Title
Studies of the development and mature function of a candidate layer 4 cortical circuit

Author
Kayser, Andrew,

Publication Date
1999

Peer reviewed|Thesis/dissertation

Studies of the Development and Mature Function of a Candidate Layer 4 Cortical Circuit

by

Andrew S. Kayser

DISSERTATION

Submitted in partial satisfaction of the requirements for the degree of

DOCTOR OF PHILOSOPHY

in

Neuroscience

in the

GRADUATE DIVISION

of the

UNIVERSITY OF CALIFORNIA SAN FRANCISCO

Date University Librarian

Degree Conferred: .....

To my family.

Acknowledgements

Obviously, no thesis is written in isolation. Thanks go first to Ken, whose mentorship during this project was invaluable (as goes without saying). His willingness to sit down with me to explain even the most basic of concepts was incredibly helpful; by no means is it guaranteed that a graduate student get so much attention from his PI, given as freely. In fact, his acceptance of a faculty position at UCSF in 1993 was one of my prime motivations for coming here, and it's a decision I have never regretted. Simply put, I owe him a great debt for all the time and energy he has invested in me.

Thanks too to my thesis committee. Steve Lisberger has been a great chairman. His timely and constructive feedback has been invaluable from my orals onward (as has his penchant for quick and punctual meetings). David Copenhagen opened up the topic of my return to medical school, for which I am very grateful. Steve and David also allowed me to rotate through their labs as I was immersing myself in neuroscience labwork for the first time. David, Michael Stryker and David Heeger have provided both stimulating discussions and great feedback.

I'd like to thank Nick Priebe, with whom I collaborated on the second half of this thesis. It was a pleasure to work with him; writing the paper represented by that work was a bear, and his contributions were absolutely essential, and much appreciated. The work that Todd Troyer and Anton Krukowski did with the mature circuit was also essential; having Anton, the other graduate student in the lab, to compare notes with on all kinds of lab-related and other issues was fantastic.

A couple of mentors from my past deserve mention as well. Paul Macdonald gave me my first exposure to research by allowing me to work in his lab at Stanford, and he deserves a great deal of thanks for taking on an undergraduate with little previous research experience. John Mulligan gave me the flexibility to work within his lab while I applied to medical school.

Graduate school can be an occasionally lonely and oftentimes frustrating existence. Thanks to my officemates – David Blake, Sri Nagarajan, and Ed Erwin – who made the office a place I often looked forward to being, and to Dave especially, for all the long bike rides and for associated instruction in bike maintenance. Also, to my friends – Cal, Ali, Dave, and others – who kept me on an even keel throughout.

And of course, without my family this thesis would have been an impossibility. Both of my parents have been very supportive, and I have been incredibly lucky to study at the same institution where my father is a professor. And to my brothers Beau, Chris, and Danny – thanks for helping me to keep my perspective through this whole thing.

Contributions

The material in the second chapter of the thesis resulted from a collaboration with Nick Priebe. In addition to beginning the work with synaptic depression and suggesting the role of spike-rate adaptation in contrast-dependent phase advance, he was responsible for all data derived from the spiking model. Thus, the data shown in figures 34, 41, and 43 represent his work alone, as do the spiking data found in figures 39, 42, and 44. He and I gathered the data in figures 35, 36, and 37 together. Ken and I collaborated on the writing, and thanks go to Bill Bialek and Aaron Hoffman for their invaluable help with the work shown in the appendix. The second chapter has been submitted for publication in essentially the same form as shown here.

Abstract

Studies of the Development and Mature Function of a Candidate Layer 4 Cortical Circuit

Andrew S. Kayser

How does circuitry within the cortex form? What mechanisms are responsible for the circuitry's function? Here we seek to understand the coordinated development of geniculocortical and intracortical connections within layer 4 of cat primary visual cortex, and the subsequent mature functioning of these connections. The mature geniculocortical connections are those of simple cells: projections from ON (response to light onset/dark offset) and OFF (response to dark onset/light offset) LGN neurons are organized into non-overlapping, oriented subregions (Reid and Alonso, 1995; Ferster et al., 1996). The mature intracortical connections are correlation-based (Troyer et al., 1998): excitatory cells send connections to other cells with similar receptive fields (RFs) – *i.e.* with a similar absolute spatial phase, or arrangement of ON and OFF subregions in visual space – while inhibitory cells send connections to other cells with anti-similar RFs. We show that given the presence of appropriate patterns of LGN activity during development (Miller, 1994), the model explored here can robustly give rise to this connectivity, as well as to an explanation for local invariance of RF properties in which simple cells in the same column display invariant orientation preferences, but varying spatial phases.

Building on the results of Troyer et al. (1998), we next extend the model to explain a number of "nonlinear" effects seen in mature V1, including contrast-dependent phase advance, temporal frequency tuning, and contrast saturation. Contrast-dependent phase advance is a change in the timing of the cortical response with respect to an input stimulus: as the contrast of the input increases, the cortical response moves earlier in the stimulus cycle. Temporal frequency tuning curves of V1 cells show amplification of higher frequency responses with contrast; at higher temporal frequencies the ratio between high and low contrast responses increases with increasing temporal frequency of the input. Lastly, contrast saturation curves display later saturation of cortical responses at higher temporal frequencies. By examining the effects of temporal nonlinearities, including synaptic depression and spike rate adaptation, within the context of the developed circuit, we account for these phenomena without violating experimental constraints, and in a circuit of plausible developmental origin that also displays contrast-invariant orientation tuning.

Stephen G. Lisberger

Contents

List of Tables

List of Figures

Chapter 1

Development of the Model Circuit

1.1 Abstract

In this paper we seek to understand the coordinated development of both geniculocortical and intracortical connections within layer 4 of cat primary visual cortex. We take the mature geniculocortical connections to be those of simple cells: projections from ON (response to light onset/dark offset) and OFF (response to light offset/dark onset) LGN neurons are organized into non-overlapping subregions that confer the simple cell with a well-defined orientation preference (Reid and Alonso, 1995; Ferster et al., 1996). We take the mature intracortical connections to be correlation-based (Troyer et al., 1998): excitatory cells send connections to other cells with similar receptive fields (RFs) – *i.e.* with a similar absolute phase, or arrangement of ON and OFF subregions in visual space – while inhibitory cells send connections to other cells with anti-similar RFs. We show that given the presence of appropriate patterns of LGN activity during development (Miller, 1994), the model can robustly give rise to this connectivity, as well as to local groups of layer 4 cells ("columns") in which the orientations of simple cells are invariant, but their spatial phases – the positions of their ON and OFF subregions with respect to the RFs themselves ("relative" spatial phase) or with respect to visual space ("absolute spatial phase") – are not. The model posits that this local invariance of orientation, but not of phase, results simply from Hebbian learning rules and the strength with which developing groups of cells are connected. These results hold even when multiple columns are linked together; however, the model cannot yet accurately reproduce the smooth variation of orientation across the cortical surface. Experimental predictions, as well as extensions of the model, are suggested.

1.2 Introduction

1.2.1 Overview

The development of specific connections within the primary visual cortex of many mammals depends critically on the firing patterns of neurons within the visual system. As demonstrated by studies of the map-like organization of central visual pathways, quite different patterns of connectivity emerge when changes in either the external or the internal environment induce changes in neural activity. In layer 4 of the primary visual cortex of the cat, for instance, cells that respond preferentially to inputs from one eye are organized into a number of stripes across the cortical surface (the so-called "ocular dominance columns"). Experiments in which one eye is closed during a critical period in development have shown that the identities of individual receptive fields are apparently malleable; relative to normal animals, the width of the open eye stripes increases, while that of the closed eye stripes decreases, implying that individual cells can potentially be driven by either eye (Wiesel and Hubel, 1965). Other monocular deprivation studies have argued that dendritic and axonal arbors display changes in extent and complexity (Kossel et al., 1995; Antonini and Stryker, 1996), and that these changes only occur when levels of activity in the other eye are normal (Guillery, 1972).

As the above example suggests, two broad forces seem to be at work in the visual cortex: cooperation, in that the inputs to a given cell tend to carry information from the same eye, and competition, in that afferents carrying input from the different eyes vie for cortical territory in a yet-unspecified way based upon their relative activities. Mathematically, a few simple expressions can describe these concepts. Cooperation is embodied by a form of learning proposed many years ago by Hebb – that is, that a neural network might learn appropriate correlations in its inputs by strengthening connections between co-active neurons (Hebb, 1949). Out of this idea of cooperation arises the justification for the competition: if the firing of most pre- and post-synaptic cells is somewhat correlated (as would be expected if both eyes share portions of the visual field, for instance), a purely Hebbian rule would imply that most synapses grow irrespective of patterned input. Thus, there must also exist some mechanism through which synapses become weaker, and as the experimental data suggests, this weakening should be tied to changes in other synapses. All theoretical attempts to understand the formation of both individual receptive fields and cortical maps based on afferent correlations have relied on a balance of both cooperative and various competitive effects.

Such theoretical efforts have been brought to bear not only on the development of ocular dominance, but also on the development of oriented receptive fields. In addition to preferring input from a particular eye, many neurons in visual cortex prefer stimuli (such as a bar of light, or a grating) presented at a certain angle. These RFs are hypothesized to arise from a competition not between left and right eyes, but between ON and OFF inputs (Miller, 1994); and, as for neurons with different eye preferences, orientation-selective neurons tend to be organized in a map-like fashion. Almost all models of the development of these properties, however, have assumed that intracortical connections are fixed (but see Sirosh and Miikkulainen, 1994; Rao and Ballard, 1999), an assumption that is known to be violated by the biology (Callaway and Katz, 1992, for example). The first project described here will explore whether a model with plastic intracortical connections informed solely by the afferent correlation structure can give rise to realistic oriented receptive fields and plausible intracortical circuitry, as well as to orientation maps. (Because the development of the orientation

map appears to precede development of the ocular dominance map (Crair et al., 1997), and because specific predictions exist (Troyer et al., 1998) for the correlation-based connectivity of simple cells (a subset of oriented neurons), the focus will be on development of orientation, and not of ocular dominance.) Moreover, by virtue of creating this circuit, the model will also generate a hypothesis for the development of cortical columns.

For this project to be successful, a number of issues must be addressed. First and foremost, the experimental constraints must be recognized and incorporated – these include, among others, the known anatomical extent of various horizontal and vertical connections within layer 4 (*e.g.* Callaway and Katz, 1992), the types of learning rules that operate between different neurons (eg. Komatsu and Iwakiri, 1993), and the patterns of input that might be present during the period orientation-selective responses are forming. Additionally, a number of thorny theoretical issues must be explored, including the appropriate levels of model abstraction, the appropriate ways to constrain weight development, and the best ways to characterize model outcomes. Ultimately, any model must satisfy two criteria: it must account for the experimental data, and it must generate testable experimental predictions.

1.2.2 Experimental Background

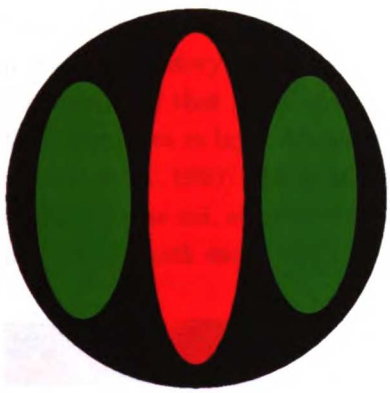


Figure 1. A cartoon of a simple cell receptive field. The figure shows the receptive field of a simple cell that prefers vertically-oriented stimuli. ON subregions are indicated by red, OFF subregions by green. As suggested by Hubel and Wiesel (1962) and later confirmed by Chapman et al. (1991) and Reid and Alonso (1995), the orientation selectivity of simple cells originates in the arrangement of their geniculocortical afferents.

The effectiveness of any developmental model must first be evaluated with reference to the desired developmental outcome – in this case, the mature physiology and anatomy of the primary thalamorecipient layer (layer 4) of the cat striate cortex. The hallmark of this layer in the cat is the simple cell, a type of neuron, either excitatory or inhibitory (Ahmed et al., 1997), that prefers stimuli of a particular orientation and phase within its receptive field (figure 1). In other words, a simple cell responds preferentially to a dark stimulus of a particular orientation (in figure 1, a vertical stimulus) flashed in an OFF subregion of its receptive field, and preferentially to a light stimulus of that same orientation flashed in an ON subregion. Due to this subregion structure, light or dark stimuli at orthogonal angles to the cell's preferred orientation produce very little response; such stimuli overlap both light- and dark-preferring areas of the receptive field, thereby recruiting less excitation (and more inhibition; see below) than optimally-aligned stimuli. Moreover, an optimally-oriented stimulus

with an inappropriate spatial phase will also suppress excitation and recruit inhibition (Ferster, 1988; Hirsch et al., 1998) – for example, an optimally-oriented light bar will inhibit the simple cell if it is placed over a dark-preferring (OFF) subregion. (One of the main differences between simple cells and another type of oriented cortical cell, the complex cell, is this phase requirement: complex cells respond well to properly-oriented bars no matter where the bars are placed in the receptive field.) Interestingly, the inhibition recruited by both optimal and non-optimal stimuli is thought to come primarily from cortical cells of similar orientation specificity (Ferster, 1986), a result that will provide a valuable constraint for modeling intracortical circuitry. Lastly, although how such orientation specificity is generated remains somewhat controversial, these properties are thought to arise from the organization of the neuron's received thalamocortical afferents (Chapman et al., 1991; Reid and Alonso, 1995), which form oriented ON (response to light onset/dark offset) and OFF (response to dark onset/light offset) receptive field subregions, and which give rise to orientation-selective responses even in the absence of almost all cortical activity (Ferster et al., 1996). In other words, the initial orientation tuning of simple cells appears to arise from the spatial organization of their inputs from the lateral geniculate nucleus (LGN), rather than from intracortical processing.

The evidence for the anatomical substrate underlying the intracortical physiology (and thus underlying a model of intracortical circuitry) is somewhat sparser than for the geniculocortical afferents. Combined electrophysiological and morphological studies of spiny stellate cells, the primary excitatory cells in layer 4, show that their dendrites and axons arborize within $150\mu m$ of the cell body[1] (Martin and Whitteridge, 1984). Likewise, the dendritic and axonal fields of the clutch cells, a major inhibitory cell in layer 4, are local (radius $\leq 200\mu m$; Kisvarday, 1992). Other work has demonstrated that spiny stellates at the border of layers 4/5 and in lower layer 3 tend to restrict their dendrites to layer 4/lower layer 3, consistent with the distribution of thalamocortical afferents (Kossel et al., 1995). All in all, these data, when combined with other anatomical and physiological findings in the cat, support two ideas: the primacy of local connectivity, and the importance of local processing (both excitatory and inhibitory) within layer 4.

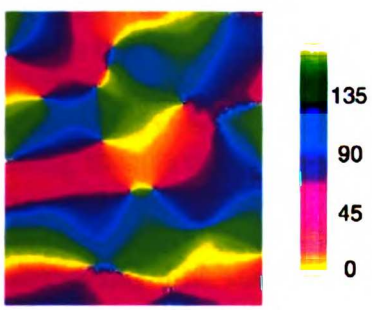


Figure 2. An example of an optically-recorded orientation map from P28 cat visual cortex, courtesy of J. Trachtenberg and M. Stryker. The bar to the right of the map shows the angle corresponding to each color. Orientation varies smoothly and periodically across the surface of the cortex, with occasional discontinuities.

In addition to being locally ordered, the cat striate cortex is also ordered on a larger scale – in particular, area 17 contains a map of orientation preference (see figure 2 for an example from cat, courtesy of the Stryker lab). As evidenced by functional mapping and electrode studies, the preferred orientation of cortical sites changes gradually across the cortical surface, with occasional

[1] For comparison, the width of a cortical hypercolumn – the cortical area occupied by left- and right-eye ocular dominance columns, or by 180 degrees of orientation – is approximately 1mm.

orientation "jumps," or discontinuities. This organization of orientation at each cortical position is columnar in nature; not only layer 4, but also the upper and lower layers show the same preferred orientation, so the map is not specific to any one cortical layer. However, not all properties are constant: while the orientation at a given point is constant within a column, the spatial phase of cells within a column varies (DeAngelis et al., 1999) – another valuable modeling constraint. The planar organization of the orientation map is mirrored by the horizontal arrangement of structures outside of layer 4; significantly, when labelled with dye and compared to optical orientation maps, long horizontal connections ($> 1mm$) from cells in layers 2/3 tend to give rise to periodic "clusters" of label centered over columns of similar orientation preference (Malach et al., 1993). Besides accounting for the existence of simple cells and local connectivity, then, a developmental model should also attempt to account for the existence of a map-like structure.

Modeling how these outcomes arise requires an understanding of the developmental process, but studying the physiological development of orientation selectivity in the cat has been somewhat problematic. Reports differ widely as to the fraction of cells displaying orientationally-selective responses at post-natal day 6 (P6), the earliest that visually-evoked responses can be measured; estimates range from 0 to 100%, with a more recent estimate falling in the neighborhood of 15% (Albus and Wolf, 1984). This discrepancy may be due largely to the delicate physiological state of the feline cortex at this age, as electrode penetrations can produce quite unstable recordings. Given these experimental problems, many researchers have turned to the ferret as a more suitable model of orientation development – while the time from conception to eye-opening is 72-75 days in both ferret and cat, the ferret is born three weeks earlier (embryonic day 41±1) than is the cat (E62±2)[2], so its physiological state tends to be much more robust while orientation selectivity is developing (see, for example, figure 7 of Issa et al., 1999). In the ferret, visual cortical cells first show visual responses at P23 (Chapman and Stryker, 1993), at which time the eyes are not yet open and most cells are not tuned for orientation. Orientation selectivity develops most quickly during week six (P35-P41), and about the beginning of week 6, if not slightly earlier, robust orientation maps can be seen in layers 2-3 (Chapman et al., 1996). In cat, laminar analyses suggest that adult-like orientation tuning evolves first in layers 4 and 6 (Albus and Wolf, 1984; Braastad and Heggelund, 1985), but one study reports that this tuning appears earliest in layer 5 (Tsumoto and Suda, 1982).

What are the inputs to these cortical systems during the time orientation develops? Broad spontaneous waves of activity (Feller et al., 1996) are present up until approximately P21 (figure 1, Wong, 1999) in the developing ferret retina, but they appear to be too broad (Miller et al., 1999) and too early (Wong, 1999) to induce development of simple cells. Alternatively, our lab has proposed that a Mexican hat correlation structure in the firing of ON and OFF neurons in the lateral geniculate nucleus may drive the formation of oriented receptive fields (Miller, 1994). [In other words, if the firing of nearby same-type LGN cell pairs (ON-ON, OFF-OFF) and more distant opposite-type LGN cell pairs (ON-OFF, OFF-ON) is best correlated within a cell's receptive field, simple cells can develop (see figure 3, in Methods).] The most direct evidence for any difference in ON-OFF firing comes from three studies. In the ferret retina, starting at P14 (and distinct from the waves mentioned above), OFF-center cells start to burst in synchrony with other OFF-center cells at

[2]Thus, eye-opening in the ferret commonly occurs between P31-P35, while in cat it often occurs between P8-P10 (Issa et al., 1999).

approximately four times the rate of ON cells (Wong and Oakley, 1996). In the adult cat retina, the firing of same-center type cells is correlated, and that of opposite-center type cells anti-correlated, out to separations at which center overlap is small (Mastronarde, 1989). (Additionally, evidence for a Mexican hat structure has been found in salamander retina (Meister et al., 1995), in which such correlations were seen between different ON-type and OFF-type ganglion cells.) In a less direct approach, Gödecke and Chapman (1998) provided evidence for a role for ON-OFF competition by showing that pharmacological inhibition of ON inputs from the retina disrupts orientation tuning – a finding consistent with experiments demonstrating a degradation of the cortical orientation tuning when ferret optic nerve was repeatedly shocked, thereby artificially correlating the output of ON and OFF cells (Weliky and Katz, 1997). Perhaps arguing against such a Mexican hat structure is a recent study by Weliky and Katz (1999), who examined the correlations between ON and OFF LGN layers in the ferret and did not find a Mexican hat relationship; but they also did not know the retinotopic locations of their recording electrodes with enough precision to define such a relationship. Thus, the existence of the hypothesized correlations is plausible but undetermined.

Other important phenomena in the developing cortical physiology have been even less well-explored with respect to the maturation of orientation tuning: these include the transient appearance of the subplate (Ghosh and Shatz, 1992); the types of oriented cells (simple or complex) joined by supragranular lateral connections; the correlation of the ease of LTP induction with the critical period (Crair and Malenka, 1995); the influences of neuromodulators (Bear and Singer, 1986) and neurotrophins (Cabelli et al., 1995); the presence of gap junctions early in postnatal life (Yuste et al., 1992); and the development of cortical inhibition (Luhmann and Prince, 1991), among other topics. Some of these topics will be discussed in more detail as they become relevant to the model (see Results).

