-excitatory synapses.

Exploration of the model parameter space and model selection

We simulated more than 50 000 network models with different combinations of synaptic depression strengths, and selected approximately 1000 models that were consistent with the physiological and psychophysical experimental data. Using connected component analysis, we showed that this set of admissible models forms a single connected "island" in parameter space. This means that the model class is robust against changes in the synaptic depression parameters: given an admissible network model its neighbors in parameter space are most likely also part of the "island". Thus, when the depression parameters are altered slightly, the behavior of the network will be preserved and, importantly, remain consistent with the experimental data. Moreover, the finding that there is only one connected component of admissible models in parameter space suggests an underlying relationship between the four parameters governing synaptic depression. Indeed, we found that the model parameters are correlated. For instance, the strengths of synaptic depression of excitatory synapses targeting excitatory and inhibitory neurons can vary in a broad range, but in order to achieve tuning curve changes consistent with experimental data, the synapses targeting inhibitory neurons must undergo less depression than the synapses targeting excitatory neurons. This finding is in agreement with *in vitro* data showing that connections from pyramidal cells to fast spiking interneurons show little depression or even facilitation (Thomson and Deuchars, 1997).

Linear regression analysis provides a compact description of how the model properties vary as a function of the model parameters. For instance, it shows that increasing the depression strength of excitatory-to-excitatory synapses leads to larger repulsive tuning curve shifts, whereas increasing depression strength of excitatory-to-inhibitory depression leads to smaller shifts. This analysis complements the intuitive visualization

provided by dimensional stacking plots, and helps to understand how each model parameter influences processing in the recurrent network, giving rise to tuning curve changes. In principle, this description could also be used to predict the changes we would expect for altered synaptic plasticity in real cortical networks.

Our systematic exploration of the network model parameter space is in contrast to the traditional hand tuning of model parameters. Of course, the robustness of hand-tuned models can be investigated by performing a sensitivity analysis around the chosen parameter combination; however, our approach directly yields the complete space of models able to account for the experimental data. This allows determining the robustness (i. e., how narrowly parameters have to be tuned to produce a specific behavior) and sensitivity to parameter changes of the network under study. Similar parameter explorations have been used to characterize the conductance space of detailed single cell models (Goldman et al., 2001; Prinz et al., 2003; Taylor et al., 2009) or the space of synapse strengths and neuron properties of a small networks composed of three neurons (Prinz et al., 2004; Marder et al., 2007), but applications to larger networks are rare.

Interestingly, our set of admissible models shares a common property with the parameter spaces of single cell modeling studies: in both single cell and network models many different parameter combinations can give rise to similar patterns of neural activity. This has been confirmed experimentally in the case of the pyloric network of the crustacean stomatogastric ganglion, where identified neurons have similar response properties from animal to animal although the conductance densities of specific ion channels show variations as large as 2–5 fold from animal to animal (Marder et al., 2007). Together, these results suggest that the parameters of neurons and networks need to co-vary, but they need not be tightly tuned to specific values. This is important for understanding the robustness and sensitivity of neural systems.

Sensitivity to modulations of the balanced recurrent operating regime

We showed that a network model operating in a highly recurrent regime – with strong, balanced recurrent excitation and inhibition – reproduces the experimentally observed, adaptation-induced changes in orientation tuning. But is this particular operating regime really a necessary precondition – or could a network operating in a different regime, with appropriately chosen synaptic depression parameters, lead to similar results? We repeated the parameter exploration and the selection of admissible models for networks with weaker recurrent excitatory connections and for networks with stronger recurrent inhibitory connections. A decrease of the excitatory connection strength S_{EE} by 10 % leads to an average tuning curve shift of 3.5° for the optimal choice

of synaptic depression parameters (i. e., for the depression parameters yielding the highest tuning curve shifts for the network with reduced S_{EE}) as compared to 5.5° for the balanced network. Proportional to a decrease of S_{EE}, the tuning curve shifts decrease further (20 % decrease of S_{EE}: 2.2° average shift, 30 % decrease of S_{EE}: 1.3° average shift). A similar relationship holds for an increase of the connection strength from excitatory to inhibitory neurons (S_{IE}): an increase of S_{IE} of 10 % (20 %, 30 %) leads to an average tuning curve shift of 3.8° (2.9°, 2.7°). Thus, only the balanced recurrent regime can replicate the magnitude of the experimentally observed tuning curve shifts (average shift in the experimental data: 5.8°, balanced recurrent network: 5.5°).

We have shown in the two previous chapters (see also Wimmer et al., 2009; Stimberg et al., 2009) that only the balanced recurrent operating regime is consistent with experimentally measured subthreshold and superthreshold orientation tuning properties and with the temporal response characteristics of V1 neurons. Recurrent processing in this regime leads to a slight sharpening of orientation tuning and to an amplification of the firing rate response. One might argue that this is a rather modest computational contribution of the extensive recurrent network. However, here we demonstrated a further benefit of the highly recurrent operating regime: it makes network processing sensitive to modulation of intracortical synaptic strengths. In fact, the modification of synaptic strengths, caused by adaptation, can be seen as a local alteration of the operating regime. Through this alteration, excitation and inhibition are no longer in balance, the recurrent synaptic inputs to V1 neurons are reorganized, and orientation tuning properties modified. Based on this, one may expect the local recurrent circuit to play a critical role in mediating other forms of bottom-up or top-down modulations, such as attention (McAdams and Maunsell, 1999), spatial context (Seriès et al., 2003) or internal states (Sharma et al., 2003) as well.

Biological plausibility of the model

One limitation of our study concerns the experimental data used to constrain the parameters of the network model: the tilt aftereffect was measured in human psychophysics experiments, the physiological tuning curve changes, on the other hand, were measured in V1 of anesthetized cats. However, orientation adaptation has also been documented in monkey (Müller et al., 1999; Dragoi et al., 2002), which, given the similarities of cortical organization across species, makes it likely that orientation adaptation of V1 neurons also occurs in humans.

An assumption of our model is that adaptation causes depression of intracortical synapses. There is ample experimental evidence for synaptic plasticity *in vitro* (Abbott et al., 1997; Tsodyks and Markram, 1997; Varela et al., 1997). Furthermore, both short-

term synaptic depression and orientation adaptation occur at multiple time scales, from tens of milliseconds to seconds (synaptic depression: Galarreta and Hestrin, 1998; Varela et al., 1999; Zucker and Regehr, 2002; orientation adaptation: Felsen et al., 2002; Müller et al., 1999; Dragoi et al., 2000, 2002). However, the strength of intracortical synaptic depression *in vivo* is unclear: Reig et al. (2006) found that synaptic depression in cat V1 *in vivo* is comparable to synaptic depression in slowly oscillating slices but weaker than in silent slices. While a study in rat barrel cortex found no evidence for depression of intracortical synapses (Chung et al., 2002), another recent *in vivo* study in cat visual cortex is consistent with depression of intracortical synapses (Boudreau and Ferster, 2005).

In order to account for the adaptive changes in orientation tuning, the depression parameters of different types of synapses have to obey certain relationships in our model. First, in almost all admissible models, depression must be stronger for the excitatory synapses onto excitatory neurons than for excitatory synapses onto inhibitory neurons. This is in agreement with *in vitro* data (Thomson and Deuchars, 1997). Second, depression of excitatory synapses onto excitatory neurons is stronger than depression of inhibitory synapses onto excitatory neurons in all admissible models. This differential (asymmetric) depression of excitatory and inhibitory synapses is again consistent with experimental data (Galarreta and Hestrin, 1998; Varela et al., 1999). Thus, the network models that achieve plausible tuning curve changes rely on plausible underlying synaptic depression strengths. However, we cannot rule out that other mechanisms, such as reduction in the neuronal excitability (Carandini and Ferster, 1997; Sanchez-Vives et al., 2000a,b) also contribute to orientation adaptation in V1 neurons (as in the case of contrast adaptation).

For simplicity, afferent synapses do not undergo synaptic depression in our model. This can be justified by experimental data indicating that synaptic depression of thalamocortical synapses is weaker than depression of intracortical synapses (Yoshimura et al., 2000). Furthermore, there is experimental evidence (*in vivo*) that thalamocortical synapses are already maintained at high levels of depression by spontaneous activity (i. e., there is little additional depression; Boudreau and Ferster, 2005).

Our general assumption that the intracortical circuit mediates the dynamic modification of orientation tuning is supported by experimental evidence. First, adaptation-induced tuning curve shifts are not spatial-phase dependent (Felsen et al., 2002; Dragoi et al., 2000). This is inconsistent with a feed-forward orientation tuning model in which adaptation reduces the afferent input (e. g., by suppression of the responses of LGN cells or by depression of thalamocortical synapses), because this model predicts that changes in orientation tuning of V1 cells depend on the spatial phase of the stimulus (Felsen et al., 2002). Second, adaptation-induced shifts of cells close to pinwheel centers are

larger compared to cells in orientation domains (Dragoi et al., 2001). It is unclear how this map-dependence could be explained in a feed-forward model that does not take into account lateral intracortical interactions. Together, these findings strongly suggest that the intracortical circuit provides the substrate for orientation adaptation.

Conclusions

Depending on the strength of the synaptic depression for the different types of connections (excitatory and inhibitory synapses targeting excitatory and inhibitory neurons), adaptation in our network model leads to suppression or enhancement of firing rates, broadening or sharpening of tuning widths, and repulsive or attractive shifts in preferred orientation. For a given parameter regime, the exact modification that a tuning curve undergoes is determined by the difference between the preferred orientation of the neuron and the orientation of the adapting stimulus, as well as by the position of the neuron in the orientation preference map. We demonstrated that a range of models – with realistic synaptic depression strengths – is consistent with the experimental data on orientation tuning changes in V1; specifically, the models are consistent with repulsive post-adaptation shifts of orientation tuning curves. However, differently parameterized models can also reproduce attractive shifts of orientation tuning curves, as have been observed in area MT (Kohn and Movshon, 2004). We thus predict that if synaptic depression underlies orientation adaptation in V1 and in MT, synaptic depression strengths and their relationship should be different in both areas.

We also investigated how the tuning curve shifts of model neurons depend on a neuron's position in the orientation preference map. As has been found experimentally (Dragoi et al., 2001), adaptation-induced shifts of tuning curves are more pronounced close to pinwheel centers. In the model, the larger shifts naturally arise from the spatial isotropy of local recurrent connections (Mariño et al., 2005): Since cells close to pinwheel centers have neighboring cells covering the whole range of preferred orientations they receive broadly tuned recurrent inputs, while cells in orientation domains receive sharply tuned recurrent inputs. The broadly tuned recurrent inputs make pinwheel cells more sensitive to modulations. Through adaptation, synaptic inputs of pinwheel cells decrease for test orientations close to the adapting orientation and increase for orientations away from the adapting orientation, causing strong repulsive shifts of the firing rate tuning curves. Orientation domain cells, however, receive strong recurrent inputs only from cells with similar orientation preferences; adaptation can therefore not lead to strong inhomogeneities in synaptic inputs, and the main effect of adaptation is response suppression rather than large tuning curve shifts. Our results suggest that top-down influences and spatial context effects too will be modulated through the local cortical

circuit and therefore be related to a neuron's position in the orientation map.

In order to determine how the adaptive reorganization of tuning curves influences the "perceived" orientation, we analyzed the population responses of the network after adaptation. We confirmed previous results from a theoretical study (Jin et al., 2005), showing that response suppression around the adapting stimulus and repulsive tuning curve shifts have opposite effects on the perceptual bias: suppression leads to repulsive perception shifts, and repulsive tuning curve shifts lead to attractive perception shifts. An interesting consequence of this antagonistic effect is that if we read out the perceived orientation taking into account only cells close to pinwheel centers, the perceptual bias (the tilt aftereffect) is weaker than if we take into account only cells in orientation domains. Thus, while individual pinwheel cells show larger shifts than orientation domain cells, they are actually contributing less to the change in perception.

Recapitulation

We set up a two-dimensional firing rate model of orientation tuning in V1 and asked how the strength of depression of the different types of intracortical synapses (excitatory and inhibitory synapses onto excitatory and inhibitory neurons) influences recurrent processing. We simulated orientation adaptation and compared the changes in orientation tuning predicted by the model to physiological and psychophysical data.

Our modeling results support the following conclusions: (i) Synaptic depression is a plausible mechanism underlying orientation adaptation in V1. We found that networks in a compact region of the model parameter space give rise to changes in orientation tuning and perceptual read-out that are consistent with experimental findings. The synaptic depression parameters within these region are consistent with differential synaptic depression found *in vitro*. (ii) The magnitude of experimentally observed shifts in orientation tuning provides strong support for a highly recurrent, balanced operating regime of the network, in which adaptation perturbs the balance between excitation and inhibition. We showed that for network models with decreased excitation or increased inhibition the adaptive shifts are diminished. (iii) The model also provides a simple, mechanistic explanation for the larger shifts close to pinwheel centers: pinwheel cells receive recurrent inputs tuned to a broader range of orientations than orientation domain cells. They are thus more sensitive to response changes of their neighboring cells. This result also stresses the importance of taking into account the local cortical architecture through which response modulations in individual neurons, be it through intrinsic, feed-forward, long-range, or top-down mechanisms, can give rise to a reorganization of the responses of a whole neural population.

Here, we have focused on a mechanistic explanation of adaptive changes in orientation tuning and on the tilt aftereffect. It has recently been found (Gutnisky and Dragoi, 2008) that adaptation also influences the structure of noise correlations across the whole population of V1 neurons. Thus adaptation affects population coding beyond a mere reorganization of firing rate tuning curves. The models that we have developed here (see also Chapter 3 and Chapter 4) are consistent with a large range of experimental data and can serve as a solid starting point for investigating population coding, including the role of adaptive changes in correlations and the underlying mechanisms.

6. Adaptation and selective information transmission in the cricket auditory neuron AN2

This chapter is based on Wimmer et al. (2008).

Abstract

Sensory systems adapt their neural code to changes in the sensory environment, often on multiple time-scales. Here, we report a new form of adaptation in a first order auditory interneuron (AN2) of crickets. We characterize the response of the AN2 neuron to amplitude-modulated sound stimuli, and find that adaptation shifts the stimulus-response curves towards higher stimulus intensities, with a time constant of 1.5 seconds for adaptation and recovery. We then address the question whether adaptation leads to an improvement of the signal's representation and compare the experimental results with the predictions of two competing hypotheses: infomax, which predicts that information conveyed about the entire signal range should be maximized, and selective coding, which predicts that "foreground" signals should be enhanced while "background" signals should be selectively suppressed. We test how adaptation changes the input-response curve when presenting signals with two or three peaks in their amplitude distributions, for which selective coding and infomax predict conflicting changes. By means of Bayesian data analysis, we quantify the shifts of the measured response curves and also find a slight reduction of their slopes. These decreases in slopes are smaller, and the absolute response thresholds are higher than predicted by infomax. Most remarkably, and in violation of the infomax principle, adaptation actually reduces the amount of encoded information when considering the whole range of input signals. The response curve changes are also not fully consistent with the selective coding hypothesis: while less information is transmitted about signals with lower intensity the amount of information conveyed about the loudest part of the signal does not increase as predicted but remains nearly constant.

6.1. Introduction

Efficient encoding of natural signals is one of the major tasks sensory pathways have to accomplish. In order to do this, neural representation should be matched to the

relevant part of incoming signals. Statistical properties of incoming signals are highly variable in a natural environment (e. g., the mean light level changes dramatically from a sunny region to a dark forest) but are mostly changing slowly over time (Nelken et al., 1999). Since the neural representation in sensory cells is limited to a certain range and resolution, the principle of efficient coding suggests that the nervous system should continually adapt its responses to changing statistical properties of the stimuli (Barlow, 1961). Firing rate adaptation changes the input-response curves of neurons in sensory pathways and has been shown to provide a mechanism for the adjustment of the encoding scheme in multiple systems (Baccus and Meister, 2002; Brenner et al., 2000; Dean et al., 2005; Fairhall et al., 2001; Maravall et al., 2007; Nagel and Doupe, 2006; Solomon et al., 2004). How the input-response curve is altered in response to a given stimulus should depend on what the relevant information is in the given context.

Here, we want to explore the response properties of a single cell (AN2) in the auditory pathway of crickets and test for two different principles that have been proposed to underlie adaptation of the input-response curve: the principle of maximum information preservation (infomax, Linsker, 1988) and that of selective coding (Sobel and Tank, 1994). The AN2 neuron provides an ideal model for studying the computational principles underlying adaptation, since (i) it receives direct input from auditory receptors and local interneurons at the first processing level (Hennig, 1988), (ii) on present evidence, it constitutes the only ascending representation of the auditory environment in the high frequency channel and thus a bottleneck for information transmission to higher centers (Horseman and Huber, 1994a,b; Wohlers and Huber, 1982), and (iii) it has a clear behavioral role because it is intimately involved in evasive behavior in response to ultrasonic signals (Marsat and Pollack, 2006; Nolen and Hoy, 1984; Pollack and Martins, 2007). Several time constants of adaptation in the range from below 100 ms to several seconds are known for the receptor cells (Givois and Pollack, 2000), local interneurons (Pollack, 1988) and the ascending neurons (Benda and Hennig, 2008; Samson and Pollack, 2002) in this model system. Since auditory processing at the stage of the AN2 neuron is mainly feed-forward, adaptation is likely driven by the stimulus only rather than by task-dependent top-down processes.

The above mentioned principles lead to conflicting hypotheses about changes of the input-response curve when more than one "signal" is present in an environment (Figure 6.1). Following the infomax principle, the input-output transformation (the neuronal response curve) should maximize the information transmission between the neural representation and the stimulus. The optimal response curve depends on the statistical properties of the input signals, but internal noise and constraints on the possible changes limit the amount of information that can be conveyed.

The infomax principle leads to the theoretical result that the derivative of the response

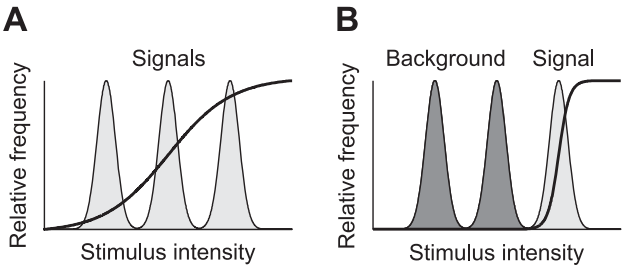


Figure 6.1 | Optimal sigmoidal stimulus response curves (solid lines) for a stimulus distribution consisting of three peaks (shaded areas) as predicted by two coding hypotheses. **A** | Infomax: the dynamic range of the adapted response curve covers the whole range of input signals. Note that the optimal *sigmoidal* response curve is shown; generic optimal transmission would be attained by a response curve that has a derivative proportional to the local stimulus distribution. Such a response curve would be steep within peaks of the stimulus distribution and much flatter in between, thus it would be more staircase-like. **B** | Selective coding: the response function optimally represents the most intense signal (light gray) whereas other signals (dark gray) are suppressed.

curve should be proportional to the probability distribution of the stimuli, so that all available signals in a given environment are represented and every possible output rate occurs with equal probability. Laughlin (1981) tested this prediction and showed that contrast response curves in the fly visual system are matched to the statistics of natural images in order to maximize information transmission. Similar results have been reported for contrast response curves of retinal and LGN neurons in cat and monkey (Tadmor and Tolhurst, 2000). Information maximization can explain retinal coding in the spatial, temporal and chromatic domain (Atick, 1992).

A consequence of the infomax principle is that a change in the statistics of the sensory input must be compensated by a change in the input-response curve. Experimental evidence from the motion-sensitive H1 neuron in the fly supports this hypothesis: this neuron adapts its response curve to changing statistics of stimuli on several time scales (Brenner et al., 2000; Fairhall et al., 2001), in a way which is compatible with the infomax prediction. Experiments have also shown that adaptation enhances information transmission in visual cortex (Sharpee et al., 2006). Sharpee et al. estimated neural filters for the responses to natural inputs and to noise inputs matched for luminance and contrast, showing that neural filters adaptively changed with higher order statistics of input signals, so as to increase the mutual information between stimulus and neural response. Theoretical work also suggested that contrast adaptation in the mammalian visual system (Baccus and Meister, 2002; Solomon et al., 2004; Carandini, 2000; Kohn and Movshon, 2003; Ohzawa et al., 1985) can be understood as a consequence of the infomax principle (Adorján et al., 1999b; Schwabe and Obermayer, 2003). However, it

is difficult to quantify the role of adaptation in enhancing coding efficiency at higher stages in vertebrate sensory pathways since in these, coding is distributed among large populations of neurons and their responses are modulated by the activity of other neural populations or brain areas. Thus, simple sensory networks of invertebrates, whose representation are not heavily influenced by feedback signals, may provide a more suitable model to understand the computational principles underlying sensory adaptation.

An alternative principle that may underlie adaptation is selective coding (or "background suppression"), a form of temporal inhibition in which a loud sound suppresses the response to subsequent sounds. This could serve to segregate a single, most important signal from other signals or background noise. It has been shown that an auditory interneuron (ON1) represents mainly the louder part of a stimulus with a bimodal intensity distribution (Pollack, 1988). Calcium aggregation in the omega neuron is a possible mechanism underlying this background suppression (Sobel and Tank, 1994; Baden and Hedwig, 2007). Similar findings have been made in bushcrickets: while multiple songs in choruses of singing males are present, only the most intense song was found to be represented in the auditory pathway (Römer and Krusch, 2000). These previous studies, however, address the phenomenon only qualitatively and not under the viewpoint of an encoding scheme and information transfer. Segregation of different auditory objects into different channels has also been studied in vertebrate hearing (Las et al., 2005; Nelken, 2004; Ulanovsky et al., 2003). In vertebrates, however, modulation of carrier frequency is assumed to play a crucial role in this stream segregation (Joris et al., 2004), complicating a detailed analysis. Information conveyed by carrier frequency modulation is very limited in crickets, as they possess only two broadly tuned frequency channels, one in the range around the carrier frequency of the calling songs (5 kHz) and another one mainly for frequencies above (12 kHz). Thus, crickets provide an ideal model system to study object-background segregation – in this simple auditory system, for a given frequency range, an auditory object can simply be seen as the loudest peak in the entire stimulus distribution.

