ATTENTION: A COMPLEX SYSTEM

FROM THE INTRICATE MODULATION OF TUNED RESPONSES TOWARDS A LAYERED CORTICAL CIRCUIT MODEL

Dissertation

zur Erlangung des mathematisch-naturwissenschaftlichen Doktorgrades
"Doctor rerum naturalium"
der Georg-August-Universität Göttingen

vorgelegt von

MARKUS HELMER

aus Ulm

Göttingen 2015

Betreuungsausschuss

- Prof. Dr. Theo Geisel (Referent)
  Abteilung für nichlineare Dynamik
  Max-Planck-Institut für Dynamik und Selbstorganistation
- Prof. Dr. Fred Wolf (Referent)
  Abteilung für nichlineare Dynamik
  Max-Planck-Institut für Dynamik und Selbstorganistation
- Prof. Dr. Stefan Treue
  Abteilung Kognitive Neurowissenschaften
  Deutsches Primatenzentrum

CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACRONYMS

BC	biased competition
BOLD	blood oxygen level-dependent
CFC	cross-frequency coupling
CTC	communication-through-coherence
FB	feedback
FF	feedforward

FLNe extrinsic fraction of labeled neurons

fMRI functional magnetic resonance imaging

FS fast-spiking

FSGM feature similarity gain model

FEF frontal eye field

HRP horseradish peroxidase

IB intrinsically bursting

LIP lateral intraparietal area

LFP local field potential

LGN lateral geniculate nucleus of the thalamus

MGM multiplicative gain modulation

MT middle temporal area

MUA multi-unit activity

PCA principal component analysis

PING pyramidal interneuron network gamma

RDP random dot pattern

RF receptive field

ROC receiver-operator-characteristic

RS regular-spiking

SFC spike field coherence

SI slow-spiking inhibitory

SLN supragranular labeled neurons

STA spike-triggered average

0 OVERVIEW

Experiments have elucidated the neuronal correlates of attention in the primate brain, and have led to a number of models capturing various aspects of these experiments. However, these model propose qualitatively different ways of how attention acts, and often stay unclear as to how attention emerges from the biological constituents. In the first part of this work we have further demonstrated novel attentional modulation patterns, that are highly cell- and stimulus-specific and lead to—over the population—complex, non-multiplicative shape changes of tuning curves, that don't seem to be compatible with any current phenomenological (like the biased competition or the feature-similarity-gain model) or circuit model (like the Ardid-Wang-Compte model) of attention.

Whereas phenomenological models fit experimental observations into an abstract, high-level description, circuit models aim at describing these data as an emergent property of the interaction of suitably chosen low-level constituents. These interactions can be constrained in a principled way through better and better fine-grained connectivity data and based on this structural skeleton complex dynamics might emerge. For example, oscillations and their interdependence have been hypothesized to play a role in the coordination between brain network constituents and, moreover, given a fixed structural skeleton, the circuit might possess a multitude of states, due to multistability or, more profanely, due to a variation in parameters like coupling efficacy or neuromodulators.

We hypothesize, thus, that if the circuit model is sufficiently good, it will possess states with dynamical fingerprints resembling functional neuronal correlates, like those occuring during attention. While we are far from such a brain-wide circuit model for attention, we have investigated in the first part of this work a multi-ring circuit model to reproduce the attentional population effects mentioned above, without yet, however, achieving satisfactory results.

Moreover, given the prominent role that cortical rhythms are hypothesized to play in interareal coordination, in the second part of this work, we investigate oscillations in a simple rate model for a cortical column with realistic interlayer connectivity, observing complex layer-specific multi-frequency oscillations, with upper and lower layers oscillating predominantely at fast (gamma-like) and slow (beta-like) frequencies, in line with experimental findings and suggesting that the cortical column might form an important building block for communication-through-coherence processes which are modulated

by attention. We show further, that this pattern of oscillations depends crucially on this, or some structurally degenerate, connectomes, arguing against the arbitrariness of structure for brain function.

Moreover, when two columns at different hierarchical levels are coupled, we obtained preliminary results indicating that a self-organized directed coupling can emerge that is "feedforward" in the gamma- but "feedback" in the beta-band, in line with a currently discussed role of cortical oscillations.

In the future, building on a more systematic understanding of the two-columns system, we aim to study emerging multi-frequency oscillations in a brain-wide model with realistic, layer-resolved topology and their potential for interareal coordination and information processing in order to eventually obtain a better understanding of how the capacity to pay attention, and its neuronal correlates, emerges in the brain.

1 INTRODUCTION

Attention is one key constituent of higher brain functions. It enables us to filter sensory input and focus on what is relevant. Apart from stimulus-driven "bottom-up" saliency cues, attention can also be directed voluntarily in a "top-down" fashion to intrinsically selected targets and attention thus involves both sensory and control processes in the brain. While the latter are thought to be situated in frontal and parietal areas, sensory systems occupy distinct areas, for example primary visual and auditory cortex are located in the occipital and temporal lobe, respectively [Kandel et al., 2012]. Paying attention thus involves the dynamic coordination of a distributed network of areas, and it cannot be explained in terms of single neurons or areas. But how, then, can "attention" be explained? Thousands of research papers [Carrasco, 2011] elucidate various aspects of this question, yet, at this point, no all-embracing answer is in sight. What we are after, ultimately, is a mechanistic understanding of the inner workings of the attentional system. That is, we would like to explain the filtering and focusing of sensory input in terms of the neuronal substrate: the involved neurons, their connections and dynamics. Such a fundamental understanding is, of course, interesting in its own right, but beyond that it could serve as a guide in the exploration of other cognitive functions and give important clues for the treatment of attentional disorders. We will now set the stage for this endeavor by discussing the phenomenology that we want to explain, as well as the framework of the model—each of which comes with its own specific challenges. In order to limit the scope we will restrict ourselves to the visual system and "top-down" attention.

The hallmark neuronal correlate associated with attention is the modulation of firing rates of those neurons processing attended versus unattended stimuli. Generally, the former are found to have, on average, higher firing rates than the latter.

A large body of experiments along this line has been summarized in phenomenological models of attention. These models provide a concise account of many datasets, and can also simply be extrapolated to situations for which no experiment have been performed yet. This extrapolation is, of course, a hypothesis until experimentally confirmed. Beyond that, importantly, phenomenological models are inherently abstract high-level descriptions of these data and do not explain the mechanism by which the modulation of firing rates emerges from the anatomical substrate, for this purpose, circuit models are better suited. Two examples are discussed in this work. In chapter 3

we present a framework of multiple interacting hypercolumns (rings) constrained by experimental data in order to describe the net effect of attention on area MT. In chapter 4 we develop a multi-layer circuit model for a cortical column which we think might constitute a fundamental building block for a brain-wide model of attention, and in chapter 5 we discuss briefly how two such columns could be coordinated dynamically.

A fine, but necessary distinction concerns the target property of attentional selection. Attention has been known to operate both with spatial locations, features and even objects as targets [Treue, 2003] and these studies were compatible with the hypothesis that "space" is just another "feature", i. e. that spatial and feature attention might work by the same principles. A systematic comparison of the two, comprising at the same time a wide range of stimuli, is yet missing however. It is understood that such a dataset would also be a novel and most interesting testbed for all models of attention. We have obtained and analyzed such a dataset and will discuss its implications in chapters 2 and 3.

Neuronal responses commonly show a large variance and the trial-to-trial variability between pairs of neurons (the noise correlations) might even be affected by attention. Noisiness and correlated noisiness thus constitute a noteworthy constraint for models. In addition, they make it necessary to critically assess any data analysis method in order to determine in how far trial-averaged results and their derivatives—that are most routinely reported and on which, hence, models are mostly based—might be misleading and representative for the full dataset. Using the above mentioned dataset, that we have acquired, we will present a method to extract features from highly variable data and ways to harness the trial-to-trial variability in order to gain a better understanding of attention in chapters 2 and 3.

Beyond modulations of firing rates, attention was also found to modulate correlations between local and global neuronal assemblies in a frequency dependent way. Specifically, more correlated and more phase-locked activity in the gamma-band (i. e. around 40 Hz) is associated with attention. At the same time, studies report an involvement of lower frequencies (e. g. alpha/beta-band oscillations around 10-20 Hz). Building on the influential communication-through-coherence hypothesis [Fries, 2005], these observations have led to the suggestion that beta and gamma oscillations are, respectively, associated with bottom-up and top-down processes [Wang, 2010; Bastos et al., 2015b]. Concomitantly, the dissociation of beta and gamma frequency band seems to extend to a layer specificity—while superficial layers tend to express oscillations in the gamma band, deep layers seem to prefer slower frequencies [Roopun et al., 2006; Buffalo et al., 2011; Kerkoerle et al., 2014]–and this seems to make sense given that anatomical

studies revealed that feedforward and feedback connections are inclined to originate from superficial and deep layers, respectively. It seems likely, thus, that the layered structure of the cortex is an important substrate for any model of attention [Wang, 2010; Arnal and Giraud, 2012; Siegel et al., 2012; Jensen et al., 2015]. Following up on these ideas, we present a circuit model for a cortical column featuring emerging layer-dependent multi-frequency oscillations in chapter 4 and present preliminary results on how two such columns interact in chapter 5.

In the following sections I will survey these topics that I just briefly touched upon in more detail—namely: attentional modulation of firing rates, phenomenological models of attention, attentional modulation of noise correlations, synchrony, inhomogeneity of attentional effects, the columnar microcircuit, laminar dependence of large-scale cortical connectivity, and models of attention—with a special emphasis on those aspects that will be treated in this work.

1.1 NEURAL CORRELATES OF ATTENTION

It has long been known that attention increases psychophysical performance (e. g. Posner et al. [1980]; Felisberti and Zanker [2005]). But what happens on a neuronal level enabling this? Physiological measurements shedding light on these *neural correlates of attention* [Bushnell et al., 1981; Mountcastle et al., 1981; Moran and Desimone, 1985; Goldberg and Segraves, 1987] will be summarized in the following (for a recent review see, for example, [Noudoost et al., 2010]).

1.1.1 *Modulations of Firing Rate & Phenomenological Models of Attention*

Experiments established that firing rates of cells processing visual stimuli can be modulated by attention, when a location (*spatial attention*, [Moran and Desimone, 1985; Mountcastle et al., 1987; Motter, 1993; Connor et al., 1996, 1997; Luck et al., 1997; McAdams and Maunsell, 2000; Hayden and Gallant, 2009; Reynolds et al., 2000; Martínez-Trujillo and Treue, 2002; Williford and Maunsell, 2006; Ghose and Maunsell, 2008; Ghose, 2009; Khayat et al., 2010]) a property of a stimulus like color or orientation (*feature attention*, [Haenny et al., 1988; Maunsell et al., 1991; Motter, 1994a; McAdams and Maunsell, 1999a; Treue and Trujillo, 1999; McAdams and Maunsell, 2000; Martínez-Trujillo and Treue, 2004; Hayden and Gallant, 2009; Khayat et al., 2010]) or even whole objects [Roelfsema et al., 1998; Blaser et al., 2000; Mitchell et al., 2004; Khayat et al., 2006; Wannig et al., 2007; Fallah et al., 2007] were cued. Effects could occur in various visual areas, down to V1 [Motter, 1993; Roelfsema et al., 1998; Roberts et al., 2007; Chen et al., 2008; Thiele et al., 2009; Buffalo et al., 2010], but tended to be stronger [Maunsell and Cook, 2002; Buffalo et al., 2010] and ear-

lier [Buffalo et al., 2010] in higher visual areas. A functional magnetic resonance imaging (fMRI) study in humans even reported attentional effects on oriented stimuli in the lateral geniculate nucleus of the thalamus (LGN) [Ling et al., 2015]. Furthermore, the strength of attentional modulations seemed to be correlated with the unmodulated response itself [McAdams and Maunsell, 1999a; Treue and Trujillo, 1999; Williford and Maunsell, 2006; Cohen and Maunsell, 2009] and to depend on contrast [Reynolds et al., 2000; Martínez-Trujillo and Treue, 2002; Williford and Maunsell, 2006]. Other factors determining attentional effects on firing rates might include the cell's receptive field (RF)'s eccentricity [Roberts et al., 2007] or task difficulty [Spitzer et al., 1988; Boudreau et al., 2006; Chen et al., 2008], they seem to be cell-type specific [Mitchell et al., 2007], but layer-independent [Buffalo et al., 2011]. Furthermore, attentional modulations of firing rates evolve on a time course on the order of 100 msec [Motter, 1994b; Luck et al., 1997; Busse et al., 2008; Lee and Maunsell, 2010] and can be enhanced through muscarinic, but not nicotinic cholinergic receptors in V1 and V4 [Herrero et al., 2008; Gieselmann and Thiele, 2010; Harris and Thiele, 2011].

To test whether attention also had an influence on the selectivity of a cell, McAdams and Maunsell [1999a] measured the response of neurons in macaque area V1 and V4 to gratings of various orientations during a delayed match-to-sample task, in which the monkey had to attend to either an oriented Gabor or an isoluminant colored patch presented at two spatially separated locations, one of them in the measured cell's RF, the other one outside. Fitting Gaussians to the bell-shaped tuning curves and analyzing the fit parameters the authors observed that, on average, attention caused amplitudes to increase by approximately the same factor irrespective of the grating's orientation, with stronger modulations in V4 compared to V1. Importantly, the selectivity of the cell, as measured by the width of the tuning curve, was, on average, unaffected by attention, as was the undriven response of the neurons. It should be noted though that others did report changes of selectivity [Spitzer et al., 1988], as well as attentional modulations of both spontaneous firing rates [Colby et al., 1996; Luck et al., 1997; Chelazzi et al., 1998] and BOLD responses [Kastner et al., 1999; Chawla et al., 1999; Shulman et al., 1999; Ress et al., 2000]

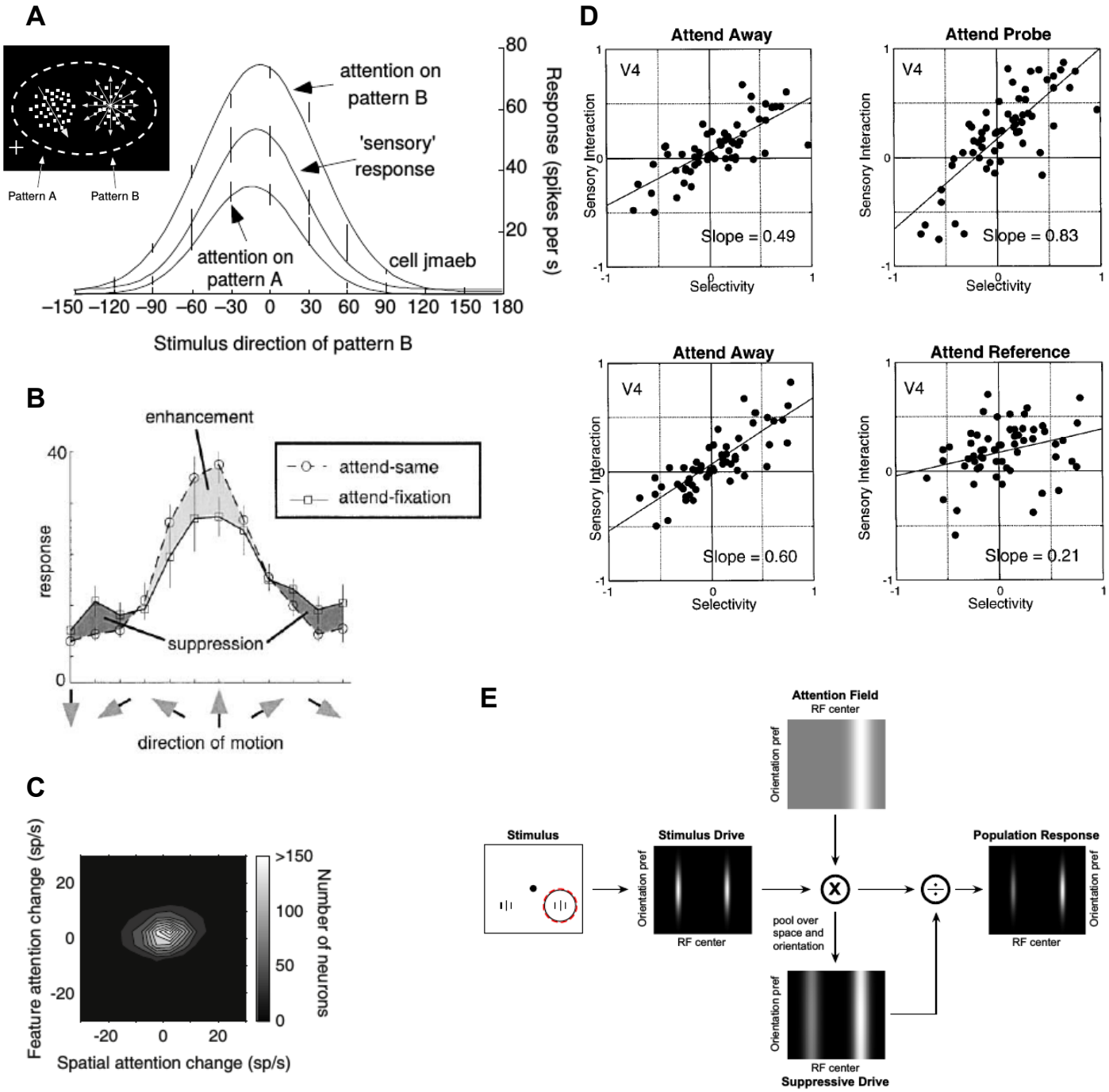
As the stimuli in the experiment by McAdams and Maunsell [1999a] differed in both location and feature (color or orientation) attention could have been allocated on the basis of both of these properties. In a related study Treue and Maunsell [1999] disentangled the two using random dot stimuli and measuring responses in macaque area MT. They found a multiplicative modulation of the tuning curve when switching attention between the stimulus inside and outside a cell's RF (spatial attention), a pure feature-based attentional modulation of the same order of magnitude when attention was switched between

a preferred and a non-preferred stimulus outside a cell's RF, and that combining spatial and feature attention yielded a modulation strength that was approximately the sum of the two individual modulation strengths (which was also reported by Patzwahl and Treue [2009] and for V4 neurons by Hayden and Gallant [2009]; more generally, Cohen and Maunsell [2011] observed rate changes for spatial attention that were completely independent of those for feature attention in the same V4 neuron, but the deviations could have been due to the fact that stimulus placement and selection was optimized for only few of their recorded neurons which might have caused differences over the population as a whole). Finally, they observed, using two spatially separated stimuli within one RF, that the multiplicative scaling factor describing the effect of attention on the tuning curve can also be smaller than one when attention is directed to an anti-preferred stimulus (see figure 1.1A). In conclusion of their results they coined the feature similarity gain model (FSGM) hypothesizing that the attentional gain factor should depend on the difference between the attended and the preferred feature of a cell—which was explicitly found in later experiments for direction sensitive cells in MT [Martínez-Trujillo and Treue, 2004] (see figure 1.1B) as well as for orientation and spatial frequency sensitivity in V4 [Cohen and Maunsell, 2011]. Given evidence from experiments in which attention was directed towards rotating surfaces [Wannig et al., 2007], a generalized FSGM was suggested [Treue and Katzner, 2007] allowing the attended feature to vary between neurons. Moreover, given that both forms of attention seem to act independently [Martínez-Trujillo and Treue, 2004; Hayden and Gallant, 2009; Cohen and Maunsell, 2011] (compare figure 1.1C) it was suggested [Martínez-Trujillo and Treue, 2004; Boynton, 2005] to consider spatial location as a feature on equal footing with others like color or orientation and that similar mechanisms are employed to mediate attentional effect for space and other features. However, the result that feature, but not spatial attention, can alter the tuning in V4 [David et al., 2008] seems in contrast to this view.

As natural scenes are cluttered with objects moving next to and even over or through each other, attention can be expected to play an important role in precisely these situations. When two stimuli are presented within the same RF (with attention directed elsewhere) neurons' responses are typically intermediate between their responses to either of the two stimuli alone [Qian and Andersen, 1994; Recanzone et al., 1997; Chelazzi et al., 1998; Britten and Heuer, 1999; Reynolds et al., 1999; Zoccolan et al., 2005; Armstrong et al., 2006]. The combined response has sometimes [Britten and Heuer, 1999; Ghose and Maunsell, 2008; Lee and Maunsell, 2010] been described in a *normalization model* framework [Carandini and Heeger, 2012], which can, but need not agree with an intermediate response to a pair of stimuli. In

normalization models, attention is typically introduced as an additional parameter, acting, for example, on the gain [Ghose and Maunsell, 2008] or the normalization [Lee and Maunsell, 2009]. Reynolds and Heeger [2009] suggested a very general equation, the *Normalization Model of Attention* (compare figure 1.1E), with which to predict the responses of a field of neurons parameterized by RF center and orientation preference. For each parameter, the response depended on an equally parameterized field of inputs corresponding to the presented stimulus, modulated by an attention field representing top-down input, and normalized by a suppression field to describe stimulus competition. While indeed able to reproduce attentional rate effects the model does not have much predictive power due to the high number of free parameters, that is it might overfit the data. Moreover, the biological basis of the attention field is not explained by the model.

Figure 1.1 *(facing page)*: Overview of attentional firing rate modulations and phenomenological models of attention. A) Varying direction of RDP B while RDP A is moving in the anti-preferred direction—both within the RF of an MT cell—results in a tuning curve which is differentially modulated depending on the attentional target (adapted by permission from Macmillan Publishers Ltd: Nature [Treue and Trujillo, 1999], copyright 1999) **B)** The attentional modulation (attend-same) compared to a reference condition (attend-fixation) of the population response of direction-tuned MT cells (preferred direction on the x-axis) depends on the angular distance between preferred and attended (here: upwards) direction (figure reprinted from [Martínez-Trujillo and Treue, 2004], copyright 2004, with permission from Elsevier). **C)** Spatial (quantified as the difference in firing rates when attention was switched between stimuli in opposite hemispheres) and feature (response modulation when either an orientation or a spatial frequency change needed to be detected) attentional modulation were independent in V4 neurons (fig. reprinted from [Cohen and Maunsell, 2011], copyright 2011, with permission from Elsevier). **D)** Selectivity denotes change in normalized firing rates when two stimuli (termed probe and reference) are presented independently at two separate positions within a V4 cell's RF, sensory integration is the response difference of the simultaneously presented pair to the reference alone. Upper (lower) row contain only stimulus pairs that were significantly modulated my when attention was directed to the probe (reference) stimulus. When attention was directed outside the RF (left column) paired responses were about half-way between single-stimulus responses, and attention (right column) to one of the stimuli made responses on average more similar to the corresponding single-stimulus response (fig. reproduced with permission of Society for Neuroscience from [Reynolds et al., 1999]). **E)** The *Normalization model of attention* predicts the response R of a cell parameterized by $\vec{x}$ (RF position, direction preference, ...) to be $R(\vec{x}) = |A(\vec{x})E(\vec{x})/(\sigma + s(\vec{x}) \star A(\vec{x})E(\vec{x}))|_T$ where E, A, σ, s and T denote, respectively, stimulus drive, attentional field, a non-negative parameter, suppressive field and rectification threshold (figure reprinted from [Reynolds and Heeger, 2009], copyright 2009, with permission from Elsevier).


A
Pattern A
Pattern B
attention on pattern B
'sensory' response
attention on pattern A
cell jmaeb
Response (spikes per s)
80
60
40
20
−150 −120 −90 −60 −30 0 30 60 90 120 150 180
Stimulus direction of pattern B
B
enhancement
attend-same
attend-fixation
response
40
0
suppression
direction of motion
C
Feature attention change (sp/s)
20
0
-20
-20 0 20
Spatial attention change (sp/s)
>150
100
50
0
Number of neurons
D
Attend Away
V4
Sensory Interaction
Selectivity
Slope = 0.49
Attend Probe
Slope = 0.83
Attend Away
Slope = 0.60
Attend Reference
Slope = 0.21
E
Attention Field
RF center
Orientation pref
Stimulus
Stimulus Drive
pool over space and orientation
Population Response
Suppressive Drive

Yet another alternative account of attentional effects emphasizes the mutual influence of distinct stimuli on each other. According to the biased competition (BC) model [Desimone and Duncan, 1995; Reynolds et al., 1999] (compare figure 1.1D) when two stimuli are presented within the same RF (with attention directed elsewhere) they "compete" for representation in the neurons' responses which are predicted to be intermediate between their responses to either of the two stimuli alone, in agreement with many experimental studies [Qian and Andersen, 1994; Recanzone et al., 1997; Chelazzi et al., 1998; Britten and Heuer, 1999; Reynolds et al., 1999; Zoccolan et al., 2005; Armstrong et al., 2006]. Attention then biases this intermediate response towards the response that would be elicited if the attended stimulus were presented alone in the absence of attention. [Moran and Desimone, 1985; Chelazzi et al., 1998; Reynolds et al., 1999]. Importantly, though, sensory (attention directed elsewhere) responses to combined stimuli were also observed to be less than either of the single-stimulus responses alone [Snowden et al., 1991]. In fact, a stimulus' contribution to the overall response depended on its exact location within the RF [Snowden et al., 1991; Jancke et al., 1999; Majaj et al., 2007] hinting at a highly nonlinear summation of RF subparts even in the absence of attention. Likewise, attentional modulation of spontaneous firing rates were reported to depend on the subpart of the RF that is behaviorally relevant [Luck et al., 1997], and generally a difference can also be expected if stimulus pairs within one RF are presented spatially separated or transparently overlaid, although the experiments by Patzwahl and Treue [2009] suggest that spatial and feature attention effects which are both at work in such a setting are correlated in case the stimulus pair consists of a preferred and an anti-preferred stimulus.

In summary, the BC model emphasizes the role of space for selection of relevant stimuli and is often paraphrased as a selective RF shrinkage around the attended stimulus [Moran and Desimone, 1985], but such an account cannot explain attentional effects on transparently presented stimuli [Patzwahl and Treue, 2009]. Nevertheless, it is conceivable that not only spatial but also more general features can "compete" for the dominant influence on a neuron such that the pattern of firing rate modulations predicted by the BC model for the case when two stimuli are combined with and without attention also hold for transparent stimuli. Furthermore, the BC model has been tested so far only with a limited number of stimulus pairs; the FSGM was found to hold for direction, orientation and spatial frequency tuning, but it is unclear if it also applies to stimulus pairs within the RF as has been predicted computationally for the case of opposite motion directions [Ardid et al., 2007]. In generalization of the experiments by Patzwahl and Treue [2009] it is also currently unclear if and how whole tuning curves in response to either spatially separated or transparently

overlaid stimulus pairs within one RF are differentially modulated by attention, if the BC or the FSGM could explain that or if conceptually new approaches are necessary. Given nonlinearities in the RF summation and the attentional effects, no clear predictions can be made, and individual cells might well behave differently in these respects. Therefore, specific care will have to be taken when analyzing data from corresponding experiments, and also models need to be flexible enough to accommodate quite heterogeneous effects. All these questions will be addressed in chapters 2 and 3.

1.1.2 *Response Variability*

Firing rates of single neurons are only one aspect of brain dynamics. In the same way as neuron's selectivities (circular variances) can vary over wide ranges [Ringach et al., 2002; Gur et al., 2005; Niell and Stryker, 2008] attentional modulations are highly variable between neurons, often spanning a wide range from positive to negative modulations [Moran and Desimone, 1985; Treue and Maunsell, 1996; Reynolds et al., 1999; McAdams and Maunsell, 1999a; Treue and Trujillo, 1999; Maunsell and Cook, 2002], often even being absent altogether: Indeed, Roelfsema et al. [2004] and Poort and Roelfsema [2009] report that 40 % of their studied cells were not significantly affected at all by object-based attention in V1. This heterogeneity in individual responses might even be beneficial for the information encodable in the population activity [Chelaru and Dragoi, 2008; Padmanabhan and Urban, 2010]. Beyond that, higher order temporal statistics, like the variance of the spike discharge rate, and correlations between neurons have been neglected so far in this review, and will be discussed in the following.

Neuronal spiking is typically variable over trials, with spike count variances roughly proportional to the mean spike counts, and the thus approximately constant ratio between the two (called the Fano factor) having values around 1-2 [Tolhurst et al., 1983; Shadlen and Newsome, 1998; McAdams and Maunsell, 1999b; Churchland et al., 2010] which is commonly reduced by stimulus onset [Churchland et al., 2010]. It has been suggested [Osborne et al., 2008] that this time-dependent variability in spiking is not independent between neurons (but see [Rolls and Treves, 2011]) so as to increase the information content in the combined spike trains of these neurons. Several studies observed that attention, too, significantly reduced the Fano factor in area MT [Niebergall et al., 2011; Zénon and Krauzlis, 2012], MST [Zénon and Krauzlis, 2012], V4 [Mitchell et al., 2007, 2009; Cohen and Maunsell, 2009] and V1 [Herrero et al., 2013], but a counterexample was also reported [McAdams and Maunsell, 1999b]. The attentional reduction of the Fano factor might be cell-type specific: Mitchell et al. [2007] observed that action potential widths showed a bimodal

distribution and identified narrow- and broad-spiking cells as presumably being inhibitory and excitatory neurons, respectively. Firing rates, while higher for narrow-spiking cells, were modulated by attention for both cell classes, but Fano factors were significantly reduced only for the narrowly-spiking putative inhibitory cells. At the receptor level, pharmacological blockade of NMDA- or AMPA/kainate-receptors by application of antagonists APV and CNQX, respectively, did not significantly alter attentional rate modulation in macaque V1, but the attentional reduction of Fano factors was abolished by both APV and CNQX [Herrero et al., 2013]. The same study also claims, without proof, that the cholinergic system does not affect Fano factors.

Rate covariations (also called *spike count* or *noise correlations*) measure the extent to which the observed spike counts of two neurons co-vary across trials and have typical values of 0.1-0.5 when observed [Zohary et al., 1994; Bair et al., 2001; Kohn and Smith, 2005; Cohen and Newsome, 2008; Mitchell et al., 2009; Poort and Roelfsema, 2009; Cohen and Maunsell, 2009; Cohen and Kohn, 2011; Ruff and Cohen, 2014a,b], but extremely low values, often statistically indistinguishable from zero, even between neighboring cells recorded from the same tetrode, have also been reported [Ecker et al., 2010]. In macaque V1, noise correlations were found to be higher in superficial and deep layers, and to almost vanish in between [Smith et al., 2013]; they were observed to vary with the task in macaque area MT [Cohen and Newsome, 2008], and were observed to decrease with firing rate [Cohen and Maunsell, 2009], task difficulty [Ruff and Cohen, 2014b] and attention in macaque area MT and MST [Zénon and Krauzlis, 2012], V4 [Mitchell et al., 2009; Cohen and Maunsell, 2009; Ruff and Cohen, 2014a] and V1 [Herrero et al., 2013], but the effect of attention seems to depend on the strength of rate modulations [Cohen and Maunsell, 2011], if these are significantly modulated by attention [Roelfsema et al., 2004] and on how similar the preferences of the cells are [Ruff and Cohen, 2014a]: for negative firing rate modulations or for very dissimilar cells attention could also increase noise correlations. Moreover, Poort and Roelfsema [2009] report the absence of attentional modulations of noise correlations for pairs of V1 neurons whose RFs fell either both on a target or both on a distractor curve during a curve-tracing task.

NMDA-receptor antagonist APV was reported to abolish attentional modulation of noise correlations, whereas AMPA receptor antagonist CNQX increased them in both attention conditions without abolishing the attentional modulation, and the cholinergic system was claimed to have no effect [Herrero et al., 2013]. Finally, attentional effects on noise correlations seem to be layer-independent [Buffalo et al., 2011].

The effect of noise correlations on down-stream processing critically depends on the coding scheme [Oram et al., 1998; Abbott and Dayan, 1999; Zhang and Sejnowski, 1999; Pouget et al., 1999; Averbeck et al., 2006; Ecker et al., 2011] such that their potential functional role remains speculative.

In a series of studies Cohen and coworkers also analyzed firing rate correlations between trials for more than two neurons. For each recorded trial, they treated the simultaneously recorded spike counts of n neurons as a point in n-dimensional space, and then compared the clouds consisting of all n-dimensional points, one for each trial, between two conditions by projecting all points onto the axis between the cloud centers and computing d' [Stanislaw and Todorov, 1999], as a measure for the separation of the two resulting distributions. Cohen and Maunsell [2009] observed that d' for the responses before and after a change of the stimulus orientation (which the monkey had to detect in order to receive a reward) was increased by attention and correlated with the behavioral improvement of orientation changes that the monkey could detect (shift to the left of the psychometric function: performance vs orientation change) and that were caused by attention. In simulations they tested the influence of firing rate increases, Fano factor decreases and noise correlation decreases on d' and thereby attributed the observed behavioral improvement predominantly to the reductions in noise correlations. In a similar vein they found that the similarity between the population response in a given trial and the mean population response in two spatial attention conditions (attend-left and attend-right) was an indicator of a monkey's performance to detect an orientation change [Cohen and Maunsell, 2010]. Cohen and Maunsell [2011] studied the effect of spatial and feature attention onto population responses in V4 employing a task in which the monkey had to detect a cued feature change (orientation or spatial frequency) at a cued location (left or right). First, they established that neurons, generally, are modulated by both forms of attention, and that noise correlations between pairs of neurons depend on the average rate modulation in the pair in essentially the same way for spatial and feature attention. Then, to analyze population responses, these were projected onto two axes: first, for trials with a given feature change, the average responses of trials in which the monkey correctly detected a stimulus change on the left and on the right were computed and the axis through these two average responses was termed the "spatial attention axis"; the reason is that if the population response in a given trial is projected onto this axis, its position on the axis quantifies its similarity with the expected response when attention was directed either to the left or to the right, and can, therefore, be used to decode the locus of attention. This measure correlated with behavioral performance as projections onto the axis for missed trials fell, on average, in between the two

points spanning the axis. In a completely analogous way they also defined a "feature attention axis" between the average responses to correctly detected orientation and spatial frequency changes (given a location change). Neurons in their experiments were recorded simultaneously from both hemispheres and they were interested if there was any difference between the employment of attention between the two hemispheres. To that end, they defined the two axes separately for populations from the two hemispheres and analyzed correlations in the projections onto these axes (from neurons that were simultaneously recorded in both hemispheres). They observed correlations between projections onto the two hemisphere's feature attention axes, but not the spatial attention axes. On the other hand, when they divided neurons from within a hemisphere in two groups and analyzed projections they found correlations between the projections for both the spatial and feature attention axis. Finally, projections onto the spatial and feature attention axes were not correlated, independent on if they were based on neurons from just one or from both hemispheres. It appeared, thus, that neurons can be modulated by both spatial and feature attention, but they become coupled (that is correlated) only locally for attentional selection related to location, but over a wider cortical range when the target property is a feature.

In summary, the studies presented in this section highlight how variable individual responses can be. It was suggested that a large part of these fluctuation can be attributed to "ongoing activity" [Arieli et al., 1996; Tsodyks et al., 1999; Kenet et al., 2003; Fiser et al., 2004]. Correspondingly, the momentary "brain state" [Gilbert and Sigman, 2007] is an important, possibly continuous [Harris and Thiele, 2011] variable that is, as seen above, influenced by attention. Moreover, populations of neuron can reveal information beyond that present in single neurons suggesting possible differences between spatial and feature attention in the coordination of widely distributed neurons and, thus, calling for an approach to an understanding of attention that takes these distributed networks into account. Finally, while it has been shown that attention leads to increased discriminability of opposite features, it is currently unclear how that discriminability is interpolated for a continuous feature category. These issues will be revisited in chapters 2, 4 and 5.

1.1.3 *Role of Oscillations*

The tendency to fire synchronously [Eckhorn et al., 1988; Singer, 1993; Vaadia et al., 1995; deCharms and Merzenich, 1996] constitutes another aspect of neuronal dynamics which might occur uncorrelated with firing rates [Vaadia et al., 1995; deCharms and Merzenich, 1996; Tiesinga and Sejnowski, 2004; Buehlmann and Deco, 2008] Synchronization can occur between [Gregoriou et al., 2009] spike timings of

neurons, between local field potentials (LFPs) of neuronal groups, and also between spike timings and LFPs—a proxy for population activity [Yu and Ferster, 2010], is often, but not always accompanied by oscillations [König et al., 1995] and the rhythms were observed in multiple frequency bands [Buzsáki and Draguhn, 2004; Roopun, 2008b]. The presence of synchronous or oscillatory activity can be determined using auto- and cross-correlograms (e. g. Gray et al. [1989], but note that peaks in the cross-correlograms not necessarily imply synchrony [Brody, 1998, 1999; Shadlen and Movshon, 1999]) or the coherence measure (e. g. Fries et al. [2008])[1]. Coherent activity might occur in short episodes ($<$ 100 ms) intermittent with asynchronous activity [Burns et al., 2011; Xing et al., 2012a] and with variable peak frequency [Burns et al., 2011; Xing et al., 2012a; Roberts et al., 2013], and this might be a reason why it is relatively weak, if present, with coherence values of 0.1-0.3 [Wang, 2010]. Synchrony and oscillations were observed in a wide range of situations, for instance during stimulus presentation [Gray and Prisco, 1997], working memory tasks [Lee et al., 2005; Salazar et al., 2012], memory recall [Tort et al., 2009], sleep [Crunelli and Hughes, 2010], motor preparation [Sanes and Donoghue, 1993].

Beyond that, Fries et al. [2001] discovered that the LFP coherence spectrum in V4 was modulated when the corresponding neuronal group was processing an attended stimulus compared to an unattended one (compare figure 1.2A). Several other studies later confirmed that LFP and spike field coherence (SFC) in the gamma-frequency range are increased by attention in V4 [Fries et al., 2001, 2008; Taylor et al., 2005; Bichot et al., 2005; Womelsdorf et al., 2006; Gregoriou et al., 2009; Chalk et al., 2010; Buffalo et al., 2011; Bosman et al., 2012;

1 The *coherency* C between two signals X and Y is defined as their normalized cross-spectral density S_{XY},

$$C_{XY}(f) = \frac{S_{XY}(f)}{\sqrt{S_{XX}(f)S_{YY}(f)}},$$

where the cross-spectral density is the Fourier transform of the cross-correlation function of X and Y,

$$S_{XY}(f) = \mathcal{F}\left[\int_{-\infty}^{\infty} x^*(\tau)y(t+\tau)d\tau\right](f).$$

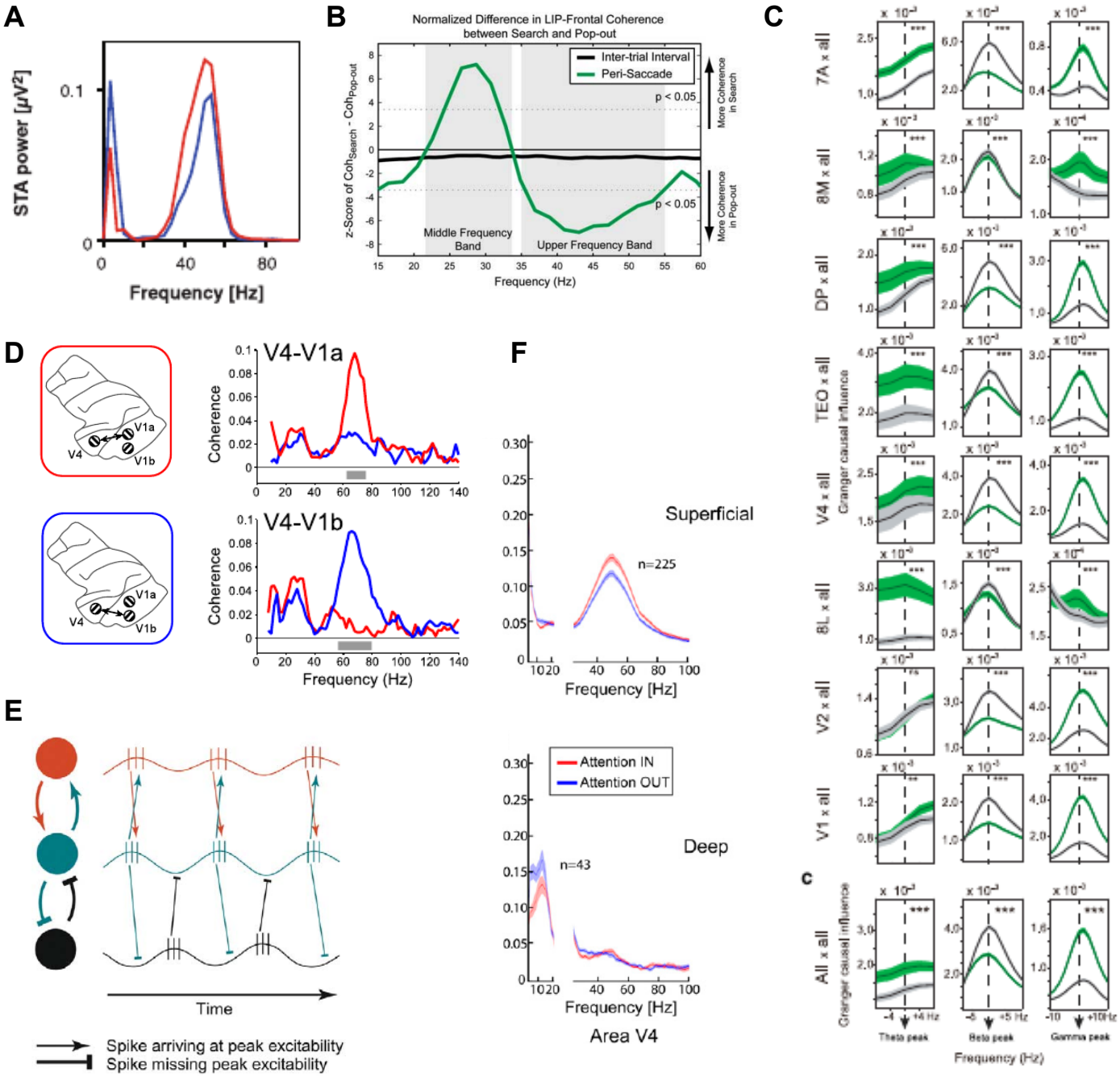
and the absolute value of $C_{XY}(f)$ is called *coherence* taking on values between 0 and 1 [Bendat and Piersol, 2011; Challis and Kitney, 1991; Jarvis and Mitra, 2001; Fries et al., 2008]. If multiple samples are available for the signals X and Y then $S_{XY}(f)$ has to be averaged over the available samples before calculating the coherency. In that case $|C_{XY}(f)|$ depends on how consistent the phases of each individual cross-spectral density are between samples. Nevertheless it also always depends on the power and therefore coherence is no unambiguous measure for fixed phase relationships between X and Y across samples, which complicates its interpretation. It is possible that the coherence exhibits a peak in the absence of oscillations. These, and a number of other complications concerning the detection of synchronization and rhythms are neatly discussed by Wang [2010].

Grothe et al., 2012]. Concomitantly, some studies found a reduction of alpha-band[2] coherence due to attention in various visual cortical areas [Fries et al., 2001; Buffalo et al., 2011; Bollimunta et al., 2011]. Notably, Roelfsema et al. [2004] observed only a general increase of the multi-unit activity (MUA) power spectrum in V1 for the attended compared to the unattended stimulus that was not frequency specific, and Chalk et al. [2010] found that attention increased LFP gamma power and gamma-range SFC in V4, but *decreased* it in V1. Furthermore, modulations of synchrony might be cell-type specific with narrow-spiking, putative interneurons being more strongly gamma-synchronized than broad-spiking, putative excitatory neurons, which also lag behind the former, and with attention enhancing and reducing, respectively, the gamma-synchronization of putative inhibitory and excitatory neurons [Vinck et al., 2013]. Finally, modulations of synchrony have also been directly linked to behavior by Womelsdorf et al. [2006] who observed that the degree of gamma synchronization could predict monkeys' reaction times.

Synchrony was also observed between cortical areas. Simultaneous recordings in V4 and frontal eye field (FEF) [Gregoriou et al., 2009] as well as in MT and LIP [Saalmann et al., 2007] demonstrated that the SFC between these areas increased as a result of attention. Moreover, comparing LFP coherence between FEF and LIP during a search

2 In this work we will treat the alpha- and beta-frequency band as one, as both have been implicated in similar top-down processes, see [Bressler and Richter, 2015; Zheng and Colgin, 2015].

Figure 1.2 ***(facing page)*****: Overview of the role of oscillations in attention. A)** Attending to a stimulus within (red) compared to outside (blue) a V4 neuron's RF increased (decreased) power of the STA in the gamma (alpha) band (fig. from [Fries et al., 2001], reprinted with permission from AAAS). **B)** Coherence between LIP and frontal cortex was higher in a middle and high frequency band during a search—involving top-down attention—compared to a pop-out—involving bottom-up attention—task (fig. from [Buschman and Miller, 2007], reprinted with permission from AAAS). **C)** Frequency resolved Granger causality is shown for theta, beta and gamma band (columns). Green (black) spectra are averaged over all lower (higher) hierarchical areas (according to Markov et al. [2014b]'s hierarchy), given the target area indicated on the y-axis. (figure reprinted from [Bastos et al., 2015b], copyright 2015, with permission from Elsevier). **D)** Two stimuli falling within the RFs of different V1, but the same V4 cell were presented simultaneously. When either V1a (red, upper row) or V1b (blue, lower row) was attended, coherence between V4 and the attended, but not the unattended, V1 cell was increased in the gamma band (figure reprinted from [Bosman et al., 2012], copyright 2012, with permission from Elsevier). **E)** The CTC hypothesis posits that incoming spike efficacy is enhanced during more excitable periods (figure reprinted from [Fries, 2005], copyright 2005, with permission from Elsevier). F) Attention enhances SFC in the gamma- (alpha-) band only in upper (lower) layers (fig. rep. f. [Buffalo et al., 2011]).


A
STA power [μV²]
0.1
0
0
40
80
Frequency [Hz]
B
Normalized Difference in LIP-Frontal Coherence
between Search and Pop-out
Inter-trial Interval
Peri-Saccade
p < 0.05
p < 0.05
More Coherence in Search
More Coherence in Pop-out
z-Score of Coh_Search - Coh_Pop-out
Middle Frequency Band
Upper Frequency Band
Frequency (Hz)
C
7A x all
8M x all
DP x all
TEO x all
V4 x all
8L x all
V2 x all
V1 x all
All x all
Granger causal influence
Theta peak
Beta peak
Gamma peak
Frequency (Hz)
D
V1a
V4
V1b
V4-V1a
V4-V1b
Coherence
Frequency (Hz)
E
Time
Spike arriving at peak excitability
Spike missing peak excitability
F
Superficial
n=225
Deep
n=43
Attention IN
Attention OUT
Frequency [Hz]
Area V4

task requiring top-down attention with a pop-out condition involving bottom-up attention revealed that the former was stronger in a frequency band around 25-30 Hz whereas the latter dominated for frequencies around 35-55 Hz [Buschman and Miller, 2007] (compare figure 1.2B). Moreover, Saalmann et al. [2012] demonstrated that attention increased LFP-LFP- and spike-LFP-coherence in the alpha-band between the pulvinar on the one and and cortical areas V4 and TEO on the other hand, and suggested, based on a Granger causality analysis, that the pulvinar might be a regulator for synchrony between cortical areas.

Beyond that, several studies have now also demonstrated large-scale coherence networks in multiple frequency bands under various circumstances, arguing for their relevance for large-scale cortical processing: Hipp et al. [2011] found them during rest, Hipp et al. [2012] demonstrated that they were associated with perception, Salazar et al. [2012] that they were predictive for working memory, Dotson et al. [2014] showed that relative phase-relationships at the recording sites were task-dependent—we will discuss in section 1.3.2 that such phase-locking patterns might be associated with different states of information processing—, and Bastos et al. [2015b] suggested that theta-/gamma- and beta-rhythms might subserve feedforward and feedback processing, respectively, thereby multiplexing interareal coordination through frequency bands (compare figure 1.2C). More generally Siegel et al. [2012] have suggested that specific frequency bands could identify canonical cortical oscillations independent of their occurrence sites. Yet, how and why specifically the observed networks occur in the circuits of the brain and if they are determined by anatomy or are, on the other hand, dynamical in nature, remains to be further elucidated, and we will come back to these questions later in this work.

Importantly, attention not only modulated coherences, it did so selectively when multiple stimuli had to be processed [Bosman et al., 2012; Grothe et al., 2012] During simultaneous presentation of two stimuli that were spatially separated such that they would fall into the non-overlapping RFs of two cells in V1, both cells increased their gamma-band LFP power, but the gamma band coherence and SFC between V1 and V4 was significantly stronger between the V4 cell and only the relevant V1 cell containing the attended stimulus in its RF, not, however, with the irrelevant V1 cell. Beyond that, when the luminance of the two stimuli presented in the same V4 RF were independently and stochastically modulated in time, the spectral coherence between the recorded neuron's LFP and the attended, but not unattended, luminance-time course (which was task-irrelevant) were elevated for low frequencies (<20 Hz). [Grothe et al., 2015] (compare figure 1.2D). Together, these studies suggest that oscillations might

play a role in interareal coordination and that attention might selectively open pathways for preferential routing of information.

A framework in which to understand this potential functional role of oscillations is provided by the CTC hypothesis [Fries, 2005] (compare figure 1.2E). Based on findings that spike generation in the postsynaptic neuron is particularly effective during synchronous input [Azouz and Gray, 2000, 2003] and that postsynaptic neurons are also more responsive to (particularly synchronous) input during excited phases of subthreshold-membrane oscillations [Volgushev et al., 1998] the influential CTC hypothesis posits that coherent fluctuations in the excitability of two neuron groups create "windows of opportunity" at the common peak of the oscillation in which spike transmission between the groups is particularly effective. In that scheme, bi-directional communication would require zero-time-lag synchronization as was observed in some experiments [Gray et al., 1989; Roelfsema et al., 1997]. Other studies report, however, non-zero phase differences between distant areas oscillating at similar gamma frequencies [Grothe et al., 2012; Jia et al., 2013; Gregoriou et al., 2009] and, consequently, it has been proposed [Bastos et al., 2015a] that, as feedforward and feedback pathways have a different laminar distribution [Markov et al., 2014b], and oscillations might acquire systematic phase lags while transmitted across layers [Livingstone, 1996; Kerkoerle et al., 2014], taking into account the layered structure of the cortex can reconcile bidirectional communication with the CTC-hypothesis. Another, complementary, idea [Bastos et al., 2015a] is to postulate different dominant frequencies for feedforward and feedback processes, for which there is now also direct evidence (based on a Granger causality analysis) [Bastos et al., 2015b]. The two frequency channels could interact through some form of cross-frequency coupling (CFC) [Jensen and Colgin, 2007; Roopun, 2008b; Canolty and Knight, 2010]—which might even be layer-dependent [Spaak et al., 2012]—or, more generally, in the form of a complex column-wide oscillation spanning multiple frequencies, as we will point out later.

Experimental evidence exists both in favor and against CTC. Womelsdorf et al. [2007] found that the Spearman-rank-correlation between 60 Hz-power of the MUA recorded with two electrodes tended to be high, whenever the phase relation between the two signals was near their mean value, and a relative phase close to the mean preceded strong power-correlations by a few milliseconds, suggesting a "good" (mean) phase relationship as a mechanistic cause for effective interactions (for which the power correlation is a presumed proxy). Moreover, it was demonstrated that the timing of an incoming sensory response relative to an ongoing gamma-oscillation influences neuronal [Cardin et al., 2009] and behavioral [Siegle et al., 2014] responses and that the phase of spikes relative to ongoing rhythms can carry information [O'Keefe and Recce, 1993; Lisman, 2005; Montemurro

et al., 2008; Vinck et al., 2010]. Furthermore, models have demonstrated [Cannon et al., 2014] that two pyramidal interneuron network gamma (PING) rhythms with slightly different frequencies might self-organize, when their excitatory populations are coupled, such that the faster one produces spikes driving the slower, but not vice versa as the spikes of the slower PING rhythm arrive during the faster rhythm while its inhibition is still strong. Finally, increased coherence in the gamma-range during attention [Fries et al., 2001; Gregoriou et al., 2009] that is selectively enhanced between areas only for an attended stimulus [Bosman et al., 2012; Grothe et al., 2012] is compatible with the CTC.

On the other hand, Jia et al. [2013] find, in anesthetized macaques, that spikes in V2 that immediately follow spikes in V1 do not occur at a phase of the local V2 rhythm that is most likely to elicit spikes; Chalk et al. [2010] report a reduction of spike-field coherence in V1 for attended stimuli arguing against a general role of increased coherence for enhanced processing; Burns et al. [2011] and Xing et al. [2012a] show that gamma-oscillation occur in brief bursts of varying length and frequency; Ray and Maunsell [2010]; Roberts et al. [2013] demonstrate that gamma peak frequency is contrast dependent; Maris et al. [2013] find phase-relationships in pairs of simultaneously recorded LFP and MUA are spatially diverse (but consistent over time); similarly Dotson et al. [2014] find task-dependent in- or anti-phase relationships between LFPs in frontal and parietal areas of monkeys that were recorded simultaneously from electrodes of varying spatial separation. The task-dependent bimodal phase distribution over a range of spatial separations is not consistent with the CTC. Instead they suggest the observed patterns of oscillations emerge as a result of metastability [Bressler and Kelso, 2001; Tognoli and Kelso, 2014]. Note however, that Palmigiano et al. [2015] demonstrated computationally that short transient gamma-bursts of varying frequency and varying phase relationships might nevertheless be compatible with the CTC-hypothesis and could be harnessed for flexible information routing.

Finally, like noise correlations [Smith et al., 2013], oscillations are also layer dependent. For example, in slices, pharmacologically evoked oscillations in the alpha/beta-range, generated in layer 5 [Silva et al., 1991; Flint and Connors, 1996; Roopun et al., 2006] as well as both very slow (1-5 Hz) [Flint and Connors, 1996] and gamma range [Roopun et al., 2006] oscillations generated in superficial layers were observed. Xing et al. [2012b] reported elevated levels of LFP gamma power in superficial layers of anesthetized macaques, in awake behaving monkeys Smith et al. [2013] and Kerkoerle et al. [2014] observed a peak in the alpha-/beta-range LFP power in deep layers and a concomitant peak in the gamma-range in the superficial layers. Buffalo et al. [2011] found a similar laminar-dependent frequency pattern for the

SFC in macaque V1, V2 and V4, although the slow and fast peak coherences were of comparable strength. Strikingly Buffalo et al. [2011] also reported that attention might act on oscillatory synchrony in a corresponding laminar-dependent way (compare figure 1.2F): in V2 and V4 gamma-SFC in the superficial layers was increased, whereas alpha-/beta-SFC in deep layers was decreased for an attended stimulus. Moreover, Bastos et al. [2015b] suggested—in line with previous studies [von Stein et al., 2000; Buschman and Miller, 2007; Arnal et al., 2011]—, based on a Granger causal analysis of ECoG data from a wide range of simultaneously recorded areas, that feedforward and feedback information processing might be subserved by theta-/gamma- and alpha-/beta-frequency bands, respectively. Furthermore, as will be discussed in more detail in section 1.2.2 feedforward and feedback connections in the cortex form largely non-overlapping, layer-dependent counterstreams [Markov et al., 2014b]. Altogether, these studies could thus indicate that feedforward and feedback communication is multiplexed in lamina- and frequency-dependent pathways, which are integrated at the column-level [Wang, 2010; Arnal and Giraud, 2012; Siegel et al., 2012; Jensen et al., 2015].

In summary, oscillations are ubiquitous in the cortex, can be coherent in and between areas (to a degree at least), are laminar dependent, are modulated by attention, also in a laminar dependent way and are hypothesized to subserve interareal coordination. Phase lags between oscillations have been suggested to be important for the communication between circuits, which will be further elucidated in section 1.3.2. In this work, we will discuss how multi-frequency multi-layer oscillations might emerge in a self-organized manner in a local circuit—if only its connectivity is sufficiently constrained in chapter 4, and start to analyze phase-relationships between two interacting local circuits in chapter 5.

1.1.4 *The Network of Areas Involved in Attention*

Attention can not only affect the processing of visual, but also for example auditory or tactile stimuli, can influence behavior, and we have already discussed that it modulates coherence in areas in the frontal and parietal lobe. These facts suggest that its computation does not originate in sensory visual cortical areas on which we have focused so far. Instead, we will see now that a distributed set of areas seems to be involved.

Based on fMRI studies two distinct networks have been distinguished (see Corbetta et al. [2008] for a review). The *dorsal frontoparietal network*, involving the intraparietal sulcus, superior parietal lobule and dorsal frontal cortex including the FEF, was found to be activated during the presentation of target stimuli [Shulman et al., 1999, 2003; Corbetta et al., 2000; Hampshire et al., 2007], but also, notably, dur-

ing attentional orienting (cue and/or delay period of an attention experiment) [Shulman et al., 1999; Kastner et al., 1999; Corbetta et al., 2000; Hopfinger et al., 2000], suggesting it as a source for top-down signals. The *ventral frontoparietal network*, on the other hand, is right-lateralized including the temporoparietal junction and the ventral frontal cortex, and shows elevated blood oxygen level-dependent (BOLD) signals during the detection of or reorienting towards relevant (i. e. not necessarily salient) stimuli [Arrington et al., 2000; Corbetta et al., 2000; Downar et al., 2001; Macaluso et al., 2002; Kincade et al., 2005; Indovina and Macaluso, 2007; Hampshire et al., 2007], but has been reported to be suppressed during search [Shulman et al., 2003] or working-memory maintenance of a stimulus [Todd et al., 2005]. Resting state fMRI has additionally revealed that regions within each of the two networks exhibit correlated activity, and that some parts of prefrontal cortex have correlated activity with both networks which could, thus, form a hub for the coordination of the two systems [Fox et al., 2006]. More generally, it has been suggested that the *rich club* [van den Heuvel and Sporns, 2011], i. e. high-degree nodes which are strongly interconnected amongst themselves, might play an important role in coordinating activity between different networks [van den Heuvel and Sporns, 2013a,b].

In addition to these rather coarse-scale imaging studies specific attentional source areas have also been studied electrophysiologically, two prominent ones will be presented in more detail in the following. The involvement of—at least vision-related neurons [Thompson et al., 2005; Gregoriou et al., 2012] in—FEF, for the generation of attentional effects is well-documented (see [Squire et al., 2013] for a recent review). In both humans [Rueckert and Grafman, 1996] and monkeys [Gregoriou et al., 2014] lesions of the FEF can lead to attentional deficits. In macaques, microstimulation of the FEF below the threshold inducing saccades, can improve luminance detection thresholds [Moore and Fallah, 2001], if the stimulation occurs close enough in time to the luminance change that is to be detected [Moore and Fallah, 2004]. In addition, Moore and Armstrong [2003] obtained increased responses in V4 when they stimulated FEF at retinotopically corresponding sites. The response enhancement was stronger for preferred than for non-preferred stimuli, and, in the case of two spatially separated stimuli in the V4 neuron's RF [Armstrong et al., 2006] obtained responses that were reminiscent of the BC phenomenology: without microstimulation, responses to the stimulus pair were intermediate between responses to each of the stimuli when presented alone, and microstimulation of FEF at a site that would lead, if the stimulation were above threshold, to saccade towards one of the two stimuli, changed the intermediate response towards the one elicited when the corresponding stimulus was presented alone (without microstimulation). A receiver-operator-characteristic (ROC) analysis showed fur-

thermore that a V4 neuron's responses could discriminate single oriented bar stimuli better when FEF was microstimulated at a corresponding site [Armstrong and Moore, 2007]. However, as microstimulation could influence antidromically neurons projecting to FEF or axons that just happen to pass by the electrode tip [Clark et al., 2011] it was not certain that FEF was the source of the described attention-like effects in V4. However, similar attention-like effects in V4 could also be obtained through injection of a dopamine D1 (but not D2) receptor antagonist if the injection sites overlapped the V4 neuron's RF [Noudoost and Moore, 2011a]; notably, D1 receptors (but not D2 receptors) are present in the superficial layers of FEF [Lidow et al., 1991; Santana et al., 2009] where projections to V4 originate [Pouget et al., 2009]. Finally, to achieve the same goal, Schafer and Moore [2011] exploited the fact that monkeys could be trained to voluntarily up- (UP trials) or down-regulate activity in FEF neurons and reported attention-like effects (increased search performance and better decodability of stimulus identity from neuron responses) during UP trials.

Electrophysiological evidence also implicates LIP into attention. Gottlieb et al. [1998] report that stimuli that are targets for saccades elicit responses in LIP neurons only if they are relevant, either because of the task or because they pop out. But the absolute value of the activity of a LIP neuron is not enough to decode the locus of attention, instead the findings of Bisley and Goldberg [2003] suggested that the ensemble of LIP neurons representing the entire visual field need to be taken into account and the current peak activity (whatever its absolute value) signals the focus of attention. Moreover, a monkey's ability to ignore a pop-out stimulus was found to be inversely correlated to the activity of LIP neurons [Ipata et al., 2006]. Together, these studies led to the suggestion that LIP contains a "priority map" which combines both bottom-up and top-down influences into a unified representation of behavioral relevance [Bisley and Goldberg, 2010].

Finally, also subcortical areas like the thalamus (including the pulvinar) [Robinson and Petersen, 1992; Bender and Youakim, 2001; O'Connor et al., 2002; Shipp, 2003, 2004; Saalmann et al., 2012] and the superior colliculus [Ignashchenkova et al., 2004; Cavanaugh and Wurtz, 2004; Müller et al., 2005; Zénon and Krauzlis, 2012; Lovejoy and Krauzlis, 2010; Krauzlis et al., 2013] have been implicated in attention.

In summary, attention is the result of a complicated interaction between a widespread network of areas. Apart from the network itself experiments have elucidated specific interactions within this network that will have to be respected by models attempting to understand attention as a whole. While we have not yet reached that point, we will begin to explore the interaction between two regions in chapter 5.

1.2 STRUCTURAL CONSTRAINTS: CORTICAL CONNECTIVITY

An elementary ingredient for circuit models of attention, that we aim to construct, is the network topology which we take to be the anatomical connectivity. How important the precise knowledge of connections is, remains to be determined, on the other hand, and later in this work we will come back to this question. Our guiding hypothesis, based on the evidence presented in section 1.1.3, will be that the appropriate level of description for a mechanistic understanding of attention needs to take the laminar structure of the cortex into account. Experimental datasets for local and global connectivity, that will be needed to test this hypothesis, have recently become available and will be described in the following. Importantly, they reveal a highly recurrent network topology, making a strong argument for the importance of cooperative effects, in contrast to uni-directionally driving processing modes.

1.2.1 *The Local Circuit Connectome: Connectivity Within a Column*

The cortical column with its six-layer-structure is often considered a "canonic" circuit which is repeated over and over in the cortex [Mountcastle, 1997; Douglas and Martin, 2007; Heinzle et al., 2007]. This view has been heavily criticized, though, due to large differences of local circuits found between areas and species in terms of cell size, composition and expressed signaling molecules, as well as synaptic organization, function, and relative abundance of functional cell types [Swindale, 1990; Horton and Adams, 2005; Rakic, 2008]. Due to this heterogeneity Barkow et al. [1992] suggested that columns might be functionally highly specialized, but Hooser [2007] pointed out the possibility that, although different, they might still have similar functions. Moreover, it seems that across areas and species, on average about 80 % of neurons are excitatory and form 85 % of synapses [Douglas and Martin, 2007] so that at this coarse level of description the columnar circuit might, indeed, be "canonic". Unfortunately though, while envisaged for the future [Briggman and Denk, 2006] detailed anatomical wiring diagrams for this circuit are not yet available—with one exception:

Binzegger et al. [2004] perfused horseradish peroxidase (HRP) into neurons in cat area 17. HRP is a dye which is transported both antero- and retrogradely, but not across synapses, and thereby labels the entire dendritic and axonal trees [Calvet and Calvet, 1979; Martin and Whitteridge, 1984] After fixating and slicing the brain, they were able to reconstruct under a light microscope 39 neurons, count their boutons and assign them to connection types using an extension of "Peter's rule" [Braitenberg and Schüz, 1991], that they develop (compare figure 1.3A). In their extended rule, they distinguish between

synapses from neuron type j onto cell bodies in layer u, occurring with probability ρ_j^u, and synapses onto processes (dendrites, axons), occuring with probability $1 - \rho_j^u$, where S_j^u is the total number of synapses that neuron type j forms in layer u. All synapses onto cell bodies are assumed to be evenly distributed among all N^u cell bodies in layer u, whereas all synapses onto processes are distributed in proportion to the length of available processes. Accordingly, they estimate the average number of synapses $\bar{s}_{ij}^u$ from presynaptic neuron type j onto postsynaptic neuron type i in layer u as

$$\bar{s}_{ij}^u = (1 - \rho_j^u) S_j^u \cdot \frac{p_{ij}^u}{\sum_k n_k p_{kj}^u} + \rho_j^u S_j^u \cdot \frac{\delta_i^u}{N^u},$$

where p_{ij}^u is the typical length of a process, n_i is the number of neurons of type i, and δ_i^u is 1 if cell bodies of type i are present in layer u and 0 otherwise. From their reconstructed neurons (and filling in missing data from the literature) they were able to estimate the laminar pattern of boutons—which they confirmed by combined light and electron microscopy to be generally similar to the number of synapses S_j^u—the laminar pattern of processes p_{ij}^u, and the number of neurons per cell type n_i, whereas the ρ_j^u are set by assumptions and data from the literature.

Taking axon terminals or boutons as proxies for synapses thus hinges on the used literature data as well as their reconstructed neurons which were all part of the same local circuit. An independent laminar-resolved stereological estimate of the number of synapses in a few cortical columns of cat area 17, on the other hand, had previously [Beaulieu and Colonnier, 1985] resulted in diverging numbers for $\bar{s}_{ij}^u$. While the match was good for layers 2 and 5, the discrepancies were particularly pronounced in layers 1, 4 and 6, where a high number of both symmetric and asymmetric synapses was missing in Binzegger's compared to Beaulieu's dataset. The reason for the missing symmetric synapses was attributed to not having available neurons to reconstruct, whereas missing asymmetric synapses might have stemmed from afferents outside of area 17.

In summary, the Binzegger et al. [2004] connectome is the "best currently available data for a 'canonic circuit' in the cortex" [Henry Kennedy, personal communication]. Accordingly, it is the best structural constraint to date to understand the dynamics in a column. This is of relevance here, as oscillations, that are a common "dynamical motif" in the cortex, are layer dependent, and attention can modulate these oscillations also with a laminar resolution. These ideas will be refined in chapters 4 and 5.

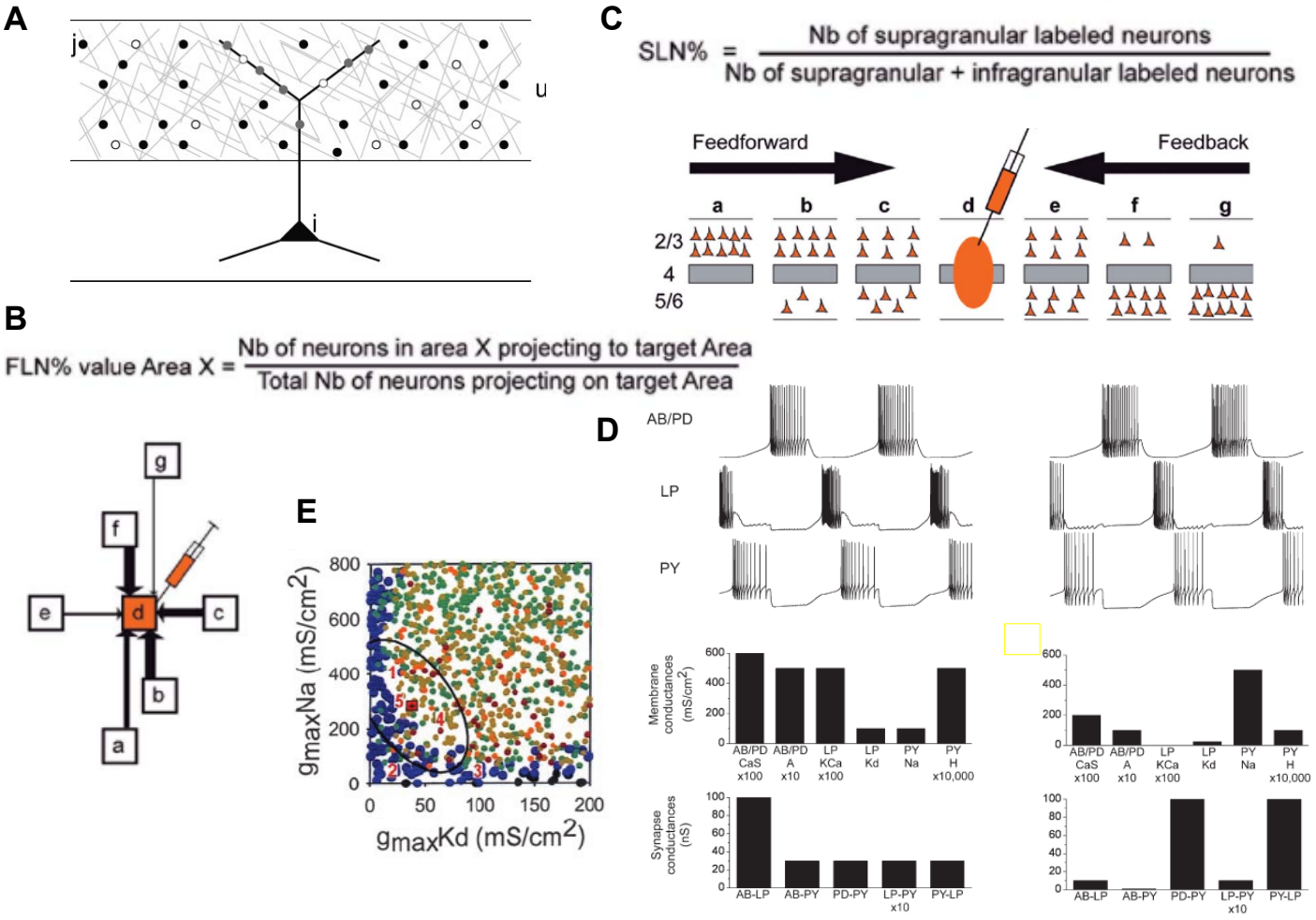


Figure 1.3: Anatomy constraints but does not determine function of neural circuits. A) Presynaptic neurons of type i form synapses (circles) onto dendrites (lines) in layer u, for example onto a postsynaptic neuron of type i (fig. reproduced with permission of Society for Neuroscience from [Binzegger et al., 2004]). **B)** Fractions of neurons outside and targeting a given area (FLNe-value), measured by retrograde tracer injections, is a natural connection-weight measure. (fig. rep. f. [Markov et al., 2014b]). **C)** Fractions of neurons originating in supragranular layers (SLN-value) varies systematically from downstream- to upstream areas. (fig. rep. f. [Markov et al., 2014b]). **D)** Similar dynamics (top row) can be obtained for strongly differing parameters in a model for the pyloric circuit of the crustacean stomatogastric ganglion (fig. reprinted by permission from Macmillan Publishers Ltd: Nature Neuroscience [Prinz et al., 2004], copyright 2004). **E)** Number of spikes per burst in a model-neuron (colors), depending on two of the model's conductance parameters. The parameter region in which one spike per burst is generated (blue) does not include its mean and only part of its 1-standard-deviation covariance (black ellipse). (fig. rep. f. [Golowasch et al., 2002]).

1.2.2 *Long-range Connectivity*

In a seminal study Felleman and Van Essen [1991] assembled anatomical data from several sources into a systematic interareal connection map. Nevertheless such collated datasets can be problematic due to interexperimental differences in tracer sensitivity or definition of areas [Kennedy et al., 2013]. Alternatively, long-range connectivity has been established using tractography, a magnetic resonance imaging technique [Jirsa and McIntosh, 2007]. While non-invasive and thus also applicable to humans, it has the disadvantage of relatively bad spatial resolution ($0.4\,mm^2$ voxel size), such that no laminar information of connection origin or destination can be determined, and neither can be the direction of a connection [Kennedy et al., 2013]. Moreover, it is known to underestimate cross-hemispheric fibers [Kinoshita et al., 2005; Hahn et al., 2006] and allows only a rough quantification of the connection weight.

Recently, however, [Markov et al., 2013, 2014a,b] published a long-range connectivity dataset based on retrograde tracer injections into 29 areas of 26 macaque monkeys thereby obtaining 1615 pathways (a third of which was previously unknown), that is two thirds of all possible connections between the 29 areas were found [Markov et al., 2014a]. They could also determine a natural "weight" for each connection, the FLNe [Markov et al., 2011], which is the fraction of all labeled neurons (which must thus project to the target area) relative to all labeled neurons outside the injection area (compare figure 1.3B), and found that this weight decays exponentially with cortical distance (shortest path via white matter) [Ercsey-Ravasz et al., 2013]. Beyond that, their experiments allowed to determine SLN values [Barone et al., 2000], the fraction of labeled neurons originating in supragranular layers (compare 1.3C). As it has been shown before that laminar pattern of connections between areas can determine a hierarchy between cortical areas [Felleman and Van Essen, 1991] they assigned a (continuous) rank, β_i, to each area i, and found that there existed an approximately monotonous relationship between $\mathrm{SLN}_{i\to j}$ values of connections originating in area i and terminating in area j, and the differences in rank of these areas, $\beta_j - \beta_i$. Appealingly, this statistical model also allows to extrapolate SLN values for pairs of areas whose connections have not been measured directly, but whose position in the hierarchy has been determined through the involvement in other connections.

The resulting order of areas in the cortical hierarchy was remarkably similar to the one determined by Felleman and Van Essen [1991] and is useful for tagging the nature of a given connection between areas as either feedforward (FF) or feedback (FB). Further analysis revealed that supragranular (infragranular) neurons project longer in FF (FB) direction. They also observed that FB pathways are more numer-

ous than FF pathways, whereas the weight of (i. e. the number of neurons involved in) short connections is stronger for FF pathways; there was no difference in weight for longer connections. FF and FB connection streams in supragranular layers were separated and labeled neurons tended to originate from a small region, while the opposite was the case in infragranular layers, where both pathways were intermingled and labeled neurons originated from a wider region. Finally, in simultaneous injections in V1 and V4 with different dyes they hardly found any double-labeled neurons in V2 and V3, whereas these existed in higher areas, suggesting that neurons participate in either the FF or the FB pathway.

What is crucially missing, is a corresponding dataset for the laminar termination patterns of connections, so far only collated datasets exist, which are summarized in [Douglas and Martin, 2004; Markov et al., 2014b].

In summary, these data by Markov et al. [2014b] are a major refinement compared to the previously available knowledge on long-range connectivity in the cortex and will therefore serve as an integral building block for a global model of attention. While we don't have, currently, such a model, we will outline corresponding ideas in chapter 5.

1.2.3 *Structural Degeneracy*

Binzegger et al. [2004]'s local connectome is based on reconstructed neurons from a single cat; Markov et al. [2014b]'s global connectome is assembled from tracer injections in 26 monkeys, but no comparison between individual monkeys was provided. The variability in both of these connectomes remains, thus, unknown. However, we can make an educated guess looking at other systems.

Structurally different systems can produce very similar outputs and this "degeneracy" is a widespread phenomenon in biology [Edelman and Gally, 2001]. One extensively studied example of it is the pyloric circuit of the stomatogastric ganglion in the crab *Cancer Borealis*. When one out of several distinct cell types from the ganglion was substituted in a circuit, similar activity could be obtained, despite large variations in the cells' intrinsic properties like slope of frequency-current curve or spike height [Grashow et al., 2010]. Measuring cellular and synaptic properties in the pyloric circuit, pronounced preparation-to-preparation variability as well as complex relations between these parameters and circuit performance were observed, pointing out the importance of measuring as many parameters as possible in the same preparation [Goaillard et al., 2009]. Thus, an intricate homeostatic mechanism seems to be at work, allowing the circuit to compensate changes in certain parameters by maintaining its essential dynamical repertoire despite variations in the anatomy.

Moreover, that a circuit's reliability, in the presence of external input or neuromodulators, can be maintained, is surprising [Nowotny et al., 2007; Marder, 2011], but consistent with experiments: for example, the relative phases of neurons in the pyloric rhythm remained relatively unaffected despite vast temperature differences [Tang et al., 2010].

Similar results to the experiments have been obtained in models [Goldman et al., 2001; Prinz et al., 2003, 2004; Taylor et al., 2009; Gutierrez et al., 2013] (compare figure 1.3D) in which qualitatively similar dynamics was found mostly in connected parts of parameter space [Goldman et al., 2001; Prinz et al., 2003; Taylor et al., 2009; Marder and Taylor, 2011], but supposedly also in distinct parts, separated by parameter phases of qualitatively different dynamics [Gutierrez et al., 2013]. Due to the intricate relation between all relevant parameters and behavior it is also not surprising that a circuit composed with average parameters could behave differently from one instantiated with one particular sample of the parameters [Golowasch et al., 2002; Marder and Taylor, 2011; Marder, 2011] (compare figure 1.3E).

In summary, we can expect (but do not know) that the column circuit and also the long-range circuit are not unique but might tolerate variations to a degree. What is unclear, in addition, is if they possess a specific internal structure that predisposes them for certain dynamics, and what kind of dynamics these are. These problems will be commented upon in chapters 4 and 5.

1.3 REVIEW OF RELEVANT MODELS

After summarizing biological facts that are relevant for an understanding of attention we will now survey existing models and how they strive to explain it. We will see that the distributed network of areas involved in attention is vastly neglected by most of the models and we will therefore also present ideas how these areas, or circuits in general, could be coordinated.

1.3.1 *Models for Visual Attention*

The described aspects of attentional response modulations in visual cortical areas have inspired quite a number of modeling studies at various levels of detail. In the following I will survey some of them, focusing on models of top-down attention. For a review of models of bottom-up attention see [Borji and Itti, 2013].

Attention is one specific case of interareal coordination. Accordingly, the CTC hypothesis—which we have discussed above in section 1.1.3 as an experimentally-grounded model for how the ubiquitous cortical oscillations might be harnessed for exerting control—generally applies, in particular it has been shown that phase relation-

ships between populations might be advantageous for flexible information routing [Battaglia et al., 2012; Kirst, 2012], which will be the topic of the following section.

First, however, we present computational models specifically designed to reproduce attentional effects. In section 1.1.1 we have already discussed phenomenological models of attention, including the BC model, FSGM, and normalization model. Other models, on the other hand, have been designed with a circuit architecture in mind. Some "early" studies aim to reproduce BC phenomenology [Niebur et al., 1993; Niebur and Koch, 1994; Reynolds et al., 1999; Deco and Rolls, 2005; Hamker, 2005], while others also consider effects of synchrony [Ardid et al., 2007, 2010; Tiesinga and Sejnowski, 2004; Tiesinga, 2005; Tiesinga and Sejnowski, 2009; Buia and Tiesinga, 2006, 2008; Mishra et al., 2006; Börgers et al., 2005; Börgers and Kopell, 2008; Börgers et al., 2008; Lee et al., 2013; Buehlmann and Deco, 2008; Zeitler et al., 2008; Paik et al., 2009; Gielen et al., 2010; Wagatsuma et al., 2011, 2013]. In all these studies external stimuli are introduced as a drive—which might be a Poisson spike train or a constant current—to the lowest hierarchical circuit element, affecting either excitatory or both excitatory and inhibitory populations, a distinction that is relevant for the mechanism of rhythm generation [Tiesinga and Sejnowski, 2009]. Moreover, the internal architecture as well as the mechanism in which attention is operating differs strongly between the models.

In a series of papers Tiesinga and coworkers explored the hypothesis that attentional effects are mediated via inhibitory interneuronal networks. They were able to reproduce competition and synchrony effects by modeling attention as an oscillatory inhibitory input to a Hodgkin-Huxley-like neuron [Tiesinga et al., 2004], a decrease of the constant current input to the inhibitory population of a densely connected spiking network [Buia and Tiesinga, 2006] or an increase in the constant current drive to a subpopulation of neurons [Buia and Tiesinga, 2008]. Tiesinga [2005] simulated the effect of attention by modulating the delay between two inhibitory inputs to an excitatory neuronal population, which also received excitatory feedforward input from a "poor" and a "good" stimulus. When the delay between the two arrival times of the two inhibitory populations was small (long) the output rate was close to the rate in response to the preferred (anti-preferred) stimulus alone. Mishra et al. [2006] proposed yet another mechanism for the mediation of attentional effects. In a compartmental model of a reconstructed V4 neuron they found that sparsely synchronized excitatory, in combination with fully synchronized inhibitory oscillatory spiking inputs led to increased firing rates in the target neuron over a wide range of input frequencies as long as the excitatory and inhibitory spike volleys were sufficiently out of phase.

The possible way in which gamma rhythms might subserve stimulus competition and selection was elucidated by Börgers and co-workers [Börgers et al., 2005, 2008; Börgers and Kopell, 2008]. They studied a conductance-based spiking network of excitatory and inhibitory cells and tuned its parameters to obtain PING[3] oscillations. When presenting bottom-up input in the form of a constant current injection to a subset of cells weak PING rhythms were strengthened to PING rhythms, an effect that could be abolished by disturbing the network coherence [Börgers et al., 2005]. In another study [Börgers et al., 2008], when two stimuli targeting different subsets of excitatory cells were presented simultaneously, they disrupted the PING rhythm present in the network causing asynchronous inhibition to hit the network and thus reducing firing rates compared to the singular presentation of one stimulus alone. Attention was modeled as a constant inhibitory current additionally injected to either the inhibitory or both the inhibitory and excitatory cells while concomitantly increasing the drive of the attended stimulus. While the latter was necessary to break the symmetry between the two stimuli in the simulation, the bath of inhibition was thought to be mediated by attention-induced cholinergic activation of a second not explicitly modeled class of inhibitory interneurons. Together, these two manipulations attributed to attention were able to restore the network rhythm and thereby the original firing rate of the targeted excitatory cells on each cycle of the gamma rhythm. Finally, when a constant input creates oscillatory output more downstream areas should correspondingly be studied with an oscillatory drive. The presence of a gamma rhythm in such a downstream area also mediated competition between concomitantly arriving inputs [Börgers and Kopell, 2008]. When one of them was more coherent than the other (even though the second might, on average, be stronger), more in phase with the network than the other, or had a frequency in the gamma range whereas the other hadn't, then this stimulus entrained the network. These three studies illustrated why gamma rhythms might be advantageous to mediate attentional effects, but, on the other hand, the fact that they only consider a small, fully connected network of excitatory and inhibitory cells, that their PING rhythms despite being "weak" seem appear pretty regular (compared to, for example, the model by [Brunel and Wang, 2003] or experimentally determined Fano factors around 1 [Churchland et al., 2010]) and model attention by introducing ad hoc manipulations makes them somewhat unrealistic.

The relationship between rate and synchrony modulations is further addressed in a model by Deco and co-workers. They used a spiking network consisting of two pools of neurons receiving selective

3 Pyramidal neurons, firing either on every or stochastically on the rhythm beat, thereby driving and synchronizing inhibitory neurons, which in turn prevent the pyramidal neurons from spiking before the next clock beat are termed, respectively, *strong* and *weak* PING rhythms [Whittington et al., 2000; Börgers et al., 2005].

input, as well as one unspecific and one inhibitory neuron pool. Together these neurons represent an early visual area (V1 or V2) which is interacting with a V4 module of the same design. Attentional bias is introduced ad hoc through an additional Poisson current to one of the selective neuron pools in the early area. This network was able to reproduce BC effects [Deco and Rolls, 2005] and, as found by Fries et al. [2001], attentional modulations of both beta and gamma frequency power [Buehlmann and Deco, 2008] (compare figure 1.4A). In the latter study, rate modulations could occur without gamma power modulations, but not vice versa, and gamma modulations could be manipulated without rate changes; the two attentional effects were, thus, independent of each other. Interestingly, though, for network parameters that led to both rate and synchrony modulations due to attention, reaction times (i. e. the time until the mean rate in a pool is reached) were faster in the neuron pool representing the attended stimulus. Finally, attentional effects were stronger in the downstream area (V4), in agreement with Buffalo et al. [2010].