1.2.3 Modeling Background

Despite (or perhaps, because of) the limited extent of both the adult and developmental data, a number of computational models have explored the development of orientation selectivity in the cortex. Unfortunately, many of them have incorporated biologically implausible elements. A significant proportion of these models assume the existence of oriented inputs (Bienenstock et al., 1982, for example), for which there is currently little evidence. In fact, in cats the normal development of orientation selectivity is unaffected by the presence or absence of visual activity until approximately P18-P20 (Fregnac and Imbert, 1984; Crair et al., 1998). Few other models consider the development of orientation tuning without oriented input, and the ones that do tend to make other unsupported assumptions. Linsker's models assume the presence of both excitatory and inhibitory inputs to the presumptive cortical layers (Linsker, 1986); experimental evidence suggests all inputs from the lateral geniculate are excitatory (Ferster and Lindstrom, 1983). Another model assumes that individual cortical cells receive only one connection – either ON or OFF – and that cortical circuitry produces orientation tuning (Miyashita and Tanaka, 1992). However, this hypothesis contradicts the experimental evidence for the organization of both ON and OFF thalamocortical inputs into segregated subregions (Chapman et al., 1991; Reid and Alonso, 1995).

Many models are more successful at explaining the development of orientation maps, as opposed to oriented receptive fields. As mentioned above, in an orientation map the preferred orientation of

a given cortical position/column changes gradually across the surface of the cortex, with occasional abrupt jumps (figure 2). One class of models creates realistic such maps through the use of a Mexican hat intracortical interaction function in orientation space, in which nearby orientations excite each other, and orientations further apart inhibit each other (*e.g.* Swindale, 1982). This type of relationship ensures that nearby cortical positions prefer similar orientations, while positions farther apart prefer different orientations. As Niebur and Worgotter (1994) have shown, the relevant aspect of this interaction may lie in its bandpass properties; the requirement that a given orientation appear with a certain frequency across the cortex gives rise to visually realistic cortical maps. Another class of models makes use of Kohonen's feature map algorithm (Ritter et al., 1992). In this subset of orientation theories, "low-dimensional" stimuli – numbers describing the angle and retintopic position of a stimulus, for instance – are "shown" sequentially to the model cortex. A Gaussian bubble of activity is placed over the cortical cell that is most active in response to the input, while cells outside this bubble are silenced. The response properties of active cells are updated, inducing cells nearby in cortex (which are co-active because of the Gaussian activity bubble) to gradually respond to similar patterns in the input space – in other words, to respond in a locally ordered, map-like fashion. This approach has recently been extended to a more realistic input representation (*e.g.* Riesenhuber et al., 1998), in which cortical cells receive inputs from multiple LGN cells ("high-dimensional inputs") rather than a set of numbers describing response properties ("low-dimensional inputs"); but the maps produced by these simulations do not look as realistic.

In the models of Miller, two of the above features are knitted together – the development of oriented receptive fields from non-oriented, correlated input, and the development of cortical maps via a Mexican hat-shaped (or purely excitatory) cortical interaction (Miller, 1994). One inherent strength of this and similar models is their susceptibility to linear analysis; an understanding of the early dynamics of the model leads to an accurate picture of the final results. However, the use of a fixed intracortical interaction function is potentially at odds with the aforementioned experimental data – functional interactions during development could possibly give rise to Mexican hat interactions, but anatomical evidence does not make a strong prediction one way or the other. The use of a two-dimensional cortex also neglects the complications due to simple-cell spatial phase in column and map formation, since each cortical location can have at most one phase. What drives the development of iso-orientation, and not iso-phase, columns, for instance? These questions are especially pertinent given the goal of the model to reproduce the phase-based connectivity shown by Troyer et al. (1998) to account for a number of mature visual cortical properties.

Of the few models that do allow for plastic intracortical connections, perhaps the most prominent are those of Sirosh (*e.g.* Sirosh and Miikkulainen, 1994). However, these models suffer from two serious problems: not only do individual cortical cells make both excitatory and inhibitory synapses onto other cortical cells, but the inhibitory connections also extend three times as far as the excitatory connections. In my own simulations, such intracortical connectivity readily gives rise to ocular dominance columns (data not shown), but the anatomical evidence does not support such a connectivity scheme. Moreover, Sirosh's model explicitly includes anti-Hebbian inhibitory synapses, and it does not allow the exploration of columnar differences in RF phase. Of course, Sirosh does not adopt such constraints haphazardly; certain nonlinearities drive these choices (see Discussion).

My working hypothesis will be that the core predictions of correlation-based models still hold

when intracortical connections are plastic. Because, for example, the refinement of lateral connections in layer 4 is known to occur concomitantly with the sharpening of orientation tuning (Nelson and Katz, 1995), and both excitatory and inhibitory neurons preferentially synapse onto cells of the same orientation, correlation-based learning in layer 4 could account for the form of this layer's circuitry. The aims of this portion of the thesis are thus to investigate the form of the geniculocortical and intracortical connectivity resulting from correlation-based development of neurons in layer 4 of cat primary visual cortex. Additionally, this work will give rise to an explanation for the development of cortical columns.

1.3 Motivation

In this section we motivate the hypothesis that appropriate correlations in the inputs to cortex are sufficient to generate not only simple cell receptive fields, but also a functional cortical circuit. We begin with two previously-established results: (1) that appropriate correlations in inputs can give rise to simple-cell receptive field structure, and (2) that a mature cortical circuit with correlation-based connectivity can account for a number of mature cortical properties.

1.3.1 Development of Simple Cell RFs

In previous work, Miller (1994) demonstrated that appropriate correlations in the firing of ON and OFF cells in the LGN, along with a Hebbian developmental rule and some form of competition, could give rise to simple cell receptive field structure in layer 4 of V1. This segregation of ON and OFF inputs occurs because of the "fire together, wire together" concept embodied by the Hebb rule. As shown in the upper left of figure 3, if ON LGN neurons, for example, tend to fire with other ON LGN neurons at smaller retinotopic distances, but with OFF LGN neurons at larger distances (where "small" and "large" are both defined relative to the overlap of a single cortical cell's dendritic arbor with the axonal fields of LGN cells), ON connections received by a given cortical cell from similar retinotopic locations will cluster into ON subregions, and OFF connections will likewise form OFF subregions. Additionally, if the scale of the Mexican hat correlation is such that each RF develops 2-3 subregions, these subregions tend to break symmetry and align in an oriented, rather than in a center-surround, fashion. The firing structure required in the LGN afferents is encapsulated by the correlation function shown in the lower left of the figure.

1.3.2 An Explanation for Contrast-Invariance

The contrast invariance of orientation tuning, as first demonstrated by Sclar and Freeman (1982), is simply the observation that the width of the orientation tuning of a cell does not change as the contrast of an oriented (grating) input increases. Naively, one might think that if cortical cells, and simple cells in particular, linearly integrate their inputs and apply a threshold, the width of the cells' orientation tuning *should* increase: higher contrast should lead previously subthreshold inputs to become suprathreshold. The fact that this type of behavior is not seen has led to a number of theories. Some models (Sompolinsky and Shapley, 1997) assume that the shape of the orientation

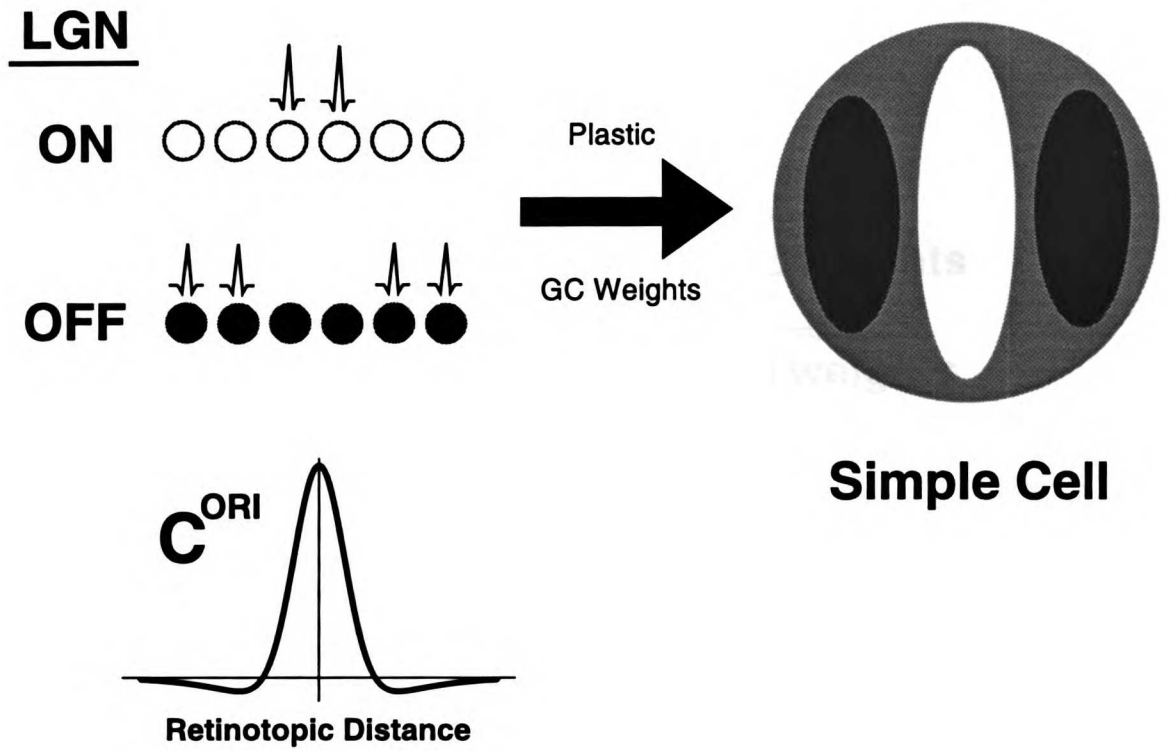


Figure 3. Development of geniculocortical (GC) connections. As demonstrated in Miller (1994), the correlation structure shown on the lower left is sufficient to generate simple cell receptive fields (upper right). The shape of C^{ORI} reflects the average firing pattern of the LGN neurons in the upper left: when ON neurons are firing, they tend to fire with other ON neurons at small retinotopic distances, but with OFF neurons at larger distances. Simple cells then arise naturally through a Hebbian learning rule.

tuning curve is generated primarily by intracortical connections, and that the cortex thus favors a response with a certain orientation tuning half-width irrespective of the input. However, because the width of tuning does depend on the spatial frequency of the input grating (Vidyasagar and Sigüenza, 1985; Webster and De Valois, 1985; Jones et al., 1987; Hammond and Pomfrett, 1990), for example, perhaps the most parsimonious explanation is that of Troyer et al. (1998). In figure 4 is shown the structure of their model circuit. As would be expected from a Hebbian developmental rule, excitatory connections are made between neurons with similar receptive fields (since those neurons would tend to fire in response to the same stimuli). Likewise, as expected from a Hebbian version of inhibition – in which an inhibitory synapse is strengthened if the postsynaptic cell is *silent* – inhibitory connections are made between pre-synaptic inhibitory cells and other cells whose RFs are of the opposite polarity – *i.e.* that have OFF subregions where the first cell has ON subregions, and vice versa. This circuit does indeed account for contrast invariance (see mature circuit modeling chapters for a more detailed explanation), and serves as the goal for the developmental work described here.

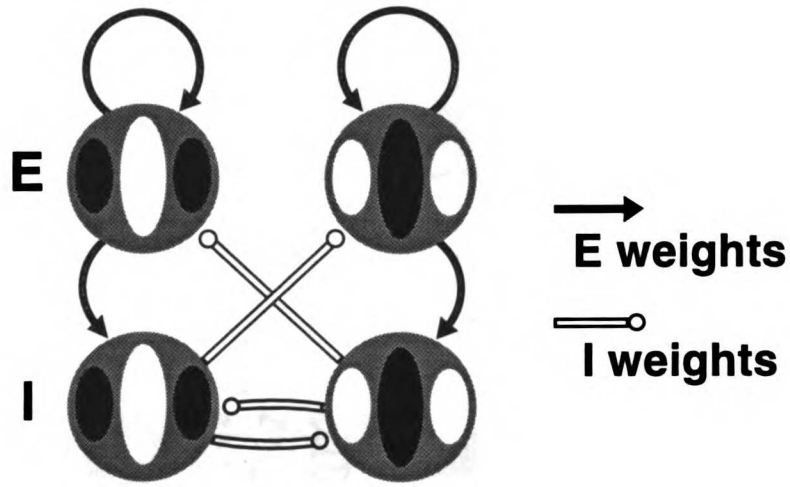


Figure 4. The full mature circuit. Excitatory weights link neurons of the same spatial phase, while inhibitory weights connect neurons of the opposite spatial phase. Troyer et al. (1998) showed that such a circuit architecture, with dominant inhibition, could account for the contrast-invariance of orientation tuning (though they did not examine inhibitory-to-inhibitory connections). This circuit forms the goal of the current developmental modeling.

1.3.3 An Explanation for Cortical Columns

One last explanation that results from this work is one for the development of cortical columns. Neurons in layer 4 share similar orientations but can vary in spatial phase (DeAngelis et al., 1999). Since simple Hebbian learning rules tend to lead strongly connected neurons to become both well correlated and well anti-correlated, and we predict that the most correlated and anti-correlated RFs share the same orientation but different spatial phases, simple Hebbian learning rules and local connectivity may be sufficient to drive development of local invariance of orientation. (One note: throughout this chapter we will often use the word "column" to refer to local groups of cells within layer 4. Of course, the simulations only model layer 4, not the full depth of cortex that the use of "column" might imply. Since layer 4 is the input layer to cortex, however, it is possible that the development of locally invariant properties in layer 4 will propagate to other layers, and lead to "columnar" invariance. In any case, we use "column", and occasionally "mini-column", as convenient short-hand for "strongly connected local group of layer 4 cells.")

1.4 Methods

In order to examine the aforementioned issues, we define a network as follows.

1.4.1 Architecture

In the model two arrays of LGN cells, representing ON and OFF LGN neurons, send connections to a series of cortical columns containing both excitatory and inhibitory cells. These columns are

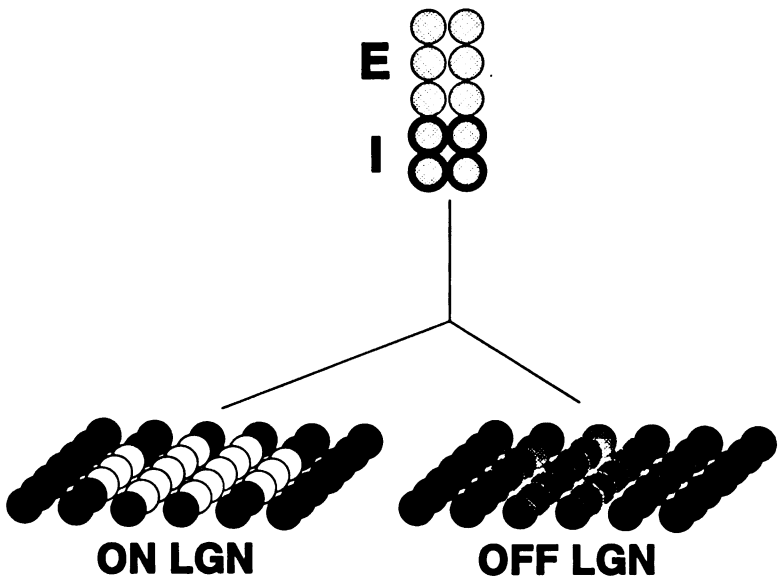


Figure 5. The structure of the model – a single column in layer 4. Each of the 10 cortical cells in the column receives afferent inputs from both ON and OFF layers of the LGN. The six excitatory cells are indicated by the thin black outlines, while the four inhibitory cells are indicated by the thick black outlines. The arrangement of the cells in the picture (in a 5 × 2 array) is arbitrary; in the model all cells effectively occupy the same retinal location and are connected all-to-all, with no cells "above" or "below" any other cells.

arranged in one of three ways: singly (*i.e.* in one column – figure 5), in a row (*i.e.* in a set of columns in one dimension), or in a grid (*i.e.* in a set of columns in two dimensions – figure 7). The number of cells in the LGN can vary with the size of the simulation, but for most single-column simulations each of the ON and OFF layers comprises a 16 × 16 grid of cells. When a single line of n cortical columns is examined (a "1-D" simulation), the LGN grid is generally $n \times 16$, though if $n < 16$, the LGN is set to 16 × 16. For a simulation in which a sheet of cortical columns is used (a "2-D" simulation), the LGN grid is either 12 × 12 or 16 × 16. Note that the number of cortical cells and the number of LGN cells need not (and in the case of single-column or 1-D simulations, cannot) be equal. When a cortical dimension is greater than one but less than the LGN dimension, the cortical columns are spaced evenly across that dimension, meaning there are retinotopic gaps between cortical columns. Within a given column, all cells are positioned at the same retinotopic point unless simulations are run with retinotopic scatter of the geniculocortical receptive fields; in the latter case, RF position is randomly chosen from a uniform distribution within an area half the radius of the geniculocortical arbor function (Hubel and Wiesel (1962); but see Hetherington and Swindale (1999)).

The cortical columns we study correspond most closely in breadth to the minicolumn, as defined by the ramification of layer 5 pyramidal cell dendrites (Peters and Yilmaz, 1993). The anatomical extent of these minicolumns parallel to the cortical surface ($\sim 100 \mu m$) also matches well the extent of individual orientation columns (Hubel and Wiesel, 1962). In the model, our approximations to these

columns generally consist of 6 excitatory cells and 4 inhibitory cells. (Columns with 16 excitatory cells and 4 inhibitory neurons can also be used, in order to preserve a 4:1 ratio of excitatory:inhibitory cells (Hendry et al., 1987), but results are no different (data not shown).) The number of inhibitory cells is never less than 4. Because we expect the presence of inhibitory cells to induce at least two spatial phases of receptive field per column (see below), any fewer than 4 cells could induce a large relative imbalance in the number of cells of each phase. Although such an imbalance can arise with 4 neurons even when no more than two phases are present (presumably through statistical fluctuations), the larger numbers of inhibitory neurons per column in the real cortex would likely make such an imbalance improbable, so we opt not to bias for such a result. Throughout, LGN cells will be represented by Greek letters (eg. α, β) whereas cortical cells will be represented by Roman letters (eg. x, y).

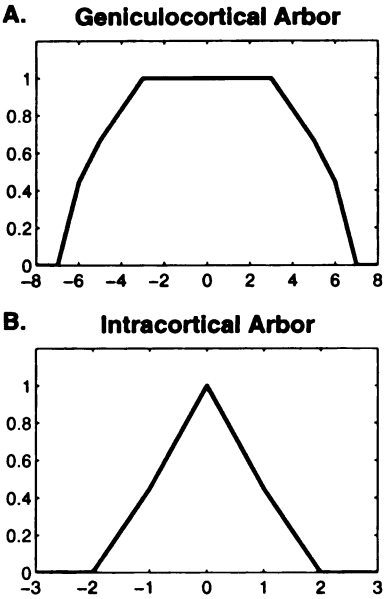


Figure 6. Default arbor functions. Retinotopic distance relative to a center point of zero is shown on the x-axis. **A.** Geniculocortical arbor function. **B.** Intracortical arbor function. Note the different scales on each abscissa.

Connectivity is represented by static arbor functions A^Y, where Y represents one of the 5 types of weight: geniculocortical (G), intracortical excitatory-to-excitatory ($E \rightarrow E$), inhibitory-to-excitatory ($I \rightarrow E$), excitatory-to-inhibitory ($E \rightarrow I$), or inhibitory-to-inhibitory ($I \rightarrow I$). The arbors define the possible connectivity; areas for which the arbor function is zero make no anatomical (and thus no functional) connections to the cell in question. Within the arbor, weights are dynamically modified (see below). The form of the arbor function for the geniculocortical weights (figure 6A) represents the overlap of a smaller cortical dendritic arbor with a larger LGN axonal arbor (Miller, 1990). The intracortical arbor functions vary in form, but are always of more limited range, and tail off more rapidly (figure 6B) – because of computational limitations, they extend no more than a radius of one column, which is slightly narrower than the anatomical constraints on local connectivity would require (see Background). These functions represent the influence of activity-independent signalling mechanisms that presumably set the scope of the arbor early in development.