The questions we sought to answer were: How does the neural response curve adapt to the statistics of the acoustic environment? Can this sensory system be characterized as a communication channel optimized for coding the inputs such that as much information as possible is preserved (infomax principle)? Or does the system perform a preprocessing that leads to a high fidelity representation of only the loudest part of the stimuli (selective coding)? To address these questions, we measured the neural response curve of AN2 neurons after adaptation to sound stimuli with either two or three peaks in their intensity distribution, depicted in Figure 6.1. The two principles studied here predict conflicting changes of the form of the input-response curve when presenting a stimulus composed

of more than one signal. Optimal selective coding should lead to a shift of the response curve in a way that only the peak with the highest intensity is represented (Figure 6.1B). If infomax is the underlying principle, adaptation pursues the objective to maximize the information that the neuron's output conveys about its sensory input. Adaptation should thus change the response curve in a way that the whole stimulus range is encoded reliably (Figure 6.1A).

Firing rate adaptation can change the stimulus-response curve basically in two ways (Benda and Herz, 2003): shifting the threshold to larger intensities and changing the slope of the curve. We first compare the experimentally observed changes in the slope and the shift in response curves to the optimal changes predicted by the two competing hypotheses. Differences between model prediction and data, however, do not necessarily imply that a particular hypothesis is unlikely to be true, because additional constraints may limit the potential of tuning curve changes. Therefore, in a second step, we calculate the mutual information between the sensory input and the neuronal response using the measured response curves. The infomax principle predicts that the mutual information between a particular stimulus distribution and the response should be highest for the response curve that is adapted to the stimulus distribution. The response curve adapted to the stimulus with three peaks should encode the three-peak stimulus better than the response curve adapted to the stimulus with two peaks. Selective coding, on the other hand, predicts, that the mutual information should decrease for the "background" signals and should increase for the most intense peak.

6.2. Experimental methods

All experiments were performed by K. Jannis Hildebrandt and R. Matthias Hennig. Here we briefly summarize the experimental methods that are necessary as a background for this chapter. For a more detailed description we refer to Hennig (1988).

Animal preparation

Crickets of the species *T. oceanicus* and *T. leo* were used in the experiments to characterize the time course of adaptation. For the experiments with the multimodal stimuli, mainly *T. leo* individuals were used. All animals were laboratory reared. For preparation, both pairs of wings and the meso- and metathoracic legs were removed. The animal was fixed ventral side up to a small platform and the prothoracic legs with the ears were waxed to pins at the coxae and the tarsi in a normal walking position. Ascending and descending connectives from the prothoracic ganglion were cut in order to reduce neuronal background activity.

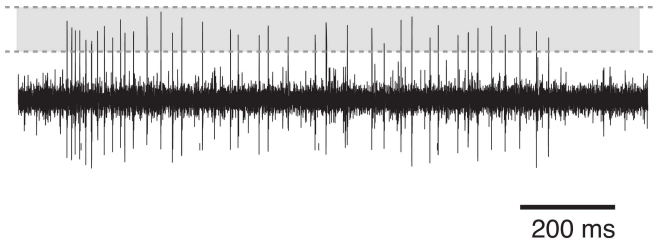


Figure 6.2 | Typical recording trace from a cricket AN2 neuron (*T. oceanicus*). The figure shows the voltage trace during constant stimulation (duration 1 s) with a sinusoidal tone of 16 kHz frequency. The shaded area depicts the spike detection window, bounded by the lower and upper threshold.

Recordings and acoustical stimulation

Two extra-cellular hook-electrodes were made from tungsten wire and placed in parallel around one of the two connectives ascending from the prothoracic ganglion. These connectives contain the axon of the ascending interneuron we wanted to record from (AN2). Vaseline was placed around connectives and hooks in order to isolate the electrodes electrically and keep the connective from drying out. Spikes of the AN2 were detected on the basis of the amplitude peaks of the voltage trace using a lower and upper threshold. Figure 6.2 shows an example recording and the spike detection window.

The recording set-up was lined with sound-absorbing foam to reduce echoes. Acoustic stimuli were presented through a loudspeaker positioned ipsilaterally to the recorded connective at a distance of 36 cm. The main input to the AN2 neuron comes from receptors ipsilateral to the connective that holds its axon.

Stimulus protocols

Figure 6.3 shows the protocol used for characterizing the adaptation process in the ascending AN2 neuron. The different ensembles of auditory stimuli consist of an adapting stimulus, a silent interval, and a test stimulus. The intensity of the adapting stimulus was adjusted in the beginning of the recording depending on the response strength of the neuron. Normally, the base line intensity of the adapting stimulus had a sound pressure level of 84 dB or 87 dB. With the term *relative intensity* we refer to the stimulus intensity relative to this base line intensity. Adapting stimuli are 16 kHz signals that were amplitude-modulated by bandpass-filtered Gaussian white noise with 100 Hz cut-off frequency. The Gaussian noise had a variance $\sigma^2 = 1.38\ \mathrm{dB}^2$ and a mean relative intensity $\mu = 0$ dB. Test stimuli were pure sinusoidal tones with a frequency of 16 kHz. To characterize adaptation, we used adapting stimuli with durations 75 ms, 150 ms, 300 ms, 600 ms, 1200 ms, 2400 ms and 4800 ms (Figure 6.3A). For testing recovery

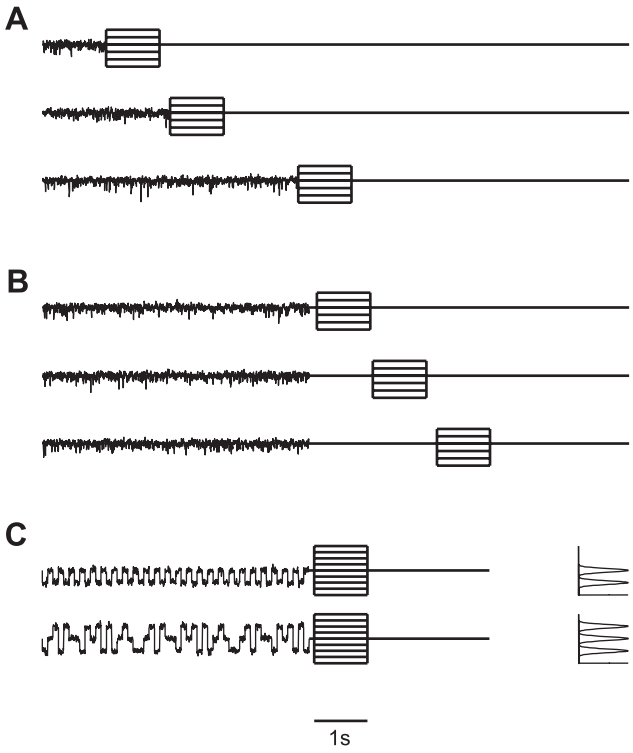


Figure 6.3 | Summary of the experimental protocols. **A** | Adaptation protocol. Amplitude-modulated noise signals (adapting stimuli with 0 dB average relative intensity) of variable duration (from 75 ms to 4800 ms) are followed by a test stimulus (16 kHz sinusoidal tone) with a duration of 1000 ms and a relative intensity ranging from −9 dB to +6 dB (several test stimuli are plotted on top of each other). **B** | Recovery protocol. Amplitude-modulated noise signals (adapting stimuli) of 5 second duration are followed by a pause of variable length (from 75 ms to 4800 ms) and a test stimulus as in (A). **C** | Adaptation protocol for amplitude-modulated noise stimuli drawn from a bimodal and a trimodal distribution (the corresponding amplitude distributions are shown in the right panel). Relative intensities of the test stimuli range from −6 dB to +6 dB.

from adaptation, the stimuli had a 5 s adaptation phase followed by pauses of varying durations from 75 ms to 4800 ms (Figure 6.3B).

Motivated by the competing coding hypotheses, we wanted to examine the consequences of adaptation to different ensembles of auditory stimuli demanding different changes in the stimulus-response curve. We designed multimodal noise-like stimuli whose amplitude distribution had two or three modes, mimicking auditory scenes with multiple signals. The amplitude distribution of the bimodal stimulus is composed of two Gaussian distributions, with mean relative intensity $\mu_1 = -3\,\mathrm{dB}$, $\mu_2 = 0\,\mathrm{dB}$ and variance $\sigma^2 = 0.2\,\mathrm{dB}^2$. The trimodal stimulus has an additional peak modeled by a third Gaussian distribution with mean $\mu_1 = +3\,\mathrm{dB}$ and $\sigma^2 = 0.2\,\mathrm{dB}^2$. An example of these stimuli is shown in Figure 6.3C, together with the respective amplitude distributions. The adaptation time was 5 s in these experiments and the silent interval before the test stimulus was 100 ms.

In all experiments, the intensity response curves were determined by sinusoidal test stimuli with a frequency of 16 kHz and duration of 1 s, following the respective adaptation stimulus. The relative intensities of the test stimuli were −9 dB to +6 dB. Each stimulus was presented at least five times.

6.3. Bayesian data analysis

We constructed intensity-response curves to quantify the neural response, as shown in Figure 6.4 and Figure 6.5. Therefore, we used the spike count in a 200 ms time window beginning 100 ms after test stimulus onset. The window was chosen such that the influence of the fast adaptation process (time constant of about 40 ms, similar to the one described for the AN1 neuron by Benda and Hennig, 2008) is minimized. In the context of this separation of time scales, we are interested only in the coding of slower stimulus dynamics. Hence we consider responses to unmodulated test stimuli and measured spike counts within a 200 ms – rather than a short – time window.

A common methodology to construct neuronal response curves is repeating a single experimental condition several times and then computing the mean of the observed spike counts and their variance. In a second step, a parametric model is fit to these data, typically using least-square approximation. Often it is interesting, however, how the parameters of the response curve change with different experimental conditions, but the confidence intervals for the model parameters and tests for the significance of parameter changes are difficult to establish with traditional statistical methods.

Here, we use a Bayesian analysis (Gelman et al., 1995; MacKay, 2003), to account for the statistics of each trial to estimate the parameters accurately and to quantify the confidence limits of the parameter estimation. The method allows estimating the full

probability distribution of the response curve parameters rather than only the mean value as with traditional methods. Similar techniques have been applied successfully to the analysis of intracellular membrane potential recordings (Gillespie et al., 2001).

Modeling the sound-pressure-level to spike-count relation

The analysis is based on the assumption that spikes are Poisson-distributed and that individual trials are independent of each other (i. e., their joint probability is equal to the product of their individual probabilities).

Let x_i denote the ith out of m stimulus intensities and n_i the number of times a stimulus with this intensity is presented. The corresponding number of spikes from an AN2 neuron is denoted by $y_{i,j}$, where j is the jth out of the n_i repetitions. If spikes are Poisson distributed, we obtain

$$P\left(y_{i,j}|r_i\right) = \frac{r_i^{y_{i,j}} e^{-r_i}}{y_{i,j}!}, \tag{6.1}$$

where r_i is the average spike count underlying the neuron's response at the ith stimulus intensity. For a set $y_i = (y_{i,1}, \ldots y_{i,n_i})$ of spike counts of n_i independent and identically distributed observations, the likelihood $P(y_i|r_i)$ of r_i being the underlying average spike count becomes

$$P\left(y_i|r_i\right) = \prod_{j=1}^{n_i} \frac{1}{y_{i,j}!} r_i^{y_{i,j}} e^{-r_i} \tag{6.2}$$

$$\propto r_i^{t(y_i)} e^{-n_i r_i}, \tag{6.3}$$

where the likelihood function is determined, up to a constant factor, by the sufficient statistic

$$t\left(y_i\right) = \sum_{j=1}^{n_i} y_{i,j}. \tag{6.4}$$

We assume a sigmoid response curve, relating stimulus intensity to spike counts as

$$r_i = f\left(x_i\right) = \frac{A}{1 + \exp\left(-\frac{x_i - B_{50}}{C}\right)}, \tag{6.5}$$

where r_i is the underlying average spike count of the neuron, x_i is the stimulus intensity, A is maximum response of the cell, B_{50} is the stimulus intensity at 50 % of maximum

response, and C is a slope factor. Inserting this relationship into Equation 6.3, we obtain the likelihood $P(y_i|r_i)$ in terms of the response curve parameters A, B_{50} and C:

$$P\left(y_i|A, B_{50}, C, x_i\right) \propto \left[\frac{A}{1+\exp\left(-\left(x_i - B_{50}\right)/C\right)}\right]^{t(y_i)} \exp\left(-n_i\left[\frac{A}{1+\exp\left(-\left(x_i - B_{50}\right)/C\right)}\right]\right). \quad (6.6)$$

Let $y = (y_1, \ldots, y_m)$ be the set of responses to stimuli with different intensities x_i, where $i = 1 \ldots m$. Applying Bayes' rule we obtain the joint posterior distribution

$$P\left(A, B_{50}, C|y, x\right) \propto P\left(A, B_{50}, C\right) P\left(y|A, B_{50}, C, x\right) \quad (6.7)$$

$$\propto P\left(A, B_{50}, C\right) \prod_{i=1}^{k} P\left(y_i|A, B_{50}, C, x_i\right) \quad (6.8)$$

of the parameters A, B_{50} and C, given the observations, where $P(A, B_{50}, C)$ is the prior distribution of the response curve parameters A, B_{50}, C. In the following, we will use a noninformative, uniform prior distribution $P(A, B_{50}, C) = \text{constant}$.

Calculating the joint posterior distribution

Following Gelman et al. (1995), the posterior was calculated for a range of A, B_{50} and C values using a grid of $200 \times 200 \times 200$ points and normalized across this grid. Initially a large parameter space was sampled (e. g., values for the parameter A in the range from o to three times the maximum observed spike count of the neuron) that was narrowed to allow finer sampling in the region of non-zero posterior values. To simplify further analysis we then draw 10 000 independent and identically distributed random samples $(A_i, B_{50,i}, C_i)$, where $i = 1 \ldots 10000$ from the joint posterior probability distribution. From these samples, we can estimate the posterior distribution of any quantity of interest, e. g., the posterior distribution of the response curve parameters "location" B_{50} or of the "slope" at half of the maximum response:

$$S_{50} = \frac{1}{4C}. \quad (6.9)$$

The slope S_{50} does not depend on the maximum response A, in order to be able to compare the response curve slopes from different neurons (i. e., for calculating the slope the neural responses are normalized to the interval between 0 and 1).

To summarize the results for all the recorded AN2 cells, we combined the samples from the posterior distributions of individual cells to obtain a "combined posterior distribution" (assuming independence of individual experiments).

When reporting experimental results, we will in most cases characterize the corresponding posterior distributions by their mean values (i. e., the expected values of the parameters, given the data). We also use these expected parameter values to illustrate the estimated sigmoid response curve. In most cases, the expected parameters and the parameters with the maximum posterior probability (i. e., the maximum a posteriori estimate) had very similar values.

Significance testing

To test for statistical significance we use the samples from the posterior distribution. Consider one of the parameters of interest, e. g., B_{50}, and its posterior distributions $P_1\left(B_{50}^1|y^1\right)$ and $P_2\left(B_{50}^2|y^2\right)$, for two stimulus conditions 1 and 2. Bayesian analysis provides us with samples from these distributions P_1 and P_2. To determine if B_{50}^1 is significantly different from B_{50}^2, we calculate the posterior distribution P_d of the difference $B_{50}^2 - B_{50}^1$. This is done by repeatedly taking one sample b_1 from the distribution $P_1\left(B_{50}^1|y^1\right)$ and one sample b_2 from $P_2\left(B_{50}^2|y^2\right)$ and calculating the difference $b_2 - b_1$, giving one sample from the distribution P_d.

To determine a significant difference, we calculate the 95 % posterior interval $[i_1, i_2]$ of P_d, defined as the range of values above and below which lie 2.5 % of the samples. The values i_1 and i_2 can be directly estimated from the samples: i_1 corresponds to the 2.5th and i_2 to the 97.5th percentile. If the 95 % posterior interval of P_d includes zero, the difference between B_{50}^2 and B_{50}^1 is not statistically significant. On the other hand, if the 95 % posterior interval excludes zero we regard the difference as significant. To test if an estimated parameter is significantly larger (smaller) than a certain value x, we calculate the right-tailed (left-tailed) posterior interval. If the right-tailed (left-tailed) posterior interval excludes the value x, i. e., less than 5 % of the corresponding samples are smaller (larger) than x, we regard the parameter significantly larger (smaller) than the value x.

Time course of adaptation

We used a single exponential decay model to characterize the time course of adaptation

$$y(t) = y_{\min} + (y_{\max} - y_{\min}) \exp\left(\frac{-t}{\tau_a}\right), \tag{6.10}$$

where $y(t)$ is the neural response at time t, $y_{\min}$ and $y_{\max}$ are minimum and maximum response, and τ_a is the decay time constant. A similar single exponential model

$$y(t) = y_{\min} + (y_{\max} - y_{\min}) \left[1 - \exp\left(\frac{-t}{\tau_r}\right)\right] \tag{6.11}$$

is used for describing the recovery from adaptation, where τ_r is the recovery time constant. Using the Bayesian approach, we calculate the posterior densities of the parameters of Equation 6.10 and Equation 6.11 in a similar manner as for the sigmoid response curve (Equation 6.5).

Numerical estimation of mutual information

The mutual information $I[Y;X]$ between the sensory signal X and the neural response Y specifies how much information is conveyed on average about all possible signals. In order to compute the mutual information numerically, taking into account the influence of discrete, Poisson distributed spike counts, we first construct the joint probability distribution

$$P(y,x) = P(y|x)P(x). \tag{6.12}$$

For each stimulus intensity x_i, we calculated the corresponding average spike count r_i using Equation 6.5. The distribution $P(y|x = x_i)$ is then given by a Poisson distribution with mean r_i (Equation 6.1). For all simulations, the stimulus X was discretized into bins of size 0.01 dB. At this resolution, the results did not depend on the bin size.

To measure the information that is associated with specific sensory signals, we define the stimulus-specific information (Butts, 2003; Butts and Goldman, 2006):

$$i_{\mathrm{SSI}}(x) = \sum_y P(y|x)\left(H[X] - H[X|Y = y]\right), \tag{6.13}$$

where $H[X] = -\sum_x P(x)\log_2 P(x)$ is the entropy of the sensory signal X, and the conditional entropy of a particular response y is $H[X|Y = y] = -\sum_x P(x|y)\log_2 P(x|y)$. Stimulus-specific information can be interpreted as the average reduction of uncertainty about the sensory signal gained from one measurement given the stimulus x. Taking the weighted average over the stimulus-specific information for all possible signals we obtain the mutual information between stimulus and response:

$$I[Y;X] = \sum_x P(x) i_{\mathrm{SSI}}(x). \tag{6.14}$$

To determine the information associated with a certain stimulus range $[x_1, x_2]$, we evaluate the sum in Equation 6.14 from x_1 to x_2.

6.4. Time course of adaptation

We first studied the effects of prolonged auditory stimulation in recordings of 6 AN2 neurons of *T. oceanicus* and 7 AN2 neurons of *T. leo*. Previously, an adaptation process

operating on a time-scale of 40 ms had been characterized for the AN1 neuron (Benda and Hennig, 2008). Here, we investigate whether adaptation also occurs on a slower time-scale, better matched to changes in the acoustic environments.

Adaptation and recovery

We recorded the responses of AN2 neurons to test stimuli of different intensities, after adaptation to noise stimuli of varying duration (see Section 6.2, Stimulus protocols). A typical example for the neural responses of an AN2 cell of *T. leo* is shown in Figure 6.4. The spike rates after an adaptation period of 4800 ms are always lower than the corresponding responses after 600 ms and 75 ms adaptation time. Responses declined with prolonged stimulation during the test interval for the applied intensities that were higher than the intensities of the adapting stimuli (Figure 6.4C–D), a phenomenon which we observed in all the recorded cells. The rapid initial change, which is most pronounced for high intensities of the test stimulus (Figure 6.4C–D) is caused by the fast firing-rate adaptation (similar to the adaptation in the AN1 neuron; Benda and Hennig, 2008). To minimize an influence of the fast and the slow adaptation occurring during test, only spikes occurring between 100 ms and 300 ms after test stimulus onset were used for further analysis (see Section 6.3, Bayesian data analysis). Figure 6.5A shows the stimulus response curves constructed from the spike counts within the abovementioned interval. Prolonged stimulation shifted the stimulus-response curves towards higher stimulus intensities. In the example shown, adapting for 4800 ms virtually eliminated the response to low relative intensities from −9 dB to −3 dB. Adaptation changes the range of relative intensities over which the cell responds, but has little effect on the maximal firing rate. Figure 6.5B shows data from the same cell when using stimuli for testing the recovery from adaptation (see Section 6.2, Stimulus protocols). Adapting stimuli were always 5 s long, followed by a silent interval of varying duration and a test stimulus. After a recovery period of 4800 ms the neuron has almost recovered its state prior to adaptation. Hence, adaptation and recovery from adaptation operate on a similar time scale.

Time constants of adaptation

To quantify the time course of adaptation and recovery we analyzed the neural responses to test stimuli that had the same relative intensity as the adapting stimuli (0 dB). Additional cells were recorded with a reduced version of the stimulus protocol that only included these 0 dB test stimuli (the total number of cells available for each species and each stimulus protocol is stated in Table 6.1). In order to determine the adaptation and

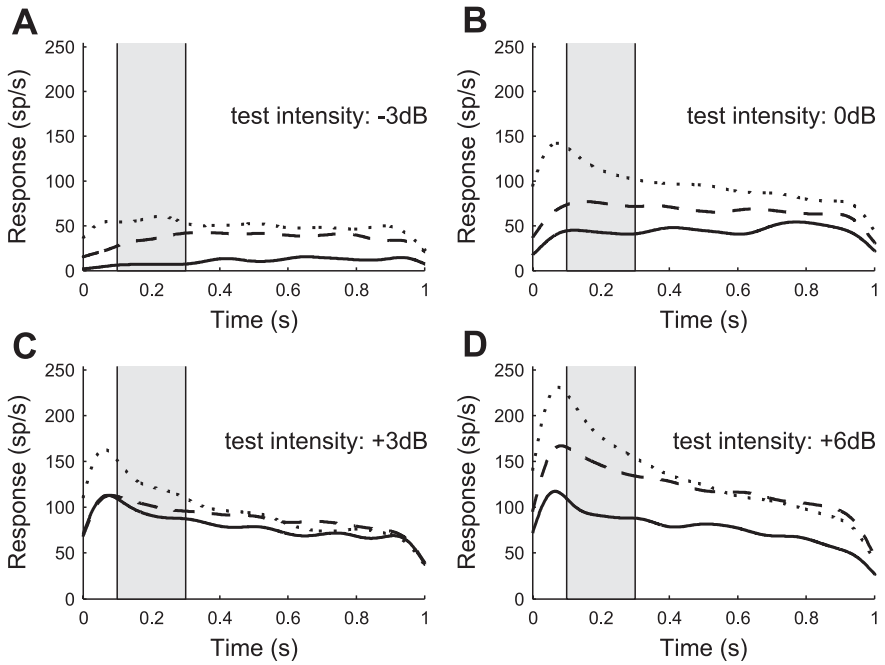


Figure 6.4 | Representative examples of the neural response (AN2 neuron from a *T. leo*) after adaptation to noise stimuli of duration 75 ms (dotted line), 600 ms (dashed line) and 4800 ms (solid line). Responses (spike rates) during a test stimulus of 1 s duration (cf. protocol of Figure 6.3A). Relative intensities of the test stimuli range from −3 dB (A) to +6 dB (D); the average relative intensity of the adapting stimulus was 0 dB. Each stimulus was presented 5 times and the recorded spike trains (1 ms resolution) were convolved with a Gaussian kernel (σ = 50 ms). The instantaneous spike rates were estimated by averaging over the 5 repetitions. The increase of the estimated rate during the first 50 ms is an artifact introduced by filtering the neural response with the Gaussian kernel. Note that the onset latency of the AN2 neuron is in the range of 15 ms to 18 ms. The spike counts during the sample period (shaded) from 100 ms to 300 ms are used to construct neural response curves.

