

Rapid long-range disynaptic inhibition explains the formation of cortical orientation maps.

Ján Antolík^{1,*}

¹ Unité de Neurosciences, Information et Complexité, CNRS UPR 3293, Gif-sur-Yvette, France

Correspondence*:

Ján Antolík
antolikjan@gmail.com

2 ABSTRACT

3

4 Competitive interactions are believed to underlie many types of cortical processing, ranging
5 from memory formation, attention and development of cortical functional organization (e.g.
6 development of orientation maps in primary visual cortex). In the latter case, the competitive
7 interactions happen along the cortical surface, with local populations of neurons reinforcing each
8 other while competing with those displaced more distally. This specific configuration of lateral
9 interactions is however in stark contrast with the known properties of the anatomical substrate,
10 i.e. excitatory connections (mediating reinforcement) having longer reach than inhibitory ones
11 (mediating competition). No satisfactory biologically plausible resolution of this conflict between
12 anatomical measures, and assumed cortical function has been proposed. Recently a specific
13 pattern of delays between different types of neurons in cat cortex has been discovered, where
14 direct mono-synaptic excitation has approximately the same delay, as the combined delays of
15 the disynaptic inhibitory interactions between excitatory neurons (i.e. the sum of delays from
16 excitatory to inhibitory and from inhibitory to excitatory neurons). Here we show that this specific
17 pattern of delays represents a biologically plausible explanation for how short-range inhibition
18 can support competitive interactions that underlie the development of orientation maps in primary
19 visual cortex. We demonstrate this statement analytically under simplifying conditions, and
20 subsequently show using network simulations that development of orientation maps is preserved
21 when long-range excitation, direct inhibitory to inhibitory interactions, and moderate inequality in
22 the delays between excitatory and inhibitory pathways is added.

23 **Keywords:** primary visual cortex, rate model, cortical functional development, Hebbian learning, cortical horizontal connectivity,
24 orientation map

25 None

1 INTRODUCTION

Competition between populations of neurons has been proposed as one of the canonical computations of cortical networks, and has been hypothesized to underly a range of brain functions including working memory (Amit and Brunel, 1995; Durstewitz et al., 2000), orientation tuning (Somers et al., 1995; Ben-Yishai et al., 1995), and functional map development (von der Malsburg, 1973; Miikkulainen et al., 2005; Antolík and Bednar, 2011). In developmental models of functional cortical organization the competition occurs between populations of neurons spatially offset along the cortical surface, whereby local populations mutually reinforce each other via excitatory connections (short-range excitation) while long-range inhibition facilitates competition between the local populations and stabilizes the activity in the network. Such so called Mexican-hat arrangement of recurrent interactions (figure 1A), however, is in stark contrast to the known anatomical arrangement of cortical circuitry: the excitatory neurons (especially in superficial layers) tend to form long-range arborizations spanning multiple columns while the axons of a majority of inhibitory neurons are confined to an area only several hundreds of micrometers in diameter (Buzás et al., 2006; Budd and Kisvárdy, 2001) (figure 1B).

As pointed out by Muir *et al.* in their recent study (Muir and Cook, 2014), to solve this apparent conflict, previous topologically organized models of cortical competition have either relied on the anatomically unsupported Mexican-hat profile of lateral interactions (von der Malsburg, 1973; Miikkulainen et al., 2005), or relied on other biologically unrealistic properties, such as selective targeting of inhibitory neurons by long range excitatory connections (Law, 2009; Rutishauser et al., 2012) or instantaneous synaptic transmission coupled with omission of recurrent inhibition (Kang et al., 2003; Levy and Reyes, 2011; Grabska-Barwinska and von der Malsburg, 2008).

A common hypothesis about how to reconcile Mexican-hat lateral interactions with anatomical reality is that the range of the effective inhibitory influence of an excitatory neuron onto other excitatory neurons (via the disynaptic pathway from excitatory to inhibitory to excitatory neurons - the disynaptic inhibition) will correspond to the cumulative reach of the direct excitatory and the direct inhibitory connections. The effective inhibitory interactions between excitatory neurons will thus have longer range than the direct excitatory connections, supporting the use of Mexican-hat lateral interactions in population models (i.e. models without explicit separation of excitatory and inhibitory populations). Even though very intuitive, this explanation has never been explicitly demonstrated, and it omits the fact that under the reasonable null hypothesis of equal transmission delays for all connections, the recurrent disynaptic inhibition will lag the direct recurrent excitation. As it turns out this is crucial, as recent model analysis by Muir *et al.* (Muir and Cook, 2014) shows that, under the assumption of uniform transmission delays, the presence of competition across the cortical surface is predicted well by the anatomy of direct excitatory and inhibitory coupling and that multi-synaptic network effects are negligible, effectively rejecting the disynaptic explanation behind Mexican-hat interaction. In conclusion, currently no satisfactory explanation of how topological functional organization develops in cortical networks that is consistent with the present anatomical findings exists.

As we will show in this study, the nature of the transmission delays between different neural type (excitatory,inhibitory) combinations holds the key to resolving this long standing open question. A recent study by Ohana *et al.* (Ohana et al., 2012) revealed a specific pattern of transmission delays between different neural type combinations. Specifically they found that the excitatory to excitatory connections are slow while the excitatory to inhibitory connections are fast. In this study we explore the possibility that this specific transmission delay pattern is the missing link that can explain how short range inhibition can lead to effective cortical competition. We employ computational models to show that the fast excitatory-to-inhibitory-to-excitatory pathway allows for the disynaptic inhibition to generate the effective Mexican-hat

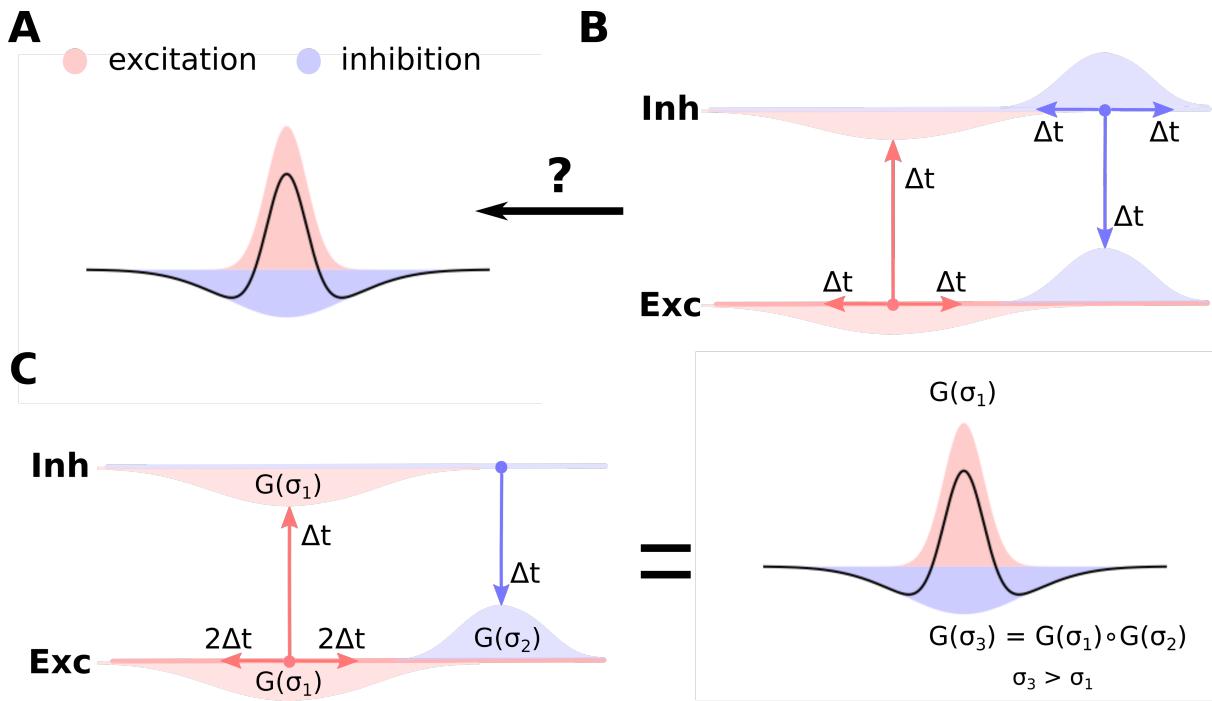


Figure 1. Cortical anatomy and effective lateral interactions. (A) Short-range excitation (red) and long-range inhibition (blue) leads to effective Mexican-hat lateral interactions (black curve). (B) Anatomical evidence indicates the opposite organization of lateral connectivity in cortex, whereby excitatory neurons send long-range connections to other neurons, while inhibitory neurons only local ones. Furthermore, uniform delays across the intra-cortical projections is typically assumed. It is not clear how such anatomical configuration can support the Mexican-hat effective interaction across cortical surface. (C) Under the assumption of slow excitatory-to-excitatory and fast excitatory-to-inhibitory and inhibitory-to-excitatory pathway (and disregarding the inhibitory-to-inhibitory interactions), the effective lateral inhibitory interactions correspond to the convolution of the excitatory-to-inhibitory and inhibitory-to-excitatory connection kernels. Under the assumption that these kernels are Gaussian ($G(\sigma)$ in the figure), the effective inhibitory interactions will be Gaussian with variance (space-constant) greater than the excitatory-to-excitatory one, thus forming Mexican-hat lateral interaction profile.

69 like lateral interactions (figure 1C), thus for the first time explaining competition in topologically organized
 70 cortical networks with no biologically implausible assumptions.

71 To demonstrate the proposed implementation of competitive interactions on a specific feature of the
 72 V1 processing, in this study we focus on models of development of functional cortical organization.
 73 Functional properties of neurons in primary visual cortex, such orientation, color and frequency preference
 74 is not randomly distributed along the cortical surface but rather form smooth topological maps (Swindale,
 75 1996; Goodhill, 2007), whereby nearby neurons prefer similar stimulus features. Such smooth topological
 76 mapping of function on cortical surface is ubiquitous throughout the cortex (Huth et al., 2016), found in
 77 multiple species, and implicated in a number of functional properties of V1, and of cortex in general. Here
 78 we restrict our attention to orientation preference, as it is the most well explored example of competition
 79 driven functional cortical organization development, but these results generalize to development of other
 80 cortical topological properties, as well as potentially other competition based computations.

2 MATERIALS AND METHODS

81 Here we describe the two models and their variants used in this study and finish with a description of a
82 measure for assessing the extent to which model orientation maps resemble their biological counterparts.
83 Since this study heavily leans on methodology developed in our previous studies, we offer here a short
84 description and refer readers to the original articles for the details.