Ardid et al. [2007, 2010] conceived a model (compare figure 1.4B) that explicitly incorporated a source area of attention such that all attentional effects are emergent phenomena. The conductance-based spiking network model consisted of a sensory area (for example MT) and a control area (for instance FEF). After tuning the network parameters the model was able to reproduce attentional firing rate modulations as captured by phenomenological models of attention, that is, tuning curves were multiplicatively scaled by attention (multiplicative gain modulation—MGM, compare figure 1.4C), the network showed BC phenomenology (compare figure 1.4D), and also exhibited FSGM behavior. Notably, it was tested in the simulation if FSGM also holds for simultaneous presentation of a preferred and anti-preferred stimulus in the same RF of a cell—which has not been done experimentally up to now. The model predicted that, indeed, the hypotheses of FSGM generalize to this situation and that the monotonic, $\theta_{\text{preferred}} - \theta_{\text{attended}}$-dependent population-modulation function was remarkably similar for both single and dual stimulus-presentations (compare figure 1.4E).

Apart from capturing, thus, all major rate effects the model also reproduced attentional gamma-band coherence modulations, both locally in MT and FEF, as well as across areas between MT and FEF. The latter selectively occurred only between neurons whose preferred stimulus matched both the attended and the stimulus direction (compare figure1.4F): When either a trial without attention was considered, or attention was directed to an anti-preferred stimulus, or attention was directed to a cell's preferred direction but the presented stimulus had an opposite direction, no increase of gamma-band coherence occurred.

The marked increase of synchrony due to attention was seemingly in contrast to the findings of high spiking variability that might even

be reduced by attention [Mitchell et al., 2007, 2009; Cohen and Maunsell, 2009]. The model also captured this aspect, in that single neurons fired irregularly and bursty with Fano factors around 0.8 while populations as a whole were engaged in rhythmic firing. Attention, thus, did not turn neurons into oscillators.

In total, this model captures a great number of experimental findings and due to this realism is a powerful tool to test in how far oscillations are functionally relevant to mediate attentional effects. To this end, random delays for the FEF→MT conductances were introduced that destroyed coherence between these areas. In addition, these random delays mildly affected attentional rate enhancements, as well as the BC phenomenology and the multiplicative gain modulation (MGM), and slightly increased Fano factors. It was also observed that decreased attentional rate enhancements could be restored to their original value observed without random delays, if either FEF firing rates were artificially enhanced or MT received an additional suitable zero-mean input oscillating in the gamma-band (which did not restore MT-FEF coherence). As, hence, attentional effects could be obtained without MT-FEF synchrony, oscillations did not seem to constitute an additional, parallel mechanism for the encoding of information. Rather, rate and synchrony seemed to be tightly intermingled. Gamma rhythms were interpreted as a more metabolically efficient replacement of higher firing rate, with the additional advantage of greater down-stream impact [Salinas and Sejnowski, 2001], and with a modest contribution to variability reduction. On the other hand, oscillations might play a much more prominent role for the encoding of which stimulus was attentionally selected [Grothe et al., 2012; Bosman et al., 2012]. In particular, information could be decodable from the phase of population rhythms [Montemurro et al., 2008; Kayser et al., 2009; Siegel et al., 2009]

This model is impressive in that it incorporates attentional rate effects, variability modulations and gamma-band coherence modulations and, notably, produces attentional effects as an emergent phenomenon. What is missing, though, is the additional modulation of low-frequency coherence. Moreover, by design, the model cannot capture the progressive increase of firing rates effects in visual cortex observed by Buffalo et al. [2010], as well as the layer-dependent frequency specific coherence modulations reported by Buffalo et al. [2011], more specifically it is unclear what role their employed connectivity profiles between neuron populations plays.

Lee et al. [2013] simulated a conductance-based spiking network consisting of two connected columns of the same internal architecture (compare figure 1.4G), each of which contained four different classes of excitatory and five different classes of inhibitory neurons which were interpreted to be distributed over cortical layers 2/3, 4 and 5. In their model, asynchronous bottom-up inputs generated PING os-

cillations in layer 4 through interactions of an excitatory and a FS inhibitory neuron population. Attention was modeled as synchronous 20 Hz spike trains targeting layer 5 excitatory neurons of one of the two columns. In this model, attention was found to increase (decrease) firing rates of L2/3 RS cells in the "attended" ("unattended") column, and to enhance (weaken) the difference in gamma (beta) range synchrony between these cells in the attended and the unattended column. A number of manipulations of network suggested a prominent role for the ascending inhibition from layer 5 SI to layer 4 FS cells in order to maintain layer 4 excitatory spike rates and the differential modulation of gamma synchrony between columns by attention; for cross-columnar connections and layer 2/3 SI cells in order to obtain desired rate and beta synchrony modulations between columns; and for FS interneurons across layers to uphold attentional gamma frequency modulation between columns. Importantly, the observed effect depended critically on the top-down input to be oscillatory. When, instead, an asynchronous top-down signal was sent to layer 5 beta-band synchrony differences between the columns vanished and firing rates as well as gamma-band synchrony even became stronger in the unattended than in the attended column. The model hence predicts a deep-layer interneuronal population resonating at beta-frequencies to top-down input. On the other hand, two important neglected aspects in this study include—as we will demonstrate in chapter 4—that oscillations depend on the collective state of the network, and are also strongly dependent on the connectome.

Finally, Wagatsuma et al. [2011, 2013] follow up the hypothesis that the layered structure might be an important substrate for attentional effects by implementing a large-scale conductance-based spiking network model which comprises excitatory and inhibitory integrate-and-fire neurons distributed over layers 2/3, 4, 5 and 6 with a realistic, experimentally determined interlayer connectivity based on data by Thomson et al. [2002]; Thomson and Morris [2002]; Binzegger et al. [2004] (compare figure 1.4H). Bottom up and top-down inputs were modeled as Poisson spike trains to both excitatory and inhibitory neurons targeting layer 4 with rates between 2 and 20 Hz (depending on the column's preference), and targeting layers 2/3 & 5 with a rate of 5 Hz, respectively. In simulations of two connected columns of opposite orientation preference neurons in L2/3 and L5 exhibited rates following BC predictions when one or two stimuli, with or without attention were considered. Notably, the simulations predicted that L4 cells deviate from the BC model [Wagatsuma et al., 2011]. To study both spatial and feature attention eight columns of gradually varying orientation preference were connected. In this setting, excitatory long-range connections only reached until the neighboring column in features space, whereas inhibitory connections targeted all eight columns. Likewise, the top-down signal representing

feature attention targeted only the column of corresponding feature selectivity, whereas a spatial attention signal reached all columns. The bell-shaped orientation tuning curves were differentially modulated by spatial and feature attention. In the former case, L2/3E and L5E tuning curves appeared multiplicatively scaled, though in L5E there was an additional increase in baseline activity that was absent in L2/3E. On the other hand, feature attention resulted in narrower tuning curves, with increased amplitude and decreased baseline for both L2/3E and L5E neurons. L4E and L6E tuning curves were inverted, with smaller responses for preferred stimuli. Interestingly, not all neurons were affected by attention. Only around 60 % to 90 % of neurons in layers 2/3 and 5 were significantly modulated, and the effects of attention—even when significant—were opposite to the paradigmatic case in a considerable fraction among them [Wagatsuma et al., 2013].

In summary, neural correlates of attention have been found in sensory areas. They comprise firing rate modulations, changes of synchrony and variability. Some phenomenological models of attention have striven to capture the "average" effects on firing rates, but cannot concomitantly explain effects on synchrony and variability. A variety of more detailed circuit models, on the other hand, reproduced more and varying parts of the experimental findings in the physiological attention literature. It stands out, furthermore, that most models do not explicitly contain the source areas of attention; instead the attentional signal is ad hoc injected from the outside. This seems unsatisfactory, not only because of a lack of elegance. Ardid et al. [2010] have shown that, indeed, a source area of attention can successfully be incorporated in a model—in a rather straightforward way. When, on the other hand, attention is injected externally, it is not clear in what way the attentional signal should be modeled, and, indeed, models have suggested a variety of possibilities: constant current injections, either positive or negative, sent to excitatory and / or inhibitory cells, synchronous or asynchronous inputs, delays between two inputs sent to different cell classes or the arrival phase of the input. Moreover, in this case no reciprocal influence on the "source" is possible. In the model of Lee et al. [2013], for example, it is assumed that a neuron class resonates to an extrinsic beta-frequency input. However, it has been pointed out that the mechanism by which oscillations are generated in the cortex might well be of considerable importance for the interplay of different neuron classes in a network [Cannon et al., 2014]. That raises the question how the network would behave if not resonance but interaction between source and target area of attention generated the oscillation intrinsically through mutual interaction. Next, the finding that coherence and its attentional modulation occur differentially in cortical layers [Buffalo et al., 2011] suggests that the cortical column might be an important building block to understand attention in a circuit model. Correspondingly, the model of Lee et al.

[2013] was designed with a laminar structure in mind, though the connectivity between cell classes of different "layers" lacks biological underpinning, so that a "layer" in the model is essentially only a terminology assigned to a group of neurons. Wagatsuma et al. [2011, 2013], on the other hand, used experimentally determined data for their column connectomes, though, in contrast to Lee et al. [2013]'s model, they used only one neuron type (integrate and fire). Moreover, due to the computational complexity of the large-scale network (about 160000 neurons in [Wagatsuma et al., 2013]) the model does not scale well to incorporate a source area of attention. Also, they did not yet study attentional coherence effects (but are planning to [Schmidt et al., 2013]).

Altogether, it would be desirable to construct circuit-based models of attention which explicitly contain the source area of attention and possesses a laminar structure at the same time. For biological plausibility both the columnar and interareal connectome should be realistic and based on experimentally determined data. It would then be of interest to see how the connectome constrains such a circuit model, in what way the dynamics of both the attentional and non-attentional state emerge from the same underlying structure and what mechanism induces the shift between the two states. These ideas will be continued in chapters 3, 4 and 5.

1.3.2 *Interareal Coordination*

Of all the models discussed above, only the model by Ardid et al. [2007, 2010] contains a non-visual area. While the authors did not carry out a detailed analysis of the the interaction between the two areas, the mechanism is very likely similar to ideas presented by Compte et al. [2000]; Wang [2001]; Wong and Wang [2006]: the system is multistable, possessing at least two attracting states, an essentially complete "quiescent" one, and a "sustained" one characterized by spiking in the control (in correspondence with experiments [Armstrong et al., 2009]) and no spiking in the sensory area. Presenting the cue-stimulus modifies the phase-space of the system in such a way that a third, "active" state is stabilized in which both the sensory and control area spike. When the cue is switched off again, the system phase-space reverts back to its previous configuration, the "active" state loses stability, and the quiescent and sustained state regain it. However, as the now unstable "active" state lay in the basin of attraction of the "sustained" state, the system does not shut down again, but is attracted towards the stable "sustained" state. When the stimulus is switched on again, mimicking the test-period of an attention-experiment, the system trajectory evolves from this sustained-spiking attractor and thus differently, compared to when it would start in the quiescent state in which none of the system's neurons is in firing

mode. Due to the multistability specific mechanisms to coordinate sensory and control areas are not necessary because they emerge from the mutual coupling making the system evolve as a *whole* between global, system-wide limit states.

More generally, it is not necessary for this scheme of interareal coordination that the system settles or even possesses limit states. It could, in principal, also work in frameworks of transient, metastable states like those described by Tognoli and Kelso [2014]; Rabinovich et al. [2008]. In fact, the essence of the idea is to consider the mutually coupled system as a *whole* thereby rationalizing the need for a cooperative mechanism—it "simply" emerges from the coupling. Unfortunately, this approach does not reveal if there exist underlying, possibly universal mechanisms at a finer than global scale of the system's sub-components.

Neural circuits are generally multifunctional [Briggman and Kristan, 2008; Kelso, 2012], that is, can support various dynamical patterns given the same structural connectivity and, even in the absence of a specific task, due to noise [Deco et al., 2011, 2013b] or inherent metastability of the state [Tognoli and Kelso, 2014; Rabinovich et al., 2008] the brain continuously explores different states [Honey et al., 2007; Hansen et al., 2015], a process that might be related to mind-wandering [Mason et al., 2007]. Importantly, the precise dynamical state might have consequences for the effective network topology. Indeed, in general neurons in the brain form a *recurrent* network [Binzegger et al., 2004; Sporns and Kötter, 2004; Song et al., 2005; Perin et al., 2011]. Nevertheless, Battaglia et al. [2012] and Kirst [2012] demonstrated that couplings between nodes are more or less effective (measured by transfer entropy [Schreiber, 2000] and delayed mutual information [Kirst, 2012], respectively) such that, in the extreme case, nodes can become *effectively uncoupled*. Interestingly, the determinant of link efficacy was the dynamical state—more precisely the relative *phase* of oscillations in interacting network nodes—and switching dynamics led to different effective network topologies. Moreover, Battaglia et al. [2012] demonstrated that information transfer (measured by the mutual information, cf. e. g. [Cover and Thomas, 2006]) contained in spiking patterns that are transferred between groups of neurons was likewise lower for less efficient connections. In that view, switching dynamics provides a way to either segregate, direct or integrate processing between sub-components on the fly.

In summary, a number of more or less abstract ideas for the dynamical coordination of cortical circuits have been proposed. In relation to the relevance of oscillatory synchrony—discussed above in section 1.1.3—the demonstration that the dynamical state and their "spectral fingerprints" [Siegel et al., 2012] (be it preferred phase relationships, patterns of cross-frequency coupling or metastable attractors) might

be determinants for information transfer is most intriguing and we will therefore revisit these ideas in chapters 4 and 5.

1.4 SUMMARY & OUTLINE

In the above survey we have identified several issues that are necessary for a deeper understanding of attention. In the remainder of this work we will address them by analysis of experimental data and modeling.

Chapter 2: feature extraction method for noisy data, differences between spatial & feature attention, highly cell- and stimulus specific effects

In **chapter 2**, we will start with the analysis of data. Due to the yet unclear differences between spatial and feature attention discussed in section 1.1.1 we analyzed a dataset recorded in macaque area MT while the monkeys were viewing one out of twelve possible pairs of directional stimuli, either spatially separate or transparently overlaid, in the same neuron's RF. As anticipated in section 1.1.2, recorded neurons had a high trial-to-trial variability. Therefore, when idealized model functions were fit to these data—as is commonly done—we found that the best model function had to be chosen on a cell-by-cell basis because each model led to statistically significant predictions for the effects of attention on the tuning curves. To circumvent the model-selection altogether we first suggest instead an alternative method for the analysis of noisy tuned data: to base the analysis of tuning curve features on algorithmic descriptors that can be calculated solely on the basis of the measured data, independent of any model function. While methodologically free of the ambiguities involved in model selection we also find that these descriptors agree well with predictions that are based on fitting to models—as long as the models agree amongst each other. Second, when applied to our data, this method revealed qualitatively different attentional modulation patterns for spatially separate and transparent stimulus pairs. Third, a cell-by-cell analysis revealed that attentional effects were highly specific. Tuning curves of roughly half of the recorded cells were not at all significantly modulated by attention, and those that were only for a small number of stimuli that were also seemingly randomly distributed over the whole tuning curve, indicating that the attended feature cannot be reliably decoded from a single cell.

Chapter 3: phenomenological models of attention problematic, rate model of multiple interacting rings

In **chapter 3** we sought a single model able to describe the population-averaged effects of attention on spatially separate and transparent stimuli at the same time. First, we tested if phenomenological models of attention, the BC and the FSGM discussed in section 1.1.1, were compatible with our dataset and found that both of them had deficiencies. Second, as an alternative, we attempted to find a minimalist description in terms of a circuit model in which attention was modeled as an external contribution to the network modulating the interactions occurring within the circuit. As the two spatially separated stimuli in the experiment presumably fell into the RFs of distinct V1 neuron

populations the model architecture consisted of three hypercolumns, one representing MT, the other two V1. The model was constrained by the measured MT responses from both paradigms of our dataset and predicts that a given response in the MT hypercolumn can be associated with a variety of qualitatively different circuits and V1 responses, and that the resulting circuits can also be qualitatively different when average compared to individual cell's MT responses are used for constraining. These results are reminiscent of homeostasis or structural degeneracy mechanisms that we discussed in section 1.2.3. On the other hand, many of the solutions obtained from the model seem counterintuitive and we will discuss possible reasons for that.

Chapter 4: Local circuit model, layer-specific context and connectome dependent multi-frequency oscillations

In **chapter 4**, given evidence for large-scale coherence networks in the cortex (cf. section 1.1.3) and the question what role the underlying anatomy plays in shaping them (cf section 1.2), we constructed a model for a local circuit based on experimentally determined interlayer connectivity data. Similar to experiments, we observed (for certain parameters) oscillations in neuronal populations in the column that were layer dependent, with a tendency of fast ("gamma") oscillations to occur in upper and slow ("beta") oscillations in deep layers. The predisposition of the column to show such a pattern of oscillation could be influenced by phenomenological contextual influences. Moreover, in this model populations of fast oscillators were present in the upper layers (causing the "gamma"-tendency there), but the slow oscillations emerged through the interactions between the layers, with characteristic frequencies that were indicative for a quasi-periodic mechanism, although artificial introduction of a slow oscillator in the deep layers led to qualitatively similar results. Furthermore, by scrambling the interlayer connectivity we showed that a very similar dynamical repertoire for the column could be obtained for qualitatively different connection patterns—but these were exceedingly rare, suggesting that the knowledge of the connectome is advantageous for the study of neural circuits. Altogether, the complexity of yet meaningful results reminiscent of experimental findings casts doubt that these effects could be captured by abstract phenomenological models. Furthermore, an interesting continuation would be to study the principles of how multiple such columns interact.

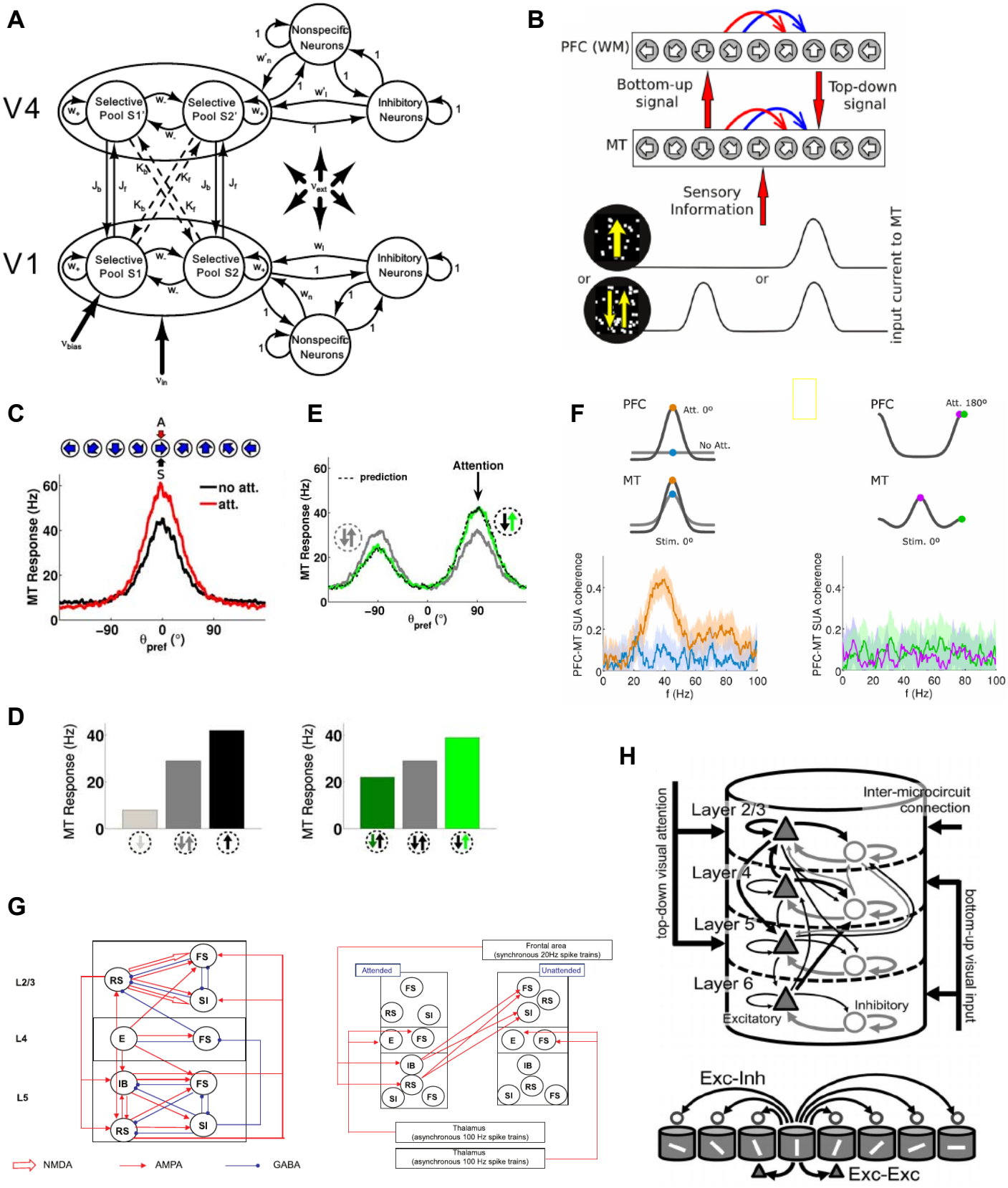
Chapter 5: frequency dependent feedforward/feedback configuration (according to oscillation phases) in 2 coupled columns

In **chapter 5**, we report preliminary results for a system consisting of two coupled columns situated at different levels of the cortical hierarchy. When parameters are chosen such that complex layer-specific multi-frequency oscillations occur, these can self-organize such that oscillations in the lower column lag behind those in the upper column in the absence of specific input, compatible—according to computational studies Battaglia et al. [2012]; Kirst [2012]—with a feedback-dominated configuration. Moreover, when the system received bottom-up input the lower column's oscillations became phase-leading—compatible with a feedforward processing mode—and when an additional top-

down drive arrived at the upper column, the system could be "feed-forward" for high (gamma-like), but "feedback" for slow (beta-like) frequencies, in line with a currently discussed role for cortical oscillations.

Finally, in **chapter 6** we conclude with a summary of results and a discussion of our approach towards a brain-wide model of attention.

Figure 1.4 ***(facing page)*****: Overview of models for attention. A)** Each area (V1, V4) contains four pools of integrate-and-fire neurons: two selective excitatory ones, another, non-selective excitatory, as well as an inhibitory pool. All inputs are modeled as Poisson spike trains, stimulus-related input, of rate ν_{in} is received only by the selective pools (the other two represent surrounding brain regions), attention is modeled as an additional bias of rate ν_{bias} to one of the selective pools in the lower area, and all pools receive unspecific input of rate ν_{ext}. Recurrent connection weights are indicated by w, strong and weak interareal connections by J and K. The model reproduces BC phenomenology and attentional modulation of synchrony, shows that the two can occur independently of each other, but indicates that the combined occurrence might reduce reaction times and might thus be advantageous. (fig. reproduced with permission of Society for Neuroscience from [Buehlmann and Deco, 2008]). **B)** Two connected direction-selective hypercolumns, situated in a sensory (e. g. MT) and control (e. g. **PFC!**) area, respectively, and consisting of locally coupled excitatory and inhibitory integrate-and-fire neurons, receive either uni- or bidirectional input, represented as a current to one or two groups of neighboring neurons in the sensory hypercolumn. Attentional effects are modeled by presenting a corresponding cue stimulus prior to the actual test stimulus presentation. (fig. reproduced with permission of Society for Neuroscience from [Ardid et al., 2010]). **C)** Attentional modulation of the population response depends on the angular distance between neurons' preferred and the attended direction, in line with the FSGM (fig. reproduced with permission of Society for Neuroscience from [Ardid et al., 2007]). **D)** Presenting uni- and bi-directional stimuli, without (left) and with (right) attention (green) reproduces BC phenomenology (fig. reproduced with permission of Society for Neuroscience from [Ardid et al., 2007]). **E)** The model predicts that the FSGM also holds for simultaneous preferred+anti-preferred stimulus presentations, i.e. attentional enhancement is weaker for neurons whose preferred direction is far from the attended direction (fig. reproduced with permission of Society for Neuroscience from [Ardid et al., 2007]). **F)** Coherence between MT and **PFC!** neurons with preferred directions 0° is selectively increased when the attended direction matches the stimulus direction (orange), but not if no attentional cue was presented before the stimulus (blue), nor when attention was directed to the opposite direction (purple), nor when the 0°-neuron receives no input and attention was directed else-where. (fig. reproduced with permission of Society for Neuroscience from [Ardid et al., 2010]). **G)** Various neuron types—FS, RS, SI, E(xcitatory), IB—ad-hoc attributed to "layers" to form a "column". Two such columns are interconnected and receive asynchronous bottom-up input in their middle layers. If attention is modeled as oscillatory input to the deep layers, BC phenomenology, and attentional enhancement (suppression) of gamma (beta) synchrony is reproduced (fig. rep. f. [Lee et al., 2013]). **H)** 8 columns, consisting of a large number of integrate-and-fire neurons situated in 4 layers that are connected according to experimental data, are placed on a ring to form a hypercolumn with angular-distance dependent connection strengths. Bottom-up and top-down inputs are modeled as Poisson spike trains. Spatial and feature attention were distinguished by targeting all or just one column, respectively, and gave rise to differential modulation of tuning curves. Moreover, there was a huge heterogeneity in single-neuron modulations. (fig. rep. f. [Wagatsuma et al., 2013].)


A
V4
V1
Nonspecific Neurons
Selective Pool S1'
Selective Pool S2'
Inhibitory Neurons
Selective Pool S1
Selective Pool S2
Nonspecific Neurons
B
PFC (WM)
Bottom-up signal
Top-down signal
MT
Sensory Information
or
or
input current to MT
C
A
S
no att.
att.
MT Response (Hz)
θpref (°)
D
MT Response (Hz)
MT Response (Hz)
E
prediction
Attention
MT Response (Hz)
θpref (°)
F
PFC
Att. 0°
No Att.
MT
Stim. 0°
PFC-MT SUA coherence
f (Hz)
PFC
Att. 180°
MT
Stim. 0°
PFC-MT SUA coherence
f (Hz)
G
L2/3
L4
L5
FS
RS
SI
E
IB
NMDA
AMPA
GABA
Frontal area (synchronous 20Hz spike trains)
Attended
Unattended
Thalamus (asynchronous 100 Hz spike trains)
Thalamus (asynchronous 100 Hz spike trains)
H
Inter-microcircuit connection
Layer 2/3
Layer 4
Layer 5
Layer 6
top-down visual attention
bottom-up visual input
Inhibitory
Excitatory
Exc-Inh
Exc-Exc

2

MODEL-FREE ESTIMATION OF TUNING CURVES AND THEIR ATTENTIONAL MODULATION, BASED ON SPARSE AND NOISY DATA

This chapter, together with appendix A, was submitted for publication [?]. I acknowledge my co-authors Vladislav Kozyrev, Valeska Stephan, Stefan Treue, Theo Geisel and Demian Battaglia. Experiments were performed by VK, VS and Anja Lochte.

2.1 ABSTRACT

Tuning curves are the functions that relate the responses of sensory neurons to various values within one continuous stimulus dimension (such as the orientation of a bar in the visual domain or the frequency of a tone in the auditory domain). They are commonly determined by fitting a model e.g. a Gaussian or other bell-shaped curves to the measured responses to a small subset of discrete stimuli in the relevant dimension. However, as neuronal responses are irregular and experimental measurements noisy, it is often difficult to determine reliably the appropriate model from the data. We illustrate this general problem by fitting diverse models to representative recordings from area MT in rhesus monkey visual cortex during multiple attentional tasks involving complex composite stimuli. We find that all models can be well-fitted, that the best model generally varies between neurons and that statistical comparisons between neuronal responses across different experimental conditions are affected quantitatively and qualitatively by specific model choices. As a robust alternative to an often arbitrary model selection, we introduce a model-free approach, in which features of interest are extracted directly from the measured response data without the need of fitting any model. In our attentional datasets, we demonstrate that data-driven methods provide descriptions of tuning curve features such as preferred stimulus direction or attentional gain modulations which are in agreement with fit-based approaches when a good fit exists. Furthermore, these methods naturally extend to the frequent cases of uncertain model selection. We show that model-free approaches can identify attentional modulation patterns, such as general alterations of the irregular shape of tuning curves, which cannot be captured by fitting stereotyped conventional models. Finally, by comparing datasets across different conditions, we demonstrate effects of attention that are cell- and even stimulus-specific. Based on these proofs-of-concept, we conclude that our data-driven methods can reliably extract relevant tuning information from neuronal recordings, including cells whose seemingly haphazard response curves defy conventional fitting approaches.

2.2 INTRODUCTION

Tuning curves represent a sensory neuron's response profile to a continuous stimuli parameter (such as orientation, direction of motion, or spatial frequency in the visual domain) and are an ubiquitous tool in neuroscience. In order to describe such selectivities in simple terms, tuning curves are commonly modeled by fitting suitable shape functions to the data, such as, for example, Gaussian distributions [Albright, 1984; Maldonado and Gray, 1996; McAdams and Maunsell, 1999a; Cronin et al., 2010], arbitrary polynomials [Etzold et al., 2004] or generic Fourier series [Wörgötter and Eysel, 1987] for orientation and direction tuning curves, or other smooth functions like splines [DeAngelis and Newsome, 1999] for non bell-shaped tuning profiles.

However, the measured tuning curve shapes are tremendously variable and they often deviate from the assumed reference shape [De Valois et al., 1982; Swindale, 1998; Amirikian and Georgopulos, 2000; Ringach et al., 2002]. More than thirty years ago, De Valois, Yund and Hepler already observed that no single function could adequately describe the tuned response of all cells [De Valois et al., 1982] they had recorded. A few years later, Swindale found that several models fit his data equally well and further questioned the existence of a single all-encompassing model, observing, moreover, that the preferred stimulus deduced from a fitted tuning curve depended on the chosen model [Swindale, 1998]. These and other studies thus manifest that the choice of a model to fit can affect the conclusions reached about tuning properties and their contextual modulations.

Here, we will first systematically investigate the problem of ambiguous model selection, highlighting the possible consequences of the choice of a "wrong" model. We will then introduce alternative methods which allow the extraction of features of interest directly from the measured neuronal responses, without the need of fitting any model to the empirical neuronal responses to a small subset of possible values from the continuous stimulus dimension. To illustrate the applicability and the heuristic power of our data-driven methods, we will carry out analyses of attentional modulation effects in single unit recordings in the middle temporal visual area (MT) of four rhesus monkeys, where neurons exhibit characteristic direction selective responses to moving visual stimuli [Dubner and Zeki, 1971; Born and Bradley, 2005]. We will consider responses to stimuli consisting of either one or two random dot patterns (RDPs) in the receptive field where the two RDPs could be either spatially separated or transparently superimposed. In these experimental paradigms, tuning curves are expected to display either one or two peaks. The expected effects of attention include gain modulations leading to changes in the amplitude of these response peaks. However, due to the high number of experimental conditions and the difficulty of the animal's behavioral

task, only relatively few trials could be recorded for each stimulus. This limited sampling, combined with the heterogeneity of response profiles, make the measured tuning curves very "noisy". The dataset, thus, besides its intrinsic interest, provides a perfect test-bed to reveal the drawbacks of fitting techniques and to benchmark alternative methods. We will discuss two complementary strategies in detail.

First, we will parse the trial-averaged responses of single neurons to obtain, through a set of algorithmic rules, a list of features characterizing the neuron's response profile. For instance, a set of rules will be used to estimate the direction of a cell's preferred stimulus, solely based on the data points themselves. Analogously, other rules will be used to capture into suitable index quantities generic variations of the average response profile of a cell across different experimental conditions (e.g. different types of stimuli, or targets of attention). As a shared prerequisite, all these rules must be able to operate by just receiving as input the set of average responses to each of the different stimuli from the small set used in the experiment. Although such an approach might seem too coarse compared to continuous interpolations, we will show an excellent agreement between the conclusions reached by feature extraction and fitting methods, whenever reliable model-based estimates can be derived. In addition, the same feature extraction rules will straightforwardly generalize even to the most irregular tuning curves, for which model fitting would be questionable. By the same token, the unbounded flexibility in rule design will allow the extraction of ad hoc features, revealing aspects of shape and shape change which elude a parameterization in terms of conventional fitted models.

Second, we will make use of the full information conveyed by the stochasticity of individual trials. For each neuron and for each different stimulus direction we will quantify the distribution of responses across trials, and compare them across different experimental conditions. This approach will show that attention frequently significantly modulates the response of a cell only for specific stimulus directions rather than a general modulation across the whole tuning profile.

Thus, through proof-of-concept analyses, we demonstrate the potential of data-driven methods for harnessing the heterogeneity of tuned responses. We prove that our approach robustly captures complex effects of attention, avoiding excessive reliance on illustrative well-behaved cases and circumventing the narrow constraints exerted by a fitting framework.

2.3 RESULTS

2.3.1 *The experiment: attentional influences on single-cell responses to composite stimuli*

To illustrate drawbacks with fitting approaches and test novel methodology on a concrete representative example, we focused on extracellular recordings of single neurons from area MT in the visual cortex of rhesus monkeys [Kozyrev et al., 2009; Lochte et al., 2009]. The aim of these recordings was to investigate how attention affects the tuning curves obtained by the simultaneous presentation of two directions of motion within the receptive field of a given neuron.

Neuronal responses were recorded under different conditions (Fig. 2.1A–E; see Methods for details). In the first experimental paradigm (Fig. 2.1A–B), random dot patterns (RDPs) moving within two *spatially separated* stationary apertures were used. They were sized and positioned for each recorded neuron to fit within its classical receptive field (RF). In a second experimental paradigm (Fig. 2.1C–D), both RDPs were fully overlapping. This single aperture contained two sets of random dots, moving transparently in two directions of motion and covered most of the RF. In all conditions a fixed angular separation of 120° between the two RDPs was used.

In each of the two paradigms, attention could be directed either to a fixation spot outside the recorded RF (condition *"afix"*), or to one of the RDPs inside the RF (condition *"ain"*). In addition, the study included conditions in which only one of the two RDPs was shown while attention was directed to the fixation spot (condition *"uni"*; see Fig. 2.1E for an explanatory cartoon of a spatially separated *"uni"* trial). Overall, data from six different experimental configurations (spatially separated afix, ain and uni; transparent afix, ain and uni) were analyzed (see also Materials and Methods).

Based on previous studies the *"uni"* experiments should generate unimodal tuning curves, while 120° separation between the spatially separated and transparent *"afix"* and *"ain"* stimuli should result in bimodal tuning curves [Treue et al., 2000], with the peaks occurring whenever one of the two stimulus components moved in the cell's preferred direction. Fig. 2.1F shows an example tuning curve from the spatially separated afix condition. Responses were sampled using 12 different directions across the full 360° range with an angular resolution of 30°. This sampling resolution is typical, as direction-tuning curves are most frequently assessed using 8 or 12 evenly spaced directions to account for the constraints of a behavioral paradigm where the number of trials that can be run is limited. Note that for plotting, stimulus directions were aligned for each cell such that the attended directino in the 240° stimulus was moving in the cell's preferred direction.

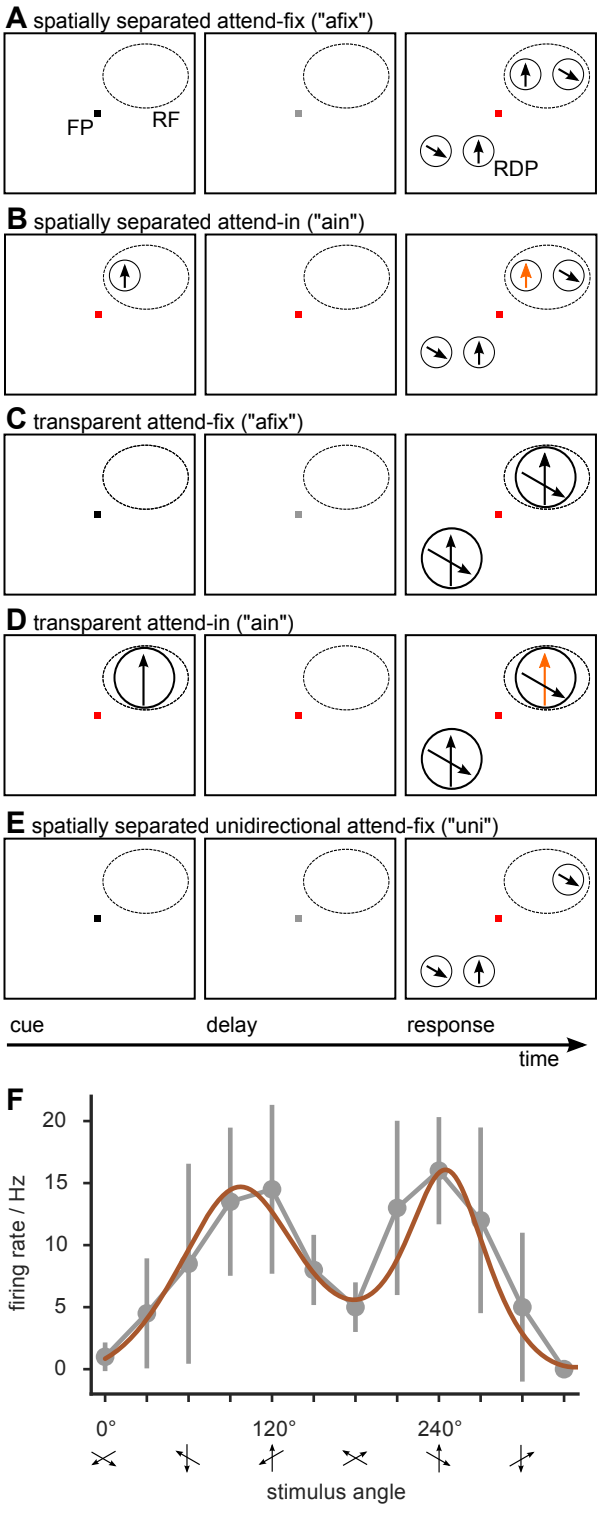


Figure 2.1: Attentional experiments. Direction selective responses of MT cells were measured using different direction combinations of stimuli and different attentional conditions. The stimuli in the receptive field (RF) of the recorded cell were either two random dot patterns (RDP) moving in directions 120° apart and placed in spatially separated (panels A–B) or overlapping (panels C–D) apertures or just one single (unidirectional) RDP (panel E). A cue instructed the monkey to attend to either: a luminance change of the fixation point (FP), in the attend-fix condition (afix, panels A and C) and single stimulus (uni, panel E) conditions; or to changes of the direction or velocity of the cued RDP (orange) in the RF, in the attend-in conditions (ain, panels B and D). The transparent uni condition was taken to be the cue-period of the ain condition (panel D). F: Example of a "well-behaved" tuning curve from the spatially separated paradigm in the afix condition. Gray circles denote trial-averaged firing rates and error bars their standard deviation. A sum-of-two-gaussians fit is also shown (brown). The stimulus directions are aligned for each cell, so that the attended direction corresponds to the preferred direction in the uni condition at 240° (see Materials and Methods for details).

2.3.2 *"Noisy" tuning curves: not one model to rule them all*

To model bell-shaped tuning curves, Gaussian curves are commonly used [Albright, 1984; Maldonado and Gray, 1996; McAdams and Maunsell, 1999a; Cronin et al., 2010] and they are usually wrapped, due to the circular nature of the fitted data [Mardia and Jupp, 1999]. This is exemplified in Fig. 2.1F where we fitted a sum of two Gaussians (brown curve) to a bimodal dataset. In this case, the fit looks adapted to the data, at least according to visual inspection. However, not all the cases are equally "well-behaved", as evinced, e.g., by comparing the tuning curve of Fig. 2.1F with a second example in Fig. 2.2A.

Two aspects need to be emphasized. First, the coarse sampling of the response profile because of the 30° separation between measurement points and second, the large trial-by-trial response variability to a given stimulus, visualized by the error bars. Indeed, the estimated coefficient of variation (i.e. the standard deviation in units of the mean) has a median of 38 % for the spatially separated paradigm and of 72 % for the transparent paradigm, as shown by Fig. 2.2B. Given this large uncertainty in the data it is thus not surprising that a "well fitted" linear combination of Gaussians lies within the error bars. However, within these large ranges of uncertainty, fitted curves obtained from model functions other than Gaussians could be accommodated as well, and there is no general *a priori* argument that a Gaussian model is the best suited model for such tuning curves.

To corroborate our intuition, we fitted eight different model functions to our data, testing for their compatibility with Gaussian and non-Gaussian shapes (Fig. 2.3). The used functions were a wrapped Gaussian (brown color) [Swindale, 1998], a wrapped Cauchy function (yellow) [Mardia and Jupp, 1999], a symmetric Beta function (violet) [Charalambides et al., 2000], a wrapped generalized bell-shaped membership function (pink) [Übeyli, 2009], a von Mises function [Swindale, 1998; Mardia and Jupp, 1999] (orange), and Fourier series of order 2 (red), 3 (blue) and 4 (green) [Wörgötter and Eysel, 1987]. Details on the used models are provided in the Materials and Methods section. In Fig. 2.3A, we show an example neuron for which all tested models provided reasonably looking fits at visual inspection (in the transparent attend fix condition).

We rigorously quantified goodness-of-fit by evaluating a score Q , giving the null hypothesis probability to observe by mere chance a sum of squared errors larger than in our fit. Small values of Q thus indicate poor fits (see Materials and Methods). We termed a fit "good" whenever $Q > 0.1$ but even lower values have been considered acceptable elsewhere [Press, 2002]. We computed goodness-of-fit for every model type and for every cell, and we assessed the fraction of cells for which each given model type provided a fit deemed to be good. Fig. 2.3B shows that—at least according to the Q measure—nearly

all cells could be well fitted by every model function, in every experimental condition. This statement still holds even when adopting more conservative thresholds for goodness-of-fit testing. As detailed in Fig. 2.4, even for threshold criteria as stringent as $Q > 0.7$, nearly all models provided good fits for more than 80% of the cells in most conditions.

Thus, for a majority of cells, goodness-of-fit alone was not enough to select the best model. We therefore adopted a model comparison approach and calculated the Akaike information criterion AIC [Akaike, 1974] for different models (see Materials and Methods). Smaller AIC values indicate better reproduction of the data by the model. Differences between AIC values for different models quantify how much information is lost describing the data with the model with a higher AIC value compared to the model with a smaller AIC value. Let, for a given cell, AIC_{min} be the minimum AIC value across all tested models. We then compute for each model m the relative information loss $\Delta AIC(m) = AIC(m) - AIC_{min}$. Hence, for the best model $\Delta AIC = 0$. However—as a conventional rule of thumb—models with a small increase ($\Delta AIC < 1$) should not be ruled out by model comparison but rather considered as equally good contenders for the "best fit" [Burnham and Anderson, 2002].

We show in Fig. 2.3C, the fraction of cells for which every tested model was evaluated as the "relative best", i.e. obtained a value of $\Delta AIC = 0$, as well as the fraction of cells in which it scored as an "equally good contender" with $\Delta AIC < 1$. The outcome was qualitatively similar across all conditions. There was not a single model which scored systematically at the top for all cells, but each of the eight tested model types scored $\Delta AIC = 0$ at least for a fraction of neurons. Even considering the softer criterion of $\Delta AIC < 1$, none of the models appeared to be good enough to be used to fit all cells. Interestingly, for all experimental paradigms, Gaussian fits only quite rarely scored as the relative best (6–15 %, depending on paradigm and condition). On the contrary, Fourier series fits were the more frequent winners (72–100 %, taken together Fourier series of all used orders).

Some studies suggest that when the number of available samples is small, the corrected Akaike information criterion AICc [Burnham and Anderson, 2002] (see Materials and Methods) should be preferred to the AIC. This AICc penalizes models with a larger number of parameters more then the AIC already does (i.e. it implements a sharper "Occam's razor"). We thus repeated the same analysis of Fig. 2.3C replacing the AIC with the AICc. Results are presented in Fig. 2.5A, which also presents the full statistical distributions of the observed AIC and AICc values (Figs. 2.5B-C). It turned out that the best model according to AICc was always one with few parameters: either four or five in the uni condition and five in the afix and ain conditions. This indicates that models with only few parameters are enough to

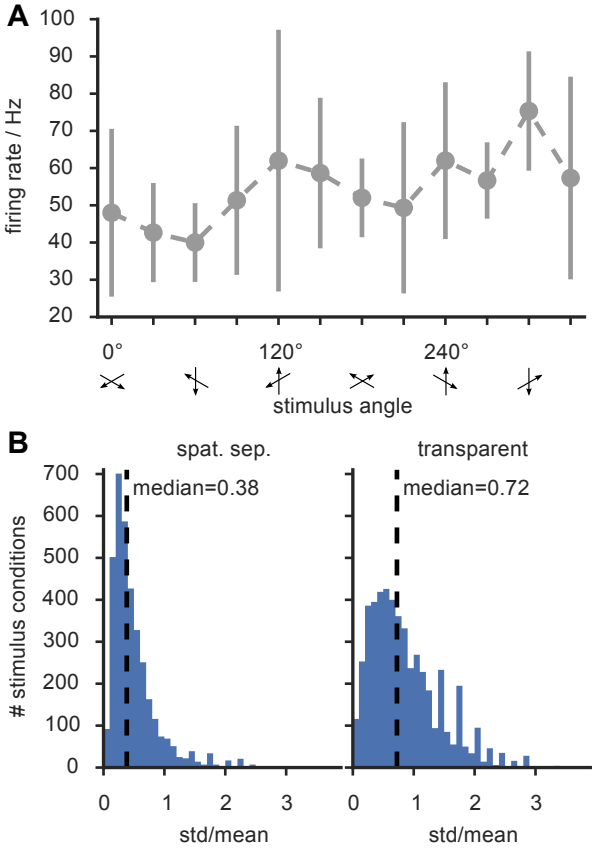


Figure 2.2: Many tuning curves are not "well-behaved". A: typical example of tuning curve from the spatially-separated afix condition (compare with Fig. 2.1F). The shape of the curve—including the position of the two peaks that should be elicited by the composite RDP stimulus—cannot reliably be inferred due to large error bars (std.). B: Histogram of estimated firing rate standard deviations (expressed in relative units, as ratios between std. and a matching mean), obtained by lumping together all stimulus directions and attentional conditions, for the spatially separated (left) and the transparent (right) paradigms. Both these histograms are strongly right-skewed, denoting the existence of cells with highly variable responses to certain stimuli.

describe our highly irregular data. While no clear winner emerged in the unidirectional paradigm, the second order Fourier series fit model clearly outperformed the other models in the bidirectional paradigms, due to its reasonable fidelity in rendering the shapes of the measured tuning curves, combined with a smaller number of parameters.

In summary, model comparisons show that no single model can fit all cells equally well and that more than one model should be used when looking for the continuous interpolation of discretely sampled noisy tuning curves. Among parametric models tested (using both the AIC and the AICc criterion), Fourier series (rather than the commonly used Gaussian curves) tended to be the relative best in a larger number of cases. This reflects the substantial diversity of tuning curve shapes present in our representative dataset, since Fourier series do not have a single shape, but can faithfully render very dissimilar circularly wrapped profiles.

2.3.3 *Intermezzo: how to compare the shapes of different parametric fits*

When fitting a model to data-points $(\mathbf{X}, \mathbf{Y})$ (such as a Gaussian profile $\mathbf{Y}(\mathbf{X}, a, b, c, d) = a \exp(-.5(\mathbf{X} - c)^2/b^2) + d$), the set of the parameters of the model (a, b, c, d in this case) form a vector of descriptors of

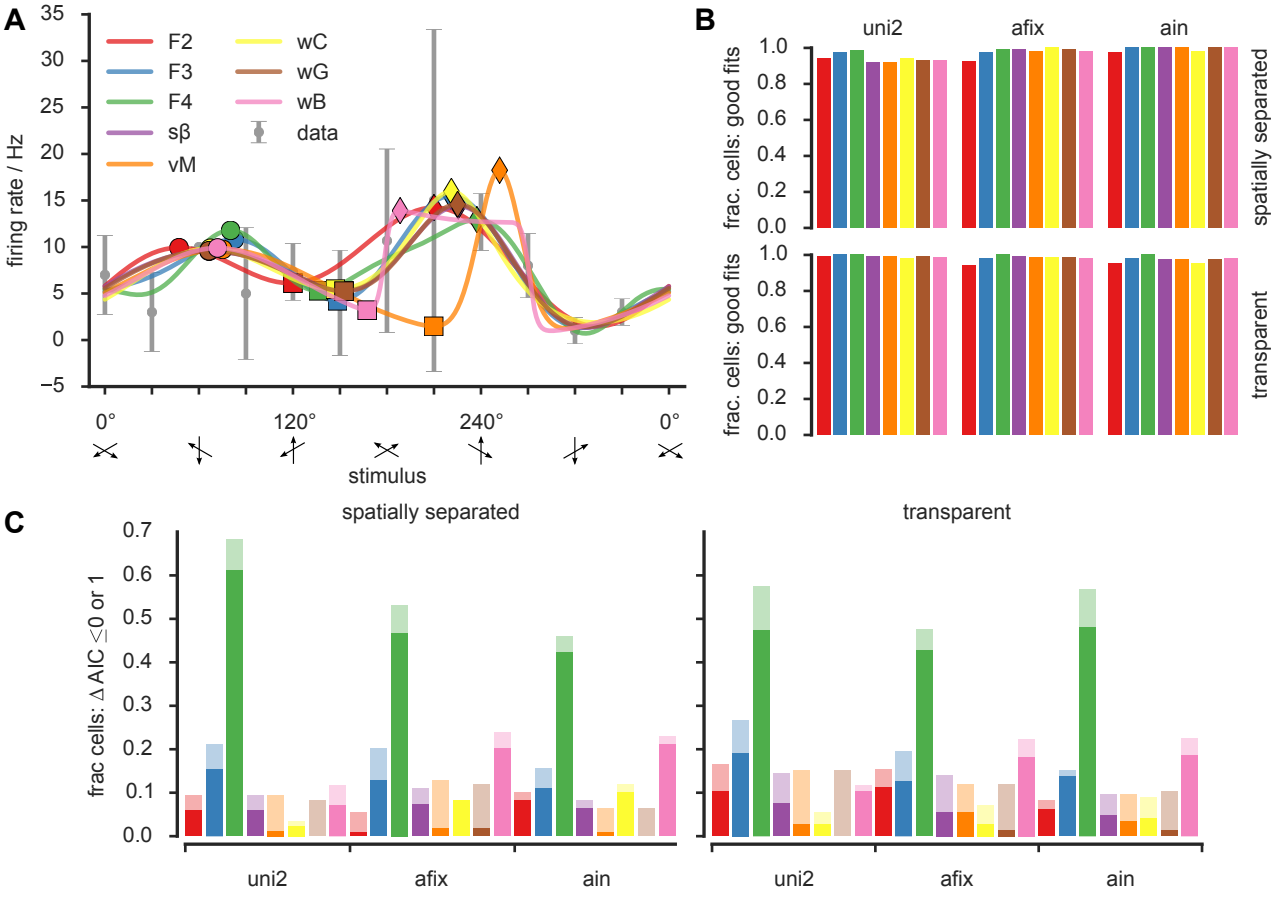


Figure 2.3: Many models are consistent with the data. A) Eight model functions were fitted to an example tuning curve (gray error bars denote std.) from the transparent afix condition. Due to the tuning curve's large error bars all models provided good fits even though they clearly differ. **B)** According to a goodness-of-fit score (see main text) all eight models provided good fits for almost all cells, independent of experimental condition and paradigm. **C)** The Akaike Information Criterion AIC was thus employed to select the best (ΔAIC=0) or at least close to best (ΔAIC$\leqslant$ 1) model for each cell. The fraction of cells for which each model constitutes the respective best or almost best model is illustrated with full and light bars. No model was chosen for all cells, still the most widely selected model was the fourth order Fourier series (F4). Both of these facts mirror the high heterogeneity in the data that is hard to capture in a single tuning curve shape. **Color code** red: 2nd order Fourier (F2); blue: 3rd order Fourier (F3); green: 4th order Fourier (F4); violet: symmetric Beta (sβ); orange - von Mises (vM); yellow: wrapped Cauchy (wC); brown: wrapped Gaussian (wG); pink: wrapped generalized bell-shaped membership function (wB).

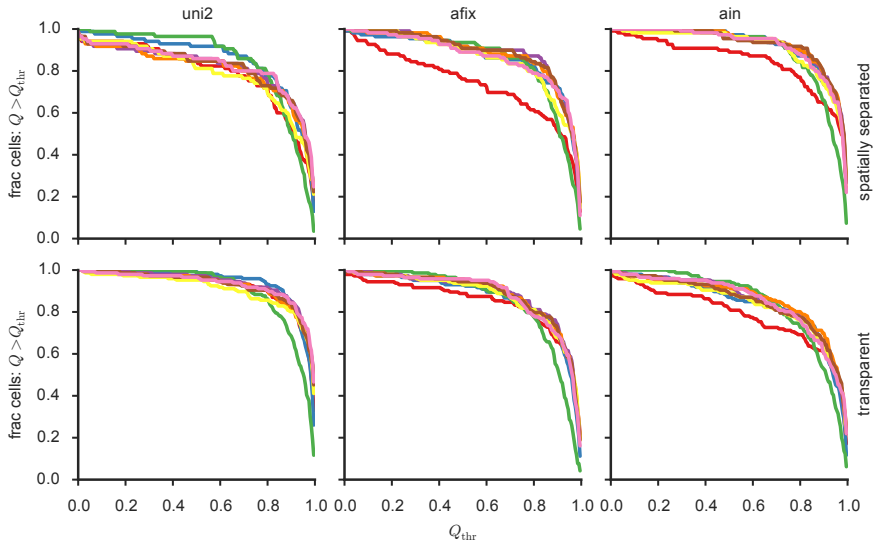


Figure 2.4: The threshold value for Q above which to accept a fit is not critical. Even for $Q_{thr} > 0.7$ more than 80 % of all cells lie above this threshold indicating a good fit.

the shape of the relation linking **X** and **Y**. In the case of the Gaussian, indeed, the parameter a represents the peak amplitude, b the peak width, c the x-position of the peak amplitude and d the baseline level. For this reason, the shape of tuning curves and its alterations have often been analyzed in terms of the values of the parameters of a fitted model [Treue and Trujillo, 1999; McAdams and Maunsell, 1999a]. However comparisons between fitted parameters are feasible only if a same underlying parametric model is used to fit tuning curves for all cells and in all experimental conditions. We have seen, on the contrary, that selecting a unique all-encompassing model to describe tuning might not be the best choice. Furthermore, fitted parameters do not have a direct geometric interpretation for every model. For instance, for Fourier series models, small changes of the internal parameters can lead to large changes in the resulting shape.

In order to compare between tuning curve shapes generated by different models in an intuitive way, we introduced generalized descriptive features that do not correspond to model parameters, but are extracted directly from the fitted curves through appropriate algorithmic rules (Fig. 2.6A).

The first step for the extraction of descriptive features is to select a certain number of points sampled along a fitted tuning curve profile and to note their coordinates $(\mathbf{X}, \mathbf{Y})$. Since the fitted profile is *continuous*, the number of selected points can be made arbitrarily large (unlike the number of actual measured data-points), but it is important to stress that the feature extraction approach that we introduce here always operates on a *discrete* set of points (a fact that will later allow us to apply it directly to the data-points themselves).

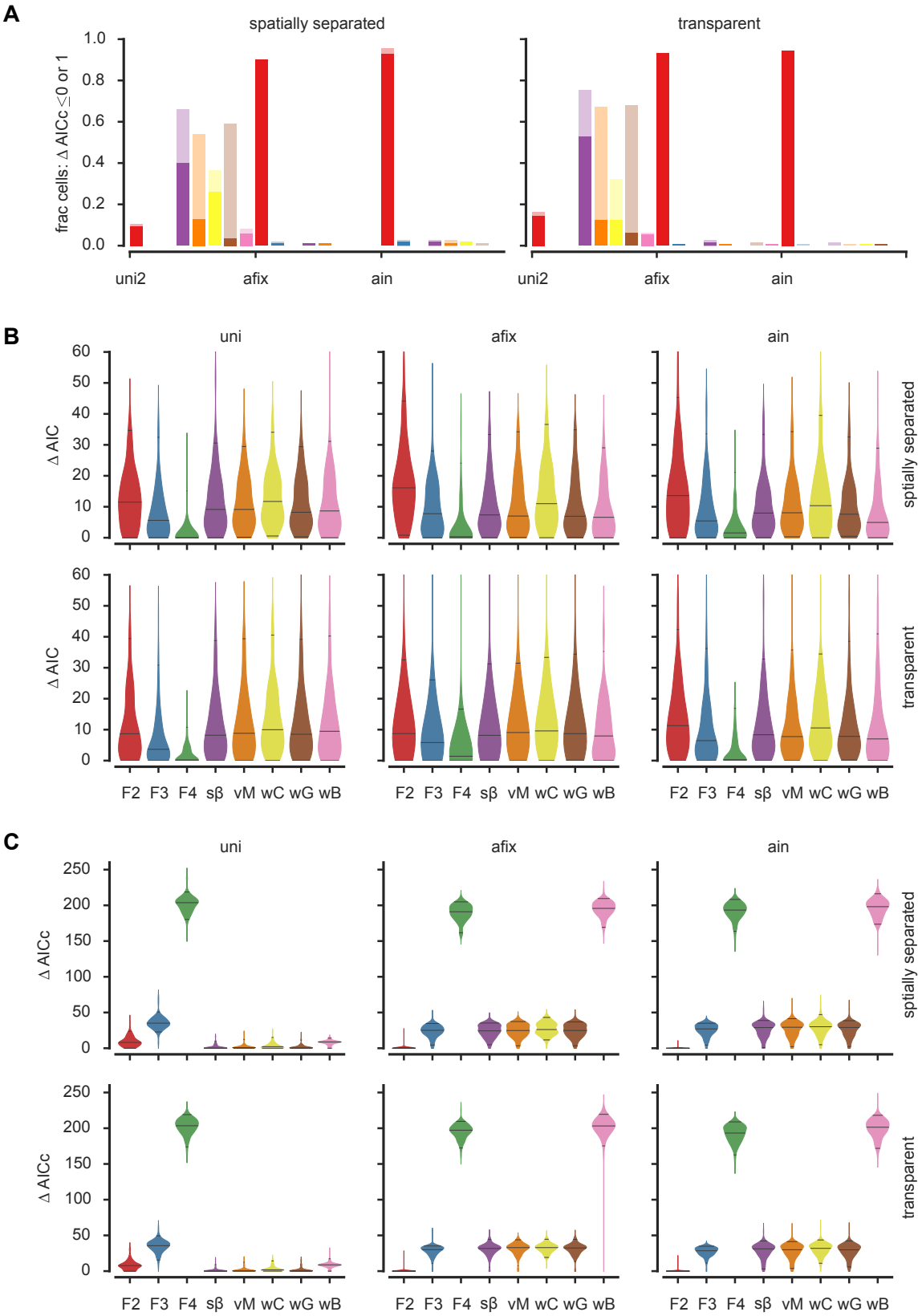


Figure 2.5: Model selection. A) Layout as in Fig. 2.3C, but showing ΔAICc instead of ΔAIC. Also for this criterion, none of the models is always selected, although for afix and ain conditions second order Fourier (F2) clearly performs best. **B,C)** Violinplots illustrating the distributions of ΔAIC (**B**) and ΔAICc (**C**).

The second step is to apply the desired feature extraction rule to the sampled points. Consider, for example, a unimodal tuning curve, for which we could extract a feature GlobalMaximum, by finding the maximum *Y* coordinate among the points sampled along the profile. Analogously, we could introduce a feature MaximumAngle, corresponding to the preferred direction, by finding the *X* coordinate of the point whose *Y* coordinate corresponds to the feature GlobalMaximum. These and other features can easily be generalized to the case of bimodal tuning curves (cf. Fig. 2.6B). In this case the evaluation rules would be modified to limit the search just to the right (left) half of the sampled points to identify peak positions and amplitude of the right (left) peak respectively. Note that these features are based on points sampled along the fitted profiles, rather than on the *parameters* of these fitted profiles, which they reflect therefore only in a highly indirect manner.

Describing the extraction of a peak amplitude and position as a feature extraction procedure could be seen as a (generally non-linear) projection method [Guyon and Elisseeff, 2003] through which a continuous—or, at least, finely discretized—tuning curve shape is converted into a vector with a much smaller finite number of entries. This might seem unnecessarily complex, however there are several good reasons for doing so.

First, computed features may but do not need to mirror classic model parameters. For example, while the feature MaximumAngle would be equivalent to the parameter c in the case of a Gaussian fit, Fourier series don't have any parameter directly reflecting peak position. More generally, all kind of convenient features can be designed, independent of the underlying model and tailored to describe specific shape aspects, such as, e.g., in a bimodal tuning curve, the position InnerMinimumAngle of the location of the lowest response between the two maxima —not necessarily centered between the two peaks—or the shape of the peak themselves whose OuterWidth and InnerWidth could differ, indicative of skewness or asymmetries. Table 2.1 gives a list of selected features, Tables A.1-A.4 provide the full list of features that we used as well as the detailed algorithm by which they were computed.

Second, as features are detached from the model itself, they allow a straightforward comparison of aspects of the tuning curves between parametrically incompatible models, which was a central aim in our study.

Third, as already anticipated, feature extraction rules can be applied directly—or with only minor modifications—to the observed data points themselves (Fig. 2.6C), without need of previously interpolating any continuous model curve. We will apply this direct approach, after completing our discussion of drawbacks inherent to

Table 2.1: Selected features for the description of tuning curve shape. The complete list of features that we used as well as the algorithm for their computation can be found in Tables A.1-A.4

Feature name	Description
GlobalMinimum [Hz]	lowest firing rate of all samples
GlobalMaximum [Hz]	highest firing rate of all samples
Maximum^{left,right} [Hz]	highest firing rate in left, right peak
PeakToPeak^{left,right} [Hz]	Maximum^{left,right} - GlobalMinimum
InnerMinimum [Hz]	minimum firing rate occuring at an angle between the left and right peak
InnerMinimumAngle [°]	stimulus angle at which firing rate is InnerMinimum
OuterMinimumAngle [°]	minimum firing rate right of right peak and left of left peak
GlobalMinimumAngle [°]	stimulus angle at which firing rate is GlobalMinimum
MaximumAngle^{left,right} [°]	angle of left, right peak at which firing rate assumes the value Maximum^{left,right}
InnerWidth^{left,right} [°]	InnerMinimumAngle - MaximumAngle^{left}
	MaximumAngle^{right} - InnerMinimumAngle
ΔInnerWidth [°]	InnerWidth^{right} - InnerWidth^{left} [°]
OuterWidth^{left,right} [°]	MaximumAngle^{left} - OuterMinimumAngle,
	OuterMinimumAngle - MaximumAngle^{right}
ΔOuterWidth [°]	OuterWidth^{right} - OuterWidth^{left}
Bandwidth$^{left,right}_{X\,\%}$ [°]	distance between left, right peak's angles at which baseline-subtracted firing rate drops below X % of PeakToPeak^{left,right}
Skewness^{right}	skewness of right peak
minusSkewness^{left}	negative of skewness of left peak
ΔSkewness	Skewness^{right} - minusSkewness^{left}
CircularVariance	circular variance of tuning

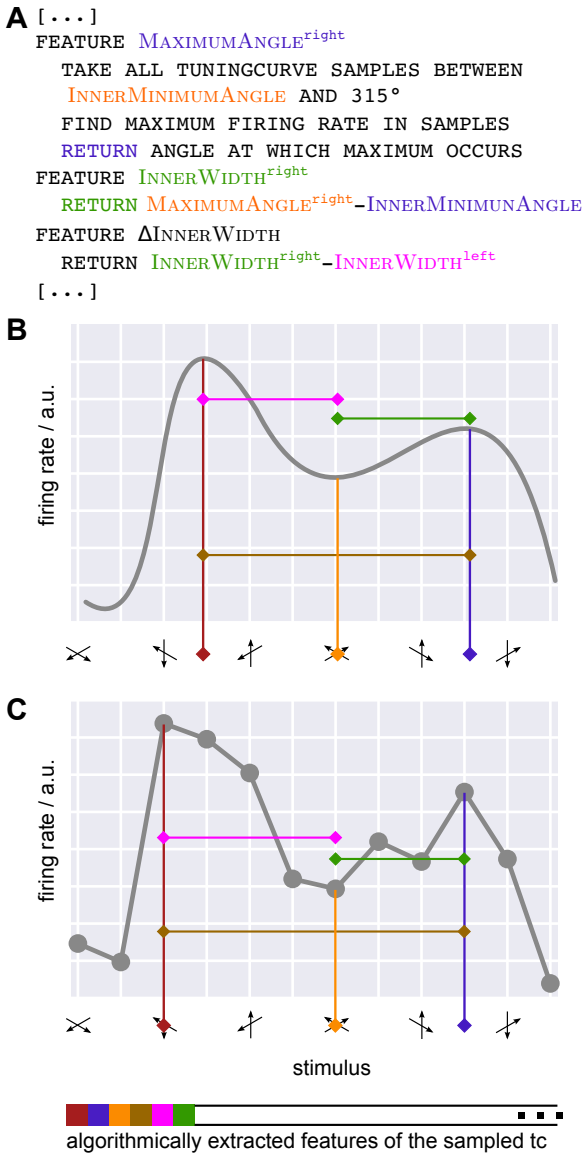


Figure 2.6: Alternative to fitting: algorithmic features. A) We describe all tuning curve properties of interest by algorithmically extracted features, as opposed to model parameters. The panel gives pseudocode for three selected features, illustrated also in panels B,C in corresponding colors. **B**) Feature extraction algorithms take as input only the sampled fitted tuning curve. Features are thus defined independent of a model function, allowing for their comparison between models. Furthermore, all aspects of tuning curves can be described by suitably chosen features such as Maximum^{left} (dark red), InnerMinimum (orange), Maximum^{right} (blue), (occuring at positions MaximumAngle^{left}, InnerMinimumAngle and MaximumAngle^{right}, respectively), InnerWidth^{left} (pink), InnerWidth^{right} (green). **C)** Due to their algorithmic nature feature extraction rules can equally well be applied directly to the coarse measured trial-averaged tuning curve. Thereby, tuning curve properties can be described and analyzed without referring to the fit.

model selection, by using feature extraction as a tool for comparing different fits.

2.3.4 *Different models can lead to different quantitative and qualitative results*

As indicated above, Fig. 2.3A shows an example for which all eight model functions could be reasonably well fitted to the measured neuronal data. We parsed these eight fitted curves based on a common set of feature extraction rules (see Tables 2.1 and A.1-A.4). We then compared the extracted shape features between the models. In Fig. 2.3A the position of the right and left peak and the inter-peak minimum are highlighted, respectively, by circle, diamond and square symbols. Across the model fits the positions of the left peak differed by up to

35° (10 % of 360°) and the corresponding firing rates by up to 2 Hz (15 % of the maximum trial-averaged firing rate for the afix condition in this cell); the position of the inter-peak minimum varied by 90° (25 %), their firing rates by 5 Hz (31 %); the positions of the right peak by 64° (18 %), their firing rates by 5 Hz (35 %). The example neuron of Fig. 2.3A thus shows substantial differences between the shape features inferred when applying different models.

These differences at the level of a single cell generalized to a large fraction of cells in the dataset, creating significant differences between models at the population level. For each of the eight models we computed the distributions of the values of different features across cells. In addition we extracted features from a ninth model, denoted the best model (bM). In this bM case, we selected the best among the eight tested fits, as indicated by the $\Delta\text{AIC} = 0$ criterion on a cell-by-cell basis, and for each cell we extracted features out of its specific best fit. Comparisons among the eight models and the ninth bM model are shown in Fig. 2.7A for nine selected features. In these nine matrices, a red entry indicates that the statistical comparison between the median values of the extracted features are significantly different between the two corresponding models (Kruskal-Wallis test, $p < 0.05$), while a blue entry denotes an agreement between the two models. Entries below and above the diagonal refer to feature comparisons for the spatially separated and the transparent paradigms, respectively.

For some features, the median value did not change significantly for most model comparisons (for example the global minimum—GlobalMinimum—in all conditions of the spatially separated paradigm). For other features, there were marked differences between Fourier series and bM on the one hand, and the remaining models on the other hand (for example the circular variance, CircularVariance in the uni condition; but also the feature GlobalMinimum for the transparent paradigm). For yet other features, almost every model yielded different results (for example: the differences between the skewnesses, ΔSkewness, of the left and right peak in afix and ain; the 75 %-bandwidth of the left peak, $\text{Bandwidth}^{\text{left}}_{75\,\%}$, in afix; and the 75 %-bandwidth in the uni condition $\text{Bandwidth}_{75\,\%}$). Importantly, in many cases there was a difference between bM and some of the other models.

Altogether, Fig. 2.7A demonstrates that the choice of model instead of another makes a difference, since it may lead to *quantitative* changes in the evaluated features. However, it would be even more severe if these quantitative changes in the evaluation of specific features led to divergent *qualitative* conclusions on the comparison between experimental conditions. For instance, when studying attentional modulation effects, it is important to compare tuning curve peak amplitudes among, e.g. an afix and an ain condition. Say, for illustration, that we estimate in the afix condition, a median Maximum feature value

of 30 Hz based on fits with the i-th model function, and a different median value of 40 Hz based on fits with the j-th model function. Let's suppose that the median values for the i-th and j-th model in the ain condition read 42 and 45 Hz, respectively. Besides quantitative differences, it might happen that based on the i-th model we conclude that attention has led to a significant increase of the MAXIMUM feature, but that this same comparison between the afix and ain conditions is not significant based on the j-th model. Thus, we would reach different conclusions on the effects of attention, depending on the chosen model. To check systematically for qualitative deviations in inter-condition comparisons, we performed comparisons between a large number of relevant feature pairs estimated from the nine different models (the eight tested model fit functions, supplemented by the bM model). We analyzed inter-model consistency for three different categories of comparisons: a feature from the spatially separate paradigm and the same feature from the transparent paradigm (this is possible for all defined features); a same feature taken from two conditions, e.g. uni vs afix, or afix vs ain (viable whenever the feature is defined for both conditions); or, two comparable features from a same condition (a list is given in Table A.5), e. g. $\text{MAXIMUM}^{\text{left}}$ vs $\text{MAXIMUM}^{\text{right}}$ for the peak firing rates of the two peaks of a bidirectional condition. For each tested feature pair we counted the number of fitted models for which the comparison was significant.

The histogram of these counts is reported, for different categories of feature pairs, in Fig. 2.7B. All histograms display a marked bimodal structure with two modes at the zero and nine counts values. These modes correspond, respectively, to the cases of complete agreement between models, i.e of a comparison which is never or always significant. Since both these two cases were the most frequent, there was a robust tendency toward a qualitative agreement between the conclusions of different models. Crucially though, the gap between the two modes of these histograms was not empty, but there were frequent cases in which the significance of comparisons between two features in a pair depended on the adopted model. Thus, for all these feature pairs, the choice of a specific model for fitting tuning curves would have led to qualitatively divergent conclusions about the effects of attention. In particular, the reached conclusion might differ from the one drawn from the bM model, the one which was constructed as optimal on a cell-by-cell basis.

This makes it advisable to always fit tuning curves based on a bM mixture of models. However, the bM approach is particularly cumbersome to calculate. Furthermore, it is ill-defined. Indeed, given the high heterogeneity in the data, it is plausible that adding even more models to the list of candidates among which to perform the bM choice would lead to further qualitative differences. It seems there-

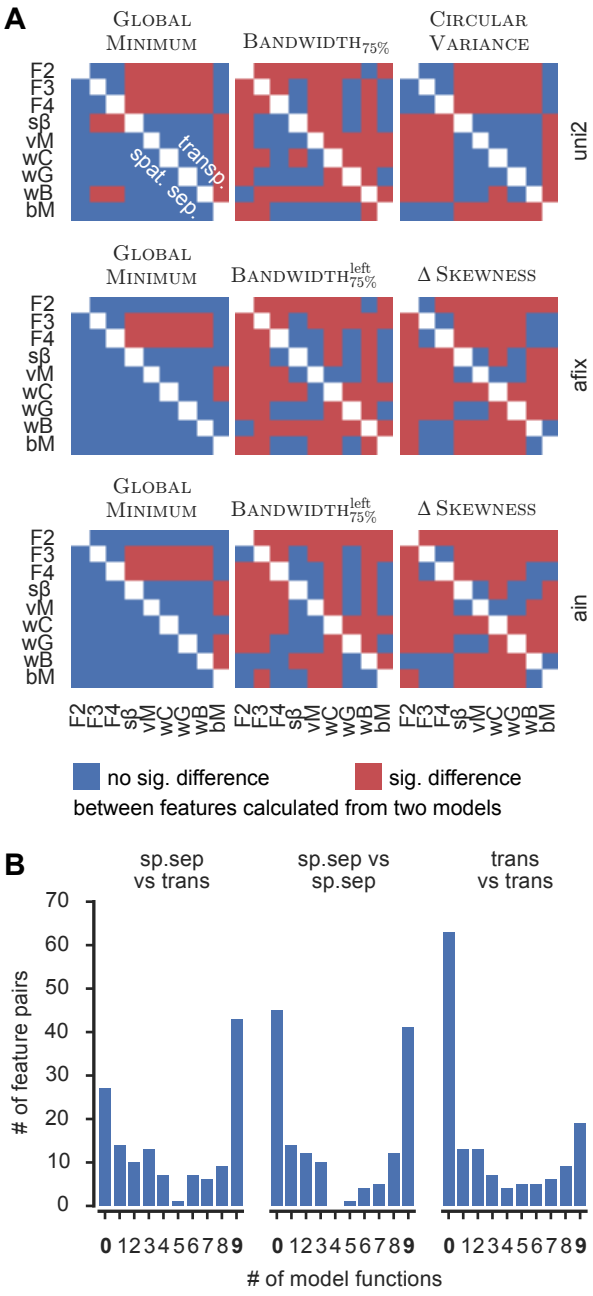


Figure 2.7: Effects found in the data depend on model. A) Nine features, measuring aspects of firing rate (first column), width (second column) and global shape (third column) in all three conditions were calculated for each cell on the basis of eight fitted models (model abbreviations are as in Fig. 2.3) and the "best model" (according to the ΔAIC=0 criterion, see main text; abbreviated "bM"). Red (blue) indicates a statistically significant (not sig.) difference between two models' values of that feature. Results from the spatially separated and transparent paradigm are plotted below and above the diagonal, respectively. The panels indicate that, in general, models disagree on the value of a feature, and, in particular, might contradict the optimal (bM) model. **B)** Histograms count for all feature pairs (depending on their category "sp.sep vs trans", "sp.sep vs sp.sep" or "trans vs trans") the number of model functions that find a significant difference ("effect") between the pair. While mostly all models agree (counts 0 and 9) there are also numerous cases in which the presence of an effect depends on the chosen model.

fore necessary to devise alternative strategies which completely avoid the questionable step of model selection itself.

2.3.5 *Feature extraction revisited: the direct method*

Rather than relying on the extraction of tuning parameters from fitted data as illustrated above, rules for feature extraction can be generalized to operate on the experimental data points themselves. The main difference between a fitted profile and the empirical data points is a coarse angular resolution of experimental measurements that might

potentially lead to a loss of precision of the extracted feature values. However, this is a quantitative, not a qualitative difference, that does not prevent the application of the rule, as illustrated by Fig. 2.6C. We therefore extracted shape-describing features directly from the data for all cells in all experimental conditions and compared them with matching features estimated from different model-based fits.

Fig. 2.8 depicts the results of this comparison and Table 2.2 reports selected feature values obtained from the direct method and the corresponding values from the bM model (see Tables A.6-A.7 for a comprehensive list). We first focus on qualitative differences between the direct and the model-based approaches (Fig. 2.8A), before delving into quantitative differences (Fig. 2.8B). Fig. 2.8A follows an approach similar to Fig. 2.7B, however we now built distinct histograms for feature pairs which are significantly different based on the direct method (blue histogram) and feature pairs which are not significantly different based on the direct method (green histogram). As in Fig. 2.7B, we counted the number of fitted models with significant with feature differences. The blue (green) histograms—peaking at the maximum (minimum) model value count—indicate that when the direct method found a feature comparison to be significant (not significant), the most frequent case was that all nine (none of the) tested models also reached the same conclusion. Concomitantly, the left (right) tails of these histograms fell off quickly.

Table 2.2: Median values and significant differences for selected features. Numbers (numbers in parentheses) denote the median value over the population of cells from the corresponding paradigm and condition, calculated with the direct method (best model "bM"). Significant differences between conditions or paradigms (abbreviated here as "sp" and "tr") are summarized (if present) in the line below the medians. $<, \ll, \lll$ denotes a p-value $< 0.05, 0.01, 0.001$, respectively, of a Kruskal-Wallis test applied to the feature values of the direct method. The condition listed before the relation sign had a smaller median than the one after it (if the medians were identical, means were compared). "NA" denotes features that were not defined for the corresponding condition.

	spatially separated			transparent		
Feature	uni	afix	ain	uni	afix	ain
GlobalMinimum	4.50 (4.14)	7.14 (6.78)	7.71 (8.09)	1.96 (1.86)	2.00 (1.53)	1.93 (1.44)
	sp uni $<$ sp afix, sp uni $\lll$ sp ain; tr uni $\ll$ sp uni, tr afix $\lll$ sp afix, tr ain $\lll$ sp ain					
GlobalMaximum	31.00 (30.62)	34.50 (35.15)	38.50 (37.87)	14.19 (15.01)	14.00 (12.94)	15.75 (15.93)
	sp uni $<$ sp ain; tr uni $\lll$ sp uni, tr afix $\lll$ sp afix, tr ain $\lll$ sp ain					
Maximum^{left}	NA (NA)	29.50 (27.77)	27.11 (27.89)	NA (NA)	11.00 (10.37)	12.22 (11.84)
	tr afix $\lll$ sp afix, tr ain $\lll$ sp ain					
Maximum^{right}	31.00 (30.62)	29.60 (30.34)	36.29 (35.76)	14.19 (15.01)	12.00 (11.57)	13.50 (12.80)
	tr uni $\lll$ sp uni, tr afix $\lll$ sp afix, tr ain $\lll$ sp ain					
PeakToPeak^{left}	NA (NA)	21.00 (20.89)	17.14 (16.69)	NA (NA)	9.08 (9.12)	10.00 (10.06)
	tr afix $\lll$ sp afix, tr ain $\lll$ sp ain					
PeakToPeak^{right}	25.50 (24.52)	22.90 (23.58)	26.60 (25.40)	11.50 (11.09)	9.50 (8.70)	10.00 (9.96)
	tr uni $\lll$ sp uni, tr afix $\lll$ sp afix, tr ain $\lll$ sp ain					
InnerMinimum	NA (NA)	16.00 (15.29)	18.80 (17.35)	NA (NA)	4.00 (4.12)	5.00 (3.68)
	tr afix $\lll$ sp afix, tr ain $\lll$ sp ain					
InnerMinimumAngle	NA (NA)	180.00 (174.00)	150.00 (165.50)	NA (NA)	180.00 (180.60)	180.00 (188.35)

Continued on next page

Table 2.2 – *Continued from previous page*

	spatially separated			transparent		
Feature	uni	afix	ain	uni	afix	ain
GlobalMinimumAngle	60.00 (67.30)	0.00 (12.80)	0.00 (18.10)	60.00 (52.10)	30.00 (26.40)	30.00 (21.10)
	sp afix $\lll$ sp uni, sp ain $\lll$ sp uni; tr afix $<$ tr uni, tr ain $\ll$ tr uni; sp afix $<$ tr afix					
MaximumAngle^{left}	NA (NA)	120.00 (115.80)	120.00 (120.10)	NA (NA)	120.00 (113.70)	120.00 (117.15)
MaximumAngle^{right}	240.00 (241.20)	240.00 (240.20)	240.00 (242.50)	240.00 (244.35)	240.00 (247.50)	240.00 (247.10)
	sp uni $<$ sp ain; tr uni $\ll$ tr ain; sp uni $\lll$ tr uni					
ΔInnerWidth	NA (NA)	0.00 (8.70)	30.00 (23.80)	NA (NA)	0.00 (7.00)	0.00 (3.45)
	sp afix $\ll$ sp ain; tr ain $\ll$ sp ain					
ΔOuterWidth	NA (NA)	30.00 (32.60)	0.00 (21.30)	NA (NA)	0.00 (2.10)	0.00 (4.65)
Bandwidth$^{left}_{75\,\%}$	NA (NA)	90.00 (53.10)	60.00 (50.50)	NA (NA)	60.00 (47.90)	60.00 (48.55)
	sp ain $<$ sp afix;					
Bandwidth$^{right}_{75\,\%}$	90.00 (61.60)	90.00 (54.80)	90.00 (70.20)	90.00 (60.90)	60.00 (50.00)	60.00 (50.40)
	sp afix $<$ sp uni, sp afix $\ll$ sp ain; tr afix $<$ tr uni; tr uni $<$ sp uni, tr afix $<$ sp afix, tr ain $\lll$ sp ain					
ΔSkewness	NA (NA)	0.11 (-0.09)	-0.13 (-0.36)	NA (NA)	0.58 (-0.10)	0.42 (0.00)
	sp ain $<$ sp afix; sp afix $\ll$ tr afix, sp ain $\lll$ tr ain					
CircularVariance	0.46 (0.47)	NA (NA)	NA (NA)	0.62 (0.59)	NA (NA)	NA (NA)
	sp uni $\lll$ tr uni					

Beyond the qualitative agreement, we also checked for quantitative agreements between features extracted by the direct and the model-based methods. We computed for each model i and for each feature F a z-score variable $z_i(F) = (\text{mean}(F)_i - \text{mean}(F)_{\text{direct}}) / \text{std}(F)_{\text{direct}}$. The distributions of these z-scores over all the different features, for each different model function are plotted in Fig. 2.8B. For different experimental conditions and for all the model functions, the z-score distributions are centered on zero, indicating quantitative agreement between the direct and model-based methods. For the spatially separated paradigm all models gave results particularly close to the "direct" method (the difference was smaller than one standard deviation in almost all cases, i.e. z lay between -1 and 1, for at least 95 % of the extracted features). A relatively weaker quantitative agreement was observed for the transparent paradigm. For this paradigm, the Fourier series and the von Mises models gave the best agreement to the direct method. However, even in the case of the wrapped Cauchy model, which gave the worst agreement with the direct method, 88 % of the features had z-scores between -1 and 1.

In conclusion, we observed a qualitative and quantitative agreement between the direct method and the tested models. But note that the direct method makes less assumptions on the expected shape of tuning curves and does not conceal their heterogeneity.

2.3.6 *The direct method in action: Tuning curve modulations*

After focusing on methodological aspects, we will now turn to the effects of different experimental paradigms, attentional conditions, and number of stimuli on the tuning. We will concentrate on a narrow selection of significant feature variations (Kruskal-Wallist test with $p < 0.05$) revealed by the direct method, performing comparisons between conditions both within the same experimental paradigm and between different paradigms (Table 2 and Fig. 2.9). Tables S8-S9 provide then a complete list of significantly different feature pairs evaluated with the direct method and the best model.

First, we evaluated tuning curves when one or two unattended stimuli were present in the receptive field, i.e. in the afix and uni conditions where attention was directed outside the receptive field (RF).