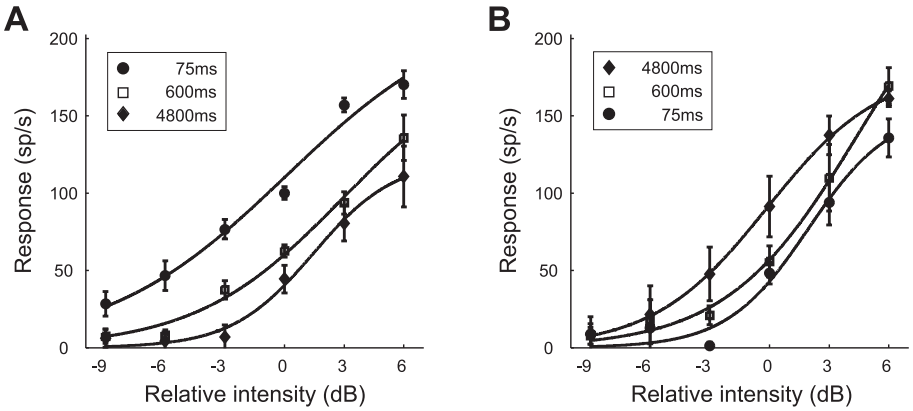


Figure 6.5 | Representative example of response curves for different adaptation (A) and recovery times (B) (cf. protocols of Figure 6.3A, B). The average relative intensity of the adapting stimulus was 0 dB. Symbols denote the average spike counts during the sample period (cf. Figure 6.4) for different test intensities. Solid lines indicate the expected response curve, i. e., the response curve with the set of parameters with the mean value of the posterior distribution (see Section 6.3, Bayesian data analysis). Each stimulus protocol was repeated 5 times (the error bars indicate the standard deviation). The data shown was obtained from a *T. leo* (the same preparation as used in Figure 6.4).

recovery time constants τ_a and τ_r, we fitted an exponential decay model to the neural responses (see Section 6.3, Bayesian data analysis). Figure 6.6A, B show examples of recorded data and exponential fits for a *T. oceanicus* and a *T. leo* cell. Both time constants lie in the range of 1 second for both of these cells. This is considerably longer than the short-term firing rate adaptation, which operates on a time scale of 40 ms.

Table 6.1. | Summary of the adaptation (τ_a) and recovery (τ_r) time constants for the *T. oceanicus* and *T. leo* AN2 cells. See Figure 6.3 for the adaptation protocols. SD is the standard deviation across the *n* cells.

Species	Adaptation			Recovery from Adaptation		
	τ_a (ms)	SD (ms)	n	τ_r (ms)	SD (ms)	n
T. oceanicus	1202	558	6	1947	1155	6
T. leo	1828	939	9	1674	582	11

The values of the adaptation and recovery time constants are summarized in Table 6.1 for both species; additionally, Figure 6.7 shows the combined posterior distributions (cf. Section 6.3, Bayesian data analysis). Comparing the time constants between *T. oceanicus* and a *T. leo* cells, we did not find significant differences, as reflected by the overlapping 95 % posterior intervals in Figure 6.7. Furthermore, adaptation and recovery time constants have similar values.

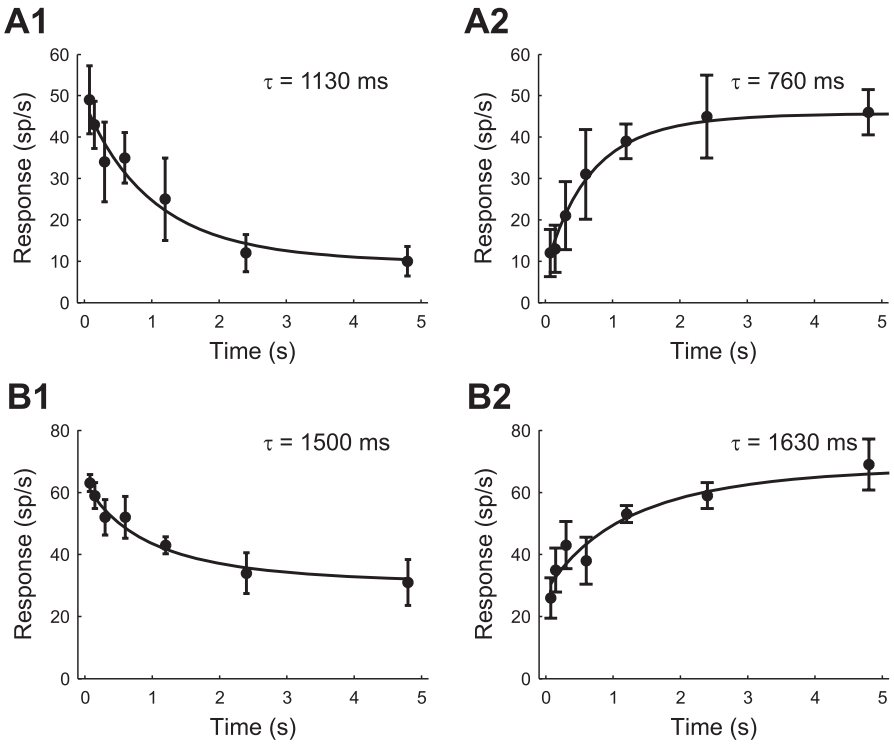


Figure 6.6 | Time course of adaptation and recovery of a *T. leo* cell (A1, A2) and of a *T. oceanicus* cell (B1, B2). The response to the test stimulus is plotted against the duration of the adapting stimulus (A1, B1) and the delay between the adapting and the test stimulus (A2, B2). Displayed are the average spike counts in the 200 ms time window of the test stimulus (cf. Figure 6.4). The intensity of the test stimulus was equal to the average intensity of the adapting stimulus (0 dB relative intensity). The error bars denote the standard deviation. Solid lines indicate the exponential function with the set of parameters with the highest value of the posterior distribution (see Section 6.3, Bayesian data analysis).

We conclude that the neuronal responsivity of AN2 neurons is affected significantly by adaptation and that the adaptation process operates on a time scale of seconds. The primary effect is a change in the range of stimulus intensities over which the cell responds. To put to test our hypothesis that this adaptation serves for adjusting the stimulus-response curve to the current acoustic environment, we first formalize the infomax and selective coding principle and then assess the experimentally observed response curve changes.

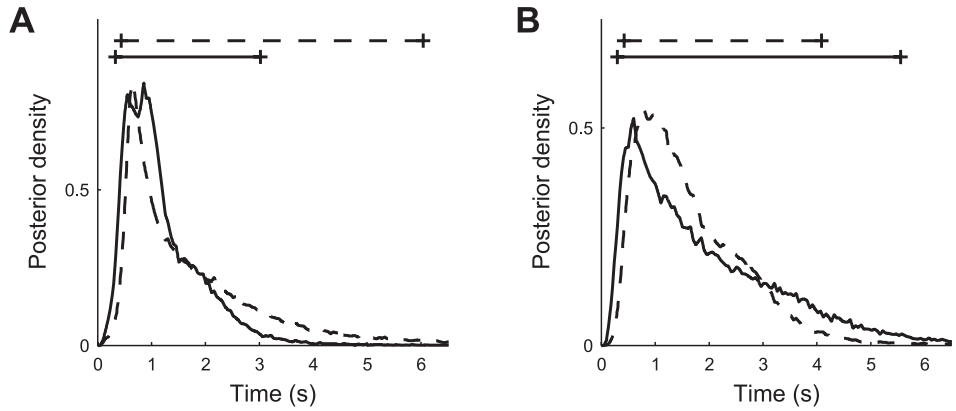


Figure 6.7 | Combined posterior distribution (cf. Section 6.3, Bayesian data analysis) of the adaptation time constants τ_a (A) and the recovery time constants τ_r (B) for the *T. oceanicus* (solid line) and *T. leo* (dotted line) AN2 cells. Solid (dotted) lines on top of the figures depict the 95 % posterior intervals.

6.5. Quantitative predictions of the coding hypotheses

The infomax principle and the selective coding hypothesis both predict how the neural response curve should optimally change in response to a change in the statistics of the environment. In order to assess the response curve changes quantitatively, we first compute the parameters of the optimal response curve under either hypothesis as well as the mutual information between stimulus and neural response.

Infomax principle

If infomax (Linsker, 1988; Atick, 1992; Adorján et al., 1999b) is the underlying principle, adaptation pursues the objective to maximize the information that the neuron's output conveys about its sensory input. Formally, the goal is maximizing the mutual information $I[R;X]$ between the sensory sound signal X and the neuronal output firing rate R as a function of the response curve parameters. This is achieved by maximizing the output entropy $H[R]$ while minimizing the uncertainty $H[R|X]$ of the output once the input is fixed

$$I[R;X] = H[R] - H[R|X]\,. \tag{6.15}$$

We first computed the optimal response curve parameters for a given signal distribution and a sigmoid response function analytically, assuming additive noise. Next, we estimated the mutual information numerically in order to account for multiplicative (Poisson) noise. If we assume only additive noise, with a probability distribution $P(n)$,

the mutual information can be written as (Bell and Sejnowski, 1995; Nadal and Parga, 1994)

$$I[R;X] = H[R] - H[N]. \tag{6.16}$$

Maximization of the mutual information is then equivalent to the maximization of the entropy of the output distribution, because the noise entropy $H[N]$ does not depend on the input-output-mapping, i. e., the neural response curve $r(x)$. Thus, we have to maximize

$$H[R] = -\sum_r P(r)\log_2 P(r), \tag{6.17}$$

where the sum goes over all possible discrete response levels r (here, the spike counts in a 200 ms window; cf. Section 6.3, Bayesian data analysis). Formally, we can treat the response as a continuous variable (Atick, 1992), i. e., as firing rate, and using the relationship between differential and discrete entropy we approximate the sum by an integral (Cover and Thomas, 1991):

$$H[R] = -\sum_r P(r)\log_2 P(r) \rightarrow \\ -\int P(r)\log_2 P(r)\mathrm{d}r - \log_2 \Delta r, \quad \text{as } \Delta r \rightarrow 0 \tag{6.18}$$

Here, Δr is the limit on the resolution with which the firing rate can be measured (the length of the bins of the discrete response levels). Note that in the limit $\Delta r \rightarrow 0$ the entropy $H[R]$ diverges (i. e., the information capacity of a continuous variable is unlimited). In the case $\Delta r \rightarrow 0$, and in the absence of noise, the sensory signal X could be recovered perfectly from the firing rate R and thus any set of response curve parameters would be "optimal". However, if we assume a finite maximum of the response curve the additive noise provides a resolution scale on the output and we can ask for an optimal response curve $f(x)$. In the low-noise limit we obtain (Nadal and Parga, 1994)

$$H[R] = -\int P(x)\log_2 \frac{P(x)}{\left|\frac{\mathrm{d}f(x)}{\mathrm{d}x}\right|}\mathrm{d}x - \log_2 \Delta r \tag{6.19}$$

$$= \int P(x)\log_2 \left|\frac{\mathrm{d}f(x)}{\mathrm{d}x}\right| \mathrm{d}x - \int P(x)\log_2 P(x)\mathrm{d}x - \log_2 \Delta r \tag{6.20}$$

Since the second term of Equation 6.20 only depends on the signal distribution and the third term only depends on the resolution Δr, they are constant, and we have to maximize:

$$\int_{-\infty}^{\infty} P(x)\log_2 \left|\frac{\mathrm{d}f(x)}{\mathrm{d}x}\right| \mathrm{d}x \rightarrow \max. \tag{6.21}$$

To compare how well a given sensory signal X with distribution $P(x)$ is encoded by response functions with different parameterizations (r_{I} and r_{II}), we compute the difference in mutual information ΔI. Under the assumption of small additive noise and for a fixed resolution Δr, this difference in mutual information is given by

$$\Delta I = \int_{-\infty}^{\infty} P(x) \log_2 \left| \frac{\mathrm{d}r_{\mathrm{II}}(x)}{\mathrm{d}x} \right| \mathrm{d}x - \int_{-\infty}^{\infty} P(x) \log_2 \left| \frac{\mathrm{d}r_{\mathrm{I}}(x)}{\mathrm{d}x} \right| \mathrm{d}x. \tag{6.22}$$

Assuming the sigmoid transfer function of Equation 6.5 and the bimodal or trimodal stimulus distribution (see Section 6.2, Stimulus protocols), we obtain the optimal values for the response curve parameters A, B_{50} and C using Equation 6.21; the optimal value for the slope S_{50} is then computed using Equation 6.9. Figure 6.8A shows the predicted response curves for both stimulus distributions, under the assumption that the response curve parameters B_{50} and S_{50} can be optimally adjusted. The optimal value for B_{50} is −1.50 dB for the bimodally distributed stimulus and 0.00 dB for the trimodally distributed stimulus, corresponding to a response curve shift of +1.50 dB. To cover the whole stimulus range, the slope should decrease for the trimodally distributed stimulus compared to the bimodally distributed stimulus by −35.3 %, from 0.25 dB^{-1} to 0.16 dB^{-1}. If we assume that the neural system can only adjust B_{50} and the slope S_{50} is constant, the infomax principle would still predict a shift of the response curve of +1.50 dB. Evaluating Equation 6.22, the infomax principle then predicts that information transmitted about the trimodal stimulus will improve by 0.61 bit for the trimodally adapted response curve compared to the bimodally adapted response curve. Note that this calculation involves the assumption of low additive noise and a fine resolution Δr. Therefore, we also calculated the predicted increase in information transmission numerically (see Section 6.3, Numerical estimation of mutual information), assuming discrete, Poisson distributed spike counts. In this case, the improvement in information transmission depends on the maximum spike count, defined by the response curve parameter A. For the experimentally observed maximum spike counts in the range of 20 spikes to 55 spikes we obtain an increase from 0.12 bit (20 spikes) to 0.25 bit (55 spikes).

Selective coding

Selective coding is a concept which is less well defined than the infomax principle, because it involves an assumption about the "signal" vs. the "background" part of a complex stimulus. In the following we assume, that the loudest signals of artificial environments, i. e., the "loudest" Gaussian distributions of the multimodal stimulus distributions (see Figure 6.8B) are encoded in an optimal way while the other ("background") signals are suppressed. To compute the optimal response curve for the bimodal (trimodal) stimulus

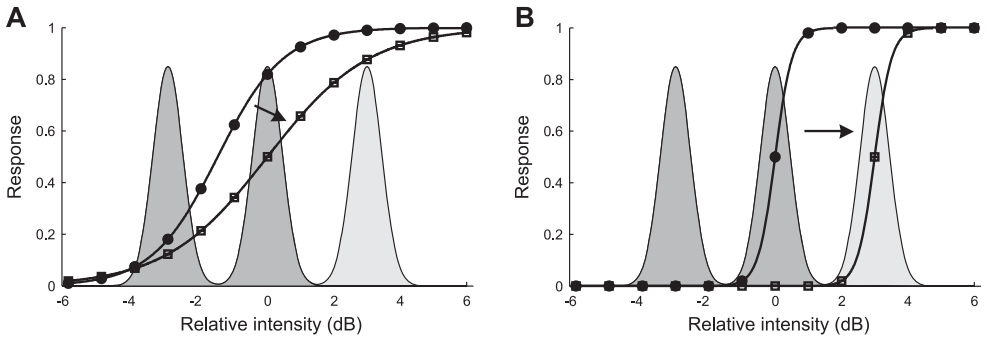


Figure 6.8 | Optimal response curves for the bimodal (circles) and trimodal (squares) stimulus distribution predicted by the infomax principle (A) and the selective coding hypothesis (B). The figures show the predicted relationship between the response variable (spike rate) and the stimulus intensity. The Gaussian curves depict the probability distributions of stimulus intensity, where the dark shaded areas under the curve denote the bimodal stimulus distribution and the light shaded area under the curve the additional peak of the trimodal stimulus distribution (cf. Figure 6.1).

distribution the objective is to maximize Equation 6.21 taking the Gaussian with μ = 0 dB (μ = 3 dB) as the "signal" part. Figure 6.8B shows the predicted response curves when we assume that the loudest signal should be encoded reliably and other signals should be suppressed. The predicted difference between the response curve optimized for the bimodal and trimodal stimulus is a shift by +3.00 dB (from B_{50} = 0.00 dB to B_{50} = 3.00 dB). The slope S_{50} does not change and remains at 0.98 dB^{-1}.

Surely, the response curves shown in Figure 6.8B are idealized but they illustrate the consequences of the selective coding vs. the infomax principles: according to the infomax principle, information transmission is optimized for the whole stimulus range, while selective coding implies a selective enhancement or a selective suppression of the transmitted information for certain kinds of stimuli. To quantify this selective stimulus encoding, we calculated the information associated with parts of the stimulus range numerically (see Section 6.3, Numerical estimation of mutual information), for the trimodal stimulus and the predicted response curves (Figure 6.8B). The maximum responses A were determined by the experimental data (from 20 spikes to 55 spikes), and noise was Poisson distributed. Next, the mutual information is evaluated for the stimulus range of −4.5 dB to 1.5 dB (background signals) and for the range of 1.5 dB to 4.5 dB (loudest signal) using Equation 6.14. We find, that the information transmitted about the loudest peak of the trimodal stimulus distribution is enhanced by 0.51 bit (0.70 bit) for A = 20 (55) spikes when using the optimal response curve for the trimodal stimulus, while at the same time less information (−0.811 bit for A = 20 spikes; −1.07 bit

bit for $A = 55$ spikes) is conveyed about the first and second peak.

Optimality vs. improvement

The infomax and selective coding hypotheses (as specified above) both make quantitative predictions of the optimal response curve parameters and how these parameters should change to optimally adjust the response curve to a changed stimulus statistics. If selective coding is the underlying principle, the response curve should be steeper than for the infomax principle. When the environment changes from a bimodal to a trimodal input distribution, selective coding predicts a large shift of the response curve towards higher stimulus amplitudes, whereas the slope should remain constant. The infomax principle, on the other hand, predicts a less pronounced shift and a decrease in slope.

However, the underlying neural architecture might impose constraints on possible response curve changes and might thus prevent the AN2 neuron from achieving the theoretically optimal response curve. It is conceivable that, for example, the neural gain cannot increase such that the slope of the stimulus-response curve would be optimal for encoding only the loudest peak of the stimulus distribution as required by "optimal" selective coding. How can we quantify the improvement in neural coding according to the one or the other hypothesis without requiring optimality?

Both the infomax principle and selective coding also predict characteristic changes in mutual information between the stimulus and the response for a change from the bimodal to the trimodal environment. Following the infomax principle, the associated response curve change leads to an increase in mutual information. Selective coding, on the other hand, leads to a selective decrease (increase) of the mutual information for the stimuli with low (high) intensities. Thus, even if architectural constraints might prevent the AN2 neuron from achieving the theoretically optimal response curve, the selective increase (decrease) in mutual information provides a test for the infomax (selective coding) hypothesis.

6.6. Adaptation to the statistics of the acoustic environment

Figure 6.9 shows example traces and the amplitude distribution of the bimodal and trimodal sound stimuli together with the corresponding neural responses of a typical AN2 cell (instantaneous firing rate). Adaptation leads to a decrease of the neural responses to 0 dB peak signals (drawn from the high amplitude and intermediate amplitude peak for the bimodal and trimodal distribution) with time.

We recorded responses from 25 AN2 cells for the two stimulus paradigms, 12 cells from *T. oceanicus* and 13 cells from *T. leo*. Since we found no significant differences in

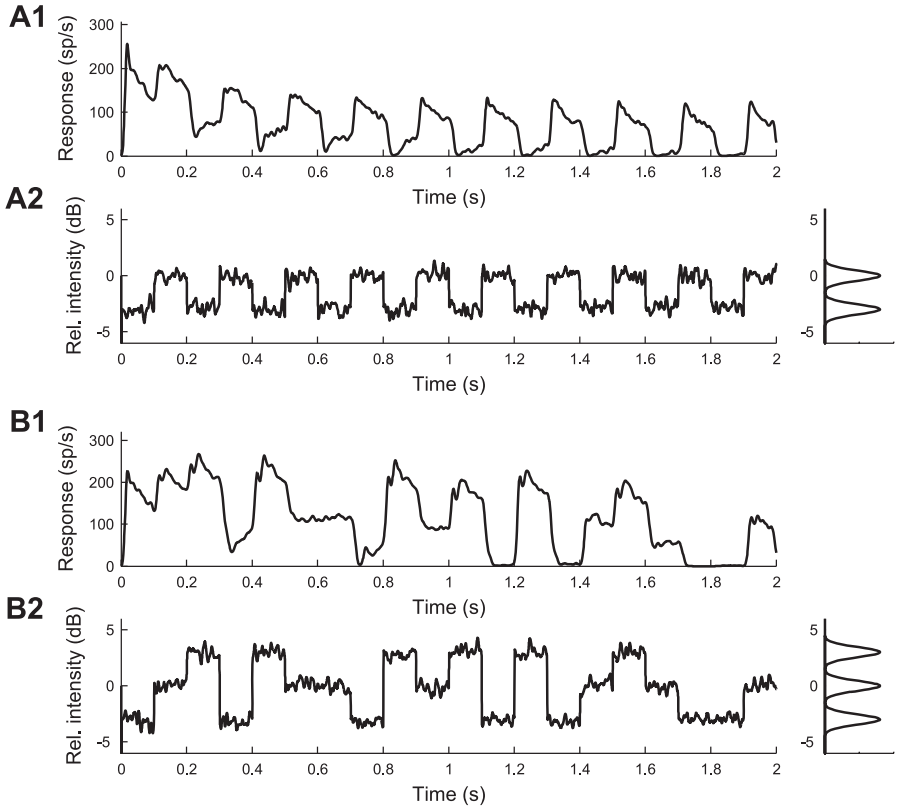


Figure 6.9 | Representative responses of an AN2 cell (*T. leo)* to the amplitude-modulated noise stimuli of Figure 6.3C. **A1, A**2 | Bimodal stimulus distribution. The envelope of an amplitude-modulated stimulus and the distribution of the stimulus amplitude are shown in (A2), the corresponding instantaneous spike rate is shown in (A1). **B1, B**2 | Trimodal stimulus distribution. The envelope of an amplitude-modulated stimulus and the distribution of the stimulus amplitude are shown in (B2), the corresponding spike rate is shown in (B1). The stimuli were presented 45 times and the recorded spike trains (1 ms resolution) were convolved with a Gaussian kernel (σ = 5 ms). The instantaneous spike rates were estimated by averaging over the 45 repetitions.

the adaptation and recovery time constants between the two species, we pooled the data from both species together for further analysis.

