

Dissertation
in partial fulfilment of the degree
Doctor rerum naturalium (Dr. rer. nat.)

Functional Role of Critical Dynamics in Flexible Visual Information Processing

submitted by
Nergis Tömen, M.Sc.
on
14.06.2019

Universität Bremen
Fachbereich für Physik und Elektrotechnik
Institut für Theoretische Physik

Gutachter:
Dr. Udo A. Ernst
Prof. Dr. Andreas Kreiter

Dissertation *Functional Role of Critical Dynamics in Flexible Visual Information Processing*, submitted by Nergis Tömen (born in Istanbul, Turkey - 26.05.1987) on 14.06.2019.

But although, as a matter of history, statistical mechanics owes its origins to investigations in thermodynamics, it seems eminently worthy of an independent development, both on account of the elegance and simplicity of its principles, and because it yields new results and places old truths in a new light in departments quite outside of thermodynamics.

–Josiah Willard Gibbs

The large brain, like large government, may not be able to do simple things in a simple way.

–Donald O. Hebb

Abstract

Recent experimental and theoretical work has established the hypothesis that cortical neurons operate close to a critical state which signifies a phase transition from chaotic to ordered dynamics. Critical dynamics are suggested to optimize several aspects of neuronal information processing. However, although signatures of critical dynamics have been demonstrated in recordings of spontaneously active cortical neurons, little is known about how these dynamics are affected by task-dependent changes in neuronal activity when the cortex is engaged in stimulus processing. In fact, some *in vivo* investigations of the awake and active cortex report either an absence of signatures of criticality or relatively weak ones. In addition, the functional role of criticality in optimizing computation is often reported in abstract theoretical studies, adopting minimalistic models with homogeneous topology and slowly-driven networks. Consequently, there is a lack of concrete links between information theoretical benefits of the critical state and neuronal networks performing a behaviourally relevant task.

In this thesis we explore such concrete links by focusing on the visual system, which needs to meet major computational challenges on a daily basis. Among others, the visual system is responsible for the rapid integration of relevant information from a large number of single channels, and in a flexible manner depending on the behavioral and environmental contexts. We postulate that critical neuronal dynamics in the form of cascades of activity spanning large populations of neurons may support such quick and complex computations.

Specifically, we consider two notable examples of well-known phenomena in visual information processing: First the enhancement of object discriminability under selective attention, and second, a feature integration and figure-ground segregation scenario. In the first example, we model the top-down modulation of the activity of visuocortical neurons in order to selectively improve the processing of an attended region in a visual scene. In the second example, we model how neuronal activity may be modulated in a bottom-up fashion by the properties of the visual stimulus itself, which makes it possible to perceive different shapes and objects.

We find in both scenarios that the task performance may be improved by employing critical networks. In addition, we suggest that the specific task- or stimulus-dependent modulations of information processing may be optimally supported by the tuning of relevant local neuronal networks towards or away from the critical point.

Thus, the relevance of this dissertation is summarized by the following points: We formally extend the existing models of criticality to inhomogeneous systems subject to a strong external drive. We present concrete functional benefits for networks operating near the critical point in well-known experimental paradigms. Importantly, we find emergent critical dynamics only in the parts of the network which are processing the behaviourally relevant information. We suggest that the implied locality of critical dynamics in space and time may help explain why some studies report no signatures of criticality in the active cortex.



Contents

I	Introduction	
1	Motivation and Outline	13
1.1	Critical phenomena: From equilibrium models of physical systems to neuronal dynamics	13
1.2	Signatures of criticality in the brain and links to function	15
1.3	Outline	17
2	Criticality in Dynamical Systems	19
2.1	Continuous phase transitions and the Ising model	19
2.1.1	Phase transitions	19
2.1.2	The Ising model	21
2.2	Self-organized criticality	25
3	Criticality in the Brain	29
3.1	Theoretical work	29
3.2	Experimental evidence for criticality	32
4	The Visual System	35
4.1	Basics of the visual system	35
4.2	Feature integration	41
4.3	Attention and gamma oscillations	42

II

Attention and Tuning to Criticality

5	Attentional Modulations of Object Representation . . .	47
5.1	Introduction	47
5.1.1	Motivation and previous experiments	48
5.2	Results	50
5.2.1	Reproduction of key experimental findings	50
5.2.2	Enhancement of stimulus discriminability	53
5.2.3	Characterization of dynamical network states	55
5.2.4	Criticality of dynamics	56
5.2.5	Information Entropy	61
5.3	Methods	63
5.3.1	Network Model	63
5.3.2	Analysis of Network Dynamics	65
5.3.3	Neuronal Avalanches	66
5.3.4	Computation of Information Entropy	68
5.4	Discussion	69
5.4.1	Robustness of results	70
5.4.2	Physiological plausibility	70
5.4.3	Functional role of local and downstream structure	72

III

Feature Integration in Critical Networks

6	Figure-ground Segregation in a Critical Network . . .	75
6.1	Introduction	75
6.2	Methods	76
6.2.1	The EHE network	76
6.2.2	Leaky network model	81
6.3	Results	82
6.3.1	The EHE network	82
6.3.2	Leaky network	88
6.4	Discussion	93
6.4.1	Robustness of results	94
6.4.2	Physiological Plausibility	95
7	Contour Integration in a Realistic LIF Model	97
7.1	Introduction	97

7.2	Results	100
7.2.1	Generation of stimuli	100
7.2.2	Network model and single neuron dynamics	101
7.2.3	Network dynamics	105
7.2.4	Task performance	107
7.3	Methods	112
7.3.1	Generation of stimuli	112
7.3.2	Network structure and feedforward input	112
7.3.3	Recurrent connectivity	114
7.3.4	Neuron model	116
7.3.5	Avalanche statistics	116
7.3.6	Read-out and task performance	117
7.4	Discussion	118
7.4.1	Robustness and interpretation of results	118
7.4.2	Physiological Plausibility	120

IV

Conclusion and Discussion

8	Summary and Conclusion	125
9	Discussion and Outlook	129
9.1	Generality and Realism of the Models	129
9.2	Critical or Near-Critical Dynamics?	131
9.3	Criticality, Structure, and Function	134
9.4	Outlook	136
	Bibliography	138



Introduction

1	Motivation and Outline	13
1.1	Critical phenomena: From equilibrium models of physical systems to neuronal dynamics	
1.2	Signatures of criticality in the brain and links to function	
1.3	Outline	
2	Criticality in Dynamical Systems . .	19
2.1	Continuous phase transitions and the Ising model	
2.2	Self-organized criticality	
3	Criticality in the Brain	29
3.1	Theoretical work	
3.2	Experimental evidence for criticality	
4	The Visual System	35
4.1	Basics of the visual system	
4.2	Feature integration	
4.3	Attention and gamma oscillations	



1. Motivation and Outline

1.1 Critical phenomena: From equilibrium models of physical systems to neuronal dynamics

Phase transitions, such as the freezing of a liquid or the condensation of a gas, are common in nature and represent a dramatic change in observable physical properties of a system. A subset of phase transitions known as continuous, or second-order, phase transitions have historically been an interesting subject of study due to their exciting mathematical properties. For example the Ising model of ferromagnetism, which is one of the oldest examples of models exhibiting continuous phase transitions [Ons44], has a correlation length which diverges at a critical value of the temperature. This means that, at the critical temperature, nonvanishing correlations between nodes separated by arbitrarily large distances emerge through only local interactions. Furthermore, we know that as a system approaches a continuous phase transition, many of its observables will assume a power-law distribution, with so-called critical exponents [Sta71]. Such distributions are said to be scale-free and indicate the self-similarity of the underlying system. Strikingly, the critical exponents are universal for a given set of models, that is, they do not change when certain features of the model are completely altered. Thus, critical exponents characterize a universality class of models which encompasses a large variety of phase transitions in different systems.

Due to their distinctive properties, systems exhibiting critical phenomena have been investigated in detail throughout the 20th century (for example in the scope of the Ising model or percolation theory). However, such systems undergo a phase transition as a function of a control parameter, such as the temperature in the Ising model, which are often continuously valued or have distributions supported on infinite intervals. Hence, the critical state of the system corresponds to the control

parameter assuming a specific value out of infinitely many possibilities. Therefore, in an uncontrolled environment, one would expect to find systems displaying critical phenomena only infrequently [Gra02].

Nevertheless, power-law behaviour, such as $1/f$ noise is abundant in nature. Based on this observation, a potential link between the ubiquity of scale-free behaviour and critical phenomena was suggested by Bak, Tang and Wiesenfeld in 1987 [BTW87]. Their work established the so-called Abelian sandpile model, which is a dynamical system of coupled nonlinear elements perpetually tuning itself to a critical point. Specifically, such a system will evolve from any initial condition towards an attractor at a critical point at which scale-free dynamics emerge. Due to the absence of the need to fine-tune a control parameter, this property is called self-organized criticality (SOC), and was proposed as an organizational principle underlying many complex systems. In fact, the proposed framework was so influential, that in the first decade following its inception, more than 2000 publications presented studies of SOC [BG97].

In the original Abelian sandpile model, dynamics consisted of cascades of activity propagating over lattice sites, so-called ‘avalanches’ of sand grains, generated by slow external perturbation. The number of sites which participated in an avalanche, as well as the avalanche lifetimes displayed scale-free distributions. Further analysis revealed that the sandpile SOC models can be described as a critical branching process, where the avalanche size and lifetime distributions have the critical exponents $\tau = 3/2$ and $\alpha = 2$ respectively [ZLS95].

Experimental studies have since then shown that a variety of physical systems such as sandpiles [HSS+90], water droplets [PNB93], superconductors [FWN+95] and earthquakes [BP04] exhibit SOC. At the same time, a diverse range of additional critical models were discovered, such as the forest-fire model describing percolating systems [DS92] and the Olami-Feder-Christensen (OFC) model describing earthquake dynamics [OFC92].

Similarly, scientists have long hypothesized that SOC might also be a feature of biological systems [BS93] and that criticality of dynamics is relevant for performing complex computations [CY89; Lan90]. This prompted a natural extension of criticality research to neural systems. Some earlier theoretical studies of critical phenomena in neuronal networks included a rewiring rule which could generate critical topology in neural networks [BR00] and the Eurich–Herrmann–Ernst (EHE) model—a model of neural networks displaying critical phenomena, based on the OFC model [EHE02]. Notably, the EHE model predicted that critical neuronal systems belong to the same universality class as a critical branching process, with avalanche sizes obeying a power-law with the critical exponent $\tau = 3/2$.

The first experimental evidence of critical dynamics in the brain was presented in 2003 by Beggs and Plenz. Specifically, their study has shown that neuronal avalanche size distributions in spontaneous activity follow a power-law with $\tau = 3/2$ in organotypic cultures as well as in acute slices of rat cortex [BP03]. The observed avalanche size distributions nicely matched the closed-form expressions derived for EHE networks of finite size [EHE02]. The analytical insight provided by the

EHE model is one of the main reasons why it is one of the two models used in the investigations presented in this thesis.

The demonstration of power-law scaling in cortical neurons sparked an interest in critical phenomena in the context of neuronal dynamical systems. Subsequently, scale-free distributions as well as other signatures of criticality were reported in many different experimental studies of cortical dynamics, and in many different scales, some of which will be presented in the next section.

1.2 Signatures of criticality in the brain and links to function

Considered a signature of neural networks poised at criticality, scale-free organization of neuronal activity has been observed in a wide range of data including in LFP and spiking dynamics in acute slices and organotypic cultures of the rat cortex [BP03; FIB+12], dissociated cultures [PMB+08], *in vivo* in awake monkeys [PTL+09], human MEG and EEG [PZH+13; SAC+13] and human intracranial depth recordings [PVW+13]. In addition, recent studies address increasingly specific questions relating to, for example, the rigorousness of statistical analysis [KYP11], subsampling [PMW09] and resolution restraints as well as exponent relations [FIB+12] in experimental investigations of critical dynamics.

What makes the idea of critical neural networks especially compelling is the suggestion that a functional relationship may exist between critical dynamics and optimal information processing [SP13]. It has been suggested that critical dynamics may play a functional role in supporting complex computations [BN04; Lan90], maximizing information diversity [NPL+08] and the number of metastable states [HB05], and improving object representation [TRE14] with favourable dynamic range characteristics [GHM+15; KC06; SYP+09].

The abundance of power-law scaling in the brain, taken together with the theoretical functional benefits of networks operating close to a critical state forms the basis of the so-called "criticality hypothesis" [Beg08; Chi04; FBF+09; Ple14]—the suggestion that critical dynamics is a fundamental emergent property of cortical networks, supporting optimal information processing [BL98; Chi10; CB99a; HG14].

While the enhanced information processing capabilities close to a phase transition are an important argument for this hypothesis, demonstrations of them often rely on rather 'abstract' theoretical measures, without realistic encoding or read-out scenarios in mind. Furthermore, such demonstrations employ relatively homogeneous models without structured functional connectivity or strong external drive. This creates an exciting opportunity to explore more concrete links between the abstract computational scenarios and the biologically realistic cortical dynamics under cognitive load.

Unfortunately, the majority of the experimental studies reporting avalanche dynamics in neuronal networks analyze recordings of spontaneous activity *in vitro* or of the resting state *in vivo*, although recent studies attempt to understand the emergence of criticality in the cortex under sensory stimulation [AGS15; SCP+15]. Nevertheless, due to substantial experimental challenges, no empirical studies exist

to date which explore the criticality of *in vivo* neuronal dynamics in conjunction with a specific behavioral task and under changing task demands. In addition, perfect scale-free organization does not seem to emerge under all conditions, especially in the strongly driven regime. Previous findings indicate that signatures of criticality are progressively disturbed during sustained wakefulness [MOS+13] and may diminish with increasing cognitive load [TV14]. Similarly, avalanche dynamics *in vivo* exhibit differences between wakefulness and deep sleep [PVW+13] and seem to be poised at a slightly subcritical state [PWV+14]. Hence, there is some controversy about whether the observed scale-free behaviour is just an artifact of imprecise statistical analysis, or an epiphenomenon not related to an underlying critical state [DHH+12; SP12; TD10; TD17].

Taken together, the abstract nature of theoretical investigations, the inconsistency of signatures of criticality in experimental studies and the scarcity of investigations in the active cortex make it difficult to draw direct conclusions concerning links between critical dynamics and cortical function. In other words, existing criticality literature lacks concrete examples of networks operating close to a critical point contributing to information processing in cortical networks and the exact functional role of scale-free behaviour remains poorly understood.

In this thesis we are aiming for this gap of concrete examples in the literature. Specifically, we will focus on the visual system, which needs to perform particularly rapid and complex computations continuously. Notably, flexible and dynamic information processing in the visuocortical areas requires fast responses to both varying sensory inputs as well as quick top-down modulations due to changing task requirements. In the upcoming chapters, we are going to present computational models of two essential phenomena of the visual system which are both rigorously explored, in typical experimental paradigms, in studies of visual information processing: First the enhancement of object representation and discriminability under selective visual attention, and second, the modulation of feature integration and figure-ground segregation tasks by stimulus salience. In the first example, we model the top-down modulation of the activity of visuocortical neurons in order to selectively improve the processing of information incoming from an attended region in a visual scene. In the second example, we model how neuronal activity may be modulated in a bottom-up fashion by the properties of the visual stimulus itself, which makes it possible to perceive different shapes and objects in a robust manner.

We postulate that critical neuronal dynamics in the form of avalanches of activity occurring on all spatial scales and regularly propagating over large populations of neurons may support the speed and complexity requirements of the computations performed by the visual system. In particular, we believe that the emergence of nonvanishing correlations may help in the integration of spatially and temporally extended visual information. We explore the functional role of critical dynamics in structured, heterogeneous and stimulus-driven networks in order to make the link to the biological systems more explicit.

Within this framework, we investigate the hypothesis that in the active brain engaged in a cognitive task not all neuronal populations are tuned to a perfect critical

state at all times. Instead, the local activity regime can be actively modulated by properties of the external input or changing task demands, thus promoting rapid and robust processing of dynamically changing information streams. We present a paradigm in which avalanche dynamics of local ‘subnetworks’, heavily involved in processing of a sensory stimulus, may be actively tuned towards the critical point, ultimately contributing to the detection or discrimination of the stimulus by higher visuocortical areas. We are going to argue that regardless of the specific mechanism, local emergence of critical dynamics in the active brain—either in a stimulus driven (bottom-up) manner or through attentional (top-down) modulations—may assist in selectively improving information processing. We hypothesize that functional benefits of critical dynamics might be optimally exploited when the emergence of critical dynamics is local in space and time, restricted to only small, local subnetworks in the brain which are involved in performing a behaviourally relevant task. Imposing such a locality constraint on critical dynamics would also help explain some of the apparent inconsistencies in experimental findings.

This framework brings together the concepts of phase transitions and critical phenomena, oscillations in population activity, synchronization at the spiking level, and suggests a universal functional role for correlated activity in the cortex: flexible processing of information in a dynamically changing world.

1.3 Outline

Since the investigations presented in this dissertation are considerably interdisciplinary and our motivations and hypotheses are inherently dependent on the history of multiple subjects of ongoing research, we will next give a brief summary of the already established concepts and models in an effort to provide a detailed background for this work. In chapters 2 to 4, we will introduce the relevant literature, terminology and concepts relating to critical phenomena in physical systems, critical phenomena in neuronal systems and information processing in the visual system, respectively.

In chapter 5, we will present the findings of our model of selective attention and demonstrate how tuning appropriate ensembles of neurons to a critical point may improve object representation and discrimination. In chapter 6, we will report the results obtained in a simple, but analytically treatable feature-integration model and show that local emergence of critical dynamics may maximize the figure-ground segregation task performance of a coincidence detector. In chapter 7, we will expand the model from chapter 6, such that the biological realism is improved upon, although the analytical understanding is lost. We will show, nevertheless, that the more detailed model also exhibits critical dynamics modulated by stimulus salience in a typical contour integration setting. Finally, we will conclude this thesis with a summary and outline of the implications of our results in chapter 8 and a discussion of how our findings fit into the scope of the existing literature as well as their robustness and significance in chapter 9.



2. Criticality in Dynamical Systems

Many past and ongoing scientific studies in computational neuroscience borrow ideas from the field of thermodynamics in order to gain insights into the way the brain is organized and how it performs certain functions. Some of these ideas and their previous applications to neuronal dynamical systems are integral for the rest of this thesis. Thus, in this chapter we are going to discuss basic terminology and concepts from thermodynamics and phase transitions in physical systems which have been relevant in neuroscientific literature constituting the background for the work presented here.

2.1 Continuous phase transitions and the Ising model

2.1.1 Phase transitions

A portion of a thermodynamical system which is homogeneous in terms of its physical properties is called a ‘phase’ [Pla03]. For example, a closed space full of liquid water, ice and water vapour would form a thermodynamic system of three phases. Transitions of physical systems between different phases are commonly found in nature. Phase transitions such as the evaporation of water or melting of ice are probably the most frequently observed ones in our every day lives. However, there are many different types of phase transitions which are of interest to physicists, as they seemingly occur in exceedingly different systems, and yet may share a surprising number of characteristics.

A formal, thermodynamic definition of phase transitions involves the free energy (or the energy available in a thermodynamic system to do work), which is non-analytic for a system which displays a phase transition (see e.g. chapter 10 in [KK80] for a detailed description). In other words, the thermodynamic free energy of a

system at the critical point, is not infinitely differentiable, but either exhibits a finite discontinuity or tends to infinity at some derivative. For the examples of melting and vapourization as a function of temperature under constant pressure, the first derivative of the Gibbs free energy displays a finite discontinuity, and this discontinuity manifests as latent heat. However, it is important to note that many phase transitions do not display such discontinuities in the first derivative of the free energy, and those which do are called first-order phase transitions after the, now antiquated, Ehrenfest classification of phase transitions [Jae98]. First-order phase transitions may also show hysteresis, as observed for example in superheating or supercooling [KK80].

The second type of phase transitions do not exhibit latent heat or hysteresis, as they do not involve a discontinuity in the first derivative, and are accordingly called continuous or second-order phase transitions. However, a discontinuity is still observed at higher order derivatives of the free energy at the critical point, where, in contrast to first-order transitions, correlation length diverges and many observables of the system obey power-law distributions [KK80]. The ferromagnetic transition, which will be described in detail in the rest of this chapter, is an example of a continuous phase transition.

When describing a system which undergoes a phase transition, it is often meaningful and useful to define a so-called ‘order parameter’, which will assume a value of zero when the system is in one phase and non-zero values in the other phase. For example magnetization is the order parameter for ferromagnetic systems in the absence of an external magnetic field, and spontaneous magnetization emerges in the ordered phase, below a critical point called the Curie temperature [KK80; Sta71]. The type of the phase transition is, in general, also apparent in the behaviour of the order parameter around the critical point, where it will be discontinuous for a first-order transition and vanish continuously for a continuous phase transition.

The value of the order parameter, which indicates the state of the system, is characteristically dependent on an independent variable, sometimes referred to as the control parameter. Since phase transitions were first discovered and investigated in thermodynamic systems, the control parameter is often the temperature. It is important to keep in mind, however, that there are many systems which display these two types (i.e. first-order or continuous) of phase transitions but are not thermodynamic systems, including cellular automata [Lan90; OFC92], percolating systems [SA94] and neuronal networks [BR00; EHE02]. In such cases, critical points exist as a function of different control parameters, which may, for example, be parameters determining topology and the connectivity structure or the interaction strength between the elementary units making up the system.

As mentioned before, at the critical point of a continuous phase transition the correlation length diverges and power-law distributions characterize the system. These signify that in a system poised at criticality, there are no characteristic length scales, the system exhibits self-similarity and global properties, such as order, can emerge from local interactions, all of which have interesting implications for biological systems.

Next, we will demonstrate some of these exciting mathematical properties of continuous phase transitions in detail in the context of a well-known model: the Ising model. We will focus on the Ising model because it is one of the oldest and most rigorously investigated models of continuous phase transitions, because it has been a point of inspiration for the dynamical models which we will introduce in the following section, and finally because it has been used extensively (both in phenomenological analogies as well as in data analysis) in investigations of neuronal systems.

2.1.2 The Ising model

The Ising model is a historically influential model of spin interactions describing magnetization which answered a very important question: how can short-range interactions give rise to long-range correlations? It therefore constitutes a simple, but elegant example in which to explore critical phenomena. In this section, we are first going to briefly introduce the Ising model, use its formal definition to visualize and illustrate the specifics of a continuous phase transition and present in detail the properties associated with such a phase transition, such as the critical exponent relations and universality classes, which will be relevant for the results presented in this thesis.

The Ising model is a model of how ferromagnetism emerges due to spontaneous symmetry breaking at a critical temperature. It is defined for an undirected graph with a d -dimensional lattice structure. The state of the whole system is given by a spin configuration σ , where on each lattice site i , there is a magnetic moment generated by spin states $\sigma_i \in \{-1, +1\}$. The model considers only short-range nearest neighbor interactions between the lattice sites. In its simplest form, the energy of a spin configuration σ in the Ising model is denoted by the Hamiltonian

$$H(\sigma) = -J \sum_{\langle i, j \rangle} \sigma_i \sigma_j - h \sum_i \sigma_i \quad (2.1)$$

where h is an external magnetic field acting on the system, J is the interaction energy between lattice sites, and $\langle i, j \rangle$ indicates that the sum is only over pairs of adjacent lattice sites i, j [Gal99]. Positive interactions ($J > 0$) between the spin sites are called ferromagnetic and are responsible for the spontaneous emergence of magnetism. For the purpose of demonstrating phase transitions in an example model, it is sufficient to consider the case $h = 0$ [Sta71]. Therefore, for simplicity, we will only examine the system without an external field.

In general, one can employ the Hamiltonian $H(\sigma)$ to compute the probability $P(\sigma)$ that the system at equilibrium will assume the state σ , using the Boltzmann distribution

$$P(\sigma) = \frac{e^{-\beta H(\sigma)}}{Z_\beta} \quad (2.2)$$

where $\beta = (k_B T)^{-1}$ is the inverse temperature and Z_β is the partition function

$$Z_\beta = \sum_{\sigma} e^{-\beta H(\sigma)}. \quad (2.3)$$

For $d > 1$ the model exhibits a continuous phase transition at a critical temperature T_C [Gal99]. For $T > T_C$, the system is disordered due to thermodynamic fluctuations and is paramagnetic. For $T < T_C$, the model system transitions to an ordered phase and is ferromagnetic [Bin11]. In order to quantify the phase transition in the Ising model, the net magnetization M is used as the order parameter, and it is defined as the mean spin value

$$M = \frac{1}{N} \sum_{i=1}^N \sigma_i \quad (2.4)$$

in a system with a total size of N lattice sites. As mentioned in section 2.1.1, in the thermodynamic limit, M takes on non-zero values only in the ordered phase ($T < T_C$) and, as is typical in a continuous phase transition, it vanishes continuously as temperature is increased (see, for example, [Ons44] or [Yan52] for the exact solutions to the 2-dimensional zero-field model).

Using the configuration probability $P(\sigma)$ we can simulate the 2-dimensional Ising model via the Metropolis–Hastings algorithm [NB99] and plot the resulting spin configurations for the disordered, ordered and critical states (Fig. 2.1). In the

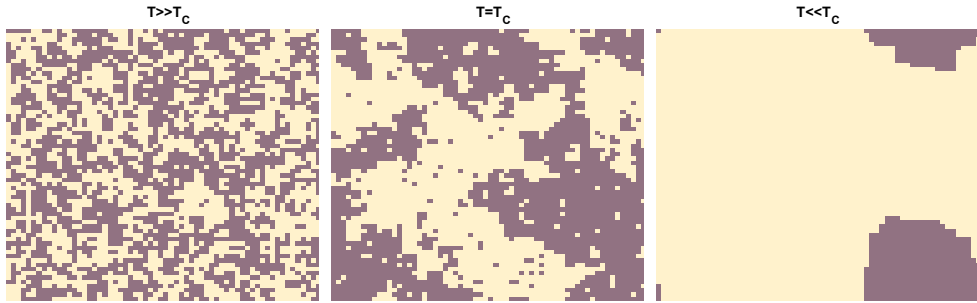


Figure 2.1: Spin configurations σ in a single simulation of the 2-dimensional Ising model on a square lattice, in the disordered (left), and critical (middle) and ordered (right) states. The binary color map represents the up and down spins at each lattice site.

disordered phase, we observe that the system is dominated by random fluctuations and the resulting spin configuration resembles salt-and-pepper noise. As we decrease the temperature, the positive nearest-neighbor interactions lead to the formation of ‘domains’ or clusters of nearby spins which all point in the same direction, until at $T = T_C$ the symmetry in the system is broken and we get net magnetization $M > 0$. This was an important discovery because it provided an explanation for how large masses of, for example, iron could be ferromagnetic, with the majority of spins directed in the same direction, even though the single atoms could only interact with their nearest neighbors.

Putting it more formally, we say that at the critical point between order and disorder, nonvanishing correlations emerge at large distances. For example, in order to quantify how often the spin states $\sigma_{i,j}$, separated by a distance $|i - j|$ (where bars

denote the Euclidean distance), flip in a coordinated fashion, we can compute their covariance $C = \langle (\sigma_i - \langle \sigma_i \rangle)(\sigma_j - \langle \sigma_j \rangle) \rangle$. Interestingly, C decays exponentially with distance as $C \sim \exp(-|i - j|/\xi)$ at all temperatures T , except at $T = T_C$ it assumes a power-law decay $C \sim (|i - j|)^{-\eta}$, and the correlation length ξ diverges [BT12; Gal99].

The pairwise spin correlation as a function of distance, however, is only one of the several observables of the zero-field Ising model which assumes a power-law scaling at the critical point, as mentioned in section 2.1.1. For example, if we define a reduced temperature difference $t = (T - T_C)/T_C$, we see that the correlation length as a function of temperature also diverges obeying a power-law $\xi \sim |t|^{-\nu}$. Similarly, the specific heat, which is the second derivative of the free energy with respect to the inverse temperature, $C_T = \frac{\partial^2 f}{\partial \beta^2} \sim |t|^{-\alpha}$ also diverges following a power-law, and the decay of the magnetization to zero, as temperature is increased, can be described by $m \sim (-t)^{\beta_C}$. Finally, exactly at $T = T_C$ the magnitude of the magnetization is related to the external field by $|m| \sim |h|^{1/\delta}$ [Gal99].

The exponents η , ν , α , β_C and δ given above are examples of so-called ‘critical exponents’ as they are central in characterizing the behaviour of the system at the critical point. By definition of the power-law function, one can, for example, scale an independent variable, such as $|i - j|$ or $|t|$ by a factor, and the behaviour of the system doesn’t change, in the sense that the power-law dependence and the exponents stay the same, except the function is now also scaled by a factor. As such, a power-law relationship is called scale-free, is indicative of self-similarity in the system, and is believed to be a common property of systems undergoing continuous phase transitions. The thought-provoking observation about the critical exponents is that, they tend to not change when some properties of the underlying system are dramatically altered. For example, in the Ising model, one can move away from the square lattice, to a triangular or hexagonal lattice structure, or instead of considering only nearest-neighbor interactions, one can include other short-range interactions (e.g. involving second or third nearest-neighbors) in the Hamiltonian, and the critical exponents stay the same for the same number of dimensions [Sta71]. As such, a set of exponents can describe the system behaviour in a large class of models. This is called a universality class, which is both an elegant way of categorizing phenomena in physical systems based on shared properties and is helpful in drawing links between established models and newly discovered systems belonging to the same universality class. For example, the Ising model is considered to be in the same universality class as the liquid-gas phase transition, which exhibits critical opalescence at the critical point.

Another line of analytical investigations, starting in the 1960s and initially mainly focused on the properties of the critical point of the Ising model, revealed that there exist specific relationships between the critical exponents within a universality class of models. For example, the universal scaling hypothesis predicts, among others, relations such as $\alpha + \beta_C(\delta + 1) = 2$ to hold for the above given critical exponents [Kad66]. Such exponent relationships give support to the idea of universality

of behaviour based on thermodynamic considerations on the microscopic level for a wide variety of systems.

The idea that the advances made in thermodynamics and statistical physics in the 19th and early 20th centuries may be applicable to more complex biological systems is an old idea, going back at least to Niels Bohr, who predicted early on the experimental difficulties posed by *in vivo* research [Boh33]. In the modern era, which gave us the more detailed studies of critical exponents and scaling laws, it became intuitively more conceivable that the complex statistical behaviours that emerge around a continuous phase transition may be analogous to the complexity observed in biological systems. This intuition is supported on a phenomenological level by power-law scaling of a wide range of observables in biological systems. An extensive example is the allometric scaling laws, which is the scale-free distribution of diverse quantities, including metabolic rate, life span, growth rate, heart rate and mass of grey matter in the cerebrum as a function of the mass of an organism (for a review see [WB04]).

One of the earliest suggestions that neuronal networks can be described by interactions between binary units which yield an energy function isomorphic to the Ising model was by Hopfield in 1982 [Hop82], and the Hopfield network model remains as one of the most influential examples of how neuronal networks can learn, store and recall patterns. More recently, it has been shown that using maximum entropy models to describe experimental neuronal data yields an ‘energy landscape’ similar to the Ising model. The details of such models are beyond the scope of this thesis, however, fitting Ising-like models to neuronal data has shown that multi-neuron firing patterns and correlations in the whole network are explained accurately by simply considering the pairwise interactions, for example, in the primate retina [SFG+06], in the salamander and guinea pig retinas as well as in cultured cortical networks [SBS+06; TMM+13]. Importantly, such studies suggest that the investigated neuronal networks display critical phenomena, as demonstrated by the divergence of the specific heat [MB11; MDM15; TMM+15], although there is some skepticism about the validity of the used methods [NBB+17].

In general, there are two important short-comings of such statistical approaches. First, they use Ising-like equilibrium models which don’t explicitly consider the dynamics of the system, whereas the central nervous system in vertebrates is most often described as a dynamical system, capable of real-time computation in a perpetually changing environment [BT12]. Therefore, a model which considers not just equilibrium statistics, but dynamics of the response of the system to external perturbations would be more informative when investigating information processing in neuronal networks. And second, the Ising model as described in this section only undergoes a phase transition at a specific value of the control parameter $T = T_C$, and the interesting properties, such as the divergence of the correlation length or specific heat, decay relatively fast, following a power-law, as one moves away from the critical point. Ising-like models used to explain correlations in neuronal data do not consider how the investigated systems are tuned to the critical state, which typically represents a tiny portion of the phase space. We will focus on these two

points in the next section, in the context of a mechanism suggested to push dynamical systems towards the critical point without the need to fine-tune a control parameter.

2.2 Self-organized criticality

As we have seen in the previous section, power-law scaling is a characteristic behaviour which emerges at the critical point of a continuous phase transition. Independently, it has long been observed that many physical systems display flicker noise with $1/f$ -like scaling, such as the electrical noise in resistors and transistors, ocean current velocities, variations in temperature and light intensities in astronomy as well as the thickness of glacial varves [Pre78]. However, in the absence of a precisely regulated control parameter, it would not seem feasible to link flicker noise to critical phenomena.

This view was challenged by Bak, Tang and Wiesenfeld in 1987, when they presented the Abelian sandpile model whose dynamics were constantly drawn toward a critical point at which a phase transition occurs [BTW87]. Termed self-organized criticality (SOC), due to the emergence of power-law scaling in the absence of fine-tuning in the parameter space, such attractor dynamics were suggested to comprise an explanation for a variety of empirical $1/f$ noise observations in natural and human made systems.

The original Abelian sandpile model described a finite grid on which each site was assigned a value representing the slope of a sandpile. The system was perturbed by the addition of a single sand grain at one site on the grid at a time, which increased the slope at that location. The slope exceeding a specific threshold after the addition of a sand grain would cause the sands at that site to topple over to neighboring sites, effectively resetting the slope at the toppling location and increasing the slope at adjacent locations. The toppling was repeated iteratively until no sites exceeded the threshold, terminating an ‘avalanche’ which began with the external perturbation. The number of sites which participated in an avalanche, as well as the avalanche lifetimes displayed scale-free distributions [BTW87]. This kind of slow external drive is one of the distinguishing properties of many SOC models, and is a cause for concern when one tries to discern distinct avalanches in both empirical data and in models with strong external drive. In order to avoid overlapping avalanches in our models, we employ a separation of time scales, which will be explained in detail in the corresponding methods sections in parts II and III.

There is experimental and theoretical evidence that diverse physical systems including sandpiles [HSS+90], ricepiles [FCMS+96], water droplets [PNB93], superconductors [FWN+95], variations in soil moisture and landslides [PMB+97] and earthquakes [BP04] exhibit SOC [Jen98]. Such observations led to the inception of further models displaying critical dynamics with power-law scaling such as the forest-fire (FF) model describing percolating systems [DS92] or the Olami-Feder-Christensen (OFC) model describing earthquake dynamics [OFC92].

Despite the claims that the Abelian sandpile model is self-organizing without the need to fine-tune a free parameter, there have been discussions regarding the fact

that the local conservation of the number of sand grains constitutes, in a sense, a free parameter. Indeed, in the sandpile model, when one site relaxes, an equal amount of total energy is distributed to the adjacent sites. It has been shown that in a dissipative regime, where some particles are annihilated during the relaxation process, SOC is lost [TK00], and that local conservation laws are necessary, at least on average, to observe SOC [MKK90]. Similarly, the FF model has an explicit control parameter θ , for which the system displays critical dynamics in the limit $\theta \rightarrow \infty$ [Gra02], which holds sufficiently well in models of rare events. Whether SOC could be observed in the dissipative regime in other models was investigated rigourously in the context of the OFC model, which also involves a free parameter α_{OFC} . Initially introduced as a model which displayed critical phenomena for a large range of α_{OFC} below the critical point [LJ96; OFC92], since then it has been shown that the OFC model is probably not critical in the dissipative regime [BG97]. These findings, however, do not directly obstruct claims about self-organizing physical systems, as conservation laws are common in nature. Nevertheless, care must be taken when SOC models are used to describe systems, such as neuronal networks, which are known to be locally non-conservative, and mechanisms other than self-organization should be explored as candidates for generating scale-free dynamics. In our models, we will consider an objective function, such as the task performance in an experimental scenario, and investigate whether their landscape can be indicative of the critical point.

As we have seen for the example of the Ising model, critical exponents play an important role in characterizing a system. However, they are often difficult to analytically calculate in complex, multi-body systems exactly. However, a generalized approach to sandpile-like models revealed that SOC dynamics can be described as a critical branching process, where the avalanche size and lifetime distributions follow power-laws with critical exponents $\tau = 3/2$ and $\alpha = 2$ respectively [ZLS95]. Outside of physical systems, the idea of SOC has also been applied to dynamical systems in life sciences and other fields, demonstrating critical phenomena in models of evolution [BS93], epidemics [RA96], econophysics [BCS+93; PP13], human closed loop control [PRE+07] and neuronal networks [BR00; EHE02]. In particular, the Eurich–Herrmann–Ernst (EHE) model of neuronal networks presented formal derivations of the avalanche size and duration distributions, and predicted a critical exponent of $\tau = 3/2$, in line with a branching process, which was later confirmed by experiments [BP03]. The experimental observations of critical dynamics in the brain will be presented in detail in chapter 3, however, it is important to mention here that such observations led to the hypothesis that the brain, as a dynamical system, might be poised at a critical point and that SOC might be a fundamental organizational principle for neuronal networks [Beg08].

Even though they are often quoted together as one hypothesis [MB11; NBB+17], the suggestion that the brain is a dynamical system with SOC-like properties, and the suggestion that correlation patterns in neuronal dynamics can be explained by Ising-like equilibrium models close to a critical point, as we’ve seen in section 2.1.2, should probably be thought of as distinct hypotheses. As of yet, there have been no explicit or formal links established between the two approaches, or between the two

mechanisms which are both capable of generating scale-free correlation structures. Nevertheless, there have been attempts to bring the two concepts together in the context of neuronal systems. For example, hidden Markov models were used in order to extend the maximum entropy Ising model to predict spatio-temporal patterns in neuronal activity [MEBF+09], which accounts for the time dimension, but such models still do not explicitly account for external perturbations or non-equilibrium dynamics. Thus it has both been proposed that it may [LD09; MDM15] and may not [SHH19] be possible to relate equilibrium models to dynamical SOC models.

Still, it is conceivable to link the two approaches more closely in the context of ferromagnetism. Continuous phase transitions, such as spontaneous magnetization, are sometimes thought of as phenomenon of macroscopic instability [Gal99]. This means that small external changes are capable of inducing dramatic changes on the system. Therefore, there have been efforts to extend the equilibrium models of ferromagnetism to dynamical models, by perturbing the system out of its equilibrium state using a changing external magnetic field. Under a slow but continuous increase of the external field, the magnetization of the system increases in discrete steps, which is called crackling or Barkhausen noise. The discontinuity in magnetization as a function of time emerges as domains of spins change their direction rapidly and collectively, analogous to an avalanche. Models of crackling noise, however, typically consider systems at zero temperature, away from the paramagnetic to ferromagnetic transition, and instead introduce domains to the system by using a local field with a disorder parameter R . Magnetization in such a system changes as the spins orient themselves to a changing magnetic field via avalanches, however the avalanches show true power-law scaling only for a critical value of the control parameter $R = R_c$ [PDS95]. As such, the relationship to an equilibrium Ising model is not explicit but some links have been drawn between the two [LD09]. Nevertheless, studies of crackling noise have revealed signatures of criticality apart from just power-law scaling [PBS+11] which may be easily applied to empirical data, where it's often difficult to fit power-laws accurately [CSN09]. Among these signatures are data collapse, as well as critical exponent relations analogous to the Ising model, which were confirmed to hold in neuronal data [FIB+12]. One relevant exponent relationship links the exponents τ and α of avalanche size and duration distributions via

$$\frac{\alpha - 1}{\tau - 1} = \frac{1}{\sigma \nu z} \quad (2.5)$$

where the average avalanche size $\langle s \rangle$ conditioned on avalanche duration T is described by the power-law distribution $\langle s \rangle(T) \sim T^{1/\sigma \nu z}$ [PDS95; SDM01]. We will also use Eq. 2.5 in the investigations presented in this thesis to demonstrate universal scaling in our models.

Finally, the dynamics of avalanches taking place at all spatial and temporal scales at the critical point has significant implications for many biological systems. Especially in neuronal systems, avalanche dynamics may help explain how neurons communicate with each other over long distances and how they collectively behave in response to external input. Continuous phase transitions as seen in these SOC

models are often described as the border between chaos and order [Kau93]. Conceptually chaotic systems contain little redundancy in their dynamics, which boosts information entropy, but are characterized by instable trajectories, which is thought to be unfavorable for recall, or object representation. This might indicate the border between order and disorder might offer some optimal computational properties. In fact, it has been shown in cellular automata that such a critical point maximizes the complexity of computations [Lan90]. In parts II and III of this dissertation, we will try to make such links to computation more concrete in the context of the cortical neuronal networks performing behaviourally relevant tasks.



3. Criticality in the Brain

In this chapter we are going to introduce the general ideas surrounding the applicability of models displaying critical phenomena to neuronal systems. In particular, we are going to talk about the existing literature which presents evidence, both from theoretical studies, and from experiments, that the brain might be poised at a critical point, which is optimal for information processing and performing complex computations.

3.1 Theoretical work

SOC models have been used to describe the observations of scale-free distributions in a variety of biological systems, including hair cells in the inner ear [CDJ+00], the collective motion of flocks of birds [TT95], human closed loop control [PRE+07], and the thermoregulatory huddling behaviour in rodents [Wil17]. Arguably, all of these systems are examples of optimization problems, implying critical states might represent—at least local—minima to such problems.

Early theoretical work has related this idea to computation, and has shown that the border between chaos and order might be optimal when performing complex computations using data streams [CY89] or cellular automata [Lan90]. This prompted a natural extension of SOC models to neuronal networks [BR00; CPDG+95; EHE02] which carry on rapid, complex computations continuously and robustly. Such studies suggested that emergent phenomena commonly observed in neuronal activity, such as synchronization and pattern formation, could be explained by a phase transition [Bot95; HH95; USO95]. The Eurich–Herrmann–Ernst (EHE) model formally adapted the OFC model in the conservative regime to a globally connected topology, and calculated a closed-form expression for avalanche size distributions in finite size systems as a function of the control parameter, the interaction strength, α . This study

also predicted that avalanches of cortical neurons may belong to the same universality class as a critical branching process with the power-law exponent $\tau = 3/2$ [EHE02]. The analytical knowledge of the critical value of the control parameter α_c and the dependence of the avalanche behaviour on α made it possible for us to extend the EHE model to a heterogeneous topology and run simulations without extensive parameter searches in our feature integration model (chapter 6).

The EHE model is, therefore, one of the earliest models concerned with bringing together the interesting ideas surrounding SOC and neuronal systems. However, the elementary units in the EHE model are non-leaky integrators, and leak currents are an integral part of biological neurons, which is why leaky integrator units are used as the neuron model in our investigations relating to contour integration (chapter 7) and selective attention (chapter 5), where a higher degree of biological realism was necessary to link our results to experimental observations. Nevertheless, it has been shown that the EHE model also displays approximate power-law scaling in the leaky regime [EHE02], and that in general, leaky integrate-and-fire units may display phase transitions that have been linked to SOC [CPDG+95; USO95].

As stated before, one of the biggest motivations for the work presented in this thesis is the hypothesis, supported by a large number of previous studies, that a functional relationship exists between the dynamics of a system close to a phase transition and optimal information processing in the brain. Following up on earlier work done in more abstract models such as cellular automata, criticality studies centered around neuronal computation have suggested that critical dynamics may play an important role in supporting complex computations [BN04], maximizing information diversity [NPL+08] and the number of metastable states [HB05], and improving object representation [TRE14] with favourable dynamic range characteristics [GHM+15; KC06; SYP+09].