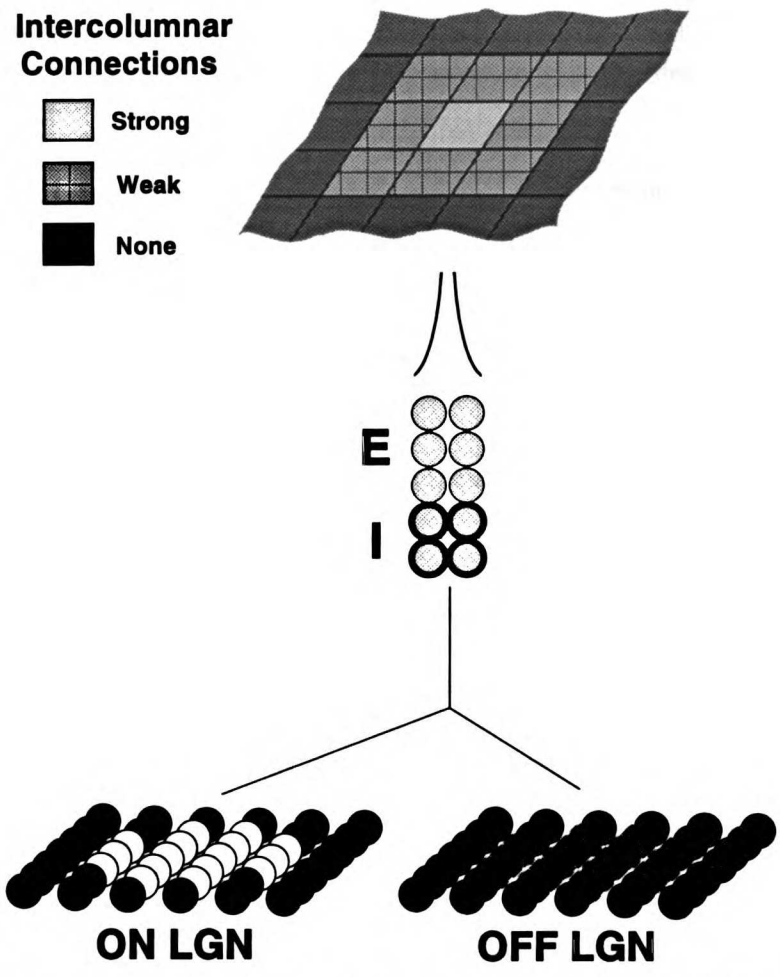


Figure 7. The structure of the model – multiple columns in layer 4. Each square at top represents a cortical column; the neurons in the central column are shown below. Adjacent columns receive afferent inputs from adjacent ON and OFF layers of the LGN. Intracortical connections to and from the central column are illustrated by the gray shading and cross-hatching; all other columns have similar connections to their neighbors (not shown). As denoted by the legend at top, these connections are strong within a column, weaker to adjacent columns, and absent at larger distances.

1.4.2 LGN Inputs

In previous models from Miller et al. (Miller et al., 1989; Miller, 1994; Erwin and Miller, 1998), LGN inputs were represented by correlation functions, so the moment-to-moment activities of LGN neurons could be neglected. The need to represent cortical activities explicitly in the current model, however, also requires the use of explicit LGN neuronal activities. To construct such activity patterns we randomly assign each LGN neuron an initial value of -0.5 or 0.5, following the work of Goodhill (1993). These random activities $r_{rnd}^{N/F}$ (where N refers to ON LGN cells, F to OFF) are then correlated between ON and OFF layers at each LGN position α, as also defined by Goodhill (1993):

$$r_{cor}^{N/F}(\alpha) = (1-h)r_{rnd}^{N/F}(\alpha) + hr_{rnd}^{F/N}(\alpha) \tag{1.1}$$

where h is a number between 0 and 0.5, inclusive[3], and $r_{cor}^{N/F}(\alpha)$ represents the new, correlated value of r at position α.

These values $r_{cor}^{N/F}$ are then convolved with a function C – applied directly to the ON LGN, multiplied first by -1 for the OFF LGN – that embodies our hypothesis about the second-order structure of LGN activity patterns: namely, that ON (OFF) cells tend to fire with other ON (OFF) cells at short distances, but to fire with OFF (ON) cells at larger distances (figure 3). C takes the form of a difference-of-gaussians (Mexican hat):

$$C(\alpha) = e^{-\alpha^2/\sigma^2} - (1/9)e^{-\alpha^2/(3\sigma)^2} \tag{1.2}$$

for which $C(0)$ is set to 1, and the convolution is

$$r_{sm}^{N/F}(\alpha) = \sum_{\beta} C(\alpha-\beta)r_{cor}^{N/F}(\beta) \tag{1.3}$$

where $r_{sm}^{N/F}(\alpha)$ is the value of r at position α after smoothing by C. Note that periodic boundary conditions are used in the LGN (as in the cortex) – *i.e.* the left edge of the LGN grid is assumed to be directly adjacent to the right edge, and the top edge is assumed to be directly adjacent to the bottom edge (giving the grid a toroidal topology). The resulting correlation structure is very similar in shape to the Mexican hat convolution function, but with a slightly broader central peak and slightly deeper adjacent troughs.

Finally, the rms (root-mean-square) value of the activities r_{rms} is computed and added to each cell to represent background activity, after which all activities are rectified at zero to give rise to LGN spike rates $s_{\alpha}^{N/F}$:

$$s_{\alpha}^{N/F} = \left[r_{sm}^{N/F}(\alpha) + r_{rms}\right]^{+} \tag{1.4}$$

where $[x]^{+} = x,\ x > 0;\ = 0$, otherwise. A sample LGN activity pattern is shown in figure 8.

1.4.3 Cortical Activities

Cortical cells receive geniculocortical input g_x

$$g_x = \sum_{\alpha,I} w_{x\alpha}^{I} s_{\alpha}^{I} \tag{1.5}$$

[3]Note that such a correlation preserves the mean of the activities, but not the variance. We could also conserve the variance, in which case the $1-h$ term becomes $\sqrt{1-h^2}$.

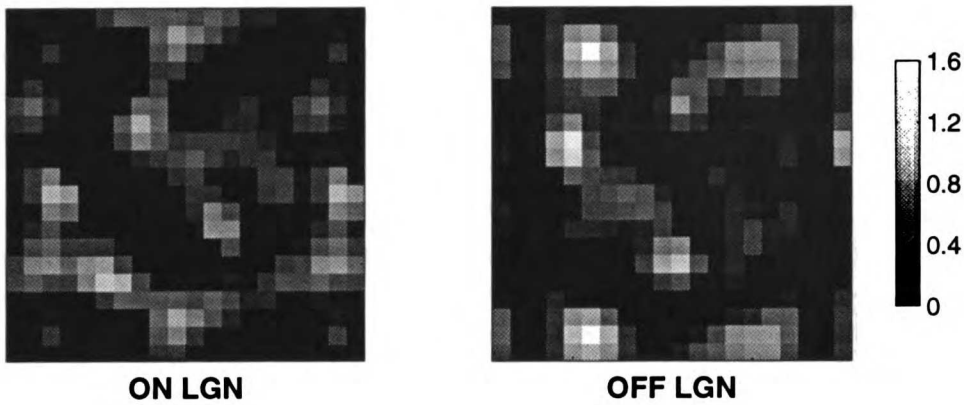


Figure 8. Sample LGN input patterns at a random time step. h is set to 0.5, so the ON and OFF patterns are strongly anti-correlated. Each pattern is also partially reproduced at the boundaries to emphasize the wrap-around boundary conditions; thus, the size of the LGN patterns shown is 20 × 20, rather than 16 × 16. The activities of individual LGN cells are indicated by the scale to the right of the figure.

where $I = N/F$, and $w^I_{x\alpha}$ is the afferent weight from ON/OFF LGN cell α to cortical cell x. These inputs are then iterated through the cortical weights t_{xy} to produce cortical activities a_x:

$$\frac{da_x}{dt} = -a_x + \sum_j t_{xj} f^E(a_j) - \sum_k t_{xk} f^I(a_k) + g_x \tag{1.6}$$

where

$$f^E(a_x) = \begin{cases} 0, & a_x \leq 0 \\ a_x, & 0 < a_x < 1 \\ 1, & a_x \geq 1 \end{cases} \qquad f^I(a_x) = \begin{cases} 0, & a_x \leq 0 \\ 1.5a_x, & 0 < 1.5a_x < 2 \\ 2, & 1.5a_x \geq 2 \end{cases} \tag{1.7}$$

The functions f, as linear approximations to a sigmoid, represent limits on the firing of cortical cells: firing rates cannot fall below zero, nor rise above a peak value. The factor of 1.5 in the definition of f^I represents one of the assumptions of the model: in addition to their ability to fire at higher rates, inhibitory cells also have higher gains (see Assumptions). At every time step a new LGN image is presented. The image is held on until the cortical activities have converged. Empirically, the discretized form of equation 1.6 converges within 40 timesteps under a simple Euler integration scheme, as accurately as and more quickly than competing methods. As implemented here, a fourth-order Runge-Kutta integration of similar accuracy is roughly twice as slow, and a Newton-Raphson method (Press et al., 1992) requires computing the Jacobian whenever the weight matrix changes, overwhelming any time savings achieved elsewhere in the routine.

1.4.4 Learning Rules

Learning rules implemented in the model are generally of a Hebbian form: $\delta w = \lambda(post \times pre)$, where λ is a function to be defined. Geniculocortical and excitatory intracortical weights follow a covariance rule; for a given intracortical weight w, for instance, if z_y is the activity of the postsynaptic cortical cell y, z_x is the activity of the presynaptic cell x, and $\bar{z}$ is the recent time-average of the activity z,

$$\delta w = \begin{cases} z_y > \bar{z}_y \text{ or } z_x > \bar{z}_x & : \quad (z_y - \bar{z}_y)(z_x - \bar{z}_x) \\ \text{otherwise} & : \quad 0 \end{cases} \tag{1.8}$$

(In the simulations described in Results, z in the above rule can represent either the membrane voltage a or the firing rate $f^{E/I}(a)$.) In words, the above equation states that if both pre- and postsynaptic neurons are more active than their recent respective time averages, the weight between them is strengthened. If one neuron is more active than its recent time average but the other is less active, that weight becomes weaker. Lastly, also embodied by the above equation is our assumption that when neither the pre- nor the postsynaptic neuron is active, there is no change in the weight. Thus, neurons that fire together, wire together; and neurons that do not fire together should share weak or no connections.

All inhibitory weights learn under a different learning rule, the exact form of which can vary from simulation to simulation. One is a variant of equation 1.8; if excitatory weights increase in strength when they successfully activate the postsynaptic cell, and decrease when they do not, then perhaps inhibitory weights should increase when they successfully inhibit the postsynaptic cell, and decrease when they do not. Such a rule takes the following form:

$$\delta w = \begin{cases} z_y > \bar{z}_y \text{ or } z_x > \bar{z}_x & : \quad -(z_y - \bar{z}_y)(z_x - \bar{z}_x) \\ \text{otherwise} & : \quad 0 \end{cases} \tag{1.9}$$

and will subsequently be referred to as an inhibitory covariance rule. For the most part, however, inhibitory weights learn under a slightly different rule inspired by the experimental work of Komatsu (Komatsu and Iwakiri, 1993; Komatsu, 1994; Komatsu, 1996). He found the following in rat visual cortical slice:

1. In the presence of blockers of excitatory neurotransmission (DNQX and APV), a stable increase in the initial slope of the IPSP occurs in response to brief tetani (Komatsu and Iwakiri, 1993)
2. This increase is not dependent on post-synaptic voltage (Komatsu, 1994)
3. LTP induction is prevented by the application of $GABA_B$, but not $GABA_A$, antagonists (Komatsu, 1996)
4. If APV is removed from the bath solution, tetanic stimulation induces a decrease in the initial slope of the IPSP (Komatsu and Iwakiri, 1993)
5. If APV is removed, and NMDA is applied to the post-synaptic (presumably excitatory) neuron, the initial slope of the IPSP decreases (Komatsu and Iwakiri, 1993)

Translating these findings into a computationally convenient form, we suggest that for inhibitory synapses

1. LTP occurs when the postsynaptic cell is both receiving more inhibition than its recent time average, irrespective of the membrane voltage, and the presynaptic input is greater than its recent time average.
2. LTD occurs when the postsynaptic cell is depolarized, and the presynaptic input is greater than its recent time average.
3. Under other conditions, no change occurs.

Of course, there are potential problems with this formulation. Still unclear is whether presynaptic activity is required for LTD of inhibitory connections; the application of NMDA is sufficient to induce LTD, yet the activity of the presynaptic cells is unknown. Assuming relative quiescence of these cells, the above learning rule should require no presynaptic component. However, this alternative – that LTD of inhibitory synapses occurs whenever the postsynaptic voltage is greater than its recent time average – seems implausible, if only because synapses from presynaptic inhibitory cells that were not firing would also be affected. Additionally, under this "Komatsu" formulation, if a cell simultaneously receives a large number of excitatory and inhibitory inputs, both LTP and LTD can be active (since inhibitory LTP does not depend on post-synaptic activity). Even more fundamentally, the identity (excitatory or inhibitory) of the postsynaptic neurons is not reported, so the above results may only apply, for example, to inhibitory connections onto excitatory cells.

Notwithstanding these caveats, we formulate the above rules into the following equation:

$$\delta w(t) = -[z_y - \bar{z}_y]^+[z_x - \bar{z}_x]^+ + \beta[\bar{a}_y^I - a_y^I]^+[z_x - \bar{z}_x]^+ \tag{1.10}$$

where z_y is the activity of the postsynaptic cell y, a_y^I is the inhibitory input received by the postsynaptic cell, z_x is the activity of the presynaptic (inhibitory) cell x, $\bar{z}$ once again represents the recent time-average of z, and β is set to 1. (Note that the weight is a positive quantity, while the inhibitory input a_y^I is negative, and that we use the variable a rather than z for the inhibitory input term to emphasize that this quantity is always a voltage, never a firing rate.) The first term decreases the strength of the inhibition: if both pre- and postsynaptic neurons are active, the magnitude of the weight should decrease. The second term is responsible for increases in the strength of inhibition: if the presynaptic neuron is firing, and the postsynaptic neuron is receiving inhibitory input greater than its recent time average, the magnitude of the weight should increase. Note that this rule is very close to Hebbian: when inhibition successfully silences the postsynaptic cell (an event that should correlate quite well with the amount of inhibition received by that cell), the inhibition becomes stronger; when inhibition is unsuccessful in preventing postsynaptic activity, it becomes weaker. Given the uncertainties in this formulation, we explore simulations using both the inhibitory covariance rule and the Komatsu rule, as described in the Results.

One unresolved issue remains the form of the postsynaptic activity incorporated into the learning (for both LTP and LTD of excitatory weights, and for LTD of inhibitory weights). The presynaptic contribution to the learning rule should be a function of the spiking activity of the presynaptic neuron, presumably because the synapse is activated by the arrival of presynaptic spikes. However, whether membrane voltage or spiking is the relevant postsynaptic variable is unclear: removing the magnesium block on the NMDA channel through membrane depolarization may be the necessary first step for excitatory LTP (Bliss and Collingridge, 1993) or inhibitory LTD (Komatsu and Iwakiri, 1993), but this result does not argue for (or necessarily against) spiking. To cover both these

possibilities, we examine the effects of learning on either postsynaptic voltage or postsynaptic spiking in the model.

Lastly, in practice all of the above weight changes are accumulated over some number of images before we actually update the values of the weights (see Assumptions). The number of images, referred to as a "batch," varies, but is usually set to 40. Additionally, before the constraints described below are applied, for the first 200 batches of each simulation we separately normalize (1) all the geniculocortical and (2) all the intracortical weight changes such that their respective rms values match preset learning rates (generally an rms value of 0.001 for each). This normalization controls the speed with which the simulations develop, and allows us to look either more finely or more coarsely at the time course of development. After the first 200 batches, the last factor required to normalize the weight changes to the desired learning rate is applied for the remainder of the simulation (see Assumptions).

1.4.5 Competitive Constraints

The above learning rules are not sufficient to allow the model to develop stably. Suppose, for instance, that the excitation to a cell overcomes the inhibition and causes the cell to fire. By the above learning rules, the excitatory synapses should strengthen, while the inhibitory synapses should weaken. With the next inputs, the cell should be even more prone to firing, leading the excitation to strengthen, the inhibition to weaken, and the cell to become yet more prone to firing. Such a positive feedback loop must be prevented by some form of competition. Some experimental data about compensatory mechanisms are beginning to emerge (Rutherford et al., 1997; Turrigiano et al., 1998; Steele and Mauk, 1999), but more work is necessary before they provide a strong guide for theoretical studies (Miller, 1996).

To implement competition, we subtractively constrain the sums (Miller and MacKay, 1994) of various sets of weights to match fixed, predetermined values after weight updates are applied. The sums of each of the 5 postsynaptic weight types onto a given cell are conserved: one sum for the geniculocortical weights (GSUM), and one for each of the $E \rightarrow E$, $I \rightarrow E$, $E \rightarrow I$, and $I \rightarrow I$ weights (EESUM, IESUM, EISUM, and IISUM, respectively, as appropriate for excitatory versus inhibitory cells). The total sum of the weights to a given cell i, therefore, remains constant over time:

$$\frac{d}{dt} \sum_j w_{ij}^Y = 0 \tag{1.11}$$

where w_{ij}^Y is the weight of type Y from presynaptic cell j to postsynaptic cell i. Complicating the application of this constraint is the fact that the weight values themselves are bounded (saturation constraints). Weights cannot change sign, so they are bounded below by 0, and above by a maximum proportional to the value of the arbor function at that weight position:

$$0 \leq w_{ij}^Y \leq wmax_{ij}^Y \tag{1.12}$$

where $wmax_{ij}^Y \propto kA_{ij}^Y$. For geniculocortical weights, $k = 0.028$ for an arbor of 5, $k = 0.02$ for an arbor of 6; for all intracortical weights, $k = 0.5$. (The different values of k were not extensively explored; see Miller (1994) for a study of the effects of a similar parameter – s_{max} – for the geniculocortical weights.) Presynaptic weight sums are of two types – projected excitatory and projected

inhibitory – and are constrained similarly. We assume, in our formulation of the Komatsu learning rule, that presynaptic projections need not know the identitity of their postsynaptic targets (*i.e.* whether those targets are excitatory or inhibitory), but there is some evidence against this position (Rutherford et al., 1997). Presynaptic constraints are *not* applied to the geniculate projections, for the following reasons. In all simulations, because central weights experience a greater total correlation with the rest of the receptive field (Miller, 1990), geniculocortical weights tend to cluster in the middle of the arbor function. If presynaptic weight sums are held fixed, those geniculate cells that do not project to the center of any receptive field (due to the large relative number of geniculate versus cortical cells) will force the cortical cell to unduly favor the edges of its receptive field. When the dimensions of cortex and LGN are exactly equal, presynaptic weight sums can be applied, but this scenario has not yet been well-explored.

The algorithm for implementing these constraints after weight changes have been calculated is as follows (similar to Erwin and Miller, 1998):

1. Apply weight cutoffs to all weights of a particular weight type, then sum the new weights.

$$\begin{aligned} sum_i^Y &= \sum_j P(w_{ij}^Y + \delta w_{ij}^Y) \\ \text{where } P &= \begin{cases} w_{ij} \leq 0 & : \quad 0 \\ 0 < w_{ij} < wmax_{ij}^Y & : \quad w_{ij} \\ w_{ij} \geq wmax_{ij}^Y & : \quad wmax_{ij}^Y \end{cases} \end{aligned} \tag{1.13}$$

2. Compute the difference between this sum and the desired (constraint) sum, divide this difference by the sum of the arbor function (the number of weights), and subtract this fraction from each weight. Then apply weight cutoffs again, and compute the new sum.

$$sum_i^Y = \sum_j P(w_{ij}^Y + \delta w_{ij}^Y - \epsilon), \text{ where } \epsilon = \frac{sum_i^Y - sum^Y}{\sum_j A_{ij}} \tag{1.14}$$

3. (First pass only) The difference between the actual sum and the desired sum is a monotonic function of ϵ. If this difference is greater than zero, bound with the current fraction ϵ and the largest possible negative weight change; if the difference is less than zero, bound with the current fraction ϵ and the largest possible positive weight change. (The largest possible positive and negative bounds can be calculated because of a bound on the weight change at each time step; see Assumptions.)
4. Take the mean of the bounds as the next upper bound if the difference is greater than zero, or as the next lower bound if the difference is less than zero.
5. Calculate the new weight sum using this mean as the fraction ϵ to be subtracted from each weight.
6. Repeat the above steps until the difference in the sums is below a predetermined maximum value (1×10^{-6} in the simulations presented here), and set the new weight value to $w + \delta w - \epsilon$.