All response curves were quantified using sigmoid input-response functions (cf. Equation 6.5), and a Bayesian approach was used to determine the distribution of the corresponding parameters A, B_{50} and C (see Section 6.3, Bayesian data analysis). Some cells did not show response saturation in the trimodal stimulus condition within the range of stimulus intensities we tested. In these cases, the uncertainty of the estimate of parameter A is high, and is reflected by a broad posterior distribution for this parameter. For most of the cells the test stimuli were strong enough to drive the cell to its maximum rate in both conditions. Although the response maximum occurs at higher stimulus intensities in the trimodal condition, we did not observe a systematic change of the saturation response. To fit the response curves, we assumed that the response maximum A has the same value for both stimulus conditions. Five cells were excluded from further analysis because the response curve corresponding to the expected parameter values (posterior means) did not provide a good fit to the data($R^2 < 0.95$). The further analysis is based on the remaining 20 cells.

A representative example of adapted response curves of an AN2 neuron is shown in Figure 6.10A1, where the input-response function is plotted for the parameters A, B_{50} and C, which correspond to the expected parameter values (posterior means). After adaptation to the bimodally distributed stimulus (filled symbols), the cell fired with 50 % of its maximal rate (parameter B_{50}) at about 1.75 dB. Adaptation to the trimodally distributed stimulus (open symbols) shifted the response curve to higher stimulus intensities while the slope of the response curve changed only slightly. In fact, the results of the Bayesian parameter estimation, depicted in Figure 6.10A2, revealed that the response curve parameter B_{50} significantly increased for the trimodal stimulus distribution. The mean of the posterior density changed from 1.74 dB to 3.23 dB (see Section 6.3, Bayesian data analysis for the definition of statistical significance using Bayesian posterior intervals), while there was no significant change for the slope S_{50} (14 % decrease from 0.160 dB^{-1}). Figure 6.10B shows data from a second cell. The mean value of the parameter B_{50} is 3.47 dB for the bimodally adapted response curve, and increased by 1.86 dB through adaptation to the trimodally distributed stimulus. The increase of B_{50} was again significant. The slope increased by 15 % (from 0.161 dB^{-1} for adaptation to the bimodal stimulus) but Bayesian analysis revealed that the increase in slope was not significant. Figure 6.10C shows data from a third cell. This cell showed a significant albeit less pronounced change in parameter B_{50} of +1.06 dB accompanied by a significant decrease in the slope S_{50} (decrease of the posterior mean by 23.5 %).

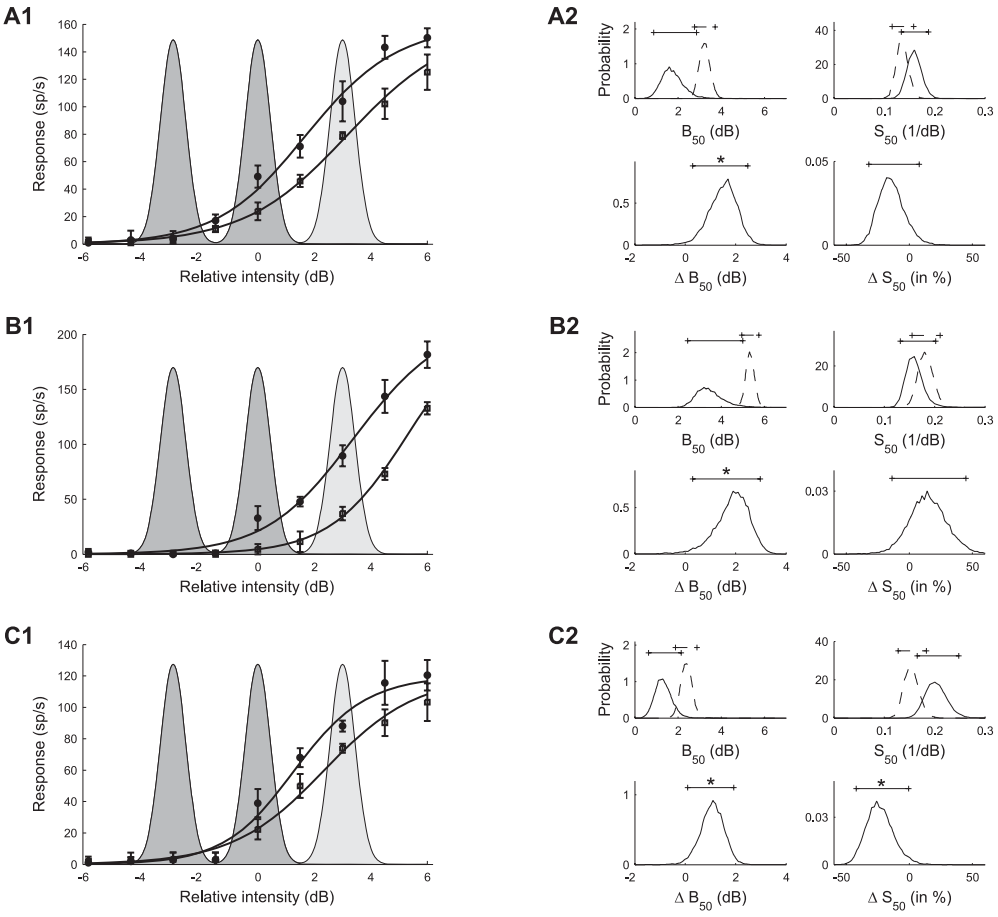


Figure 6.10 | Typical examples of stimulus response curves after adaptation to the bimodal and to the trimodal stimulus distributions (A1, B1, C1) and posterior densities of the corresponding response curve parameters (A2, B2, C2). Results for AN2 cells of *T. leo* (A, C) and *T. oceanicus* (B). **A1, B1, C1** | Circles and squares denote the mean spike counts in a 200 ms time window of the test stimulus after adaptation to the bimodal and trimodal distributions, measured for 9 different relative intensities of the test stimulus (cf. protocol of Figure 6.3C). Error bars denote the standard deviation. Solid lines indicate the expected response curve, i. e., the response curve with the set of parameters with the mean value of the posterior distribution (see Section 6.3, Bayesian data analysis). The shaded areas depict the intensity distribution of the stimuli (dark: bimodal stimulus distribution, light: additional peak of the trimodal stimulus distribution). **A2, B2, C2** | Marginal posterior densities (cf. Section 6.3, Bayesian data analysis) of the response curve parameters B_{50} (location) and S_{50} (slope). The posterior densities after adaptation to the bimodal (solid lines) and trimodal (dotted lines) stimulus distributions are shown in the top panels and the corresponding posterior densities of the changes (ΔB_{50}, ΔS_{50}) between stimulus conditions in the bottom panels. Solid (dotted) lines on top of the figures depict the 95 % posterior intervals. Significant changes between stimulus conditions are indicated by a star.

Adaptation induced changes in the response curve parameters B_{50} and S_{50}

Figure 6.11 summarizes the mean values of the posterior densities of the B_{50} parameters for all 20 AN2 cells. Figure 6.11A1 shows the values of parameter B_{50} after adaptation to the bimodal stimulus. The median value in the population is 2.34 dB (mean: 2.43 dB) and 2.02 dB (mean: 2.07 dB) for cells in which adaptation to the trimodal stimulus led to individual statistically significant changes in parameter B_{50} compared to adaptation to the bimodal stimulus (black distribution). The optimal B_{50} value predicted by the infomax principle is −1.5 dB (star), while selective coding predicts a B_{50} value of 0 dB (circle). The combined posterior distribution of B_{50} (cf. Section 6.3, Bayesian data analysis) is shown in Figure 6.11A2 (mean: 2.43 dB). The measured B_{50} values are significantly larger than the values predicted by either hypotheses (infomax: −1.5 dB, selective coding: 0 dB). Figure 6.11B1 shows the histogram of B_{50} values after adaptation to the trimodal stimulus (median: 3.92 dB, mean: 4.04 dB; individually significant cells: median: 3.57 dB, mean: 3.69 dB). These values are significantly larger than the infomax prediction, but similar to the selective coding prediction (Figure 6.11B2). Figure 6.11C1 quantifies the difference of the parameter B_{50} between the two adaptation conditions. The median of the distribution of differences is 1.53 dB (mean: 1.61 dB). The right-tailed posterior interval in Figure 6.11C2 excludes the value 0 dB, indicating that adaptation to the trimodal stimulus significantly shifts the distribution of response curves towards higher signal intensities. Individual differences are statistically significant in 8 of 20 cells (see Section 6.3, Bayesian data analysis); the median of the changes in these cells is 1.46 dB (mean: 1.62 dB). The observed shifts are smaller than expected for optimal selective coding (predicted shift: 3 dB), but compatible with the infomax principle (predicted shift: 1.5 dB). Due to the high absolute values of the thresholds, however, the response curves do not allow for reliable encoding of the whole stimulus range.

Figure 6.12 summarizes the mean estimates of the slope S_{50}, for all 20 AN2 cells. The slopes in the bimodal adaptation paradigm (shown in Figure 6.12A1) have a median value of 0.16 dB^{-1} (mean: 0.17 dB^{-1}), and are significantly smaller than the value of 0.98 dB^{-1} predicted by the selective coding hypothesis (Figure 6.12A2). The observed slopes S_{50} after adaptation to the trimodal stimulus are shown in Figure 6.12B, and the relative change of the slope compared to the bimodal paradigm is quantified in Figure 6.12C. The slope decreased for most cells (median: −15.1 %, mean: −15.6 %). Significant changes in S_{50} were found individually in 5 of 20 cells, and all of those cells showed decreases in slope. However, the changes are less pronounced than predicted by the infomax principle.

We conclude that the main difference between the response curves adapted to the bimodal vs. the trimodal stimulus distribution is the shift towards higher stimulus inten-

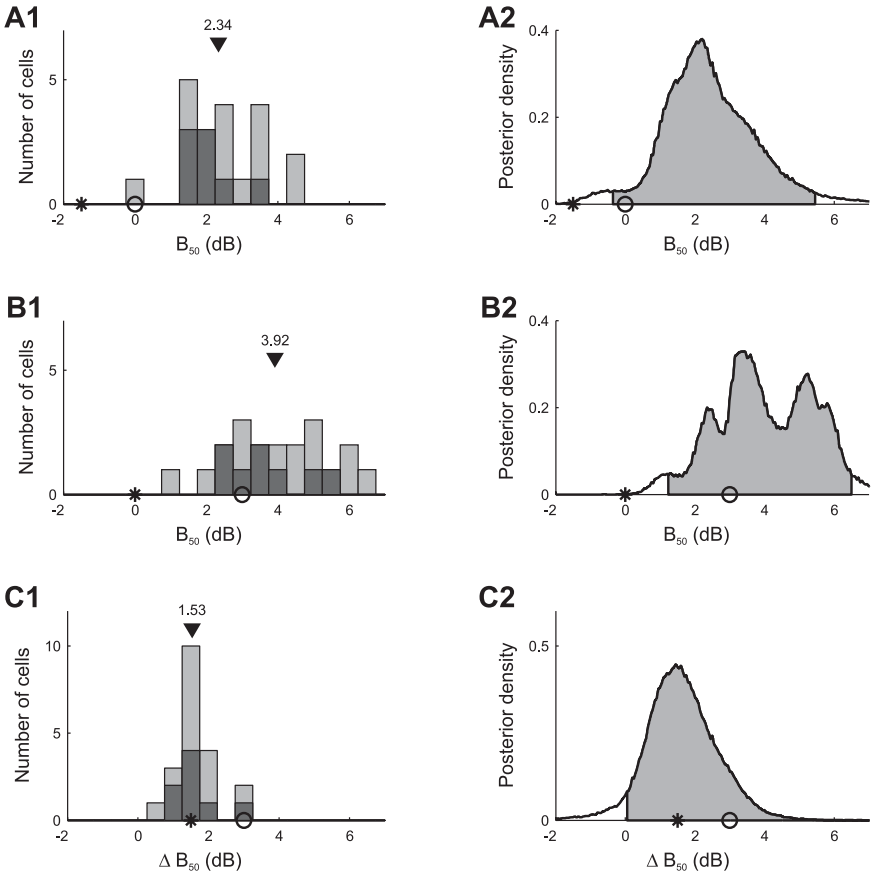


Figure 6.11 | Summary of adaptation induced changes of the response curve parameter B_{50} for all 20 AN2 cells. **A1, A**2 | Distribution of the mean values of the parameters B_{50} for individual cells (A1) and combined posterior density (see Section 6.3, Bayesian data analysis) over all cells (A2) after adaptation to the bimodal stimulus distribution. **B1, B2** | Distribution and combined posterior density of the parameter B_{50} after adaptation to the trimodal stimulus distribution. **C1, C**2 | Distribution and combined posterior density of the change of the parameter B_{50} between the two stimulus distributions. Symbols depict the values predicted by infomax (stars) and the selective coding hypothesis (circles). Triangles denote the median value. The distribution of cells that showed changes in B_{50} that were significant (Bayesian posterior intervals, see Section 6.3, Bayesian data analysis) is marked black in (A1, B1, C1). Shaded areas depict the two-tailed 95 % posterior intervals in (A2, B2) and the right-tailed 95 % posterior interval in (C2).

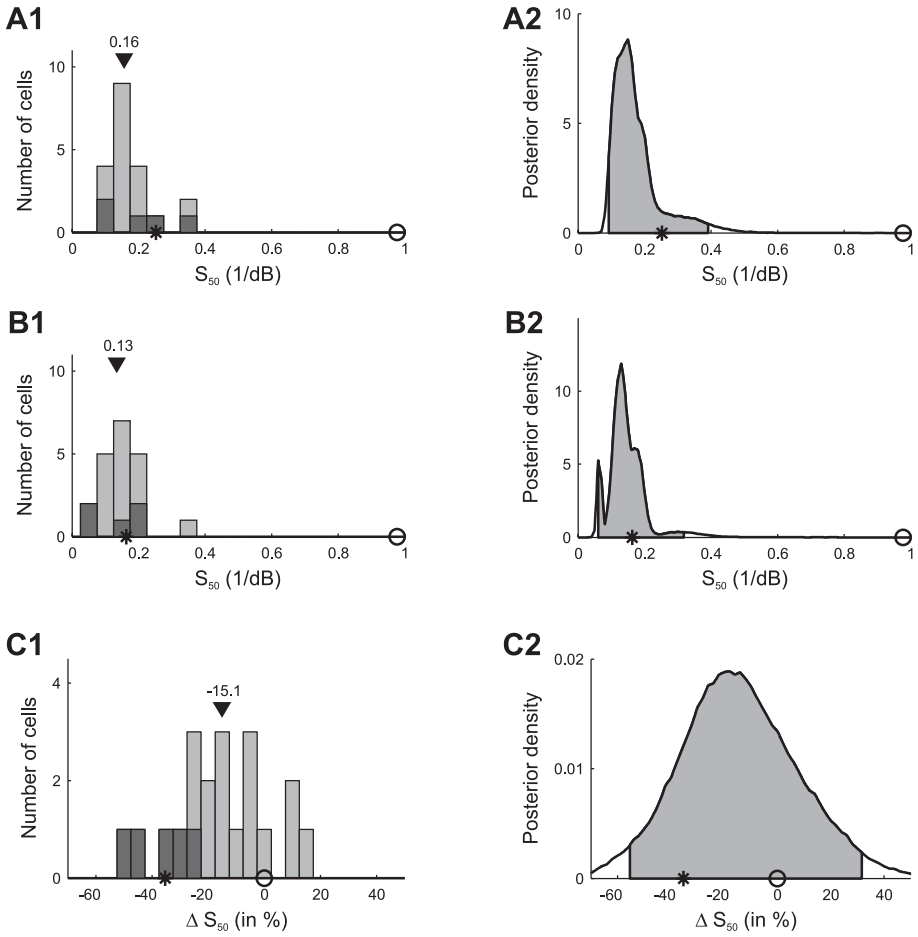


Figure 6.12 | Summary of adaptation induced changes of the slope S_{50} of the response curves. **A1, A**2 | Distribution of the mean values of the parameters S_{50} for individual cells (A1) and combined posterior density (cf. Section 6.3, Bayesian data analysis) over all cells (A2) after adapting to the bimodal stimulus distribution. **B**1**, B**2 | Distribution and combined posterior density of the parameter S_{50} after adapting to the trimodal stimulus distribution. **C**1**, C**2 | Distribution and combined posterior density of the relative change of S_{50} between the two stimulus distributions. Symbols depict the values predicted by infomax (stars) and the selective coding hypothesis (circles). Triangles denote the median value. The distribution of cells that showed changes in S_{50} that were significant (Bayesian posterior intervals, Methods, Bayesian data analysis) is marked black in (A1, B1, C1). Shaded areas in (A2, B2, C2) depict the 95 % posterior intervals.

sities and a reduction in slope. This shift, however, is less pronounced than predicted by optimal selective coding, and the observed decrease in slopes is smaller than predicted by the infomax principle and larger than expected by selective coding. Together with the fact, that the absolute thresholds are too high, these results seem not to favor either of the two coding hypotheses, if optimality is required.

Reliability of stimulus encoding

Adaptation in a biological system, which is constrained in multiple ways, may fall short of achieving the theoretical optimum, but may still lead to an improved representation according to the one or the other principle. In order to test for this, we calculate the mutual information between the stimulus and the neural response for the whole and for the high intensity part of the stimulus range. Therefore, 10 000 samples were drawn from the joint posterior for the parameters A, B_{50}, C, for each cell and for each stimulus condition, and the corresponding response curves were calculated using Equation 6.5. For each response curve, the joint distribution of stimulus and spike count was calculated assuming that spike counts are Poisson distributed with the underlying average spike count given by the response curve (cf. Section 6.3, Numerical estimation of mutual information). Each of these joint distributions determines the mutual information (see Equation 6.13 and Equation 6.14) that corresponds to a particular response curve.

We first consider the whole stimulus range from −4.5 dB to 4.5 dB and calculate the mutual information between the stimulus (trimodal distribution) and the neural response, for the response curves obtained after adaptation to the bimodal and trimodal stimulus distributions. According to the infomax principle the purpose of adaptation is to reliably encode the whole stimulus range and thus, the mutual information between the trimodal stimulus and the neural response should increase for the trimodally compared to the bimodally adapted response curve (predicted increase between 0.12 bit and 0.25 bit, depending on the maximum spike count; see Quantitative predictions).

For the example neurons in Figure 6.10, however, we observed a significant decrease in mutual information, varying from a mean value of −0.183 bit (Figure 6.10A) to −0.372 bit (Figure 6.10B) and −0.187 bit (Figure 6.10C). This trend is confirmed by a full analysis of all 20 recorded AN2 cells (Figure 6.13A), which shows that mutual information decreased for all cells. The median is −0.21 bit (mean: −0.21 bit), and this decrease is significant (the left-tailed 95 % posterior interval in Figure 6.13A2 excludes the value 0 dB). 15 of 20 cells showed an individually statistically significant decrease in mutual information (median −0.24 bit, mean −0.24 bit; black distribution in Figure 6.13A2). These findings provide strong evidence against the infomax principle.