For the spatially separated paradigm, the peaks in the afix condition were smaller than in the uni condition. We monitored peak elevation over the baseline using the ad hoc engineered features PeakToPeak$^{\text{right}}$ and normalizedPeakToPeak$^{\text{right}}$ (analogous results hold for the left peak). These features quantify for each cell the variation between the right peak's maximum and the response's global minimum, which is normalized for the latter feature by the maximum firing rate in the uni condition (which is aligned, by convention, such that its

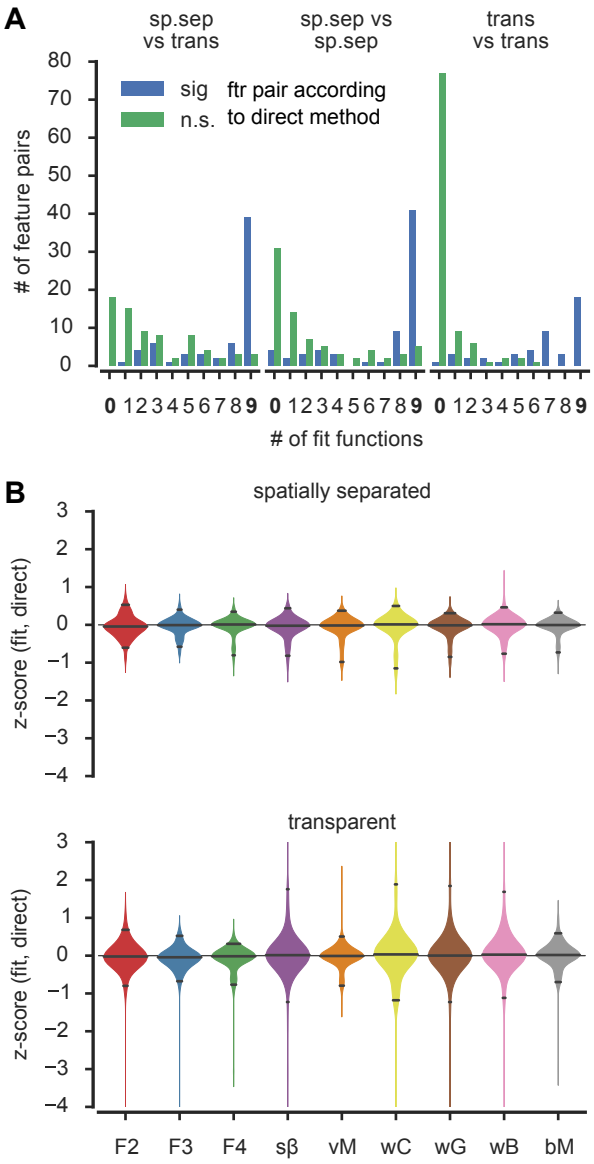


Figure 2.8: Direct method yields very similar results as fits. A) Layout is similar to figure 2.7B. Each bar from there was split into two, depending on if a considered feature pair was judged significantly different (blue) or not (green) when evaluated with the direct method. The panel illustrates a strong tendency to find a significant effect with either both the direct method and all nine models, or with neither the direct method and none of the models. **B)** z-scored quantitative differences between direct and fitted method's feature values is less than one standard deviation for almost all features independent of the model indicating a considerable quantitative agreement between the methods. Solid lines in violines mark 2.5 %, 50 % and 97.5 % quantiles. Color code and model abbreviations are as in Figs. 2.3 and 2.7

peak overlaps the right peak in the afix condition). This NORMALIZED-PEAKTOPEAK^{right} feature decreased from 0.86 in the uni condition to 0.68 in the bidirectional stimulus afix condition (we report, here and in the following, sample median values). The same trend also held in the transparent condition, where NORMALIZEDPEAKTOPEAK^{right} decreased from 0.84 to 0.55, when superposing a second stimulus within the RF.

As detailed in the Methods section, in the spatially separate paradigm the uni condition was measured with attention directed to the fixation spot, whereas in the transparent case it was taken to be the cue-period of the ain condition, that is, attention was directed to the stimulus during the measurement. Accordingly, features regarding uni conditions cannot be compared between the two paradigms. Please note that this

is due to the experimental design and does not limit the applicability of the method.

No significant differences were found between the amplitudes of the two peaks present in the afix condition, within both the spatially separated and the transparent paradigms, as monitored by the features $\text{MAXIMUM}^{\text{left}}$ and $\text{MAXIMUM}^{\text{right}}$.

We then evaluated the effects on tuning curves when deploying attention into the RF, i.e. in the ain condition.

We first monitored the emergence of amplitude differences between the two peaks of the tuning curve, computing, for instance, the feature $\Delta\text{PEAKTOPEAK}$, i.e. the difference between $\text{PEAKTOPEAK}^{\text{right}}$ and $\text{PEAKTOPEAK}^{\text{left}}$. In the spatially separated paradigm, there was a significant increase of the amplitude difference between the attended and unattended peaks, with $\Delta\text{PEAKTOPEAK}$ rising from 4 Hz in the afix up to 9 Hz in the ain condition, as a combined effect of a decrease of the left peak (NORMALIZEDPEAKTOPEAK changes from 0.6 to 0.5) and an increase of the right peak (from 0.68 to 0.76). In contrast, in the transparent paradigm, both peaks increased as an effect of attention (median NORMALIZEDPEAKTOPEAK changed from 0.51 to 0.64 and from 0.55 to 0.70, for the left and right peaks, respectively). The different effect of attention on peak amplitudes in the spatially separate and transparent paradigms is also illustrated by scatter plots of the NORMALIZEDPEAKTOPEAK in the ain vs the afix condition (Figs. 2.9A-B), where the cloud of points lies slightly below the diagonal for the unattended (left) peak in the spatially separated paradigm and slightly above it for the attended (right) peak in the spatially separated paradigm and for both peaks in the transparent paradigm.

Besides analyses of the amplitude and width of tuning curve peaks, the feature extraction approach allows the investigation of more general alterations in the response profile. The general shape of the attended peak differed between the spatially separate and the transparent paradigms. In particular the right peak was flatter (more platykurtic) in the spatially separate than in the transparent paradigm, as revealed by the median values of the feature $\text{BANDWIDTH}^{\text{right}}_{75\,\%}$ (see Table 1 for its definition), respectively, of 90° versus 60°. A usually unreported effect of attention is illustrated in Fig. 2.9C, where we analyze variations of the INNERWIDTH features, i.e. the (absolute values of the) angular distance between each peak and the minimum between the peaks. Usually similar for both peaks in the afix condition, differences between the INNERWIDTH feature for the attended and the unattended peaks may signal interaction phenomena, such as, e.g., a tendency for the attended peak to re-absorb the unattended peak. For the spatially separated paradigm, the difference between the left and right INNERWIDTH, given by the compound feature $\Delta\text{INNERWIDTH}$, increased significantly from 0° to 30°. On the contrary, no significant change was observed for the transparent paradigm. Note that the

values of ΔInnerWidth are discretely quantized due to the coarse angular resolution of our measurements and the lack of interpolation in the direct method.

Conversely, the firing rate at the minimum between the peaks, monitored by the ad hoc feature normalizedInnerMinimum, increased significantly from 0.32 to 0.41, only for the transparent paradigm. For the spatially separated paradigm a trend in the same direction was also present, but was not significant.

Together these effects denote different shape alteration typologies for the two paradigms, which represent an asymmetric expansion of the attended at the expense of the unattended peak in the spatially separated paradigm and a symmetric, growth of both peaks for the transparent paradigm, increasing responses in the inter-peak dip.

In conclusion, the composition of multiple stimuli and the attentional state affected general global aspects of tuning curves, inducing characteristic and significant patterns of changes. The direct method allowed to isolate known effects of attention without need to resorting to any fit, and identified different patterns of attentional modulation for the two tested experimental paradigms. It also cast light on usually neglected aspects of tuning curve shapes such as peak asymmetries, which experimental condition and attention can also modulate, besides the most commonly studied effects on peak amplitude and width.

2.3.7 *Cell- and stimulus-specific aspects of attentional modulation*

So far, the analyses have been based on responses averaged across all trials available for a given stimulus. While such an approach is very common, it is a simplification. Indeed, neuronal responses fluctuate strongly from trial to trial, which may be functionally relevant [Arieli et al., 1996; Chelaru and Dragoi, 2008; Padmanabhan and Urban, 2010; McDonnell and Ward, 2011]. We therefore compared the distributions of responses across different attentional conditions, in a cell-by-cell and stimulus angle-by-stimulus angle fashion.

The cartoon in Fig. 2.10A illustrates this approach for a single cell. In the plot, each dots corresponds to the response of the cell in a given trial. The two experimental conditions (e.g., afix vs ain in the spatially separated paradigm) are represented in different colors. Comparing responses between any two conditions for matching stimuli, the large trial-to-trial variability stands out, as evident by scanning vertically the clouds of colored dots for any given fixed position on the horizontal axis (stimulus configurations). Accordingly, for some stimuli, the trial ensembles of responses may be significantly different between the two conditions (for example, in the cartoon of Fig. 2.10A, at the 60° stimulus), while for other stimulus configurations, the trial ensembles will not (for example at the 240° stimulus, in the cartoon).

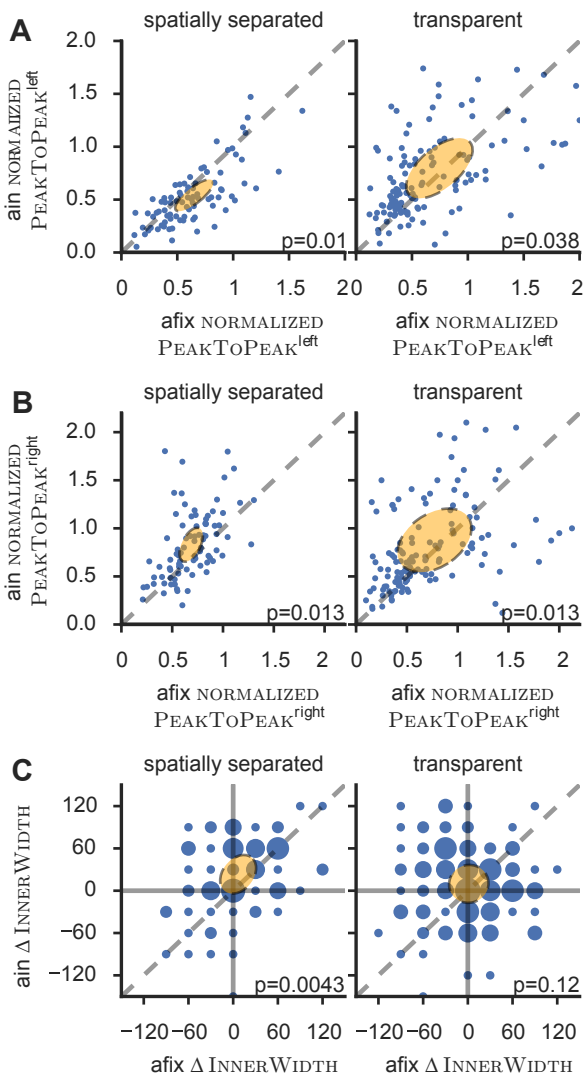


Figure 2.9: Effects of attention on tuning curves. Scatter plots of various features in the afix versus the ain condition. Orange error ellipses are centered on the mean feature values with half-axes corresponding to eigenvalues and -vectors of the feature-pair's covariance matrix. Some outliers were omitted for better visualization. The indicated p-value in each panel corresponds to a Kruskal-Wallis test. **A)** Attention decreased (increased) the left peak—as measured by the feature NORMALIZEDPEAKTOPEAK^{left}—in the spatially separated (transparent) paradigm. **B)** Attention increased the right peak in both paradigms according to the feature NORMALIZEDPEAKTOPEAK^{right}. **C)** Attention significantly increased the difference between left and right peak's inner width—ΔINNERWIDTH—only for the spatially separated paradigm. Size of circles in panel C illustrates density of points at each particular coordinate (note that values of ΔINNERWIDTH from the direct method are quantized in steps of 30°due to the design of experimentally used stimuli). Altogether panels indicate that attention asymetrically expanded the right at the expense of the left peak for the spatially separated paradigm, but increased both peaks similarly for the transparent paradigm.

For each given cell we identified subsets of stimulus directions for which attention caused a significant response modulation ($p < 0.05$, two-sample Kolmogorov-Smirnov test). We found that these stimulus subsets were highly cell-specific, with different cells exhibiting robustly significant attentional modulations at different angles, not necessarily concentrated in proximity of a specific attended direction, but scattered for each cell over the entire range of possible stimuli (see Fig. 2.11A). Correspondingly, we will refer to these stimulus-resolved significant differences, assessed at the single cell level as *specific effects*. While the statistical power of this analysis at the single cell level was limited by the small number of trials (see Fig. 2.11B), the population level showed narrow stimulus ranges for which the fraction of significant specific effects were larger.

Despite the irregularity and large inter-cell variability of the significance patterns of specific effects, some weak overlap at the population level could still be identified, identifying narrow stimulus ranges

for which the fraction of cells manifesting a significant specific effect were larger. Fig. 2.10B,C and D show how the response profile of a cell was altered when adding a second stimulus component within its RF, i.e. when going from a uni to an afix condition. The green histograms in Fig. 2.10B show the frequency distribution of the number of stimuli with significant changes, for the spatially separated and the transparent paradigm. In the spatially separated paradigm, 29 % of cells showed significant changes for four or more stimulus directions and 16 % of the cells did not show significant specific effects at any angle. The corresponding numbers in the transparent paradigm are 25 % and 23 %, respectively. This means that, for around a fifth of all cells, the entire tuning curves for the uni and afix conditions where statistically indistinguishable.

The cells for which the addition of the second stimulus caused no significant response modulation tended to be poorly tuned already in the uni condition (see Figs. 2.12A,B). On the other hand, some cells with equally poor tuning in the uni condition nevertheless displayed significant modulations of their firing rate when adding the second stimulus component.

Fig. 2.10B,E and F show cell- and stimulus-specific effects of attention, by comparing the afix and the ain conditions. The pink histograms in Fig 2.10B show the results of such a comparison for the spatially separated and the transparent paradigms. 7 % of the cells in the spatially separated paradigm and 10 % in the transparent paradigm showed significant specific effects of attentional modulation in four or more stimulus directions. The majority of cells in the spatially separated paradigm (60 %) and 50 % of the cells in the transparent paradigm showed a significant specific effect of attention for *at least* one stimulus. Note that most cells for both paradigms showed a clear tuning profile. Only 23 % (36 %) of the cells lacking any significant attentional modulation for the spatially-separated (transparent) paradigm were among the cells irresponsive to a second stimulus (Fig. 2.12C,D).

Figs. 2.10C-F also show the stimulus directions for which significant specific effects of the addition of a stimulus component or of the allocation of attention were more frequently observed. The blue bars indicate the fraction of cells without a significant response modulation for a given direction, while the green and pink bars indicate significant increases and decreases of responses, respectively. Not surprisingly, the most frequent significant response enhancement occurred for the direction bins around 90° in the spatially separate uni vs afix comparison, corresponding to the cases where the preferred or a similar direction was added to a single stimulus moving about 120° away from the preferred direction.

For significant attentional modulations between the afix and ain conditions response increases were more frequent than response de-

creases at all stimulus directions, except 90° and 120° in the spatially separate paradigm (where response decreases were more frequent) and 60° in the transparent paradigm (where response increases and decreases were equally rare). In addition, significant attentional modulations, occurred mostly for stimulus angles between 180° and 330° (in the spatially separate paradigm) and 90° and 270° (in the transparent paradigm). These observations on stimulus-specific effects are compatible with the previous observation based on the direct feature extraction method, in that, for the transparent paradigm, both peaks are positively enhanced, while, for the spatially separate paradigm, only the attended peak is boosted, but the unattended one tends to be depressed. In this way the analysis of specific effects can shed light on the cell-level genesis of global shape changes of the average tuning curves. The trial ensemble comparison analyses presented in this section manifest how significant gain modulations at the level of a cell population may arise from the contribution of specific effects which are only rarely significant at the single cell level.

2.4 DISCUSSION

We showed that the commonly used approach to analyze tuning curve by fitting an idealized model function to the trial-averaged data may be more problematic than usually thought. Indeed, when adopting a model-based approach, there is a clear danger to reach model-specific conclusions (cf. [Swindale, 1998]), which would not be confirmed by selecting different, equally viable models and which may be spurious. Here, going beyond model fitting and remaining within a purely data-driven framework, we extracted information about tuning and its modulations directly from the measured data points, through the application of rules for the extraction of suitable features. The high flexibility in feature design provided an antidote against over-constrained angles of view, which may be inherited by the adoption of narrow models.

Previous works already explored possible improvements on conventional least-squares fitting when dealing with noisy tuning curve data [Etzold et al., 2004; Cronin et al., 2010] and a wide alternative of possible functional models to fit, not only Gaussians [Albright, 1984; Maldonado and Gray, 1996; McAdams and Maunsell, 1999a; Cronin et al., 2010], but also typical circular statistics distributions [Swindale, 1998; Amirikian and Georgopulos, 2000], as well as Fourier series [Wörgötter and Eysel, 1987; Swindale, 1998]. Even the most sophisticated techniques, however, are not immune to the drawbacks inherent to any procedure assuming a common underlying statistical model. On the contrary, as already pointed out long ago [De Valois et al., 1982] and further confirmed by our analyses, the "best model" may

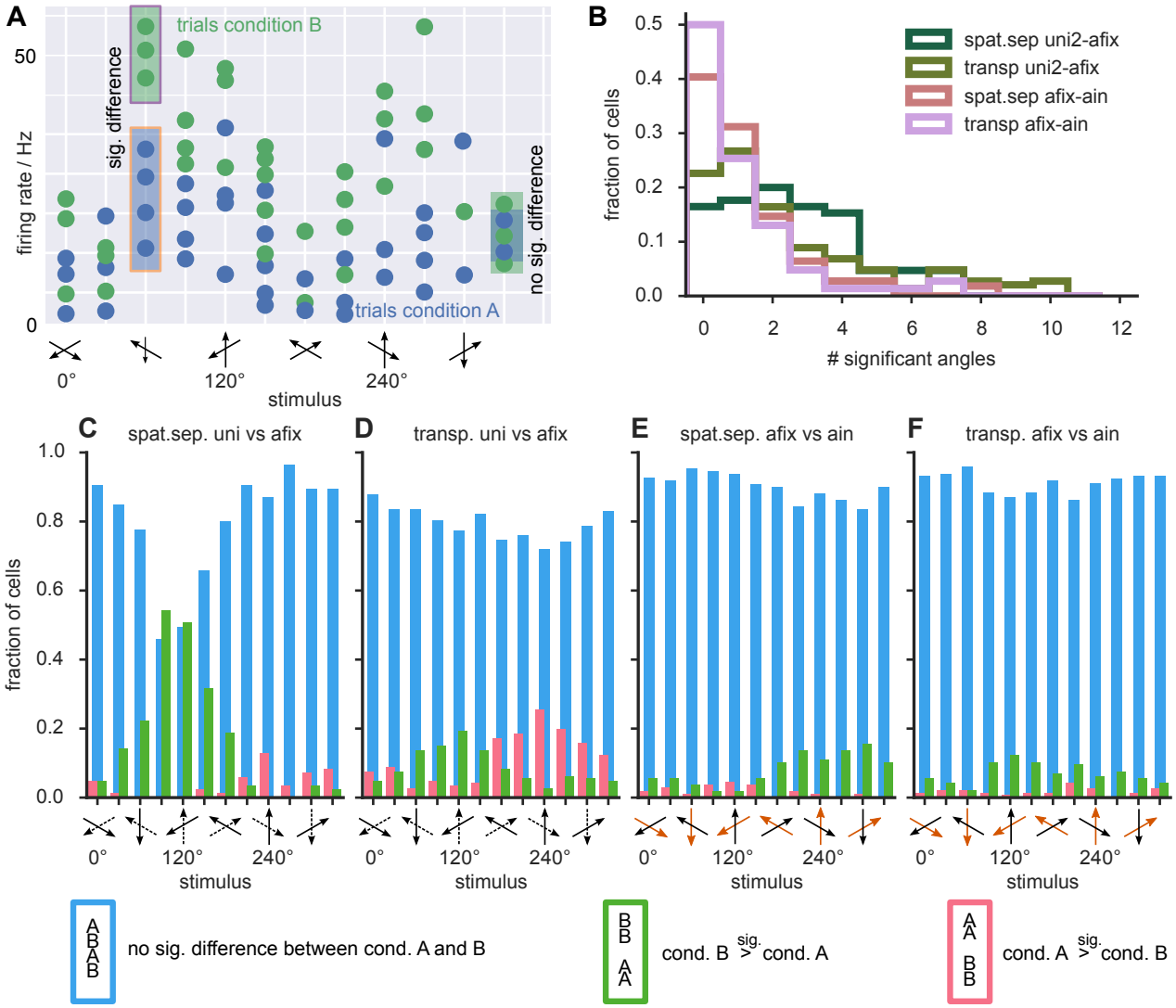


Figure 2.10: Effects of adding a second stimulus or attention to the receptive field. A) Each dot in this cartoon (not based on measured data) represents the observed spike count in one trial. For a given stimulus, spike count distributions can differ between experimental conditions either significantly (e. g. at 60°) or not (e. g. at 240°). **B)** Distribution of the proportion of cells with a significant difference between conditions for a given number of stimuli (maximum 12). The green histograms represent the two conditions where a second stimulus was added and pink histograms the conditions where attention was switched. **C-F)** Histograms show the stimulus-dependent fraction of cells with a non-significant response modulation (blue), a significant response enhancement (green) or response suppression (pink). The dotted and orange arrows along the x-axes in E and F indicate the RDP direction not present in the uni condition and the attended RDP in ain condition, respectively. Across the population a second stimulus tended to increase firing rates around 120°(C,D) and to decrease them around 240°. Attention asymmetrically affected the left and right peak in the spatially separated paradigm (E) whereas it symmetrically increased both peaks for the transparent paradigm (F). These stimulus-specific changes were compatible with the results of the direct method discussed in the text.

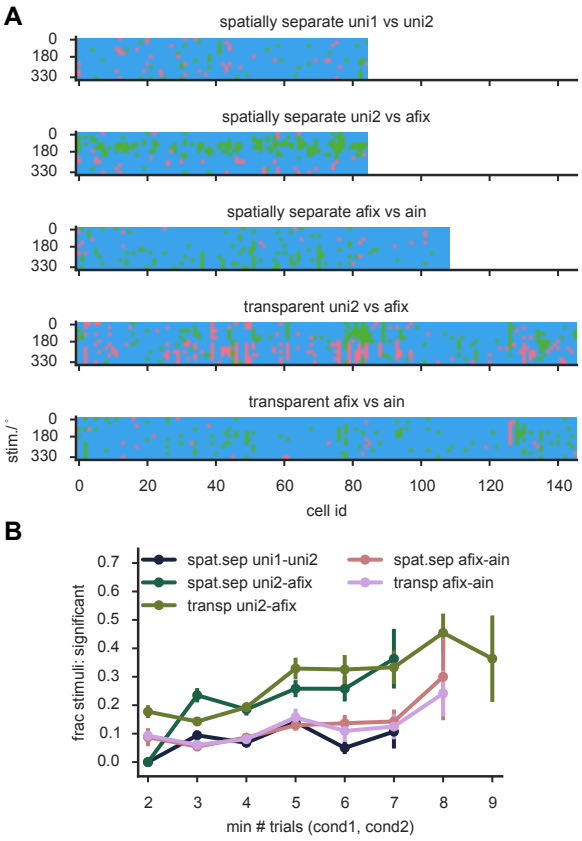


Figure 2.11: Analysis of the statistical power for specific effects. A) For each cell (x-axis) and stimulus (y-axis) colors indicate if there was a significant difference between the trials of the conditions marked in the title of each subplot. Colors are as in Fig. 2.10. **B)** We determined the impact of the number of trials available for our various conditions on the number of cells exhibiting significant changes between conditions. The plot shows how the number of cells exhibiting significant changes between conditions varied as a function of the number of trials included in the analysis (using the smaller of the two ensemble sizes for the x-value). The fraction of significant changes for all the tested condition changes showed a clear trend to increase with the number of included trials, possibly saturating when the number of trials reached about 8. Error bars denote standard-error of the mean. Note that points are only shown in this plot when we had a minimum of 10 samples, on average we had between 170 and 250 samples (depending on condition).

vary from cell to cell, making the problem of its selection conceptually ill-posed.

Yet, fitting still remains a practical tool to inspect tuning behavior in data, abstracting, at least as a first step, from the variety of tuning curve shapes present in any dataset. Although the tested models all give rise to bell-shaped tuning profiles, they differ in the geometry of the bells' flanks and these differences might be relevant for fine stimulus discrimination [Treue et al., 2000; Butts and Goldman, 2006]. Therefore, whenever fitting is used, one should carefully explore the entire *set* of candidate models, rather than of a single model, as a defense against excessive model bias. Results from our direct method itself could be included as well in the tested mix of analyses. A common set of features could then be extracted through a set of shared operational rules to systematically identify patterns of (or lack of) consistency between the diverse considered approaches. Comparisons between experimental conditions which are found to be significant only for a narrow subset of methods should then be looked at with suspicion, and confirmed by additional independent verifications.

The novelty of our data-driven approach is, however, more substantial than just providing yet another "model-less model". First, if the correct model function cannot be certified with certainty, much of the seemingly high precision achieved by model-based interpolation

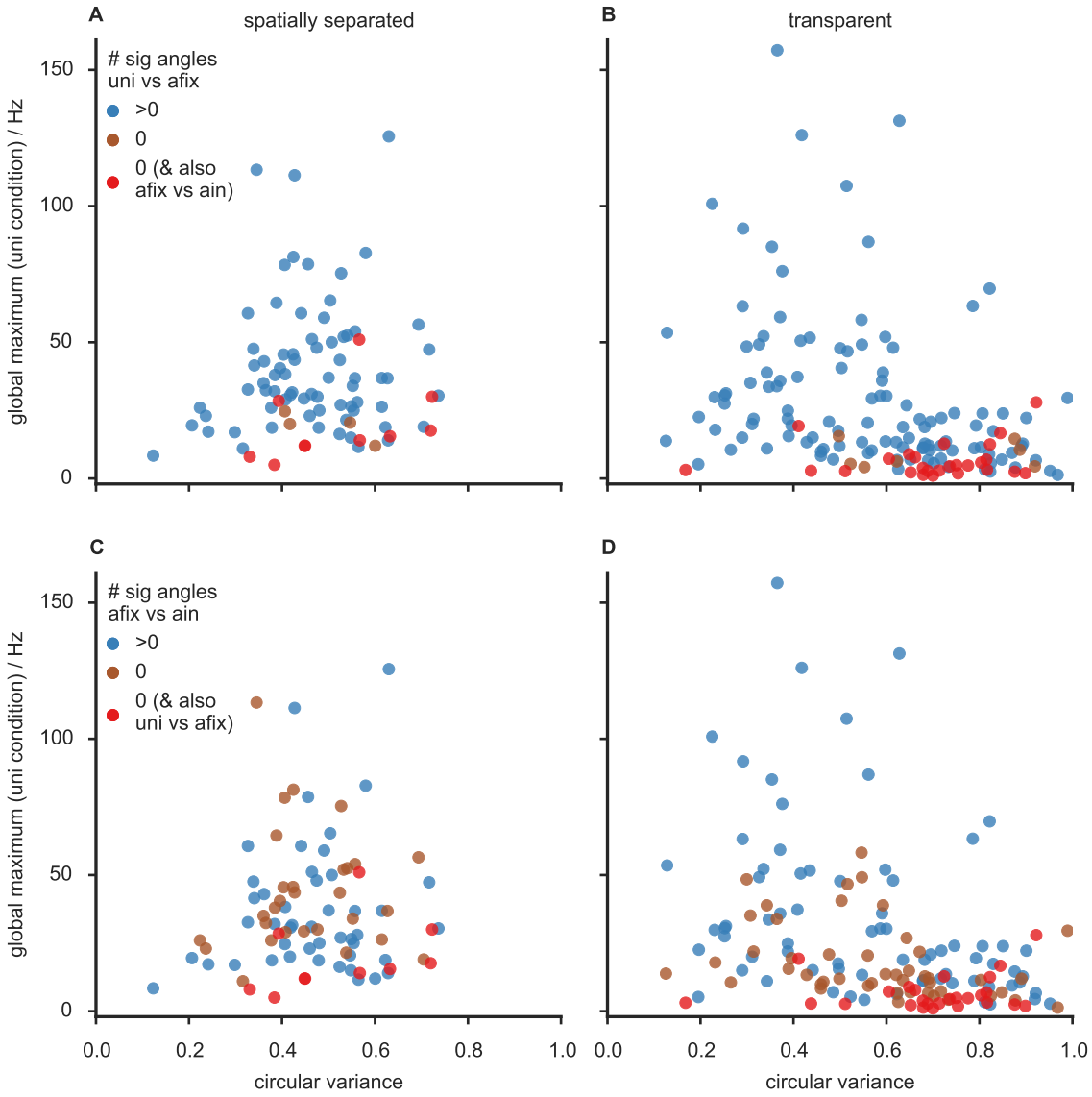


Figure 2.12: Cells not signficantly modulated by the addition of a second stimulus tended to be badly tuned but not vice versa whereas attentional modulation was unrelated to tuning. We compared trial ensembles for a given stimulus between conditions. Dots mark tuning properties of cells with at least one (blue) and zero (red, brown) significantly different stimuli between **A,B)** uni and afix condition, and **C,D)** afix and ain condition. Red dots mark cells without any significant change in both comparisons. These cells (A,B) tended to be badly tuned, but there were also equally badly tuned cells sensitive to this manipulation. Directing attention to the receptive field (C,D) had no clear relation to tuning properties.

may be just an illusion. Looking at the data in an agnostic and democratic manner, our data-driven methods could assess the statistical significance of attentional effects, strongly localized in both stimulus- and neuronal spaces. In particular they highlighted that only about 40 %-50 % of cells—similar to some previous reports of object-based attention in V1 [Roelfsema et al., 2004; Poort and Roelfsema, 2009]—were significantly modulated attention, among them some virtually unresponsive cells which would often be discarded in conventional model-based studies. It remains an open question whether these specific effects are an artifact due to the limited availability of information (too sparse sampling of stimuli, limited number of available trials, etc.) or if they can be related to the fine-scale synaptic structure of top-down inputs. Indeed, at the local circuit level there is evidence for an extreme functional specificity of wiring [Ko et al., 2011; DeBello et al., 2014] and the frontal eye field, one of the assumed source areas of attention [Moore, 2006], might provide not more than a couple of synapses to excitatory (but not inhibitory) neurons in V4 [Anderson et al., 2011]. In addition, models have shown that random and sparse recurrent network architectures are compatible with highly heterogeneous tuning curves [Battaglia and Hansel, 2011; Hansel and Vreeswijk, 2012]. In the context of the present study, it is enough to stress that such fine-grained specific attentional effects would remain hidden to any approach based on the fitting of a stereotyped smooth model to cell responses. Adopting a model-free characterization of neuronal responses may thus well be necessary to relate advances in connectomics with cell-level modulations of functional activation.

Another potential application in which data-driven approaches could prove to be qualitatively superior to model-based approaches is the study of how attention affects complex population codes [Saproo and Serences, 2010] of tuned responses. Indeed, by comparing trial ensembles of dozens of simultaneously recorded neurons previous studies already suggested that noise correlations were essential for the attentional performance enhancements [Cohen and Maunsell, 2009] and that feature attention is coordinated across hemispheres whereas spatial attention correlates only local groups of neurons [Cohen and Maunsell, 2011]. We, on the other hand, had only single cell recordings available, but they revealed a high degree of heterogeneity in tuning which may be functional, not merely reflecting noise, but carrying relevant information [Pouget et al., 1999; Seriès et al., 2004; Averbeck et al., 2006; Butts and Goldman, 2006; Chelaru and Dragoi, 2008; McDonnell and Ward, 2011; Padmanabhan and Urban, 2010]. In particular, such single-cell "weird" modulations may build up in a coordinated manner to give rise to population-level representations of the attended stimulus with a higher quality of encoding or with better and faster decodability properties [Tkačik et al., 2010, 2014]. Until now only very few studies have addressed the recording of the

tuned response of many cells simultaneously [Maynard et al., 1999; Stevenson et al., 2012; Cohen and Maunsell, 2009, 2011] but the fast pace of growth of the number of simultaneously recorded neurons [Stevenson and Kording, 2011] will certainly call for more detailed characterizations of tuned responses, such as the ones that our methods begin to provide.

Conventional model fitting methodology is restricted to the analysis of a model's parameters thereby potentially overlooking some features of the tuning with high discriminatory power. We have circumvented this problem in that we analyzed a set of features describing a wider range of aspects of the data. In the extreme case one could set up an all-encompassing feature library and programmatically mine for the most relevant ones. A possible drawback of massive feature libraries may be the feature selection analogue of over-fitting, i.e. the inevitability that some statistical comparison will appear to be spuriously significant just in virtue of multiple comparison issues. However, even this "data dredging" [Smith and Ebrahim, 2002] is legitimate when used as an explorative technique for the generation of hypotheses to be verified by further studies on independently acquired data-sets. As a matter of fact, with twenty-first-century neuroscience entering an age of "big data" and large-scale cooperation [Sejnowski et al., 2014], feature selection [Guyon and Elisseeff, 2003], in which features with optimized classification relevance are engineered in a (semi-)unsupervised manner, will increasingly become a method of choice for machine-augmented data-set parsing and knowledge discovery.

To conclude, although we are still far from understanding the intricate circuit mechanism through which attention influences information representation, routing and processing in the brain, we hope that our general methodology will assist the interpretation and inspire the design of future experiments necessary to advance this research endeavor.

2.5 METHODS

2.5.1 *Experimental procedures*

All experimental procedures were approved by the regional ani- mal welfare office and complied with relevant laws and insti- tutional guidelines.

Single-unit action potentials were recorded extracellularly from extrastriate cortical area MT of four male rhesus monkeys (Macaca mulatta), using two sets of covert attention tasks. Two of the animals were performing the "spatially separated" paradigm, the other two the "transparent" paradigm. For the duration of every trial the monkeys were required to maintain their gaze on a fixation point in the

middle of a computer monitor, placed at a viewing distance of 57 cm. While the animal maintained fixation, either one or two moving RDPs appeared in apertures in the receptive field (RF) of a given cell, as well as in the opposite hemifield outside the RF. In case of two RDPs the direction of the RDP in aperture 2 was always shifted clockwise from the direction of the other RDP by 120°. The direction of motion of the RDPs were varied in steps of 30° to obtain a tuning curve. The angular difference of 120° was selected as bi-directional tuning curves are expected to have two peaks in this case [Treue et al., 2000]. In the transparent paradigm the two RDPs were fully overlapping, crating just one aperture, covering most of the RF whereas in the spatially separate paradigm the two apertures were smaller and non-overlapping, but both still fully contained in the RF.

In the transparent condition there always existed just one aperture resulting in a single "uni" response profile. In the spatially separate paradigm, on the other hand, presenting the stimulus in either one of the two apertures gave rise to different responses, "uni1" and "uni2", where the latter refers to the condition in which the stimulus appeared in the to-be-attended aperture (in the ain condition). If not noted otherwise, we always analyzed the "uni2" condition in the data from the spatially separate paradigm, and for simplicity also refer to it as just "uni".

For the spatially separate unidirectional attend-fix condition (Fig. 1E) the monkeys were instructed to direct attention to the fixation spot, after a delay one RDP appeared in one of the two non-overlapping apertures and the monkey needed to detect a change of color of the fixation spot in order to receive a liquid reward. The spatially separate attend-fix condition (Fig. 1A) was similar, but RDPs were presented in both of the apertures. In the spatially separated attend-in condition (Fig. 1B) a RDP in one of the apertures was presented as a cue (of 500 or 600 msec duration), indicating to the monkey the location and the motion direction of a stimulus to be attended in the course of the trial. After a delay (800 ms) RDPs appeared in both apertures, and the monkey had to detect a transient change of motion velocity in the cued aperture at a random time point till maximally 2.5 s after the stimulus onset while ignoring possible changes in the other (distracting) RDPs.

The transparent attend-fix (Fig. 1C) and transparent attend-in (Fig. 1D) differed from the corresponding spatially separate conditions only in that the two apertures in which the RDPs were presented overlapped.

We generally analyzed data from the response period, which was defined as the time window 200-700 msec after onset of the stimulus in the RF. However, as no distinct unidirectional condition was recorded for the transparent paradigm, we used the cue period of the attend-in condition (50-500 msec after the cue onset) as a proxy

for the uni condition in this case. That means that in the transparent uni condition attention was directed to the stimulus, whereas in the spatially separate uni condition attention was directed to the fixation spot. Accordingly, the uni conditions cannot be directly compared between the two paradigms.

We had 109 and 146 cells in the spatially separate and transparent paradigm, respectively. Uni conditions were recorded for 85 out of the 109 cells in the spatially separate paradigm. In 3 cells of the transparent paradigm the afix rates were not recorded and, therefore, our analysis disregarded the afix condition of those cells.

2.5.2 *Tuning data pre-processing*

Data analysis was performed using custom-written software in Python (available on request). We did not perform spike-density estimation, but all analyses of tuning responses were based on raw firing rates, either averaged over trials (for model fitting and data-driven feature extraction) or estimated within each trial independently (for trial ensemble comparisons). Cells were included in the analysis only if at least two trials were available for every recorded condition. For some cells of the spatially separated paradigm no uni conditions were recorded. These cells were generally included in the analysis and exempted only in calculations concerning the uni condition.

All tuning curves were conventionally aligned, such that the maximum firing rate of the uni condition corresponded to the angular coordinate 240°. Whenever uni conditions had not been recorded (this was the case for some cells of the spatially separated paradigm) the angular position of the maximum firing rate of the right peak in the afix condition was defined to be 240°.

2.5.3 *Fitted models*

To analyse tuning curves we fitted several model functions to the trial-averaged firing response data. Each fit to each specific cell was fully determined by the chosen parametric model and by a vector $\vec{p}$ of model parameters. The wrapped Gaussian (wG) was given by:

$$w\mathrm{G}(\theta, \vec{p}) = a \sum_{i=-N}^{N} \exp\left(-\frac{1}{2}\left(\frac{\theta - c + 360i}{b}\right)^2\right) + d$$

where $\vec{p} = (a, b, c, d)$. We always assumed $N = 4$ for wrapping.

The wrapped Cauchy (wC) function was given by:

$$w\mathrm{C}(\theta, \vec{p}) = a\frac{\sinh(b)}{\cosh(b) - \cos(\Omega(\theta - c))} + d$$

where $\vec{p} = (a, b, c, d)$ and $\Omega = 2\pi/360$.

The (modified) von Mises function (νM) was given by:

$$\nu M(\theta, \vec{p}) = a\left(e^{k\cos\left(\frac{2\pi}{360}(\theta-c)\right)} - e^{-k}\right)/(e^{k} - e^{-k}) + d$$

where $\vec{p} = (a, k, c, d)$.

The symmetric Beta function ($s\beta$) was given by:

$$s\beta(\theta, \vec{p}) = a(4x(1-x))^{b} + d$$

where $x = (2\pi/360(\theta - c) + \pi)/(2\pi) \mod 1$ and $\vec{p} = (a, b, c, d)$.

The wrapped generalized bell-shaped membership function (wB) was given by

$$wB(\theta, \vec{p}) = a\left(\sum_{i=-N}^{N} \frac{1}{1 + \left|\frac{\theta - c + 360i}{b}\right|^{2s}} - \beta\right)/(\alpha - \beta) + d$$

where $\alpha = \sum_{i=-N}^{N} 1/\left(1 + \left|\frac{360i}{b}\right|^{2s}\right)$, $\beta = \sum_{i=-N}^{N} 1/\left(1 + \left|\frac{180+360i}{b}\right|^{2s}\right)$, $\vec{p} = (a, b, c, d, s)$. We always assumed $N = 4$ for wrapping.

All these functions are illustrated in Fig. 2.3A. In their basic form, they give rise to unimodal tuning profiles, as in the uni condition. To fit bimodal responses to composite stimuli, in the afix and ain conditions, we used a sum of two (identical) model functions, i.e.:

$$f(\theta, \vec{p}) = g(\theta, \vec{p}_1) + g(\theta, \vec{p}_2)$$

where g is either one of $wG, wC, s\beta, wB$ and the total parameter vector is $\vec{p} = (\vec{p}_1, \vec{p}_2)$. There was some redundancy between the parameter sets for the two peaks. For the first four functions $p_1 = (a_1, b_1, c_1, d/2)$ and $p_2 = (a_2, b_c, c_2, d/2)$, while, for the wB model, both $\vec{p}_1$ and $\vec{p}_2$ contained an additional component, s_1 and s_2, respectively. Hence, f had overall seven ($wG, wC, \nu M, s\beta$) or nine (wB) free parameters.

In addition we also fitted Fourier series of order $n = 2, 3, 4$, given by:

$$Fn(\theta, \vec{p}) = a_0 + \sum_{i=1}^{n} (a_i \cos(\Omega\theta) + b_i \sin(\Omega\theta))$$

where $\Omega = 2\pi/360$ and $\vec{p} = (a_0, a_1, b_1, \ldots, a_n, b_n)$. These Fourier series were fully determined by five, seven or nine parameters respectively, depending on their order $n = 2, 3, 4$. Fourier fits of unimodal or bimodal tuning curves shared a common functional form.

2.5.4 *Fitting methods*

We used standard weighted non-linear least square fitting, relying on routines within Python's SciPy package (`http://www.scipy.org`, sequential quadratic programming) for the minimization of the χ^2

statistics. The applied initial conditions and boundaries therefore are listed in Table A.10. An exception was given by Fourier series. Let p_i be the ith component of the Fourier series' parameter vector $\vec{p} = (a_0, a_1, b_1, \ldots, a_n, b_n)$ and $X_i(\theta)$ be the ith compoment of $(1, \cos(\Omega\theta), \sin(\Omega\theta), \ldots, \cos(n\Omega$ Then, an exact analytical solution to the least squares problem exists, which can be straightforwardly derived to be:

$$\vec{p} = \sum_{m=1}^{2n+1} \left(\frac{\vec{U}_{(m)} \cdot \vec{b}}{w_m} \right) \vec{V}_{(m)}$$

where $\vec{b}$ has components $b_i = y_i/\sigma_i$, $\vec{U}_{(m)}, \vec{V}_{(m)}$ denote the mth column of U and V, respectively. The matrices U, V and w_m form the singular value composition of matrix A, s.t. $A = U\mathrm{diag}(w_1, \ldots, w_{2n+1})V^T$. The matrix A, in turn, has components $A_{im} = X_m(\theta_i)/\sigma_i$.

To quantify goodness-of-fit we also used a standard framework, as laid out in [Press, 2002]. We assume that measurement errors in y_i are normally distributed. For model functions that are linear in their parameters —note, that in our library of models, this assumption holds only for the Fourier series Fn—, the null hypothesis probability that the sum of squared errors is equal or larger than the observed χ^2 is given by $Q(K - (2n+1)/2, \chi^2/2)$ where $Q(a, x) = \int_x^\infty e^{-t} t^{a-1} dt/\Gamma(a)$ is the incomplete gamma function and K is the number of independent samples (here, $K = 12$ tested stimulus directions). We use this quantity Q as measure for goodness of fit. If the probability Q is $\leqslant$ 10% we term the quality of the fit "bad", otherwise we cannot rule out the hypothesis that the fit is an appropriate statistical models for our observations. For general non-linear models—for which the sum of squared errors cannot be expected to follow a conventional χ^2 distribution—we evaluated approximately the goodness-of-fit Q statistics through a Montecarlo resampling approach (10000 replicas, cf. [Press, 2002] for details).

2.5.5 *Model selection*

We performed model selection based on the Akaike information criterion (AIC) [Burnham and Anderson, 2002, 2004]. The information-theoretic quantity AIC gives the expected increase in uncertainty when using a certain model to describe the data rather than the "true" model. It can be computed from the sum of squared errors in the least-squares procedure according to

$$\mathrm{AIC} = K \log \left(\sum_{i=1}^{K} (y_i - f(\theta_i, \vec{p}))^2 / K \right) + 2M$$

where $K = 12$ is once again the number of independent samples and M is the number of free parameters of the model function. Importantly, this formula is only valid in the limit of large K. Some studies

[Burnham and Anderson, 2002, 2004] therefore recommend to use a correction factor in the case in which $K/M \lessapprox 40$. This corrected Akaike information criterion reads:

$$\mathrm{AICc} = \mathrm{AIC} + \frac{2M(M+1)}{K-M-1}$$

Such AICc converges to AIC for large K and mainly differs for it by applying a stronger penalty to models with larger number of free parameters.

2.5.6 *Features*

Given a tuning curve $tc(\theta)$—either as a discretized version of a fitted tuning profile, or directly by the vector of empirically observed average responses to stimuli with different directions—we characterized its shape in a non-parametric manner by calculating general *features*. A feature F is any map of the graph of the tuning curve $tc(\theta)$ onto some scalar number $F : tc \mapsto \mathbb{R}$. Examples of features are the maximum of the tuning curve or the preferred direction (see Fig. 2.6 for an illustration). The complete list of features that we used is given in Tables A.1-A.4. Although we did not perform any feature clustering or redundancy elimination through e.g. factor analysis, willing in reality to maintain a library of features as wide as possible, we verified that most of the features bear complementary information, as hinted to by a sample distribution of pairwise correlations between feature values strongly peaked around zero (not shown).

Angular features were measured in degrees, ranging from $0°$ to $360°$ (with the exception of the feature GlobalMinimumAngle, for which we used the range $-120° \leqslant \text{GlobalMinimumAngle} \leqslant 240°$) and coarsely quantized at just $K = 12$ equally spaced angular values. For each feature measured in units of Hz we also computed a normalized counterpart, denoted with the prefix NORMALIZED, by dividing the feature value by the global maximum firing rate found in the uni condition of the cell (if available).

We run statistical tests between *feature pairs* to search for effects of changes of experimental condition (transparent vs spatially separated, afix vs ain, etc.). For each of the two compared conditions we evaluated the values of the tested feature for each cell. We then performed two-way Kruskal-Wallis testing and dubbed a comparison as significant, whenever the p-value of this Kruskal-Wallis test was smaller than 0.05. All found significantly different feature pairs are listed in Tables A.8 and A.9, an excerpt in Table 2.2. The feature pairs reported herein are also the ones used in the systematic counting of significant comparisons reported in Figs. 2.7 and 2.8.

2.5.7 *Violin plots*

Violin plots were calculated using Gaussian kernel density estimations with Scott's rule (as implemented by *www.scipy.org*; [Scott, 1992]) for bandwidth estimation. Highlighted horizontal lines within the violin-shaped plot elements denote 2.5 %, 50 % and 97.5 % quantiles.

2.5.8 *Trial ensemble comparison*

Beyond feature extraction we also compared directly vectors of firing rates measured across different trials for a same common cell and a same common stimulus. We then compared firing rate ensembles over trials for matching stimulus directions and cells across different experimental conditions, by means of a between-sample two-way Kolmogorov-Smirnov test. As for pairwise feature comparisons, we deemed a comparison between firing rate trial ensembles to be significant, whenever the p-value of this Kolmogorov-Smirnov test was smaller than 0.05. We call *specific effects* such stimulus- and cell-dependent effects of a significant change in condition revealed by trial ensemble comparison.

3

RING-ARCHITECTURE RATE MODEL DESCRIBES TUNED RESPONSES TO COMPOSITE STIMULI AND THEIR ATTENTIONAL MODULATION, IN CONTRAST TO EXISTING PHENOMENOLOGICAL MODELS OF ATTENTION

In the last chapter, we discussed how attention modulates response patterns under various composite stimulus configurations. Now, we will see that existing phenomenological models of attention, the BC model and the FSGM, cannot well describe the data as a whole. As an alternative, we will develop a non-phenomenological model which is conceptually based on the interactions between visual areas, in which the effect of attention—which we model as external to the circuit—is shaped by interactions within the circuit. While this model can account for the data it also produces counter-intuitive predictions, like an average decrease of firing rates in V1 due to attention. Consequently, we will argue that the model is ill-constrained and uses a too naive, system-extrinsic description of attention.

3.1 INTRODUCTION

Numerous studies have found modulations of a cell's firing rate when attention was directed to a stimulus falling within its RF (e. g. [Moran and Desimone, 1985; Reynolds et al., 1999; McAdams and Maunsell, 1999a; Treue and Trujillo, 1999]) and to summarize such data, phenomenological models of attention have been developed. The BC model [Desimone and Duncan, 1995; Reynolds et al., 1999], for instance, stipulates that when two stimuli fall within a cell's RF the response to the pair of stimuli will lie in between the response each of the stimuli would elicit if presented alone, and when one of the two is attended while presented simultaneously, then the response will be biased towards the corresponding single-stimulus response. The FSGM, another example, states that neurons undergo a response gain when an attended stimulus falls within their RF and the gain depends on the similarity between the attended and the cell's preferred stimulus, so that the cell's activity can even decrease when a non-preferred stimulus is attended. Computationally, it has been demonstrated that these two models are not mutually exclusive [Ardid et al., 2007].

Like any model, these need to withstand tests with novel datasets. Specifically, the BC has been tested most convincingly with only few different combinations of oriented and colored stimuli where one of them was always eliciting a strong response [Reynolds et al., 1999]

and the FSGM was tested, so far, with only single stimuli in the RF [Treue and Maunsell, 1999; Martínez-Trujillo and Treue, 2004], although simulations suggest that it also holds for combinations of preferred and anti-preferred stimuli. An important novel test case for both of these models constitutes, therefore, composite stimuli where neither of the two needs to be a preferred stimulus.

We have analyzed responses of neurons in macaque area MT to the simultaneous presentation of two RDP. These two stimuli always had an angular distance of 120° and were co-varied in steps of 30°, so that, in general, none of the two was a preferred stimulus. Moreover, while the two stimuli were always presented within the recorded cell's RF, our collaborators conducted two separate experiments in which the two were either presented in two spatially separated apertures each covering only a small part of the RF, or transparently overlaid in an aperture that spanned a great part of the cell's RF. We found that, for both paradigms, neither the average nor the majority of individual tuning curves did correspond to the BC model. The FSGM which makes a prediction for the population response curve (i. e. the average tuning curve [Treue et al., 2000]) did probably not hold for the *transparent* paradigm, but, might have for the *spatially separate* paradigm; due to the size of the error bars, a decisive conclusion could not be drawn.

Importantly, though, even if the BC or the FSGM held, none of them would be able to distinguish the *spatially separate* from the *transparent* paradigm. The attentional modulations of responses to bidirectional stimuli were, however, qualitatively different in the *spatially separate* and in the *transparent* paradigm. Most prominently, while for both paradigms, the tuning curves were bimodal with approximately equally high peaks when attention was directed to a fixation point and one of the peaks was increased when attention was directed to one of the two stimuli, the other peak was decreased in the *spatially separate*, but enhanced in the *transparent* paradigm. Hence, a novel approach is necessary to describe these data.

Ardid et al. [2007, 2010] have described a spiking-network model of two hypercolumns, one representing a sensory area (e. g. MT), the other a control area (e. g. FEF), with which they could reproduce a variety of neural correlates of attention, among them the BC and FSGM. The model in the presented form would, however, not be able to distinguish spatially separate from transparent stimulus presentations. This could be achieved by extending the model, incorporating two V1 hypercolumns—corresponding to two subparts of the MT RFs—interacting with each other and with the already present sensory area (MT). While such an approach would be most interesting, it would also be computationally expensive, as the number of units to simulate would double, and, furthermore, the interactions between the

V1 and MT hypercolumns could be modeled in a number of ways, increasing the number of free parameters of the model.

Given the wide range of attentional effects that can be described with the Ardid-model, we hypothesized that our data can be described in the framework of a circuit model, but we strove for a simpler description. We therefore opted for a rate model with ring-architecture, introduced by Ben-Yishai et al. [1995] to describe orientation tuning arising from within the cortex even in absence of tuned input. As already anticipated above our data necessitates three hypercolumns, or rings, representing two V1 locations whose RFs are both contained with the RFs of a MT population. For biological plausibility, we assume that each location actually contains two population of neurons, one excitatory, the other inhibitory. Attention is postulated to act additively as an external current source. The model will be constrained by the measured responses for MT cells which we assume to have been excitatory; as the model is underconstrained by the available data, some V1 responses will be treated as free variables, all remaining variables, including the remaining V1 responses as well as, importantly, connection strengths between the populations and the external inputs corresponding to the stimulus and attention will arise as solutions of the model.

Interestingly, the model possesses numerous and qualitatively different solutions, indicating that a given response in area MT is compatible with a variety of different circuit mechanisms and V1 response patterns, as long as these are suitably matched. We found that attentional currents, representing the net-effect of attention from all control areas, often, but not always were negative, at least for some stimuli. Moreover, the model predicts that only few of the possible V1 response pattern are compatible with the BC model. On the other hand, due to the variety in individual solutions, average response profiles in V1 over all these solutions showed somewhat counterintuitive effects (and therefore they were not compatible with FSGM) indicating that the average solution is not yet physiological and that complementary constraints or hypothesis regarding the action of attention are required.

3.2 MATERIALS AND METHODS

3.2.1 *Model architecture and assumptions*

The model architecture is illustrated in Figure 3.1. Activity is determined by an interplay of two V1 locations, $V1^a$ and $V1^b$, and one MT location, MT, and each location is modeled as one excitatory and one additional inhibitory direction hypercolumn. Note however, that in the *transparent* paradigm all presented stimuli always cover most of the RF. We make the assumption that all subparts of the RF are ho-

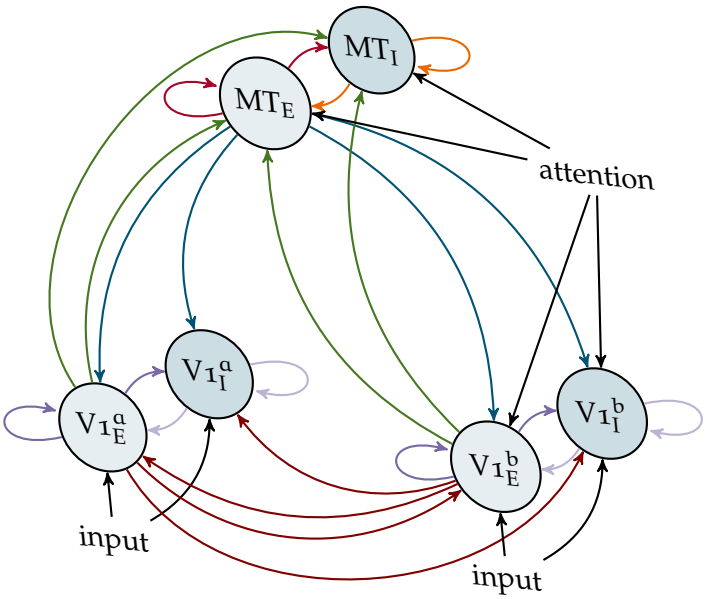


Figure 3.1: Circuit diagram

mogeneous and consequently $V1^a$ and $V1^b$ are indistinguishable in the *transparent* paradigm, and we will therefore consider only one of the two referring to it as simply V1. Hypercolumns within each location are reciprocally connected, and receive additional input from the excitatory hypercolumns of all other locations. In addition to these inputs from within the network, locations in V1 receive external input. In the *spatially separate uni* condition this input, $I^0(\theta)$, is sent only to one of the two V1 locations, $V1^b$, whereas the other, $V1^a$ receives no input and we assume that $I^0(\theta)$ is symmetric around (the stimulus direction) $\theta = 0$. In the *transparent uni* condition, it is assumed that the *same* input $I^0(\theta)$ arrives at both V1 locations. In the *afix* and *ain* conditions, we assume the stimulus directions to be $\pm\mu = \pm 120°$ and that, consequently, the corresponding inputs are $I^0(\theta+\mu)$, $I^0(\theta-\mu)$ and $I^0(\theta+\mu)+I^0(\theta-\mu)$ for the $V1^a$ (spatially separate paradigm), $V1^b$ and V1 (transparent) hypercolumn. Additionally, we make the assumption that an attentional top-down current arrives at only $V1^b$ in the *spatially separate ain* condition, and at MT in the *transparent ain* condition, referred to as, respectively, $K(\theta)$ and $M(\theta)$. All described internal connection strengths and external inputs shall be identical for an excitatory and a corresponding inhibitory target hypercolumn. Altogether, we will model neuronal activity using the equations

$$r^{(\aleph\alpha)} = f_\alpha\left(I^{(\aleph\alpha)}(\theta) + \sum_\beta \int_{-\pi}^{\pi} J^{(\aleph\alpha\beta)}(\theta-\theta')r^{(\aleph\beta)}(\theta')d\theta'\right) \tag{3.1}$$

where $\aleph$ will refer to the experimental paradigm and condition takin on values $\{su, sf, si, tu, tf, ti\}$ where s and t denote the paradigm, either *spatially separate* or *transparent*, and u, f and i denote the condition, either *uni*, *afix* or *ain*. Moreover, α (and also β) will denote populations, $V1_{E,I}^{a,b}$ or $MT^{E,I}$, where E and I abbreviate the excitatory and inhibitory hypercolumn at each location, respectively, and the superscript a or b will be dropped for the *transparent* paradigm as explained above.

The sum of all inputs a neuron receives is assumed to be related to its output firing rate via a non-linear sigmoid transfer function f_α, modeled as a hyperbolic ratio

$$f_\alpha(x) = R_{max}\frac{x^n}{x^n + C_{50}^n}$$

We will only use two different f_α, one for excitatory (E) and one for inhibitory (I) populations, independent of the population location (but use the more general subscript α for notational reasons) and the parameters R_{max}, n and C_{50} will be different for E and I populations.

We had access to an experimental dataset for responses of MT neurons in all six mentioned conditions $\aleph$. In these experiments it was not determined if a measured neuron was excitatory or inhibitory; as excitatory neurons have been reported to have larger spikes and are easier to detect [McCormick et al., 1985; Connors and Gutnick, 1990], we assumed, therefore, that they were always excitatory and fixed the values of $r^{(\aleph,MT_E)}$ with the values of these measurements (see below for details).

Even with these constraints, we had more variables than equations and treated the six parameters R_{max}, n and C_{50} (for E and I), as well as the firing rates $r^{(\aleph\alpha)}$ for $\aleph\alpha \in \{suV1_E^a, suV1_E^b, tuV1_E, tfV1_E\}$ as free variables.

3.2.2 *Solution*

As the sigmoid transfer functions f_α are strictly monotonous we can invert them and then, using the convolution theorem for periodic functions [Zygmund, 2002, p. 36], express (3.1) in terms of Fourier coefficients

$$f_\alpha^{-1}(r^{(\aleph\alpha)}(\theta)) = I^{(\aleph\alpha)}(\theta) + \sum_\beta \int_{-\pi}^{\pi} J^{(\alpha\leftarrow\beta)}(\theta-\theta')r^{(\aleph\beta)}(\theta')d\theta'$$

$$\left[f_\alpha^{-1}(r^{(\hat{\aleph}\alpha)}(\cdot))\right]_n = \hat{I}_n^{(\aleph\alpha)} + 2\pi\sum_\beta \hat{J}_n^{(\alpha\leftarrow\beta)}\hat{r}_n^{(\aleph\beta)}$$

which we will rewrite using the abbreviation $\hat{\jmath}_n^{(\aleph\alpha)} = \left[f_\alpha^{-1}(r^{(\hat{\aleph}\alpha)}(\cdot))\right]_n$ as

$$\hat{\jmath}_n^{(\aleph\alpha)} = \hat{I}_n^{(\aleph\alpha)} + 2\pi \sum_\beta \hat{J}_n^{(\alpha\leftarrow\beta)} \hat{r}_n^{(\aleph\beta)} \tag{3.2}$$

As noted above we make the assumption that excitatory connections have the same strength as corresponding inhibitory connections, that is $\hat{J}_n^{(X_E\leftarrow\beta)} = \hat{J}_n^{(X_I\leftarrow\beta)}$ for all source populations β and for $X \in \{MT, V1, V1^a, V1^b\}$. Furthermore, we assume that input currents to corresponding excitatory and inhibitory populations are identical, that is $\hat{I}_n^{(\aleph X_E)} = \hat{I}_n^{(\aleph X_I)}$ ($X \in \{MT, V1, V1^a, V1^b\}$). This last assumption could be neglected at the expense of introducing additional free parameters.

A direct consequence of these assumptions is that firing rates of all inhibitory populations are directly related to the firing rates of the corresponding excitatory populations, mediated only via the non-linearities $f_{E,I}$. Indeed, the right-hand-sides of (3.7) coincide whenever $\alpha = X_E$ and $\alpha = X_I$. Hence, we obtain

$$\hat{\jmath}_n^{(\aleph,X_I)} = \hat{\jmath}_n^{(\aleph,X_E)} \tag{3.3}$$

for all possible $\aleph$ and all possible X. Instantiating the Fourier series from these Fourier coefficients,

$$f_\alpha^{-1}(r^{(\aleph\alpha)}(x))) = \sum_n \left[f_\alpha^{-1}(r^{(\hat{\aleph}\alpha)}(\cdot))\right]_n e^{inx} = \sum_n \hat{\jmath}_n^{(\aleph,\alpha)} e^{inx}$$

we obtain, for $\alpha = X_I$, and using (3.3)

$$r^{(\aleph X_I)}(x) = f_{X_I}\left(\sum_n \hat{\jmath}_n^{(\aleph,X_E)} e^{inx}\right);$$

the corresponding Fourier coefficients, which we are seeking, read

$$\hat{r}_n^{(\aleph X_I)} = \frac{1}{2\pi}\int f_{X_I}\left(\sum_m \hat{\jmath}_m^{(\aleph,X_E)} e^{inx}\right) e^{inx} dx. \tag{3.4}$$

Note that, even if $\hat{\jmath}_n^{(\aleph,X_E)} = 0$ for $|n| > n_0$ ($n_0 \in \mathbb{N}$), in general $\hat{r}_n^{(\aleph X_I)} \neq 0$ for $|n| > n_0$ due to the non-linearity f_{X_I}. In the computer implementation of this solution algorithm we only keep track of Fourier coefficients up to a fixed order (we will use up to order 2). Therefore, our obtained solution for $r^{(\aleph X_I)}$ will, in general, not be completely compatible with (3.3), so that the solution cannot be exact up to machine precision.

For $X = MT$ the Fourier coefficients of our measured data determine $\hat{r}_n^{(\aleph X_E)}$. Moreover, we assume that $\hat{r}_n^{(\aleph X_E)}$ are free parameters for $\aleph X \in \{suV1^a, suV1^b, tuV1, tfV1\}$. In all these cases (3.4) can be applied directly, in all other cases (3.4) remains among the equations to solve. All other equations, complementing these cases, read

$$\begin{pmatrix} \hat{\jmath}_n^{(su,MT_E)} \\ \hat{\jmath}_n^{(tu,MT_E)} \\ \hat{\jmath}_n^{(sf,MT_E)} \\ \hat{\jmath}_n^{(tf,MT_E)} \\ \hat{\jmath}_n^{(si,MT_E)} \\ \hat{\jmath}_n^{(ti,MT_E)} \end{pmatrix} = \hat{I}_n^0 \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ \hat{M}_n \end{pmatrix} + 2\pi\cdot \tag{3.5}$$

$$\begin{pmatrix} \hat{r}_n^{(su,MT_E)} & -\hat{r}_n^{(su,MT_I)} & \hat{r}_n^{(su,V1_E^a)} + \hat{r}_n^{(su,V1_E^b)} \\ \hat{r}_n^{(tu,MT_E)} & -\hat{r}_n^{(tu,MT_I)} & 2\hat{r}_n^{(tu,V1_E)} \\ \hat{r}_n^{(sf,MT_E)} & -\hat{r}_n^{(sf,MT_I)} & \hat{r}_n^{(sf,V1_E^a)} + \hat{r}_n^{(sf,V1_E^b)} \\ \hat{r}_n^{(tf,MT_E)} & -\hat{r}_n^{(tf,MT_I)} & 2\hat{r}_n^{(tf,V1_E)} \\ \hat{r}_n^{(si,MT_E)} & -\hat{r}_n^{(si,MT_I)} & \hat{r}_n^{(si,V1_E^a)} + \hat{r}_n^{(si,V1_E^b)} \\ \hat{r}_n^{(ti,MT_E)} & -\hat{r}_n^{(ti,MT_I)} & 2\hat{r}_n^{(ti,V1_E)} \end{pmatrix} \begin{pmatrix} \hat{\jmath}_n^{(MT\leftarrow MT_E)} \\ \hat{\jmath}_n^{(MT\leftarrow MT_I)} \\ \hat{\jmath}_n^{(MT\leftarrow V1_E)} \end{pmatrix}$$

$$\begin{pmatrix} \hat{\jmath}_n^{(su,V1_E^a)} \\ \hat{\jmath}_n^{(su,V1_E^b)} \\ \hat{\jmath}_n^{(tu,V1_E)} \\ \hat{\jmath}_n^{(sf,V1_E^a)} \\ \hat{\jmath}_n^{(sf,V1_E^b)} \\ \hat{\jmath}_n^{(tf,V1_E)} \\ \hat{\jmath}_n^{(si,V1_E^a)} \\ \hat{\jmath}_n^{(si,V1_E^b)} \\ \hat{\jmath}_n^{(ti,V1_E)} \end{pmatrix} = \hat{I}_n^0 \begin{pmatrix} 0 \\ 1 \\ 1 \\ e^{i\mu n} \\ e^{-i\mu n} \\ e^{-i\mu n} + e^{i\mu n} \\ e^{i\mu n} \\ e^{-i\mu n} \\ e^{-i\mu n} + e^{i\mu n} \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ \hat{K}_n \\ 0 \end{pmatrix} + 2\pi\cdot \tag{3.6}$$

$$\begin{pmatrix} \hat{r}_n^{(su,V1_E^a)} & -\hat{r}_n^{(su,V1_I^a)} & \hat{r}_n^{(su,V1_E^b)} & \hat{r}_n^{(su,MT_E)} \\ \hat{r}_n^{(su,V1_E^b)} & -\hat{r}_n^{(su,V1_I^b)} & \hat{r}_n^{(su,V1_E^a)} & \hat{r}_n^{(su,MT_E)} \\ \hat{r}_n^{(tu,V1_E)} & -\hat{r}_n^{(tu,V1_I)} & \hat{r}_n^{(tu,V1_E)} & \hat{r}_n^{(tu,MT_E)} \\ \hat{r}_n^{(sf,V1_E^a)} & -\hat{r}_n^{(sf,V1_I^a)} & \hat{r}_n^{(sf,V1_E^b)} & \hat{r}_n^{(sf,MT_E)} \\ \hat{r}_n^{(sf,V1_E^b)} & -\hat{r}_n^{(sf,V1_I^b)} & \hat{r}_n^{(sf,V1_E^a)} & \hat{r}_n^{(sf,MT_E)} \\ \hat{r}_n^{(tf,V1_E)} & -\hat{r}_n^{(tf,V1_I)} & \hat{r}_n^{(tf,V1_E)} & \hat{r}_n^{(tf,MT_E)} \\ \hat{r}_n^{(si,V1_E^a)} & -\hat{r}_n^{(si,V1_I^a)} & \hat{r}_n^{(si,V1_E^b)} & \hat{r}_n^{(si,MT_E)} \\ \hat{r}_n^{(si,V1_E^b)} & -\hat{r}_n^{(si,V1_I^b)} & \hat{r}_n^{(si,V1_E^a)} & \hat{r}_n^{(si,MT_E)} \\ \hat{r}_n^{(ti,V1_E)} & -\hat{r}_n^{(ti,V1_I)} & \hat{r}_n^{(ti,V1_E)} & \hat{r}_n^{(ti,MT_E)} \end{pmatrix} \begin{pmatrix} \hat{\jmath}_n^{(V1^i\leftarrow V1_E^i)} \\ \hat{\jmath}_n^{(V1^i\leftarrow V1_I^i)} \\ \hat{\jmath}_n^{(V1^i\leftarrow V1_E^j)} \\ \hat{\jmath}_n^{(V1\leftarrow MT_E)} \end{pmatrix}$$

These equations constitute a high-dimensional nonlinear system Therefore, a brute-force approach to solve them would probably be problematic and we undertook the following steps to divide them into uncoupled, lower-dimensional subsystems in order to increase the chances to find solutions.

First, as we assumed that $\hat{r}_n^{(\aleph X_E)}$ are free parameters for $\aleph X \in \{suV1^a, suV1^b, tuV1, tfV1\}$, the only unknowns in the first, second and fourth equation of (3.5) are the Fourier coefficients of the connection kernels for which can thus solve, provided that the 3×3 matrix of firing rate Fourier coefficients from these three rows in (3.5) is invertible. Hence, we have $\hat{J}_n^{(MT \leftarrow MT_E)}$, $\hat{J}_n^{(MT \leftarrow MT_I)}$ and $\hat{J}_n^{(MT \leftarrow V1_E)}$.

Therewith, we obtain $\hat{r}_n^{(sf,V1_E^a)} + \hat{r}_n^{(sf,V1_E^b)}$ and $\hat{r}_n^{(si,V1_E^a)} + \hat{r}_n^{(si,V1_E^b)}$ from the third and fifth equation in (3.5).

The only equation left now in (3.5) is the sixth which contains two unknowns, $\hat{r}_n^{(ti,V1_E)}$ and $\hat{M}_n$. As the latter appears only here, this equation must be used to fix $\hat{M}_n$ in dependence on $\hat{r}_n^{(ti,V1_E)}$ which will be determined later.

We now turn to the 9 equations of (3.6) with the 11 remaining unknowns $\hat{I}_n^0$, $\hat{J}_n^{(V1^i \leftarrow V1_E^i)}$, $\hat{J}_n^{(V1^i \leftarrow V1_I^i)}$, $\hat{J}_n^{(V1^i \leftarrow V1_E^j)}$, $\hat{J}_n^{(V1 \leftarrow MT_E)}$, $\hat{K}_n$, as well as $\hat{r}_n^{(\aleph X_{E,I})}$ for $\aleph X \in \{sfV1^a, sfV1^b, siV1^a, siV1^b, tiV1\}$; note however, that we have determined $\hat{r}_n^{(sf,V1_E^a)} + \hat{r}_n^{(sf,V1_E^b)}$ and $\hat{r}_n^{(si,V1_E^a)} + \hat{r}_n^{(si,V1_E^b)}$ before.

The first, second, third and sixth equation of (3.6) contain five of these unknowns, $\hat{I}_n^0$, $\hat{J}_n^{(V1^i \leftarrow V1_E^i)}$, $\hat{J}_n^{(V1^i \leftarrow V1_I^i)}$, $\hat{J}_n^{(V1^i \leftarrow V1_E^j)}$, $\hat{J}_n^{(V1 \leftarrow MT_E)}$. Assuming that the 4×4 matrix of firing rate Fourier coefficients made from these four rows is invertible we can straightforwardly solve for the four $\hat{J}_n$s, in dependence of $\hat{I}_n^0$.

We plug the corresponding formula for the $\hat{J}_n$s in the fourth and fifth equation of (3.6). Together with (3.3) for $\aleph X = sfV1^{a,b}$, and the already determined value for $\hat{r}_n^{(sf,V1_E^a)} + \hat{r}_n^{(sf,V1_E^b)}$ we have five equations to determine the five unknowns $\hat{I}_n^0$, and $\hat{r}_n^{(sf,V1_{E,I}^{a,b})}$. This system needs to be solved numerically. As a corollary, we can now also determine the values of the connection kernels, $\hat{J}_n^{(V1^i \leftarrow V1_E^i)}$, $\hat{J}_n^{(V1^i \leftarrow V1_I^i)}$, $\hat{J}_n^{(V1^i \leftarrow V1_E^j)}$, $\hat{J}_n^{(V1 \leftarrow MT_E)}$, which we knew so far only in dependence of $\hat{I}_n^0$.

Note also that while the spatially separated afix activity in $V1^a$ or $V1^b$ need not be symmetric around $0°$, but their sum must be for symmetry reasons ($V1^a$ and $V1^b$ can be interchanged), likewise the difference must be antisymmetric around $0°$.

The seventh equation of (3.6), together with (3.3) for $\aleph X = siV1^a$ and the value of $\hat{r}_n^{(si,V1_E^a)} + \hat{r}_n^{(si,V1_E^b)}$ that we obtained already above, can be solved numerically for $\hat{r}_n^{(si,V1_{E,I}^a)}$ and $\hat{r}_n^{(si,V1_E^b)}$. As a consequence, we obtain also $\hat{r}_n^{(si,V1_I^b)}$ from (3.6).

This leaves $\hat{K}_n$ as the only unknown left in the eighth equation of (3.6) which is then straightforwardly determinable.

As a last step, we solve the last equation of (3.6) as well as (3.3) for $\aleph X = tiV1$ for $\hat{r}_n^{(ti,V1_{E,I})}$. Therewith we can fix the value of $\hat{M}_n$ which we determined before only in dependence of these two.

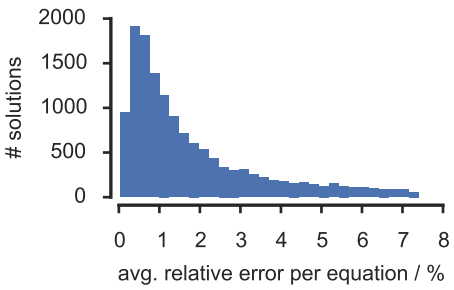


Figure 3.2: All found solutions satisfy model equations with relatively good numerical accuracy. Note that due to the non-linearities f_α (3.4) can only be solved approximately. Nonetheless, for all solutions found the average relative error per model equation is relatively small.

Note that, the described solution algorithm critically depended on the knowledge of $r^{(\aleph\alpha)}$ for $\aleph\alpha \in \{suV1_E^a, suV1_E^b, tuV1_E, tfV1_E\}$ which we assumed were free variables. In case we make any additional assumptions on the structure of the solution that would eliminate some of the free variables, we would not any more in the position to pick three equations in (3.5) and four equations in (3.6) that allowed us to determine the connection strengths $\hat{J}_n^{(\alpha\leftarrow\beta)}$ via matrix inversion (see above). In such a case, thus, we would have to solve numerically higher-dimensional subsystems than we have in the case described here.

3.2.3 *Accuracy of solutions*

As described after the derivation of (3.4), we cannot expect to find solutions with a numerical accuracy up to machine precision due to the non-linearities f_α. We therefore calculated for all n ($n = 0, 1, 2$) the relative error for each of the equations in (3.5), (3.6) as well as the corresponding equations for the inhibitory populations, by subtracting the right-hand-sides from the left-hand-sides, dividing by left-hand sides (in case the absolute values of the left-hand-sides were not greater than 10^{-8} the right-hand sides were used for normalization, and in case the absolute values of these these were not greater than 10^{-1} no normalization was performed), and taking the absolute value. The histogram of the average of these relative errors over all equations is shown in Figure 3.2 and illustrates that despite the necessary approximation in (3.4) all found solutions satisfy the model equations with only a few percents of error per equation.

3.2.4 *Numerical root-finding for non-linear subsystems*

In all cases where we have to solve subsystems numerically we employ a modified version of the Cuckoo Search algorithm, which was originally described by Yang and Deb [2010] and which aims at minimizing a given objective function; the objective function used here is the sum of squared residuals of all the equations in the subsystem to be solved. The algorithm starts from a set of candidate solutions

(we use 5) and, in each iteration, makes two attempts to improve them. First, each candidate solution is tentatively shifted towards the currently known best solution where the step size is proportional (we take the proportionality constant to be 0.01) to a random variable with Levy distribution (calculated using the algorithm described by Mantegna [1994] with parameter $\alpha = 3/2$). Whenever this leads to a better solution it is accepted, otherwise the previous candidate solution is kept. Second, a proportion $1 - p$ (we take $p = 0.2$) of randomly chosen candidate solutions is moved by a randomly chosen fraction of the vector connecting two randomly chosen candidate solutions, where, within this iteration and over all candidate solutions, every candidate solution will be used at most once as a starting point, and at most once as an endpoint of these vectors. If the updated solution outperforms its predecessor it is accepted, otherwise rejected. We modified this algorithm by executing a Newton-Krylow solver [Hindmarsh et al., 2005], whenever the cuckoo search finds an improvement for one of the candidate solution. The algorithm was run for only 5000 iterations for performance reasons. Importantly, when the system has not converged at that point, this cannot be interpreted as an unsolvability of the system. During the solution process, we monitored if any of the tuning curves would drop below zero, and if so abandoned the corresponding solution immediately.

3.2.5 *Incorporating measured values for MT neurons*

We constrained the model by measurements of MT neurons in the six considered conditions $\aleph$ in which responses were obtained for 12 stimulus directions. In the spatially separate paradigm two uni conditions were recorded in which the stimulus was presented two one of two spatially separate subparts of the cell's RF. In the following we used only the data from the condition in which the RF's subpart matched the location to be attended in the *ain* condition (but the average tuning curve using just this condition, and the average tuning curves over both these uni conditions were very similar, so we don't expect a substantial bias through this choice) and refer to it as *uni*.

The approximate solution of the model equations derived above is expressed in terms of Fourier coefficients. We therefore fitted a Fourier series (of order 2) to the trial-averaged response data in each condition and used the resulting Fourier coefficients as values for $\hat{r}^{(\aleph,\mathrm{MT_E})}$ (i. e. we always assumed that the measured neurons were excitatory).

Furthermore, the experimental raw data was aligned such that the preferred direction of the cell was at 0° (see section 2.5.2 for details). In the solution of the model equation we assumed, however, for convenience, that the stimulus directions for bidirectional conditions are at $\pm\mu$ ($\mu = 120°$) and for that reason we shifted the *afix* and *ain* tun-

ing curves relative to the raw data alignment by 60° to accommodate this choice. On average, tuning curves in the *uni* condition should be symmetric around the direction of the preferred stimulus direction (0° with our alignment) as there is no apparent reasons why a certain direction with a clockwise shift from the preferred one should be processed any different, on average, than the corresponding direction with the same angular shift in the anti-clockwise direction. The same argument holds true for the *afix* condition around the average direction of the two stimuli (also 0° with our alignment). This was, however, not strictly true. Therefore, as for real symmetric (around 0°) Fourier series $\sum_n c_n e^{inx}$ it holds that $c_n^* = c_n$ (where * denotes complex conjugation)[1], we symmetrized *uni* and *afix* tuning curves by rounding the phases of the corresponding Fourier coefficients to multiples of π.

3.2.6 *Solution filtering*

We required all solutions to fulfill the following properties in order to be deemed admissible. First, we required the Fourier series for all tuning curves, the connection kernels and the input current I^0 to have positive values below 100. Second, for the attentional currents K and M we allowed negative values, but stipulated that their absolute values be smaller than the absolute value of the summed recurrent input from within the network in the *ain* condition to $V1^b_{E,I}$ and $MT_{E,I}$, respectively, so that the activity in these areas would not just be imposed by the attentional currents due to their strengths but would still be determined in large part by the local network itself. Third, we required that, in the *uni* conditions of both paradigms, tuning curves at V1 locations receiving input (i. e. $V1^b$), as well as the input I^0 itself had their global maximum either around 0 rad ($\pm\pi/6$ rad) which, by design, is the position of the *uni* tuning curves in the excitatory MT population, or around π, ($\pm\pi/6$ rad) in which case we could just "rename" the preferred direction of the cell from π to 0 and thereby restitute the consistency between MT and V1. Note that, by design, *uni* tuning curves were symmetric around 0, so that whenever the global maximum was at any other position, there had to be another corresponding global maximum at the position mirrored at 0. Such a case would contrast the classical view that tuning curves are, on average, symmetric around their preferred direction and we therefore decided to exclude these cases.

1 Let $f \sim \sum_n c_n e^{inx}$ and note that $\{e^{inx}/\sqrt{2\pi}\}$ form an orthonormal system for $n \in \mathbb{Z}$ (follows from [Zygmund, 2002, p. 6]). Then, from $f^*(x) = f(x)$ it follows that $c_{-n} = c_n^*$ and, in the next step, we get $c_n^* = c_n$ from the condition that $f(-x) = f(x)$.

3.2.7 *Data analysis*

All data analysis was performed in Python.

3.2.8 *Violin plots*

Violin plots were calculated using Gaussian kernel density estimations with Scott's rule (as implemented by *www.scipy.org*; Scott [1992]) for bandwidth estimation. Highlighted horizontal lines within the violin-shaped plot elements denote 2.5 %, 50 % and 97.5 % quantiles.

3.3 RESULTS

3.3.1 *Probing phenomenological models of attention*

This section, originally conceived to appear in a separate manuscript, was written in collaboration with Demian Battaglia.

Several phenomenological models of attention have been introduced, which try to predict the tuned response in presence of attention by applying a transformation to the the tuning curve in absence of attention. Commonly discussed models include, e.g. the biased competition (BC) model [Reynolds et al., 1999], the feature similarity gain model (FSGM) [Treue and Trujillo, 1999; Martínez-Trujillo and Treue, 2004] or the normalization model [Reynolds and Heeger, 2009].

As a further application of the trial ensemble comparison method introduced in section 2.3.7, we tested the compatibility of our measurements with the BC model. According to the BC model, whenever two stimuli are presented in the RF of the same neuron, they compete to get maximally represented in the output firing rate of the cell. Here, the two stimuli correspond to the two RDPs that the *afix* condition is composed of, and we will tell them apart by calling them "*uni1*" and "*uni2*", the latter one being the RDP that receives attention in the *ain* condition. As in the *spatially separated* paradigm the two RDPs are presented in different parts of the RF, they will have different tuning curves, whereas in the transparent paradigm such a distinction cannot be made and *uni1* and *uni2* tuning curves are, thus, identical (even if we still plot them apart as peaking at two different positions, reflecting the fixed separation of 120° between the two presented stimuli).

Let us suppose, now, that the responses of a given cell to each of the RDPs alone are given by r_{uni1} and r_{uni2}, in the *uni1* and *uni2* condition, respectively and let us assume for the moment that $r_{uni1} \leqslant r_{uni2}$. Let also denote as r_{afix} the response to the same cell to the presentation of the composite stimulus in the *afix* condition. Under the BC model, competition between the stimuli would translate into the constraint $r_{uni1} \leqslant r_{afix} \leqslant r_{uni2}$. In the *ain* condition, the cell additionally receives an "attentional spotlight", which modulates the response to the composite stimulus toward a novel value r_{ain}, which,

according to the BC model tenets, should be closer to the original response to the attended stimulus (always the *uni2* stimulus, in our data-set) when presented alone. In equations, this constraint reads $r_{uni1} \leqslant r_{afix} \leqslant r_{ain} \leqslant r_{uni2}$. In the complementary case in which $r_{uni1} \geqslant_{uni2}$, we could follow a similar derivation to obtain the alternative constraint $r_{uni1} \geqslant r_{afix} \geqslant r_{ain} \geqslant r_{uni2}$. If either one of these two constraints is satisfied by the cell responses in the different conditions, then the BC model holds. These predictions have found previous confirmation in experimental studies [Reynolds et al., 1999] and we checked to which extent they can be considered to hold even for our dataset.

Figure 3.3 shows sample-averaged tuning curves in the *uni1*, *uni2*, *afix* and *ain* conditions, for the spatially separate (panel A) and the transparent (panel B) paradigms. The predictions of the BC model clearly did not hold for these average tuning curves. However, as we have seen, modulations of tuning are highly cell- and stimulus-specific. Therefore we estimated the fraction of individual cells for which the BC condition was fulfilled, for each stimulus direction separately. As plotted in Figure 3.3C,D (blue curves), the fraction of cells fulfilling the BC conditions depended on the stimulus direction and never rose above 30% (for the spatially separate paradigm) or 20% (for the transparent paradigm). These maximum fractions were achieved in alignment with the tuning curve peaks. Thus, we expect the probability of fulfilling BC conditions to be maximized when one of the presented stimuli matches the preferred orientation of the cell we recorded from. This might explain, at least in part, why previous experiments using only preferred stimuli found evidence for the BC model [Moran and Desimone, 1985; Treue and Maunsell, 1996]. However, for general stimuli, the fraction of cells fulfilling BC conditions was even smaller, down to close to 10% (for the spatially separate paradigm) or less (for the transparent paradigm).