To prevent any one cell from losing all of its incoming or outgoing weights, subtractive constraints are applied either pre- or post-synaptically, at alternating time points. This alternating procedure stems

from practical considerations: simultaneously satisfying both pre- and postsynaptic constraints exactly is computationally difficult, but with a small learning rate, applying the constraints successively is a good approximation.

One small modification of this scheme occurs when inhibition is allowed to increase slowly across the time course of the simulation. This "inhibitory ramp" affects the performance of the model under certain learning rules, and has some experimental justification: slice recordings from layer II/III regular spiking cells, taken from rats of different ages, show that the IPSP reversal potential grows more negative, and the IPSP peak conductance increases, as rats mature (Luhmann and Prince, 1991). In the model, we modify the above constraint algorithm to incorporate the ramp by constraining inhibitory weights to gradually increasing values of the inhibitory weight sums over time.

1.4.6 Mature Properties: Contrast Invariance

Once simulations finish developing, we examine whether the developed circuits demonstrate contrast-invariant orientation tuning (because the goal of the project is not just to define the circuitry, but to show that the developed circuitry also functions appropriately). To this end we display fixed, oriented gratings of 36 different orientations and 8 different spatial phases to mature networks. The response of each neuron to each orientation is determined by the same cortical dynamics applied during development; final responses are averaged across spatial phase; and the mean and standard deviation of each neuron's orientation tuning curve is then determined. Additionally, it is in terms of contrast invariance that the weight sums used to constrain each of the types of intracortical weights become important. Inhibition must be strong enough to reduce/remove responses to the null orientation, yet excitation must be strong enough to produce firing rates of reasonable magnitude. Our requirements are (1) minimal responses at the null orientation, (2) orientation tuning width as small as possible, and (3) strength of inhibition as small as possible. To meet these requirements with minimal computational expense (see Results, page 34), we started with the final state of a simulation run with EESUM = EISUM = IESUM = IISUM = 1.0, scaled the values of the weights to give all combinations of EESUM = 0.125, 0.25, 0.375, 0.5; EISUM = 0.125, 0.375, 0.625, 0.875; IESUM = 1.5, 1.75, 2, 2.25, 2.5, 2.75, 3, 3.25; and IISUM = 0.125, 0.25, 0.375, 0.5; and displayed gratings as described above. We examined the outputs of this search for combinations that appeared to meet these criteria, and arrived at the following weight sums: EESUM = 0.125, EISUM = 0.5, IESUM = 2.25, and IISUM = 0.25. For columns of 6 excitatory and 4 inhibitory cells, these postsynaptic sums correspond to projected excitatory and inhibitory sums of approximately 0.458 and 3.625, respectively. Because GSUM is important only in setting the scale of the cortical activies, it was not explored, and was always set to 1.0.

One potential source of ambiguity in the examination of contrast invariance is the magnitude of the LGN inputs at different contrasts. To deal with this issue, we take experimentally-determined F1 responses to 4 different contrasts (10, 20, 40, and 80%) for gratings drifting at 8 Hz (Sclar, 1987). (The 8Hz temporal frequency is arbitrary; it is used solely because it gives the largest LGN responses, and we wanted to test the model with the largest contrast changes possible.) We then compare these LGN responses to an ON background rate of 15 Hz (see Methods in the next chapter) to get a multiplicative factor for each contrast response relative to background.

To determine the appropriate background rate for the grating inputs, we note that the mean of the LGN ON input in the developmental simulations is roughly 0.3, in arbitrary units. Presumably these developmental inputs are of low "contrast"; thus we set the background rate such that this rate, plus our guess for the (unrectified) amplitude of the sinusoidal oscillation across the LGN corresponding to a 10% contrast grating, gives a mean LGN input of roughly 0.3 after LGN rectification. Because we are free to set the LGN rates directly – *i.e.* as simply a mean plus a modulation – our typical background rate is 0.2, which then constrains the amplitude values at 10, 20, 40, and 80% contrast to be 0.5728, 0.8748, 1.1827 and 1.4477, respectively.

1.4.7 Assumptions

A number of assumptions are either explicitly or implicitly used in the model.

LGN

1. Fixed arbor functions determine the largest extent of connectivity. This assumption is not too onerous; weights are free to refine within a relatively large area.
2. LGN inputs are determined by Mexican-hat within-layer correlations, and by between-layer anticorrelations. The biological existence of this correlation structure is a prediction of the model (Miller, 1994).
3. Ergodicity is assumed in the calculation of the mean LGN activity (*i.e.* the average of all LGN activities across space at one time is equal to the average of one LGN cell's activity across time). Because our LGN activity statistics remain constant, this assumption is valid within the model, and it reduces the number of computations that must be performed at each time step. That LGN statistics may not remain constant is certainly possible, but given the lack of experimental evidence, we opt for the null hypothesis of no, or insignificant, change.

Cortex

1. Cortical cells make no self connections. This assumption was made initially to prevent the model cortex from becoming unstable. It may be possible to relax it.
2. Simulations begin with all-to-all intracortical connectivity. Experimentally, initial axonal arbors from layer 4 neurons to other layer 4 cells are unstructured and sparse (Callaway and Katz, 1992), rather than unstructured and all-to-all. However, it is computationally convenient to use all-to-all connectivity, which can be thought of as representing all potential intracortical connections. This assumption is nonetheless relaxed at times by the "probability" parameter, which establishes intracortical connections with a probability less than 1 (Results, page 31).
3. Ramps of inhibition generally start at 20% of their final value, and increase linearly. The value of 20% is varied (Results, page 30). The linearity of the ramp is chosen for simplicity; inhibitory ramps with nonlinear functional forms are certainly possible, but the experimental data do not argue one way or the other.
4. Cortical inputs are converted to cortical activities through a linear approximation to a sigmoid function (f^E/f^I). Again, this assumption improves the speed of the simulation.

5. Inhibitory cells have greater gains and maximum firing rates than do excitatory cells (McCormick et al., 1985). We therefore set the inhibitory gain to be 1.5 times that of the excitatory cells, and the peak inhibitory firing rate to be twice that of the excitatory cells.
6. Running averages of cortical activities (necessary for use of the learning rules) are exponential.

Weights

1. Both pre- and postsynaptic weights are subtractively constrained. As mentioned in the Background and Discussion, subtractive constraints are an approximation to a yet-undetermined competitive mechanism.
2. Postsynaptic weight sums for all cortical cells are equal across cells. This constraint is likely not necessary, as the presynaptic weight sums used in all the simulations are different for excitatory versus inhibitory cells; but it is convenient. Adding a random factor ϵ to each weight sum would not be expected to alter model outcomes.
3. Individual intracortical weights have a maximum value of $k = 0.5$ times the allotted weight sum; individual geniculocortical weights have a maximum value of $k = 0.02$ (GC arbor = 6) or $k = 0.028$ (GC arbor = 5) times the allotted weight sum. Variation in this parameter for the geniculocortical weights was examined in (Miller, 1994), but no further variation has been explored here.
4. The learning rates (*i.e.* the rms values to which the summed geniculocortical and intracortical weight changes are normalized after each batch) are only enforced over the first 200 batches. This assumption is an artifact of the fact that different cortical columns, in a multi-column simulation, will learn at different rates. If the learning rates are strictly enforced throughout the simulation, the global normalizing factors required to enforce them will grow large enough, because of the larger weight changes in the more developed columns, that the weight changes in less developed columns become tiny. Consequently, less developed columns never fully develop. Letting the learning rate vary somewhat by fixing the normalizing factor after an arbitrary period of time (here, 200 batches) eliminates this problem. We have not yet tried fixing the normalizing factor after less than 200 batches.
5. Maximum weight *changes* are capped at the maximum *weight* value. As above, this assumption helps less developed columns to "catch up" to more developed columns.
6. Excitatory and inhibitory learning rules are not fully constrained by experiment (see Learning Rules).
7. Weight changes are averaged over some number of image presentations (the aforementioned "batch", generally 40 images) before weights are updated. This assumption follows from the idea that weights change on a slow time scale relative to neural activities. In Parameters, this assumption is examined directly.
8. Pre- and postsynaptic constraints on the intracortical weight sums are applied alternately (see Competitive Constraints).

1.4.8 Analysis

After simulations have developed, determining the orientation and phase of the receptive fields forms the basis for much later analysis. To find these parameters, we analyze RFs via 2-dimensional Fourier transform (Miller, 1994). The point in Fourier space with maximum absolute value ($\sqrt{Re^2 + Im^2}$) determines the preferred orientation and spatial frequency, as well as the spatial phase, of the neuron, all of which accord well with the perceived values.

To determine the orientation selectivity of individual geniculocortical RFs, we find the largest absolute value of the above 2-D Fourier transform in each of 18 10° bins from 0° to 180°. This binning gives rise to a histogram of "responses." In a manner similar to that of Chapman and Stryker (1993), we then take the Fourier transform of the histogram; well-oriented RFs should show a single peak in this histogram – *i.e.* a well-defined preferred orientation – and we should therefore find a strong F1 (first harmonic) component. (Chapman and Stryker use the F2 (second harmonic) because they examine orientations over 360°.) We then normalize by the square root of the power in all Fourier components, including the DC:

$$selectivity = \frac{\sqrt{2}F_1}{\sqrt{\sum_a {F_a}^2}} \tag{1.15}$$

where F_a is the amplitude of the Fourier component at frequency a. The factor of $\sqrt{2}$ in the numerator arises because the Fourier transform has both positive- and negative-frequency components. If the RF were perfectly selective – *i.e.* all the power were found in the first harmonic F_1 – the denominator would give rise to the term $\sqrt{2}F_1$; thus, we also multiply the numerator by $\sqrt{2}$ to give a selectivity scale from 0 to 1. In practice, such selective RFs are not found; the ubiquitous presence of a DC component leads to orientation selectivities of approximately 0.6 for RFs that are strongly-oriented to the eye.

Correlations between receptive fields are computed by taking a point-by-point multiplication of the ON-OFF difference (*i.e.* ON weight - OFF weight) at each point within the neurons' receptive fields. The resulting sum of products is multiplied by the strength of the connection between the cells:

$$C_{ij} = I_{ij} \sum_{\alpha} S_{i\alpha} S_{j\alpha} \tag{1.16}$$

where $S_{i\alpha}$ is the RF weighting at the retinotopic position α for the cortical neuron at position i, I_{ij} is the intracortical weight from j to i, and C_{ij} is the correlation between neurons i and j.

1.4.9 Monte Carlo Simulations

If the model described to this point captures some of the principles underlying the developmental process, other models which also capture those principles should provide an independent check of the results. Thus, to better understand the basis for the activity-dependent developmental model, we also run so-called "Monte Carlo" simulations. These simulations develop under qualitatively different conditions than the previous ones and are less explicitly linked to the biological data, but nonetheless approximate the same endpoints (as will become evident in the discussion of energy functions below).

Neurons in these simulations are initially assigned random simple cell receptive fields (see below) and intracortical weights. Thus, there is no LGN input to the model. At each time step, the value of one parameter of the model is simply changed, and the new "energy" of the network is calculated. If the energy increases, the change is automatically accepted. If the energy decreases, the change is accepted with a probability that depends on the difference between the new and old energies, E_{new} and E_{old}, divided by a slowly-increasing normalizing factor T referred to as the temperature ($P(acceptance) = exp((E_{new} - E_{old})/T)$) (Bishop, 1995). Thus, changes in the direction of lower energy in this method – a simulated annealing version of the Metropolis algorithm – become more difficult as time progresses. As in the developmental model, weights are subject to both pre- and post-synaptic constraints applied in alternating order, as well as to saturation limits.

In practice, the geniculocortical weights are assumed to give rise to oriented receptive fields because of the correlation structures used as inputs to the developmental model, and thus these RFs are parametrized by 2 variables: an orientation θ, where $0 \leq \theta < 180$, and a phase ϕ, where $0 \leq \phi < 360$. At each time step, either a new receptive field parameter is chosen, or an intracortical weight is changed, and we calculate the following energy:

$$E = \sum_{i,j} I_{ij} \sum_{\alpha} S_{i\alpha} S_{j\alpha} \tag{1.17}$$

In words, this function simply sums the ON/OFF overlap of every possible pair of receptive fields, weighted by the strength of the connection between them.

The origin of this energy function is intuitively, if not mathematically, straightforward; it is inspired by the energy that can be derived from the models of Miller (Miller et al., 1989; Miller, 1994), and thus provides the link between the more detailed synaptic models previously described both here and by Miller, and the Monte Carlo approach. In Miller's models, the energy function, disregarding the arbor function, can be formulated as

$$E = \frac{1}{2} \sum_{i,j,\alpha,\beta} S_{i\alpha} I_{ij} C_{\alpha\beta} S_{j\beta} \tag{1.18}$$

where $C_{\alpha\beta}$ is the correlation between different LGN positions α and β and, as defined for the Monte Carlo simulations, $S_{i\alpha}$ is the geniculocortical weight between cortical cell i and LGN cell α, and I_{ij} represents the (fixed) intracortical weight between cortical cells i and j. This energy function arises from the learning rules in Miller's models – specifically, if I and C are symmetric, the learning rule is the derivative of the energy with respect to each geniculocortical weight (*i.e.* the gradient vector):

$$\frac{dS_{i\alpha}}{dt} = \frac{dE}{dS_{i\alpha}} = \sum_{j,\beta} I_{ij} C_{\alpha\beta} S_{j\beta} \tag{1.19}$$

As mentioned above, because we specify the orientation structure of the receptive fields in the Monte Carlo formulation, we already implicitly include the correlation structure in the RF variable S in the Monte Carlo equations. More fundamentally, the energy function for the new model is not formally identical to the one of Miller because integrating the learning rule to give this energy function would require that the I matrix be symmetric. In the models with a fixed cortical interaction, the matrix I can indeed be symmetric; but when intracortical connections can vary, the I matrix is almost certainly not symmetric. Thus, the energy function in equation 1.17 is only an approximation.

Multi-column Monte Carlo simulations induce some complications. Because the parameter space to explore increases dramatically with the number of cells, we only use one-dimensional cortices, and develop either the weights or the receptive fields, but not both simultaneously. Other changes include adding a spatial offset between receptive fields of adjacent columns, and including a simple between-column arbor function (1 for the center column, 0.5 for the two adjacent columns). Lastly, with Monte Carlo techniques, deciding on an appropriate stopping criterion for each simulation can be difficult. We run all simulations for a defined number of timesteps, rather than searching for a robust stopping criterion, but this issue is not critical to our results.

1.5 Results

We first examine the performance of the model in single-column simulations, and then move on to multiple columns.

1.5.1 Single-Column Simulations

The performance of the model for a single column of 10 cells – 6 excitatory, 4 inhibitory – is illustrated in figure 9. Initially, as shown on the left of the figure, the weights in the model are randomly distributed. Geniculocortical weights show no evidence of ON and OFF subregion segregation, and intracortical weights are distributed all-to-all (with the exception of self-connections, which are not allowed; see Methods, page 21). The model is then allowed to develop with a combination of learning rules – Hebbian covariance for the excitatory weights, a Komatsu-inspired (Komatsu and Iwakiri, 1993; Komatsu, 1994; Komatsu, 1996) rule for the inhibitory weights – and subtractive constraints (see Methods). Over time, both simple cells and appropriate intracortical connectivity develop, as shown on the right of the figure. For cell 1, both excitatory and inhibitory connections are phase appropriate, with the exception of some small weights received by cell 1 from inhibitory cells 7 and 8. Moreover, orientations are identical, while spatial phases are not; appropriate local invariance has also appeared. Here, then, is the first evidence that the appropriate LGN correlations can give rise to both simple cells, appropriate intracortical circuitry, and appropriate columnar organization.

Intuitively, this result follows from the operation of Hebbian learning rules, which predict that the network seek out the most correlated set of geniculocortical and intracortical weights. For excitatory cells, this set should include strong positive intracortical connections between neurons of the same afferent orientation and spatial phase. For inhibitory cells, this set should include strong negative intracortical connections from inhibitory cells, to other cells of the same orientation but *opposite* spatial phase. Note that this second intuition represents a departure from previous models. When only orientation is considered, independent of spatial phase, the most anti-correlated RF is one of orthogonal orientation, not one of the same orientation but opposite phase. This issue of anti-correlation and spatial phase will recur in the multi-column simulations, for which it becomes significant in relation to the development of orientation maps.

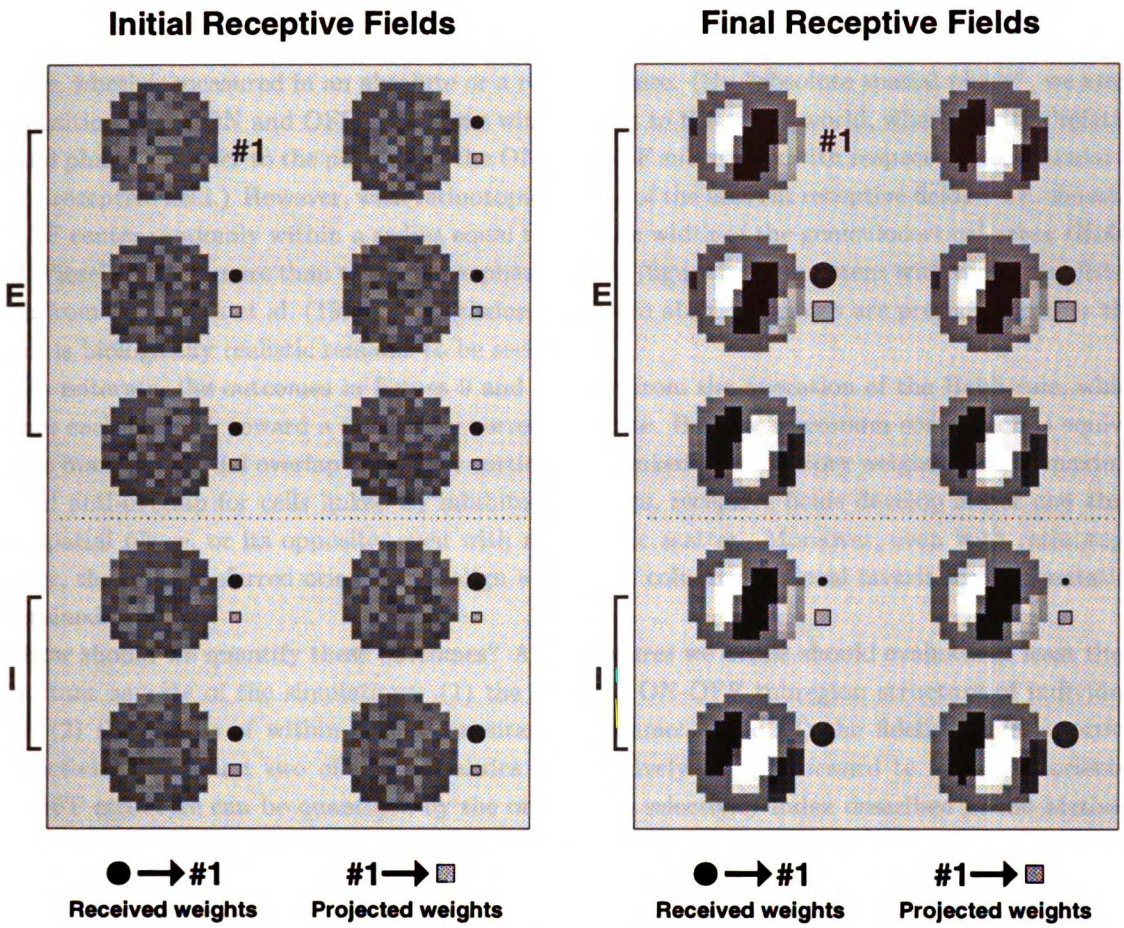


Figure 9. Geniculocortical weights for each of the ten cells in the column, at two different times, are shown by the circular afferent receptive fields. Strongly ON and strongly OFF regions of the RFs are colored white and black, respectively; regions for which ON and OFF weights are roughly equal are shown in gray. Note that this scale confounds positions that have both ON and OFF weights where both are roughly equal, but small, or roughly equal, but large. In the initial RFs, ON and OFF weights are roughly equal and of moderate size; however, in the final RFs the segregation of ON and OFF weights is quite strong, and gray regions are areas with very small weights. Colors cover the range between maximum and minimum afferent weights across all cells, where these extrema are recomputed each time afferent weights are printed. ***Initial condition*** of the column (left). The salt-and-pepper appearance of these RFs is due to the fact that the weights have not yet segregated, so ON and OFF connections are both present at most retinotopic positions. Because displaying all the intracortical weights simultaneously is difficult, connections for a representative cell (#1, an excitatory neuron) are shown next to the afferent RFs. Cell #1 receives an intracortical weight from every cell whose RF has an adjacent circle; cell #1 projects a weight to every cell whose RF has an adjacent square. The sizes of the weights are represented by the sizes of the respective symbols. Note that connections are initially all-to-all, and that the sizes of the weights are random, and roughly equal within a weight class ($E \rightarrow E$, $I \rightarrow E$, etc.). For display purposes, the weights are normalized by the respective maximum possible weight for that weight class, so the relative magnitudes of weights *between* classes are not accurately portrayed. ***Final condition*** of the column (right). The 10 cells in the column have now all developed simple-cell geniculocortical structure, and the intracortical weights for cell 1 are mostly, but not completely, phase-specific.