In order to test the selective coding hypothesis, we calculated the mutual informa-

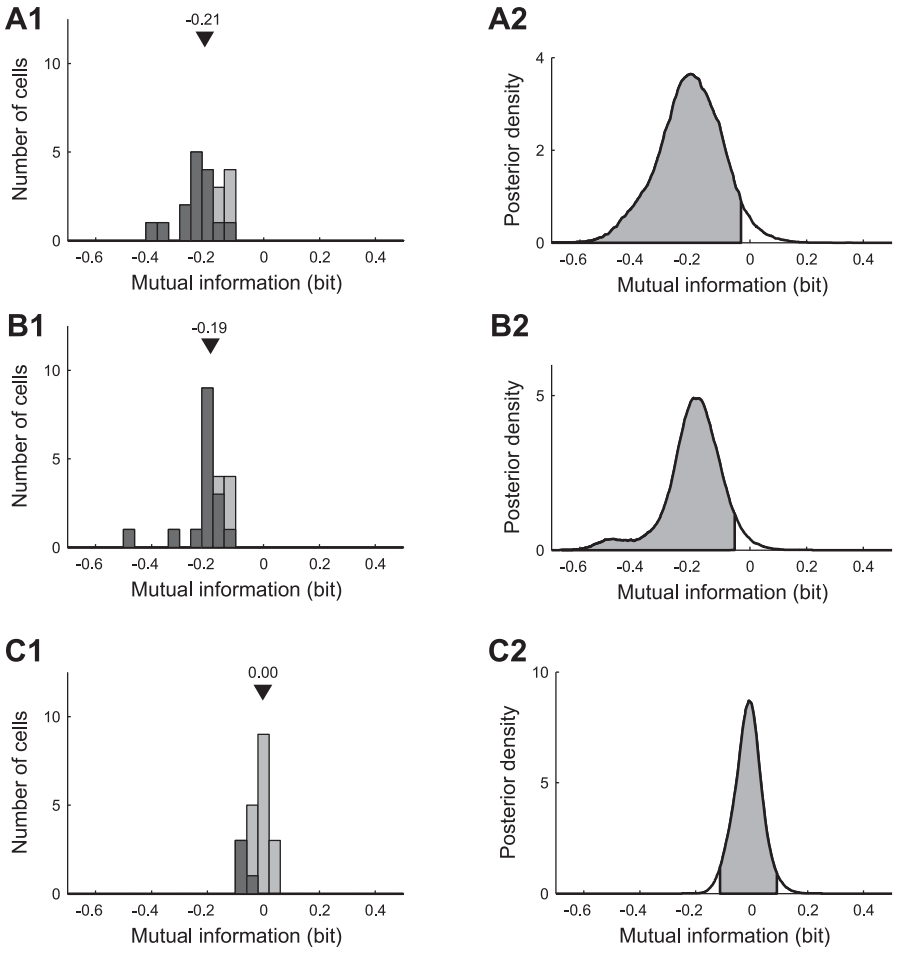


Figure 6.13 | Adaptation induced changes in the mutual information between the stimulus and the neural response. **A1, A**2 | Distribution and combined posterior density of changes in the mutual information when considering the whole stimulus range (relative intensity from −4.5 dB to 4.5 dB) and the trimodal amplitude distribution. For each cell the change of the mutual information is calculated as the difference of the mutual information for the "trimodal" (neural response adapted to the trimodal stimulus) and the "bimodal" (neural response adapted to the bimodal stimulus) response curve. The distribution in (A1) is based on the mean values of changes in mutual information for individual cells. **B1, B**2 | Distribution and combined posterior density of changes in the transmitted mutual information when considering the stimulus range from −4.5 dB to 1.5 dB (including only the two low-intensity peaks of the trimodal stimulus distribution). **C1, C**2 | Distribution and combined posterior density of changes in the transmitted mutual information when considering the stimulus range from 1.5 dB to 4.5 dB (including only the high-intensity peak of the trimodal stimulus distribution). Triangles denote the median value. The distribution of cells that changed significantly is marked black in (A1, B1, C1). Shaded areas depict the left-tailed 95 % posterior intervals in (A2, B2) and the two-tailed 95 % posterior interval in (C2).

tion separately for the stimulus range from 1.5 dB to 4.5 dB (high-intensity peak, "foreground") and from −4.5 dB to 1.5 dB (low-intensity peaks, "background"; see Quantitative predictions) using Equation 6.14. For the cells shown in Figure 6.10, the mutual information decreased significantly by −0.184 bit (Figure 6.10A), −0.335 bit (Figure 6.10B) and −0.182 bit (Figure 6.10C) for the stimulus range from −4.5 dB to 1.5 dB. While the mutual information for the peak of the distribution with the highest intensity increased slightly by +0.018 bit for the cell shown in Figure 6.10A, in other cells, such as the ones shown in Figure 6.10B and 10C, the mutual information decreased not only for the "background" but also for the loudest signal (−0.038 bit vs. −0.005 bit). However, these changes in encoding of the loudest signal were not statistically significant. Figure 6.13B summarizes the change in mutual information for the range from −4.5 dB to 1.5 dB for all 20 AN2 cells. Mutual information decreased significantly (the left-tailed 95 % posterior interval in Figure 6.13B2 excludes the value 0 dB; median −0.19 bit, mean −0.20 bit), and the decrease was individually significant for 16 of the 20 cells. The information transmitted about the "loudest peak" (Figure 6.13C), in the interval from 1.5 dB to 4.5 dB, remained constant (median 0.00 bit, mean −0.01 bit) and is not significantly different from zero (the 95 % posterior interval in Figure 6.13C2 includes the value 0 dB). Although these results are consistent with the selective decrease of information for low-intensity stimuli, they contradict the selective coding hypothesis because information does not increase for high-intensity stimuli, as would be required for an improvement of the neural representation.

6.7. Discussion

Neurons in the auditory pathway of crickets adapt on several time scales

In the cricket auditory system, time scales of adaptation observed at first level interneurons range from short (below 100 ms; AN1: Benda and Hennig, 2008) over intermediate (ca. 300 ms; AN2: Samson and Pollack, 2002, receptors: Givois and Pollack, 2000) to long time constants (ca. 10 s; AN2: Samson and Pollack, 2002, receptors: Givois and Pollack, 2000). In the present study, we report firing rate adaptation with a time constant of about one second not reported before in the auditory ascending neuron AN2 of *T. oceanicus* and *T. leo* (Figure 6.6 and Table 6.1).

At present the origin of adaptation in this small network is not known. There is likely a contribution to adaptation from the receptor neurons (Givois and Pollack, 2000). At the level of interneurons, both local cells (ON1, Sobel and Tank, 1994; Pollack, 1988) as well as the ascending interneuron AN2 (Samson and Pollack, 2002) exhibit long-lasting hyperpolarizations with intermediate time-constants (approx. 5 s) that

may reflect adaptation processes at the level of the spike-generator in these cells. The primary task of the first stage of auditory processing in crickets is to maintain and possibly to condense relevant information for object localization and recognition for higher computational centers in the brain. In this context, it is remarkable to note that already at the first synapse several time scales of adaptation can be observed, similar to those reported from vertebrate systems (Ulanovsky et al., 2003) as well as other sensory modalities (Fairhall et al., 2001). The adaptation time scale we report here seems ideal for an adjustment of the coding scheme to the current sensory environment. However, we find no enhancement of information transfer in the neuron under study. We thus report the unusual case that adaptation seems to rather selectively suppress sensory coding instead of improving it.

Bayesian parameter estimation

We characterized the neural response using a sigmoid response curve and estimated the model parameters from the measured spike counts using a Bayesian approach. An important feature of the analysis is that it yields the joint posterior distribution of the model parameters which allows us to calculate the posterior distribution and precise confidence limits of adaptive changes as well as of derived quantities such as mutual information. The Bayesian framework can also be applied to other experimental paradigms and sensory systems in which neural responses can be described in terms of tuning functions. For instance, it can directly be used to quantify orientation-tuned responses of V1 neurons, by exchanging the sigmoid response curve used here with a circular Gaussian tuning function.

One limitation of the approach presented here is that the calculation of the joint posterior density at a grid of points is only feasible for models with few parameters. However, for more complex models, approximations of the posterior densities can be obtained using Markov chain simulation (Gelman et al., 1995).

Adaptation and the infomax principle

We measured how adaptation changes the response curves of AN2 cells depending on different stimulus conditions (Figure 6.10) and found that the changes are not compatible with infomax (optimal coding of the entire stimulus range). In general, our results indicate a response threshold that is too high to allow for reliable encoding of the whole stimulus range (Figure 6.10 and Figure 6.11). Most remarkably, adaptation reduced the amount of encoded information about the stimulus when considering the whole range of input signals (Figure 6.13A). This is in contrast to other studies (fly visual system: Fairhall

et al., 2001; Laughlin, 1981; midbrain of guinea pigs: Dean et al., 2005; inferior colliculus of cats: Kvale and Schreiner, 2004; songbird auditory forebrain: Nagel and Doupe, 2006; rat barrel cortex: Maravall et al., 2007) that reported that stimulus encoding is compatible with infomax. However, the infomax principle, which considers sensory systems as communication channels that are optimized for preserving all information from the sensory input, may fail to explain neural coding when considering stages where actual processing of information takes place (instead of mere transmission). Indeed, a recent study by Ringach and Malone (2007) has shown that neurons in the primary visual cortex of macaque maintain an operating point that does not maximize information transmission but is tuned to the detection of signals in background noise.

Adaptation and the selective coding hypothesis

We also tested if the response curve changes induced by adaptation are compatible with selective coding (reliable coding of the most intense signal while suppressing the "background"). Selective coding can be seen as the simplest form of separating neural representations of discrete objects in multiple channels or "streams" that has been found in higher auditory processing levels in vertebrates (Nelken, 2004). Neurons in the inferior colliculus of cats display the same firing pattern when a stimulus composed of a signal with or without background noise is presented, indicating a representation of the signal only (Chechik et al., 2006), and in auditory cortex, neurons show locking to the amplitude modulations of a low level sound but not to the noise it is embedded in (Las et al., 2005). In insects, the principle has been found in bushcrickets, separating single males from background choruses (Römer and Krusch, 2000) and has been suggested to be at work in crickets as well (Pollack, 1988). However, the selective coding principle is not clearly defined in an information theoretical framework. Here, we formalize the hypothesis and accordingly make two predictions about the change of information transfer with adaptation: (i) information conveyed about the "background" should decrease, while (ii) transmission for the high-intensity signals should increase. Indeed, we find that the neural output of the AN2 conveys less information about the first two peaks ("background") of the stimulus distribution (Figure 6.13B), supporting the selective coding hypothesis. However, if we take only the mutual information between the loudest part of the signal and the neural response into account, information rate remains nearly constant (Figure 6.13C). This contradicts the selective coding hypothesis.

The computational role of adaptation

Surprisingly, our results suggest that instead of improving sensory coding, adaptation in the AN2 decreases information transmission and leaves higher processing centers in the cricket brain with less (or at most equal) information, regardless of what part of the stimulus is considered. What could be the reason for this?

Adaptation leads to a selective decrease in the mutual information for the low-intensity sounds, mainly by shifting the stimulus-response curves towards higher stimulus intensities. As a consequence, spike counts are reduced for low-intensity signals (cf. Figure 6.10). Through adaptation to the trimodal stimulus, average spike counts in response to the 0 dB test stimulus decreased for 18 of 20 cells (mean decrease: 42 %, standard deviation: 31 %) compared to the spike counts after adaptation to the bimodal stimulus. Thus, in an ecological setting, where background signals are present and foreground signals are changing their presence dynamically, background signals transmitted to downstream neurons by the AN2 will be reduced. We speculate that this might reduce the potential interference of "background" and "foreground" spikes in downstream processing. However, the observed response curves do not represent an optimal solution for the task of filtering out the most intense part of the stimulus.

Additionally, the algorithm behind adaptation could serve the goal of enhancing the representation of even louder signals, occurring with less probability. Examples, in which optimal coding is not used to maximize the average information gained about high probability stimuli include auditory receptors of locusts which seem to maximize the information gained about specific, but less often occurring aspects of the stimuli (Machens et al., 2005) and stimulus specific adaptation in single neurons of auditory cortex that leads to an enhanced representation of low-probability sounds deviating from the distribution of the surrounding signal (Ulanovsky et al., 2003). This can be seen as a form of novelty detection, where part of the dynamic range is preserved for even louder sounds in a way that the sensory pathway is always able to detect brief, transient high-intensity signals (Ringach and Malone, 2007). In order to test if the representation of loud signals is enhanced, we calculated information transmission for a signal distribution that has an additional peak (modeled by a fourth Gaussian distribution with mean $\mu_4 = +6$ dB). Indeed, we find that adaptation increases mutual information in the stimulus range from 4.5 dB to 7.5 dB for all cells. This increase is significant in 6 of 20 cells, but it is not statistically significant on the population level (the 95 % interval of the combined posterior density ranges from −0.015 bit to 0.243 bit).

In this context, it should also be noted that the AN2 neuron in crickets may serve several functions. Under most stimulus conditions, relatively low firing rates will likely monitor slowly changing signals as observed in the present study (up to about 5 Hz).

The AN2, however, can also operate in a burst mode with high intra-burst firing rates for the detection of bat calls (Marsat and Pollack, 2006) for which our analysis is not appropriate. Nevertheless, low firing rates are likely to transmit relevant information since input-response curves built from spike counts similar to those in the present study are maintained at somewhat higher thresholds in wingless cricket morphs that are not at risk from bat predation (Pollack and Martins, 2007). In addition to its relevance for slow signal features, the adaptation time course reported here is likely to adjust the operating point of the faster response dynamics (i. e., bursts). Apart from possible physiological limitation, the findings we report here could be the result of a trade-off between setting the operating regime for the bursting mode on the one hand and suppression of background noise on slower time scales on the other hand.

Generally, a neural system may achieve improved performance by means of different mechanisms as has been shown in a recent modeling study (Schwabe and Obermayer, 2005). Depending on the specific physiological constraints, the resulting neural representation can be optimal or a trade-off between optimality and the flexibility of the neural circuit. Indeed, we found that, for example, the slope of the stimulus-response curve is not steep enough for optimal encoding of only the loudest peak of the bimodal or trimodal stimulus distribution. Possibly the neural gain can only increase to a limited value, leading to a decrease – rather than an increase – in information transmission for the chosen experimental paradigm.

Recapitulation

We tested whether one of two computational principles may underlie adaptive sensory processing in an auditory interneuron (AN2) of crickets: infomax and selective coding. Infomax, a form of optimal coding, predicts that adaptation should maximize the transmitted information, taking into account the (changing) statistics of sensory stimuli. In contrast, selective coding predicts that encoding of relevant signals should increase while background signals should be suppressed.

Our results lead to the following conclusions about the computation performed by the neuron under study: (i) Adaptation does not maximize the amount of transmitted information when considering the whole range of input signals. This finding is inconsistent with the infomax hypothesis. (ii) Adaptation decreases the amount of information that is transmitted about background signals. This might facilitate the detection of signals in background noise. However, we do not find an increase of information transfer for the relevant (loudest) part of the signal range. Our results are thus also not fully consistent with the selective coding hypothesis.

Although we cannot give a conclusive answer on what the adaptation-induced selective suppression in the AN2 serves for yet, the paradigm we propose here is rather general and applicable to other sensory systems. Importantly, our approach allows quantifying the improvement in neural coding without requiring that neural response curves achieve optimality. Measuring the information between the proposed relevant stimulus and the neural output allows testing for different hypotheses on what a sensory pathway actually adapts to. Ultimately, testing various hypotheses on different stimulus ensembles will yield important insights on what is or what is not the relevant part of a sensory environment for a given sensory unit.

A. Orientation selectivity index and preferred orientation

Orientation selectivity index (OSI)

Orientation tuning was analyzed using the orientation selectivity index (Swindale, 1998), which is given by

$$\mathrm{OSI} = \frac{\sqrt{a^2 + b^2}}{\sum_{i=1}^{N} R(\phi_i)}, \tag{A.1}$$

where the components a and b are defined by

$$a = \sum_{i=1}^{N} R(\phi_i)\cos(2\phi_i); \quad b = \sum_{i=1}^{N} R(\phi_i)\sin(2\phi_i). \tag{A.2}$$

$R(\phi_i)$ is the value of the quantity whose tuning is to be analyzed (e. g., the spiking activity) in response to a grating stimulus of orientation ϕ_i. For all measurements, the stimulus orientations ϕ_i, $i = 1 \ldots N$, are uniformly distributed over −90° to +90°. Then the OSI is a measure of tuning sharpness ranging from 0 (unselective) to 1 (perfectly selective).

Map OSI

In addition, the OSI was used to characterize the sharpness of the recurrent input a cell receives based on the orientation preference map. To calculate this *map OSI*, we estimate the local orientation preference distribution by binning the orientation preference of all pixels within a radius of 250 µm around a cell into bins of 10° size; the number of cells in each bin replaced the quantity $R(\phi_i)$. Figure B.1B shows the map OSI of our artificial orientation preference map (Figure B.1A). The map OSI ranges from almost 0 for cells close to pinwheel centers to almost 1 in the linear zones of the iso-orientation domains.

Preferred orientation

Preferred orientation is calculated from the vector average of the responses (Swindale, 1998). Using the components a and b, the preferred orientation is given by

$$\mathrm{PO} = \frac{1}{2}\arctan(b/a). \tag{A.3}$$

B. Orientation preference map for the network models

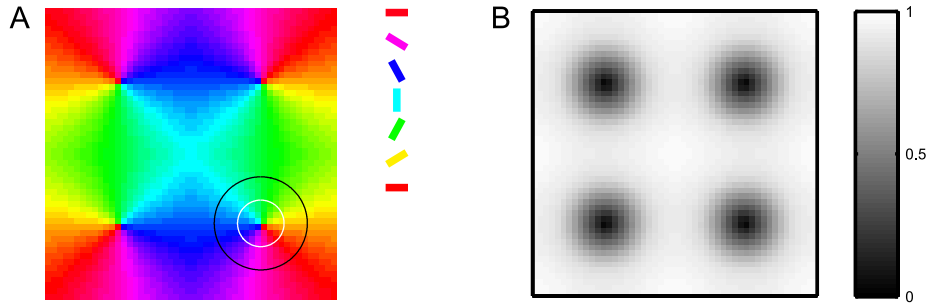


Figure B.1 | Artificial orientation preference map and map OSI. **A** | Artificial orientation map with four pinwheels of alternating handedness arranged on a 2-dimensional grid. The white (black) circle denotes the one-(two-) σ-area corresponding to the radial Gaussian synaptic connection profile ($\sigma_E = \sigma_I = 125\,\mu m$). **B** | Map OSI of the artificial orientation map. Pinwheel centers appear in black.

All orientation maps used in the computer simulations were composed of four pinwheels with alternating handedness, arranged on a grid of 50×50 (Hodgkin-Huxley network) or 64×64 (firing rate network) pixels (Figure B.1A). Periodic boundary conditions were used. The spatial scale was calibrated against experimental orientation maps (i) by matching the relationship between the orientation selectivity index (OSI) of the local orientation map (map OSI) and the OSI of recurrent synaptic inputs (which are drawn from a spatially Gaussian distribution) and (ii) by matching the overall distributions of local map OSIs. Details on the map calibration can be found in Stimberg et al. (2009). The best match between experimentally measured orientation maps and the artificial map consisting of 50×50 pixels is achieved with $\sigma_{rec} = 4$ pixels (corresponding to 125 µm, Mariño et al., 2005). This choice corresponds to an average nearest neighbor pinwheel distance of 781 µm and a pinwheel density of 1.64 pinwheels/mm^2 for the artificial map. The analysis of the tuning properties of the model cells was restricted to excitatory cells with a map OSI < 0.9 (the highest map OSI in the experimental data of Mariño et al., 2005 is 0.86). This was done because OSI values greater than 0.9 are only possible for cells which lie in the corners of the square around a pinwheel center in the artificial orientation map (cf. Figure B.1B), resulting in a bias for the orientation preference of such cells. However, all our results remain valid if the cells with map OSI > 0.9 are included in the analysis.

C. The Hodgkin Huxley network model

The network consists of a two-dimensional layer of excitatory and inhibitory neurons, which receive orientation tuned feed-forward input. Figure C.1 shows the general architecture of the network model. 2500 excitatory neurons are arranged in a 50×50 grid, while the 833 inhibitory neurons are placed at random grid locations. The model thus contains 75 % excitatory and 25 % inhibitory cells. While some experimental evidence (Beaulieu et al., 1992) points to a lower fraction of inhibitory neurons (20 %), we have chosen 25 % to ensure a smooth covering of the orientation map. All model cells receive input via afferent, recurrent and background synaptic currents. Orientation tuning is determined according to the calibrated artificial map shown in Figure B.1A.

C.1. Single cell model

For each model neuron, the dynamics of the membrane potential V_m is described by

$$C_\mathrm{m} \frac{dV_\mathrm{m}}{dt} = -g_\mathrm{L} \left(V_\mathrm{m} - E_\mathrm{L} \right) - \sum_\mathrm{int} I_\mathrm{int} - I_\mathrm{syn} - I_\mathrm{bg}, \tag{C.1}$$

where I_syn, I_int, and I_bg denote the voltage-dependent synaptic, intrinsic and background currents, g_L and E_L denote the leak conductance and its reversal potential, C_m denotes the membrane capacitance, and t the time (for parameters see Table C.1).

Each current I_int is described by an expression

$$I_\mathrm{int}(t) = \bar{g} m^\mathrm{M}(t) \, h^\mathrm{N}(t) \left(V_\mathrm{m}(t) - E \right), \tag{C.2}$$

where $\bar{g}$ is the peak conductance, E is the reversal potential, $m(t)$ and $h(t)$ are the activation and inactivation variables, and M, N are natural numbers. We included three voltage dependent currents: a fast Na^+ current and a delayed-rectifier K^+ current for the generation of action potentials, and a slow non-inactivating K^+ current responsible for spike frequency adaptation. For parameters and the functional forms of the activating and inactivating variables, see Destexhe and Paré (1999). The peak conductance of the non-inactivating K^+-current is multiplied by the factor 0.1 for inhibitory neurons, in order to reduce their spike-frequency adaptation compared to the excitatory neurons.

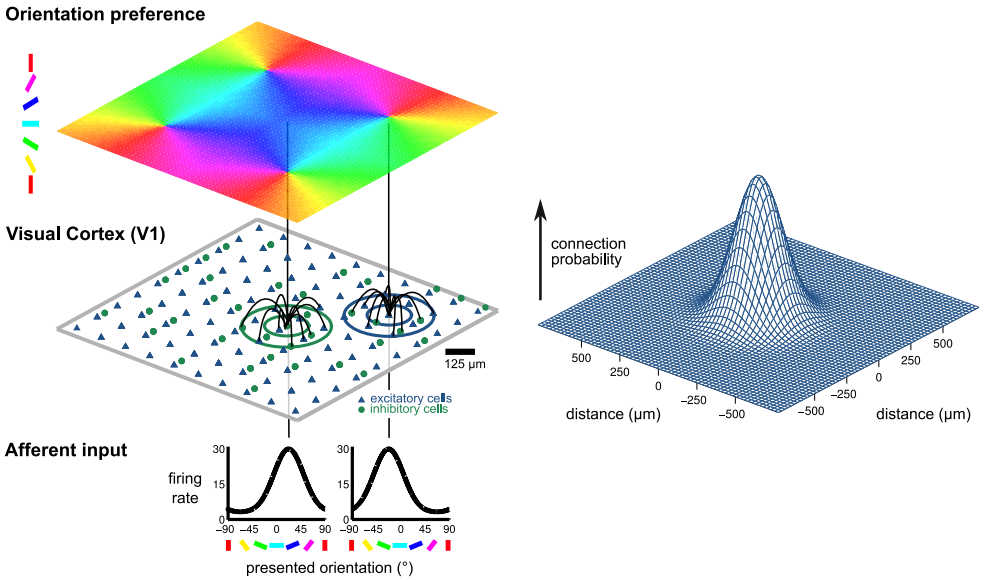


Figure C.1 | Network architecture. The cartoon shows the general architecture of the Hodgkin Huxley network model: A layer of excitatory neurons (blue triangles) and inhibitory neurons (green circles) receives afferent as well as lateral input. Cells are placed on a grid, where the fewer inhibitory neurons occupy random positions (for simplicity, only 10 × 10 excitatory neurons are shown). This network models a patch of cortex 1.56 mm × 1.56 mm in size (see scale bar). Examples for lateral connections are indicated for an inhibitory neuron in an iso-orientation domain (lines connecting to the neuron in the center) and an excitatory cell close to a pinwheel center (lines connecting to the neuron at the right). The default values for the connection probabilities are given by the same circular Gaussian with a standard deviation corresponding to 125 µm (right) for all types of connections. The preferred orientation of each neuron is assigned according to its position in an artificial orientation map with four pinwheels (top). Circular Gaussian tuning curves with standard deviation 27.5° (bottom) determine the input firing rate for each neuron, depending on the presented orientation and the orientation preference of the cell.

