



**University of  
Zurich**<sup>UZH</sup>

**Zurich Open Repository and  
Archive**

University of Zurich  
University Library  
Strickhofstrasse 39  
CH-8057 Zurich  
[www.zora.uzh.ch](http://www.zora.uzh.ch)

---

Year: 2014

---

## **Anatomical constraints on lateral competition in columnar cortical architectures**

Muir, Dylan R ; Cook, Matthew

DOI: [https://doi.org/10.1162/NECO\\_a\\_00613](https://doi.org/10.1162/NECO_a_00613)

Posted at the Zurich Open Repository and Archive, University of Zurich

ZORA URL: <https://doi.org/10.5167/uzh-107834>

Journal Article

Published Version

Originally published at:

Muir, Dylan R; Cook, Matthew (2014). Anatomical constraints on lateral competition in columnar cortical architectures. *Neural Computation*, 26(8):1624-1666.

DOI: [https://doi.org/10.1162/NECO\\_a\\_00613](https://doi.org/10.1162/NECO_a_00613)

## Anatomical Constraints on Lateral Competition in Columnar Cortical Architectures

Dylan R. Muir

*dylan.muir@unibas.ch*

Matthew Cook

*cook@ini.phys.ethz.ch*

*Institute of Neuroinformatics, University of Zürich and ETH Zürich,  
Winterthurerstrasse 190, 8057 Zürich, Switzerland*

Competition is a well-studied and powerful mechanism for information processing in neuronal networks, providing noise rejection, signal restoration, decision making and associative memory properties, with relatively simple requirements for network architecture. Models based on competitive interactions have been used to describe the shaping of functional properties in visual cortex, as well as the development of functional maps in columnar cortex. These models require competition within a cortical area to occur on a wider spatial scale than cooperation, usually implemented by lateral inhibitory connections having a longer range than local excitatory connections. However, measurements of cortical anatomy reveal that the spatial extent of inhibition is in fact more restricted than that of excitation. Relatively few models reflect this, and it is unknown whether lateral competition can occur in cortical-like networks that have a realistic spatial relationship between excitation and inhibition. Here we analyze simple models for cortical columns and perform simulations of larger models to show how the spatial scales of excitation and inhibition can interact to produce competition through disynaptic inhibition. Our findings give strong support to the direct coupling effect—that the presence of competition across the cortical surface is predicted well by the anatomy of direct excitatory and inhibitory coupling and that multisynaptic network effects are negligible. This implies that for networks with short-range inhibition and longer-range excitation, the spatial extent of competition is even narrower than the range of inhibitory connections. Our results suggest the presence of network mechanisms that focus on intra-rather than intercolumn competition in neocortex, highlighting the need for both new models and direct experimental characterizations of lateral inhibition and competition in columnar cortex.

---

D.R.M. is now at Biozentrum, University of Basel, Basel, Switzerland.

## 1 Introduction

---

How can we expect to tease apart the mechanisms of neocortex? The only justification for our hubris is the observation that each area of cortex exists not as a solitary and unique design, but instead adopts a variation on a shared but elusive theme (Ramón y Cajal, 1892; DeFelipe & Jones, 1998; Mountcastle, 2003; Douglas & Martin, 2007; Muir et al., 2011). Known as the *canonical cortical microcircuit* (Douglas, Martin, & Whitteridge, 1989), the notion that every cortical area reproduces a common network motif kindles a hope that each cortical area might also perform its computational role using a common form of computational dynamics.

Competition between the activity of several neurons is a well-studied mechanism that has been suggested as a canonical computation for cortex due to the useful theoretical properties of competitive interactions and the relative simplicity of implementing competition with neuronal elements. Two neurons are said to be in competition with each other if the activity of one of the neurons directly or indirectly reduces the activity of the other. Although two cross-connected inhibitory neurons have this property, more attention is usually paid to the information-encoding properties of excitatory neurons of cortex. These neurons form the vast majority of projections to and from other cortical areas and subcortical nuclei, and so they could be considered to embody the result of a cortical area's computation. The simplest networks that implement competition consist of two or more excitatory neurons coupled to a single common inhibitory neuron (Coultrip, Granger, & Lynch, 1992; Douglas, Mahowald, & Martin, 1994; Douglas & Martin, 2007). Depending on the parameters of the network, the excitatory neurons can be placed in a competitive regime. The excitatory neuron that receives the strongest external input will then effectively suppress the activity of other excitatory neurons through disinaptic inhibition via the shared inhibitory neuron. In extreme cases a single excitatory neuron—the “winner”—will be active, and all other excitatory neurons will be inactive. This network behavior is known as hard winner-take-all (WTA) behavior.

The set of excitatory neurons can be placed in a geometric (or topological) space, with a distance-based neighborhood function defining cooperative connections among the excitatory neurons. The WTA then behaves as an associative memory network, where fixed points shaped by the excitatory neighborhood function are placed in competition. In this regime, a “winner” is no longer a single neuron but a set of cooperating excitatory neurons. The autoassociative function of these networks brings desirable information processing properties such as noise rejection and analog signal restoration (Douglas & Martin, 2007). This network architecture has been proposed as a model for cortical computation by interpreting each excitatory neuron as representing a cortical column. For example, ring models of orientation tuning for primary visual cortex assign a different preferred orientation to each excitatory neuron (Douglas et al., 1994, Ben-Yishai, Bar-Or, & Sompolinsky,

1995; Somers, Nelson, & Sur, 1995). More elaborated models consist of multiple subnetworks, each spanning the full range of preferred orientations (a hypercolumn), with global competition within each hypercolumn and feedforward inhibition between competing hypercolumns (Lundqvist, Rehn, Djurfeldt, & Lansner, 2006; Lundqvist, Compte, & Lansner, 2010). Models for working memory in prefrontal cortex have been proposed using similar cooperative and competitive mechanisms (Amit & Brunel, 1995; Durstewitz, Kelc, & Güntürkün, 1999; Compte, Brunel, Goldman-Rakic, & Wang, 2000; Miller, Brody, Romo, & Wang, 2003). For a review, see Durstewitz, Seamans, & Sejnowski, 2000). In these models, a pattern of activity is embedded in the configuration of recurrent excitatory connections. This pattern of activity is then self-sustaining once activated by external input, and wide-ranging inhibitory feedback is used to ensure stability and robustness of the stored activity pattern.

Defining WTA models in these ways makes several assumptions about the anatomy and physiology of inhibition in columnar cortex. First, inhibitory projections are wide ranging in the WTA models described. Global inhibition can be softened by adopting a Mexican hat network connectivity profile, whereby a point in the network sends spatial inhibitory connections extending over a longer range than excitatory connections (Somers et al., 1995; Sperling, 1970) (see Figure 1A). However, inhibitory neurons in neocortex are mostly limited in their lateral extent, making projections either vertically between cortical layers or proximal to their somata (Lund, 1987; Lund & Wu, 1997; Lund & Yoshioka, 1991; Lund, Hawken, & Parker, 1988; DeFelipe, 2002; Douglas & Martin, 2004; Markram et al., 2004; Douglas & Martin, 2009). Some inhibitory neurons are coupled via electrical synapses called gap junctions, providing an effective excitatory coupling across an inhibitory population (Galarreta & Hestrin, 1999; Gibson, Beierlein, & Connors, 1999). This could theoretically serve to widen the effective spatial extent of inhibition. Unfortunately, the electrical connections are weak (Galarreta & Hestrin, 1999; Gibson et al., 1999), sparse compared with the number of chemical synapses made by a neuron (Fukuda, Kosaka, Singer, & Galuske, 2006), and mostly absent in adult animals (Connors, Bernardo, & Prince, 1983; Peinado, Yuste, & Katz, 1993). For these reasons, gap junctions cannot generally be relied on as a substrate for long-range spreading of inhibitory influences. Some models address this concern through long-range excitatory projections that selectively target inhibitory neurons (Li, 1998, 2002; Rutishauser, Slotine, & Douglas, 2012) or by including instantaneous disinaptic inhibition while neglecting f-inhibitory recurrence (Pinto & Ermentrout, 2001; Kang, Shelley, & Sompolinsky, 2003; Levy, Reyes, & Alex, 2011). However, reconstructions of cortical neurons that engage in long-range excitatory projections do not reveal evidence for neuron-class-specific connections (Kisvárdy et al., 1986; but see Bock et al., 2011).

Second, the physiology of inhibition is either untuned or broadly tuned in WTA models, so that inhibition is activated similarly by any input to

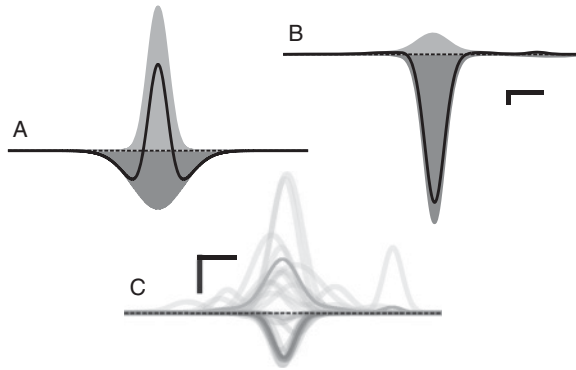


Figure 1: Mexican hats versus cortical anatomy. (A) A classical Mexican hat profile of lateral connectivity, with short-range excitation (light gray; positive) and longer-range inhibition (dark gray; negative) leading to a net profile (black line) that is vaguely reminiscent of a sombrero. (B) The effective lateral coupling for a slice through layer 2/3 of cat visual cortex (Binzegger, Douglas, & Martin, 2007). Strong local inhibition coupled with long-range excitation results in a net profile with local inhibitory dominance and much weaker lateral excitatory dominance. Inhibitory synapses were estimated to be 10 times stronger than excitatory synapses (Binzegger, Douglas, & Martin, 2009). (C) Raw synaptic profiles of excitation (positive curves) and inhibition (negative curves) for layer 2/3 of cat visual cortex (Binzegger et al., 2007). Faint curves show data from individual reconstructed neurons; darker curves indicate the average over the set of reconstructed neurons. Scale bars in panels B & C: horizontal: 500  $\mu\text{m}$ , vertical: 2 synapses per  $50 \times 50 \times 50 \mu\text{m}$  volume.

cortex, a stance that is not supported by *in vivo* single-cell electrophysiology (Mariño et al., 2005).

Finally, in the simplest WTA networks, inhibitory neurons receive no external input. In columnar cortex, inhibitory neurons certainly receive input from outside the layer their soma resides in, from both other layers and other cortical and subcortical structures (Binzegger, Douglas, & Martin, 2004). There is no evidence that feedforward inputs specifically target excitatory or inhibitory classes (Freund, Martin, Somogyi, & Whitteridge, 1985; Freund, Martin, & Whitteridge, 1985; Anderson, Dehay, Friedlander, Martin, & Nelson, 1992).

To seriously consider competition as a canonical computational mechanism for cortex, this potential conflict between model assumptions and cortical anatomy must be resolved. When is it possible for two points in columnar cortex to be in competition? In this letter, we study this question in both very small networks that can be analyzed mathematically and in larger networks via simulations.

In section 2 we present linear-threshold network models for groups of two or three cortical columns and determine analytically the conditions under which two columns can be in competition through disynaptic or multisynaptic inhibition. In addition to examining these simple models that are tractable for direct analysis, we also present simulations in larger 1D and 2D models in section 3. The parameters in these models are designed to capture the anatomical issue of the relative extent of lateral excitatory and inhibitory projections. Through piecewise linear systems analysis of the tractable models, we obtain bounds on parameter regimes that permit disynaptic inhibitory competition between two cortical columns, and we then compare these results with the larger simulation models.

Surprisingly, we find that the presence and strength of cooperation or competition between two columns in a network is determined primarily by the direct excitatory and inhibitory coupling between the two columns, with indirect network effects only weakly modulating this direct cooperation or competition. We refer to this phenomenon as the direct coupling effect. Our results provide a simple intuitive rule of thumb for understanding cooperation and competition between two columns in a large network: that cooperative and competitive effects arise primarily from the direct influence of one column on another.

## 2 Analytical Models

---

**2.1 A Cortical "Column."** The concept of a cortical column is primarily functional (Mountcastle, Berman, & Davies, 1955). In cat, monkey, ferret, tree shrew, and many other higher mammals, neurons existing on a line perpendicular to the pia share many commonalities in their function. Aside from the canonical example of cat somatosensory cortex (Mountcastle et al., 1955), neurons in visual cortex exhibit this strong columnar organization by sharing the orientation preference of their vertically adjacent neighbors (Hubel & Wiesel, 1968). However, this fact should not be interpreted to mean that a "column" is an isolated unit, either functionally or anatomically. A lateral displacement of even the width of a neuron's soma is sufficient to record a measurable difference in orientation preference in visual cortex, implying that a functional column is about as small as it can possibly be (Hubel & Wiesel, 1968). Anatomically, projections from the neurons in a column are diffuse. Although many intrinsic cortical projections are made across laminae, they nevertheless span a horizontal distance much larger than the size of a single soma in columnar and rodent cortex (Weliky & Katz, 1994; Hellwig, 2000; Lund, Angelucci, & Bressloff, 2003; Thompson & Bannister, 2003; Holmgren, Harkany, Svennenfors, & Zilberter, 2003; Boucsein, Nawrot, Schnepel, & Aertsen, 2011). Input projections to cortex also do not treat single columns as independent entities; single input fibers projecting from the LGN cover large areas in primary visual cortex (Lund et al., 2003).

The notable exception is rodent somatosensory cortex, where input fibers carrying information from single whiskers project to large, nonoverlapping regions within layer 4 known as barrels.

In this letter, we take a column to be a small region within a neocortical area of a higher mammal, of the minimum size such that the function of each column is homogeneous but that neighboring columns can have different functions. This allows us to simplify the neurons in a column to a small population of interacting excitatory and inhibitory units. However, our simulations incorporate the fact that single columns make lateral projections to a large number of neighboring columns and receive input from a similar large number of neighbors. The function of our column model is discrete, but the virtual anatomical inputs and outputs of our columns are highly overlapping.

**2.2 Model Simplifications.** We assume that a column of cortical tissue can be reduced to a population of excitatory neurons and a population of inhibitory neurons. We model the average activity of these two classes with two linear-threshold units, which are known to be a good approximation to the I–F (current to firing rate) curve for an adapted cortical neuron (Ermentrout, 1998a). The differing proportions of inhibitory and excitatory neurons in cortex are modeled by adding a factor to our synaptic weights to correct for this. Although different neuron classes may have different time constants of activation, we will show that the possibility of competition is independent of these time constants.

We assume that neurons connect to each other based on opportunity and without bias, an assumption known as Peters' rule (Peters, 1979; Braitenberg & Schüz, 1991). This implies that an excitatory projection to a point in cortex forms synapses with both excitatory and inhibitory neurons at that location, without preference for a particular neuron class. This is the most conservative assumption to make regarding neural connectivity. Although some specific connections are known to exist in cortex (Fairén & Valverde, 1980; Somogyi, Freund, & Cowey, 1982; Stepanyants, Tamás, & Chklovskii, 2004; Morishima, Morita, Kubota, & Kawaguchi, 2011), the majority of local and lateral connections do not show evidence of class-specific targeting (Kisvárdy et al., 1986; Binzegger et al., 2004). We further assume that input to a cortical column targets both excitatory and inhibitory populations, without bias (Freund, Martin, Somogyi et al., 1985; Freund, Martin, & Whitteridge, 1985; Kisvárdy et al., 1986; Anderson et al., 1992; Keller & Asanuma, 1993).