As a specific example, Bertschinger and Natschl ger proposed a measure of ‘computational capability’, which quantified the separation of network states as a function of the distance between different driving inputs, in a recurrent, randomly coupled network of binary threshold units, driven by time-varying input [BN04]. Using a mean-field approach, they identified the boundary between chaotic and ordered dynamics, and demonstrated that in a liquid state machine (LSM) framework, the memory capacity is maximized close to the phase transition. Depending on the task, and the complexity of the function being computed by the LSM, the memory capacity could display a very sharp peak around the critical point, leading the authors to conclude that complex computations could only be performed near the edge of chaos.

Such a model is a good example of what we mean when we say many theoretical investigations of criticality employ abstract measures or models lacking in biological realism. In terms of its details, the model described above assumed random (Gaussian) weights for the recurrent connectivity, where the in-degrees for all units were identical, making it a prime candidate for analytical treatment while lacking in biological complexity. Similarly, the performance of the LSM was evaluated on the basis of parity and randomly drawn Boolean functions, so as to ensure the task

is not linear separable. While this is a decent representation of complex nonlinear functions computed by the cortex for such a simple model, how the maximized memory capacity demonstrated in a model of binary units, for the computation of boolean functions, might play a role in sensory input processing in the brain is not immediately clear.

Another interesting approach by Haldeman and Beggs employed a branching process to model suprathreshold LFP activity recorded by a microelectrode array [HB05]. They found that a critical branching process, where the branching parameter is unity, successfully reproduced the correlation structures of the dynamics measured in cultures of the rat cortex. In addition, they demonstrated that the branching parameter which maximizes the number of metastable states in such a model approached unity with increasing system size. Once again, such a model provides strong insights about the emergence of correlation structures measured by electrode arrays, and how operating close to criticality—where the activity from different channels are neither completely independent, nor mostly redundant—may enrich neuronal dynamics. However, it is still an indirect model of the underlying neuronal system and does not address the question of how the maximized number of metastable states might be beneficial for a concrete function or for potential read-out schemes.

Several other frequently cited papers in literature report that critical dynamics and neuronal avalanches maximize measures such as dynamic range, which are more relatable to sensory information processing and possible to quantify in control experiments with stimulus-evoked activity [GHM+15; KC06; SYP+09]. However, while such models often incorporate certain biological details (such as refractoriness [KC06] or short term depression [GHM+15]) which help capture the main characteristics of the neuronal system under investigation (such as global oscillations [KC06]), such models still leave a lot to be desired in terms of biological realism. For example, they often employ binary neurons [GHM+15; KC06; NPL+08; SYP+09], consider homogeneous models, with uniform topological properties (e.g. all-to-all connections [BN04; GHM+15; SYP+09]), whereas cortical neuronal networks often display robustly structured functional connectivity. Similarly, most SOC models depend on a slow external drive in order to observe phase transitions at all [PWV+14], whereas the awake cortex engaged in a task, where the need for fast information processing is at its highest, may be bombarded with parallel streams of external input. Finally, as elaborated above, demonstrations of a functional role for criticality often rely on rather abstract theoretical measures, without realistic encoding or read-out scenarios in mind. Addressing these points in the context of critical dynamics in neuronal networks under sensory stimulation is one of the main goals of the results presented in this dissertation.

Of course, modelling alone is insufficient to formalize claims about criticality in the brain. Due to experimental challenges surrounding measurements of electrical activity from a large number of neurons simultaneously (in order to, e.g. quantify neuronal avalanche statistics), it has only been possible in the last few decades—thanks to advances in recording techniques such as multielectrode arrays—to support the theoretical claims with experimental evidence, some of which will be discussed

in the next section.

3.2 Experimental evidence for criticality

Initial experimental evidence for the existence of critical dynamics in cortical neuronal networks was reported in 2003 by Beggs and Plenz [BP03], which had a similar effect on criticality literature as the Abelian sandpile model, in that it opened the way to a large number of research questions relating to critical phenomena in the brain. The Beggs and Plenz study demonstrated avalanche dynamics in the local field potentials (LFPs) in organotypic cultures as well as in acute slices of rat cortex, whose size distributions followed a power-law with the critical exponent $\tau = 3/2$. As mentioned before, this placed the neuronal avalanche dynamics in the same universality class as a critical branching process [ZLS95] as well as the EHE neuronal network model [EHE02].

A multitude of experimental investigations followed which demonstrated power-law scaling in the brain in a large range of species and temporal and spatial scales. For example, it was shown that the fine scale *in vitro* spiking activity in dissociated and cultured cortical networks from the rat cortex displayed dynamics which were organized into scale-free avalanches [PMB+08]. Similarly, it was reported that the avalanche size distributions of spontaneous *in vivo* LFP activity in awake monkeys displayed power-law scaling [PTL+09]. On a larger spatial scale, power-law behaviour was presented in resting state MEG and EEG in humans [PZH+13; SAC+13]. Also in humans, intracranial depth recordings displayed signatures of critical avalanches [PVW+13]. In addition, recent experimental studies address more specific questions relating to, for example, the rigorousness of statistical analysis [KYP11], subsampling [PMW09] and resolution restraints [FIB+12] in investigations of neuronal avalanches. In terms of signatures of criticality beyond simple power-law scaling, it has also been shown that neuronal data exhibits data collapse and obeys the critical exponent relation given in Eq.2.5 [FIB+12].

Combined with the theoretical findings of improved information processing presented in section 3.1, such experimental demonstrations of dynamical criticality constitute the so-called ‘criticality hypothesis’ that cortical neuronal networks operate near a phase transition [Beg08; Chi04; FBF+09; Ple14]. The hypothesis posits, that either through self-organization, through regulation of neuromodulators or via adaptation to external input, the dynamics of neuronal networks are tuned to the critical point, supporting optimal information processing. As such it has both been proposed that self-organized criticality is a fundamental organizational principle in the brain, and that criticality emerges as the brain optimizes certain computational properties [BL98; Chi10; CB99a; HG14].

However, the majority of neuronal avalanche observations are made in recordings of either spontaneous activity *in vitro* or the resting state *in vivo*, in the absence of an actual sensory stimulus being processed by the cortex. Some recent studies address this gap in the literature and investigate the emergence of criticality in the cortex under sensory stimulation [AGS15; SCP+15]. However, one of those experiments reports

findings in the disembodied turtle cortex and the other one demonstrates scaling in MEG data, therefore, little is known so far about how the neuronal dynamics change in the awake, behaving state on most spatial scales. Furthermore, due to substantial experimental challenges, no empirical studies exist to date which explore criticality in local neuronal dynamics in relation to a direct behavioural task and how it might vary under changing task demands.

In addition, as mentioned in chapter 1, scale-free organization does not seem to consistently emerge under different conditions. For example, it has been hypothesized that since fundamental brain functions such as responsiveness to stimuli or learning tends to be impaired by sleep deprivation, a functional role in restoring optimal information processing, and thus critical brain dynamics, may be attributed to sleep. Experimental evidence indicates that signatures of criticality, in the form of precise power-laws in EEG recordings, are progressively disturbed through the course of 40 hours of sustained wakefulness in humans [MOS+13]. Similarly, it is not clear whether scale-free dynamics are present in the active brain when it's processing sensory information in order to perform a behaviourally relevant task, which would constitute a strongly driven regime either by external (e.g. thalamic) input or by interareal connections or both. In fact, in an experimental study conducted with children, signatures of criticality in MEG recordings were shown to diminish with increased cognitive load (defined by the authors as the difficulty of the task which correlated with attention allocation) during executive function tests in different modalities [TV14].

In a complementary line of thinking, it has also been suggested that the ground state for brain dynamics is not perfectly critical, but slightly subcritical. For example, it has been shown that avalanche dynamics of intracranial depth recordings in humans exhibit differences between different vigilance states (wakefulness, slow wave sleep and REM sleep) [PVW+13]. An SOC model of these observations indicate that the brain in general operates in the subcritical regime, with small, global changes to the effective synaptic strength mediating the variations with vigilance states. A more detailed follow-up study, using data from awake rats and monkeys, anesthetized cats as well as humans presents further support for the subcriticality hypothesis [PWV+14].

Based on the inconsistencies between observations of criticality, the criticality hypothesis still remains controversial in the field. This is evidenced by several studies suggesting the experimentally reported scale-free behaviour is just an artifact of incorrect approaches to data or imprecise statistical analysis. For example, one influential study shows no power-law scaling in single unit activity in recordings of a large range of species and vigilance states (cat, monkey, human; wakefulness, slow wave sleep and REM sleep) [DHH+12]. Interestingly, they report that apparent power-law scaling for the LFPs in the same data set will not survive rigorous statistical analysis, where alternative, non-power-law hypotheses (such as multiple exponential) provide better fits to the empirical distributions. Similarly, other authors have questioned what level of statistical analysis is appropriate to conclude that apparent power-laws are indeed precise power-laws [SP12; TD10], whether power-laws are representative of underlying critical dynamics or whether they can emerge as

an epiphenomenon in the macroscopic scale in systems of independent stochastic units [TD17], and whether useful new insights can at all be gained by rigorously identifying power-laws[SP12].

To summarize all the points discussed in this chapter, the abstract nature of theoretical investigations, as well as the lack of biological realism in many of the models, make it difficult to relate the computational benefits of criticality to concrete brain function. Similarly, the scarcity of studies in the active cortex and the inconsistency of experimental findings of signatures of criticality make it a controversial subject of study, which calls for a theoretical framework which, while modelling critical dynamics, can also closely reproduce other aspects of neuronal activity, ideally on a range of scales from the single neuron all the way up to systems and behavioural level. To that end, we believe it is necessary to first try and narrow the gap between theory and experiment. In other words, we believe there is a significant lack, in existing literature, of critical models with relevant biological details, which showcase optimal information processing properties in behaviourally relevant tasks. We believe tackling the issues of biological realism and physiologically plausible encoding and read-out mechanisms will help us gain better insights about the underlying system, how and when it might display critical dynamics and potentially provide an explanation for the inconsistencies in experimental findings.

With the results which will be presented in this thesis, we are aiming for this apparent gap of concrete, behaviourally anchored examples in the criticality literature. As a model system which has both been rigorously investigated, and is known for efficient and rapid processing of parallel streams of information, we will specifically focus on the visual system. As stated both in the abstract and in chapter 1, we will explore how the remarkable speed and robustness with which the visual system modulates its responses with varying sensory input and changing task requirements may be supported by neuronal avalanches. We will elaborate on some of these basic but crucial functions of the visual system in the next chapter.



4. The Visual System

In this chapter, we will present a brief introduction to the visual system, its dynamics, and the typical computational problems it has to solve, as well as establish the general concepts and terminology which will be relevant for the detailed examples, given in the following chapters, demonstrating how critical dynamics might support cortical function.

4.1 Basics of the visual system

In the next sections we will introduce some of the basic properties of the biological components which make up the visual system in the brain. We will intentionally focus only on the properties which are relevant for the results presented in this dissertation, which will inevitably mean leaving out some interesting details. We refer the curious reader to [KSJ+00] for further information on the anatomy and physiology of the nervous system and to [DA01] for a detailed overview of different computations performed by the brain.

The neuron

The neuron is a type of cell which makes up the nerve tissue, which, in turn, makes up the nervous system found in most members of the animal kingdom. The lipid bilayer membrane of the neuron is relatively special, in that through the function of so-called ion channels and ion pumps, it retains a difference between the concentrations of different ions inside and outside of the cell membrane. With the largest contributors being K^+ , Na^+ and Cl^- ions, these different ion concentrations generate an electric potential difference across the membrane, which underlies the ability of neurons to communicate and perform computation.

The basic anatomical structures that broadly make up a neuron are the cell body

(soma), where the cell nucleus and most of the organelles are typically located, the dendrites, which are branching extensions from the cell body which often assume the role of the receiver in neuronal communications, and the axon, which is a cable like extension from the cell body which generally functions as the transmitter, to continue the analogy. Different neurons typically communicate via chemical synapses found at the axon terminal, which, depending on the activity within the cell, may release neurotransmitters into the gap between the axon terminal and the dendrite of a second neuron, termed the synaptic cleft. The neurotransmitters would then bind to receptors found in the dendrite, potentially eliciting a response in the second neuron.¹

When a neurotransmitter activates a receptor protein in the dendrite, it may become transiently permeable to certain types of ions, leading to the flow of current in or out of the cell. Physiologically, this leads to a local change of voltage in the dendritic branch, which travels along the length of the dendritic tree. However, in a greatly simplified view of the neuron, its state can be parametrized by the single value of the membrane potential at the soma, which is modulated by such influx of ions up or down the dendritic tree. The membrane potential at the soma often assumes values close to -65 mV, however, through excitatory (inhibitory) synaptic input it may increase (decrease), and the cell becomes depolarized (hyperpolarized). A certain amount of depolarization can push the neuronal dynamics into a regime of positive feedback, where the membrane potential will increase, and subsequently decrease, very rapidly, displaying a voltage spike. Such a voltage spike is then capable of travelling down the axon of the neuron and triggering the release of neurotransmitters to further synaptic clefts.²

The visual pathway

Primates, including humans, depend strongly on their sense of vision when interacting with the world around them. The importance of vision is evidenced by the number and complexity of the structures in the central nervous system (CNS) dedicated to visual information processing [KSJ+00]. The first of such structures is the retina [Dow09], where a two-dimensional projection of the three-dimensional world in your visual field is processed. In the initial stage, photoreceptor cells in the retina perform phototransduction, i.e. convert the electromagnetic radiation in the visible spectrum to electrical signals, which can then be transferred to and further processed by downstream areas in the CNS.

The electrical signals generated by the photoreceptor cells then activate bipolar cells, where information processing already begins to take place through feature extraction, based on the on-center and off-center receptive fields of the bipolar cells. Signals from the retina are then passed on to the retinal ganglion cells, and through their axonal connections, travel down the optic tract to the lateral geniculate nucleus (LGN) of the thalamus [Dow09]. The LGN serves as a relay station which provides

¹For simplicity, other forms of communication between neurons, such as gap junctions, axosomatic or e.g. dendro-dendritic synapses are not introduced here.

²See [KSJ+00] for further biological details and [DA01] for interesting biophysical models of single neuron dynamics such as the influential Hodgkin–Huxley neuron model.

the feedforward input connections afferent to the first cortical area in the visual pathway, the primary visual cortex (V1). V1 is located in the occipital lobe of the cerebral cortex, which, in healthy individuals, is entirely dedicated to the processing of visual information [CM90]—another testament to the importance of vision in the mammalian CNS.

Visual information, after some processing by the local network structures found in V1, is then passed on to downstream areas in the visual pathway. The dominant view of the visual system is that after exiting occipital areas, streams of visual information branch into two pathways, sometimes also referred to as the "what" and "where" pathways [UH94].

The first of these branches is the dorsal stream (the "where" pathway), which extends to the parietal cortex, and includes areas such as middle temporal area (MT) and medial superior temporal area (MST) where neurons selectively respond to spatial and dynamic properties of a stimulus, such as direction and speed of motion [UH94]. Taken together with the proximity of these structures to the sensorimotor areas, the dorsal stream is thought to aid in behavioural tasks which require collaboration of vision and motor skills, such as visually tracking, reaching out and grabbing objects. Since the dynamic aspects of a stimulus is beyond the scope of this thesis, we do not consider the dorsal stream in any of our models.

The second branch of the cortical visual system is the ventral stream (the "what" pathway), which extends to the temporal cortex, and after V1, traverses visual area 2 (V2), visual area 4 (V4) and inferior temporal cortex (IT). Along the ventral stream, the neurons respond preferentially to different properties of a stimulus which may play an important role in object recognition, such as color or features related to outlines or texture [UH94]. In general, when we talk about downstream areas in this thesis, we refer to the ventral stream (and in chapter 5, we explicitly model the activity of V1 and V4 populations). While we often think of the stream of information as a one-way street, and consider only lateral (within an area) and feedforward (between areas) connections in our models, feedback projections between visual areas also play a crucial role when performing behaviourally relevant tasks. We will elaborate on potential contributions from feedback connections, and how they may be incorporated into our models, in the corresponding discussion sections of each chapter.

Response properties and structural organization

Due to the structure of the retina, and the way it receives light, the visual field is projected onto the lattice of cells in the retina rather directly. The mapping from the retina to V1 conserves the spatial relationships between the inputs to different neurons to a good extent, which is called retinotopic mapping [Hor06]. In other words, the visual information contained within a local patch in the visual field will be processed by V1 neurons in close proximity to each other in the cortex. This also means, that just like the retinal cells, each cortical neuron in the visual system has access to information from a small part of the visual field only, termed the size of its receptive field (RF).

Further organization of the neurons in the cortex involves looking at their re-

sponse properties. As mentioned above, given enough input, a neuron may display a "spike" in its membrane potential. Sometimes these spikes are referred to as the currency unit of information processing in the brain, in analogy to the binary representation of information (in bits) in digital systems. A simple way to quantify the response of a neuron to a stimulus within its receptive field is to average the number of spikes it elicits over time. This is called the firing rate and is often given in units of Hz [DA01].

In order to characterize the response properties of a neuron, its firing rate in response to a stimulus (such as a bright bar of light on a black background) can be plotted as a function of changing stimulus parameters (such as the orientation or movement direction of said bar) [HW68]. Such a plot is called a tuning curve, and it is a good way to visualize the selectivity of a neuron for certain features in a stimulus. For certain properties (such as orientation or spatial frequency of a grating stimulus), the tuning curve of neurons selective for those properties will be unimodal, and the x-axis value which elicits the largest response is called the preferred value (e.g. preferred orientation or preferred frequency) for that neuron. Due to such response properties, in its simplest form, a neuron can be thought of as feature detector and can be modelled as a linear filter, with the filter kernel described by its receptive field, which can assume, for example, the shape of an oriented Gabor function if the neuron is orientation selective [BG06; DA01].

Going back to cortical anatomy, V1 is further organized into so-called orientation columns, which refer to the grouping of neurons with similar preferred orientations and nearby receptive field centers into circular structures in the cortex [BS86]. The physical proximity of neurons with similar response patterns in V1 has implications for their functional connectivity to each other, which will be appropriately elaborated on in the introduction sections of the next chapters.

Due to the retinotopic mapping, neurons in different visual areas (such as V1) have full coverage of the visual field [WDB07]. In other words, although the level of detail in the processed information differs between the fovea and its periphery, information from the complete visual field is retained in the neuronal representations in each area. However, the retinotopic maps become more complex as we move away from V1 to higher areas. Similarly, RF sizes become bigger and RF shapes become more complex [MD85; WDB07]. This happens through a pooling of information from upstream areas to down stream areas. For example, a V2 neuron typically receives synaptic input from several V1 neurons, with varying RF centers and preferred orientations. Such pooling of information is thought to aid in bringing together information about local features (i.e. feature integration) and segregating global shapes (such as the outline of an object) from the background in a visual scene. Next we will discuss the relevant concepts regarding how the visual system may tackle the enormous task of combining, and flexibly coordinating responses of local feature detectors in a way which leads to perception and meaningful behaviour.

Visual scene segmentation

A visual scene is usually composed of a multitude of ‘objects’ such as four animals in front of a house with a window (Fig. 4.1). However, as mentioned above, their representations on the retina or the primary visual cortex is distributed and fragmented: Every neuron has access to only a tiny part of the visual field (small blue circles), while objects typically occupy a much larger region (large blue frame). Hence for making sense of a visual scene, the brain has to process these elementary pieces of information in relation to each other, thus enabling perception and, ultimately, a successful interaction with the environment.



Figure 4.1: A natural scene from northern Germany. For segmenting this image into meaningful parts, our brain has to integrate local feature information into global figure representations: For example, orientation-selective cells in V1 might detect the presence of horizontally oriented bars inside their classical receptive fields (blue circles), which are then linked into a representation of the whole window (blue rectangle) by subsequent stages in visual cortex. Thereby representations of different “figures” (such as the dog and donkey) might share subsets of features (contours, outlined in dark and bright red). More complex feature combinations such as whole shapes can only be successfully integrated if attention is directed towards the corresponding location in the visual field (dashed blue circles, head of dog and cat) [Mac03]. (Image adapted from the book chapter by the author and Udo Ernst, titled "The role of criticality in flexible visual information processing", as featured in the book "The Functional Role of Critical Dynamics in Neural Systems", Springer-Verlag, 2019.)

Nevertheless, information integration in vision is computationally demanding. Objects in a scene might be partially occluded or possess fragmented outlines,

thus complicating the integration process. Scenes can be cluttered, requiring to search for the behaviorally relevant information in a background of noise. Low level representation of an object (e.g. in pixels) can vary dramatically, for example trees might seem different when viewed from varying angles, or under different lighting conditions, while still belonging to the same class of objects. Moreover, scenes are dynamic: their content can change rapidly, making a continuous and fast reassessment of the current visual input necessary. Realizing these different computational functions requires many different neuronal processes to contribute, and despite the amazing progress on deep convolutional neural networks in computer vision, it is still largely unclear how the human visual system solves all of these challenges.

Similarly, as described above, the retina of an individual with open eyes receives a large number of streams of information continuously. It is clear both intuitively, and looking at the anatomy and functional connectivity in the visual system, that certain information bottlenecks exist at different stages along the visual pathway, in order to gate out irrelevant information, as well as amplify and bring attention to behaviourally relevant input [VEOA+91]. In particular, the visual system resolves complex questions relating to motion detection, face detection and face recognition (such as those of family members or celebrities) rapidly and robustly on a daily basis. The independence of the success of such tasks from background clutter or partial occlusion suggests that the bottlenecks arising from pooling of information (which lead to the emergence of more complex selectivity patterns and RF shapes, as described before) along the visual hierarchy play an important role in information routing. Such a hierarchy is capable of generating remarkably invariant representations in higher areas, such as the medial temporal lobe (MTL), where neurons selectively respond to strikingly different photographs of a given individual (such as Halle Berry), but not to other faces, and even to images of strings with their name, but not to other text images [QRK+05].

In addition, varying estimates show that less than 10%, up to at most 20%, of the afferent input to V1 neurons stem from the LGN, meaning a big part of information processing and feature extraction in early visual areas potentially depend on lateral and feedback connections [PPB94]. Similarly, experimental studies have shown that suppression of inhibition in V1 may degrade orientation tuning of neurons, and leading models for the emergence of orientation columns assume a strong role for recurrent interactions in V1 [BYBOS95]. Combined, this implies how information is processed in the visual system, and what information is gated out along the bottlenecks is determined to a large extent by recurrent processing in local networks within different areas.

Finally, evidence suggests that both information gating by bottlenecks as well as recurrent interactions are dynamically modulated by top-down processes such as selective attention [VEOA+91]. For example, it has been shown that neurons in downstream areas along the ventral pathway, which are presented with two distinct stimuli within their receptive fields with attention directed to only one of them, may respond almost as if only the attended one was present [MD85]. A potent functional

role is thus attributable to selective visual attention in routing of information through the bottlenecks along the visual pathway. The role of attention in real-time information gating and associated mechanisms will be described in detail in the next sections.

Taken together, we have, on the one hand potentially rigid anatomical connections which points towards the formation of robust, complex and invariant internal representations of the visual world through bottlenecks of purely feedforward processing. On the other hand, crucial components of visual information processing depend critically on recurrent interactions, and also, need to be rapidly modulated by task demands and behavioural goals. We hypothesize that modulations of activity in a single area can efficiently take place, and relevant information can robustly move downstream through the bottlenecks, by flexible variations of correlated activity in local networks. In line with the maximized dynamic range suggested for critical networks, we believe the ability to engage a very large number of neurons in dependence on external evidence and behavioral context, and to do this very quickly and flexibly, forms the backbone of a promising ‘application’ for critical networks.

In the next two sections, we will give a broad overview of how feature integration dynamically takes place in the bottom-up direction of information processing, leading to the detection and recognition of global percepts, such as objects and faces, in a continuously changing visual world. We will then talk about how this information processing may be modulated in the reverse direction, by top-down mechanisms such as selective attention, and attempt to build a comprehensive framework for how, independent of the processing direction or specific mechanisms, the visual system may benefit from the local emergence of critical dynamics.

4.2 Feature integration

Physiological studies have established that, apart from the relatively local processing in the retina and LGN, information integration is mainly performed in the cortical domain. As explained above, single neurons in the primary visual cortex (V1) preferentially respond to certain basic features of a local image patch, such as its orientation or spatial frequency [DA01; HW68]. Since these neurons thus signal the presence or absence of specific features in a visual scene, information integration in vision is also termed feature integration. In general, perceptual grouping of visual features take place based on classical Gestalt principles, where factors such as proximity, similarity or good continuation of features determine whether they will collectively be processed as belonging together [WEK+12], thus leading to a salient percept. Salient figures in a cluttered scene are robustly and easily segregated from the background, in what is termed a “pop-out” effect [KJ93]. For example in contour integration, fragmented outlines of potential object boundaries are bound into coherent percepts [FHH93]. Such enhanced processing of features due to stimulus properties is often described as a bottom-up effect [ZZZ+12].

So what might be the neuronal correlate of feature integration? In many different settings, feature integration is accompanied by firing rate modulations (e.g.,

see [CYG+14; PS93; Roe06]). Neurons which are stimulated with their preferred feature centered inside their classical receptive field (cRF) have been shown to exhibit lower or higher firing rates in dependence on a surrounding, contextual stimulus outside their cRF [LL97; SGJ+95]—in particular when the center-surround configuration matches Gestalt principles such as colinearity of oriented line segments [LPG06]. The caveat of these rate modulations serving as an explanation for feature integration is both their magnitude and their temporal characteristics: for example, in contour integration the rate enhancement is often small [BH02], and occurs with a large latency (150 ms) after stimulus onset [GMS13]. In contrast, other experiments show that human image classification can be extremely rapid, leading to consistent differences in EEG traces in an animal/no-animal discrimination task (where contour integration would only be one of the first steps in a chain of multiple processing stages) after only 150 ms [TFM96].

As an alternative, information integration might rely on the relative timing of action potentials: Spatially and temporally correlated neuronal activity has long been hypothesized as a putative mechanism for representing parts or features of an object which belong together [Heb49; vdM94]. In general, mutual synchronization between two neurons tends to become stronger if the stimulus components within their receptive fields are more likely to belong to one object [KS96]. Models using oscillatory synchrony for binding features have been proposed [FK14; RGF+94], but establishing oscillations takes time and the emerging neural activity is too regular in comparison to synchronization observed in physiological studies of the visual system [SK93].

4.3 Attention and gamma oscillations

To address the problem that a visual scene typically contains a substantial amount of information at any given time, the visual system dynamically and flexibly allocates its limited resources, in a context and task-dependent manner. For example, attentional mechanisms may selectively enhance the processing of information about specific features or locations in the visual field. Electrophysiological studies have shown that rate modulations can gradually boost the representation of behaviourally relevant stimuli at the expense of irrelevant stimuli (e.g., see [LCH+97; MD85; TM96]). In contrast to feature integration, mechanisms of selective attention such as the enhancement of stimulus representations and the preferential routing of information are top-down effects [The10].

Other than spiking rates, another important feature of electrophysiological recordings is the low-pass filtered components of the recorded electrical activity. Termed the local field potential (LFP), such low frequency activity is thought to be mainly generated through the integration of synaptic transmembrane currents within the postsynaptic neurons and is therefore representative of oscillatory activity [BAK12]. Starting from the early days of EEG research, such oscillatory components of brain activity has been classified into different frequency bands [Ber29], including the so-called γ -band, which is often reported to be in the range 30–150 Hz (see also

the introduction and discussion of chapter 5 for more information on LFPs and γ -oscillations).

Selective visual attention and the accompanying enhancements in perception and behaviour have been experimentally linked to an increase in oscillatory cortical activity in the γ -band [FRR+01; TMF+05]. In this context, γ -oscillations have been proposed to be the essential mechanism for information routing regulated by attention [Fri05a; GNM+12a]. Such oscillation based modulations provide the cortex with an effective mechanism capable of efficiently gating information flows [HEP15] which would be difficult to achieve with the often much weaker rate modulations [GRN+18].

In a similar line of thought, criticality describes the border between asynchronous and substantially correlated dynamics, and in the field of vision research, synchronization has been studied extensively as a putative mechanism for information processing [vdM94]. Experimental studies demonstrated that in early visual areas, oscillations in the γ -range occur during processing of a visual stimulus [EBJ+88; GS89]. Hereby mutual synchronization between two neurons tends to become stronger if the stimulus components within their receptive fields are more likely to belong to one object [KS96], thus potentially supporting feature integration. Moreover, recent studies have demonstrated links between synchronized activity in the form of oscillations in MEG [PHM+12] and LFP recordings [GP08] and in the form of neuronal avalanches.

In this thesis, we are going to focus on the general idea that near-critical dynamics may promote rapid and robust processing of dynamically modulated information streams. In order not to be constrained by the specifics of a certain mechanism, we are going to discuss two distinct examples, the first a top-down enhancement of object discriminability under selective attention (Part II) and the second a bottom-up figure-ground segregation scenario (Part III) (for an overview of the two approaches see Fig. 4.2). We are going to argue that regardless of the specific mechanism, local emergence of critical dynamics in the active brain, either in a stimulus driven manner or through attentional modulations, may assist in selectively improved processing of local information. This is broadly achieved by the activity of local subnetworks, representing different objects in the visual field, being pushed towards a critical point and away from a subcritical regime (leading to the emergence or an amplification of oscillations in the γ -band through mutual synchrony in the functionally connected subnetworks, Fig. 4.2). This framework brings together the concepts of oscillations in population activity, synchronization at the spiking level and phase transitions, and suggests a universal functional role for correlated activity: flexible processing of information in a dynamically changing world.

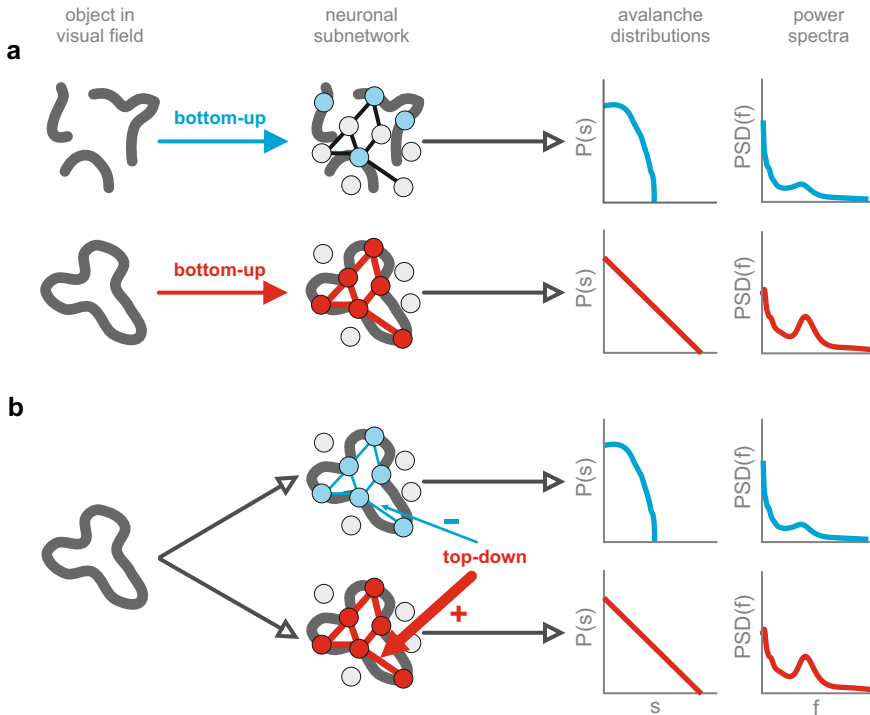
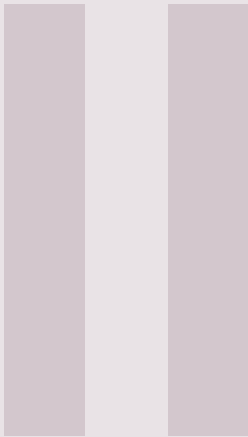


Figure 4.2: Critical dynamics exhibited by local cortical subnetworks may be instrumental in flexible modulations of object representation. **a** Bottom-up modulation of the neuronal response: A complete set of features which constitute an object in the visual field is a more salient stimulus than the random co-occurrence of a few of its features. A subnetwork poised at the critical point will display scale-free dynamics over the full system size, only if all of its units are externally driven (red circles and links). Similarly, power spectral distribution of local population activity will display stronger oscillatory dynamics when the stimulus is more salient. **b** Top-down modulation of the neuronal response. According to changing behavioural needs, mechanisms such as selective visual attention may dynamically shift the subnetwork towards or away from the critical point. A shift towards the ordered regime is accompanied by enhanced oscillatory power, potentially assisting in enhancing the communication between different areas. (Image adapted from the book chapter by the author and Udo Ernst, titled "The role of criticality in flexible visual information processing", as featured in the book "The Functional Role of Critical Dynamics in Neural Systems", Springer-Verlag, 2019.)



Attention and Tuning to Criticality

5	Attentional Modulations of Object Representation	47
5.1	Introduction	
5.2	Results	
5.3	Methods	
5.4	Discussion	



5. Attentional Modulations of Object Representation

In this part, we will talk about how operating close to a phase transition may benefit top-down modulations of information processing. In a specific example, we will focus on how the experimentally observed phenomena accompanying visual selective attention may be reproduced in a simple network model poised close to a transition boundary. We will also discuss how such an operational regime may support rapid and flexible enhancement of stimulus representation, as well as represent the information entropy maxima given certain read-out restrictions. All the figures and text in this chapter have been adapted from [TRE14].

5.1 Introduction

In this chapter we are going to explore potential links between synchronization, cortical information processing, attentional modulations of oscillatory states and criticality in the underlying networks in the visual system. In particular, we will investigate the criticality hypothesis in the context of γ -oscillations induced by selective visual attention. We will posit that if visual cortical networks indeed assume a critical state in order to optimize information processing, such a state should be prominent during the processing of an attended stimulus, since attention is known to improve perception [Car11] and enhance stimulus representations [RTE+09]. In particular, we focus on recent findings that local field potentials (LFPs) in macaque area V4 demonstrate an increase in γ -band synchrony and a simultaneous enhancement of object representation with attention.

Specifically, we will explore a structurally simple network model of leaky integrate-and-fire neurons for population activity in visual area V4. We will first demonstrate that our model reproduces key dynamical features of cortical activation patterns including the increase in γ -oscillations under attention observed in experiments [FRR+01; TMF+05]. In particular, we will explain how attention enhances the

representation of visual stimuli, thus allowing to classify the brain state corresponding to a particular stimulus with higher accuracy [RTE+09], and we will identify mutual synchronization as the key mechanism underlying this effect.

Construction of this model allowed us to analyze dependencies between network states and stimulus processing in a parametric way. In particular, we were interested in whether such a network displayed critical dynamics, and how they relate to cognitive states. We inquired: Is criticality a ‘ground state’ of the cortex which is assumed in the absence of stimuli, and helps process information in the most efficient way as soon as a stimulus is presented? Or is the cortex rather driven towards a critical state only when there is a demand for particularly enhanced processing, such as when a stimulus is attended?

For answering these questions, we (a) characterized the network state based on neuronal avalanche statistics (subcritical, critical, or supercritical), (b) quantified stimulus discriminability, and (c) analyzed the richness of the dynamics (information entropy of spike patterns) in the two-dimensional phase space spanned by excitatory and inhibitory coupling strengths. Within this coupling space, we identified a transition region where the network undergoes a phase transition from subcritical to supercritical dynamics for different stimuli. We found that the onset of γ -band synchrony within the transition region is accompanied by a dramatic increase in discriminability. At supercritical states epileptic activity emerged, thus indicating an unphysiological regime, and both information entropy and discriminability values exhibited a sharp decline.

Our main finding is that cortical networks operating at marginally subcritical states provide the best explanation for the experimental data [FRR+01; RTE+09; TMF+05]. At such states, small modulations of network excitability are sufficient for significant increases in discriminability. This establishes a concrete link between flexible modulation of visual information processing via selective attention, and neuronal networks poised close to a critical point, and we suggest an emergent functional role for critical dynamics in efficient top-down modulations of neural activity.

5.1.1 Motivation and previous experiments

As described in section 4.3, selective visual attentional mechanisms are associated with the emergence of γ -oscillations in visuocortical areas. In addition, attention modulated increases in γ -band power have also been observed accompanying increases in spike-field coherence [FRR+01], supporting our hypothesis that cascades of rapid spiking activity may play a role in mechanisms underlying selective visual attention. In particular, a top-down modulation of spike time correlation seems to be necessary to dynamically resolve which information will be gated in and which will be gated out by downstream visual areas [Fri05a; GRN+18].

In this chapter we will probe in which ways improved synchrony in the γ -band under selective attention may contribute to an improvement in information processing in the visual cortex, and how such an improvement may coincide with a phase transition boundary in system dynamics.

Our model is motivated by an electrophysiology experiment [TMF+05] where a rhesus monkey (*Macaca mulatta*) had to perform a demanding delayed-match-to-sample task on one of two simultaneously presented, dynamic visual stimuli. These stimuli consisted of morphing shape sequences covering about $4^\circ \times 4^\circ$ of visual angle in the left and right visual hemifields (for an example see Fig. 5.1). The monkey was instructed to attend to one of the stimuli, memorize the initial shape, and release a lever as soon as the initial shape reappeared in the morphing sequence. In parallel to this task, local field potential (LFP) recordings were performed with an epidural multielectrode array implanted over visual areas V1 and V4.

For three reasons, this experiment is interesting for our framework: First, recognizing the target shape required the monkey's visual system to integrate local information into a global percept. This had to be done rapidly and continuously, since the shapes were morphing and the target could appear at different times during a trial. Second, neural activity was recorded in an area (V4) which is believed to be central for shape perception, and with a method (epidural recordings) where a single electrode has access to a large population of neurons and is especially sensitive to large synchronous events. Third, the task was cognitively demanding and was not possible to perform without directing attention towards the behaviourally relevant stimulus.

Indeed the power spectra of the wavelet-transformed LFPs exhibited strong synchronization with a characteristic peak at γ -range frequencies between 35 and 80 Hz superimposed on a $1/f$ offset (Figure 5.2A). The amplitude of the γ -oscillations were visibly larger when the stimulus inside the RF of the corresponding neural population was attended [TMF+05]. Although LFP recordings have a coarse spatial resolution, the γ -activity proved to be specific enough to successfully be employed in discriminating between different shapes in the RF [RTE+09].

For assessing stimulus discriminability, Rotermund et al. (2009) used a support vector machine (SVM) on the LFP spectral power distributions in order to classify the stimuli on a single trial basis [RTE+09]. A total of six different visual stimuli (complex shapes) were used in the experiments, therefore, the chance level was close to 17%. This analysis yielded two results which are central for the model we will present in this chapter:

1. Stimulus classification performance was significantly above chance level even in the absence of attention (35.5% for the V4 electrode with maximum classification performance).
2. Under attention, differences in synchronous activity between shapes were enhanced, hence improving classification performance of different visual stimuli by an SVM significantly (by 6.7% for the V4 electrode with maximum classification performance) when the monkey attended the stimulus inside the receptive field (RF) of the recorded neuronal population.

This prompted us to ask: can these observations of neural representations of complex shapes and their modulation by selective attentional mechanisms be linked to correlated activity in the form of neuronal avalanches?

In this chapter we reproduce these results using a minimalistic model of leaky

integrate-and-fire (LIF) neurons, which allows us to investigate putative neuronal avalanches underlying the observed data, and in which attention increases synchrony by enhancing the efficacy of recurrent interactions. In the phase space spanned by excitatory and inhibitory coupling strengths, we identify critical points and regions of enhanced discriminability. Furthermore, we quantify encoding capacity using information entropy. We find a rapid enhancement of stimulus discriminability with the emergence of synchrony in the network. Strikingly, only a narrow region in the phase space, at the transition from subcritical to supercritical dynamics, supports the experimentally observed discriminability increase. At the supercritical border of this transition region, information entropy decreases drastically as synchrony sets in. At the subcritical border, entropy is maximized under the assumption of a coarse observation scale. Our results suggest that cortical networks operate at such near-critical states, allowing minimal attentional modulations of network excitability to substantially augment stimulus representation in the LFPs.

5.2 Results

5.2.1 Reproduction of key experimental findings

The spectra recorded in the experiment are consistent with neural dynamics comprising irregular spiking activity (the $1/f$ -background) and oscillatory, synchronized activity in the γ -band. In order to realize such dynamics in a structurally simple framework, we considered a recurrently coupled network of leaky integrate-and-fire (LIF) neurons which is driven by Poisson spike trains. We assume that the epidural recording electrode averages over a large population of N neurons. Each of these neurons can be, part of a subnetwork specialized in the processing of a particular shape. Presentation of a visual stimulus will activate a subset of $N_{active} < N$ neurons (Fig. 5.1A) responding to the presence of the particular features contained in a specific shape, while the neurons representing features not contained in the current shape will remain silent. The network consists of both excitatory and inhibitory neurons interacting via a sparse, random coupling matrix with a uniform probability of a connection between two neurons (for details see section 5.3.1). The most important parameters determining the network's dynamics are the synaptic input strengths J_{inh} and J_{exc} of inhibitory and excitatory recurrent couplings, which are homogeneous for a given network. While oscillatory activity is generated as a consequence of the recurrent excitatory interactions, the stochastic external input and inhibitory couplings induce irregular spiking, thus providing a source for the observed background activity.

We consider this network as a simplified model of a neuronal population represented in LFP recordings of area V4 and the external Poisson input as originating from lower visual areas such as V1. One specific visual stimulus activates only a subset of V4 neurons by providing them with a strong external drive while the remaining V4 neurons receive no such input (Figure 5.1A). We drove a different, but equally sized subset of V4 neurons for each stimulus. Hence in a recording of summed population activity (e.g. LFPs), where the identity of activated neurons is

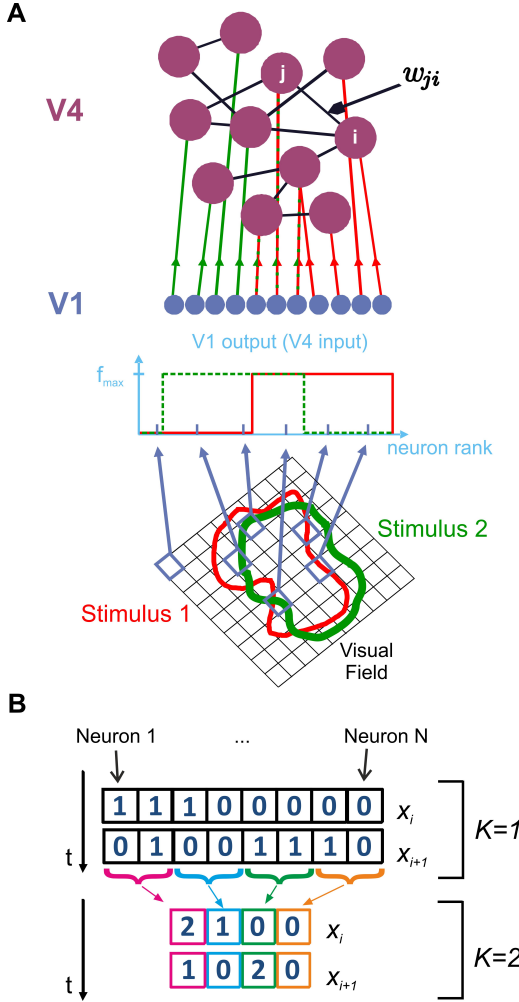


Figure 5.1: V4 populations are modelled as randomly coupled recurrent networks of mixed excitatory (80%) and inhibitory (20%) LIF neurons. **A** Depending on their receptive field properties, a different set of V1 neurons are externally activated by different stimuli. Activated V1 neurons provide feedforward input to V4 neurons j in the form of Poisson spike trains with rate f_{max} . Consequently, a different, randomly coupled, subnetwork of V4 neurons are driven by external input for each stimulus. Recurrent connections within V4 are represented by the random, non-symmetric coupling matrix w_{ji} . Finally, action potentials generated in V4 are observed by an (epidural) electrode, yielding local field potentials (LFPs) modeled as a sum over the pre- and postsynaptic network activity convolved with an exponentially decaying kernel. **B** Information entropy of the spike patterns generated by area V4 is calculated using state variables x_i . At the finest observation scale ($K = 1$), x_i consist of N -dimensional binary vectors, which represent whether each neuron j fired a spike (1) or not (0) at a given point in time. For larger K , the activity of K adjacent cells is summed to construct x_i .

lost, stimulus identity is represented in the particular connectivity structure of the activated V4 subnetwork. We simulated a total of $N = 2500$ neurons but kept the number of activated V4 neurons fixed at $N_{active} = 1000$ since every stimulus in the experiment was approximately the same size. With this setup, we ensured that the emerging stimulus-dependent differences in the network output are a consequence of stimulus identity and not of stimulus amplitude.