C.2. Synaptic background inputs

The synaptic background current I_{bg} is determined, independently for each cell, from one excitatory and one inhibitory background conductance,

$$I_{\mathrm{bg}} = g^{\mathrm{e}}_{\mathrm{bg}}\left(V_{\mathrm{m}}(t) - E^{\mathrm{e}}_{\mathrm{bg}}\right) + g^{\mathrm{i}}_{\mathrm{bg}}\left(V_{\mathrm{m}}(t) - E^{\mathrm{i}}_{\mathrm{bg}}\right), \tag{C.3}$$

which follow a stochastic process similar to an Ornstein-Uhlenbeck process (Destexhe et al., 2001). g_{bg} is given by

$$g_{\mathrm{bg}}(t+\Delta t) = g^{0}_{\mathrm{bg}} + \left[g_{\mathrm{bg}}(t) - g^{0}_{\mathrm{bg}}\right]\exp\left(-\Delta t/\tau_{\mathrm{bg}}\right) + \sqrt{1-\exp\left(-2\Delta t/\tau_{\mathrm{bg}}\right)}N(0,\sigma_{\mathrm{bg}}), \tag{C.4}$$

Table C.1. | Parameters for the Hodgkin-Huxley neuron model.

Parameter	Description	Value
Cell properties		
C_m	Membrane capacity	0.35 nF
g_L^E	Leak conductance of excitatory cells	15.7 nS
g_L^I	Leak conductance of inhibitory cells	31.4 nS
E_L	Leak reversal potential	−80 mV
Background activity		
$g_\mathrm{bg}^\mathrm{e0}$	Mean excitatory background conductance	$0.56 \cdot g_\mathrm{L}$
$g_\mathrm{bg}^\mathrm{i0}$	Mean inhibitory background conductance	$1.84 \cdot g_\mathrm{L}$
$\tau_\mathrm{bg}^\mathrm{e}$	Excitatory time constant	2.7 ms
$\tau_\mathrm{bg}^\mathrm{i}$	Inhibitory time constant	10.5 ms
$\sigma_\mathrm{bg}^\mathrm{e}$	Standard deviation of excitatory conductance	$0.01 \cdot g_\mathrm{L}$
$\sigma_\mathrm{bg}^\mathrm{i}$	Standard deviation of inhibitory conductance	$0.01 \cdot g_\mathrm{L}$
E_bg^e	Reversal potential of excitatory conductance	−5 mV
E_bg^i	Reversal potential of inhibitory conductance	−70 mV

where g_bg^0 is the average conductance, τ_bg is the background synaptic time constant, and $N(0, \sigma_\mathrm{bg})$ is a normally distributed random number with zero mean and standard deviation σ_bg. For the parameter values, see Table C.1.

C.3. Afferent and recurrent synaptic connections

The synaptic current for neurons of the population $j \in \{\mathrm{E,I}\}$ is composed of the input received from $N_\mathrm{Aff} = 20$ afferent excitatory, $N_\mathrm{E} = 100$ recurrent excitatory and $N_\mathrm{I} = 50$ recurrent inhibitory neurons. The currents are given by

$$I_\mathrm{syn}^j(t) = \Bigg[\underbrace{\frac{\overline{g}_j^\mathrm{Aff}}{N_\mathrm{Aff}} g^\mathrm{Aff}(t)}_{\text{afferent excitatory}} + \underbrace{\frac{\overline{g}_{j\mathrm{E}}}{N_\mathrm{E}} \big(\alpha g_\mathrm{E1}(t) + (1-\alpha) g_\mathrm{E2}(t)\big)}_{\text{recurrent excitatory}} \Bigg] (V_\mathrm{m} - E_\mathrm{e}) \\ \underbrace{- \frac{\overline{g}_{j\mathrm{I}}}{N_\mathrm{I}} g_\mathrm{I}(t)(V_\mathrm{m} - E_\mathrm{i})}_{\text{recurrent inhibitory}}, \quad \text{(C.5)}$$

where $g(t)$ are the time-dependent conductances, E_e and E_i are the synaptic reversal potentials, $\alpha = 0.7$ determines the number of fast (70 %) versus slow (30 %) receptors of

recurrent excitatory synapses, and $\bar{g}$ are scale factors (peak conductances; the values are summarized in Table C.2).

Table C.2. | Parameters for connectivity patterns and synaptic properties.

Parameter	Description	Value
Connectivity		
N_E	Excitatory synaptic connections per cell	100
N_I	Inhibitory synaptic connections per cell	50
$\sigma_\mathrm{E} = \sigma_\mathrm{I}$	Spread of recurrent connections (SD)	4 pixels (125 µm)
Synaptic properties		
E_e	Reversal potential excitatory synapses	0 mV
E_i	Reversal potential inhibitory synapses	−80 mV
τ_E	Time constant of AMPA-like synapses	5 ms
τ_I	Time constant of $\mathrm{GABA_A}$-like synapses	5 ms
τ_1	Time constant of NMDA-like synapses	80 ms
τ_2	Time constant of NMDA-like synapses	2 ms
$\mu_\mathrm{E}^\mathrm{delay}$	Mean excitatory synaptic delay	4 ms
$\sigma_\mathrm{E}^\mathrm{delay}$	Standard deviation of excitatory synaptic delays	2 ms
$\mu_\mathrm{I}^\mathrm{delay}$	Mean inhibitory synaptic delay	1.25 ms
$\sigma_\mathrm{I}^\mathrm{delay}$	Standard deviation of inhibitory synaptic delays	1 ms
Afferent synaptic strengths		
$\overline{g}_\mathrm{E}^\mathrm{Aff}$	Afferent peak conductance to excitatory cells	$9 \cdot g_\mathrm{L}^\mathrm{E}$
$\overline{g}_\mathrm{I}^\mathrm{Aff}$	Afferent peak conductance to inhibitory cells	$0.73 \cdot \overline{g}_\mathrm{E}^\mathrm{Aff}$
Recurrent synaptic strengths		
$\overline{g}_\mathrm{II}$	Peak conductance from inh. to inh. cells	$1.33 \cdot \overline{g}_\mathrm{E}^\mathrm{Aff}$
$\overline{g}_\mathrm{EI}$	Peak conductance from inh. to exc. cells	$1.33 \cdot \overline{g}_\mathrm{E}^\mathrm{Aff}$

Recurrent connections to a given neuron were drawn randomly from the Gaussian probability distribution:

$$P_j(x) = \begin{cases} 0 & \text{for } x = 0 \text{ (no self-connections)} \\ 1/\sqrt{2\pi\sigma_j}\exp\left(-x^2/2\sigma_j^2\right) & \text{otherwise,} \end{cases} \tag{C.6}$$

where x is the distance to the presynaptic neuron (in pixels) and $\sigma_\mathrm{E} = \sigma_\mathrm{I} = 4$ (in pixels), corresponding to 125 µm. Periodic boundary conditions were used. If a presynaptic neuron generated a spike, this spike was transferred to the postsynaptic neuron with a certain delay. This delay was determined for every cell (i. e., the delay to all postsynaptic

targets was identical) by drawing a random number from a normal distribution with mean $\mu_{\mathrm{E}}^{\mathrm{delay}}$ ($\mu_{\mathrm{I}}^{\mathrm{delay}}$) and standard deviation $\sigma_{\mathrm{E}}^{\mathrm{delay}}$ ($\sigma_{\mathrm{I}}^{\mathrm{delay}}$) for the excitatory (inhibitory) population (see Table C.2). Delays that were smaller than a single integration step dt were set to dt.

C.4. Synapse model

Three different kinds of synapses are included: $\mathrm{GABA_A}$-like inhibitory, fast AMPA-like excitatory, and slow NMDA-like excitatory synapses. Fast excitatory (g^{Aff} and g_{El}) and inhibitory synaptic (g_{I}) conductances arising from a single synaptic event are modeled as an exponential postsynaptic conductance, multiple events are added linearly. The synaptic conductance in response to spikes at time t_j^k is thus given by

$$g_j(t) = \frac{1}{\tau_j} \sum_{t^k<t} \exp\left(-(t^k-t)/\tau_j\right), \tag{C.7}$$

where τ_j with $j \in \{\mathrm{E}, \mathrm{I}\}$ is the synaptic time-constant. Synaptic events occurring at the slow NMDA-like synapses are modeled by a difference of two exponentials with multiple events again adding linearly:

$$g_{\mathrm{E2}}(t) = \frac{1}{\tau_1 - \tau_2} \sum_{t^k<t} \left(\exp(-(t^k-t)/\tau_1) - \exp(-(t^k-t)/\tau_2)\right), \tag{C.8}$$

where τ_1 and τ_2 are time constants and t^k denotes the time of an incoming spike. Both types of conductances are normalized so that for a single incoming spike, the integral over time is 1. Parameters are summarized in Table C.2.

C.5. Afferent input

The afferent input neurons were not modeled explicitly, but represented by Poisson spike trains generated from an afferent firing rate f_{Aff}.

Time-invariant afferent input

In simulations with a time-invariant input, for every network parameterization, the stimuli $\theta_{\mathrm{stim}} \in \{-90°, -67.5°, -45°, -22.5°, 0°, 22.5°, 45°, 67.5°\}$ were presented to the network. The afferent firing rate is then given by

$$f_{\mathrm{Aff}}(t, \theta_{\mathrm{stim}}) = 30\,\mathrm{Hz} \cdot \left(r_{\mathrm{Aff}}(\theta_{\mathrm{stim}}) + r_{\mathrm{base}}\right), \tag{C.9}$$

$$r_{\mathrm{Aff}}(\theta_{\mathrm{stim}}) = (1 - r_{\mathrm{base}}) \cdot \exp\left(-\frac{(\theta_{\mathrm{stim}} - \theta)^2}{(2\sigma_{\mathrm{Aff}})^2}\right), \tag{C.10}$$

where θ is the preferred orientation chosen according to the neuron's location in the artificial orientation map (Figure B.1A), $\sigma_{\text{Aff}} = 27.5°$ is the orientation tuning width, and $r_{\text{base}} = 0.1$ is a baseline response.

Dynamics of the afferent input

For simulations under the reverse correlation paradigm, the stimulus orientation θ_{stim} varies with time and is one of 16 orientations equidistantly spaced between −90° and 88.75°, or a "blank" stimulus. The afferent firing rate is given by

$$f_{\text{Aff}}(t, \theta_{\text{stim}}) = \left[30\,\text{Hz}\left(\int_0^t r_{\text{Aff}}\big(\theta_{\text{stim}}(t')\big) \cdot h(t - t')dt' + r_{\text{base}}\right)\right]^+, \tag{C.11}$$

where $\theta_{\text{stim}}(t)$ is the orientation of the stimulus presented at time t and h is the temporal response envelope. The afferent firing rate for the "blank" stimulus is just the baseline response, i. e., r_{Aff} ("blank") = 0. For each cell, h is chosen randomly with probability 0.3, 0.3, 0.2 and 0.2 as one of four temporal kernels h_p, $p = 1 \ldots 4$ (Figure C.2), in order to account for the variability observed in the temporal responses of LGN cells and V1 simple cells in cat (Alonso et al., 2001; Wolfe and Palmer, 1998). The four kernels have a temporal profile

$$h(t) = \begin{cases} \Gamma(s)^{-1}(\tau_{\text{Aff}})^{-s} t^{s-1} \exp\left(-1/t\right) \cos(2\pi \cdot f_\omega\, t + \Phi) & t \geq 0, \\ 0 & t < 0, \end{cases} \tag{C.12}$$

which closely resembles the profiles found in V1 simple cells (Chen et al., 2001). $\Gamma(s)$ denotes the gamma function; parameters are summarized in Table C.3. Each kernel was scaled such that, if the neurons were driven by afferent input of that kernel alone, this neuron would fire at 6 Hz for the preferred orientation stimulus.

Table C.3. | Parameters for the temporal input kernels.

Parameter	Description	h_1	h_2	h_3	h_4
τ_{Aff}	Time constant	16 ms	16 ms	16 ms	30 ms
s	Skewness of the gamma distribution	2	2	2	1
f_ω	Frequency	7.2 Hz	7.2 Hz	7.2 Hz	5 Hz
Φ	Phase shift	0.1	−0.15	0.4	−0.5

In the simulations with more uniform afferent input (Figure 4.4), the temporal kernels were assigned randomly to the individual afferent input synapses instead of assigning one temporal kernel to all synapses of a given cell. Therefore, each cell received input

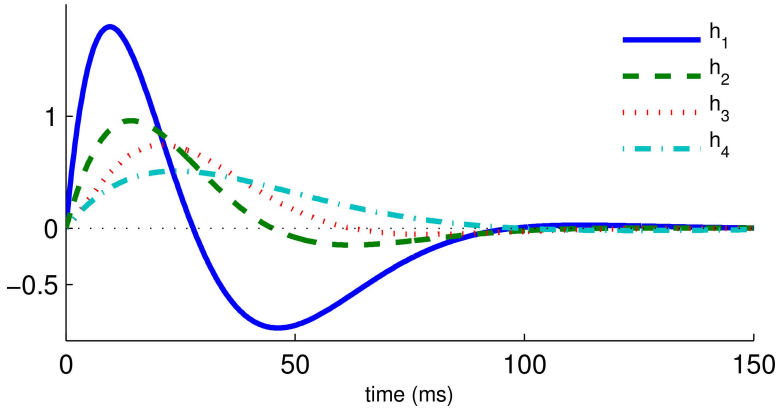


Figure C.2 | Temporal kernels used for modeling the input in the reverse correlation simulations. The graphs show the temporal kernels used in the reverse correlation experiments for a stimulus with the cell's preferred orientation. For parameters see Table C.3; kernels h_1 (solid line) and h_2 (dashed line) where each used with probability 0.2, kernels h_3 (dotted line) and h_4 (dash-dotted line) where each used with probability 0.3.

from afferent synapses with different temporal behavior. This had the effect of making the effective afferent input time course more similar across cells.

C.6. Simulating the network

All simulations were performed in Matlab (Mathworks, Natick, MA) with a fixed time step of $dt = 0.25$ ms, employing an integration scheme that alternately updates the membrane potential and the gating variables. Such a scheme is accurate to second order in dt (Mascagni and Sherman, 1998). The network was simulated for 1.5 s, and the mean conductances (g_e and g_i), the mean membrane potential (V_m) and the firing rate (f) were then calculated for every cell from the last 1 s. For calculating the mean membrane potential, the values from 4 ms before until 7 ms after every spike were first removed. From the resulting values we then calculated the mean value and subtracted the resting potential of the cell ($V_m = -64.5$ mV) in order to calculate the OSI. The total excitatory conductance g_e is the sum of the afferent conductance, the recurrent excitatory conductance of the slow and the fast excitatory synapses, and the excitatory background conductance. The total inhibitory conductance g_i is the sum of the recurrent inhibitory conductance, the conductance of the non-inactivating K^+ current, and the inhibitory background conductance.

For analyzing the dynamics of orientation tuning, the network was simulated for at least 100 s; the input orientation was every 20 ms randomly chosen as one of 16 orienta-

tions or a "blank" stimulus. The spike output of the excitatory cells was then analyzed using the reverse-correlation technique as described in previous reports (Ringach et al., 1997; Schummers et al., 2007).

A network model is termed "unstable" if – for the combination of values for $\overline{g}_{\mathrm{EE}}$ and $\overline{g}_{\mathrm{IE}}$– model neurons showed strong responses (firing rate across cells of more than 100 Hz) and remain at these high firing rates if the afferent input is turned off; i. e., the network shows self-sustained activity. In such a network, all cells display untuned firing (the OSI of the firing rate is < 0.3 for all cells).

D. The firing rate network model

The firing rate (mean-field) model consists of an excitatory and an inhibitory population of threshold-linear neurons, each arranged in a two-dimensional grid of 64 × 64 cells. Lateral connection strengths between the neurons are weighted according to a Gaussian distribution with the same spatial extent ($\sigma = 125$ µm) for excitation and inhibition (Mariño et al., 2005; Stimberg et al., 2009); periodic boundary conditions were used. Both excitatory and inhibitory cells receive identically tuned feed-forward input, determined by an artificial orientation preference map consisting of four pinwheels (Figure B.1).

D.1. Dynamics of the firing rate

The network model is similar to the models of Kang et al. (2003) and Stimberg et al. (2009), but was extended to include dynamic synaptic connections. The dynamics of the firing rates of the excitatory (m_E) and inhibitory (m_I) population are given by

$$\begin{aligned}\tau_\mathrm{E}\frac{dm_\mathrm{E}(\vec{r},t)}{dt} = -m_\mathrm{E}(\vec{r},t) + \Bigg[I_\mathrm{Aff}(\vec{r}) + \sum_{\vec{r}\,'} \Phi_\mathrm{EE}(\vec{r},\vec{r}\,',t) m_\mathrm{E}(\vec{r}\,',t) \\ + \sum_{\vec{r}\,'} \Phi_\mathrm{EI}(\vec{r},\vec{r}\,',t) m_\mathrm{I}(\vec{r}\,',t) \Bigg]^+ ,\end{aligned} \tag{D.1}$$

$$\begin{aligned}\tau_\mathrm{I}\frac{dm_\mathrm{I}(\vec{r},t)}{dt} = -m_\mathrm{I}(\vec{r},t) + \Bigg[I_\mathrm{Aff}(\vec{r}) + \sum_{\vec{r}\,'} \Phi_\mathrm{IE}(\vec{r},\vec{r}\,',t) m_\mathrm{E}(\vec{r}\,',t) \\ + \sum_{\vec{r}\,'} \Phi_\mathrm{II}(\vec{r},\vec{r}\,',t) m_\mathrm{I}(\vec{r}\,',t) \Bigg]^+ ,\end{aligned} \tag{D.2}$$

where $\vec{r}$ is a vector describing a neuron's position in the layer (i. e., the 2D cortical coordinate), $\tau_\mathrm{E} = 12$ ms ($\tau_\mathrm{I} = 4$ ms) denotes the excitatory (inhibitory) synaptic time constant, and $[\dots]^+$ is the rectification function. The functions $\Phi_{ij}(\vec{r},\vec{r}\,',t)$ denote the time-dependent weights of the recurrent connections between a presynaptic neuron from population j at position $\vec{r}\,'$ and a postsynaptic neuron from population i at position $\vec{r}$, where $i, j \in \{\mathrm{E},\mathrm{I}\}$ denote the excitatory and inhibitory populations. These recurrent weights are given by

$$\Phi_{ij}(\vec{r},\vec{r}\,',t) = S_{ij} \cdot \varphi(\vec{r},\vec{r}\,') \cdot x_{ij}(\vec{r},\vec{r}\,',t), \tag{D.3}$$

where the S_{ij} denotes the connection strength, φ denotes the spatial connectivity profile, and x_{ij} governs synaptic depression.

D.2. Recurrent connectivity

The factors S_{ij} denote the summed strength of the connections between the presynaptic population j and one neuron from the postsynaptic population j ($i, j \in \{\mathrm{E}, \mathrm{I}\}$). In contrast to Kang et al. (2003) we include self-inhibition ($S_{\mathrm{II}} \neq 0$). We used the following standard values for the connections strengths: $S_{\mathrm{EE}} = 3.4$, $S_{\mathrm{IE}} = 4.7$, $S_{\mathrm{II}} = 1.3$, and $S_{\mathrm{EI}} = 1.4$. These values correspond to a strongly recurrent, balanced operating regime of the network, in line with the results from Chapter 3 (see also Mariño et al., 2005; Schummers et al., 2007; Stimberg et al., 2009; Wimmer et al., 2009). The function $\varphi(\vec{r}, \vec{r}\,')$, denoting the normalized spatial profile of cortical interactions, is defined by

$$\varphi(\vec{r}, \vec{r}\,') = \frac{1}{\sqrt{2\pi\sigma}} \exp\left(-\frac{(\vec{r} - \vec{r}\,')^2}{2\sigma^2}\right), \tag{D.4}$$

and depends only on the cortical distance of interacting neurons. The standard deviation of the spread of recurrent connections is $\sigma = 8$ pixels (corresponding to $\sigma = 125\,\mu\mathrm{m}$). This spatial connectivity profile was the same for all types of connections.

D.3. Synaptic depression model

Intracortical synaptic depression was described by the mean-field equation derived by Tsodyks et al. (1998):

$$\frac{dx_{ij}(\vec{r}, \vec{r}\,', t)}{dt} = \frac{1 - x_{ij}(\vec{r}, \vec{r}\,', t)}{\tau_{\mathrm{rec}}} - U_{ij} x_{ij}(\vec{r}, \vec{r}\,', t) m_j(\vec{r}\,', t) \tag{D.5}$$

where τ_{rec} is the recovery time constant, m_j is the presynaptic firing rate and U_{ij} is the utilization of synaptic efficacy. In the detailed model of synaptic depression, involving the arrival times of individual spikes (Abbott et al., 1997; Tsodyks and Markram, 1997), each presynaptic spike activates a fraction U_{ij} of resources, which then quickly inactivate (with time constant of few milliseconds) and recover with a larger time constant of τ_{rec}. We used the same time constant $\tau_{\mathrm{rec}} = 300\,\mathrm{ms}$ for all the connections. The mean-field equation was derived under the assumption that the presynaptic firing rate follows a Poisson process. Besides the absolute connection strengths S_{ij}, the major parameters of the model are the U_{ij}'s, which determine the dynamics of the synaptic response. Experimentally, U values were found to cover almost the entire range from

0 to 1 (Tsodyks and Markram, 1997). In agreement with experimental data (Galarreta and Hestrin, 1998; Varela et al., 1999), we assume that both excitatory and inhibitory recurrent connections undergo depression. Thus, in the exploration of the phase space, we systematically vary U_{EE}, U_{IE}, U_{II} and U_{EI}. For simplicity, afferent synapses do not undergo synaptic depression in our model. This can be justified by experimental data indicating that synaptic depression is weaker for feed-forward synapses compared to intracortical synapses (Yoshimura et al., 2000), and by a modeling study (Chelaru and Dragoi, 2008) showing that the major factor for adaptation-induced response changes is intracortical synaptic depression. Furthermore, there is experimental evidence (in-vivo) that thalamocortical synapses are already maintained at high levels of depression by spontaneous activity (i. e., there is little additional depression; Boudreau and Ferster, 2005).