We assume that connections between columns in cortex are arranged predominantly spatially, such the coupling strength between two points decreases monotonically with distance. This is of course not true for a single cortical neuron, but is a reasonable aggregate assumption based on Peters' rule (Binzegger et al., 2004; Perin, Berger, & Markram, 2011).

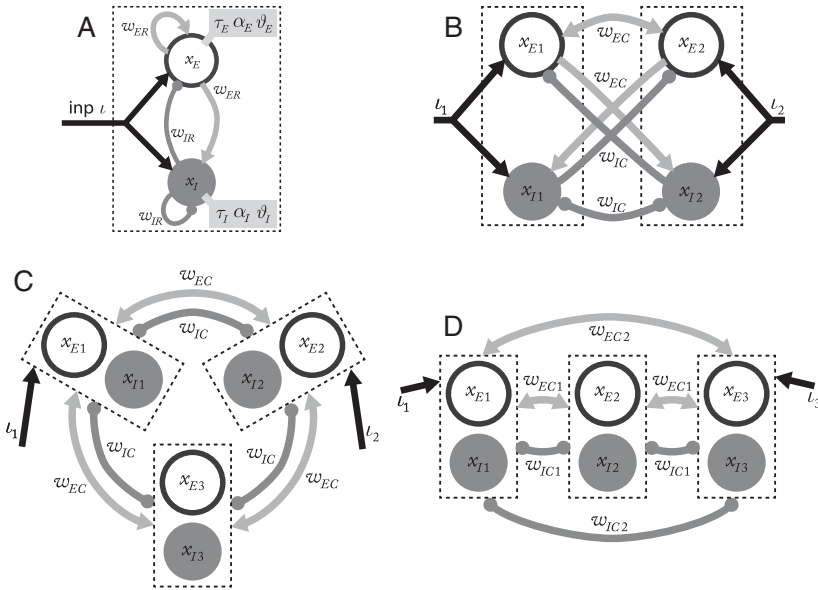


Figure 2: Columnar models analyzed in this letter. All analytical models are constructed of column elements (A), each composed of an excitatory (outlined circles) and an inhibitory (filled circles) linear-threshold unit (Wilson & Cowan, 1973). A single column is internally coupled with recurrent excitatory (pointed arrowheads) and inhibitory (circular arrowheads) weights. Excitatory input ( $l_n$ ) is provided equally to all units in a column. (B) Two interacting columns (internal connections within a column not shown). (C) A ring composed of three columns with identical connections between columns. (D) A chain of three columns with different weights for short and long connections. Parameters are defined in Table 1.

**2.3 Basic Column Model.** The foundation of the analytical models presented here is a simplified version of a cortical column, consisting of a coupled pair of an excitatory and an inhibitory linear-threshold unit (Wilson & Cowan, 1973; Landsman, Neftci, & Muir, 2012; see Figure 2A). These units are designed to correspond in behavior to the average excitatory neuron and average inhibitory neuron in the small population of neurons within a single cortical column of very narrow width. The excitatory and inhibitory pair of units are assumed to exist at the same point on a cortical sheet, so that each unit has the same average self-connectivity as with the other unit of the pair. In this letter, when we refer to “self-excitation” and “self-inhibition,” we mean recurrent excitation within the population of neurons that is represented by a single excitatory or inhibitory unit.



Table 1: Analytical Model Parameters.

$w_{ER}$	Recurrent synaptic weight from an excitatory unit to the units in the same column
$w_{ECm}$	Synaptic weight from an excitatory unit to the units in another cortical column $m$ steps away
$w_{IR}$	Recurrent synaptic weight from an inhibitory unit to the units in the same column
$w_{ICm}$	Synaptic weight from an inhibitory unit to the units in another cortical column $m$ steps away
$\tau_n$	Time constant of unit $n$
$\alpha_n$	Activation gain of unit $n$
$\vartheta_n$	Activation threshold of unit $n$
$\iota_n$	External input current to column $n$

The column dynamics are governed by the system of equations,

$$\tau_E \cdot \dot{x}_E + x_E = w_{ER} \cdot \alpha_E [x_E - \vartheta_E]^+ - w_{IR} \cdot \alpha_E [x_I - \vartheta_I]^+ + \iota, \quad (2.1)$$

$$\tau_I \cdot \dot{x}_I + x_I = w_{ER} \cdot \alpha_E [x_E - \vartheta_E]^+ - w_{IR} \cdot \alpha_E [x_I - \vartheta_I]^+ + \iota, \quad (2.2)$$

where  $x_E$  and  $x_I$  are the internal state of the excitatory and inhibitory unit in the pair;  $[x]^+$  denotes the linear-threshold transfer function  $[x]^+ = \max(x, 0)$ ; and other parameters are as described in Table 1.

**2.4 Summary of Analytical Method.** The details of our analysis are presented in appendix A. Briefly, we construct a set of differential equations embodying one of the columnar network models shown in Figure 2. Since the systems are piecewise-linear, a Jacobian of the system can be constructed for each linear partition in the state space defined by the activity of all units (Hahnloser, 1998b). The real parts of the eigenvalues and trace of the Jacobians determine when the system is stable in a bounded input-bounded output (BIBO) sense. The BIBO stability criterion guarantees that the system will not approach infinite activity for a finite input. For the simple systems shown in Figure 2, the set of eigenvalues can be described analytically. This allows constraints on each of the system parameters to be found that guarantee BIBO stability.

To determine whether two columns in a model are in competition, we measure the activity increase or decrease of activity in the excitatory unit in column 2 produced by an increase in the input to column 1 (i.e.,  $\partial \bar{x}_{E2} / \partial \iota_1$ ). The value of this partial derivative depends on the system parameters, including the weights between the two columns. When the partial derivative is negative, increasing the input to column 1 leads to a decrease in the activity of column 2 via disinaptic inhibition or other network effects. Due to the symmetric nature of our models, the same interaction would also occur

in the reverse direction from column 2 to column 1. If increasing the input to either column decreases the activity of the other, we say the columns are in competition. Again, for our simple models, we can find closed analytical forms for the partial derivative  $\partial \bar{x}_{E2}/\partial \iota_1$  and so can solve for simple conditions on each of the system parameters corresponding to competitive interactions.

By combining the conditions for BIBO stability and for competition, we can determine what parameter constraints ensure that a model operates in a stable winner-take-all (WTA) mode. Our method for evaluating competition operates on system fixed points and does not take into account transient modes. However, we also identify when a system is expected to operate in a nonoscillatory mode such that transient dynamics can be ignored.

**2.5 Two-Column Analytical Model.** Analysis of the two-column model is described in detail in section A.2. This model examines two points in a columnar cortical system in an abstract form, including only the direct excitatory and inhibitory connections between the two columns (see Figure 2B;  $w_{EC}$  and  $w_{IC}$ , respectively). More complex network interactions contributed by intermediate columns are excluded in this minimalistic model.

The question explored by the simple two-column model is this: When can two points in columnar cortex be in direct competition, disregarding network connectivity external to the columns in question? To answer that question, we examine the fixed-point solutions of the two-column model to determine its behavior and examine the Jacobian of the model network to determine its stability properties (see Figure 3). For the two columns to be in competition, an input given only to column 1 should reduce the activity of column 2, and this must occur in a network that is stable in a bounded-input, bounded-output sense (BIBO stability).

We found that two points in a columnar system can be in competition only when the inhibitory coupling  $w_{IC}$  between the two columns is stronger than the excitatory coupling  $w_{EC}$ . This is a strong result that does not depend on the thresholds for excitation and inhibition or on the time constants of excitation and inhibition (see section A.2). We found also that hard-WTA competitive behavior (i.e., one column is silenced by the other) can occur only for a certain range of input differentials between the two columns. This result implies that for non-saturating columnar systems, there is no parameter regime that guarantees hard-WTA operation regardless of the network input; networks operate in a soft- or hard-WTA regime depending on the difference in input between two columns. We also found that nonzero thresholds for excitation and inhibition cannot introduce or abolish competition. However, they can establish a memory state in an already competitive network. A network in this regime can maintain suprathreshold activity without input once a winner has been determined.

The two-column model described here ignores contributions from other columns across the cortical surface. We considered whether multisynaptic

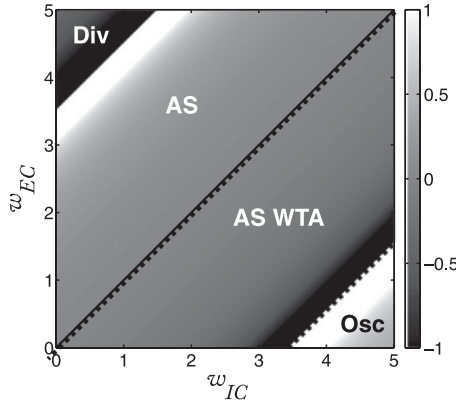


Figure 3: Stability and competitive parameter regimes for the two-column model. Shown is the derivative  $\partial x_{E2}^- / \partial u_1$ , the net effect on column 2 for an increase in the input to column 1. Competition is possible when the derivative is negative (indicated within dashed boundaries). This condition is satisfied when the mutual inhibitory coupling  $w_{IC}$  is stronger than the mutual excitatory coupling  $w_{EC}$ . Shown is the Part[11] derivative; the parameter regimes that ensure competition for Part[10] are identical, but the shape of the derivative surface is linear. Div: BIBO unstable (divergent); AS: Asymptotically stable, no competition; AS WTA: asymptotically stable, competitive regime; Osc: oscillatory dynamics caused by overly strong inhibition.  $w_{ER} = 2.5$ ;  $w_{IR} = 5$ . Other parameters are not relevant for the presence of competition.

inhibition provided by intermediate columns could be strong enough to drive competition between two points by exploring more elaborate models that include intermediate columns, described in the following sections.

**2.6 Three-Column Ring Analytical Model.** The two-column model neglects the effect of network interactions that might be mediated by additional columns. For example, competition between two distant points in a columnar system could be mediated by a third column placed at an intermediate location. We explored this possibility by designing networks containing three columns. The first such network had three columns arranged in a ring (see Figure 2C). The connections in the model are homogeneous, such that every column is equivalent. Competition in this network is sought between two of the three columns.

This model is analyzed in detail in section A.3. We found, just as for the two-column model described above, that competition can occur only when the direct inhibitory coupling between the two columns is stronger than the direct excitatory coupling. The third column cannot provide a sufficient indirect inhibitory contribution to mediate competition. We call this the

direct coupling effect: the interaction between two columns is primarily determined by direct excitatory and inhibitory coupling.

However, since the three columns in the model examined here were arranged in a ring, it is possible that the direct excitatory and inhibitory connections between the two columns that should compete were unrealistically strong. We therefore examined another three-column model with the columns arranged in a line rather than a continuous ring.

**2.7 Three-Column Chain Analytical Model.** The direct connections between two distant columns in cortex may be weak; certainly two proximal columns are expected to have stronger coupling than two distant columns. We examined a more general form of the three-column network, where three columns are arranged in a linear chain (see Figure 2D). Competition was sought between the columns at the two ends of the chain (edge columns). As for the previous model, analysis of this network indicated whether competition between two distant columns in cortex (represented by the edge columns) could be driven by the activity of an intermediate column (represented by the central column). The principal difference from the previous model was in the structure of the connections between the two edge columns. These columns shared symmetric mutual coupling weights ( $w_{EC2}$  and  $w_{IC2}$ ), which were not constrained to be equal to the weights between the central and edge columns ( $w_{EC1}$  and  $w_{IC1}$ ). The three-column chain model therefore approximated the physical arrangement of three equally spaced columns, such that the two edge columns were further apart and therefore more weakly connected.

This model is analyzed in detail in section A.4. Surprisingly, despite the potentially weaker coupling between the edge columns, the central column was still not able to drive competition between them. This appears unintuitive, but is caused by the assumption of homogeneous local connections between neighboring columns. For one edge column to indirectly inhibit the other, it must first activate the central column. This implies that the excitatory coupling from edge to center columns should be stronger than the inhibitory coupling. Likewise, since connections in cortex are assumed to be homogeneous, the connections from the center to both edge columns are then also dominated by excitation. This implies that driving an edge column will recruit both excitation and inhibition in the central column, but that driving the central column will also activate the edge columns. It is therefore not possible to indirectly activate the central column by driving an edge column and have a net suppressive effect on the opposite edge column.

We also examined the conditions required for competition between an edge column and the center column. In this configuration, the two end columns could be positioned close together in cortical space with the central column equidistant (but far) from both end columns. Coupling between the edge columns could be arranged to be dominated by inhibition, with

excitatory coupling between edge and center columns, which one might assume would lead to indirect competition between edge and center columns. However, for the indirect competition to outweigh direct excitation, direct inhibition between edge columns would have to be so strong that it would in fact lead to complete suppression of one edge column, thereby eliminating the effect. Thus the direct coupling effect applies also to this configuration; coupling between center and edge columns must be dominated by inhibition for competition to be present between them.

Once again, we must conclude from this analytical model that for competition to occur between two columns, we need consider only the direct column-to-column coupling, which must be dominated by inhibition.

### 3 Simulation Models

---

The simple analytical models we have described had only a few units and directly modeled at most three columns. The constraints for stability and competition were remarkably similar from the simplest to the most complex analytical model, implicating the direct excitatory and inhibitory coupling over multisynaptic network interactions. But how predictive are these simple models for a larger-scale 1D or 2D simulation composed of many columns, and with realistic spatial profiles of connectivity? The models discussed so far treated a cortical column as an isolated entity; the interactions between several columns were divorced from the remainder of a cortical area. We would like to understand how competition is mediated across a homogeneous cortical surface. We would also like to address the possibility that the summed effect of inhibition from many columns across a larger model might succeed in driving competition where a single intermediate column cannot.

To answer these questions, we simulated linear and two-dimensional models composed of columns with the same structure as the basic analytical column (see Figure 2A). In place of simple point-to-point connectivity, we introduced spatial profiles of synaptic connections based on gaussian fields (see Figures 4A and 4B) with synaptic parameters estimated from the experimental literature (see appendix B).