85 2.1 GCAL model

86 In sections 3.1 through 3.3 (model 1, 2 and 3; see figure 3) we use the GCAL model (Stevens et al.,
87 2013) which is the most advanced variant of the LISSOM (Laterally Interconnected Synergetically Self-
88 Organizing Map) algorithm introduced by Miikkulainen *et al.* (Miikkulainen et al., 2005), which itself
89 is based on earlier Self-Organizing Map models (Kohonen, 1982). Several mechanistic explanations of
90 how orientation maps can develop in primary visual cortex have been proposed, but most, including the
91 LISSOM family of models, involve two key ingredients: (1) stimulus driven Hebbian learning on the
92 thalamo-cortical synapses, that ensures the formation of afferent connectivity pattern inducing Gabor
93 like RFs (and consequently orientation, frequency and phase preference) to V1 neural units; and (2) a
94 Mexican-hat-like effective lateral interactions within the cortical population of neural units, that induce
95 co-activation among proximate units while competition between more distal units.

96 To understand how these two mechanisms lead to development of orientation maps, consider first the
97 initial state of the model with isotropic afferent connectivity. When stimulus is presented to the model
98 for the first time, the activity of the cortical population will after few simulation steps settle into a pattern
99 that can be best described as random placement of 'blobs', whereby nearby neurons tend to be either
100 co-active or silent. At this point, the placement of the local activity centers (the 'blobs') is essentially
101 random, determined by the interplay of the competitive influence of the lateral interactions with whatever
102 source of variability present in the model (e.g. random noise in the initial connections or content of the
103 stimulus). The Hebbian learning will ensure, that all neurons within the active local populations will adjust
104 their afferent weights slightly towards the stimulus pattern appearing within their RFs. They will thus in
105 subsequent iterations be slightly more activated by similar patterns.

106 This iterative process of stimulus presentation, activity pattern formation and Hebbian adjustment of
107 the thalamo-cortical weights will keep repeating. However note, that in the subsequent iterations, neurons
108 whose afferent weights will be more similar with the currently presented stimulus falling within their
109 RF will tend to be more active than those whose RF at that point differs from the presented stimulus.
110 Importantly even small differences of the initial activations will be magnified by the lateral competitive
111 interactions thus overall driving nearby neurons to over time develop sharp selectivity for similar stimulus
112 features while more distal neurons will be driven to develop selectivity to other stimulus features. Which
113 features will be mapped onto cortical surface will depend on the exact statistics of the stimuli shown during
114 training. It turns out that if the stimuli are natural images, or other artificial stimuli with strong oriented
115 components, smooth representation of orientation preference across the cortical surface (the orientation
116 maps) will develop. In this study we will focus on this specific feature (orientation), as it is the most
117 well studied one, but it is important to emphasize that given the right stimulation conditions, multiplexed
118 representation of other features, such as ocular dominance, disparity, spatial frequency and others can
119 develop (Miikkulainen et al., 2005).

120 In the reminder of this section we will describe one specific implementation of model with overall
121 dynamics following the broad description outlined in the previous paragraphs. For clarity and consistency

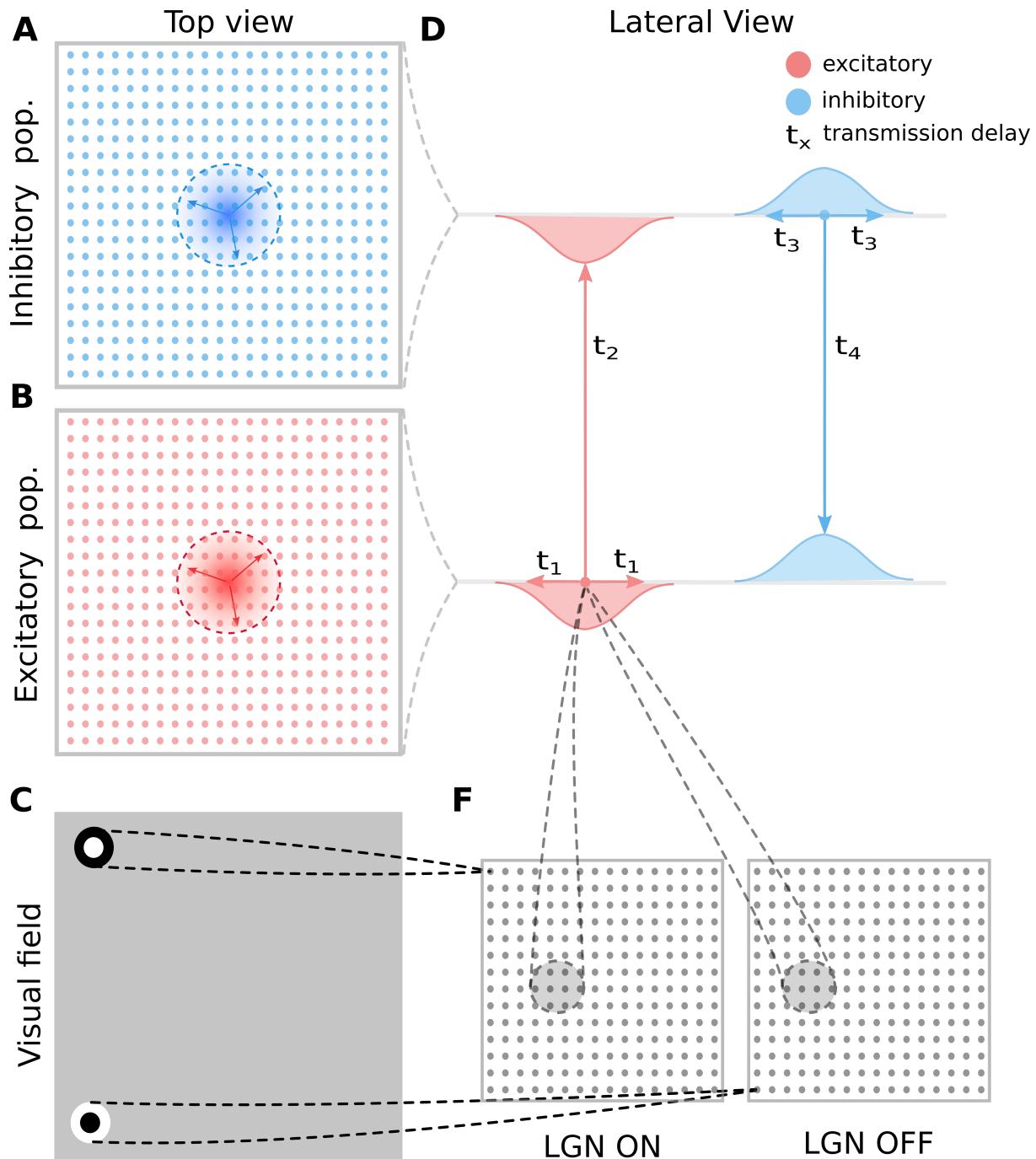


Figure 2. General model architecture. (A) The inhibitory cortical population corresponds to a regular lattice of units in cortical space. (B) As in A but for the excitatory population. (C) Thalamic neurons are modeled as simple difference-of-Gaussian filters followed by a threshold-linear transfer function. (D) The intra-cortical connectivity between the excitatory and inhibitory populations and their transmission delays. The intra-cortical connectivity is the only aspect of the general architecture presented here that changes between the different model variants explored in this study and each variant is detailed in figure 3. (F) The RF centers of the LGN neurons form a regular lattice across the visual space covered by the model. This retinotopic mapping of connection fields between layers of the model is also maintained in the thalamo-cortical projection (lines between D and F). These thalamo-cortical projections are the only connections in the model that undergo Hebbian adaptation during simulated visually driven development.

122 the following model description closely follows the methodology sections in our previous work (Stevens
 123 et al., 2013; Antolík and Bednar, 2011). The architecture of the 3 GCAL variants presented in this paper is
 124 depicted in figure 2 and figure 3A-C. The models are implemented in the Topographica simulator, freely
 125 available at www.topographica.org. Each of the GCAL models consists of a set of sheets of model neural
 126 units, corresponding to: (1) the photo-receptors, (2) the combined effects of the retinal ganglion cells (RGC)
 127 and ON and OFF LGN cells. Furthermore each model has one V1 sheet of units with direct excitatory
 128 and inhibitory interactions (model 1; see figure 3), or two V1 sheets, one corresponding to excitatory and
 129 one to inhibitory neurons (model 2 and 3; see figure 3). The model sheet is a 2D array of computational
 130 elements (called units or, loosely, neurons), with activation and plasticity equations as described below and
 131 referenced by a coordinate system we will refer to as sheet coordinates, where the center of the sheet is
 132 designated (0.0,0.0). The number of units simulated in each sheet is determined by the density of units
 133 per unit length in both sheet dimensions. All cortical sheets have nominal dimensions 1.0×1.0 in sheet
 134 coordinates. The sizes of the RGC/LGN (1.5×1.5) and photo-receptor (3.5×3.5) sheets were chosen to
 135 ensure that each unit in the receiving sheet has a complete set of connections, thus minimizing edge effects
 136 in the RGC/LGN and V1 sheets. The density of units per 1.0×1.0 area is 96×96 for the photo-receptor and
 137 RGC/LGN ON and OFF sheets, and 96×96 for both cortical sheets.

138 2.1.1 Simulation run-time and stimuli

139 As a simplification, GCAL ignores the detailed temporal properties of the sub-cortical neural responses
 140 and of signal propagation along the various types of connections. Instead, the model ON/OFF units have a
 141 constant, sustained output, and all connections have a constant delay, independent of the physical length of
 142 that connection. The simulator operates in discrete time steps. Retinal input changes every 16 time steps
 143 (and during this period is kept constant), and therefore afferent inputs to the V1 sheet(s) are effectively
 144 updated every 16 steps. This process is a discrete simulation of an otherwise continuous series of changes
 145 in membrane potential due to incoming spikes and consequent generation of spikes. One such training
 146 iteration (16 steps) in the model represents one visual fixation i.e., an iteration consists of a constant retinal
 147 activation, followed by processing at the ON/OFF and cortical levels.

148 The activation value ψ_i of unit i in the photo-receptor sheet (P) is given by the gray-scale value in the
 149 chosen image at that point. In this study we use two-dimensional elongated Gaussian patterns, whose
 150 center coordinates and orientation are sampled from a uniform distribution that covers the area of the photo-
 151 receptor sheet and the full range of orientations. Every 16 time steps 2 Gaussian patterns are superimposed
 152 to form the visual input.