D.4. Afferent input

The afferent input I_{Aff} is given by $I_{Aff} = A + B\cos\big(2(\theta - \theta_{stim})\big)$, where θ_{stim} is the orientation of the stimulus presented to the network, θ is the preferred orientation given by the corresponding location in the orientation map, A is the mean input and B is the orientation modulation amplitude. For all simulations we chose $A = B = 20$, which results in a steady state firing rate of 47 Hz in the control condition.

D.5. Simulating the network

The differential equations were numerically integrated, using an explicit Runge-Kutta (4, 5) method (implemented by the function `ode45`; Matlab; Mathworks, Natick, MA). In the simulation of adaptation experiments we distinguish control and test condition. In order to measure the control tuning curves, synaptic depression was turned off by setting $U_{EE} = U_{IE} = U_{II} = U_{EI} = 0$ so that the synaptic connections are time-invariant with a summed strengths of S_{EE}, S_{EI}, S_{II} and S_{EI}. The network was then simulated for stimuli of different orientations and the firing rates of the excitatory neurons were recorded. For obtaining the tuning curves in the test condition, the network was first adapted for 1000 ms to a stimulus of fixed orientation 0°. Synaptic depression was then kept fixed, by setting $U_{EE} = U_{IE} = U_{II} = U_{EI} = 0$ and holding x_{ij} at their adapted values, before the firing rates of excitatory neurons as well as their excitatory and inhibitory inputs were recorded. Assuming an asynchronous state, in the steady state limit, the inputs are proportional to the total mean conductance values in a conductance based neuron model (Shriki et al., 2003). For obtaining tuning curves, we recorded every unit's response to stimuli with one of 32 equidistantly spaced orientations θ_{stim} between

−90° and +84.38°. All numerical simulations were performed using the 4-pinwheel map (Figure B.1) with periodic boundary conditions.

A network model is termed "unstable" if – for the combination of values for U_{EE}, U_{IE}, U_{II} and U_{EI}– either the network activity diverges, the network activity does not reach a stationary state (because of oscillatory behavior), or the model neurons showed elevated firing rates. Specifically, the unstable parameter region was determined numerically using the following criteria: First, the network model was adapted for 4 s (for each parameter combination) and the standard deviation of the depression variables x_{EE}, x_{IE}, x_{II} and x_{EI} was calculated across the time points from 2 s to 4 s. We defined a network model as unstable if any of these standard deviations exceeds the value 0.001. This criterion excludes divergent and oscillatory regimes. Second, we also excluded network models in which model neurons showed strong responses (firing rates of more than 100 Hz) in either the adapting or the test condition. Note that 100 Hz correspond to a more than two-fold increase in firing rate compared to the control condition, which is higher than the experimentally observed changes (Dragoi et al., 2000).

D.6. Analytical solution of the mean-field equations

The mean-field equations can also be solved analytically if (i) the interactions between different pinwheels are neglected and (ii) recurrent synapses do not undergo synaptic depression ($U_{EE} = U_{IE} = U_{II} = U_{EI} = 0$). Expressing the 2D position of a unit in polar coordinates $\vec{r} = (r, \phi)$, centered at the pinwheel, the stationary activity profile of the excitatory population is then given by (cf. Kang et al., 2003)

$$m_E(r, \theta) = Aa + Bb(r)\cos\big(2(\theta - \theta_{stim})\big), \tag{D.6}$$

where $\theta = \phi/2$ denotes the unit's preferred orientation. Please note that this notation differs from Kang et al. (2003), where θ corresponds to ϕ in our notation. We chose to consistently use θ for (preferred or presented) orientations.

The factor a is

$$a = \frac{1 - S_{EI} + S_{II}}{1 - S_{EE} + S_{EI}S_{IE} + S_{II} - S_{II}S_{EE}} \tag{D.7}$$

and determines the cortical gain of the mean activity. The expression $b(r)$ represents the effect of the cortical interactions on orientation tuning and is given by

$$b(r) = \int_0^\infty \frac{dk}{k}\left[\frac{1 - (S_{EI} + S_{II})\tilde{\varphi}(k)}{1 - D(k)}\right] J_1(kr), \tag{D.8}$$

where $\tilde{\varphi}$ is the Fourier transform of the connectivity profile φ, and $J_1(kr)$ is the Bessel function of order one. Following Kang et al. (2003), we call $D(k)$ the "cortical feedback kernel",

$$D(k) = S_{EE}\tilde{\varphi}(k) - S_{EI}S_{IE}\tilde{\varphi}^2(k) - S_{II}\tilde{\varphi}(k) + S_{II}S_{EE}\tilde{\varphi}^2(k). \tag{D.9}$$

Note that this kernel differs from the one analyzed in Kang et al. (2003) for two reasons: we include self-inhibition ($S_{II} \neq 0$); and we have identical connectivity profiles φ for the excitatory and the inhibitory populations.

The different parameter regimes ("phases") of the firing rate model are characterized by certain properties of the cortical feedback kernel $D(k)$:

- Unstable if $D(k) > 1$ for any k.
- Feedforward (FF) if $|D(k)| < 0.5$ for all k.
- Excitatory dominated (EXC) if $D(k) > 0$ for all k and $D_{max} = D(0)$.
- Recurrent (REC) if $D_{max} > 0.5$ and $1 - D(0) > 2(1 - D_{max})$.
- Inhibitory dominated (INH) if $D(k) < 0$ for some k and $D_{max} < 0.5$.

With the chosen values for the connections strengths (S_{EE}, S_{IE}, S_{II}, S_{EI}), the network operates in the recurrent (REC) regime.

Bibliography

Abbott, L. F., Varela, J. A., Sen, K., and Nelson, S. B. (1997). Synaptic depression and cortical gain control. *Science*, 275(5297):220–224.

Adelson, E. H. and Bergen, J. R. (1985). Spatiotemporal energy models for the perception of motion. *J Opt Soc Am A*, 2(2):284–299.

Adorján, P., Levitt, J. B., Lund, J. S., and Obermayer, K. (1999a). A model for the intracortical origin of orientation preference and tuning in macaque striate cortex. *Vis Neurosci*, 16(2):303–318.

Adorján, P., Piepenbrock, C., and Obermayer, K. (1999b). Contrast adaptation and infomax in visual cortical neurons. *Rev Neurosci*, 10(3-4):181–200.

Alonso, J. M., Usrey, W. M., and Reid, R. C. (2001). Rules of connectivity between geniculate cells and simple cells in cat primary visual cortex. *J Neurosci*, 21(11):4002–4015.

Atick, J. J. (1992). Could information theory provide an ecological theory of sensory processing. *Network: Computation in Neural Systems*, 3(2):231–251.

Averbeck, B. B., Latham, P. E., and Pouget, A. (2006). Neural correlations, population coding and computation. *Nat Rev Neurosci*, 7(5):358–366.

Baccus, S. A. and Meister, M. (2002). Fast and slow contrast adaptation in retinal circuitry. *Neuron*, 36(5):909–919.

Baden, T. and Hedwig, B. (2007). Neurite-specific Ca^{2+} dynamics underlying sound processing in an auditory interneurone. *Dev Neurobiol*, 67(1):68–80.

Barlow, H. B. (1961). Possible principles underlying the transformation of sensory messages. In Rosenblith, W. A., editor, *Sensory Communication*, pages 217–234. MIT Press, Cambridge, MA.

Barlow, H. B. (2001). Redundancy reduction revisited. *Network*, 12(3):241–253.

Beaulieu, C., Kisvarday, Z., Somogyi, P., Cynader, M., and Cowey, A. (1992). Quantitative distribution of GABA-immunopositive and -immunonegative neurons and synapses in the monkey striate cortex (area 17). *Cereb Cortex*, 2(4):295–309.

Bednar, J. A. and Miikkulainen, R. (2000). Tilt aftereffects in a self-organizing model of the primary visual cortex. *Neural Comput*, 12(7):1721–1740.

Bell, A. J. and Sejnowski, T. J. (1995). An information-maximization approach to blind separation and blind deconvolution. *Neural Comput*, 7(6):1129–1159.

Ben-Yishai, R., Bar-Or, R. L., and Sompolinsky, H. (1995). Theory of orientation tuning in visual cortex. *Proc Natl Acad Sci U S A*, 92(9):3844–3848.

Benda, J. and Hennig, R. M. (2008). Spike-frequency adaptation generates intensity invariance in a primary auditory interneuron. *J Comput Neurosci*, 24(2):113–136.

Benda, J. and Herz, A. V. M. (2003). A universal model for spike-frequency adaptation. *Neural Comput*, 15(11):2523–2564.

Binzegger, T., Douglas, R. J., and Martin, K. A. C. (2004). A quantitative map of the circuit of cat primary visual cortex. *J Neurosci*, 24(39):8441–8453.

Blakemore, C., Nachmias, J., and Sutton, P. (1970). The perceived spatial frequency shift: evidence for frequency-selective neurones in the human brain. *J Physiol*, 210(3):727–750.

Boudreau, C. E. and Ferster, D. (2005). Short-term depression in thalamocortical synapses of cat primary visual cortex. *J Neurosci*, 25(31):7179–7190.

Brenner, N., Bialek, W., and de Ruyter van Steveninck, R. (2000). Adaptive rescaling maximizes information transmission. *Neuron*, 26(3):695–702.

Butts, D. A. (2003). How much information is associated with a particular stimulus? *Network*, 14(2):177–187.

Butts, D. A. and Goldman, M. S. (2006). Tuning curves, neuronal variability, and sensory coding. *PLoS Biol*, 4(4):e92.

Cai, D., DeAngelis, G. C., and Freeman, R. D. (1997). Spatiotemporal receptive field organization in the lateral geniculate nucleus of cats and kittens. *J Neurophysiol*, 78(2):1045–1061.

Carandini, M. (2000). Visual cortex: Fatigue and adaptation. *Curr Biol*, 10(16):R605–R607.

Carandini, M., Demb, J. B., Mante, V., Tolhurst, D. J., Dan, Y., Olshausen, B. A., Gallant, J. L., and Rust, N. C. (2005). Do we know what the early visual system does? *J Neurosci*, 25(46):10577–10597.

Carandini, M. and Ferster, D. (1997). A tonic hyperpolarization underlying contrast adaptation in cat visual cortex. *Science*, 276(5314):949–952.

Carandini, M. and Ringach, D. L. (1997). Predictions of a recurrent model of orientation selectivity. *Vision Res*, 37(21):3061–3071.

Cardin, J. A., Palmer, L. A., and Contreras, D. (2007). Stimulus feature selectivity in excitatory and inhibitory neurons in primary visual cortex. *J Neurosci*, 27(39):10333–10344.

Celebrini, S., Thorpe, S., Trotter, Y., and Imbert, M. (1993). Dynamics of orientation coding in area V1 of the awake primate. *Vis Neurosci*, 10(5):811–825.

Chalupa, L. M. and Werner, J. S., editors (2003). *The Visual Neurosciences*. MIT Press.

Chechik, G., Anderson, M. J., Bar-Yosef, O., Young, E. D., Tishby, N., and Nelken, I. (2006). Reduction of information redundancy in the ascending auditory pathway. *Neuron*, 51(3):359–368.

Chelaru, M. I. and Dragoi, V. (2008). Asymmetric synaptic depression in cortical networks. *Cereb Cortex*, 18(4):771–788.

Chen, G., Dan, Y., and Li, C.-Y. (2005). Stimulation of non-classical receptive field enhances orientation selectivity in the cat. *J Physiol*, 564(Pt 1):233–243.

Chen, Y., Wang, Y., and Qian, N. (2001). Modeling V1 disparity tuning to time-varying stimuli. *J Neurophysiol*, 86(1):143–155.

Chung, S., Li, X., and Nelson, S. B. (2002). Short-term depression at thalamocortical synapses contributes to rapid adaptation of cortical sensory responses in vivo. *Neuron*, 34(3):437–446.

Clatworthy, P. L., Chirimuuta, M., Lauritzen, J. S., and Tolhurst, D. J. (2003). Coding of the contrasts in natural images by populations of neurons in primary visual cortex (V1). *Vision Res*, 43(18):1983–2001.

Clifford, C. W. (2002). Perceptual adaptation: motion parallels orientation. *Trends Cogn Sci*, 6(3):136–143.

Clifford, C. W. and Rhodes, G. (2005). *Fitting the mind to the world: Adaptation and after-effects in high-level vision*. Oxford University Press, USA.

Clifford, C. W., Webster, M. A., Stanley, G. B., Stocker, A. A., Kohn, A., Sharpee, T. O., and Schwartz, O. (2007). Visual adaptation: neural, psychological and computational aspects. *Vision Res*, 47(25):3125–3131.

Clifford, C. W., Wenderoth, P., and Spehar, B. (2000). A functional angle on some after-effects in cortical vision. *Proc Biol Sci*, 267(1454):1705–1710.

Clifford, C. W., Wyatt, A. M., Arnold, D. H., Smith, S. T., and Wenderoth, P. (2001). Orthogonal adaptation improves orientation discrimination. *Vision Res*, 41(2):151–159.

Cohen, M. R. and Newsome, W. T. (2009). Estimates of the contribution of single neurons to perception depend on timescale and noise correlation. *J Neurosci*, 29(20):6635–6648.

Cover, T. M. and Thomas, J. A. (1991). *Elements of information theory*. Wiley, New York.

Creutzfeldt, O. D. (1977). Generality of the functional structure of the neocortex. *Naturwissenschaften*, 64(10):507–517.

Crook, J. M., Kisvárday, Z. F., and Eysel, U. T. (1997). GABA-induced inactivation of functionally characterized sites in cat striate cortex: effects on orientation tuning and direction selectivity. *Vis Neurosci*, 14(1):141–158.

de Lafuente, V. and Romo, R. (2006). Neural correlate of subjective sensory experience gradually builds up across cortical areas. *Proc Natl Acad Sci U S A*, 103(39):14266–14271.

Dean, I., Harper, N. S., and McAlpine, D. (2005). Neural population coding of sound level adapts to stimulus statistics. *Nat Neurosci*, 8(12):1684–1689.

Deneve, S., Latham, P., and Pouget, A. (1999). Reading population codes: a neural implementation of ideal observers. *Nat Neurosci*, 2(8):740–745.

Destexhe, A., Mainen, Z. F., and Sejnowski, T. J. (1998). Kinetic models of synaptic transmission. In Koch, C. and Segev, I., editors, *Methods in neuronal modeling: from ions to networks*, pages 1–26. MIT Press, Cambridge, Mass, 2nd edition.

Destexhe, A. and Paré, D. (1999). Impact of network activity on the integrative properties of neocortical pyramidal neurons in vivo. *J Neurophysiol*, 81(4):1531–1547.

Destexhe, A., Rudolph, M., Fellous, J., and Sejnowski, T. (2001). Fluctuating synaptic conductances recreate in vivo-like activity in neocortical neurons. *Neuroscience*, 107(1):13–24.

Douglas, R. J., Koch, C., Mahowald, M., Martin, K. A., and Suarez, H. H. (1995). Recurrent excitation in neocortical circuits. *Science*, 269(5226):981–985.

Douglas, R. J. and Martin, K. A. C. (2004). Neuronal circuits of the neocortex. *Annu Rev Neurosci*, 27:419–451.

Douglas, R. J. and Martin, K. A. C. (2007). Mapping the matrix: the ways of neocortex. *Neuron*, 56(2):226–238.

Dragoi, V., Rivadulla, C., and Sur, M. (2001). Foci of orientation plasticity in visual cortex. *Nature*, 411(6833):80–86.

Dragoi, V., Sharma, J., Miller, E. K., and Sur, M. (2002). Dynamics of neuronal sensitivity in visual cortex and local feature discrimination. *Nat Neurosci*, 5(9):883–891.

Dragoi, V., Sharma, J., and Sur, M. (2000). Adaptation-induced plasticity of orientation tuning in adult visual cortex. *Neuron*, 28(1):287–298.

Dragoi, V. and Sur, M. (2006). Image structure at the center of gaze during free viewing. *J Cogn Neurosci*, 18(5):737–748.

Eagleman, D. M., Jacobson, J. E., and Sejnowski, T. J. (2004). Perceived luminance depends on temporal context. *Nature*, 428(6985):854–856.

Fairhall, A. L., Lewen, G. D., Bialek, W., and de Ruyter Van Steveninck, R. R. (2001). Efficiency and ambiguity in an adaptive neural code. *Nature*, 412:787–792.

Felleman, D. J. and Essen, D. C. V. (1991). Distributed hierarchical processing in the primate cerebral cortex. *Cereb Cortex*, 1(1):1–47.

Felsen, G., Shen, Y., Yao, H., Spor, G., Li, C., and Dan, Y. (2002). Dynamic modification of cortical orientation tuning mediated by recurrent connections. *Neuron*, 36(5):945–954.

Felsen, G., Touryan, J., and Dan, Y. (2005). Contextual modulation of orientation tuning contributes to efficient processing of natural stimuli. *Network*, 16(2-3):139–149.

Ferster, D., Chung, S., and Wheat, H. (1996). Orientation selectivity of thalamic input to simple cells of cat visual cortex. *Nature*, 380(6571):249–252.

Ferster, D. and Miller, K. D. (2000). Neural mechanisms of orientation selectivity in the visual cortex. *Annu Rev Neurosci*, 23:441–471.

Finn, I. M., Priebe, N. J., and Ferster, D. (2007). The emergence of contrast-invariant orientation tuning in simple cells of cat visual cortex. *Neuron*, 54(1):137–152.

Galarreta, M. and Hestrin, S. (1998). Frequency-dependent synaptic depression and the balance of excitation and inhibition in the neocortex. *Nat Neurosci*, 1(7):587–594.

Gelman, A., Carlin, J., Stern, H., and Rubin, D. (1995). *Bayesian Data Analysis*. Chapman & Hall, 1st edition.

Gibson, J. J. and Radner, M. (1937). Adaptation, after-effect, and contrast in the perception of tilted lines. I. Quantitative studies. *Journal of Experimental Psychology*, 20:453–467.

Gilbert, C. D., Sigman, M., and Crist, R. E. (2001). The neural basis of perceptual learning. *Neuron*, 31(5):681–697.

Gillespie, D. C., Lampl, I., Anderson, J. S., and Ferster, D. (2001). Dynamics of the orientation-tuned membrane potential response in cat primary visual cortex. *Nat Neurosci*, 4(10):1014–1019.

Girardin, C. C. and Martin, K. A. C. (2009). Cooling in cat visual cortex: stability of orientation selectivity despite changes in responsiveness and spike width. *Neuroscience*.

Givois, V. and Pollack, G. S. (2000). Sensory habituation of auditory receptor neurons: implications for sound localization. *J Exp Biol*, 203:2529–2537.

Goldman, M. S., Golowasch, J., Marder, E., and Abbott, L. F. (2001). Global structure, robustness, and modulation of neuronal models. *J Neurosci*, 21(14):5229–5238.

Gutnisky, D. A. and Dragoi, V. (2008). Adaptive coding of visual information in neural populations. *Nature*, 452(7184):220–224.

Haider, B., Duque, A., Hasenstaub, A. R., and McCormick, D. A. (2006). Neocortical network activity in vivo is generated through a dynamic balance of excitation and inhibition. *J Neurosci*, 26(17):4535–4545.

Hammond, P., Mouat, G. S., and Smith, A. T. (1985). Motion after-effects in cat striate cortex elicited by moving gratings. *Exp Brain Res*, 60(2):411–416.

Hansel, D. and Sompolinsky, H. (1996). Chaos and synchrony in a model of a hypercolumn in visual cortex. *J Comput Neurosci*, 3(1):7–34.

Hübener, M., Shoham, D., Grinvald, A., and Bonhoeffer, T. (1997). Spatial relationships among three columnar systems in cat area 17. *J Neurosci*, 17(23):9270–9284.

Hennig, R. M. (1988). Ascending auditory interneurons in the cricket teleogryllus commodus (Walker): comparative physiology and direct connections with afferents. *J Comp Physiol [A]*, 163(1):135–143.

Hirsch, J. A. and Martinez, L. M. (2006). Laminar processing in the visual cortical column. *Curr Opin Neurobiol*, 16(4):377–384.

Hirsch, J. A., Martinez, L. M., Pillai, C., Alonso, J. M., Wang, Q., and Sommer, F. T. (2003). Functionally distinct inhibitory neurons at the first stage of visual cortical processing. *Nat Neurosci*, 6(12):1300–1308.

Horseman, G. and Huber, F. (1994a). Sound localisation in crickets. I. Contralateral inhibition of an ascending auditory interneuron (AN1) in the cricket gryllus bimaculatus. *J Comp Physiol A*, 175:389–398.

Horseman, G. and Huber, F. (1994b). Sound localisation in crickets. II. Modelling the role of a simple neural network in the prothoracic ganglion. *J Comp Physiol A*, 175:399–413.

Hubel, D. H. and Wiesel, T. N. (1962). Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. *J Physiol*, 160:106–154.

Hubel, D. H. and Wiesel, T. N. (1968). Receptive fields and functional architecture of monkey striate cortex. *J Physiol*, 195(1):215–243.

Jazayeri, M. and Movshon, J. A. (2006). Optimal representation of sensory information by neural populations. *Nat Neurosci*, 9(5):690–696.

Jin, D. Z., Dragoi, V., Sur, M., and Seung, H. S. (2005). Tilt aftereffect and adaptation-induced changes in orientation tuning in visual cortex. *J Neurophysiol*, 94(6):4038–4050.

Joris, P. X., Schreiner, C. E., and Rees, A. (2004). Neural processing of amplitude-modulated sounds. *Physiol Rev*, 84(2):541–577.

Kandel, E., Schwartz, J., and Jessell, T. (2000). *Principles of Neural Science*. McGraw-Hill Medical, 4th edition.

Kang, K., Shapley, R. M., and Sompolinsky, H. (2004). Information tuning of populations of neurons in primary visual cortex. *J Neurosci*, 24(15):3726–3735.