The variability of the couplings in our network mimics the structure of cortical

couplings, which are believed to enhance certain elementary feature combinations (such as edge elements aligned to the populations' RF features [KTR+97]) while suppressing others. Consequently, there will be stimuli activating subsets of V4 neurons which are strongly interconnected, while other stimuli will activate subsets which are more weakly connected.

We simulated the network's dynamics in response to $N_a = 6$ different stimuli in $N_{tr} = 20$ independent trials. Comparable to the experiments, LFP signals were generated by low-pass filtering the summed pre- and postsynaptic V4 activity (section 5.3.1). We computed the spectral power distributions using the wavelet transform of LFP time series.

As one would expect, for combinations of sufficiently large J_{exc} , and sufficiently small J_{inh} the neurons in the V4 population are mutually synchronized. The power spectra display a characteristic peak at γ -range frequencies, which grows with enhanced synchrony, as well as a $1/f$ offset, characteristic of neuronal background activity. The average frequency of the emergent oscillations depends mainly on the membrane time constant τ for the particular choice of external input strength.

Averaged over trials, these power spectra reproduced all the principal features displayed by the experimental data (Figure 5.2). In particular, spectra for individual stimuli differed visibly, with largest variability observed in the γ -range. Since the identity of activated neurons is lost in the population average, any differences in strength of the observed γ -oscillations can only be attributed to subnetwork connectivity. This result has a natural explanation because connection strength and topology strongly determine synchronization properties in networks of coupled oscillatory units (see for example [GDGL+00; LFHC+00; NML+03]). In other words, due to our choice of random connectivity between the different neuronal units activated, according to their receptive field properties, by different stimuli in the visual field, the stimulus-specific subnetworks display coupling structures which slightly differ in graph theoretical measures. This leads to slight, but discriminable differences in the frequency components between the dynamics of each subnetwork.

Differences in power spectra become even more pronounced if a stimulus is attended. We modeled attention by globally enhancing excitability in the V4 population by multiplicative modulation of the gain of all neurons. This can be realized either by increasing the efficacy of excitatory interactions, or by decreasing efficacy of inhibition. In this way, the gain of the V4 neurons is increased [BFL+10; FRR+01; RPD00; Tre01], and synchronization in the γ -range gets stronger and more diverse for different stimuli while the $1/f$ -background remains largely unaffected (Figure 5.2B). For visualizing the effect of attention, single trial LFP signals corresponding to attended and non-attended conditions for a specific stimulus are given in Figure 5.2C. Note that the change induced by attention does not need to be large; in the example in Figure 5.2 inhibitory efficacy was reduced by 10% from $j_{inh} = 0.80$ to 0.72.

The observed changes in the power spectra with attention can be interpreted in terms of the underlying recurrent network dynamics: each activated subnetwork has a particular composition of oscillatory modes, and enhancing excitability in such a non-linear system will activate a larger subset of these modes more strongly. This

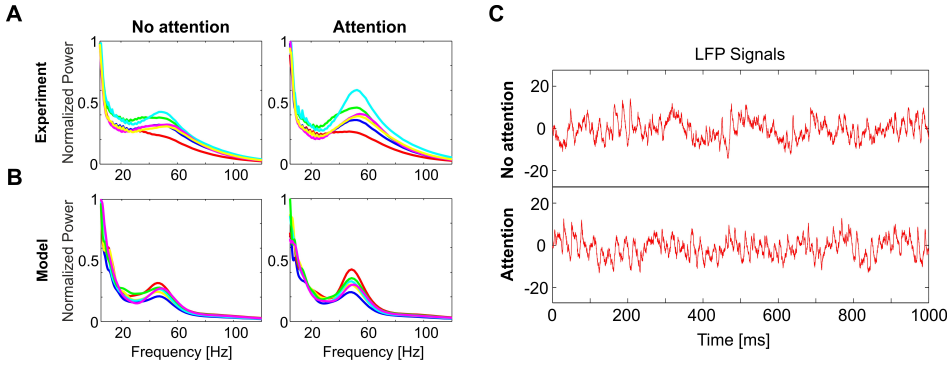


Figure 5.2: Comparison of model dynamics to experimental recordings. Trial-averaged spectral power distributions of the LFP signals in **A** the experiment and **B** the model for non-attended (left) and attended (right) conditions. In each case, the spectra recorded during presentation of 6 different shapes is shown (indicated by different colors). In both **A** and **B** the spectra for each stimulus is normalized to its respective maximum in the non-attended case. Model spectra reproduce the stereotypical $1/f$ background as well as the γ -peaks observed in the experimental spectra. Under attention, γ -band oscillations become more prominent and spectra for different stimuli become visibly more discriminable. **C** Single trial LFP time series from the model, illustrating the analyzed signals in the non-attended (top) and attended (bottom) conditions. The neuronal gain is increased in the model in order to model the enhanced synchrony in the attended condition. (Experimental data shown in **A** is courtesy of Dr. Andreas Kreiter and Dr. Sunita Mandon and Katja Taylor [TMF+05].)

effect is enhanced by synchronization emerging at different coupling strengths for different stimuli. With a further increase in the coupling, however, groups of neurons oscillating at different frequencies will become synchronized at a single frequency (Arnold tongues [CB99b]), which ultimately decreases the diversity of power spectra.

5.2.2 Enhancement of stimulus discriminability

The spectra in Figure 5.2B were generated using coupling parameters J_{exc} and J_{inh} specifically tuned for reproducing the experimental data. However, the basic phenomenon is robust against large changes in the parameters: Discriminability increase is coupled to the emergence of strong γ -oscillations. To show this, we varied the excitatory and inhibitory coupling strengths independently, and quantified stimulus discriminability in the LFP spectra from short, single trials, using SVM classification for every parameter combination. We define J_{exc} as the absolute strength of an excitatory synapse and, when varying the inhibitory efficacies, we used a step size that is proportional to the excitatory efficacy: $J_{inh} = \varepsilon \cdot J_{exc} \cdot j_{inh}$ for every point in the coupling space where j_{inh} is the inhibitory scaling factor and ε is the ratio of the number of inhibitory neurons to the number of excitatory neurons.

We set the upper bound of excitation and the lower bound of inhibition so as to avoid unphysiologically high firing rates due to the activation of all neurons, including those that did not receive external input. Figure 5.3A shows the classification results in coupling space, averaged over $N_w = 5$ independently realized random connectivity architectures of the V4 network. The coupling values used for generating the spectra in Figure 5.2B are indicated by white markers. Classification performance is 24.2% in the non-attended (white cross) condition (significantly above chance level, $\sim 17\%$, via a one-tailed binomial test with $p < 0.005$) and 32.8% in the attended (white circle) condition. Notably, discriminability is significantly above chance level only in a bounded region of the parameter space. Within this region, relatively small increases in excitatory, or decreases in inhibitory coupling strengths lead to an acute discriminability enhancement.

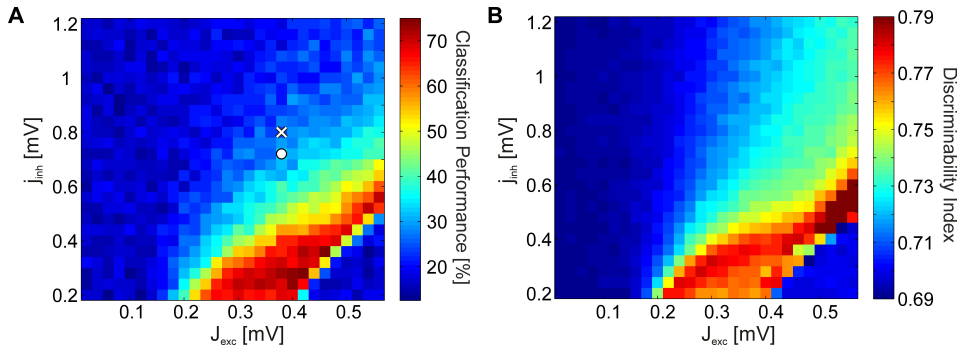


Figure 5.3: SVM test results and the discriminability index. (A) SVM classification performance as a function of the excitatory coupling strength J_{exc} and the inhibitory coupling scaling factor j_{inh} (obeying $J_{inh} = \varepsilon \cdot J_{exc} \cdot j_{inh}$). The coupling values representing the non-attended and attended conditions in Figure 5.2B are marked by a cross and a circle respectively. (B) Discriminability index in the coupling space for the same spectra. For both A and B, the strength of the background noise was $c_{mix} = 0.2$.

This effect comes about in the following way: In networks with low excitation and high inhibition, the dynamics are asynchronous and the LFP spectra are dominated by the $1/f$ -noise. In this case, every stimulus input is mapped to a network output with similar spectral components and with a large trial-to-trial variance. This severely impedes the ability to classify stimuli correctly. On the other hand, in networks with very high excitation and low inhibition, synchronous activity dominates the dynamics and epileptic behavior is observed. Mutual synchronization of the activated V4 neurons leads to co-activation of the otherwise silent V4 neurons which do not receive external input. This means that every stimulus input is mapped to spike patterns where almost all neurons are simultaneously active at all times. The corresponding spectra have reduced trial-to-trial variability but are almost identical for different stimuli. Consequently, stimulus discriminability reaches a maximum only in a narrow region of the parameter space which is associated with the onset of

synchrony.

It is necessary to point out that the absolute magnitude of the SVM performance depends strongly on the background noise (i.e. on the value of c_{mix}) which constitutes the $1/f$ -background in the spectra. For example, without the addition of the background noise (i.e. $c_{mix} = 0$), SVM classification performance is 36.67% for the non-attended and 43.83% for the attended spectra in Figure 5.2B. Nevertheless, the observation of a bounded region of enhanced discriminability persists even in the absence of $1/f$ -noise. This finding has an important consequence: It allows us to identify coupling parameters which cannot explain the experimental data regardless of the ‘real’ noise level. Thus, it outlines a specific working regime in which the model can reproduce both of the experimental findings described in section 5.1.1.

5.2.3 Characterization of dynamical network states

Our findings indicate that a significant discriminability increase correlates implicitly with the onset of synchronous dynamics. In the following, we will focus on this network effect in more detail, and investigate its ramifications for information processing in the visual system.

In order to obtain a better understanding of the behavior of the system, we implemented certain reductions to our simulations. First, we excluded regions in parameter space where all neurons not receiving external input became activated. For most of the phase space, recurrent excitation is not strong enough to activate these stimulus-nonspecific neurons. At the supercritical regions, where excitation is strong and neurons are firing synchronously, however, these silent neurons become activated. This effect further increases the average excitatory input strength in the recurrent V4 population, leading to epileptic activity at very high (biologically implausible) frequencies.

Such a regime would be highly unrealistic, since neurons in V4 populations have well-structured receptive fields and are only activated by specific stimuli [DHG06; DS87]. Therefore, we proceeded to isolate the activity of externally driven subnetworks and focused our analysis on their output. This was realized by limiting the number of neurons in the network to $N = N_{active} = 1000$ and by assigning different random coupling matrices to simulate different stimulus presentations. Thus, distinct network architectures stand for distinct stimulus identities.

When constructing the output signal, we now excluded the background noise induced by the V1 afferents (i.e. we set $c_{mix} = 0$), but note that the V4 neurons were still driven by this stochastic, Poisson input. This segregation of V4 activity from background noise was necessary for the analysis of network dynamics, in order to ensure that the observed variance of the LFP spectra across trials originated in the V4 population.

In the reduced simulations, spikes propagated and impacted the postsynaptic neurons’ membrane potentials instantaneously (see section 5.3.3). We also prevented neurons from firing twice during an avalanche. These latter changes were introduced for inspecting criticality in the system dynamics (as described in detail in chapters 2 and 3), allowing us to quantify the number of neurons involved in an avalanche

event accurately. This was necessary in order to establish a link between dynamical criticality in the brain and a putative functional role in object representation.

In order to make this link concrete, we introduce the discriminability index (DI) as a simplified measure of how well stimulus identity is represented in the synchronous activity in our model, since SVM classification is a comparatively indirect method for quantifying discriminability, employing classifiers which are difficult to interpret. The DI represents by how much, averaged over frequencies, the distributions of LFP spectra over trials overlap for each stimulus pair. In essence, it quantifies the discriminability of each stimulus based on the difference between the trial-averaged LFP spectra of different stimuli, normalized by the trial-to-trial variance of the spectra for each, individual stimulus (see section 5.3.2). This gives us a measure of how much the LFP spectra vary between presentations of different stimuli as compared to how much the spectra vary between different trials using the same stimulus. As oscillations emerge in network dynamics, trial-to-trial variability of the spectra decrease (i.e. width of the distributions become narrower), and the average spectra for each stimulus is more distinct (i.e. the means of the distributions disperse). Hence, DI provides us with a meaningful approximation of the SVM classification performance. We find that the DI yields a phase space portrait (Figure 5.3B) similar to the SVM classification result (Figure 5.3A) for the full network simulations.

In order to compute discriminability in the reduced simulations, we used $N_{tr} = 36$ trials from each of the $N_a = 20$ different stimuli. Simulations with the reduced network produce the same qualitative behavior in phase space (Figure 5.4A), in the sense that a dramatic discriminability increase is only observed in a narrow region in the recurrent coupling phase space, located in the border between regimes with and without strongly synchronous activity. Discriminability is maximized as oscillations emerge, and decays quickly in the regions where epileptic behavior is observed as all neurons fire simultaneously. Combined with the experimental evidence, our findings suggest that the cortex operates near a particular state where small modifications of excitability lead to substantial changes in its collective dynamics.

However, time-averaged power spectra of local field potentials are not well suited for characterizing different aspects of this state. Since epidural LFPs are signals averaged over large neuronal populations, dynamic features in spiking patterns become obscured, and temporal variations in the network dynamics are lost in the averaging process. In the following, we will go beyond LFPs and focus on (a) the size distribution of synchronized events (avalanche statistics), and (b) on the diversity and richness of patterns generated by the network (measured by information entropy).

5.2.4 Criticality of dynamics

The transition to oscillatory activity, as we've seen in LFP spectra, is well reflected in the avalanche dynamics of the network. The network dynamics can be classified into three distinct regimes of activity characterized by their avalanche size distributions: subcritical, critical, and supercritical (Figure 5.4B). In the subcritical state spiking activity is uncorrelated, events of large sizes are not present and the probability distributions $P(s)$ of observing an avalanche event of size s exhibit an exponential decay.

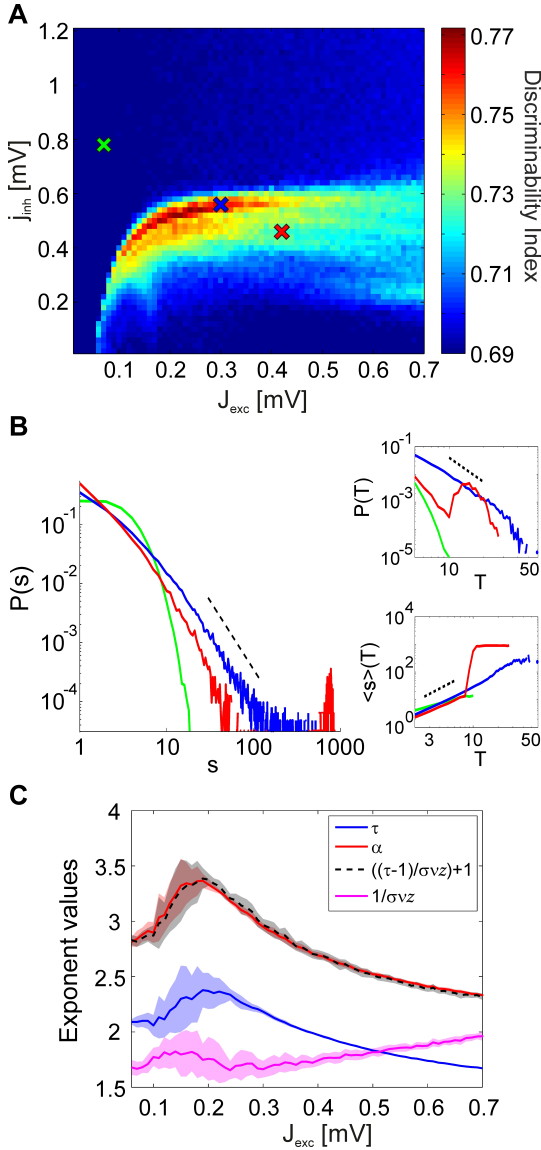


Figure 5.4: **Discriminability of the LFP spectra in relation to the avalanche statistics.** **A** Discriminability index based on the LFP spectra of each stimulus in the reduced simulations. As in the full simulations (Figure 5.3B), stimulus discriminability peaks rapidly in a narrow region of the coupling space. **B** Avalanche size distributions $P(s)$ for three different synaptic strength values in the subcritical (green), critical (blue) and supercritical (red) regimes for a single stimulus. Insets show how the corresponding avalanche duration distributions $P(T)$ and the mean avalanche sizes $\langle s \rangle$ conditioned on the avalanche duration T behave in the three distinct regimes. Corresponding coupling parameter values are marked with crosses of matching colors in **A**. **C** The values of the estimated power-law exponents τ , α and $1/\sigma v_z$ for each value of the excitatory coupling strength J_{exc} . The lines mark the mean exponent at the critical point for each stimulus and the corresponding colored patches represent the standard deviation over the stimuli. The black dashed line shows the value of α computed using equation 5.3, by plugging in the other two exponents.

In the supercritical state, spiking activity is strongly synchronous and avalanches spanning the whole system are observed frequently. This behavior is represented in the avalanche size distributions by a characteristic bump at large event sizes. The critical state signifies a phase transition from asynchronous to oscillatory activity and the corresponding avalanche size distributions $P(s)$ display scale-free behavior.

$$P(s) \propto s^{-\tau} \quad (5.1)$$

If we look at subnetworks with high inhibition and low excitation, we have exponential avalanche size distributions (green curve in Fig. 5.4B), signifying, on average, small avalanche sizes and low overall correlated activity. As we increase excitation or decrease inhibition in the subnetwork, we observe supercritical size distributions (red curve in Fig. 5.4B), characterized by a large bump at large avalanche sizes, signifying highly correlated activity. A similarly abrupt transition in dynamics is also apparent in the distributions of avalanche duration $P(T)$, and average avalanche size $P(S)$ for a given duration (Fig. 5.4B insets).

Even though power-law scaling of the avalanche size distributions, combined with the sudden emergence of oscillatory behavior in the system strongly suggest a phase transition in network dynamics, it is not sufficient to definitively conclude that the system is critical [BT12; FIB+12]. Therefore, for inspecting criticality in the network dynamics, we have investigated the behavior of two other, relevant avalanche statistics: the distribution $P(T)$ of avalanche durations T and the mean avalanche size $\langle s \rangle$ given the avalanche duration T , $\langle s \rangle(T)$. We find that both of these distributions follow a power-law for intermediate values of T at the critical points (Figure 5.4B, insets).

$$P(T) \propto T^{-\alpha} \quad (5.2)$$

$$\langle s \rangle(T) \propto T^{1/\sigma\nu_z} \quad (5.3)$$

We observe that the behavior of $P(T)$ within the phase space is similar to that of $P(s)$. In the subcritical regime, there are only avalanches of short durations, and $P(T)$ has a short tail. In the supercritical regime, $P(T)$ displays a bump at large event durations. For $\langle s \rangle(T)$, we observe scale-free behavior of the distributions in both subcritical and critical regimes. Again a bump appears for large T at the supercritical regimes. In order to quantify the power-law scaling of the avalanche size and duration distributions we applied a maximum-likelihood (ML) fitting procedure [CSN09] and obtained an ML estimation of the power-law exponent for every stimulus. We obtained the power-law exponent of the mean size distributions conditioned on the avalanche duration using a least squares fitting procedure [Wei02]. Notably, the exponents obtained from the simulated dynamics fulfill the exponent scaling relationship (Figure 5.4C)

$$\frac{\alpha - 1}{\tau - 1} = \frac{1}{\sigma\nu_z} \quad (5.4)$$

as predicted by universal scaling theory [FIB+12; SDM01].

We next identified where exactly in phase space model dynamics transition from asynchronous to regular oscillatory activity, using the avalanche distributions. As a goodness-of-fit measure for the avalanche size distributions, we employed the Kolmogorov–Smirnov (KS) statistic. The KS statistic D averaged over all stimuli (i.e., network architectures) is given in Figure 5.5A. However, for identifying points in the phase space at which the network dynamics are critical, the KS statistic is ineffective: Even in the transition region from subcritical to supercritical behavior, the avalanche size distributions rarely display a perfect power-law which extends from the smallest to the largest possible event size. Therefore, we introduced lower

and upper cut-off thresholds on the avalanche sizes during the fitting process (see section 5.3.3). While this procedure allowed us to do better fits, it also led to a large region of subcritical states which had relatively low (and noisy) D values. This presents a predicament for automatically and reliably detecting the critical points by searching for minima in the D -landscape. Furthermore, we found that avalanche size distributions become scale-free at different points in phase space for different stimuli (Figure 5.5B). Therefore, the minima of the average KS statistic in Figure 5.5A are not representative of the critical points of the system.

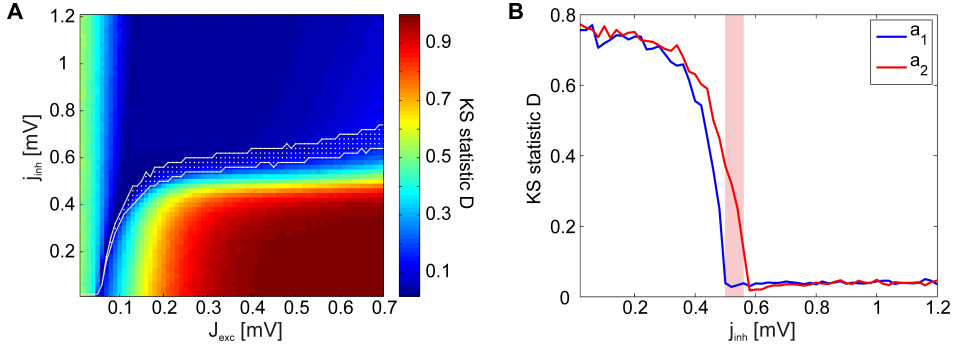


Figure 5.5: KS statistic as a measure of criticality. (A) KS statistic D of the avalanche size distributions in the reduced network, averaged over all stimulus presentations. Visual inspections reveal that the avalanche size distributions $P(s)$ are characteristically subcritical (exponential) for most points in the coupling space with low D values. The transition region calculated using the γ measure is given in white. (B) KS statistic D as a function of inhibitory coupling scaling factor J_{inh} for two exemplary stimuli, a_1 (blue) and a_2 (red), illustrating how the D minima occur at different points in the phase space for different stimuli ($J_{exc} = 0.2$ mV). The γ -transition region is given in magenta.

Visual inspections revealed that the subcritical avalanche size distributions converge slowly to a power-law as inhibition is decreased. At a critical value of inhibition, a phase transition occurs and the bump characteristic of supercritical distributions appears abruptly. Consequently, it is trivial to determine the transition regions graphically. We automatized this procedure by introducing a binary variable γ , which is a function of the relative weight of the tail of the size distributions close to the system size N . The γ -measure assumes a value of 1 if this relative weight exceeds a given threshold, in other words if a bump is detected in the avalanche size distributions (indicating supercritical dynamics), and 0 otherwise (indicating subcritical dynamics). The γ -measure averaged over all stimuli $\langle \gamma \rangle$ is given for each point in the phase space in Figure 5.6A, visualizing what percentage of the subnetworks have transitioned to the supercritical regime. We observe that there are clearly defined regions of sub- and supercritical dynamics, where γ is 0 or 1 for all stimuli respectively. The points for which $0 < \langle \gamma \rangle < 1$ define the transition region, where synchronization builds up rapidly for different stimuli. This shows that between excitation-dominated and

inhibition-dominated dynamics, there is a narrow transition region in which some subnetworks are already supercritical while others exhibit asynchronous dynamics.

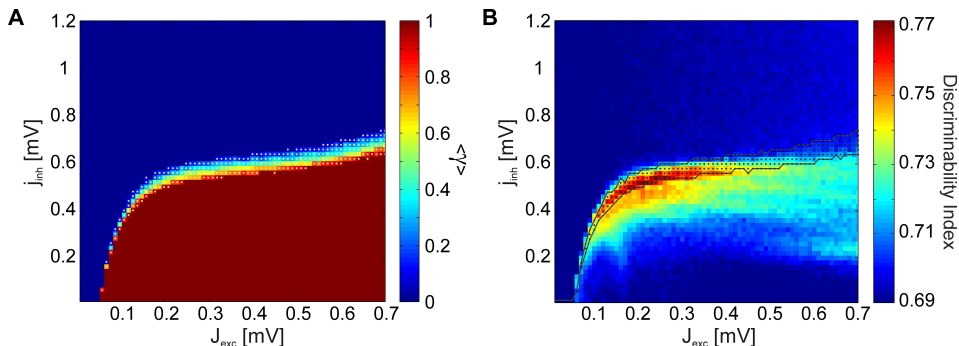


Figure 5.6: Discriminability is enhanced in the region defining a phase transition from subcritical to supercritical avalanche statistics. (A) γ measure averaged over all stimulus presentations as a function of the excitatory and inhibitory coupling strengths. Network dynamics are asynchronous (subcritical) for all subnetworks, representing all stimuli, in the regions of the phase space where $\langle \gamma \rangle = 0$, and strongly correlated (supercritical) in the regions where $\langle \gamma \rangle = 1$. A phase transition from subcritical to supercritical dynamics takes place between these two regions, at different points for different stimuli. This transition region where $0 < \langle \gamma \rangle < 1$ is indicated by white dots. (B) Comparison of the discriminability index from Figure 5.4A and the transition region (black lines and dots). The increase in stimulus discriminability agrees well with the transition region up to large values of the excitatory coupling.

In Figure 5.6B the transition region is plotted together with the discriminability index for comparison. We observe that the points at which discriminability is enhanced are confined to the neighborhood of the transition region. Discriminability is maximized within the transition region, where the network dynamics are supercritical for a subset of architectures and subcritical for the remaining ones. This means that if cortical neurons were to maximize discriminability, a set of stimulus inputs would effectively map to epileptic output activity. Such a scenario is not only physiologically implausible, but actually pathological. Taken together, these findings suggest that only marginally subcritical points, and not ones within the transition and supercritical regions, qualify for explaining the experimental observations.

These results suggest that networks operating close to the transition boundary are capable of rapidly enhancing stimulus representation in their dynamics, with only small modulations to the neuronal input gain. If true, this means that flexible tuning of information processing by attentional mechanisms is optimal when the network employs a ‘ground state’ close to or within the transition region. Therefore, we propose that the cortex operates at near-critical states, at the subcritical border of the transition region. Such near-critical states are unique in their ability to display significant discriminability enhancement under attention while avoiding

pathologically oscillatory dynamics.

However, these findings raise two important points to consider:

1. In our model it is so far not clear how or whether downstream visual areas can make use of the stimulus information contained in the LFP spectral power representation. This is because, neither our discriminability measure of the LFP spectra, nor the avalanche statistics consider putative, neurophysiologically plausible decoding schemes used by visuocortical areas.
2. While the slope of the stimulus discriminability index is favorable for rapid modulations of object representation at the boundaries of the transition region, its absolute magnitude is maximized within the transition region. However, within the transition region, a large number of subnetworks already display supercritical activity, which is strongly synchronous and exhibits a high degree of correlations. Such level of redundancy in spiking activity is associated with encoding limitations. Ideally, we would like to have a better understanding of the trade-off between enhanced stimulus discriminability, and the loss of information at the single spike level due to increasing correlations, within and around the transition boundaries.

To address this issue, we next inspected the diversity of spike patterns generated in the V4 network, and how this diversity behaves in the neighborhood of the transition region.

5.2.5 Information Entropy

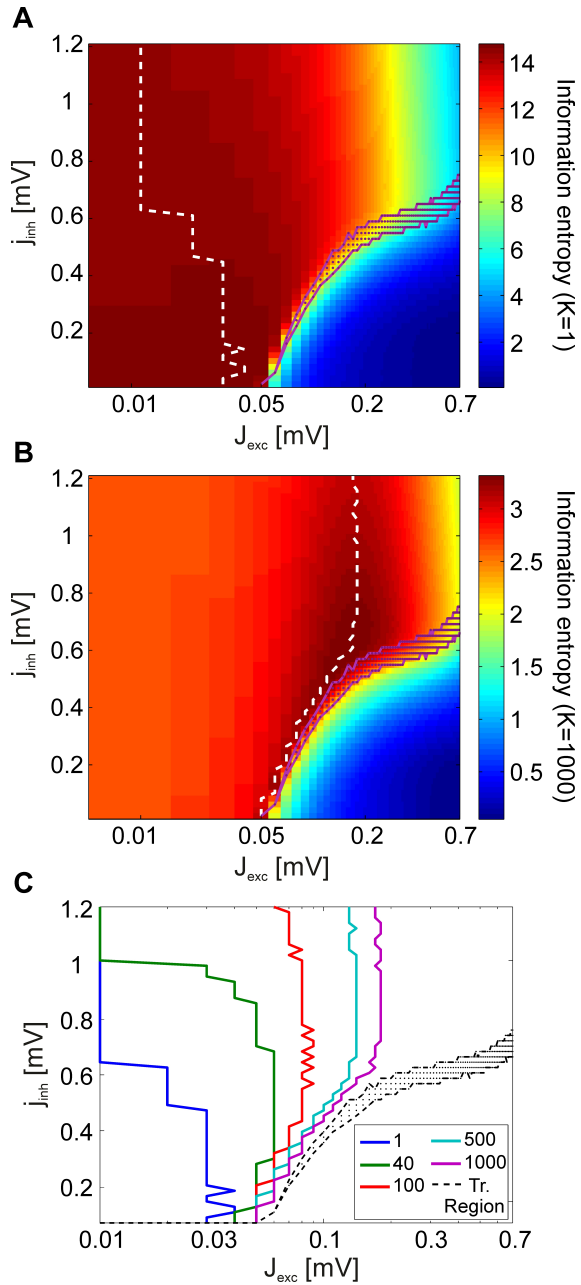
We computed information entropy [Sha48] in order to assess the diversity of V4 spike patterns generated in response to stimuli within the coupling space. In doing so, we considered different scales of observation on which read-out of these patterns, e.g. by neurons in visual areas downstream of V4, might take place.

At the finest scale of observation, parametrized by $K = 1$, the read-out mechanism has access to complete information about the spiking activity in the V4 subnetwork. In this case, it can discriminate between spikes originating from distinct presynaptic V4 neurons. At the coarsest observation scale, $K = N$, the read-out mechanism is not capable of observing every individual neuron, but rather integrates the total spiking activity of all neurons in the V4 subnetwork at each discrete simulation time step.

We then calculate the information entropy $H_K(\mathbf{X})$ of the spike patterns \mathbf{X} , where each state X_i is based on a binary word of length N formed by the spiking activity at a given time step, for each K . We define the observation scale parameter such that for $K = 1$, the state variable \mathbf{X} consists of N channels: each channel assumes a value of 1 if the corresponding neuron generated an action potential at time t , and 0 otherwise. The scale parameter K then reduces a spike pattern comprising spikes from N neurons to a representation of N/K channels with each channel containing the summed activity of K neurons (Fig. 5.1B).

Quantified this way, we find that information entropy declines drastically within the transition region for all K (Fig. 5.7). As we increase K , however, the entropy maxima in the phase space are shifted towards the transition region and to higher coupling strengths. Figures 5.7A-B show how information entropy compares with

Figure 5.7: Analysis of information entropy in V4 spike patterns. Information entropy of the spike patterns X_i in coupling space for the finest observation scale ($K = 1$) **A** and the coarsest observation scale ($K = N = 1000$) **B** averaged over all stimuli. At the finest observation scale ($K = 1$), X_i consist of N -dimensional binary vectors, which represent whether each neuron fired a spike (1) or not (0) at a given point in time. For larger K , the activity of K adjacent cells is summed to construct X_i . In **A** and **B**, the dashed white lines indicate the entropy maxima for each value of the inhibitory coupling scaling factor j_{inh} . The magenta dots mark the transition region as identified in Fig. 5.6A. **C** The maxima of information entropy for different observation scales K . Entropy maxima converge towards the transition region (black) as K is increased.



the transition region of the system for $K = 1$ (full representation) and for $K = N$ (summed activity over whole network). For each inhibitory coupling, the value of the excitatory coupling which maximizes information entropy is marked with a dashed line. For both conditions, we see that information entropy displays a sharp decline near the transition region. This behavior is consistent with a phase transition towards a regime of synchronous activity as the emergence of strong correlations attenuate entropy by severely limiting the maximum number of possible states. In comparison to the finest scale of observation ($K = 1$), we find that the maxima of information

entropy are shifted to greater values of excitation at the coarsest scale of observation ($K = N = 1000$). Figure 5.7C shows how the maxima of information entropy evolve as a function of observation scale K , converging onto near-critical points. This effect arises because, as K increases, the points with the greatest number of states in the network activity are shifted towards the transition region. By construction, the number of possible states of X is finite, and the uniform distribution has the maximum entropy among all the discrete distributions supported on the finite set $\{x_1, \dots, x_n\}$. Hence, information entropy of the spike patterns increases with both an increase in the number of observed states and an increase in the flatness of the probability mass function $P(\mathbf{X})$ of the states.

For the coarsest scale of observation, $P(\mathbf{X})$ is equivalent to the avalanche size distributions, and it is clear that a power-law scaling of these distributions cover the largest range of states (Figure 5.4B). However, for large j_{inh} ($j_{inh} \gtrsim 0.6$), entropy maxima persist at moderately subcritical regions. For large K , these regions are characterized by $P(\mathbf{X})$ with smaller supports but more uniform shapes than the $P(\mathbf{X})$ near the transition region. The flatness of these distributions, especially at small event sizes, causes the entropy maxima to appear around $J_{exc} = 1.8$ mV, instead of being located at higher values of excitation.

Combined, our results can be interpreted in the following way for the two extreme conditions discussed:

1. If neurons in higher areas of the visual system perform a spatial integration of the neuronal activity in the lower areas (K large), V4 networks operating at near-critical regimes both maximize information entropy and achieve significant discriminability enhancement under attention.
2. If V4 neurons employ a more efficient encoding strategy, where both spike times and neuron identities contain meaningful information for higher areas (K small), entropy is maximized by subcritical states with asynchronous dynamics. In such a scenario near-critical states represent a best-of-both-worlds optimization. At the subcritical border of the transition region, onset of oscillations and discriminability enhancement can manifest while avoiding a drastic loss in information entropy.

5.3 Methods

5.3.1 Network Model

Structure and Dynamics

The V4 network consists of N recurrently coupled leaky integrate-and-fire neurons $i = 1, \dots, N$ described by their membrane potential $V(t)$:

$$\begin{aligned} \tau_{mem} \frac{dV_i(t)}{dt} = & -(V_i(t) - V_R) + J_{ext} \sum_k \delta(t - t'_{ik}) \\ & + J_{exc} \sum_{j=1}^{N_{exc}} w_{ij} \sum_k \delta(t - t'_{jk}) - J_{inh} \sum_{j=N_{exc}+1}^N w_{ij} \sum_k \delta(t - t'_{jk}) \quad (5.5) \end{aligned}$$

The membrane potential evolves according to equation 5.5 where every V4 neuron i has a resting potential $V_R = -60$ mV and generates an action potential when V crosses a threshold $V_\theta = -50$ mV. After spiking, $V(t)$ is reset back to V_R . We picked the parameters to be representative of those of an average cortical neuron [KSJ+00; NSD+05]. We used a membrane time constant of $\tau_{mem} = 10$ ms. In equation 5.5, t'_{jk} denotes the timing of the k -th spike from V4 neuron j , and t'_{ik} the k -th spike from V1 (external input) to V4 neuron i .

V4 neurons are primarily driven by the external (feedforward) input once a stimulus is presented (see section 5.3.1). Presynaptic V1 spikes have an external input strength $J_{ext} = 0.1$ mV.

N_{inh} V4 neurons are inhibitory (interneurons) and the remaining N_{exc} are excitatory cells (pyramidal neurons). We assumed a fixed ratio of $\varepsilon = N_{exc}/N_{inh} = 4$ [Abe91]. The neurons are connected via a stimulus-specific random coupling matrix with connection probability $p = 0.02$ (Erdős-Rényi graph). Connections are directed (asymmetrical), and we allow for self-connectivity. w_{ij} assumes a value of 1 if a connection exists from neurons j to i , and is 0 otherwise. Global coupling strengths can independently be varied by changing J_{inh} and J_{exc} .

Simulations were performed with an Euler integration scheme using a time step of $\Delta t = 0.1$ ms. Membrane potentials of V4 neurons were initialized such that they would fire at random times (pulled from a uniform distribution) when isolated and driven by a constant input current (asynchronous state). We simulated the network's dynamics for a period of $T_{total} = 2.5$ seconds and discarded the first, transient 500 ms before analysis.

Stimulus and External Input

For comparison with the experimental data, we drove our network using N_a different stimuli. Specifically, we assumed that each stimulus activates a set of neurons in a lower visual area such as V1 or V2 whose receptive fields match (part of) the stimulus (Figure 5.1A). These neurons in turn provide feedforward input to a subset of N_{active} neurons in the V4 layer. We realized this input as independent homogeneous Poisson processes with rate $f_{max} = 10$ kHz. This situation is equivalent to each activated V4 neuron receiving feedforward input from roughly 1000 neurons, each firing at about 10 Hz during stimulus presentation.

Since stimuli used in the experiment had similar sizes, we assumed the subset of activated V4 neurons to have constant size $N_{active} = 1000$ for all stimuli. For each stimulus, we randomly choose the subset of V4 neurons which were activated by feedforward input. With a total of $N = 2500$ neurons, these subsets were not mutually exclusive for different stimuli. The remaining $N - N_{active}$ neurons received no feedforward input. Each stimulus was presented to the network in N_{tr} independent trials, and the simulations were repeated for N_w independent realizations of the V4 architecture w_{ij} .

With a finite N_{active} , such randomly generated couplings yield different graph theoretical parameters, such as mean in-degrees, mean out-degrees, mean betweenness centrality and so on, for each subnetwork which is activated during a stimulus

presentation. The varying properties of the different subnetworks give rise to different dynamics, and are responsible for different stimuli producing different LFP spectra in our model (Fig. 5.2B).

Local Field Potentials (LFPs)

In the experiments motivating this work, spiking activity was not directly observable. Only neural population activities (LFPs) were measured by epidural electrodes. Likewise, using our model we generated LFP signals $U(t)$ by a linear superposition of spiking activities of all neurons j in layer V4 and spiking activities of V1 neurons presynaptic to V4 neurons i , scaled by a mixing constant of $c_{mix} = 0.2$. This was followed by a convolution with an exponential kernel K_{exp} (low-pass filter). In our network, this is a close approximation of summing the postsynaptic transmembrane currents of the V4 neurons [BAK12; LTP+11; MOK+14].

$$U(t) = K_{exp}(t, \tau_k) \otimes \left(\sum_{jk} \delta(t - t_{jk}) + c_{mix} \sum_{ik} \delta(t - t'_{ik}) \right) \quad (5.6)$$

$$K_{exp}(t, \tau_k) = \frac{1}{\tau_k} e^{-t/\tau_k}. \quad (5.7)$$

We used a time constant of $\tau_k = 15$ ms for the kernel and discarded a period of 50 ms ($\sim 3.3 \tau_k$) from both ends of the LFP signal in order to avoid boundary effects.

5.3.2 Analysis of Network Dynamics

Spectral analysis

Mirroring the experiments, we performed a wavelet transform using complex Morlet's wavelets $\psi(t, f)$ [KMMG87] for time-frequency analysis. We obtained the spectral power of the LFPs via

$$p(t, f) = \left| \int_{-\infty}^{+\infty} \psi(\tau_w, f) U(t - \tau_w) d\tau_w \right|^2. \quad (5.8)$$

In order to exclude boundary effects, we only took wavelet coefficients outside the cone-of-influence [TC98]. Finally, we averaged the power $p(t, f)$ over time to obtain the frequency spectra $\bar{p}(f)$. This method is identical to the one used for the analysis of the experimental data [RTE+09]. The power $p(t, f)$ of the signal was calculated in $N_f = 20$ different, logarithmically spaced frequencies f , in the range from $f_{min} = 5$ Hz to $f_{max} = 200$ Hz.

Support Vector Machine classification

In order to assess the enhancement of stimulus representation in the LFPs, we performed SVM classification using the libsvm package [CL11]. The SVM employed a linear kernel function and the quadratic programming method to find the separating hyperplanes. We implemented a leave-one-out routine, where we averaged over N_{tr} results obtained by using $N_{tr} - 1$ randomly selected trials for each stimulus for training and the remaining trial for testing.

Discriminability Index

The discriminability index $DI(J_{exc}, j_{inh})$ was defined as

$$DI = \frac{1}{N_a(N_a - 1)/2} \frac{1}{N_f} \frac{1}{N_{tr}} \sum_{i=1}^{N_a-1} \sum_{j=i+1}^{N_a} \sum_f \sum_{tr} \frac{\text{erf}(Z_{DI}(f, tr, i, j)/\sqrt{2})}{2} + \frac{1}{2} \quad (5.9)$$

with

$$Z_{DI}(f, tr, i, j) = \frac{|\bar{p}_i(f, tr) - \bar{p}_j(f, tr)|}{\sigma_{tr}(\bar{p}_i(f, tr)) + \sigma_{tr}(\bar{p}_j(f, tr))} \quad (5.10)$$

where σ_{tr} is the standard deviation of frequency spectra \bar{p} over different trials tr and $\text{erf}(\cdot)$ is the error function. The assumption underlying the DI measure is that, at a given frequency f , the magnitude of the LFP power distribution for different trials tr is normally distributed. Discriminability of two stimuli thus depend on how much the areas under their corresponding distributions overlap. DI represents the mean pairwise discriminability of unique stimulus pairs $\{i, j\}$, averaged over frequencies and trials. For one particular frequency band, the DI measure is related to the area-under-the-curve of a receiver-operator-characteristic of two normal distributions. By this definition, DI is normalized between 0.5 and 1, a higher DI indicating better discriminability. Because of trials having a finite duration, however, DI has a bias which took an approximate value of 0.69 in our simulations (Figures 5.3B, 5.4A, 5.6B). In addition, since there are typically frequencies which carry no stimulus information (e.g. the 110 Hz-band, see Figure 5.2B), DI is confined to values smaller than 1.