**3.1 Presence of Competition in Simulated Networks.** The series of analytical models described in section 2 suggest that competition through disinaptic inhibition can occur between two columns only when the direct inhibitory coupling between those columns is stronger than the direct excitatory coupling—the direct coupling effect. In this section, we explore how well that prediction applies to networks that include spatial profiles of synaptic weights from excitatory and inhibitory units that extend across many columns. Connections in these models were made homogeneously, meaning that every point in the network had the same spatial profile of excitatory and inhibitory coupling. For networks with this structure,

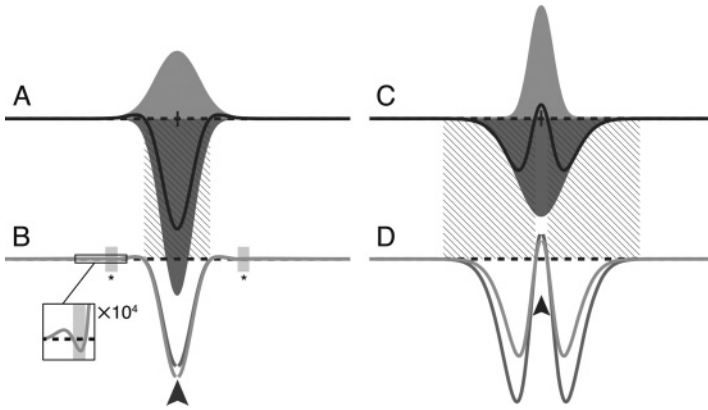


Figure 4: The direct coupling effect predicts the presence and strength of competition in homogeneous linear networks. Linear networks were designed with gaussian synaptic profiles of excitation (light gray areas in A and C) and inhibition (dark gray areas in A and C) projecting identically from each point. In the first network (left, A and B), the spread of inhibition was narrower than that of excitation, indicated by the difference between the profiles of excitation and inhibition (black curves). A second network was constructed with a wider profile of inhibition (right, C and D). The profile of competition was probed, using point stimuli, by injecting positive current into both units of a single column (arrowheads in B and D). The net current received by each unit is shown for excitatory units (light curves in B and D) and inhibitory units (dark curves in B and D), once the network has reached fixed-point equilibrium. When the internal state is negative, that unit is effectively suppressed by the point stimulus, implying competition between that unit and the stimulated point. Hatching indicates regions of the linear network where the inhibitory coupling with the stimulated location is stronger than the excitatory coupling—the locations for which competition should be possible according to our analytical predictions. Asterisks and shading in B indicate regions where multicolumnar interactions cause weak competition, an effect that is not present in the analytical networks of section 2. The inset in B shows a heavily magnified version of this region (vertical magnification  $\times 10^4$ ). Parameters for these simulations are given in Table 3 in appendix B.

stability was predicted well by the behavior of a single column, under the parameter transformation that the sum of the weights from a single point was equivalent to the weights in the single-column model (Landsman et al., 2012),

$$w_{E,\text{eff}} = \int_j w_{Eji} \quad \text{and} \quad w_{I,\text{eff}} = \int_j w_{Iji}, \quad (3.1)$$

where  $w_{Eji}$  and  $w_{Iji}$  are the excitatory and inhibitory projections from point  $i$  to point  $j$ , respectively. We find that the stability criteria given for our networks hold regardless of the spatial pattern of a stimulus. In other words, local columnar inhibitory feedback is able to stabilize local excitatory activity, even in the presence of wide-ranging excitatory input to a column in the model.

We examined the presence and absence of competition in these linear models by injecting a point excitatory stimulus into a single column of a quiescent network with stable, nonoscillatory dynamics (see section 2). Once the network reached the stable fixed point, we measured the net current arriving at each column in the network, provoked by the point stimulus passing through the entire network. Two locations are in competition if providing a positive input current to a source column results in a net suppressive effect on a target column, indicated by a net negative current arriving at the target column. Since the coupling patterns of our networks are homogeneous and symmetric, the effect of injecting a point excitatory stimulus is identical between any two locations on the network when using either location as the source. Therefore, locations across the network for which the effect of a point stimulus is to provide a net negative input current are in mutual competition with the stimulated column.

We designed linear networks with spatial profiles of lateral connectivity encompassing lateral excitation and lateral inhibition (see Figures 4A and 4C). Each network consisted of 360 columns spaced at a  $12.5 \mu m$  pitch. The spatial range of lateral excitation for both models shown in Figure 4 was  $\sigma_E = 750/4 \mu m$ ; the spatial range of inhibition was  $\sigma_I = 550/4 \mu m$  for the network with local inhibition (see Figure 4A) and  $\sigma_I = 800/4 \mu m$  for the network with lateral inhibition. Total synaptic strength for each neuron was estimated to be realistic for cat visual cortex (see Table 3) synaptic coupling between columns was determined by the mean field estimate under the assumption of gaussian connectivity profiles, normalized to the total estimated synaptic strength. Injecting excitatory input currents into single columns of these models produced regions across the networks that received net excitatory and inhibitory currents at a steady state through the combined interactions of many columns of the networks.

We found that under realistic spatial profiles of lateral excitation and short-range inhibition (see Figures 4A and 4B), and under a Mexican hat arrangement with lateral inhibition (see Figures 4C and 4D), the direct coupling effect predicted a central region of competition that matched the simulation results to within the spatial resolution of the simulation. However, in the case of lateral excitation, a region of competition mediated by multicolumnar interactions emerged (asterisks and inset in Figure 4B). This competition occurred because a column activated by lateral excitation distant from the point stimulus can suppress activity locally through short-range inhibitory connections. However, since the gain of single synaptic connections is low, an effect relying on three or more synapses must also

be comparatively weak. Under the realistic parameters simulated here, the scale of the multisynaptic effect was at least four orders of magnitude weaker than that produced by the direct coupling effect.

We performed equivalent experiments in two-dimensional networks with symmetric gaussian profiles of lateral excitation and inhibition, with other parameters identical to the one-dimensional models. The overall patterns of competition and facilitation were qualitatively the same as for the one-dimensional linear networks (not shown).

**3.2 Accuracy of Analytical Predictions.** The direct coupling effect predicted competition for the particular weight parameters simulated in Figure 4. To determine how well the two-column analytical predictions hold for an arbitrary homogeneous model, we directly compared the numerical predictions between a linear model and our two-column model configured with identical coupling strengths. We simulated 2500 linear models with gaussian profiles of excitatory and inhibitory coupling (such as those shown in Figure 4), built with random and independent excitatory and inhibitory spatial ranges and total synaptic strengths as given in Table 3. Each model was composed of 400 columns (400 excitatory and 400 inhibitory units). We injected current into 50 pairs of columns in each model, taken in turn and spanning a range of spatial separations, and numerically computed the resulting activation fixed point. We then injected a step current into one column in the pair and numerically computed the partial derivative  $\partial \bar{x}_{E2} / \partial \iota_1$  to measure the presence and strength of competition between the pair of columns, as for the analytical models described above (see section 2.4). We then reduced the linear model to a two-column configuration by removing all weights except those within and between the units in the pair of driven columns. The derivative  $\partial \bar{x}_{E2} / \partial \iota_1$  was again computed numerically for the two-column model. Cases where the sign of the predicted and measured strengths of competition did not match indicated weight configurations where the analytical predictions did not hold.

Figure 5 shows the comparison between the strength of competition predicted under the two-column model and the strength of competition measured in the line model. The derivatives computed for the two-column model showed an impressive predictive power for the line model, such that most prediction and measurement pairs lay close to a 45 degree line passing through the origin. A small gain factor difference between predicted and measured facilitation was apparent due to the effect of recurrent amplification in network interactions in the line model. A very small proportion of simulated line models exhibited competition when the two-column model predicted facilitation, and vice versa (highlighted points in Figure 5). However, all mismatches between two-column and line model results occurred close to the origin, where interactions between the two tested columns were very weak.



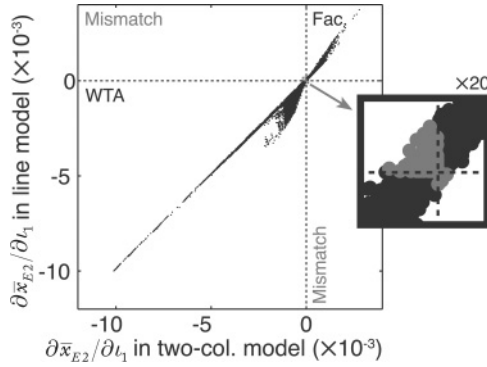


Figure 5: The two-column model is a good predictor of facilitation and competition in a line model. Each point corresponds to a numerical computation of the partial derivative  $\partial \bar{x}_{E2}/\partial l_1$ , which determines the presence and strength of competition between two columns (see section 2.4). Predictions of competition (WTA; negative  $\partial \bar{x}_{E2}/\partial l_1$ ) and predictions of facilitation (Fac.; positive  $\partial \bar{x}_{E2}/\partial l_1$ ) matched well between the two-column and line models, indicating that the direct coupling effects in the line model dominated over multicolumnar network interactions. Configurations where the two-column model predicted competition but the line model exhibited facilitation (or vice versa) are indicated by highlighted points (Mismatch). Mismatched predictions were confined to regions close to the origin, where only weak interactions between the two columns were present (i.e., neither strong competition nor facilitation). The inset shows the central region magnified 20 times to highlight the region of prediction mismatch.

#### 4 Discussion

We explored the possibility of competition between columns in simple models for columnar cortex that allow the relationship between competition and the spatial profiles of excitation and inhibition to be examined directly. Networks composed of up to three columns were analytically tractable and could be solved exactly. In this way we obtained closed-form constraints on the model parameters that permit competition to exist between two columns, which we found to involve only the direct lateral coupling between the columns. In a columnar model with homogeneous connectivity, the direct inhibitory coupling between two columns must be stronger than the direct excitatory coupling to permit competition to emerge.

In our analyses described here, we found that our toy analytical models provided a great deal of insight into the behavior of larger systems that are not tractable for analysis. In particular, we found that conditions for stability and competition are remarkably insensitive to the size of the analyzed model and continue to apply even in the context of increasingly

complex network interactions (see also Landsman et al., 2012). Surprisingly, we found that the presence and strength of competition or cooperation between two columns was primarily determined by the direct excitatory and inhibitory coupling between those columns. We observed very slight deviations from our analytical expectations in 2D and 1D models. However, the deviations due to multicolumnar network interactions were considerably weaker than the direct coupling effects predicted by our analytical models. We therefore expect that in a biologically realistic network or in cortex itself, the first-order direct coupling effects are likely to remain, while the small deviations from these effects are unlikely to be a significant factor in the face of the many noisy phenomena that influence a biological network.

We found the constraint relating inhibitory and excitatory coupling to be independent of the time constants and thresholds of excitatory and inhibitory elements in a network. However, positive excitatory thresholds introduce a subtractive influence on the fixed point of a network. This can introduce the appearance of competition if the internal state of the network is not accessible and instead the output firing rate gains  $g_1 = [x_{E1}]_+ / \iota_1$  and  $g_2 = [x_{E2}]_+ / \iota_2$  are used to evaluate the presence of competition. If both columns are driven with unequal inputs  $\iota_1$  and  $\iota_2$ , a subtractive threshold will result in the gains  $g_1$  and  $g_2$  being unequal, even if the derivatives  $\partial x_{E1} / \partial \iota_2$  and  $\partial x_{E2} / \partial \iota_1$  are equal. The difference in gains does not indicate the presence of competition through recurrent network interactions in this case, and the ratio  $g_1 / g_2$  will converge to 1 as the overall strength of input increases. Illusory competition can also occur if the inputs to the network are appropriately structured. For example, Mexican hat-shaped input can induce lateral cooperative and competitive interactions in a network without lateral inhibition (Linkser, 1986).

Increasing the length of the inhibitory time constant can lead to oscillatory dynamics (Wilson & Cowan, 1973; Hahnloser, 1998a; Tang & Tan, 2005; Landsman et al., 2012). This does not change whether the fixed points of the network express competition between columns, but can cause the network to oscillate around the fixed point. In this case, the fixed point will not be informative of the dynamics of the network and may not accurately reflect the relationship between the activity of two columns.

**4.1 Implications for Cortical Models.** Our results show that the possibility and lateral extent of disinaptic competition in cortical field models with homogeneous, nonspecific connectivity is accurately predicted by the direct difference between the spatial profiles of excitation and inhibition emerging from a point. The predictions for lateral excitation and inhibition architectures are illustrated in Figure 6. Classical lateral-inhibition architectures produce an annulus of competition surrounding a core of facilitation, depending on the relative strengths of the excitatory and inhibitory components (see Figures 6A–6C). This mechanism has been used via

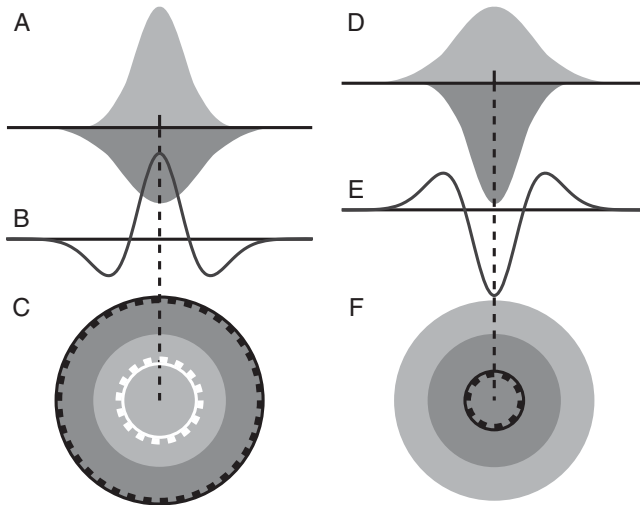


Figure 6: Lateral competition is highly spatially constrained by local inhibition. (A) If the profile of excitation (light shading; positive) arising from a point is narrower than the profile of inhibition (dark shading; negative), then the net influence over space adopts the classical “mexican hat” profile (B). Seen from above C, competition is possible in an annulus (between the white and black dashed circles in C) surrounding a core of cooperation. However, if the profile of inhibition is narrower than that of excitation (D), the net influence adopts an inverse Mexican hat (E), and the spatial occurrence of competition is dramatically reduced. Strong competition through disynaptic inhibition can occur only in the region local to a driving neuron (within the black dashed circle in F)—indeed, narrower than the range of inhibitory projections. The result is independent of scale.

lateral-inhibition neighborhood functions in developmental models of cortical areas to provide local spatial grouping of function and medium-range decorrelation of function, and to therefore reproduce some of the form of functional maps in visual cortex (von der Malsburg, 1973; Swindale, 1982; Grabska-Barwińska & von der Malsburg, 2008; Antolík & Bednar, 2011; Plebe, 2012). The same mechanism can be used to describe pattern formation during ongoing activity in columnar cortex (Ernst, Pawelzik, Sahar-Pikielny, & Tsodyks, 2001; Pinto & Ermentrout, 2001; Blumenfeld, Bibitchkov, & Tsodyks, 2006; Baker & Cowan, 2009).

Broadly tuned or untuned inhibitory feedback has also been used in abstract competitive models to explain the intracortical emergence of sharp orientation tuning in primary visual cortex (Douglas et al., 1994; Ben-Yishai et al., 1995; Somers et al., 1995; Li, 1998, 2002). If we are to interpret these models as applying to columnar visual cortex (e.g., cat, tree shrew, ferret,

monkey), where orientation is smoothly mapped to space across the surface of area 17, then these models require competition over long distances across the cortical surface.

In contrast, the extent of competition in lateral-excitation models is expected to be even narrower than the range of local inhibition (see Figure 6F). Note that this is true in spite of the presence of widespread disynaptic inhibition in the models. Since the measured cortical architecture appears to be of this type, our results raise serious questions for all cortical models that rely on lateral or global inhibition.