Kang, K., Shelley, M., and Sompolinsky, H. (2003). Mexican hats and pinwheels in visual cortex. *Proc Natl Acad Sci U S A*, 100(5):2848–2853.

Kohn, A. (2007). Visual adaptation: physiology, mechanisms, and functional benefits. *J Neurophysiol*, 97(5):3155–3164.

Kohn, A. and Movshon, J. A. (2003). Neuronal adaptation to visual motion in area MT of the macaque. *Neuron*, 39(4):681–691.

Kohn, A. and Movshon, J. A. (2004). Adaptation changes the direction tuning of macaque MT neurons. *Nat Neurosci*, 7(7):764–772.

Kvale, M. N. and Schreiner, C. E. (2004). Short-term adaptation of auditory receptive fields to dynamic stimuli. *J Neurophysiol*, 91(2):604–612.

Las, L., Stern, E. A., and Nelken, I. (2005). Representation of tone in fluctuating maskers in the ascending auditory system. *J Neurosci*, 25(6):1503–1513.

Laughlin, S. (1981). A simple coding procedure enhances a neuron's information capacity. *Z Naturforsch [C]*, 36(9-10):910–912.

LeBlanc, J., Ward, M. O., and Wittels, N. (1990). Exploring n-dimensional databases. In *Proc. Visualization '90*, pages 230–239.

Lennie, P. (2003). Receptive fields. *Curr Biol*, 13(6):R216–R219.

Lennie, P. and Movshon, J. A. (2005). Coding of color and form in the geniculostriate visual pathway (invited review). *J Opt Soc Am A Opt Image Sci Vis*, 22(10):2013–2033.

Levinson, E. and Sekuler, R. (1976). Adaptation alters perceived direction of motion. *Vision Res*, 16(7):779–781.

Linsker, R. (1988). Self-organization in a perceptual network. *Computer*, 21(3):105–117.

Lund, J. S., Angelucci, A., and Bressloff, P. C. (2003). Anatomical substrates for functional columns in macaque monkey primary visual cortex. *Cereb Cortex*, 13(1):15–24.

Machens, C. K., Gollisch, T., Kolesnikova, O., and Herz, A. V. M. (2005). Testing the efficiency of sensory coding with optimal stimulus ensembles. *Neuron*, 47(3):447–456.

MacKay, D. J. C. (2003). *Information Theory, Inference and Learning Algorithms*. Cambridge University Press.

Maffei, L., Fiorentini, A., and Bisti, S. (1973). Neural correlate of perceptual adaptation to gratings. *Science*, 182(116):1036–1038.

Maldonado, P. E., Gödecke, I., Gray, C. M., and Bonhoeffer, T. (1997). Orientation selectivity in pinwheel centers in cat striate cortex. *Science*, 276(5318):1551–1555.

Maravall, M., Petersen, R. S., Fairhall, A. L., Arabzadeh, E., and Diamond, M. E. (2007). Shifts in coding properties and maintenance of information transmission during adaptation in barrel cortex. *PLoS Biol*, 5(2):e19.

Marder, E., Tobin, A.-E., and Grashow, R. (2007). How tightly tuned are network parameters? insight from computational and experimental studies in small rhythmic motor networks. *Prog Brain Res*, 165:193–200.

Mariño, J., Schummers, J., Lyon, D. C., Schwabe, L., Beck, O., Wiesing, P., Obermayer, K., and Sur, M. (2005). Invariant computations in local cortical networks with balanced excitation and inhibition. *Nat Neurosci*, 8(2):194–201.

Marsat, G. and Pollack, G. S. (2006). A behavioral role for feature detection by sensory bursts. *J Neurosci*, 26(41):10542–10547.

Martinez, L. M., Alonso, J. M., Reid, R. C., and Hirsch, J. A. (2002). Laminar processing of stimulus orientation in cat visual cortex. *J Physiol*, 540:321–333.

Mascagni, M. V. and Sherman, A. S. (1998). Numerical methods for neuronal modeling. In Koch, C. and Segev, I., editors, *Methods in neuronal modeling : from ions to networks*, pages 569–606. MIT Press, Cambridge, Mass., 2nd edition.

Mazer, J. A., Vinje, W. E., McDermott, J., Schiller, P. H., and Gallant, J. L. (2002). Spatial frequency and orientation tuning dynamics in area V1. *Proc Natl Acad Sci U S A*, 99(3):1645–1650.

McAdams, C. J. and Maunsell, J. H. (1999). Effects of attention on orientation-tuning functions of single neurons in macaque cortical area V4. *J Neurosci*, 19(1):431–441.

McLaughlin, D., Shapley, R., Shelley, M., and Wielaard, D. J. (2000). A neuronal network model of macaque primary visual cortex (V1): orientation selectivity and dynamics in the input layer 4Calpha. *Proc Natl Acad Sci U S A*, 97(14):8087–8092.

Meister, M. and Berry, M. J. (1999). The neural code of the retina. *Neuron*, 22(3):435–450.

Mitchell, D. E. and Muir, D. W. (1976). Does the tilt after-effect occur in the oblique meridian? *Vision Res*, 16(6):609–613.

Müller, J. R., Metha, A. B., Krauskopf, J., and Lennie, P. (1999). Rapid adaptation in visual cortex to the structure of images. *Science*, 285(5432):1405–1408.

Monier, C., Chavane, F., Baudot, P., Graham, L. J., and Frégnac, Y. (2003). Orientation and direction selectivity of synaptic inputs in visual cortical neurons: a diversity of combinations produces spike tuning. *Neuron*, 37(4):663–680.

Movshon, J. A., Thompson, I. D., and Tolhurst, D. J. (1978a). Receptive field organization of complex cells in the cat's striate cortex. *J Physiol*, 283:79–99.

Movshon, J. A., Thompson, I. D., and Tolhurst, D. J. (1978b). Spatial summation in the receptive fields of simple cells in the cat's striate cortex. *J Physiol*, 283:53–77.

Myme, C. I. O., Sugino, K., Turrigiano, G. G., and Nelson, S. B. (2003). The NMDA-to-AMPA ratio at synapses onto layer 2/3 pyramidal neurons is conserved across prefrontal and visual cortices. *J Neurophysiol*, 90(2):771–779.

Nadal, J.-P. and Parga, N. (1994). Nonlinear neurons in the low-noise limit: a factorial code maximizes information transmission. *Network: Comput. Neural Syst.*, 5:565–581.

Nagel, K. I. and Doupe, A. J. (2006). Temporal processing and adaptation in the songbird auditory forebrain. *Neuron*, 51(6):845–859.

Nauhaus, I., Benucci, A., Carandini, M., and Ringach, D. L. (2008). Neuronal selectivity and local map structure in visual cortex. *Neuron*, 57(5):673–679.

Nelken, I. (2004). Processing of complex stimuli and natural scenes in the auditory cortex. *Curr Opin Neurobiol*, 14(4):474–480.

Nelken, I., Rotman, Y., and Yosef, O. B. (1999). Responses of auditory-cortex neurons to structural features of natural sounds. *Nature*, 397:154–157.

Nelson, S., Toth, L., Sheth, B., and Sur, M. (1994). Orientation selectivity of cortical neurons during intracellular blockade of inhibition. *Science*, 265(5173):774–777.

Nishimoto, S., Arai, M., and Ohzawa, I. (2005). Accuracy of subspace mapping of spatiotemporal frequency domain visual receptive fields. *J Neurophysiol*, 93(6):3524–3536.

Nolen, T. G. and Hoy, R. R. (1984). Initiation of behavior by single neurons: the role of behavioral context. *Science*, 226:992–994.

Nowak, L. G., Sanchez-Vives, M. V., and McCormick, D. A. (2008). Lack of orientation and direction selectivity in a subgroup of fast-spiking inhibitory interneurons: cellular and synaptic mechanisms and comparison with other electrophysiological cell types. *Cereb Cortex*, 18(5):1058–1078.

Obermayer, K. and Blasdel, G. G. (1993). Geometry of orientation and ocular dominance columns in monkey striate cortex. *J Neurosci*, 13(10):4114–4129.

Ohki, K., Chung, S., Kara, P., Hübener, M., Bonhoeffer, T., and Reid, R. C. (2006). Highly ordered arrangement of single neurons in orientation pinwheels. *Nature*, 442(7105):925–928.

Ohzawa, I., Sclar, G., and Freeman, R. D. (1985). Contrast gain control in the cat's visual system. *J Neurophysiol*, 54(3):651–667.

Pollack, G. (1988). Selective attention in an insect auditory neuron. *J Neurosci*, 8(7):2635–2639.

Pollack, G. S. and Martins, R. (2007). Flight and hearing: ultrasound sensitivity differs between flight-capable and flight-incapable morphs of a wing-dimorphic cricket species. *J Exp Biol*, 210:3160–3164.

Pouget, A., Dayan, P., and Zemel, R. (2000). Information processing with population codes. *Nat Rev Neurosci*, 1(2):125–132.

Pouget, A., Dayan, P., and Zemel, R. S. (2003). Inference and computation with population codes. *Annu Rev Neurosci*, 26:381–410.

Priebe, N. J. and Ferster, D. (2008). Inhibition, spike threshold, and stimulus selectivity in primary visual cortex. *Neuron*, 57(4):482–497.

Prinz, A. A., Billimoria, C. P., and Marder, E. (2003). Alternative to hand-tuning conductance-based models: construction and analysis of databases of model neurons. *J Neurophysiol*, 90(6):3998–4015.

Prinz, A. A., Bucher, D., and Marder, E. (2004). Similar network activity from disparate circuit parameters. *Nat Neurosci*, 7(12):1345–1352.

Regan, D. and Beverley, K. I. (1985). Postadaptation orientation discrimination. *J Opt Soc Am A*, 2(2):147–155.

Reid, R. C. and Alonso, J. M. (1995). Specificity of monosynaptic connections from thalamus to visual cortex. *Nature*, 378(6554):281–284.

Reid, R. C. and Alonso, J. M. (1996). The processing and encoding of information in the visual cortex. *Curr Opin Neurobiol*, 6(4):475–480.

Reig, R., Gallego, R., Nowak, L. G., and Sanchez-Vives, M. V. (2006). Impact of cortical network activity on short-term synaptic depression. *Cereb Cortex*, 16(5):688–695.

Rhodes, G., Jeffery, L., Watson, T. L., Clifford, C. W. G., and Nakayama, K. (2003). Fitting the mind to the world: face adaptation and attractiveness aftereffects. *Psychol Sci*, 14(6):558–566.

Ringach, D. L., Hawken, M. J., and Shapley, R. (1997). Dynamics of orientation tuning in macaque primary visual cortex. *Nature*, 387(6630):281–284.

Ringach, D. L., Hawken, M. J., and Shapley, R. (2003). Dynamics of orientation tuning in macaque V1: the role of global and tuned suppression. *J Neurophysiol*, 90(1):342–352.

Ringach, D. L. and Malone, B. J. (2007). The operating point of the cortex: neurons as large deviation detectors. *J Neurosci*, 27(29):7673–7683.

Ringach, D. L., Shapley, R. M., and Hawken, M. J. (2002). Orientation selectivity in macaque V1: diversity and laminar dependence. *J Neurosci*, 22(13):5639–5651.

Römer, H. and Krusch, M. (2000). A gain-control mechanism for processing of chorus sounds in the afferent auditory pathway of the bushcricket Tettigonia viridissima (Orthoptera; Tettigoniidae). *J Comp Physiol [A]*, 186(2):181–191.

Roudi, Y., Nirenberg, S., and Latham, P. E. (2009). Pairwise maximum entropy models for studying large biological systems: when they can work and when they can't. *PLoS Comput Biol*, 5(5):e1000380.

Rust, N. C. and Movshon, J. A. (2005). In praise of artifice. *Nat Neurosci*, 8(12):1647–1650.

Samson, A.-H. and Pollack, G. S. (2002). Encoding of sound localization cues by an identified auditory interneuron: effects of stimulus temporal pattern. *J Neurophysiol*, 88(5):2322–2328.

Sanchez-Vives, M. V., Nowak, L. G., and McCormick, D. A. (2000a). Cellular mechanisms of long-lasting adaptation in visual cortical neurons in vitro. *J Neurosci*, 20(11):4286–4299.

Sanchez-Vives, M. V., Nowak, L. G., and McCormick, D. A. (2000b). Membrane mechanisms underlying contrast adaptation in cat area 17 in vivo. *J Neurosci*, 20(11):4267–4285.

Schoups, A., Vogels, R., Qian, N., and Orban, G. (2001). Practising orientation identification improves orientation coding in V1 neurons. *Nature*, 412(6846):549–553.

Schummers, J., Cronin, B., Wimmer, K., Stimberg, M., Martin, R., Obermayer, K., Koerding, K., and Sur, M. (2007). Dynamics of orientation tuning in cat V1 neurons depend on location within layers and orientation maps. *Frontiers in Neuroscience*, 1:145–159.

Schummers, J., Mariño, J., and Sur, M. (2002). Synaptic integration by V1 neurons depends on location within the orientation map. *Neuron*, 36(5):969–978.

Schummers, J., Mariño, J., and Sur, M. (2004). Local networks in visual cortex and their influence on neuronal responses and dynamics. *J Physiol Paris*, 98(4-6):429–441.

Schwabe, L. and Obermayer, K. (2003). Modeling the adaptive visual system: a survey of principled approaches. *Neural Netw*, 16(9):1353–1371.

Schwabe, L. and Obermayer, K. (2005). Adaptivity of tuning functions in a generic recurrent network model of a cortical hypercolumn. *J Neurosci*, 25(13):3323–3332.

Schwartz, O., Hsu, A., and Dayan, P. (2007). Space and time in visual context. *Nat Rev Neurosci*, 8(7):522–535.

Seriès, P., Lorenceau, J., and Frégnac, Y. (2003). The "silent" surround of V1 receptive fields: theory and experiments. *J Physiol Paris*, 97(4-6):453–474.

Shapiro, L. G. and Stockman, G. C. (2001). *Computer Vision*. Prentice Hall.

Shapley, R. and Enroth-Cugell, C. (1984). Visual adaptation and retinal gain controls. *Progress in Retinal Research*, 3:263–346.

Sharma, J., Dragoi, V., Tenenbaum, J. B., Miller, E. K., and Sur, M. (2003). V1 neurons signal acquisition of an internal representation of stimulus location. *Science*, 300(5626):1758–1763.

Sharon, D. and Grinvald, A. (2002). Dynamics and constancy in cortical spatiotemporal patterns of orientation processing. *Science*, 295(5554):512–515.

Sharpee, T. O., Sugihara, H., Kurgansky, A. V., Rebrik, S. P., Stryker, M. P., and Miller, K. D. (2006). Adaptive filtering enhances information transmission in visual cortex. *Nature*, 439:936–942.

Sherman, S. M. and Guillery, R. W. (2000). *Exploring the Thalamus*. Academic Press, 1st edition.

Shevelev, I. A., Sharaev, G. A., Lazareva, N. A., Novikova, R. V., and Tikhomirov, A. S. (1993). Dynamics of orientation tuning in the cat striate cortex neurons. *Neuroscience*, 56(4):865–876.

Shriki, O., Hansel, D., and Sompolinsky, H. (2003). Rate models for conductance-based cortical neuronal networks. *Neural Comput*, 15(8):1809–1841.

Sobel, E. C. and Tank, D. W. (1994). In vivo Ca^{2+} dynamics in a cricket auditory neuron: An example of chemical computation. *Science*, 263:863–826.

Solomon, S. G., Peirce, J. W., Dhruv, N. T., and Lennie, P. (2004). Profound contrast adaptation early in the visual pathway. *Neuron*, 42(1):155–162.

Somers, D., Nelson, S., and Sur, M. (1995). An emergent model of orientation selectivity in cat visual cortical simple cells. *J Neurosci*, 15(8):5448–5465.

Sompolinsky, H. and Shapley, R. (1997). New perspectives on the mechanisms for orientation selectivity. *Curr Opin Neurobiol*, 7(4):514–522.

Spitzer, H. and Hochstein, S. (1985). A complex-cell receptive-field model. *J Neurophysiol*, 53(5):1266–1286.

Stimberg, M., Wimmer, K., Martin, R., Schwabe, L., Mariño, J., Schummers, J., Lyon, D. C., Sur, M., and Obermayer, K. (2009). The operating regime of local computations in primary visual cortex. *Cereb Cortex*, 19(9):2166–2180.

Stocker, A. and Simoncelli, E. (2006). Sensory adaptation within a bayesian framework for perception. In Weiss, Y., Schölkopf, B., and Platt, J., editors, *Advances in Neural Information Processing Systems 18*, pages 1289–1296, Cambridge, MA. MIT Press.

Supèr, H., Spekreijse, H., and Lamme, V. A. (2001). A neural correlate of working memory in the monkey primary visual cortex. *Science*, 293(5527):120–124.

Swindale, N. V. (1998). Orientation tuning curves: empirical description and estimation of parameters. *Biol Cybern*, 78(1):45–56.

Tadmor, Y. and Tolhurst, D. J. (2000). Calculating the contrasts that retinal ganglion cells and lgn neurones encounter in natural scenes. *Vision Res*, 40(22):3145–3157.

Taylor, A. L., Goaillard, J.-M., and Marder, E. (2009). How multiple conductances determine electrophysiological properties in a multicompartment model. *J Neurosci*, 29(17):5573–5586.

Taylor, A. L., Hickey, T. J., Prinz, A. A., and Marder, E. (2006). Structure and visualization of high-dimensional conductance spaces. *J Neurophysiol*, 96(2):891–905.

Teich, A. F. and Qian, N. (2003). Learning and adaptation in a recurrent model of V1 orientation selectivity. *J Neurophysiol*, 89(4):2086–2100.

Teich, A. F. and Qian, N. (2006). Comparison among some models of orientation selectivity. *J Neurophysiol*, 96(1):404–419.

Thomson, A. M. and Deuchars, J. (1997). Synaptic interactions in neocortical local circuits: dual intracellular recordings in vitro. *Cereb Cortex*, 7(6):510–522.

Thomson, A. M. and Lamy, C. (2007). Functional maps of neocortical local circuitry. *Front Neurosci*, 1(1):19–42.

Troyer, T. W., Krukowski, A. E., Priebe, N. J., and Miller, K. D. (1998). Contrast-invariant orientation tuning in cat visual cortex: thalamocortical input tuning and correlation-based intracortical connectivity. *J Neurosci*, 18(15):5908–5927.

Tsodyks, M., Pawelzik, K., and Markram, H. (1998). Neural networks with dynamic synapses. *Neural Comput*, 10(4):821–835.

Tsodyks, M. V. and Markram, H. (1997). The neural code between neocortical pyramidal neurons depends on neurotransmitter release probability. *Proc Natl Acad Sci U S A*, 94(2):719–723.

Ulanovsky, N., Las, L., and Nelken, I. (2003). Processing of low-probability sounds by cortical neurons. *Nat Neurosci*, 6(4):391–398.

Varela, J. A., Sen, K., Gibson, J., Fost, J., Abbott, L. F., and Nelson, S. B. (1997). A quantitative description of short-term plasticity at excitatory synapses in layer 2/3 of rat primary visual cortex. *J Neurosci*, 17(20):7926–7940.

Varela, J. A., Song, S., Turrigiano, G. G., and Nelson, S. B. (1999). Differential depression at excitatory and inhibitory synapses in visual cortex. *J Neurosci*, 19(11):4293–4304.

Vogels, R. (1990). Population coding of stimulus orientation by striate cortical cells. *Biol Cybern*, 64(1):25–31.

Volgushev, M., Vidyasagar, T. R., and Pei, X. (1995). Dynamics of the orientation tuning of postsynaptic potentials in the cat visual cortex. *Vis Neurosci*, 12(4):621–628.

Wainwright, M. J. (1999). Visual adaptation as optimal information transmission. *Vision Res*, 39(23):3960–3974.

Webster, M. A., Georgeson, M. A., and Webster, S. M. (2002). Neural adjustments to image blur. *Nat Neurosci*, 5(9):839–840.

Webster, M. A., Kaping, D., Mizokami, Y., and Duhamel, P. (2004). Adaptation to natural facial categories. *Nature*, 428(6982):557–561.

Wielaard, D. J., Shelley, M., McLaughlin, D., and Shapley, R. (2001). How simple cells are made in a nonlinear network model of the visual cortex. *J Neurosci*, 21(14):5203–5211.

Wimmer, K., Hildebrandt, K. J., Hennig, R. M., and Obermayer, K. (2008). Adaptation and selective information transmission in the cricket auditory neuron an2. *PLoS Comput Biol*, 4(9):e1000182.

Wimmer, K., Stimberg, M., Martin, R., Schwabe, L., Mariño, J., Schummers, J., Lyon, D. C., Sur, M., and Obermayer., K. (2009). Dependence of orientation tuning on recurrent excitation and inhibition in a network model of V1. In Koller, D., Schuurmans, D., Bengio, Y., and Bottou, L., editors, *Advances in Neural Information Processing Systems 21*, pages 1769–1776, Cambridge, MA. MIT Press.

Wohlers, D. W. and Huber, F. (1982). Processing of sound signals by six types of neurons in the prothoracic ganglion of the cricket, gryllus campestris l. *J Comp Physiol A*, 146:161–173.

Wolfe, J. and Palmer, L. A. (1998). Temporal diversity in the lateral geniculate nucleus of cat. *Vis Neurosci*, 15(4):653–675.

Wörgötter, F. and Koch, C. (1991). A detailed model of the primary visual pathway in the cat: comparison of afferent excitatory and intracortical inhibitory connection schemes for orientation selectivity. *J Neurosci*, 11(7):1959–1979.

Yoshimura, Y., Sato, H., Imamura, K., and Watanabe, Y. (2000). Properties of horizontal and vertical inputs to pyramidal cells in the superficial layers of the cat visual cortex. *J Neurosci*, 20(5):1931–1940.

Yousef, T., Tóth, E., Rausch, M., Eysel, U. T., and Kisvárday, Z. F. (2001). Topography of orientation centre connections in the primary visual cortex of the cat. *Neuroreport*, 12(8):1693–1699.

Yu, H., Farley, B. J., Jin, D. Z., and Sur, M. (2005). The coordinated mapping of visual space and response features in visual cortex. *Neuron*, 47(2):267–80.

Zucker, R. S. and Regehr, W. G. (2002). Short-term synaptic plasticity. *Annu Rev Physiol*, 64:355–405.