One aspect of the column that becomes immediately apparent is the positioning of the ON and OFF subregions of the cells. The model reliably gives rise to two (and only two) distinct spatial phases, whether measured in an absolute or a relative sense. (By "absolute spatial phase", we mean the position of the ON and OFF subregions with respect to the visual world, whereas with "relative spatial phase" we refer to the position of the ON and OFF subregions with respect to the boundaries of the receptive field.) However, with retinotopic scatter of the afferent receptive fields – *i.e.* choosing the RF center randomly within a radius equal to half the width of the geniculocortical arbor (Hubel and Wiesel, 1962) – more than two relative phases result (figure 10), consistent with recent published work from DeAngelis et al. (1999). Nonetheless, only two absolute phases are present; whether this result is biologically realistic remains to be seen.

To reiterate, the outcomes in figures 9 and 10 arise from the operation of the Hebb rule, which pushes each network toward a maximally correlated state. Because maximum correlation is equivalent to maximal spatial overlap of RFs for cortical cells linked by excitatory weights, and to maximal spatial anti-overlap for cells linked by inhibitory weights, receptive fields develop either one absolute spatial phase, or its opposite, even with retinotopic scatter. Moreover, even with retinotopic scatter, the cells' preferred orientations align within the column; the local invariance of orientation is retained.

How should we quantify these outcomes? Any measures we define should evaluate at least three important aspects of the simulations: (1) the oriented ON-OFF subregion structure of individual RFs; (2) the degree of within-column orientation alignment; and (3) the fidelity of intracortical connectivity. The first two of these qualities are relatively straightforward to measure: oriented ON-OFF structure can be quantified by the orientation selectivity index described in the Methods (equation 1.15), and within-column orientation alignment can be measured by simply taking the standard deviation of the preferred orientations of all cells within the column. The third concept is slightly more involved. We want to determine that connectivity is appropriate – *i.e.* that a given weight, if excitatory, links cells of the same orientation and phase, and, if inhibitory, connects cells of the same orientation and opposite spatial phase. To capture this behavior, we calculate the correlation between every pair of connected RFs, and weight it by the strength of the connection between them (Methods, 1.16). We then sum these quantities for each of the four types of weights – $E \rightarrow E$, $E \rightarrow I$, $I \rightarrow E$, and $I \rightarrow I$ – as well as for all the weights together (the "total" weights), and normalize by the appropriate weight sum to produce five quantities (EE corr, EI corr, IE corr, II corr, and total corr) that range from -1 to 1. For all cases, the appropriate weights, either excitatory or inhibitiory, give rise to values near one[4]; the exactly "wrong" weights, of any type, give rise to values near -1; and values near 0 indicate a lack of correlation. Thus, columns that demonstrate phase-appropriate intracortical connectivity should show positive values for all five correlations.

For the columns shown in figures 9 and 10, these measures indicate both that they provide a good marker of development *within* a simulation, and that they accord well with the perceived outcomes. As shown in table 1, all seven measures show large changes after the column develops: orientation selectivity increases, the standard deviation of the orientation tuning within the column decreases, and all the correlation measures grow stronger, consistent with the visual impression.

[4]Anti-correlated RFs have a correlation of -1. When this correlation is multiplied by an inhibitory weight, which is negative, the result is a positive quantity.

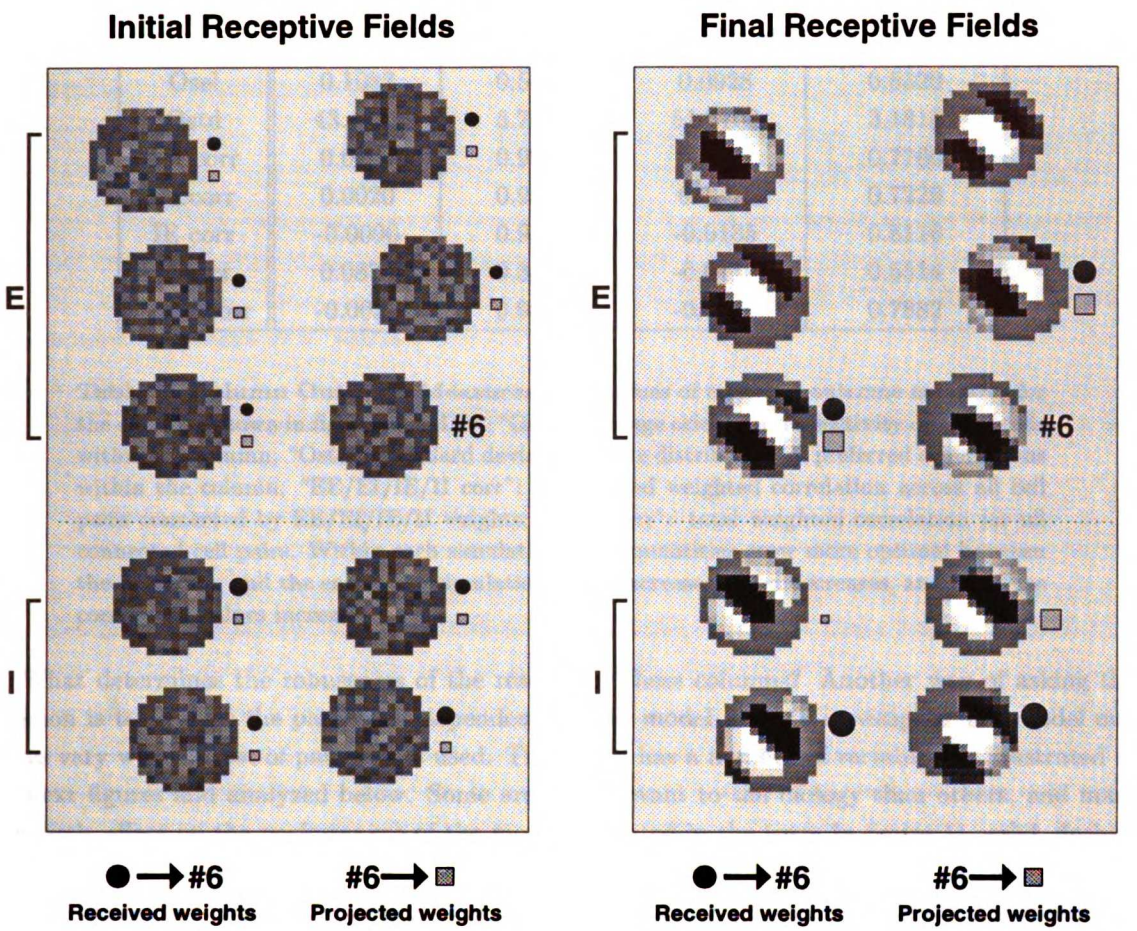


Figure 10. The conventions in this figure are the same as for the preceding one, with the exception that the position of each receptive field is shifted from its normal position by a random amount (representing retinotopic scatter of the RF centers (Hubel and Wiesel, 1962)). Connections respect absolute spatial phase (table 1).

While all measures improve, two findings should be noted: (1) the II correlations are not as strong as those of the other weight types, and (2) the correlations for the column shown in figure 10 are somewhat less than those for the column shown in figure 9. The former finding results from the fact that individual *excitatory* cells (with mature RFs) do not respond to all orientations; thus, there are orientations to which all *inhibitory* cells do respond (because all the inhibitory cells must be active to prevent excitatory cells of both spatial phases from firing). Consequently, the $I \rightarrow I$ weights learn on patterns in which cells of opposite phases both respond, and their connectivity is somewhat less specific. The latter finding is a simple consequence of retinotopic scatter: when RFs do not completely overlap, the maximal correlation between those RFs is lessened. Nonetheless, the 7 measures appear to provide reliable indicators of the quality of column outcomes.

	Fig. 9, Left	Fig. 9, Right	Fig. 10, Left	Fig. 10, Right
Osel	0.1087	0.5304	0.0928	0.5329
Ostd	43.4841	3.7455	51.5703	3.1815
EE corr	0.0104	0.9946	0.0496	0.7766
EI corr	0.0020	0.9344	0.0204	0.7229
IE corr	-0.0090	0.9330	-0.0185	0.8116
II corr	0.0870	0.3341	-0.0110	0.5514
Total corr	-0.0027	0.9089	-0.0110	0.7882

Table 1: Column Outcome Measures. The values of each of 7 outcome measures for the columns shown in figures 9 and 10. "Osel": average orientation selectivity of all 10 cells within the column; "Ostd": standard deviation of the distribution of preferred orientations within the column; "EE/EI/IE/II corr": normalized weighted correlation across all cell pairs connected by EE/EI/IE/II weights; "total corr": total weighted correlation for all connected cell pairs. Within each simulation, all 7 quantities grow more optimal between the beginning and the end of the simulation: Osel increases, Ostd decreases, and all of the correlation values increase.

What determines the robustness of the results for these columns? Another way of asking this question is to analyze the parameter-dependence of the model – i.e. to investigate how model outcomes vary with the set of parameters used. The model has a number of variables, as illustrated by the next figures and analyzed below. Some are less relevant to the biology than others, and many have little effect on the performance of the model, as noted in the text. In figure 11, solid, dashed, or dashed-dotted lines indicate that isolated variation of the parameter listed to the left produces results that do not differ significantly from outcomes using the default value of the parameter (indicated by an asterisk; as noted above, figures 9 and 10 were run with default parameters). Even where results were not significantly different, some dependence on the parameter can be seen in a statistically significant upward or downward trend, represented by the dashed and dashed-dotted lines, respectively. We again use the 7 outcome measures described above.

1. *EESUM*: This variable sets the sum of the $E \rightarrow E$ weights, which acts as a kind of positive feedback in the model. EESUM has little effect on the outcome of the simulation.
2. *EISUM*: This variable sets the sum of the $E \rightarrow I$ weights. Increasing EISUM tends to improve the specificity of inhibitory weights, but to disrupt the $E \rightarrow E$ coupling at the highest values. For the most part, however, it has little effect.
3. *IESUM*: The largest weight sum, this parameter can most severely affect the segregation of $I \rightarrow I$ weights when big. Presumably a larger IESUM is able to shut down the excitatory cells during development, and the resulting "null" activity patterns (in which all inhibitory neurons fire) do not permit robust segregation of the inhibitory weights.
4. *IISUM*: As $IISUM$ increases, the phase-specificity of the $I \rightarrow I$ weights improves. Larger relative imbalances in the weights between inhibitory cells allow one spatial phase of inhibitory cell to become more dominant for "null" inputs, and thereby permit the $I \rightarrow I$ weights to learn on more "segregated" activity patterns. While increasing this weight sum may improve $I \rightarrow I$

specificity, however, the same, more segregated activity responses are less effective at shutting down both phase of excitatory cell, and orientation tuning can broaden (see discussion of contrast invariance below).

5. *GSUM*: The sum of the feedforward weights has no effect. It gives a scale to the incoming LGN activities, and thus to the initial cortical activities, but has no further consequence.

6. *batch-size*: This parameter sets the number of LGN images over which weight changes will be accumulated before those weight changes are applied – *i.e.* it determines how much averaging of its inputs the cortex does. As shown by the graph, the model is insensitive to the exact number of images in a "batch."

7. *h*: Taken from Goodhill (1993), in this model h ultimately determines how anti-correlated LGN inputs are between ON and OFF layers. Unless h is reduced to zero, the cortex can still robustly find the second-order structure in its inputs.

8. *init-igain*: For all simulations, the strength of the inhibitory connections is ramped – *i.e.* inhibitory strength increases over time, consistent with experimental results (Luhmann and Prince, 1991). The ramp is linear in time (see Assumptions, in Methods). The value of init-igain – the initial value of the inhibitory strength before the ramp begins, where 0 is no inhibition at all, and 1 is equivalent to full strength – does not matter.

9. *lambda*: A covariance rule of the form $(z_y - \bar{z}_y)(z_x - \bar{z}_x)$, where z_y and z_x are post- and pre-synaptic activity, and $\bar{z}_y$ and $\bar{z}_x$ are their recent time-averages, requires some time scale over which the averages are determined. Roughly, the number of LGN activity patterns over which the post- and pre-synaptic activities are averaged goes as $1/lambda$. The plot shows that the exact value of *lambda* does not matter.

10. *max-count*: When activity patterns are presented to the cortex, some number of intra-cortical activity iterations is necessary before the cortex converges to a stable final state. Arbitrarily setting a value for this number of iterations has very little effect.

11. *neg-ramp*: As discussed for the parameter init-igain, the strength of inhibition in the model increases across development, consistent with experimental data from rat cortical slice (Luhmann and Prince, 1991). The number of time steps (defined as weight updates) over which the strength of the inhibition increases to the final values described by IESUM and IISUM is denoted by *neg-ramp*, and has a default value of 6000 batches. (For comparison, the total length of most simulations is 15100 batches.) Whether this ramp is necessary to generate the model circuit depends on the form of the learning rule used. Intracortical excitatory synapses learn under a Hebbian covariance rule: $\delta w = (z_y - \bar{z}_y)(z_x - \bar{z}_x)$, where δw is set to zero if both terms in parentheses are negative (see Methods). We consider two choices for the measure of activity in this rule: the raw activity a (Eq. 1.5), which can be positive or negative and is analogous to a voltage deviation from rest; or the spike rate $f^{E/I}(a)$ (Eq. 1.6), which is proportional to a for $a \geq 0$ and is 0 otherwise. In the default case, the presynaptic activity measure is spiking – since, presumably, spiking is essential for generating significant transmitter release at the neuron's terminals – and the postsynaptic measure is voltage (Bliss and Collingridge, 1993). We consider the following alternatives, which also hold for the Komatsu rule and the *inhibitory* covariance rule when they are used (Eq. 1.9, and next paragraph): "SS", for which all

excitatory learning rules (geniculocortical and intracortical) learn on post- and pre-synaptic spiking; "SV", for which all intracortical excitatory weights learn on post-synaptic spiking and pre-synaptic voltage; and "VV", for which all intracortical excitatory weights learn on post- and pre-synaptic voltage. For all simulations, the presynaptic component of the geniculocortical learning rule depends on LGN spiking; the postsynaptic component follows that for the intracortical learning rule.

For inhibitory synapses, either a rule inspired by the work of Komatsu (see Methods) or an inhibitory covariance rule is used. Unlike the other terms in the rule, one term in the inhibitory LTP portion of the Komatsu rule (see Methods) depends primarily on the amount of inhibition received by a postsynaptic cell, not on either the postsynaptic membrane voltage or the firing rate. Thus, as shown in the figure, under this learning rule development is normal whether the post-synaptic terms in the excitatory covariance rules are converted into spike rates or not, and no matter what the actual value of *neg-ramp*. However, if the inhibitory covariance rule is used instead (denoted by "inhebb = 0"), simulations in which inhibitory learning depends on post-synaptic spiking (SS, SV) fail for low *neg-ramp* values. The reason for this failure is straightforward. Inhibition is still strong, even as it ramps. Initially, because the afferent receptive fields are not yet organized, the inhibition shuts off the excitatory cells. When the Komatsu rule is used, even with an excitatory rule that depends on post-synaptic spiking the $I \rightarrow E$ weights can learn on the differences in inhibition received by each (silent) excitatory cell, and drive segregation of the network. When an inhibitory covariance rule is used, however, and the ramp increases quickly, the zero activity of each excitatory cell renders them all equivalent, and weight segregation cannot occur.

12. *prob*: This parameter describes the probability that a post-synaptic cell receives a connection from a given pre-synaptic cell. *prob* has strong effects on model outcomes, as connectivity that is too sparse (low *prob*) can essentially isolate cells, or groups of cells, within the network. Larger columns (e.g. of 100 cells) can, of course, tolerate lower *prob* values (data not shown).

We also examined the robustness of the model in response to changes in two other factors: retinotopic scatter of the geniculocortical RFs (the "scatter" parameter), and variations in the relative learning rates of geniculocortical and intracortical weights (the "wt-factor" and "wt-factx" parameters, respectively).

1. *scatter*: When receptive fields are permitted to have retinotopic scatter, the basic results of the simulation are unchanged (figure 12). The correlation values drop (as compared to the no-scatter case) because the amount of overlap between RFs decreases, but orientation selectivity and orientation alignment remain strong.
2. *wt-factor* and *wt-factx*: These two parameters control the learning rates of the afferent and intracortical weights, respectively. As shown by the figure, these variables do not affect development of well-aligned columns unless they differ widely – by several orders of magnitude. Illustrated are two cases in which the simulations were run for either 21100 batches (top) or 30000 batches (bottom). Some changes are evident as the simulations run longer, especially for cases in which both learning rates are quite low. At the longer time, it appears that columns with a high intracortical learning rate but a low geniculocortical rate can nonetheless align

in orientation, but that the opposite values – low intracortical learning rate, but high geniculocortical rate – cannot. In the latter case, the geniculocortical RFs apparently grow strong before the slowly-changing intracortical weights can align them.

What is the biological meaning of this parameter exploration? For the most part, the model develops the appropriate architecture independent of many of the parameters. This independence suggests that the circuitry will arise without perfect anticorrelation between ON and OFF LGN inputs, without any particular local initial connectivity between cells, and somewhat irrespective of the relative strengths of excitation and inhibition. Some of these factors will affect other properties (see Contrast Invariance), but the main result is that the model can tolerate "noise" in the parameters, as would be expected in a biological system. Additionally, the necessity of the experimentally-determined ramp of inhibition depends on the form of the learning rule. That any network employing the Komatsu rule remains unaffected by the inhibitory ramp may indicate that an explanation for the inhibitory ramp lies elsewhere; alternatively, it may indicate that further clarification of the Komatsu rule may alter its behavior in certain cases (see Learning Rules, in Methods).