Our results do not mathematically prohibit competition between columns in cortex. Not only are there extremely weak deviations from the direct coupling effect in the models we examine, but it is also certainly possible to hard-wire a model of arbitrary and asymmetric connections between columns to provide multicolumnar competition. Our models examine the expectation for homogeneous and symmetric cortical networks, reflecting the minimal assumption of opportunistic connectivity between neurons. Our results show that the baseline expectation for competition in cortex can be estimated by the direct coupling between points in cortex. Searching for competition in cortex must be a search for deviations from nonspecific, homogeneous, and symmetric connectivity.

Accordingly, we considered whether effective lateral inhibitory profiles (and thereby lateral competition) might be obtained in a network with lateral excitatory projections and only local inhibitory projections, through specificity of where on a respective axonal and dendritic tree synaptic connections were formed. For example, synapses on an inhibitory axonal arbor that are distal to the soma of the source neuron might be biased to contact the distal segments of its targets (see Figures 7A–7C). This effectively widens the spatial range of inhibition without requiring long-range inhibitory projections, and under the direct coupling effect therefore permits lateral competition to occur (see Figure 7C). The opposite mode of synapse location specificity would also support lateral competition (see Figures 7D–7F). This hypothesis is consistent with the assumptions made for our analytical models, consistent with the known spatial ranges of excitatory and inhibitory axonal projections and consistent with the absence of neuron class projection bias described in the literature (Kisvárdy et al., 1986).

The question of dendritic location specificity is difficult to tackle experimentally, and so has been only sparsely examined. In the mammalian hippocampus, both long-range and local projections are laminar specific, which due to the highly ordered radial arrangement of Purkinje and granule cell dendrites implies that individual pathways are highly selective for particular dendritic (and somatic) domains (Blackstad, 1956, 1958; Ribak & Seress, 1983; Soriano & Frotscher, 1989; Han, Buhl, Lörinczi, & Somogyi, 1993; Deller, Martinez, Nitsch, & Frotscher, 1996). Lamination is also a striking feature of the neocortex, and there is some evidence that afferent projections to cortex are also laminar-specific. Petreanu and colleagues

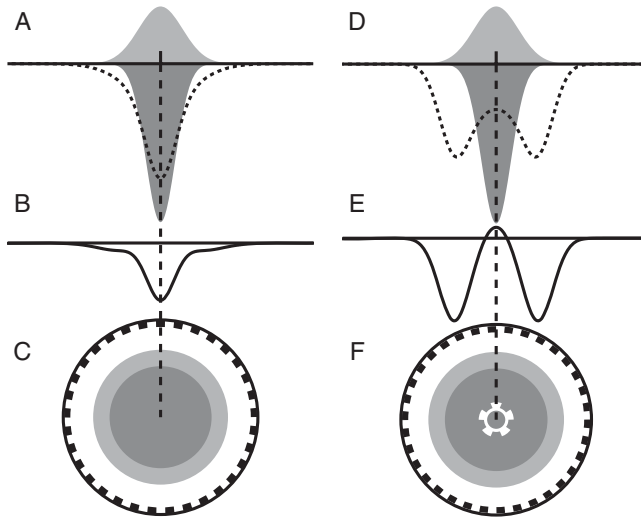


Figure 7: Dendrite location specificity modifies the effective range of inhibition and competition. (A–C) Estimated profiles of effective inhibition and competition, under the assumption that inhibitory synapses are made onto excitatory neuron dendrites in a biased manner, such that synapses on a distal inhibitory axonal segment contact only distal dendritic segments (and proximal contacting proximal). The effective range of inhibition (dashed gray curve in A) can be wider than the spatial range of individual inhibitory axons (dark shaded curve in A; difference between excitation and inhibition shown in B). Under the direct coupling effect, this would permit disynaptic competition to occur over larger spatial distances (C; see Figure 6F). (D–F) If distal inhibitory synapses are made onto proximal excitatory dendritic segments and vice versa (D), then side lobes of effective inhibition arise (dashed gray curve in D; difference between excitation and inhibition shown in E) and the range of competition is again increased with respect to nonspecific connectivity (F; see Figure 6F). Conventions as in Figure 6. Scale is in proportion with synaptic strength estimates for cat visual cortex (see appendix B).

investigated whether individual pathways targetting rodent barrel cortex, arising from other cortical areas and subcortical structures, formed synapses on specific dendritic segments of excitatory neurons (Petreanu, Mao, Sternson, & Svoboda, 2009). They found that long-range projections to neurons in layers 2, 3, and 5 targeted specific dendritic domains ranging in depth from basal to apical dendrites. In contrast, local excitatory projections from layers 2 and 3 to neurons in layer 5 did not show a preference for a particular dendritic location. However, the results in hippocampus and barrel cortex do not measure preference for lateral dendritic location of the form

we discussed in Figure 7, but only for vertical dendritic location within a cortical column. In principle, the experimental technique of Petreanu and colleagues could be applied to explore lateral dendritic specificity, but this remains an unexplored hypothesis.

**4.2 Intracolumnar Competition.** Our results indicate that while lateral competition is difficult to justify in columnar cortical architectures, competition could nevertheless occur between neighboring cortical columns over short distances (see Figures 4A, 4B, and 6D, 6F). Within a single column, the machinery required for competition—recurrent excitatory and inhibitory connections—is readily available without making unreasonable assumptions about the cortical architecture. Indeed, responses of neighboring neurons in cat visual cortex are highly decorrelated, over and above what is expected from differences in their respective receptive fields (Yen, Baker, & Gray, 2007; Tolhurst, Smyth, & Thompson, 2009; Ecker et al., 2010; Martin & Schröder, 2013). This surprising lack of correlation between neurons with similar orientation preference and similar retinotopic location could occur through local competition between neurons within a cortical column. Decorrelation of neurons that receive similar inputs would increase the information coding capacity of single neurons and populations in cortex (Shamir & Sompolinsky, 2004; Averbeck, Latham, & Pouget, 2006).

Some existing models for learning receptive field properties in cortex are defined without an explicit mapping to cortical space, but are nevertheless compatible with the concept of strong competition within a column of visual cortex (Olshausen & Field, 1996; Bell & Sejnowski, 1997; Perrinet, 2004; Rehn & Sommer, 2007). These models seek to learn maximally sparse cortical representations by providing negative feedback between neurons with similar receptive fields. Neurons in strongest competition would therefore represent similar locations and preferred orientations in visual space, and consequently map to similar locations in cortical space.

Recent work exploring competition and information processing in non-columnar (mouse visual) cortex (Muir, Molina-Luna, Helmchen, & Kampa, 2014), competition and learning within local populations (Jug, Cook, & Steger, 2012) and dynamics of cortical columns with local inhibition (Landsman et al., 2012) show that local excitatory connectivity can provide a rich repertoire of complex dynamics and competitive behaviour for information processing in cortex.

## Appendix A: Detailed Analysis

---

**A.1 Analytical System Definition.** The differential equation for a single unit in the system is given by

$$\tau_n \cdot \dot{x}_n + x_n = \mathbf{W} (\mathbf{a} \odot [\mathbf{x} - \mathbf{v}]^+) + \iota_n, \quad (\text{A.1})$$

Table 2: Variables in the System Differential Equations.

$\tau_n$	Time constant for unit $n$
$\mathbf{T}$	Matrix composed of all network time constants $\tau_n$
$x_n$	Activation value of unit $n$
$\mathbf{x}$	Vector composed of unit activations $x_n$
$\alpha_n$	Activation gain of unit $n$ (slope of the linear-threshold transfer function)
$\mathbf{a}$	Vector composed of unit gain factors $\alpha_n$
$\vartheta_n$	Activation threshold of unit $n$
$\mathbf{v}$	Vector composed of unit activation thresholds $\vartheta_n$
$w_{ij}$	Synaptic weight from unit $j$ to unit $i$
$\mathbf{W}$	Matrix composed of all network weights $w_{ij}$
$\mathbf{W}^+$	Matrix composed of network weights, with rows and columns corresponding to inactive units set to zero, that is, the weight matrix corresponding to the active network partition
$\iota_n$	External instantaneous input current injected into unit $n$
$\mathbf{J}^+$	Jacobian of the system for the active network partition
Part [ $p$ ]	Nomenclature for referring to a particular partition $p$ of the network, where $p$ is a Boolean vector indicating which columns of the network are active in the partition
$\lambda^p$	Set of eigenvalues of the system Jacobian $\mathbf{J}$ , in partition $p$
$N$	Number of units in the network

where  $x_n$  is the activation of unit  $n$ ;  $\tau_n$  is the time constant of unit  $n$ ;  $\mathbf{W}$  is the matrix composed of the individual weights  $w_{ij}$  of the network;  $\mathbf{a}$  is the vector of activation gains  $\alpha_n$  of the network;  $\mathbf{v}$  is the vector of activation thresholds  $\vartheta_n$ ;  $\iota_n$  is the current injected into unit  $n$ ; and with

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_N \end{bmatrix}; \mathbf{a} = \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_N \end{bmatrix}; \mathbf{v} = \begin{bmatrix} \vartheta_1 \\ \vartheta_2 \\ \vdots \\ \vartheta_N \end{bmatrix}; \mathbf{W} = \begin{bmatrix} w_{11} & w_{12} & \dots & w_{1N} \\ w_{21} & w_{22} & & w_{2N} \\ \vdots & & \ddots & \vdots \\ w_{N1} & w_{N2} & \dots & w_{NN} \end{bmatrix}. \tag{A.2}$$

In this notation,  $[x]^+$  is the linear-threshold transfer function given by  $[x]^+ = \max(x, 0)$ , and  $\mathbf{a} \odot \mathbf{b}$  denotes the element-wise product of the vectors  $\mathbf{a}$  and  $\mathbf{b}$ . The activation gains  $\alpha_n$  can be absorbed into the weights arising from each unit without loss of generality; for further analysis, we take  $\forall n : \alpha_n = 1$  and omit the vector  $\mathbf{a}$  from equation A.1. All parameters except the weights  $\mathbf{W}$  are constrained to be non-negative. The definition of all parameters is given in Table 2.

The local stability and behavior of a linear-threshold network can be determined by examining the eigenvalues and the trace of the system Jacobian, under the assumption of a specified active network partition (Hahnloser,

1998a). This Jacobian is given by  $\mathbf{J}^+ = (\mathbf{W}^+ - \mathbf{I}) \oslash \mathbf{T}$ , where  $\mathbf{I}$  is the identity matrix;  $\mathbf{A} \oslash \mathbf{B}$  denotes element-wise division of the matrices  $\mathbf{A}$  and  $\mathbf{B}$ ;  $\mathbf{W}^+$  is the network weight matrix, with rows and columns corresponding to inactive units set to zero; and  $\mathbf{T}$  is the  $N \times N$  square matrix composed of all unit time constants  $\tau_N$ :

$$\mathbf{T} = \begin{bmatrix} \tau_1 & \cdots & \tau_1 \\ \vdots & \ddots & \vdots \\ \tau_N & \cdots & \tau_N \end{bmatrix}. \quad (\text{A.3})$$

A partition is stable in the bounded-input bounded-output (BIBO) sense when the eigenvalues of  $\mathbf{J}^+$  have no positive real components, and  $\text{Trace}(\mathbf{J}^+) \leq 0$ . Note that the full system can have a mixture of stable and unstable partitions and that the system can be globally stable if all unstable partitions result in a transition to stable partitions (Hahnloser, 1998a). Partitions that contain large eigenvalues with complex components have oscillatory dynamics, which lead to either stable or unstable spirals depending on the magnitude of the real component of the eigenvalues.

**A.2 Stability and Behavior of Two Columns.** Here we describe in detail the analysis of the two-column network presented in the body of the letter (section 2.5; see Figure 2B). This network consists of two excitatory and two inhibitory units  $x_{En}$  and  $x_{In}$ , respectively, where  $n$  is the column number. The time constants for the system are defined by the class of each unit, with a class time constant  $\tau_E$  for the excitatory units and another  $\tau_I$  for the inhibitory units. Activation thresholds are similarly defined by class, giving  $\vartheta_E$  and  $\vartheta_I$ . The pair of units in a column receive a common input  $\iota_n$ . The system weights are as shown in Figure 2B; the weight matrix is therefore given by

$$\mathbf{W} = \begin{bmatrix} w_{ER} & w_{EC} & -w_{IR} & -w_{IC} \\ w_{EC} & w_{ER} & -w_{IC} & -w_{IR} \\ w_{ER} & w_{EC} & -w_{IR} & -w_{IC} \\ w_{EC} & w_{ER} & -w_{IC} & -w_{IR} \end{bmatrix}. \quad (\text{A.4})$$

In this work, partitions are denoted by superscripts indicating which columns of the network are active. For example,  $\lambda^{11}$  denotes the set of eigenvalues in the network partition when columns 1 and 2 have nonzero activity. The network partition itself is denoted  $\text{Part}[11]$ . Sets of equations



that apply to a given partition are grouped with a vertical line as shown here:

$$\left. \begin{matrix} A \\ B \end{matrix} \right|^{11}$$

*A.2.1 Simplifying Substitutions.* We define two values,  $\Lambda_R$  and  $\Lambda_C$ , which correspond to the excess of inhibition over excitation within a column (i.e. recurrent connections) and between columns, given by

$$\begin{aligned} \Lambda_R &= 1 + w_{IR} - w_{ER} \text{ and} \\ \Lambda_C &= w_{IC} - w_{EC}. \end{aligned} \tag{A.5}$$

*A.2.2 Zero Thresholds; Equal Time Constants.* We begin by examining the system where  $\vartheta_E = \vartheta_I = 0$  and  $\tau_E = \tau_I = \tau$ . Under these simplifying assumptions, the system eigenvalues are given by

$$\begin{aligned} \lambda^{11} &= \left\{ \frac{-1}{\tau}, \frac{-1}{\tau}, \frac{\Lambda_C - \Lambda_R}{\tau}, \frac{-\Lambda_C - \Lambda_R}{\tau} \right\}, \\ \lambda^{10} &= \left\{ \frac{-1}{\tau}, \frac{-1}{\tau}, \frac{-1}{\tau}, \frac{-\Lambda_R}{\tau} \right\}. \end{aligned} \tag{A.6}$$

Note that the system can never have complex eigenvalues, and so single partitions can never have oscillatory dynamics. The system is globally BIBO stable under the condition

$$\Lambda_R > \text{abs}(\Lambda_C), \tag{A.7}$$

where  $\text{abs}(\Lambda_C)$  is the absolute value of  $\Lambda_C$ .

In many of the conditions for stability that will follow, the constraint  $\Lambda_R > 0$  (or similar), implying that  $w_{ER} < 1 + w_{IR}$ , appears often. With all other weights set to zero and the excitatory gain  $\alpha_E = 1$ , a value of  $w_{ER} = 1$  implies that if the excitatory unit has a net activity of  $r$ , then the recurrent excitatory input current supplied back to the same excitatory unit will also be  $r$ . In other words, the open-loop gain of the recurrent excitatory connection is unitary. If  $w_{ER} > 1$ , implying that the open-loop gain of the recurrent excitatory connection is greater than unitary, it is easy to see that the activity of the excitatory unit will grow without bound (in the absence of any network stability mechanism such as recurrent inhibition, or single-unit stability mechanism such as a saturating transfer function). If  $w_{ER} < 1$ , the open-loop gain of the recurrent excitatory connection is less than 1, implying that for an activity of  $r$ , the recurrent excitatory input will be less than  $r$ .