153 2.1.2 The LGN/RGC ON and OFF sheets

154 The ON/OFF units are called RGC/LGN units because they represent the entire pathway between the
 155 retinal photo-receptors and V1, including the retinal ganglion cells, LGN cells, and the connection pathways.
 156 The activation level for a unit at position j in an RGC/LGN sheet at time t is defined as:

$$\eta_j(t) = f \left(\frac{A_j(t)}{c + \gamma_L \sum_k A_k(t) l_{kj}} \right) \quad (1)$$

157 where

$$A_j(t) = \gamma_F \sum_{i \in F_j} \Psi_i(t) \omega_{ij} \quad (2)$$

158 The activation function f is a half-wave rectifying function that ensures positive activation values, constant
 159 $\gamma_F = 1.5$ defines the overall strength of the afferent connections from the retina, constant $\gamma_L = 0.6$ defines
 160 the strength of the lateral connections within the RGC/LGN sheet, constant $c = 0.11$ defines the slope of
 161 the gain, Ψ_i is the activation of unit i taken from the set of photo-receptors from which RGC/LGN unit
 162 j receives input (its connection field F_j), ω_{ij} is the connection weight from unit i in the retina to unit j
 163 in the RGC/LGN, and l_{kj} is the lateral connection weight from RGC/LGN unit k to RGC/LGN unit j .
 164 Weights from the photo-receptors to units in the ON and OFF channels are set to fixed strengths with a
 165 difference-of-Gaussians kernel ($\sigma_{center} = 0.07385$, $\sigma_{surround} = 0.2954$, in sheet dimensions), with ON
 166 connection fields having a positive center and a negative surround and vice versa for OFF. The lateral
 167 RGC/LGN weights are 2D Gaussians with kernel size $\sigma = 0.25$. The center of the afferent and lateral
 168 connection field of each ON/OFF unit is mapped to the location in the photo-receptor and LGN sheet
 169 corresponding to the location of that unit in sheet coordinates, making all these projections retinotopic.

170 2.1.3 Cortical model

171 Units in the cortical sheets each receive three types of projections represented as matrices of weights:
 172 afferent excitatory ($p = A$), lateral excitatory ($p = E$) and lateral inhibitory ($p = I$). The contribution C_{jp}
 173 to the activation of unit j in a cortical sheet from each projection p at time t is given by:

$$C_{jp}(t + \delta t) = \sum_{i \in F_{jp}} \Psi_i(t) \omega_{pij} \quad (3)$$

174 where $\Psi_i(t)$ is the activation of unit i taken from the set of units in the input sheet of projection p from
 175 which unit j receives input (its connection field F_{jp}), and ω_{pij} is the connection weight from unit i in the
 176 input sheet of projection p to unit j in the output sheet of projection p . All connection field weights are
 177 initialized with uniform random noise multiplied by a 2D Gaussian profile, cut off at the distance specified
 178 below. Contributions from each projection are weighted and summed and passed via a non-linearity f to
 179 calculate the activation of a cortical neuron i at time t :

$$\Psi_i(t) = f\left(\sum_p \gamma_p C_{ip}(t)\right) \quad (4)$$

180 where γ_p is a constant determining the sign (negative for inhibitory) and strength of projection p . The
 181 transfer function f is a half-wave rectifying function that ensures positive activation values. It has a variable
 182 threshold point (θ) dependent on the average activity of the unit as described in the next subsection, but
 183 in all cases the gain is fixed at unity. The projection strength scaling factor of the afferent projection γ_A
 184 was set to 1.5 based on Stevens *et al.* (Stevens et al., 2013) while the values of the lateral excitatory and
 185 inhibitory scaling factors, γ_E and γ_I respectively, were varied (see figure 4) to find a balance between
 186 excitation and inhibition, and between afferent and lateral influences, to provide robust formation of activity
 187 bubbles that facilitates the formation of smooth maps.

188 Once all 16 settling steps are complete, the settled cortical activation pattern is deemed to be the response
 189 of cortical sheets to the presented pattern. At this point we use the response of cortical neurons to update
 190 their threshold point (θ) (using the adaptation process described below) and to update the afferent weights
 191 via Hebbian learning. Cortical activity is then reset to zero, and a new pattern is presented. Note that both
 192 adaptation and learning could instead be performed at every settling step, but this would greatly decrease
 193 computational efficiency.

194 2.1.4 Homeostatic adaptation

195 The threshold θ of all cortical excitatory units is updated at the end of each settling phase based on the
 196 following equations:

$$\theta_{t+1} = \theta_t + \xi(\tilde{\Psi}(t) - \mu) \quad (5)$$

197 where $\xi = 0.01$ is the time constant of the threshold adaptation, $\mu = 0.24$ is a constant defining the target
 198 average activity, and $\tilde{\Psi}$ is the recent average activity of the unit:

$$\tilde{\Psi}(t) = (1 - \chi)\Psi(t) + \chi\tilde{\Psi}(t - 1) \quad (6)$$

199 where $\Psi(t)$ is the output of the unit at time t and $\chi = 0.991$ is a time constant controlling the decay of
 200 the influence of the past activities. The effect of this scaling mechanism is to bring the average activity
 201 of each cortical unit closer to the specified target. If the activity in a V1 unit moves away from the target
 202 during training, the threshold for activation is thus automatically raised or lowered to bring it closer to the
 203 target. Note that an alternative rule with only a single smoothing parameter (rather than ξ and χ) could be
 204 formulated, but the rule as presented here makes it simple for the modeler to set a desired target activity.

205 2.1.5 Hebbian adaptation

206 The initial connection field weights are isotropic 2D Gaussians for the lateral excitatory projection and
 207 uniformly random within a Gaussian envelope for afferent and lateral inhibitory projections. Specifically, a
 208 neuron located at (i, j) will have the following weights in projection p :

$$\omega_{ijp} = \frac{1}{Z_p} u \exp\left(-\frac{x^2 + y^2}{2\sigma_p^2}\right) \quad (7)$$

209 where (x, y) is the sheet-coordinate location of the presynaptic neuron, $u = 1$ for the lateral excitatory
 210 projection ($p=E$), and u is a scalar value drawn from a uniform random distribution for the afferent and
 211 lateral inhibitory projections ($p = A, I$), σ_p determines the width of the Gaussian in sheet coordinates
 212 ($\sigma_A = 0.27, \sigma_E = 0.035, \sigma_I = 0.035\sqrt{2}$), and Z_p is a constant normalizing term that ensures that the total
 213 of all weights ω_{ijp} to neuron j in projection p is 1.0. Weights for each projection are only defined within a
 214 specific maximum circular radius r_p ($r_A = 0.27, r_E = 0.15, r_I = 0.15\sqrt{2}$).

215 In the model, as images are presented to the photo-receptors, the cortical afferent connection weights
 216 $\omega_{i,j,A}$ from the ON/OFF sheets are adjusted once per iteration (after cortical settling is completed) using
 217 a simple Hebbian learning rule. This rule results in connections that reflect correlations between the
 218 presynaptic ON/OFF unit activities and the postsynaptic cortical response. Hebbian connection weight
 219 adjustment at each iteration is dependent on the presynaptic activity, the postsynaptic response, and the
 220 Hebbian learning rate:

$$\omega_{ijA}(t) = \frac{\omega_{ij}(t - 1) + \beta_p \Psi_j(t) \Psi_i(t)}{\sum_{p \in \{ON, OFF\}} \sum_k (\omega_{kj,p}(t - 1) + \beta_p \Psi_j(t) \Psi_k(t))} \quad (8)$$

221 where β_p is the Hebbian learning rate for the connection fields in the two afferent projections from
 222 RGC/LGN $p \in \{ON, OFF\}$. I.e., the afferent weights from RGC/LGN are normalized jointly. Learning

rate parameters are specified as a fixed value $\iota_p = 0.2$ for each projection, and then the unit-specific values used in the equation above are calculated as $\beta_p = \frac{\iota_p}{v_p}$, where v_p is the number of connections per connection field in projection p . The base parameters described here correspond to the first model variant (figure 4). Any modifications of these base parameters in the other two GCAL model variants 2 and 3 (figure 5,6) examined in sections 3.2 and 3.3 are then reported in the respective sections.

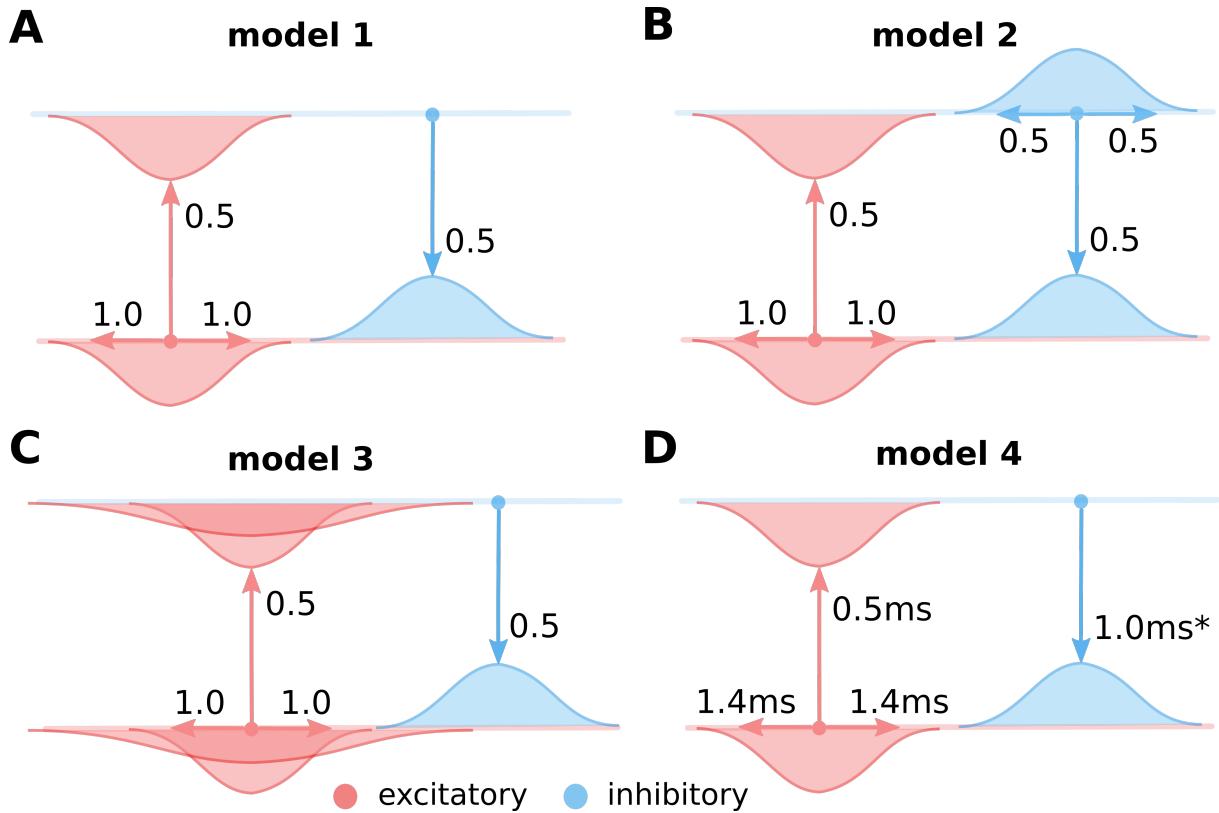


Figure 3. The four considered intra-cortical connectivity variants. As in figure 2, in all panels, the top line represents the excitatory and the bottom line the inhibitory population of neurons. Models (A-C) are using firing rate models of neural units with instantaneous translation of inputs into membrane-potential (see section 2.1). (A) Variant 1, assuming only local equally wide excitatory and inhibitory connectivity, and ignoring direct inhibitory to inhibitory interactions. (B) Variant 2, as in A but with added direct inhibitory to inhibitory connections of equal extent as the other excitatory and inhibitory projections. (C) Variant 3, as in A but with added long-range excitatory connections represented based on Buzás *et al.* (Buzás *et al.*, 2006) as a second wider Gaussian. (D) The same as A but modeled with neural units that take into account membrane time-constant (see section 2.2). The transmission delay of the inhibitory to excitatory projections in model 4 was varied in the experiments presented in section 3.4 (figure 7).