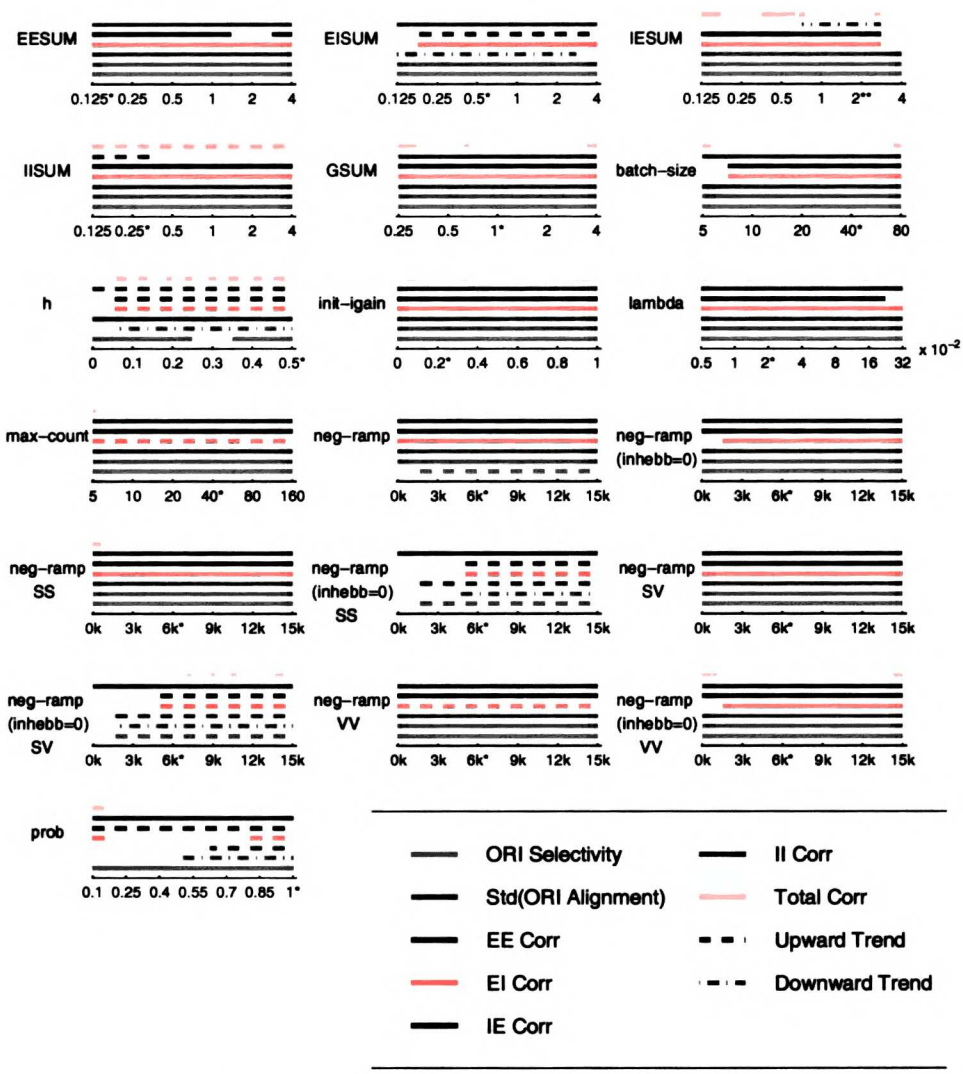


Figure 11. Parameter exploration of the single-column model for the variables noted to the left of each figure. The default value of the parameter on the x-axis is highlighted by an asterisk. The different colored lines correspond to the different outcome measures averaged over 5 simulations for each parameter value (except for EISUM and GSUM, whose outcome measures are averaged over 10 simulations); this number of simulations represents an attempt to balance statistical power with computational expense. Solid, dashed, or dashed-dotted colored lines indicate parameter values for which the result did not differ significantly from the default. An absent line, or absent portion of a line, identifies a result that did differ significantly from default ($p < 0.05$, one-way ANOVA followed by a Tukey post-hoc procedure). Isolated gaps in the middle of lines (eg. as seen in the EESUM figure) are likely due to the relatively small number of simulations contributing to each point. Dashed lines indicate significant (i.e. $p < 0.05$) upward trends in the data, while dashed-dotted lines indicate significant downward trends, as determined by a Spearman rank-order correlation coefficient. Each parameter is varied *in isolation*; when one parameter varies, the others are set to their default values, unless otherwise noted in the figure.

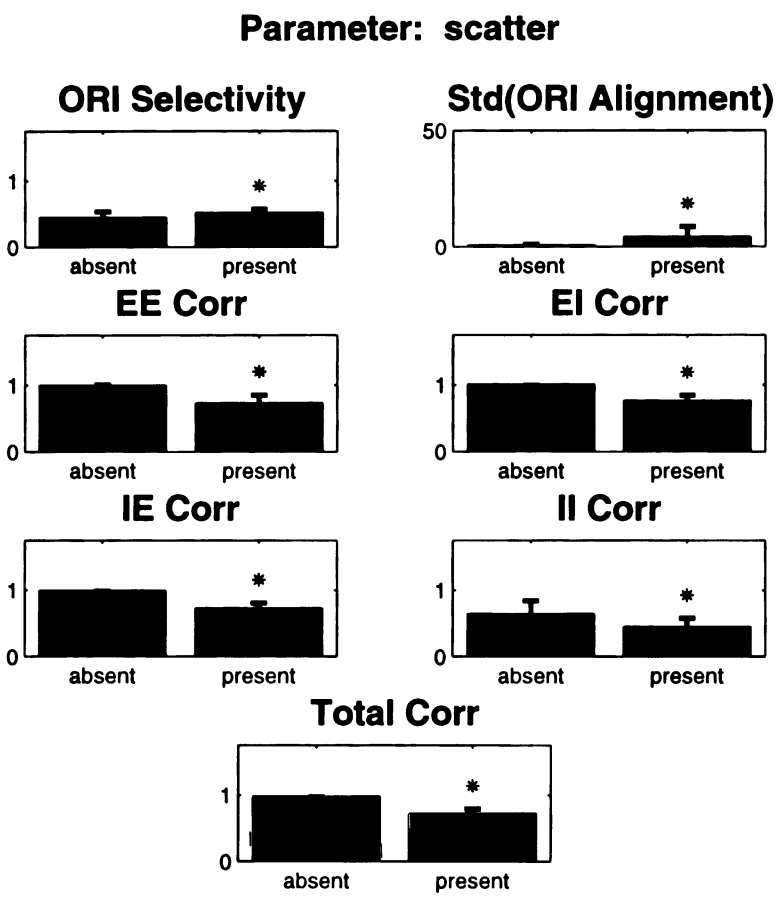


Figure 12. Effects of retinotopic scatter on column outcomes. The seven variables examined here are the same as those in the preceding figure. Asterisks indicate significant differences ($p < 0.05$) between columns, as determined by an analysis of variance over the 25 simulations for each condition (*i.e.* scatter "absent" or "present").

In sum, the simulations can robustly develop the appropriate columnar structure, independently of individual variations in a large number of parameters. The next question is whether they also display appropriate mature properties – *i.e.* contrast-invariant orientation tuning. In previous work (Troyer et al., 1998), our lab has shown that correlation-based connectivity, in conjunction with strong feedforward inhibition, can lead to contrast-invariant responses. The above models develop the appropriate connectivity; how should the feedforward inhibitory strength be determined?

To answer this question, we studied a single-column simulation in which all the intracortical weight sums (EESUM, EISUM, IESUM, IISUM) were set to 1. (The alternative – running a large number of simulations in this 4-dimensional parameter space – is both time- and CPU-intensive, so we opted for this more approximate, but more efficient, tack.) For weight sums of 1, both aligned simple cell RFs and appropriate intracortical connections resulted. For the mature network, we then artificially set each of the weight sums to a number of different values (thereby altering both the absolute and relative strength of excitation versus inhibition), and tested the network with gratings at the preferred and orthogonal orientations, to determine appropriate values for the weight sums. Our requirements were (1) minimal responses at the null orientation, (2) orientation tuning width as small as possible, and (3) strength of inhibition as small as possible. Of course, this approach neglects the nonlinearities inherent in *developing* a column with those weight sums, including the

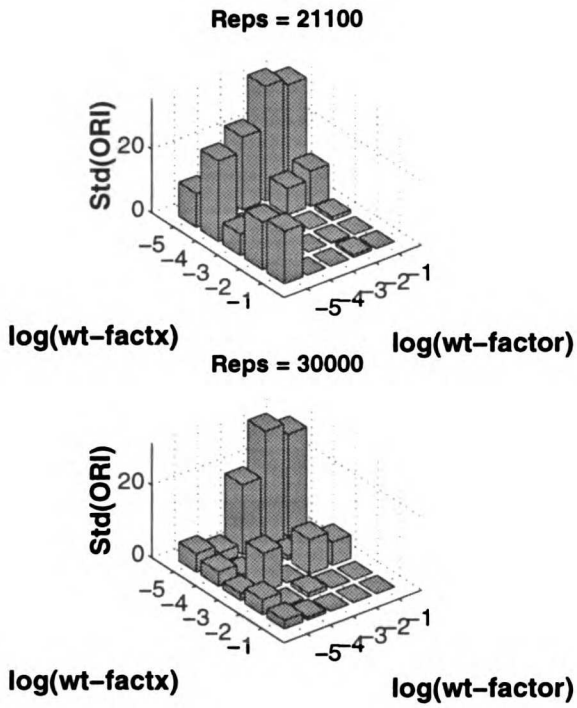


Figure 13. Effects of varying geniculocortical (wt-factor) and intracortical (wt-factx) learning rates on the development of well-aligned columns. Top: 21100 batches. Bottom: 30000 batches. Each bar represents the average of the standard deviation of the preferred orientations for the 10 cells within the column over 5 simulations; error bars show standard deviations. **Top.** For small learning rates, either geniculocortical or intracortical, the orientation alignment of the column suffers (higher standard deviation of orientation tuning). Moreover, this misalignment appears larger as the mismatch between geniculocortical and intracortical rates increases. **Bottom.** Some of this variability can be explained by the duration of the simulation: if the geniculocortical learning rate is small, longer simulations show better alignment. However, if the intracortical learning rate is small while the geniculocortical rate is large, longer run times still give larger alignment errors.

fact that developing a column under vastly different weight sums may disrupt the connectivity, but figure 11 suggests that this problem is minimal. Additionally, requirements (2) and (3) can conflict; decreasing the strength of the inhibition can increase the width of tuning. We balance these criteria by striving to keep the standard deviation of the orientation tuning curve to approximately 20°-25°, in rough agreement with Ferster et al. (1996), as shown in figure 5 of Troyer et al. (1998).

This approach produces the following default values (as described in the Methods): EESUM = 0.125, EISUM = 0.5, IESUM = 2.25, IISUM = 0.25. Intuitively, these values make sense. IESUM should be large, in order to silence excitatory cells when the null orientation is shown. IISUM acts as a "gain control", in that it reduces inhibitory cell firing and thus trades off with IESUM: larger values of IISUM require larger values of IESUM, if the excitatory cells are to "see" the same level of inhibition. Thus, IISUM is relatively small. EISUM is not strongly constrained since, as shown by Troyer et al. (1998), these connections have little effect on contrast-invariant-tuning. Lastly, EESUM governs the amplification of cortical responses, and should be set to a level that allows robust, but not saturating, excitatory firing rates.

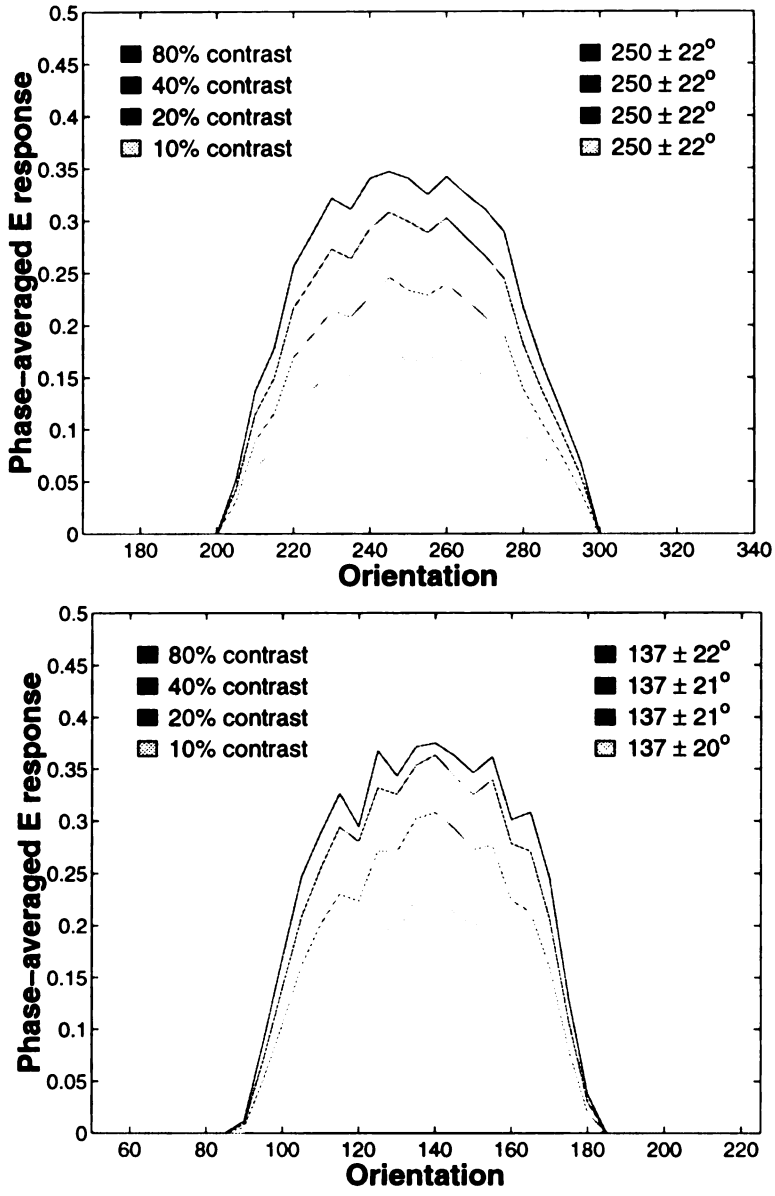


Figure 14. Two good examples of contrast-invariant orientation-tuning across 3 octaves of contrast. **Top.** Response of cell 1 from figure 9 to gratings shown at 10%, 20%, 40%, and 80% contrast (see Methods), as indicated by increasing gray-level. **Bottom.** Response of cell 6 from figure 10 to the same contrast levels. Numbers at top right in each case indicate mean and standard deviation of best Gaussian fit to each response distribution, showing that the tuning widths are relatively invariant with contrast.

Consistent with the robustness of contrast-invariant orientation tuning shown in Troyer et al. (1998, figure 13), these weight levels do indeed give rise to contrast-invariant responses, as shown in figure 14 for two excitatory cells taken from figures 9 and 10. As parametrized by the mean and standard deviation of the responses (*i.e.* by a Gaussian fit), the cells are essentially contrast-invariant. As in Troyer et al. (1998), inhibitory cells do not display contrast-invariant orientation tuning (not shown); instead, they show a tuned component to their orientation response riding on a large untuned response.

One prominent feature of the orientation tuning curves in figure 14 is their width, as measured by the standard deviation of the best-fitting Gaussian. In our search for appropriate weight sums, we tried to constrain this measure to approximately 20° while keeping inhibition as small as possible. However, these cells (and others; see below) show somewhat broader tuning. One simple way to decrease this tuning width would be to increase the magnitude of IESUM; increasing this value should still allow for normal columnar development (figure 11), but the stronger inhibition should reduce the range of orientations to which the cells respond. We strove to keep inhibition from becoming too powerful since the experimental data do not suggest vast differences in excitatory-inhibitory strength (see Discussion), but this bias may not be justified, so increasing inhibition still further may be reasonable. Alternatively, the broader tuning may result from the simple activity model we use; correlation-based connectivity within a model with more realistic LGN inputs and a more realistic implementation of synaptic integration, for instance, can give rise to tighter orientation tuning (Troyer et al., 1998) than does the model described here. Lastly, the intracortical connectivity itself may be responsible. The connections explored here may not be as specific as those simply set (rather than developed) in the mature modeling, and the $I \rightarrow I$ connections, which reduce the firing of inhibitory cells at off orientations, may very well reduce inhibition enough to broaden orientation tuning. The mature models (Troyer et al., 1998, and the second chapter of this thesis) have not yet explored $I \rightarrow I$ connections.

Summary orientation tuning data for 384 excitatory cells are shown both in figure 15, for simulations in which retinotopic scatter is absent, and figure 16, for simulations in which retinotopic scatter is present. In the left columns of both figures, the histograms show the tuning widths of all 384 cells at 4 different levels of contrast; the relatively constant shape of the histogram indicates that orientation tuning is essentially contrast-invariant (though as in figure 14, many of these cells have relatively broad ($> 20°$) tuning). In the right columns this contrast invariance is examined more closely: the values for the standard deviations of orientation tuning at 20%, 40%, and 80% contrast are compared, for each cell, to that same cell's orientation tuning at 10% contrast. These ratios are about 1, confirming the impression derived from the left columns. However, a broadening of approximately 5% can be seen as the contrast increases. No such effect is seen in the experimental data; if anything, a slight narrowing of the orientation tuning may be seen experimentally with increasing contrast. We have not explored the cause(s) of this effect. As for the above discussion of the potential causes for the breadth of the orientation tuning, the values of the weight sums and the impoverished activity model we use to allow simulation over developmental times may affect tuning width as contrast changes.

One comment on these results concerns responses at the null orientation (*i.e.* the orientation perpendicular to the preferred). In the experimental work (Sclar and Freeman, 1982; Skottun et al.,

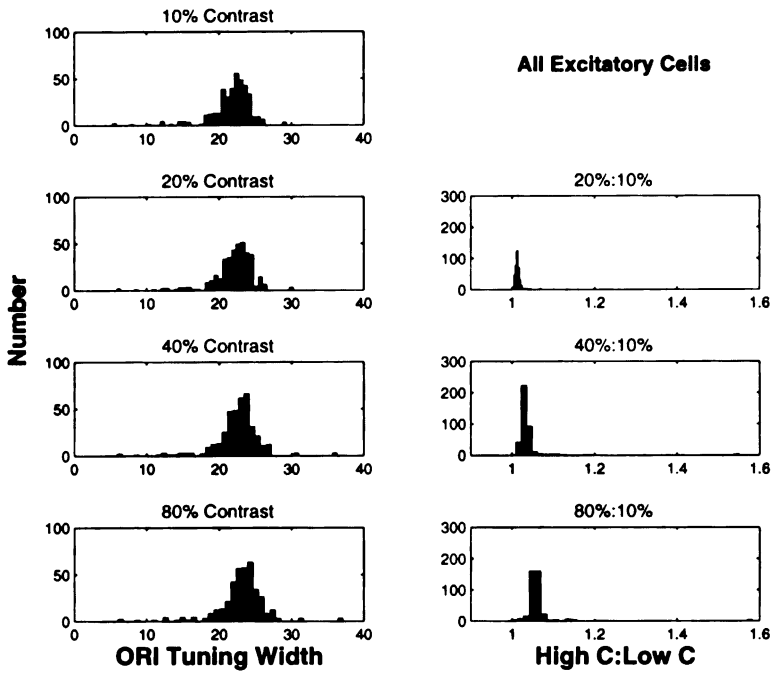


Figure 15. Left column: histograms of standard deviations of the central peaks of the orientation tuning curves for all excitatory cells in 64 different columns, shown at 10%, 20%, 40%, and 80% contrast. Right column: ratio of high contrast:low contrast standard deviations of orientation tuning for all excitatory cells. In this set of simulations, all without retinotopic scatter, the tuning of the central peaks is relatively tight, and the tuning width increases by 5% at the highest contrast. By "central peak", we refer to the response peak about the preferred orientation. Many of the cells in this simulation (48%) will give responses to the null orientation (see text).

1987), only the central peak of the orientation tuning curve was examined. Many of the cells represented in figures 15 and 16 (48% and 28%, respectively) produce a small second response peak around the null orientation which, if included in the calculation of the best Gaussian fit to the orientation tuning curve, would substantially widen the tuning. These null responses could arise from three main sources: (1) the form of the afferent receptive field; (2) the form of the intracortical connectivity; (3) the strength of intracortical inhibition. Replacing all of the receptive fields with 16×16 sinusoidal gratings of the same orientation and phase (positive/negative values giving strengths of ON/OFF connections respectively) does indeed remove the null responses. Replacing only the intracortical weights with the most phase-appropriate connections does not, and doubling the amount of intracortical $I \rightarrow E$ inhibition decreases the number of neurons with null responses, but does not remove those responses entirely. Thus, it appears that the main reason for a second response peak around the null orientation involves the shape of the RFs that develop here.

The most likely cause is the RF aspect ratio. As shown in figures 9 and 10, the RFs found here are close to circular. For cells with 3 subregions in particular (of which there are many), the grating of optimal spatial frequency, when presented at the "null", will often drive responses. The reason for these responses is as follows: suppose the geniculocortical RF is vertically oriented with a central ON subregion and adjacent, slightly smaller OFF subregions. The spatial frequency of a horizontal grating can be such that at the null orientation, a dark bar that overlaps both the ON and OFF subregions may contribute very little to the cell's activity, but the grating's adjacent light bars may overlap the ends of the RF's central ON subregion, thereby providing a net positive input. If this account were correct, increasing the spatial frequency of the grating should remove the null

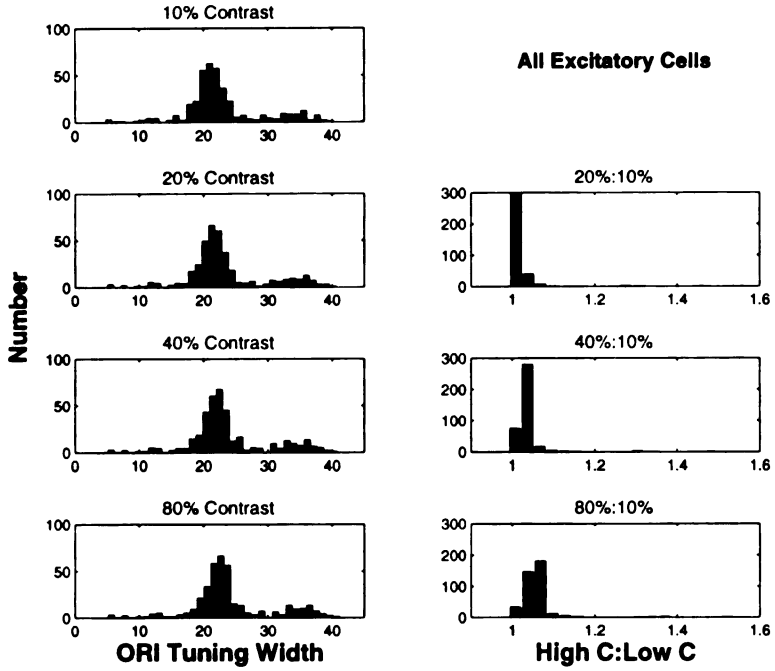


Figure 16. Same as the preceding figure, except that now all 64 columns are run with retinotopic scatter. Some broadening of the orientation tuning can be seen at all contrasts, but the responses remain contrast-invariant. 28% of the cells show a null orientation response.

responses, and it does.