The discriminability index was further averaged over N_w independent realizations of the coupling matrix in the full simulations. In the reduced model, we ran the simulations for an extended duration of $T_{total} = 12$ seconds. For computing DI, we then divided the LFP time series into $N_{tr} = 36$ trials.

5.3.3 Neuronal Avalanches

Separation of Time Scales

A neuronal avalanche is defined as the consecutive propagation of activity from one unit to the next in a system of coupled neurons. The size of a neuronal avalanche is equal to the total number of neurons that are involved in that avalanche event, which starts when a neuron fires, propagates through generations of postsynaptic neurons, and ends when no neurons are activated anymore. Avalanche duration is then defined as the number of generations of neurons an avalanche event propagated through. In such a system, the critical point is characterized by a scale-free distribution of avalanche sizes and durations.

In simulations assessing avalanche statistics, recurrent spikes were delivered instantaneously to all postsynaptic neurons for proper separation of two different avalanches. This means that as soon as an avalanche event started, action potentials were propagated to all the generations of postsynaptic spikes within the same time step, until the avalanche event ended. This corresponds to a separation of timescales between delivery of external input and avalanche dynamics. In this way we could

determine the avalanche sizes precisely, by ‘following’ the propagation of every spike through the network.

In addition, we implemented a basic form of refractoriness which prevented a neuron from firing more than once during an avalanche event (holding its membrane potential at V_R after it fired). Since each avalanche event took place in a single time step of the simulations, this corresponded to each neuron having an effective refractory period equivalent to the integration time step Δt .

Analysis of Criticality of Dynamics

For each network realization, we obtained the probability $P(s) \propto s^{-\tau}$ of observing an avalanche of size s by normalizing histograms of avalanche sizes. For every distribution $P(s)$ obtained from our simulations, we calculated a maximum-likelihood estimator $\hat{\tau}$ for the power-law exponent τ using the statistical analysis described in [CSN09] for discrete distributions. For a comprehensive account of the fitting method please see [CSN09]. To explain the procedure briefly, we started by defining a log-likelihood function $\mathcal{L}(\tau)$. This quantifies the likelihood that the n empirical avalanche size observations s_i ($i = 1, \dots, n$), which were recorded during our simulations, were drawn from a perfect power-law distribution with exponent τ .

$$\mathcal{L}(\tau) = -n \ln \zeta(\tau, s_{\min}) - \tau \sum_{i=1}^n \ln s_i \quad (5.11)$$

where

$$\zeta(\tau, s_{\min}) = \sum_{n=0}^{\infty} (n + s_{\min})^{-\tau} \quad (5.12)$$

is the Hurwitz zeta function. For a set of τ values in the interval $[1.1, 4]$, we computed $\mathcal{L}(\tau)$ (using equation 5.11) and the value of τ which maximized the log-likelihood was taken as the exponent $\hat{\tau}$ of the power-law fit $P_{\text{fit}}(s) \propto s^{-\hat{\tau}}$. During the fitting procedure, we used a lower cut-off threshold $s_{\min} = N/100 = 10$ and an upper cut-off threshold $s_{\max} = 0.6N = 600$. In other words, we fit a power-law to the set of empirical observations in the interval $s_{\min} \leq s_i \leq s_{\max}$. We repeated this fitting procedure to obtain power-law exponents α for the avalanche duration distributions $P(T) \propto T^{-\alpha}$, using $T_{\min} = 5$ and $T_{\max} = 30$.

For clarity, it is important to point out that the ML analysis described in [CSN09] does not take into consideration an upper cut-off in the empirical power-law distributions. One of the reasons we used an upper cut-off threshold during fitting is that the automated detection of critical points using the γ measure required us to fit a power-law exponent also to subcritical and supercritical avalanche size distributions. Using the complete tail of the distribution during the fitting procedure, for example in supercritical regimes, would yield a bias toward lower exponent estimates which would make it difficult to reliably detect the bump at large event sizes. This would hinder the detection of critical points using the γ measure, as it depends on an exponent which reliably represents the behavior of the distribution in the medium range of event sizes. More importantly, most of the size and duration distributions we observed at critical points displayed an exponential upper cut-off, as

also observed in other experimental and theoretical work [Beg08; BP03; DAH12; KYP11; PTL+09]. In statistics of neuronal avalanches, the exact location of the cut-off threshold depends strongly on system size and the duration of observations, and increasing either will increase the number of sampled avalanches and shift the cut-off threshold to higher values, but not make it vanish. In addition, excluding the observations above a cut-off threshold reduced the absolute magnitude of the log-likelihood function for all values of τ (equation 5.11), but the value of τ which maximized the log-likelihood provided us with a better estimate of the exponent for the middle range of the distributions where power-law scaling was prominent.

We used a least squares fitting procedure to find the power-law exponents for $\langle s \rangle(T)$ [Wei02], as it is not a probability distribution, using $T_{\min} = 2$ and $T_{\max} = 20$. In this procedure, the exponent $1/\sigma v_z$ of the function $\langle s \rangle(T) \propto T^{1/\sigma v_z}$ is given by the closed expression

$$\frac{1}{\sigma v_z} = \frac{m \sum_{i=1}^m (\ln T_i \ln \langle s \rangle_i) - \sum_{i=1}^m (\ln T_i) \sum_{i=1}^m (\ln \langle s \rangle_i)}{m \sum_{i=1}^m (\ln T_i)^2 - (\sum_{i=1}^m (\ln T_i))^2} \quad (5.13)$$

where m is the total number of points on the function $\langle s \rangle(T)$, T_i are the duration values of the points and $\langle s \rangle_i$ are the corresponding $\langle s \rangle$ values.

The KS statistic D was computed using

$$D = \max_{s \geq N/100} |F(s) - F_{fit}(s)| \quad (5.14)$$

where $F(s)$ and $F_{fit}(s)$ are the cumulative distribution functions (CDFs) of $P(s)$ and $P_{fit}(s)$, respectively.

We defined the transition region where the network dynamics switch from sub-critical to super-critical statistics using the binary variable indicator function γ .

$$\gamma = \begin{cases} 1 & \text{if } F(N) - F(0.6N - 1) > F'_{fit}(N) - F'_{fit}(0.6N - 1) \\ 0 & \text{else} \end{cases} \quad (5.15)$$

In equation 5.15, $F'_{fit}(s) = F_{fit}(s) \frac{F(N/100)}{F_{fit}(N/100)}$. γ assumes a value of 1, signifying super-critical statistics, if the tail of the empirical avalanche size distributions $P(s > 0.6N)$ is heavier than that of the fit. Additionally, we visually verified that the indicator γ works well for describing the behavior of the distributions in coupling space. The region in which its mean $\langle \gamma \rangle$ over N_a different stimuli lies between 0 and 1 was termed the transition region.

5.3.4 Computation of Information Entropy

We quantified information entropy $H(\mathbf{X})$ using a state variable \mathbf{X} which represents the spiking patterns of V4 neurons at a given time point t (Figure 5.1B). We construct the probability $P(\mathbf{X} = x_i)$ of observing a spike pattern x_i using the $T_{total}\Delta t$ spike patterns observed in one trial.

$$H(\mathbf{X}) = -\sum_i P(x_i) \log_2 P(x_i) \quad (5.16)$$

Considering different read-out strategies of the information encoded by V4 neurons in the higher visual areas, we computed information entropy in different scales of observation K . These scales were defined as follows (Figure 5.1B):

For the finest observation scale, $K = 1$, the state variable \mathbf{X} consists of N channels, representing N V4 neurons. Each channel assumes a value of 1 if the corresponding neuron generated an action potential at time t , and 0 otherwise. We randomly picked the order in which different neurons were represented in \mathbf{X} .

As we increase the observation scale K , \mathbf{X} comprises N/K channels, and each channel represents the sum of spikes from K different neurons. For $K > 1$, we constructed \mathbf{X} by adding up the spiking activity of K consecutive neurons, while conserving the aforementioned random order of neurons over the channels. At the coarsest scale of observation, we sum over the activity of the whole network (i.e. for $K = 1000$, \mathbf{X} is a scalar in the interval $[0, 1000]$).

5.4 Discussion

In this chapter we addressed the criticality hypothesis in the context of task-dependent modulations of neuronal stimulus processing. We focused, in particular, on changes in cortical activity induced by selective visual attention. We considered recent findings that γ -band oscillations emerge collectively with an enhancement of object representation in LFPs in macaque area V4 under attention [RTE+09]. We reproduced these results using a model of a visual area V4 population comprising IAF neurons recurrently coupled in a random network, where attention induces synchronous activity by modulating the efficacy of recurrent interactions. In the model, we investigated the link between experimentally observed enhancement of stimulus discriminability, scale-free behavior of neuronal avalanches and encoding properties of the network quantified by information entropy.

We found that the emergence of γ -band synchrony is strongly coupled to a rapid discriminability enhancement in the phase space. Notably, we observed that discriminability levels comparable to the experiments appear exclusively in the neighborhood of the transition region, where network dynamics transition from subcritical to supercritical for consecutive values of excitation for different stimuli. This effect arises because synchronizability of the network depends inherently on its connectivity structure, and the strength of synchrony for different stimuli is most diverse near and within the transition region. However, this also means that information entropy displays a sharp decline as network activity becomes strongly correlated for some stimuli, beginning within the transition region and reaching a minimum in the supercritical region. Therefore, we propose that cortical networks operate at near-critical states, at the subcritical border of the transition region. Such marginally subcritical states allow for fine modulations of network excitability to dramatically enhance stimulus representation in the LFPs. Finally, for a putative encoding scheme in which higher area neurons integrate over the spiking activity in local V4 populations (coarse observation scale), we find that near-critical states maximize information entropy.

In summary, our study establishes several, novel links between criticality, gamma-synchronization and task requirements (attention) in the mammalian visual system. Our model predicts that the cortical networks, specifically in visual area V4, operate at marginally subcritical regimes; quick and efficient task-dependent (e.g. attention induced) modulations of neuronal activity may push network dynamics toward a critical state; and the experimentally observed discriminability increase in LFP spectra can be attributed to differences in the network structure across different stimulus-specific populations, operating close to a phase transition.

5.4.1 Robustness of results

In this work we aimed to reproduce the characteristic features of the experimental findings with an uncomplicated model, in part because we did not want to overextend the model and be unable to appropriately sample the parameter space and in part due to considerations of computational expense. Therefore, we are confident that our qualitative results do not depend on a specific selection of parameter values. In addition, the conclusions of this chapter depend mainly on the facts that in our model: (1) the emergence of synchronous spiking activity can be described by a phase transition as a function of an excitability parameter, and (2) synchronizability of the network depends implicitly on the topography of its connections. Therefore, we believe that as long as these requirements are met, discriminability enhancement will correlate with a narrow choice of parameters which generate near-critical dynamics. This will also be the case in more complex and biologically plausible models which detail different synchronization mechanisms which might be responsible for generating neural γ -activity (see, for example, the reviews [BW12] and [TS09]).

In fact, recent modeling work by [PHM+12], which employed a network consisting of IAF neurons with stochastic spiking and local connectivity, reported a result which nicely parallels our findings. For random realizations of their network architecture, the greatest variance of the power-law scaling of the avalanche size distributions was found near the critical points. In this framework, different random realizations of network connectivity were used to describe differences between human subjects, and the authors concluded that their findings provide an explanation for interindividual differences in α -oscillations in human MEG.

5.4.2 Physiological plausibility

We simulated the structure of cortical subnetworks employing a random network of finite size, thus our model had a connectivity structure which varied for different subpopulations of activated neurons. This setting spared us any particular assumptions about the connection topology of V4 neurons, which is still subject of extensive anatomical research. In the brain, variability in connectivity of neurons in a local population is not random, but signifies a highly structured global network. Such functional connectivity is exemplified in the primary visual cortex by long-range connections between neurons with similar receptive field properties such as orientation preference [KTR+97]. These connections are thought to serve feature integration

processes such as linking edge segments detected by orientation-selective neurons in V1 or V2 into more complex shapes, thus giving rise to the array of receptive field structures found in V4 [DHG06; DS87]. In consequence, connection variability in the brain is significantly higher than random. Specifically, the variance of degree distributions is higher, the synaptic weights are heterogeneous, and the coupling structures are more anisotropic than in our simulations. Hence connection variability across different local networks is not decreased as drastically when the number of neurons is increased. In fact, assuming random variability implied a trade-off in our simulations: On the one hand, increasing the number of neurons decreased diversity in activation patterns and pattern separability, while on the other hand, it improved the assessment of criticality by increasing the range over which avalanche events could be observed.

In addition, in our model, we posited that attention modulates the efficacy of interactions, in order to reproduce the attention induced gain modulation and γ -synchrony using a reductionist approach. In biological networks, these effects may originate from more complicated mechanisms. For example, previous studies have shown that such an increase in gain [CAR02] as well as synchronous activity [BT06] can be achieved by modulating the driving background current. However, as explained in section 5.4.1, we expect our results will persist in other models where the network dynamics undergo a phase transition towards synchronous dynamics as a function of the responsiveness of neurons which is enhanced by attention. As an alternative to enhancing synaptic efficacy, we also tested a scenario in which attention provided an additional, weak external input to all neurons (results not shown). This led to qualitatively similar findings, with a quantitatively different discriminability boost.

Lastly, our current understanding of cortical signals strongly suggests that LFPs are generated mainly by a postsynaptic convolution of spikes from presynaptic neurons [LTP+11; MOK+14] and that even though other sources may contribute to the LFP signal, they are largely dominated by these synaptic transmembrane currents [BAK12]. We generate the LFP signal through a convolution of the sum of appropriately scaled recurrent and external spiking activity. In our model, this closely approximates the sum of postsynaptic currents to V4 neurons: We are considering a very simple model of a small V4 population in which the postsynaptic potentials are evoked solely by these recurrent and external presynaptic spikes; degree distributions in the connectivity structure of the network has a small variance; the recurrent synaptic strengths are homogeneous; and there is no stochasticity in the recurrent synaptic transmission (i.e. every V4 spike elicits a postsynaptic potential in the V4 neurons it is recurrently coupled to). In addition, there is no heterogeneity in cell morphologies or the location of synapses, which are believed to influence the contribution of each synaptic current to the LFP signal in cortical tissue [LPE10]. Combined, this means that each spike elicited by a model V4 neuron has a similar total impact on the postsynaptic membrane potentials, and the low-pass filtered spiking activity represents the postsynaptic currents well. Furthermore, even though our model does not incorporate the full biological complexity of cortical neurons,

we believe that the particular choice of constructing the LFP signal in our model is not consequential for our results. The increase in discriminability of the LFP spectra originate primarily in the γ -band (both in the model and the experimental data), and we assume that correlated synaptic currents emerge simultaneously with correlated spiking activity, as there is experimental evidence that spiking (multi-unit) activity is synchronized with the LFP signal during attention-induced γ -oscillations [FRR+01].

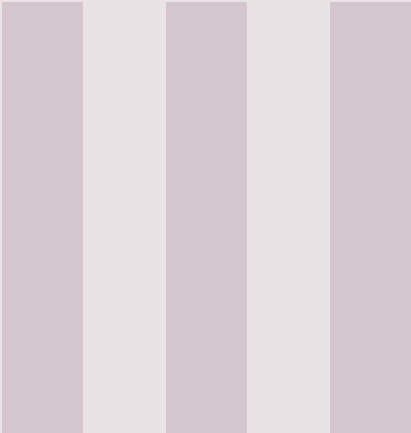
5.4.3 Functional role of local and downstream structure

In order to scrutinize the role of synchrony in enhancing stimulus representations, we considered an idealistic scenario: Each stimulus activates a different set with an *identical* number of neurons, so that without synchronization stimulus information encoded in activated neuron identities would be lost in the *average* population rate. By means of the different connectivities within different sets, however, this information becomes re-encoded in response amplitude and γ -synchrony. In principle, this concept is very similar to the old idea of realizing binding by synchrony [vdM94], namely, using the temporal domain to represent information about relevant properties of a stimulus, for example, by tagging its features as belonging to the same object or to different objects in a scene.

However, strong synchronization hurts encoding by destroying information entropy. This is visible in the dynamics in the supercritical regime where ultimately all neurons do the same: fire together at identical times. Therefore, synchronization is only beneficial for information processing if additional constraints exist: for example, a neural bottleneck in which some aspect of the full information available would be lost, or a certain robustness of signal transmission against noise is required and can be realized by the synchronous arrival of action potentials at the dendritic tree.

In our setting, this bottleneck is the coarse observation scale where neuron identity information is lost by averaging over all neural signals. In such a case, information entropy is maximized as oscillations emerge at near-critical points. Although this situation is most dramatic for epidural LFPs that sum over thousands of neurons, it may also arise in more moderate scales if neurons in visual areas downstream of V4 have a large fan-in of their presynaptic connections. Naturally, this does not exclude the possibility that such a bottleneck may be absent and that cortical encoding can make use of spike patterns on finer spatial scales. This would shift the optimal operating regime ‘deeper’ into the subcritical regime, and away from the transition region. Nonetheless, for this finer scale assumption, marginal subcriticality might represent a best-of-both-worlds approach. In particular, a penalty in information entropy may be necessary to ensure a certain level of synchronous activity required for other functionally relevant aspects of cortical dynamics, such as information routing regulated by attention via ‘communication through coherence’ [Fri05b; GNM+12b].

In general, coding schemes being optimal for information transmission and processing always depends strongly on neural constraints and read-out schemes. Nevertheless, specific assumptions about stimulus encoding do not influence our conclusion that the experimentally observed effects are unique to near-critical dynamics.



Feature Integration in Critical Networks

6	Figure-ground Segregation in a Critical Network	75
6.1	Introduction	
6.2	Methods	
6.3	Results	
6.4	Discussion	
7	Contour Integration in a Realistic LIF Model	97
7.1	Introduction	
7.2	Results	
7.3	Methods	
7.4	Discussion	



6. Figure-ground Segregation in a Critical Network

In this part we will investigate the idea that the correlations in the activity of local, functionally connected subpopulations of cortical neurons may be modulated in a bottom-up manner by the presence of objects in the visual field. In particular we will focus on how critical dynamics may emerge in a subnetwork with the presentation of a salient stimulus, contributing to enhanced figure-ground segregation performance when using a coincidence detector as a read-out mechanism. Parts of chapter 6 and figures 6.1-6.7 have been adapted from the book chapter “The role of criticality in flexible visual information processing” by Nergis Tömen and Udo Ernst, as featured in the book *The Functional Role of Critical Dynamics in Neural Systems* (in press, eds. Tömen N., Herrmann, M., and Ernst U., Springer-Verlag). doi:10.1007/978-3-030-20965-0

6.1 Introduction

As described in chapter 4, particular feature combinations (‘figures’) are processed by our visual system selectively, or in a privileged manner, even in the absence of top-down instructions to do so. Such stimuli are commonly accompanied by a “pop-out” effect, making these feature combinations highly salient in perception even if they are embedded into a “background” of distractor elements (for an example, see Fig. 6.1a). “Pop-out” is generally described as a stimulus-directed bottom-up effect. For example, a single local feature, such as a line segment whose orientation deviates from all other segments present in a scene will “pop out” to an observer [TG80]. Interestingly, evidence suggests that both context [JGW+01] as well as the global stimulus structure [KJ93] play a role in modulating the processing of an object by low-level, local feature detectors, thus requiring rapid co-processing and integration of multiple streams of information on different spatial scales.

Considering both the general notion that a dynamical regime close to a phase transition may boost the capacity of a network to sustain rapid and efficient mod-

ulations, as well as the dependence of perceptual grouping tasks on integration of spatially extended information, we hypothesize that spatially and temporally correlated neuronal activity may be conducive to bottom-up “pop-out” effects, which are helpful in figure-ground segregation.

How local features are integrated to form global percepts, in order to segregate relevant figures from a background, is extensively studied, especially in psychophysical investigations on Gestalt perception [WEK+12]. As a specific example, research on contour integration focuses on how the visual system identifies the outlines of shapes or figures by grouping colinearly or cocircularly aligned edge elements into coherent percepts (Fig. 4.1, blue circles and red outlines). For illustrating the modeling paradigm in this chapter, we will use contour integration as a proxy for all processes where binding of local features ultimately contributes to visual perception.

A typical visual stimulus employed in psychophysical contour integration experiments consists of oriented Gabor patches. In addition to a large number of randomly oriented distractor elements in the background, a stimulus may contain a target contour made up of appropriately placed and mutually aligned edge elements [FHH93] (Fig. 6.1a). Due to the adherence of the Gabor patches to cRF shapes reflecting the orientation selectivity of neurons in the early visual system, we think of each edge element as driving a population of local feature detectors. Contour integration has been shown to be an efficient and fast process, which is robust within a large range of stimulus parameters both in humans [EMS+12; HHF03] and in animal experiments [MK05]. Overall, the speed and accuracy with which contours can be detected in psychophysical experiments lead us to believe that a quick spread of activity and a subsequent swift employment of a large number of feature detectors which are involved in encoding the figure (or alternatively a swift suppression of background units) is necessary for effective contour integration. Therefore, we next investigate whether and how large scale avalanches spanning subnetworks processing a figure input may be beneficial in figure-ground segregation tasks.

6.2 Methods

In this section, we will first introduce the basic network and neuron models, with and without inhibitory connections, which we adopted for performing feature integration, followed by the task description and the introduction of the read-out scheme. Finally, we will explain how we introduced leak currents into the neuron model, before presenting the results.

6.2.1 The EHE network

Model description

As the network model, we adapt the Eurich–Herrmann–Ernst (EHE) model, which is an analytically well-understood network model displaying critical dynamics [EHE02]. The units in the original network are globally connected ($w_{ji} = 1 \forall j, i$) non-leaky integrators and the coupling strength α between the units determines the distance of the dynamics from the critical state (see [EHE02] for details).

In particular, each unit $i = 1, \dots, N$ is described by a state variable $u_i \in [0, 1)$. At each discrete time step k , a randomly selected unit u_i is perturbed by external input Δu :

$$u_i(k+1) = u_i(k) + \Delta u. \quad (6.1)$$

If its state variable u_i exceeds a threshold, in this case 1, then unit i is reset, and all units j in the network including i receive recurrent input, scaled by the weight matrix w_{ji} and the coupling strength α :

$$\begin{aligned} u_i(k) \geq 1 &\iff u_i \rightarrow u_i - 1 \\ u_j &\rightarrow u_j + \frac{\alpha w_{ji}}{N}. \end{aligned} \quad (6.2)$$

If any unit j then exceeds the threshold succeeding the recurrent input, the procedure in Eq. 6.2 is iterated until the avalanche terminates. Here we assume a separation of time scales, such that an avalanche of spikes is infinitely fast compared to the time scale on which the external, driving input arrives. This means that each avalanche is initiated by a single unit firing due to external input, and is completed within one discrete simulation time step. Such an implementation allows us to precisely quantify the avalanche size and duration distributions.

Couplings and subnetworks

For studying feature integration, we now assume that each unit represents a feature detector. If the corresponding feature, for example an oriented line segment, is present in a visual stimulus, the unit is activated by some external input. Detectors for features not present in the current visual stimulus receive no external drive. A figure is defined by an ensemble of features, which is a subset of all features represented in the network. We further require our network to be able to represent multiple figures, and that different figures might share subsets of features (Fig. 6.1B).

Unlike in the original model [EHE02], we use a systematically structured coupling matrix for representing figures in network topology. We begin with a zero matrix of size $M \times M$ and successively embed N_{sub} subnetworks, each containing N mutually and excitatorily coupled units (Fig. 6.1B, regions shaded in light orange). We posit that each subnetwork overlaps with exactly two other subnetworks by sharing N_{shr} units (feature detectors), which yields a coupling structure resembling a one-dimensional figure ‘chain’ with periodic boundary conditions, with a total network size of $M = N_{sub} \cdot (N - N_{shr})$ units. In this scheme, the units in each subnetwork collectively encode the presence of a figure in the visual field (such as a well-defined contour line as in Fig. 6.1A), and neighboring subnetworks may be activated by figures which share local features in their neuronal representation (Fig. 4.1, red contours).

Addition of inhibition

Since in a setting with shared figure representations we expect some degree of activity to spread from externally driven units to the rest of the network, we would ideally like

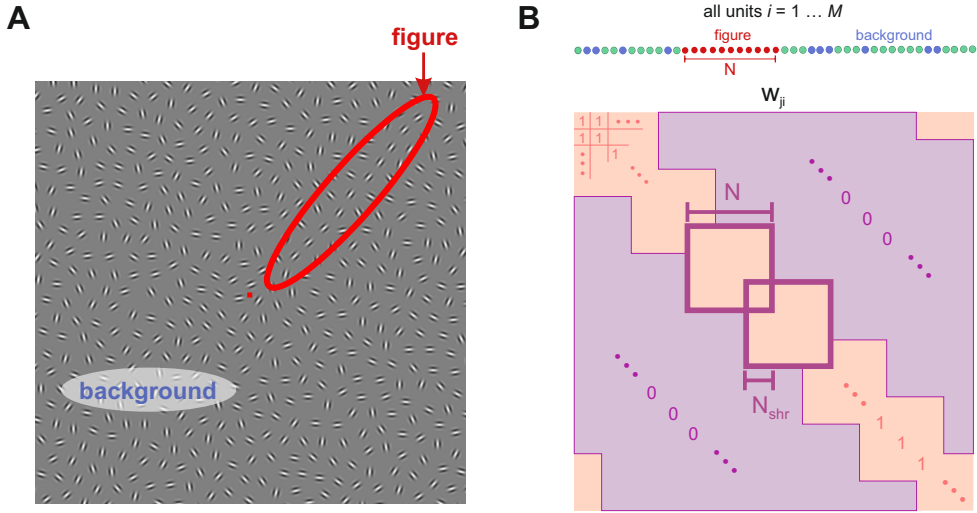


Figure 6.1: **A** A typical stimulus from a contour integration experiment. It contains one figure consisting of colinearly aligned, oriented Gabor patches, which are hidden among randomly oriented background elements. **B** The coupling matrix for the purely excitatory network model. This example consists of $N_{sub} = 7$ subnetworks of N mutually excitatorily coupled EHE units each (regions shaded in light orange). Every subnetwork shares N_{shr} units with its two nearest neighbors. The stimulus externally drives all N units in one subnetwork (filled red circles) and $N - N_{shr}$ background units randomly selected from the remaining subnetworks (filled blue circles). In the dynamics of the model (Eq. 6.2), the value of w_{ji} displayed in this example is scaled with the coupling strength α , meaning that the absolute efficacies of all excitatory connections are equal.

to assess the extent to which it would affect network dynamics, and ultimately task performance. To address this point, we extend our paradigm by introducing inhibitory connections between neighboring subnetworks, with the intention to alleviate cross-talk between them through the shared units, and compare the results to the original, purely excitatory setup.

Inhibitory synapses are introduced into the coupling matrix w_{ji} by the addition of $(N - N_{shr}) \times (N - N_{shr})$ blocks of negatively valued connections (Fig. 6.2, regions shaded in blue color). In general, runaway activity may be circumvented by the addition of global inhibitory connections, i.e. by replacing all zero-valued connections with inhibitory ones. However, it is both more faithful to biology to implement strong, local inhibition as well as more meaningful for the simulation of a behavioural task where we want to conserve, to some degree, the activity of the background units, instead of completely suppress them. This allows us to model noise in neuronal activity of the whole network, decreasing signal to noise ratio, and helps emulate the role of the distractors when evaluating figure detection performance. The structured, local inhibition is thus implemented by feature detector units belonging to one figure

inhibiting only those units which belong to directly neighboring figure ensembles. For the units which belong to two adjacent figures at the same time (e.g. figures F_2 and F_3), the inhibitory connections reach out to the two nearest neighbors of the ensemble formed by those two figures (e.g. to figures F_1 and F_4) (Fig 6.2).

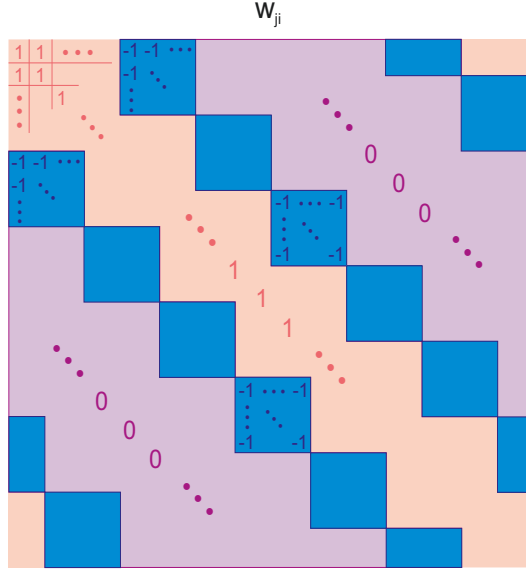


Figure 6.2: The coupling matrix w_{ji} as shown in Fig. 6.1, with the addition of inhibitory synaptic connections. We introduce the inhibitory couplings only between subnetworks sharing units (regions shaded in blue), such that feature detectors which belong to one or two figures inhibit other feature detectors belonging to directly neighboring figures. In our investigations, we consider both, purely excitatory and mixed excitatory-inhibitory networks. As in the purely excitatory case, the weight matrix w_{ji} is scaled with the coupling strength α (Eq. 6.2), hence all excitatory and inhibitory connections have the same absolute value.

Stimulus presentation and network activity

In order for the network to perform a figure detection task, we present it with a stimulus containing one whole figure and random background elements. Such a stimulus drives all of the N units of one subnetwork, and $N - N_{shr}$ randomly selected units from the remaining $M - N$ units with external input (filled circles at the top of Fig. 6.1B). The $N - N_{shr}$ units are selected randomly with uniform probability. All other units receive no driving input. The single unit i which receives the external drive according to Eq. 6.1 is selected at every discrete time step with uniform probability from the $N_{act} = 2N - N_{shr}$ externally activated units which make up the figure and background subpopulations.

For evaluating the dynamics of different subpopulations, we separately record the activity of a ‘target’ and a ‘distractor’ ensemble with N units each. By keeping the size of the recorded subpopulations the same, we control for the number of units

in both ensembles, which will be observed by neurons in a higher visual area in order to perform a figure detection task. The target ensemble is composed of the N activated figure units. The distractor ensemble is composed of the $N - N_{shr}$ externally activated background units plus N_{shr} units of the currently activated figure which are shared with one of the other, not activated figures. This is in consideration of the idea that units which belong to more than one figure may also be sampled as part of the background by higher areas. For each ensemble, we record the output spikes fired at each simulation time step.

Due to the separation of time scales, the total population activity in one ensemble at a given discrete time step k describes the size s of an avalanche within that ensemble at k . After running the simulations for a total of K time steps, we can construct frequency histograms for the avalanche sizes, which approximate the avalanche size distributions. Since we are interested in the power-law behaviour of the network, we compute the Kolmogorov–Smirnov (KS) statistic [Smi48] of avalanche size distributions $P(s)$ with respect to a reference power-law function $P_{ref}(s)$

$$KS = \max_s |F(s) - F_{ref}(s)| \quad (6.3)$$

using $F(s)$ and $F_{ref}(s)$, which are the cumulative distribution functions (CDFs) of $P(s)$ and $P_{ref}(s)$, respectively. For the calculations, we used a reference power-law distribution with exponent -1.43 ($P_{ref}(s) \propto s^{-1.43}$). This was the exponent which yielded the minimum KS distance for a control network: a globally connected EHE network of size N , simulated for a total of K time steps. The control network is simulated at its analytically known critical point with $\alpha_{crit} = 1 - 1/\sqrt{N}$.

Since in the model a figure is described by a subnetwork of feature detectors with strongly excitatory recurrent connectivity, we expect that the synchronous spiking in the target ensemble will be a powerful signal that a complete figure is present in the visual field. In contrast, we expect that the firing of the distractor ensemble will be relatively asynchronous, or at least randomly correlated in the purely excitatory case, and somewhat suppressed in the inhibitory case. Therefore, our model encodes information about the presence of a salient stimulus in the time domain, where relative spike times in an ensemble contains the relevant information. For this reason, we employ coincidence detection as the read-out mechanism in a figure-ground segregation task.

Specifically, we use two leaky integrate-and-fire (LIF) neurons for discriminating whether a figure is observed by the target ensemble or in the distractor ensemble. We feed the population activity of the target population to one LIF neuron, and activity of the distractor population to a second LIF neuron. In order to employ the LIF neurons as coincidence detectors and exclude information contained in the firing rates, we normalize the total input to each LIF neuron over the duration of a trial by its sum (Eq. 6.5). Formally, the membrane potentials $V_{tar}(t)$ and $V_{dis}(t)$ of the respective read-out neurons evolve according to

$$\tau \frac{dV_p(t)}{dt} = -V_p(t) + f \sum_{k=1}^K \bar{s}_p(k) \delta(t - k\Delta t) \quad (6.4)$$

where $p \in \{tar, dis\}$, τ is a time constant with $\tau = 10$ ms, and f denotes an arbitrary gain factor. The ODE was numerically integrated using the Euler method, with the spiking threshold at $V_\theta = 1$. In addition, for the numerical integration, we picked a time step Δt , such that we could conserve the biological realism in the firing rates of both the lower area EHE units and the higher area LIF neurons (see section 6.3.1 for details).

Inputs $\bar{s}_p(k)$ are derived from the discrete population spike time series $s_p(k)$ which contain the size of the avalanche which took place in population p at simulation time step $k = 1 \dots K$ of the discretized dynamics (Eq. 6.1). The avalanche size is defined as the number of units which elicited a spike in the same time step k due to recurrent input. Since we separately record the population activity of the target and distractor populations, $s_{tar}(k)$ and $s_{dis}(k)$, respectively, and each population contains N units in total, s assumes integer values in the interval $[0, N]$. As mentioned before, in order to eliminate a performance increase due to rate differences between the populations, $\bar{s}_p(k)$ is obtained by normalizing the population activity by its sum over time.

$$\bar{s}_p(k) = \frac{s_p(k)}{\sum_{k'=1}^K s_p(k')} \quad (6.5)$$

where once again $p \in \{tar, dis\}$.

Figure detection performance P_{fig} is then defined as the difference between the output rates of the two read-out neurons. However, we find that in addition to α and N_{shr} , the task performance also depend on the gain factor f (Eq. 6.4). In other words, how the task performance behaves, and where in the phase space it exhibits a maximum, changes depending on the value of f (see section 6.3.1, specifically Fig. 6.6, for details).

In order to isolate the effects of N_{shr} and α on performance, we first assess whether a global maximum can be located for the performance as a function of f . This is carried out by first computing the task performance $P_{fig}(N_{shr}, \alpha, f)$ in the whole phase space spanned by N_{shr} and α for a given value of f , and then computing the maximum

$$\Sigma_{perf}(f) = \max_{N_{shr}, \alpha} [P_{fig}(N_{shr}, \alpha, f)] \quad (6.6)$$

for each f . We then locate the factor f_{opt} at which $\Sigma_{perf}(f)$ reaches its maximum. In order to achieve the desired degree of precision with lower computation time, we implement a golden section search on the function $\Sigma_{perf}(f)$ in the interval $f \in [0, 0.02]$.

6.2.2 Leaky network model

Since the EHE network discussed so far is composed of non-leaky units, a simple way to improve the biological realism in the model is the addition of leak currents to the neuron model. As mentioned in chapter 3, and demonstrated in chapter 5, leaky neuron models can display phase transitions [CPDG+95; TRE14; USO95]. In addition, the avalanche size distributions in the EHE model with small leaks exhibit a typical transition from exponential to heavier tailed distributions [EHE02].

In the leaky network model, we use the non-leaky EHE network which evolves according to equations 6.1 and 6.2 and we implement the leaks by simply enforcing that all membrane potentials u_i decay back to zero by a portion of their current magnitude

$$u_i(k+1) = u_i(k)(1 - C_{leak}) \quad (6.7)$$

at every simulation time step k . Since with large leaks, and small driving current, the network will be completely silenced, we picked a relatively small leak constant $C_{leak} = 5 \cdot 10^{-5}$. In order to define a time axis with units for the read-out procedure, we follow the same thinking as before, and pick the Δt , which yields a mean target population firing rate around 50 Hz, around the critical value of the coupling strength, averaged over all N_{shr} . As before, this also corresponds to periodic, oscillatory activity in the low gamma-range. Determined this way, the specific values of $\Delta t = 0.004$ ms and the leak constant C_{leak} correspond to an exponential decay $e^{-t/\tau}$ over time with $\tau = 80$ ms.

For the coupling structure, we reuse the coupling matrix shown in Fig. 6.2, and, as before, we have excitatorily coupled figure subpopulations with structured, local inhibition. Since neither α_{crit} , nor the exponent values for the avalanche distributions are theoretically known for the leaky system, we fit exponents to the avalanche size distributions by searching in the range $[-4, -1.1]$ for the exponent value which minimized the KS distance of the observed distributions to an ideal power-law. In the fitting procedure, we excluded avalanches of size $s < 3$, which yielded an exponent of -2.3 for the distributions closest to a power-law as observed in visual inspections. We defined the critical coupling strength $\alpha_{crit}(N_{shr})$ as the value of α which produced the lowest minimum KS distance for each N_{shr} .

The simulation methods, sampling of the target and distractor populations as well as the read-out methods are identical to those presented in the previous section for the non-leaky network.

6.3 Results

In the following sections, we present the results of the above described feature integration and figure detection task, first using the basic EHE network with and without inhibition, and afterwards using the slightly more realistic network model including leak currents.

6.3.1 The EHE network

For the results presented here, we use $N_{sub} = 7$ subnetworks, each consisting of $N = 225$ EHE units, and we drive the network with an external input of $\Delta u = 0.022$ for a total of $K = 10^7$ time steps. The spiking activity of the network is recorded, and the dynamics of different populations characterized as we vary the overlap between the subnetworks N_{shr} ¹ in addition to the coupling strength α . This gives us an idea

¹The maximum N_{shr} scanned for the excitatory network is larger than the maximum N_{shr} for the network with inhibition due to the restrictions imposed on the structure of the coupling matrix by the

about the plausibility of such a network topology under increasingly heavy load (i.e. where many units need to be shared between internal representations of different figures) as well as about the robustness of dynamics as we move closer to and beyond the critical state.

Synchronization and avalanches

As one would expect, when α is increased, in other words when the recurrent excitatory couplings within a subnetwork is increased, the target ensemble is strongly synchronized and we observe consecutively larger avalanches. Interestingly, we find that large sized avalanches are generated by the target ensemble in a relatively periodic manner, whereas large avalanches generated by the distractor ensemble are relatively rare and not as predictable (Fig. 6.3A). In addition, the target ensemble elicits strongly synchronized spikes more regularly as α is increased. We illustrate this increasing regularity in oscillatory subnetwork activity by the coefficient of variation of the inter-avalanche time intervals (population CV) for avalanches with size $s > 56$ (Fig. 6.3B). This enhancement of periodic activity as we get closer to the critical point is also reminiscent of the emergence of gamma-oscillations close to the transition boundary presented in chapter 5.

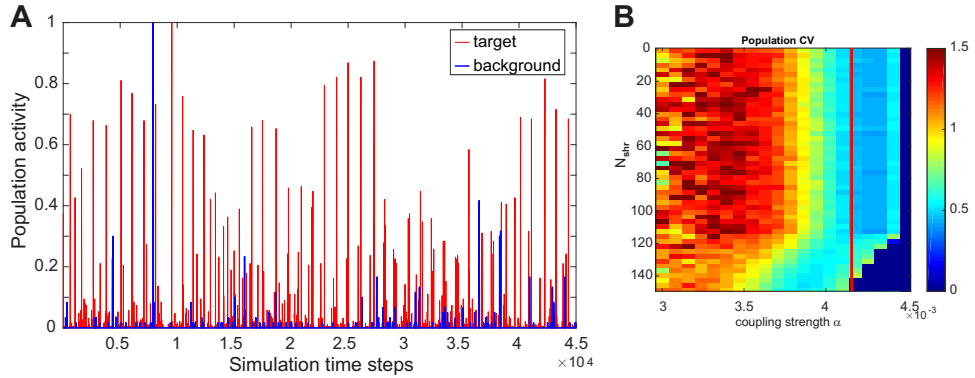


Figure 6.3: **A** Population activity of the target ensemble (red) and the remaining background units (blue) in the network with inhibition as a function of time. The population spike activity is obtained using $N_{shr} = 73$ and $\alpha = 4.168 \cdot 10^{-3} \approx \alpha_{crit}$, for which the target ensemble generates strongly synchronous activity with regular intervals, whereas the dynamics in the background are relatively asynchronous and irregular. For plotting, both activities are normalized by their respective maxima, which were avalanches of size 190 and 60 for the figure and background respectively. **B** Coefficient of variation for the inter-avalanche intervals (population CV) for avalanches of large size ($s > 56$) in the target ensemble. The regularity with which the network elicits large avalanches increases dramatically around the critical value of the coupling strength (red line).

In the purely excitatory model, when we investigate the avalanche dynamics in periodic boundary conditions.

terms of the KS statistic of the size distributions, we find, trivially, that the activity of the target ensemble becomes critical and subsequently supercritical with increasing α . Eventually, firing rates grow without bound and the network gets stuck in a regime of infinitely long avalanches. We identify the α for which subnetwork activity is critical by finding the avalanche size distributions for which the KS statistic is minimized (Fig. 6.4A, white circles).

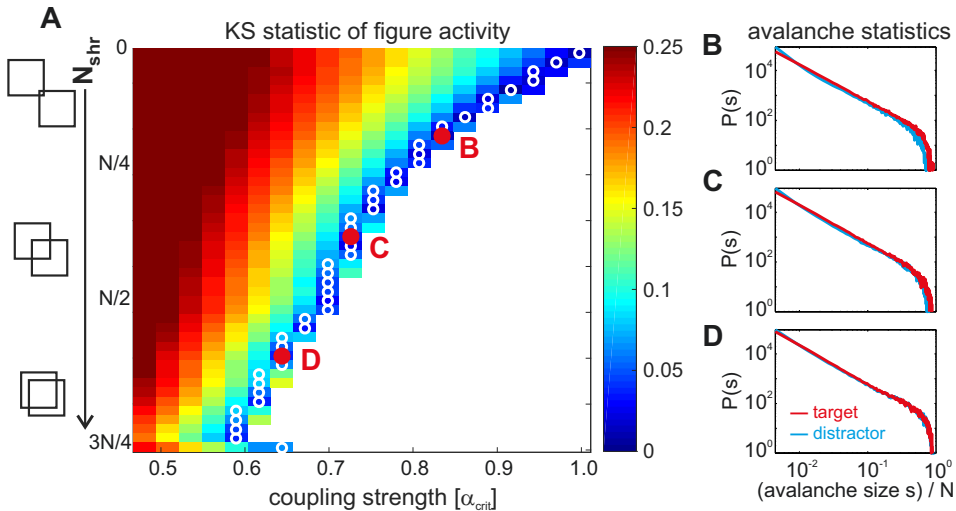


Figure 6.4: **A** The KS statistic of the avalanche size distributions in the target population as a function of the subnetwork overlap N_{shr} and the coupling strength α , in the purely excitatory network. In the bottom-right white region, infinite avalanches emerge. The numerically calculated critical point for each N_{shr} is marked with a white circle. Coupling strength α is displayed in units of the theoretical critical value $\alpha_{crit} = 1 - 1/\sqrt{N}$ for a globally connected network of N units [EHE02]. **B**, **C** and **D** show the avalanche size distributions in the target and distractor populations for the corresponding parameter combinations marked in red in **A**.