228 2.2 Rate model with membrane time constant

The architecture of this model (model variant 4; figure 2 and 3D) is identical to the GCAL models (model variants 1 through 3; figure 3A-C) with the exception of the equations 3 and 4, which were replaced with an equation taking into account the membrane time constant:

$$\tau_z \frac{\delta \Psi_i}{\delta t} = -\Psi_i + \sum_p \gamma_p \sum_{j \in F_{ip}} f(\Psi_j(t - \nu_p)) \omega_{pji} \quad (9)$$

232 where τ_z ($z \in \{E, I\}$) is the membrane time constant of excitatory and inhibitory neurons ($\tau_E = 2$ ms, τ_I
233 = 0.5 ms), and ν_p is the transmission delay for projection p ($\nu_{EE} = 1.4$ ms, $\nu_{EI} = 0.5$ ms, ν_{IE} was varied
234 (see section 3.4)). As a consequence the evolution of the network dynamics has to be simulated at higher
235 resolution. We chose the update step to be 0.1 ms, and we let the activity in the model settle for 150 ms, thus
236 resulting in 1500 settling steps as opposed to the 16 of the GCAL model. The sub-cortical parametrization
237 of the model is identical to the model variants 1-3 (figure 3A-C), but because the addition of membrane
238 time constants substantially changes the cortical dynamics the cortical parameters had to be re-adjusted.
239 Specifically the strength of the thalamo-cortical connections was set to $\gamma_A = 0.5$, the strength of the lateral excitatory-
240 to-excitatory connections was set to $\gamma_{EE} = 3.5$, the strength of the lateral excitatory-
241 to-inhibitory connections was set to $\gamma_{EI} = 1.0$, and the strength of the lateral inhibitory-to-excitatory
242 connections was set to $\gamma_{IE} = 2.94$.

243 2.3 Orientation map analysis

244 Model orientation maps are calculated based on the vector average method (Miikkulainen et al., 2005).
245 We first determine the preferred frequency of neurons across the map. Due to the simplified stereotypical
246 stimulus used in this study (elongated Gaussian inputs) the spatial frequency preference of all neurons lies
247 in a very narrow band, and we thus use the mean preferred spatial frequency across all cortical neurons as
248 the value for the spatial frequency parameter across all subsequent analysis. Next, sinusoidal grating inputs
249 that cover the full range of remaining parameter values (combinations of all orientations and phases) are
250 presented, and for each orientation, the peak response of the neuron is recorded. The orientation preference
251 is calculated by constructing a vector for each orientation θ (between 0 and 180°), with the peak response
252 as the length and θ as its orientation. These vectors are summed and the preferred orientation is calculated
253 as half of the orientation of the summed vector. The selectivity is given by the magnitude of the summed
254 vector.

255 2.4 Orientation map quality measure

256 In order to assess whether the proposed model develops orientation maps that match the structure of those
257 found in real animals we need to utilize an automatic metric that tells how close the maps are to animal
258 data. To this end we will utilize a map-quality metric that we have recently developed (Stevens et al., 2013)
259 based on the empirical observation that pinwheel count in biological orientation maps scales linearly with
260 hypercolumn size across many different species (Kaschube et al., 2010). Specifically the pinwheel density
261 per hypercolumn area (Λ^2) converges to π , when averaged across a sufficiently large cortical surface. For a
262 detailed description of the procedure for calculating this metric we refer the reader to our previous work
263 (Stevens et al., 2013), but briefly, its calculation involves three steps. First, the locations of the pinwheels
264 in the orientation map are determined as the intersections of the zero contours of the real and imaginary
265 components in the polar representation of the maps, thus yielding the total pinwheel count in the map.
266 Second the hypercolumn size is determined as the peak in the isotropic ring-like Fourier transform of the
267 orientation maps. Third, using these two numbers we can derive the pinwheel density, but to transform
268 it to a useful metric between unity (high-quality map) and zero (low quality map) we pass it through a
269 normalized Gamma distribution. We have shown that this metric reliably distinguishes low and high quality
270 maps (Stevens et al., 2013) and is a valid measure for assessing how well the model orientation maps match
271 animal data.

3 RESULTS

272 In this article we will proceed through multiple gradually more complex models of orientation development,
273 addressing several of the major issues with modeling this phenomenon, eventually demonstrating that the
274 experimentally identified fast inhibitory loop is a satisfactory explanation for how short-range inhibition
275 can support the development of cortical functional organization. We will use two different computational
276 abstractions to explore the questions at hand. In the fist part of the study we will use a computational
277 model that is derived from the LISSOM family of models (Miikkulainen et al., 2005). This choice has three
278 advantages. It allows for a very straightforward explanation of why a fast inhibitory loop enables short-
279 range inhibition to induce competition. It makes our explanation directly comparable to the extensive set of
280 published LISSOM family models, and thus demonstrates that the solution proposed here generalizes to a
281 range of other functional properties. And finally the LISSOM abstraction enables very fast simulations, thus
282 allowing us to perform a parameter search analysis that would otherwise be computationally prohibitive.
283 However, as we will explain further in section 3.4, some simplifications made by the LISSOM abstraction,
284 specifically the instantaneous translation of the neuronal input into its activity, will leave certain questions
285 unanswered. These will be addressed in section 3.4 using more detailed rate model framework.

286 **3.1 A fast excitatory to inhibitory to excitatory loop enables competition in networks**
287 **with short-range inhibitory connections.**

288 Ohana *et al.* (Ohana et al., 2012) have shown that transmission delays between different types of cortical
289 neurons are not uniform, specifically they found that on average the transmission delays between excitatory
290 neurons are ~ 1.4 ms, from excitatory to inhibitory neurons are ~ 0.5 ms, and from inhibitory to excitatory
291 neurons are ~ 0.98 ms. The sample size for connections between inhibitory cells in the study was not large
292 enough to be quantitatively reliable. The key observation here is that the combined di-synaptic delay from
293 excitatory to inhibitory and inhibitory to excitatory cells is approximately as long as the mono-synaptic
294 delay of the excitatory to excitatory connections. Drawing from this, the core insight of this study is
295 that under such a pattern of delays the effective inhibitory interactions are well approximated by the
296 convolution of the excitatory and inhibitory connection kernels, as we will show below. This gives the
297 effective inhibition longer range, consequently fulfilling the essential requirement for cortical competition
298 to occur. For didactic purposes, and to simplify analytical treatment, let us first explore one specific set
299 of conditions under which the above statement holds exactly, we will explore the situation when these
300 conditions are relaxed in subsequent sections:

- 301 1. The sum of the excitatory to inhibitory and inhibitory to excitatory delays is exactly the same as the
302 excitatory to excitatory delay
- 303 2. There is no inhibitory to inhibitory interaction
- 304 3. The synaptic inputs into the excitatory and inhibitory neurons are instantaneously translated into their
305 rate response via a positive rectified transfer function.
- 306 4. The connection kernels of all neurons are Gaussian kernels with spatial constant σ_e for excitatory
307 neurons and σ_i for inhibitory neurons.
- 308 5. We will assume only local connectivity (disregarding long-range excitatory connections), and assume
309 that both excitatory and inhibitory neurons have the same extent, thus $\sigma_e = \sigma_i$

These conditions lead to the following set of equations governing the cortico-cortical interaction.

$$\begin{aligned} R_e(x, t) &= \left[\sum_y N_{\sigma_e}(\|y - x\|) R_e(y, t - \theta_{EE}) - N_{\sigma_i}(\|y - x\|) R_i(y, t - \theta_{IE}) + I_{aff}(x, t) \right]^+ \\ R_i(x, t) &= \left[\sum_y N_{\sigma_e}(\|y - x\|) R_e(y, t - \theta_{EI}) \right]^+ \end{aligned} \quad (10)$$

310 where $R_p(x, t)$ is the response of a neuron of type p located at position x at time t , $I_{aff}(x, t)$ is the afferent
 311 input to a neuron at position x at time t , N_σ is a normal distribution of variance σ , corresponding to the
 312 lateral connection kernel, and θ_{ab} is a delay on connections from neural type a to neural type b . When we
 313 expand for R_i we obtain:

$$\begin{aligned} R_e(x, t) &= \left[\sum_y N_{\sigma_e}(\|y - x\|) R_e(y, t - \theta_{EE}) - \sum_y N_{\sigma_i}(\|y - x\|) \left[\sum_z N_{\sigma_e}(\|z - y\|) R_e(z, t - \theta_{EI} - \theta_{IE}) \right] \right. \\ &\quad \left. + I_{aff}(x, t) \right]^+ \end{aligned} \quad (11)$$

314 Because $\theta_{EI} + \theta_{IE} = \theta_{EE}$ (assumption 1) and because the inhibitory neurons receive only excitatory
 315 connections (assumption 2) we can further simplify the above as:

$$R_e(x, t) = \left[\sum_y N_{\sigma_e}(\|y - x\|) R_e(y, t - \theta_{EE}) - \sum_z \sum_y N_{\sigma_i}(\|y - x\|) N_{\sigma_e}(\|z - y\|) R_e(z, t - \theta_{EE}) + I_{aff}(x, t) \right]^+ \quad (12)$$

316 and thus due to the symmetry of the normal distribution:

$$R_e(x, t) = \left[\sum_y N_{\sigma_e}(\|y - x\|) R_e(y, t - \theta_{EE}) - \sum_z N_{\sigma_i} * N_{\sigma_e}(\|z - x\|) R_e(z, t - \theta_{EE}) + I_{aff}(x, t) \right]^+ \quad (13)$$

317 Because convolution of two Normal distributions of variance σ_i and σ_e is a normal distribution with
 318 variance $\sqrt{\sigma_i^2 + \sigma_e^2}$ and because $\sigma_i = \sigma_e$ we can simplify to:

$$R_e(x, t) = \left[\sum_y N_{\sigma_e}(\|y - x\|) R_e(y, t - \theta_{EE}) - \sum_z N_{\sqrt{2}\sigma_e}(\|z - x\|) R_e(z, t - \theta_{EE}) + I_{aff}(x, t) \right]^+ \quad (14)$$