Note that we generally ignore the partition where all columns are switched off (Part[00] or Part[000] for a three-column network). This partition is never unstable (under the reasonable assumption of bounded weights), never oscillatory, cannot exhibit competitive behavior, and is guaranteed to transition to another partition for inputs greater than the excitatory threshold  $\vartheta_E$ .

To determine when and whether this two-column network can display competitive behavior, we examine the fixed points  $\bar{x}_n$  of the system (Rutishauser & Douglas, 2009). These are found by solving the system of differential equations (see equation A.1) by setting  $\dot{x}_n \cdot \tau_n = 0$ . The two columns are in competition if the net action of the network is such that an increase in the input to column 1 results in a decrease in the input to column 2. The fixed points of the two-column network are given by

$$\begin{aligned} \bar{x}_{E1} = \bar{x}_{I1} &= \left. \frac{\iota_2 \cdot \Lambda_C - \iota_1 \cdot \Lambda_R}{\Lambda_C^2 - \Lambda_R^2} \right|^{11} & \text{and} & \quad (A.8) \\ \bar{x}_{E2} = \bar{x}_{I2} &= \left. \frac{\iota_2 \cdot \Lambda_C - \iota_1 \cdot \Lambda_R}{\Lambda_C^2 - \Lambda_R^2} \right|^{11} \\ \bar{x}_{E1} = \bar{x}_{I1} &= \left. \frac{\iota_1}{\Lambda_R} \right|^{10} \\ \bar{x}_{E2} = \bar{x}_{I2} &= \left. \iota_2 - \frac{\iota_1 \cdot \Lambda_C}{\Lambda_R} \right|^{10} . \end{aligned}$$

We now calculate the partial derivatives  $\partial \bar{x}_{E2} / \partial \iota_1$  for Part[11] and Part[10], which are given by

$$\begin{aligned} \frac{\partial \bar{x}_{E2}}{\partial \iota_1} &= \left. \frac{\Lambda_C}{\Lambda_C^2 - \Lambda_R^2} \right|^{11} \text{ and} & \quad (A.9) \\ \frac{\partial \bar{x}_{E2}}{\partial \iota_1} &= \left. -\frac{\Lambda_C}{\Lambda_R} \right|^{10} . \end{aligned}$$

For competitive behavior to exist between the two columns, we require that the partial derivative  $\partial \bar{x}_{E2} / \partial \iota_1 < 0$  for both partitions, so that regardless of the initial conditions of the network, column 2 is inactivated by an input only to column 1. This criterion implicitly assumes that the excitatory unit in column 1 is above threshold, since a subthreshold input to column 1 could not have any effect on the activity of column 2. The condition above is satisfied when

$$\wedge \left[ \Lambda_C < \Lambda_R \right. \\ \left. [(\Lambda_R > 0) \wedge (\Lambda_C > 0)] \vee [\Lambda_R < 0] \right] . \quad (A.10)$$

Combining these conditions with the global conditions for stability (see equation A.7) gives the global conditions for stable WTA system behavior, namely,

$$0 < \Lambda_C < \Lambda_R. \tag{A.11}$$

The factor  $\Lambda_R > 0$  in equation A.11 is required for the network to be globally stable. The factor  $\Lambda_C > 0$ , however, speaks directly to the possibility of lateral competition in this simple network, since it reduces to  $w_{IC} > w_{EC}$ . The direct inhibitory connection between the two columns,  $w_{IC}$ , must be stronger than the corresponding excitatory connection,  $w_{EC}$ . This is a strong result, since the derivatives of the fixed points (see equation A.9) do not depend on the respective time constants of inhibition and excitation. For any combination of  $\tau_E > 0$  and  $\tau_I > 0$ , differential amplification can occur only if the network is dominated by lateral inhibition.

*A.2.3 Hard Winner-Take-All Behavior.* A network is said to display hard WTA behavior if it permits only the strongest-driven unit to remain active, while suppressing the activity of all more weakly driven units to zero. Interpreted for our simple two-column network, if column 1 and 2 are driven with a related input  $\iota_2 = \delta \cdot \iota_1$  where  $0 < \delta < 1$ , then for hard WTA behavior, the fixed-point activity of column one should be positive ( $\bar{x}_{E1} > 0$ ), while column 2 should be silenced ( $\bar{x}_{E2} \leq 0$ ). Solving the fixed points in equation A.8 for these conditions, in addition to the requirements for BIBO stability, gives the following constraints on the system parameters:

$$[\Lambda_R > 0] \wedge [\Lambda_C < \Lambda_R] \wedge \left[ \delta \leq \frac{\Lambda_C}{\Lambda_R} \right]. \tag{A.12}$$

The first two terms come directly from the conditions for stable WTA behavior, equation A.11. The third term dictates whether the network can display hard WTA behavior. Unfortunately, the factor  $\delta$ , which governs the difference between the inputs to columns 1 and 2 is constrained by the weights in the network. For a given set of weights, the network will produce hard winner-take-all behavior only for sufficiently different values of the input. For columnar systems of this form, a network that displays hard WTA behavior for any differential input cannot exist. There is no regime that operates exclusively in a hard WTA mode.

*A.2.4 Gain of the Winning Column.* For completeness, we derive the gain of the two-column system. When operated in a competitive regime, the gain of the winning column is obtained by examining the steady-state equations in equation A.8. When the network is operating in a soft WTA regime, competition can occur without inactivating column 2. We provide a

differential input to columns 1 and 2 such that  $\iota_1 = \iota_C + \iota_\delta$  and  $\iota_2 = \iota_C$ , where  $\iota_C$  is a common-mode component and  $\iota_\delta$  is the differential component, with  $\iota_\delta > 0$ . Since the network is in Part[11], the gain of the winning column (assumed to be column 1) is given by

$$G = \frac{\bar{x}_{E1}}{\iota_1} = \frac{\iota_C \cdot \Lambda_C - (\iota_C + \iota_\delta) \Lambda_R}{(\iota_C + \iota_\delta) (\Lambda_C^2 - \Lambda_R^2)}. \quad (\text{A.13})$$

The gain of the winning column has both a common-mode and differential component.

If the network is operating in a hard WTA regime the system must be in Part[10], therefore, the gain of the winning column is given by (see Rutishauser & Douglas, 2009)

$$\frac{\bar{x}_{E1}}{\iota_1} = \frac{1}{1 + w_{IR} - w_{ER}} = \frac{1}{\Lambda_R}. \quad (\text{A.14})$$

*A.2.5 Unequal Time Constants.* The time constants of excitation and inhibition have no effect on the fixed points of the two-column network. However, the fixed points are a useful description of the network activity only to the extent that they help to predict the response of the system to a given input. Depending on the network parameters, the fixed points can be exponentially unstable (if the system has unbounded behavior) or can provide a focus around which the network activity oscillates.

For a single column, the parameter constraints that ensure oscillatory behavior implicate slow inhibition as the mechanism for generating oscillations (Wilson & Cowan, 1973; Hahnloser, 1998a; Ermentrout, 1998b; Landsman et al., 2012). In all cases, oscillations are not possible unless the time constant for inhibition  $\tau_I$  is longer than the time constant for excitation  $\tau_E$ . The general constraints relating  $\tau_I$  and  $\tau_E$  are given by

$$\frac{(a - b)\tau_E}{(w_{ER} - 1)^2} < \tau_I < \frac{(a + b)\tau_E}{(w_{ER} - 1)^2}, \quad \text{where} \quad (\text{A.15})$$

$a = 1 + w_{ER}(w_{IR} - 1) + w_{IR}$  and  $b = 2\sqrt{w_{ER} \cdot w_{IR}(w_{IR} - w_{ER} + 1)}$ . The constraint in equation A.15 requires that  $\tau_I$  is longer than  $\tau_E$ , since the factor of  $\tau_E$  in the first term is always larger than 1 as long as  $w_{ER} \leq 1 + w_{IR}$  (a general constraint for stability similar to that given in equation A.7). If  $w_{ER} = 1$ , then the relationship constraining  $\tau_I$  and  $\tau_E$  is given by the simpler form,

$$\tau_I > \frac{(1 + w_{IR})^2 \tau_E}{4 \cdot w_{IR}}, \quad (\text{A.16})$$

which likewise constrains  $\tau_I$  to be longer than  $\tau_E$  for oscillatory dynamics.

For the two-column model, the constraints are obtained by examining the eigenvalues for the system where  $\tau_I \neq \tau_E$ , which have the form

$$\lambda = \frac{1}{2\tau_I \cdot \tau_E} \begin{cases} -p + q \pm \sqrt{r} \\ -p - q \pm \sqrt{s} \end{cases}, \text{ where} \tag{A.17}$$

$$p = \tau_I - w_{ER} \cdot \tau_I + \tau_E + w_{IR} \cdot \tau_E; \quad q = w_{IC} \cdot \tau_E - w_{EC} \cdot \tau_I; \quad r = -4(1 + w_{EC} - w_{ER} - w_{IC} + w_{IR})\tau_I \cdot \tau_E + ((1 + w_{EC} - w_{ER})\tau_I + (1 - w_{IC} + w_{IR})\tau_E)^2; \quad \text{and} \\ s = 4(-1 + w_{EC} + w_{ER} - w_{IC} - w_{IR})\tau_I \cdot \tau_E + ((-1 + w_{EC} + w_{ER})\tau_I - (1 + w_{IC} + w_{IR})\tau_E)^2.$$

Unfortunately, the simplifying substitutions of  $\Lambda_R$  and  $\Lambda_C$  (see equation A.5) do not help here. The system is oscillatory when either of  $r$  or  $s$  becomes negative. The parameter constraints obtained by expanding these inequalities have a similar form to equation A.15, but are long, complicated, and not included here. Nevertheless, for the two-column network as for the single column,  $\tau_I$  must be longer than  $\tau_E$  for oscillatory dynamics to be present.

*A.2.6 Nonzero Thresholds.* The thresholds for excitation and inhibition  $\vartheta_E$  and  $\vartheta_I$  do not enter the expressions for the system eigenvalues and therefore cannot have an effect on the stability or oscillatory dynamics of the network. However, the thresholds do play a role in determining the fixed points of the network, and so might determine whether competitive interactions can occur. If nonzero thresholds are allowed separately for the excitatory and inhibitory units, then the fixed points for the two-column network (for  $\iota_2 = 0$ ) are given by

$$\begin{aligned} \bar{x}_{E1} = \bar{x}_{I1} &= \frac{\iota_1 \cdot \Lambda_R}{-\Lambda_C^2 + \Lambda_R^2} + f \Bigg|^{11} \\ \bar{x}_{E2} = \bar{x}_{I2} &= \frac{\iota_1 \cdot \Lambda_C}{\Lambda_C^2 - \Lambda_R^2} + f \Bigg| \text{ and (A.18)} \\ \bar{x}_{E1} = \bar{x}_{I1} &= \frac{\iota_1}{\Lambda_R} + \frac{w_{IR} \cdot \vartheta_I - w_{ER} \cdot \vartheta_E}{\Lambda_R} \Bigg|^{10} \\ \bar{x}_{E2} = \bar{x}_{I2} &= -\frac{\iota_1 \cdot \Lambda_C}{\Lambda_R} + \frac{w_{IC} \cdot \vartheta_I - w_{EC} \cdot \vartheta_E + (w_{ER} \cdot w_{IC} - w_{EC} \cdot w_{IR})(\vartheta_E - \vartheta_I)}{\Lambda_R} \Bigg| \end{aligned}$$

where  $f = \frac{(w_{IC} + w_{IR})\vartheta_I - (w_{EC} + w_{ER})\vartheta_E}{\Lambda_C + \Lambda_R}$ .

The fixed points in equation A.18 show that the response of the network contains a component that depends on the thresholds of excitation and inhibition  $\vartheta_E$  and  $\vartheta_I$  but not on the input  $\iota_1$  and a separate component that depends on the input but not on the thresholds. Therefore, thresholds for excitation and inhibition that are identical between the two columns can modify the fixed points only in a manner independent of the input to the

network. This implies that the partial derivatives  $\partial \bar{x}_{E2}/\partial \iota_1$  in both partitions are independent of the activation thresholds, and in fact they have the same form as in equation A.9. Setting a nonzero threshold for either excitation or inhibition therefore has no effect on the existence of competition between columns.

*A.2.7 Memory State with Nonzero Thresholds.* Winner-take-all networks can support the existence of a memory state, where activity persists in the absence of external input (Rutishauser & Douglas, 2009). The stability and configuration of this memory state can be explored by examining the steady-state network activity equations, with input terms  $\iota_1$  and  $\iota_2$  set to zero. For the two-column network presented here, equation A.18 reveals that for Part[11], the common-mode term  $f$  in equation A.18 will completely determine the network response. If  $f$  is positive, a stable memory state will exist; however, this memory state is identical for both columns, and so the activity of both columns will become equal. If  $f$  is negative, the memory state in Part[11] is unstable, and one or both columns will become inactive.

For the memory state to operate in a competitive switchable mode, where the activity in the network can be nudged from one column to the other, the two-column model must be able to operate in a hard-WTA regime in the absence of input. This is unrelated to the condition in section A.2.3, which applies for nonzero input. For the memory state to be stable, the steady-state solutions given in equation A.18 for Part[10] must be positive for the winning column (assumed to be column 1) and negative for the losing column (column 2).