For very small N_{shr} , the presentation of a stimulus with one complete figure embedded into random background elements splits the network into two distinct groups: the dynamics of the subnetwork representing the target figure quickly become synchronous as we increase the coupling strength, whereas the dynamics of the rest of the network as well as the distractor population lag slightly behind the target population in the sense that they comparatively lack large avalanches and highly correlated activity. Background activity instead exhibits a slightly larger amount of small sized avalanches (Fig. 6.4B).

As mentioned before, due to our choice of sampling of target and distractor activity, the dynamics of the two populations converge naturally as N_{shr} approaches N . Therefore, any existing discriminability of target avalanche activity from distractor activity is quickly diminished with increasing N_{shr} (Fig. 6.4B-D). However, this convergence is bolstered by runaway activity, or activity spreading out from the

strongly synchronous target units. This cross-talk between the subnetworks is also a factor in why the critical value of α decreases as we increase N_{shr} .

Addition of inhibition

One way to circumvent runaway activity and alleviate the co-activation of background populations is to employ inhibitory interactions. Indeed, in the network with structured inhibition, we find that the coupling strength for which the target activity is critical is now robust against changes in N_{shr} (Fig. 6.5A), as well as against the number of subnetworks embedded in the network N_{sub} (not shown) up to an overlap of $N_{shr} = N/2$. After this point, non-neighboring subnetworks begin to share units, which is an effect our choice of inhibitory topology cannot offset. We can verify that the cross-talk between subnetworks is negligible by plotting the analytical value of the critical α for a globally connected, stand-alone network of N units in comparison to the KS statistic minima (Fig. 6.5A, red line). This also confirms that in general changing network size, i.e. varying N_{shr} or increasing N_{sub} , should not change the results reported here, and the only parameter that the critical value of α depends on in our model is the subnetwork size N .

The reduction in activity spill-over from the target population as well as the suppression of the externally activated distractor units is also apparent in the corresponding avalanche size distributions (Fig. 6.5B-D). In comparison to the dynamics in the absence of inhibition, we find that synchrony in the distractor population is further quenched while overall background activity levels are lowered. This across-the-board decrease in background activity also contributes to the formation of a larger buffer region of supercriticality between the critical point, and the transition to the infinite avalanche regime. While an average α increase of about 4% is sufficient to bring the purely excitatory network from critical dynamics into the infinite avalanche regime, an average increase of about 8% in α is required for the network with inhibition.²

Task performance

The discrepancy between the synchronization levels of target and distractor populations seems promising for putative read-out mechanisms to successfully perform figure-ground segregation. To capitalize on this finding, we use leaky integrate-and-fire (LIF) neurons, acting as coincidence detectors, for distinguishing between figure and background. In order for the results to only be sensitive to coincidences, and not the firing rates of the two ensembles, we normalize the total input into the LIF neurons over time (see Eq. 6.5 and section 6.2.1 for details).

For feeding the EHE network outputs $\bar{s}_p(k)$ into the LIF neurons, we choose the width of our simulation time step $\Delta t = 0.02$ ms, such that the average firing rate of a target figure unit, when the whole figure subnetwork was activated at the critical point, was approximately 40-50 Hz (depending on N_{shr}). Furthermore, the EHE network also displayed relatively regular activity in the form of large sized

²The reported ‘average increase’ is averaged over N_{shr} and is much smaller than 4% in the purely excitatory network for some N_{shr} .

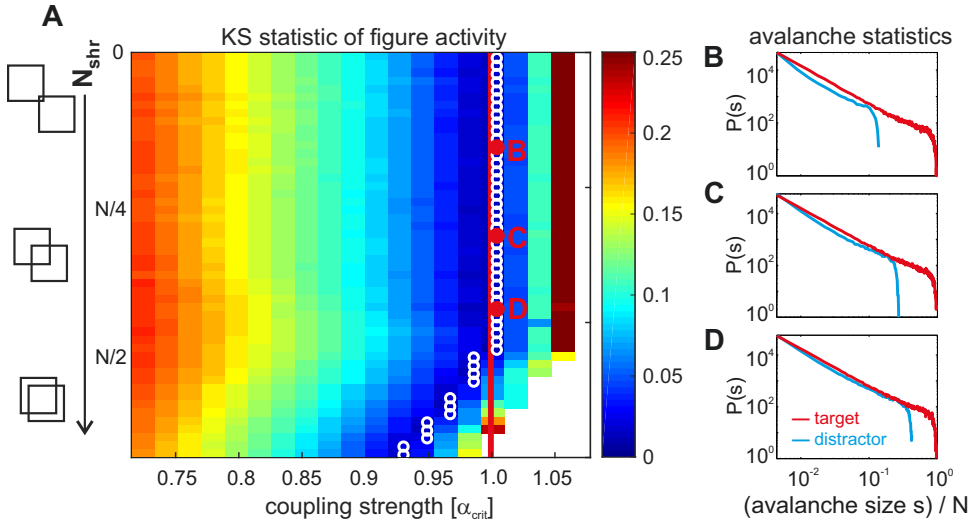


Figure 6.5: **A** The KS statistic of the avalanche size distributions in the target population as a function of the subnetwork overlap N_{shr} and the coupling strength α , in the network with mixed excitation and inhibition. The red line marks the critical point α_{crit} of a globally coupled network with N units. Other annotations are the same as in figure 6.4. **B**, **C** and **D** show the avalanche size distributions in the target and distractor populations for the corresponding parameter combinations marked in red in **A**.

avalanches, as shown in Fig. 6.3. Interestingly, this choice of Δt additionally revealed a periodic, or oscillatory behaviour, in the γ -frequency range – a realistic setup for a local subnetwork in the visual cortex [EBJ+88]. These values correspond to low frequency gamma-oscillations which we expect to emerge strongly around the critical point for small subnetworks processing a figure, as explained in chapters 4 and 5.

Using this scheme, we compute the figure detection performance P_{fig} , defined as the difference between the output rate of the target read-out neuron and the distractor read-out neuron, in the phase space as shown in Fig. 6.5A. We find that in addition to α and N_{shr} , P_{fig} also depends on the LIF neuron gain factor f (Eq. 6.4). Figure 6.6 shows how the P_{fig} landscape changes as we vary the value of f in the range $[0.0018 - 0.0076]$.

For the small set of f values considered, we find that the behaviour of P_{fig} within our phase space changes with varying f , and in particular, that the α which allows for maximum task performance for a given N_{shr} is in general shifted to subcritical values with increasing f (Fig. 6.6). However, we also note that the maximum performance we find for each f is different. This raises the question of whether there might be a global maximum for P_{fig} in the three-dimensional phase space spanned by N_{shr} , α and f . Nevertheless, scanning the whole three-dimensional space with sufficient resolution would be inefficient, therefore, we implemented a golden section search on the function $\Sigma_{perf}(f)$ (Eq. 6.6), and evaluated P_{fig} in our standard phase space for

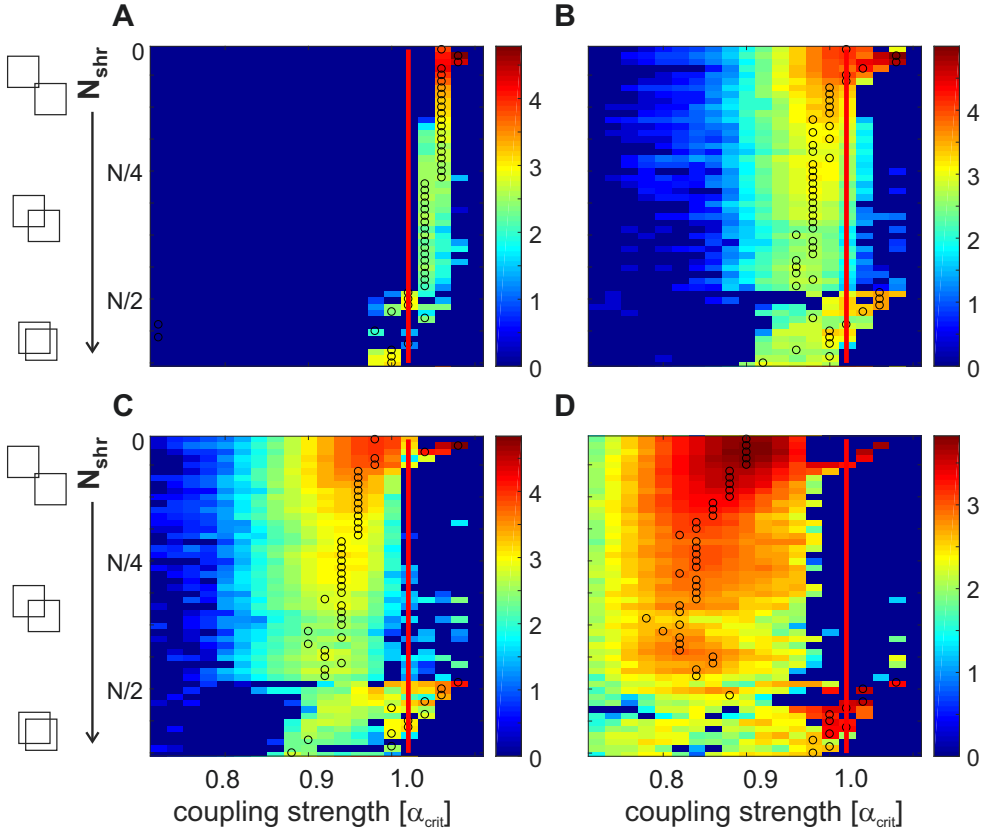


Figure 6.6: Logarithm of the figure detection task performance $\ln(P_{fig})$ in the network with inhibition as we vary the read-out neuron gain factor f . For all figures, the phase space is identical to that in Fig. 6.5. For **A** to **D**, the f values are 0.0018, 0.0029, 0.0047, 0.0076 respectively. Black circles mark the maximum performance for each value of N_{shr} and the red lines. A careful look at the colorbars reveals that the maximum performance value in the phase space varies with f , along with the location of the maxima.

a relatively small number of f values, with increasing precision around the global maximum.

This allowed us to demonstrate that $\Sigma_{perf}(f)$ has a clear global maximum at f_{opt} , with the performance falling off monotonically around f_{opt} (Fig. 6.7A). Typically the performance maxima are found at smallest values of N_{shr} and large to supercritical values of α , where target activity is high and distractor activity is almost completely suppressed (see e.g. Fig. 6.6A-D). The global maximum shown here is close to a rate difference of 150 Hz between the two read-out neurons, but the exact value of the performance will vary depending on N_{shr} and read-out parameters such as the time constant τ of the LIF neurons (see section 6.4 for a more in-depth discussion).

The value of f which maximizes $\Sigma_{perf}(f)$ appears to be relatively robust against changes in other parameters such as the number of externally driven background

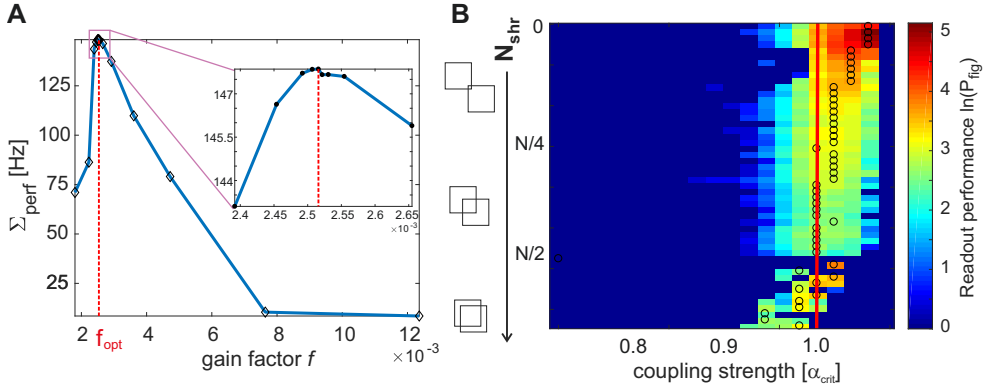


Figure 6.7: **A** The maxima Σ_{perf} of the figure-ground segregation task's performance P_{fig} as a function of f displays a clear global maximum at f_{opt} . Since P_{fig} is defined as the output rate difference of the read-out LIF neurons, its units are in Hz. **B** Logarithm of P_{fig} as a function of subnetwork overlap N_{shr} and coupling strength α using the optimal gain factor f_{opt} . Black circles mark the performance maxima for each value of N_{shr} . The red line shows the theoretical value of the critical coupling strength α_{crit} as in Figures 6.5 and 6.6. For the medium range of N_{shr} where the performance maxima converge to the critical point, the read-out rate difference P_{fig} is in the range of 12-20 Hz.

units and the read-out neuron time constant τ , when chosen within a biologically plausible range (see section 6.4). Interestingly, when adopting the value $f = f_{opt}$, figure detection performance in the phase space peaks close to the critical α for most medium-sized values of N_{shr} (Fig. 6.7B).

We find, as would be expected, that for a small number of shared units between subnetworks, where the risk of activity leak is low, it is most advantageous for the local figure population to display supercritical, strongly synchronous activity, in order to maximize coincidence detection. Each figure which may appear in our visual field having its own isolated, dedicated population of neurons, however, is both unrealistic and arguably a suboptimal use of available resources. As N_{shr} , along with the overlap of the target and distractor populations, increases, figure detection performance P_{fig} drops and the location of the peak shifts towards the critical α . These findings suggest that the critical state provides the optimal regime in which to perform feature integration in our model for a range of medium to large subnetwork overlap N_{shr} .

6.3.2 Leaky network

For the results presented in this section, we use the same parameters and the coupling matrix as for the non-leaky network with inhibition: $N_{sub} = 7$ subnetworks, each consisting of $N = 225$ leaky EHE units, and we drive the network with an external input of $\Delta u = 0.022$. Due to an overall lower firing rate, we simulate the network for longer, for a total of $K = 3 \cdot 10^7$ time steps.

As in the non-leaky network, we observe that oscillatory activity, in the form of periodically elicited medium to large size avalanches, also emerges in the leaky network as α approaches the critical point. The emergence of oscillations, as well as the transition to supercritical avalanche distributions, however, are more abrupt in the leaky model, in line with the suggestion that LIF neurons exhibit first order phase transitions [SAM+18]. Figure 6.8 shows how the network dynamics in the target population (red) evolve from relatively low, asynchronous activity to regular, medium-sized avalanches. Similarly, the activity in the distractor population undergoes a dramatic change, from almost completely asynchronous to periodic avalanche activity.

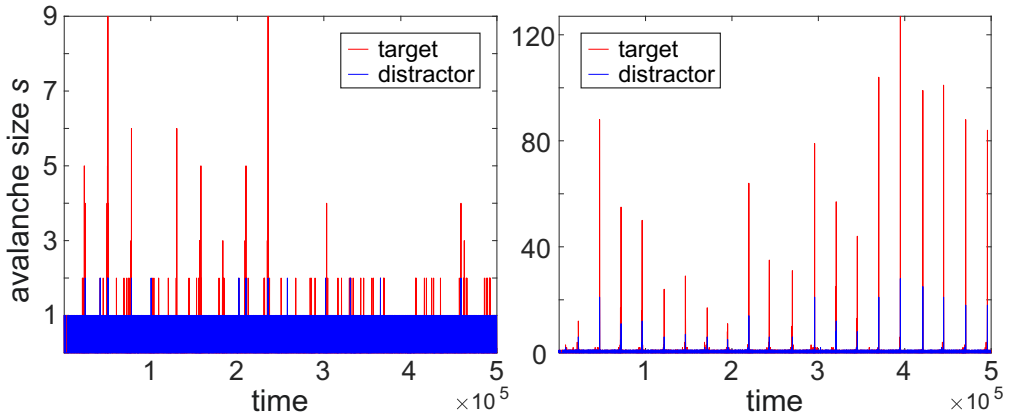


Figure 6.8: **A** Population spiking activity in the target and distractor populations as a function of simulation time steps for coupling strength $\alpha = 1.47 \cdot 10^{-3}$. **B** Same as **A**, but with $\alpha = 1.63 \cdot 10^{-3}$. This is a very slight increase in the coupling efficacy considering the epileptic regime is only reached at $\alpha \approx 4.48 \cdot 10^{-3}$ (see, for example, figure 6.9A). The change in the dynamics in terms of avalanche sizes and the regularity of population activity is dramatic.

Part of the increase in oscillatory activity we see in the distractor population, in comparison to the non-leaky network, is due to the fact that in the non-leaky network inactive units could be infinitely hyperpolarized. In the non-leaky network, even the units which received external drive could have a negative membrane potential, averaged over the total duration of the simulations, due to strong recurrent inhibition, and units which were inhibited by the target population were strongly hyperpolarized (with $u_i < -5$ at the end of simulations). With the addition of leaks, even units which receive no external drive will have their membrane potentials converge to zero over time in periods with no input. Such units may now elicit a spike through synchronous, excitatory recurrent input, for example, in the form of a large avalanche taking place in the target population spilling over to units in neighboring figures.

Due to this increase in activity spill-over from the target population to background units, we find that the large differences between the avalanche size distributions of target and distractor populations are, to some degree, lost in the leaky network.

Figures 6.9B-D show how the avalanche size distributions behave in the leaky network, where, interestingly, we find that the power-law exponent for the target population is larger in magnitude than in the network without leaks. This change may be attributed to both the addition of the leaks (as also shown for the original EHE model [EHE02]), as well as the extra inhibition received by the target population due to the increased activity of the background units.

In figure 6.9A we show the KS statistic calculated using the avalanche size distributions and an ideal power-law with the fitted exponent -2.3 (see section 6.2.2 for details), and identify $\alpha_{crit}(N_{shr})$ for each value of N_{shr} (white crosses). We find that in the leaky network the independence of α_{crit} on N_{shr} is lost, although the interval of α in which we find $\alpha_{crit}(N_{shr})$ is still relatively small, compared to the total range from $\alpha = 0$ to $\alpha \approx 4.48 \cdot 10^{-3}$, where infinite avalanches emerge.

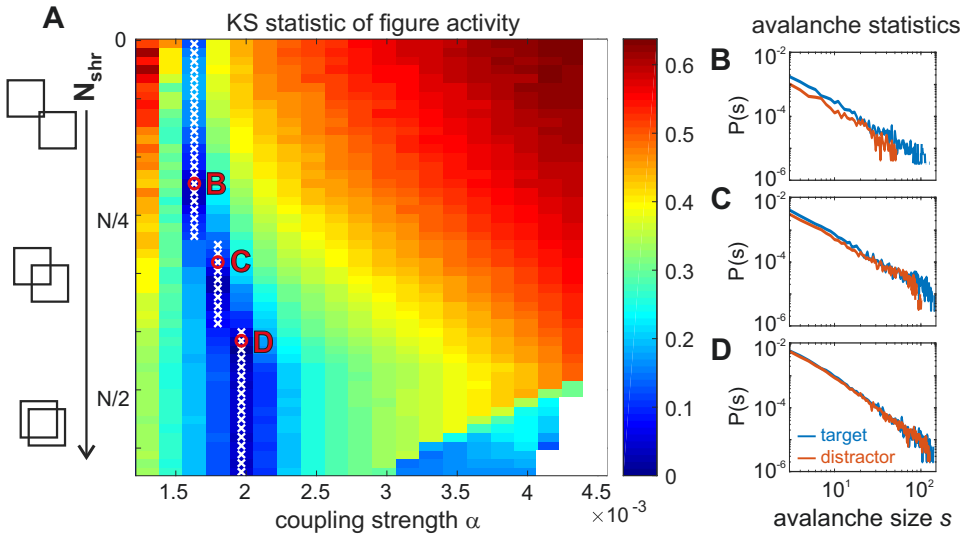


Figure 6.9: **A** The KS statistic of the avalanche size distributions in the target population as a function of the subnetwork overlap N_{shr} and the coupling strength α , in the leaky network with mixed excitation and inhibition. The white crosses mark the $\alpha_{crit}(N_{shr})$, defined as the numerically computed minimum of the KS statistic for each value of N_{shr} . Other annotations are the same as in figure 6.5. **B**, **C** and **D** show the avalanche size distributions in the target and distractor populations for the corresponding parameter combinations marked in red in **A**.

Interestingly, with the addition of the leaks, the network exhibits a very large buffer zone between the critical dynamics and the epileptic regime (Fig 6.9A). In comparison with the inhibitory, non-leaky EHE network, where an 8% increase in α_{crit} would place the dynamics in the infinite avalanche activity zone (Fig. 6.5), in the leaky network an increase of almost 150% is required. This is potentially an important property for the brain where, during learning or development, neurons in a network may experience large fluctuations in their synaptic input, without displaying epileptic behaviour.

Finally, performing the same LIF neuron read-out method as before (as described in section 6.2.1) in our phase space spanned by α and N_{shr} , we find that the leaky network displays a certain level of robustness in the location of the maxima of the task performance P_{fig} with respect to the read-out neuron gain constant f (Fig. 6.10). As before, P_{fig} is defined as the rate difference between the target and distractor read-out neurons, and is given in units of Hz, however, for visibility we plot the logarithm $\ln(P_{fig})$. In contrast to the non-leaky network (Fig. 6.6), the P_{fig} landscape in the leaky network varies relatively little with changing f . The location of task performance maxima (Fig. 6.10, white crosses) also correlate reasonably well with the numerically identified α_{crit} (Fig. 6.10, black circles), for all values of f evaluated in the interval $[0.0112, 0.044]$.

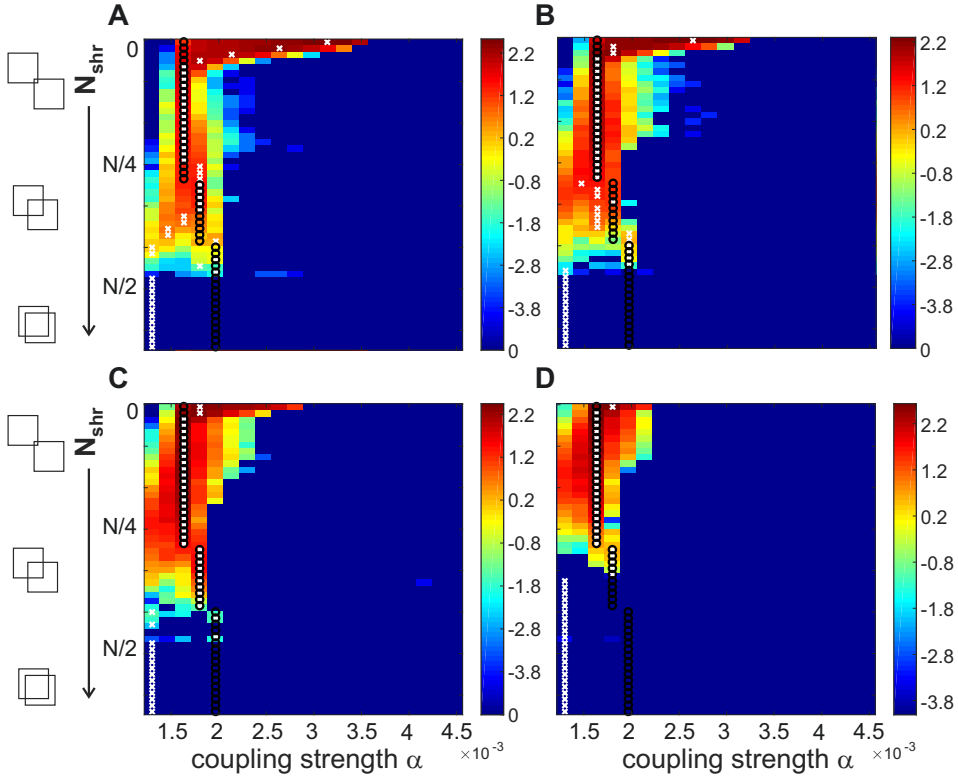


Figure 6.10: Logarithm of the figure detection task performance $\ln(P_{fig})$ in the leaky network with inhibition as we vary the read-out neuron gain factor f . For all figures, the phase space is identical to that in Fig. 6.9A. For A to D, the f values are 0.016, 0.02, 0.024, 0.032 respectively. White crosses mark the maximum performance for each value of N_{shr} and the black circles correspond to $\alpha_{crit}(N_{shr})$ as identified in figure 6.9A. The location of the performance maxima correlate reasonably well with the critical coupling strength values for each f .

As in the non-leaky network, we find that, when using the value of the read-out gain factor f which globally maximizes task performance, the maxima of P_{fig}

correlate strongly with the critical points (Fig. 6.11). In the leaky network, we did not perform a golden section search, but simply sampled 12 different values of f in the range given above, since we did not need a very high f resolution, as the P_{fig} landscape did not vary as dramatically with small changes in f (Fig. 6.10).

Due to the excitation spill over, and the consequent increased firing rates and synchrony of the background units in the leaky network, we find that the maximum read-out performance is considerably lower than in the non-leaky network: it is in the range of 2 – 9 Hz for varying N_{shr} . Since the activity levels in the distractor subpopulation catches up relatively quickly with the target population with increasing N_{shr} (Fig. 6.9B-D), figure detection performance converges quickly to zero around $N_{shr} = N/2$. Similarly, since both the activity of the target population and the distractor population become abruptly more synchronous at the critical value of α (Fig. 6.8 and Fig. 6.11D), and the differences in the maximum avalanche size are compensated by rate normalization, task performance also quickly converges to zero for the supercritical values of α .

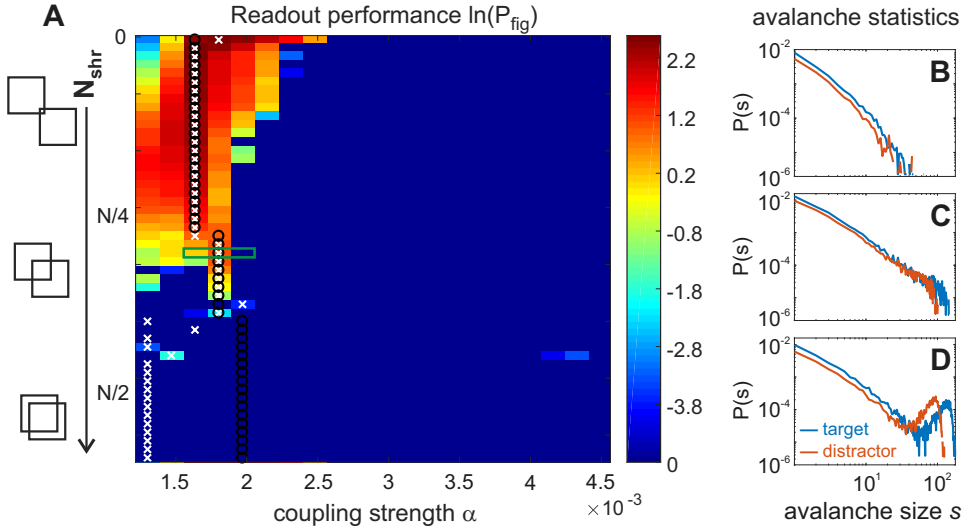


Figure 6.11: **A** The logarithm of the task performance $\ln(P_{fig})$ for the value of the read-out gain factor which globally maximizes task performance $f_{opt} = 0.028$. White crosses mark the performance maxima for each N_{shr} . For most large values of N_{shr} , $P_{fig} = 0$ for all values of α , and for these rows, the white crosses appear at the smallest value of α . The black circles mark $\alpha_{crit}(N_{shr})$ as identified in figure 6.9A. The location of the performance maxima correlate strongly with the critical coupling strength values. **B**, **C** and **D** show the avalanche size distributions in the target and distractor populations for $N_{shr} = 76$ and three consecutive values of α around the critical point (framed by the green rectangle).

In summary, we have found that also in the leaky EHE network, oscillatory activity emerges at the critical point, although more abruptly than in the non-leaky network, resembling a first-order phase transition. We find, nevertheless, the typical

transition of the avalanche size distributions from exponential to supercritical as we increase the coupling strength α , although with larger exponents. Due to the lack of strong hyperpolarization in the background units, the cross-talk between the target and distractor populations is strengthened with the addition of the leaks, which leads to decreased figure-ground segregation performance. In addition, the leaky network displays a large buffer zone between the critical point and the epileptic regime. With the addition of leak currents, we find that, although the independence of the critical value of α in the subnetworks on the overlap between the subnetworks N_{shr} is lost, the critical points of the system still correlate well with the location of task performance maxima, which are more robust against changes in the read-out neuron gain factor f in comparison to the non-leaky network.

6.4 Discussion

In this section we have inspected the dynamics of an analytically well-understood network of EHE units, where mutually excitatorily coupled subnetworks are embedded into the coupling matrix. Each subnetwork can be interpreted as representing a visual object, or figure, defined by an ensemble of local features. As soon as these features are collectively present in a stimulus, the subnetwork will engage in synchronous activity, thus enabling coincidence detectors to optimally signal the presence of the figure when the dynamics are poised close to the critical point. Implementing structured inhibition between subnetworks which share units efficiently prevents cross-talk between neighboring populations, and makes the critical value of our control parameter α robust against changes in the number of shared units N_{shr} as well as the total number of subnetworks N_{sub} . Thus inhibition serves to make computation with critical dynamics structurally invariant, since little to no changes in the existing network topology would be necessary as the network expands and scales up.

This robustness is lost to some extent with the addition of leak currents into the neuron model, due to the increase in cross-talk between subnetworks. However, the maxima of the figure detection task performance display enhanced robustness against changes in read-out parameters such as the LIF neuron gain constant f in the leaky model. Interestingly, the correlation of task performance maxima and the critical points of the system persist, even with the addition of leaks, as the critical point is where oscillatory activity emerges, similar to the leaky model presented in chapter 5.

Taken together, our results suggest that a task requiring fast and accurate integration of spatially extended information may benefit from figure representations in the form of critical subnetworks. When individual feature detectors in such subnetworks are simultaneously driven by a stimulus, large avalanches can cascade through them rapidly and in a synchronous manner. Capitalizing on this, our findings indicate that the critical state will provide the best figure-ground segregation results in a synchrony coding scenario, given realistic resource constraints such as a limited number of neurons, with or without leaks. Overall, these results imply that bottom-up, stimulus-driven modulations of visual information processing, and well-known psychophysical phenomena such as the “pop-out” of more salient stimuli

and suppression of distractors, may benefit from local subnetworks operating close to the critical state.

6.4.1 Robustness of results

As in chapter 5, we limited the models presented in this chapter to minimalistic systems capable of demonstrating the relevant effects. The resulting lack of a large number of free parameters facilitated confirming that our results don't critically depend on a very specific choice of parameter combinations among many.

In the feature integration model, we have run parameter scans over most of the free parameters. We have confirmed that the robustness of the critical coupling parameter against subnetwork overlap N_{shr} in the network with inhibition does not depend on the number of externally activated units N_{act} , subnetwork size N or the number of subnetworks N_{sub} . Concerning the read-out results, we have checked that varying the LIF neuron time constant τ within a biologically plausible range (5-20 ms) does not alter our conclusions qualitatively. However, we found the position of the task performance maxima in the phase space spanned by N_{shr} and α to strongly depend on our choice of Δt , when Δt is varied in isolation. Nevertheless, changes in Δt also heavily influence the value of f_{opt} . Therefore, repeating the golden section search and using the appropriate value of f_{opt} for a given Δt , we observed the performance maxima always close to the critical point (within the range $\Delta t = 0.01$ - 0.02 ms). Task performance results in the leaky network is also qualitatively robust against changes in the leak constant C_{leak} below some upper limit. Since increasing the leak constant, or decreasing the membrane time constant of the simulated units, increases the spill-over of activity from the target population to the background population, the overall task performance decays to zero with increasing C_{leak} .

We have initially used the non-leaky integrator units due to the analytical understanding of how the coupling strength modulates the avalanche size distributions. However, we have confirmed in section 6.3.2, that the addition of leak currents do not change our conclusion that the critical coupling strength for the subnetworks represent the figure-ground segregation performance maxima for a coincidence detection read-out scheme. More complex neuron models (e.g. with non-conservative reset of the membrane potential or conductances) may lead to varying results, although network behaviour may be reproduced qualitatively given small leaks and adjusted parameters [EHE02]. Nevertheless, with the appropriate choice of network topology, which ensures synchronous activity in subnetworks and properly structured inhibition to abate cross-talk between subnetworks, we would expect similar results for the task performance. Our specific choice of the read-out mechanism (LIF with normalized input) is arbitrary and we believe any coincidence detector scheme should produce a similar effect.

Finally, for the quantification of figure detection performance, one can also use a more standard measure such as the d' -value, which would take into account the mean as well as the standard deviation of the rates of the two read-out neurons within discrete time bins. Instead, we use the rate difference between the two neurons

within the $K\Delta t = 200$ s total duration of a simulation. Such a simple approach was preferred in order to avoid introducing new parameters (such as the width of the time bin required for computing the d') into our model. In general, however, when using the read-out gain factor f which globally maximizes performance, one large sized avalanche sweeping the figure subnetwork within an observation time window is sufficient to ensure effective figure detection. Since with our choice of Δt large sized avalanches take place in an oscillatory manner, with roughly 40 – 50 Hz frequency, choosing an observation time window appropriately (with a minimum width of 20 – 25 ms) should reveal optimal d' -values.

6.4.2 Physiological Plausibility

As stated above, in this chapter we restricted ourselves to relatively simple models in order to ensure the demonstrated effects do not require fine-tuning of a large number of free parameters. In general, many modelling studies which similarly put forward computational benefits of criticality rely on comparatively abstract measures or are restricted to non-spiking neuron models or homogeneous, unstructured network models [BN04; KC06; Lan90; SP13]. In order to better understand the dynamics of the awake, active cortex, where such computational benefits will presumably be most conspicuous, we explore the dynamics of highly structured networks of spiking neurons, with and without inhibition and leaks, while performing realistic tasks, with biologically plausible read-out mechanisms.

As a read-out mechanism in the feature integration model, we used coincidence detection. We normalized the inputs into the read-out neurons such that the input from the neuronal population processing a figure and the input from the neuronal population processing the distractors in the background would be indistinguishable using simply the rate information. This is equivalent to adjusting stimulus contrast, which contour detection performance is known to be resistant against [HHF03]. In addition, although firing rate is a good discriminator in our specific, simple model, in other models which incorporate higher biological complexity, such as detailed balance [VSZ+11] or inhibitory normalization [CHM97], rates might carry little to no information useful for figure detection. Taken together with our results, this may imply that the cortex transmits local stimulus information using rates, and retains information about global object configurations in the time domain, thus implementing a multiplexed coding scheme.

As already mentioned in chapter 4 and section 5.4, the structure of neuronal connectivity in the brain is not as homogeneous or symmetric as considered in the models presented in this chapter. Nevertheless, there is evidence that the statistics of correlations between edge elements found in natural images capture human contour detection performance reasonably well [GPS+01]. Therefore, we believe modelling individual, local feature detectors collectively processing an object, with features which are likely to co-occur in the visual field, as integrator units with strong excitatory lateral connections to each other is fair in terms of the behavioural outcome. In addition, the use of local inhibition in order to circumvent strong global synchrony resembling seizure dynamics is in line with neuroscientific literature,

where it is believed that seizures emerge as a consequence of the breakdown of surround inhibition in the cortex (see e.g. [KSJ+00] Chapter 46). Still, in our models, the 1-dimensional and symmetric coupling structure may be replaced by a more random assignment of units to multiple subnetworks or learned dynamical attractors with strong recurrent excitation within a subnetwork. With this altered topology, we would still expect a coincidence detector to display high performance around the critical point, as long as there's mutual inhibition between units belonging to different subnetworks. This is because the most important ingredient for the maximized figure detection performance with coincidence detection is synchronous activity within a target subnetwork which is absent between random background units.

Finally, our investigations raise one important question: Is our scheme only useful for a single stage of processing, or would it still be applicable in a multilayer model? This was not explored in this chapter, however, one important theoretical property of critical networks is their branching parameter being unity, often thought of as being beneficial for dynamics where activity should neither die out nor explode over time. This idea is reinforced by recent investigations which show that arbitrarily deep neural networks may only be trained close to criticality [SGG+16].

Contour

7. Contour Integration in a Realistic LIF Model

In chapters 5 and 6, we have presented rather minimalistic models, where we were able to scan a large portion of the parameter space. However, one of the main goals of this dissertation, as also stated in the abstract and the introduction (Chapter 1) is to improve the biological realism in models of criticality in the cortex, while establishing more explicit links to function. In this chapter we aim to achieve this both by presenting a model with improved biological details, and by making more explicit links to a well-known psychophysical, behavioural task. In terms of the network model, we will use a biologically inspired connectivity structure, and demonstrate phase transition dynamics in local, functional subnetworks in the presence of a suitable stimulus. While in the previous chapter we have focused on feature integration as a general mechanism, and mentioned contour integration as only one example that our model might apply to, in this chapter we are going to focus concretely and particularly on contour integration, using realistic stimuli resembling those from experimental studies.

7.1 Introduction

The main research question which motivated the work presented in this chapter is whether or not the functional benefits of critical dynamics presented in the previous chapter would generalize to more realistic settings. Specifically, we strive to improve the biorealism of the model by employing

- a large scale network of leaky integrate-and-fire (LIF) neurons,
- more realistic anatomical structures, such as columnar organization involving retinotopic hypercolumns,
- structured recurrent connectivity inspired by Gestalt laws and the association field idea, and
- a concrete behavioural task, using realistic stimuli similar to those used in

psychophysics experiments (Fig. 6.1A).

In general, this chapter expands upon the idea of bottom-up modulation of activity in a feature integration task, as presented in the previous chapter. To that end, chapters 6 and 7 share many of their basic assumptions. These assumptions include each neuron acting as a feature detector and the large neuronal avalanches taking place within a subnetwork being a neural correlate of the "pop-out" effect, as described in chapters 4 and 6.1.

We hypothesize that as long as certain aspects of the model are maintained—including emergent critical activity within functional subnetworks co-processing an object in the visual field, mutual synchronization in subnetworks modulated by stimulus properties, and a read-out based on correlated activity in subpopulations—it should be possible to qualitatively replicate our results from chapter 6.

In terms of the neuron model, we improve upon the EHE units from the previous chapter in terms of biological complexity, by using LIF neurons, similar to the model presented in chapter 5. In terms of the network, we will present the largest model we've used in this dissertation so far, with system size approaching 100,000 units. Similarly, we improve upon the implementation of inhibition from the previous chapter: the model in this chapter obeys Dale's law and consists of local interneurons along with pyramidal neurons responsible for the long-range lateral connections [KSJ+00]. In terms of the network model, we employ a biologically inspired model of columnar organization, where neurons are grouped into orientation columns based on their orientation selectivity [DA01; HW68; KSJ+00], as described in chapter 4.

Recurrent connectivity of the network is highly structured and inspired by the 'association field' (AF) idea from psychophysics. Initially proposed by Field, Hayes and Hess in 1993 [FHH93], AF hypothesis proposes a functional relationship between orientation selective neurons with colinearly or cocircularly aligned receptive fields (Fig. 7.1), which aids in integration of local features into contours.

The main suggestion of the AF hypothesis is that there exists mechanisms in the visual system, which facilitate the perceptual grouping of features based on Gestalt principles (see section 4.2), including whether edge elements are located close together in the visual field or show good continuity [WEK+12]. The AF hypothesis has gained support through the AF-like organization of edge elements in natural images [GPS+01] and has been investigated in many experimental follow-up studies (see e.g. [HHF03] for an overview), as well as in theoretical models of contour integration [EMS+12; Li98; YF98]. In this chapter, we adopt an approach similar to those in the referenced modelling studies and define our connectivity structure based on the AF hypothesis. Different to previous modelling studies, we stochastically sample individual synaptic connections. The randomly assigned couplings give rise to a directed graph, and improves upon the symmetrically coupled subnetworks from the previous chapter, while still maintaining functionally grouped subpopulations. In comparison to the previous chapter, in this chapter, these randomized subpopulations are considered to be the neuronal representations of 'figures' or more specifically contours which can appear in a visual stimulus.

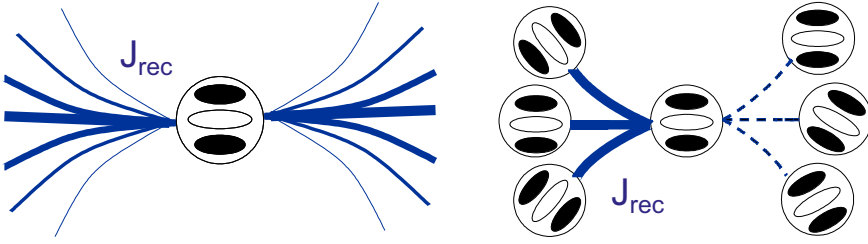


Figure 7.1: Schematic of the ‘association field’ type connectivity. The strength of lateral connections between feature detectors (blue lines) depend on whether the classical receptive fields of the detectors are placed in a colinear or cocircular fashion (left). For example, the detectors of features which are more likely to co-occur in a contour strongly interact with each other (right, solid lines), whereas for features which are less likely to encode a shape together, interaction strength decreases (right, dashed lines). See also [FHH93] and [HHF03] for further visual descriptions and in depth information.

Historically, band-pass elements, such as Gabor patches, have been used as edge elements in psychophysical studies of contour integration [FHH93; GGE17; HHF03; KJ93; MW11; MK05; MF07]. This is because, based on the response properties of cells in the primary visual cortex (V1), the use of Gabor patches serves two important purposes. First, they strongly excite neurons which are known to have Gabor-shaped receptive fields and are orientation selective [HW68]. Second, the use of band-pass elements limits the spatial scale, along with the spatial spectrum, of the investigations, making it easier to ascertain that the observed effects are not due to a complex interaction of cells which are responsive to input from a wide range of orientations and spatial frequencies [FHH93]. In order to closely follow the experimental procedures, we also use a realistic stimulus to drive our network in this chapter. In addition, V1 cells are often thought of and modelled as linear filters [DA01]. Similarly, we model the external drive as the result of a filtering operation with an appropriately oriented filter.

As mentioned in section 4.2, certain figures in a cluttered scene will “pop-out” to the observer due to its characteristics. This is sometimes also termed exogenous attention due the bottom-up nature of the effect and how such figures tend to draw the attention of the observer [ZZZ+12]. In general, how strong such bottom-up “pop-out” effects are and how well an observer can detect the presence of a figure is determined by the so-called salience of the figure [YF98]. Many psychophysical variables of salience, such as inter-element distance and contour closure, which strongly modulate contour detection have been identified in experimental studies [HHF03; KJ93; LPG06; MF07]. Specifically, contour detection performance have been shown to improve, trivially, with increasing contour length [LG02; LPG08] in humans and monkeys. Similarly, it has been observed that contours with large orientation jitter (which quantifies how well the edge elements forming a contour are aligned along the contour path) are harder to detect for subjects [FHH93; MW11; MK05].

Motivated by such studies, we will also investigate the effects of these two salience parameters, namely, contour length and orientation jitter, on contour integration in our model. We hypothesize that if emergent criticality in subnetworks plays a functional role in facilitating contour integration, there should be a monotonous dependency between contour salience and correlated activity within the subnetwork processing the contour.

Finally, in accordance with the observation that contour integration, both in humans and monkeys, is known to take place rapidly [EMS+12; GGE17; HHF03; MK05](e.g. correct response rates can reach 80% in a 4AFC task, for a stimulus presentation duration of only 70 ms in macaque monkeys), we will also aim in our model to reach high contour detection performance with short stimulus presentation times.