319 This shows that under these specific assumptions the effective inhibitory interactions between excitatory
 320 neurons are $\sqrt{2}$ longer than the excitatory ones and thus follow the Mexican-hat like profile required
 321 for the lateral cortical competition underlying functional map development. Even though this essential
 322 condition of cortical competition is fulfilled in this model configuration, important constraints on the
 323 extent of the effective lateral interactions (relative to lateral excitation) remain. It is thus still unclear

324 whether development of high-quality orientation maps as observed experimentally is supported under these
325 conditions. Conveniently, the lateral interaction in the LISSOM family of models is governed by the same
326 equations as 10. In the following we will use the GCAL model (Stevens et al., 2013), the latest and most
327 robust addition to the LISSOM family of models to demonstrate that the above specific configuration of
328 effective lateral excitation and inhibition permits the development of high-quality orientation maps. GCAL
329 is the only model of stimulus dependent functional development which achieves emergence of biologically
330 realistic orientation maps in terms of pinwheel density, a signature that is a useful objective measure of
331 orientation map quality. We will utilize this map quality measure to reliably detect model configurations
332 which permit the successful emergence of orientation maps. To facilitate a comparison we will use exactly
333 the same model configuration as in Stevens *et al.* (Stevens et al., 2013) (see Materials and Methods), except
334 for three modifications necessitated by the analysis above (see figure 3A):

- 335 1. We will change the spread of lateral inhibition to be $\sqrt{2}$ longer than lateral excitatory spread, such that
336 the resulting lateral interactions conform with equation 14.
337 2. The change in point 1 will result in a change in the balance of overall excitation and inhibition in the
338 model which is critical for successful functional development in the model. We will thus modify the
339 strength of the lateral inhibition to compensate for changes due to modification 1.
340 3. Analogously to 2, the changes in 1 change the overall balance between feed-forward and lateral
341 contributions to model a cortical neuron's activity, which we will compensate by changing the overall
342 strength of the lateral interactions.

343 In order to find a working combination of parameters in points 2 and 3, and also to show that the model is
344 robust to a certain level of changes in these two parameters, we have performed a parameter search across
345 these two parameters and evaluated the quality of the orientation map (see Materials and Methods) for each
346 parameter combination. As figure 4 shows, under a range of values of both parameters the model develops
347 high-quality orientation maps indistinguishable from their experimental counterparts, thus concluding our
348 first step towards showing that a fast excitatory-to-inhibitory-to-excitatory loop can explain how short-range
349 inhibition can induce cortical competition and consequently the development of topological organization
350 of functional properties.

351 Furthermore, note that the GCAL model used in figure 4 only explicitly models excitatory neurons and
352 assumes both direct excitatory and inhibitory interactions between them, thus corresponding to equation 14.
353 Even though above we have shown that equation 14 is equivalent to equation 10 to verify the correctness
354 of our analysis we have run a single simulation of the GCAL model corresponding to the parameter
355 combination with the highest map quality found in figure 4, but with an explicitly simulated inhibitory
356 population (figure 4E). In this model we thus do not model direct inhibitory interactions between excitatory
357 neurons, but instead add excitatory to inhibitory and inhibitory to excitatory connections of the same extent
358 as those of the excitatory to excitatory pathway (see the assumption #5). As we have shown above this
359 model should be mathematically equivalent to the simulation shown in figure 4D. Note, however, that the
360 GCAL simulations represent a discrete approximation in both time and space of equations 10,14 and we
361 thus expect small numerical discrepancies. Indeed, orientation maps shown in figure 4E are nearly identical
362 to those in the figure 4D, with only barely perceptible numerical differences, confirming the validity of our
363 approach.

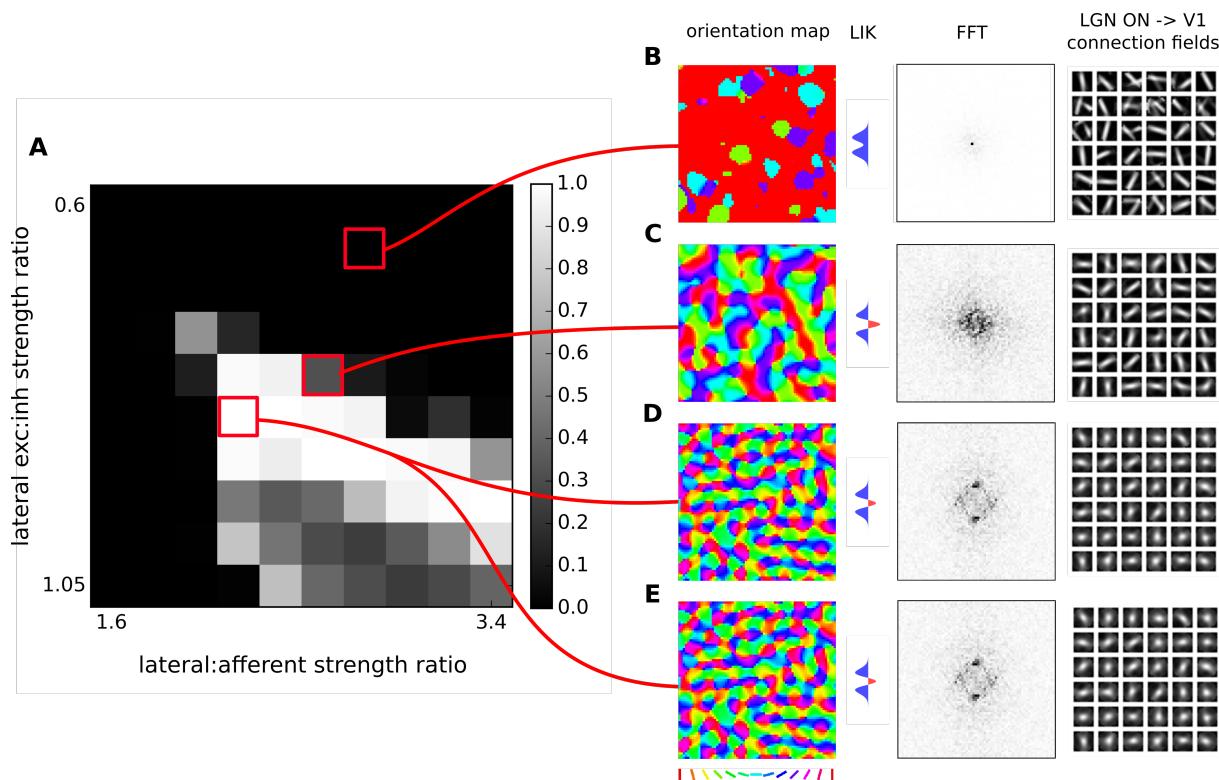


Figure 4. Orientation map development with short range inhibition and a fast excitatory-to-inhibitory-to-excitatory loop. (A) Map quality (see Materials and Methods) at a range of lateral excitatory vs inhibitory projection strength ratios and afferent vs lateral projection strength ratios. (B-D) Functional organization in 3 example parameter configurations indicated by the red marks. From left to right, the orientation map, the lateral interaction kernel (LIK; this is calculated based on eq. 14), fast-Fourier transform of the orientation map and afferent connection fields from the ON LGN model sheet for 25 example model V1 neurons. (B-C) Two examples of sub-optimal orientation maps. (D) Model configuration with the highest quality map found in this parameter search. (E) The same model configuration but in this case it was run with explicit simulation of inhibitory neurons and corresponding connections. E and D are nearly identical confirming the correctness of our analysis.

364 **3.2 Inhibitory to inhibitory connections are consistent with development of functional
365 organization**

366 In the previous section we have shown that under a set of specific assumptions that facilitate analytical
367 treatment, short-range inhibition can induce effective Mexican-hat like interaction and thus support the
368 development of orientation maps. However, not all assumptions we made were in line with experimental
369 evidence. In this section we will show, now only numerically, that one of these assumptions is not
370 necessary, specifically that the addition of inhibitory to inhibitory connections does not prevent emergence
371 of orientation maps.

372 To this end, we will use the exact GCAL model configuration that we have found in the previous section
373 to possess the highest quality orientation map (see figure 4). We will use the GCAL configuration in which
374 we will explicitly model the inhibitory neurons (see figure 3B, and equation 10) and consequently also
375 explicitly the excitatory to inhibitory and inhibitory to excitatory connections. Furthermore, we will add
376 inhibitory to inhibitory connections to the model with the same extent as that of inhibitory to excitatory
377 connections ($\sigma_i = \sigma_e$ in eq 10).

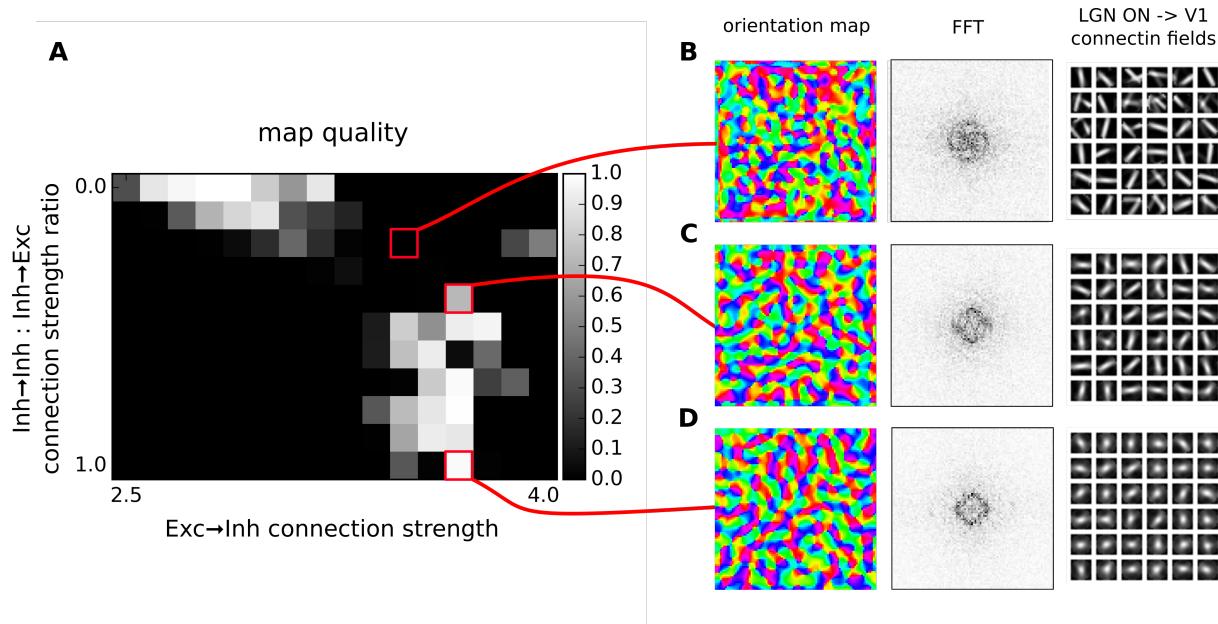


Figure 5. Orientation map development with short range inhibition, a fast excitatory-to-inhibitory-to-excitatory loop and inhibitory to inhibitory connections. (A) Map quality (see Materials and Methods) at a range of inhibitory to inhibitory vs inhibitory to excitatory projection strength ratios and excitatory to inhibitory connection strengths. (B-D) Functional organization in 3 example parameter configurations indicated by the red marks. From left to right, the orientation map, the lateral interaction kernel (LIK; this is calculated based on eq. 14), fast-Fourier transform of the orientation map and afferent connection fields from the ON LGN model sheet for 25 example model V1 neurons.