**A.3 Stability and Behaviour of a Three-Column Ring.** Here we describe the analysis of a three-column ring network, with 50% more brevity than for the two-column network. This network consists of three identical columns, each with the same parameters as for the two-column network described above (see Figure 2C). The columns are arranged in a linear ring, with each column connecting symmetrically to its nearest neighbors. The dynamic equations for the system are as in equation A.1, with the system weight matrix given by

$$\mathbf{W} = \begin{bmatrix} w_{ER} & w_{EC} & w_{EC} & -w_{IR} & -w_{IC} & -w_{IC} \\ w_{EC} & w_{ER} & w_{EC} & -w_{IC} & -w_{IR} & -w_{IC} \\ w_{EC} & w_{EC} & w_{ER} & -w_{IC} & -w_{IC} & -w_{IR} \\ w_{ER} & w_{EC} & w_{EC} & -w_{IR} & -w_{IC} & -w_{IC} \\ w_{EC} & w_{ER} & w_{EC} & -w_{IC} & -w_{IR} & -w_{IC} \\ w_{EC} & w_{EC} & w_{ER} & -w_{IC} & -w_{IC} & -w_{IR} \end{bmatrix}. \quad (\text{A.19})$$

The eigenvalues for this network are given by

$$\begin{aligned} \lambda^{111} &= \frac{1}{2 \cdot \tau_I \cdot \tau_E} \left\{ p_1 \pm \sqrt{r_1} \right. \\ &\quad \left. (2 \cdot w_{EC} + w_{ER} - 1) \tau_I + (1 + 2 \cdot w_{IC} + w_{IR}) \tau_E - \sqrt{r_2} \right. \\ \lambda^{110} &= -\frac{1}{\tau_E}, -\frac{1}{\tau_I}, \frac{1}{2 \cdot \tau_I \cdot \tau_E} \left\{ p_1 \pm \sqrt{r_1} \right. \\ &\quad \left. p_2 \pm \sqrt{r_1} \right. \\ \lambda^{100} &= -\frac{1}{\tau_E}, -\frac{1}{\tau_I}, \frac{1}{2 \cdot \tau_I \cdot \tau_E} (-1 + w_{ER}) \tau_I - (1 + w_{IR}) \tau_E \pm \sqrt{r_3}, \text{ where} \end{aligned} \tag{A.20}$$

$$\begin{aligned} r_1 &= -4(1 + w_{EC} - w_{ER} - w_{IC} + w_{IR})\tau_I \cdot \tau_E + ((1 + w_{EC} - w_{ER})\tau_I + (1 - w_{IC} \\ &+ w_{IR})\tau_E)^2; r_2 = 4(2 \cdot w_{EC} + w_{ER} - 2 \cdot w_{IC} - w_{IR} - 1)\tau_I \cdot \tau_E + ((2 \cdot w_{EC} + w_{ER} \\ &- 1)\tau_I - (2 \cdot w_{IC} + w_{IR} - 1)\tau_E)^2; r_3 = 4(-1 + w_{ER} - w_{IR})\tau_I \cdot \tau_E + ((w_{ER} - 1) \\ &\tau_I - (1 + w_{IR})\tau_E)^2; \quad p_1 = (w_{ER} - w_{EC} - 1)\tau_I + (w_{IC} - w_{IR} - 1)\tau_E; \quad \text{and} \\ p_2 &= (w_{EC} + w_{ER} - 1)\tau_I - (1 + w_{IC} + w_{IR})\tau_E. \end{aligned}$$

Again, the thresholds for excitation and inhibition do not enter the eigenvalues and so cannot have an effect on the stability of the system. Examining the case frequently used in the literature where  $\tau_E = \tau_I = \tau$ , the eigenvalues of the system have the much simpler form

$$\begin{aligned} \lambda_{\tau_E=\tau_I}^{111} &= \left\{ -\frac{1}{\tau}, -\frac{1}{\tau}, -\frac{1}{\tau}, -\frac{2\Lambda_C + \Lambda_R}{\tau}, \frac{\Lambda_C - \Lambda_R}{\tau}, \frac{\Lambda_C - \Lambda_R}{\tau} \right. \tag{A.21} \\ \lambda_{\tau_E=\tau_I}^{110} &= \left\{ -\frac{1}{\tau}, -\frac{1}{\tau}, -\frac{1}{\tau}, -\frac{1}{\tau}, -\frac{\Lambda_C + \Lambda_R}{\tau}, \frac{\Lambda_C - \Lambda_R}{\tau} \right. \\ \lambda_{\tau_E=\tau_I}^{100} &= \left\{ -\frac{1}{\tau}, -\frac{1}{\tau}, -\frac{1}{\tau}, -\frac{1}{\tau}, -\frac{1}{\tau}, -\frac{\Lambda_R}{\tau} \right. \end{aligned}$$

Under the assumption of equal time constants, the parameter bounds for global BIBO stability are given by

$$[2\Lambda_C + \Lambda_R > 0] \wedge [\Lambda_C < \Lambda_R]. \tag{A.22}$$

To determine the conditions for competitive interaction between columns 1 and 2, we examine the steady-state solutions for the system as for the two-column model. Here the steady-state solutions for driven columns 1

and 2 are given by the exhaustive and exhausting set of equations

$$\left. \begin{aligned} \bar{x}_{E1} = \bar{x}_{I1} = \vartheta_I - \frac{-2\iota_1 + \iota_2}{3(\Lambda_R - \Lambda_C)} - h_1 \\ \bar{x}_{E2} = \bar{x}_{I2} = \vartheta_I - \frac{\iota_1 - 2\iota_2}{3(\Lambda_R - \Lambda_C)} - h_1 \\ \bar{x}_{E3} = \bar{x}_{I3} = \vartheta_I - \frac{\iota_1 + \iota_2}{3(\Lambda_R - \Lambda_C)} - h_1 \end{aligned} \right|^{111} \tag{A.23}$$

$$h_1 = \frac{3(2 \cdot w_{EC} + w_{ER} - 1) \vartheta_I - 3(2 \cdot w_{EC} + w_{ER})}{3(2 \cdot \Lambda_C + \Lambda_R)}$$

$$\left. \begin{aligned} \bar{x}_{E1} = \bar{x}_{I1} = \frac{\iota_1 - \iota_2}{2(\Lambda_R - \Lambda_C)} + \frac{\iota_1 + \iota_2}{2(\Lambda_R + \Lambda_C)} + h_2 \\ \bar{x}_{E2} = \bar{x}_{I2} = \frac{-\iota_1 + \iota_2}{2(\Lambda_R - \Lambda_C)} + \frac{\iota_1 + \iota_2}{2(\Lambda_R + \Lambda_C)} + h_2 \\ \bar{x}_{E3} = \bar{x}_{I3} = \frac{(\iota_1 + \iota_2)\Lambda_C}{-\Lambda_C - \Lambda_R} + h_3 \end{aligned} \right|^{110}$$

$$h_2 = \vartheta_I + \frac{(w_{EC} + w_{ER} - 1) \vartheta_i - (w_{EC} + w_{ER}) \vartheta_E}{(\Lambda_R + \Lambda_C)}$$

$$h_3 =$$

$$\frac{2((-1 + w_{ER}) w_{IC} - w_{EC} \cdot w_{IR}) \vartheta_I + 2(w_{EC} - w_{ER} \cdot w_{IC} + w_{EC} \cdot w_{IR}) \vartheta_E}{-\Lambda_C - \Lambda_R}$$

$$\left. \begin{aligned} \bar{x}_{E1} = \bar{x}_{I1} = \frac{\iota_1 \cdot \Lambda_R}{\Lambda_R^2 - \Lambda_C^2} + h_4 \\ \bar{x}_{E2} = \bar{x}_{I2} = \iota_2 + \frac{\iota_1 \cdot \Lambda_C}{-\Lambda_C - \Lambda_R} + h_3 \\ \bar{x}_{E3} = \bar{x}_{I3} = \frac{\iota_1 \cdot \Lambda_C}{\Lambda_C^2 - \Lambda_R^2} + h_4 \end{aligned} \right|^{101}$$

$$h_4 = \frac{(w_{EC} + w_{ER}) \vartheta_E - (w_{IC} + w_{IR}) \vartheta_I}{-\Lambda_C - \Lambda_R}$$

$$\left. \begin{aligned} \bar{x}_{E1} = \bar{x}_{I1} = \frac{\iota_1}{\Lambda_R} + \frac{w_{IR} \cdot \vartheta_I - w_{ER} \cdot \vartheta_E}{\Lambda_R} \\ \bar{x}_{E2} = \bar{x}_{I2} = \iota_2 + \frac{(w_{EC} - w_{IC}) \iota_1}{\Lambda_R} + h_5 \\ \bar{x}_{E3} = \bar{x}_{I3} = \frac{(w_{IC} - w_{EC}) \iota_1}{\Lambda_R} + h_5 \end{aligned} \right|^{100}$$

$$h_5 = \frac{(w_{IC} - w_{ER} w_{IC} + w_{EC} \cdot w_{IR}) \vartheta_I + (w_{ER} \cdot w_{IC} - w_{EC} \cdot w_{IR} - w_{EC}) \vartheta_E}{\Lambda_R}$$

As in the two-column analysis, we examine the partial derivatives  $\partial \bar{x}_{E2} / \partial \iota_1$  for each partition to find the parameter bounds that guarantee competitive interactions between columns 1 and 2. As we saw previously,



the activation thresholds  $\vartheta_E$  and  $\vartheta_I$  drop out of the derivatives, leaving the simple form

$$\begin{aligned} \frac{\partial \bar{x}_{E2}}{\partial \iota_1} &= \frac{\Lambda_C}{(\Lambda_C - \Lambda_R)(2 \cdot \Lambda_C + \Lambda_R)} \Big|^{111} & (A.24) \\ \frac{\partial \bar{x}_{E2}}{\partial \iota_1} &= \frac{\Lambda_C}{\Lambda_C^2 - \Lambda_R^2} \Big|^{110} \\ \frac{\partial \bar{x}_{E2}}{\partial \iota_1} &= \frac{-\Lambda_C}{\Lambda_C + \Lambda_R} \Big|^{101} \text{ and} \\ \frac{\partial \bar{x}_{E2}}{\partial \iota_1} &= \frac{-\Lambda_C}{\Lambda_R} \Big|^{100} . \end{aligned}$$

For the three-column ring network, the parameter bounds that ensure competition between columns 1 and 2 are given by

$$[\Lambda_C < \Lambda_R] \wedge [((\Lambda_R > 0) \wedge (\Lambda_C > 0)) \vee (\Lambda_R < 0)] . \quad (A.25)$$

When combined with the conditions for BIBO stability (see equation A.22), these bounds reduce to the wonderfully simple  $[\Lambda_C > 0] \wedge [\Lambda_C < \Lambda_R]$ , implying once again that  $w_{IC} > w_{EC}$ .

Perhaps surprisingly, the third column cannot mediate competition between columns 1 and 2 by providing disynaptic inhibition. For competition to occur, the direct inhibitory coupling between columns 1 and 2 must be stronger than the direct excitatory coupling.

*A.3.1 Oscillatory Behavior.* The three-column ring has oscillatory dynamics if any of the roots  $r_1$  through  $r_3$  from the system eigenvalues (see equation A.20) are negative. The full parameter bounds are not included here, but oscillatory dynamics are again possible only if the inhibitory time constant  $\tau_I$  is longer than the excitatory time constant  $\tau_E$ .

**A.4 Stability and Behavior of a Three-Column Chain.** Here we describe the analysis of a three-column chain network. This network is similar to the three-column ring, but with the columns conceptually placed on a line rather than a circle (see Figure 2D). The connectivity between the outer two columns is modified such that they are more weakly coupled than in the ring model. This model is designed to explore whether competition between two distant columns (columns 1 and 3) can be driven by a spatially intermediate column (column 2). The dynamic equations for the system are

as in equation A.1, with the system weight matrix given by

$$\mathbf{W} = \begin{bmatrix} w_{ER} & w_{EC1} & w_{EC2} & -w_{IR} & -w_{IC1} & -w_{IC2} \\ w_{EC1} & w_{ER} & w_{EC1} & -w_{IC1} & -w_{IR} & -w_{IC1} \\ w_{EC2} & w_{EC1} & w_{ER} & -w_{IC2} & -w_{IC1} & -w_{IR} \\ w_{ER} & w_{EC1} & w_{EC2} & -w_{IR} & -w_{IC1} & -w_{IC2} \\ w_{EC1} & w_{ER} & w_{EC1} & -w_{IC1} & -w_{IR} & -w_{IC1} \\ w_{EC2} & w_{EC1} & w_{ER} & -w_{IC} & -w_{IC1} & -w_{IR2} \end{bmatrix}. \tag{A.26}$$

For this network, we modify our simplifying assumptions from equation A.5 slightly, to incorporate the difference in nearest and distant column connections, and so use

$$\begin{aligned} \Lambda_R &= 1 + w_{IR} - w_{ER}, \\ \Lambda_{C1} &= w_{IC1} - w_{EC1}, \\ \Lambda_{C2} &= w_{IC2} - w_{EC2}. \end{aligned} \tag{A.27}$$

The eigenvalues for this network are given by

$$\lambda_{\tau_E=\tau_I}^{111} = \frac{-1}{\tau}, \frac{-1}{\tau}, \frac{-1}{\tau}, \frac{\Lambda_{C2} - \Lambda_R}{\tau}, \frac{-1}{2\tau} \left( \Lambda_{C2} + 2 \cdot \Lambda_R \pm \sqrt{8 \cdot \Lambda_{C1}^2 + \Lambda_{C2}^2} \right) \tag{A.28}$$

$$\lambda^{110} = \frac{-1}{\tau_E}, \frac{-1}{\tau_I}, \frac{-1}{2 \cdot \tau_E \cdot \tau_I} \left\{ \begin{aligned} &q_1 + q_2 \pm \sqrt{(q_1 + q_2)^2 - r_4} \\ &q_3 + q_4 \pm \sqrt{(q_3 + q_4)^2 - r_5} \end{aligned} \right.$$

$$\begin{aligned} \lambda^{100} &= \frac{-1}{\tau_E}, \frac{-1}{\tau_E}, \frac{-1}{\tau_I}, \frac{-1}{\tau_I}, \frac{-1}{2 \cdot \tau_E \cdot \tau_I} \\ &\times \left\{ p_4 \pm \sqrt{p_4^2 - 4(1 - w_{ER} + w_{IR}) \tau_E \cdot \tau_I} \right. \end{aligned}$$

$$\lambda_{\tau_E=\tau_I}^{110} = -\frac{1}{\tau}, -\frac{1}{\tau}, -\frac{1}{\tau}, -\frac{1}{\tau}, -\frac{\Lambda_{C1} + \Lambda_R}{\tau}, \frac{\Lambda_{C1} - \Lambda_R}{\tau}$$

$$\lambda_{\tau_E=\tau_I}^{100} = -\frac{1}{\tau}, -\frac{1}{\tau}, -\frac{1}{\tau}, -\frac{1}{\tau}, -\frac{1}{\tau}, -\frac{\Lambda_R}{\tau}$$

where  $q_1 = (1 - w_{IC1} + w_{IR})\tau_E$ ;  $q_2 = (1 + w_{EC1} - w_{ER})\tau_I$ ;  $q_3 = (1 + w_{IC1} + w_{IR})\tau_E$ ;  $q_4 = (1 - w_{EC1} - w_{ER})\tau_I$ ;  $r_4 = 4(1 + w_{EC1} - w_{ER} - w_{IC1} + w_{IR})\tau_E \cdot \tau_I$ ;  $r_5 = 4(1 - w_{EC1} - w_{ER} + w_{IC1} + w_{IR}) \tau_E \cdot \tau_I$ ; and  $p_4 = (1 + w_{IR}) \tau_E + (1 - w_{ER}) \tau_I$ .

The eigenvalues for Part[111] above are given only for the simplifying case where  $\tau_E = \tau_I$ , as the general case is overly complex. The eigenvalues for Part[101] are identical to those for Part[110], but depend on the weights  $w_{EC2}$  and  $w_{IC2}$  (as well as on  $\Lambda_{C2}$ ) in place of  $w_{EC1}$ ,  $w_{IC1}$ , and  $\Lambda_{C1}$ , respectively.