The existence of null responses, of course, may not be a problem. Azouz et al. (1997, see Discussion) state that in one of their samples, up to 70% of cells show a null response, though they do not state whether this null response represents a second peak centered on the null orientation, or a very broad central peak centered on the preferred orientation that extends to the null. (We tend to see the former.) In any case, as shown none of the model responses are obviously inconsistent with the experimental data; if they are, a model of the afferent receptive field that includes a larger aspect ratio or finer detail in the subregion boundaries may improve the tuning (as could the stronger $I \rightarrow E$ inhibition previously mentioned).

1.5.2 Why the Model Works

The intuition for why the model works – *i.e.* for why the orientations are aligned, and the intracortical connections are phase-specific – can be understood using a different type of simulation. In these Monte Carlo simulations, the network is parametrized by three sets of variables: the orientation of each of the geniculocortical receptive fields, the spatial phase of each of the receptive fields, and the strength of the weight between each pair of cells. Thus, there is no development of the RFs, and no LGN; we assume that the correlations present in the LGN would give rise to simple cell structure. The new model otherwise preserves the architecture of the old model, but it will be trained in an entirely different way. Rather than "developing", in the sense of the previous simulations, variables in the new network are changed randomly throughout the simulation, subject to the condition that they increase some measure of the "energy" of the network (at least, increase that energy most of the time – see Methods). The question then becomes one of defining what this energy should be.

In the models of Miller (Miller et al., 1989; Miller, 1994), the learning rule, ignoring the arbor function, is $dS_{i\alpha}/dt = \sum_{j,\beta} I_{ij} C_{\alpha\beta} S_{j\beta}$, in which $S_{j\beta}$ is the geniculocortical weight connecting LGN cell β to cortical cell j, $C_{\alpha\beta}$ is the correlation structure present in the activity between LGN cells α and β, and I_{ij} is the intracortical connection between cortical cells i and j.) If we can define an energy measure for which this learning rule performs gradient-descent/ascent (Cohen and Grossberg, 1983), the model can be said to minimize/maximize that energy. Fortunately, for this equation, if the matrices I and C are symmetric, the type of energy function is relatively straightforward (see Methods): $E = \frac{1}{2} \sum_{i,j,\alpha,\beta} S_{i\alpha} I_{ij} C_{\alpha\beta} S_{j\beta}$. Intuitively, this equation states that the energy is proportional to the correlation between pairs of geniculocortical receptive fields, times the strength of the intracortical weight connecting them, where inhibitory weights are taken as negative numbers. In reality, intracortical connections are not symmetric in our developed columns, but given that the energy's intuitive meaning, we will guess that the network actually maximizes something like this energy. We also simplify by assuming that the correlation structure C is incorporated into the simple cell structure we take as given. Thus, our new energy equation $E = \sum_{i,j} I_{ij} \sum_{\alpha} S_{i\alpha} S_{j\alpha}$ represents the sum of the dot products of each pair of RFs, weighted by the strength of the connection between them. If this energy is maximized, what do columns looks like?

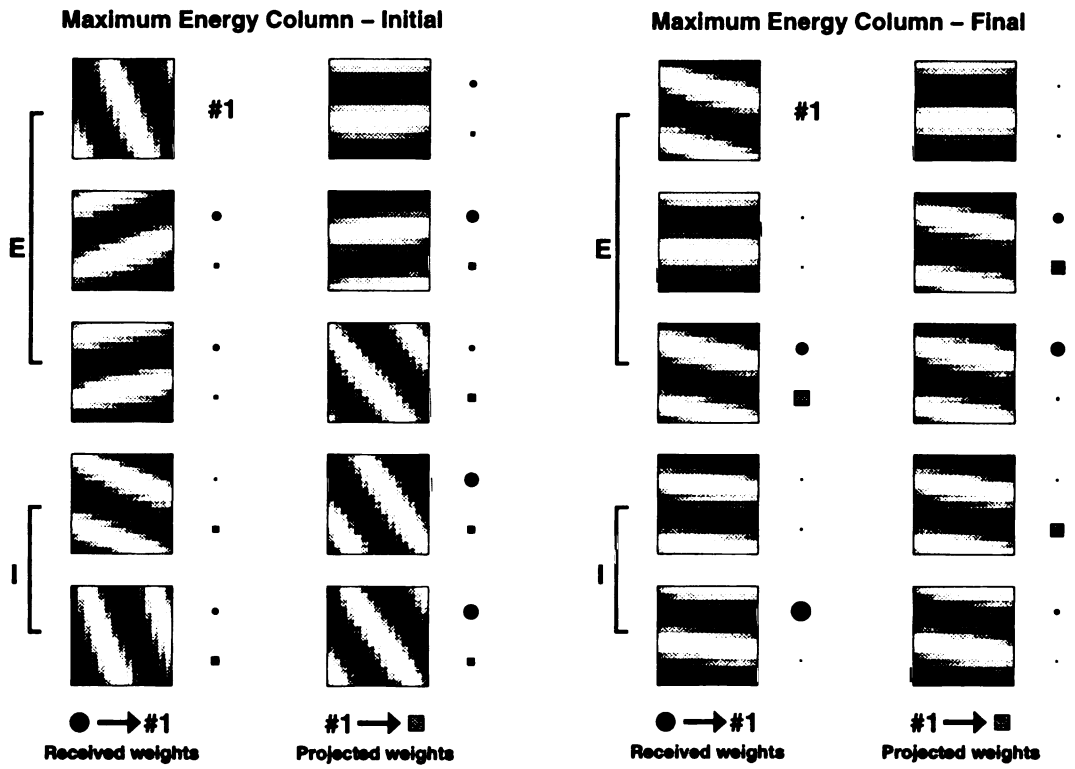


Figure 17. The highest energy network resulting from the Monte Carlo simulations described in the text. The initial state is shown on the left; note that weights are small and fairly evenly distributed. The final state is shown on the right; note that orientations are now aligned, and intracortical connections are now large and phase-specific.

On the left, figure 17 shows the initial state of a single-column simulation. The RFs are randomly oriented, and of random phases, and the randomly-selected intracortical weights are roughly equal and distributed among all cells. (As in previous simulations, self connections are not permitted.) The same simulation after 10000 time steps is shown on the right. The RF orientations are now well-aligned, the phases have segregated into two, and the connections (shown here for cell #1) are phase-specific. In fact, for the set of 100 simulations analyzed in the next figures, this simulation produced the maximum value of the energy.

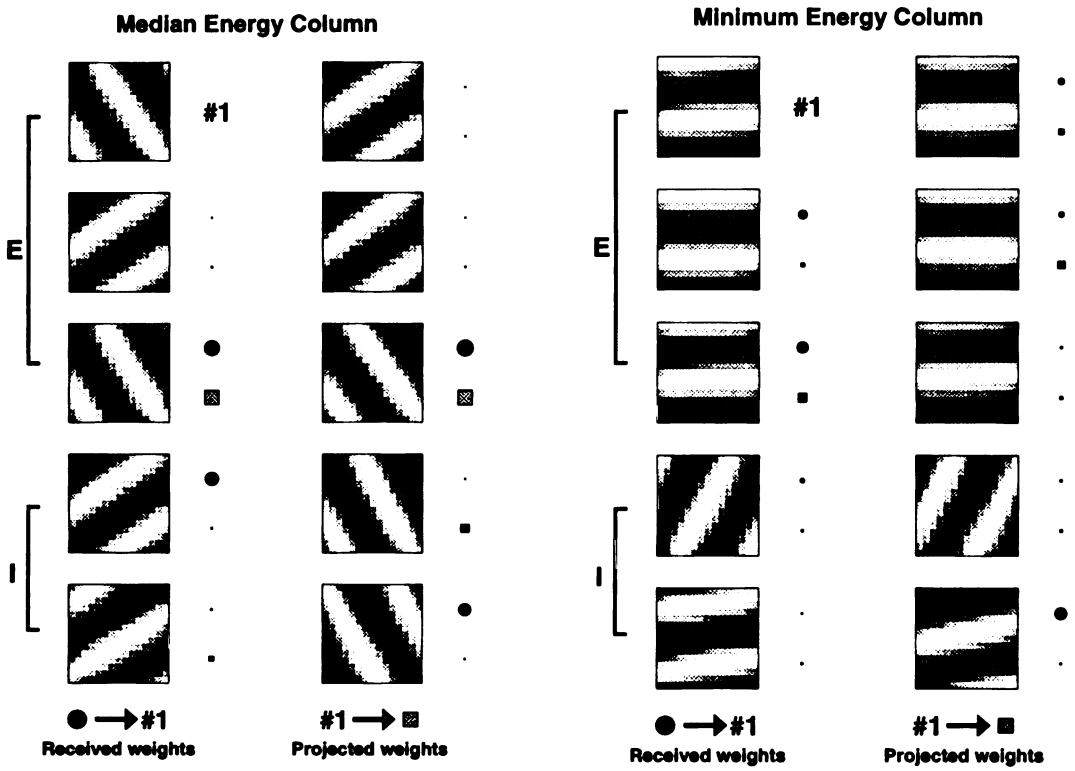


Figure 18. The median (left) and minimum (right) energy states resulting from the Monte Carlo simulations described in the text. The median energy outcome attempts to create independent, uncorrelated subsets of cells, but the single excitatory cell phase in each subset necessitates lower-energy $I \rightarrow E$ projections (not shown) from un-, rather than anti-, correlated inhibitory neurons. The minimum energy network, on the other hand, finds a state in which all excitatory cells share the same phase. Consequently, only a single antiphase inhibitory cell can make phase-appropriate excitatory connections, either projected or received.

Two other simulations from this batch of 100 are shown in figure 18: the median energy simulation, on the left, and the minimum energy simulation, on the right. The median energy shows a very different outcome; neurons segregate into one of two orientation subsets, within each of which all excitatory cells have identical phase. Because the sums of the weights are constrained, the way in which the phases have developed ensures that inappropriate connections will be made. Inhibitory cells of the same orientation and phase as one set of excitatory cells, for instance, will be constrained to synapse onto neurons of the opposite orientation – cells with which they are uncorrelated, rather

than anticorrelated. These connections will make little net contribution to the energy. In the minimum energy network on the right, the excitatory cells are all of the same phase and orientation, while two of the inhibitory cells are of a different orientation. As in the median energy case, the connections tend not to anticorrelate, but to *decorrelate* excitation and inhibition.

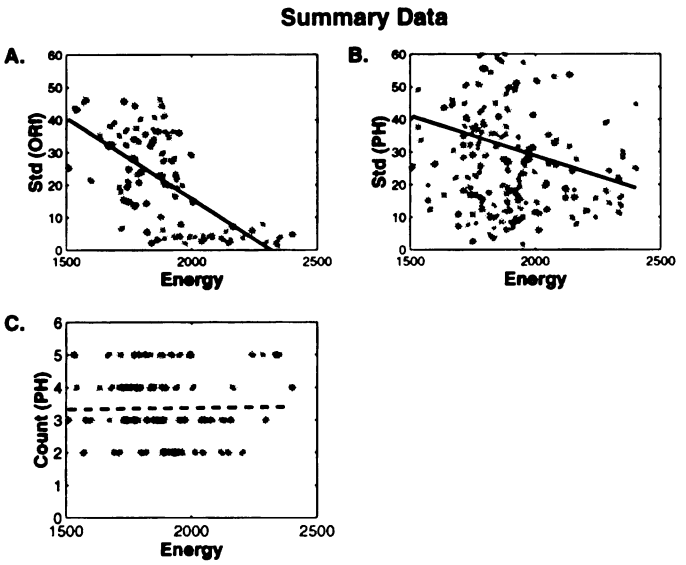


Figure 19. **A.** Standard deviation of orientation tuning within a column, versus energy. As the energy of the column increases, the standard deviation of the ORI tuning decreases – *i.e.* the columns become better aligned. This trend, illustrated by the solid black line, is highly significant ($p = 2.7 \times 10^{-12}$). **B.** Standard deviation of phase within a column, versus energy. As shown in the figure, there is a significant ($p = 0.0022$) negative correlation between the standard deviation of the spatial phase, and the energy of the column. **C.** Number of cells within each column that have the less common of the two spatial phases. The correlation between phase number and energy is not significant ($p = 0.2391$), and a one-way ANOVA shows that the mean energy between different phase groups is not significantly different ($p = 0.1507$).

The summary of all of these simulations can be seen in figure 19. To quantify the degree of alignment of orientation, we simply calculate the standard deviation in the preferred orientations for the ten neurons in each simulation. Thus, better-aligned columns will have smaller standard deviations, and worse-aligned columns will have larger. The results are shown in figure 3A: a higher-energy state correlates very significantly ($p = 2.7 \times 10^{-12}$) with more uniform orientation tuning. Similarly, to quantify the degree of phase segregation for each column, we assume that two spatial phases will form. We find the mean of the complete phase distribution, then calculate the standard deviation in phase separately for the subset of phases below the mean, and the subset of phases above the mean. Thus, each simulation contributes two values to the plot shown in figure 19B. Larger energy is significantly correlated with a reduced variation in spatial phase – *i.e.* with a strongly bimodal phase distribution. Lastly, we find the number of cells within each column that show the less common of the two spatial phases. No correlation exists between phase number and energy, and the mean energy for each phase count is not significantly different.

These Monte Carlo simulations suggest that the true developmental simulations described earlier

actually do seek out a maximally-correlated, or high-energy, state that is a consequence of Hebbian learning rules and constraints. Having rendered this intuition more formal will prove to be useful in understanding multi-column behavior.

Summary of Single-Column Results

The single-column simulations answer all of the questions posed at the outset: (1) the simultaneous development of appropriate geniculocortical and intracortical connections is not only possible, but robust; (2) the development of the columnar invariance of orientation, but the variation of spatial phase (both relative and absolute), is also quite robust; and (3) the resulting feedforward and intracortical circuitry is in accord with the results of Troyer et al. (1998), and capable of producing contrast-invariant, orientation-tuned responses. Across a range of parameters, under different learning rules, and under two very different styles of simulation, the same single-column cortical structure results. The next task, then, is to examine the issue of orientation maps, for which we expand the model to include multiple cortical columns.

1.5.3 Multi-Column Simulations

For multi-column simulations, three outstanding questions are the following:

1. Can columns form within a larger network?
2. If so, can orientation vary smoothly across a map?
3. If so, can orientation vary periodically across a map?

The first issue has yet to be addressed; the single-column simulations suggest that strong connectivity can organize columns, but this connectivity has not been investigated in the context of a larger network with retinotopic extent. The next two issues – smoothness/continuity and periodicity/diversity – frame the discussion in numerous models of orientation maps, as well as of ocular dominance (Erwin et al., 1995) and other maps. These questions are obviously related: a periodic map is most likely to be a smooth one, though a smooth map may not be periodic. *A priori*, how to establish smoothness and periodicity is not obvious from the framework developed for the single-column simulations. Periodicity would appear to require a long-range interaction to ensure that columns at a certain separation prefer different orientations. On the other hand, smoothness might only require the presence of strong weights between adjacent columns, in order to couple their developing orientation preferences.

To examine these issues, we first explore how the model works when we extend the intracortical arbor outside a single column. Initially, it seems plausible that establishing weak coupling between columns may permit them to develop similar orientations. These weaker connections could allow for smooth, linked change in orientation while the stronger vertical connectivity drives the development of iso-orientation, phase-varying columns. Thus, we begin with all-to-all connectivity between columns out to an arbor-distance of one, with weaker connections between columns than within columns (figure 7).

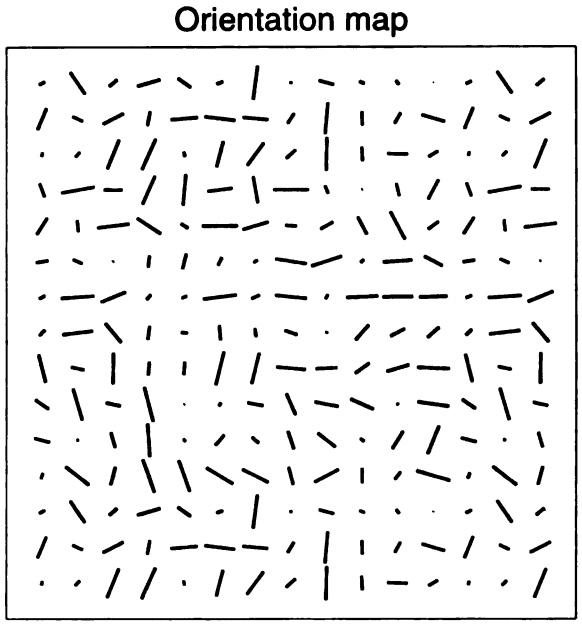


Figure 20. Final results of a simulation with 12 × 12 cortical columns, an arbor of 1, and an arbor scale factor of 0.25 (see text). Each column of 10 cells is represented by a single oriented line segment. The angle of the segment represents the mean preferred orientation of the column, and the length of the segment represents how well-aligned the cells in that column are: a longer line indicates better orientation alignment. Because there are no boundaries to the map – the columns on the right edge of the 12 × 12 map are adjacent to the columns on the left edge – columns from the left and top edges are reproduced on the bottom and the right, rendering the grid 15 × 15.

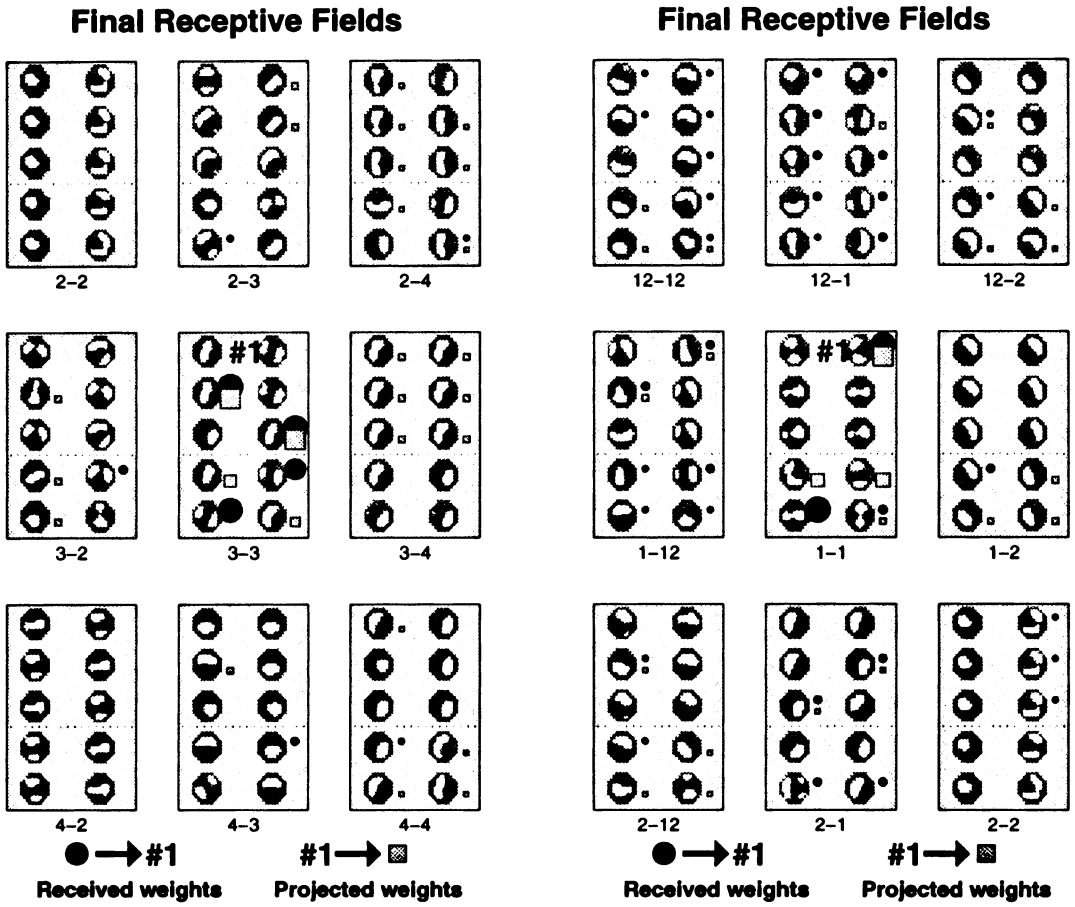


Figure 21. Two representative columns (central square in each 3x3 set) from the map, along with their eight neighboring columns. The numbers beneath each column show the row and column, respectively, where each column can be found in the previous figure, with 1-1 indicating the top left column. Connections of cell 1 in each representative column are shown. The central column on the left is well-oriented. The central column on the right is not well-oriented. Note that each column can be found twice in the preceding figure due to the display of the wrap-around boundary conditions. Column 3-3, for instance, can also be seen at the lower right corner of the 15×15 grid.