378 By explicitly modeling the inhibitory neurons in this model we have replaced a single parameter governing
 379 the strength of inhibitory lateral interactions in the model from the previous section with three new
 380 parameters that set: (1) the strength of excitatory to inhibitory, (2) inhibitory to excitatory and (3) inhibitory
 381 to inhibitory connections (all other parameters remained the same as in the best parameterization found
 382 in the previous section). Note that in principle, there is redundancy in these parameters as the overall
 383 strength of the projections from inhibitory neurons onto both excitatory and inhibitory populations is scaled
 384 by the excitatory to inhibitory projection strength. Therefore in figure 5 we have systematically varied the
 385 strength of the added inhibitory to inhibitory projection expressed relatively to the strength of the inhibitory
 386 to inhibitory connections (which was set to 1), while also varying the strength of the excitatory to inhibitory
 387 projection. We have investigated the quality of the orientation maps that developed under these different
 388 levels of inhibitory to inhibitory interaction strengths. As can be seen, high quality maps can develop under the full
 389 range of the inhibitory to inhibitory interaction strengths, depending on the overall excitatory to inhibitory
 390 drive. This shows that inclusion of direct inhibitory to inhibitory interactions does not invalidate the results
 391 of section 3.1 (figure 4).

392 3.3 Long-range excitation

393 In model variants 1 and 2 (figure 4 and 5) we have only assumed local connectivity by setting both
 394 excitatory and inhibitory interactions to have the same spatial extent. However experimental evidence shows
 395 that excitatory cells send longer connections compared to inhibitory cells (Buzás et al., 2006; Budd and
 396 Kisvárdy, 2001). In this section we will explore what happens if we add long-range excitatory connectivity
 397 into model variant 1. Buzás *et al.* (Buzás et al., 2006) have shown that the lateral connectivity in layer 2/3
 398 can be best described as superimposition of two gabor connectivity likelihoods, one short-range but not

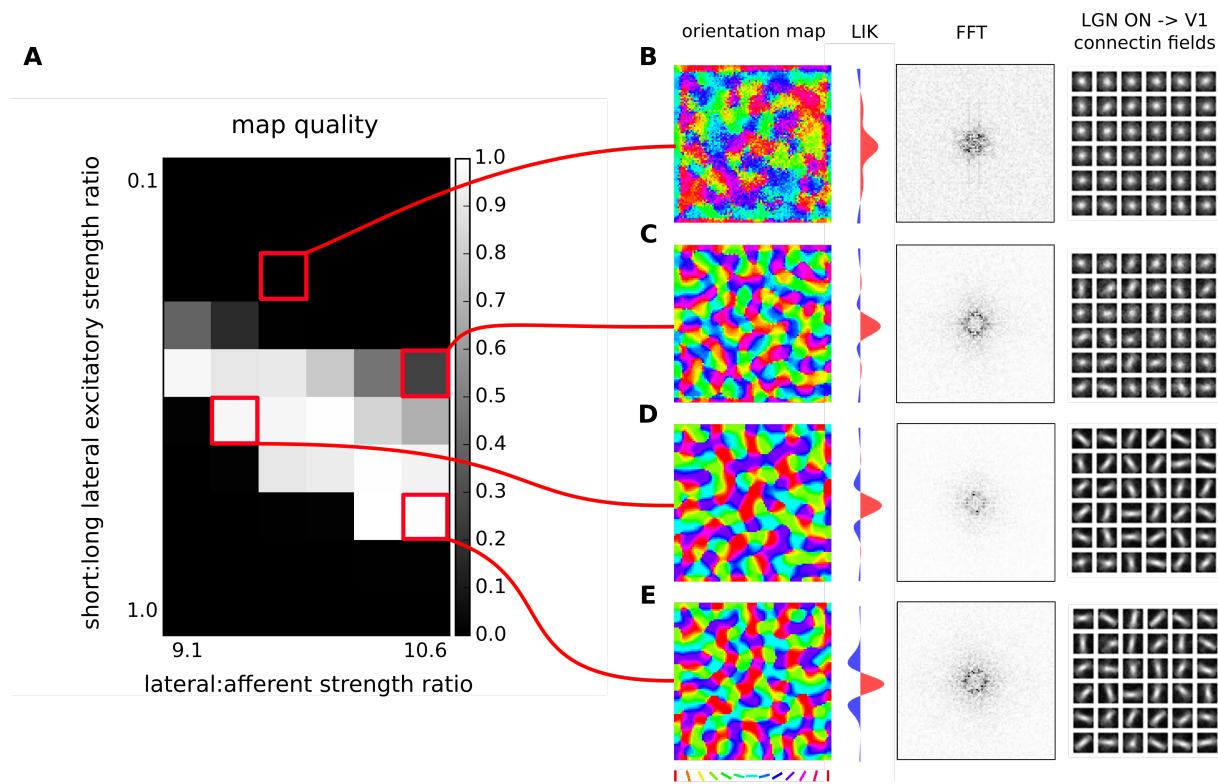


Figure 6. Orientation map development with short range inhibition, a fast excitatory-to-inhibitory-to-excitatory loop and long-range excitation. (A) Map quality (see Materials and Methods) at a range of short vs. long range excitatory connection strength ratios (y-axis) and a range of afferent vs. lateral connection strength ratios (x-axis). (B-E) Functional organization in 4 example parameter configurations indicated by the red marks. From left to right, the orientation map, fast-Fourier transform of the orientation map and afferent connection fields from the ON LGN model sheet for 25 example model V1 neurons.

399 orientation specific and one long-range and orientation specific. Here we will assume such dual structure,
 400 leading us to add a second excitatory to excitatory and excitatory to inhibitory projection into model variant
 401 1, but with a space constant that is 3 times larger (see figure 3C), in line with Buzás *et al.* (Buzás et al.,
 402 2006) quantitative findings.

403 In figure 6 we will examine what strength of the long-range excitatory connections, relative to the
 404 short-range excitatory ones (y axis in figure 6A), leads to development of a high-quality orientation map.
 405 Adding the long-range excitation changes the spatial configurations over which excitatory interactions win
 406 over inhibitory ones, as well as the overall magnitude of the resulting net local excitation. Consequently,
 407 the proportion between the magnitude of the net local excitation due to the lateral interactions and the
 408 excitation due to afferent inputs is changed, which is a crucial parameter for map development. To
 409 compensate for these changes, we also systematically explore the ratio of the overall magnitudes of
 410 the lateral and afferent interactions (x axis of figure 6). To make the parameter search computationally
 411 feasible, we perform the parameter search only in the region of parameters that allow for sufficiently strong
 412 long-range excitatory projections. As can be seen in figure 6D and E, under appropriately strong lateral
 413 interactions, substantial long-range excitatory connections still permit the development of high quality
 414 orientation maps, demonstrating that the proposed model is consistent with the experimentally identified
 415 long-range excitatory connectivity.

416 3.4 Non-equal effective excitatory and inhibitory delays

417 In all model variants examined so far we have made the key assumption that the delay on the excitatory
418 to excitatory connections is exactly equal to the sum of excitatory to inhibitory and inhibitory to excitatory
419 delays. This assumption is approximately supported by the experimental evidence (Ohana et al., 2012), but
420 we cannot assume it holds exactly in a real biological substrate. However, we hypothesize, that the small
421 discrepancies between the delay of the mono-synaptic excitatory connections and the cumulative delay of
422 the bi-synaptic inhibitory interactions can be absorbed into the membrane time-constant of the neurons.
423 In this section we will verify this hypothesis by extending the modeling framework used thus far with a
424 finite membrane time-constant (see Materials and Methods) and proceed to determine the magnitude of the
425 discrepancy in the delays between the excitatory and inhibitory interactions that can be managed by the
426 model without impairing the resulting orientation map quality.

427 We use a model parameterization similar to those determined for model variant 1 (section 3.1; figure 3A).
428 For simplicity and computational efficiency we omit the inhibitory to inhibitory and long-range excitatory
429 connections that have already been investigated with model variants 2 and 3. The membrane time-constant
430 of excitatory neurons was set to 2ms while those of inhibitory neurons to 0.5ms. These faster inhibitory
431 dynamics are necessary to prevent oscillations in the system (Kang et al., 2003). We set the excitatory to
432 excitatory delay to 1.4 ms and excitatory to inhibitory delay to 0.5 ms based on Ohana *et al.* (Ohana et al.,
433 2012). In order to understand how closely the cumulative bi-synaptic inhibition delay has to match that of
434 the direct excitatory to excitatory delay, we will vary the delay on the inhibitory to excitatory projection
435 (note that we could achieve the same by varying the excitatory to inhibitory delays, and this choice was
436 arbitrary).

437 Figure 7 shows the resulting orientation maps and associated map quality measures of models with a
438 range of differences between the delays of monosynaptic excitatory and bi-synaptic inhibitory interactions,
439 that are in the figure expressed as the sum of the excitatory to inhibitory and inhibitory to excitatory
440 delays minus the excitatory to excitatory delay (i.e. the figure shows how much slower the inhibitory
441 disynaptic interactions were in comparison with the monosynaptic excitatory ones). As can be seen when
442 the differences between the excitatory delay (1.4ms) and cumulative inhibitory delay is small (<0.8 ms)
443 high quality orientation maps develop in the model, confirming that sufficiently large discrepancy between
444 the direct excitatory and bi-synaptic inhibitory delays can be accommodated in the model. However, as
445 expected, as the difference between the delays increases the ability of the model to learn a topologically
446 organized representation of orientation preference diminishes. Crucially, if the delays across all the
447 projections were equal (figure 7), as is typically assumed, the model fails to develop orientation maps
448 in line with the analysis by Muir *et al.* (Muir and Cook, 2014), thus confirming that the specific delay
449 pattern between neural types identified by Ohana *et al.* (Ohana et al., 2012) is key to achieving competitive
450 dynamics in topologically organized neural models.