The global BIBO stability of the system is guaranteed, for the case where  $\tau_E = \tau_I$ , by the parameter bounds

$$\wedge \left[ \begin{array}{l} -\Lambda_R < \Lambda_{C1} < \Lambda_R \cdot \\ \frac{2 \cdot \Lambda_{C1}^2}{\Lambda_R} - \Lambda_R < \Lambda_{C2} < \Lambda_R \end{array} \right. \quad (A.29)$$

The steady-state equations for the general case of  $\tau_E \neq \tau_I$  are overly complex and are not included here. In the case where excitatory and inhibitory thresholds  $\vartheta_E = 0$  and  $\vartheta_I = 0$ , the steady-state equations are given by

$$\begin{array}{l} \bar{x}_{E1} = \bar{x}_{I1} = \frac{\iota_2 (-\Lambda_{C1}^2 + \Lambda_{C2} \cdot \Lambda_R) + \iota_1 (\Lambda_{C1}^2 - \Lambda_R^2)}{(\Lambda_{C2} - \Lambda_R) (-2 \cdot \Lambda_{C1}^2 + \Lambda_R (\Lambda_{C2} + \Lambda_R))} \\ \bar{x}_{E2} = \bar{x}_{I2} = \frac{(\iota_1 + \iota_2) \Lambda_{C1}}{2 \cdot \Lambda_{C1}^2 - \Lambda_R (\Lambda_{C2} + \Lambda_R)} \\ \bar{x}_{E3} = \bar{x}_{I3} = \frac{\iota_1 (-\Lambda_{C1}^2 + \Lambda_{C2} \cdot \Lambda_R) + \iota_2 (\Lambda_{C1}^2 - \Lambda_R^2)}{(\Lambda_{C2} - \Lambda_R) (-2 \cdot \Lambda_{C1}^2 + \Lambda_R (\Lambda_{C2} + \Lambda_R))} \end{array} \quad (A.30)$$

$$\begin{array}{l} \bar{x}_{E1} = \bar{x}_{I1} = \frac{\iota_1 \cdot \Lambda_R}{-\Lambda_{C1}^2 + \Lambda_R^2} \\ \bar{x}_{E2} = \bar{x}_{I2} = \frac{\iota_1 \cdot \Lambda_{C1}}{\Lambda_{C1}^2 - \Lambda_R^2} \\ \bar{x}_{E3} = \bar{x}_{I3} = \frac{\iota_1 (-\Lambda_{C1}^2 + \Lambda_{C2} \cdot \Lambda_R) + \iota_2 (\Lambda_{C1}^2 - \Lambda_R^2)}{\Lambda_{C1}^2 - \Lambda_R^2} \end{array} \quad (110)$$

$$\begin{array}{l} \bar{x}_{E1} = \bar{x}_{I1} = \frac{\iota_2 \cdot \Lambda_{C2} - \iota_1 \cdot \Lambda_R}{\Lambda_{C2}^2 - \Lambda_R^2} \\ \bar{x}_{E2} = \bar{x}_{I2} = -\frac{(\iota_1 + \iota_2) \Lambda_{C1}}{\Lambda_{C2} + \Lambda_R} \\ \bar{x}_{E3} = \bar{x}_{I3} = \frac{\iota_1 \cdot \Lambda_{C2} - \iota_2 \cdot \Lambda_R}{\Lambda_{C2}^2 - \Lambda_R^2} \end{array} \quad (101)$$

$$\begin{array}{l} \bar{x}_{E1} = \bar{x}_{I1} = \frac{\iota_1}{\Lambda_R} \\ \bar{x}_{E2} = \bar{x}_{I2} = -\frac{\iota_1 \cdot \Lambda_{C1}}{\Lambda_R} \\ \bar{x}_{E3} = \bar{x}_{I3} = \iota_2 - \frac{\iota_1 \cdot \Lambda_{C2}}{\Lambda_R} \end{array} \quad (100)$$

For the general case of  $\vartheta_E, \vartheta_I \neq 0$ , the activation thresholds  $\vartheta_E$  and  $\vartheta_I$  once again drop out of the steady-state derivatives  $\partial \bar{x}_{E2} / \partial \iota_1$ , leaving the form

$$\begin{aligned} \frac{\partial \bar{x}_{E2}}{\partial \iota_1} &= \left. \frac{\Lambda_{C1}}{2 \cdot \Lambda_{C1}^2 - \Lambda_R (\Lambda_{C2} + \Lambda_R)} \right|^{111} & (A.31) \\ \frac{\partial \bar{x}_{E2}}{\partial \iota_1} &= \left. \frac{\Lambda_{C1}}{\Lambda_{C1}^2 - \Lambda_R^2} \right|^{110} \\ \frac{\partial \bar{x}_{E2}}{\partial \iota_1} &= \left. \frac{-\Lambda_{C1}}{\Lambda_{C2} + \Lambda_R} \right|^{101} \\ \frac{\partial \bar{x}_{E2}}{\partial \iota_1} &= \left. \frac{-\Lambda_{C1}}{\Lambda_R} \right|^{100}. \end{aligned}$$

Stable WTA behavior is guaranteed under the conditions

$$\left[ \frac{\Lambda_{C1}^2}{\Lambda_R} < \Lambda_{C2} < \Lambda_R \right] \wedge [\Lambda_R > \text{abs}(\Lambda_{C1})]. \tag{A.32}$$

Since  $\Lambda_{C1}^2$  must be nonnegative, and  $\Lambda_R$  is constrained to be positive by the second condition in equation A.32, then  $\Lambda_{C1}^2 / \Lambda_R$  must also be nonnegative, and so  $\Lambda_{C2} > 0$  by the first condition in equation A.32. Similar to the previous models, this implies that  $w_{IC2} > w_{EC2}$ . We find that once again, for competition to occur between columns 1 and 3, the direct inhibitory coupling between those columns must be stronger than the direct excitatory coupling.

Similar conditions hold for competition between columns 1 and 2, with input applied to column 1. Stable competition can exist under the conditions

$$[0 < \Lambda_{C1} < \Lambda_R] \wedge \left[ \frac{2\Lambda_{C1}^2}{\Lambda_R} - \Lambda_R < \Lambda_{C2} < \Lambda_R \right]. \tag{A.33}$$

Even more starkly than in equation A.32 above, we have the condition  $\Lambda_{C1} > 0$ , which implies that  $w_{IC2} > w_{EC2}$ . Coupling between columns 1 and 2 must be dominated by inhibition for competition to occur. A coupling regime that supports competition in Part [111] requires the network dynamics in that partition to be unstable, leading to a transition to another partition where one edge column is inactive, thus removing the possible effect of indirect competition mediated by that column.

**Appendix B: Parameters for the Simulation Models**

Table 3: Estimation of Network Parameters.

	Value	Formula	Estimate	Units	References
1	Prop. of excitatory neurons		0.82	(Proportion)	(Gabott & Somogyi, 1986; Martin & Whitteridge, 1984)
2	Input to pyramidal cell (total)		7000	Synapses	(Binzegger et al., 2004)
3	(exc. synapses)	$R2 \times R1$	5740	Synapses	
4	(inh. synapses)	$R2 \times (1 - R1)$	1260	Synapses	
5	(exc. from other L2/3 pyr.)		3500	Synapses	(Binzegger et al., 2004)
6	Input to basket (inh.) cell (total)		4000	Synapses	(Binzegger et al., 2004)
7	(exc. synapses)	$R6 \times R1$	3280	Synapses	
8	(inh. synapses)	$R6 \times (1 - R1)$	720	Synapses	
9	Synapses per L2/3 pyr. cell axon		5000	Synapses	(Binzegger et al., 2004)
10	Synapses per basket cell axon		4200	Synapses	(Binzegger et al., 2004)
11	L2/3 pyr. local/total boutons		0.5	(Proportion)	(Binzegger et al., 2007)
12	Average spontaneous firing rate		7.56	Hz	(Noda, Freeman Jr, Gies, & Creutzfeldt, 1971)
13	exc. spikes per pC input		0.066	spikes/pC	(Ahmed, Anderson, Douglas, Martin, & Whitteridge, 1998)
14	inh. spikes per pC input		0.310	spikes/pC	(Nowak et al., 2003)
15	exc. PSP charge		0.1	pC/spike	(Binzegger et al., 2009)
16	inh. PSP charge (basket)		0.365	pC/spike	(Binzegger et al., 2009)
17	syn. release probability		0.1	(probability)	(Binzegger et al., 2009)
18	exc. synapse strength per syn.	$R15 \times R17$	0.01	pC/spike/syn.	
19	inh. synapse strength per syn.	$R16 \times R17$	0.0365	pC/spike/syn.	
20	exc. gain multiplier per syn.	$R13 \times R15 \times R17$	$660 \times 10^{-6}$	pC/pC/syn.	
21	inh. gain multiplier	$R14 \times R16 \times R17$	$11 \times 10^{-3}$	pC/pC/syn.	
22	inh. syn. strength delta (est.)	$R21/R20$	17.14	(Proportion)	
23	inh. syn. strength delta		10.00	(Proportion)	(Binzegger et al., 2009)

Table 3: Continued.

Value	Formula	Estimate	Units	References
Spontaneous input				
24 exc. spikes into L2/3 pyr. cell	$R3 \times R12 \times R17$	4339	Hz	
25 inh. spikes into L2/3 pyr. cell	$R4 \times R12 \times R17$	953	Hz	
26 exc. spikes into basket cell	$R7 \times R12 \times R17$	2480	Hz	
27 inh. spikes into basket cell	$R8 \times R12 \times R17$	544	Hz	
Estimated effective lumped output weights				
28 L2/3 pyr cell $\sum w_E$	$R1 \times R9 \times R13 \times R18$	2.71	pC/pC	
29 Basket cell $\sum w_I$ (delta)	$(1 - R1) \times R10 \times R20 \times R23$	4.99	pC/pC	
30 Basket cell $\sum w_I$ (delta est.)	$(1 - R1) \times R10 \times R21$	8.55	pC/pC	

Note: exc: excitatory; inh: inhibitory; prop: proportion; pyr: pyramidal; syn: synapses.

## Acknowledgments

---

We gratefully acknowledge Tom Binzegger, Kevan Martin and Rodney Douglas for providing the data in Figure 1 (Binzegger et al., 2007). We thank Rodney Douglas for spurring this work on in its early stages. We also thank the participants of the Winner-Take-All and Neural Computation work groups at the Capo Caccia meeting (<http://capocaccia.ethz.ch>), who provided a stimulating environment for discussion of this work and on cortical computation in general. This work was funded by a John Crampton Travelling Fellowship to D.R.M., by the European Commission (FP6-2005-015803 DAISY), by the Velux Stiftung, and by CSN fellowships to D.R.M.

## References

---

- Ahmed, B., Anderson, J. C., Douglas, R. J., Martin, K. A., & Whitteridge, D. (1998). Estimates of the net excitatory currents evoked by visual stimulation of identified neurons in cat visual cortex. *Cerebral Cortex*, *8*, 462–476.
- Amit, D. J., & Brunel, N. (1995). Learning internal representations in an attractor neural network with analogue neurons. *Network: Computation in Neural Systems*, *6*(3), 359–388.
- Anderson, J. C., Dehay, C., Friedlander, M. J., Martin, K. A., & Nelson, J. C. (1992). Synaptic connections of physiologically identified geniculocortical axons in kitten cortical area 17. *Proc. Biol. Sci.*, *250*(1329), 187–194.
- Antolík, J., & Bednar, J. A. (2011). Development of maps of simple and complex cells in the primary visual cortex. *Frontiers in Computational Neuroscience*, *5*.
- Averbeck, B. B., Latham, P. E., & Pouget, A. (2006). Neural correlations, population coding and computation. *Nat. Rev. Neurosci.*, *7*(5), 358–366.
- Baker, T. I., & Cowan, J. D. (2009). Spontaneous pattern formation and pinning in the primary visual cortex. *Journal of Physiology*, *103*(1–2), 52–68. doi: 10.1016/j.physparis.2009.05.01
- Bell, A. J., & Sejnowski, T. J. (1997). The independent components of natural scenes are edge filters. *Vision Research*, *37*(23), 3327–3338.
- Ben-Yishai, R., Bar-Or, R. L., & Sompolinsky, H. (1995). Theory of orientation tuning in visual cortex. *Proc. Natl. Acad. Sci. USA*, *92*, 3844–3848.
- Binzegger, T., Douglas, R. J., & Martin, K. A. C. (2004). A quantitative map of the circuit of cat primary cortex. *Journal of Neuroscience*, *24*(39), 8441–8453.
- Binzegger, T., Douglas, R. J., & Martin, K. A. C. (2007). Stereotypical bouton clustering of individual neurons in cat primary visual cortex. *Journal of Neuroscience*, *27*(45), 12242–12254.
- Binzegger, T., Douglas, R., & Martin, K. A. C. (2009). Topology and dynamics of the canonical circuit of cat V1. *Neural Networks*, *22*, 1071–1078.
- Blackstad, T. W. (1956). Commissural connections of the hippocampal region in the rat, with special reference to their mode of termination. *J. Comp. Neurol.*, *105*(3), 417–537.

- Blackstad, T. W. (1958). On the termination of some afferents to the hippocampus and fascia dentata: An experimental study in the rat. *Acta Anat. (Basel)*, 35(3), 202–214.
- Blumenfeld, B., Bibitchkov, D., & Tsodyks, M. (2006). Neural network model of the primary visual cortex: From functional architecture to lateral connectivity and back. *Journal of Computational Neuroscience*, 20, 219–241.
- Bock, D. D., Lee, W.-C. A., Kerlin, A. M., Andermann, M. L., Hood, G., Wetzel, A. W., . . . Reid, R. C. (2011). Network anatomy and in vivo physiology of visual cortical neurons. *Nature*, 471, 177–182.
- Boucsein, C., Nawrot, M. P., Schnepel, P., & Aertsen, A. (2011). Beyond the cortical column: Abundance and physiology of horizontal connections imply a strong role for inputs from the surround. *Front. Neurosci.*, 5, 32.
- Braitenberg, V., & Schüz, A. (1991). *Anatomy of the cortex: Statistics and geometry*. New York: Springer-Verlag.
- Compte, A., Brunel, N., Goldman-Rakic, P. S., & Wang, X.-J. J. (2000). Synaptic mechanisms and network dynamics underlying spatial working memory in a cortical network model. *Cerebral Cortex*, 10(9), 910–923.
- Connors, B. W., Bernardo, L. S., & Prince, D. A. (1983). Coupling between neurons of the developing rat neocortex. *Journal of Neuroscience*, 3(4), 773–782.
- Coultrip, R., Granger, R., & Lynch, G. (1992). A cortical model of winner-take-all competition via lateral inhibition. *Neural Networks*, 5, 47–54.
- DeFelipe, J. (2002). Cortical interneurons: From Cajal to 2001. *Progress in Brain Research*, 136, 215–238.
- DeFelipe, J., & Jones, E. G. (1998). From: A new concept of the histology of the nerve centers. In *Cajal on the cerebral cortex*. New York: Oxford University Press.
- Deller, T., Martinez, A., Nitsch, R., & Frotscher, M. (1996). A novel entorhinal projection to the rat dentate gyrus: Direct innervation of proximal dendrites and cell bodies of granule cells and GABAergic neurons. *J. Neurosci.*, 16(10), 3322–3333.
- Douglas, R. J., Mahowald, M. A., & Martin, K. A. C. (1994). Hybrid analog-digital architectures for neuromorphic systems. *IEEE Transactions on Neural Networks*, 5, 1848–1853.
- Douglas, R. J., & Martin, K. A. C. (2004). Neuronal circuits of the neocortex. *Annual Review of Neuroscience*, 27, 419–451.
- Douglas, R. J., & Martin, K. A. C. (2007). Recurrent neuronal circuits of the neocortex. *Current Biology*, 17(13), R496–R500.
- Douglas, R. J., & Martin, K. A. C. (2009). Inhibition in cortical circuits. *Current Biology*, 19(10), R398–R402.
- Douglas, R., Martin, K. A. C., & Whitteridge, D. (1989). A canonical microcircuit for neocortex. *Neural Computation*, 1, 480–488.
- Durstewitz, D., Kelc, M., & Güntürkün, O. (1999). A neurocomputational theory of the dopaminergic modulation of working memory functions. *Journal of Neuroscience*, 19(7), 2807–2822.
- Durstewitz, D., Seamans, J. K., & Sejnowski, T. J. (2000). Neurocomputational models of working memory. *Nature Neuroscience*, 3, 1184–1191.
- Ecker, A. S., Berens, P., Keliris, G. A., Bethge, M., Logothetis, N. K., & Tolias, A. S. (2010). Decorrelated neuronal firing in cortical microcircuits. *Science*, 327(5965), 584–587.