7.2 Results

7.2.1 Generation of stimuli

In line with psychophysical experiments of contour integration, we begin by generating a realistic stimulus for a contour integration task. Similar to our approach in chapter 6, the stimulus we present to our model network consists of one target ‘figure’, which throughout this chapter is a contour made-up of smaller, colinearly aligned edge elements, and some randomly oriented distractor elements. As is typical in psychophysics [HHF03], the edge elements i comprising the contour and background are generated using a 2-dimensional Gabor function

$$G_i(x, y) = A_G \exp\left(-\frac{r}{2\sigma_G^2}\right) \cos\left(\phi_i + 2\pi \frac{x'}{\lambda}\right) \quad (7.1)$$

with

$$r = (x - k_x^i)^2 + (y - k_y^i)^2 \quad (7.2)$$

and

$$x' = (x - k_x^i) \cos(\theta_i) + (y - k_y^i) \sin(\theta_i); \quad (7.3)$$

where A_G is the amplitude and σ_G is the standard deviation of the Gaussian envelope of the Gabor function, ϕ_i is the randomly assigned spatial phase, λ the spatial frequency and θ_i the orientation of each element. To generate the full stimulus $S(x, y)$, a total of n_e Gabor patches i centered at (k_x^i, k_y^i) were added together

$$S(x, y) = \sum_{i=1}^{n_e} G_i(x, y) \quad (7.4)$$

and the resulting sum defined the brightness of a pixel at a given position (x, y) on the stimulus. Finally, the stimulus was rendered in grayscale (Fig. 7.2).

Due to computational concerns, we chose a diagonal square lattice placement of the Gabor elements on the full stimulus (also called a discrete hexagonal grid in [Li98]). This allowed us to place more non-overlapping elements in total and

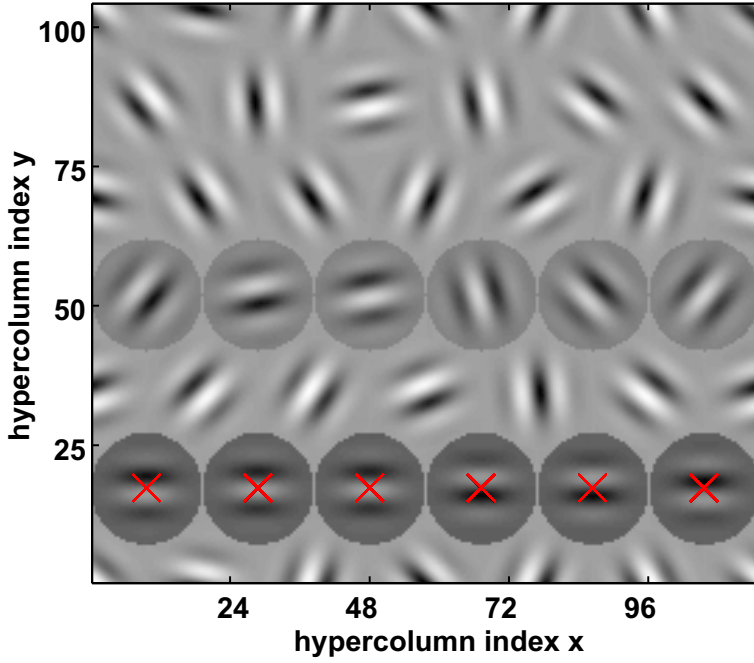


Figure 7.2: An example of a realistic visual stimulus generated using edge elements defined by Gabor patches. The stimulus contains a colinearly aligned contour, formed by 6 edge elements marked with red crosses, and randomly oriented background elements. We use such stimuli to externally drive a network of neurons organized into hypercolumns with classical receptive fields centered on each pixel of the stimulus. As in chapter 6, we record the activity of a target ensemble, with RFs located on the contour elements (circles shaded in dark gray), and a distractor ensemble, with RFs located on an equal number of randomly oriented background elements (circles shaded in light gray).

experiment with longer contour lengths, while keeping the pixel count of the stimulus, and consequently the size of the model network, in check. Thus, the final stimulus consisted of $n_x = 120$ by $n_y = 104$ pixels and $n_e = 36$ edge elements in total, organized into 6 rows and placed with periodic boundary conditions (as shown in Fig. 7.2).

The orientations of all the edge elements constituting the ‘target’ contour are assigned, without loss of generality, to be horizontal ($\theta_{i \in \text{contour}} = 0$) and for the background elements the orientations are allocated randomly. As mentioned above, the spatial phase ϕ_i for each element, including the contour elements, is picked at random.

7.2.2 Network model and single neuron dynamics

In terms of the network topology, we have so far considered random graphs in chapter 5 and constructed a more structured, but very simple and highly symmetric

model of subnetworks in chapter 6. In order to improve the biorealism, in this chapter we introduce a network structure consisting of $H = n_x \cdot n_y = 12480$ orientation hypercolumns, each of which consists of a population of neurons with their classical receptive fields (RFs) centered on one pixel of the visual stimulus (see Fig. 7.2). Each hypercolumn is made up of $M = 32$ orientation columns, where the preferred orientation of the neurons in each orientation column is equidistantly distributed over the half-circle. However, in order to be able to sweep a large parameter space within a reasonable amount of time, we did not sample neurons from all M orientation columns in each hypercolumn. Instead we randomly sampled $n_o = 8$ neurons with different preferred orientations in each hypercolumn, taking care that each of the 32 possible orientations are equally represented in the total network (see section 7.3 for details).

The neurons in each orientation column j are externally driven by the stimulus, based on the location of the center of their RF and their preferred orientation. In order to implement the external drive, we first perform a linear filtering operation on the stimulus M times, using complex Gabor energy filters (see e.g. [GPW03]) with orientations θ_j matching those of the preferred orientations of the M columns, and obtain a 3-dimensional filtered stimulus $S_{filt}(x, y, \theta_j)$.

The filtering with the Gabor energy filters (described in detail in section 7.3) is equivalent to summing the squared responses of a pair of quadrature Gabor filters (separated in phase by $\pi/2$) and yields a result similar to the complex cell energy model [HH02; SB13], which is phase invariant. The filtered stimulus is then normalized between 0 and 1.

$S_{filt}(x, y, \theta_j)$ is subsequently used to define the strength $J_{ext}(x, y, \theta_j)$ of the external input to the neurons with their classical RFs centered on (x, y) and with preferred orientation θ_j .

$$J_{ext}(x, y, \theta_j) = S_{filt}(x, y, \theta_j)(1 - J_{thr}) + J_{thr} \quad (7.5)$$

where J_{thr} is an offset used to mimic ongoing activity and increase background noise in the system by making sure all neurons in the network receive non-zero activating input on average.

The stochastic feedforward input I_{ext} is finally modelled as a normally distributed random variable $\xi \sim N(\mu_{ext}, \sigma_{ext}^2)$ with mean $\mu_{ext} = 0.7$ and standard deviation $\sigma_{ext} = 20$, scaled by the external input strength J_{ext}

$$I_{ext} = J_{ext} \cdot \xi. \quad (7.6)$$

With a very large variance (σ_{ext}^2) and the distribution being symmetric around a relatively small mean (μ_{ext}), the external input was frequently inhibitory, mimicking the noisy but balanced input reported for cortical neurons [OL08].

The network is made up of 20% inhibitory (interneurons) and 80% excitatory (pyramidal) neurons. All excitatory synaptic connections had the uniform efficacy w_{exc} and all inhibitory connections had the uniform, on average balancing efficacy $w_{inh} = 4w_{exc}$. However, the topology of the connections differed greatly between the excitatory and inhibitory neurons: while the excitatory couplings were heavily

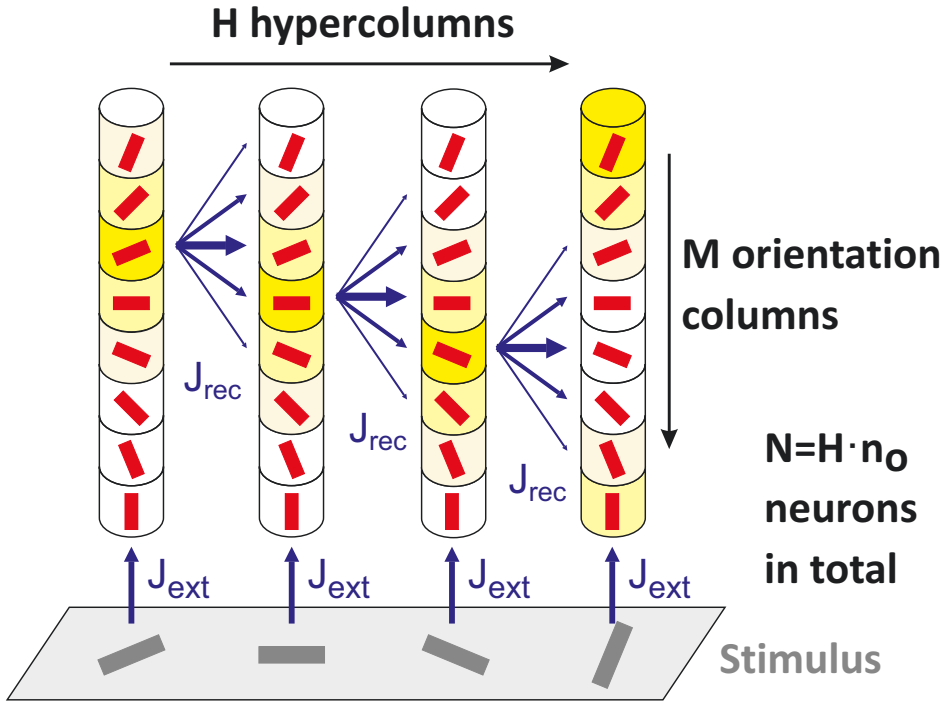


Figure 7.3: Simple sketch of the network structure. Neurons in each orientation column receive external input according to the location of their receptive field and preferred orientation (the magnitude of the external input represented by the yellow shading). The lateral coupling structure is based on the association field idea, as explained in section 7.1, which preferentially connects orientation columns with colinearly and cocircularly aligned RFs.

structured and preferentially connected neurons which would be co-activated by the presence of a contour, the inhibitory couplings were local and decayed with distance between the RF locations and difference of preferred orientations (see section 7.3 for details).

The recurrent connectivity structure for the excitatory neurons is based on the association field idea, as described in section 7.1. The coupling matrix was constructed using a fixed number $n_s = 500$ of outdegrees per neuron (both excitatory and inhibitory). The n_s outgoing synaptic connections each excitatory neuron made were stochastically sampled using a probability distribution p_s which took into account the distance between the receptive fields of pairs of neurons, as well as their relative location and orientation (for formal details, see section 7.3). The probability of connections were therefore translationally, rotationally and reflectionally invariant similar to other models of contour integration (see e.g. [Li98]), however, asymmetry was introduced to the topology by the random sampling of synapses for each neuron. The resulting structure of the excitatory couplings preferentially connected neurons with colinearly or cocircularly organized receptive fields (Fig. 7.4).

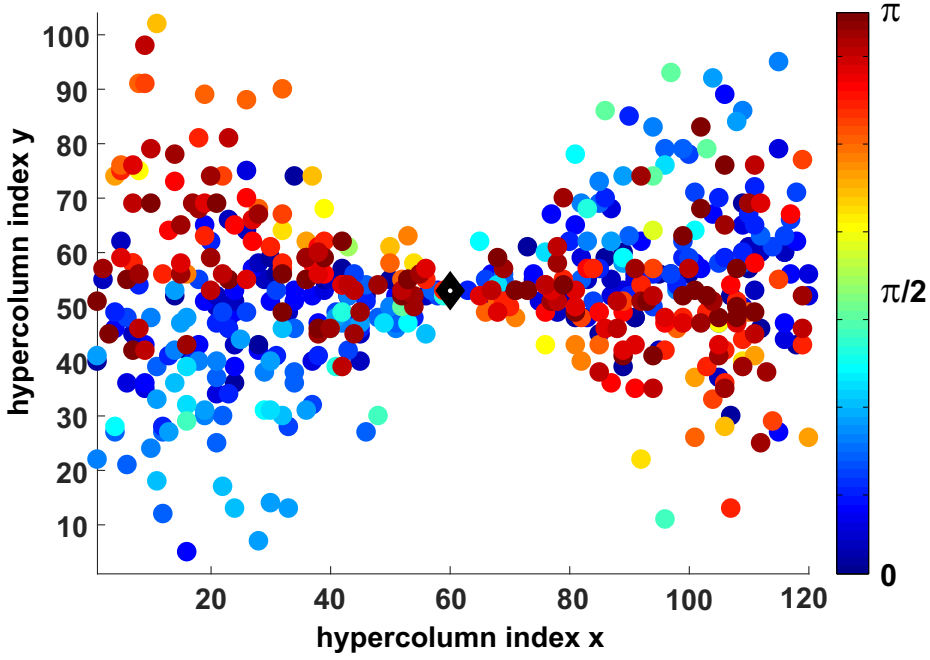


Figure 7.4: The stochastically sampled synaptic connectivity for one excitatory neuron with its RF center (black diamond) close to the center of the stimulus and preferred orientation $\theta_j = 0$. The colored circles mark the locations of the RF centers of all the $n_s = 500$ neurons postsynaptic to the center neuron. The color axis shows the preferred orientation for each postsynaptic neuron. (See section 7.3 for formal details of the probability distribution used for sampling the shown connections.)

Similar to the excitatory neurons, n_s synaptic connections each inhibitory neuron made were stochastically sampled using a probability distribution p_s^{inh} . For the inhibitory neurons, the connectivity structure was local, and decayed relatively fast with increasing distance between RFs of neurons and difference of preferred orientations (see section 7.3 and figure 7.12 for details). In accordance with the stimulus, network connections assumed periodic boundary conditions.

As the neuron model, we used the leaky integrate-and-fire (LIF) model (Eq. 7.21) similar to the model in chapter 5. However, in contrast to the previous model, where the external input was a Poisson process with nonnegative values, in this model the stochastic external input is a normally distributed random variable with high variance and a mean close to zero (Eq. 7.6). Hence, the evolution of the membrane potential was noisier and the spiking activity much more irregular (Fig. 7.5), consistent with experimental recordings of V1 and MT neurons [SK93].

Finally, similar to the approach in chapter 6, we separately recorded the spiking activity of a target ensemble, which included all the neurons in the hypercolumns with RFs located on the contour elements, and a distractor ensemble, with RFs located on an equal number of randomly oriented background elements (visualized

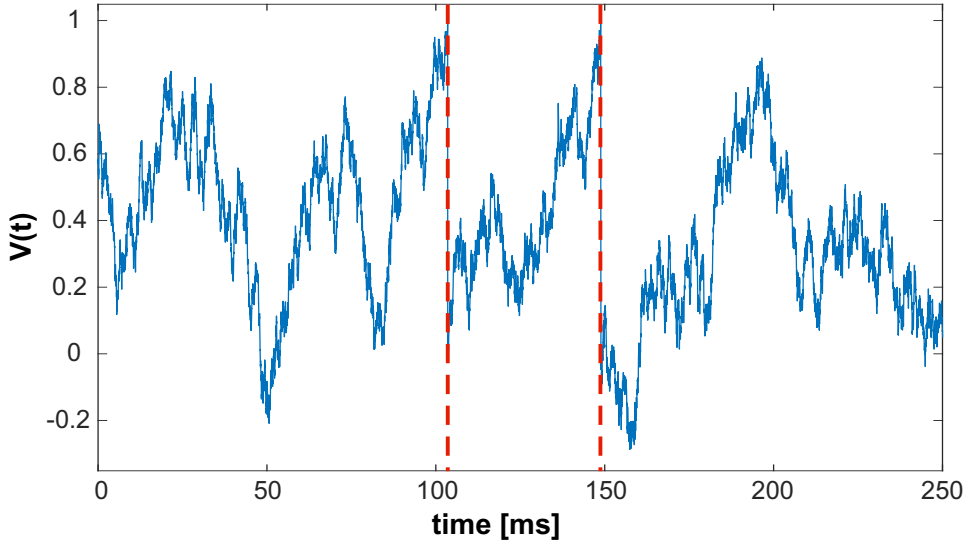


Figure 7.5: Example of a membrane potential time series $V(t)$, for a neuron with its receptive field centered on a contour element. Spike times are marked by the red dashed lines. Because of the noisy external input, as well as the local recurrent inhibition, spiking activity in the network was relatively irregular. In the low coupling strength (subcritical) regime (as illustrated in this figure), the average Fano factor of spiking events of all neurons within the duration of a trial (255 ms excluding the initial transient period) was $FF_{\text{sub}} = 2.75$. In the high coupling strength (supercritical) regime, it could decrease down to a minimum of $FF_{\text{sub}} = 0.5$ depending on the neuron.

by the gray shaded circles in Fig. 7.2).

7.2.3 Network dynamics

We initially investigate network dynamics during the presentation of a stimulus containing one contour, made up of six Gabor elements, and 30 randomly oriented background elements (Fig. 7.2). We inspect network activity as we vary the synaptic strengths, and for each set of coupling parameters, we run the simulation for $N_{\text{stim}} = 100$ independently generated stimuli. As expected, we observe enhanced synchrony in the target ensemble as we increase the coupling strengths in the balanced regime $w_{\text{inh}} = 4w_{\text{exc}}$. Figure 7.6A shows how the avalanche size distributions, averaged over all stimuli, transition from exponential to supercritical with increasing synaptic efficacy.

As in chapter 5, we numerically identify the critical value of the excitatory coupling strength $w_{\text{crit}} \approx 0.042$ by finding the value of w_{exc} which minimizes the KS statistic of the avalanche size distributions. Figure 7.6B shows that in contrast to the target population, the avalanche size distributions of the distractor population do not reach a local minimum within the range of w_{exc} scanned in our simulations.

Avalanche statistics suggest that this is because the dynamics in the distractor

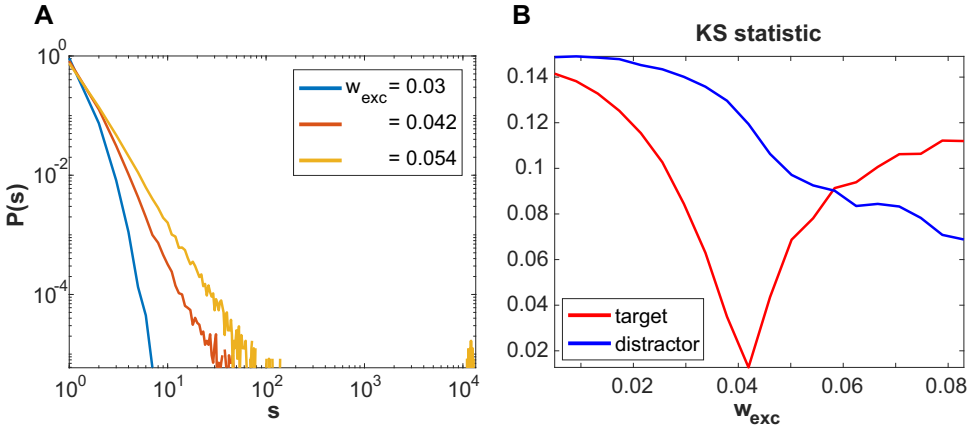


Figure 7.6: **A** Distributions of avalanche sizes s in the target ensemble with increasing inhibitory and excitatory synaptic strength $w_{inh} = 4w_{exc}$. With low coupling strength (blue) the distributions are exponential. Increasing the synaptic strength leads to increased correlations (red) and finally to the emergence of a bump at the tail of the distribution (yellow). **B** KS statistic for the avalanche size distributions display a clear local minimum for the target ensemble (red), but not for the distractor population (blue), within the scanned range of coupling strengths.

population are relatively asynchronous, even when synaptic strengths are large (Fig. 7.7A). In other words, similar to our results in chapter 6, we find that, also in this more complex setting, the dynamics of the units which are processing the colinearly aligned contour elements are more strongly correlated than the dynamics of the units processing the random background elements. Consequently, the avalanche size distributions of the target ensemble approach power-law scaling sooner than the distributions of the distractor ensemble, as we globally increase the coupling strength w_{exc} (Fig. 7.7).

This result is purely a consequence of our stimulus statistics appropriately matching the internal subnetwork structure. In other words, had we presented the network with a contour spanning the RF locations of the distractor population, the results would be reversed. This flexible (re)organization of correlation structures within subnetworks is what we call the top-down modulation of activity and it allows for accurate and dynamic detection of contours by downstream visual areas.

We see a similar effect in the dynamics as we vary the salience of the presented contour, by tweaking two psychophysical parameters: contour length and orientation jitter. We vary the size of the contour from a minimum of two adjacent, colinearly aligned Gabor elements up to the full size of six elements. For contour lengths smaller than six, the target population was composed only of the neurons in the hypercolumns with RFs centered on the contour elements. Similarly, the number of neurons in the distractor population was reduced to match that of the target population. Orientation jitter was defined, as is typical in psychophysics experiments [FHH93; MK05; MF07], as a $\pm\alpha_{jitter}$ deviation of the orientations of

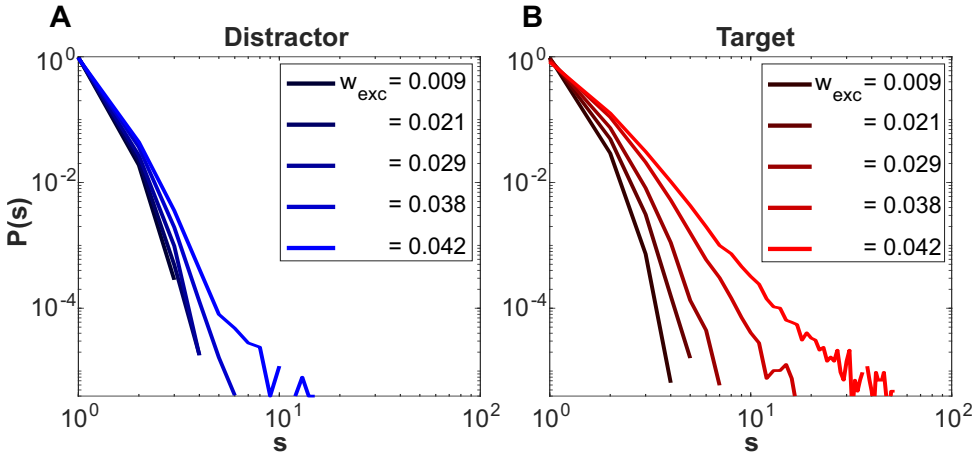


Figure 7.7: Emergence of distinct dynamics in **A** the distractor ensemble (blue) and **B** the target ensemble (red), as we vary the coupling strength w_{exc} (different color shades). The target subnetwork becomes mutually synchronized and its avalanche size distributions $P(s)$ approach a power-law faster than those of the distractor ensemble as we increase w_{exc} , as a consequence of stimulus statistics.

individual Gabor elements from the contour path. In our case the contour path was a horizontal line, and the orientations of the contour elements were randomly jittered by $\alpha_{jitter} \in \{7.5, 15, 22.5, 30, 37.5, 45\}$ degrees.

In experimental studies, increasing the length of the contour or decreasing the orientation jitter of the contour elements is known to increase the salience of the contour, and improve target detection performance in human and primate observers [FHH93; KJ93; LPG06; LPG08; MK05]. We find that, in our model, the level of correlated activity within the target subnetwork is directly modulated by the salience of the target contour. Figure 7.8 shows that the avalanche size distributions in the target ensemble approach a power-law, both with increasing number of contour elements, and with decreasing orientation jitter.

7.2.4 Task performance

Next, we investigate two putative read-out schemes for figure-ground segregation, one using the relative rates of the two ensembles, and one using the level of synchrony. The rate-based task performance is given simply by the difference of the average firing rate in the target and distractor populations (averaged over all neurons in each population and all N_{stim} stimulus presentations). For computing the synchrony-based task performance, we feed the output spikes from the target and distractor populations separately to two different LIF read-out neurons (see section 7.3 for details). Similar to our approach in chapter 6, we use the difference of the output rates of the read-out neurons to quantify the synchrony-based read-out performance. Hence, both measures of the task performance are given in units of Hz.

Figure 7.9 shows the contour detection task performance, using both the rate

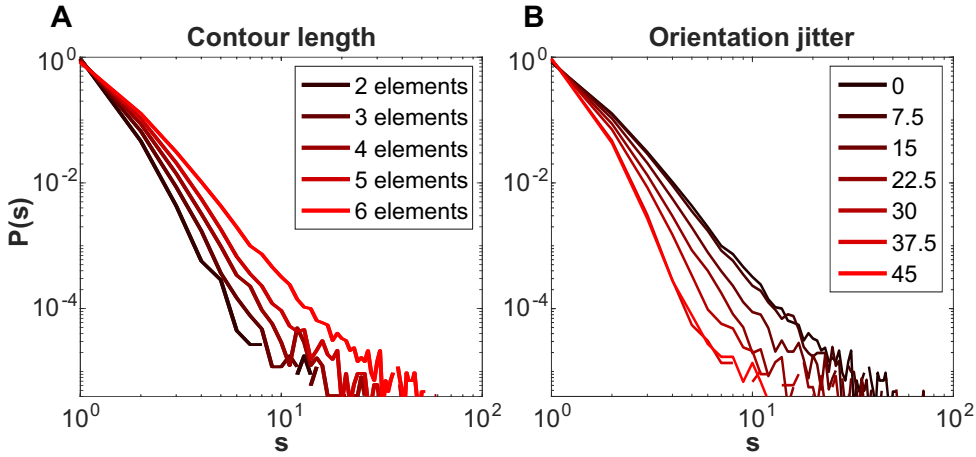


Figure 7.8: **A** Avalanche size distributions in the target ensemble with varying contour length, as measured by the number of colinearly aligned Gabor elements (different color shades). **B** Avalanche size distributions in the target ensemble with varying orientation jitter α_{jitter} of the individual contour elements (different color shades). In both **A** and **B** $w_{exc} = w_{crit}$, and we find that synchronous activity is enhanced with increased contour salience.

code (left) and the synchrony code (right) measures, as a function of the coupling strength w_{exc} and contour length. The rate code does not display a clear dependence on either parameter, except in the subcritical region and for contour lengths larger than 3 elements. Part of the reason for this is that, due to the strong local inhibition, the rates of the target and distractor populations were very similar to each other and differed on average by less than 1.5 Hz for all parameter combinations (Fig. 7.9A, y-axis). The firing rate of an individual neuron i with its RF close to the center of a contour element, and preferred orientation close to $\theta_i = 0$ varied between 15 and 40 Hz for individual stimulus presentations. The noisiness of the task performance in the supercritical regime stems from the fact that with larger w_{exc} and w_{inh} , the irregularity of firing in both populations increased and the stimulus presentation time of $T_{total} = 280$ ms was not long enough to obtain a reliable average. Therefore, the population rate contained little information about the presence of a contour within the RFs of a given population.

In contrast, the task performance based on the LIF read-out was strongly modulated by changes in w_{exc} and contour length (Fig. 7.9B). Interestingly, for the full length contour made up of six Gabor elements, the task performance was maximized at the critical coupling strength w_{exc} . For shorter contours, the maximum performance was slightly shifted to supercritical values. This is expected since the level of synchrony in the target population decreases with contour length (Fig. 7.8A) and therefore larger values of w_{exc} are required for a power-law scaling of avalanche dynamics.

Even though the given rate differences for the LIF read-out neuron are unnaturally

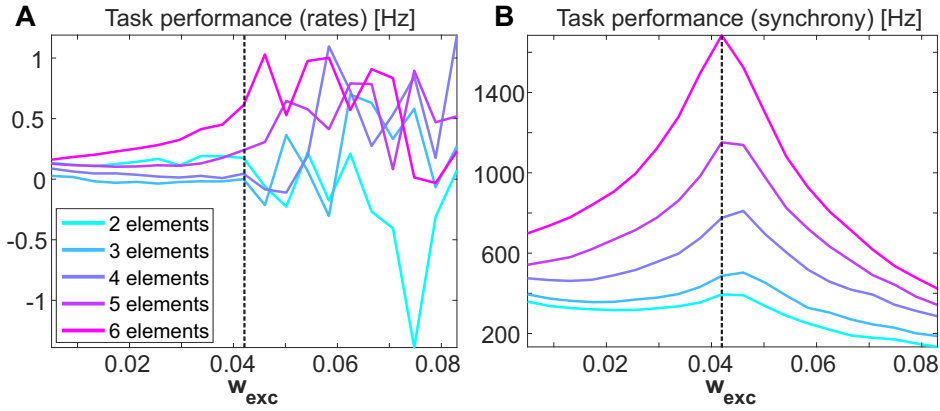


Figure 7.9: **A** Contour detection task performance using the rates of the target and distractor populations. The average rate difference between the two populations is small, and the dependence of task performance on the coupling strength w_{exc} or on the contour length is ambiguous. **B** Contour detection task performance using the level of synchrony in the target and distractor populations. There is a clear dependence of the synchrony-based contour detection performance on the contour length (different colors) and w_{exc} . For the full contour size of six Gabor elements, performance is maximized at the critical value of the coupling strength $w_{exc} = w_{crit}$. In both **A** and **B**, the dashed black lines mark the critical point w_{crit} .

large, this is because, in order to be consistent with chapter 6, we use a gain factor f_g for the read-out neurons, and report here the results obtained for the value of f_g which globally maximizes task performance. However, we have confirmed that the read-out neuron rates can be lowered to physiologically meaningful levels by using smaller values of f_g , and the qualitative results presented here, such as maximized task performance around w_{crit} , persist (not shown). Similarly, task performance landscapes almost identical to the ones shown in Fig. 7.9 are obtained when the task performance is plotted for different orientation jitters instead of contour length (not shown).

Next we obtain the psychometric curves based on our two (rate-based and synchrony-based) task performance measures, using the critical coupling strength w_{crit} . Figure 7.10 shows the psychometric curves for both contour length and orientation jitter. For the synchrony-based read-out we find that the task performance as a function of contour length resembles the relatively linear increase found in experiments (see, for example, figures 2 and 5 in [LPG06], and figure 5 in [LPG08] for psychometric curves obtained in macaque monkeys, as well as figure 4 in [LG02] for human subjects). Even though such psychometric curves typically display a relatively linear dependence on contour length for a large range of medium contour sizes, they are usually fitted using sigmoid functions [LG02] due to the decreasing slope at very small and very large contour lengths. The increase of slope we observe at full contour length in Fig. 7.10A can be attributed to the effects of contour

closure. The observation that as the shape of a contour approaches a closed circle, it becomes more detectable than open contours with similar curvatures has been well-documented in literature [KJ93; MF07] and with psychometric curves displaying a similar increase in task performance slope for large contour lengths (as an example, see figure 4 in [KJ93]).

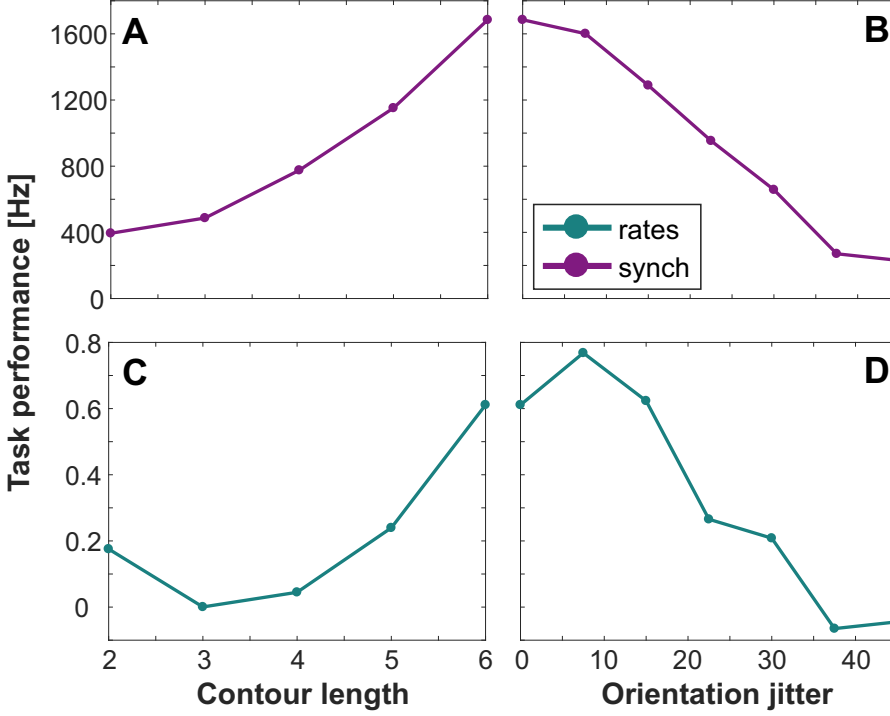


Figure 7.10: Psychometric curves in dependence on the saliency parameters contour length (A and C) and orientation jitter (B and D), using the read-out schemes based on rates (blue) and based on synchrony (purple). In all plots, task performance is displayed for the critical coupling strength w_{crit} . As expected, task performance generally increases with increasing contour salience, and the synchrony-based read-out scheme reproduces the experimentally observed curves reliably.

Similarly, the synchrony-based psychometric curve for orientation jitter (Fig. 7.10B) assumes a sigmoidal shape, as typically reported in experiments (for a comparison see, for example, figure 3 in [MK05] for curves obtained in macaque monkeys and figure 7 in [FHH93] and figure 5 in [MW11] for human subjects).

The psychometric curves obtained using the rate-based read-out (Fig. 7.10C-D) display similar overall shapes for both saliency parameters. However, as mentioned before, due to the noisy external input and the strong local inhibition in the recurrent connectivity, the rate differences between the target and distractor populations are small. Hence, the trial-to-trial variance of this measure is relatively large and it does not contain information about the presence of a contour as reliably as the relative spike times of the neurons within a population. Therefore, the psychometric curves

obtained using the rates of the simulated populations is somewhat noisy, although this might be improved with longer stimulus presentation times.

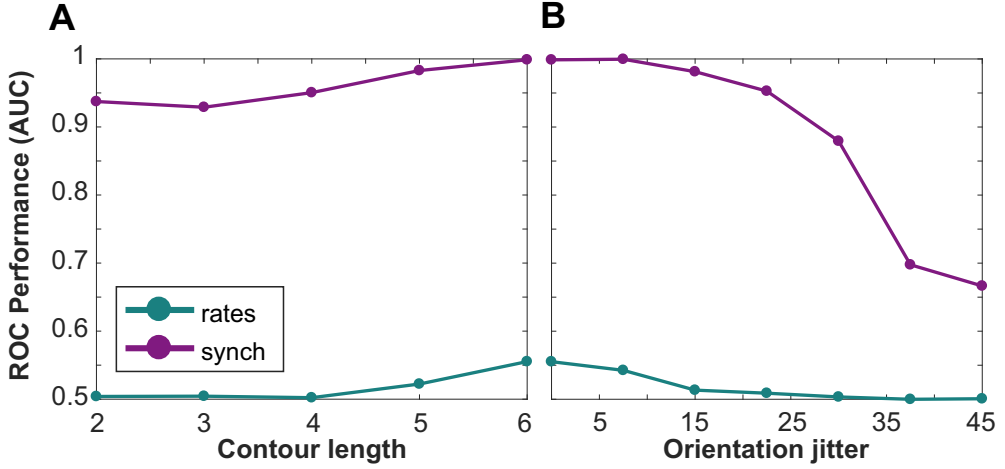


Figure 7.11: Area under the curve (AUC) of the trial-based ROC analysis using the read-out schemes based on rates (blue) and based on synchrony (purple). The ROC performance is plotted in dependence on the salience parameters **A** contour length and **B** orientation jitter. In both plots, task performance is displayed for the critical coupling strength w_{crit} . Synchrony-based read-out outperforms contour detection based on rates for all parameters.

Finally, in order to compare the rate- and synchrony-based contour detection performance measures to each other, we perform an ROC analysis. The ROC analysis using the rate code evaluates the distributions of the mean firing rates within the target and distractor populations over N_{stim} stimulus presentations. The ROC analysis using the synchrony code instead evaluates the output rates of the dedicated LIF read-out neuron of each population.

Figure 7.11 shows that at the critical point ($w_{exc} = w_{crit}$), the synchrony code greatly outperforms the rate code for all salience values. For both read-out schemes, we find that, as before, contour detection performance generally increases with increasing contour salience. For the synchrony-based scheme, we see that the task performance curves have somewhat different shapes than those shown in Fig. 7.10. However, we find a good match between our quantitative results for varying orientation jitter and human performance reported in experiments (see Fig. 5 in [MW11]).

The exact contour detection performance values as a function of contour length, however, are much larger than what is typically measured in experiments. This might be due to the smaller size of our stimulus and our consequent parameter choices, such as the ratio between the inter-element distance of the Gabor patches and the width of the association field used to determine recurrent connectivity. Similarly, as mentioned above, we use here the value of the read-out neuron gain factor f_g which globally maximizes task performance, and yields biologically unrealistic firing rates

for the read-out neurons, which may contribute to the inflated task performance at small contour lengths. Therefore, by tweaking such parameters, it may be possible to find a biologically realistic regime which may also quantitatively reproduce the task performance for varying contour length.

7.3 Methods

7.3.1 Generation of stimuli

In order to have some scaling consistency in our model between

- the visual stimulus,
- the receptive field kernels used to filter the stimulus, and
- the retinotopic mapping of the orientation hypercolumns

we have defined all our parameters in units of the visual angle. The elementary distance between two neighboring pixels of the stimulus in x and y dimensions was $dx = dy = 0.0275$ degrees of visual angle.

For the generation of stimuli, we used Gabor patches as given in Eq. 7.1, with parameter values $A_G = 127$, $\sigma_G = 0.1/dx$, and $\lambda = 0.2/dx$. A_G is chosen such that the background of the stimulus would have the center value of a grayscale colormap in RGB format, with values in the range $[0, 255]$. The spatial phase of the Gabor patches are randomly and independently assigned for each Gabor element i in each realization of a stimulus, using a uniform distribution on $\phi_i \in [0, 2\pi]$. Similarly, the orientation of the background elements are pulled randomly and independently $\theta_{i \in \text{background}} \in [0, \pi]$.

The individual Gabor patches were placed on the complete stimulus using the sum given in equation 7.4. For display purposes with the MATLAB `imagesc` function, a constant value of $c_{img} = 128$ was added to each pixel value $S(x, y)$. As explained in section 7.2.1, the resulting stimulus $S(x, y)$ consisted of one colinear contour (made up of 6 Gabor elements with horizontal orientation) and 30 randomly oriented background elements, rendered with periodic boundary conditions in $n_x = 120$ by $n_y = 104$ pixels. This corresponds to a stimulus size of 3.3 by 2.86 degrees of visual angle.

7.3.2 Network structure and feedforward input

As mentioned in section 7.2.2, the network was made up of $H = n_x \cdot n_y = 12480$ orientation hypercolumns. Even though we assume the existence of $M = 32$ orientation columns in each hypercolumn, for computational efficiency, we only simulated the activity of one neuron from $n_o = 8$ randomly sampled columns in each hypercolumn. This was achieved by first dividing the half-circle $[0, \pi]$ range of all possible orientations into n_o slices. With our choice of M and n_o , this would yield 4 evenly distributed preferred orientations in each slice. We then randomly picked a neuron to sample from each slice (i.e. with one of the 4 possible preferred orientations) with uniform probability. This process was carried out independent of the orientations in the stimulus and ensured that different orientation preferences are relatively evenly represented in our model. Sampling n_o neurons in each hypercolumn resulted in

a total network size of $N = n_o H = 99840$ neurons. After the sampling of neurons from different orientation columns, 20% of the neuron indices ($N_{inh}=19968$) were randomly assigned to the inhibitory population, with uniform probability.

In order to define the strength of the external input to each neuron in the presence of a stimulus, we assume that the classical receptive field of a neuron in orientation column j is Gabor shaped and aligned to its preferred orientation θ_j . We further assume, that the response of a neuron is phase invariant, and, in the absence of recurrent connections, can be given by a linear filtering of the stimulus $S(x,y)$ with a complex Gabor energy filter g_j .

$$S_{filt}(x,y,\theta_j) = |\hat{S}_{filt}(x,y,\theta_j)| \quad (7.7)$$

$$\hat{S}_{filt}(x,y,\theta_j) = \iint S(\hat{x},\hat{y}) g_j(x-\hat{x},y-\hat{y}) d\hat{x} d\hat{y} \quad (7.8)$$

with

$$g_j(x,y) = Z(x,y) (\exp(id_j) - \exp(i\Phi) \exp(-2(\pi\sigma_f/\lambda)^2)), \quad (7.9)$$

$$d_j(x,y) = \frac{2\pi}{\lambda} (x \cos(\theta_j) + y \sin(\theta_j)), \quad (7.10)$$

and

$$Z(x,y) = \exp\left(-\frac{x^2+y^2}{2\sigma_f^2}\right) \quad (7.11)$$

where i is the imaginary unit, $Z(x,y)$ is the Gaussian envelope of the filter with standard deviation $\sigma_f = 0.2/dx$ and spatial phase $\Phi = 0$. The spatial frequency $\lambda = 0.2/dx$ is the same as the one used for generating the edge elements in the stimuli (Eq. 7.1) and the absolute value operation in Eq. 7.7 denotes the complex modulus. As before, (x,y) denote coordinates on the stimulus with periodic boundary conditions, e.g. the Gaussian envelope given in Eq. 7.11 is centered at $(0,0)$, but wraps around the coordinate space. For the numerical implementation of the filtering operation, we took advantage of the convolution theorem and performed a multiplication in frequency domain using 2-dimensional FFT.

When defining the strength of the external input $J_{ext}(x,y,\theta_j)$, we used the offset parameter $J_{thr} = 0.3$ (Eq. 7.5). It is important to note that because the phases of the Gabor elements are allocated independently in different random realizations of the stimulus, even if the orientations of the elements remained constant, the filtering operation could return different results, especially for the pixels located further away from any of the element centers (k_x^i, k_y^i) . For example, the filtering operation will return a high value at pixel location $((k_x^1 + k_x^2)/2, k_y^1 = k_y^2)$ in between two Gabor elements 1 and 2, both with orientations $\theta_{1,2} = 0$, if the phases ϕ_1 and ϕ_2 are close to each other, and a low value at the same location if the phases are separated by a value close to π . This means that, even for the neurons processing the contour elements, the strength J_{ext} of the external input could vary non-negligibly between presentations of different stimuli.

7.3.3 Recurrent connectivity

The connection probabilities for the excitatory neurons were defined similarly to the association field describing the conditional link probabilities in [EMS+12] (see, for example, Figure 2 in [EMS+12] for a visualized description). For two neurons i, j with their receptive fields centered on (x_i, y_i) and (x_j, y_j) , we first define the distance between their RFs as the Euclidean distance r_{ij}

$$r_{ij} = \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2}, \quad (7.12)$$

the relative position of the RF of neuron i with respect to the position and preferred orientation of neuron j as α_{ij}

$$\alpha_{ij} = \text{atan2}(y_i - y_j, x_i - x_j) - \theta_j, \quad (7.13)$$

and the difference between the preferred orientations θ_i and θ_j of the two neurons as β_{ij}

$$\beta_{ij} = \theta_i - \theta_j \quad (7.14)$$

where $\text{atan2}(a, b)$ is the two-argument arctangent function of a/b .