A possible reason for these small fractions could lie in the use of too stringent conditions and we therefore also checked if weaker constraints increased the fraction of conforming cells. In particular, in these relaxed constraints, we considered the BC model to hold when r_{ain} was biased away from the competition value r_{afix} towards r_{uni2}, but was also allowed to go beyond, to values higher or lower than r_{unid2}. The fraction of cells fulfilling this weaker conditions rose up to 20 % higher (figure 3.3C,D, pink curves), however they still remained low. As a matter of fact, even the weaker competition condition alone —i.e. either $r_{uni1} \leqslant r_{afix} \leqslant r_{uni2}$ or $r_{uni2} \leqslant r_{afix} \leqslant r_{uni1}$— was generally satisfied by only a minority of cells, apart from the stimuli including a component with the preferred direction (figure 3.3C,D, green curves). Thus, the BC model could not be considered to hold in general, when analyzing cell responses and their modulations at the level of specific effects.

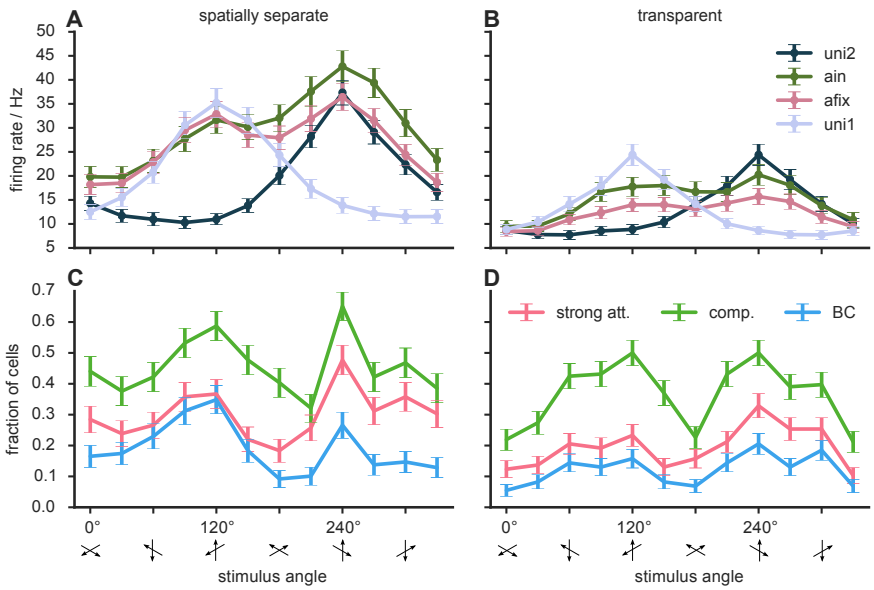


Figure 3.3: Biased Competition model holds only for a subset of cells and stimuli. Mean tuning curves from spatially separated (**A**) and transparent (**B**) paradigm. Error bars denote standard error of the mean. The biased competition (BC) model posits that firing rates be ordered (either increasing or decreasing) as uni1, afix, ain, uni2 (see main text). This holds robustly only for a small subset of stimuli in the average tuning curves, namely the right peak in the transparent paradigm. **C,D)** Likewise, BC holds only for a small fraction of cells when tested on each cell's tuning curves individually (blue) and even weakened conditions—"strong att." (pink): attention biases response to go beyond uni2; "comp." (green): only competition, i.e. the firing rate order uni1, afix, uni2 is tested for—fail often. Error bars denote standard error of the mean of each points binary distribution "BC holds" vs "doesn't hold".

The FSGM [Treue and Trujillo, 1999; Martínez-Trujillo and Treue, 2004] stipulates that attention multiplies a cell's tuning curve by a certain factor. The value of this factor, in turn, depends monotonically on the angular distance between the attended and the preferred direction of the cell. Cells were included in the current analysis only if their firing rates were above 0 Hz for all stimuli in afix and ain. The average gain factor $\langle \text{ain}/\text{afix} \rangle$, dependent on stimulus, is plotted in figure 3.4A. The average gain factors vary strongly with stimulus angle. Figure 3.4B shows the histogram over cells of gain factor ranges $\Delta g = \max_\theta \text{afix}/\text{ain} - \min_\theta \text{afix}/\text{ain}$. The first bin, collecting values $0 \leqslant \Delta g \leqslant 0.25$ is essentially empty, so that almost all cells have values of Δg that are greater than 25 %. As the values of afix/ain have typical values on the order of 1 (figure 3.4A) that implies that gain factors vary strongly with stimulus for bidirectional stimuli, casting doubt on the MGM as a good description of our dataset.

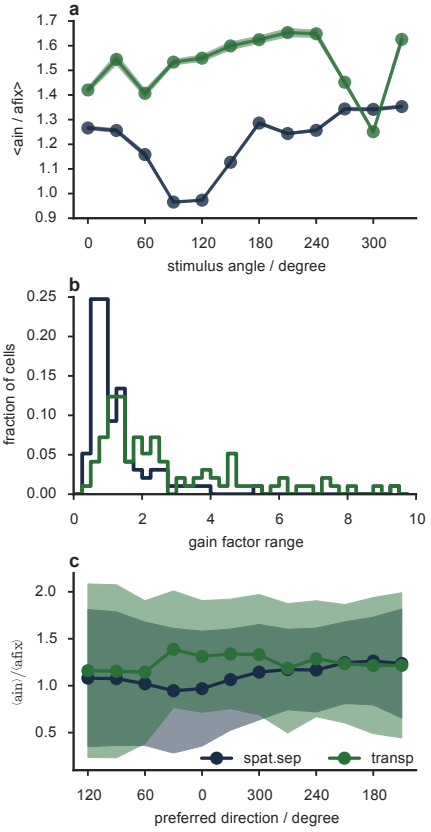


Figure 3.4: Multiplicative gain modulation and feature similarity gain modulation model don't always hold. A) The attentional gain factor ain / afix (due to the division only cells with firing rates above 0 Hz for all stimuli are included in this analysis: $n_{spat.sep} = 97$, $n_{transp} = 96$) is averaged across cells depending on stimulus and paradigm (circles). Shaded areas give standard error of the mean, but note that these errors don't take trial-to-trial variability into account, assuming, rather, a unique value of the tuning curve at each angle. The average gain factor is not constant over stimuli. **B)** Histogram over cells of $\max_\theta$ afix/ain $-$ $\min_\theta$ afix/ain. Irrespective of condition, the first bin with an appreciable number of entries is the second, collecting gain factor ranges from 0.25 to 0.5. As gain factors for our data have an absolute value on the order of magnitude of one, the range corresponds to 25% - 50% variation, that is they vary strongly within cells. **C)** Averaging tuning curves over cells we obtain estimates of population tuning curves $\langle ain \rangle$ and $\langle afix \rangle$, their ratio gives the attentional gain depending on preferred direction. For a given preferred direction, shaded areas indicate mean$\pm$std estimates, obtained via error propagation according to $std = std_{afix}/mean_{ain} + mean_{afix} std_{ain}/mean_{ain}^2$, and (for $x \in \{afix, ain\}$) $std_x = \sqrt{\sum_{i=1}^{N_x} var_{x,i}/N_x}$, N_x is the number of cells in population x and $var_{x,i}$ is the variance of cell i in population x over trials. The ratio $\langle ain \rangle / \langle afix \rangle$ is predicted to depend monotonically on the distance between preferred direction (x-axis) and attended direction (240°). The curve for the mean of the spatially separate paradigm is, indeed, monotonic but centered around 180° instead of 240°, the transparent paradigm's mean curve does not seem to follow the prediction at all.

To test the FSGM we approximate the population tuning curve by averaging tuning curves over all cells. In that way we get an estimate of how firing rate depends on preferred direction over the population of cells [Treue et al., 2000]. The attentional gain factor $\langle ain \rangle / \langle afix \rangle$ of the population tuning curve (figure 3.4C) has a broad peak around 120° in the transparent paradigm which stands in contrast to the FSGM. In the spatially separated paradigm, on the other hand, it is monotonic with maximum at a preferred direction of 180° and minimum at 30°. This is similar to what we would expect according to the FSGM, just, strictly, the maximum should be at 240° and the minimum at 60°.

3.3.2 *There exists a variety of solutions*

As an alternative for existing phenomenological models of attention we studied if a circuit model with ring architecture, as described in the Materials and Methods section, allows an appropriate description of the data.

First, we performed a principal component analysis (PCA) on all found solutions after all variables have been z-scored to make them

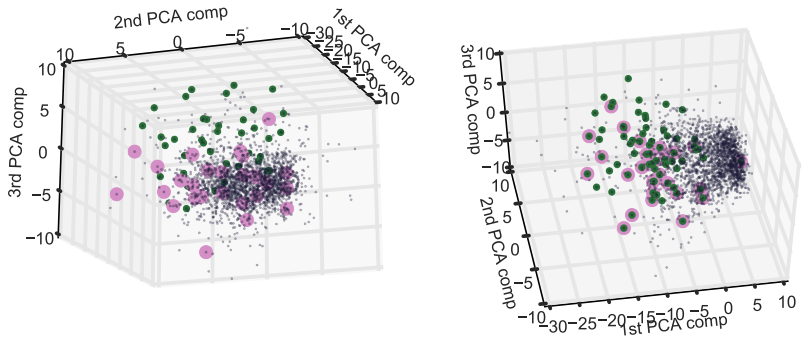


Figure 3.5: Admissible solutions separated from other solutions and from each other. Projection of solutions to first three principal components. Both panels show the same data from different perspectives. Solutions having admissible values for the attentional currents have green halos, in addition, those whose uni-peaks match the input-positions, green halos.

comparable in size and variability. Figure 3.5 shows the solutions in the PCA-coordinate system projected to the first three components. Those solutions that have admissible values for the attentional currents (see Materials and Methods) are marked in green, those that have additionally their peaks in the *uni* conditions matching the input peak (see Materials and Methods) in pink. Interestingly, these admissible solutions seem to form a cluster that is separated from the bulk of the non-admissible solutions. Moreover, while the cloud of unphysiological solutions is relatively densely populated, the cluster of admissible solutions contains only few points, seemingly separated from each other. This could be due to limited sampling, but also because there really do exist non-admissible parameter sets in between.

We then looked at the admissible solutions and found, in support of the latter view, that there exist qualitative differences between them. Indeed, Figure 3.6A, the distribution of Fourier coefficients for all profiles, illustrates the wide range of values they take on, including sign changes for first and second order coefficients, leading, thus, potentially to shape changes of the profiles. To confirm that, we calculated the number of peaks of each profile and sorted them by the number of occurring local maxima and according to whether the global maximum is at 0 ($\pm\pi/6$), π ($\pm\pi/6$) or in between (the latter possibility does not exist in case of only one local maximum due to the symmetry of the profile around 0). The distribution into these five categories is shown in Figure 3.6B for 14 variables. Apparently, the distributions differ strongly between these variables, indicating that there must exist various combinations of peak position profiles among the admissible solutions. In particular this has the effect that the average profiles (depicted for V1 responses in Figure 3.6C) do not represent, in terms of number and location of peaks, all solutions. In the extreme case, occurring for the V1 response in the *transparent uni* condition

($\aleph\alpha = \mathrm{tuV1_E}$) the majority of solutions has only one peak, but their average has two.

The average tuning curves for V1 responses also point to a qualitative difference between the spatially separate and transparent case. As illustrated in Figure 3.6C, the strength of V1 responses tends to be similar in all three conditions in the *transparent* paradigm, comparable to the strength of the *uni* responses in the *spatially separate* paradigm. *afix* responses, and even more *ain* responses, on the other hand, seem to undergo a significant baseline reduction. This prediction would need to be tested experimentally; nevertheless the pattern is clearly different from what happens in MT (cf. section 2.3.6).

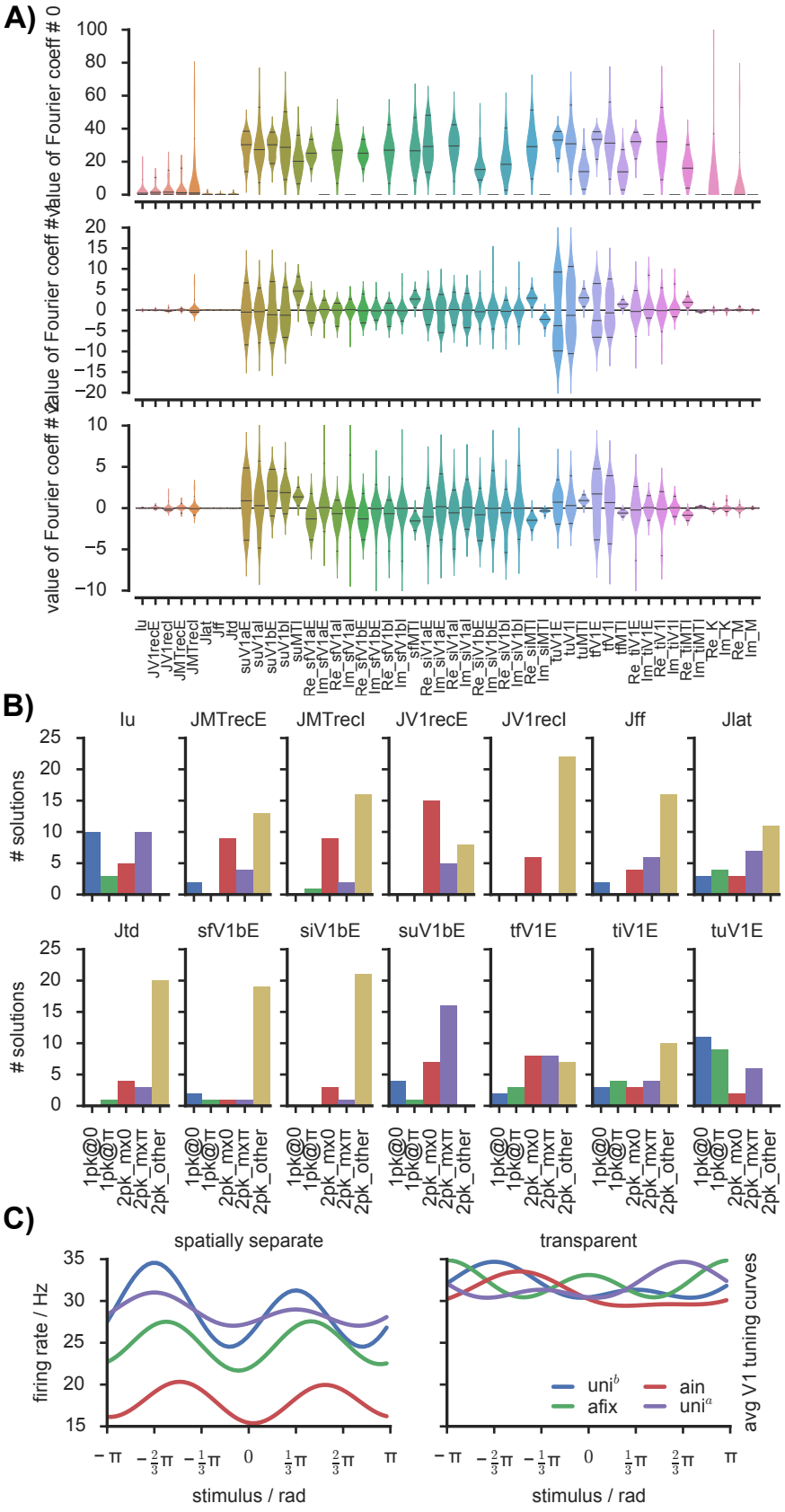
3.3.3 *Effects of attention*

As described in Materials and Methods we allowed only solutions where the attentional currents did not exceed the recurrent inputs to a population. Figure 3.7A shows the average strength over stimulus angles of both the attentional current in the *spatially separate* paradigm (sent to $\mathrm{V1^b}$), K, and in the *transparent* paradigm (sent to MT), in units of the maximum of the input, I which was always positive. For most admissible solutions, at least one of the two lies close to an axis, indicating that it was vanishingly small, on average, compared to the input strength, I. Moreover, while there exist solutions in which the attentional currents were positive for all stimulus angles, in most cases at least one of the two, K or M had inhibitory effects for at least some angles.

We used the responses at V1 locations for admissible solutions as proxies in order to determine putative attentional modulations. Figure 3.7B shows the dependence of attentional indices ai[2] on stimulus angles for all these individual solutions. The first row shows, for each stimulus angle, the median, as well as the 2.5 and 97.5 quantiles of responses. In accordance with Figure 3.6C that attention tends

2 $\mathrm{ai}(\mathrm{i},\mathrm{f}) = (\mathrm{i}-\mathrm{f})/(\mathrm{i}+\mathrm{f})$ where i and f are the responses in the attended (*ain*) and unattended (*afix*) condition, respectively.

Figure 3.6 (*facing page*): Variety of qualitatively different solutions. A) Distribution of Fourier coefficients of admissible solutions. Broad distribution, including sign changes, suggests many qualitatively different profiles. **B)** Number of admissible solutions for which the tuning curve of the variable indicated in the title of each panel had either one peak at 0 rad (blue), one peak at π rad (green), 2 peaks with the higher one at $0 \pm \pi/6$ rad (red), 2 peaks with the higher one at $\pi \pm \pi/6$ rad (purple), or 2 peaks (of the same height) neither of which around 0 and π rad. As the distributions vary between variables, this indicates that there's a wide range of qualitatively different solutions. **C)** V1 tuning curves averaged over all admissible solutions. They look unintuitive.


A)
value of Fourier coeff # 0
value of Fourier coeff # 1
value of Fourier coeff # 2
Iu
JV1recE
JV1recI
JMTrecE
JMTrecI
Jlat
Jff
Jtd
suV1aE
suV1aI
suV1bE
suV1bI
suMTI
Re_sfV1aE
Im_sfV1aE
Re_sfV1aI
Im_sfV1aI
Re_sfV1bE
Im_sfV1bE
Re_sfV1bI
Im_sfV1bI
Re_sfMTI
Im_sfMTI
Re_siV1aE
Im_siV1aE
Re_siV1aI
Im_siV1aI
Re_siV1bE
Im_siV1bE
Re_siV1bI
Im_siV1bI
Re_siMTI
Im_siMTI
tuV1E
tuV1I
tuMTI
tfV1E
tfV1I
tfMTI
Re_tiV1E
Im_tiV1E
Re_tiV1I
Im_tiV1I
Re_tiMTI
Im_tiMTI
Re_K
Im_K
Re_M
Im_M
B)
solutions
Iu
JMTrecE
JMTrecI
JV1recE
JV1recI
Jff
Jlat
Jtd
sfV1bE
siV1bE
suV1bE
tfV1E
tiV1E
tuV1E
1pk@0
1pk@π
2pk_mx0
2pk_mxπ
2pk_other
C)
spatially separate
transparent
firing rate / Hz
stimulus / rad
avg V1 tuning curves
uni^b
afix
ain
uni^α

to decrease responses for all stimulus angles in the *spatially separate* paradigm, but has no consistent effect in the *transparent* paradigm. The stimulus dependencies of individual $\alpha$is, on the other hand, are highly variable as shown in the second row of Figure 3.7C featuring strong qualitative differences between solutions in terms of number, width and location of peaks.

Finally, we checked if responses in V1 are compatible with phenomenological models of attention. To probe the FSGM we need the population response function, which is calculated by averaging individual tuning curves [Treue et al., 2000]. However, as discussed above in the context of Figure 3.6C, the average tuning curve is not indicative of individual cells and, moreover, features non-intuitive effects, like the strong baseline reduction from *afix* to *ain* condition in the *spatially separate* paradigm, as well as, in the *transparent* paradigm, the absence of a peak in the *afix* condition around $\pi/3$, the stimulus direction that became attended, and the much smaller right compared to the left peak in the *ain* condition. The FSGM predictions therefore clearly fail for the found solutions in V1.

The BC model requires that response strengths, for a given stimulus, can be ordered, either increasing or decreasing, from uni^b, over *ain* and *afix* to uni^a, as described in section 3.3.1. For the average profiles (see Figure 3.6C) this was nowhere the case for the *spatially separate*, and essentially nowhere the case in the *transparent* paradigm. Checking the condition solution by solution, the conditions were satisfied for a small fraction of cells of around 0 to 30 % in the *spatially separate* paradigm (similar to what we found in experimental data for MT cells (compare figure 3.3), and for a somewhat higher number of cells (0-70 %, dependent on the stimulus) in the *transparent* paradigm (not so similar to our analyzed experimental data in MT cells).

3.3.4 *Variability in the model with respect to constraining MT responses*

To test both the robustness of results, as well as the variability in the results that is due to the variability in the measured data, it is of interest to compare the solutions of the model when it is constrained not by the tuning curves that are averaged over all cells, but instead by the tuning curves of individual cells. Pilot runs with three different combinations of cells from the *spatially separate* and *transparent* paradigm did, however, not yield any solutions. It is unclear if this is due to the model really not having any solutions in that case, or if they are simply harder to find for the employed algorithm. For one other combination of cells form the *spatially separate* and *transparent* paradigm solutions could be found, on the other hand and they are clearly separated from the solutions that were constrained by the population average tuning curve (see Figure 3.8).

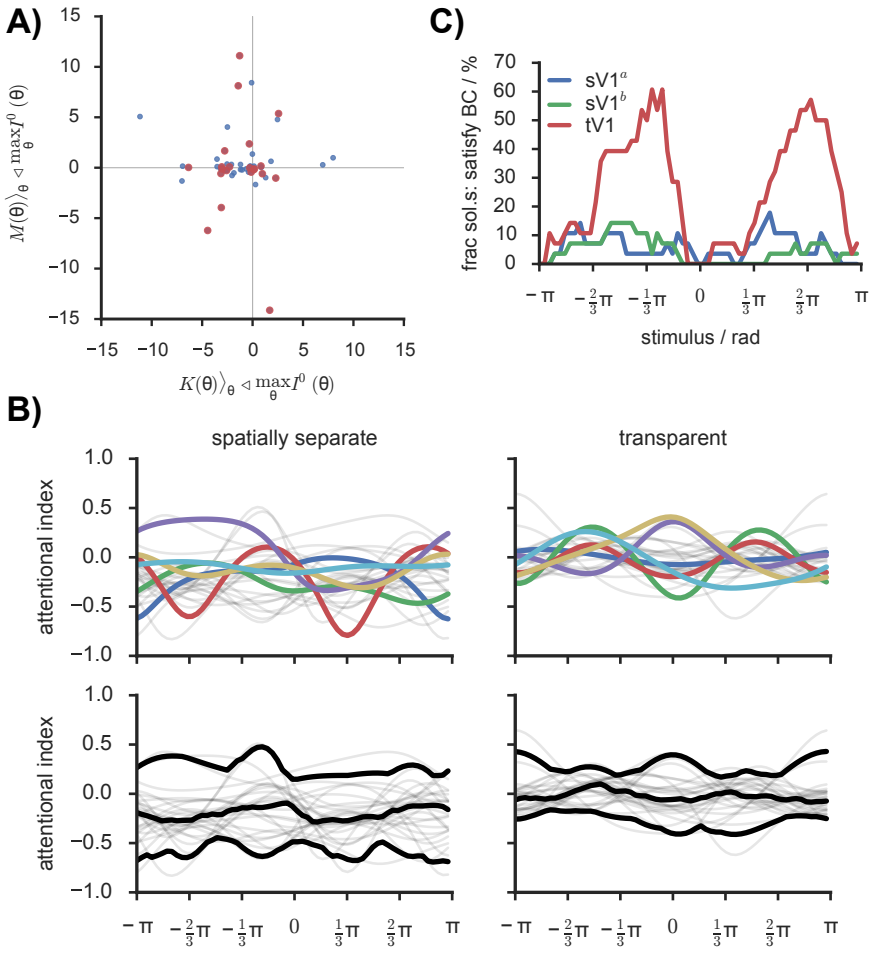


Figure 3.7: Moderately sized attentional currents lead to cell-specific effects. A) For each admissible solution, the strength of attentional currents in the spatially separate (K, x-axis) and transparent (M, y-axis) averaged over all stimuli are shown in units of the maximum strength (over stimuli) of the input I^0 of the corresponding solution. As the latter was always positive, negative values in the plot indicate that attentional currents were inhibitory. We found only two solutions for which both K and M were positive for all stimuli. **B)** Attentional modulation of V1 tuning curves for all admissible solutions (light black), expressed as an attentional index ai between -1 (no response in attend-in condition) and 1 (no response in attend-fix condition). The first row highlights qualitatively ai profiles, the second rows shows the median, as well as 2.5% and 97.5 % confidence intervals, none of which is representative of individual solutions. Strangely, the median is negative in the *spatially separate* paradigm (but in line with Figure 3.6C). **C)** Fraction of V1 solutions satisfying the BC model, depending on stimulus is low, in line with our results for MT cells which constrained the model (cf. section 3.3.1)

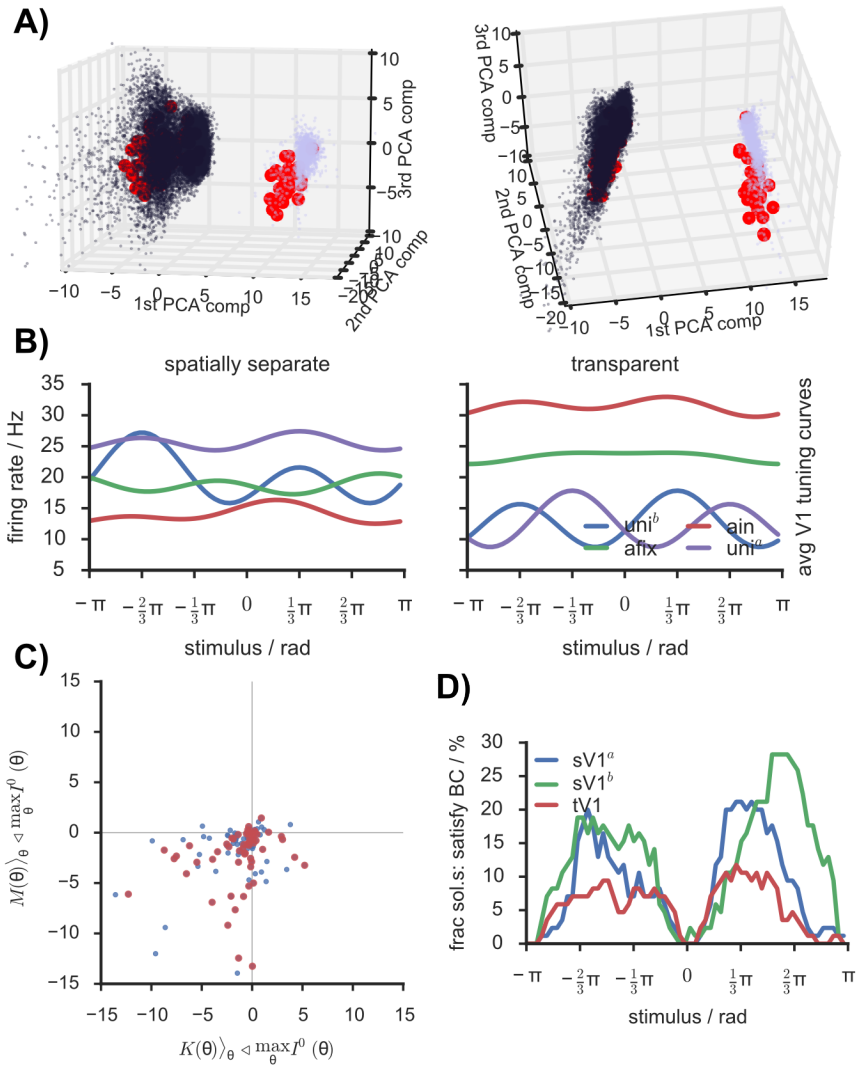


Figure 3.8: Distinct solutions when model is constrained by individual cells' rather than population averaged tuning curves. A) Scatter plot of first three PCA components showing separate clusters of solutions obtained when the model was constrained by the population averaged tuning curve (purple) or one particular combination of individual tuning curves from the *spatially separate* and *transparent* paradigm. Red points in each cluster mark admissible solutions. **B)** V1 tuning curves averaged over all admissible solutions of the black cluster from A). Tuning curve modulations are very different compared to Figure 3.6C. Also, *afix* curves are only weakly angle-dependent, and in the *spatially separate* paradigm peaks are phase-shifted with respect to input positions ($\pm\pi/3$) **C)** Analogous to Figure 3.7A for the black cluster. Attentional currents had an even stronger tendency to be negative than before. Only one of the solutions had attentional currents completely positive for all stimuli. **D)** Fraction of solutions for which V1 tuning curves satisfy BC model very low, now also for *transparent* case.

3.4 DISCUSSION

Having identified problems with existing phenomenological models of attention, we have developed and analyzed a model that aims to explain tuned responses in MT as a result of circuit interactions with a hierarchically neighboring area, V1. Admissible solutions of the model varied qualitatively between each other in terms of circuit interaction profiles, resulting tuning curves and attentional effects. Nevertheless, a majority of them showed baseline reductions in the *afix* and even stronger in the *ain* condition of the *spatially separate* paradigm.

3.4.1 *Phenomenological models of attention*

We found a relatively low number of cells 5% to 30% (depending on stimulus angle and paradigm) that confirm to the biased competition model. While our data might be uncertain to some extent, we would nevertheless expect a higher number of matches if that model provided a *robust* description of the underlying mechanism.

Previously, responses to single and two spatially separated stimuli within the same V4 Ghose and Maunsell [2008] and MT [Lee and Maunsell, 2010] RF were studied whereby the orientation and direction, respectively, of each stimulus was independently chosen to be one out of three possibilities (preferred, anti-preferred or intermediate) so that in total nine stimulus pairs were tested. When attention was directed to either one of the two, the cell's response was similar to the response when the attended stimulus was presented alone. They do not, unfortunately, check quantitatively if their data satisfies the BC conditions, neither for the average data nor for individual cells; for the average data, their figure 4 and 2, respectively, hints, that it does not.

As the stimulus resolution in the dataset was 30°there is some uncertainty when aligning tuning curves to their preferred direction. This potentially affects the determination of the population response curves ⟨ain/afix⟩ and thus our results regarding the FSGM. In the *transparent* condition it was violated clearly, but in the *spatially separate* condition it held approximately, some of the deviation could, thus, be explained by aligning uncertainty. Moreover, due to the variance over trials there is also a large uncertainty in the population response, which could also affect results, in the worst case a large trial-to-trial variability might render analyses based on average effects inappropriate [Golowasch et al., 2002; Marder and Taylor, 2011; Marder, 2011].

3.4.2 *About the model architecture*

The model could be considerably simplified by neglecting the non-linearities f_α. This, is, however, not possible. To see why, consider (3.7) which, without the non-linearities, would read

$$\hat{r}_n^{(\aleph\alpha)} = \hat{I}_n^{(\aleph\alpha)} + 2\pi \sum_\beta \hat{J}_n^{(\alpha\leftarrow\beta)} \hat{r}_n^{(\aleph\beta)} \tag{3.7}$$

Let $\hat{R}_n^{(\aleph)} = (\hat{r}^{(\aleph\alpha)})_\alpha$, be the vector of $\hat{r}_n^{(\aleph\alpha)}$ for all α, likewise $\hat{I}_n^{\aleph} = (\hat{I}^{(\aleph\alpha)})_\alpha$ and let $\underline{J}_n = (\hat{J}_n^{(\alpha\leftarrow\beta)})_{\alpha,\beta}$ the matrix with α counting rows and β columns. The solution can then be written as $\hat{R}_n^{(\aleph)} = (1 - 2\pi\underline{J}_n)^{-1}\hat{I}_n^{(\aleph)}$. Note then that in the spatially separate uni^a condition only the V1^a condition receives an input I^0, such that $\hat{I}_n^{(su\alpha)} = \hat{I}_n^0\delta_{\alpha,V1^a}$ where δ denotes the Kronecker symbol; we will call the corresponding solution $\hat{r}_n^{(su^a)}$ (note that up to here in this chapter V1^a never received an input, so there was no need, until now, to distinguish $r^{(su^a)}$ from $r^{(su^b)}$). Likewise, we have $\hat{I}_n^{(su\alpha)} = \hat{I}_n^0\delta_{\alpha,V1^b}$ for the spatially separate uni^b condition, leading to the solution $r^{(su^b)}$, and $\hat{I}_n^{(tu\alpha)} = \hat{I}_n^0\delta_{\alpha,V1^a} + \hat{I}_n^0\delta_{\alpha,V1^b}$ for both of the transparent uni conditions giving rise to $r^{(tu)}$. With these notations it follows that $\hat{r}_n^{(tu\alpha)} = \hat{r}_n^{(su^a)} + \hat{r}_n^{(su^b)}$. This equation means, that, when tuning curves from all *uni* conditions are aligned to a common input direction, then the two *uni* curves from the spatially separate paradigm would need to add up to the *uni* tuning curve from the transparent paradigm. Moreover, the same reasoning leads to the conclusion that when the uni^a and uni^b curves are aligned with the first and second stimulus component of the *afix* tuning curve, respectively, then the sum of the two *uni* curves must, according to the linear model, necessarily equal the responses of the *afix* curve. These conditions were, however, not fulfilled in our experimental data and we had thus no choice but to incorporate a non-linearity in the model.

A feasible simplification would be, however, to lump together excitatory and inhibitory populations at each locations thereby reducing the number of equations by a factor of two. As, in the current implementation of the model, long-range interaction kernels are independent of the target population being excitatory or inhibitory, the two populations at each locations differ anyway solely due to the non-linearity (see (3.3)), so this simplification seems viable—the cost being some biological realism.

We modeled attentional currents additively and assumed they target only V1 in the *spatially separate* paradigm, and only MT in the *transparent* paradigm. None of these three assumptions is compellent. Indeed, as summarized in the introduction, attention has been modeled in a huge variety of ways, most convincingly in the model by Ardid et al. [2007] where it emerges as a property of the network

due to long-range interactions between areas. Here, we did not attempt to include attentional source areas in the model and in such a case attention has to be included in an ad-hoc fashion—as in most models. While evidence exists that both MT and V1 are anatomically connected to attentional source areas [Markov et al., 2014b] and that at least V4 and MT are functionally affected by these source areas [Squire et al., 2013; Gregoriou et al., 2014], how these can effectively be replaced in a non-recurrent way, is an open question and the answer quite possibly depends on the problem at hand. Therefore, a principled approach to model attention likely needs to consider as well these source areas.

We had six experimental conditions to constrain the model and four free variables (per Fourier order) to solve the equations (3.7). Consequently, if we were to obtain measurements for five or more different stimulus combinations, or measurements of V1 responses under the same stimulus combinations that were used here, the system of equations would be *overdetermined*. Of course, if the model is suitable to describe the experimental data, the system of equations should still be solvable. Consequently, *if* there still existed solutions in that case that would constitute a strong argument in favor of the model.

3.4.3 *What can we predict?*

Any solution of the model fixes a potential circuit which could then be used to predict responses for other stimuli, for example a pair of preferred and anti-preferred directions, which has been studied previously both experimentally [Patzwahl and Treue, 2009] and computationally [Ardid et al., 2007]. Unfortunately, the algorithm that we used to partially decouple the model equations (3.7) cannot be applied in this case and we would have to solve the complete system numerically which is challenging due to its high dimensionality. Indeed, pilot attempts converged rarely and, if so, led to unphysiological results.

The model predicts that numerous response patterns in V1 can be related to the same MT response. A majority of them showed a decrease in activity for all stimuli in the *afix* and even more so in the *ain* condition of the spatially separate paradigm, but not for the *transparent* paradigm where baseline responses were similar in all conditions. We have not observed baseline decreases due to a second stimulus in the RF or due to attention in our MT-data. Moreover, while attentional modulations do vary considerably between individual cells, generally there seems to be a trend to increase rather than to consistently decrease activity.

3.4.4 *Open questions*

Individual solutions showed a large variability. This suggests that a given response (the one measured in MT that constrained the model) can be obtained through a variety of circuit mechanisms which are potentially qualitatively different, i. e. are unconnected in parameter space. Such built-in compensatory mechanisms would allow, for example, a certain leeway in the circuit wiring as long as other connections co-vary. In a different context, Gutierrez et al. [2013] pointed out that qualitatively similar patterns of oscillations could be obtained in a model for the crab stomatogastric ganglion and that the corresponding model parameters are possibly qualitatively distinct, separated in parameter space by phases of qualitatively different dynamics. Here, more systematic attempts to obtain solutions in the parameter region between these solutions are necessary to determine if there exist connections between the regions of parameter space corresponding to the admissible solutions.

On the other hand, V1 is not only connected to MT but also to a variety of other areas [Markov et al., 2014b]. When a variety of response patterns in V1 can give rise to the same response in MT, this does not mean that also all other downstream areas (which we did not model) would respond equally. Moreover, as we and others [Golowasch et al., 2002; Marder, 2011] have pointed out, individual behavior can differ substantially from average behavior. Possibly, then, the cortex needs to have a mechanism to deal with the large individual variability of responses—either compensatory ones in case the variability is just noise, or integrating ones, in case it is informative as has been suggested [Chelaru and Dragoi, 2008; Padmanabhan and Urban, 2010].

Mitchell et al. [2007] measured attentional effects in macaque V4 and, distinguishing putative excitatory and inhibitory neurons on the basis of their spake-waveforms, found that putative interneurons had higher firing rates as well as larger attention-induced increases in absolute firing rate and decreases in Fano factors; Cohen and Maunsell [2009] found, in addition, that noise correlations were more strongly reduced for neurons with higher firing rates, which could, correspondingly, also mean primarily for interneurons. These results could be used to benchmark our model, but not, due to the non-linear recurrent circuitry, to deduce the target of attentional currents.

In our model, attentional currents, K and M were often inhibitory. They targeted, however, both excitatory and inhibitory neuron populations, so that it their net-effect in the circuitry is difficult to predict. Indeed, while in the *transparent* paradigm average responses were not much affected by attention, they were reduced in the majority of solutions in the *spatially separate* paradigm. The differences between the paradigms likely stem from non-linear summation properties of subparts of a cell's RF. Phenomenologically, we capture these by the non-

linearities f_α in our model. But how could an inhibitory attentional current arise? On the one hand, Anderson et al. [2011] found that connections from FEF to V4 target almost exclusively excitatory pyramidal neurons, arguing against such a possibility. However, while FEF is a prominent "source area" for attentional modulations [Moore, 2006], it is only one in a whole network of areas [Corbetta et al., 2008; Markov et al., 2014b], including subcortical regions [Zénon and Krauzlis, 2012], and I am not aware of comparable data for the target specificity of projections from these areas, leaving open the possibility that other areas exert inhibitory influences. Another potential player is neuromodulation which has been demonstrated to be involved in attention [Himmelheber et al., 2000; Herrero et al., 2008; Noudoost and Moore, 2011a,b] and which can exert inhibitory effects (possibly layer-dependent) [Gulledge and Stuart, 2005; Gulledge et al., 2007; Eggermann and Feldmeyer, 2009]. The attentional currents in the model describe the net effect of all attentional source areas and neuromodulation. It is therefore rather difficult to constrain them with other data, pinpoint their best target location (V1 and/or MT, excitatory and/or inhibitory population) or exclude that they could be inhibitory.

3.5 CONCLUSION

It is tempting to relate the outcomes of this model to our previously discussed experimental results (sections 2.3.7 and 3.3.1). Indeed, the high individual variability and the low percentage of solutions that satisfy the BC model strongly resemble our findings in MT and might be taken as predictions for experiments measuring V1 responses for the same stimulus combinations. Note however, that experimentally we determined single-cell tuning curves which were themselves highly variable, while the model predicts population responses for cells with identical tuning curves.

Moreover, the V1 tuning curves in the *spatially separate paradigm* averaged over all solutions show an unintuitive decrease of responses with attention, likewise the average *ain* tuning curve in the *transparent* paradigm features a barely visible peak where attention is directed to the preferred stimulus. These two findings cast doubt on the validity of the model. We already acknowledged that not all solutions of the model are plausible and we filtered these out. This rather inelegant property of the model might have to be extended in order to keep only solutions which do not show this attention-induced decrease in responses.

Moreover, while the model architecture is straightforward, it has the unfortunate property that even once we have determined the circuitry it is technically challenging to obtain predictions for other stimulus combinations as a the roots of a high-dimensional non-linear system have to be determined and we cannot proof its existence nor

its uniqueness, nor guarantee the convergence of an appropriate root-finding algorithm.

Next, it would also be necessary to investigate how robust the results are with respect to the constraining experimental data, as MT responses are themselves highly variable.

Finally, the way we modeled attention is just one ad-hoc possibility and we need to complement it by studying various other ways in which it could influence each of the populations. However, when we allow attention on more than one population that just increases the number of free variables, so we gain limited insight by this choice; when we enforce additional constraints, these possibly hinder the partial decoupling of the system (3.7) so that, potentially, we would have to solve the full 135-dimensional system numerically—a daunting task. A more principle concern is that the visual cortex is *not* a closed system on which attention acts from the outside. In such a closed-system approach, as we have used it, the only insight we can hope to obtain is on the net-effect of attention on the system under consideration, but due to the recurrent circuitry that effect is likely dependent on numerous properties and, thus, of little generality. To understand how neuronal correlates of attention can emerge from within the system we need to consider models that include interactions between sensory and control areas. In the next chapter we will make a first step along the complex path towards such a model.

4

CONNECTOME OF A MODEL LOCAL CORTICAL CIRCUIT FLEXIBLY SHAPES LAYER-DEPENDENT MULTI-FREQUENCY OSCILLATIONS

This chapter is being prepared for submission as a separate publication [?]. I acknowledge my co-authors Xue-Jie Chen, Wei Wei, Fred Wolf and Demian Battaglia.

4.1 ABSTRACT

Lamination is a landmark feature of cortical architecture, but the role played by interlayer connections in shaping local responses and their long-range interaction has not yet been fully elucidated. Here, we analyze a rate model of a canonic local circuit, embedding realistic interlayer connections. Systematically varying the efficacy of excitation and inhibition in the model we find a great diversity of possible dynamical states that can be influenced through feed-forward or feed-back inputs and which include, in particular, oscillatory regimes with layer-specific patterns of phase relations and frequencies. In agreement with experimental observations, in extended regions of the phase diagram gamma-band frequencies predominate in supragranular layers while beta-like or slower frequencies predominate in infragranular layers. While this tendency is often attributed to different cortical sources to different layers or to layer-specific intrinsic resonances, here it arises as a result of inter-layer connectivity. Comparisons with an alternative model including as well a slow-oscillating inhibitory population in infragranular layer 5 show that both intrinsically and network-generated slow oscillations could co-exist in the same circuit and may thus be hard to disentangle. Finally, by randomizing inter-layer connectivity, we reveal that only a tiny minority of rewired connectomes leads to oscillatory behaviors compatible with experimental observations. At the same time, we find that remarkably similar target oscillatory repertoires may stem from well distinct connectome classes. These results suggest that functional homeostasis mechanisms may operate to drive inter-layer connectivity toward one out of many possible target templates, making the local circuit equally fit to integrate and multiplex signals from several sources in multiple frequency bands.

4.2 INTRODUCTION

The microcircuits of mammalian brain cortex present prominent hallmark features such as a six-layered architecture [Brodmann, 1909] and an organization into vertical columns [Mountcastle, 1957]. Remarkable regularities in the wiring between neuronal populations in different layers, both within and between columns [Maunsell and

van Essen, 1983; Gilbert and Wiesel, 1989; Douglas et al., 1989; Douglas and Martin, 1991, 2004; Kätzel et al., 2011] have led to propose the existence of a "canonical local microcircuit", providing a building block for larger scale cortical networks. This hypothesis, culminated in the compilation of quantitative inter-layer wiring diagrams [Thomson and Bannister, 2003; Binzegger et al., 2004; Haeusler and Maass, 2007], despite its approximations and limits [Douglas and Martin, 2007], has proved particularly attractive, bolstering theories about specialized roles of different layers [Raizada and Grossberg, 2003; Bastos et al., 2012; Self et al., 2013] as well as motivating computational investigations of the local microcircuit function [Traub et al., 2005; Markram, 2006; Haeusler and Maass, 2007; Wagatsuma et al., 2011, 2013; Potjans and Diesmann, 2014].

The layer structure of cortical circuits also affects the generation and the properties of brain oscillations. Oscillations of neural activity in different layers have different spatial spread and power [Xing et al., 2009, 2012b] and display spectral resonances at layer-dependent frequencies [Roopun et al., 2006; Sun and Dan, 2009; Buffalo et al., 2011; Ainsworth et al., 2011]. These findings have been put in relation with the proposal that inter-areal communication in different directions exploits functional interactions mediated by oscillatory coherence in different frequency bands [Saalmann et al., 2012; Kerkoerle et al., 2014; Bastos et al., 2014, 2015b]. For instance, the fact that bottom-up interactions occur in the gamma band and that top-down interactions occur in the beta or the alpha band [Kerkoerle et al., 2014; Bastos et al., 2015b] may be explained in terms of the different laminar origins of ascending and descending structural connections, which proceed respectively from more superficial and deeper layers [Douglas and Martin, 2004], in combination with the characteristic oscillatory resonances of these different source layers, with gamma band oscillations arising in more superficial layers and slower oscillations dominant in deeper layers [Roopun et al., 2006; Buffalo et al., 2011].

These correlational observations do not allow recognizing whether local microcircuit connectivity plays an actual causal role in shaping layer-specific properties of neural oscillations and of their flexible modulation by context. To evince the existence of a direct influence we adopt here a computational approach. Through systematic simulations of a rate model embedding an anatomically realistic multi-layer connectivity [Binzegger et al., 2004], we reveal that inter-layer interactions give rise to a rich repertoire of possible oscillatory modes. This dynamical repertoire includes robust regimes in which the laminar separation between slower and faster oscillations emerges purely through network mechanisms, not necessarily requiring the introduction of neuronal populations with distinct intrinsic resonance frequencies, unlike in other models [Lee et al., 2013].

Finally, we inquire whether the observed effect of inter-layer connectivity on multi-frequency oscillatory activity is to be attributed to unique properties of the adopted connectome, or it is on the contrary a general outcome of unspecific inter-layer interactions. To do so, we compare the self-organized oscillatory dynamics of our model with realistic anatomy with the one generated by models with randomized connectomes. We find that random connectomes do not lead to a natural layer separation between slower and fast oscillations. On the contrary, we are able to generate such a dynamical regime only by carefully selecting the inter-layer wiring among a highly restricted set of connectivity configurations, which include the reference connectome by [Binzegger et al., 2004], but other very diverse connectomes as well. Therefore, despite this lack of uniqueness, we establish evidence that the empirical canonical microcircuit belongs at least to a very exclusive club of degenerate structures all achieving a common target behavior, in a way reminiscent of variability compensation and homeostasis mechanisms in other neuronal systems [Prinz et al., 2004; Marder and Goaillard, 2006; Goaillard et al., 2009]. This suggests that peculiar micro-circuit patterns are selectively developed and maintained through evolution and development, because of the functional advantages for inter-areal communication that the associated multi-frequency oscillatory "dynome" [Kopell et al., 2014] confers.

4.3 RESULTS

4.3.1 *Rate model of the canonical local circuit*

We analyzed the dynamics of a rate model of a local cortical circuit, whose realistic connectivity, illustrated in Figure 4.1, was inspired from anatomical studies. It consisted of five layers (L1, L2/3, L4, L5 and L6), each containing one excitatory and one inhibitory unit which represented the mean activity in the corresponding population. The connection weights between all populations were taken to be proportional to the relative number of synapses between excitatory (E) and inhibitory (I) populations as measured by Binzegger et al. [2004] (see Figure 4.1A–D). The actual strengths of inter-population connections were then obtained by multiplying these relative numbers by two phenomenological parameters, K_E and K_I, indicating global scales of the strengths of excitatory and inhibitory connections, respectively. We assumed here, that these parameters depended only on the type of connection (E or I) but were otherwise independent from the target population. Note that Binzegger et al. [2004] did not report outgoing connections from L1. Therefore, we ignored L1 in the following, given its lack of influence on the dynamics of the other layers, reducing correspondingly the number of neuronal populations explicitly included in the model to eight.

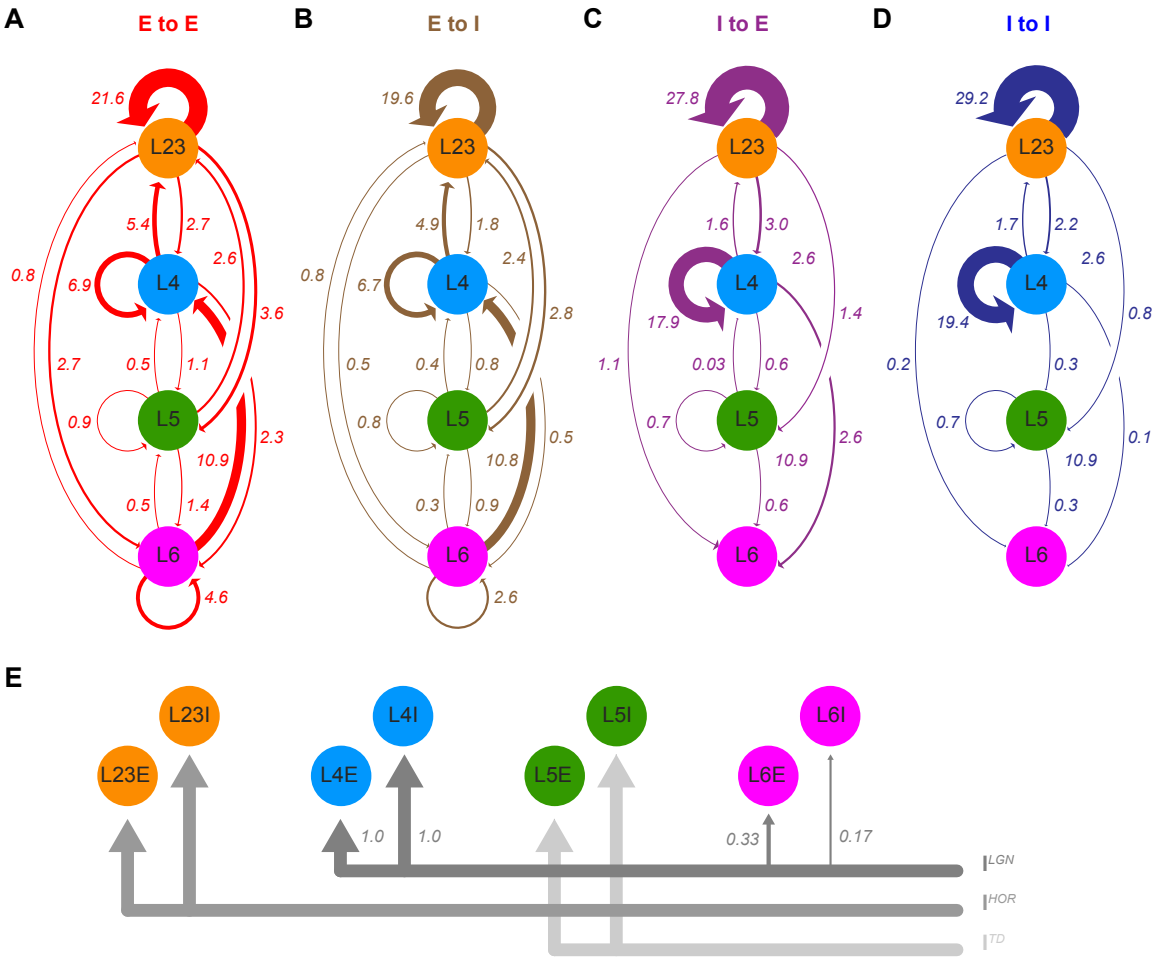


Figure 4.1: Local circuit connectivity. All layers contained one excitatory and one inhibitory rate unit. Relative strength (arrow labels) of **(A)** excitatory to excitatory, **B)** excitatory to inhibitory, **C)** inhibitory to excitatory and **D)** inhibitory to inhibitory recurrent connections was taken from [Binzegger et al., 2004, Fig. 12], while their absolute strength was set by two phenomenological parameters $K_{E,I}$ representing the efficacy of excitatory and inhibitory spikes, respectively. All connections were delayed. **E)** Extrinsic connections were sent to specific layers only. Bottom-up input from LGN targeted layers 4 and to a lesser extent 6 (numbers indicate their relative weight), horizontal connections layer 2/3 and top-down connections layer 5.

Connections in our model did not give rise to instantaneous interactions, but were delayed by an amount D, assumed for simplicity to be identical for intra- and inter-layer monosynaptic connections. The inclusion of delays favored the emergence of oscillations in the firing rates when the column was stimulated by a constant baseline background current I^{bg} (received by all populations). In addition, other external inputs were targeting specific layers (Figure 4.1E). As the anatomical connections were measured in area 17, bottom-up input I^{LGN} from LGN was sent to L4 and, to a lesser extent, L6, con-

sistent with the literature [Felleman and Van Essen, 1991; Douglas and Martin, 2004; Binzegger et al., 2004; Sincich and Horton, 2005]. Likewise, we mimicked horizontal connections from other columns within the same cortical region via an input I_{HOR} specific to L2/3 and "top-down" connections from extra-striate cortices via an input I^{TD} specific to L5. By varying the levels of these inputs and their relative balance, different perceptual and cognitive contexts can be phenomenologically emulated. For instance, increasing I^{LGN} can represent an increase of contrast of a centrally presented stimulus, and enhanced I^{HOR} or I^{TD} the presence of modulatory signals from, respectively, the classical or the extra-classical surround of the local circuit receptive field [Angelucci and Bressloff, 2006]. Alternatively, top-down modulatory inputs from higher order cortical areas, such as, e.g., an "attentional spotlight" sent by prefrontal areas [Moore, 2006] may be represented by a simultaneous increase of both I^{TD} and I^{HOR} (analogously to [Wagatsuma et al., 2011]).

4.3.2 *Model supports a rich dynamical repertoire*

For each fixed value of the background and external inputs, and depending on the efficacies of excitatory and inhibitory interconnections, the model local circuit gave rise to a large repertoire of different possible dynamical states, including steady, oscillatory or chaotic firing modes. To assess the behavior for different values of K_E and K_I, we simulated the activity of the model in different regions of the parameter space and systematically extracted four summary statistics (see Materials and Methods) for each layer which, taken together, provided a qualitative profile of the dynamical regime. The first metric was the mean firing rate. The second was the relative fraction of low power, defined as a layer's integrated power spectrum below 30 Hz, P_{lo}, divided by the summed power spectrum over all frequencies, P_{tot}. The amplitude and the lag of the peaks (excluding the central, zero-delay one) of the autocorrelation of each layer's trace provided then the two last metrics. The lag of the first peak measured the period of the fastest appreciable oscillatory structure in the time trace (for easier assessment we present it as a multiple of L4E's delay), whereas the value of the highest peak in the autocorrelation quantified the degree to which the trace was either more periodic (autocorrelation has values close to 1) or more chaotic (autocorrelation values close to 0). Plots of these quantities in dependence of K_E and K_I, which we will refer to as *dynamic regime profiles*, summarize the behavior of the model circuit.

We first studied the behavior of the column in a condition of exclusive bottom-up drive ($I^{LGN} = 2$, $I^{hor} = I^{td} = 0$), besides the normal background inputs. The dynamic regime profiles of Figure 4.2A (see Figures 4.3 and 4.4 for other input configurations) clearly show

the existence of homogeneous phases, i.e. regions of the parameter space with qualitatively distinct but internally homogeneous dynamics. While firing rates varied smoothly with K_E and K_I, sharp transitions were revealed by the three other monitored metrics.

For weak inhibition K_I the dynamics settled in a regime characterized by constant levels of activity in all layers (representative traces are shown in Figure 4.2B, ♦ marker), as it would be observed in the case of asynchronous neuronal population firing with homogeneous rate. A rate instability line was crossed for too strong values of K_E beyond which firing rates diverged (white region at the top right of the dynamic regime profiles). The homogeneous rate regime lost its stability as well for larger inhibition strengths. When the absolute value of K_I was gradually increased periodic oscillations emerged (traces in Figure 4.2B, ● marker), in which all the layers oscillated with a fast frequency (in the gamma range) and similar relative amplitudes. When further increasing K_I, more complex oscillatory patterns were observed (example traces in Figure 4.2B, ■ marker and Figure 4.5). Among them were apparent chaotic and periodic oscillations distinguished by, respectively, low and high values of the autocorrelation's highest peak (fourth row of Figure 4.2A), and which appeared to be locked across layers at various frequency ratios as indicated by the delay of the autocorrelation's first peak (third row of Figure 4.2A).

Note that the parameters K_E and K_I, fixed for each given simulation, could in reality be seen as effectively fluctuating around an average working-point, phenomenologically representing the effects of, e.g., neuronal adaptation [Ulanovsky et al., 2004]), short term plasticity [Markram and Tsodyks, 1996] or neuromodulation [Brunel and Wang, 2001]. In that sense, the complete dynamic regime profile is indicative of the diverse dynamical repertoire that the activity of a local circuit may sample at different times, especially if the working point is close to different phase boundaries, which can then be crossed by only slight changes of parameters (see Discussion).

For better visualization, we summarized all the qualitatively different dynamical modes in a cartoon regime profile, sketched in Figure 4.2C. Even more phase subdivisions could be generated by inspecting the relative phase of oscillation of the different layers. Before analyzing inter-layer phase differences, however, we will first study the dominant frequencies of oscillations, which happen to be strongly dependent on the considered layers and dynamical regimes.

4.3.3 *Fast and slow oscillations dominate respectively in superficial and deep layers*

The ■-marker oscillatory trace in Figure 4.2B clearly shows that spectrally rich oscillatory patterns can be generated by our model, simultaneously involving multiple faster and slower frequencies which

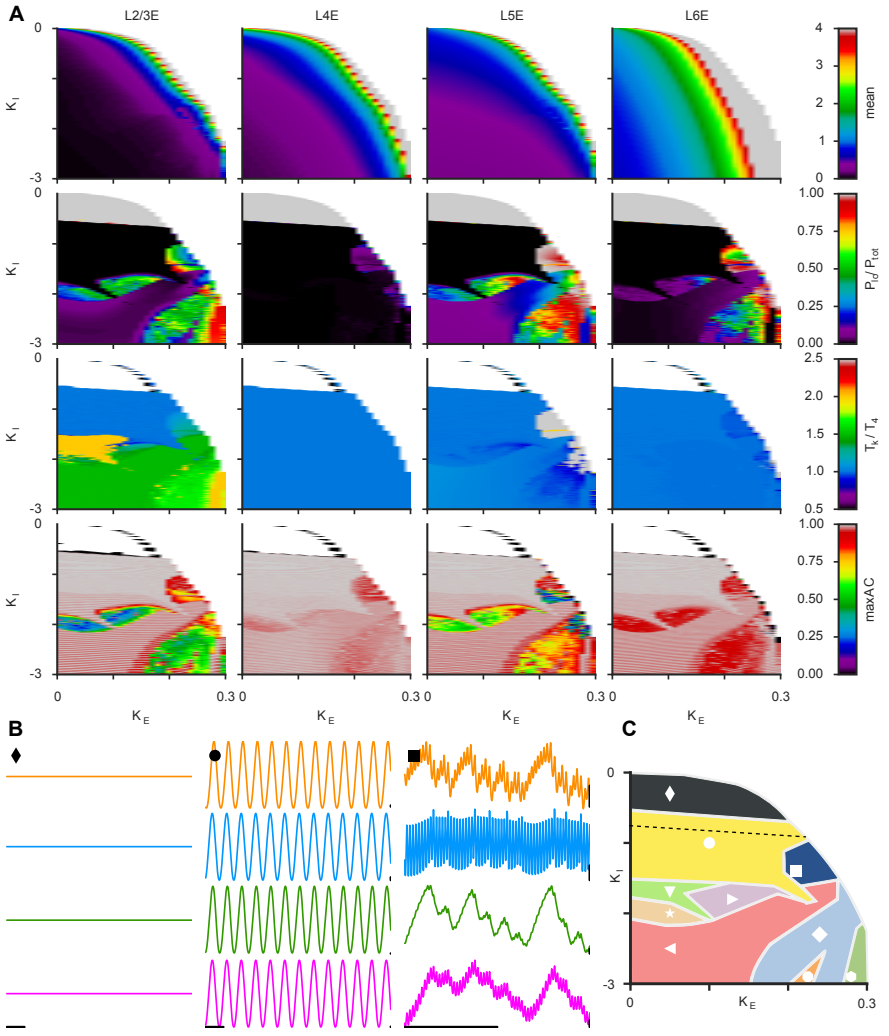


Figure 4.2: Model yields various qualitatively different kinds of dynamics. **A)** Average activity (first row), fraction of low power (second row), delay of autocorrelation's first peak in units of this quantity for layer 4 (third row), and value of autocorrelation's highest peak (fourth row) provide summary statistics—shown for all excitatory layers (column)—to assess the possible dynamics of the local circuit for various values of excitatory and inhibitory efficacies (x and y axis). Summary statistics taken together clear boundaries between different regimes can be made out. **B)** Exemplary traces for 3 working points (indicated by symbols in upper left) for oscillations that are asynchronous, fast in all layers, and fast and slow predominantly in upper and lower layers, respectively. **C)** Summary phase cartoon indicating regions of qualitatively similar dynamics. Working points from B are marked, as are additional ones the traces of which are shown in 4.5 Fig.)

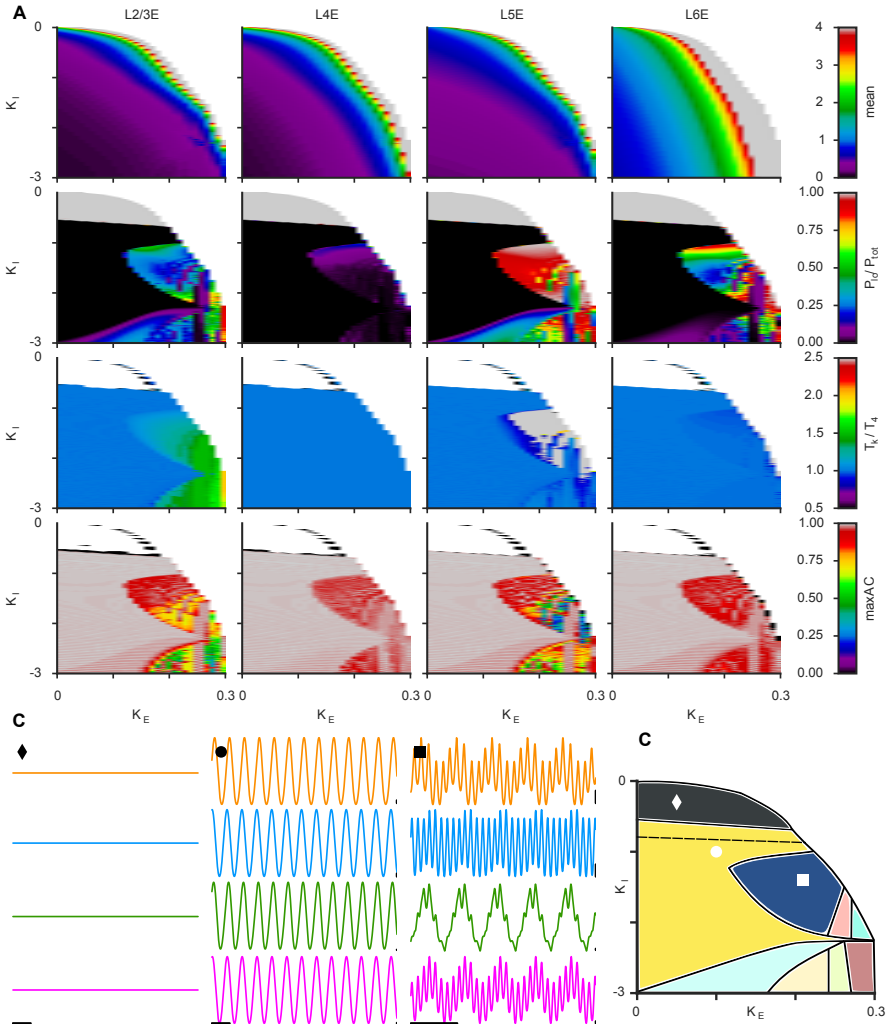


Figure 4.3: Dynamical regime profiles under bottom-up and horizontal stimulation. Like fig. 4.2, but for bottom-up + horizontal input ($I^{LGN} = I^{hor} = 2$, $I^{td} = 0$). Asynchronous, all-fast and mixed fast and slow oscillatory phase are still visible, and are marked in the same color (dark gray, yellow and dark blue, respectively) in **C**).

are expressed more or less prominently depending on the considered layer. Notably, in the specific case of the ■-marker working point (briefly referred to in the following as the "fast/slow working point"), while L4 activity is dominated by a fast gamma-band oscillatory component around 71 Hz, which is also strong in L2/3, other layers like L5 and L6 are clearly dominated by slower oscillatory frequencies, e.g. in the alpha or beta range. This fact is intriguing, since as, previously mentioned, such dichotomy in the frequency of neuronal rhythms between superficial and deeper layers occurs also experimentally [Roopun et al., 2006; Sun and Dan, 2009; Buffalo et al., 2011; Ainsworth et al., 2011].

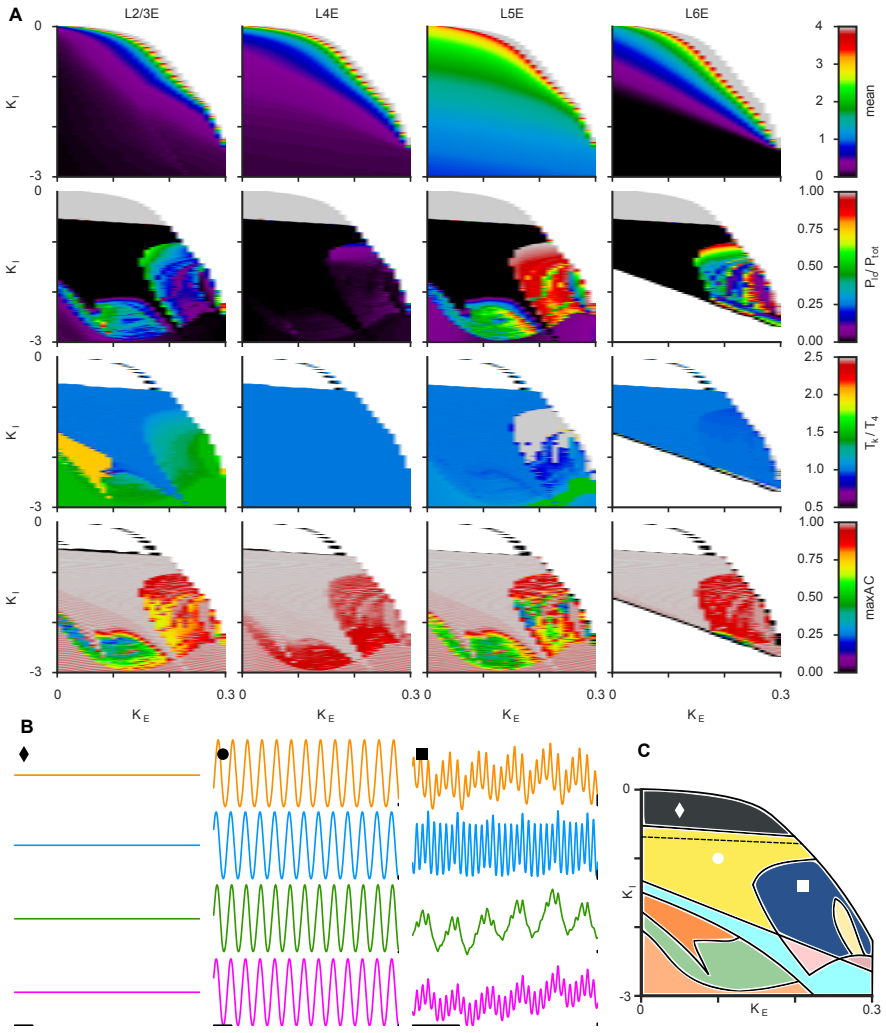


Figure 4.4: Dynamical regime profiles under bottom-up and top-down stimulation. Like fig. 4.2, but for bottom-up + top-down input ($I^{LGN} = I^{td} = 2$, $I^{hor} = 0$). Asynchronous, all-fast and mixed fast and slow oscillatory phase are still visible, and are marked in the same color (dark gray, yellow and dark blue, respectively) in **C**).

The pattern of layer-specificity of frequency shown by the "■" fast/slow working point is not obtained by a careful tuning of parameters, but represents a feature shared by an entire region in the K_E/K_I plane, which we call the "fast/slow region" and which is represented in dark blue color in Figure 4.2C . For each point belonging to this region (see Materials and Methods for details about the algorithm for the determination of its boundaries), we evaluated the relative fractions of low power P_{lo}/P_{tot}, separately for each layer. As shown by Figure 4.6A (computed for the same bottom-up input context as in figure 4.2), within this region, deep layers' activities—particularly L5's—was generally dominated by slow oscillations, whereas L4 developed almost exclusively fast frequency oscillations. L2/3 manifested an in-

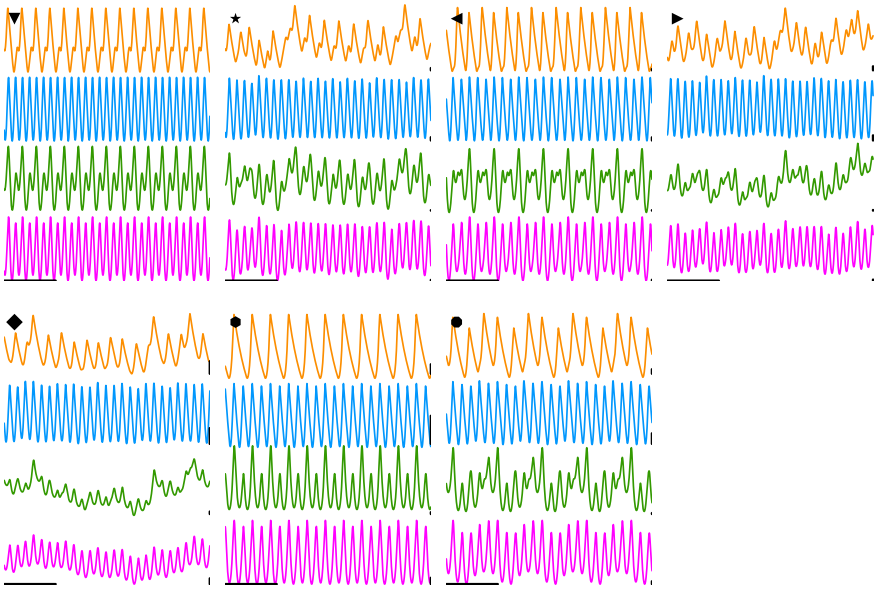


Figure 4.5: Additional example traces. Like fig. 4.2C, but for the remaining working points marked in fig. 4.2C.

termediate behavior, with a balanced median value of P_{lo}/P_{tot} across the whole region of approximately 50 %.

The size and the specific frequency distribution of the fast/slow region depended on the specific context, determined by the applied values of the bottom-up I^{LGN}, horizontal I^{HOR} and top-down I^{TD} driving inputs. In absence of horizontal or top-down drive, the size of this region was relatively constant for a wide range of contrasts, I^{LGN} (Figure 4.6B), and for each of them the average distribution of high and low power over the layers (Figure 4.6C) followed the same pattern as in Figure 4.6A.

Stronger effects were observed by changing horizontal and top-down drives. Dynamic regime profiles computed for different contexts (see Figures 4.3 and 4.4 for applied horizontal and top-down input, respectively), showed that qualitatively similar phases and equivalently rich dynamical repertoires continued to exist. Nevertheless, phase boundaries were distorted, producing variations of the size and the position of the phases summarized by Figure 4.2C. The fast/slow phase was very robust and still continued to include the ■ marker point on the K_E/K_I plane. The layer-dependent pattern of P_{lo}/P_{tot}, however, did change quantitatively. The distributions shown in Figure 4.6A were altered into the ones displayed by Figure 4.6D (by adding an horizontal input $I^{HOR} = 2$) or in Figure 4.6G (by adding a top-down input $I^{TD} = 2$). In both these cases, the power balance of the more superficial L2/3 was shifted toward the faster frequencies of L4, while L5 continued to be dominated by slow frequencies. The dominance pattern of fast and slow frequencies was essentially maintained—despite some non monotonic changes in L6—over the

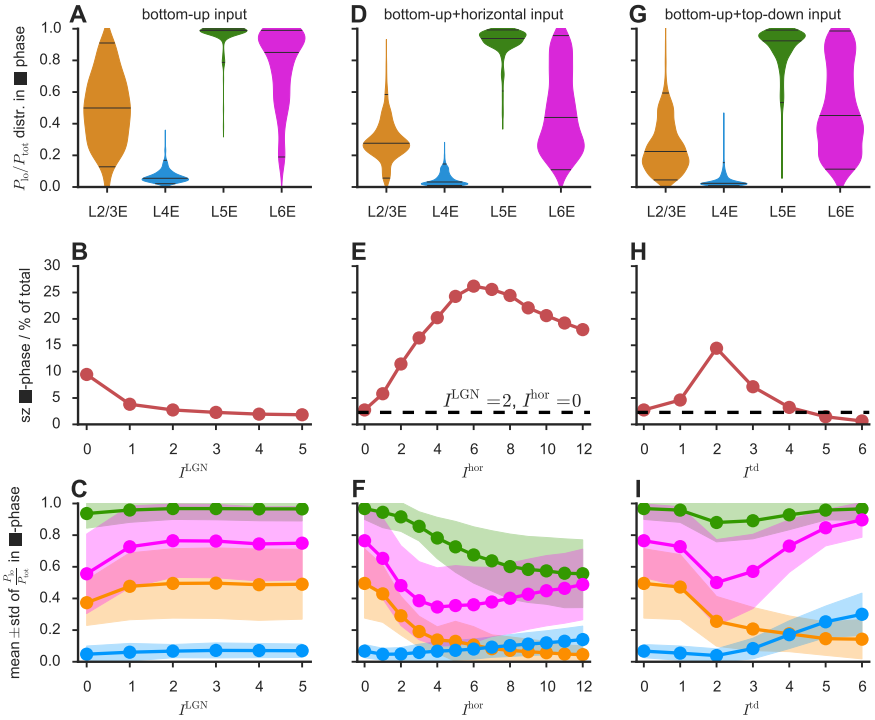


Figure 4.6: Dependence of slow/fast-region on context. A) Distribution of relative amount of low power P_{lo}/P_{tot} over all pixels of the ■-region for bottom-up input. Deep layers were dominated by slow, L4E by fast, and L2E expressed both fast and slow oscillations in this region. **B)** Size and **C)** average (shaded areas denote ± one standard deviation) amount of P_{lo}/P_{tot} of the ■-region were relatively constant with the strength of the input. **D-F)** Like A-C but for simultaneous bottom-up and horizontal input ($I^{LGN} = I^{hor} = 2$). The size of the region grew markedly and for strong I^{HOR} layer 2/3,5 and 6 acquired more gamma power, L2/3 even more than layer 4. **G-I)** Like A-C but for simultaneous bottom-up and top-down input. ($I^{LGN} = I^{TD} = 2$). The size of the phase grew for certain values of I^{TD} and for strong I^{TD} layer 2/3 became more gamma-dominated than layer 4 while the deep layers could retain their preference for slow frequencies.

entire explored ranges of I^{HOR} and I^{TD}, and the segregation of fast and slow frequencies in superficial and deep layers became even stronger for strong top-down inputs (see Figures 4.6F,I). In addition, changes of horizontal and top-down drive also affected the size of the fast/slow phase. Unlike in the case of bottom-up drive I^{LGN}, changes in size with I^{HOR} and I^{TD} were non monotonic, with the fast/slow region reaching a maximum size for intermediate values of these contextual inputs. Overall—even if the details of the dependency on context depend most likely on a variety of factors which in our model have been fixed somewhat arbitrarily—these analyses show a tendency of the fast/slow region to expand in presence of applied inputs to L2/3 and L5 (as expected in presence of attentional modulation), increasing the likelihood that the system dynamics enters in the fast/slow regime if initially tuned to a baseline working point close to its critical boundaries.

4.3.4 *Dynamic phase leadership hierarchy between cortical layers*

Besides variations in their spectral amplitudes at different frequency bands, the oscillations generated by our model display also a remarkable diversity of possible inter-layer phase locking modes as a function of both the strengths of excitation and inhibition and of context, as we analyze in Figure 4.7. Given the multi-frequency nature of the generated oscillations, we analyzed average phase differences in different frequency bands, first for the slow oscillatory components (between 10 and 20 Hz) and then for the fast gamma oscillatory components (between 50 and 60 Hz). In both cases, the inspection of phase differences on the K_E-K_I-plane allows recovering a phase structure largely matching the one extracted from the four metrics of Figure 4.2 (and Figures 4.3 and 4.4).

Interestingly, an additional subdivision is manifested within the periodic oscillation phase (denoted by a ● marker in Figure 4.7). In Figure 4.7A, we consider first the case of exclusive bottom-up drive, in absence of other contextual modulations ($I^{LGN} = 2$, $I^{hor} = I^{td} = 0$). Average phase-differences relative to a reference layer are shown as a function of K_E and K_I. For sufficiently low inhibition (small absolute value of K_I) the oscillations of all layers were in-phase in all frequency bands. This is visible also from the detailed plot of the phase-locking-value (average of $\exp(i\Delta\phi)$ over the frequency band, see [Lachaux et al., 1999] and Materials and Methods) for the ♦-marked working point. However, increasing inhibition, while remaining within the boundary of the ●-marked periodic oscillatory phase, out-of-phase inter-layer phase relations emerged. This finding is in line with rigorous theoretical results about the phase-locking behavior of simpler models with symmetrically interacting oscillatory neuronal populations [Battaglia et al., 2007]. In the periodic out-of-phase regime, L2/3

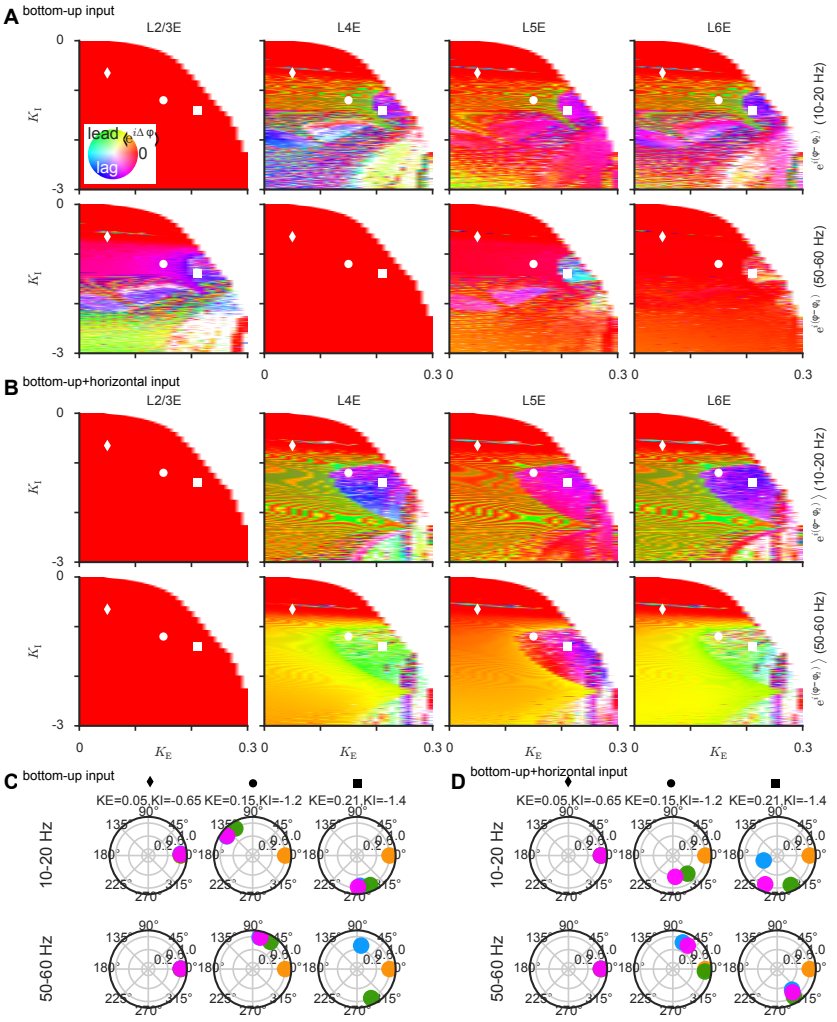


Figure 4.7: Inter-layer phase differences are dynamic. A) Plots of the phase difference of oscillations within each layer (L2/3 to L6, from left to right columns) relatively to a reference layer, averaged over different frequency bands: slow oscillations (top row, 10–20 Hz band, L2/3 used as reference layer) and fast oscillations (bottom row, 50–60 Hz band, L4 used as reference layer). Relative phase differences are evaluated for different strengths of excitations and inhibition, in presence of bottom-up input only ($I^{LGN} = 2$, $I^{hor} = I^{td} = 0$). Yellow to green color denotes phase leadership with respect to reference, blue to violet color phase lagging. Phase relationships between layers strongly depend on the dynamical regime. **B)** Same information as in panel (A), but in presence of an additional horizontal input ($I^{LGN} = 2$, $I^{hor} = 2$, $I^{td} = 0$). Phase relationships between layers are also modified by context. **C)** Polar plots of the phase locking values of different layers (color codes as in Figure 4.1) with L2/3, in different frequency bands (top, 10–20 Hz; bottom, 50–60 Hz) and for different working points (in the positions denoted by the corresponding symbol in panels (A–B)), evaluated in presence of bottom-up input only (as in panel (A)). D) Same information as in panel (C), but in presence of an additional horizontal input (as in panel (B)).

always phase-lagged with respect to L4, who was the absolute phase leader among all layers for both slow and fast oscillation frequencies (cf. also the detailed coherency plots for the ●-point in Figure 4.7C, middle column).

In the slow/fast-region (including the ■-marked working point), phase relationships were different for different frequency bands. As evident from the right column of Figure 4.7C, L2/3 was the leader layer in the slow frequency band, while L4 continued to lead in the fast gamma band.

Addition of a horizontal input ($I^{LGN} = 2$, $I^{hor} = 2$, $I^{td} = 0$), lead, as previously mentioned, to an expansion of the fast/slow region. As a result, the ●-marked working point, previously well within the periodic out-of-phase locking regime was incorporated into the periphery of this expanded fast/slow phase, causing L2/3 to switch from a phase laggard into a phase leader role in the slow frequency band (compare the middle column of Figure 4.7D with Figure 4.7C). The leadership of L2/3 over the other layers was even more distinct at the ■-marked working point, for which, in presence of horizontal inputs, L2/3 led in phase over L4 even in the fast gamma band and not only in the slow frequency band, as in exclusive presence of a bottom-up input (compare the right column of Figure 4.7D with Figure 4.7C).

A more systematic study (with improved evaluators) of the dependency of inter-layer phase differences on context is deferred to future studies.

4.3.5 *Layer interactions cause frequency segregation*

After characterizing the fast/slow phase and its robustness under variations of the input configuration, we investigated possible dynamical mechanisms that may fundamentally explain its origin. In particular, we probed the causal role played by inter-layer interactions. In order to prove their importance, we tuned systematically the strength of all excitatory and inhibitory inter-layer couplings, by multiplying them by a factor Γ, varying between 0 (no interlayer connections) and 1 (default interlayer connection strength as above). While rescaling inter-layer interactions, we maintained an unaltered connectivity within each layer, in order to retain specific properties like the intrinsic oscillatory frequencies of each of them. We focused on the ■-working point, in exclusive presence of bottom-up input, as in Figure 4.2.