- Ermentrout, B. (1998a). Linearization of F-I curves by adaptation. *Neural Computation*, 10, 1721–1729.
- Ermentrout, B. (1998b). Neural networks as spatio-temporal pattern-forming systems. *Reports on Progress in Physics*, 61, 353–430.
- Ernst, U. A., Pawelzik, K. R., Sahar-Pikielny, C., & Tsodyks, M. V. (2001). Intracortical origin of visual maps. *Nature Neuroscience*, 4(4), 431–436.
- Fairén, A., & Valverde, F. (1980). A specialized type of neuron in the visual cortex of cat: A Golgi and electron microscope study of chandelier cells. *J. Comp. Neurol.*, 194(4), 761–779.
- Freund, T. F., Martin, K. A., Somogyi, P., & Whitteridge, D. (1985). Innervation of cat visual areas 17 and 18 by physiologically identified *x*- and *y*-type thalamic afferents. II. Identification of postsynaptic targets by GABA immunocytochemistry and Golgi impregnation. *J. Comp. Neurol.*, 242(2), 275–291.
- Freund, T. F., Martin, K. A., & Whitteridge, D. (1985). Innervation of cat visual areas 17 and 18 by physiologically identified *x*- and *y*-type thalamic afferents. I. Arborization patterns and quantitative distribution of postsynaptic elements. *J. Comp. Neurol.*, 242(2), 263–274.
- Fukuda, T., Kosaka, T., Singer, W., & Galuske, R.A.W. (2006). Gap junctions among dendrites of cortical GABAergic neurons establish a dense and widespread intercolumnar network. *Journal of Neuroscience*, 26(13), 3434–3443.
- Gabott, P.L.A., & Somogyi, P. (1986). Quantitative distribution of GABA-immunoreactive neurons in the visual cortex (area 17) of the cat. *Experimental Brain Research*, 61, 323–331.
- Galarreta, M., & Hestrin, S. (1999). A network of fast-spiking cells in the neocortex connected by electrical synapses. *Nature*, 402(6757), 72–75.
- Gibson, J. R., Beierlein, M., & Connors, B. W. (1999). Two networks of electrically coupled inhibitory neurons in neocortex. *Nature*, 402(6757), 75–79.
- Grabska-Barwińska, A., & von der Malsburg, C. (2008). Establishment of a scaffold for orientation maps in primary visual cortex of higher mammals. *Journal of Neuroscience*, 28(1), 249–257.
- Hahnloser, R.H.R. (1998a). *Computation in recurrent networks of linear threshold neurons: Theory, simulation, and hardware implementation*. Doctoral dissertation, Swiss Federal Institute of Technology Zürich.
- Hahnloser, R.H.R. (1998b). On the piecewise analysis of networks of linear threshold neurons. *Neural Networks*, 11, 691–697.
- Han, Z.-S., Buhl, E. H., Lörinczi, Z., & Somogyi, P. (1993). A high degree of spatial selectivity in the axonal and dendritic domains of physiologically identified local-circuit neurons in the dentate gyrus of the rat hippocampus. *European Journal of Neuroscience*, 5, 395–410.
- Hellwig, B. (2000). A quantitative analysis of the local connectivity between pyramidal neurons in layers 2/3 of the rat visual cortex. *Biological Cybernetics*, 82(2), 111–121.
- Holmgren, C., Harkany, T., Svennenfors, B., & Zilberter, Y. (2003). Pyramidal cell communication within local networks in layer 2/3 of rat neocortex. *Journal of Physiology*, 551(1), 139–153.
- Hubel, D. H., & Wiesel, T. N. (1968). Receptive fields and functional architecture of monkey striate cortex. *Journal of Physiology (London)*, 195, 215–243.

- Jug, F., Cook, M., & Steger, A. (2012). Recurrent competitive networks can learn locally excitatory topologies. In *Proceedings of the International Joint Conference on Neural Networks*. Piscataway, NJ: IEEE.
- Kang, K., Shelley, M., & Sompolinsky, H. (2003). Mexican hats and pinwheels in visual cortex. *Proc. Natl. Acad. Sci. USA*, 100(5), 2848–2853.
- Keller, A., & Asanuma, H. (1993). Synaptic relationships involving local axon collaterals of pyramidal neurons in the cat motor cortex. *Journal of Comparative Neurology*, 336, 229–242.
- Kisvárdy, Z. F., Martin, K. A. C., Freund, T. F., Maglóczy, Z., Whitteridge, D., & Somogyi, P. (1986). Synaptic targets of HRP-filled layer III pyramidal cells in the cat striate cortex. *Experimental Brain Research*, 64, 541–552.
- Landsman, A., Neftci, E., & Muir, D. R. (2012). Noise robustness and spatially-patterned synchronisation of cortical network oscillators. *New Journal of Physics*, 14(12), 123031.
- Levy, R. B., Reyes, & Alex, A. (2011). Coexistence of lateral and co-tuned inhibitory configurations in cortical networks. *PLoS Comput. Biol.*, 7(10), e1002161.
- Li, Z. (1998). A neural model of contour integration in the primary visual cortex. *Neural Computation*, 10(4), 903–940.
- Li, Z. (2002). A saliency map in primary visual cortex. *Trends in Cognitive Sciences*, 6(1), 9–16.
- Linker, R. (1986). From basic network principles to neural architecture: Emergence of orientation columns. *Proc. Natl. Acad. Sci. USA*, 83, 8779–8783.
- Lund, J. S. (1987). Local circuit neurons of macaque monkey striate cortex: I. Neurons of laminae 4c and 5a. *Journal of Comparative Neurology*, 257, 60–92.
- Lund, J. S., Angelucci, A., & Bressloff, P. C. (2003). Anatomical substrates for functional columns in macaque monkey primary visual cortex. *Cerebral Cortex*, 13(1), 15–24.
- Lund, J. S., Hawken, M. J., & Parker, A. J. (1988). Local circuit neurons of macaque monkey striate cortex: II. neurons of laminae 5b and 6. *Journal of Comparative Neurology*, 276, 1–29.
- Lund, J. S., & Wu, C. Q. (1997). Local circuit neurons of macaque monkey striate cortex: IV. neurons of laminae 1–3a. *Journal of Comparative Neurology*, 384, 109–126.
- Lund, J. S., & Yoshioka, T. (1991). Local circuit neurons of macaque monkey striate cortex: III. neurons of laminae 4b, 4a and 3b. *Journal of Comparative Neurology*, 311, 234–258.
- Lundqvist, M., Compte, A., & Lansner, A. (2010). Bistable, irregular firing and population oscillations in a modular attractor memory network. *PLoS Comput. Biol.*, 6(6), e1000803.
- Lundqvist, M., Rehn, M., Djurfeldt, M., & Lansner, A. (2006). Attractor dynamics in a modular network model of neocortex. *Network: Computation in Neural Systems*, 17(3), 253–276.
- Mariño, J., Schummers, J., Lyon, D. C., Schwabe, L., Beck, O., Wiesing, P., . . . Sur, M. (2005). Invariant computations in local cortical networks with balanced excitation and inhibition. *Nature Neuroscience*, 8(2), 194–201.
- Markram, H., Toledo-Rodriguez, M., Wang, Y., Gupta, A., Silberberg, G., & Wu, C. (2004). Interneurons of the neocortical inhibitory system. *Nature Reviews Neuroscience*, 5.

- Martin, K. A. C., & Schröder, S. (2013). Functional heterogeneity in neighboring neurons of cat primary visual cortex in response to both artificial and natural stimuli. *Journal of Neuroscience*, *33*(17), 7325–7344.
- Martin, K. A. C., & Whitteridge, D. (1984). Form, function and intracortical projections of spiny neurones in the striate visual cortex of the cat. *Journal of Physiology (London)*, *353*, 463–504.
- Miller, P., Brody, C. D., Romo, R., & Wang, X.-J.J. (2003). A recurrent network model of somatosensory parametric working memory in the prefrontal cortex. *Cerebral Cortex*, *13*(11), 1208–1218.
- Morishima, M., Morita, K., Kubota, Y., & Kawaguchi, Y. (2011). Highly differentiated projection-specific cortical subnetworks. *Journal of Neuroscience*, *31*(28), 10380–10391.
- Mountcastle, V. B. (2003). Introduction (special issue on cortical computation). *Cerebral Cortex*, *13*(1), 2–4.
- Mountcastle, V. B., Berman, A. N., & Davies, P. W. (1955). Topographic organization and modality representation in first somatic area of cat's cerebral cortex by method of single unit analysis. *American Journal of Physiology*, *183*, 646–647.
- Muir, D. R., Da Costa, N.M.A., Girardin, C., Naaman, S., Omer, D. B., Ruesch, E., Grinvald, A., Martin, K. A., & Douglas, R. J. (2011). Embedding of cortical representations by the superficial patch system. *Cerebral Cortex*, *21*(10), 2244–2260.
- Muir, D. R., Molina-Luna, P., Helmchen, F., & Kampa, B. (2014). *Specific excitatory connectivity for feature integration in primary visual cortex*. Manuscript in preparation.
- Noda, H., Freeman Jr, R. B., Gies, B., & Creutzfeldt, O. D. (1971). Neuronal responses in the visual cortex of awake cats to stationary and moving targets. *Experimental Brain Research*, *12*, 389–405.
- Nowak, L. G., Azouz, R., Sanches-Vivez, M. V., Gray, C. M., & McCormick, D. A. (2003). Electrophysiological classes of cat primary visual cortical neurons in vivo as revealed by quantitative analyses. *Journal of Neurophysiology*, *89*, 1541–1566.
- Olshausen, B. A., & Field, D. J. (1996). Emergence of simple-cell receptive field properties by learning a sparse code for natural images. *Nature*, *381*, 607–609.
- Peinado, A., Yuste, R., & Katz, L. C. (1993). Extensive dye coupling between rat neocortical neurons during the period of circuit formation. *Neuron*, *10*(1), 103–114.
- Perin, R., Berger, T. K., & Markram, H. (2011). A synaptic organizing principle for cortical neuronal groups. *Proceedings of the National Academy of Sciences*, *108*(13), 5419–5424.
- Perrinet, L. (2004). Finding independent components using spikes: A natural result of Hebbian learning in a sparse spike coding scheme. *Natural Computing*, *3*, 159–175.
- Peters, A. (1979). Thalamic input to the cerebral cortex. *Trends in Neurosciences*, *2*, 183–185.
- Peteanu, L., Mao, T., Sternson, S. M., & Svoboda, K. (2009). The subcellular organization of neocortical excitatory connections. *Nature*, *457*(7233), 1142–1145.
- Pinto, D. J., & Ermentrout, G. B. (2001). Spatially structured activity in synaptically coupled neuronal networks: II. lateral inhibition and standing pulses. *SIAM Journal of Applied Mathematics*, *62*(1), 226–243.

- Plebe, A. (2012). A model of the response of visual area V2 to combinations of orientations. *Network*, 23(3), 105–122.
- Ramón y Cajal, S. (1892). A new concept of the histology of the nerve centers. *Revista De Ciencias Médicas De Barcelona*, 18, 457–476.
- Rehn, M., & Sommer, F. T. (2007). A network that uses few active neurones to code visual input predicts the diverse shapes of cortical receptive fields. *Journal of Computational Neuroscience*, 22(2), 135–146.
- Ribak, C. E., & Seress, L. (1983). Five types of basket cell in the hippocampal dentate gyrus: A combined Golgi and electron microscopic study. *J. Neurocytol.*, 12(4), 577–597.
- Rutishauser, U., & Douglas, R. J. (2009). State-dependent computation using coupled recurrent networks. *Neural Computation*, 21(2), 478–509.
- Rutishauser, U., Slotine, J.-J. J., & Douglas, R. J. (2012). Competition through selective inhibitory synchrony. *Neural Computation*, 24(8), 2033–2052.
- Shamir, M., & Sompolinsky, H. (2004). Nonlinear population codes. *Neural Computation*, 16, 1105–1136.
- Somers, D. C., Nelson, S. B., & Sur, M. (1995). An emergent model of orientation selectivity in cat visual cortical simple cells. *Journal of Neuroscience*, 15(8), 5448–5465.
- Somogyi, P., Freund, T. F., & Cowey, A. (1982). The axo-axonic interneuron in the cerebral cortex of the rat, cat and monkey. *Neuroscience*, 7(11), 2577–2607.
- Soriano, E., & Frotscher, M. (1989). A GABAergic axo-axonic cell in the fascia dentata controls the main excitatory hippocampal pathway. *Brain Res.*, 503(1), 170–174.
- Sperling, G. (1970). Model of visual adaptation and contrast detection. *Perception and Psychophysics*, 8(3), 143–157.
- Stepanyants, A., Tamás, G., & Chklovskii, D. B. (2004). Class-specific features of neuronal wiring. *Neuron*, 43, 251–259.
- Swindale, N. V. (1982). A model for the formation of orientation columns. *Proceedings of the Royal Society of London. Series B, Biological Sciences*, 215, 211–230.
- Tang, H. J., & Tan, K. C. (2005). Analysis of cyclic dynamics for networks of linear threshold neurons. *Neural Computation*, 17, 97–114.
- Thompson, A. M., & Bannister, A. P. (2003). Interlaminar connections in the neocortex. *Cerebral Cortex*, 13(1), 5–14.
- Tolhurst, D. J., Smyth, D., & Thompson, I. D. (2009). The sparseness of neuronal responses in ferret primary visual cortex. *Journal of Neuroscience*, 29(8), 2355–2370.
- von der Malsburg, C. (1973). Self-organization of orientation sensitive cells in the striate cortex. *Kybernetik*, 14, 85–100.
- Weliky, M., & Katz, L. C. (1994). Functional mapping of horizontal connections in developing ferret visual cortex: Experiments and modelling. *Journal of Neuroscience*, 14(12), 7291–7305.
- Wilson, H. R., & Cowan, J. D. (1973). A mathematical theory of the functional dynamics of cortical and thalamic nervous tissue. *Kybernetik*, 13, 55–80.
- Yen, S. C., Baker, J., & Gray, C. M. (2007). Heterogeneity in the responses of adjacent neurons to natural stimuli in cat striate cortex. *Journal of Neurophysiology*, 97(2), 1326–1341.