The connection probability p_s (Fig. 7.12A) between excitatory presynaptic and all other neurons is then parametrized by r , α and β , and is defined as the product of a radial part p_s^d and an angular part p_s^θ

$$p_s(r, \alpha, \beta) = c_s p_s^d(r) p_s^\theta(\alpha, \beta) \quad (7.15)$$

with a normalization constant c_s . The radial part of the excitatory connection probability is given by the Gaussian

$$p_s^d(r) = \exp\left(\frac{-r^2}{2\sigma_d^2}\right) \quad (7.16)$$

with standard deviation $\sigma_d = 0.825$ degrees of visual angle. The angular part of the excitatory connection probability is given by

$$\begin{aligned} p_s^\theta(\alpha, \beta) &= 0.5 \left[M\left(\frac{\beta}{2}, \alpha, \frac{1}{\sigma_\alpha^2}\right) M\left(\frac{\beta}{2}, 0, \frac{4}{\sigma_\beta^2}\right) \right. \\ &\quad \left. + M\left(\frac{\beta}{2}, \alpha + \pi, \frac{1}{\sigma_\alpha^2}\right) M\left(\frac{\beta}{2}, \pi, \frac{4}{\sigma_\beta^2}\right) \right] \\ &= \cosh\left(\frac{1}{\sigma_\alpha^2} \cos\left(\frac{\beta}{2} - \alpha\right) + \frac{4}{\sigma_\beta^2} \cos\left(\frac{\beta}{2}\right)\right) / \left(4\pi^2 I_0\left(\frac{1}{\sigma_\alpha^2}\right) I_0\left(\frac{4}{\sigma_\beta^2}\right)\right) \end{aligned} \quad (7.17)$$

where $M(v, \mu, \kappa)$ is the von Mises distribution, which is a bell-shaped distribution supported on a periodic domain

$$M(v, \mu, \kappa) = \frac{e^{\kappa \cos(v - \mu)}}{2\pi I_0(\kappa)}, \quad (7.18)$$

and where I_0 is the modified Bessel function of order 0. For defining the length scale of the von Mises distribution, we used the parameter values $\sigma_\alpha = 0.25$ and $\sigma_\beta = 0.57$ degrees of visual angle. Finally, the normalization constant c_s (Eq. 7.15) is computed numerically

$$c_s = \left[\sum_{r,\alpha,\beta} p_s^d(r) p_s^\theta(\alpha, \beta) \right]^{-1}. \quad (7.19)$$

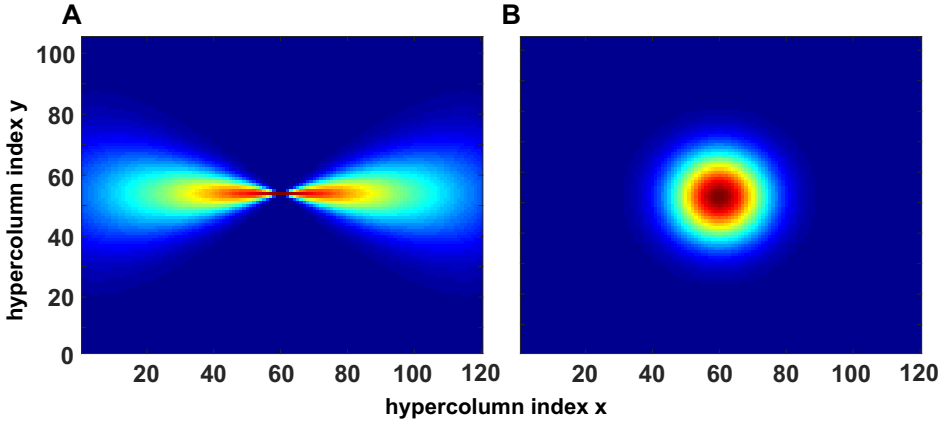


Figure 7.12: Visualization of the probability distributions used in sampling the synaptic connections. **A** The association field structure of the excitatory synaptic connection probabilities p_s . The image shows p_s between an excitatory neuron with its RF at the center of the stimulus and orientation preference $\theta_j = 0$, and all other neurons with orientation preference $\theta_i = 0$. Probability distribution p_s was used in the sampling of connections shown in Fig. 7.4. **B** The Gaussian dependence of the local inhibitory connection probabilities p_s^{inh} on RF distance. The image shows p_s^{inh} between an inhibitory neuron with its RF at the center of the stimulus and orientation preference $\theta_j = 0$, and all other neurons with orientation preference $\theta_i = 0$.

In contrast to the highly structured connection probabilities for the excitatory presynaptic neurons (Fig. 7.12A), the axonal connections made by the inhibitory neurons were local (Fig. 7.12B) and were parametrized simply as a function of the distance between the RFs r_{ij} (Eq. 7.12) and the difference of preferred orientations β_{ij} (Eq. 7.14). The probability p_s^{inh} that an inhibitory presynaptic neuron was synaptically coupled to another neuron

$$p_s^{inh} = \exp\left(\frac{-r^2}{2\sigma_{inh}^2}\right) \exp\left(\frac{-1.7\pi}{4} \beta\right) \frac{1.7\pi}{4\sigma_{inh}\sqrt{2\pi}} \quad (7.20)$$

decayed as a Gaussian with RF distance and as an exponential with preferred orientation difference. For the local inhibitory connections, we used a relatively small

length scale $\sigma_{inh} = 0.275$ degrees of visual angle (compare, for example, Fig. 7.12A and B). Finally, p_s and p_s^{inh} were used to sample $n_s = 500$ outgoing connections independently for each excitatory and inhibitory neuron respectively.

7.3.4 Neuron model

The membrane potential $V_i(t)$ of the leaky integrate-and-fire neurons $i = 1 \dots N$ evolved according to

$$\begin{aligned} \tau_m \frac{dV_i(t)}{dt} = & -(V_i(t) - V_R) + I_{ext} \sum_m \delta(t - m\Delta t) \\ & + w_{exc} \sum_{j \in \Omega_{exc}} w_{ij} \sum_k \delta(t - t'_{jk}) - w_{inh} \sum_{j \in \Omega_{inh}} w_{ij} \sum_k \delta(t - t'_{jk}) \end{aligned} \quad (7.21)$$

where $\tau_m = 10$ ms is the membrane time constant, $V_R = 0$ is the normalized (unitless) resting potential, I_{ext} is the stochastic external input (Eq. 7.6) delivered independently to each neuron at every discrete simulation time step $m \in \mathbb{Z}^+$, $w_{inh} = 4w_{exc}$ are the inhibitory and excitatory synaptic strengths and Ω_{exc} and Ω_{inh} are the set of indices of the excitatory and inhibitory neurons respectively. The sum over k denotes the spike train (sum of Dirac delta functions at spike times t'_{jk}) of neuron j . The coupling matrix w_{ij} assumes a value of 1 if a connection exists from neuron j to i as determined by the random sampling described above, and is 0 otherwise.

In general, the parameters of the model were chosen to be representative of those of an average cortical neuron [KSJ+00; NSD+05]. For simplicity, the membrane potential was normalized between 0 and 1, and neuron i generated an action potential if $V_i(t)$ crossed a threshold $V_\theta = 1$. After spiking, $V_i(t)$ was reset back to V_R . The differential equation in 7.21 was numerically integrated using the Euler method with a time bin $\Delta t = 0.0035$ ms. As in the previous two chapters, we have employed a separation of time scales in order to quantify the sizes of individual avalanches precisely (see, for example, section 6.2.1 for more information). We simulated the network's dynamics for a period of $T_{total} = 280$ ms, for each of the $N_{stim} = 100$ randomly generated stimuli and for each parameter combination. We discarded the transient activity in the initial 25 ms before analysis.

7.3.5 Avalanche statistics

The avalanche size distributions shown in section 7.2.3 were averaged over all N_{stim} stimuli. Similar to the approach used in section 6.2.2, we first visually identified the values of w_{exc} which were most different from exponential, but not supercritical. In the neighborhood of these values of w_{exc} , an exponent δ was fitted to the empirical avalanche size distributions $P(s)$

$$P(s) \sim s^{-\delta} \quad (7.22)$$

by searching in the range $[1.1, 4]$ for the exponent value which minimized the Kolmogorov–Smirnov (KS) statistic

$$KS_{\hat{\delta}} = \sup_s |F_{\hat{\delta}}(s) - F(s)| \quad (7.23)$$

where $F(s)$ is the cumulative distribution of $P(s)$ and $F_{\hat{\delta}}(s)$ is an ideal power-law distribution with exponent $\hat{\delta}$. This yielded a critical exponent of $\delta = 3$, which is also the value used in Fig. 7.6B.

In order to modify the salience of the contour, we varied the contour length and the orientation jitter. In the original stimulus, the contour was made up of six colinearly aligned Gabor elements (Fig. 7.2). For generating the modified stimuli with contours of varying length N_{cont} , we followed the same procedure as before, however, the stimulus contained $36 - N_{cont}$ randomly oriented background elements instead of the previous 30. The contour elements were always adjacent and horizontally oriented ($\theta_i = 0$). As mentioned in section 7.2.3, the orientation of the contour elements were jittered by diverting the individual elements from the horizontal contour path by a fixed amount $\theta_i \rightarrow \theta_i \pm \alpha_{jitter}$, either clockwise or counter-clockwise (assigned randomly). For each contour length, the hypercolumns which were included in the target ensemble were selected from a circular area centered on the N_{cont} contour elements as shown in Fig. 7.2.

7.3.6 Read-out and task performance

The rate-based read-out scheme is calculated by averaging the firing rates r_i^S of all the neurons i within each population (target and distractor) over all stimulus presentations S and then computing their difference

$$\frac{1}{N_{stim}} \sum_S \frac{1}{N_t} \sum_{i \in \text{target}} r_i^S - \frac{1}{N_{stim}} \sum_S \frac{1}{N_d} \sum_{i \in \text{distractor}} r_i^S \quad (7.24)$$

where $N_t = N_d$ are the number of neurons in each population.

The calculation of the synchrony-based task performance is identical to the read-out process described in section 6.2.1. The only difference is that, in this chapter, the output spikes of the simulated network, which are fed into the LIF read-out neurons, are not normalized. In other words, the population spike time series $s_p(k)$ are used directly for the read-out, as given in Eq. 6.4, without the normalization step given in Eq. 6.5. The normalization step was introduced in the previous chapter in order to exclude the contour information contained in the rates of the two populations and was avoided in this chapter in line with our aim of improving the biological realism. However, since rates contain little information about the presence of a contour in the model presented here, including normalization does not alter our results dramatically. The task performance with the normalized input displays peaks at slightly supercritical values, but other observations remain the same (not shown).

For the LIF read-out neurons, we used the same temporal parameters $\tau = 10$ ms and $\Delta t = 0.0035$ ms as with the neurons in the model network. For the gain factor f_g , we have scanned 40 different values in the range $[10, 10000]$ and picked the value which globally maximized task performance for the results shown in Fig. 7.9.

7.4 Discussion

In this chapter, we have presented a biologically realistic network performing a concrete behavioural task, namely, contour integration. We started by generating a stimulus containing a ‘target’ contour and randomly oriented background elements, similar to the ones used in psychophysical contour integration experiments. We used a biologically inspired network structure composed of orientation columns, organized into hypercolumns based on the location of their receptive fields. We have adopted the ‘association field’ idea, which is based on the Gestalt principles of perceptual grouping, in order to determine the recurrent connectivity. Finally, we have investigated network dynamics and contour detection task performance under different stimulus conditions.

Compared to chapter 6, we have thus greatly improved the biorealism of the network model in this chapter. Nevertheless, we find that our results in this more complex setting closely resemble the previously presented qualitative results. Even in the absence of a strongly symmetrical coupling matrix with explicitly embedded subnetworks, as in chapter 6, we still find evidence of functional subnetworks, collectively processing global shape information by integrating local features, such as the orientation of edge elements. This is illustrated by the dynamics of the neuronal population processing the ‘target’ contour becoming mutually synchronized and displaying power-law scaling while the population processing the background elements is relatively asynchronous, with exponential avalanche size distributions.

In addition, we observe that the correlation structure in the target ensemble depends strongly on contour salience. We show that this dependence can be exploited by using a synchrony-detector to quantify the contour detection performance. Interestingly, we find that such a read-out scheme displays maximized task performance at the critical value of the coupling strength w_{crit} . Finally, we show that synchrony-based read-out produces psychometric curves which resemble those observed in experimental studies, and, in a trial-based ROC analysis, greatly outperforms rate-based read-out.

Taken together, our results suggest that a synchrony coding scheme can effectively be used for binding local features together into coherent percepts of shapes and objects. The performance of this binding process decreases with object salience similar to what was observed in psychophysics experiments, and is maximized close to a phase transition in the dynamics. This implies that critical networks and spike-timing based encoding of global shape information can be efficiently used in realistic settings, where there is a large overlap between neural representations of different objects, in order to implement flexible top-down modulation of neuronal activity by stimulus properties.

7.4.1 Robustness and interpretation of results

Due to the enhanced biorealism, and the consequent increased complexity of the model presented in this chapter, large scale parameter scans were not performed. In terms of robustness against parameters, it was confirmed that the qualitative results

did not change with changing read-out neuron gain parameter f_g in the synchrony-based read-out scheme, and the peak task performance was still observed close to the critical point w_{crit} . In addition, we expect that rescaling the stimulus, while keeping the relative scaling of the stimulus elements, receptive fields and network size to each other constant, should not alter our results. Similarly, we would anticipate little to no change in our qualitative observations when changing the stimulus size by expanding the stimulus (i.e. adding more Gabor elements and using longer contours), instead of rescaling it. In this scenario, if the network size is scaled up proportionately to the stimulus, and the retinotopic mapping is conserved, we would expect quantitative changes to the task performance, but no dramatic changes with respect to our findings related to the critical point and salience, due to the periodic boundary conditions of the network.

We find that our network displays rather similar behaviour in response to changing two basic psychophysical parameters, namely, contour length and orientation jitter. In both cases, the decreased stimulus salience leads to a decrease in correlated activity within the target population, which results in decreased contour detection performance. Due to our association field inspired connectivity pattern, we believe, a similar behaviour would emerge in the network when changing other classic salience parameters such as interelement distance or contour curvature [FHH93; HHF03]. However, some more complex interactions, such as spatial phase selectivity of the neurons or dependence of synaptic connectivity on spatial frequency were not built into the model, and it is not clear whether or how experimental observations with respect to such parameters could be reproduced in our model. Similarly, due to time and computational expense constraints, we did not perform investigations involving systematically varying the contrast of the Gabor elements or stimulus presentation time. However, such inspections are possible within the scope of our model and we hope that it will constitute a test bed for future investigations.

In addition, when we inspect the contour detection performance quantified by the ROC analysis, we find that the performance is unrealistically high for short contour lengths [LG02; LPG06; LPG08] as well as for large orientation jitter [FHH93; MW11; MK05]. These effects can at least partially be explained by the discretization of the orientation preference. As the background elements are randomly assigned orientations in a continuous spectrum (within double computer precision), it is likely that, on average, the feedforward input the distractor population receives will be slightly weaker than the target population, where the orientations of the contour elements $\theta_{i \in \text{contour}} = 0$ exactly matches the preferred orientation of at least one column. This is also true, for example, for presentations of contours with orientation jitter $\alpha_{jitter} = 45$ degrees, and may account for the inflated performance at extreme parameter values.

Finally, we find an increase in the slope of the psychometric curves shown in Fig. 7.10A and C, at large contour lengths (5 and 6 elements). As mentioned before, this can be attributed to contour closure, as experimental studies report improved contour detection performance for closed circles as compared to open contours with identical length and curvature [KJ93; MF07]. Due to the length scale of the

association field we used for defining the excitatory connectivity, circle closure is a factor contributing to the increase in correlated activity in the target population from very small contour lengths. It may, therefore, also play a role in the elevated ROC performance we report. Nevertheless, such effects are consistent with findings in similar models, which use connectivity structures inspired by the association field, and which have replicated experimental studies of contour closure [YF98].

7.4.2 Physiological Plausibility

For the single neuron dynamics, we have made sure that temporal values such as the membrane time constant and firing rates of the model neurons were within the physiological range [KSJ+00]. Each randomly generated stimulus was presented to the network for a total of $T_{total} = 280$ ms, and, in order to exclude transient activity, only the last 255 ms of each trial was considered for analysis. The relatively short stimulus presentation time was used based on the speed with which contour integration is known to take place in experiments [EMS+12; FHH93; HHF03; MK05].

With respect to the network connectivity, evidence for anisotropy of anatomical connections, which could support an association field, is limited (see, for example, [KTR+97]) and mainly isotropic connections have been reported for the macaque area V1 [ALW+02]. However, it is not known how much of the anatomical structure of connections translates directly into functional connectivity. Therefore, we believe that, taken together, the experimental support for the role of Gestalt principles in contour integration [FHH93; HHF03], as well as the edge co-occurrence statistics in natural images [GPS+01] make a good case for an AF-like functional connectivity in the cortex. Furthermore, we implement local lateral inhibition [KSJ+00], and the stochastic sampling of synaptic connections in our model results in a directed graph with sparse recurrent connectivity, consistent with experimental observations [SCD+18].

Based on both the retinotopic organization of the orientation hypercolumns, as well as the complex cell energy nature of the feedforward input, our network can be interpreted as a model of an early visuocortical area such as V1 [DA01; KSJ+00]. Evidence for the involvement of upstream areas like V1 in contour integration have been presented in various studies, ranging from psychophysical experiments in humans [KIG+95; LG02], psychophysics and electrophysiology in behaving monkeys [KIG+95; LPG06] to injection labelling in monkeys [SDB+02]. Although it has been suggested that feedback projections from downstream areas such as V2 to V1 may play an important role in contour integration as well [ALW+02], findings such as lack of orientation specificity in V2 to V1 feedback [SDB+02] as well as the speed with which successful contour integration can take place [EMS+12; MK05] make it hard to completely rule out V1 involvement.

The interpretation of our recurrently connected layer as an early cortical area such as V1 presents another reason for choosing a distance-dependent connectivity pattern like the association field, instead of comparatively unrealistic global couplings. Under this light, the read-out neurons in the synchrony-based read-out scheme can be

interpreted as neurons in downstream areas of the visual pathway, such as V4. This is also reflected in the pooling of input from a large number of neurons in the recurrent layer (i.e. pooling of input from all contour neurons, with potentially distant RF centers) to the read-out neurons, which results in a much larger classical RF size for the read-out neurons, which is still consistent with biological RF sizes [Mot09]¹. As explained in chapter 4, such feedforward bottlenecks not only provide an anatomical mechanism for contour integration based on local feature detectors, but also suggest a functional role for critical dynamics, where asynchronous activity in subcritical regimes do not elicit strong responses in higher areas and runaway activity in supercritical regimes lead to a loss of specificity in the responses of higher areas (due to increasing false positives).

For the display of our psychometric curves as given in Fig. 7.10, we chose to use the rate differences (either directly of model populations, or difference of the rates of the read-out neurons), as well as the ROC performance. This is both in order to be consistent with the previous chapter, but also because it is not immediately clear whether or how the brain might make use of a linear classifier based on mean population rates. However, difference of rates is utilized often in predictive coding schemes, where an error signal is generated by the subtraction of the rates of a prediction generating population from the rates of feedforward input. In hierarchical models of the visual cortex which perform predictive coding (see, e.g. [HR11] and [BUA+12] for an overview), this subtraction takes place in populations which receive feedforward input from downstream areas and predictive input from upstream areas. Our model contains only a single layer, bar the two read-out neurons, and in its current form cannot accommodate a predictive circuitry. However, the resemblance of our synchrony-based read-out psychometric curves (Fig. 7.10A-B) to experimentally obtained ones [FHH93; LG02; LPG06; LPG08; MW11; MK05] suggests that a possible extension of the model to multiple layers, including read-out and feedback populations, may be useful in quantitatively replicating experimental results.

¹While our longest contour is 3.3° long, V4 neurons are known to have cRF sizes ranging from 2° to 12° depending on eccentricity [Mot09].

IV

Conclusion and Discussion

8	Summary and Conclusion	125
9	Discussion and Outlook	129
9.1	Generality and Realism of the Models	
9.2	Critical or Near-Critical Dynamics?	
9.3	Criticality, Structure, and Function	
9.4	Outlook	
	Bibliography	138



8. Summary and Conclusion

As already stated in the abstract and introduction, this thesis was motivated by recent experimental and theoretical work which has established the hypothesis that cortical neurons operate close to a critical state which signifies a phase transition from chaotic to ordered dynamics.

We have discussed in Part I that when a system with multiple nonlinear elements interacting with each other assumes dynamical states close to such a phase transition, it displays certain interesting properties. Including a diverging correlation length, where global patterns can emerge in system dynamics through strictly local interactions, such properties decay and vanish rapidly as one moves away from the critical point of the phase transition (Ch. 2). We have elaborated, in particular, on existing evidence which suggests that near-critical dynamics optimize several aspects of information processing. This body of ideas and theoretical investigations have collectively formed the backbone of the criticality hypothesis in the central nervous system, which we presented in chapter 3, due to the efficiency with which the vertebrate brain can perform computations relevant for behaviour.

In addition, we have presented examples from the vast experimental literature which have reported both direct and indirect signatures of critical dynamics in brain activity in a large range of spatial and temporal scales (Ch. 3). We have identified that, nevertheless, much of the experimental work is based on *in vitro* investigations, or otherwise, examine the dynamics of spontaneous neuronal activity. Therefore, little is known about how such signatures are modified by sensory input or how they are affected by task-dependent changes in neuronal activity when the cortex is engaged in stimulus processing.

Similarly, we made a case for the need for biologically inspired (Ch. 3) models tackling more concrete behavioural tasks in theoretical literature. In constructing

our models, we have aimed to bridge the gap between the original ideas of maximized information theoretical measures, as demonstrated in abstract models, and the observation of near-critical dynamics in the brain. To that end, we have focused on information processing in the visual system, and examined both the emergence and disappearance of critical dynamics in visuocortical circuitry under changing task demands, as well as their contribution to function in biologically plausible settings.

In this dissertation, we have investigated a specific role for critical dynamics in the cortex, namely, the rapid dynamic modulation of visual information processing. In order to demonstrate the versatility of our approach, we have focused on two distinct examples: one where top-down selective attentional mechanisms regulate oscillatory activity in local circuits (Part II), and one where changing stimulus salience modulates local subnetwork dynamics in a bottom-up manner (Part III). In both cases we have shown that system dynamics transition rapidly as networks get closer to a critical point, and thus being poised at the critical state provides us with an optimal regime in which to flexibly modulate network activity, with either changing task demands or small modifications of stimulus properties.

Specifically, in part II we investigated the emergence of criticality in local visuocortical networks in the context of task-dependent modulations of stimulus processing. We focused specifically on changes induced by selective visual attention, as an example of an experimentally rigorously studied mechanism. We reproduced, in a network of leaky integrate-and-fire (LIF) neurons, the empirical observation that LFPs in visual areas display stronger γ -band oscillations under directed attention, which has been related to the enhancement of object representation. We have shown that, similar to the experimental findings, γ -oscillations and enhanced stimulus discriminability co-emerge in the same region of our model phase space, and that the border of this transition region can be identified by functionally connected local populations of neurons displaying scale-free dynamics. We suggested, therefore, that a near-critical state might be beneficial for rapid and flexible top-down modulations of stimulus processing, where small, local modulations of system parameters, such as neuronal gain, might lead to large qualitative changes in perceptual outcomes, such as object representation and discriminability.

In part III, we investigated the functional role of critical dynamics in the context of local modulations of network activity by stimulus properties. In chapter 6, we constructed a critical network model of analytically tractable Eurich–Herrmann–Ernst (EHE) units, in which we embedded functional subnetworks co-processing an object in the visual field. We have shown that the parameter regime in which such subnetworks display scale-free dynamics, while structured local inhibition helps avoid runaway activity, leads to maximum performance in a figure-ground segregation task, quantified by a coincidence detector. In addition, we have shown that the result can be qualitatively reproduced to a large extent in a model with EHE units including small leak currents (section 6.3.2). In chapter 7, we expanded our feature integration model from chapter 6 to the more specific, and experimentally carefully studied task of contour integration. Accompanied by a more realistic network model with increased biocomplexity, we have shown that the contour detection performance,

using synchrony-based read-out, was once again maximized near a phase transition in this new setting. Thus we suggested, that the local emergence of critical dynamics, within a slightly subcritical larger network, might be a favorable framework for dynamic bottom-up modulations of stimulus processing. The localization of highly correlated activity in space and time allows for the level of synchrony within a given, functionally coupled population to reliably and flexibly signal the presence of global shapes in the visual field.

Hence, we believe with the three distinct models presented here, we have addressed our initial goals of formally extending existing models of critical dynamics to inhomogeneous systems subject to a strong external drive, as well as demonstrating concrete functional benefits for the critical state in network models of well-known experimental paradigms.

Altogether, we have shown that both attentional shifts and an increase in stimulus salience may push local subnetworks, processing a behaviourally relevant figure or object in the visual field, towards the critical state. In the top-down scenario such a push is accompanied by an enhancement of object representation in the oscillatory dynamics of these subnetworks (e.g. figures 5.4 and 5.3), and in the bottom-up example by an improvement in figure detection task performance (e.g. Fig. 6.7). Taken together, the two distinct scenarios, both benefiting from critical dynamics, hint at a generic correlation between fast and flexible information processing and an operating regime close to a phase transition, which is not restricted to a single mechanism. Importantly, these results provide us with concrete examples of how local emergence of criticality may play a functional role in thoroughly studied visuocortical phenomena.

In light of these findings we argue that the active cortex is not globally operating in a critical regime at all times but that the differential in correlations, both over time and space, is an effective tool for the cortex to selectively enhance the processing of relevant sensory information. In general, our findings imply that selective, preferential or otherwise dynamic processing of visual information in a continuously changing visual environment may be supported by local subnetworks operating close to a critical state, and our results show that in models under realistic constraints, the critical state may yield optimal performance for downstream read-out mechanisms.



9. Discussion and Outlook

9.1 Generality and Realism of the Models

The models we presented in chapters 5 and 6, are relatively minimalistic models, capable of demonstrating the intended effects with a small number of free parameters. In addition, in both chapters we have reported the results of extensive parameter scans and confirmed, where applicable, on which parameter values our results critically depended on and what systematic changes they are robust against. In general, we made sure that our models can be interpreted as basic mechanisms which can plausibly underlie experimental observations, without the need to pick an unjustifiably specific combination of parameter values.

In addition, any free parameter values which could be reasonably approximated by real-world values (such as time constants) were chosen appropriately in order to be consistent with the demonstrated physiological dynamics (such as γ -oscillations). In general, all of our parameters values are either within the biological range, or can assume realistic values under appropriate normalization without loss of generality (such as membrane potentials). Where exhaustive parameter scans were not possible, as in chapter 7, we have used a combination of known physiological parameter values and, as estimates for more complex anatomical variables such as the distribution of synaptic connections, values which yielded biologically plausible dynamics (e.g. in terms of firing rates). For an in depth discussion of particular parameter choices, please see the discussion of the corresponding chapter.

Nevertheless, using simplified models in order to identify generally applicable mechanisms is widely used in computational neuroscience in general and in criticality research in particular. As we have discussed in the introduction of this thesis, the majority of theoretical studies concerned with demonstrating computational benefits of criticality rely on rather abstract models, which are hard to interpret from a

biological perspective, or reveal maximized values for abstract measures, which may be hard for physiological neurons to directly and effectively make use of [BN04; KC06; Lan90; SP13]. Similarly, whether such results would translate to network and neuron models with enhanced biorealism is not always clear. In that sense, here we have improved upon previous work by employing non-homogeneous models, biologically inspired anatomical connections and read-out mechanisms and concrete behavioural tasks performed by a strongly driven cortical network actively engaged in processing sensory information.

As stated in section 5.4, the variability of neuronal connectivity in a local population in the brain is not random or homogeneous, but signifies a highly structured global network. As such, the biological cortex is not a random graph as was used in our attention model (Part II), nor is it strongly deterministic and symmetrical as in our idealized feature integration model (chapter 6). However, both models capture the general idea that local feature detectors collectively process an object in the visual field via the use of long-range connections [KTR+97], and that local inhibition plays a role in circumventing runaway or epileptic activity [KSJ+00; PHM+12; SYP+09]. Nevertheless, we have also shown that our results can be reproduced in an anatomically more detailed and stochastic model of synaptic connections in chapter 7.

To address our findings in relation to the topology in our attention model (chapter 5), the fact that different subnetworks have different critical points in our phase space originates from modest topological differences between the randomly generated coupling matrices. This means that the demonstrated transition region shrinks and ultimately approaches a transition ‘line’ at the limit $N \rightarrow \infty$, while discriminability decreases to chance level. One may alternatively choose to extend the model to encompass a more structured topology established, for example, by a biologically inspired learning rule. Hence, enhancing topological differences between subnetworks in the attention model would increase the size of the transition region and make our results more robust against increasing system size. A similar effect would be observed if one were to use a deterministically structured topology with stochastic sampling of synaptic connections, as we used in chapter 7. In other words, our columnar model can also be used as a test bed for attentional modulations in a contour integration task, bringing the top-down and bottom-up aspects of our investigations together. Unfortunately, such investigations could not be performed within the time limit of this thesis work, and remains a topic for future research. Finally, the effects observed in our attention model are compatible with other tentative mechanisms relating to how attention may intervene with feedforward processing such as rate increases [Fri05a; TMF+05].

In all our models, we have considered functionally connected subnetworks, representative of a collective internal representation of shapes or objects in the visual field. There is evidence for such functional groups in the visual system, which potentially help with feature integration processes, including contour integration, by pooling responses of orientation-selective neurons in upstream areas (such as V1 or V2), and giving rise, through feedforward connections, to a wide range of larger and

more complex receptive field structures in downstream areas (such as V4) [DHG06; DS87]. In addition, for our purposes in this thesis, it was sufficient to exclusively consider lateral and feedforward connections. In general, however, a more complete understanding of both feature integration and attentional modulations of sensory processing can only be reached via the simultaneous consideration of feedback connections to upstream areas (see e.g. [The10] or [ALW+02]). Even though explicit modelling of the interactions between feedforward, lateral and feedback connections in a multilayered model would lead to the emergence of richer and more complex control mechanisms, we believe that the input from feedback connections can be successfully integrated into the processes which tune local networks towards or away from criticality. Therefore, even though the specifics of such a model are speculative and beyond the scope of this thesis, we would predict that feedback connections would not necessarily change our conclusions about improved object representation and figure-ground segregation under critical dynamics (see also the discussion in section 7.4.2).

Finally, we used a variety of different read-out measures. In chapter 5, we have demonstrated the enhancement of stimulus information contained in LFP signals, using classification of oscillatory modes and discrimination of different figures. In chapter 6 we used a coincidence detector and in chapter 7, we have compared the performance of rate-based and synchrony-based schemes. What we find in common in all of these different scenarios is a link between criticality and improved task performance when using synchrony-based read-out mechanisms. This demonstrates the generality of our findings independent of the exact mechanism or direction of modulations (i.e. task or context-dependent). It is also indicative of a functional role for criticality in a cortex which communicates through mutual synchronization or oscillations, or otherwise, of a functional role for oscillatory dynamics in a cortex which is poised at criticality.

9.2 Critical or Near-Critical Dynamics?

In part II we have shown that in a network of LIF neurons, the transition region between asynchronous and highly correlated dynamics, identified on the basis of neuronal avalanches, correlates well with a peak in stimulus discriminability. However, such enhancement of object representation comes at the cost of dramatically reduced information entropy in spike patterns as a larger number of subnetworks begin to display supercritical dynamics with increasing relative excitation. This dynamical effect implies the existence of an optimum between linking features together and losing detailed stimulus information.

We conclude, in chapter 5, that even if we were to assume no bottlenecks existed in the flow of information within the visual system, and a ‘fine scale of observation’ was feasible (through, for example, dendritic integration [Mag00]), marginal subcriticality might still represent a best-of-both-worlds approach. In particular, a penalty in information entropy may be necessary to establish a certain level of synchronous activity, which may be required for other functionally relevant

aspects of cortical dynamics, such as information routing regulated by attention via ‘communication through coherence’ [Fri05b; GNM+12b]. Therefore, coding schemes being optimal for information transmission and processing always depends strongly on neural constraints and read-out schemes. Similarly, how far from criticality this optimal working point for the brain is depends strongly on the nature of the function being computed, and the corresponding read-out mechanisms. In particular in a passive cognitive state, the cortex might be further from criticality, but ready for top-down mechanisms of selective attention to kick in and rapidly enhance object representation as necessary.

This suggests a more interesting functional role for near-critical dynamics, in terms of flexible modulations of local activity, than the static role which can be attributed to the precise critical point. Such an observation is in line with the experimental findings that cortical networks might be operating at a slightly subcritical state [LP17; PWV+14]. Our suggestion that the distance of neuronal dynamics from the critical point might be actively modulated is also consistent with changes in brain states under sustained wakefulness [MOS+13] and cognitive load [TV14].

Nevertheless, in our feature integration models (Part III), we find maximized task performance at precisely the critical point. However, in these models our objective function is always figure-ground segregation, where we comparatively evaluate the activity of subnetworks dedicated to processing a figure and subnetworks dedicated to behaviourally irrelevant background elements. The detection of the presence of a figure or a contour, however, is a fundamentally different task from, for example, identifying, recognizing or otherwise discriminating between two stimuli. Arguably, in such tasks, retaining the distributed information about specific local features or details of a shape might be more advantageous than integrating over the entirety of local information and strongly signalling the presence of a global shape. In such cases, one might assume *a priori* that a subcritical state might provide the optimal encoding.

In our simple feature integration model (chapter 6), the coupling matrix is symmetric and all the subnetworks are identical, which makes it impossible to perform an object discriminability experiment with it. In fact, if a stimulus discrimination task is attempted on the basis of oscillatory nodes, as in chapter 5, one would find that the only difference between the dynamics of any two subnetworks stem from the different initial conditions and the random selection of the external input location (given that the coupling and external input strengths are kept constant). This is because the initial conditions and the external input are the only sources of randomness or noise in the original EHE model [EHE02].

A good way to bring together the ideas of top-down and bottom-up modulation, as well as to see whether different performance maxima may be identified for different behavioural tasks, would be to extend our realistic network model (chapter 7). As also mentioned in the previous section, the stochastic sampling of the synaptic connections lead to a much more natural formation of functional subnetworks, which can nevertheless be very different from each other in terms of graph theoretical measures. In this case, we would expect to be able to qualitatively reproduce our

results from chapter 5 in our realistic network model, however, how far the stimulus discriminability optimum would be located from the critical point is still an open and interesting question.

Another important question is, of course, whether our models are really ‘critical’ at the points which we identify as critical points. In our model networks, we have employed two different neuron models: EHE units and LIF units. The EHE model is based on a critical model of earthquake dynamics [OFC92] and its critical point in a system of finite size N can be analytically derived as a function of N [EHE02]. Due to the analytical understanding of the original EHE model, we do know that it is a reliable example of dynamical criticality [SHH19], which displays a continuous phase transition characterized by power-law event distributions. The connection between LIF neurons, power-laws and how to identify the critical point, however, is a little more unclear.

Several studies have suggested that networks of LIF units undergo phase transitions, which have been linked to SOC [CPDG+95; USO95; ZG13]. However, all of the cited studies consider LIF neurons interacting on a lattice structure, and the exact relationship of a phase transition and the emergence of scale-free distributions is unclear (for example, in [ZG13], they conclude that power-law distributions appear in the supercritical regime, while other properties of the dynamics determine the critical point). This makes it hard to judge whether different kinds of LIF networks may be able to display continuous phase transitions.

A sparsely connected network of mixed inhibitory and excitatory LIF neurons (such as the networks we used in chapters 5 and 7) is known to undergo a bifurcation as a function of its coupling strengths [Bru00]. In order to slightly bridge the gap between the LIF and the non-leaky EHE model, we have investigated, the performance of our feature integration model both with and without leak currents (chapter 6). Such addition of leaks constituted a middle step between the two models, as we retained the post-spike reset schemes of the EHE model. In the ‘leaky’ EHE model, we observed that the emergence of oscillations, as well as the transition to supercritical avalanche distributions, were already visibly more abrupt. This abruptness of the transition was also true, if not enhanced, for the dynamics around the critical points we observed in our LIF models in chapters 5 and 7.

Our observations about the leaky systems are in line with the suggestion that LIF neurons exhibit first order phase transitions [SAM+18]. Nevertheless, we do not believe the exact nature of the transition is critically important for our work. It is known, and can also be seen throughout this dissertation, that both the EHE model with small leaks [EHE02] and the LIF models [CPDG+95; TRE14; ZG13] display approximate power-law scaling around the critical point. In addition, when working with the LIF models, we did not identify the critical point based on a precise power-law fit at the critical point, but rather relied on the sudden nature of the transition to identify the abrupt appearance of supercritical distributions, with large bumps at large avalanche sizes. Essentially, our investigations about flexible information processing in the cortex attribute a special function to the emergence of strong mutual synchrony or oscillations at a critical point, which can approximately

be characterized by scale-free dynamics. Although relying on elegant ideas, such as SOC to explain our observations is tempting, whether such correlated activity appears as a result of a first or second order phase transition is ultimately inconsequential.

Finally, the brain itself is not a static but a dynamic structure. Changes take place in the brain constantly, through, for example, learning and memory [DA01; Hop82; VSZ+11], perceptual learning [GGE17; LPG08] and sleep [DHH+12; MOS+13; PVW+13]. Although homeostatic mechanisms on many different spatial and temporal scales help regulate that neuronal activity remains in a biologically and behaviourally nondetrimental regime [KSJ+00; VSZ+11], fluctuations of synaptic input, on short time scales, or of synaptic weights, on medium to long time scales, may shift network dynamics away from the critical point. Therefore, it would be nonsensical to talk about a network poised at the critical point, if arbitrarily small shifts towards the supercritical state is enough to push it into a regime of positive feedback, resulting in infinite avalanches or epileptic activity. This idea, that cortical networks may operate at the subcritical region, in order to maintain a buffer zone from epileptic dynamics, has been suggested before [PWV+14]. In our investigations, we show that both addition of local, structured inhibition (to a small extent) as well as leak currents (to a much larger extent) contributes to the formation of such a buffer zone. This implies that in a brain known to both sustain a highly detailed balance between inhibition and excitation [OL08], and where leak currents play a big part in single neuron dynamics [DA01; KSJ+00], small excursions to the supercritical state might be well tolerated and it is not necessarily self-destructive that cortical networks employ a ground state close to the critical point.

9.3 Criticality, Structure, and Function

Within the current neuroscientific literature, there is growing interest in bridging the gap between the abstract concepts surrounding critical phenomena in the thermodynamical sense and biological concepts such as structure and function which are of interest to neuroscientists. For example, new studies are aiming to establish links between more subtle effects of different connectivity structures on criticality [MPM15] as well as investigating the capacity of networks to support replay of learned spatio-temporal patterns near a phase transition [SC13].

In this work, we intend to contribute to this growing body of literature, by providing examples which directly relate to experimental settings at the behavioural level. We suggest that critical dynamics may be favorable for supporting rapid changes in network activity, necessary for flexibly modulating information processing based on changing stimulus properties and task requirements. In order to achieve this flexibility, one factor plays a crucial role in our models: that the active cortex, strongly stimulated by sensory input, is not globally and constantly at a critical point. Similar observations were made previously, demonstrating progressive disturbance of critical dynamics with sustained wakefulness [MOS+13] and cognitive load [TV14]. Similarly, it has been suggested that avalanche dynamics *in vivo* display differences between wakefulness and deep sleep [PVW+13] and may be poised at a slightly

subcritical state [PWV+14]. In an attempt to explain many of these observations, we suggest the tuning of local subnetworks towards and away from criticality as a generic mechanism to modulate sensory information processing.

Some recent literature suggests that a shift in neuronal dynamics away from the critical state may be linked to focused attention as well as active preparation for stimulus processing, seemingly at odds with our predictions. For example, it has been shown that a scale-free distribution of neuronal activity cascades in human EEG is associated with the resting state while focused attention shifts the dynamics towards a subcritical regime [FLS+15]. The conclusions we present in this chapter, however, are based on dynamics of local subnetworks, spanning much smaller spatial and temporal scales than a whole-brain EEG. Furthermore, cascade analysis of continuous data, such as EEG signals, typically depends on a thresholding of the signal in time domain, and will mainly capture the high amplitude oscillatory activity in the lower frequency bands, such as α -waves, which are known to decrease with more vigilant mental states. Similarly, power-law distributions in spiking data have been linked in awake monkeys with eye closure and asynchronous dynamic states have been observed when the subjects had their eyes open in a dark setting [HPAM+17]. However, due to the lack of stimulus evoked activity in the experimental paradigm, the presented dataset lacked discernible oscillations in the γ -band, which is the dominant synchronization frequency in our models. Therefore, we do not believe these results are in contradiction with our predictions. In fact, inhibitory suppression of cortical subnetworks which are irrelevant for the processing of a visual stimulus is a fundamental part of our feature integration models (Part III), which may be related to the asynchronous dynamics observed in these studies.

In order to scrutinize the functional role of synchrony in top-down and bottom-up modulations of cortical activity, we considered an idealistic scenario in all our models, where each given shape or figure in the visual field activates a given set of neurons (which we call subnetworks). Such subnetworks are always either of identical (chapters 5 and 6) or very similar size (chapter 7) and consist of identical units. This has the following implication for our attention model: without synchronization, stimulus information encoded in activated neuron identities would be lost in the *average* population rate. By means of the different connectivities within different sets, however, this information becomes re-encoded in response amplitude and γ -synchrony. For purposes of figure-ground segregation in our feature models, the similar size and structure of subnetworks can be interpreted in a similar manner: the presence of a global shape or object is signalled not by the simple external coactivation, or rates, of a group of feature detectors (which may not collectively represent a shape, but just the presence of random background elements), but by the correlated spike times which is a result of the functional connectivity forming the behaviourally relevant subnetworks.

In principle, this idea is very similar to the well-known concept of binding by synchrony [vdM94], namely, using the temporal domain to represent information about relevant properties of a stimulus, for example, by tagging its features as belonging to the same object or to different objects in a scene. Taken together with

our results, this may imply that the cortex transmits local stimulus information using rates, and retains information about global object configurations in the time domain, thus implementing a multiplexed coding scheme.

Such a scheme, however, employs mutual synchrony or correlated activity to encode information, which leads to greatly decreased information entropy, as we have explicitly shown in chapter 5. In addition, a high level of correlation between spike times is contrary to the popular ideas of sparse coding and efficient coding (see e.g. [HH02; SB13]), where information entropy and transmission may be maximized with minimal effort and energy expenditure.

As such, synchronization is only beneficial for information processing if additional constraints exist: for example, a neural bottleneck in which some aspect of the full information available would be lost or actively gated out, as mentioned in chapter 4. In addition, correlated activity emerges naturally if a certain degree of robustness of signal transmission against noise is actualized through increased redundancy, or a gating mechanism such as ‘communication through coherence’ [Fri05b; GNM+12b] is realized by the synchronous arrival of action potentials at the dendritic tree.

In our investigations, these bottlenecks are represented by a finite number of neurons in a single layer, each of which may be tasked with contributing to the neural representation of many different objects (i.e. large overlaps between many different functional subnetworks), as well as the integration of activity from a large number of neurons within a subnetwork by a small number of read-out neurons in downstream layers, resembling the pooling of receptive fields at downstream visual areas.

9.4 Outlook

The most straightforward expansion of this work would be to bring together the top-down and bottom-up aspects of neural modulation presented in this dissertation in a single model. How this may be established (using the realistic model presented in chapter 7 as a test bed) and what hypotheses might be investigated are discussed in the previous sections of this chapter.

Another interesting expansion of our models would involve investigating the validity of our approach in multilayer models. Consistent amount of activity observed in multiple layers is an oft-cited and important theoretical property of critical feedforward networks, as their branching parameter is equal to one [BP03]. This is a characteristic often thought of as beneficial for dynamics where activity should neither die out nor explode over time. The idea is reinforced by recent investigations which show that arbitrarily deep feedforward neural networks may only be trained close to criticality [SGG+16], however, whether or how the property would carry over to networks with recurrently interacting layers is still an open question.