At $\Gamma = 0$ (Fig. 4.8A, top panel) layers 2 and 4 oscillated periodically at around 60 Hz and 70 Hz, respectively. Layers 5 and 6, on the other hand, didn't oscillate as inhibition was too weak in layer 5, and not fed back to the excitatory population in layer 6 [Battaglia et al., 2007]

Once interlayer interactions were turned on ($\Gamma = 0.1$, Fig. 4.8A second panel), power spectra showed a multitude of peaks in all layers

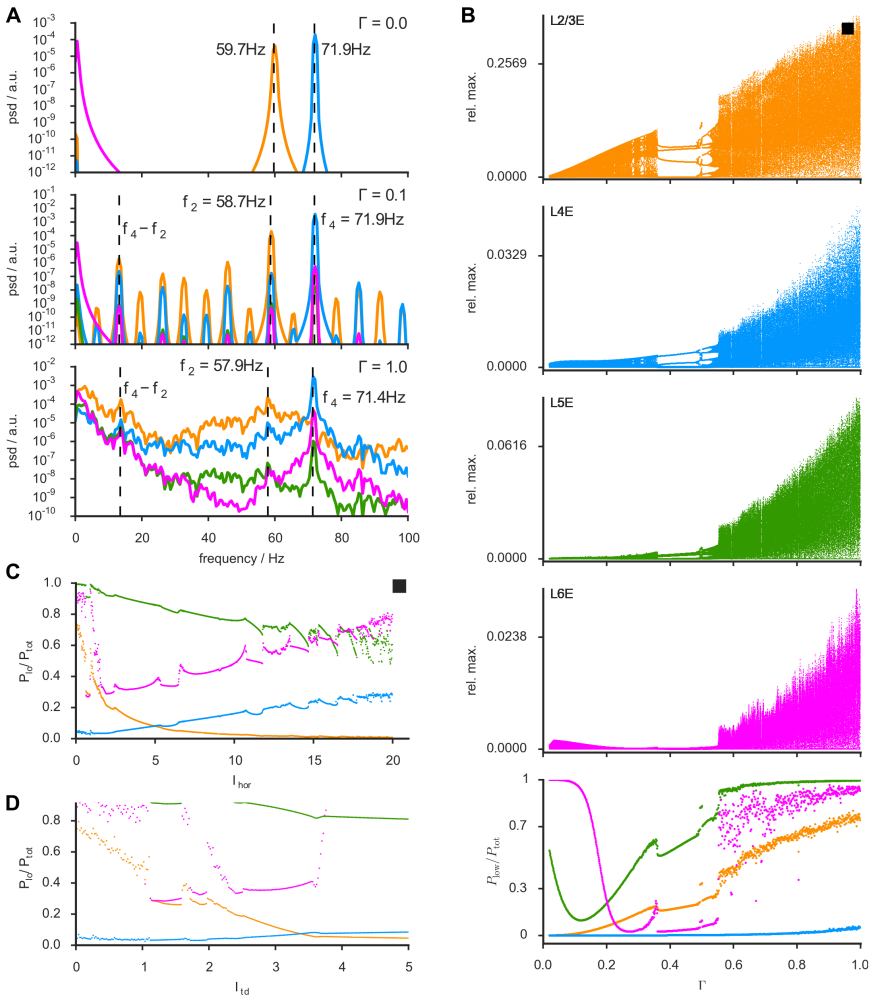


Figure 4.8: Layer interactions cause frequency segregation. A) Power spectra for decoupled ($\Gamma = 0$), slightly coupled ($\Gamma = 0.1$) and fully coupled ($\Gamma = 1$) layers at the ■-working point. Intrinsic ($\Gamma = 0$) layer 2 and 4 frequencies, f_2 and f_4, respectively, exist, slightly changed, until full coupling. Deep layers don't oscillate intrinsically, but only due to layer interactions. A peak in the power spectrum at $f_4 - f_2$, contributing to P_{lo}, indicates that the system undergoes ($\Gamma = 0.1$) or is at least prone to undergo ($\Gamma = 1$) quasi-periodic oscillations and this mechanism might thus decisively contribute to slow oscillations in the model column. **B)** Bifurcation diagrams (first 4 rows) depicting all local maxima in a layer's activity trace and fraction of low power (last row), depending on interlayer coupling multiplier Γ, show that, at the ■-working point, the system undergoes a transition to chaos at around $\Gamma = 0.5$ and concomitantly develops a predominance of slow oscillations in layer 5E and 6E. **C,D)** Like in B, where Γ was varied, gradually increasing horizontal (C) or top-down (D) inputs (under full interlayer-coupling $\Gamma = 1$) leads to apparently discontinuous jumps in the relative amount of low power.

that continued to exist, at slightly varying frequencies, for higher values of Γ, until $\Gamma = 1$ (figure 4.8A third panel). Among them were peaks very near both the intrinsic layer 2 and 4 frequencies, f_2 and f_4, indicating that fast oscillations in the model were related in origin to the inherent dynamics of the upper layers. In addition, the power spectra developed a peak in the alpha/beta-range at the difference of these two frequencies, at $f_{diff} = f_4 - f_2$ suggesting that the oscillations were quasi-periodic. Moreover, as this peak was absent at $\Gamma = 0$ we conclude that slow oscillations in the model occurred due to interactions between the layers.

At about $\Gamma = 0.5$ oscillations became apparently chaotic as can be seen from the broadening of the power spectrum (not shown) and the bifurcation diagrams in Fig. 4.8B which depict all local maxima of each layer's firing rate trace depending on Γ. This transition to chaos was accompanied by a sudden increase of low power in layers 2/3, 5 and 6 (Fig. 4.8B, last row). As Γ was further increased the range of firing rates at which local maxima occurred in the layers' traces broadened with layer 2/3 being much more variable than other layers, which reflects the superposition of similarly strong slow and fast oscillatory components in this layer.

Finally, the power spectrum at $\Gamma = 1$ (Fig. 4.8A, last row), shows in detail the full model's dynamics in the frequency domain. Interestingly, it still exhibited peaks at f_2, f_4 and $f_4 - f_2$ within a broad-band power spectrum, so although the dynamics was chaotic and not quasi-periodic any more, a "ghost" of quasi-periodicity continued to shape the behavior.

Experimentally, it would be hard to probe the dynamics at various strengths of the interlayer coupling and therefore the results presented in this section so far might not be measurable. Accessible, on the other hand, could be contextual modulations of a column's dynamics. Figs. 4.8C,D show the relative amount of low power for increasing values of I^{HOR} and I^{TD} (note that in contrast to Figs. 4.2C,F,I these are now not averaged over a whole phase). These plots feature several apparently discontinuous jumps indicative of regularity windows (see figure 4.9). These windows might be "Arnold tongues" which would be expected to appear if our hypothesis of a mechanism that generates quasi-periodic oscillations was true and might even be visible in experiments that systematically stimulate the column with a fine step size.

In summary, the presence of a peak in the power spectrum at the difference frequency, f_{diff}, as well as the recurring windows of periodicity of varying cycle length in the bifurcation diagrams (Fig 4.8B) suggest that the system's phase space possesses an invariant torus, or a ghost thereof, that is, the system undergoes or is prone to undergo quasi-periodic oscillations and it depends on the exact parameters if the oscillations are indeed quasi-periodic, chaotic or even periodic.

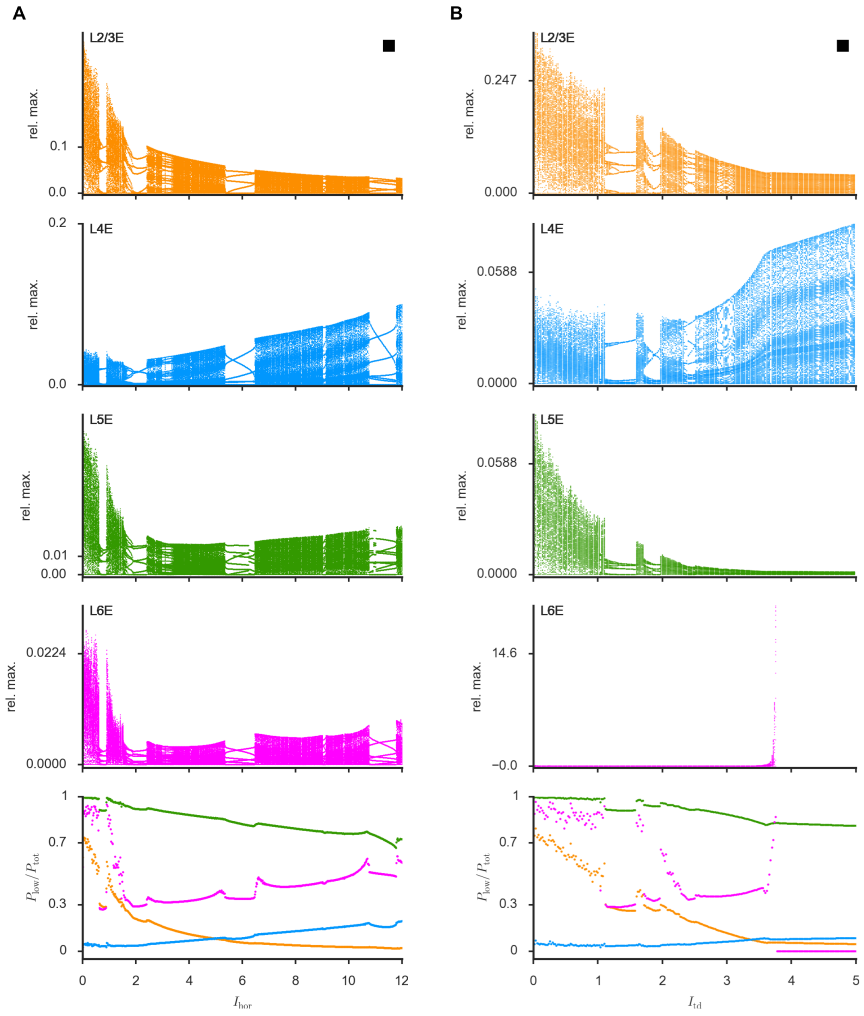


Figure 4.9: Bifurcation diagrams for contextual inputs. Like fig. 4.8B, but for varying horizontal (A) and top-down (B), instead of bottom-up, input. The last panels in each column are identical to fig. 4.8C,D and are repeated for clarity. Several regularity windows are apparent, for example around $I^{HOR} = 2$, in which the oscillations become quasi-periodic.

Moreover, the sequence of power spectra and their properties show that the dynamics in the column is tightly linked to the coupling between layers demonstrating that the oscillations in all layers should be considered holistically rather than independently.

4.3.6 *Intrinsically and network generated slow oscillations can co-exist in same circuit*

It has been suggested that slow oscillations in deep layers are generated by intrinsic oscillators in these layers [Silva et al., 1991; Flint and Connors, 1996; Roopun et al., 2006]. As our analysis, on the other hand, indicated that slow oscillations were a network effect we studied how the model column would behave in the presence of an intrinsic slow oscillator in layer 5. Such an oscillator, with a frequency incommensurate with the intrinsic frequencies of layers 2/3 and 4, could be obtained in the model by multiplying the strength of connections 5I→5E and 5I→5I by a factor of 400 compared to the original model. The resulting dynamic regime profiles for bottom-up input ($I^{LGN} = 2$) are shown in Fig. 4.10A and Fig. 4.11. Notably, regions reminiscent of those in the original model (compare Figs. 4.2A,C)—and in particular the ■-marked region—are still clearly recognizable in the traces and cartoon regime profiles (Figs. 4.10B,C) neglecting layer 5 which is now predominantly oscillating at its intrinsic, slow frequency irrespective of $K_{E,I}$. Nevertheless, due to the coupling, layer 5 was influenced by and in turn influenced other layers as can be seen in the power spectra for various values of the interlayer coupling strength multiplier Γ. For $\Gamma = 0$ (Fig. 4.10D, left panel), the spectrum was line-like featuring the three intrinsic frequencies of layer 2/3, 4 and 5 (layer 6 was not oscillating due to missing inhibitory-to-excitatory connections in the connectome), expressed only in those layers themselves. Already for weak coupling ($\Gamma = 0.1$, see Fig. 4.10D middle panel), however, the spectra for *all* layers became broader, while nonetheless each layer retained peaks at its intrinsic frequency. When the coupling was further strengthened to $\Gamma = 1$ (Fig. 4.10D, right panel) the spectra became broad, indicating chaotic dynamics. Interestingly, a peak at the difference frequency of layer 4's and layer 2's main frequencies was also visible, at least in layers 2/3, 4 and 6, indicating that the oscillations might still be generated in an approximately quasi-periodic scenario. In our specific example the two mechanisms to produce slow oscillations, quasi-periodicity and intrinsic resonance, could hence co-exist, albeit in different layers. Moreover, modifying slightly the parameters that create the intrinsic oscillator in layer 5, could change its frequency to overlap with the difference frequency. In that case it would be impossible to distinguish the different mechanisms from the power spectra.

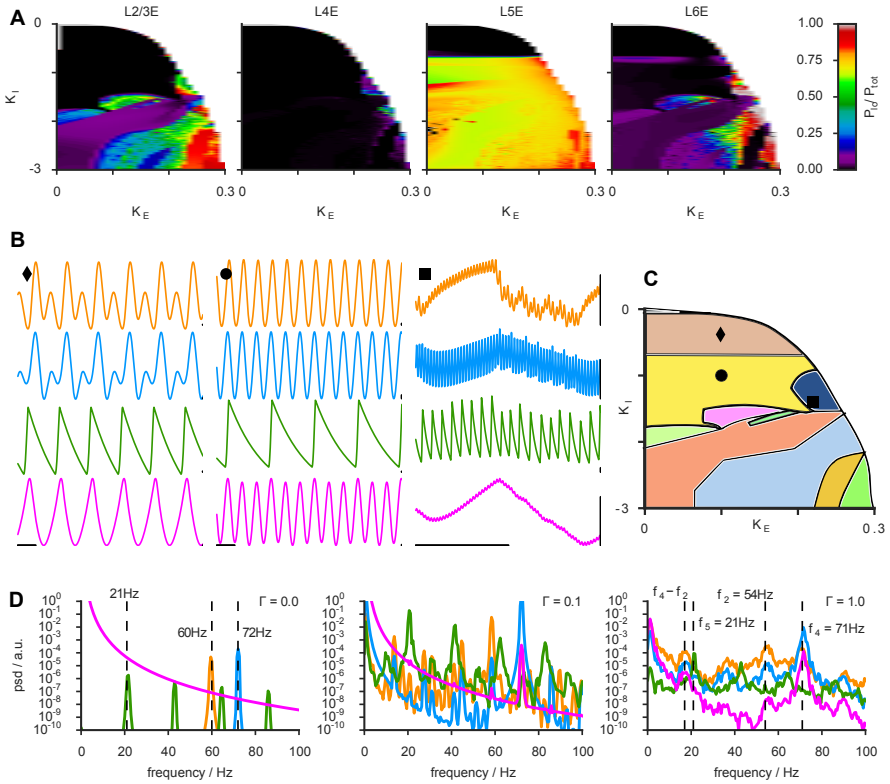


Figure 4.10: Similar behavior even in presence of an intrinsic slow oscillator in L5 A) Dynamic regime profiles for fraction of low power, **B)** example traces, and **C)** summary regime profile cartoon bore markable resemblance—apart from layer 5—to the original model when we artificially introduced a slow oscillator in layer 5. Asynchronous, all-fast, and mixed fast-slow phase in C are marked with same symbols and in same colors as in Fig. 4.2B,C (but note that working points are different here). **D)** Progression of power spectra for various values of interlayer coupling multiplier Γ also resembled those in the original model. For full coupling both the intrinsic slow layer 5 oscillation (with frequency peak at f_5) and the slow quasi-periodic oscillation (peak at $f_4 - f_2$) co-existed.

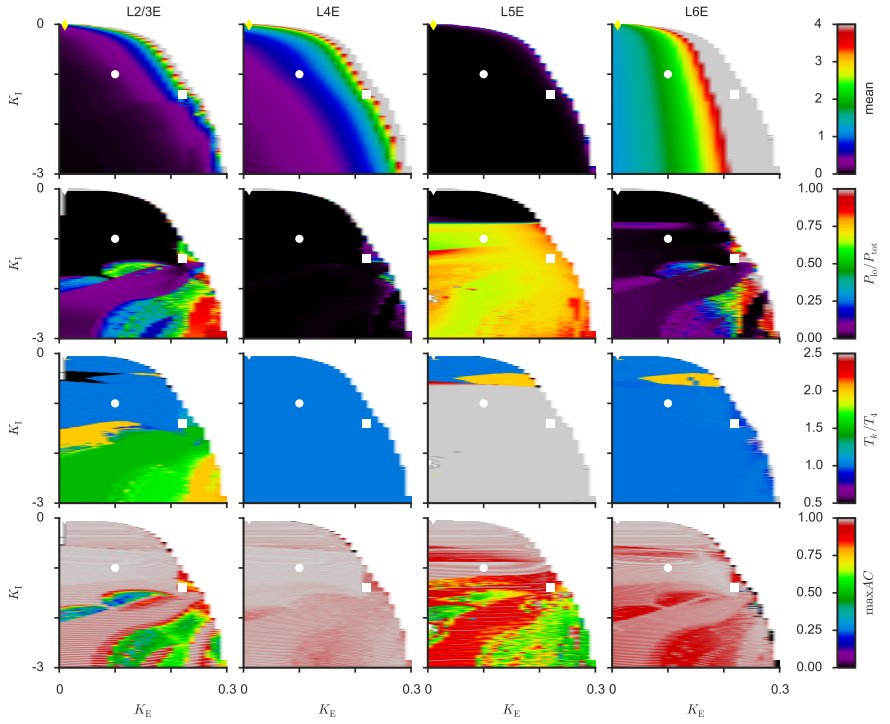


Figure 4.11: Dynamic regime profiles for modified model with intrinsic slow oscillator in layer 5. Like fig. 4.2A, but for the modified Binzegger matrix which gives rise to an intrinsic slow oscillator in layer 5. Second row is identical to Fig. 4.10A and repeated for clarity.

4.3.7 *Only few random connectomes support layer-specific frequency segregation*

The rate model equations governing the present column model are experimentally constrained by the Binzegger connectome [Binzegger et al., 2004], and that this model can give rise to a region of layer-specific fast and slow oscillations is intriguing. To see how tightly the occurrence of such a region depended on the connectome, we searched the K_E-K_I parameter space of 100000 randomized connectomes during bottom-up drive for working points in which there existed both a high and a low power dominated layer while the other layers were only required to oscillate at all (see Materials and Methods for details).

Intriguingly, in only 125 of the 100000 randomized connectomes all three conditions were satisfied for at least one $K_{E,I}$ pair that we tested (Fig. 4.12A). Moreover, in almost all (119) of these "good" connectomes, at the first pair of $K_{E,I}$ that the algorithm found to satisfy the conditions, either L2/3E or L4E was fast dominated, and either L5E or L6E was slow-dominated. In the remaining six connectomes, while the power was distributed over the layers in a different way at the algorithmically found working point, it could still be possible that

there exists another working point in which the paradigmatic case of fast- (slow-)dominated oscillations in upper (lower) layers was satisfied. As the connectomes were random with the exception of the couplings within layers, the upper layers always contained intrinsic fast resonances as in the original model, whereas the deep layers did not contain intrinsic resonances. Presumably, thus, in all "good" connectomes the intrinsic resonances must have biased the dynamics to be predominantly fast in upper layers, and the slow oscillations in slow layers must have emerged through inter-layer interactions.

The existence of at least the slow oscillations, thus, was depending on the randomized connection strengths between layers. Some excitatory connections, as well as inhibitory connections to layer 6I and in particular the connection 5I→4E, were relatively unconstrained, varying over a wide range of values in the set of "good" connectomes (Fig. 4.12B). Others, mostly inhibitory connections (with the noted exceptions), on the other hand, took on only relatively specific values among the "good" connectomes. Conspicuously, connection strengths within the artificial "good" connectomes tended to deviate to higher values compared to the original Binzegger connectome. That only around 0.1 % of connectomes are "good", indicates that a delicate interplay between the interlayer connections is necessary in order to achieve a fast/slow-phase. However, the corresponding patterns were hard to identify: PCA yielded no obvious separation of "good" from "bad" connectomes (Fig. 4.12C shows scatter plot of the first three PCA components, figure 4.13 shows a comparison of the distributions of all relevant PCA components between "good" and "bad" connectomes), neither did other clustering attempts (not shown).

We next looked at the dynamics produced by randomized connectomes. Figure 4.12D illustrates exemplarily how pronounced the differences between a randomized "good" and the original Binzegger connectomes can be. Surprisingly though, the dynamic regime profiles (figure 4.12E,F) still had features resembling the original one (Fig. 4.2A,B). Indeed, as before, firing rates remained bounded only when excitation was not too strong (right upper corners of phase diagrams) and in that case, when (the absolute value of) inhibition was weak, the dynamics was completely asynchronous, that is firing rates were constant in all layers (dark gray in cartoon diagrams). Moreover, when inhibition was increased, dynamics suddenly switched to fast oscillations in all layers (yellow phase in cartoon diagrams), and for even stronger inhibition and specific values of excitation there emerged a phase that tended to have fast and slow dominated oscillations in upper and lower layers, respectively. The remaining phases were somewhat more "blurry" compared to the Binzegger connectome and we did not attempt to match them up, as also they were not our main interest. We emphasize though that the described example and its re-

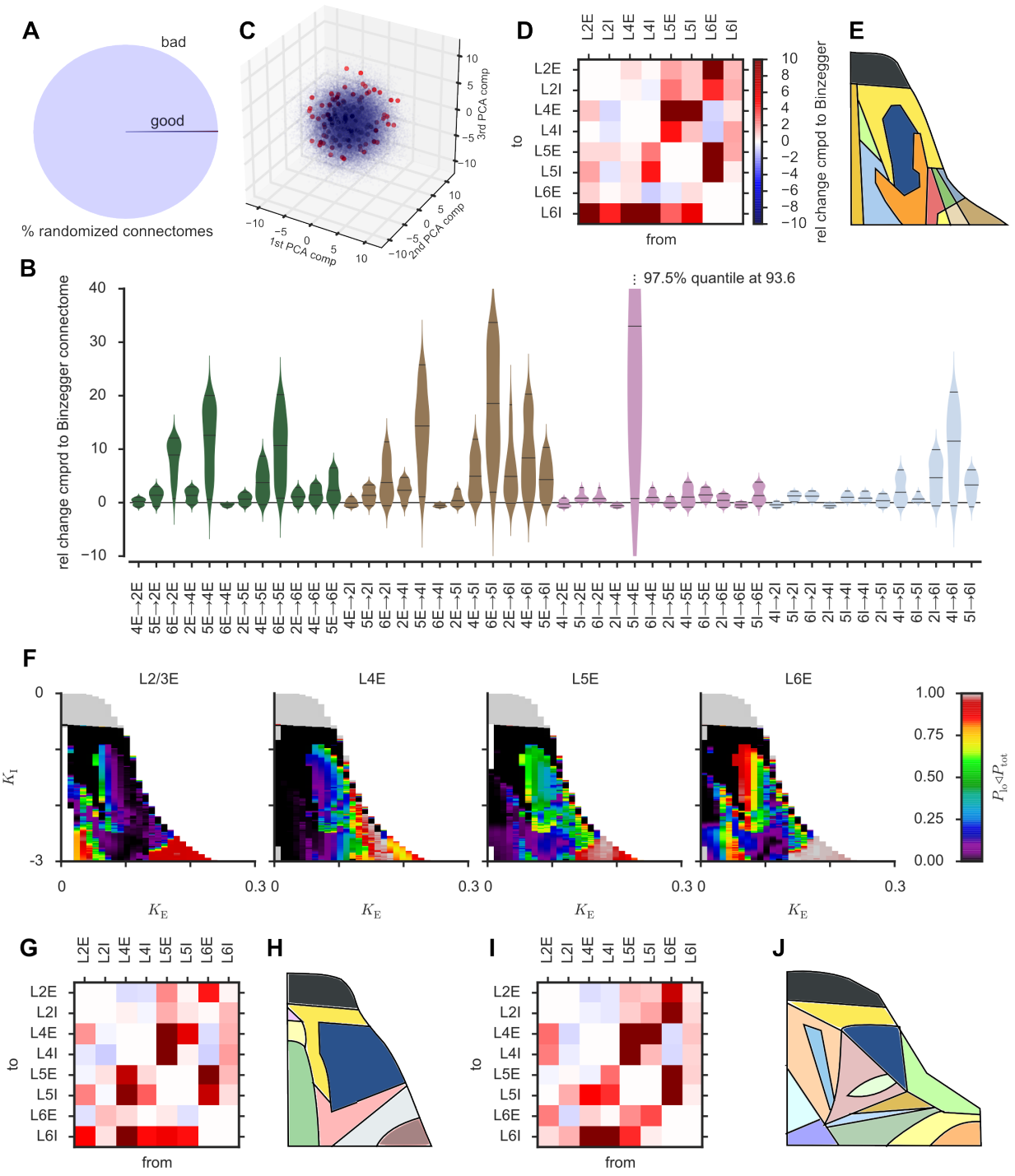


Figure 4.12: Connectome is pivotal for frequency segregation. A) In only 0.1 % of randomized connectomes we found a "good" working point such that all layers were oscillating, and at least one layer was dominated by slow, and one by fast oscillations. **B)** Distributions of connection strengths for all found good connectomes. Some components needed to take on relatively precise values, whereas others were largely unconstrained. **C)** PCA was unable to separate "good" (red) from "bad" (blue) connectomes. **D)** Relative difference between exemplary "good" and the Binzegger connectome. **E)** Corresponding summary regime profile cartoon and **F)** regime profile for fraction of low power revealed similarities (asynchronous, all-fast, and mixed slow/fast-region, marked in same colors as in Fig. 4.2C) to their original counterparts despite these large differences. **G,H)** and **I,J)** Like D,E but for two additional "good" connectomes.

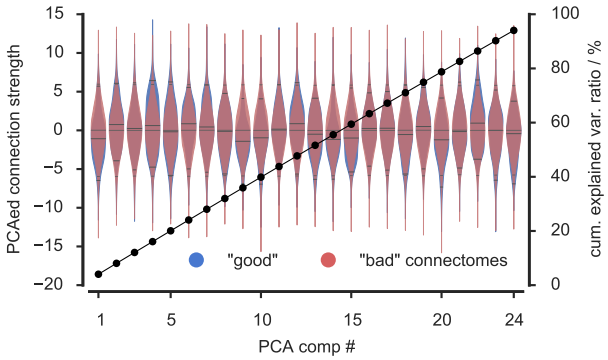


Figure 4.13: Distributions of relevant PCA components for "good" and "bad" connectomes overlap. Thickness of blue and red violins, along the left y-axis, illustrate distributions of the first 24 PCA components (x-axis). Together, these 24 PCA components capture most of the variance, as indicated by the black circles along the right y-axis.

semblances to the original case were not unique and could also be found in (at least some of the) other randomized connectomes, see figures 4.12G-J for two additional examples.

4.3.8 *Slow/fast-phase occurs robustly.*

Finally, in order to see how small changes to the connection matrix would affect our results, we studied the occurrence of layer-specific oscillations in the dynamic response profiles for slightly modified connection matrices. For computational reasons it was impossible to exhaustively explore the 100-dimensional space of disturbed matrices around Binzegger's matrix. Instead, we compared this matrix with another local circuit connectome from [Haeusler and Maass, 2007], identified the matrix elements with the biggest differences, and studied the column dynamics when one or several of these were multiplied by a modification factor α (see Materials and methods for details). The fraction of the phase diagram where the relative amount of low power, P_{lo}/P_{tot} was, respectively, greater than 50 % in deep (L5E, L6E) and smaller than 50 % in superficial (L2/3E, L4E) layers changed with the strength of the modification in all cases that we tested, but only gradually so (figure 4.14), that is, a segregation of dominant frequencies between upper and lower layers could also be found for slightly disturbed connectomes suggesting that this region occurs robustly.

4.4 DISCUSSION

We showed that layer-specific fast and slow oscillations reminiscent of experimental findings can occur in a model for s local cortical circuit if only the connections between the layers follow—according to current knowledge—a realistic pattern. This segregation of frequencies was context dependent and matching a hypothesis of the differential involvement of upper and lower layers in top-down and bottom-up processing, subserved by fast and slow frequencies, respectively [Wang, 2010; Bastos et al., 2015b]. The segregation emerged mecha-

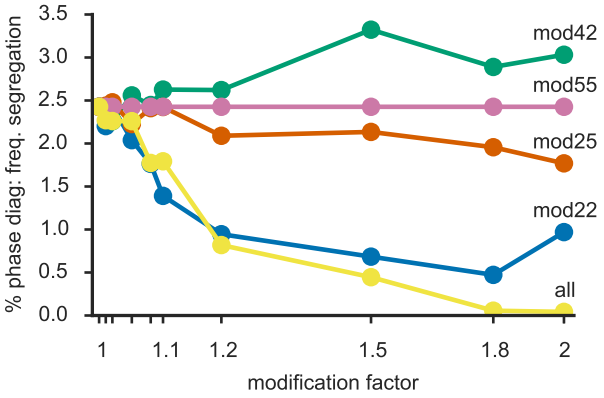


Figure 4.14: Frequency segregation is robust against small changes of the connectome The size of the slow/fast-region in the dynamic regime profiles changed only gradually when the Binzegger et al. [2004] connectome was increasingly disturbed in five different ways, indicating that the presence of this region is a robust feature of this connectome.

nistically through the interactions between layers. Although it is commonly believed that slow oscillations in deep layers are intrinsically generated [Silva et al., 1991; Flint and Connors, 1996; Roopun et al., 2006; Lee et al., 2013], this is not in contradiction with our finding, as we also show that both intrinsically and network generated oscillations may co-exist in the same circuit and that they may be hard to disentangle. Finally, we found that "good" connectomes giving rise to layer-segregated predominant frequencies were although not unique but yet exceedingly rare, and thus far from random, making their knowledge decisive for an informed local circuit model.

4.4.1 *Reasons for studying a model of a local cortical circuit*

We studied a model for a local circuit containing six layers which are vertically interconnected. It has been argued that such a circuit is "canonic", repeated over and over in the cortex [Mountcastle, 1997; Douglas and Martin, 2007; Heinzle et al., 2007]. This view has been heavily criticized, though, due to large differences of local circuits found between areas and species [Swindale, 1990; Horton and Adams, 2005; Rakic, 2008] and it has been pointed out that due to this heterogeneity circuits might be highly specialized [Barkow et al., 1992] or that although different, they might have similar function [Hooser, 2007]. Our results support this last view in that several—although highly particular—structures led to similar properties (layer-specific segregation of frequencies, in our case). This "functional homeostasis" has been documented extensively in the crab stomatogastric ganglion. In models for this system, depending on the value of a electrical and chemical conductance, various phases of oscillatory states occurred within each of which the dynamics was qualitatively similar [Gutierrez et al., 2013]. Beyond that, the model suggested that such regions of qualitatively similar behavior might exist that are unconnected in parameter space, separated from each other through regions of qual-

itatively different dynamics—much like what we observed when we randomized the connectomes.

Here, we consider a model for an isolated column which, in reality, is highly interconnected. Indeed, Binzegger and colleagues acknowledge that their analysis might have missed a considerable number of synapses in layer 4 and particularly layers 1 and 6 compared to an independent stereological study [Beaulieu and Colonnier, 1985] which might have been—at least in part—due to extra-columnar input, possibly even from a quite large number of source areas [Douglas and Martin, 2007; Markov et al., 2014b]. As a result, the emergent dynamics in such a highly interconnected network could be much different from the one in an isolated column. The other case, in which the overall dynamics is reminiscent of its constituent's dynamics, is, however, also conceivable, and the two extreme cases as well as their intermediates could maybe even be found in the same system. The fact that we demonstrated layer-specific predominant frequencies similar to what has been observed in an experimental setting where columns are not isolated [Buffalo et al., 2011; Kerkoerle et al., 2014; Bastos et al., 2015b] could indicate that the global dynamics is heavily influenced by local circuits (although within a column the opposite might be true and dynamics tends to be emergent, as we have indicated). But even in the other case, the isolated column, being a building block of the large-scale network, necessarily constrains the global dynamics and is, therefore, informative in itself.

4.4.2 *About the model*

We used a coarse-grained rate model including only one excitatory and inhibitory unit per layer, as well as a realistic connectivity between layers. As a result, we model population activity neglecting the heterogeneity and variability of single, as well as the interdependence of small groups of neurons [Ringach et al., 2002; Gur et al., 2005; Niell and Stryker, 2008; Goaillard et al., 2009; Cohen and Maunsell, 2009; Churchland et al., 2010; **?**]. On the other hand, Binzegger et al. [2004]'s connectome is based on the reconstruction of only few neurons and it is doubtful that they are representative for each neuron, given the heterogeneity of synapses [Dobrunz and Stevens, 1997; Craig and Boudin, 2001], bouton clustering [Binzegger et al., 2007] and postsynaptic areas [Anderson et al., 2011]. Including all these details—even assuming they were known—in a model, as was suggested [Markram, 2006], would result in a wealth of data in which essential details would possibly be quite hard to unveil and also requiring tremendous computing power. On the other hand, a more minimalist rate model for the average population activity seems both more adapted to the sparseness of available data and apt to explore the dependence on parameters. Moreover, computational studies us-

ing both rate and spiking network models have demonstrated that generally a good match between the two could be expected [Battaglia et al., 2007; Deco and Jirsa, 2012; Deco et al., 2013a]. Finally, even though probably as simple as possible, the biologically realistic circuit model possesses already a degree of complexity which makes it difficult to analyze thoroughly. It is therefore noteworthy that we were able to indicate the probable underlying quasi-periodic mechanism of the occurring oscillation, which could be probed for in experiments.

The most vital ingredient of the model is the Binzegger connectome [Binzegger et al., 2004] which is an estimate of the number of synapses between various neuronal types, based on anatomical reconstruction of 39 neurons from cat area 17, literature data on the frequency and laminar distribution of neuronal types and estimates of properties of those neuron types that were not among the reconstructed ones but known to exist. An alternative existing connectome was obtained from around 1000 paired electrophysiological recordings of cat and rat slices [Thomson et al., 2002], a third connectome integrates the anatomic and physiological data into a common connectivity map [Potjans and Diesmann, 2014]. We did try also these other two connectomes in our model but did not find parameters $K_{E,I}$ for which oscillations with layer-specific predominant frequencies like those presented occurred. One possible reason for the discrepancy could be that the physiological dataset is quite sparely sampled considering that the number of synapses per neuron is on the order of magnitude of 10^3 to 10^4. Another reason could be that data from several species and sources were collocated in these alternative connectomes. This could be problematic as the apparently low probability of a random connectome to give rise to layer-specific oscillations that we found, could indicate that specific necessary patterns might have been destroyed due to the data collocation. It would therefore be most insightful to analyze the model for additional local circuit connectivity maps from different brain regions and species [Kandel et al., 2013].

Slow oscillations in the model were generated through interactions between layers and this seemed to be related to the propensity of the model to undergo quasi-periodic oscillations. These slow oscillations occurred in all layers although to a varying extent and, as a matter of fact, the quasi-periodicity scenario does not explain why upper layers were predominantly fast while lower layers were predominantly slow oscillating. A possible reason could be the different dispositions of the layers to undergo fast oscillations, as remarked when we randomized the connectomes. Indeed, while the upper layers contained intrinsic fast oscillators, the lower ones did not and as the main peaks of the fast intrinsic oscillations in layers 2/3 and 4 survived the interlayer-coupling the absence of such oscillators in deep layers might make them, overall, predominantly slower.

4.4.3 *About other models*

The column being a suspected building block of cortical architecture has made it a worthwhile target for modeling. At the coarse-grained end of the spectrum "layer"-specific temporal and spatial integration of incoming excitation [Krone et al., 1986], and emerging gamma oscillations [Traub et al., 2005; Battaglia and Hansel, 2011] have been observed even without explicitly including a realistic connectome. Indeed, the model V1 column by Battaglia and Hansel [Battaglia and Hansel, 2011] consisted of only two putative layers of excitatory and inhibitory neurons and they found that the power spectra of simulated LFPs showed broad band peaks in the gamma range that were due to the system being in a chaotic state, as well as indications for quasi-periodic oscillations for reduced interlayer coupling strength These properties already hint at our findings albeit layer-specific differences and their relation to a realistic connectome could of course not be determined with a simplified network topology.

Binzegger and colleagues suggested in a model based on their connectome that a realistic working point for a system should be close to a rate instability [Binzegger et al., 2009]—a prediction that is compatible with our ■-working point that we discussed in the results section.

According to our results the column connectome is highly specific. Along this line, it has been suggested before, that a column's internal wiring is beneficial (compared to random connections) for local computations, like pattern classification of current and past inputs, and non-linear calculations on the basis of input spike trains when either the Binzegger connectome [Binzegger et al., 2004] or one based on electrophysiological data [Thomson et al., 2002] was used [Haeusler and Maass, 2007; Haeusler et al., 2009]. Moreover, in a layered model for the frontal eye field that was based on the Binzegger connectome electrophysiological as well as behavioral findings were simulated and could be replicated [Heinzle et al., 2007]. These results have been questioned, however, in that a network, wired like in [Haeusler and Maass, 2007], did not perform better on two pattern recognition tasks than a random network [Stoop et al., 2013]. While, thus, a column might be more or less suited for a particular local input-output-computation, one of its virtues might be to integrate and multiplex incoming signals with different frequencies as suggested by our and other's [Silva et al., 1991; Flint and Connors, 1996; Roopun, 2008a; Sun and Dan, 2009; Oke et al., 2010; Ainsworth et al., 2011; Buffalo et al., 2011; Xing et al., 2012b; Roberts et al., 2013; Kerkoerle et al., 2014] finding of layer-specific frequency prevalences.

Yet another spike-based model analyzed spontaneously occurring population-bursts in a model column [Kremkow et al., 2007].

For a large-scale simulation of a cortical column Potjans and Diesmann [2014] first integrated data from several sources, both anatomical and physiological (among them [Thomson et al., 2002; Binzegger et al., 2004]) into one common connectivity map and showed that a column of integrate and fire neurons connected according to this map reproduced realistic, generally very low, spontaneous firing rates for all layers which increase from superficial to deeper layers with the exception of L6E. In comparison, firing rates in our model also tended to follow this pattern (Fig. 4.2A, first row for bottom-up activity; similar for spontaneous activity, not shown), except that firing rates in L6E were much higher. A potential reason could be that the Binzegger connectome—possibly erroneously—does not include feedback connections from L6I to L6E.

Some other models were specifically concerned with oscillations. In the model column of Du et al. [2012] the power spectra had a $1/f$ (power-law-like) decay, which we also see in our model for large frequencies. There were, however, no pronounced differences between layers except in the absolute amount of power. A possible reason for this discrepancy could be a different working point. Neymotin and colleagues found an emergent broad low power peak in the simulated spontaneous MUA for excitatory neurons that remained at the same frequency in response to stimulation of L2E, whereas the inhibitory neurons' **PSD!** (PSD!) peak was rather flat and shifted to higher frequencies when a stimulus was presented to L2E [Neymotin et al., 2011]. Our model, in contrast, showed rather similar spectra for excitatory and inhibitory populations, but we did find that dominant peak frequencies do not necessarily change when stimulations or other manipulations were applied to the circuit. An important confounding factor was, though, that the MUA in Neymotin's model was calculated from neurons in all layers. They did not, unfortunately, perform a layer-specific analysis of their simulations.

Beyond that, the behavior of more than one column with [Thomas et al., 1991; Patton et al., 1992; Hill and Tononi, 2005; Markram, 2006; Izhikevich and Edelman, 2008; Rasch et al., 2011; Wagatsuma et al., 2011, 2013; Potjans and Diesmann, 2014] realistic connectomes has been explored and it would be interesting to see and compare the emergent dynamics when we couple several of our model columns.

4.4.4 *Relation of the model to interareal coordination*

Slow oscillations in deep layers are commonly attributed to specific oscillators within these layers [Silva et al., 1991; Flint and Connors, 1996; Roopun et al., 2006]. Nevertheless network-dependent oscillations and their principal compatibility with intrinsic oscillators have been observed both experimentally [Roopun, 2008a], and now in our model and could subserve the integration of oscillations at various

frequencies. Indeed, as separate groups of neurons tend to mediate feedforward and feedback processing [Markov et al., 2014b] one might speculate that one intrinsically oscillating group listens to oscillatory input at its resonance frequency [Lee et al., 2013; Kopell et al., 2014], and then activates the remaining circuit, up to the output neurons which might be either intrinsically oscillating at the same or at a different frequency or just network-participators. In any case, the network generated multi-frequency oscillations which presumably comprise all neurons of the circuit might underlie the coordination of all inputs and outputs at their various frequencies through their inherent cross-frequency-coupling.

4.4.5 *Conclusion*

"Connectomics" stipulates that the structural connections in the brain constrain its function [Sporns et al., 2005]. In line with this idea, our analysis indicated that a given connectome might support several functional states and that a random connectome, in general, will lead to qualitatively different effects. Moreover, although not just one connectome was able to produce layer-specific multi-frequency oscillations in our model, admissible connectomes were so rare, that they could not be obtained by chance, arguing for the importance of collecting fined-grained collectivity data [Kandel et al., 2013].

4.5 MATERIALS AND METHODS

4.5.1 *Model*

We simulate the rate response $r_{k\alpha}$ of a population $\alpha \in \{E(\text{xcitatory}), I(\text{nhibitory})\}$ in layer $k \in \{1, 2/3, 4, 5, 6\}$. Each layer receives recurrent inputs with a delay $D = 0.1$ time units from within the column. The strength of these inputs, denoting the total charge induced into the postsynaptic population $l\beta$ ($l \in \{1, 2/3, 4, 5, 6\}$, $\beta \in \{E, I\}$) over all existing connections $l\beta \to k\alpha$ is written as $\gamma^{l\beta}_{k\alpha} K^{l\beta}_{k\alpha}$ where $\gamma^{l\beta}_{k\alpha}$ is the fraction of all synapses from population β to population α that are formed between layers l and k; values for $\gamma^{l\beta}_{k\alpha}$ were taken from Fig. 12 in [Binzegger et al., 2004]. $K^{l\beta}_{k\alpha}$ then gives the total number of synapses between populations β and α times the average charge induced in the postsynaptic population due to a single spike in the presynaptic population. We used $K^{l\beta}_{k\alpha}$ as parameters and explored their influence on the behavior of the model. To make this feasible, we made the simplifying assumption that $K^{l\beta}_{k\alpha} \approx K_\beta$, thus reducing the effective number of parameters to two, K_E and K_I, which then describe phenomenologically the efficacy of excitatory and inhibitory connections, respectively.

Apart from recurrent input the column is driven by four kinds of constant external currents. All layers receive the same background

current I^{bg} modeling diffuse neuronal noise. Bottom-up stimuli influence the model column via currents I^{LGN}, $I^{LGN}/3$ and $I^{LGN}/6$ sent to layers L4E&I, L6E and L6I, respectively, in accordance with [Binzegger et al., 2004]; the strength of I^{LGN} can be interpreted as the contrast of the stimulus. Contextual influences on the column are divided into horizontal, I^{HOR}, and top-down, I^{TD}, currents, targeting layer L2/3E&I and L5E&I, respectively. A cartoon illustrating the complete circuit is shown in 4.1.

The total input into each layer, I^{tot}, from both intrinsic and extrinsic sources, is assumed to activate the layer after transients have passed to a level given by the $f-I$-curve $F(I^{tot}) = \max(0\,A, I^{tot})\frac{1}{A\cdot\text{time unit}}$. In summary, our model column was governed by the equations

$$\tau_{k\alpha}\dot{r}_{k\alpha}(t) = -r_{k\alpha}(t) + F\left(I^{bg}_{k\alpha} + I^{LGN}_{k\alpha} + I^{hor}_{k\alpha} + I^{td}_{k\alpha} + \sum_{l\in\text{Layers}}\sum_{\beta=E,I}\gamma^{l\beta}_{k\alpha}K_\beta r_{l\beta}(t-D)\right) \tag{4.1}$$

For simplicity, the relaxation time constants $\tau_{k\alpha}$ were taken to be 1 time unit for all k and α. The rate equation has not specifically been obtained through a mean-field reduction, but it does resemble a simplified Wilson-Cowan equation [Wilson and Cowan, 1972]. It was integrated with a custom written C code using a Runge-Kutta algorithm with time step 0.0001 time units. To overcome initial transients the first 200 time units (i. e. 2 000 000 time steps) were discarded and data analysis was performed on the subsequent 2^{20} time steps.

4.5.2 *Data analysis*

All data analysis was performed in Python. Power spectra were calculated using Welch's method with a window size of 2^{19} time steps, a window overlap of 2^{18} time steps and with the time series' means subtracted. To make frequencies concrete, we set the transmission delay D to 10/3 msec, i.e. the time unit to 1/30 sec, i.e. the sampling frequency to $3\cdot10^5$ Hz. Therewith, we defined low power P_{lo} as the summed power spectral density between 0 Hz and 30 Hz, and the total power P_{tot} as the sum of the power spectrum over all frequencies.

Cross-correlations were estimated by $\mathcal{F}^{-1}[\mathcal{F}r_{k\alpha}\cdot\overline{\mathcal{F}r_{l\beta}}]$ where $\mathcal{F}$ denotes the fast Fourier transform, $\mathcal{F}^{-1}$ the inverse fast Fourier transform, and $\overline{\mathcal{F}r_{l\beta}}$ the complex conjugate of $\mathcal{F}r_{l\beta}$.

Cartoon regime profiles were obtained by semi-transparently overlying all layers' profiles for the rate, fraction of low power, first delay and highest value of the autocorrelation and then tracing "obvious" boundaries. While some judgment calls had to be made this procedure is in no way critical as it only serves to highlight qualitative differences and similarities of the dynamics for different values of $K_{E,I}$.

To determine the extent of the slow/fast-region we operationally defined this region as those points in the dynamic response profile that are "neighbors" of the point $(K_E, K_I) = (0.24, -1.4)$, neighbors of neighbors, ... For that matter, a point on the grid of sampled $K_{E,I}$ values, reachable with one step either orthogonally or diagonally from a given source point is called a "neighbor" if it fulfills the condition that the fraction of low power was smaller than 0.999999 in all layers and greater or equal than 0.6 in at least one layer.

4.5.3 *Randomized connectomes*

We randomized the Binzegger connectome to determine the relevance of connections for local circuit processing. To that end each connection—with the exception of those to and from layer 1, as well as those within a layer—was assigned a random value from a uniform distribution that was specific to the type of connection, $E \to E$, $E \to I$, $I \to E$ or $I \to I$. The lower and upper bounds of these distributions were the minimum and maximum values of all matrix elements in the original Binzegger connection matrix belonging to the respective connection type.

For each randomized connectome we searched a grid of working points, $K_E = 1/30 \ldots 1$, $K_I = -10 \ldots 0$ with a step size of $\Delta K_{E,I} = 1/30$ which was traversed in order of increasing distance to the reference working point $(K_E, K_I) = (0.21, -1.4)$ during bottom-up drive ($I^{bg} = 1$, $I^{LGN} = 2$, $I^{HOR} = I^{TD} = 0$). To speed up the search, we first run a coarse-grained simulation for each working point, with a temporal step size of 0.001 and analyzed 2^{16} time steps (after the initial transients). If total power P_{tot} in all layers was above 10^{-5} and there existed at least one layer with more than 25 %, and at least one other layer with less than 75 % of relative amount of low power P_{lo}/P_{tot}, we run another simulation for the same working point with the same, more fine-grained parameters that we used in the rest of the analyses, that is with a time step of 0.0001 and with 2^{20} time steps (after the initial transients). If we found that $P_{tot} > 10^{-5}$ in all layers, and that there existed at least one layer with more, and at least one other layer with less than 75 % of P_{lo}/P_{tot} we stopped the search for the given connectome and termed it "good".

4.5.4 *Robustness of results against small changes of connectome*

We compared the connectome from [Binzegger et al., 2004, Fig. 12] to the one from [Haeusler and Maass, 2007, Fig. 1] by normalizing each of them separately with its respective maximum value and looking for the six biggest absolute values in the difference which were divided into four groups. The matrix elements of each group were then modified in order to determine their effects on the dynamics

of the model column. The groups that we manipulated were: first, we changed the couplings $C_{L2I\to L2E} \to C_{L2I\to L2E}/\alpha$ and $C_{L2I\to L2I} \to C_{L2I\to L2I}/\alpha$ (we termed this modification "mod22"); second, $C_{L4E\to L2E} \to \alpha C_{L4E\to L2E}$ and $C_{L4I\to L2E} \to \alpha C_{L4I\to L2E}$ ("mod42"); third $C_{L2I\to L5E} \to \alpha C_{L2I\to L5E}$ ("mod25"); fourth, $C_{L5I\to L5I} \to \alpha C_{L5I\to L5I}$ ("mod55"); and fifth, all of the above four modifications at the same time ("all"); we run simulations for all of these five modifications for values $\alpha \in \{1, 1.01, 1.025, 1.05, 1.125, 1.5, 2\}$ on a grid $K_E = 0 \ldots 0.3$ and $K_I = -3 \ldots 0$ with step size $\Delta K_{E,I} = 0.01$ and assessed how the fraction of working points (K_E, K_I) (excluding those with diverging rates $r_{k\alpha}$) where the relative amount of low power P_{lo}/P_{tot} was above and below 50 % in superficial and deep layers, respectively, changed with α.

5

DYNAMIC COORDINATION IN A MODEL OF TWO COUPLED COLUMNS

5.1 INTRODUCTION

Information processing in the brain is believed to occur at multiple levels, from the molecular, to the sub-cellular, cellular, multi-cellular to the brain-wide. Nevertheless, in all cases the computation requires the coordinated action of several of the system's components at the given scale and disentangling this interplay is paramount for understanding the system's functioning.

Experiments have established that the efficacy of spikes depends on the timing of these spikes relative to ongoing subthreshold oscillations in the target neurons [Volgushev et al., 1998], and the CTC hypothesis, one existing theory for how coordination could work, posits that this fact might be harnessed to rhythmically gate communication [Fries, 2005] between neurons. Computational models [Battaglia et al., 2012; Kirst, 2012] have further established that the relative phase of oscillations in connected neuronal populations determines the "effective connectivity" between these populations, that is, the uncertainty in the prediction of the activity of one population, given its past activity, is reduced when also the history of the other population is considered, and this reduction is stronger for the "leader", than for the "laggard" population, leading to effectively directed couplings. Moreover, the information transmitted by the identity of those neurons that are spiking, and those that are not spiking, in a modular network of neurons firing irregularly with a population gamma rhythm follows the same asymmetry pattern as the directionality established by the effective connectivity.

That "leader" and "laggard" phase relationships establish an effective directionality in the coupling was studied for small network motifs of rate units [Battaglia et al., 2012] and pulse-coupled oscillators of arbitrary topology [Kirst, 2012]. Here, we begin to study the phase-locking between two populations which internally possess a realistic topology of a cortical column [Binzegger et al., 2004]. These columns are then coupled, assuming they are situated at different stages of the cortical hierarchy, in order to determine if feedforward- (lower column leading) or feedback- (upper column leading) phase-relationships can occur, possibly in multiple frequency bands, just by virtue of the local-circuit-connectome equipped with simple local dynamics.

We report preliminary findings, suggesting that when the two column's don't receive specific input there exists a parameter regime for which the columns operate largely in a feedback-mode (upper column leading), which switches to a feedforward-mode (lower column leading) when bottom-up input is provided. When in addition to the bottom-up input the upper column receives top-down input (from another area extrinsic to the two-column-system), then, for certain parameters, the system is "feedforward" in a fast gamma-frequency band, but top-down for slow (beta-like) frequencies. Moreover, we found indications, that this pattern of phases might depend on the mechanism by which these oscillations are created: in our model slow oscillations emerge through the interplay of layers, whereas in an alternative model with a dedicated slow oscillator in layer 5, the pattern of phase-relationship was largely different. Yet, specific randomized connectomes might also support the same pattern.

These preliminary findings would be the first—to the best of our knowledge—to reproduce experimental evidence suggesting that feedforward and feedback processes are subserved by oscillations in fast (gamma) and slow (beta) frequencies, respectively [Kerkoerle et al., 2014; Bastos et al., 2015b].

5.2 RESULTS

5.2.1 *Model*

We coupled two model columns in order to see possible ways in which they can coordinate their activity. Each column was modeled precisely as described in the previous chapter with an excitatory and inhibitory population per layer and the internal wiring determined by the Binzegger et al. [2004] connectome. The columns are envisaged to be situated at different stages in the cortical hierarchy, and, accordingly, we adopted a simple coupling scheme in which layer 2E of the lower column projected to the upper column's layer 4E and 4I, and activity from the upper column's layer 5E was fed back to the lower column's layer 2E, 2I as well as 5E and 5I (see figure 5.1). As short-range connections, long-range connections in the model were delayed and we assumed the long-range delay to be ten times as long as the short-range delay. In all simulations every layer was stimulated with a constant current mimicking unspecified background "noise" (*spontaneous* condition), and, depending on the situation of interest, further constant currents were injected to the lower column's L4 and L6 (*bottom-up* stimulation: I_{LGN}) and L2 (*horizontal* stimulation: I_{hor}). A detailed description of the model is given in Materials and Methods.

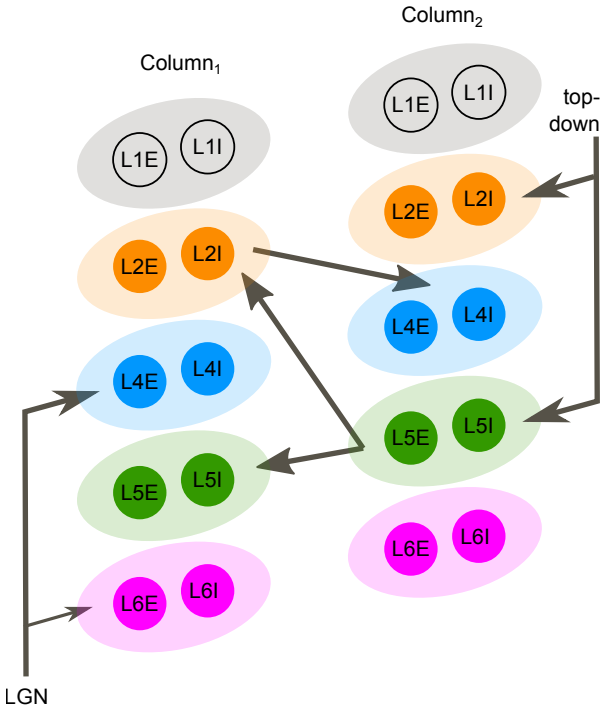


Figure 5.1: Illustration of model. We model a cortical column by an excitatory and inhibitory rate unit per layer that are coupled internally according to experimentally determined data [Binzegger et al., 2004]. Two of such columns, representing different stages of the cortical hierarchy, are coupled by feedforward projections from layer 2E to layers 4E,I of the upper column, and by feedback projections from layer 5E to layers 2E,I as well as 5E,I in the lower-column. Each layer is stimulated with a constant current representing unspecific connections, as well as, depending on the studied situation, also by bottom-up drive to layers 4 and 6 of the lower column, and horizontal drive to layer 2 in the lower column.

5.2.2 *Interdependence of columns grows with long-range coupling*

Dynamic response profiles for an uncoupled columns were presented in the previous chapter in figures 4.2, 4.3 and 4.4. When two such columns at different stages of the cortical hierarchy are coupled weakly ($K_{lr} = 0.01$) mean activity, fraction of low power and autocorrelation structure stay qualitatively similar over the $K_E - K_I$-parameter space (cf. figure 5.2A). On the other hand, when coupling strength is increased to $K_{lr} = 1$, the more "chaotic" region of the parameter-space—characterized by max AC $\lessapprox 0.95$ and the presence of slow oscillations around the ■-working point introduced in the previous chapter—expands. Nevertheless, coarse-grained features of this phase, like the approximate 3:2 coupling in the layers 2E (third row), or patterns in the fraction of low power (second row) are recognizable even then (cf. figure 5.2B). A more systematic study of the effects of long-range coupling, as well as the effect of the long-range delay has yet to be undertaken.

"weak" coupling: $K_{lr} = 0.01$, *fig. 5.2A*

"strong" coupling: $K_{lr} = 1$, *fig. 5.2B*

5.2.3 *Collective multi-frequency oscillations with layer-dependent phase-shifts*

As phase differences of oscillations in various populations might be indicative for the efficacy of information transfer between the populations [Battaglia et al., 2012] we studied such phase differences in our model. Fig. 5.4 illustrated the definition and calculation of

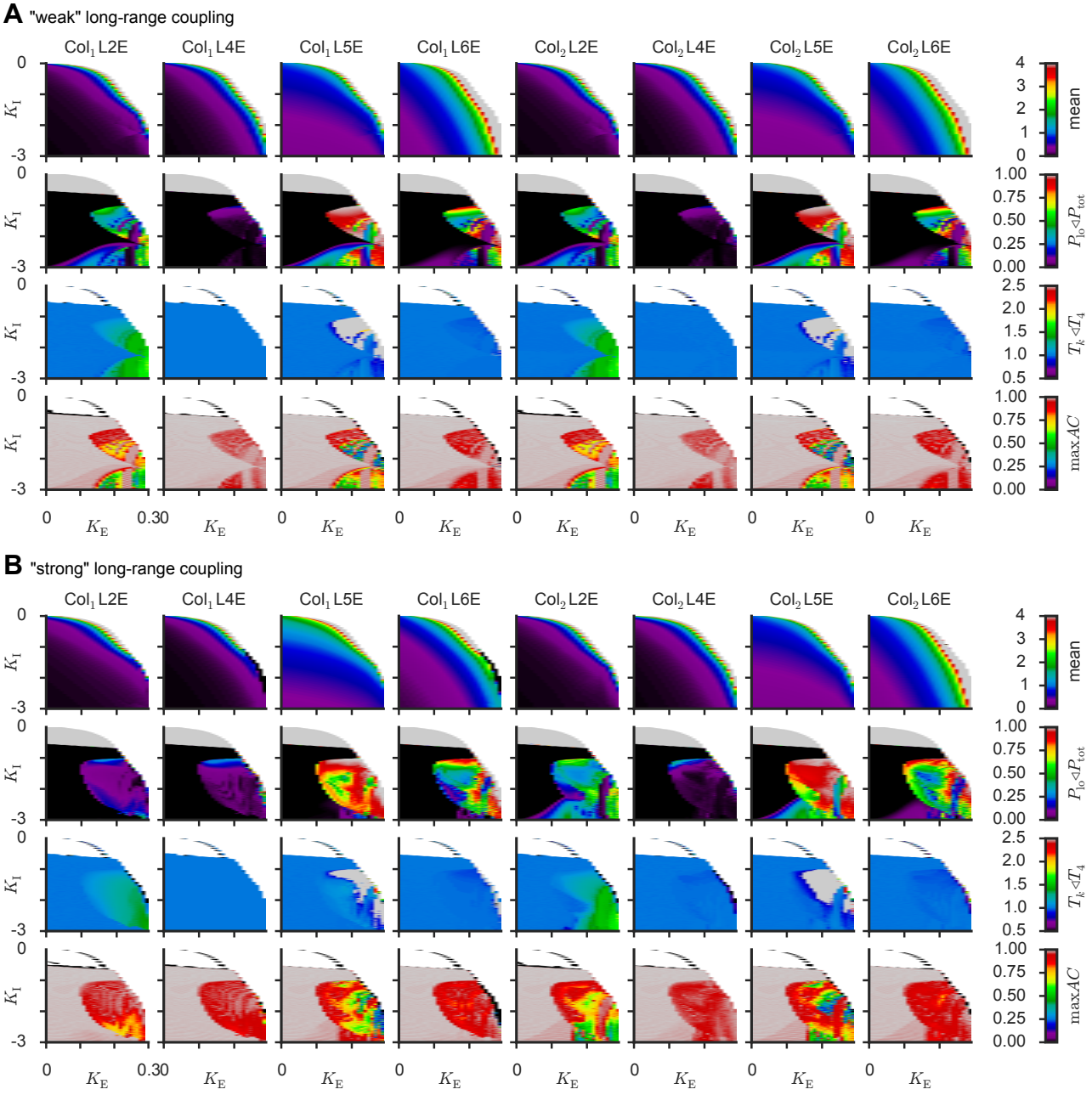


Figure 5.2: Dynamic response profiles are (are not) qualitatively affected by strong (weak) long-range coupling. Response profiles are organized as in the previous chapter, with the left (right) four columns depicting the lower (upper) column, respectively, and the rows showing, respectively, mean activity, fraction of low power, delay of first peak in autocorrelation in unit's of that of the lower column's layer 4E, and value of highest peak in autocorrelation for non-zero delays. Here, the column receives only homogeneous background drive ($I^{BG} = 1$). **A)** For weak coupling ($K_{lr} = 0.01$) response profiles essentially indistinguishable from non-coupled case ($K_{lr} = 0$), cf. figure 5.3. **B)** For stronger coupling ($K_{lr} = 1$) response profiles change, but coarse-grain structure stays comparable.

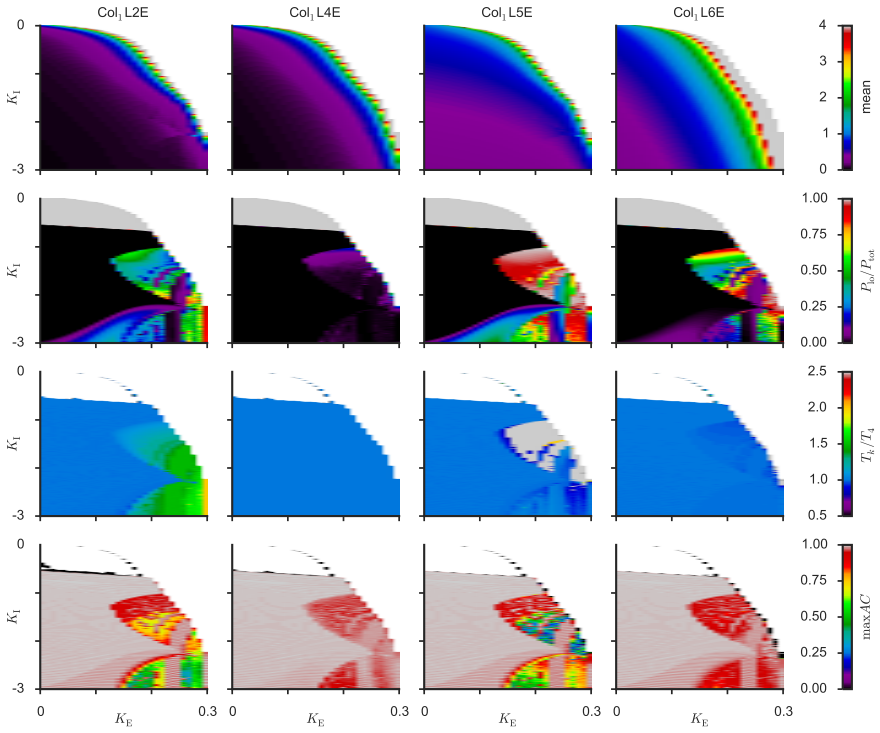


Figure 5.3: Dynamic response profiles for a single isolated column without specific drive. When two columns are coupled weakly (compare figure 5.2A) the response profiles are not much affected.

these phases for a given oscillation with predominant frequency. For a given value of local coupling efficacies ($K_E = 0.21$, $K_I = -1.4$, corresponding to the working point used in the previous chapter) we calculated the position of the global maximum in the power-spectrum of the L2E time series in the 8-30 Hz frequency range ("L2E-β peak"), in the 30-90 Hz frequency band ("L2E-γ peak"), and of the L4E time series in the (30-90 Hz) frequency band ("L4E-γ peak"). For each of these peak frequencies we applied a narrow band-pass filter to the complete time series for all populations and determined the time stamps of the local maxima. Results for the L2E-β- and L4E-γ-band are depicted in figure 5.5 (left column).

baseline condition: $I^{BG} = 1$, *fig. 5.5A*

If both columns receive only an unspecific homogeneous background drive ($I^{BG} = 1$) the time stamps of the maxima of the lower column (colored dots) relative to the upper column (grey dots) drift gradually in both frequency bands from slightly leading to slightly lagging, and that trend continues beyond the time-excerpt shown in the figure. Indeed, when the position of the upper column's maxima relative to those in the lower column are quantified (5.5A right column), then—although there are preferential values—all possible "phase" relationships (see Materials and Methods for the exact definition of the "phase" used here) do occur, from 0 (maxima in upper and lower column occur simultaneously, "in-phase"-mode), over slightly positive

values (upper column reaches maximum just before lower column, "feedback"-mode), 0.5 (upper column reaches maximum just in between the maxima of the lower-column, "anti-phase"-mode), to values just below 1 (lower column reaches maximum just before upper column, "feedforward"-mode). The two columns thus explore both feedforward- and feedback-modes over the course of time, when they are not driven by a "sensory" input.

bottom-up condition: $I^{BG} = 1$, $I^{LGN} = 2$, *fig. 5.5B*

When, however, a bottom-up stimulation is sent to the lower column's layer 4 and 6 ($I_{bg} = 1$, $I^{LGN} = 2$), the relative timing of the maxima in the two columns stabilizes, with all layers of the lower column slightly leading the upper column in both frequency bands, consistent with a feedforward processing mode (figure 5.5B).

response-period condition: $I^{BG} = 1$, $I^{LGN} = I^{Col_2L2} = I^{Col_2L5} = 2$, *fig. 5.5C*

An additional top-down stimulation of the upper column on top of the bottom-up stimulation ($I^{LGN} = I^{Col_2L2} = I^{Col_2L5} = 2$)—mimicking the response period of an attention experiment (cf. figure 2.1)—retains the consistent phase relationships over time, but makes them frequency- and layer-dependent (figure 5.5C): in layer 4, the maxima in the lower column still occur slightly before those in the upper column ("feedforward") for both frequency bands; likewise layer 2 upholds its feedforward-mode for fast frequencies. However, in the slow-frequency band, layers 2 and 5 now switch to a feedback mode. Also note that the upper column's layer 6 appears to stay mostly silent (which also happened for a single column under simultaneous bottom-up and top-down drive, cf. Fig. 4.4), which is puzzling and needs to be addressed by future work.

alternative model with slow oscillator, figure 5.5D

Next, we compare our model, in which slow oscillations in a single isolated column are generated due to the interactions between layers, with an alternative one possessing an internal slow oscillator in layer 5 (figure 5.5D). When a bottom-up input is applied (I^{LGN}) the relative positions of the maxima in the two columns are relatively inconsistent in the slow frequency band, but for fast frequencies the upper column is leading, corresponding to a feedback mode during bottom-up stimulation which is in contrast to our original model (figure 5.5B). Importantly though, it is possible that a behavior consistent with the original model can be achieved by modifying parameters appropriately. We have not yet studied this.

Finally note, that—apart maybe from the alternative model in figure 5.5D)—the order in which layers achieve the maxima, and the delays between the maxima in a given column, appeared relatively fixed. This could indicate that oscillations in a column should be considered holistic such that phase differences between the various layers in the columns are not varying in time but stable.

So far, we have studied phase-locking between columns for only a single value of the local coupling efficacies $K_{E,I}$, however, the results obtained therewith largely generalize to a set of other values. To illustrate this, we collapsed—for a range of $K_{E,I}$-values—the histograms

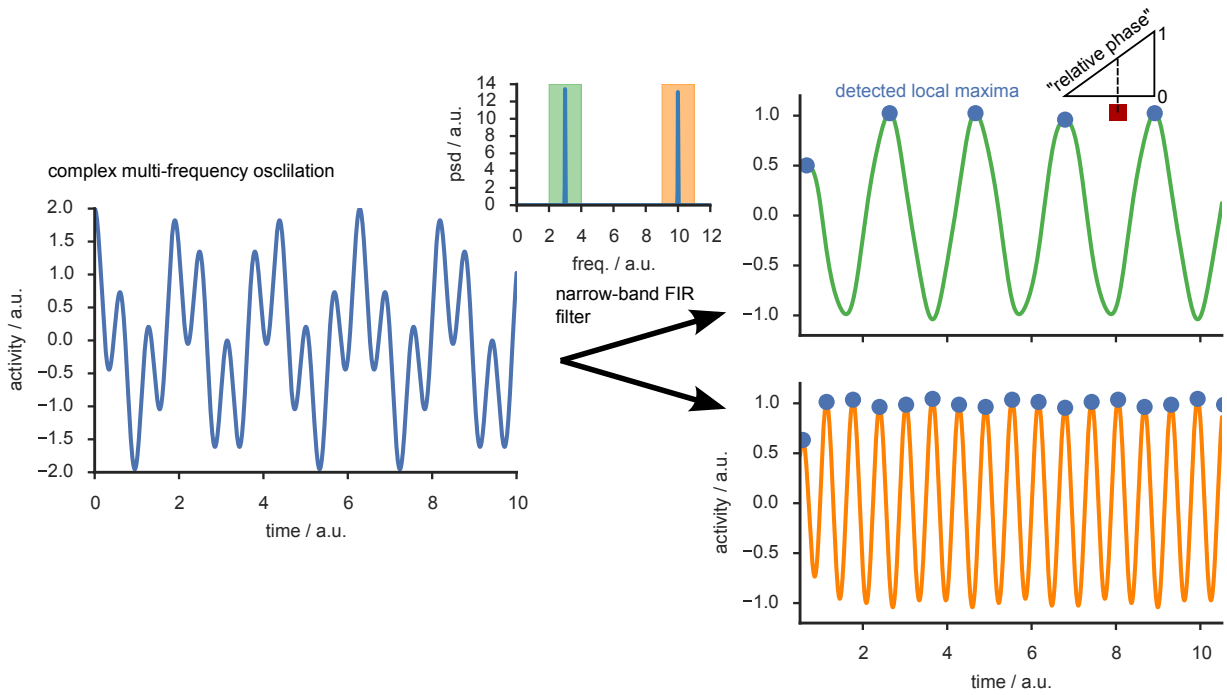


Figure 5.4: Cartoon illustrating the calculation of phases. After subsampling the original time series (for performance reasons) we apply a finite impulse response filter with a pass-band of ± 2 Hz around a predominant frequency of the power spectral density, detect the local maxima in the filtered time series and interpolate linearly between the maxima to assign a "relative phase" to a target event, like a local maximum detected in the same way in another time series.

on the right of figure 5.5 in a single number, by calculating a complex phase-locking index $\Phi^{(l)} = \langle \exp(2\pi i \phi_i^{(l)}) \rangle_i$, where the $\phi_i^{(l)}$ denote the histogram for layer l with bins numbered by i, and $\langle \cdot \rangle_i$ denotes averaging over bins.

Figure 5.5 ***(facing page)*****: Depending on context, simultaneous feedforward and feedback modes in different frequency bands are possible. Left)** time stamps of local maxima (color: lower column, grey: upper column) in filtered time series (beta: upper panel, fast gamma: lower panel). **Right)** Distributions of phases of maxima in upper column relative to phases in lower column (cf. fig. 5.4 and Materials and Methods). **A)** In the baseline condition all phases occur. **B)** Under bottom-up stimulation the lower column is leading the upper column, consistent with a "feedfoward" configuration. **C)** Under both bottom-up and top-down stimulation ("response-period" condition) the lower column is leading in the gamma-band (feedforward), but lagging (with the exception of layer 4E) in the beta-band (feedback). **D)** When a dedicated slow oscillator is introduced in the lower columns, the phase-relationships in the beta-band become inconsistent over time, and switch to a feedback configuration in the gamma-band.

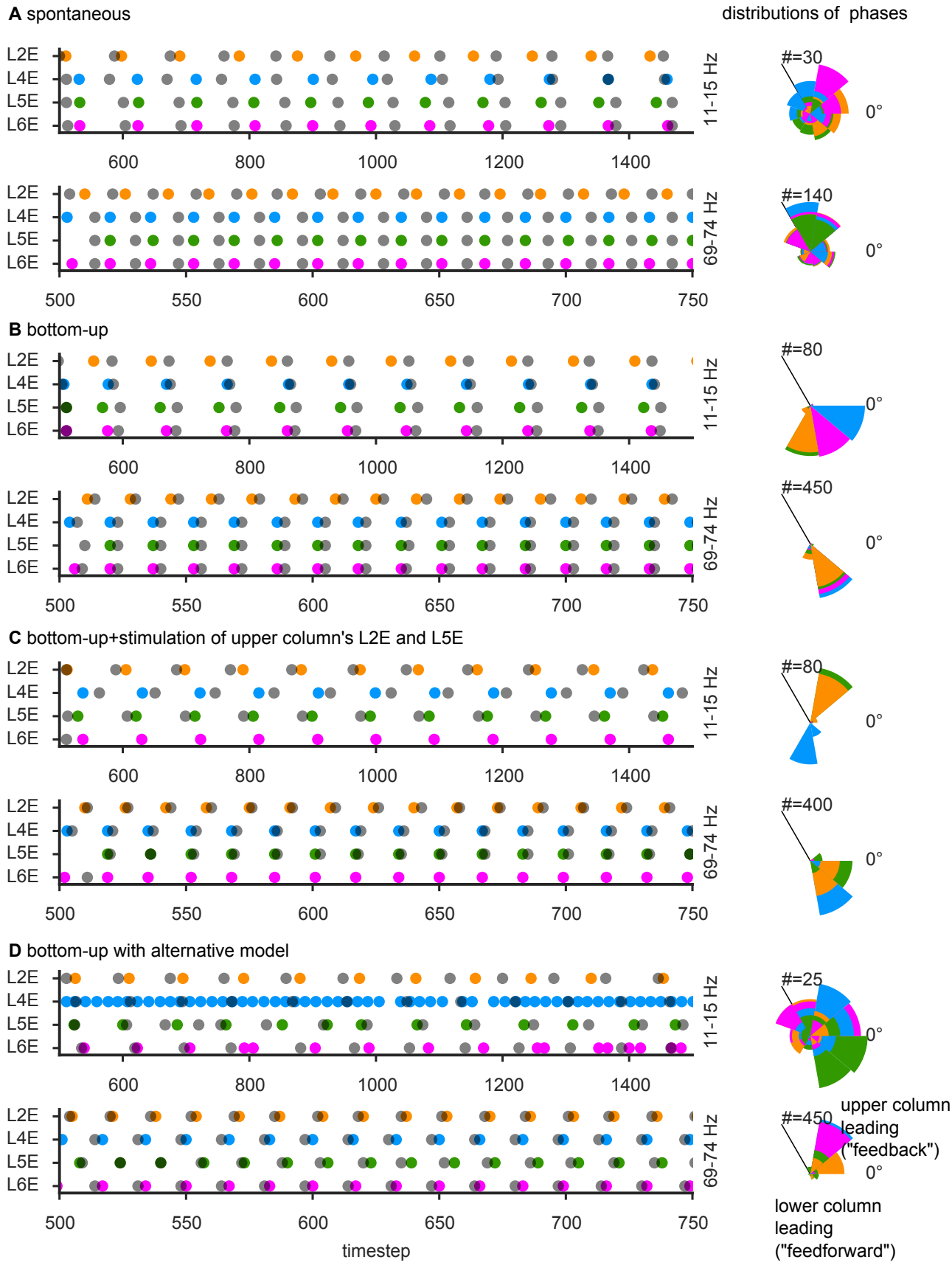

A spontaneous
distributions of phases
L2E
L4E
L5E
L6E
11-15 Hz
69-74 Hz
600
800
1000
1200
1400
500
550
600
650
700
750
#=30
#=140
0°
B bottom-up
#=80
#=450
C bottom-up+stimulation of upper column's L2E and L5E
#=80
#=400
D bottom-up with alternative model
#=25
#=450
upper column leading ("feedback")
lower column leading ("feedforward")
timestep

Figure 5.6A shows dynamic response profiles for Φ in the baseline condition ($I^{bg} = 1$). Just below the asynchronous region (top red regions for $K_i \gtrapprox -0.5$), where oscillations set in, the system switches in a "feedback mode" which gradually changes towards a "feedforward" mode for decreasing K_I as long as excitation is not too strong. In the clearly distinguished working-point phase for stronger excitation, upper layers are more feedforward than feedback for the L2E-β-band, feedback for all layers in the L4E-γ-band, and show a inhibition dependent phase relationship in the L2E-γ-band. Also note, that often no consistent phase-relationship is reached in the former two (bleached out colors).

baseline condition: $I^{bg} = 1$, *fig. 5.6A*

When the lower column is additionally stimulated with bottom-up input ($I^{bg} = 1$, $I^{LGN} = 2$, see figure 5.6B), the parameter region just below the asynchronous phase gives rise, again, to a feedback-mode, but this time only for very limited range of inhibition-efficacies, beyond which the oscillations are in-phase. The working-point phase, for $K_I \approx -.5$, just to the left of the rate instability where time series diverge (white) has shrunk considerably compared to spontaneous (baseline) case and is now predominantly feedforward in all layers and frequency-bands.

bottom-up condition: $I^{bg} = 1$, $I^{LGN} = 2$, *fig. 5.6B*

Stimulation of the upper column's layer 2 and 5 ($I^{Col_2L2} = I^{Col_2L5} = 2$, see figure 5.6C), mimicking a pure top-down signal, like for example in the cue-period of an attention-experiment (cf figure 2.1), silences layer 6 in a great part of the parameter space. The remaining layers show phase relationships similar to the spontaneous case, except that the working-point phase appears smaller.

cue-period condition: $I^{bg} = 1$, $I^{Col_2L2} = I^{Col_2L5} = 2$, *fig. 5.6C*

When both bottom-up and top-down stimulation is applied (figure 5.6D), mimicking the response period of an attention experiment (cf figure 2.1), the phase-diagrams become clearly layer- and frequency-dependent. Just below the asynchronous region there is, again, a small band of K_I-values, for which the system is in feedback-mode. Beyond that, the lower column's upper layers don't exhibit a consistent phase-relationship in the L2E-β-band, whereas the deep layers are phase-lagging, consistent with a feedback configuration. For the γ-bands, layers 2 and 5 become feedforward, layer 4, on the other hand, is in-phase, with a slight tendency to feedback. In the working-point phase for slow frequencies layers 2 and 5 are feedback, and layer 4 assumes all possible phase-relationships depending on $K_{E,I}$, likewise phases vary with $K_{E,I}$ in the working-point phase of the L2E-γ-band, but for the L4E-γ-band the columns are in feedforward-mode.

response period condition: $I^{bg} = 1$, $I^{LGN} = I^{Col_2L2} = I^{Col_2L5} = 2$, *fig. 5.6D*

We computed these dynamic response profiles also for the alternative model possessing an artificially introduced slow oscillator in layer 5 (figures 5.7A,B) . There, the phase-relationships are highly variable for $K_I \gtrapprox -1$ and often inconsistent (bleached out colors) for $K_I \lessapprox -1.5$. In between ($-1.5 \lessapprox K_I \lessapprox -1$) the system is in feedforward

alternative model with slow oscillator: fig. 5.7A,B

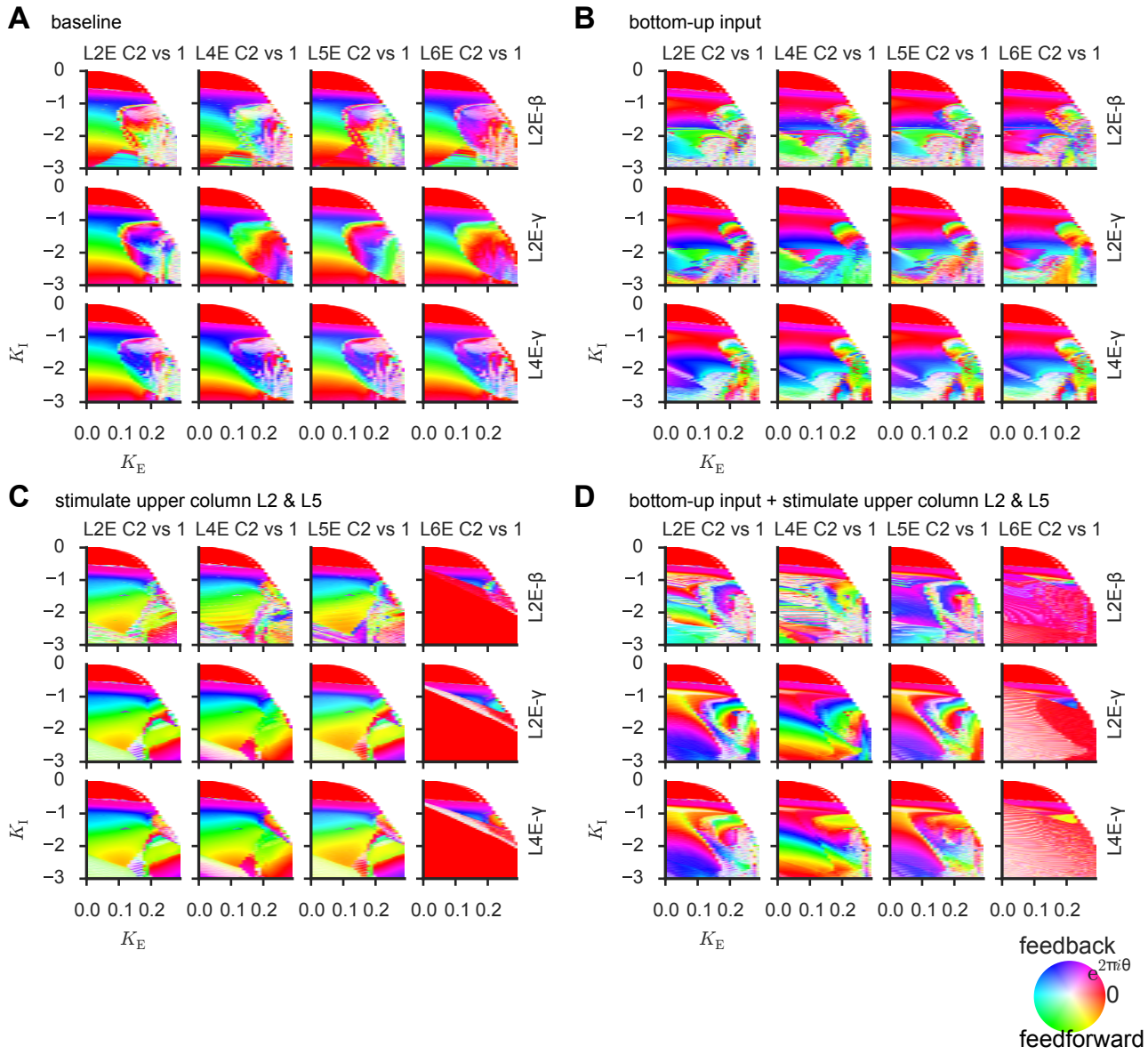


Figure 5.6: Dynamic response profiles for relative oscillation phases, dependent on context. A) During spontaneous exploration the working-point phase, around $K_I \approx -2$ and near the rate white diverging region, is predominantly in feedback-mode (upper column's maxima leading those of the lower column) for the L2E-β- and L4E-γ-band, whereas in the L2E-γ-band, the phase-relationship is inhibition-dependent. **B)** For all three frequency-bands, during bottom-up stimulation, this working-point phase shrinks, and switches mainly into feedforward mode (lower column's maxima leading those of the upper column). **C)** During top-down stimulation of the upper column, the working-point phase has also shrunk. **D)** When both bottom-up and top-down stimulation is applied, processing modes become layer- and frequency dependent. In the L4E-γ-band it is feedforward, in the L2E-γ-band the mode is inhibition-dependent, and in the L2E-β-band it is feedback, with the exception of a parameter region in layer 4, where it is feedforward.

mode for the γ-bands when stimulated with bottom-up input (5.7B), but in the β-band, the phase-relationships are messy.