4 DISCUSSION

451 In this study we have shown how recent findings on dependence of neural transmission delays on the type
452 of pre- and post-synaptic neurons (Ohana et al., 2012) can resolve a long-standing question on how short-
453 range inhibition can support cortical competition and consequently the development of functional cortical
454 topological organization. Under simplifying assumptions, we have analytically shown how disynaptic
455 inhibition that is as fast as mono-synaptic excitation can extend the effective range of inhibitory interactions,
456 in contrast to the recent analytical results showing that in the case of equal synaptic delays on all connections

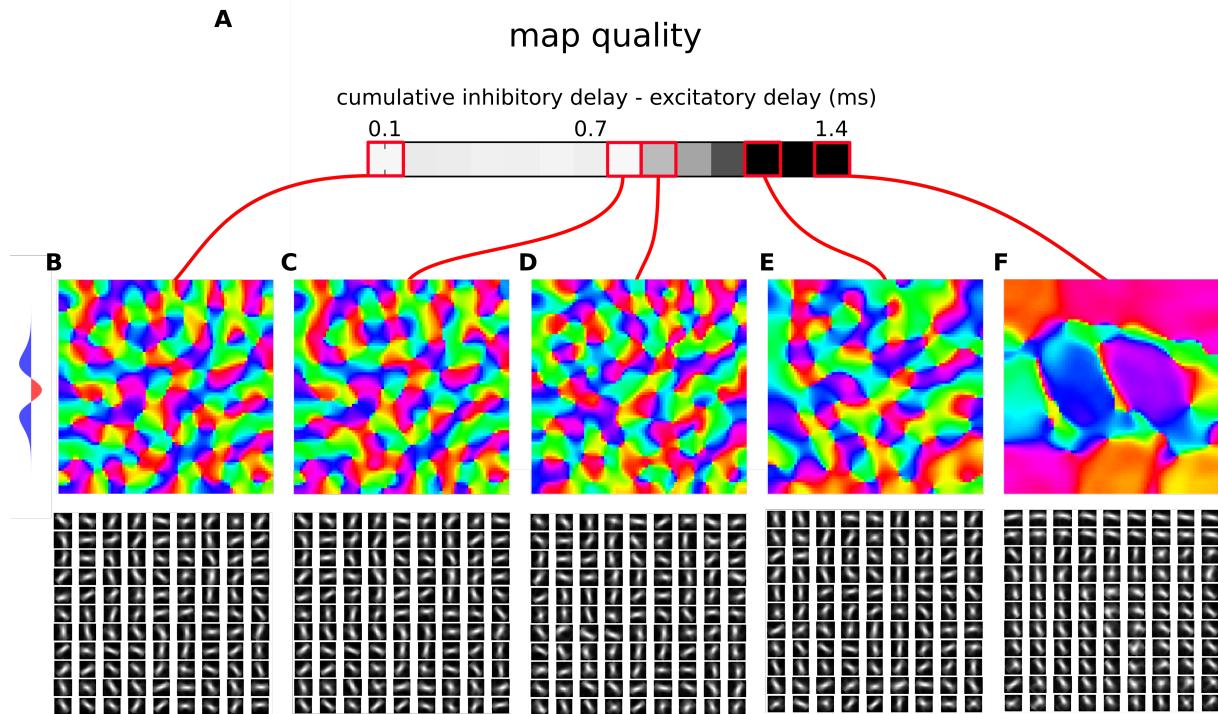


Figure 7. Orientation map development in a rate model with a synaptic time constant. (A) Map quality (see Materials and Methods) at a range of cumulative inhibitory delays expressed as the difference from the direct excitatory to excitatory delays. (B-F) Functional organization in 5 parameter configurations indicated by the red marks. The orientation map (top), and afferent connection fields from the ON LGN model sheet for 81 example model V1 neurons (bottom). (B-D) Three examples of configurations where good quality orientation maps develop. (E) If the cumulative inhibitory delay is longer by more than approximately 0.8ms in comparison to the direct excitatory delay the map quality starts to drop. (F) In the configuration corresponding to the case where delays on all connections are equal (i.e. the cumulative delays of the inhibitory interactions is twice as long as on the direct excitatory ones) orientation maps fail to develop. The lateral interaction kernel is indicated on the left, but note that it is valid only for the first parametrization, corresponding to equal delays on the mono-synaptic excitation and dy-synaptic inhibition pathway.

457 the disynaptic inhibition has negligible effects (Muir and Cook, 2014). We have also shown that these
 458 findings are applicable to the problem of functional development in primary visual cortex. We have then
 459 proceeded to show using computational methods that the proposed models are robust to the addition
 460 of other well established features of cortical anatomy, commonly ignored by similar studies, including
 461 the long-range excitatory connections and mutual inhibition among inhibitory neurons. Finally, we have
 462 shown that the proposed mechanisms are robust to the variations of the exact delay ratio between the
 463 mono-synaptic excitation and di-synaptic inhibition. Overall, this study represents an important advance in
 464 our understanding of how orientation map development can be supported by the cortical neural substrate.

465 Using neuron reconstruction data of the recorded neurons, Ohana *et al.*,(Ohana et al., 2012) identified the
 466 placement of the different pre-post synaptic combinations on the dendrites of the target cell as the likely
 467 origin of this transmission delay heterogeneity. They found that the excitatory-to-inhibitory synapses were
 468 closer to the soma of the post-synaptic neuron than the excitatory-to-excitatory and inhibitory-to-excitatory

469 synapses. An axonal origin for the transmission delay heterogeneity is unlikely, as Ohana *et al.*, (Ohana
470 et al., 2012) found that all the pre-post combinations were, on average, equally distant from the pre-synaptic
471 cell body. They further supported the dendritic origin of the observed delay inhomogeneities by showing in
472 a computational model that the observed delay magnitudes can be explained by these anatomical findings.
473 Overall this indicates a mechanism for the generation of the differences in inter-neuron transmission delays
474 that is likely to generalize beyond the primary visual cortex. We thus suggest that the results presented in
475 this study generalize to the development of other functional features and other cortical competition based
476 mechanisms whose origin is outside of primary visual cortex.

477 In this study we have focused on the development of the global organization of the thalamo-cortical and
478 cortico-cortical connectivity and as a consequence the global organization of V1's most salient functional
479 property - orientation tuning. How do, however, the individual model cortical units correspond to individual
480 biological V1 neurons, e.g. do they have matching tuning properties? The detailed systematic investigation
481 of this question is outside the scope of this study; however, to offer at least a basic view of how the
482 explored models behave at single-cell level, we show in figure 8 representative orientation tuning curves
483 for all model parametrizations for which we have shown a detailed view of their properties (i.e. we have
484 presented their orientation map etc.). Figure 8 shows that orientation tuning of individual neurons varies
485 between the models and parametrizations, but all the models with good quality orientation maps also have
486 reasonably realistic orientation tuning of single units. Several of these model parametrizations achieve
487 sharp realistic contrast invariance of the orientation tuning width. At the same time, figure 8 shows that
488 the relationship between the orientation maps and afferent connectivity patterns and the orientation tuning
489 is not straightforward. For example some parametrizations with inferior maps achieve sharper tuning
490 (e.g. figure 8 A2) in comparison with model parametrizations exhibiting high-quality orientation maps
491 (e.g. figure 8 A3). Ultimately, further, more rigorous, quantification and systematic investigation of the
492 single cell tuning properties and their dependence on the model parametrization will be necessary to fully
493 understand these relationships.

494 The most related past explanation of how cortical competition can arise under short range inhibition is
495 that of Kang *et al.* (Kang et al., 2003), who have shown that under the assumption of a faster inhibitory
496 time constant (as opposed to excitatory), the effective excitatory and inhibitory interactions will follow
497 the Mexican hat profile and thus support competition along the cortical surface. Indeed, in our final
498 model variant 4 that explicitly considers membrane time-constant presented in section 3.4, we assume that
499 inhibitory neurons have faster membrane time constants than excitatory ones, as otherwise we observe
500 oscillatory behavior in line with the analytical findings of Kang *et al.* (Kang et al., 2003). Crucially, Kang *et*
501 *al.* (Kang et al., 2003) assumed instantaneous neural transmission, and when this biologically implausible
502 assumption is rectified by addition of transmission delays that are uniform across the connections between
503 the different pre- and post-synaptic neural types, we find that the competitive dynamics in the neural
504 model break (figure 7F) in line with the analytical and computational results of Muir *et al.* (Muir and Cook,
505 2014). However, when we replace the transmission delays with the neural-type specific pattern uncovered
506 by Ohana *et al.* (Ohana et al., 2012) the competitive dynamics in the model are rescued and we observe
507 development of high quality orientation maps (section 3.4AB), in line with the analytical results under
508 simplified conditions (section 3.1).

509 The analytical results in this study were obtained only under simplifying assumptions, specifically
510 instantaneous translation of inputs to membrane potential, equal extent of excitatory and inhibitory
511 connections and a lack of inhibitory-to-inhibitory interactions. Even though we have shown computationally
512 that these assumptions are not necessary for achieving the cortical competition and the consequential