In addition, we have not explicitly addressed the question of how our networks may grow to be critical. Since our main research goal was to investigate concrete functional benefits in critical networks, it was implicitly implied that either evolutionary or physiological optimization mechanisms may tune cortical subnetworks to

this demonstrably flexible regime. However, several studies suggested that critical dynamics may be a consequence of homeostatic structural plasticity (see e.g. [KGA+18; OBO19]) or activity-dependent depressive synapses [LHG07]. However, despite showcasing analytically very interesting systems, such studies lack explicit links to functional benefits. A more interesting approach would be to demonstrate a learning rule which tunes a network to the critical point while minimizing an appropriately chosen cost function. Therefore, potential future work on our models would include investigations of whether critical dynamics would emerge as an optimal solution to a behaviourally relevant function. Whether optimizing parameters with respect to synaptic weights would be enough to tune the subnetworks to the critical point, or otherwise, whether changes to structural connectivity, parameters of neuron models (such as time constants) or synapse models (going from static to dynamic synapses) would be necessary are interesting and open questions.

Finally, as stated in the previous section, the main experimental hypothesis which emerge from our models is, that signatures of criticality in stimulus-evoked activity and under attention may be readily and consistently found *in vivo* only when analyzing the activity of functionally connected local subnetworks, during the presentation of a suitable stimulus.

It remains for future studies to explore these predictions in more detail in an experimental setup. Currently, both subsampling of relevant populations as well as the separation of avalanche and external drive time scales present challenges for experimental tests of this hypothesis. However, we believe that both the efforts to correct for subsampling effects in data analysis [LP17] and recent advances in multielectrode recording techniques and optogenetic methods will be instrumental in the progress towards a more complete understanding of critical dynamics in the cortex.

Final notes

No copyright infringement is intended by the author throughout the entirety of this thesis. Proper permissions are obtained beforehand where applicable.

The L^AT_EX template used for this dissertation is adapted from the Legrand Orange Book template as found in <https://www.latextemplates.com/template/the-legrand-orange-book>.

On page 6, the quote attributed to J. W. Gibbs can be found in the preface to his book "Elementary Principles in Statistical Mechanics" (Cambridge University Press, 2010; originally published in 1902). The quote attributed to D. O. Hebb is reported by Daniel C. Dennett in his book "Consciousness explained" (Little, Brown and Co., 1991).

Parts of this thesis have been adapted from author's previously published own work. Chapter 5 is largely based on the published paper [TRE14] (covered under the Creative Commons license CC-BY), reprinted here with permission from the co-authors Udo Ernst and David Rotermund. Chapter 6 is largely based on the book chapter by the author and Udo Ernst, titled "The role of criticality in flexible visual information processing", as featured in the book co-edited by the author, Michael Herrmann and Udo Ernst titled "The Functional Role of Critical Dynamics in Neural Systems" (Springer-Verlag, 2019). Parts of the introduction and discussion sections have also been adapted from these two publications.

The author would like to thank Udo Ernst for permission to use a photo taken by him for figure 4.1. The cover art as well as chapter images for chapters 6 to 8 are the author's own work. Artwork for the headers of chapters 3, 5 and 9 are plots obtained in computer simulations by the author.

Chapter header image for the contents and bibliography is original artwork by Sandra Strait (<https://lifeimitatesdoodles.blog>), taken from www.flickr.com/photos/aceo/47549012861/in/photostream with her permission. For the remaining chapter headers I've used:

chapter 1-microscopy image of three neurons, credit: lab of Dr.Chris McBain, NICHD/NIH (courtesy of Dr. McBain, science.nichd.nih.gov/confluence/display/mcbain/Home),

chapter 2-taken from [commons.wikimedia.org/wiki/File:Snowflake_macro_\(Unsplash\).jpg](https://commons.wikimedia.org/wiki/File:Snowflake_macro_(Unsplash).jpg) (under CC0 license),

chapter 4-from Descartes' "Treatise of Man" (1664), taken from stanford.library.sydney.edu.au/archives/spr2010/entries/mental-imagery.



Bibliography

- Abe91 Abeles, M. *Corticonics: Neural Circuits of the Cerebral Cortex* (Cambridge University Press, New York, 1991).
- ALW+02 Angelucci, A., Levitt, J. B., Walton, E. J., *et al.* Circuits for local and global signal integration in primary visual cortex. *Journal of Neuroscience* **22**, 8633–8646 (2002).
- AGS15 Arviv, O., Goldstein, A. & Shriki, O. Near-Critical Dynamics in Stimulus-Evoked Activity of the Human Brain and Its Relation to Spontaneous Resting-State Activity. *The Journal of Neuroscience* **35**, 13927–42 (2015).
- BP04 Baiesi, M. & Paczuski, M. Scale-free networks of earthquakes and aftershocks. *Phys. Rev. E* **69**, 066106 (2004).
- BCS+93 Bak, P., Chen, K., Scheinkman, J., *et al.* Aggregate fluctuations from independent sectoral shocks: self-organized criticality in a model of production and inventory dynamics. *Ricerche Economiche* **47**, 3–30 (1993).
- BS93 Bak, P. & Sneppen, K. Punctuated equilibrium and criticality in a simple model of evolution. *Phys. Rev. Lett.* **71**, 4083–4086 (1993).
- BTW87 Bak, P., Tang, C. & Wiesenfeld, K. Self-organized criticality: An explanation of the 1/f noise. *Physical Review Letters* **59**, 381–384 (1987).
- BUA+12 Bastos, A. M., Usrey, W. M., Adams, R. A., *et al.* Canonical microcircuits for predictive coding. *Neuron* **76**, 695–711 (2012).
- BH02 Bauer, R. & Heinze, S. Contour integration in striate cortex. Classic cell responses or cooperative selection? *Experimental Brain Research* **147**, 145–152 (2002).
- Beg08 Beggs, J. M. The criticality hypothesis: how local cortical networks might optimize information processing. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* **366**, 329–43 (2008).
- BP03 Beggs, J. M. & Plenz, D. Neuronal avalanches in neocortical circuits. *The Journal of Neuroscience* **23**, 11167–77 (2003).

- BT12 Beggs, J. M. & Timme, N. Being Critical of Criticality in the Brain. *Frontiers in Physiology* **3**, 163 (2012).
- BYBOS95 Ben-Yishai, R., Bar-Or, R. L. & Sompolinsky, H. Theory of orientation tuning in visual cortex. *Proceedings of the National Academy of Sciences* **92**, 3844–3848 (1995).
- Ber29 Berger, H. Über das Elektrenkephalogramm des Menschen. *Archiv für Psychiatrie und Nervenkrankheiten* **87**, 527–570 (1929).
- BN04 Bertschinger, N. & Natschläger, T. Real-Time Computation at the Edge of Chaos in Recurrent Neural Networks. *Neural Computation* **16**, 1413–1436 (2004).
- BL98 Bienenstock, E. & Lehmann, D. Regulated Criticality in the Brain? *Advances in Complex Systems* **01**, 361–384 (1998).
- Bin11 Binder, Kurt. *Ising model*, *Encyclopedia of Mathematics* [Online; accessed 5.12.2018] https://www.encyclopediaofmath.org/index.php/Ising_model. 2011.
- BS86 Blasdel, G. G. & Salama, G. Voltage-sensitive dyes reveal a modular organization in monkey striate cortex. *Nature* **321**, 579 (1986).
- Boh33 Bohr, N. Light and Life. *Nature* **131**, 421–423 (1933).
- BR00 Bornholdt, S. & Rohlf, T. Topological Evolution of Dynamical Networks: Global Criticality from Local Dynamics. *Physical Review Letters* **84**, 6114–6117 (2000).
- Bot95 Bottani, S. Pulse-coupled relaxation oscillators: from biological synchronization to self-organized criticality. *Physical Review Letters* **74**, 4189 (1995).
- BG97 Bröker, H.-M. & Grassberger, P. Random neighbor theory of the Olami-Feder-Christensen earthquake model. *Physical Review E* **56**, 3944–3952 (1997).
- Bru00 Brunel, N. Dynamics of sparsely connected networks of excitatory and inhibitory spiking neurons. *Journal of Computational Neuroscience* **8**, 183–208 (2000).
- BFL+10 Buffalo, E. A., Fries, P., Landman, R., *et al.* A backward progression of attentional effects in the ventral stream. *PNAS* **107**, 361–365 (2010).
- BT06 Buia, C. & Tiesinga, P. Attentional modulation of firing rate and synchrony in a model cortical network. *Journal of Computational Neuroscience* **20**, 247–264 (2006).
- BG06 Butts, D. A. & Goldman, M. S. Tuning curves, neuronal variability, and sensory coding. *PLoS biology* **4**, e92 (2006).
- BAK12 Buzsáki, G., Anastassiou, C. A. & Koch, C. The origin of extracellular fields and currents—EEG, ECoG, LFP and spikes. *Nature Reviews Neuroscience* **13**, 407–420 (2012).
- BW12 Buzsáki, G. & Wang, X.-J. Mechanisms of gamma oscillations. *Annual review of neuroscience* **35**, 203–225 (2012).
- CDJ+00 Camalet, S., Duke, T., Jülicher, F., *et al.* Auditory sensitivity provided by self-tuned critical oscillations of hair cells. *PNAS* **97**, 3183–3188 (2000).

- CHM97 Carandini, M., Heeger, D. J. & Movshon, J. A. Linearity and normalization in simple cells of the macaque primary visual cortex. *The Journal of Neuroscience* **17**, 8621–44 (1997).
- Car11 Carrasco, M. Visual attention: the past 25 years. *Vision Research* **51**, 1484–1525 (2011).
- CAR02 Chance, F. S., Abbott, L. & Reyes, A. D. Gain Modulation from Background Synaptic Input. *Neuron* **35**, 773–782 (2002).
- CL11 Chang, C.-C. & Lin, C.-J. LIBSVM: A library for support vector machines. *ACM Transactions on Intelligent Systems and Technology* **2**. Software available at <http://www.csie.ntu.edu.tw/~cjlin/libsvm>, 27:1–27:27 (2011).
- CYG+14 Chen, M. G., Yan, Y., Gong, X. J., *et al.* Incremental Integration of Global Contours through Interplay between Visual Cortical Areas. *Neuron* **82**, 682–694 (2014).
- Chi04 Chialvo, D. R. Critical brain networks. *Physica A: Statistical Mechanics and its Applications* **340**, 756–765 (2004).
- Chi10 Chialvo, D. R. Emergent complex neural dynamics. *Nature Physics* **6**, 744–750 (2010).
- CB99a Chialvo, D. & Bak, P. Learning from mistakes. *Neuroscience* **90**, 1137–1148 (1999).
- CM90 Clarke, S. & Miklossy, J. Occipital cortex in man: Organization of callosal connections, related myelo-and cytoarchitecture, and putative boundaries of functional visual areas. *Journal of Comparative Neurology* **298**, 188–214 (1990).
- CSN09 Clauset, A., Shalizi, C. R. & Newman, M. E. Power-law distributions in empirical data. *SIAM review* **51**, 661–703 (2009).
- CB99b Coombes, S. & Bressloff, P. C. Mode locking and Arnold tongues in integrate-and-fire neural oscillators. *Phys. Rev. E* **60**, 2086–2096 (1999).
- CPDG+95 Corral, A., Pérez, C. J., Díaz-Guilera, A., *et al.* Self-Organized Criticality and Synchronization in a Lattice Model of Integrate-and-Fire Oscillators. *Phys. Rev. Lett.* **74**, 118–121 (1995).
- CY89 Crutchfield, J. P. & Young, K. Inferring statistical complexity. *Phys. Rev. Lett.* **63**, 105–108 (1989).
- DHG06 David, S. V., Hayden, B. Y. & Gallant, J. L. Spectral Receptive Field Properties Explain Shape Selectivity in Area V4. *Journal of Neurophysiology* **96**, 3492–3505 (2006).
- DA01 Dayan, P. & Abbott, L. F. *Theoretical neuroscience : computational and mathematical modeling of neural systems* 1st (Massachusetts Institute of Technology Press, 2001).
- DAH12 De Arcangelis, L. & Herrmann, H. J. Activity-dependent neuronal model on complex networks. *Frontiers in Physiology* **3** (2012).
- DHH+12 Dehghani, N., Hatsopoulos, N. G., Haga, Z. D., *et al.* Avalanche Analysis from Multielectrode Ensemble Recordings in Cat, Monkey, and Human Cerebral Cortex during Wakefulness and Sleep. *Frontiers in Physiology* **3**, 302 (2012).

- DS87 Desimone, R. & Schein, S. Visual properties of neurons in area V4 of the macaque: sensitivity to stimulus form. *Journal of neurophysiology* **57**, 835–68 (1987).
- Dow09 Dowling, J. in *Encyclopedia of Neuroscience* (ed Squire, L. R.) 159–169 (Academic Press, Oxford, 2009).
- DS92 Drossel, B. & Schwabl, F. Self-organized critical forest-fire model. *Physical Review Letters* **69**, 1629–1632 (1992).
- EBJ+88 Eckhorn, R., Bauer, R., Jordan, W., *et al.* Coherent oscillations: A mechanism of feature linking in the visual cortex? *Biological Cybernetics* **60**, 121–130 (1988).
- EMS+12 Ernst, U. A., Mandon, S., Schinkel–Bielefeld, N., *et al.* Optimality of Human Contour Integration. *PLoS Computational Biology* **8**, e1002520 (2012).
- EHE02 Eurich, C. W., Herrmann, J. M. & Ernst, U. A. Finite-size effects of avalanche dynamics. *Physical Review E* **66**, 066137 (2002).
- FLS+15 Fagerholm, E. D., Lorenz, R., Scott, G., *et al.* Cascades and cognitive state: focused attention incurs subcritical dynamics. *The Journal of Neuroscience* **35**, 4626–34 (2015).
- FHH93 Field, D. J., Hayes, A. & Hess, R. F. Contour integration by the human visual system: Evidence for a local “association field”. *Vision Research* **33**, 173–193 (1993).
- FWN+95 Field, S., Witt, J., Nori, F., *et al.* Superconducting Vortex Avalanches. *Phys. Rev. Lett.* **74**, 1206–1209 (1995).
- FK14 Finger, H. & König, P. Phase synchrony facilitates binding and segmentation of natural images in a coupled neural oscillator network. *Frontiers in Computational Neuroscience* **7**, 195 (2014).
- FBF+09 Fraiman, D., Balenzuela, P., Foss, J., *et al.* Ising-like dynamics in large-scale functional brain networks. *Phys. Rev. E* **79**, 061922 (2009).
- FCMS+96 Frette, V., Christensen, K., Mølle-Sørensen, A., *et al.* Avalanche dynamics in a pile of rice. *Nature* **379**, 49 (1996).
- FIB+12 Friedman, N., Ito, S., Brinkman, B. A. W., *et al.* Universal Critical Dynamics in High Resolution Neuronal Avalanche Data. *Physical Review Letters* **108**, 208102 (2012).
- Fri05a Fries, P. A mechanism for cognitive dynamics: neuronal communication through neuronal coherence. *Trends in Cognitive Sciences* **9**, 474–480 (2005).
- Fri05b Fries, P. A mechanism for cognitive dynamics: neuronal communication through neuronal coherence. *Trends in Cognitive Sciences* **9**, 474–480 (2005).
- FRR+01 Fries, P., Reynolds, J. H., Rorie, A. E., *et al.* Modulation of oscillatory neuronal synchronization by selective visual attention. *Science* **291**, 1560–3 (2001).
- Gal99 Gallavotti, G. *Statistical mechanics: A short treatise* (Springer-Verlag, Berlin, 1999).
- GHM+15 Gautam, S. H., Hoang, T. T., McClanahan, K., *et al.* Maximizing Sensory Dynamic Range by Tuning the Cortical State to Criticality. *PLOS Computational Biology* **11**, e1004576 (2015).

- GPS+01 Geisler, W. S., Perry, J. S., Super, B., *et al.* Edge co-occurrence in natural images predicts contour grouping performance. *Vision Research* **41**, 711–724 (2001).
- GMS13 Gilad, A., Meirovithz, E. & Slovin, H. Population responses to contour integration: Early encoding of discrete elements and late perceptual grouping. *Neuron* **78**, 389–402 (2013).
- GP08 Gireesh, E. D. & Plenz, D. Neuronal avalanches organize as nested theta- and beta/gamma-oscillations during development of cortical layer 2/3. *PNAS* **105**, 7576–7581 (2008).
- Gra02 Grassberger, P. Critical behaviour of the Drossel-Schwabl forest fire model. *New Journal of Physics* **4** (2002).
- GS89 Gray, C. M. & Singer, W. Stimulus-specific neuronal oscillations in orientation columns of cat visual cortex. *PNAS* **86**, 1698–1702 (1989).
- GPW03 Grigorescu, C., Petkov, N. & Westenberg, M. A. Contour detection based on nonclassical receptive field inhibition. *IEEE Transactions on Image Processing* **12**, 729–739 (2003).
- GNM+12a Grothe, I., Neitzel, S. D., Mandon, S., *et al.* Switching neuronal inputs by differential modulations of gamma-band phase-coherence. *The Journal of Neuroscience* **32**, 16172–80 (2012).
- GNM+12b Grothe, I., Neitzel, S. D., Mandon, S., *et al.* Switching neuronal inputs by differential modulations of gamma-band phase-coherence. *The Journal of Neuroscience* **32**, 16172–16180 (2012).
- GRN+18 Grothe, I., Rotermund, D., Neitzel, S. D., *et al.* Attention Selectively Gates Afferent Signal Transmission to Area V4. *The Journal of Neuroscience* **38**, 3441–3452 (2018).
- GGE17 Grzymisch, A., Grimsen, C. & Ernst, U. A. Contour Integration in Dynamic Scenes: Impaired Detection Performance in Extended Presentations. *Frontiers in Psychology* **8**, 1501 (2017).
- GDGL+00 Guardiola, X., Díaz-Guilera, A., Llas, M., *et al.* Synchronization, diversity, and topology of networks of integrate and fire oscillators. *Phys. Rev. E* **62**, 5565–5570 (2000).
- HPAM+17 Hahn, G., Ponce-Alvarez, A., Monier, C., *et al.* Spontaneous cortical activity is transiently poised close to criticality. *PLOS Computational Biology* **13**, e1005543 (2017).
- HB05 Haldeman, C. & Beggs, J. M. Critical Branching Captures Activity in Living Neural Networks and Maximizes the Number of Metastable States. *Physical Review Letters* **94**, 058101 (2005).
- HEP15 Harnack, D., Ernst, U. A. & Pawelzik, K. R. A model for attentional information routing through coherence predicts biased competition and multistable perception. *Journal of Neurophysiology* **114**, 1593–1605 (2015).
- Heb49 Hebb, D. O. *The organization of behavior: A neuropsychological theory* (Wiley, New York, 1949).
- HSS+90 Held, G. A., Solina, D., Solina, H., *et al.* Experimental study of critical-mass fluctuations in an evolving sandpile. *Phys. Rev. Lett.* **65**, 1120–1123 (1990).

- HH95 Herz, A. V. & Hopfield, J. J. Earthquake cycles and neural reverberations: collective oscillations in systems with pulse-coupled threshold elements. *Physical Review Letters* **75**, 1222 (1995).
- HHF03 Hess, R., Hayes, A & Field, D. Contour integration and cortical processing. *Journal of Physiology-Paris* **97**, 105–119 (2003).
- HG14 Hesse, J. & Gross, T. Self-organized criticality as a fundamental property of neural systems. *Frontiers in Systems Neuroscience* **8**, 166 (2014).
- Hop82 Hopfield, J. J. Neural networks and physical systems with emergent collective computational abilities. *PNAS* **79**, 2554–2558 (1982).
- Hor06 Horton, J. C. Ocular integration in the human visual cortex. *Canadian Journal of Ophthalmology* **41**, 584–593 (2006).
- HH02 Hoyer, P. O. & Hyvärinen, A. A multi-layer sparse coding network learns contour coding from natural images. *Vision Research* **42**, 1593–1605 (2002).
- HR11 Huang, Y. & Rao, R. P. Predictive coding. *Wiley Interdisciplinary Reviews: Cognitive Science* **2**, 580–593 (2011).
- HW68 Hubel, D. H. & Wiesel, T. N. Receptive fields and functional architecture of monkey striate cortex. *The Journal of Physiology* **195**, 215–243 (1968).
- Jae98 Jaeger, G. The Ehrenfest Classification of Phase Transitions: Introduction and Evolution. *Archive for History of Exact Sciences* **53**, 51–81 (1998).
- Jen98 Jensen, H. J. *Self-organized criticality: emergent complex behavior in physical and biological systems* (Cambridge university press, Cambridge, UK, 1998).
- JGW+01 Jones, H. E., Grieve, K. L., Wang, W., *et al.* Surround Suppression in Primate V1. *Journal of Neurophysiology* **86**, 2011–2028 (2001).
- Kad66 Kadanoff, L. P. Scaling laws for Ising models near T_c . *Physics Physique Fizika* **2**, 263 (1966).
- KSJ+00 Kandel, E. R., Schwartz, J. H., Jessell, T. M., *et al.* *Principles of Neural Science, Fourth Edition* (McGraw-Hill Companies, Incorporated, New York, 2000).
- KIG+95 Kapadia, M. K., Ito, M., Gilbert, C. D., *et al.* Improvement in visual sensitivity by changes in local context: parallel studies in human observers and in V1 of alert monkeys. *Neuron* **15**, 843–856 (1995).
- Kau93 Kauffman, S. A. *The origins of order: Self-organization and selection in evolution* (Oxford University Press, New York, 1993).
- KC06 Kinouchi, O. & Copelli, M. Optimal dynamical range of excitable networks at criticality. *Nature Physics* **2**, 348–351 (2006).
- KTR+97 Kisvarday, Z., Tóth, E., Rausch, M., *et al.* Orientation-specific relationship between populations of excitatory and inhibitory lateral connections in the visual cortex of the cat. *Cerebral Cortex* **7**, 605–618 (1997).
- KK80 Kittel, C. & Kroemer, H. *Thermal Physics* (WH Freeman and Co., New York, 1980).
- KYP11 Klaus, A., Yu, S. & Plenz, D. Statistical analyses support power law distributions found in neuronal avalanches. *PLoS ONE* **6**, e19779 (2011).

- KGA+18 Kossio, F. Y. K., Goedeke, S., van den Akker, B., *et al.* Growing critical: self-organized criticality in a developing neural system. *Physical Review Letters* **121**, 058301 (2018).
- KJ93 Kovács, I & Julesz, B. A closed curve is much more than an incomplete one: effect of closure in figure-ground segmentation. *PNAS* **90**, 7495–7 (1993).
- KS96 Kreiter, A. K. & Singer, W. Stimulus-dependent synchronization of neuronal responses in the visual cortex of the awake macaque monkey. *The Journal of Neuroscience* **16**, 2381–96 (1996).
- KMMG87 Kronland-Martinet, R., Morlet, J. & Grossmann, A. Analysis of sound patterns through wavelet transforms. *International Journal of Pattern Recognition and Artificial Intelligence* **1**, 273–302 (1987).
- LFHC+00 Lago-Fernández, L. F., Huerta, R., Corbacho, F., *et al.* Fast response and temporal coherent oscillations in small-world networks. *Phys. Rev. Lett.* **84**, 2758–2761 (2000).
- Lan90 Langton, C. G. Computation at the edge of chaos: Phase transitions and emergent computation. *Physica D: Nonlinear Phenomena* **42**, 12–37 (1990).
- LP17 Levina, A. & Priesemann, V. Subsampling scaling. *Nature Communications* **8**, 15140 (2017).
- LHG07 Levina, A., Herrmann, J. M. & Geisel, T. Dynamical synapses causing self-organized criticality in neural networks. *Nature Physics* **3**, 857 (2007).
- LL97 Levitt, J. B. & Lund, J. S. Contrast dependence of contextual effects in primate visual cortex. *Nature* **387**, 73–76 (1997).
- LG02 Li, W. & Gilbert, C. D. Global contour saliency and local colinear interactions. *Journal of Neurophysiology* **88**, 2846–2856 (2002).
- LPG06 Li, W., Piëch, V. & Gilbert, C. D. Contour Saliency in Primary Visual Cortex. *Neuron* **50**, 951–962 (2006).
- LPG08 Li, W., Piëch, V. & Gilbert, C. D. Learning to link visual contours. *Neuron* **57**, 442–451 (2008).
- Li98 Li, Z. A neural model of contour integration in the primary visual cortex. *Neural Computation* **10**, 903–940 (1998).
- LPE10 Lindén, H., Pettersen, K. H. & Einevoll, G. T. Intrinsic dendritic filtering gives low-pass power spectra of local field potentials. *Journal of Computational Neuroscience* **29**, 423–444 (2010).
- LTP+11 Lindén, H., Tetzlaff, T., Potjans, T. C., *et al.* Modeling the Spatial Reach of the LFP. *Neuron* **72**, 859–872 (2011).
- LJ96 Lise, S. & Jensen, H. J. Transitions in nonconserving models of self-organized criticality. *Physical Review Letters* **76**, 2326 (1996).
- LD09 Liu, Y. & Dahmen, K. A. Unexpected universality in static and dynamic avalanches. *Physical Review E* **79**, 061124 (2009).
- LCH+97 Luck, S., Chelazzi, L., Hillyard, S., *et al.* Neural mechanisms of spatial selective attention in areas V1, V2 and V4 of macaque visual cortex. *Journal of Neurophysiology* **77**, 24–42 (1997).

- MW11 Machilsen, B. & Wagemans, J. Integration of contour and surface information in shape detection. *Vision Research* **51**, 179–186 (2011).
- Mac03 Mack, A. Inattentional Blindness: Looking Without Seeing. *Current Directions in Psychological Science* (2003).
- Mag00 Magee, J. C. Dendritic integration of excitatory synaptic input. *Nature Reviews Neuroscience* **1**, 181 (2000).
- MOK+14 Makarova, J., Ortuno, T., Korovaichuk, A., *et al.* Can pathway-specific LFPs be obtained in cytoarchitectonically complex structures? *Frontiers in Systems Neuroscience* **8** (2014).
- MK05 Mandon, S. & Kreiter, A. K. Rapid contour integration in macaque monkeys. *Vision Research* **45**, 291–300 (2005).
- MKK90 Manna, S., Kiss, L. B. & Kertész, J. Cascades and self-organized criticality. *Journal of Statistical Physics* **61**, 923–932 (1990).
- MEBF+09 Marre, O., El Boustani, S., Frégnac, Y., *et al.* Prediction of spatiotemporal patterns of neural activity from pairwise correlations. *Physical Review Letters* **102**, 138101 (2009).
- MPM15 Massobrio, P., Pasquale, V. & Martinoia, S. Self-organized criticality in cortical assemblies occurs in concurrent scale-free and small-world networks. *Scientific Reports* **5**, 10578 (2015).
- MF07 Mathes, B. & Fahle, M. Closure facilitates contour integration. *Vision Research* **47**, 818–827 (2007).
- MOS+13 Meisel, C., Olbrich, E., Shriki, O., *et al.* Fading signatures of critical brain dynamics during sustained wakefulness in humans. *The Journal of Neuroscience* **33**, 17363–72 (2013).
- MB11 Mora, T. & Bialek, W. Are Biological Systems Poised at Criticality? *Journal of Statistical Physics* **144**, 268–302 (2011).
- MDM15 Mora, T., Deny, S. & Marre, O. Dynamical Criticality in the Collective Activity of a Population of Retinal Neurons. *Physical Review Letters* **114**, 078105 (2015).
- MD85 Moran, J. & Desimone, R. Selective attention gates visual processing in the extrastriate cortex. *Science* **229**, 782–784 (1985).
- Mot09 Motter, B. C. Central V4 receptive fields are scaled by the V1 cortical magnification and correspond to a constant-sized sampling of the V1 surface. *Journal of Neuroscience* **29**, 5749–5757 (2009).
- NB99 Newman, M. E. J. & Barkema, G. T. *Monte Carlo Methods in Statistical Physics* (Oxford University Press, New York, USA, 1999).
- NML+03 Nishikawa, T., Motter, A. E., Lai, Y.-C., *et al.* Heterogeneity in oscillator networks: Are smaller worlds easier to synchronize? *Phys. Rev. Lett.* **91**, 014101 (2003).
- NSD+05 Noback, C. R., Strominger, N. L., Demarest, R. J., *et al.* *The Human Nervous System: Structure and Function* (Humana Press, New Jersey, 2005).
- NBB+17 Nonnenmacher, M., Behrens, C., Berens, P., *et al.* Signatures of criticality arise from random subsampling in simple population models. *PLOS Computational Biology* **13**, e1005718 (2017).

- NPL+08 Nykter, M., Price, N. D., Larjo, A., *et al.* Critical Networks Exhibit Maximal Information Diversity in Structure-Dynamics Relationships. *Physical Review Letters* **100**, 058702 (2008).
- OL08 Okun, M. & Lampl, I. Instantaneous correlation of excitation and inhibition during ongoing and sensory-evoked activities. *Nature Neuroscience* **11**, 535 (2008).
- OFC92 Olami, Z., Feder, H. J. S. & Christensen, K. Self-organized criticality in a continuous, nonconservative cellular automaton modeling earthquakes. *Physical Review Letters* **68**, 1244–1247 (1992).
- Ons44 Onsager, L. Crystal Statistics. I. A Two-Dimensional Model with an Order-Disorder Transition. *Physical Review* **65**, 117–149 (1944).
- PZH+13 Palva, J. M., Zhigalov, A., Hirvonen, J., *et al.* Neuronal long-range temporal correlations and avalanche dynamics are correlated with behavioral scaling laws. *PNAS* **110**, 3585–90 (2013).
- PBS+11 Papanikolaou, S., Bohn, F., Sommer, R. L., *et al.* Universality beyond power laws and the average avalanche shape. *Nature Physics* **7**, 316–320 (2011).
- PMB+08 Pasquale, V., Massobrio, P., Bologna, L. L., *et al.* Self-organization and neuronal avalanches in networks of dissociated cortical neurons. *Neuroscience* **153**, 1354–1369 (2008).
- PP13 Patzelt, F. & Pawelzik, K. An inherent instability of efficient markets. *Scientific Reports* **3**, 2784 (2013).
- PRE+07 Patzelt, F., Riegel, M., Ernst, U., *et al.* Self-organized critical noise amplification in human closed loop control. *Frontiers in Computational Neuroscience* **1**, 4 (2007).
- PMB+97 Pelletier, J. D., Malamud, B. D., Blodgett, T., *et al.* Scale-invariance of soil moisture variability and its implications for the frequency-size distribution of landslides. *Engineering Geology* **48**, 255–268 (1997).
- PDS95 Perković, O., Dahmen, K. & Sethna, J. P. Avalanches, Barkhausen Noise, and Plain Old Criticality. *Physical Review Letters* **75**, 4528–4531 (1995).
- PTL+09 Petermann, T., Thiagarajan, T. C., Lebedev, M. A., *et al.* Spontaneous cortical activity in awake monkeys composed of neuronal avalanches. *PNAS* **106**, 15921–6 (2009).
- PPB94 Peters, A., Payne, B. R. & Budd, J. A numerical analysis of the geniculocortical input to striate cortex in the monkey. *Cerebral Cortex* **4**, 215–229 (1994).
- Pla03 Planck, M. *Treatise on Thermodynamics* (Longmans, Green and Co., London, 1903).
- Ple14 Plenz, D. in *Criticality in Neural Systems* 5–42 (Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim, Germany, 2014).
- PNB93 Plourde, B., Nori, F. & Bretz, M. Water droplet avalanches. *Phys. Rev. Lett.* **71**, 2749–2752 (1993).
- PHM+12 Poil, S.-S., Hardstone, R., Mansvelder, H. D., *et al.* Critical-State Dynamics of Avalanches and Oscillations Jointly Emerge from Balanced Excitation/Inhibition in Neuronal Networks. *The Journal of Neuroscience* **32**, 9817–9823 (2012).

- PS93 Polat, U. & Sagi, D. Lateral interactions between spatial channels: Suppression and facilitation revealed by lateral masking experiments. *Vision Research* **33**, 993–999 (1993).
- Pre78 Press, W. H. Flicker noises in astronomy and elsewhere. *Comments on Astrophysics* **7**, 103–119 (1978).
- PMW09 Priesemann, V., Munk, M. H. & Wibral, M. Subsampling effects in neuronal avalanche distributions recorded in vivo. *BMC neuroscience* **10**, 40 (2009).
- PVW+13 Priesemann, V., Valderrama, M., Wibral, M., *et al.* Neuronal Avalanches Differ from Wakefulness to Deep Sleep – Evidence from Intracranial Depth Recordings in Humans. *PLoS Computational Biology* **9**, e1002985 (2013).
- PWV+14 Priesemann, V., Wibral, M., Valderrama, M., *et al.* Spike avalanches in vivo suggest a driven, slightly subcritical brain state. *Frontiers in Systems Neuroscience* **8**, 108 (2014).
- QRK+05 Quiroga, R. Q., Reddy, L., Kreiman, G., *et al.* Invariant visual representation by single neurons in the human brain. *Nature* **435**, 1102 (2005).
- RPD00 Reynolds, J. H., Pasternak, T. & Desimone, R. Attention increases sensitivity of V4 neurons. *Neuron* **26**, 703–714 (2000).
- RA96 Rhodes, C. J. & Anderson, R. M. Power laws governing epidemics in isolated populations. *Nature* **381**, 600 (1996).
- RGF+94 Ritz, R., Gerstner, W., Fuentes, U., *et al.* A biologically motivated and analytically soluble model of collective oscillations in the cortex. II. Application to binding and pattern segmentation. *Biol Cybern.* **71**(4), 349–358 (1994).
- Roe06 Roelfsema, P. R. Cortical algorithms for perceptual grouping. *Annual Review of Neuroscience* **29**, 203–227 (2006).
- RTE+09 Rotermund, D., Taylor, K., Ernst, U. A., *et al.* Attention Improves Object Representation in Visual Cortical Field Potentials. *Journal of Neuroscience* **29**, 10120–10130 (2009).
- SAM+18 Scarpetta, S., Apicella, I., Minati, L., *et al.* Hysteresis, neural avalanches, and critical behavior near a first-order transition of a spiking neural network. *Physical Review E* **97**, 062305 (2018).
- SC13 Scarpetta, S. & de Candia, A. Neural Avalanches at the Critical Point between Replay and Non-Replay of Spatiotemporal Patterns. *PLoS ONE* **8**, e64162 (2013).
- SBS+06 Schneidman, E., Berry, M. J., Segev, R., *et al.* Weak pairwise correlations imply strongly correlated network states in a neural population. *Nature* **440**, 1007–1012 (2006).
- SGG+16 Schoenholz, S. S., Gilmer, J., Ganguli, S., *et al.* Deep Information Propagation. arXiv: 1611.01232 (2016).
- SCD+18 Seeman, S. C., Campagnola, L., Davoudian, P. A., *et al.* Sparse recurrent excitatory connectivity in the microcircuit of the adult mouse and human cortex. *eLife* **7**, e37349 (2018).
- SDM01 Sethna, J. P., Dahmen, K. A. & Myers, C. R. Crackling noise. *Nature* **410**, 242–250 (2001).

- Sha48 Shannon, C. E. A Mathematical Theory of Communication. *Bell System Technical Journal* **27**, 379–423 (1948).
- SCP+15 Shew, W. L., Clawson, W. P., Pobst, J., *et al.* Adaptation to sensory input tunes visual cortex to criticality. *Nature Physics* **11**, 659–663 (2015).
- SP13 Shew, W. L. & Plenz, D. The Functional Benefits of Criticality in the Cortex. *The Neuroscientist* **19**, 88–100 (2013).
- SYP+09 Shew, W. L., Yang, H., Petermann, T., *et al.* Neuronal avalanches imply maximum dynamic range in cortical networks at criticality. *The Journal of Neuroscience* **29**, 15595–600 (2009).
- SFG+06 Shlens, J., Field, G. D., Gauthier, J. L., *et al.* The Structure of Multi-Neuron Firing Patterns in Primate Retina. *Journal of Neuroscience* **26**, 8254–8266 (2006).
- SAC+13 Shriki, O., Alstott, J., Carver, F., *et al.* Neuronal avalanches in the resting MEG of the human brain. *The Journal of Neuroscience* **33**, 7079–90 (2013).
- SGJ+95 Sillito, A. M., Grieve, K. L., Jones, H. E., *et al.* Visual cortical mechanisms detecting focal orientation discontinuities. *Nature* **378**, 492–496 (1995).
- SB13 Sinz, F. & Bethge, M. Temporal adaptation enhances efficient contrast gain control on natural images. *PLoS Computational Biology* **9**, e1002889 (2013).
- Smi48 Smirnov, N. Table for Estimating the Goodness of Fit of Empirical Distributions. *The Annals of Mathematical Statistics* **19**, 279–281 (1948).
- SK93 Softky, W. & Koch, C. The highly irregular firing of cortical cells is inconsistent with temporal integration of random EPSPs. *Journal of Neuroscience* **13**, 334–350 (1993).
- SHH19 Sorbaro, M., Herrmann, M. J. & Hennig, M. *Statistical models of neural activity, criticality, and Zipf's law* In print, in *The Functional Role of Critical Dynamics*. 2019.
- Sta71 Stanley, H. E. *Introduction to Phase Transitions and Critical Phenomena* (Oxford University Press, New York, 1971).
- SA94 Stauffer, D. & Aharony, A. *Introduction to Percolation Theory* (Taylor and Francis Inc., Philadelphia, 1994).
- SDB+02 Stettler, D. D., Das, A., Bennett, J., *et al.* Lateral connectivity and contextual interactions in macaque primary visual cortex. *Neuron* **36**, 739–750 (2002).
- SP12 Stumpf, M. P. & Porter, M. A. Critical truths about power laws. *Science* **335**, 665–666 (2012).
- TMF+05 Taylor, K., Mandon, S., Freiwald, W., *et al.* Coherent Oscillatory Activity in Monkey Area V4 Predicts Successful Allocation of Attention. *Cerebral Cortex* **15**, 1424–1437 (2005).
- The10 Theeuwes, J. Top–down and bottom–up control of visual selection. *Acta Psychologica* **135**, 77–99 (2010).
- TFM96 Thorpe, S., Fize, D. & Marlot, C. Speed of processing in the human visual system. *Nature* **381**, 520–522 (1996).
- TS09 Tiesinga, P. & Sejnowski, T. J. Cortical enlightenment: are attentional gamma oscillations driven by ING or PING? *Neuron* **63**, 727–732 (2009).

- TV14 Tinker, J. & Velazquez, J. L. P. Power law scaling in synchronization of brain signals depends on cognitive load. *Frontiers in Systems Neuroscience* **8**, 73 (2014).
- TMM+13 Tkačik, G., Marre, O., Mora, T., *et al.* The simplest maximum entropy model for collective behavior in a neural network. *Journal of Statistical Mechanics: Theory and Experiment* **2013**, P03011 (2013).
- TMM+15 Tkačik, G., Mora, T., Marre, O., *et al.* Thermodynamics and signatures of criticality in a network of neurons. *PNAS* **112**, 11508–13 (2015).
- TRE14 Tomen, N., Rotermund, D. & Ernst, U. Marginally subcritical dynamics explain enhanced stimulus discriminability under attention. *Frontiers in Systems Neuroscience* **8** (2014).
- TT95 Toner, J. & Tu, Y. Long-range order in a two-dimensional dynamical XY model: how birds fly together. *Physical review letters* **75**, 4326 (1995).
- TC98 Torrence, C. & Compo, G. P. A practical guide to wavelet analysis. *Bulletin of the American Meteorological Society* **79**, 61–78 (1998).
- TD10 Touboul, J. & Destexhe, A. Can Power-Law Scaling and Neuronal Avalanches Arise from Stochastic Dynamics? *PLoS ONE* **5**, e8982 (2010).
- TD17 Touboul, J. & Destexhe, A. Power-law statistics and universal scaling in the absence of criticality. *Physical Review E* **95**, 012413 (2017).
- TG80 Treisman, A. M. & Gelade, G. A feature-integration theory of attention. *Cognitive Psychology* **12**, 97–136 (1980).
- TM96 Treue, S. & Maunsell, J. Attentional modulation of visual motion processing in cortical areas MT and MST. *Nature* **382**, 539–541 (1996).
- Tre01 Treue, S. Neural correlates of attention in primate visual cortex. *Trends in Neurosciences* **24**, 295–300 (2001).
- TK00 Tsuchiya, T. & Katori, M. Proof of breaking of self-organized criticality in a nonconservative abelian sandpile model. *Physical Review E* **61**, 1183 (2000).
- UH94 Ungerleider, L. G. & Haxby, J. V. ‘What’ and ‘where’ in the human brain. *Current opinion in neurobiology* **4**, 157–165 (1994).
- USO95 Usher, M., Stemmler, M. & Olami, Z. Dynamic Pattern Formation Leads to 1/f Noise in Neural Populations. *Physical Review Letters* **74**, 326 (1995).
- OBO19 Van Ooyen, A. & Butz-Ostendorf, M. *Homeostatic structural plasticity can build critical networks* In print, in *The Functional Role of Critical Dynamics*. 2019.
- VEOA+91 Van Essen, D., Olshausen, B. A., Anderson, C. H., *et al.* *Pattern recognition, attention, and information bottlenecks in the primate visual system* in *Visual Information Processing: From Neurons to Chips* **1473** (1991), 17–29.
- VSZ+11 Vogels, T. P., Sprekeler, H., Zenke, F., *et al.* Inhibitory plasticity balances excitation and inhibition in sensory pathways and memory networks. *Science (New York, N.Y.)* **334**, 1569–73 (2011).
- vdM94 Von der Malsburg, C. in *Models of neural networks* 95–119 (Springer, New York, NY, 1994).

- WEK+12 Wagemans, J., Elder, J. H., Kubovy, M., *et al.* A century of Gestalt psychology in visual perception: I. Perceptual grouping and figure-ground organization. *Psychological bulletin* **138**, 1172–217 (2012).
- WDB07 Wandell, B. A., Dumoulin, S. O. & Brewer, A. A. Visual field maps in human cortex. *Neuron* **56**, 366–383 (2007).
- Wei02 Weisstein, E. W. *The CRC Concise Encyclopedia of Mathematics, Second Edition* (Chapman & Hall/CRC Press, Boca Raton, 2002).
- WB04 West, G. B. & Brown, J. H. Life’s universal scaling laws. *Physics Today* **57**, 36–43 (2004).
- Wil17 Wilson, S. P. Self-organised criticality in the evolution of a thermodynamic model of rodent thermoregulatory huddling. *PLoS computational biology* **13**, e1005378 (2017).
- Yan52 Yang, C. N. The Spontaneous Magnetization of a Two-Dimensional Ising Model. *Physical Review* **85**, 808–816 (1952).
- YF98 Yen, S.-C. & Finkel, L. H. Extraction of perceptually salient contours by striate cortical networks. *Vision Research* **38**, 719–741 (1998).
- ZLS95 Zapperi, S., Lauritsen, K. B. & Stanley, H. E. Self-Organized Branching Processes: Mean-Field Theory for Avalanches. *Physical Review Letters* **75**, 4071–4074 (1995).
- ZG13 Zare, M. & Grigolini, P. Criticality and avalanches in neural networks. *Chaos, Solitons & Fractals* **55**, 80–94 (2013).
- ZZZ+12 Zhang, X., Zhaoping, L., Zhou, T., *et al.* Neural Activities in V1 Create a Bottom-Up Saliency Map. *Neuron* **73**, 183–192 (2012).