Moreover, we have also computed these relative phases for one randomized column-connectome (figure 5.7C,D), which was qualitatively different from the original one, but nevertheless gave rise to qualitatively similar behavior, with similar regions recognizable in the response profiles (compare figure 4.12). Remarkably, also for two coupled columns the phase-diagrams, and thus the emerging dynamics, for the selected randomized connectome are very similar to those for the original connectome, suggesting that the structural degeneracy principle extends beyond the single-column case. This needs to be studied more systematically, though.

"good" randomized connectome: fig. 5.7C,D

5.3 DISCUSSION

We have presented preliminary evidence that when two columns at different stages of the cortical hierarchy undergo spontaneously oscillations in the absence of specific bottom-up or top-down inputs, the upper column is mostly phase-leading the lower column (although, for some coupling efficacies, there is a considerable drift), consistent with a feedback-dominated processing mode. Exogenous and Endogenous input biases can switch the phase-relationships, however, so that effective connectivities become "feedforward", and when, in addition, the processing is modified through top-down signals like (like, for example, attention) then there exist a whole phase of coupling efficacies for which phase-relationships are predominantly feedforward in a fast (gamma-like), but predominantly feedback in a slow (beta-like) frequency band.

We have tagged phase-leads and -lags of the lower relative to the upper column as a directionality in coupling, feedforward and feedback, respectively. This assumption is based on computational studies which have demonstrated that such phase-differences are associated with an asymmetry in the information contained in time-series that are out of phase, as measured by transfer entropy [Battaglia et al., 2012] and delayed mutual information [Kirst, 2012]. These studies have further indicated that information contained in spikes sent between two populations might be equally asymmetric with the same directionality. If this holds true also in our setting, if information can be transferred concurrently in multiple frequency bands, and the role that each layer plays in that process, remains to be studied by implementing a spiking-network version of our model.

Evidence for frequency-dependent directionality has been found previously in experiments [Kerkoerle et al., 2014; Bastos et al., 2015b]. To the best of our knowledge, our model is the first to reproduce this pattern; moreover, a competing group's attempt are based only on a simplified, phenomenological layer structure with distinct fre-

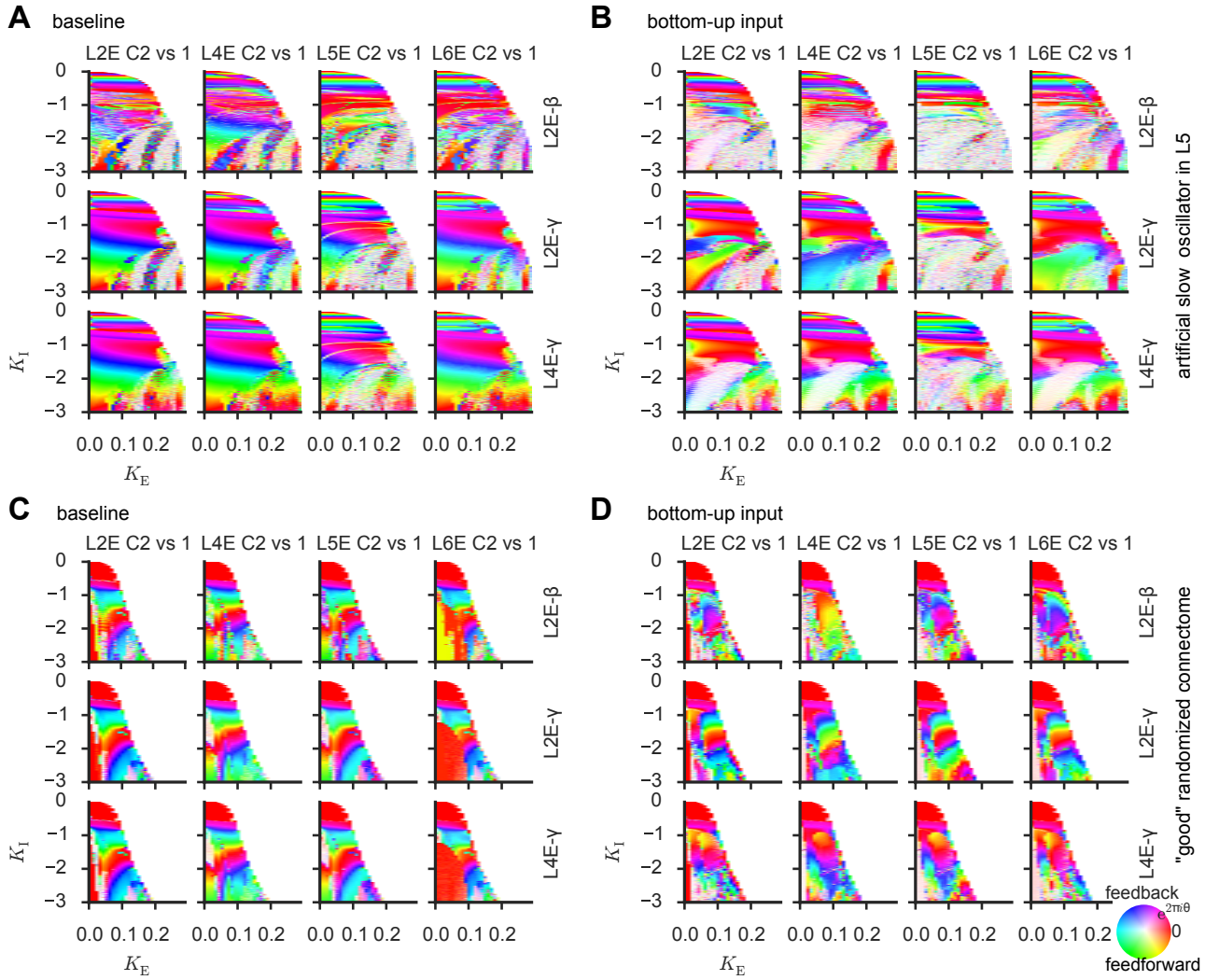


Figure 5.7: Alternative connectomes can destroy or maintain feedforward- and feedback-modes. When an intrinsic slow oscillator is introduced in layer 5 **A)** during spontaneous and **B)** during bottom-up stimulation, relative phases become strongly $K_{E,I}$-dependent and often even inconsistent. On the other hand, **C)** during spontaneous and **D)** stimulation with the column wired according to a qualitatively different connectome, which nevertheless produces predominantly fast- and slow- oscillations in a region of the $K_E - K_I$-parameters pace when one column is considered in isolation, the phase-diagrams show qualitatively similar features as those for the original column-connectome.

quency resonators [personal communication, X.J. Wang]. Our model depends on the experimentally determined local-circuit connectome measured by [Binzegger et al., 2004]. We have yet to study systematically the dependence of the pattern of phase-relationships on the connectome. However, as we have previously shown, that arbitrary connectomes do not give rise to layer-specific oscillations which are predominantly fast and slow in upper and lower layers, respectively, —another experimentally observed result [Kerkoerle et al., 2014]—, and as a randomized, qualitatively different connectome which does nevertheless exhibit the same layer-specific power-dependence shows also a similar pattern of phase-relationships, as we have reported here, we expect that only selected connectomes are capable of exhibiting that pattern.

In our model, slow oscillations are network-generated, but we have also demonstrated that an intrinsic slow oscillator in deep layers might lead to similar results when a single column is considered in isolation. Here, we go beyond these results, showing that for this alternative model qualitatively different phase relationships might occur, that are much more variable for changing coupling efficacies, and often even inconsistent over time. A somewhat hand-waving explanation is that the introduced artificial oscillator is too “strong", too “independent” from the others, whereas in our original model slow oscillations appear due to interactions and are thus much better “integrated” with the fast oscillations. While there are certainly other ways in which an intrinsic oscillator in a deep layer could be included in the model, which would interact differently with the remaining layers, our model might tentatively hint that a mechanism that produces slow oscillations through the interplay of network participants might be favorable to one relying on the presence of a dedicated oscillator.

We note that feedback connections in our model are mediated solely by L5E for simplicity, but it has been suggested that L6E is also involved [Douglas and Martin, 2004]. There is also evidence that thalamus can activate layer 5 directly [Constantinople and Bruno, 2013], which we have not considered. Moreover, the pulvinar might play a role [Shipp, 2003; Saalmann et al., 2012]. The effect of the precise pattern of feedforward and feedback connectivity, in so far as it is not precisely known, needs to be analyzed.

Other patterns of stimulation beyond those discussed, for example stimulating layers of the upper column, or periodic stimulations will be necessary to determine how the observed pattern of phase relationships can be manipulated, possibly in a frequency dependent way.

We have yet to study more systematically the effect of long-range coupling and delay. Furthermore, it would be interesting to see if the system was multistable, as this would allow processing modes in the absence of structural changes.

Understanding and being able to control the coordinated dynamics between two columns will allow us to proceed in our modeling effort to multiple coupled columns, along the lines of top-down, connectome-based whole-brain models [Deco et al., 2011]. While certainly further methodological obstacles have to be overcome in order to deal with the wealth of data, our hope is that this approach will—eventually—shed light on the genesis of large-scale multi-frequency coherence networks.

5.4 MATERIALS AND METHODS

5.4.1 *Model*

Each column is modeled exactly as described in the previous chapter. We extended that model by introducing feedforward connections from the lower column's layer 2E to the upper column's layer 4 (both the excitatory and inhibitory population), and feedback connections from the upper column's layer 5E to the lower column's layer 2 and 5 (both the excitatory and inhibitory populations). The complete equations read

$$\begin{aligned}\tau_{\aleph\alpha k}\dot{r}_{\aleph\alpha k}(t) = &- r_{\aleph\alpha k}(t) + \\ &F\Big(I^{bg}_{\aleph\alpha k} + I^{LGN}_{\aleph\alpha k} + I^{hor}_{\aleph\alpha k} + I^{td}_{\aleph\alpha k} \\ &+ \sum_{\imath\in Layers}\sum_{\beta=E,I}\gamma^{\imath\beta}_{\alpha k}K_\beta r_{\aleph\imath\beta}(t-D) \\ &+ K_{lr}\sum_{\beth=1}^{\aleph-1}\delta_{\alpha,4}r_{\beth 2E}(t-D_{lr}) \\ &+ K_{lr}\sum_{\beth=\aleph+1}^{N_{Cols}}(\delta_{\alpha,2}+\delta_{\alpha,5})r_{\beth 5E}(r-D_{lr})\Big)\end{aligned}$$

where $N_{cols} = 2$, $1 \leqslant \aleph, \beth \leqslant N_{cols}$ denote columns, ordered ascendingly according to their position in the cortical hierarchy, $\delta_{\alpha,\varkappa}$ is the Kronecker-symbol, K_{lr} scales the efficacy of the long-range connections, D_{lr} is the delay of long-range connections (we chose $D_{lr} = 10D$) and all remaining notations are as previously described.

5.4.2 *Model with intrinsic slow oscillator in layer 5*

As laid out in the previous chapter multiplying both γ^{5I}_{5E} and γ^{5I}_{5I} by 400 introduces a slow oscillator in layer 5.

5.4.3 *Oscillation phases*

The model equations are integrated as described in the previous chapter and time series for 2^20 time steps (after an initial transient) obtained. Fig. 5.4 illustrates how relative phases for oscillations of a predominant frequency are calculated: The power spectrum was estimated with Welch's method as in the previous chapter, then the global maxima in the power spectral density of layer 2 in the frequency band 0-30 Hz, of layer 2 in the frequency band 30-90 Hz, and of layer 4 in the frequency band 30-90 Hz was determined, and for each of these frequencies the time series of all populations, subsampled by a factor of 256 for performance reasons, were filtered with a finite-impulse-response filter, using a Kaiser-window of length 2^10 with parameter $\beta = 3.5$ and a pass-band of ± 2 Hz around the peak frequencies.

To determine the relative phase between the oscillations in populations a relative to population b, local maxima were detected in the filtered time series of these populations. Assume for the moment that population b possesses at most as many local maxima as population a. For each local maximum $t_i^{(b)}$ of b we then determined the closest local maximum of a occurring before at time $t_i^{(b)}$. Let $d_{\text{right}} = t_i^{(b)} - t_i^{(a)}$ and $d_{\text{left}} = t_i^{(a)} - t_{i-1}^{(b)}$. Therewith, we defined the relative phase as $\theta = \langle \exp(2\pi i d_{\text{right}}/(d_{\text{left}} + d_{\text{right}})) \rangle_i$, where $\langle \cdot \rangle_i$ denotes averaging over all local maxima of b. In case population b had more local maxima then population a we used $1 - \theta$.

6 DISCUSSION

6.1 ATTENTION: A COMPLEX SYSTEM

In this work, we have analyzed several phenomena related to neuronal correlates of attention and found, in all cases, a bewildering complexity.

We started off by analyzing experimental data recorded to elucidate differences between spatial and feature attention. Single cell responses to a given stimulus were already highly heterogeneous, potentially because only few trials per condition could be recorded—such that the heterogeneity would vanish if more recordings were made—but on the other hand this heterogeneity could also be "true"—present even if more trials could be recorded. On the next level, we saw that the response repertoire to a set of stimuli (the tuning curve) differed greatly between neurons. Together, these highly stimulus- and cell-specific responses reflect the complexity of the system, the highly non-random structural connectivity, the intricate dynamics of neurons on top it. It is conceivable that evolution might have found a way to harness the complexity for efficient information coding [Chelaru and Dragoi, 2008; Padmanabhan and Urban, 2010].

As a result of the high variability it was also necessary to develop analysis methods for tuning curves that are more flexible than conventional fitting approaches. We found that this was possible by describing features of interest algorithmically. These descriptions could be either hand-crafted as we have done here, or even computationally mined for in order to optimize a target criterion, like for example the discriminability between conditions.

Another consequence of the complexity was that phenomenological models of attention, like the BC model or the FSGM were not fully compatible with the new data. BC phenomenology held for only a small subset of cells, and the FSGM might be able to describe the responses for spatially separated stimuli within one RF, but probably not for two transparently overlaid ones. Existing circuit models often attempt to reproduce BC phenomenology, and it is consequently not straightforwardly clear if they could reproduce the new data (which none of them attempted so far).

To describe the population responses in this dataset we hypothesized that they could emerge in a model of multiple interacting rings, representing hypercolumns in V1 and V5, and that attention would reshape the local dynamics thereby biasing responses towards one out of two stimuli. Constraining the circuit with the experimental

data, we found that a whole set of circuits was compatible with them, and that this set might even contain circuits that, despite producing the experimental data measured in V5, are nevertheless qualitatively different regarding the predicted responses in V1. On the other hand, some of the V1 responses seemed unintuitive, possibly not physiological, suggesting that the circuits were not constrained enough. We suspected that one major short-coming of the model was that the effect of attention was modeled as an external contribution outside the V1-V5-network, which we did for simplicity. Instead, attention might be better modeled in a more complex circuit, including an attentional source area as a network element.

In a second line of research, we were interested in the emergence of flexible coherence networks, which are possibly involved in CTC and have been implicated in attention. There, we found that laminar patterns of oscillations—analogous to experimental observations—might not be possible if the circuit topology is random. Rather, specific (but possibly not unique) circuitries might be needed. While one of these was based on experimentally determined connectivities for a cortical column, we cannot—due to measurement errors and the limited size of the reconstructed circuit—be sure, that the "good" connectomes, are actually also realistic connectomes, and we are therefore awaiting more and better connectivity data. Moreover, we saw that a multitude of qualitatively different dynamical phases were supported by the model—phenomenologically, tentatively related to "brain states" [Gilbert and Sigman, 2007]—and that oscillations created due to the interactions between layers might be favorable to structural oscillators.

Subsequently, we were interested to see how the dynamics self-organizes when two such columns, assumed at different levels of the cortical hierarchy, are coupled. Our preliminary results indicate that, again, several phases of qualitatively different dynamics exist. Interestingly, we discovered one, in which feedforward- and feedback-processing in the gamma- and beta-band, respectively, can co-exist. That result, in line with currently discussed ideas about the role of cortical oscillations, could indicate that populations might coordinate their activity and bidirectionally transmit information through multi-frequency oscillations. However, we have not yet quantified the information transmitted by population firing patterns for which we would need a spiking network model.

We don't yet, however, have results corroborating this idea in the model. We also still need to understand more in detailed influences of the local- and long-range connectome, coupling strengths, transmission delays, contextual stimulations and if the highlighted pattern of processing modes is an outlier or emerges robustly provided "good" (which needs to be defined) circumstances.

6.2 HOW TO MODEL ATTENTION (MAYBE)

We have reviewed models for visual attention in the introduction, and have, in addition, introduced one (and a half) new models in this work. Now, we will critically revisit them in order to elucidate the most promising path towards a mechanistic model of attention.

Current models for attention aim at reproducing BC phenomenology, but our results cast doubt on BC as being generally applicable.

We have discussed in chapter 3, that the BC model doesn't seem to be valid without reservation, as in the dataset that we have analyzed it held only for very few stimulus combinations when the model was checked against the trial-averaged responses, and it held only for around 10 %-30 % of cells when they were tested individually. Yet, most models of attention, that we have discussed in section 1.3.1, attempt (among other things) to reproduce BC phenomenology. This casts serious doubt on these models but it does not necessarily invalidate them. Indeed, Ardid et al. [2007] have shown that the same model can give rise to several phenomenological descriptions at the same time, and that restricted changes to the circuit can be enough to abolish or introduce these phenomena. Accordingly it is conceivable that also other models might be able to reproduce other phenomenological descriptions of attentional firing rate modulations, or can, at least, be tweaked to do so. This would have to be studied on a case-by-case basis for each model.

Response heterogeneity is mostly neglected by current models, circuit models (like ours) could potentially address it.

Our results in chapter 3 further show that attentional firing rate modulations can be not only highly cell- (as known before, cf. introduction 1.1.2) but also highly stimulus-specific. Without knowing the "neural code" it is impossible to state with certainty the relevance of this response specificity, but it was demonstrated that it could potentially be functionally beneficial for the amount of information encoded in population responses [Chelaru and Dragoi, 2008; Padmanabhan and Urban, 2010]. Ideally, thus, models should address the heterogeneity. Of all the models discussed solely Wagatsuma et al. [2013] reported that attentional firing rate modulations occurred in only around 60 %-90 % of model cells and could have both increasing and decreasing effects, unfortunately, though, they did not document the stimulus specificity albeit their hypercolumn model generates the necessary data. Wagatsuma et al. [2013] used a large-scale spiking network model, and it is conceivable that a certain stochasticity in the circuit wiring is sufficient to obtain specific effects (cf. [Battaglia and Hansel, 2011; Hansel and Vreeswijk, 2012]. If true, then one could also expect to see such specific effects in the model of Ardid et al. [2007] who did not study it. Nevertheless, it would be instructive to contrast the occurrence of such specific effects in the model by Wagatsuma et al. [2013], who used a realistic column topology based on the Binzegger et al. [2004] data, with that in the model by Ardid et al. [2007] who represented each "column" by a single pyramidal neuron (with an additional interneuron per four pyramidal neurons), in or-

der to reveal the potential role of the column architecture. Neither the ring-model presented in chapter 3, nor the column model (chapters 4 and 5) (nor the remaining models presented in the introduction, including the phenomenological models of attention can address, in their current form, the question of specific effects. In contrast to these other models, the models presented in this work are, however, circuit models and could, thus, also be reconsidered in a spiking network form in order to analyze specific effects.

Due to firing rate heterogeneity models for attention might first focus on collective phenomena, like oscillations; rate models are then appropriate.

Firing rates and their modulations are considered a hallmark feature of attention. Nevertheless, averaged rates are not necessarily representative for the single cell behavior [Golowasch et al., 2002] and thus—potentially at least—not meaningful. Moreover, the precise nature of the deviations from the mean is not well documented. Consequently, it might be prudent at the moment, when setting up a model of attention, to first focus on other neural correlates of attention. While the ring-model from chapter 3 was specifically designed to reproduce average firing rate effects—although it can also be applied to single cells, as we have shown—(and was, historically, conceived and worked out before we discovered the specific effects) the column model focuses on the occurrence of oscillations. Oscillations and their attentional modulations in the cortex are best seen in the LFP and are, thus, collective phenomena [Yu and Ferster, 2010]. Accordingly, rate models for the average activity of a population of neurons might be suitable and minimal descriptions for these oscillations.

Replacing a highly recurrent network by an effective, unidirectional one—like most models of attention do—is generally problematic.

Most models of attention model "attention" as an external feature that has a certain effect on the (sensory) system, for example as constant current injections, either positive or negative, sent to excitatory and / or inhibitory cells, as synchronous or asynchronous inputs, as delays between two inputs sent to different cell classes or as the arrival phase of the input (e.g. [Buia and Tiesinga, 2008, 2006; Tiesinga et al., 2004; Tiesinga, 2005], cf. section 1.3.1). In all cases BC phenomenology or attentional coherence modulations can be reproduced in the model—although the descriptions of attentions are, apparently, very different. This fact demonstrates that, in general, many mechanisms are compatible with a given observed effect. (Moreover, artificial neural networks can approximate any sufficiently regular function with arbitrary precision [Cybenko, 1989], so a more or less "biologically plausible" mechanism can always be found.) Thus, there is no shortage of ideas how to model attention, at least "effectively", but "effective" descriptions might in the worst case not be very useful. Indeed, it could be that the multiplicity of mechanisms is due to truly different ways that attention could act to produce an effect and in that case combinations of more detailed analyses of the model and experiments will be able to eventually determine the real mechanism. On the other hand, due to the recurrent nature of neural networks in the cortex [Binzegger et al., 2004; Sporns and Kötter, 2004; Song et al.,

2005; Perin et al., 2011] one also has to consider the possibility that a single given mechanism possesses several "effective" manifestations when it is cut out of the recurrent system, only unidirectionally affecting the system but not, in return, being affected by the system; also, different global states of the recurrent network might need different "effective" ways to reintroduce the cut out part. It seems thus, that it might be impossible to pinpoint the mechanism in an "effective" description. Rather, it would be desirable to consider models in which attention emerges from within the system—as the one by Ardid et al. [2007, 2010].

Ardid et al. [2007, 2010]'s model reproduces many neuronal correlates of attention, but potentially not our reported tuning curves for transparent stimuli.

Indeed, in my opinion this is the most convincing among currently discussed models of attention and captures, as discussed in the introduction, a great number of neuronal correlates of attention in the firing rate, variability and the oscillation domain. The model uses a minimal circuit for a hypercolumn, consisting of only one excitatory neuron per "column" plus an additional inhibitory neuron per four excitatory ones. Due to this design, the model cannot address layer-specific coherence modulations, nor cell- nor stimulus-specific effects of attention. (Battaglia and Hansel [2011] studied a model consisting of two coupled rings, representing phenomenological layers, showing that such a design might be beneficial for fast decorrelation of oscillations and heterogeneous tuning properties, but did not address attention.). The model's circuit was, however, described by its authors as a "backbone" for the mediation of attentional effects and each "column" in the circuit could straightforwardly be equipped with a realistic layered topology to study layer-specific effects. Moreover, one could study how sufficiently strong deviations from a given connectivity scheme affect the circuit response and under which circumstances as well as how these deviations could give rise to cell- and stimulus specific effects. In a similar vein, noise correlations due to these deviations could be analyzed, on which the authors did not report. Also, when Ardid et al. considered stimulus pairs these always consisted of a preferred and anti-preferred one, and note that this might be one cause for discrepancies when comparing their results relating to the BC and FSGM with ours. In particular, as other models, also this one reproduces BC phenomenology, beyond that, however, it suggests that when the BC predictions are violated—as in the majority of cells in our dataset—the bottom-up and local excitatory drive to inhibitory interneurons in MT might be too weak. Finally, their simulations suggested that the FSGM also holds for pairs of transparently presented stimuli. The dataset presented in chapters 2 and 3 is now the first available experimental test for this prediction of the FSGM. Our analyses there revealed differences in how attentional modulation affects tuning curves for spatially separated and transparently presented stimuli within a single RF respectively. In particular, in the spatially separate paradigm, the right ("attended") peak was

increased while the left one was decreased as an effect of directing attention to one of the stimuli in the RF, and the position of the minimum between the peaks was shifted to the left as well as decreased compared to the *attend-fix* condition. In comparison, in the transparent paradigm, both peaks were increased by approximately the same amount and the minimum between the two peaks was not significantly affected by attention. Besides, we observed that the *attend-in* tuning curve in the spatially separate paradigm had a "flatter" (more platycurtic) right peak than the transparent one. The model by Ardid et al. [2007] consists of a single hypercolumn and possesses, thus, no notion of "space". Consequently, when they present two stimuli simultaneously this mimics a "transparent" stimulus display. However, the tuning curve they present features two peaks of unequal height—like the one we obtain in the spatially separate, but *not* in the transparent paradigm. On the other hand both of their peaks seem equally wide, but they don't analyze that specifically. Besides their simulated attention-induced population-response modulation ratio follows the predictions of the FSGM (i. e. is a monotonous function of $\theta_{\text{attended}} - \theta_{\text{preferred}}$) and was, in their simulations, very similar for both single and dual stimulus presentations. Our data analysis showed that the modulation response function might conform to the FSGM in the transparent, but probably not in the spatially separate paradigm. The model, thus produces a curious mix of the results for the spatially separate and transparent paradigm—without getting one completely right. It is clear that, in order to address our data, the model needs to be extended to include a notion of space, but it is not straightforwardly apparent how the complex differences between the two paradigms that we observed can be reproduced.

Model attention as dynamical states, and study how these contribute to information routing.

The model for attention, that we envisage, is different from, but not necessarily incompatible with the Ardid et al. model. Due to the multitude of potential mechanisms for attention and because recently published structural connectivity data permits us to do so, we constrain the circuitry with realistic experimental data and study the emergent mechanism(s). Equipping this structural skeleton with even very simple local dynamics and a suitable coupling between neuron populations, can give rise to a myriad of possible dynamical states, potentially resembling experimentally determined networks [Hansen et al., 2015]. These states might occur in the model either due to meta- or multistability, or due to changes in parameters, which—phenomenologically—are controlled by noise, ongoing activity, and neuromodulators, such that the "mental life" can be seen as a progression of different states [Gilbert and Sigman, 2007], amongst them "attentional" ones [Harris and Thiele, 2011]. (We don't make any statement if the states are stable, metastable or unstable, what causes transitions, what happens during transitions, and if there is structure in the order in which states are visited by the dynamics.) Our fundamen-

tal assumption is that (some of) these states possess dynamical properties resembling those seen *in vivo*, for example the default mode network, or the dorsal and ventral attention network. If we can succeed in identifying and matching simulated and experimental dynamical states, for instance on the basis of their oscillatory fingerprints, this would allow us to analyze in detail the properties and differences between the states. In particular, we would be interested in how dynamics shapes the interplay between areas, that is, how a given sensory input "flows" through the network, and which layers of which areas at what time are (co-)activated in what frequency-bands—given the global network state. The co-activation patterns of areas on a brain-wide level might ideally even resemble large-scale networks identified in fMRI experiments, like the default-mode-network, or the dorsal and ventral attention network [Corbetta et al., 2008], or large-scale coherence networks [Siegel et al., 2008; Hipp et al., 2011, 2012; Bosman et al., 2012]. At the moment it is not clear why these network exist, why for example the frontal cortex is involved in attention, but in a suitable and minimal model we could study how these networks are generated through the interplay of structural connectivity with the local dynamics and an exciting question would be if the attention-networks are associated with enhanced top-down communication. Furthermore, as attention can be voluntarily switched between features within a fraction of a second this excludes structural changes as underlying causes and suggests that the "attentive state" actually consists of a whole repertoire of "representations", each one corresponding to a particular set of attended features. We would thus aim at inventorying dynamical states whose global activation patterns are similar to the dorsal and ventral attention network (defined by fMRI studies) and search for "representations" within these states which facilitate or suppress sensory areas in various distinct ways, corresponding to all possible units that attention can select (in the framework of this model), for example, attending to a particular location or motion direction, and it would be interesting to see if and how different features can be jointly selected. Moreover, we could study how the interplay of the structural connectivity with the local dynamics in each node creates a flexible functional connectivity, try to understand how information is routed through this network, and how attention modulates this routing. Like Ardid et al. we could also perform surgical interventions on the circuitry in order to pinpoint those parts of the circuitry that are critical for the occurrence of specific attentional effects.

Do we get large-scale brain networks as "free lunch"? If so, study effectivity of interareal communication in attention networks.

A priori it is unclear what level of description should be used. Indeed, it has been argued that computations are performed at a subcellular level, for example in dendrites [London and Häusser, 2005; Branco and Häusser, 2010] or even in molecular networks [Tyson et al., 2003; Bhalla, 2014]. At such a level modeling and simulation

Rate model with layered topology might be an appropriate level of description.

analysis would, however, be very challenging, moreover it would require huge amounts of experimental data to constrain parameters and yet contain a plentitude of other parameters that (ideally) need to be scanned in the simulation. To illustrate, consider our model as a contrast. Our model includes layers because connections between areas, as well as the oscillatory fingerprints and their attentional modulations, that we aim to reproduce, are layer-dependent. Dynamics is modeled at the level of populations, i. e. with a rate model description for one excitatory and one inhibitory population per layer, and the local dynamics follows a simple scheme (a simplified Wilson and Cowan [1972] model), that can be enriched if necessary. Coupling between the populations induces additional currents in a population that are proportional to the presynaptic activity and the coupling strength which is taken to be the measured anatomical connectivity data. These coupling-currents are delayed, mathematically in order to produce oscillatory activity, and biologically, because of actually occurring transmission delays. This is probably the most minimal description possible for what we want to model. It contains the values of delays, external input currents and the factors scaling the influence of the connectome as parameters and it is still feasible to scan the corresponding parameter space given sufficient computing power. However, as each additional parameter multiplies the volume of the parameter space, any additional detail included in the model makes it vastly more complex to analyze. We therefore think that our minimal description is a good starting point because essential mechanisms cannot be clouded by a wealth of potentially unnecessary details. Moreover, several models have already established a good match between spiking and rate models [Battaglia et al., 2007; Deco and Jirsa, 2012; Deco et al., 2013a] so that transitioning to a more detailed model is possible if necessary. We will not, as already mentioned, be able to reproduce the specific effects on firing rates or noise correlations with precisely this model, but once we understand its behavior we can migrate to a more fine-grained description striving to keep the same properties while allowing specific single-neuron behavior in addition.

Realistic structural connectivity is potentially necessary for model.

Potentially, the model could be simplified if instead of the complex column and long-range circuits random or unstructured connection matrices would be used. However, connectivity in the cortex is far from random [Sporns et al., 2002; Sporns and Zwi, 2004; Song et al., 2005; Perin et al., 2011; Harris and Mrsic-Flogel, 2013], which might have functional relevance, as connected cells have an elevated probability for similar response tunings [Bosking et al., 1997; Ko et al., 2011; Li et al., 2013; Cossell et al., 2015]. Theoretical considerations have even suggested that the "real" connectome might be advantageous for efficient input-output transformations [Haeusler and Maass, 2007], although this view has been challenged [Stoop et al., 2013]. Furthermore, our simulations with randomized connectomes

in chapter 4 revealed that those connectomes, for which we found layer-specific multi-frequency oscillations like in experiments, were not unique. This could be interpreted as arguing that the precise knowledge of the connectome is not informative enough to justify the efforts needed to collect the data. On the other hand, we also saw that the number of "good" connectomes were exceedingly rare, making it essentially impossible to just guess one. Moreover, theoretical investigations have demonstrated that the balanced state [Vreeswijk and Sompolinsky, 1996] is not exactly achievable with random connection topologies [Landau et al., 2014]. Altogether it seems, thus, that precise connectivity data is valuable and necessary in order to obtain results comparable to experiments, and to study mechanisms that emerge from these structural constraints. At a later stage, it is then possible to analyze the precise role of the connectivities.

Limited structural connectivity data is available, but we expect more and improved data.

Hence, our envisaged model depends fundamentally on the quality of the available structural connectivity data, that is the column connectome from Binzegger et al. [2004] and the long-range connectivity data from Markov et al. [2014b]. The column connectome is based on the reconstructions of neuronal arborizations from 39 neurons from area 17 of one cat. The long-range connectivity stems from retrograde tracer studies in 26 macaque monkeys, such that the source but not the target layer of connections is known. Thus, we (have to) assume (for the moment at least) that the column-connectome is "canonic" over areas and species, that the measured long-range connectivities are representative and that their target layers, which we have to collocate from other studies [Markov et al., 2014b] are sufficiently stereotyped. These assumptions are potentially problematic. New connectivity data, both local and global, would therefore be most desirable to update and improve the quality of this model and several large scale projects are currently under way [Kandel et al., 2013]. In the meantime, it is important to keep in mind that due to these shortcomings we cannot expect a precise match between experiments and simulations. Instead, we have to aim at identifying robust mechanisms that give rise to the results that we obtain with the model. Moreover, we generally expect the model to be structurally stable, that is, that sufficiently small variations in the used connectome leave results qualitatively unchanged and we can also hope that, even when the connectomes are qualitatively different, there might exist limited parameter regimes for which nevertheless a qualitatively matching dynamics can be obtained (in the spirit of "structural degeneracy").

Open question: how to analyze wealth of data in such a large-scale model?

In the end, our circuit will consist of several hundred interacting populations of neurons and we need to find a way to understand the dynamics of this complex circuit. We have started in chapter 4 to analyze properties of a single column and have started in chapter 5 to study how two columns interact. Our hope is first, that the methods and results from these chapters generalize and help us in the under-

standing of more interacting columns, and that clues obtained from experiments guide us towards relevant analyses. Besides, it might also be necessary to use machine assisted methods to inventory states and discover structures in the data that go beyond intuitive or commonly used paradigms—much like we have done when extracting features from tuning curves in chapter 2.

6.3 CONCLUSION

In this work, we have contributed several (we hope helpful) pieces towards a better understanding of how attention works in the brain. We have analyzed a dataset that shed new light on how attention modulates sensory responses to complex stimuli. The identified differences between spatially separate and transparent stimulus displays, as well as the high cell- and stimulus specificity of effects proofed to be challenging for all existing models of attention. We have therefore envisaged an alternative circuit model, based on realistic local and long-range connectivity data, in which attentional effects will—we hope—eventually emerge in the form of specific dynamical states. Towards that goal, we have analyzed a model of a single cortical column which will become the building block for the global model. In that model we have observed context-dependent dynamical states, in which due to the specific structure of the interlayer connections upper and lower layers oscillate predominantly with fast (gamma-like) and slow (beta-like) frequencies (for certain parameter regimes). These column-wide complex and layer-dependent oscillations could thus form a dynamical backbone for the interaction and coordination between columns and, indeed, when coupling two such columns representing different stages of the cortical hierarchy we observe, again context- and state-dependent, phase relationships that could benefit bidirectional information transmission between columns by multiplexing in various frequency bands. The model, thus, appears flexible and powerful enough to extend it further to many interacting columns in order to eventually—having overcome all foreseeable and unforeseeable obstacles—understand global phenomena, how structural or dynamical anomalies may lead to clinically relevant symptoms, or, of course, attention, a truly complex system.

A

SUPPORTING INFORMATION FOR CHAPTER 2

Table A.1: List of features defined for all tuning curves. Each feature is calculated once for uni, once for afix and once for ain condition.

Feature name	Description
GlobalMinimumAngle [1]	$\operatorname{argmin} tc(\theta)$
GlobalMinimum	$\min tc(\theta)$
GlobalMaximumAngle	$\operatorname{argmax} tc(\theta)$
GlobalMaximum	$\max tc(\theta)$
PeakToPeak[2]	GlobalMaximum - GlobalMinimum
InnerMinimumAngle[3]	$\operatorname{argmin}_{\theta \in I} tc(\theta)$
InnerMinimum	$tc(\text{InnerMinimumAngle})$
OuterMinimumAngle[4]	$\operatorname{argmin}_{\theta \in O} tc(\theta)$
MaximumAngle^{right}[5]	$\operatorname{argmax}_{\theta \in R} tc(\theta)$
InnerWidth^{right}	(MaximumAngle^{right} - InnerMinimumAngle) mod 360
OuterWidth^{right}	(OuterMinimumAngle - MaximumAngle^{right}) mod 360
Width^{right}	InnerWidth^{right} + OuterWidth^{right}
ΔWidth^{right}	OuterWidth^{right} - InnerWidth^{right}
Maximum^{right}	$tc(\text{MaximumAngle}^{right})$

Continued on next page

1 The value of this circular variable was chosen to lie in the range $[-120°, 240°)$

2 if PeakToPeak $< 10^{-10}$ don't calculate any feature at all

3 $I = [0, 360)$ for uni conditions; for afix and ain first find local maxima for $45 \leqslant \theta \leqslant 315$ (to avoid spurious local maxima require that there's no higher point in the tuning curve within 15° on both sides of the local maxima). If there's no such local maximum $I = (120, 240)$, if there's one at 180° $I = (120, 240)$; if there's one at $\theta_m \neq 180$ $I = (\theta_m, 240)$ or $I = (120, \theta_m)$ depending on if θ_m lower or greater than 180°; if there are two $I = (\theta_0, \theta_{-1})$ where $0 \leqslant \theta_0$ and $\theta_{-1} \leqslant 360$ are the local maxima closest to 0° and 360°; if there are more than two: if all of them are $\leqslant$ or $\geqslant$ 180°also use $I = (\theta_0, \theta_{-1})$, otherwise there's at least one to the left and at least one to the right of 180°; pick the highest on either side; if, in addition, there is a local maximum at 180°, calculate it's distance d to the left and right peak according to $((x_{left/right} - x_{180})/360)^2 + ((y_{left/right} - y_{180})/ptp)^2$; replace the peak which is closer according to the measure d if the peak at 180°is higher than it

4 $O = [0, 360)$ for uni conditions; for afix and ain $O = [0, 360) \setminus I$

5 $R = [180, 300]$ for uni conditions; for afix and ain $R = [\text{InnerMinimumAngle}, 315]$; if $tc(\text{MaximumAngle}^{right}) = \text{GlobalMinimum}$ then MaximumAngle^{right} (and likewise all dependent features) is not defined

Table A.1 – *Continued from previous page*

Feature	Description
PeakToPeak^{right}	tc(MaximumAngle^{right}) - GlobalMinimum
Skewness^{right}[6][7]	$mom_3(\tilde{R})/mom_2^{1.5}(\tilde{R})$
Kurtosis^{right}	$mom_4(\tilde{R})/mom_2^2(\tilde{R})$
Dip^{right}	Maximum^{right} - InnerMinimum
InnerBandwidth$^{right}_{X\%}$ [8]	bw(MaximumAngle^{right},left,X)
OuterBandwidth$^{right}_{X\%}$	bw(MaximumAngle^{right},right,X)
Bandwidth$^{right}_{X\%}$	InnerBandwidth$^{right}_{X\%}$ + OuterBandwidth$^{right}_{X\%}$
tcSymmetryIndex	$\sqrt{\sum_{\theta}(tc(\theta) - tc(360-\theta))^2/2}$

Table A.2: List of features defined only for uni condition. Each feature is calculated only for uni condition.

Feature name	Description
CircularVariance[a]	$1 - \left\lvert \sum_{\theta_n} \tilde{tc}(\theta_n) e^{2\pi i \theta_n/360} / \sum_{\theta_n} \tilde{tc}(\theta_n) \right\rvert$

a $\tilde{tc}(\theta) = tc(\theta) - \min tc(\theta)$

6 $mom_n(X) = \sum_{x \in X} (x - \bar{X})^n / n$ and $\bar{X}$ is the mean of set X

7 if $0 \leqslant$ InnerMinimumAngle $<$ OuterMinimumAngle $\leqslant 360$ then $\tilde{R}$ = [InnerMinimumAngle, OuterMinimumAngle], otherwise $\tilde{R} = [0, 360) \setminus$ [OuterMinimumAngle,InnerMinimumAngle]

8 $bw(m, d, X)$ is defined as follows given the tuning curve tc: starting from position m it returns the first angle θ in direction $d \in \{left, right\}$ for which $tc(\theta) < \frac{X}{100}$PeakToPeak. Additionally, if an inner bandwidths (denoted "I") calculated this way is greater than the distance d of the corresponding peak to InnerMinimumAngle, the bandwidth is set to d

Table A.3: List of features defined only for afix and ain condition. Each features is calcualted once for afix and once for ain condition.

Feature name	Description
MaximumAngle$^{\text{left}}$[a]	$\underset{\theta \in L}{\operatorname{argmax}}\, tc(\theta)$
ΔMaximumAngle	MaximumAngle$^{\text{right}}$ - MaximumAngle$^{\text{left}}$
OuterWidth$^{\text{left}}$	(MaximumAngle$^{\text{left}}$ - OuterMinimumAngle) mod 360
InnerWidth$^{\text{left}}$	(InnerMinimumAngle - MaximumAngle$^{\text{left}}$) mod 360
Width$^{\text{left}}$	InnerWidth$^{\text{left}}$ + OuterWidth$^{\text{left}}$
ΔWidth$^{\text{left}}$	OuterWidth$^{\text{left}}$ - InnerWidth$^{\text{left}}$
ΔOuterWidth$^{\text{left}}$	OuterWidth$^{\text{right}}$ - OuterWidth$^{\text{left}}$
ΔInnerWidth$^{\text{left}}$	InnerWidth$^{\text{right}}$ - InnerWidth$^{\text{left}}$
Maximum$^{\text{left}}$	$tc($MaximumAngle$^{\text{left}})$
PeakToPeak$^{\text{left}}$	Maximum$^{\text{left}}$ - GlobalMinimum
ΔMaximum	Maximum$^{\text{right}}$ - Maximum$^{\text{left}}$
minusSkewness$^{\text{left}}$[b]	$-mom_3(\tilde{L})/mom_2^{1.5}(\tilde{L})$
Kurtosis$^{\text{left}}$	$mom_4(\tilde{L})/mom_2^2(\tilde{L})$
ΔSkewness	Skewness$^{\text{right}}$ - minusSkewness$^{\text{left}}$
ΔKurtosis	Kurtosis$^{\text{right}}$ - Kurtosis$^{\text{left}}$
Dip	(Maximum$^{\text{left}}$+Maximum$^{\text{right}}$)/2 - InnerMinimum
Dip$^{\text{left}}$	Maximum$^{\text{left}}$ - InnerMinimum
OuterBandwidth$^{\text{left}}_{X\,\%}$	bw(MaximumAngle$^{\text{left}}$,left,X)
InnerBandwidth$^{\text{left}}_{X\,\%}$	bw(MaximumAngle$^{\text{left}}$,right,X)
Bandwidth$^{\text{left}}_{X\,\%}$	InnerBandwidth$^{\text{left}}_{X\,\%}$ + OuterBandwidth$^{\text{left}}_{X\,\%}$

a $L = [45,$ InnerMinimumAngle$]$; if $tc($MaximumAngle$^{\text{left}}) =$ GlobalMinimum then MaximumAngle$^{\text{left}}$ (and likewise all dependent features) is not defined

b if $0 \leqslant$ OuterMinimumAngle $<$ InnerMinimumAngle $\leqslant 360$ then $\tilde{L} =$ [OuterMinimumAngle,InnerMinimumAngle], otherwise $\tilde{L} = [0, 360) \setminus$ [InnerMinimumAngle,OuteMinimumAngle]

Table A.4: List of additional features comparing two conditions. Conditions A,B can be either of uni, afix, ain.

Feature name	Description
$\Delta\text{Maximum}^{\text{left}}_{\text{A,B}}$	$\text{Maximum}^{\text{left}}_{\text{B}}$ - $\text{Maximum}^{\text{left}}_{\text{A}}$
$\Delta\text{minusSkewness}^{\text{left}}_{\text{A,B}}$	$\text{minusSkewness}^{\text{left}}_{\text{B}}$ - $\text{minusSkewness}^{\text{left}}_{\text{A}}$
$\Delta\text{Kurtosis}^{\text{left}}_{\text{A,B}}$	$\text{Kurtosis}^{\text{left}}_{\text{B}}$ - $\text{Kurtosis}^{\text{left}}_{\text{A}}$
$\Delta\text{Maximum}^{\text{right}}_{\text{A,B}}$	$\text{Maximum}^{\text{right}}_{\text{B}}$ - $\text{Maximum}^{\text{right}}_{\text{A}}$
$\Delta\text{Skewness}^{\text{right}}_{\text{A,B}}$	$\text{Skewness}^{\text{right}}_{\text{B}}$ - $\text{Skewness}^{\text{right}}_{\text{A}}$
$\Delta\text{Kurtosis}^{\text{right}}_{\text{A,B}}$	$\text{Kurtosis}^{\text{right}}_{\text{B}}$ - $\text{Kurtosis}^{\text{right}}_{\text{A}}$
$\Delta\text{GlobalMinimum}_{\text{A,B}}$	$\text{GlobalMinimum}_{\text{B}}$ - $\text{GlobalMinimum}_{\text{A}}$

Table A.5: Feature pair categories. The features in each row were compaired against each other within one condition (uni, afic or ain) and for all conditions they are defined.

$\text{OuterWidth}^{\text{left}}$, $\text{OuterWidth}^{\text{right}}$
$\text{InnerWidth}^{\text{left}}$, $\text{InnerWidth}^{\text{right}}$
$\text{Width}^{\text{left}}$, $\text{Width}^{\text{right}}$
$\Delta\text{Width}^{\text{left}}$, $\Delta\text{Width}^{\text{right}}$
$\text{minusSkewness}^{\text{left}}$, $\text{Skewness}^{\text{right}}$
$\text{Kurtosis}^{\text{left}}$, $\text{Kurtosis}^{\text{right}}$
$\text{Maximum}^{\text{left}}$, $\text{Maximum}^{\text{right}}$
$\text{PeakToPeak}^{\text{left}}$, $\text{PeakToPeak}^{\text{right}}$
Dip^{left}, $\text{Dip}^{\text{right}}$
$\text{Bandwidth}^{\text{left}}_{75\,\%}$, $\text{Bandwidth}^{\text{right}}_{75\,\%}$
$\text{InnerBandwidth}^{\text{left}}_{75\,\%}$, $\text{InnerBandwidth}^{\text{right}}_{75\,\%}$, $\text{OuterBandwidth}^{\text{left}}_{75\,\%}$, $\text{OuterBandwidth}^{\text{right}}_{75\,\%}$
$\Delta\text{OuterWidth}$,$\Delta\text{InnerWidth}$
$\Delta\text{Maximum}^{\text{left}}$,$\Delta\text{Maximum}^{\text{right}}$
$\Delta\text{minusSkewness}^{\text{left}}$,$\Delta\text{Skewness}^{\text{right}}$
$\Delta\text{Kurtosis}^{\text{left}}$,$\Delta\text{Kurtosis}^{\text{right}}$

Table A.6: Spatially separate paradigm's statistics for all features. Table lists cell count, mean, standard deviation, minimum, 25 % quantile, median, 75 % quantle and maximum for all features when evaluated with the direct method (values from best model in parentheses).

		count	mean	std	min	25%	50%	75%	max
afix	ΔInnerWidth	109	8.26 (4.31)	47.41 (43.80)	-90.00 (-133.90)	-30.00 (-12.50)	0.00 (8.70)	30.00 (24.60)	120.00 (155.90)
afix	ΔKurtosis	109	-0.15 (-0.06)	0.78 (0.99)	-2.69 (-3.61)	-0.61 (-0.35)	-0.11 (-0.06)	0.30 (0.15)	2.48 (3.85)
afix	ΔMaximum	109	2.98 (2.68)	13.98 (15.40)	-63.00 (-68.92)	-3.40 (-5.60)	3.60 (3.21)	10.67 (9.97)	38.80 (47.58)
afix	normalizedΔMaximum	85	0.03 (0.04)	0.41 (0.43)	-1.19 (-1.12)	-0.20 (-0.25)	0.08 (0.07)	0.27 (0.27)	0.98 (1.09)
afix	ΔOuterWidth	109	12.66 (22.18)	68.23 (58.93)	-180.00 (-124.10)	-30.00 (-15.90)	30.00 (32.60)	60.00 (62.60)	150.00 (126.60)
afix	ΔPeakToPeak	109	2.98 (2.68)	13.98 (15.40)	-63.00 (-68.92)	-3.40 (-5.60)	3.60 (3.21)	10.67 (9.97)	38.80 (47.58)
afix	normalizedΔPeakToPeak	85	0.03 (0.04)	0.41 (0.43)	-1.19 (-1.12)	-0.20 (-0.25)	0.08 (0.07)	0.27 (0.27)	0.98 (1.09)
afix	ΔSkewness	109	0.18 (-0.09)	0.92 (1.00)	-2.18 (-2.61)	-0.43 (-0.56)	0.11 (-0.09)	0.86 (0.41)	2.19 (2.95)
afix	ΔWidth	109	11.01 (23.18)	86.47 (84.19)	-180.00 (-166.40)	-60.00 (-23.00)	0.00 (35.60)	60.00 (78.00)	120.00 (177.80)
afix	Dip	109	14.41 (15.79)	9.20 (10.26)	2.47 (2.76)	8.75 (9.80)	12.50 (13.62)	17.30 (18.90)	56.79 (64.28)
afix	normalizedDip	85	0.42 (0.47)	0.22 (0.24)	0.05 (0.06)	0.28 (0.30)	0.39 (0.44)	0.52 (0.61)	1.47 (1.34)
afix	GlobalMaximum	109	43.41 (44.06)	30.97 (32.62)	4.00 (4.70)	22.00 (21.58)	34.50 (35.15)	55.43 (52.43)	167.50 (168.79)
afix	GlobalMaximumAngle	109	192.39 (192.80)	67.36 (65.92)	60.00 (50.30)	120.00 (124.50)	210.00 (222.90)	240.00 (246.30)	300.00 (306.40)
afix	normalizedGlobalMaximum	85	1.05 (1.10)	0.26 (0.26)	0.21 (0.25)	0.87 (0.92)	1.00 (1.08)	1.20 (1.26)	2.14 (1.97)
afix	GlobalMinimum	109	13.52 (13.40)	18.10 (18.71)	0.00 (-1.50)	2.67 (2.79)	7.14 (6.78)	15.67 (14.58)	101.50 (100.00)
afix	GlobalMinimumAngle	109	15.69 (21.42)	50.10 (56.15)	-60.00 (-64.40)	-30.00 (-15.20)	0.00 (12.80)	30.00 (41.60)	210.00 (218.90)
afix	normalizedGlobalMinimum	85	0.22 (0.22)	0.18 (0.18)	0.00 (-0.08)	0.07 (0.07)	0.20 (0.20)	0.32 (0.31)	0.81 (0.77)
afix	innerMinimumAngle	109	174.50 (175.73)	29.77 (30.29)	120.00 (69.60)	150.00 (157.50)	180.00 (174.00)	210.00 (196.20)	240.00 (250.40)
afix	innerMinimumVal	109	23.84 (22.79)	24.05 (23.58)	0.00 (-1.50)	8.80 (7.84)	16.00 (15.29)	29.50 (28.06)	137.20 (128.80)
afix	normalizedinnerMinimumVal	85	0.47 (0.46)	0.24 (0.25)	0.00 (-0.08)	0.30 (0.31)	0.49 (0.46)	0.64 (0.63)	1.09 (1.10)
afix	$\text{Bandwidth}^{\text{left}}_{75\,\%}$	109	82.57 (58.68)	27.74 (26.64)	30.00 (18.90)	60.00 (43.70)	90.00 (53.10)	90.00 (67.50)	180.00 (159.10)

Continued on next page

Table A.6 – *Continued from previous page*

		count	mean	std	min	25%	50%	75%	max
afix	ΔWidth^{left}	109	52.84 (49.12)	54.30 (45.13)	-90.00 (-65.10)	0.00 (18.00)	60.00 (50.60)	90.00 (80.00)	180.00 (152.70)
afix	Dip^{left}	109	12.92 (14.46)	11.38 (12.89)	0.00 (0.00)	5.60 (5.71)	9.80 (11.25)	16.00 (18.30)	63.00 (68.97)
afix	normalizedDip^{left}	85	0.40 (0.45)	0.34 (0.35)	0.00 (0.00)	0.19 (0.20)	0.31 (0.38)	0.52 (0.65)	2.02 (1.82)
afix	InnerWidth^{left}	109	57.52 (58.82)	28.03 (27.58)	0.00 (0.00)	30.00 (43.80)	60.00 (56.30)	90.00 (68.30)	150.00 (156.90)
afix	InnerBandwidth$^{left}_{75\%}$	109	39.36 (27.89)	18.57 (16.50)	0.00 (0.00)	30.00 (19.80)	30.00 (25.50)	60.00 (33.90)	120.00 (121.40)
afix	Kurtosis^{left}	109	-0.93 (-0.92)	0.58 (0.72)	-1.76 (-1.90)	-1.34 (-1.34)	-1.02 (-1.11)	-0.65 (-0.81)	1.32 (2.31)
afix	Maximum^{left}	109	36.76 (37.25)	29.43 (30.37)	1.00 (0.99)	18.00 (16.71)	29.50 (27.77)	47.33 (46.06)	167.50 (168.79)
afix	MaximumAngle^{left}	109	116.97 (116.91)	30.26 (31.93)	60.00 (45.50)	90.00 (97.40)	120.00 (115.80)	120.00 (138.40)	180.00 (190.30)
afix	normalizedMaximum^{left}	85	0.88 (0.91)	0.37 (0.39)	0.12 (0.12)	0.63 (0.66)	0.85 (0.91)	1.16 (1.19)	2.14 (1.97)
afix	OuterWidth^{left}	109	110.37 (107.94)	40.18 (35.31)	30.00 (38.80)	90.00 (78.10)	120.00 (106.70)	150.00 (135.10)	210.00 (191.30)
afix	OuterBandwidth$^{left}_{75\%}$	109	43.21 (30.79)	22.15 (19.38)	30.00 (9.40)	30.00 (21.10)	30.00 (25.80)	60.00 (32.30)	120.00 (124.30)
afix	PeakToPeak^{left}	109	23.24 (23.85)	15.69 (16.42)	1.00 (0.95)	12.67 (12.70)	21.00 (20.89)	29.20 (31.54)	84.50 (85.28)
afix	normalizedPeakToPeak^{left}	85	0.65 (0.69)	0.34 (0.36)	0.12 (0.12)	0.43 (0.43)	0.60 (0.64)	0.83 (0.91)	2.14 (1.96)
afix	Width^{left}	109	167.89 (166.76)	43.03 (44.48)	60.00 (62.90)	150.00 (140.30)	150.00 (160.20)	210.00 (190.80)	270.00 (286.10)
afix	maxAngleDist	109	123.30 (121.94)	36.13 (32.83)	60.00 (37.40)	90.00 (100.50)	120.00 (117.90)	150.00 (140.90)	240.00 (214.30)
afix	minusLeftSkewness	109	-0.10 (0.11)	0.66 (0.66)	-1.69 (-1.58)	-0.58 (-0.35)	-0.11 (0.14)	0.38 (0.48)	1.51 (1.89)
afix	outerMinimumAngle	109	108.99 (147.68)	135.63 (147.99)	0.00 (0.50)	0.00 (21.70)	30.00 (56.20)	300.00 (325.00)	330.00 (359.70)
afix	PeakToPeak	109	29.89 (30.67)	17.09 (18.25)	4.00 (4.34)	16.80 (16.89)	26.80 (26.89)	36.33 (35.29)	87.67 (87.61)
afix	normalizedPeakToPeak	85	0.83 (0.88)	0.28 (0.28)	0.21 (0.23)	0.62 (0.71)	0.83 (0.84)	0.94 (1.00)	2.14 (1.96)
afix	Bandwidth$^{right}_{75\%}$	109	84.22 (58.21)	25.32 (21.02)	60.00 (15.20)	60.00 (45.20)	90.00 (54.80)	90.00 (68.80)	180.00 (129.70)
afix	ΔWidth^{right}	109	60.55 (67.00)	52.60 (39.41)	-120.00 (-62.90)	30.00 (48.00)	60.00 (68.00)	90.00 (91.50)	180.00 (156.60)
afix	Dip^{right}	109	15.90 (17.13)	11.73 (12.77)	0.00 (0.00)	9.00 (8.88)	13.50 (14.56)	20.00 (23.63)	58.86 (63.92)
afix	normalizedDip^{right}	85	0.44 (0.49)	0.27 (0.29)	0.00 (0.00)	0.24 (0.33)	0.42 (0.45)	0.56 (0.60)	1.28 (1.28)

Continued on next page

Table A.6 – *Continued from previous page*

		count	mean	std	min	25%	50%	75%	max
afix	InnerWidth^{right}	109	65.78 (63.12)	31.48 (27.16)	0.00 (0.00)	30.00 (49.60)	60.00 (61.70)	90.00 (74.80)	150.00 (180.00)
afix	InnerBandwidth$^{right}_{75\,\%}$	109	39.36 (26.07)	18.57 (10.21)	0.00 (0.00)	30.00 (21.70)	30.00 (25.90)	60.00 (31.00)	90.00 (56.80)
afix	Kurtosis^{right}	109	-1.07 (-0.98)	0.55 (0.81)	-1.86 (-1.93)	-1.41 (-1.39)	-1.22 (-1.21)	-0.86 (-0.84)	1.78 (2.74)
afix	Maximum^{right}	109	39.73 (39.92)	29.57 (30.35)	3.00 (2.31)	19.33 (18.89)	29.60 (30.34)	48.00 (47.89)	152.67 (163.97)
afix	MaximumAngle^{right}	109	240.28 (238.85)	24.66 (25.48)	180.00 (158.90)	240.00 (223.40)	240.00 (240.20)	240.00 (251.40)	300.00 (306.40)
afix	normalizedMaximum^{right}	85	0.91 (0.95)	0.22 (0.24)	0.21 (0.23)	0.80 (0.82)	0.94 (0.97)	1.04 (1.10)	1.44 (1.42)
afix	OuterWidth^{right}	109	126.33 (130.12)	36.30 (32.07)	30.00 (53.60)	90.00 (109.90)	120.00 (133.90)	150.00 (151.50)	210.00 (196.80)
afix	OuterBandwidth$^{right}_{75\,\%}$	109	44.86 (32.14)	19.42 (15.44)	30.00 (8.50)	30.00 (22.90)	30.00 (28.80)	60.00 (37.30)	120.00 (103.30)
afix	PeakToPeak^{right}	109	26.22 (26.52)	16.19 (17.12)	3.00 (2.31)	14.33 (14.02)	22.90 (23.58)	34.50 (33.21)	87.67 (87.61)
afix	normalizedPeakToPeak^{right}	85	0.69 (0.73)	0.23 (0.26)	0.21 (0.19)	0.54 (0.56)	0.68 (0.74)	0.83 (0.88)	1.30 (1.31)
afix	Skewness^{right}	109	0.08 (0.02)	0.51 (0.56)	-1.86 (-1.64)	-0.21 (-0.21)	0.14 (0.06)	0.42 (0.36)	1.27 (1.46)
afix	Width^{right}	109	192.11 (193.24)	43.03 (44.48)	90.00 (73.90)	150.00 (169.20)	210.00 (199.80)	210.00 (219.70)	300.00 (297.10)
afix	tcSymmetryIndex	109	0.40 (6.86)	0.13 (2.20)	0.14 (2.10)	0.30 (5.54)	0.43 (7.08)	0.48 (8.33)	0.76 (12.45)
afix vs ain	ΔGlobalMinimum	109	1.95 (1.57)	5.26 (4.27)	-16.67 (-7.38)	-0.20 (-0.08)	1.14 (0.35)	3.36 (2.20)	30.25 (19.22)
afix vs ain	normalizedΔGlobalMinimum	85	0.03 (0.03)	0.09 (0.06)	-0.20 (-0.14)	-0.02 (0.00)	0.02 (0.02)	0.08 (0.08)	0.29 (0.17)
afix vs ain	ΔLeftKurtosis	109	-0.14 (-0.03)	0.73 (0.79)	-1.96 (-1.20)	-0.67 (-0.49)	-0.14 (-0.16)	0.28 (0.13)	2.59 (3.59)
afix vs ain	ΔLeftMaximum	109	-1.05 (-2.16)	8.95 (7.45)	-29.94 (-14.82)	-6.00 (-7.01)	-1.00 (-3.10)	2.75 (1.64)	28.90 (24.98)
afix vs ain	normalizedΔLeftMaximum	85	-0.08 (-0.14)	0.22 (0.25)	-0.60 (-0.76)	-0.21 (-0.24)	-0.05 (-0.17)	0.04 (0.01)	0.59 (0.41)
afix vs ain	ΔMinimumusLeftSkewness	109	0.08 (0.00)	0.76 (0.50)	-1.65 (-0.95)	-0.48 (-0.29)	0.10 (0.02)	0.47 (0.33)	2.43 (0.91)
afix vs ain	ΔRightKurtosis	109	0.00 (-0.03)	0.76 (0.48)	-2.87 (-1.93)	-0.30 (-0.20)	0.00 (0.01)	0.46 (0.21)	3.27 (1.02)
afix vs ain	ΔRightMaximum	109	6.37 (3.10)	9.95 (6.25)	-14.67 (-11.49)	0.07 (-0.57)	4.57 (3.05)	10.18 (6.25)	45.90 (23.82)
afix vs ain	normalizedΔRightMaximum	85	0.18 (0.09)	0.32 (0.22)	-0.50 (-0.38)	-0.04 (-0.04)	0.11 (0.07)	0.30 (0.21)	1.37 (0.62)
afix vs ain	ΔRightSkewness	109	-0.14 (-0.08)	0.60 (0.43)	-1.70 (-1.25)	-0.50 (-0.26)	-0.08 (-0.01)	0.12 (0.17)	1.87 (0.55)

Continued on next page

Table A.6 – *Continued from previous page*

		count	mean	std	min	25%	50%	75%	max
ain	ΔInnerWidth	109	25.32 (32.40)	52.31 (51.63)	-150.00 (-140.40)	0.00 (7.70)	30.00 (23.80)	60.00 (60.30)	150.00 (175.40)
ain	ΔKurtosis	109	-0.00 (-0.13)	0.82 (0.81)	-2.21 (-3.71)	-0.38 (-0.29)	-0.07 (-0.01)	0.50 (0.24)	3.93 (1.66)
ain	ΔMaximum	109	10.40 (9.91)	13.66 (14.10)	-52.11 (-53.34)	1.67 (1.68)	9.13 (8.46)	16.87 (17.92)	44.07 (39.76)
ain	normalizedΔMaximum	85	0.29 (0.27)	0.42 (0.41)	-0.86 (-0.85)	0.04 (0.02)	0.26 (0.26)	0.47 (0.48)	1.57 (1.58)
ain	ΔOuterWidth	109	12.11 (19.35)	67.34 (59.14)	-150.00 (-146.40)	-30.00 (-18.60)	0.00 (21.30)	60.00 (65.10)	150.00 (140.40)
ain	ΔPeakToPeak	109	10.40 (9.91)	13.66 (14.10)	-52.11 (-53.34)	1.67 (1.68)	9.13 (8.46)	16.87 (17.92)	44.07 (39.76)
ain	normalizedΔPeakToPeak	85	0.29 (0.27)	0.42 (0.41)	-0.86 (-0.85)	0.04 (0.02)	0.26 (0.26)	0.47 (0.48)	1.57 (1.58)
ain	ΔSkewness	109	-0.04 (-0.21)	0.87 (0.92)	-1.72 (-1.90)	-0.63 (-0.89)	-0.13 (-0.36)	0.46 (0.25)	2.46 (2.89)
ain	ΔWidth	109	1.10 (31.94)	93.09 (91.09)	-180.00 (-179.00)	-60.00 (-28.80)	0.00 (49.40)	60.00 (91.80)	120.00 (178.40)
ain	Dip	109	13.23 (14.43)	8.66 (10.16)	1.10 (0.00)	7.41 (7.30)	11.20 (12.15)	16.26 (19.51)	59.83 (66.73)
ain	normalizedDip	85	0.39 (0.42)	0.23 (0.26)	0.04 (0.00)	0.23 (0.25)	0.35 (0.38)	0.47 (0.52)	1.37 (1.29)
ain	GlobalMaximum	109	47.58 (47.18)	34.15 (33.78)	2.00 (2.00)	24.36 (23.60)	38.50 (37.87)	59.20 (62.41)	198.57 (196.01)
ain	GlobalMaximumAngle	109	223.21 (221.85)	58.66 (54.92)	60.00 (62.10)	210.00 (217.20)	240.00 (236.90)	270.00 (254.00)	330.00 (299.70)
ain	normalizedGlobalMaximum	85	1.14 (1.16)	0.32 (0.30)	0.28 (0.32)	0.91 (0.95)	1.09 (1.15)	1.29 (1.30)	2.00 (2.08)
ain	GlobalMinimum	109	15.47 (15.47)	20.27 (20.41)	0.00 (-0.37)	4.00 (4.27)	7.71 (8.09)	17.43 (16.54)	123.25 (125.63)
ain	GlobalMinimumAngle	109	18.99 (24.03)	50.77 (52.10)	-60.00 (-102.50)	0.00 (-15.00)	0.00 (18.10)	60.00 (49.10)	210.00 (202.50)
ain	normalizedGlobalMinimum	85	0.26 (0.27)	0.20 (0.21)	0.00 (-0.02)	0.11 (0.11)	0.23 (0.23)	0.37 (0.36)	0.79 (0.81)
ain	innerMinimumAngle	109	171.19 (164.07)	27.78 (36.01)	120.00 (69.40)	150.00 (139.70)	150.00 (165.50)	180.00 (187.30)	240.00 (291.40)
ain	innerMinimumVal	109	27.68 (26.18)	26.15 (25.53)	0.00 (-0.18)	9.00 (7.44)	18.80 (17.35)	38.33 (36.55)	154.50 (153.15)
ain	normalizedinnerMinimumVal	85	0.56 (0.55)	0.30 (0.34)	0.00 (-0.01)	0.35 (0.28)	0.57 (0.51)	0.76 (0.78)	1.67 (1.63)
ain	$\text{Bandwidth}^{\text{left}}_{75\%}$	109	74.59 (57.65)	25.70 (39.98)	30.00 (8.00)	60.00 (36.40)	60.00 (50.50)	90.00 (69.10)	150.00 (303.30)
ain	$\Delta\text{Width}^{\text{left}}$	109	66.61 (63.67)	61.83 (53.81)	-90.00 (-45.30)	30.00 (22.40)	60.00 (60.30)	90.00 (91.80)	210.00 (254.10)
ain	Dip^{left}	109	8.03 (9.47)	8.80 (10.07)	0.00 (0.00)	2.67 (2.62)	5.96 (6.21)	11.09 (13.96)	53.45 (58.49)

Continued on next page

Table A.6 – *Continued from previous page*

		count	mean	std	min	25%	50%	75%	max
ain	NormalizedDip^{left}	85	0.24 (0.28)	0.24 (0.27)	0.00 (0.00)	0.08 (0.08)	0.19 (0.23)	0.31 (0.39)	1.40 (1.33)
ain	InnerWidth^{left}	109	48.17 (46.88)	31.89 (25.98)	0.00 (0.00)	30.00 (35.50)	30.00 (46.60)	60.00 (58.70)	180.00 (140.40)
ain	InnerBandwidth$^{left}_{75\,\%}$	109	33.30 (23.75)	17.95 (15.30)	0.00 (0.00)	30.00 (14.80)	30.00 (23.40)	30.00 (30.70)	90.00 (94.60)
ain	Kurtosis^{left}	109	-1.07 (-0.98)	0.58 (0.75)	-2.00 (-1.88)	-1.47 (-1.43)	-1.25 (-1.18)	-0.81 (-0.88)	0.83 (2.66)
ain	Maximum^{left}	109	35.71 (35.65)	30.17 (30.19)	2.00 (1.05)	15.75 (14.98)	27.11 (27.89)	44.00 (42.97)	154.50 (156.25)
ain	MaximumAngle^{left}	109	123.03 (117.20)	32.39 (36.84)	60.00 (45.00)	120.00 (97.20)	120.00 (120.10)	150.00 (137.40)	210.00 (281.00)
ain	NormalizedMaximum^{left}	85	0.80 (0.83)	0.35 (0.38)	0.15 (0.05)	0.56 (0.56)	0.79 (0.84)	0.99 (1.11)	1.87 (1.79)
ain	OuterWidth^{left}	109	114.77 (110.55)	45.62 (47.74)	30.00 (30.30)	90.00 (81.20)	120.00 (103.40)	150.00 (135.00)	300.00 (329.20)
ain	OuterBandwidth$^{left}_{75\,\%}$	109	41.28 (33.90)	17.70 (37.14)	30.00 (8.00)	30.00 (20.90)	30.00 (26.90)	60.00 (34.50)	120.00 (283.70)
ain	PeakToPeak^{left}	109	20.24 (20.18)	15.39 (15.76)	1.33 (1.18)	9.33 (8.87)	17.14 (16.69)	26.58 (27.30)	71.78 (75.53)
ain	NormalizedPeakToPeak^{left}	85	0.54 (0.57)	0.30 (0.34)	0.05 (0.06)	0.34 (0.32)	0.50 (0.52)	0.64 (0.68)	1.56 (1.61)
ain	Width^{left}	109	162.94 (157.42)	48.71 (54.88)	60.00 (57.20)	120.00 (129.00)	150.00 (150.60)	180.00 (183.60)	390.00 (404.30)
ain	maxAngleDist	109	121.65 (126.16)	37.60 (37.70)	30.00 (20.80)	90.00 (102.30)	120.00 (120.00)	150.00 (151.00)	210.00 (238.10)
ain	minusLeftSkewness	109	-0.02 (0.14)	0.58 (0.62)	-1.59 (-1.83)	-0.31 (-0.26)	0.01 (0.20)	0.40 (0.60)	1.53 (1.81)
ain	outerMinimumAngle	109	97.43 (138.76)	129.11 (144.23)	0.00 (1.70)	0.00 (25.70)	30.00 (50.30)	90.00 (323.70)	330.00 (359.60)
ain	PeakToPeak	109	32.11 (31.71)	18.90 (18.85)	2.00 (2.00)	17.71 (16.66)	29.33 (28.10)	40.80 (40.01)	86.29 (88.97)
ain	normalizedPeakToPeak	85	0.88 (0.89)	0.35 (0.34)	0.26 (0.30)	0.61 (0.63)	0.80 (0.82)	1.05 (1.07)	1.94 (1.96)
ain	Bandwidth$^{right}_{75\,\%}$	109	94.95 (72.73)	28.44 (27.88)	60.00 (18.20)	60.00 (54.80)	90.00 (70.20)	120.00 (86.40)	180.00 (169.50)
ain	ΔWidth^{right}	109	53.39 (50.62)	51.54 (49.85)	-90.00 (-120.10)	30.00 (16.90)	60.00 (58.10)	90.00 (85.90)	150.00 (147.40)
ain	Dip^{right}	109	18.43 (19.38)	12.89 (14.31)	0.00 (0.00)	9.56 (9.19)	15.45 (15.64)	26.07 (26.99)	66.20 (74.96)
ain	NormalizedDip^{right}	85	0.53 (0.56)	0.37 (0.38)	0.00 (0.00)	0.27 (0.27)	0.45 (0.51)	0.70 (0.77)	1.65 (1.81)
ain	InnerWidth^{right}	109	73.49 (79.28)	32.53 (36.99)	0.00 (0.00)	60.00 (54.30)	90.00 (72.70)	90.00 (101.60)	150.00 (194.10)
ain	InnerBandwidth$^{right}_{75\,\%}$	109	46.24 (33.98)	22.56 (18.43)	0.00 (0.00)	30.00 (24.20)	30.00 (30.40)	60.00 (38.90)	120.00 (125.40)

Continued on next page

Table A.6 – *Continued from previous page*

		count	mean	std	min	25%	50%	75%	max
ain	Kurtosis^{right}	109	-1.07 (-1.10)	0.58 (0.45)	-1.86 (-1.93)	-1.42 (-1.36)	-1.15 (-1.21)	-0.90 (-0.92)	2.49 (0.51)
ain	Maximum^{right}	109	46.10 (45.56)	34.12 (33.60)	1.00 (0.87)	22.00 (22.59)	36.29 (35.76)	58.80 (57.95)	198.57 (196.01)
ain	MaximumAngle^{right}	109	244.68 (243.35)	26.51 (26.44)	180.00 (179.80)	240.00 (229.50)	240.00 (242.50)	270.00 (255.40)	300.00 (301.80)
ain	normalizedMaximum^{right}	85	1.09 (1.11)	0.33 (0.32)	0.20 (0.17)	0.88 (0.93)	1.04 (1.07)	1.25 (1.29)	1.93 (2.08)
ain	OuterWidth^{right}	109	126.88 (129.90)	35.97 (30.82)	60.00 (63.40)	90.00 (108.70)	120.00 (133.40)	150.00 (148.70)	240.00 (228.80)
ain	OuterBandwidth$^{right}_{75\,\%}$	109	48.72 (38.75)	23.38 (20.91)	30.00 (9.00)	30.00 (26.50)	30.00 (32.80)	60.00 (44.20)	150.00 (128.10)
ain	PeakToPeak^{right}	109	30.64 (30.09)	18.67 (18.62)	1.00 (0.87)	15.92 (15.50)	26.60 (25.40)	40.14 (39.56)	86.29 (88.97)
ain	normalizedPeakToPeak^{right}	85	0.83 (0.84)	0.35 (0.36)	0.20 (0.17)	0.58 (0.60)	0.76 (0.81)	1.04 (1.03)	1.80 (1.96)
ain	Skewness^{right}	109	-0.06 (-0.07)	0.54 (0.53)	-1.17 (-1.36)	-0.35 (-0.40)	-0.04 (-0.12)	0.25 (0.25)	2.03 (1.42)
ain	Width^{right}	109	200.37 (209.18)	45.25 (46.40)	90.00 (90.50)	180.00 (179.00)	210.00 (210.50)	240.00 (239.70)	330.00 (327.20)
ain	tcSymmetryIndex	109	0.44 (7.58)	0.15 (2.86)	0.13 (2.13)	0.31 (5.19)	0.45 (8.12)	0.55 (9.63)	0.77 (14.38)
uni	CircularVariance	85	0.47 (0.46)	0.13 (0.15)	0.12 (0.06)	0.39 (0.40)	0.46 (0.47)	0.55 (0.56)	0.74 (0.78)
uni	GlobalMaximum	85	37.34 (36.73)	23.80 (24.07)	5.00 (6.00)	20.00 (19.69)	31.00 (30.62)	47.60 (45.89)	125.60 (127.83)
uni	GlobalMaximumAngle	85	240.00 (242.81)	0.00 (16.64)	240.00 (212.90)	240.00 (232.00)	240.00 (241.20)	240.00 (251.40)	240.00 (296.80)
uni	normalizedGlobalMaximum	85	1.00 (1.00)	0.00 (0.00)	1.00 (1.00)	1.00 (1.00)	1.00 (1.00)	1.00 (1.00)	1.00 (1.00)
uni	GlobalMinimum	85	7.36 (7.25)	9.09 (9.33)	0.00 (-0.89)	1.33 (1.09)	4.50 (4.14)	9.33 (8.54)	49.60 (47.06)
uni	GlobalMinimumAngle	85	64.59 (63.41)	51.86 (52.48)	-60.00 (-70.70)	30.00 (37.40)	60.00 (67.30)	90.00 (102.70)	150.00 (156.30)
uni	normalizedGlobalMinimum	85	0.17 (0.17)	0.13 (0.15)	0.00 (-0.09)	0.06 (0.04)	0.14 (0.13)	0.24 (0.26)	0.58 (0.62)
uni	PeakToPeak	85	29.98 (29.48)	17.72 (18.22)	5.00 (6.56)	17.50 (16.76)	25.50 (24.52)	38.80 (35.27)	95.83 (93.39)
uni	normalizedPeakToPeak	85	0.83 (0.83)	0.13 (0.15)	0.42 (0.38)	0.76 (0.74)	0.86 (0.87)	0.94 (0.96)	1.00 (1.09)
uni	Bandwidth$^{right}_{75\,\%}$	85	94.24 (69.70)	32.45 (27.70)	60.00 (17.30)	60.00 (52.10)	90.00 (61.60)	120.00 (81.90)	210.00 (160.80)
uni	ΔWidth^{right}	85	9.18 (1.20)	103.72 (96.44)	-240.00 (-213.80)	-60.00 (-63.20)	0.00 (0.00)	60.00 (75.20)	180.00 (198.40)
uni	InnerWidth^{right}	85	175.41 (179.40)	51.86 (48.22)	90.00 (80.80)	150.00 (142.40)	180.00 (180.00)	210.00 (211.60)	300.00 (286.90)

Continued on next page

Table A.6 – *Continued from previous page*

		count	mean	std	min	25%	50%	75%	max
uni	InnerBandwidth$^{right}_{75\,\%}$	85	44.12 (34.48)	19.96 (15.17)	30.00 (8.60)	30.00 (25.40)	30.00 (30.20)	60.00 (37.50)	120.00 (94.80)
uni	Kurtosis^{right}	85	-0.52 (-0.48)	1.08 (1.42)	-1.79 (-1.92)	-1.21 (-1.16)	-0.74 (-0.75)	-0.09 (-0.32)	5.34 (6.66)
uni	Maximum^{right}	85	37.34 (36.73)	23.80 (24.07)	5.00 (6.00)	20.00 (19.69)	31.00 (30.62)	47.60 (45.89)	125.60 (127.83)
uni	MaximumAngle^{right}	85	240.00 (242.81)	0.00 (16.64)	240.00 (212.90)	240.00 (232.00)	240.00 (241.20)	240.00 (251.40)	240.00 (296.80)
uni	normalizedMaximum^{right}	85	1.00 (1.00)	0.00 (0.00)	1.00 (1.00)	1.00 (1.00)	1.00 (1.00)	1.00 (1.00)	1.00 (1.00)
uni	OuterWidth^{right}	85	184.59 (180.60)	51.86 (48.22)	60.00 (73.10)	150.00 (148.40)	180.00 (180.00)	210.00 (217.60)	270.00 (279.20)
uni	OuterBandwidth$^{right}_{75\,\%}$	85	50.12 (35.21)	25.52 (16.55)	30.00 (8.70)	30.00 (25.70)	30.00 (31.00)	60.00 (39.10)	150.00 (98.50)
uni	PeakToPeak^{right}	85	29.98 (29.48)	17.72 (18.22)	5.00 (6.56)	17.50 (16.76)	25.50 (24.52)	38.80 (35.27)	95.83 (93.39)
uni	normalizedPeakToPeak^{right}	85	0.83 (0.83)	0.13 (0.15)	0.42 (0.38)	0.76 (0.74)	0.86 (0.87)	0.94 (0.96)	1.00 (1.09)
uni	Skewness^{right}	85	0.69 (0.71)	0.53 (0.58)	-0.41 (-0.37)	0.31 (0.33)	0.63 (0.64)	1.04 (1.00)	2.62 (2.80)
uni	Width^{right}	85	360.00 (360.00)	0.00 (0.00)	360.00 (360.00)	360.00 (360.00)	360.00 (360.00)	360.00 (360.00)	360.00 (360.00)
uni vs afix	ΔGlobalMinimum	85	3.30 (1.69)	7.93 (2.76)	-6.80 (-2.26)	-0.17 (-0.23)	1.60 (0.98)	4.17 (3.03)	58.00 (9.68)
uni vs afix	normalizedΔGlobalMinimum	85	0.06 (0.05)	0.10 (0.09)	-0.16 (-0.13)	-0.01 (-0.02)	0.05 (0.03)	0.11 (0.13)	0.46 (0.23)
uni vs afix	ΔRightMaximum	85	-2.40 (0.35)	7.22 (4.23)	-30.40 (-8.97)	-5.33 (-2.09)	-2.00 (-0.24)	0.67 (2.22)	15.90 (8.60)
uni vs afix	normalizedΔRightMaximum	85	-0.09 (-0.00)	0.22 (0.15)	-0.79 (-0.34)	-0.20 (-0.07)	-0.06 (-0.01)	0.04 (0.09)	0.44 (0.37)
uni vs ain	ΔGlobalMinimum	85	4.23 (3.62)	8.18 (5.52)	-6.27 (-5.26)	0.00 (0.36)	2.13 (2.88)	4.76 (4.91)	55.75 (16.89)
uni vs ain	normalizedΔGlobalMinimum	85	0.09 (0.09)	0.12 (0.13)	-0.14 (-0.19)	0.00 (0.01)	0.06 (0.07)	0.16 (0.15)	0.44 (0.33)
uni vs ain	ΔRightMaximum	85	2.95 (3.41)	11.55 (12.80)	-29.00 (-33.73)	-4.00 (-2.17)	1.25 (1.87)	7.56 (7.67)	44.40 (42.74)
uni vs ain	normalizedΔRightMaximum	85	0.09 (0.13)	0.33 (0.27)	-0.80 (-0.54)	-0.12 (-0.05)	0.04 (0.13)	0.25 (0.29)	0.93 (0.84)

Table A.7: Transparent paradigm's statistics for all features. Table lists cell count, mean, standard deviation, minimum, 25 % quantile, median, 75 % quantle and maximum for all features when evaluated with the direct method (values from best model in parentheses).

		count	mean	std	min	25%	50%	75%	max
afix	ΔInnerWidth	143	0.63 (4.89)	53.34 (56.36)	-120.00 (-175.60)	-30.00 (-22.05)	0.00 (7.00)	30.00 (34.70)	150.00 (154.90)
afix	ΔKurtosis	144	-0.12 (0.07)	1.11 (1.15)	-2.85 (-4.26)	-0.77 (-0.42)	-0.09 (-0.01)	0.52 (0.45)	3.00 (4.90)
afix	ΔMaximum	143	0.97 (1.58)	7.53 (8.43)	-21.00 (-25.79)	-2.50 (-1.77)	0.00 (0.11)	3.33 (3.77)	43.00 (53.44)
afix	normalizedΔMaximum	143	0.02 (-0.16)	0.57 (2.29)	-2.00 (-26.40)	-0.19 (-0.13)	0.00 (0.02)	0.18 (0.22)	4.18 (2.82)
afix	ΔOuterWidth	143	-0.63 (7.33)	89.96 (81.96)	-180.00 (-165.60)	-60.00 (-51.80)	0.00 (2.10)	60.00 (72.80)	150.00 (178.50)
afix	ΔPeakToPeak	143	0.97 (1.58)	7.53 (8.43)	-21.00 (-25.79)	-2.50 (-1.77)	0.00 (0.11)	3.33 (3.77)	43.00 (53.44)
afix	normalizedΔPeakToPeak	143	0.02 (-0.16)	0.57 (2.29)	-2.00 (-26.40)	-0.19 (-0.13)	0.00 (0.02)	0.18 (0.22)	4.18 (2.82)
afix	ΔSkewness	144	0.51 (-0.09)	1.04 (1.20)	-2.22 (-2.99)	-0.15 (-0.79)	0.58 (-0.10)	1.19 (0.56)	3.58 (5.22)
afix	ΔWidth	143	-2.52 (-0.37)	95.77 (105.53)	-180.00 (-180.00)	-60.00 (-87.40)	0.00 (1.60)	60.00 (82.50)	120.00 (178.60)
afix	Dip	143	8.12 (8.15)	6.39 (6.90)	0.38 (0.39)	3.50 (3.21)	6.50 (6.52)	10.83 (10.93)	36.00 (38.09)
afix	normalizedDip	143	0.56 (0.72)	0.54 (1.26)	0.04 (0.03)	0.25 (0.23)	0.43 (0.45)	0.73 (0.77)	4.50 (13.20)
afix	GlobalMaximum	145	21.15 (20.66)	21.92 (22.34)	0.50 (0.76)	7.50 (6.39)	14.00 (12.94)	27.71 (26.66)	127.00 (130.39)
afix	GlobalMaximumAngle	145	173.59 (184.15)	86.76 (81.34)	0.00 (4.90)	120.00 (119.70)	180.00 (184.70)	240.00 (248.20)	330.00 (349.90)
afix	normalizedGlobalMaximum	145	1.14 (1.69)	1.10 (5.96)	0.07 (0.17)	0.62 (0.63)	0.93 (0.89)	1.43 (1.37)	11.57 (70.99)
afix	GlobalMinimum	145	5.01 (5.08)	8.56 (8.99)	0.00 (-1.38)	0.00 (0.27)	2.00 (1.53)	5.00 (5.50)	48.00 (51.41)
afix	GlobalMinimumAngle	145	38.90 (42.14)	78.11 (88.81)	-120.00 (-116.70)	0.00 (-15.75)	30.00 (26.40)	90.00 (89.35)	210.00 (238.50)
afix	normalizedGlobalMinimum	145	0.17 (0.40)	0.22 (2.57)	0.00 (-1.37)	0.00 (0.03)	0.12 (0.12)	0.28 (0.27)	2.00 (30.49)
afix	innerMinimumAngle	145	177.72 (179.87)	31.24 (34.74)	120.00 (108.20)	150.00 (154.40)	180.00 (180.60)	210.00 (203.50)	240.00 (269.50)
afix	innerMinimumVal	145	10.48 (9.94)	16.66 (16.20)	0.00 (-0.45)	1.33 (1.10)	4.00 (4.12)	12.00 (11.07)	99.43 (96.17)
afix	normalizedinnerMinimumVal	145	0.37 (0.69)	0.36 (3.73)	0.00 (-0.24)	0.15 (0.12)	0.32 (0.32)	0.51 (0.56)	3.27 (44.59)
afix	$\text{Bandwidth}^{\text{left}}_{75\,\%}$	144	76.88 (57.43)	25.51 (31.29)	30.00 (12.80)	60.00 (37.00)	60.00 (47.90)	90.00 (69.15)	180.00 (217.90)