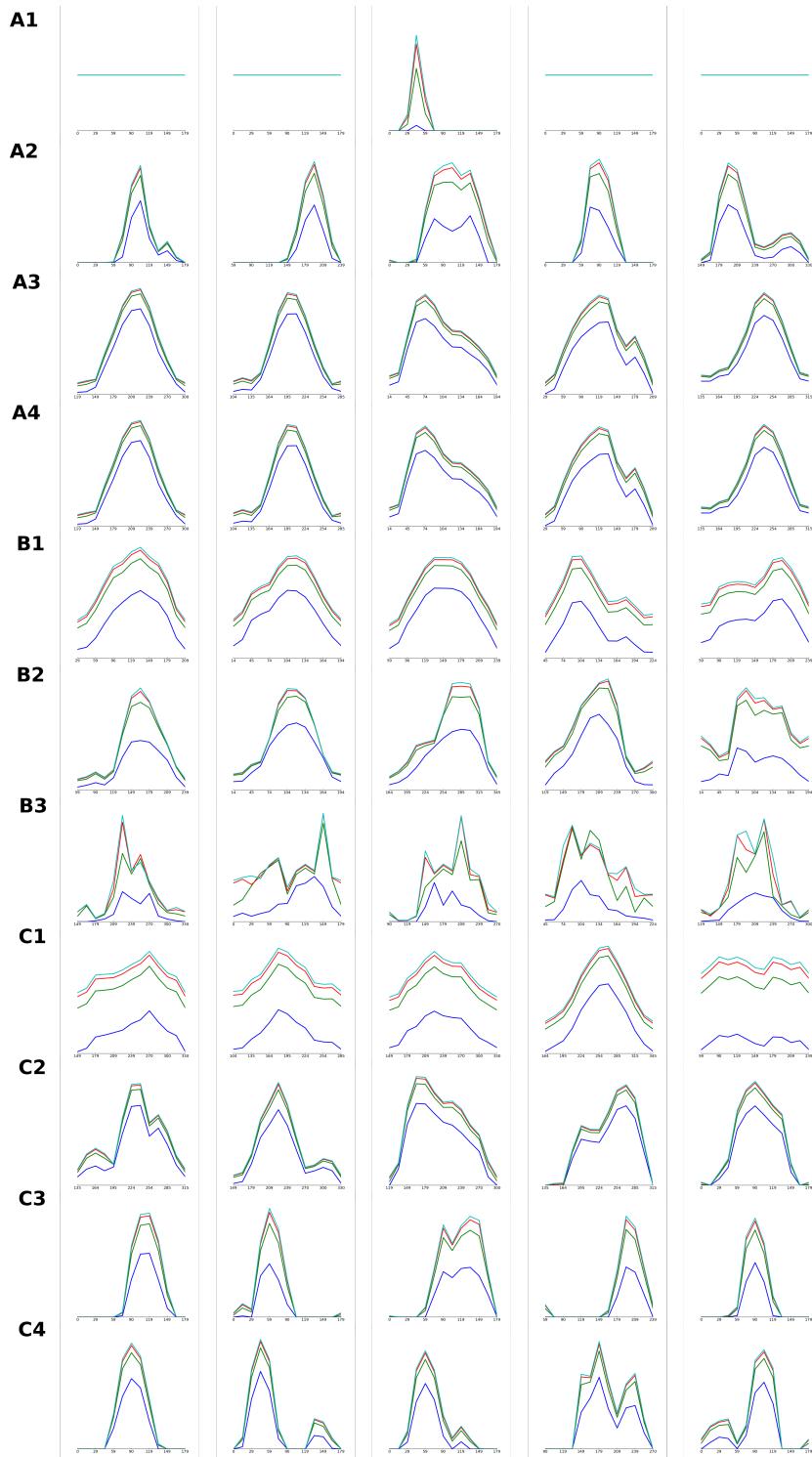


Figure 8. Orientation tuning curves of individual V1 model units in models detailed in figures 4-6. Each row of the figure shows orientation tuning curves measured at 4 different contrasts of 5 model units, one in the center of the modelled cortical area and 4 positioned in the middle between the center and one of the 4 corners of the modelled cortical area. Each row corresponds to one of the model parametrizations that were detailed (i.e. their orientation map etc. was shown) in figure 3-5: tuning curves marked with A correspond to models from figure 4, B from figure 5 and C from figure 6. The order of the rows corresponding to a given figure 4, 5 or 6 (i.e. marked with A, B or C) is the same as the order in which the model parametrization in the given figure were presented (i.e. orientation tuning curves measured in the model parametrization presented in the 2nd row of figure 5 will be displayed in the row marked as B2 in this figure).

513 orientation map development sought in this study, further analytical work then can circumvent these
514 simplifications would undoubtedly provide deeper understanding of the dynamics of the studied neural
515 system and its dependence on the various parameters. This sentiment underlies the parameter explorations
516 presented here, which show that even though the model is robust to changes in the considered parameters,
517 the existence or not of dynamics supporting development of orientation maps can form a complex pattern
518 within the explored parameter spaces. This is particularly the case for the model variant with inhibitory-
519 to-inhibitory interactions, in which (unlike in the other variants) the inhibitory population gains it's own
520 dynamics. Furthermore, the relative computational complexity of the studied models and the extensive set of
521 parameters involved preclude systematic search across the full parameter space, and we have only explored
522 parameters that we empirically found to have the biggest impact. Finally, one simplifying assumption that
523 we have not treated in this study is the lack of direct thalamic input onto inhibitory cells. Since inhibitory
524 cells in cortical layer 4 do receive thalamic input (Binzegger et al., 2004) the inclusion of external input in
525 the inhibitory population needs to be considered in the future.

526 In this paper we have decided to investigate cortical competitive mechanisms through the prism of
527 orientation map development. The advantage of this approach is that it allows us not only to show that
528 some form of competition is possible, but also that it is of the form that actually supports implementation
529 of specific cortical computations. Given that we show that our model implements effective Mexican hat
530 lateral interactions (section 3.1) and these have in the past been shown to be sufficient to explain cortical
531 organization of other functional features (i.e. retinotopy, ocular dominance, spatial frequency and color) it
532 is very likely that our results will generalize to these other dimensions of sensory input as well. Cortical
533 competition of other forms has been proposed to underly a broad variety of other cortical operations,
534 including associative memory, noise suppression, decision making, saliency detection and other forms of
535 attentional computations. Even though additional work will be required to determine if the mechanisms
536 proposed here can generalize to these other neural computations, this study offers a promising framework
537 for anatomically plausible mechanistic explanations of these important aspects of brain function.

CONFLICT OF INTEREST STATEMENT

538 The authors declare that the research was conducted in the absence of any commercial or financial
539 relationships that could be construed as a potential conflict of interest.

AUTHOR CONTRIBUTIONS

540 JA performed the calculus, simulations and wrote the manuscript.

FUNDING

541 The research leading to these results has received funding from the European Union Seventh Framework
542 Programme (FP7/2007-2013) under grant agreement no. 604102 (Human Brain Project).

ACKNOWLEDGMENTS

543 I would like to thank James A. Bednar and Andrew Davison for reviewing the original draft and providing
544 valuable comments on the presented research.

REFERENCES

- 545 Amit, D. J. and Brunel, N. (1995). Learning internal representations in an attractor neural network with
546 analogue neurons. *Network: Computation in Neural Systems* 6, 359–388. doi:10.1088/0954-898X/6/3/
547 004
- 548 Antolík, J. and Bednar, J. A. (2011). Development of maps of simple and complex cells in the primary
549 visual cortex. *Frontiers in Computational Neuroscience* 5, 1–19. doi:10.3389/fncom.2011.00017
- 550 Ben-Yishai, R., Bar-Or, R. L., and Sompolinsky, H. (1995). Theory of orientation tuning in visual cortex.
551 *Proceedings of the National Academy of Sciences of the United States of America* 92, 3844–3848.
552 doi:10.1073/pnas.92.9.3844
- 553 Binzegger, T., Douglas, R. J., and Martin, K. a. C. (2004). A quantitative map of the circuit of cat primary
554 visual cortex. *The Journal of neuroscience : the official journal of the Society for Neuroscience* 24,
555 8441–53. doi:10.1523/JNEUROSCI.1400-04.2004
- 556 Budd, J. M. and Kisvárday, Z. F. (2001). Local lateral connectivity of inhibitory clutch cells in layer 4 of
557 cat visual cortex (area 17). *Experimental Brain Research* 140, 245–250. doi:10.1007/s002210100817
- 558 Buzás, P., Kovács, K., Ferecskó, A. S., Budd, J. M. L., Eysel, U. T., and Kisvárday, Z. F. (2006).
559 Model-based analysis of excitatory lateral connections in the visual cortex. *The Journal of comparative
560 neurology* 499, 861–81. doi:10.1002/cne.21134
- 561 Durstewitz, D., Seamans, J. K., and Sejnowski, T. J. (2000). Neurocomputational models of working
562 memory. *Nature neuroscience* 3, 1184–1191. doi:10.1038/81460
- 563 Goodhill, G. J. (2007). Contributions of theoretical modeling to the understanding of neural map
564 development. *Neuron* 56, 301–311
- 565 Grabska-Barwinska, A. and von der Malsburg, C. (2008). Establishment of a scaffold for orientation maps
566 in primary visual cortex of higher mammals. *The Journal of neuroscience : the official journal of the
567 Society for Neuroscience* 28, 249–57. doi:10.1523/JNEUROSCI.5514-06.2008
- 568 Huth, A. G., Heer, W. A. D., Griffiths, T. L., Theunissen, F. E., and Jack, L. (2016). Natural speech reveals
569 the semantic maps that tile human cerebral cortex. *Nature* 532, 453–458. doi:10.1038/nature17637
- 570 Kang, K., Shelley, M., and Sompolinsky, H. (2003). Mexican hats and pinwheels in visual cortex.
571 *Proceedings of the National Academy of Sciences of the United States of America* 100, 2848–53.
572 doi:10.1073/pnas.0138051100
- 573 Kaschube, M., Schnabel, M., Löwel, S., Coppola, D. M., White, L. E., and Wolf, F. (2010). Universality
574 in the Evolution of Orientation Columns in the Visual Cortex. *Science* 330, 1113–1116. doi:10.1126/
575 science.1194869.Universality
- 576 Kohonen, T. (1982). Self-organized formation of topologically correct feature maps. *Biological Cybernetics*
577 43, 59–69. doi:10.1007/BF00337288
- 578 Law, J. S. (2009). *Modeling the Development of Organization for Orientation Preference in Primary Visual
579 Cortex*. Ph.D. thesis, School of Informatics, The University of Edinburgh, Edinburgh, UK
- 580 KEY: law:phd09
- 581 ANNOTATION: Abstract and possibly title need to be updated to match the final version filed
- 582 Levy, R. B. and Reyes, A. D. (2011). Coexistence of lateral and Co-Tuned inhibitory configurations in
583 cortical networks. *PLoS Computational Biology* 7. doi:10.1371/journal.pcbi.1002161
- 584 Miikkulainen, R., Bednar, J. A., Choe, Y., and Sirosh, J. (2005). *Computational Maps in the Visual Cortex*.
585 (Springer)
- 586 Muir, D. R. and Cook, M. (2014). Anatomical Constraints on Lateral Competition in Columnar Cortical
587 Architectures. *Neural computation* 1872, 1840–1872. doi:10.1162/NECO

- 588 Ohana, O., Portner, H., and Martin, K. a. C. (2012). Fast recruitment of recurrent inhibition in the cat
589 visual cortex. *PLoS ONE* 7. doi:10.1371/journal.pone.0040601
- 590 Rutishauser, U., Slotine, J.-J., and Douglas, R. J. (2012). Competition through selective inhibitory
591 synchrony. *Neural computation* 24, 2033–52. doi:10.1162/NECO_a_00304
- 592 Somers, D. C., Nelson, S. B., and Sur, M. (1995). An emergent model of orientation selectivity in cat visual
593 cortical simple cells. *The Journal of neuroscience : the official journal of the Society for Neuroscience*
594 15, 5448–5465
- 595 Stevens, J.-L. R., Law, J. S., Antolík, J., and Bednar, J. A. (2013). Mechanisms for stable, robust, and
596 adaptive development of orientation maps in the primary visual cortex. *The Journal of neuroscience : the
597 official journal of the Society for Neuroscience* 33, 15747–66. doi:10.1523/JNEUROSCI.1037-13.2013
- 598 Swindale, N. V. (1996). The development of topography in the visual cortex: A review of models. *Network:
599 Computation in Neural Systems* 7, 161–247
- 600 von der Malsburg, C. (1973). Self-organization of orientation sensitive cells in the striate cortex. *Kybernetik*
601 14, 85–100