Continued on next page

Table A.7 – *Continued from previous page*

		count	mean	std	min	25%	50%	75%	max
afix	ΔWidth^{left}	144	59.58 (51.84)	79.09 (64.23)	-120.00 (-153.50)	0.00 (6.75)	60.00 (52.70)	120.00 (97.55)	240.00 (187.00)
afix	Dip^{left}	144	7.59 (7.36)	6.38 (6.63)	0.00 (0.00)	2.99 (2.44)	5.83 (5.10)	10.56 (10.53)	34.00 (35.40)
afix	normalizedDip^{left}	144	0.55 (0.80)	0.53 (2.31)	0.00 (0.00)	0.19 (0.16)	0.41 (0.43)	0.70 (0.73)	3.22 (26.40)
afix	InnerWidth^{left}	144	63.33 (66.69)	36.72 (36.63)	0.00 (0.00)	30.00 (45.70)	60.00 (57.10)	90.00 (87.30)	150.00 (212.10)
afix	InnerBandwidth$^{left}_{75\%}$	144	36.04 (25.20)	16.09 (14.00)	0.00 (0.00)	30.00 (16.80)	30.00 (22.40)	30.00 (31.70)	90.00 (89.10)
afix	Kurtosis^{left}	144	-0.79 (-0.82)	0.80 (0.94)	-3.00 (-2.00)	-1.34 (-1.34)	-0.95 (-1.00)	-0.47 (-0.66)	2.17 (6.26)
afix	Maximum^{left}	144	18.14 (17.30)	18.85 (18.78)	0.25 (0.43)	5.88 (4.98)	11.00 (10.37)	25.08 (23.79)	102.00 (108.84)
afix	MaximumAngle^{left}	144	114.58 (113.18)	36.29 (37.60)	60.00 (45.00)	90.00 (84.10)	120.00 (113.70)	150.00 (145.05)	210.00 (190.70)
afix	normalizedMaximum^{left}	144	0.92 (1.49)	0.67 (5.97)	0.12 (0.11)	0.49 (0.52)	0.76 (0.75)	1.04 (1.06)	4.67 (70.99)
afix	OuterWidth^{left}	144	122.92 (118.53)	65.83 (57.63)	30.00 (39.40)	90.00 (70.05)	120.00 (114.40)	150.00 (150.85)	330.00 (313.20)
afix	OuterBandwidth$^{left}_{75\%}$	144	40.83 (32.24)	21.17 (25.55)	30.00 (0.10)	30.00 (18.65)	30.00 (23.60)	60.00 (36.60)	150.00 (187.30)
afix	PeakToPeak^{left}	144	13.09 (12.22)	12.19 (11.88)	0.25 (0.00)	4.60 (4.24)	9.08 (9.12)	18.18 (16.64)	76.00 (75.32)
afix	normalizedPeakToPeak^{left}	144	0.74 (1.09)	0.59 (3.45)	0.12 (0.00)	0.36 (0.35)	0.51 (0.53)	0.95 (0.85)	3.41 (40.50)
afix	Width^{left}	144	186.25 (185.22)	71.48 (72.11)	60.00 (53.90)	150.00 (136.10)	180.00 (179.20)	210.00 (224.50)	480.00 (469.10)
afix	maxAngleDist	143	127.76 (138.27)	46.38 (46.73)	60.00 (61.50)	90.00 (100.35)	120.00 (129.40)	150.00 (170.75)	240.00 (270.00)
afix	minusLeftSkewness	144	-0.26 (0.06)	0.70 (0.70)	-2.04 (-2.73)	-0.69 (-0.36)	-0.28 (0.04)	0.19 (0.59)	1.49 (1.90)
afix	outerMinimumAngle	145	108.41 (168.36)	124.41 (140.15)	0.00 (1.90)	0.00 (33.25)	30.00 (92.10)	240.00 (318.15)	330.00 (359.40)
afix	PeakToPeak	145	16.14 (15.58)	14.96 (15.06)	0.50 (0.77)	6.00 (5.66)	11.93 (10.43)	20.00 (20.63)	86.50 (87.22)
afix	normalizedPeakToPeak	145	0.97 (1.29)	1.07 (3.45)	0.07 (0.17)	0.44 (0.48)	0.74 (0.72)	1.14 (1.11)	11.57 (40.50)
afix	Bandwidth$^{right}_{75\%}$	144	79.17 (59.65)	28.07 (33.27)	30.00 (0.40)	60.00 (37.90)	60.00 (50.00)	90.00 (76.45)	180.00 (195.20)
afix	ΔWidth^{right}	144	68.54 (46.73)	72.16 (64.02)	-90.00 (-145.70)	22.50 (1.20)	60.00 (45.70)	120.00 (89.25)	240.00 (223.00)
afix	Dip^{right}	143	8.60 (8.94)	8.34 (9.31)	0.00 (0.00)	3.00 (2.92)	6.07 (6.59)	11.13 (10.78)	46.67 (55.35)
afix	normalizedDip^{right}	143	0.57 (0.64)	0.69 (0.69)	0.00 (0.00)	0.22 (0.21)	0.41 (0.47)	0.72 (0.82)	6.59 (4.80)

Continued on next page

Table A.7 – *Continued from previous page*

		count	mean	std	min	25%	50%	75%	max
afix	InnerWidth^{right}	144	63.96 (71.58)	33.89 (36.58)	0.00 (0.00)	30.00 (49.00)	60.00 (61.60)	90.00 (86.45)	180.00 (204.00)
afix	InnerBandwidth$^{right}_{75\,\%}$	144	35.83 (27.69)	14.75 (19.46)	0.00 (0.00)	30.00 (16.85)	30.00 (23.00)	30.00 (33.25)	90.00 (144.60)
afix	Kurtosis^{right}	145	-0.89 (-0.74)	0.74 (1.00)	-2.00 (-1.72)	-1.44 (-1.34)	-1.09 (-0.98)	-0.62 (-0.64)	1.67 (5.64)
afix	Maximum^{right}	144	19.10 (18.88)	21.80 (22.16)	0.50 (0.03)	6.00 (5.11)	12.00 (11.57)	26.00 (24.68)	127.00 (130.39)
afix	MaximumAngle^{right}	144	241.67 (251.45)	33.62 (35.63)	150.00 (168.50)	210.00 (229.30)	240.00 (247.50)	270.00 (272.55)	300.00 (315.00)
afix	normalizedMaximum^{right}	144	0.94 (1.33)	0.76 (3.78)	0.07 (0.01)	0.53 (0.58)	0.78 (0.76)	1.12 (1.22)	7.07 (44.59)
afix	OuterWidth^{right}	144	132.50 (118.31)	59.31 (53.97)	30.00 (12.80)	90.00 (69.90)	120.00 (118.00)	180.00 (148.50)	330.00 (313.50)
afix	OuterBandwidth$^{right}_{75\,\%}$	144	43.33 (31.96)	22.93 (23.33)	30.00 (0.40)	30.00 (18.95)	30.00 (25.60)	60.00 (35.80)	150.00 (150.00)
afix	PeakToPeak^{right}	144	14.05 (13.79)	14.83 (15.05)	0.50 (0.00)	4.69 (4.18)	9.50 (8.70)	17.62 (16.57)	86.50 (87.22)
afix	normalizedPeakToPeak^{right}	144	0.76 (0.93)	0.73 (1.34)	0.07 (0.00)	0.41 (0.40)	0.55 (0.62)	0.96 (1.05)	7.07 (14.11)
afix	Skewness^{right}	145	0.26 (-0.03)	0.64 (0.74)	-1.21 (-2.14)	-0.21 (-0.55)	0.24 (-0.05)	0.66 (0.44)	1.79 (2.49)
afix	Width^{right}	144	196.46 (189.89)	64.24 (66.37)	60.00 (12.80)	150.00 (144.80)	195.00 (187.80)	240.00 (237.20)	450.00 (451.80)
afix	tcSymmetryIndex	145	0.40 (6.96)	0.13 (2.41)	0.00 (0.38)	0.31 (5.29)	0.41 (6.71)	0.49 (8.73)	0.72 (12.73)
afix vs ain	ΔGlobalMinimum	145	1.07 (1.78)	6.41 (4.10)	-30.00 (-2.86)	-0.90 (-0.34)	0.13 (0.38)	1.53 (2.28)	48.57 (15.96)
afix vs ain	normalizedΔGlobalMinimum	145	0.03 (0.08)	0.24 (0.14)	-2.00 (-0.10)	-0.04 (-0.02)	0.01 (0.05)	0.13 (0.13)	0.77 (0.52)
afix vs ain	ΔLeftKurtosis	143	-0.10 (0.13)	1.07 (0.93)	-3.71 (-3.97)	-0.76 (-0.26)	-0.05 (0.04)	0.53 (0.60)	2.52 (1.63)
afix vs ain	ΔLeftMaximum	143	4.35 (8.18)	14.29 (16.93)	-63.00 (-10.86)	-1.08 (-0.31)	1.70 (3.11)	6.34 (9.36)	88.50 (76.93)
afix vs ain	normalizedΔLeftMaximum	143	0.08 (0.17)	0.60 (0.46)	-4.20 (-1.27)	-0.12 (-0.04)	0.13 (0.20)	0.31 (0.36)	2.40 (1.27)
afix vs ain	ΔMinimumusLeftSkewness	143	0.06 (0.06)	0.84 (0.69)	-1.97 (-1.65)	-0.53 (-0.27)	0.14 (0.05)	0.60 (0.36)	2.21 (2.51)
afix vs ain	ΔRightKurtosis	145	0.02 (0.14)	1.03 (1.04)	-2.65 (-1.75)	-0.49 (-0.61)	0.01 (0.16)	0.54 (0.71)	3.30 (3.46)
afix vs ain	ΔRightMaximum	144	4.65 (7.10)	13.12 (15.52)	-38.00 (-6.69)	-1.08 (0.28)	1.65 (3.22)	7.86 (7.23)	78.83 (87.33)
afix vs ain	normalizedΔRightMaximum	144	0.14 (0.18)	0.81 (0.47)	-5.38 (-1.06)	-0.11 (0.04)	0.13 (0.14)	0.44 (0.37)	4.17 (2.05)
afix vs ain	ΔRightSkewness	145	-0.03 (-0.01)	0.82 (0.84)	-1.75 (-1.93)	-0.59 (-0.51)	-0.07 (-0.12)	0.54 (0.38)	2.14 (1.81)

Continued on next page

Table A.7 – *Continued from previous page*

		count	mean	std	min	25%	50%	75%	max
ain	ΔInnerWidth	145	9.93 (-1.12)	53.74 (50.55)	-150.00 (-135.40)	-30.00 (-26.90)	0.00 (3.45)	30.00 (23.32)	150.00 (136.40)
ain	ΔKurtosis	145	0.01 (0.03)	1.06 (1.01)	-3.52 (-3.32)	-0.54 (-0.46)	-0.10 (-0.03)	0.53 (0.44)	4.21 (4.16)
ain	ΔMaximum	145	1.29 (1.18)	12.84 (12.98)	-67.20 (-39.83)	-2.45 (-3.38)	0.50 (0.40)	4.54 (3.38)	91.50 (104.98)
ain	normalizedΔMaximum	145	0.08 (0.14)	0.66 (0.99)	-3.40 (-4.07)	-0.20 (-0.19)	0.05 (0.05)	0.29 (0.27)	3.45 (6.52)
ain	ΔOuterWidth	145	-6.21 (7.12)	92.19 (81.85)	-180.00 (-170.10)	-60.00 (-50.72)	0.00 (4.65)	60.00 (63.00)	150.00 (178.20)
ain	ΔPeakToPeak	145	1.29 (1.18)	12.84 (12.98)	-67.20 (-39.83)	-2.45 (-3.38)	0.50 (0.40)	4.54 (3.38)	91.50 (104.98)
ain	normalizedΔPeakToPeak	145	0.08 (0.14)	0.66 (0.99)	-3.40 (-4.07)	-0.20 (-0.19)	0.05 (0.05)	0.29 (0.27)	3.45 (6.52)
ain	ΔSkewness	145	0.43 (-0.03)	0.95 (1.08)	-2.27 (-2.51)	-0.27 (-0.74)	0.42 (0.00)	1.02 (0.58)	3.14 (3.12)
ain	ΔWidth	145	-8.69 (-3.86)	96.43 (97.31)	-180.00 (-178.60)	-60.00 (-79.30)	0.00 (1.10)	60.00 (55.90)	120.00 (177.40)
ain	Dip	145	9.95 (10.46)	10.17 (10.54)	0.30 (0.02)	4.07 (3.67)	6.90 (7.01)	11.55 (13.12)	64.40 (57.75)
ain	normalizedDip	145	0.59 (0.87)	0.46 (2.52)	0.06 (0.02)	0.27 (0.33)	0.48 (0.57)	0.73 (0.80)	3.18 (30.32)
ain	GlobalMaximum	146	26.52 (26.18)	29.44 (29.83)	1.60 (0.45)	8.00 (6.38)	15.75 (15.93)	33.69 (33.30)	150.57 (151.46)
ain	GlobalMaximumAngle	146	180.00 (189.61)	80.73 (79.79)	0.00 (1.10)	120.00 (120.70)	180.00 (206.35)	240.00 (250.30)	330.00 (356.00)
ain	normalizedGlobalMaximum	146	1.25 (1.53)	0.71 (2.66)	0.31 (0.22)	0.83 (0.80)	1.08 (1.14)	1.44 (1.53)	4.80 (31.50)
ain	GlobalMinimum	146	6.04 (5.84)	11.48 (11.72)	0.00 (-3.49)	0.41 (0.20)	1.93 (1.44)	6.09 (5.01)	64.50 (63.63)
ain	GlobalMinimumAngle	146	31.44 (38.31)	81.67 (83.13)	-120.00 (-120.00)	-30.00 (-15.45)	30.00 (21.10)	60.00 (72.78)	210.00 (237.80)
ain	normalizedGlobalMinimum	146	0.20 (0.19)	0.21 (0.28)	0.00 (-1.55)	0.02 (0.02)	0.14 (0.14)	0.32 (0.33)	0.99 (1.13)
ain	innerMinimumAngle	146	177.33 (184.08)	33.04 (36.33)	120.00 (99.90)	150.00 (156.90)	180.00 (188.35)	210.00 (207.80)	240.00 (246.00)
ain	innerMinimumVal	146	12.94 (12.11)	21.22 (21.03)	0.00 (-3.49)	1.52 (1.03)	5.00 (3.68)	16.30 (15.32)	138.00 (134.53)
ain	normalizedinnerMinimumVal	146	0.44 (0.40)	0.31 (0.37)	0.00 (-1.53)	0.22 (0.16)	0.41 (0.38)	0.62 (0.60)	1.55 (1.64)
ain	$\text{Bandwidth}^{\text{left}}_{75\%}$	145	74.07 (56.33)	21.81 (32.39)	30.00 (16.10)	60.00 (38.30)	60.00 (48.55)	90.00 (62.88)	150.00 (268.50)
ain	$\Delta\text{Width}^{\text{left}}$	145	66.83 (50.50)	67.06 (64.97)	-90.00 (-113.30)	30.00 (9.15)	60.00 (61.90)	120.00 (95.22)	240.00 (261.20)
ain	Dip^{left}	145	9.30 (9.88)	11.55 (10.99)	0.00 (0.00)	2.67 (3.22)	5.94 (6.02)	11.00 (12.77)	98.00 (62.14)

Continued on next page

Table A.7 – *Continued from previous page*

		count	mean	std	min	25%	50%	75%	max
ain	normalizedDip^{left}	145	0.55 (0.80)	0.50 (2.30)	0.00 (0.00)	0.21 (0.25)	0.42 (0.47)	0.75 (0.80)	3.40 (27.61)
ain	InnerWidth^{left}	145	57.52 (66.95)	32.88 (36.26)	0.00 (0.00)	30.00 (46.40)	60.00 (57.95)	90.00 (79.80)	180.00 (194.90)
ain	InnerBandwidth$^{left}_{75\,\%}$	145	36.00 (25.52)	14.83 (15.28)	0.00 (0.00)	30.00 (17.05)	30.00 (22.25)	30.00 (29.67)	90.00 (102.10)
ain	Kurtosis^{left}	145	-0.88 (-0.85)	0.68 (0.69)	-1.99 (-1.92)	-1.39 (-1.34)	-1.06 (-1.01)	-0.51 (-0.64)	1.95 (1.88)
ain	Maximum^{left}	145	22.33 (21.98)	26.84 (26.92)	0.50 (0.03)	6.40 (5.72)	12.22 (11.84)	28.00 (27.63)	141.25 (150.08)
ain	MaximumAngle^{left}	145	120.00 (117.13)	33.54 (36.59)	60.00 (45.00)	90.00 (93.88)	120.00 (117.15)	150.00 (143.05)	210.00 (190.40)
ain	normalizedMaximum^{left}	145	0.99 (1.20)	0.58 (2.16)	0.09 (0.01)	0.65 (0.66)	0.87 (0.91)	1.20 (1.25)	4.80 (26.07)
ain	OuterWidth^{left}	145	124.34 (117.45)	61.03 (52.19)	30.00 (21.90)	90.00 (70.85)	120.00 (116.40)	150.00 (148.25)	330.00 (338.60)
ain	OuterBandwidth$^{left}_{75\,\%}$	145	38.07 (30.81)	16.30 (27.36)	30.00 (8.10)	30.00 (19.02)	30.00 (24.05)	30.00 (33.60)	90.00 (265.70)
ain	PeakToPeak^{left}	145	16.25 (16.14)	19.11 (18.93)	0.50 (0.13)	4.90 (4.49)	10.00 (10.06)	21.10 (21.51)	135.00 (104.42)
ain	normalizedPeakToPeak^{left}	145	0.79 (1.01)	0.57 (2.30)	0.07 (0.05)	0.44 (0.44)	0.64 (0.69)	0.96 (1.03)	4.58 (27.62)
ain	Width^{left}	145	181.86 (184.40)	71.52 (62.10)	60.00 (60.40)	150.00 (144.53)	180.00 (180.60)	210.00 (220.02)	450.00 (416.00)
ain	maxAngleDist	145	124.97 (132.78)	42.28 (44.84)	30.00 (41.50)	90.00 (100.78)	120.00 (119.90)	150.00 (166.10)	240.00 (270.00)
ain	minusLeftSkewness	145	-0.19 (0.02)	0.63 (0.64)	-1.95 (-1.57)	-0.59 (-0.41)	-0.14 (0.04)	0.22 (0.50)	1.30 (1.57)
ain	outerMinimumAngle	146	128.84 (159.96)	134.26 (145.39)	0.00 (0.40)	30.00 (25.68)	60.00 (74.60)	300.00 (321.70)	330.00 (357.50)
ain	PeakToPeak	146	20.48 (20.34)	22.04 (22.60)	1.46 (0.33)	6.60 (5.29)	12.50 (12.58)	26.30 (26.61)	135.00 (119.88)
ain	normalizedPeakToPeak	146	1.05 (1.34)	0.72 (2.80)	0.25 (0.22)	0.62 (0.60)	0.85 (0.87)	1.30 (1.35)	4.58 (33.05)
ain	Bandwidth$^{right}_{75\,\%}$	146	81.78 (62.54)	33.34 (41.07)	30.00 (7.30)	60.00 (39.50)	60.00 (50.40)	90.00 (71.95)	270.00 (274.30)
ain	ΔWidth^{right}	146	62.88 (58.74)	71.63 (64.35)	-90.00 (-83.60)	0.00 (8.53)	60.00 (62.40)	120.00 (96.83)	240.00 (239.00)
ain	Dip^{right}	145	10.59 (11.05)	12.49 (13.63)	0.00 (0.00)	3.80 (3.32)	6.80 (6.56)	12.00 (13.65)	95.50 (104.98)
ain	normalizedDip^{right}	145	0.63 (0.94)	0.62 (2.80)	0.00 (0.00)	0.26 (0.28)	0.49 (0.53)	0.82 (0.85)	4.35 (33.04)
ain	InnerWidth^{right}	146	67.60 (65.83)	35.38 (31.11)	0.00 (0.00)	30.00 (48.65)	60.00 (61.55)	90.00 (75.90)	150.00 (141.60)
ain	InnerBandwidth$^{right}_{75\,\%}$	146	36.99 (25.71)	16.54 (15.61)	0.00 (0.00)	30.00 (17.65)	30.00 (22.10)	30.00 (31.48)	120.00 (85.20)

Continued on next page

Table A.7 – *Continued from previous page*

		count	mean	std	min	25%	50%	75%	max
ain	Kurtosis^{right}	146	-0.87 (-0.83)	0.76 (0.81)	-1.92 (-1.93)	-1.46 (-1.40)	-1.02 (-0.99)	-0.67 (-0.56)	2.71 (2.76)
ain	Maximum^{right}	146	23.48 (23.16)	27.41 (28.07)	0.75 (0.45)	6.71 (5.29)	13.50 (12.80)	30.00 (30.68)	150.57 (151.46)
ain	MaximumAngle^{right}	146	244.93 (249.91)	32.87 (33.51)	150.00 (160.30)	240.00 (228.45)	240.00 (247.10)	270.00 (271.52)	300.00 (315.00)
ain	normalizedMaximum^{right}	146	1.08 (1.34)	0.66 (2.66)	0.21 (0.21)	0.68 (0.65)	0.90 (0.96)	1.30 (1.36)	4.53 (31.50)
ain	OuterWidth^{right}	146	130.48 (124.57)	61.39 (56.50)	30.00 (42.30)	90.00 (79.45)	120.00 (125.40)	172.50 (150.97)	300.00 (333.60)
ain	OuterBandwidth$^{right}_{75\,\%}$	146	44.79 (36.83)	30.15 (37.52)	30.00 (7.30)	30.00 (19.70)	30.00 (24.95)	60.00 (40.25)	240.00 (267.50)
ain	PeakToPeak^{right}	146	17.44 (17.32)	19.47 (20.74)	0.33 (0.33)	4.95 (4.36)	10.00 (9.96)	21.18 (21.35)	108.90 (119.88)
ain	normalizedPeakToPeak^{right}	146	0.87 (1.15)	0.65 (2.80)	0.12 (0.03)	0.49 (0.48)	0.70 (0.72)	1.08 (1.07)	4.53 (33.05)
ain	Skewness^{right}	146	0.23 (-0.01)	0.67 (0.69)	-1.36 (-1.83)	-0.25 (-0.42)	0.26 (0.01)	0.68 (0.43)	2.02 (1.59)
ain	Width^{right}	146	198.08 (190.40)	70.07 (64.65)	60.00 (43.60)	150.00 (145.80)	180.00 (183.60)	240.00 (235.32)	450.00 (435.50)
ain	tcSymmetryIndex	146	0.39 (6.79)	0.12 (2.53)	0.11 (0.93)	0.31 (5.10)	0.40 (6.78)	0.48 (8.42)	0.72 (15.02)
uni	CircularVariance	146	0.59 (0.57)	0.21 (0.21)	0.13 (0.10)	0.42 (0.41)	0.62 (0.59)	0.74 (0.73)	0.99 (0.98)
uni	GlobalMaximum	146	24.40 (23.88)	27.07 (27.18)	1.04 (0.89)	6.95 (6.90)	14.19 (15.01)	30.35 (29.16)	157.22 (156.52)
uni	GlobalMaximumAngle	146	233.42 (241.27)	35.84 (48.46)	0.00 (42.80)	240.00 (230.75)	240.00 (244.10)	240.00 (255.23)	240.00 (358.80)
uni	normalizedGlobalMaximum	146	1.00 (1.00)	0.00 (0.00)	1.00 (1.00)	1.00 (1.00)	1.00 (1.00)	1.00 (1.00)	1.00 (1.00)
uni	GlobalMinimum	146	5.08 (5.07)	8.24 (8.50)	0.00 (-10.71)	0.56 (0.52)	1.96 (1.86)	5.42 (5.55)	52.93 (50.08)
uni	GlobalMinimumAngle	146	58.56 (53.81)	72.92 (69.54)	-60.00 (-80.30)	0.00 (0.00)	60.00 (52.10)	120.00 (98.80)	210.00 (239.80)
uni	normalizedGlobalMinimum	146	0.20 (0.21)	0.16 (0.19)	0.00 (-0.46)	0.07 (0.05)	0.16 (0.17)	0.30 (0.34)	0.69 (0.70)
uni	PeakToPeak	146	19.32 (18.80)	22.83 (23.20)	0.74 (0.66)	5.35 (5.84)	11.50 (11.16)	25.37 (23.71)	150.56 (150.27)
uni	normalizedPeakToPeak	146	0.80 (0.79)	0.16 (0.19)	0.31 (0.30)	0.70 (0.66)	0.84 (0.83)	0.93 (0.95)	1.00 (1.46)
uni	Bandwidth$^{right}_{75\,\%}$	146	84.66 (68.48)	29.41 (36.56)	60.00 (18.60)	60.00 (45.03)	90.00 (60.90)	90.00 (80.15)	210.00 (199.40)
uni	ΔWidth^{right}	146	-2.88 (-21.02)	145.84 (134.70)	-240.00 (-262.00)	-120.00 (-100.95)	0.00 (-25.20)	120.00 (57.45)	300.00 (264.00)
uni	InnerWidth^{right}	146	181.44 (190.51)	72.92 (67.35)	30.00 (48.00)	120.00 (151.27)	180.00 (192.60)	240.00 (230.47)	300.00 (311.00)

Continued on next page

Table A.7 – *Continued from previous page*

		count	mean	std	min	25%	50%	75%	max
uni	InnerBandwidth$^{right}_{75\,\%}$	146	40.68 (33.93)	17.17 (21.25)	30.00 (9.30)	30.00 (22.53)	30.00 (29.50)	60.00 (37.58)	120.00 (159.60)
uni	Kurtosis^{right}	146	-0.55 (-0.21)	0.96 (1.99)	-1.69 (-1.93)	-1.19 (-1.19)	-0.83 (-0.85)	-0.24 (-0.18)	4.06 (6.66)
uni	Maximum^{right}	146	24.39 (23.77)	27.07 (27.23)	1.04 (0.83)	6.95 (6.87)	14.19 (15.01)	30.35 (29.16)	157.22 (156.52)
uni	MaximumAngle^{right}	146	240.00 (244.32)	0.00 (21.77)	240.00 (180.00)	240.00 (232.03)	240.00 (244.35)	240.00 (254.88)	240.00 (300.00)
uni	normalizedMaximum^{right}	146	1.00 (0.98)	0.00 (0.07)	0.96 (0.41)	1.00 (1.00)	1.00 (1.00)	1.00 (1.00)	1.00 (1.00)
uni	OuterWidth^{right}	146	178.56 (169.49)	72.92 (67.35)	60.00 (49.00)	120.00 (129.53)	180.00 (167.40)	240.00 (208.73)	330.00 (312.00)
uni	OuterBandwidth$^{right}_{75\,\%}$	146	43.97 (34.54)	21.77 (23.09)	30.00 (6.30)	30.00 (22.07)	30.00 (28.45)	60.00 (38.03)	150.00 (157.00)
uni	PeakToPeak^{right}	146	19.31 (18.70)	22.83 (23.23)	0.74 (0.25)	5.35 (5.60)	11.50 (11.09)	25.37 (23.71)	150.56 (150.27)
uni	normalizedPeakToPeak^{right}	146	0.80 (0.78)	0.17 (0.19)	0.31 (0.14)	0.69 (0.65)	0.84 (0.83)	0.93 (0.93)	1.00 (1.34)
uni	Skewness^{right}	146	0.53 (0.61)	0.57 (0.81)	-1.00 (-1.15)	0.14 (0.04)	0.55 (0.51)	0.91 (0.94)	2.17 (2.80)
uni	Width^{right}	146	360.00 (360.00)	0.00 (0.00)	360.00 (360.00)	360.00 (360.00)	360.00 (360.00)	360.00 (360.00)	360.00 (360.00)
uni vs afix	ΔGlobalMinimum	145	-0.10 (0.14)	5.42 (4.72)	-19.30 (-8.55)	-1.89 (-1.60)	-0.39 (-0.65)	1.00 (0.19)	30.00 (17.11)
uni vs afix	normalizedΔGlobalMinimum	145	-0.03 (-0.06)	0.24 (0.18)	-0.43 (-0.51)	-0.15 (-0.15)	-0.02 (-0.08)	0.07 (0.01)	2.00 (0.48)
uni vs afix	ΔRightMaximum	144	-5.58 (-7.15)	14.20 (17.87)	-59.22 (-55.90)	-8.64 (-8.68)	-2.59 (-1.66)	1.22 (1.21)	41.92 (31.60)
uni vs afix	normalizedΔRightMaximum	144	-0.06 (-0.03)	0.76 (0.74)	-0.93 (-0.90)	-0.47 (-0.50)	-0.22 (-0.22)	0.12 (0.23)	6.07 (3.20)
uni vs ain	ΔGlobalMinimum	146	0.96 (1.53)	6.04 (6.52)	-14.83 (-6.99)	-1.12 (-1.08)	0.00 (-0.09)	0.89 (1.07)	41.39 (28.01)
uni vs ain	normalizedΔGlobalMinimum	146	0.00 (0.02)	0.16 (0.20)	-0.35 (-0.32)	-0.08 (-0.09)	0.00 (-0.01)	0.08 (0.12)	0.60 (0.61)
uni vs ain	ΔRightMaximum	146	-0.91 (0.23)	15.75 (16.33)	-49.44 (-33.66)	-6.11 (-6.67)	-0.71 (-0.22)	2.93 (4.81)	76.94 (58.96)
uni vs ain	normalizedΔRightMaximum	146	0.08 (0.13)	0.66 (0.72)	-0.79 (-0.63)	-0.32 (-0.29)	-0.10 (-0.03)	0.30 (0.37)	3.53 (3.85)

Table A.8: Significantly different feature pairs based on the direct method. List of all significantly different feature pairs when evaluated with the direct method, as well as the corresponding p-value of the Kruskal-Wallis test, and medians, means and cell counts.

(a) spatially separate: uni2

None significant.

(b) transparent: uni2

None significant.

(c) spatially separate: uni2-afix

					p	median1	median2	mean1	mean2	n1	n2
x	uni2	GlobalMinimumAngle	afix	GlobalMinimumAngle	0.000	60.00	0.00	64.59	15.69	85	109
x	uni2	$\text{Bandwidth}^{\text{right}}_{75\,\%}$	afix	$\text{Bandwidth}^{\text{right}}_{75\,\%}$	0.037	90.00	90.00	94.24	84.22	85	109
x	uni2	$\Delta\text{Width}^{\text{right}}$	afix	$\Delta\text{Width}^{\text{right}}$	0.000	0.00	60.00	9.18	60.55	85	109
x	uni2	$\text{InnerWidth}^{\text{right}}$	afix	$\text{InnerWidth}^{\text{right}}$	0.000	180.00	60.00	175.41	65.78	85	109
x	uni2	$\text{OuterWidth}^{\text{right}}$	afix	$\text{OuterWidth}^{\text{right}}$	0.000	180.00	120.00	184.59	126.33	85	109
x	uni2	$\text{Width}^{\text{right}}$	afix	$\text{Width}^{\text{right}}$	0.000	360.00	210.00	360.00	192.11	85	109
y	uni2	GlobalMinimum	afix	GlobalMinimum	0.017	4.50	7.14	7.36	13.52	85	109
y	uni2	$\text{normalizedMaximum}^{\text{right}}$	afix	$\text{normalizedMaximum}^{\text{right}}$	0.000	1.00	0.94	1.00	0.91	85	85
y	uni2	$\text{normalizedPeakToPeak}^{\text{right}}$	afix	$\text{normalizedPeakToPeak}^{\text{right}}$	0.000	0.86	0.68	0.83	0.69	85	85
shape	uni2	$\text{Kurtosis}^{\text{right}}$	afix	$\text{Kurtosis}^{\text{right}}$	0.000	-0.74	-1.22	-0.52	-1.07	85	109
shape	uni2	$\text{Skewness}^{\text{right}}$	afix	$\text{Skewness}^{\text{right}}$	0.000	0.63	0.14	0.69	0.08	85	109

(d) transparent: uni2-afix

					p	median1	median2	mean1	mean2	n1	n2
x	uni2	GlobalMinimumAngle	afix	GlobalMinimumAngle	0.015	60.00	30.00	58.56	38.90	146	145
x	uni2	$\text{Bandwidth}^{\text{right}}_{75\,\%}$	afix	$\text{Bandwidth}^{\text{right}}_{75\,\%}$	0.049	90.00	60.00	84.66	79.17	146	144
x	uni2	$\Delta\text{Width}^{\text{right}}$	afix	$\Delta\text{Width}^{\text{right}}$	0.000	0.00	60.00	-2.88	68.54	146	144
x	uni2	$\text{InnerWidth}^{\text{right}}$	afix	$\text{InnerWidth}^{\text{right}}$	0.000	180.00	60.00	181.44	63.96	146	144
x	uni2	$\text{InnerBandwidth}^{\text{right}}_{75\,\%}$	afix	$\text{InnerBandwidth}^{\text{right}}_{75\,\%}$	0.006	30.00	30.00	40.68	35.83	146	144
x	uni2	$\text{OuterWidth}^{\text{right}}$	afix	$\text{OuterWidth}^{\text{right}}$	0.000	180.00	120.00	178.56	132.50	146	144
x	uni2	$\text{Width}^{\text{right}}$	afix	$\text{Width}^{\text{right}}$	0.000	360.00	195.00	360.00	196.46	146	144
y	uni2	normalizedGlobalMaximum	afix	normalizedGlobalMaximum	0.039	1.00	0.93	1.00	1.14	146	145
y	uni2	normalizedGlobalMinimum	afix	normalizedGlobalMinimum	0.023	0.16	0.12	0.20	0.17	146	145
y	uni2	$\text{normalizedMaximum}^{\text{right}}$	afix	$\text{normalizedMaximum}^{\text{right}}$	0.000	1.00	0.78	1.00	0.94	146	144
y	uni2	$\text{normalizedPeakToPeak}^{\text{right}}$	afix	$\text{normalizedPeakToPeak}^{\text{right}}$	0.000	0.84	0.55	0.80	0.76	146	144
shape	uni2	$\text{Kurtosis}^{\text{right}}$	afix	$\text{Kurtosis}^{\text{right}}$	0.000	-0.83	-1.09	-0.55	-0.89	146	145
shape	uni2	$\text{Skewness}^{\text{right}}$	afix	$\text{Skewness}^{\text{right}}$	0.000	0.55	0.24	0.53	0.26	146	145

(e) spatially separate vs transparent: uni2-afix

		p	median1	median2	mean1	mean2	n1	n2
y	normalizedΔRightMaximum	0.021	-0.06	-0.22	-0.09	-0.06	85	144

(f) spatially separate: afix

				p	median1	median2	mean1	mean2	n1	n2
x	afix	InnerBandwidth$^{\text{left}}_{75\,\%}$	afix OuterBandwidth$^{\text{right}}_{75\,\%}$	0.027	30.00	30.00	39.36	44.86	109	109
x	afix	OuterWidth$^{\text{left}}$	afix OuterWidth$^{\text{right}}$	0.002	120.00	120.00	110.37	126.33	109	109
x	afix	Width$^{\text{left}}$	afix Width$^{\text{right}}$	0.000	150.00	210.00	167.89	192.11	109	109
x	afix	InnerBandwidth$^{\text{right}}_{75\,\%}$	afix OuterBandwidth$^{\text{right}}_{75\,\%}$	0.017	30.00	30.00	39.36	44.86	109	109
y	afix	Dip$^{\text{left}}$	afix Dip$^{\text{right}}$	0.010	9.80	13.50	12.92	15.90	109	109
shape	afix	Kurtosis$^{\text{left}}$	afix Kurtosis$^{\text{right}}$	0.034	-1.02	-1.22	-0.93	-1.07	109	109

(g) transparent: afix

				p	median1	median2	mean1	mean2	n1	n2
x	afix	InnerBandwidth$^{\text{left}}_{75\,\%}$	afix OuterBandwidth$^{\text{right}}_{75\,\%}$	0.007	30.00	30.00	36.04	43.33	144	144
x	afix	Width$^{\text{left}}$	afix Width$^{\text{right}}$	0.014	180.00	195.00	186.25	196.46	144	144
x	afix	InnerBandwidth$^{\text{right}}_{75\,\%}$	afix OuterBandwidth$^{\text{right}}_{75\,\%}$	0.003	30.00	30.00	35.83	43.33	144	144

(h) spatially separate vs transparent: afix

		p	median1	median2	mean1	mean2	n1	n2
x	GlobalMinimumAngle	0.017	0.00	30.00	15.69	38.90	109	145
x	$\text{Bandwidth}_{75\,\%}^{\text{right}}$	0.032	90.00	60.00	84.22	79.17	109	144
y	ΔMaximum	0.012	3.60	0.00	2.98	0.97	109	143
y	ΔPeakToPeak	0.012	3.60	0.00	2.98	0.97	109	143
y	Dip	0.000	12.50	6.50	14.41	8.12	109	143
y	GlobalMaximum	0.000	34.50	14.00	43.41	21.15	109	145
y	GlobalMinimum	0.000	7.14	2.00	13.52	5.01	109	145
y	normalizedGlobalMinimum	0.006	0.20	0.12	0.22	0.17	85	145
y	innerMinimumVal	0.000	16.00	4.00	23.84	10.48	109	145
y	normalizedinnerMinimumVal	0.000	0.49	0.32	0.47	0.37	85	145
y	Dip^{left}	0.000	9.80	5.83	12.92	7.59	109	144
y	$\text{Maximum}^{\text{left}}$	0.000	29.50	11.00	36.76	18.14	109	144
y	$\text{PeakToPeak}^{\text{left}}$	0.000	21.00	9.08	23.24	13.09	109	144
y	PeakToPeak	0.000	26.80	11.93	29.89	16.14	109	145
y	$\text{Dip}^{\text{right}}$	0.000	13.50	6.07	15.90	8.60	109	143
y	$\text{Maximum}^{\text{right}}$	0.000	29.60	12.00	39.73	19.10	109	144
y	$\text{normalizedMaximum}^{\text{right}}$	0.021	0.94	0.78	0.91	0.94	85	144
y	$\text{PeakToPeak}^{\text{right}}$	0.000	22.90	9.50	26.22	14.05	109	144
shape	ΔSkewness	0.009	0.11	0.58	0.18	0.51	109	144

(i) spatially separate: afix-ain

				p	median1	median2	mean1	mean2	n1	n2
x	afix	ΔInnerWidth	ain	ΔInnerWidth	0.004	0.00	30.00	8.26	25.32	109	109
x	afix	$\text{Bandwidth}_{75\,\%}^{\text{left}}$	ain	$\text{Bandwidth}_{75\,\%}^{\text{left}}$	0.040	90.00	60.00	82.57	74.59	109	109
x	afix	$\text{InnerWidth}^{\text{left}}$	ain	$\text{InnerWidth}^{\text{left}}$	0.011	60.00	30.00	57.52	48.17	109	109
x	afix	$\text{InnerBandwidth}_{75\,\%}^{\text{left}}$	ain	$\text{InnerBandwidth}_{75\,\%}^{\text{left}}$	0.020	30.00	30.00	39.36	33.30	109	109
x	afix	$\text{Bandwidth}_{75\,\%}^{\text{right}}$	ain	$\text{Bandwidth}_{75\,\%}^{\text{right}}$	0.004	90.00	90.00	84.22	94.95	109	109
x	afix	$\text{InnerBandwidth}_{75\,\%}^{\text{right}}$	ain	$\text{InnerBandwidth}_{75\,\%}^{\text{right}}$	0.010	30.00	30.00	39.36	46.24	109	109
y	afix	ΔMaximum	ain	ΔMaximum	0.000	3.60	9.13	2.98	10.40	109	109
y	afix	normalizedΔMaximum	ain	normalizedΔMaximum	0.000	0.08	0.26	0.03	0.29	85	85
y	afix	ΔPeakToPeak	ain	ΔPeakToPeak	0.000	3.60	9.13	2.98	10.40	109	109
y	afix	normalizedΔPeakToPeak	ain	normalizedΔPeakToPeak	0.000	0.08	0.26	0.03	0.29	85	85
y	afix	Dip^{left}	ain	Dip^{left}	0.000	9.80	5.96	12.92	8.03	109	109
y	afix	$\text{normalizedPeakToPeak}^{\text{left}}$	ain	$\text{normalizedPeakToPeak}^{\text{left}}$	0.010	0.60	0.50	0.65	0.54	85	85
y	afix	$\text{normalizedMaximum}^{\text{right}}$	ain	$\text{normalizedMaximum}^{\text{right}}$	0.000	0.94	1.04	0.91	1.09	85	85
y	afix	$\text{normalizedPeakToPeak}^{\text{right}}$	ain	$\text{normalizedPeakToPeak}^{\text{right}}$	0.013	0.68	0.76	0.69	0.83	85	85
shape	afix	ΔSkewness	ain	ΔSkewness	0.044	0.11	-0.13	0.18	-0.04	109	109
shape	afix	$\text{Kurtosis}^{\text{left}}$	ain	$\text{Kurtosis}^{\text{left}}$	0.032	-1.02	-1.25	-0.93	-1.07	109	109
shape	afix	$\text{Skewness}^{\text{right}}$	ain	$\text{Skewness}^{\text{right}}$	0.006	0.14	-0.04	0.08	-0.06	109	109

(j) spatially separate vs transparent: afix

		p	median1	median2	mean1	mean2	n1	n2
x	GlobalMinimumAngle	0.017	0.00	30.00	15.69	38.90	109	145
x	$\text{Bandwidth}_{75\,\%}^{\text{right}}$	0.032	90.00	60.00	84.22	79.17	109	144
y	ΔMaximum	0.012	3.60	0.00	2.98	0.97	109	143
y	ΔPeakToPeak	0.012	3.60	0.00	2.98	0.97	109	143
y	Dip	0.000	12.50	6.50	14.41	8.12	109	143
y	GlobalMaximum	0.000	34.50	14.00	43.41	21.15	109	145
y	GlobalMinimum	0.000	7.14	2.00	13.52	5.01	109	145
y	normalizedGlobalMinimum	0.006	0.20	0.12	0.22	0.17	85	145
y	innerMinimumVal	0.000	16.00	4.00	23.84	10.48	109	145
y	normalizedinnerMinimumVal	0.000	0.49	0.32	0.47	0.37	85	145
y	Dip^{left}	0.000	9.80	5.83	12.92	7.59	109	144
y	$\text{Maximum}^{\text{left}}$	0.000	29.50	11.00	36.76	18.14	109	144
y	$\text{PeakToPeak}^{\text{left}}$	0.000	21.00	9.08	23.24	13.09	109	144
y	PeakToPeak	0.000	26.80	11.93	29.89	16.14	109	145
y	$\text{Dip}^{\text{right}}$	0.000	13.50	6.07	15.90	8.60	109	143
y	$\text{Maximum}^{\text{right}}$	0.000	29.60	12.00	39.73	19.10	109	144
y	$\text{normalizedMaximum}^{\text{right}}$	0.021	0.94	0.78	0.91	0.94	85	144
y	$\text{PeakToPeak}^{\text{right}}$	0.000	22.90	9.50	26.22	14.05	109	144
shape	ΔSkewness	0.009	0.11	0.58	0.18	0.51	109	144

(k) transparent: afix-ain

					p	median1	median2	mean1	mean2	n1	n2
y	afix	normalizedGlobalMaximum	ain	normalizedGlobalMaximum	0.004	0.93	1.08	1.14	1.25	145	146
y	afix	normalizedinnerMinimumVal	ain	normalizedinnerMinimumVal	0.008	0.32	0.41	0.37	0.44	145	146
y	afix	normalizedMaximum^{left}	ain	normalizedMaximum^{left}	0.013	0.76	0.87	0.92	0.99	144	145
y	afix	normalizedPeakToPeak^{left}	ain	normalizedPeakToPeak^{left}	0.038	0.51	0.64	0.74	0.79	144	145
y	afix	normalizedPeakToPeak	ain	normalizedPeakToPeak	0.015	0.74	0.85	0.97	1.05	145	146
y	afix	normalizedMaximum^{right}	ain	normalizedMaximum^{right}	0.005	0.78	0.90	0.94	1.08	144	146
y	afix	normalizedPeakToPeak^{right}	ain	normalizedPeakToPeak^{right}	0.013	0.55	0.70	0.76	0.87	144	146

(l) spatially separate vs transparent: afix-ain

		p	median1	median2	mean1	mean2	n1	n2
y	ΔLeftMaximum	0.000	-1.00	1.70	-1.05	4.35	109	143
y	normalizedΔLeftMaximum	0.000	-0.05	0.13	-0.08	0.08	85	143
y	ΔRightMaximum	0.018	4.57	1.65	6.37	4.65	109	144

(m) spatially separate: ain

		p	median1	median2	mean1	mean2	n1	n2
x ain $\text{Bandwidth}^{\text{left}}_{75\,\%}$	ain $\text{Bandwidth}^{\text{right}}_{75\,\%}$	0.000	60.00	90.00	74.59	94.95	109	109
x ain $\text{InnerWidth}^{\text{left}}$	ain $\text{InnerWidth}^{\text{right}}$	0.000	30.00	90.00	48.17	73.49	109	109
x ain $\text{InnerBandwidth}^{\text{left}}_{75\,\%}$	ain $\text{OuterBandwidth}^{\text{left}}_{75\,\%}$	0.002	30.00	30.00	33.30	41.28	109	109
x ain $\text{InnerBandwidth}^{\text{left}}_{75\,\%}$	ain $\text{InnerBandwidth}^{\text{right}}_{75\,\%}$	0.000	30.00	30.00	33.30	46.24	109	109
x ain $\text{InnerBandwidth}^{\text{left}}_{75\,\%}$	ain $\text{OuterBandwidth}^{\text{right}}_{75\,\%}$	0.000	30.00	30.00	33.30	48.72	109	109
x ain $\text{OuterWidth}^{\text{left}}$	ain $\text{OuterWidth}^{\text{right}}$	0.017	120.00	120.00	114.77	126.88	109	109
x ain $\text{OuterBandwidth}^{\text{left}}_{75\,\%}$	ain $\text{OuterBandwidth}^{\text{right}}_{75\,\%}$	0.013	30.00	30.00	41.28	48.72	109	109
x ain $\text{Width}^{\text{left}}$	ain $\text{Width}^{\text{right}}$	0.000	150.00	210.00	162.94	200.37	109	109
y ain Dip^{left}	ain $\text{Dip}^{\text{right}}$	0.000	5.96	15.45	8.03	18.43	109	109
y ain $\text{Maximum}^{\text{left}}$	ain $\text{Maximum}^{\text{right}}$	0.003	27.11	36.29	35.71	46.10	109	109
y ain $\text{normalizedMaximum}^{\text{left}}$	ain $\text{normalizedMaximum}^{\text{right}}$	0.000	0.79	1.04	0.80	1.09	85	85
y ain $\text{PeakToPeak}^{\text{left}}$	ain $\text{PeakToPeak}^{\text{right}}$	0.000	17.14	26.60	20.24	30.64	109	109
y ain $\text{normalizedPeakToPeak}^{\text{left}}$	ain $\text{normalizedPeakToPeak}^{\text{right}}$	0.000	0.50	0.76	0.54	0.83	85	85

(n) transparent: ain

		p	median1	median2	mean1	mean2	n1	n2
x ain $\text{InnerWidth}^{\text{left}}$	ain $\text{InnerWidth}^{\text{right}}$	0.014	60.00	60.00	57.52	67.60	145	146
x ain $\text{InnerBandwidth}^{\text{left}}_{75\,\%}$	ain $\text{OuterBandwidth}^{\text{right}}_{75\,\%}$	0.019	30.00	30.00	36.00	44.79	145	146
x ain $\text{Width}^{\text{left}}$	ain $\text{Width}^{\text{right}}$	0.013	180.00	180.00	181.86	198.08	145	146
x ain $\text{InnerBandwidth}^{\text{right}}_{75\,\%}$	ain $\text{OuterBandwidth}^{\text{right}}_{75\,\%}$	0.017	30.00	30.00	36.99	44.79	146	146

(o) spatially separate vs transparent: ain

		p	median1	median2	mean1	mean2	n1	n2
x	ΔInnerWidth	0.010	30.00	0.00	25.32	9.93	109	145
x	InnerWidth^{left}	0.029	30.00	60.00	48.17	57.52	109	145
x	Width^{left}	0.032	150.00	180.00	162.94	181.86	109	145
x	Bandwidth$^{right}_{75\,\%}$	0.000	90.00	60.00	94.95	81.78	109	146
x	InnerBandwidth$^{right}_{75\,\%}$	0.000	30.00	30.00	46.24	36.99	109	146
x	OuterBandwidth$^{right}_{75\,\%}$	0.010	30.00	30.00	48.72	44.79	109	146
y	ΔMaximum	0.000	9.13	0.50	10.40	1.29	109	145
y	normalizedΔMaximum	0.000	0.26	0.05	0.29	0.08	85	145
y	ΔPeakToPeak	0.000	9.13	0.50	10.40	1.29	109	145
y	normalizedΔPeakToPeak	0.000	0.26	0.05	0.29	0.08	85	145
y	Dip	0.000	11.20	6.90	13.23	9.95	109	145
y	normalizedDip	0.001	0.35	0.48	0.39	0.59	85	145
y	GlobalMaximum	0.000	38.50	15.75	47.58	26.52	109	146
y	GlobalMinimum	0.000	7.71	1.93	15.47	6.04	109	146
y	normalizedGlobalMinimum	0.009	0.23	0.14	0.26	0.20	85	146
y	innerMinimumVal	0.000	18.80	5.00	27.68	12.94	109	146
y	normalizedinnerMinimumVal	0.001	0.57	0.41	0.56	0.44	85	146
y	Maximum^{left}	0.000	27.11	12.22	35.71	22.33	109	145
y	normalizedMaximum^{left}	0.020	0.79	0.87	0.80	0.99	85	145
y	PeakToPeak^{left}	0.000	17.14	10.00	20.24	16.25	109	145
y	normalizedPeakToPeak^{left}	0.000	0.50	0.64	0.54	0.79	85	145
y	minusLeftSkewness	0.034	0.01	-0.14	-0.02	-0.19	109	145
y	PeakToPeak	0.000	29.33	12.50	32.11	20.48	109	146
y	Dip^{right}	0.000	15.45	6.80	18.43	10.59	109	145
y	Maximum^{right}	0.000	36.29	13.50	46.10	23.48	109	146
y	normalizedMaximum^{right}	0.034	1.04	0.90	1.09	1.08	85	146
y	PeakToPeak^{right}	0.000	26.60	10.00	30.64	17.44	109	146
shape	ΔSkewness	0.000	-0.13	0.42	-0.04	0.43	109	145
shape	Kurtosis^{left}	0.021	-1.25	-1.06	-1.07	-0.88	109	145
shape	Kurtosis^{right}	0.047	-1.15	-1.02	-1.07	-0.87	109	146
shape	Skewness^{right}	0.000	-0.04	0.26	-0.06	0.23	109	146
shape	tcSymmetryIndex	0.009	0.45	0.40	0.44	0.39	109	146

Table A.9: Significantly different feature pairs based on the best model. List of all significantly different feature pairs when evaluated with the best model "bM", as well as the corresponding p-value of the Kruskal-Wallis test, and medians, means and cell counts.

(a) spatially separate: uni2

None significant.

(b) transparent: uni2

None significant.

(c) spatially separate: uni2-afix

					p	median1	median2	mean1	mean2	n1	n2
x	uni2	GlobalMinimumAngle	afix	GlobalMinimumAngle	0.000	67.30	12.80	63.41	21.42	85	109
x	uni2	$\text{Bandwidth}_{75\,\%}^{\text{right}}$	afix	$\text{Bandwidth}_{75\,\%}^{\text{right}}$	0.002	61.60	54.80	69.70	58.21	85	109
x	uni2	$\Delta\text{Width}^{\text{right}}$	afix	$\Delta\text{Width}^{\text{right}}$	0.000	0.00	68.00	1.20	67.00	85	109
x	uni2	$\text{InnerWidth}^{\text{right}}$	afix	$\text{InnerWidth}^{\text{right}}$	0.000	180.00	61.70	179.40	63.12	85	109
x	uni2	$\text{InnerBandwidth}_{75\,\%}^{\text{right}}$	afix	$\text{InnerBandwidth}_{75\,\%}^{\text{right}}$	0.000	30.20	25.90	34.48	26.07	85	109
x	uni2	$\text{OuterWidth}^{\text{right}}$	afix	$\text{OuterWidth}^{\text{right}}$	0.000	180.00	133.90	180.60	130.12	85	109
x	uni2	$\text{Width}^{\text{right}}$	afix	$\text{Width}^{\text{right}}$	0.000	360.00	199.80	360.00	193.24	85	109
y	uni2	normalizedGlobalMaximum	afix	normalizedGlobalMaximum	0.003	1.00	1.08	1.00	1.10	85	85
y	uni2	GlobalMinimum	afix	GlobalMinimum	0.015	4.14	6.78	7.25	13.40	85	109
y	uni2	$\text{normalizedPeakToPeak}^{\text{right}}$	afix	$\text{normalizedPeakToPeak}^{\text{right}}$	0.001	0.87	0.74	0.83	0.73	85	85
shape	uni2	$\text{Kurtosis}^{\text{right}}$	afix	$\text{Kurtosis}^{\text{right}}$	0.000	-0.75	-1.21	-0.48	-0.98	85	109
shape	uni2	$\text{Skewness}^{\text{right}}$	afix	$\text{Skewness}^{\text{right}}$	0.000	0.64	0.06	0.71	0.02	85	109

(d) transparent: uni2-afix

					p	median1	median2	mean1	mean2	n1	n2
x	uni2	$\text{Bandwidth}^{\text{right}}_{75\,\%}$	afix	$\text{Bandwidth}^{\text{right}}_{75\,\%}$	0.012	60.90	50.00	68.48	59.65	146	143
x	uni2	$\Delta\text{Width}^{\text{right}}$	afix	$\Delta\text{Width}^{\text{right}}$	0.000	-25.20	45.70	-21.02	46.73	146	143
x	uni2	$\text{InnerWidth}^{\text{right}}$	afix	$\text{InnerWidth}^{\text{right}}$	0.000	192.60	61.60	190.51	71.58	146	143
x	uni2	$\text{InnerBandwidth}^{\text{right}}_{75\,\%}$	afix	$\text{InnerBandwidth}^{\text{right}}_{75\,\%}$	0.000	29.50	23.00	33.93	27.69	146	143
x	uni2	$\text{OuterWidth}^{\text{right}}$	afix	$\text{OuterWidth}^{\text{right}}$	0.000	167.40	118.00	169.49	118.31	146	143
x	uni2	$\text{OuterBandwidth}^{\text{right}}_{75\,\%}$	afix	$\text{OuterBandwidth}^{\text{right}}_{75\,\%}$	0.048	28.45	25.60	34.54	31.96	146	143
x	uni2	$\text{Width}^{\text{right}}$	afix	$\text{Width}^{\text{right}}$	0.000	360.00	187.80	360.00	189.89	146	143
y	uni2	normalizedGlobalMaximum	afix	normalizedGlobalMaximum	0.003	1.00	0.89	1.00	1.69	146	143
y	uni2	$\text{normalizedMaximum}^{\text{right}}$	afix	$\text{normalizedMaximum}^{\text{right}}$	0.000	1.00	0.76	0.98	1.33	146	143
y	uni2	$\text{normalizedPeakToPeak}^{\text{right}}$	afix	$\text{normalizedPeakToPeak}^{\text{right}}$	0.003	0.83	0.62	0.78	0.93	146	143
shape	uni2	$\text{Kurtosis}^{\text{right}}$	afix	$\text{Kurtosis}^{\text{right}}$	0.020	-0.85	-0.98	-0.21	-0.74	146	143
shape	uni2	$\text{Skewness}^{\text{right}}$	afix	$\text{Skewness}^{\text{right}}$	0.000	0.51	-0.05	0.61	-0.03	146	143

(e) spatially separate vs transparent: uni2-afix

			p	median1	median2	mean1	mean2	n1	n2
y	$\Delta\text{RightMaximum}$	0.042		-0.24	-1.66	0.35	-7.15	27	40

(f) spatially separate: afix

		p	median1	median2	mean1	mean2	n1	n2
x afix ΔOuterWidth	afix ΔInnerWidth	0.004	32.60	8.70	22.18	4.31	109	109
x afix $\Delta\text{Width}^{\text{left}}$	afix $\Delta\text{Width}^{\text{right}}$	0.002	50.60	68.00	49.12	67.00	109	109
x afix $\text{InnerWidth}^{\text{left}}$	afix $\text{InnerWidth}^{\text{right}}$	0.042	56.30	61.70	58.82	63.12	109	109
x afix $\text{InnerBandwidth}^{\text{left}}_{75\,\%}$	afix $\text{OuterBandwidth}^{\text{right}}_{75\,\%}$	0.017	25.50	28.80	27.89	32.14	109	109
x afix $\text{OuterWidth}^{\text{left}}$	afix $\text{OuterWidth}^{\text{right}}$	0.000	106.70	133.90	107.94	130.12	109	109
x afix $\text{Width}^{\text{left}}$	afix $\text{Width}^{\text{right}}$	0.000	160.20	199.80	166.76	193.24	109	109
x afix $\text{InnerBandwidth}^{\text{right}}_{75\,\%}$	afix $\text{OuterBandwidth}^{\text{right}}_{75\,\%}$	0.016	25.90	28.80	26.07	32.14	109	109
y afix Dip^{left}	afix $\text{Dip}^{\text{right}}$	0.037	11.25	14.56	14.46	17.13	109	109

(g) transparent: afix

		p	median1	median2	mean1	mean2	n1	n2
x afix $\text{InnerBandwidth}^{\text{left}}_{75\,\%}$	afix $\text{OuterBandwidth}^{\text{left}}_{75\,\%}$	0.036	22.40	23.60	25.20	32.24	143	143
x afix $\text{InnerBandwidth}^{\text{left}}_{75\,\%}$	afix $\text{OuterBandwidth}^{\text{right}}_{75\,\%}$	0.013	22.40	25.60	25.20	31.96	143	143

(h) spatially separate vs transparent: afix

		p	median1	median2	mean1	mean2	n1	n2
x	maxAngleDist	0.016	117.90	129.40	121.94	138.27	109	143
x	ΔWidth^{right}	0.004	68.00	45.70	67.00	46.73	109	143
x	MaximumAngle^{right}	0.005	240.20	247.50	238.85	251.45	109	143
x	OuterWidth^{right}	0.005	133.90	118.00	130.12	118.31	109	143
y	Dip	0.000	13.62	6.52	15.79	8.15	109	143
y	GlobalMaximum	0.000	35.15	12.94	44.06	20.66	109	143
y	normalizedGlobalMaximum	0.006	1.08	0.89	1.10	1.69	85	143
y	GlobalMinimum	0.000	6.78	1.53	13.40	5.08	109	143
y	normalizedGlobalMinimum	0.023	0.20	0.12	0.22	0.40	85	143
y	innerMinimumVal	0.000	15.29	4.12	22.79	9.94	109	143
y	normalizedinnerMinimumVal	0.001	0.46	0.32	0.46	0.69	85	143
y	Dip^{left}	0.000	11.25	5.10	14.46	7.36	109	143
y	Maximum^{left}	0.000	27.77	10.37	37.25	17.30	109	143
y	PeakToPeak^{left}	0.000	20.89	9.12	23.85	12.22	109	143
y	PeakToPeak	0.000	26.89	10.43	30.67	15.58	109	143
y	normalizedPeakToPeak	0.035	0.84	0.72	0.88	1.29	85	143
y	Dip^{right}	0.000	14.56	6.59	17.13	8.94	109	143
y	Maximum^{right}	0.000	30.34	11.57	39.92	18.88	109	143
y	normalizedMaximum^{right}	0.021	0.97	0.76	0.95	1.33	85	143
y	PeakToPeak^{right}	0.000	23.58	8.70	26.52	13.79	109	143
shape	Kurtosis^{right}	0.005	-1.21	-0.98	-0.98	-0.74	109	143

(i) spatially separate: afix-ain

					p	median1	median2	mean1	mean2	n1	n2
x	afix	ΔInnerWidth	ain	ΔInnerWidth	0.000	8.70	23.80	4.31	32.40	109	109
x	afix	innerMinimumAngle	ain	innerMinimumAngle	0.007	174.00	165.50	175.73	164.07	109	109
x	afix	$\text{InnerWidth}^{\text{left}}$	ain	$\text{InnerWidth}^{\text{left}}$	0.002	56.30	46.60	58.82	46.88	109	109
x	afix	$\text{Bandwidth}^{\text{right}}_{75\,\%}$	ain	$\text{Bandwidth}^{\text{right}}_{75\,\%}$	0.000	54.80	70.20	58.21	72.73	109	109
x	afix	$\Delta\text{Width}^{\text{right}}$	ain	$\Delta\text{Width}^{\text{right}}$	0.014	68.00	58.10	67.00	50.62	109	109
x	afix	$\text{InnerWidth}^{\text{right}}$	ain	$\text{InnerWidth}^{\text{right}}$	0.001	61.70	72.70	63.12	79.28	109	109
x	afix	$\text{InnerBandwidth}^{\text{right}}_{75\,\%}$	ain	$\text{InnerBandwidth}^{\text{right}}_{75\,\%}$	0.000	25.90	30.40	26.07	33.98	109	109
x	afix	$\text{OuterBandwidth}^{\text{right}}_{75\,\%}$	ain	$\text{OuterBandwidth}^{\text{right}}_{75\,\%}$	0.003	28.80	32.80	32.14	38.75	109	109
x	afix	$\text{Width}^{\text{right}}$	ain	$\text{Width}^{\text{right}}$	0.020	199.80	210.50	193.24	209.18	109	109
y	afix	ΔMaximum	ain	ΔMaximum	0.000	3.21	8.46	2.68	9.91	109	109
y	afix	normalizedΔMaximum	ain	normalizedΔMaximum	0.001	0.07	0.26	0.04	0.27	85	85
y	afix	ΔPeakToPeak	ain	ΔPeakToPeak	0.000	3.21	8.46	2.68	9.91	109	109
y	afix	normalizedΔPeakToPeak	ain	normalizedΔPeakToPeak	0.001	0.07	0.26	0.04	0.27	85	85
y	afix	Dip^{left}	ain	Dip^{left}	0.000	11.25	6.21	14.46	9.47	109	109
y	afix	$\text{PeakToPeak}^{\text{left}}$	ain	$\text{PeakToPeak}^{\text{left}}$	0.044	20.89	16.69	23.85	20.18	109	109
y	afix	$\text{normalizedPeakToPeak}^{\text{left}}$	ain	$\text{normalizedPeakToPeak}^{\text{left}}$	0.006	0.64	0.52	0.69	0.57	85	85
y	afix	$\text{normalizedMaximum}^{\text{right}}$	ain	$\text{normalizedMaximum}^{\text{right}}$	0.001	0.97	1.07	0.95	1.11	85	85
shape	afix	$\text{Skewness}^{\text{right}}$	ain	$\text{Skewness}^{\text{right}}$	0.049	0.06	-0.12	0.02	-0.07	109	109
shape	afix	tcSymmetryIndex	ain	tcSymmetryIndex	0.050	7.08	8.12	6.86	7.58	109	109

(j) transparent: afix-ain

				p	median1	median2	mean1	mean2	n1	n2
y afix	normalizedGlobalMaximum	ain	normalizedGlobalMaximum	0.002	0.89	1.14	1.69	1.53	143	146
y afix	normalizedMaximum^{left}	ain	normalizedMaximum^{left}	0.004	0.75	0.91	1.49	1.20	143	146
y afix	normalizedPeakToPeak^{left}	ain	normalizedPeakToPeak^{left}	0.011	0.53	0.69	1.09	1.01	143	146
y afix	normalizedPeakToPeak	ain	normalizedPeakToPeak	0.011	0.72	0.87	1.29	1.34	143	146
y afix	normalizedMaximum^{right}	ain	normalizedMaximum^{right}	0.023	0.76	0.96	1.33	1.34	143	146

(k) spatially separate vs transparent: afix-ain

	p	median1	median2	mean1	mean2	n1	n2
y ΔLeftMaximum	0.000	-3.10	3.11	-2.16	8.18	36	39
y normalizedΔLeftMaximum	0.000	-0.17	0.20	-0.14	0.17	31	39

(l) spatially separate: ain

		p	median1	median2	mean1	mean2	n1	n2
x ain Bandwidth$^{left}_{75\,\%}$	ain Bandwidth$^{right}_{75\,\%}$	0.000	50.50	70.20	57.65	72.73	109	109
x ain InnerWidth^{left}	ain InnerWidth^{right}	0.000	46.60	72.70	46.88	79.28	109	109
x ain InnerBandwidth$^{left}_{75\,\%}$	ain OuterBandwidth$^{left}_{75\,\%}$	0.007	23.40	26.90	23.75	33.90	109	109
x ain InnerBandwidth$^{left}_{75\,\%}$	ain InnerBandwidth$^{right}_{75\,\%}$	0.000	23.40	30.40	23.75	33.98	109	109
x ain InnerBandwidth$^{left}_{75\,\%}$	ain OuterBandwidth$^{right}_{75\,\%}$	0.000	23.40	32.80	23.75	38.75	109	109
x ain OuterWidth^{left}	ain OuterWidth^{right}	0.000	103.40	133.40	110.55	129.90	109	109
x ain OuterBandwidth$^{left}_{75\,\%}$	ain InnerBandwidth$^{right}_{75\,\%}$	0.028	26.90	30.40	33.90	33.98	109	109
x ain OuterBandwidth$^{left}_{75\,\%}$	ain OuterBandwidth$^{right}_{75\,\%}$	0.000	26.90	32.80	33.90	38.75	109	109
x ain Width^{left}	ain Width^{right}	0.000	150.60	210.50	157.42	209.18	109	109
y ain Dip^{left}	ain Dip^{right}	0.000	6.21	15.64	9.47	19.38	109	109
y ain Maximum^{left}	ain Maximum^{right}	0.005	27.89	35.76	35.65	45.56	109	109
y ain normalizedMaximum^{left}	ain normalizedMaximum^{right}	0.000	0.84	1.07	0.83	1.11	85	85
y ain PeakToPeak^{left}	ain PeakToPeak^{right}	0.000	16.69	25.40	20.18	30.09	109	109
y ain normalizedPeakToPeak^{left}	ain normalizedPeakToPeak^{right}	0.000	0.52	0.81	0.57	0.84	85	85

(m) transparent: ain

		p	median1	median2	mean1	mean2	n1	n2
x ain InnerBandwidth$^{left}_{75\,\%}$	ain OuterBandwidth$^{right}_{75\,\%}$	0.004	22.25	24.95	25.52	36.83	146	146
x ain InnerBandwidth$^{right}_{75\,\%}$	ain OuterBandwidth$^{right}_{75\,\%}$	0.003	22.10	24.95	25.71	36.83	146	146

(n) spatially separate vs transparent: ain

		p	median1	median2	mean1	mean2	n1	n2
x	ΔInnerWidth	0.000	23.80	3.45	32.40	-1.12	109	146
x	ΔWidth	0.002	49.40	1.10	31.94	-3.86	109	146
x	innerMinimumAngle	0.000	165.50	188.35	164.07	184.08	109	146
x	$\text{InnerWidth}^{\text{left}}$	0.000	46.60	57.95	46.88	66.95	109	146
x	$\text{Width}^{\text{left}}$	0.000	150.60	180.60	157.42	184.40	109	146
x	$\text{Bandwidth}_{75\,\%}^{\text{right}}$	0.000	70.20	50.40	72.73	62.54	109	146
x	$\text{InnerWidth}^{\text{right}}$	0.002	72.70	61.55	79.28	65.83	109	146
x	$\text{InnerBandwidth}_{75\,\%}^{\text{right}}$	0.000	30.40	22.10	33.98	25.71	109	146
x	$\text{OuterBandwidth}_{75\,\%}^{\text{right}}$	0.000	32.80	24.95	38.75	36.83	109	146
x	$\text{Width}^{\text{right}}$	0.001	210.50	183.60	209.18	190.40	109	146
y	ΔMaximum	0.000	8.46	0.40	9.91	1.18	109	146
y	normalizedΔMaximum	0.000	0.26	0.05	0.27	0.14	85	146
y	ΔPeakToPeak	0.000	8.46	0.40	9.91	1.18	109	146
y	normalizedΔPeakToPeak	0.000	0.26	0.05	0.27	0.14	85	146
y	Dip	0.000	12.15	7.01	14.43	10.46	109	146
y	normalizedDip	0.000	0.38	0.57	0.42	0.87	85	146
y	GlobalMaximum	0.000	37.87	15.93	47.18	26.18	109	146
y	GlobalMinimum	0.000	8.09	1.44	15.47	5.84	109	146
y	normalizedGlobalMinimum	0.007	0.23	0.14	0.27	0.19	85	146
y	innerMinimumVal	0.000	17.35	3.68	26.18	12.11	109	146
y	normalizedinnerMinimumVal	0.001	0.51	0.38	0.55	0.40	85	146
y	$\text{Maximum}^{\text{left}}$	0.000	27.89	11.84	35.65	21.98	109	146
y	$\text{normalizedMaximum}^{\text{left}}$	0.041	0.84	0.91	0.83	1.20	85	146
y	$\text{PeakToPeak}^{\text{left}}$	0.001	16.69	10.06	20.18	16.14	109	146
y	$\text{normalizedPeakToPeak}^{\text{left}}$	0.000	0.52	0.69	0.57	1.01	85	146
y	PeakToPeak	0.000	28.10	12.58	31.71	20.34	109	146
y	$\text{Dip}^{\text{right}}$	0.000	15.64	6.56	19.38	11.05	109	146
y	$\text{Maximum}^{\text{right}}$	0.000	35.76	12.80	45.56	23.16	109	146
y	$\text{PeakToPeak}^{\text{right}}$	0.000	25.40	9.96	30.09	17.32	109	146
shape	$\text{Kurtosis}^{\text{left}}$	0.031	-1.18	-1.01	-0.98	-0.85	109	146
shape	$\text{Kurtosis}^{\text{right}}$	0.029	-1.21	-0.99	-1.10	-0.83	109	146
shape	tcSymmetryIndex	0.016	8.12	6.78	7.58	6.79	109	146

Table A.10: Initial conditions and bounds for least-squares-fits. Model functions and their parameters are described in Methods subsection "Model functions", number before model indicates if the model pertained to data from the one or two stimulus conditions. $u = \min_i y_i + 1.2\,\mathrm{ptp}_i\, y_i$ where $\mathrm{ptp}_i\, y_i = \max_i y_i - \min_i y_i$ and $\{y_i\}$ is the set of all firing rates in the tuning curve. The values of 10.74, 2.41 and 100 for the width-parameter k of the $s\beta$ model correspond to a half-width-at-half-maximum of 45°, 90°and 15°. Likewise the values of $k = 2.3$, 0 and 20.34 for the vM model correspond to 45°, 90°and 15°; As $k = 0$ would make the denominator in the definition of vM zero it was replaced by 0.001.

model	initial condition	lower bounds	upper bounds
1 sβ	$(u/2, 10.74, 240, 10)$	$(0, 2.41, 180, -50)$	$(u, 100, 300, 100)$
1 vM	$(u/2, 2.3, 240, 10)$	$(0, 0.001, 180, -50)$	$(u, 20.34, 300, 100)$
1 wG	$(u/2, 45, 240, 10)$	$(0, 15, 180, -50)$	$(u, 90, 300, 100)$
1 wC	$(u/2, 45, 240, 10)$	$(0, 15, 180, -50)$	$(u, 90, 300, 100)$
1 wB	$(u/2, 45, 240, 10, 2)$	$(0, 15, 180, -50, 1)$	$(u, 90, 300, 100, 10)$
2 sβ	$(u/2, 10.74, 120, 5, u/2, 10.74, 240, 5)$	$(0, 2.41, 30, -25, 0, 2.41, 180, -25)$	$(u, 100, 180, 50, u, 100, 330, 50)$
2 vM	$(u/2, 2.3, 120, 5, u/2, 2.3, 240, 5)$	$(0, 0.001, 30, -25, 0, 0.001, 180, -25)$	$(u, 20.34, 180, 50, u, 20.34, 330, 50)$
2 wG	$(u/2, 45, 120, 5, u/2, 45, 240, 5)$	$(0, 15, 30, -25, 0, 15, 180, -25)$	$(u, 90, 180, 50, u, 90, 330, 50)$
2 wC	$(u/2, 45, 120, 5, u/2, 45, 240, 5)$	$(0, 15, 30, -25, 0, 15, 180, -25)$	$(u, 90, 180, 50, u, 90, 330, 50)$
2 wB	$(u/2, 45, 120, 5, 2, u/2, 45, 240, 5, 2)$	$(0, 15, 30, -25, 1, 0, 15, 180, -25, 1)$	$(u, 90, 180, 50, 10, u, 90, 330, 50, 10)$

BIBLIOGRAPHY

L. F. Abbott and P. Dayan. The effect of correlated variability on the accuracy of a population code. *Neural computation*, 11(1):91–101, 1999.

M. Ainsworth, S. Lee, M. O. Cunningham, A. K. Roopun, R. D. Traub, N. J. Kopell, and M. A. Whittington. Dual Gamma Rhythm Generators Control Interlaminar Synchrony in Auditory Cortex. *The Journal of Neuroscience*, 31(47):17040–17051, November 2011.

H. Akaike. A new look at the statistical model identification. *Automatic Control, IEEE Transactions on*, 19(6):716–723, 1974.

T. D. Albright. Direction and Orientation Selectivity of Neurons in Visual Area MT of the Macaque. *Journal of Neurophysiology*, 52(6):1106–1130, December 1984.

B. Amirikian and A. P. Georgopulos. Directional tuning profiles of motor cortical cells. *Neuroscience Research*, 36(1):73–79, January 2000.

J. C. Anderson, H. Kennedy, and K. A. C. Martin. Pathways of Attention: Synaptic Relationships of Frontal Eye Field to V4, Lateral Intraparietal Cortex, and Area 46 in Macaque Monkey. *The Journal of Neuroscience*, 31(30):10872–10881, July 2011.

A. Angelucci and P. C. Bressloff. Contribution of feedforward, lateral and feedback connections to the classical receptive field center and extra-classical receptive field surround of primate V1 neurons. In *Visual Perception Fundamentals of Vision: Low and Mid-Level Processes in Perception*, volume Volume 154, Part A, pages 93–120. Elsevier, 2006.

S. Ardid, X. J. Wang, and A. Compte. An integrated microcircuit model of attentional processing in the neocortex. *The Journal of neuroscience*, 27(32):8486, 2007.

S. Ardid, X.-J. Wang, D. Gomez-Cabrero, and A. Compte. Reconciling Coherent Oscillation with Modulationof Irregular Spiking Activity in Selective Attention:Gamma-Range Synchronization between Sensoryand Executive Cortical Areas. *The Journal of Neuroscience*, 30(8):2856–2870, February 2010.

A. Arieli, A. Sterkin, A. Grinvald, and A. D. Aertsen. Dynamics of ongoing activity: explanation of the large variability in evoked cortical responses. *Science*, 273(5283): 1868–1871, 1996.

K. M. Armstrong and T. Moore. Rapid enhancement of visual cortical response discriminability by microstimulation of the frontal eye field. *Proceedings of the National Academy of Sciences*, 104(22):9499–9504, May 2007.

K. M. Armstrong, J. K. Fitzgerald, and T. Moore. Changes in Visual Receptive Fields with Microstimulation of Frontal Cortex. *Neuron*, 50(5):791–798, June 2006.

K. M. Armstrong, M. H. Chang, and T. Moore. Selection and Maintenance of Spatial Information by Frontal Eye Field Neurons. *The Journal of Neuroscience*, 29(50): 15621–15629, December 2009.

L. H. Arnal and A.-L. Giraud. Cortical oscillations and sensory predictions. *Trends in Cognitive Sciences*, 16(7):390–398, July 2012.

L. H. Arnal, V. Wyart, and A.-L. Giraud. Transitions in neural oscillations reflect prediction errors generated in audiovisual speech. *Nature Neuroscience*, 14(6):797–801, June 2011.

C. M. Arrington, T. H. Carr, A. R. Mayer, and S. M. Rao. Neural mechanisms of visual attention: object-based selection of a region in space. 2000.

B. B. Averbeck, P. E. Latham, and A. Pouget. Neural correlations, population coding and computation. *Nature Reviews Neuroscience*, 7(5):358–366, May 2006.

R. Azouz and C. M. Gray. Dynamic spike threshold reveals a mechanism for synaptic coincidence detection in cortical neurons in vivo. *Proceedings of the National Academy of Sciences*, 97(14):8110–8115, July 2000.

R. Azouz and C. M. Gray. Adaptive Coincidence Detection and Dynamic Gain Control in Visual Cortical Neurons In Vivo. *Neuron*, 37(3):513–523, February 2003.

W. Bair, E. Zohary, and W. T. Newsome. Correlated Firing in Macaque Visual Area MT: Time Scales and Relationship to Behavior. *The Journal of Neuroscience*, 21(5): 1676–1697, March 2001.

J. H. Barkow, L. E. Cosmides, and J. E. Tooby. *The adapted mind: Evolutionary psychology and the generation of culture.* Oxford University Press, 1992.

P. Barone, A. Batardiere, K. Knoblauch, and H. Kennedy. Laminar Distribution of Neurons in Extrastriate Areas Projecting to Visual Areas V1 and V4 Correlates with the Hierarchical Rank and Indicates the Operation of a Distance Rule. *The Journal of Neuroscience*, 20(9):3263–3281, May 2000.

A. M. Bastos, F. Briggs, H. J. Alitto, G. R. Mangun, and W. M. Usrey. Simultaneous Recordings from the Primary Visual Cortex and Lateral Geniculate Nucleus Reveal Rhythmic Interactions and a Cortical Source for Gamma-Band Oscillations. *The Journal of Neuroscience*, 34(22):7639–7644, May 2014.

A. M. Bastos, J. Vezoli, and P. Fries. Communication through coherence with inter-areal delays. *Current Opinion in Neurobiology*, 31:173–180, April 2015a.

A. Bastos, W. M. Usrey, R. Adams, G. Mangun, P. Fries, and K. Friston. Canonical Microcircuits for Predictive Coding. *Neuron*, 76(4):695–711, November 2012.

A. Bastos, J. Vezoli, C. Bosman, J.-M. Schoffelen, R. Oostenveld, J. Dowdall, P. De-Weerd, H. Kennedy, and P. Fries. Visual Areas Exert Feedforward and Feedback Influences through Distinct Frequency Channels. *Neuron*, 85(2):390–401, January 2015b.

D. Battaglia and D. Hansel. Synchronous Chaos and Broad Band Gamma Rhythm in a Minimal Multi-Layer Model of Primary Visual Cortex. *PLoS Comput Biol*, 7 (10):e1002176, October 2011.

D. Battaglia, N. Brunel, and D. Hansel. Temporal Decorrelation of Collective Oscillations in Neural Networks with Local Inhibition and Long-Range Excitation. *Physical Review Letters*, 99(23), December 2007.

D. Battaglia, A. Witt, F. Wolf, and T. Geisel. Dynamic effective connectivity of inter-areal brain circuits. *PLoS computational biology*, 8(3):e1002438, 2012.

C. Beaulieu and M. Colonnier. A laminar analysis of the number of round-asymmetrical and flat-symmetrical synapses on spines, dendritic trunks, and cell bodies in area 17 of the cat. *The Journal of Comparative Neurology*, 231(2):180–189, January 1985.

R. Ben-Yishai, R. L. Bar-Or, and H. Sompolinsky. Theory of orientation tuning in visual cortex. *Proceedings of the National Academy of Sciences of the United States of America*, 92(9):3844, 1995.

J. S. Bendat and A. G. Piersol. *Random Data: Analysis and Measurement Procedures*. John Wiley & Sons, September 2011.

D. B. Bender and M. Youakim. Effect of Attentive Fixation in Macaque Thalamus and Cortex. *Journal of Neurophysiology*, 85(1):219–234, January 2001.

U. S. Bhalla. Molecular computation in neurons: a modeling perspective. *Current Opinion in Neurobiology*, 25:31–37, April 2014.

N. P. Bichot, A. F. Rossi, and R. Desimone. Parallel and Serial Neural Mechanisms for Visual Search in Macaque Area V4. *Science*, 308(5721):529–534, April 2005.

T. Binzegger, R. J. Douglas, and K. A. C. Martin. Topology and dynamics of the canonical circuit of cat V1. *Neural Networks*, 22(8):1071–1078, October 2009.

T. Binzegger, R. J. Douglas, and K. A. C. Martin. A Quantitative Map of the Circuit of Cat Primary Visual Cortex. *The Journal of Neuroscience*, 24(39):8441–8453, September 2004.

T. Binzegger, R. J. Douglas, and K. A. C. Martin. Stereotypical Bouton Clustering of Individual Neurons in Cat Primary Visual Cortex. *The Journal of Neuroscience*, 27 (45):12242–12254, November 2007.

J. W. Bisley and M. E. Goldberg. Neuronal Activity in the Lateral Intraparietal Area and Spatial Attention. *Science*, 299(5603):81–86, January 2003.

J. W. Bisley and M. E. Goldberg. Attention, Intention, and Priority in the Parietal Lobe. *Annual Review of Neuroscience*, 33(1):1–21, 2010.

E. Blaser, Z. W. Pylyshyn, and A. O. Holcombe. Tracking an object through feature space. *Nature*, 408(6809):196–199, November 2000.

A. Bollimunta, J. Mo, C. E. Schroeder, and M. Ding. Neuronal Mechanisms and Attentional Modulation of Corticothalamic Alpha Oscillations. *The Journal of Neuroscience*, 31(13):4935–4943, March 2011.

A. Borji and L. Itti. State-of-the-Art in Visual Attention Modeling. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 35(1):185–207, January 2013.

R. T. Born and D. C. Bradley. Structure and Function of Visual Area Mt. *Annual Review of Neuroscience*, 28(1):157–189, July 2005.

W. H. Bosking, Y. Zhang, B. Schofield, and D. Fitzpatrick. Orientation Selectivity and the Arrangement of Horizontal Connections in Tree Shrew Striate Cortex. *The Journal of Neuroscience*, 17(6):2112–2127, March 1997.

C. Bosman, J.-M. Schoffelen, N. Brunet, R. Oostenveld, A. Bastos, T. Womelsdorf, B. Rubehn, T. Stieglitz, P. DeWeerd, and P. Fries. Attentional Stimulus Selection through Selective Synchronization between Monkey Visual Areas. *Neuron*, 75(5): 875–888, September 2012.

C. E. Boudreau, T. H. Williford, and J. H. R. Maunsell. Effects of Task Difficulty and Target Likelihood in Area V4 of Macaque Monkeys. *Journal of Neurophysiology*, 96 (5):2377–2387, November 2006.

G. M. Boynton. Attention and visual perception. *Current Opinion in Neurobiology*, 15 (4):465–469, August 2005.

P. D. V. Braitenberg and P. D. D. A. Schüz. Peters' Rule and White's Exceptions. In *Anatomy of the Cortex*, number 18 in Studies of Brain Function, pages 109–112. Springer Berlin Heidelberg, 1991.

T. Branco and M. Häusser. The single dendritic branch as a fundamental functional unit in the nervous system. *Current Opinion in Neurobiology*, 20(4):494–502, August 2010.

S. L. Bressler and J. A. S. Kelso. Cortical coordination dynamics and cognition. *Trends in Cognitive Sciences*, 5(1):26–36, January 2001.

S. L. Bressler and C. G. Richter. Interareal oscillatory synchronization in top-down neocortical processing. *Current Opinion in Neurobiology*, 31:62–66, April 2015.

K. L. Briggman and W. Denk. Towards neural circuit reconstruction with volume electron microscopy techniques. *Current Opinion in Neurobiology*, 16(5):562–570, October 2006.

K. Briggman and W. Kristan. Multifunctional Pattern-Generating Circuits. *Annual Review of Neuroscience*, 31(1):271–294, 2008.

K. H. Britten and H. W. Heuer. Spatial Summation in the Receptive Fields of MT Neurons. *The Journal of Neuroscience*, 19(12):5074 –5084, June 1999.

K. Brodmann. Vergleichende Lokalisationslehre der Groshirnrinde. *Leipzig: Barth*, 1909.

C. D. Brody. Slow Covariations in Neuronal Resting Potentials Can Lead to Artefactually Fast Cross-Correlations in Their Spike Trains. *Journal of Neurophysiology*, 80 (6):3345–3351, December 1998.

C. D. Brody. Correlations Without Synchrony. *Neural Computation*, 11(7):1537–1551, October 1999.

N. Brunel and X. J. Wang. What determines the frequency of fast network oscillations with irregular neural discharges? I. Synaptic dynamics and excitation-inhibition balance. *Journal of neurophysiology*, 90(1):415, 2003.

N. Brunel and X.-J. Wang. Effects of Neuromodulation in a Cortical Network Model of Object Working Memory Dominated by Recurrent Inhibition. *Journal of Computational Neuroscience*, 11(1):63–85, July 2001.

A. Buehlmann and G. Deco. The Neuronal Basis of Attention: Rate versus Synchronization Modulation. *The Journal of Neuroscience*, 28(30):7679–7686, July 2008.

E. A. Buffalo, P. Fries, R. Landman, T. J. Buschman, and R. Desimone. Laminar differences in gamma and alpha coherence in the ventral stream. *Proceedings of the National Academy of Sciences*, 108(27):11262–11267, June 2011.

E. A. Buffalo, P. Fries, R. Landman, H. Liang, and R. Desimone. A backward progression of attentional effects in the ventral stream. *Proceedings of the National Academy of Sciences*, 107(1):361 –365, January 2010.

C. Buia and P. Tiesinga. Attentional modulation of firing rate and synchrony in a model cortical network. *Journal of Computational Neuroscience*, 20(3):247–264, June 2006.

C. I. Buia and P. H. Tiesinga. Role of Interneuron Diversity in the Cortical Microcircuit for Attention. *Journal of Neurophysiology*, 99(5):2158–2182, May 2008.

K. P. Burnham and D. R. Anderson. Multimodel Inference: Understanding AIC and BIC in Model Selection. *Sociological Methods & Research*, 33(2):261–304, November 2004.

K. P. Burnham and D. R. Anderson. *Model selection and multimodel inference: a practical information-theoretic approach*. Springer, New York, 2002.

S. P. Burns, D. Xing, and R. M. Shapley. Is Gamma-Band Activity in the Local Field Potential of V1 Cortex a "Clock" or Filtered Noise? *The Journal of Neuroscience*, 31 (26):9658–9664, June 2011.

T. J. Buschman and E. K. Miller. Top-Down Versus Bottom-Up Control of Attention in the Prefrontal and Posterior Parietal Cortices. *Science*, 315(5820):1860–1862, March 2007.

M. C. Bushnell, M. E. Goldberg, and D. L. Robinson. Behavioral enhancement of visual responses in monkey cerebral cortex. I. Modulation in posterior parietal cortex related to selective visual attention. *Journal of Neurophysiology*, 46(4):755–772, October 1981.

L. Busse, S. Katzner, and S. Treue. Temporal dynamics of neuronal modulation during exogenous and endogenous shifts of visual attention in macaque area MT. *Proceedings of the National Academy of Sciences*, 105(42):16380–16385, October 2008.

D. A. Butts and M. S. Goldman. Tuning Curves, Neuronal Variability, and Sensory Coding. *PLoS Biol*, 4(4):e92, March 2006.

G. Buzsáki and A. Draguhn. Neuronal Oscillations in Cortical Networks. *Science*, 304(5679):1926–1929, June 2004.

C. Börgers and N. J. Kopell. Gamma Oscillations and Stimulus Selection. *Neural Computation*, 20(2):383–414, 2008.

C. Börgers, S. Epstein, and N. J. Kopell. Background gamma rhythmicity and attention in cortical local circuits: A computational study. *Proceedings of the National Academy of Sciences of the United States of America*, 102(19):7002–7007, May 2005.

C. Börgers, S. Epstein, and N. J. Kopell. Gamma oscillations mediate stimulus competition and attentional selection in a cortical network model. *Proceedings of the National Academy of Sciences*, 105(46):18023–18028, November 2008.

M.-C. Calvet and J. Calvet. Horseradish peroxidase iontophoretic intracellular labelling of cultured Purkinje cells. *Brain Research*, 173(3):527–531, September 1979.

J. Cannon, M. M. McCarthy, S. Lee, J. Lee, C. Börgers, M. A. Whittington, and N. Kopell. Neurosystems: brain rhythms and cognitive processing. *European Journal of Neuroscience*, 39(5):705–719, March 2014.

R. T. Canolty and R. T. Knight. The functional role of cross-frequency coupling. *Trends in Cognitive Sciences*, 14(11):506–515, November 2010.

M. Carandini and D. J. Heeger. Normalization as a canonical neural computation. *Nature Reviews Neuroscience*, 13(1):51–62, January 2012.

J. A. Cardin, M. Carlén, K. Meletis, U. Knoblich, F. Zhang, K. Deisseroth, L.-H. Tsai, and C. I. Moore. Driving fast-spiking cells induces gamma rhythm and controls sensory responses. *Nature*, 459(7247):663–667, April 2009.

M. Carrasco. Visual attention: The past 25 years. *Vision Research*, 51(13):1484–1525, July 2011.

J. Cavanaugh and R. H. Wurtz. Subcortical Modulation of Attention Counters Change Blindness. *The Journal of Neuroscience*, 24(50):11236–11243, December 2004.

M. Chalk, J. L. Herrero, M. A. Gieselmann, L. S. Delicato, S. Gotthardt, and A. Thiele. Attention Reduces Stimulus-Driven Gamma Frequency Oscillations and Spike Field Coherence in V1. *Neuron*, 66(1):114–125, April 2010.

R. E. Challis and R. I. Kitney. Biomedical signal processing (in four parts). Part 3. *Medical and Biological Engineering and Computing*, 29(3):225–241, May 1991.

C. A. Charalambides, M. V. Koutras, and N. Balakrishnan. *Probability and Statistical Models with Applications*. Chapman & Hall, Boca Raton, FL, auflage: 2003. edition, September 2000.

D. Chawla, G. Rees, and K. J. Friston. The physiological basis of attentional modulation in extrastriate visual areas. *Nature Neuroscience*, 2(7):671–676, July 1999.

M. I. Chelaru and V. Dragoi. Efficient coding in heterogeneous neuronal populations. *Proceedings of the National Academy of Sciences*, 105(42):16344–16349, October 2008.

L. Chelazzi, J. Duncan, E. K. Miller, and R. Desimone. Responses of Neurons in Inferior Temporal Cortex During Memory-Guided Visual Search. *Journal of Neurophysiology*, 80(6):2918–2940, December 1998.

Y. Chen, S. Martinez-Conde, S. L. Macknik, Y. Bereshpolova, H. A. Swadlow, and J.-M. Alonso. Task difficulty modulates the activity of specific neuronal populations in primary visual cortex. *Nature Neuroscience*, 11(8):974–982, July 2008.

M. M. Churchland, B. M. Yu, J. P. Cunningham, L. P. Sugrue, M. R. Cohen, G. S. Corrado, W. T. Newsome, A. M. Clark, P. Hosseini, B. B. Scott, and others. Stimulus onset quenches neural variability: a widespread cortical phenomenon. *NATURE NEUROSCIENCE*, 13(3):369, 2010.

K. L. Clark, K. M. Armstrong, and T. Moore. Probing neural circuitry and function with electrical microstimulation. *Proceedings of the Royal Society of London B: Biological Sciences*, page rspb20102211, January 2011.

M. R. Cohen and A. Kohn. Measuring and interpreting neuronal correlations. *Nature Neuroscience*, 14(7):811–819, July 2011.

M. R. Cohen and J. H. R. Maunsell. Attention improves performance primarily by reducing interneuronal correlations. *Nature Neuroscience*, 12(12):1594–1600, December 2009.

M. R. Cohen and J. H. R. Maunsell. A Neuronal Population Measure of Attention Predicts Behavioral Performance on Individual Trials. *The Journal of Neuroscience*, 30(45):15241–15253, November 2010.

M. R. Cohen and W. T. Newsome. Context-Dependent Changes in Functional Circuitry in Visual Area MT. *Neuron*, 60(1):162–173, October 2008.

M. Cohen and J. R. Maunsell. Using Neuronal Populations to Study the Mechanisms Underlying Spatial and Feature Attention. *Neuron*, 70(6):1192–1204, June 2011.

C. L. Colby, J. R. Duhamel, and M. E. Goldberg. Visual, presaccadic, and cognitive activation of single neurons in monkey lateral intraparietal area. *Journal of neurophysiology*, 76(5):2841–2852, 1996.

A. Compte, N. Brunel, P. S. Goldman-Rakic, and X.-J. Wang. Synaptic Mechanisms and Network Dynamics Underlying Spatial Working Memory in a Cortical Network Model. *Cerebral Cortex*, 10(9):910–923, September 2000.

C. E. Connor, J. L. Gallant, D. C. Preddie, and D. C. V. Essen. Responses in area V4 depend on the spatial relationship between stimulus and attention. *Journal of Neurophysiology*, 75(3):1306–1308, March 1996.

C. E. Connor, D. C. Preddie, J. L. Gallant, and D. C. V. Essen. Spatial Attention Effects in Macaque Area V4. *The Journal of Neuroscience*, 17(9):3201–3214, May 1997.

B. W. Connors and M. J. Gutnick. Intrinsic firing patterns of diverse neocortical neurons. *Trends in Neurosciences*, 13(3):99–104, March 1990.

C. M. Constantinople and R. M. Bruno. Deep Cortical Layers Are Activated Directly by Thalamus. *Science*, 340(6140):1591–1594, June 2013.

M. Corbetta, J. M. Kincade, J. M. Ollinger, M. P. McAvoy, and G. L. Shulman. Voluntary orienting is dissociated from target detection in human posterior parietal cortex. *Nature Neuroscience*, 3(3):292–297, March 2000.

M. Corbetta, G. Patel, and G. L. Shulman. The Reorienting System of the Human Brain: From Environment to Theory of Mind. *Neuron*, 58(3):306–324, May 2008.

L. Cossell, M. F. Iacaruso, D. R. Muir, R. Houlton, E. N. Sader, H. Ko, S. B. Hofer, and T. D. Mrsic-Flogel. Functional organization of excitatory synaptic strength in primary visual cortex. *Nature*, 518(7539):399–403, February 2015.

T. M. Cover and J. A. Thomas. *Elements of information theory*. Wiley-Interscience, Hoboken, N.J., 2006.

A. M. Craig and H. Boudin. Molecular heterogeneity of central synapses: afferent and target regulation. *Nature Neuroscience*, 4(6):569–578, June 2001.

B. Cronin, I. H. Stevenson, M. Sur, and K. P. Körding. Hierarchical Bayesian Modeling and Markov Chain Monte Carlo Sampling for Tuning-Curve Analysis. *Journal of Neurophysiology*, 103(1):591–602, January 2010.

V. Crunelli and S. W. Hughes. The slow (<1 Hz) rhythm of non-REM sleep: a dialogue between three cardinal oscillators. *Nature Neuroscience*, 13(1):9–17, January 2010.

G. Cybenko. Approximation by superpositions of a sigmoidal function. *Mathematics of Control, Signals and Systems*, 2(4):303–314, December 1989.

S. V. David, B. Y. Hayden, J. A. Mazer, and J. L. Gallant. Attention to Stimulus Features Shifts Spectral Tuning of V4 Neurons during Natural Vision. *Neuron*, 59 (3):509–521, August 2008.

R. L. De Valois, E. William Yund, and N. Hepler. The orientation and direction selectivity of cells in macaque visual cortex. *Vision research*, 22(5):531–544, 1982.

G. C. DeAngelis and W. T. Newsome. Organization of Disparity-Selective Neurons in Macaque Area MT. *The Journal of Neuroscience*, 19(4):1398–1415, February 1999.

W. M. DeBello, T. J. McBride, G. S. Nichols, K. E. Pannoni, D. Sanculi, and D. J. Totten. Input clustering and the microscale structure of local circuits. *Frontiers in Neural Circuits*, 8:112, 2014.

R. C. deCharms and M. M. Merzenich. Primary cortical representation of sounds by the coordination of action-potential timing. *Nature*, 381(6583):610–613, June 1996.

G. Deco, A. Ponce-Alvarez, D. Mantini, G. L. Romani, P. Hagmann, and M. Corbetta. Resting-State Functional Connectivity Emerges from Structurally and Dynamically Shaped Slow Linear Fluctuations. *Journal of Neuroscience*, 33(27):11239–11252, July 2013a.

G. Deco and V. K. Jirsa. Ongoing Cortical Activity at Rest: Criticality, Multistability, and Ghost Attractors. *The Journal of Neuroscience*, 32(10):3366–3375, March 2012.

G. Deco and E. T. Rolls. Neurodynamics of Biased Competition and Cooperation for Attention: A Model With Spiking Neurons. *Journal of Neurophysiology*, 94(1): 295–313, July 2005.

G. Deco, V. K. Jirsa, and A. R. McIntosh. Emerging concepts for the dynamical organization of resting-state activity in the brain. *Nature Reviews Neuroscience*, 12 (1):43–56, January 2011.

G. Deco, V. K. Jirsa, and A. R. McIntosh. Resting brains never rest: computational insights into potential cognitive architectures. *Trends in Neurosciences*, 36(5):268–274, May 2013b.

R. Desimone and J. Duncan. Neural mechanisms of selective visual attention. *Annual review of neuroscience*, 18(1):193–222, 1995.

L. E. Dobrunz and C. F. Stevens. Heterogeneity of Release Probability, Facilitation, and Depletion at Central Synapses. *Neuron*, 18(6):995–1008, June 1997.

N. M. Dotson, R. F. Salazar, and C. M. Gray. Frontoparietal Correlation Dynamics Reveal Interplay between Integration and Segregation during Visual Working

Memory. *The Journal of Neuroscience*, 34(41):13600–13613, October 2014.

R. J. Douglas and K. A. Martin. A functional microcircuit for cat visual cortex. *The Journal of Physiology*, 440(1):735–769, August 1991.

R. J. Douglas and K. A. C. Martin. Mapping the Matrix: The Ways of Neocortex. *Neuron*, 56(2):226–238, October 2007.

R. J. Douglas and K. A. Martin. Neuronal Circuits of the Neocortex. *Annual Review of Neuroscience*, 27(1):419–451, 2004.

R. J. Douglas, K. A. Martin, and D. Whitteridge. A Canonical Microcircuit for Neocortex. *Neural Computation*, 1(4):480–488, December 1989.

J. Downar, A. P. Crawley, D. J. Mikulis, and K. D. Davis. The Effect of Task Relevance on the Cortical Response to Changes in Visual and Auditory Stimuli: An Event-Related fMRI Study. *NeuroImage*, 14(6):1256–1267, December 2001.

J. Du, V. Vegh, and D. C. Reutens. The Laminar Cortex Model: A New Continuum Cortex Model Incorporating Laminar Architecture. *PLoS Comput Biol*, 8(10): e1002733, October 2012.

R. Dubner and S. Zeki. Response properties and receptive fields of cells in an anatomically defined region of the superior temporal sulcus in the monkey. *Brain Research*, 35(2):528–532, December 1971.

A. S. Ecker, P. Berens, G. A. Keliris, M. Bethge, N. K. Logothetis, and A. S. Tolias. Decorrelated Neuronal Firing in Cortical Microcircuits. *Science*, 327(5965):584–587, January 2010.

A. S. Ecker, P. Berens, A. S. Tolias, and M. Bethge. The Effect of Noise Correlations in Populations of Diversely Tuned Neurons. *The Journal of Neuroscience*, 31(40): 14272–14283, October 2011.

R. Eckhorn, R. Bauer, W. Jordan, M. Brosch, W. Kruse, M. Munk, and H. J. Reitboeck. Coherent oscillations: A mechanism of feature linking in the visual cortex? *Biological Cybernetics*, 60(2):121–130, December 1988.

G. M. Edelman and J. A. Gally. Degeneracy and complexity in biological systems. *Proceedings of the National Academy of Sciences*, 98(24):13763–13768, November 2001.

E. Eggermann and D. Feldmeyer. Cholinergic filtering in the recurrent excitatory microcircuit of cortical layer 4. *Proceedings of the National Academy of Sciences*, 106 (28):11753–11758, July 2009.

M. Ercsey-Ravasz, N. Markov, C. Lamy, D. VanEssen, K. Knoblauch, Z. Toroczkai, and H. Kennedy. A Predictive Network Model of Cerebral Cortical Connectivity Based on a Distance Rule. *Neuron*, 80(1):184–197, October 2013.

A. Etzold, H. Schwegler, and C. W. Eurich. Coding with noisy neurons: stability of tuning curve estimation strongly depends on the analysis method. *Journal of Neuroscience Methods*, 134(2):109–119, April 2004.

M. Fallah, G. R. Stoner, and J. H. Reynolds. Stimulus-specific competitive selection in macaque extrastriate visual area V4. *Proceedings of the National Academy of Sciences*, 104(10):4165–4169, March 2007.

F. M. Felisberti and J. M. Zanker. Attention modulates perception of transparent motion. *Vision Research*, 45(19):2587–2599, September 2005.

D. J. Felleman and D. C. Van Essen. Distributed Hierarchical Processing in the Primate Cerebral Cortex. *Cerebral Cortex*, 1(1):1 –47, January 1991.

J. Fiser, C. Chiu, and M. Weliky. Small modulation of ongoing cortical dynamics by sensory input during natural vision. *Nature*, 431(7008):573–578, September 2004.

A. C. Flint and B. W. Connors. Two types of network oscillations in neocortex mediated by distinct glutamate receptor subtypes and neuronal populations. *Journal of Neurophysiology*, 75(2):951–957, February 1996.

M. D. Fox, M. Corbetta, A. Z. Snyder, J. L. Vincent, and M. E. Raichle. Spontaneous neuronal activity distinguishes human dorsal and ventral attention systems. *Proceedings of the National Academy of Sciences*, 103(26):10046–10051, June 2006.

P. Fries. A mechanism for cognitive dynamics: neuronal communication through neuronal coherence. *Trends in Cognitive Sciences*, 9(10):474–480, October 2005.

P. Fries, J. H. Reynolds, A. E. Rorie, and R. Desimone. Modulation of Oscillatory Neuronal Synchronization by Selective Visual Attention. *Science*, 291(5508):1560 –1563, February 2001.

P. Fries, T. Womelsdorf, R. Oostenveld, and R. Desimone. The Effects of Visual Stimulation and Selective Visual Attention on Rhythmic Neuronal Synchronization in Macaque Area V4. *The Journal of Neuroscience*, 28(18):4823–4835, April 2008.

G. M. Ghose. Attentional Modulation of Visual Responses by Flexible Input Gain. *Journal of Neurophysiology*, 101(4):2089–2106, April 2009.

G. M. Ghose and J. H. R. Maunsell. Spatial Summation Can Explain the Attentional Modulation of Neuronal Responses to Multiple Stimuli in Area V4. *The Journal of Neuroscience*, 28(19):5115 –5126, May 2008.

S. Gielen, M. Krupa, and M. Zeitler. Gamma oscillations as a mechanism for selective information transmission. *Biological Cybernetics*, 103(2):151–165, August 2010.

M. A. Gieselmann and A. Thiele. Cholinergic influence on attentional modulation in extrastriate cortex of the rhesus monkey. page 429.7, San Diego, 2010.

C. D. Gilbert and T. N. Wiesel. Columnar Specificity of Intrinsic Horizontal and Corticocortical Connections in Cat Visual Cortex. *The Journal of Neuroscience*, 9(7): 2432–2442, July 1989.

C. D. Gilbert and M. Sigman. Brain States: Top-Down Influences in Sensory Processing. *Neuron*, 54(5):677–696, June 2007.

J.-M. Goaillard, A. L. Taylor, D. J. Schulz, and E. Marder. Functional consequences of animal-to-animal variation in circuit parameters. *Nature Neuroscience*, 12(11): 1424–1430, November 2009.

M. E. Goldberg and M. A. Segraves. Visuospatial and motor attention in the monkey. *Neuropsychologia*, 25(1, Part 1):107–118, 1987.

M. S. Goldman, J. Golowasch, E. Marder, and L. F. Abbott. Global Structure, Robustness, and Modulation of Neuronal Models. *The Journal of Neuroscience*, 21(14): 5229–5238, July 2001.

J. Golowasch, M. S. Goldman, L. F. Abbott, and E. Marder. Failure of Averaging in the Construction of a Conductance-Based Neuron Model. *Journal of Neurophysiology*, 87(2):1129–1131, February 2002.

J. P. Gottlieb, M. Kusunoki, and M. E. Goldberg. The representation of visual salience in monkey parietal cortex. *Nature*, 391(6666):481–484, January 1998.

R. Grashow, T. Brookings, and E. Marder. Compensation for Variable Intrinsic Neuronal Excitability by Circuit-Synaptic Interactions. *The Journal of Neuroscience*, 30 (27):9145–9156, July 2010.

C. M. Gray and G. V. D. Prisco. Stimulus-Dependent Neuronal Oscillations and Local Synchronization in Striate Cortex of the Alert Cat. *The Journal of Neuroscience*, 17 (9):3239–3253, May 1997.

C. M. Gray, P. König, A. K. Engel, and W. Singer. Oscillatory responses in cat visual cortex exhibit inter-columnar synchronization which reflects global stimulus properties. *Nature*, 338(6213):334–337, 1989.

G. G. Gregoriou, S. J. Gotts, H. Zhou, and R. Desimone. High-Frequency, Long-Range Coupling Between Prefrontal and Visual Cortex During Attention. *Science*, 324(5931):1207–1210, May 2009.

G. G. Gregoriou, A. F. Rossi, L. G. Ungerleider, and R. Desimone. Lesions of prefrontal cortex reduce attentional modulation of neuronal responses and synchrony in V4. *Nature Neuroscience*, 17(7):1003–1011, July 2014.

G. Gregoriou, S. Gotts, and R. Desimone. Cell-Type-Specific Synchronization of Neural Activity in FEF with V4 during Attention. *Neuron*, 73(3):581–594, February 2012.

I. Grothe, S. D. Neitzel, S. Mandon, and A. K. Kreiter. Switching Neuronal Inputs by Differential Modulations of Gamma-Band Phase-Coherence. *The Journal of Neuroscience*, 32(46):16172–16180, November 2012.

I. Grothe, D. Rotermund, S. D. Neitzel, S. Mandon, U. A. Ernst, A. K. Kreiter, and K. R. Pawelzik. Attention selectively gates afferent signal transmission to area V4. *bioRxiv*, page 019547, May 2015.

A. T. Gulledge and G. J. Stuart. Cholinergic Inhibition of Neocortical Pyramidal Neurons. *The Journal of Neuroscience*, 25(44):10308–10320, November 2005.

A. T. Gulledge, S. B. Park, Y. Kawaguchi, and G. J. Stuart. Heterogeneity of Phasic Cholinergic Signaling in Neocortical Neurons. *Journal of Neurophysiology*, 97(3): 2215–2229, March 2007.

M. Gur, I. Kagan, and D. M. Snodderly. Orientation and Direction Selectivity of Neurons in V1 of Alert Monkeys: Functional Relationships and Laminar Distributions. *Cerebral Cortex*, 15(8):1207–1221, August 2005.

G. J. Gutierrez, T. O'Leary, and E. Marder. Multiple Mechanisms Switch an Electrically Coupled, Synaptically Inhibited Neuron between Competing Rhythmic Oscillators. *Neuron*, 77(5):845–858, June 2013.

I. Guyon and A. Elisseeff. An introduction to variable and feature selection. *The Journal of Machine Learning Research*, 3:1157–1182, 2003.

P. E. Haenny, J. H. R. Maunsell, and P. H. Schiller. State dependent activity in monkey visual cortex. II. Extraretinal factors in V4. *Exp Brain Res*, 69:245–259, 1988.

S. Haeusler and W. Maass. A Statistical Analysis of Information-Processing Properties of Lamina-Specific Cortical Microcircuit Models. *Cerebral Cortex*, 17(1):149–162, January 2007.

S. Haeusler, K. Schuch, and W. Maass. Motif distribution, dynamical properties, and computational performance of two data-based cortical microcircuit templates. *Journal of Physiology-Paris*, 103(1–2):73–87, January 2009.

H. K. Hahn, J. Klein, C. Nimsky, J. Rexilius, and H.-O. Peitgen. Uncertainty in diffusion tensor based fibre tracking. In C. Nimsky and R. Fahlbusch, editors, *Medical Technologies in Neurosurgery*, number 98 in Acta Neurochirurgica Supplements, pages 33–41. Springer Vienna, 2006.

F. H. Hamker. The Reentry Hypothesis: The Putative Interaction of the Frontal Eye Field, Ventrolateral Prefrontal Cortex, and Areas V4, IT for Attention and Eye Movement. *Cerebral Cortex*, 15(4):431–447, April 2005.

A. Hampshire, J. Duncan, and A. M. Owen. Selective Tuning of the Blood Oxygenation Level-Dependent Response during Simple Target Detection Dissociates Human Frontoparietal Subregions. *The Journal of Neuroscience*, 27(23):6219–6223, June 2007.

D. Hansel and C. v. Vreeswijk. The Mechanism of Orientation Selectivity in Primary Visual Cortex without a Functional Map. *The Journal of Neuroscience*, 32(12):4049–4064, March 2012.

E. C. A. Hansen, D. Battaglia, A. Spiegler, G. Deco, and V. K. Jirsa. Functional connectivity dynamics: Modeling the switching behavior of the resting state. *NeuroImage*, 105:525–535, January 2015.

K. D. Harris and T. D. Mrsic-Flogel. Cortical connectivity and sensory coding. *Nature*, 503(7474):51–58, November 2013.

K. D. Harris and A. Thiele. Cortical state and attention. *Nature Reviews Neuroscience*, 12(9):509–523, September 2011.

B. Y. Hayden and J. L. Gallant. Combined effects of spatial and feature-based attention on responses of V4 neurons. *Vision Research*, 49(10):1182–1187, June 2009.

J. Heinzle, K. Hepp, and K. A. C. Martin. A Microcircuit Model of the Frontal Eye Fields. *The Journal of Neuroscience*, 27(35):9341–9353, August 2007.

J. L. Herrero, M. J. Roberts, L. S. Delicato, M. A. Gieselmann, P. Dayan, and A. Thiele. Acetylcholine contributes through muscarinic receptors to attentional modulation in V1. *Nature*, 454(7208):1110–1114, August 2008.

J. Herrero, M. Gieselmann, M. Sanayei, and A. Thiele. Attention-Induced Variance and Noise Correlation Reduction in Macaque V1 Is Mediated by NMDA Receptors. *Neuron*, 78(4):729–739, May 2013.

S. Hill and G. Tononi. Modeling Sleep and Wakefulness in the Thalamocortical System. *Journal of Neurophysiology*, 93(3):1671–1698, March 2005.

A. M. Himmelheber, M. Sarter, and J. P. Bruno. Increases in cortical acetylcholine release during sustained attention performance in rats. *Cognitive Brain Research*, 9 (3):313–325, June 2000.

A. C. Hindmarsh, P. N. Brown, K. E. Grant, S. L. Lee, R. Serban, D. E. Shumaker, and C. S. Woodward. SUNDIALS: Suite of Nonlinear and Differential/Algebraic Equation Solvers. *ACM Trans. Math. Softw.*, 31(3):363–396, September 2005.

J. F. Hipp, A. K. Engel, and M. Siegel. Oscillatory Synchronization in Large-Scale Cortical Networks Predicts Perception. *Neuron*, 69(2):387–396, January 2011.

J. F. Hipp, D. J. Hawellek, M. Corbetta, M. Siegel, and A. K. Engel. Large-scale cortical correlation structure of spontaneous oscillatory activity. *Nature Neuroscience*, 15(6): 884–890, June 2012.

C. J. Honey, R. Kötter, M. Breakspear, and O. Sporns. Network structure of cerebral cortex shapes functional connectivity on multiple time scales. *Proceedings of the National Academy of Sciences*, 104(24):10240–10245, June 2007.

S. D. V. Hooser. Similarity and Diversity in Visual Cortex: Is There a Unifying Theory of Cortical Computation? *The Neuroscientist*, 13(6):639–656, December 2007.

J. B. Hopfinger, M. H. Buonocore, and G. R. Mangun. The neural mechanisms of top-down attentional control. *Nature Neuroscience*, 3(3):284–291, March 2000.

J. C. Horton and D. L. Adams. The cortical column: a structure without a function. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 360(1456):837–862, April 2005.

A. Ignashchenkova, P. W. Dicke, T. Haarmeier, and P. Thier. Neuron-specific contribution of the superior colliculus to overt and covert shifts of attention. *Nature Neuroscience*, 7(1):56–64, January 2004.

I. Indovina and E. Macaluso. Dissociation of Stimulus Relevance and Saliency Factors during Shifts of Visuospatial Attention. *Cerebral Cortex*, 17(7):1701–1711, July 2007.

A. E. Ipata, A. L. Gee, J. Gottlieb, J. W. Bisley, and M. E. Goldberg. LIP responses to a popout stimulus are reduced if it is overtly ignored. *Nature Neuroscience*, 9(8): 1071–1076, August 2006.

E. M. Izhikevich and G. M. Edelman. Large-scale model of mammalian thalamocortical systems. *Proceedings of the National Academy of Sciences*, 105(9):3593–3598, March 2008.

D. Jancke, W. Erlhagen, H. R. Dinse, A. C. Akhavan, M. Giese, A. Steinhage, and G. Schöner. Parametric Population Representation of Retinal Location: Neuronal Interaction Dynamics in Cat Primary Visual Cortex. *The Journal of Neuroscience*, 19 (20):9016–9028, October 1999.

M. R. Jarvis and P. P. Mitra. Sampling Properties of the Spectrum and Coherency of Sequences of Action Potentials. *Neural Computation*, 13(4):717–749, April 2001.

O. Jensen and L. L. Colgin. Cross-frequency coupling between neuronal oscillations. *Trends in Cognitive Sciences*, 11(7):267–269, July 2007.

O. Jensen, M. Bonnefond, T. R. Marshall, and P. Tiesinga. Oscillatory mechanisms of feedforward and feedback visual processing. *Trends in Neurosciences*, 2015.

X. Jia, S. Tanabe, and A. Kohn. Gamma and the Coordination of Spiking Activity in Early Visual Cortex. *Neuron*, 77(4):762–774, February 2013.

V. K. Jirsa and A. R. McIntosh. *Handbook of brain connectivity*. Springer, Berlin; New York, 2007.

E. R. Kandel, J. H. Schwartz, T. M. Jessell, S. A. Siegelbaum, and A. J. Hudspeth. *Principles of Neural Science*. Mcgraw-Hill Publ.Comp., New York, auflage: 5. auflage. edition, 2012.

E. R. Kandel, H. Markram, P. M. Matthews, R. Yuste, and C. Koch. Neuroscience thinks big (and collaboratively). *Nature Reviews Neuroscience*, 14(9):659–664, September 2013.

S. Kastner, M. A. Pinsk, P. De Weerd, R. Desimone, and L. G. Ungerleider. Increased Activity in Human Visual Cortex during Directed Attention in the Absence of Visual Stimulation. *Neuron*, 22(4):751–761, April 1999.

C. Kayser, M. A. Montemurro, N. K. Logothetis, and S. Panzeri. Spike-Phase Coding Boosts and Stabilizes Information Carried by Spatial and Temporal Spike Patterns. *Neuron*, 61(4):597–608, February 2009.

J. A. S. Kelso. Multistability and metastability: understanding dynamic coordination in the brain. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, 367(1591):906–918, April 2012.

T. Kenet, D. Bibitchkov, M. Tsodyks, A. Grinvald, and A. Arieli. Spontaneously emerging cortical representations of visual attributes. *Nature*, 425(6961):954–956, October 2003.

H. Kennedy, K. Knoblauch, and Z. Toroczkai. Why data coherence and quality is critical for understanding interareal cortical networks. *NeuroImage*, 80:37–45, October 2013.

T. v. Kerkoerle, M. W. Self, B. Dagnino, M.-A. Gariel-Mathis, J. Poort, C. v. d. Togt, and P. R. Roelfsema. Alpha and gamma oscillations characterize feedback and feedforward processing in monkey visual cortex. *Proceedings of the National Academy of Sciences*, 111(40):14332–14341, October 2014.

P. S. Khayat, H. Spekreijse, and P. R. Roelfsema. Attention Lights Up New Object Representations before the Old Ones Fade Away. *The Journal of Neuroscience*, 26 (1):138–142, January 2006.

P. S. Khayat, R. Niebergall, and J. C. Martinez-Trujillo. Attention Differentially Modulates Similar Neuronal Responses Evoked by Varying Contrast and Direction Stimuli in Area MT. *The Journal of Neuroscience*, 30(6):2188–2197, February 2010.

J. M. Kincade, R. A. Abrams, S. V. Astafiev, G. L. Shulman, and M. Corbetta. An Event-Related Functional Magnetic Resonance Imaging Study of Voluntary and Stimulus-Driven Orienting of Attention. *The Journal of Neuroscience*, 25(18):4593–4604, May 2005.

M. Kinoshita, K. Yamada, N. Hashimoto, A. Kato, S. Izumoto, T. Baba, M. Maruno, T. Nishimura, and T. Yoshimine. Fiber-tracking does not accurately estimate size of fiber bundle in pathological condition: initial neurosurgical experience using neuronavigation and subcortical white matter stimulation. *NeuroImage*, 25(2):424–429, April 2005.

C. Kirst. *Synchronization, Neuronal Excitability and Information Flow in Networks of Neuronal Oscillators*. PhD thesis, Universität Göttingen, 2012.

H. Ko, S. B. Hofer, B. Pichler, K. A. Buchanan, P. J. Sjöström, and T. D. Mrsic-Flogel. Functional specificity of local synaptic connections in neocortical networks. *Nature*, 473(7345):87–91, May 2011.

A. Kohn and M. A. Smith. Stimulus Dependence of Neuronal Correlation in Primary Visual Cortex of the Macaque. *The Journal of Neuroscience*, 25(14):3661–3673, April 2005.

N. J. Kopell, H. J. Gritton, M. A. Whittington, and M. A. Kramer. Beyond the Connectome: The Dynome. *Neuron*, 83(6):1319–1328, September 2014.

V. Kozyrev, A. Lochte, M. Daliri, D. Battaglia, and S. Treue. Attentional modulation of the tuning of neurons in macaque area MT to the direction of two spatially separated motion patterns. In *Society for Neuroscience*, Chicago, 2009.

R. J. Krauzlis, L. P. Lovejoy, and A. Zénon. Superior Colliculus and Visual Spatial Attention. *Annual review of neuroscience*, 36, July 2013.

J. Kremkow, A. Kumar, S. Rotter, and A. Aertsen. Emergence of population synchrony in a layered network of the cat visual cortex. *Neurocomputing*, 70(10–12): 2069–2073, June 2007.

G. Krone, H. Mallot, G. Palm, and A. Schuz. Spatiotemporal Receptive Fields: A Dynamical Model Derived from Cortical Architectonics. *Proceedings of the Royal Society of London. Series B, Biological Sciences*, 226(1245):421–444, January 1986.

D. Kätzel, B. V. Zemelman, C. Buetfering, M. Wölfel, and G. Miesenböck. The columnar and laminar organization of inhibitory connections to neocortical excitatory cells. *Nature Neuroscience*, 14(1):100–107, January 2011.

P. König, A. K. Engel, and W. Singer. Relation between oscillatory activity and long-range synchronization in cat visual cortex. *Proceedings of the National Academy of Sciences*, 92(1):290–294, January 1995.

J.-P. Lachaux, E. Rodriguez, J. Martinerie, and F. J. Varela. Measuring phase synchrony in brain signals. *Human Brain Mapping*, 8(4):194–208, 1999.

I. D. Landau, R. Egger, M. Oberlaender, and H. Sompolinsky. The relationship between microcircuit structure and dynamics in the rat barrel cortex. Washington D.C., 2014.

H. Lee, G. V. Simpson, N. K. Logothetis, and G. Rainer. Phase Locking of Single Neuron Activity to Theta Oscillations during Working Memory in Monkey Extrastriate Visual Cortex. *Neuron*, 45(1):147–156, January 2005.

J. Lee and J. H. R. Maunsell. A Normalization Model of Attentional Modulation of Single Unit Responses. *PLoS ONE*, 4(2):e4651, February 2009.

J. Lee and J. H. R. Maunsell. Attentional Modulation of MT Neurons with Single or Multiple Stimuli in Their Receptive Fields. *The Journal of Neuroscience*, 30(8): 3058–3066, February 2010.

J. H. Lee, M. A. Whittington, and N. J. Kopell. Top-Down Beta Rhythms Support Selective Attention via Interlaminar Interaction: A Model. *PLoS Computational Biology*, 9(8):e1003164, August 2013.

Y.-t. Li, L. A. Ibrahim, B.-h. Liu, L. I. Zhang, and H. W. Tao. Linear transformation of thalamocortical input by intracortical excitation. *Nature Neuroscience*, 16(9):1324–1330, September 2013.

M. S. Lidow, P. S. Goldman-Rakic, D. W. Gallager, and P. Rakic. Distribution of dopaminergic receptors in the primate cerebral cortex: Quantitative autoradiographic analysis using [3h]raclopride, [3h]spiperone and [3h]SCH23390. *Neuroscience*, 40(3):657–671, 1991.

S. Ling, M. S. Pratte, and F. Tong. Attention alters orientation processing in the human lateral geniculate nucleus. *Nature neuroscience*, 2015.

J. Lisman. The theta/gamma discrete phase code occuring during the hippocampal phase precession may be a more general brain coding scheme. *Hippocampus*, 15 (7):913–922, January 2005.

M. S. Livingstone. Oscillatory firing and interneuronal correlations in squirrel monkey striate cortex. *Journal of Neurophysiology*, 75(6):2467–2485, June 1996.

A. Lochte, V. Stephan, V. Kozyrev, A. Witt, and S. Treue. Attentional modulation of the tuning of neurons in macaque area MT to the direction of transparent motion patterns. In *Society for Neuroscience*, Chicago, October 2009.

M. London and M. Häusser. Dendritic Computation. *Annual Review of Neuroscience*, 28(1):503–532, July 2005.

L. P. Lovejoy and R. J. Krauzlis. Inactivation of primate superior colliculus impairs covert selection of signals for perceptual judgments. *Nature Neuroscience*, 13(2): 261–266, 2010.

S. J. Luck, L. Chelazzi, S. A. Hillyard, and R. Desimone. Neural Mechanisms of Spatial Selective Attention in Areas V1, V2, and V4 of Macaque Visual Cortex. *Journal of Neurophysiology*, 77(1):24–42, January 1997.

E. Macaluso, C. D. Frith, and J. Driver. Supramodal Effects of Covert Spatial Orienting Triggered by Visual or Tactile Events. *Journal of Cognitive Neuroscience*, 14(3): 389–401, April 2002.

N. J. Majaj, M. Carandini, and J. A. Movshon. Motion Integration by Neurons in Macaque MT Is Local, Not Global. *The Journal of Neuroscience*, 27(2):366–370, January 2007.

P. E. Maldonado and C. M. Gray. Heterogeneity in local distributions of orientation-selective neurons in the cat primary visual cortex. *Visual Neuroscience*, 13(03): 509–516, 1996.

R. N. Mantegna. Fast, accurate algorithm for numerical simulation of L\'evy stable stochastic processes. *Physical Review E*, 49(5):4677–4683, May 1994.

E. Marder. Variability, compensation, and modulation in neurons and circuits. *Proceedings of the National Academy of Sciences*, 108(Supplement 3):15542–15548, September 2011.

E. Marder and J.-M. Goaillard. Variability, compensation and homeostasis in neuron and network function. *Nature Reviews Neuroscience*, 7(7):563–574, July 2006.

E. Marder and A. L. Taylor. Multiple models to capture the variability in biological neurons and networks. *Nature Neuroscience*, 14(2):133–138, February 2011.

K. V. Mardia and P. E. Jupp, editors. *Directional Statistics*. Wiley Series in Probability and Statistics. John Wiley & Sons, Inc., Hoboken, NJ, USA, January 1999.

E. Maris, T. Womelsdorf, R. Desimone, and P. Fries. Rhythmic neuronal synchronization in visual cortex entails spatial phase relation diversity that is modulated by stimulation and attention. *NeuroImage*, 74:99–116, July 2013.

N. T. Markov, P. Misery, A. Falchier, C. Lamy, J. Vezoli, R. Quilodran, M. A. Gariel, P. Giroud, M. Ercsey-Ravasz, L. J. Pilaz, C. Huissoud, P. Barone, C. Dehay, Z. Toroczkai, D. C. V. Essen, H. Kennedy, and K. Knoblauch. Weight Consistency Specifies Regularities of Macaque Cortical Networks. *Cerebral Cortex*, 21(6): 1254–1272, June 2011.

N. T. Markov, M. M. Ercsey-Ravasz, A. R. R. Gomes, C. Lamy, L. Magrou, J. Vezoli, P. Misery, A. Falchier, R. Quilodran, M. A. Gariel, J. Sallet, R. Gamanut, C. Huissoud, S. Clavagnier, P. Giroud, D. Sappey-Marinier, P. Barone, C. Dehay, Z. Toroczkai, K. Knoblauch, D. C. V. Essen, and H. Kennedy. A Weighted and Directed Interareal Connectivity Matrix for Macaque. *Cerebral Cortex*, 24(1):17–36, January 2014a.

N. T. Markov, M. Ercsey-Ravasz, C. Lamy, A. R. R. Gomes, L. Magrou, P. Misery, P. Giroud, P. Barone, C. Dehay, Z. Toroczkai, K. Knoblauch, D. C. V. Essen, and H. Kennedy. The role of long-range connections on the specificity of the macaque interareal cortical network. *Proceedings of the National Academy of Sciences*, 110(13): 5187–5192, March 2013.

N. T. Markov, J. Vezoli, P. Chameau, A. Falchier, R. Quilodran, C. Huissoud, C. Lamy, P. Misery, P. Giroud, S. Ullman, P. Barone, C. Dehay, K. Knoblauch, and H. Kennedy. Anatomy of hierarchy: Feedforward and feedback pathways in macaque visual cortex. *Journal of Comparative Neurology*, 522(1):225–259, 2014b.

H. Markram. The Blue Brain Project. *Nature Reviews Neuroscience*, 7(2):153–160, February 2006.

H. Markram and M. Tsodyks. Redistribution of synaptic efficacy between neocortical pyramidal neurons. *Nature*, 382(6594):807–810, August 1996.

K. A. Martin and D. Whitteridge. Form, function and intracortical projections of spiny neurones in the striate visual cortex of the cat. *The Journal of Physiology*, 353 (1):463–504, August 1984.

J. C. Martínez-Trujillo and S. Treue. Attentional modulation strength in cortical area MT depends on stimulus contrast. *Neuron*, 35(2):365–370, 2002.

J. C. Martínez-Trujillo and S. Treue. Feature-Based Attention Increases the Selectivity of Population Responses in Primate Visual Cortex. *Current Biology*, 14(9):744–751, May 2004.

M. F. Mason, M. I. Norton, J. D. V. Horn, D. M. Wegner, S. T. Grafton, and C. N. Macrae. Wandering Minds: The Default Network and Stimulus-Independent Thought. *Science*, 315(5810):393–395, January 2007.

J. Maunsell and D. van Essen. The connections of the middle temporal visual area (MT) and their relationship to a cortical hierarchy in the macaque monkey. *The Journal of Neuroscience*, 3(12):2563 –2586, December 1983.

J. H. R. Maunsell and E. P. Cook. The role of attention in visual processing. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, 357(1424): 1063–1072, August 2002.

J. H. R. Maunsell, G. Sclar, T. A. Nealey, and D. D. DePriest. Extraretinal representations in area V4 in the macaque monkey. *Visual Neuroscience*, 7(06):561–573, 1991.

E. M. Maynard, N. G. Hatsopoulos, C. L. Ojakangas, B. D. Acuna, J. N. Sanes, R. A. Normann, and J. P. Donoghue. Neuronal Interactions Improve Cortical Population Coding of Movement Direction. *The Journal of Neuroscience*, 19(18):8083–8093, September 1999.

C. J. McAdams and J. H. R. Maunsell. Effects of Attention on Orientation-Tuning Functions of Single Neurons in Macaque Cortical Area V4. *The Journal of Neuroscience*, 19(1):431 –441, January 1999a.

C. J. McAdams and J. H. R. Maunsell. Effects of Attention on the Reliability of Individual Neurons in Monkey Visual Cortex. *Neuron*, 23(4):765–773, August 1999b.

C. J. McAdams and J. H. R. Maunsell. Attention to Both Space and Feature Modulates Neuronal Responses in Macaque Area V4. *Journal of Neurophysiology*, 83(3): 1751–1755, March 2000.

D. A. McCormick, B. W. Connors, J. W. Lighthall, and D. A. Prince. Comparative electrophysiology of pyramidal and sparsely spiny stellate neurons of the neocortex. *Journal of Neurophysiology*, 54(4):782–806, October 1985.

M. D. McDonnell and L. M. Ward. The benefits of noise in neural systems: bridging theory and experiment. *Nature Reviews Neuroscience*, 12(7):415–426, July 2011.

J. Mishra, J.-M. Fellous, and T. J. Sejnowski. Selective attention through phase relationship of excitatory and inhibitory input synchrony in a model cortical neuron. *Neural Networks*, 19(9):1329–1346, November 2006.

J. F. Mitchell, G. R. Stoner, and J. H. Reynolds. Object-based attention determines dominance in binocular rivalry. *Nature*, 429(6990):410–413, May 2004.

J. F. Mitchell, K. A. Sundberg, and J. H. Reynolds. Differential Attention-Dependent Response Modulation across Cell Classes in Macaque Visual Area V4. *Neuron*, 55 (1):131–141, July 2007.

J. F. Mitchell, K. A. Sundberg, and J. H. Reynolds. Spatial Attention Decorrelates Intrinsic Activity Fluctuations in Macaque Area V4. *Neuron*, 63(6):879–888, September 2009.

M. A. Montemurro, M. J. Rasch, Y. Murayama, N. K. Logothetis, and S. Panzeri. Phase-of-Firing Coding of Natural Visual Stimuli in Primary Visual Cortex. *Current Biology*, 18(5):375–380, March 2008.

T. Moore. The neurobiology of visual attention: finding sources. *Current Opinion in Neurobiology*, 16(2):159–165, April 2006.

T. Moore and K. M. Armstrong. Selective gating of visual signals by microstimulation of frontal cortex. *Nature*, 421(6921):370–373, January 2003.

T. Moore and M. Fallah. Control of eye movements and spatial attention. *Proceedings of the National Academy of Sciences*, 98(3):1273–1276, January 2001.

T. Moore and M. Fallah. Microstimulation of the Frontal Eye Field and Its Effects on Covert Spatial Attention. *Journal of Neurophysiology*, 91(1):152–162, January 2004.

J. Moran and R. Desimone. Selective attention gates visual processing in the extrastriate cortex. *Science*, 229(4715):782–784, 1985.

B. C. Motter. Focal Attention Produces Spatially Selective Processing in Visual Cortical Areas V1, V2, and V4 in the Presence of Competing Stimuli. *Journal of Neurophysiology*, 70(3):909–919, September 1993.

B. C. Motter. Neural correlates of attentive selection for color or luminance in extrastriate area V4. *The Journal of Neuroscience*, 14(4):2178–2189, April 1994a.

B. C. Motter. Neural correlates of feature selective memory and pop-out in extrastriate area V4. *The Journal of neuroscience*, 14(4):2190–2199, 1994b.

V. B. Mountcastle. The columnar organization of the neocortex. *Brain*, 120(4):701–722, April 1997.

V. B. Mountcastle, R. A. Andersen, and B. C. Motter. The influence of attentive fixation upon the excitability of the light- sensitive neurons of the posterior parietal

cortex. *The Journal of Neuroscience*, 1(11):1218–1225, November 1981.

V. B. Mountcastle, B. C. Motter, M. A. Steinmetz, and A. K. Sestokas. Common and differential effects of attentive fixation on the excitability of parietal and prestriate (V4) cortical visual neurons in the macaque monkey. *The Journal of neuroscience*, 7 (7):2239–2255, 1987.

V. B. Mountcastle. Modality and topographic properties of single neurons of cat's somatic sensory cortex. *J. neurophysiol*, 20(4):408–434, 1957.

J. R. Müller, M. G. Philiastides, and W. T. Newsome. Microstimulation of the superior colliculus focuses attention without moving the eyes. *Proceedings of the National Academy of Sciences of the United States of America*, 102(3):524–529, January 2005.

S. A. Neymotin, H. Lee, E. Park, A. A. Fenton, and W. W. Lytton. Emergence of Physiological Oscillation Frequencies in a Computer Model of Neocortex. *Frontiers in Computational Neuroscience*, 5, 2011.

R. Niebergall, P. S. Khayat, S. Treue, and J. C. Martinez-Trujillo. Expansion of MT Neurons Excitatory Receptive Fields during Covert Attentive Tracking. *The Journal of Neuroscience*, 31(43):15499–15510, October 2011.

E. Niebur and C. Koch. A model for the neuronal implementation of selective visual attention based on temporal correlation among neurons. *Journal of Computational Neuroscience*, 1(1-2):141–158, June 1994.

E. Niebur, C. Koch, and C. Rosin. An oscillation-based model for the neuronal basis of attention. *Vision Research*, 33(18):2789–2802, December 1993.

C. M. Niell and M. P. Stryker. Highly Selective Receptive Fields in Mouse Visual Cortex. *Journal of Neuroscience*, 28(30):7520–7536, July 2008.

B. Noudoost and T. Moore. Control of visual cortical signals by prefrontal dopamine. *Nature*, 474(7351):372–375, June 2011a.

B. Noudoost and T. Moore. The role of neuromodulators in selective attention. *Trends in Cognitive Sciences*, 15(12):585–591, December 2011b.

B. Noudoost, M. H. Chang, N. A. Steinmetz, and T. Moore. Top-down control of visual attention. *Current Opinion in Neurobiology*, 20(2):183–190, April 2010.

T. Nowotny, A. Szücs, R. Levi, and A. I. Selverston. Models Wagging the Dog: Are Circuits Constructed with Disparate Parameters? *Neural Computation*, 19(8):1985–2003, June 2007.

D. H. O'Connor, M. M. Fukui, M. A. Pinsk, and S. Kastner. Attention modulates responses in the human lateral geniculate nucleus. *Nat Neurosci*, 5(11):1203–1209, November 2002.

O. O. Oke, A. Magony, H. Anver, P. D. Ward, P. Jiruska, J. G. R. Jefferys, and M. Vreugdenhil. High-frequency gamma oscillations coexist with low-frequency gamma oscillations in the rat visual cortex in vitro. *European Journal of Neuroscience*, 31(8): 1435–1445, April 2010.

J. O'Keefe and M. L. Recce. Phase relationship between hippocampal place units and the EEG theta rhythm. *Hippocampus*, 3(3):317–330, July 1993.

M. W. Oram, P. Földiák, D. I. Perrett, M. W. Oram, and F. Sengpiel. The 'Ideal Homunculus': decoding neural population signals. *Trends in Neurosciences*, 21(6): 259–265, June 1998.

L. C. Osborne, S. E. Palmer, S. G. Lisberger, and W. Bialek. The Neural Basis for Combinatorial Coding in a Cortical Population Response. *The Journal of Neuroscience*, 28(50):13522–13531, December 2008.

K. Padmanabhan and N. N. Urban. Intrinsic biophysical diversity decorrelates neuronal firing while increasing information content. *Nature Neuroscience*, 13(10):1276–1282, October 2010.

S.-B. Paik, T. Kumar, and D. A. Glaser. Spontaneous Local Gamma Oscillation Selectively Enhances Neural Network Responsiveness. *PLoS Comput Biol*, 5(3):e1000342, April 2009.

A. Palmigiano, T. Geisel, F. Wolf, and D. Battaglia. Dynamic information routing at the edge of synchrony. *submitted*, 2015.

P. Patton, E. Thomas, and R. E. Wyatt. A computational model of vertical signal propagation in the primary visual cortex. *Biological Cybernetics*, 68(1):43–52, November 1992.

D. R. Patzwahl and S. Treue. Combining spatial and feature-based attention within the receptive field of MT neurons. *Vision Research*, 49(10):1188–1193, June 2009.

R. Perin, T. K. Berger, and H. Markram. A synaptic organizing principle for cortical neuronal groups. *Proceedings of the National Academy of Sciences*, 108(13):5419–5424, March 2011.

J. Poort and P. R. Roelfsema. Noise Correlations Have Little Influence on the Coding of Selective Attention in Area V1. *Cerebral Cortex*, 19(3):543–553, March 2009.

M. I. Posner, C. R. Snyder, and B. J. Davidson. Attention and the detection of signals. *Journal of experimental psychology*, 109(2):160–174, June 1980.

T. C. Potjans and M. Diesmann. The Cell-Type Specific Cortical Microcircuit: Relating Structure and Activity in a Full-Scale Spiking Network Model. *Cerebral Cortex*, 24 (3):785–806, March 2014.

A. Pouget, S. Deneve, J.-C. Ducom, and P. E. Latham. Narrow Versus Wide Tuning Curves: What's Best for a Population Code? *Neural Computation*, 11(1):85–90, January 1999.

P. Pouget, I. Stepniewska, E. A. Crowder, M. W. Leslie, E. E. Emeric, M. J. Nelson, J. D. Schall, P. Pouget, I. Stepniewska, E. A. Crowder, M. W. Leslie, E. E. Emeric, M. J. Nelson, and J. D. Schall. Visual and motor connectivity and the distribution of calcium-binding proteins in macaque frontal eye field: implications for saccade target selection. *Frontiers in Neuroanatomy*, 3:2, 2009.

W. H. Press. *Numerical recipes: the art of scientific computing*. Cambridge Univ. Press, Cambridge [u.a., 2002.

A. A. Prinz, C. P. Billimoria, and E. Marder. Alternative to Hand-Tuning Conductance-Based Models: Construction and Analysis of Databases of Model Neurons. *Journal of Neurophysiology*, 90(6):3998–4015, December 2003.

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

N. Qian and R. A. Andersen. Transparent motion perception as detection of unbalanced motion signals. II. Physiology. *The Journal of Neuroscience*, 14(12):7367–7380, December 1994.

M. I. Rabinovich, R. Huerta, P. Varona, and V. S. Afraimovich. Transient Cognitive Dynamics, Metastability, and Decision Making. *PLoS Comput Biol*, 4(5):e1000072, May 2008.

R. D. S. Raizada and S. Grossberg. Towards a Theory of the Laminar Architecture of Cerebral Cortex: Computational Clues from the Visual System. *Cerebral Cortex*, 13 (1):100–113, January 2003.

P. Rakic. Confusing cortical columns. *Proceedings of the National Academy of Sciences*, 105(34):12099–12100, August 2008.

M. J. Rasch, K. Schuch, N. K. Logothetis, and W. Maass. Statistical Comparison of Spike Responses to Natural Stimuli in Monkey Area V1 With Simulated Responses of a Detailed Laminar Network Model for a Patch of V1. *Journal of Neurophysiology*, 105(2):757–778, February 2011.

S. Ray and J. H. R. Maunsell. Differences in Gamma Frequencies across Visual Cortex Restrict Their Possible Use in Computation. *Neuron*, 67(5):885–896, September 2010.

G. H. Recanzone, R. H. Wurtz, and U. Schwarz. Responses of MT and MST neurons to one and two moving objects in the receptive field. *Journal of Neurophysiology*, 78 (6):2904–2915, 1997.

D. Ress, B. T. Backus, and D. J. Heeger. Activity in primary visual cortex predicts performance in a visual detection task. *Nature Neuroscience*, 3(9):940–945, September 2000.

J. H. Reynolds and D. J. Heeger. The Normalization Model of Attention. *Neuron*, 61 (2):168–185, January 2009.

J. H. Reynolds, L. Chelazzi, and R. Desimone. Competitive Mechanisms Subserve Attention in Macaque Areas V2 and V4. *The Journal of Neuroscience*, 19(5):1736 –1753, March 1999.

J. H. Reynolds, T. Pasternak, and R. Desimone. Attention Increases Sensitivity of V4 Neurons. *Neuron*, 26(3):703–714, June 2000.

D. L. Ringach, R. M. Shapley, and M. J. Hawken. Orientation Selectivity in Macaque V1: Diversity and Laminar Dependence. *The Journal of Neuroscience*, 22(13):5639–5651, July 2002.

M. Roberts, L. S. Delicato, J. Herrero, M. A. Gieselmann, and A. Thiele. Attention alters spatial integration in macaque V1 in an eccentricity-dependent manner. *Nature Neuroscience*, 10(11):1483–1491, November 2007.

M. Roberts, E. Lowet, N. Brunet, M. TerWal, P. Tiesinga, P. Fries, and P. DeWeerd. Robust Gamma Coherence between Macaque V1 and V2 by Dynamic Frequency Matching. *Neuron*, 78(3):523–536, May 2013.

D. L. Robinson and S. E. Petersen. The pulvinar and visual salience. *Trends in Neurosciences*, 15(4):127–132, April 1992.

P. R. Roelfsema, A. K. Engel, P. König, and W. Singer. Visuomotor integration is associated with zero time-lag synchronization among cortical areas. *Nature*, 385 (6612):157–161, January 1997.

P. R. Roelfsema, V. A. F. Lamme, and H. Spekreijse. Object-based attention in the primary visual cortex of the macaque monkey. *Nature*, 395(6700):376–381, September 1998.

P. R. Roelfsema, V. A. F. Lamme, and H. Spekreijse. Synchrony and covariation of firing rates in the primary visual cortex during contour grouping. *Nature Neuroscience*, 7(9):982–991, September 2004.

E. T. Rolls and A. Treves. The neuronal encoding of information in the brain. *Progress in Neurobiology*, 95(3):448–490, November 2011.

A. K. Roopun. Period concatenation underlies interactions between gamma and beta rhythms in neocortex. *Frontiers in Cellular Neuroscience*, 2, 2008a.

A. K. Roopun. Temporal interactions between cortical rhythms. *frontiers in Neuroscience*, 2(2):145–154, December 2008b.

A. K. Roopun, S. J. Middleton, M. O. Cunningham, F. E. N. LeBeau, A. Bibbig, M. A. Whittington, and R. D. Traub. A beta2-frequency (20–30 Hz) oscillation in nonsynaptic networks of somatosensory cortex. *Proceedings of the National Academy of Sciences*, 103(42):15646–15650, October 2006.

L. Rueckert and J. Grafman. Sustained attention deficits in pat ients with right frontal lesions. *Neuropsychologia*, 34(10):953–963, October 1996.

D. A. Ruff and M. R. Cohen. Attention can either increase or decrease spike count correlations in visual cortex. *Nature Neuroscience*, 17(11):1591–1597, November 2014a.

D. A. Ruff and M. R. Cohen. Global Cognitive Factors Modulate Correlated Response Variability between V4 Neurons. *The Journal of Neuroscience*, 34(49):16408–16416, December 2014b.

Y. B. Saalmann, I. N. Pigarev, and T. R. Vidyasagar. Neural Mechanisms of Visual Attention: How Top-Down Feedback Highlights Relevant Locations. *Science*, 316 (5831):1612–1615, June 2007.

Y. B. Saalmann, M. A. Pinsk, L. Wang, X. Li, and S. Kastner. The Pulvinar Regulates Information Transmission Between Cortical Areas Based on Attention Demands. *Science*, 337(6095):753–756, August 2012.

R. F. Salazar, N. M. Dotson, S. L. Bressler, and C. M. Gray. Content-Specific Fronto-Parietal Synchronization During Visual Working Memory. *Science*, 338(6110):1097–1100, November 2012.

E. Salinas and T. J. Sejnowski. Correlated neuronal activity and the flow of neural information. *Nature Reviews Neuroscience*, 2(8):539–550, August 2001.

J. N. Sanes and J. P. Donoghue. Oscillations in local field potentials of the primate motor cortex during voluntary movement. *Proceedings of the National Academy of Sciences*, 90(10):4470–4474, May 1993.

N. Santana, G. Mengod, and F. Artigas. Quantitative Analysis of the Expression of Dopamine D1 and D2 Receptors in Pyramidal and GABAergic Neurons of the Rat Prefrontal Cortex. *Cerebral Cortex*, 19(4):849–860, April 2009.

S. Saproo and J. T. Serences. Spatial Attention Improves the Quality of Population Codes in Human Visual Cortex. *Journal of Neurophysiology*, 104(2):885–895, August 2010.

R. J. Schafer and T. Moore. Selective Attention from Voluntary Control of Neurons in Prefrontal Cortex. *Science*, 332(6037):1568–1571, June 2011.

M. Schmidt, S. v. Albada, R. Bakker, and M. Diesmann. Integrating multi-scale data for a network model of macaque visual cortex. *BMC Neuroscience*, 14(Suppl 1): P111, July 2013.

T. Schreiber. Measuring information transfer. *Physical review letters*, 85(2):461–464, 2000.

D. Scott. *Multivariate Density Estimation: Theory, Practice, and Visualization*. John Wiley & Sons, Chicester, 1992.

T. J. Sejnowski, P. S. Churchland, and J. A. Movshon. Putting big data to good use in neuroscience. *Nature Neuroscience*, 17(11):1440–1441, November 2014.

M. Self, T. vanKerkoerle, H. Super, and P. Roelfsema. Distinct Roles of the Cortical Layers of Area V1 in Figure-Ground Segregation. *Current Biology*, 23(21):2121–2129, November 2013.

P. Seriès, P. E. Latham, and A. Pouget. Tuning curve sharpening for orientation selectivity: coding efficiency and the impact of correlations. *Nature Neuroscience*, 7 (10):1129–1135, October 2004.

M. N. Shadlen and J. A. Movshon. Synchrony Unbound: A Critical Evaluation of the Temporal Binding Hypothesis. *Neuron*, 24(1):67–77, September 1999.

M. N. Shadlen and W. T. Newsome. The Variable Discharge of Cortical Neurons: Implications for Connectivity, Computation, and Information Coding. *The Journal of Neuroscience*, 18(10):3870–3896, May 1998.

S. Shipp. The functional logic of cortico–pulvinar connections. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, 358(1438):1605–1624, October 2003.

S. Shipp. The brain circuitry of attention. *Trends in Cognitive Sciences*, 8(5):223–230, May 2004.

G. L. Shulman, J. M. Ollinger, E. Akbudak, T. E. Conturo, A. Z. Snyder, S. E. Petersen, and M. Corbetta. Areas Involved in Encoding and Applying Directional Expectations to Moving Objects. *The Journal of Neuroscience*, 19(21):9480–9496, November 1999.

G. L. Shulman, M. P. McAvoy, M. C. Cowan, S. V. Astafiev, A. P. Tansy, G. d'Avossa, and M. Corbetta. Quantitative Analysis of Attention and Detection Signals During Visual Search. *Journal of Neurophysiology*, 90(5):3384–3397, November 2003.

M. Siegel, T. H. Donner, R. Oostenveld, P. Fries, and A. K. Engel. Neuronal Synchronization along the Dorsal Visual Pathway Reflects the Focus of Spatial Attention. *Neuron*, 60(4):709–719, November 2008.

M. Siegel, M. R. Warden, and E. K. Miller. Phase-dependent neuronal coding of objects in short-term memory. *Proceedings of the National Academy of Sciences*, 106 (50):21341–21346, December 2009.

M. Siegel, T. H. Donner, and A. K. Engel. Spectral fingerprints of large-scale neuronal interactions. *Nature Reviews Neuroscience*, 13(2):121–134, February 2012.

J. H. Siegle, D. L. Pritchett, and C. I. Moore. Gamma-range synchronization of fast-spiking interneurons can enhance detection of tactile stimuli. *Nature Neuroscience*, 17(10):1371–1379, October 2014.

L. R. Silva, Y. Amitai, and B. W. Connors. Intrinsic oscillations of neocortex generated by layer 5 pyramidal neurons. *Science*, 251(4992):432–435, January 1991.

L. C. Sincich and J. C. Horton. The Circuitry of V1 and V2: Integration of Color, Form, and Motion. *Annual Review of Neuroscience*, 28(1):303–326, 2005.

W. Singer. Synchronization of Cortical Activity and its Putative Role in Information Processing and Learning. *Annual Review of Physiology*, 55(1):349–374, 1993.

G. D. Smith and S. Ebrahim. Data dredging, bias, or confounding. *BMJ : British Medical Journal*, 325(7378):1437–1438, December 2002.

M. A. Smith, X. Jia, A. Zandvakili, and A. Kohn. Laminar dependence of neuronal correlations in visual cortex. *Journal of Neurophysiology*, 109(4):940–947, February 2013.

R. J. Snowden, S. Treue, R. G. Erickson, and R. A. Andersen. The response of area MT and V1 neurons to transparent motion. *The Journal of Neuroscience*, 11(9):2768–2785, September 1991.

S. Song, P. J. Sjöström, M. Reigl, S. Nelson, and D. B. Chklovskii. Highly Nonrandom Features of Synaptic Connectivity in Local Cortical Circuits. *PLoS Biol*, 3(3):e68, March 2005.

E. Spaak, M. Bonnefond, A. Maier, D. Leopold, and O. Jensen. Layer-Specific Entrainment of Gamma-Band Neural Activity by the Alpha Rhythm in Monkey Visual Cortex. *Current Biology*, 22(24):2313–2318, December 2012.

H. Spitzer, R. Desimone, and J. Moran. Increased attention enhances both behavioral and neuronal performance. *Science*, 240(4850):338–340, April 1988.

O. Sporns, G. Tononi, and G. M. Edelman. Theoretical neuroanatomy and the connectivity of the cerebral cortex. *Behavioural Brain Research*, 135(1–2):69–74, September 2002.

O. Sporns and R. Kötter. Motifs in Brain Networks. *PLoS Biol*, 2(11):e369, October 2004.

O. Sporns and J. D. Zwi. The small world of the cerebral cortex. *Neuroinformatics*, 2 (2):145–162, June 2004.

O. Sporns, G. Tononi, and R. Kötter. The Human Connectome: A Structural Description of the Human Brain. *PLoS Comput Biol*, 1(4):e42, September 2005.

R. F. Squire, B. Noudoost, R. J. Schafer, and T. Moore. Prefrontal Contributions to Visual Selective Attention. *Annual Review of Neuroscience*, 36(1):451–466, 2013.

H. Stanislaw and N. Todorov. Calculation of signal detection theory measures. *Behavior research methods, instruments, & computers*, 31(1):137–149, 1999.

I. H. Stevenson and K. P. Kording. How advances in neural recording affect data analysis. *Nature Neuroscience*, 14(2):139–142, February 2011.

I. H. Stevenson, B. M. London, E. R. Oby, N. A. Sachs, J. Reimer, B. Englitz, S. V. David, S. A. Shamma, T. J. Blanche, K. Mizuseki, A. Zandvakili, N. G. Hatsopoulos, L. E. Miller, and K. P. Kording. Functional Connectivity and Tuning Curves in Populations of Simultaneously Recorded Neurons. *PLoS Comput Biol*, 8(11): e1002775, November 2012.

R. Stoop, V. Saase, C. Wagner, B. Stoop, and R. Stoop. Beyond Scale-Free Small-World Networks: Cortical Columns for Quick Brains. *Physical Review Letters*, 110 (10):108105, March 2013.

W. Sun and Y. Dan. Layer-specific network oscillation and spatiotemporal receptive field in the visual cortex. *Proceedings of the National Academy of Sciences*, 106(42): 17986–17991, October 2009.

N. V. Swindale. Is the cerebral cortex modular? *Trends in Neurosciences*, 13(12):487–492, December 1990.

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

L. S. Tang, M. L. Goeritz, J. S. Caplan, A. L. Taylor, M. Fisek, and E. Marder. Precise Temperature Compensation of Phase in a Rhythmic Motor Pattern. *PLoS Biol*, 8 (8):e1000469, August 2010.

A. L. Taylor, J.-M. Goaillard, and E. Marder. How Multiple Conductances Determine Electrophysiological Properties in a Multicompartment Model. *The Journal of Neuroscience*, 29(17):5573–5586, April 2009.

K. Taylor, S. Mandon, W. A. Freiwald, and A. K. Kreiter. Coherent Oscillatory Activity in Monkey Area V4 Predicts Successful Allocation of Attention. *Cerebral Cortex*, 15(9):1424–1437, September 2005.

A. Thiele, A. Pooresmaeili, L. S. Delicato, J. L. Herrero, and P. R. Roelfsema. Additive Effects of Attention and Stimulus Contrast in Primary Visual Cortex. *Cerebral Cortex*, 19(12):2970–2981, December 2009.

E. Thomas, P. Patton, and R. E. Wyatt. A computational model of the vertical anatomical organization of primary visual cortex. *Biological Cybernetics*, 65(3):189–202, July 1991.

K. G. Thompson, K. L. Biscoe, and T. R. Sato. Neuronal Basis of Covert Spatial Attention in the Frontal Eye Field. *The Journal of Neuroscience*, 25(41):9479–9487, October 2005.

A. M. Thomson and A. P. Bannister. Interlaminar Connections in the Neocortex. *Cerebral Cortex*, 13(1):5–14, January 2003.

A. M. Thomson and O. T. Morris. Selectivity in the inter-laminar connections made by neocortical neurones. *Journal of Neurocytology*, 31(3-5):239–246, March 2002.

A. M. Thomson, D. C. West, Y. Wang, and A. P. Bannister. Synaptic Connections and Small Circuits Involving Excitatory and Inhibitory Neurons in Layers 2–5 of Adult Rat and Cat Neocortex: Triple Intracellular Recordings and Biocytin Labelling In Vitro. *Cerebral Cortex*, 12(9):936–953, September 2002.

P. Tiesinga and T. J. Sejnowski. Cortical Enlightenment: Are Attentional Gamma Oscillations Driven by ING or PING? *Neuron*, 63(6):727–732, September 2009.

P. H. Tiesinga, J.-M. Fellous, E. Salinas, J. V. José, and T. J. Sejnowski. Inhibitory synchrony as a mechanism for attentional gain modulation. *Journal of Physiology-Paris*, 98(4–6):296–314, July 2004.

P. H. E. Tiesinga. Stimulus Competition by Inhibitory Interference. *Neural Computation*, 17(11):2421–2453, November 2005.

P. Tiesinga and T. J. Sejnowski. Rapid Temporal Modulation of Synchrony by Competition in Cortical Interneuron Networks. *Neural Computation*, 16(2):251–275, February 2004.

G. Tkačik, J. S. Prentice, V. Balasubramanian, and E. Schneidman. Optimal population coding by noisy spiking neurons. *Proceedings of the National Academy of Sciences*, 107(32):14419–14424, August 2010.

G. Tkačik, O. Marre, D. Amodei, E. Schneidman, W. Bialek, and M. J. Berry, II. Searching for Collective Behavior in a Large Network of Sensory Neurons. *PLoS Comput Biol*, 10(1):e1003408, January 2014.

J. J. Todd, D. Fougnie, and R. Marois. Visual Short-Term Memory Load Suppresses Temporo-Parietal Junction Activity and Induces Inattentional Blindness. *Psychological Science*, 16(12):965–972, December 2005.

E. Tognoli and J. A. S. Kelso. The Metastable Brain. *Neuron*, 81(1):35–48, January 2014.

D. J. Tolhurst, J. A. Movshon, and A. F. Dean. The statistical reliability of signals in single neurons in cat and monkey visual cortex. *Vision Research*, 23(8):775–785, 1983.

A. B. L. Tort, R. W. Komorowski, J. R. Manns, N. J. Kopell, and H. Eichenbaum. Theta–gamma coupling increases during the learning of item–context associations. *Proceedings of the National Academy of Sciences*, 106(49):20942–20947, December 2009.

R. D. Traub, D. Contreras, M. O. Cunningham, H. Murray, F. E. N. LeBeau, A. Roopun, A. Bibbig, W. B. Wilent, M. J. Higley, and M. A. Whittington. Single-Column Thalamocortical Network Model Exhibiting Gamma Oscillations, Sleep Spindles, and Epileptogenic Bursts. *Journal of Neurophysiology*, 93(4):2194–2232, April 2005.

S. Treue. Visual attention: the where, what, how and why of saliency. *Current Opinion in Neurobiology*, 13(4):428–432, August 2003.

S. Treue and S. Katzner. Visual attention: of features and transparent surfaces. *Trends in Cognitive Sciences*, 11(11):451–453, November 2007.

S. Treue and J. H. R. Maunsell. Attentional modulation of visual motion processing in cortical areas MT and MST. *Nature*, 382(6591):539–541, 1996.

S. Treue and J. H. R. Maunsell. Effects of Attention on the Processing of Motion in Macaque Middle Temporal and Medial Superior Temporal Visual Cortical Areas. *The Journal of Neuroscience*, 19(17):7591–7602, September 1999.

S. Treue and J. C. M. Trujillo. Feature-based attention influences motion processing gain in macaque visual cortex. *Nature*, 399(6736):575–579, June 1999.

S. Treue, K. Hol, and H.-J. Rauber. Seeing multiple directions of motion - physiology and psychophysics. *Nat Neurosci*, 3(3):270–276, March 2000.

M. Tsodyks, T. Kenet, A. Grinvald, and A. Arieli. Linking Spontaneous Activity of Single Cortical Neurons and the Underlying Functional Architecture. *Science*, 286 (5446):1943–1946, December 1999.

J. J. Tyson, K. C. Chen, and B. Novak. Sniffers, buzzers, toggles and blinkers: dynamics of regulatory and signaling pathways in the cell. *Current Opinion in Cell Biology*, 15(2):221–231, April 2003.

N. Ulanovsky, L. Las, D. Farkas, and I. Nelken. Multiple Time Scales of Adaptation in Auditory Cortex Neurons. *The Journal of Neuroscience*, 24(46):10440–10453, November 2004.

E. Vaadia, I. Haalman, M. Abeles, H. Bergman, Y. Prut, H. Slovin, and A. Aertsen. Dynamics of neuronal interactions in monkey cortex in relation to behavioural events. *Nature*, 373(6514):515–518, February 1995.

M. P. van den Heuvel and O. Sporns. Rich-Club Organization of the Human Connectome. *The Journal of Neuroscience*, 31(44):15775–15786, November 2011.

M. P. van den Heuvel and O. Sporns. An Anatomical Substrate for Integration among Functional Networks in Human Cortex. *The Journal of Neuroscience*, 33(36): 14489–14500, September 2013a.

M. P. van den Heuvel and O. Sporns. Network hubs in the human brain. *Trends in Cognitive Sciences*, 17(12):683–696, December 2013b.

M. Vinck, B. Lima, T. Womelsdorf, R. Oostenveld, W. Singer, S. Neuenschwander, and P. Fries. Gamma-Phase Shifting in Awake Monkey Visual Cortex. *The Journal of Neuroscience*, 30(4):1250–1257, January 2010.

M. Vinck, T. Womelsdorf, E. Buffalo, R. Desimone, and P. Fries. Attentional Modulation of Cell-Class-Specific Gamma-Band Synchronization in Awake Monkey Area V4. *Neuron*, 80(4):1077–1089, November 2013.

M. Volgushev, M. Chistiakova, and W. Singer. Modification of discharge patterns of neocortical neurons by induced oscillations of the membrane potential. *Neuroscience*, 83(1):15–25, 1998.

A. von Stein, C. Chiang, and P. König. Top-down processing mediated by interareal synchronization. *Proceedings of the National Academy of Sciences*, 97(26):14748–14753, December 2000.

C. v. Vreeswijk and H. Sompolinsky. Chaos in Neuronal Networks with Balanced Excitatory and Inhibitory Activity. *Science*, 274(5293):1724–1726, December 1996.

N. Wagatsuma, T. C. Potjans, M. Diesmann, and T. Fukai. Layer-dependent attentional processing by top-down signals in a visual cortical microcircuit model. *Frontiers in Computational Neuroscience*, 5:31, 2011.

N. Wagatsuma, T. C. Potjans, M. Diesmann, K. Sakai, and T. Fukai. Spatial and Feature-Based Attention in a Layered Cortical Microcircuit Model. *PLoS ONE*, 8 (12):e80788, December 2013.

X.-J. Wang. Synaptic reverberation underlying mnemonic persistent activity. *Trends in Neurosciences*, 24(8):455–463, August 2001.

X.-J. Wang. Neurophysiological and Computational Principles of Cortical Rhythms in Cognition. *Physiological Reviews*, 90(3):1195–1268, July 2010.

A. Wannig, V. Rodríguez, and W. A. Freiwald. Attention to Surfaces Modulates Motion Processing in Extrastriate Area MT. *Neuron*, 54(4):639–651, May 2007.

M. A. Whittington, R. D. Traub, N. Kopell, B. Ermentrout, and E. H. Buhl. Inhibition-based rhythms: experimental and mathematical observations on network dynamics. *International Journal of Psychophysiology*, 38(3):315–336, 2000.

T. Williford and J. H. R. Maunsell. Effects of Spatial Attention on Contrast Response Functions in Macaque Area V4. *Journal of Neurophysiology*, 96(1):40–54, July 2006.

H. R. Wilson and J. D. Cowan. Excitatory and Inhibitory Interactions in Localized Populations of Model Neurons. *Biophysical Journal*, 12(1):1–24, January 1972.

T. Womelsdorf, P. Fries, P. P. Mitra, and R. Desimone. Gamma-band synchronization in visual cortex predicts speed of change detection. *Nature*, 439(7077):733–736, February 2006.

T. Womelsdorf, J.-M. Schoffelen, R. Oostenveld, W. Singer, R. Desimone, A. K. Engel, and P. Fries. Modulation of Neuronal Interactions Through Neuronal Synchronization. *Science*, 316(5831):1609–1612, June 2007.

K.-F. Wong and X.-J. Wang. A Recurrent Network Mechanism of Time Integration in Perceptual Decisions. *The Journal of Neuroscience*, 26(4):1314–1328, January 2006.

F. Wörgötter and U. T. Eysel. Quantitative determination of orientational and directional components in the response of visual cortical cells to moving stimuli. *Biological Cybernetics*, 57(6):349–355, December 1987.

D. Xing, C.-I. Yeh, and R. M. Shapley. Spatial Spread of the Local Field Potential and its Laminar Variation in Visual Cortex. *The Journal of Neuroscience*, 29(37): 11540–11549, September 2009.

D. Xing, Y. Shen, S. Burns, C.-I. Yeh, R. Shapley, and W. Li. Stochastic Generation of Gamma-Band Activity in Primary Visual Cortex of Awake and Anesthetized Monkeys. *The Journal of Neuroscience*, 32(40):13873–13880a, October 2012a.

D. Xing, C.-I. Yeh, S. Burns, and R. M. Shapley. Laminar analysis of visually evoked activity in the primary visual cortex. *Proceedings of the National Academy of Sciences*, 109(34):13871–13876, August 2012b.

X. S. Yang and S. Deb. Engineering optimisation by cuckoo search. *International Journal of Mathematical Modelling and Numerical Optimisation*, 1(4):330–343, 2010.

J. Yu and D. Ferster. Membrane Potential Synchrony in Primary Visual Cortex during Sensory Stimulation. *Neuron*, 68(6):1187–1201, December 2010.

M. Zeitler, P. Fries, and S. Gielen. Biased competition through variations in amplitude of -oscillations. *Journal of Computational Neuroscience*, 25(1):89–107, August 2008.

K. Zhang and T. J. Sejnowski. Neuronal Tuning: To Sharpen or Broaden? *Neural Computation*, 11(1):75–84, January 1999.

C. Zheng and L. Colgin. Beta and Gamma Rhythms Go with the Flow. *Neuron*, 85 (2):236–237, January 2015.

D. Zoccolan, D. D. Cox, and J. J. DiCarlo. Multiple Object Response Normalization in Monkey Inferotemporal Cortex. *The Journal of Neuroscience*, 25(36):8150 –8164, 2005.

E. Zohary, M. N. Shadlen, and W. T. Newsome. Correlated neuronal discharge rate and its implications for psychophysical performance. *Nature*, 370(6485):140–143, July 1994.

A. Zygmund. *Trigonometric Series*. Cambridge University Press, 2002.

A. Zénon and R. J. Krauzlis. Attention deficits without cortical neuronal deficits. *Nature*, 489(7416):434–437, September 2012.

E. D. Übeyli. Adaptive Neuro-Fuzzy Inference Systems for Automatic Detection of Breast Cancer. *Journal of Medical Systems*, 33(5):353–358, October 2009.

ACKNOWLEDGMENTS

First of all, I am extremely grateful to Theo Geisel for making it all possible: for having given me the opportunity to write this thesis, for his support during the process, for providing an institute with immense freedom, for enabling my extended stay in Marseille.

I would also like to sincerely thank the other two members of my thesis committee, Fred Wolf and Stefan Treue, for their advice and comments during and beyond these meetings.

Demian Battaglia helped and guided me through this endeavour from the beginning, he introduced me to the ideas that led to this thesis, revised and refined them with me, always had a good advice, and showed amazing patience in the process. Moreover, I am indebted to him for his invaluable comments on this manuscript. Thank you very much.

In addition, I acknowledge, most gratefully, very insightful discussions with Manuel Schottdorf, Viola Priesemann, Rainer Enkelgken, Annette Witt, Timothy Proix and Enrique Hansen.

I would like to sincerely thank Ayse Bolik, Viktoryia Novak, Regina Wunderlich and Barbara Guichemer for their responsivity, helpfulness and patience in all administrative matters. Likewise, I am grateful to our former and current project managers Tobias Niemann, Kerstin Mosch and Yvonne Reimannn for all their work behind the scenes. I am also deeply indebted to the computer clusters in our department who did an immense amount of work for me (and the trees that, sadly, had to die in the process). Correspondingly, I am most grateful to Yorck-Fabian Beensen and Denny Fliegner who kept the desktop and cluster computers in such a marvellous shape and always had the right advice in case these were refusing to let me work.

STATEMENT OF ORIGINALITY

I hereby declare that I prepared this disseration entitled *Attention: A Complex System* on my own and with no other sources and aids than quoted.

Göttingen, June 2015

Markus Helmer