Figure 20 shows the final result of a simulation in which a 12 × 12 set of cortical columns is allowed to develop together. In this case, the arbor is set to 1, meaning that each column projects to the immediately adjacent columns, but no farther. The default strength of the between-column connections is arbitrarily set, as shown in figure 6, to be roughly 50% (actually 44.8%) as strong between columns as it is within columns. A new parameter – the arbor scale – then scales this value; thus, for this simulation, which has an arbor scale of 0.25, the strength of between-column connections is actually 44.8% × 0.25 = 11.2% of the within-column connection strength. Almost immediately apparent is the fact that this simulation is neither smooth nor periodic. The preferred orientation of each column, represented by the angle of the line segment at each position, does not appear to be ordered, and the degree of orientation alignment within a column, represented by the length of each line segment, is quite variable. Occasional runs of similar orientation are accompanied

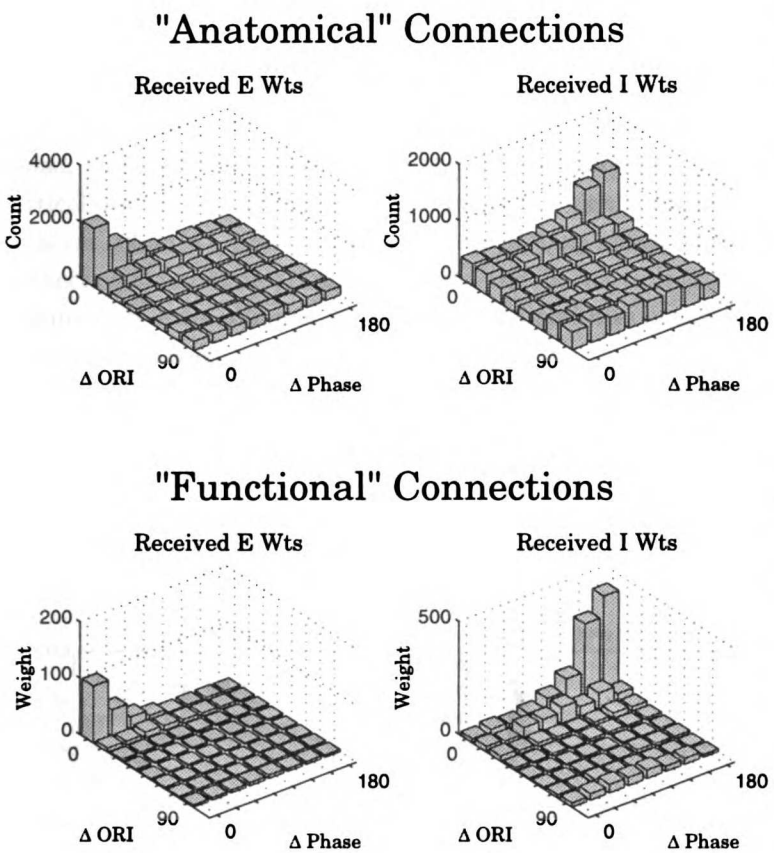


Figure 22. Summary of connection specificity for all received weights in the simulation. **Top.** "Anatomical" specificity. All weights greater than a threshold value of 0.0001 are counted equally. **Bottom.** "Functional" specificity. All weights greater than the threshold are counted, but larger weights contribute a proportionally larger amount to the total. For either method of counting, received excitatory weights arise primarily from cells of the same orientation (0° orientation difference) and the same spatial phase (0° phase difference), while inhibitory connections arise primarily from cells of the same orientation but opposite (180°) spatial phase. Spatial phases plotted here are relative spatial phases.

by abrupt changes, and no consistent progression of orientations moves across the cortical "surface."

On the other hand, other features of the map are consistent with results from the single-column simulations. Most, though certainly not all, columns develop a well-defined preferred orientation. There is no bias in the initial configuration of the model, except for the fact that within-column connections are stronger than between-column connections, but this difference is enough to ensure the vertical alignment of orientation. Thus, this local invariance is another feature that "falls out" of the model; the energy ideas that drive development of single columns hold true for multiple columns.

Another property that holds true is the phase-specificity of the synaptic weights. Figure 21 shows two examples of columns taken from the large simulation of figure 20. The one on the left is taken from the right lower corner of the orientation map, and is well-aligned; the one on the right is taken from the left upper corner of the map, and is poorly aligned. In either case, connections are predominantly phase-specific. Connection specificity is summarized for the entire simulation in figure 22; note that excitatory connections are primarily made between cells with similar orientation and spatial phase[5], but that inhibitory connections occur most often between cells of similar orientation

[5]Figure 22 actually shows relative spatial phases, which are easier to compute, and which are very similar to the absolute phases given that (1) there is no retinotopic scatter, and (2) the arbor size is small, meaning the offset between RF centers for connected cells in adjacent columns is also small.

but opposite spatial phase, as expected.

Lastly, the columns within the simulation show contrast-invariant orientation tuning, as summarized in figure 23. Gaussian fits to the orientation tuning curves of most cells show a standard deviation between 15° and 30° (left column, 4 contrasts), and the width of the tuning, though broad, stays relatively constant as the contrast increases (right column). Thus, even with intracortical connections strong enough to produce a significant degree of orientation alignment between adjacent columns – *i.e.* columns are not developing independently – columnar structure, appropriate connections, and appropriate orientation tuning can emerge. As evidenced by the poor orientation alignment of some columns and the lack of smoothness in the map, however, this development is not without some cost.

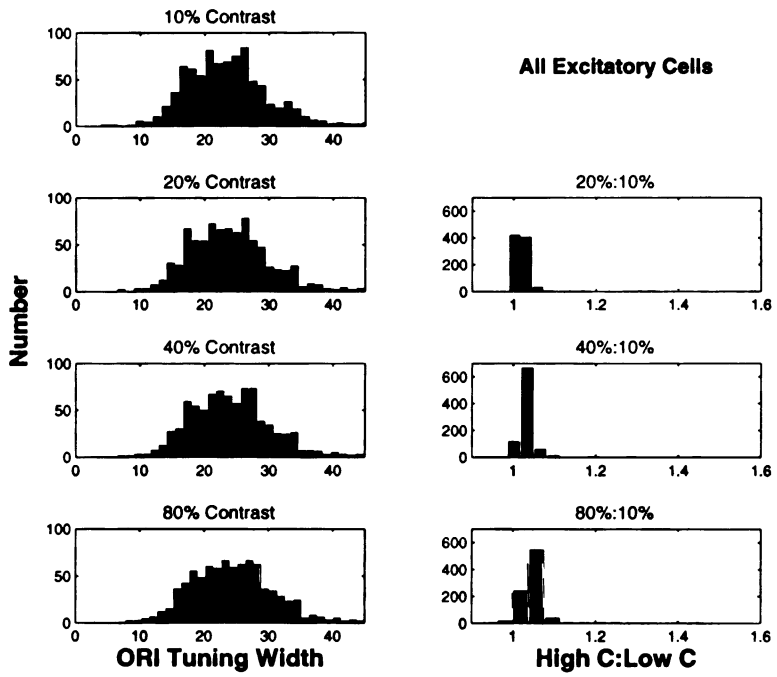


Figure 23. Contrast invariance of the central peak of the orientation tuning curve for all excitatory cells. Left column: histograms of standard deviations of the central peaks of the orientation tuning curves for all excitatory cells at 10%, 20%, 40%, and 80% contrast. Right column: ratio of high contrast:low contrast standard deviations of orientation tuning for all excitatory cells. The tuning of the central peaks is similar to that of singly-developed columns, but somewhat more broadly distributed across cells. Also as for single-column simulations, the tuning width tends to increase by about 5% at the highest contrast. 18% of the cells have null orientation responses.

Unfortunately, while retaining all these properties will be a necessary precondition for successful map development, these properties do not appear sufficient to ensure well-ordered maps. The issue of smoothness is likely tied to the strength of the between-column connections – *i.e.* to the intracortical arbor function; intuitively, more strongly connected columns should share more similar orientations (see below). Weakly-interacting columns may develop more independently, and thus more closely resemble isolated single columns, which would be expected to bear little relationship to each other. To investigate how the horizontal connection strength influences the map, we choose a smaller network – in this case, a 1-D network of 16 cortical columns – and vary the value of the arbor scale parameter between 3 and 0.11.

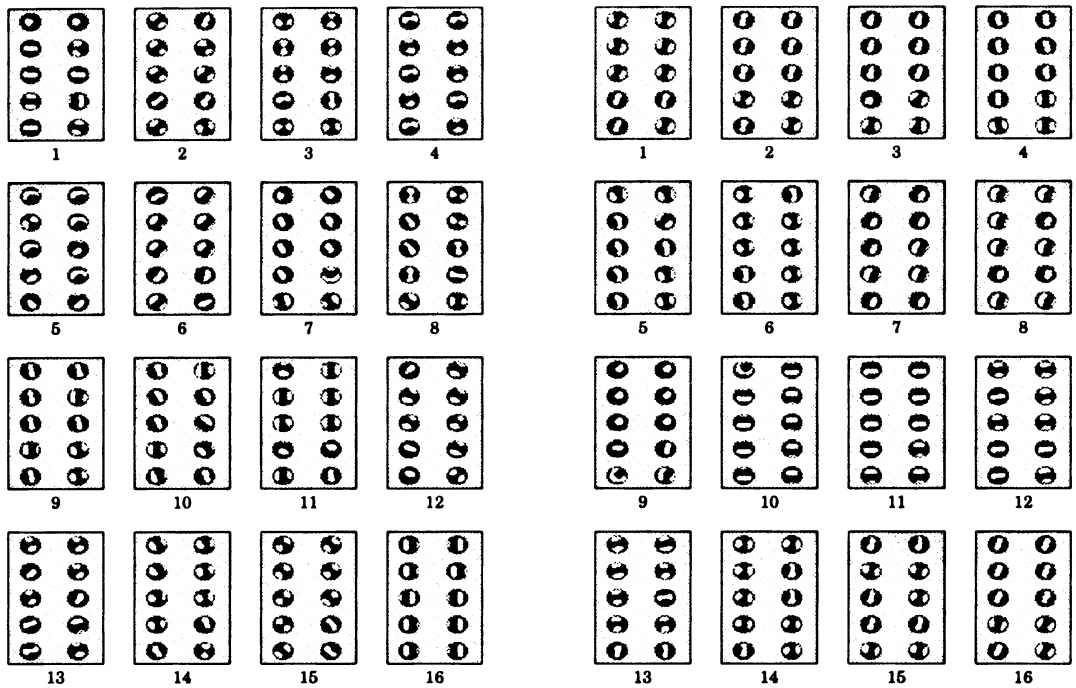


Figure 24. Results for 1-D simulations with 16 columns. **Left.** Arbor scale = 0.33. **Right.** Arbor scale = 1.0. In each case, the columns are actually organized in a single vertical line, but are presented here, numbered sequentially, in four 4-column rows.

Figure 24 shows two simulations that result from this parameter search: on the left is a simulation run with an arbor scale of 0.33, and on the right is a simulation run with an arbor scale of 1.0. The left simulation is in some ways similar to that shown in figure 20: the orientations of adjacent columns do not seem to share a defined relationship to each other, as though they were developing more independently. In the right simulation, however, columns appear to "know" about each other. Many of the columns share the same orientation – *e.g.* columns 14-16 and 1-4, which are all adjacent because of the periodic boundary conditions. Unfortunately, while this type of alignment does provide one solution to the smoothness requirement, it does not readily lend itself to a periodic solution. Columns have appeared to group into clusters of the same orientation, with large jumps in orientation, or fractures, between clusters[6]. Interestingly, even with the increased crosstalk between columns, the columnar structure of individual columns is mostly retained. Orientations align and phases do not, again suggesting that the existence of strong vertical connections is sufficient to produce columns. In some cases, though, as adjacent columns become more strongly connected, spatial phase can begin to segregate between, rather than solely within, mini-columns (*e.g.* between columns 1 and 16 in figure 24 (right), which are adjacent because of the lack of boundary conditions). This result follows from the idea that columns form from strongly-linked cells: as the between-column

[6]Despite the global change in the simulation, properties of individual cells and columns within the simulation have not changed. The simulation shown on the right, for example, still retains its contrast-invariant orientation tuning (data not shown), again indicating that this property is not solely associated with independence of columnar development.

connections become stronger, the columnar "unit" can become larger than a single mini-column. Nonetheless, most individual mini-columns still show some variation in the phase of the excitatory cells.

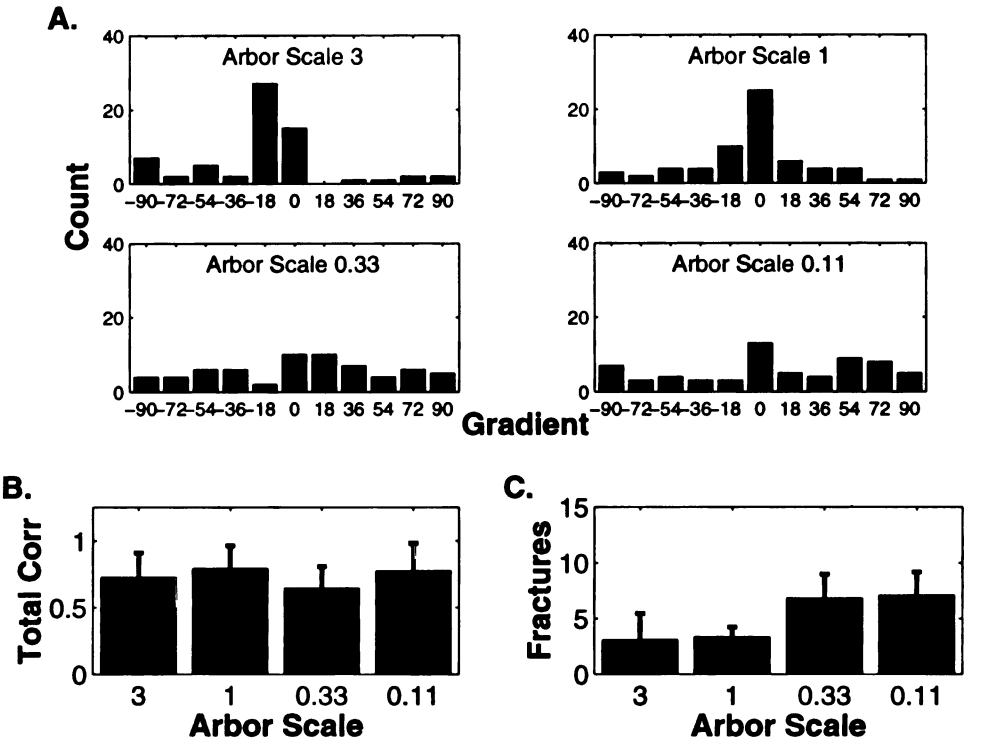


Figure 25. Summary data for 1-D simulations. **A.** Orientation differences between adjacent columns for 4 1-D, 16-column simulations each at arbor scales of 3, 1, 0.33, and 0.11. For the 2 larger arbor scales, corresponding to more strongly interconnected columns, the orientation gradient is centered at zero. For the 2 smaller arbor scales, corresponding to more weakly interconnected columns, the orientation differences are distributed more uniformly. **B.** Total correlation of each simulation, averaged across 4 simulations for each arbor scale; error bars represent standard deviations. Correlation is not affected by the arbor scale. **C.** Number of fractures, arbitrarily defined as orientation differences greater than 45°, averaged across the 4 simulations for each arbor scale; error bars represent standard deviations. A transition occurs between arbor scales of 1 and 0.33. For completely uncoupled columns we would expect to see fractures in half of the transitions – *i.e.* assuming a uniform distribution of orientation differences between 0° and 90°, where any transition greater than 45° is classified as a fracture, 8 out of 16 transitions would be so defined. For arbor scales of 0.33 and 0.11, we appear to be seeing this essentially random behavior, with approximately 8 fractures in the mean.

To quantify these ideas about smoothness, we summarize data for 4 simulations at each arbor scale in figure 25. Figure 25A shows, for all runs at each arbor scale, the combined histograms of the differences between mean orientations of adjacent columns. This difference between columns – the orientation gradient – provides a measure of the rate of orientation change. For larger arbor scales – *i.e.* stronger intracortical connectivity – the orientation differences are centered at zero,

consistent with the impression provided by the previous figure that orientation varies very little between columns. For smaller arbor scales, the orientation differences are somewhat more uniform, consistent with a more independent origin. These differences are not a consequence of problems with appropriate receptive field or weight development, as the values of "total corr" are relatively constant across arbor scales. However, the number of fractures, arbitrarily defined to be orientation differences greater than 45^o, is lower for more strongly-interconnected columns, as expected.

What is responsible for this network behavior? Given the complexity of these full simulations, we return to the Monte Carlo simulations previously used for exploring single cortical column outcomes. In these new simulations, we use a single-dimensional array of 16 columns, each of which contains two excitatory and two inhibitory cells. Our approach will be the following: if we define and fix the desired endpoint of the simulation – a smoothly varying, periodic orientation map – we should be able to understand what set of intracortical weights is consistent with that endpoint. Those weights might then be paired with plastic receptive fields to understand with what map the weights themselves are most consistent. If the weights defined by the periodic map cannot regenerate that map, the most-correlated network may not be capable of producing a biologically realistic map. The defined receptive fields are shown in figure 26. Two complete cycles in orientation – each column differing by 22.5^o from its neighbor – are used, with wrap-around boundary conditions (so that column 1 and column 16 are adjacent). With these RFs fixed, we then develop the weights.

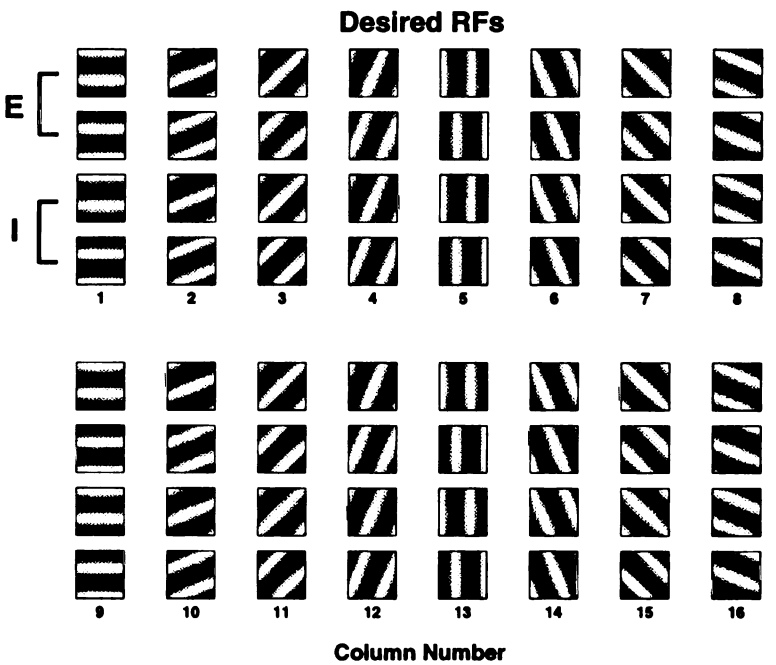


Figure 26. In multi-column Monte Carlo simulations, either the intracortical weights or the geniculocortical receptive fields are varied, but not both. As shown in the figure, to begin we fix the receptive fields for 16 columns of cells; we will first develop the weights that optimally match this well-ordered map of orientation. Between-column spatial phase is arbitrary, but within-column phase is defined to give two absolute phases.

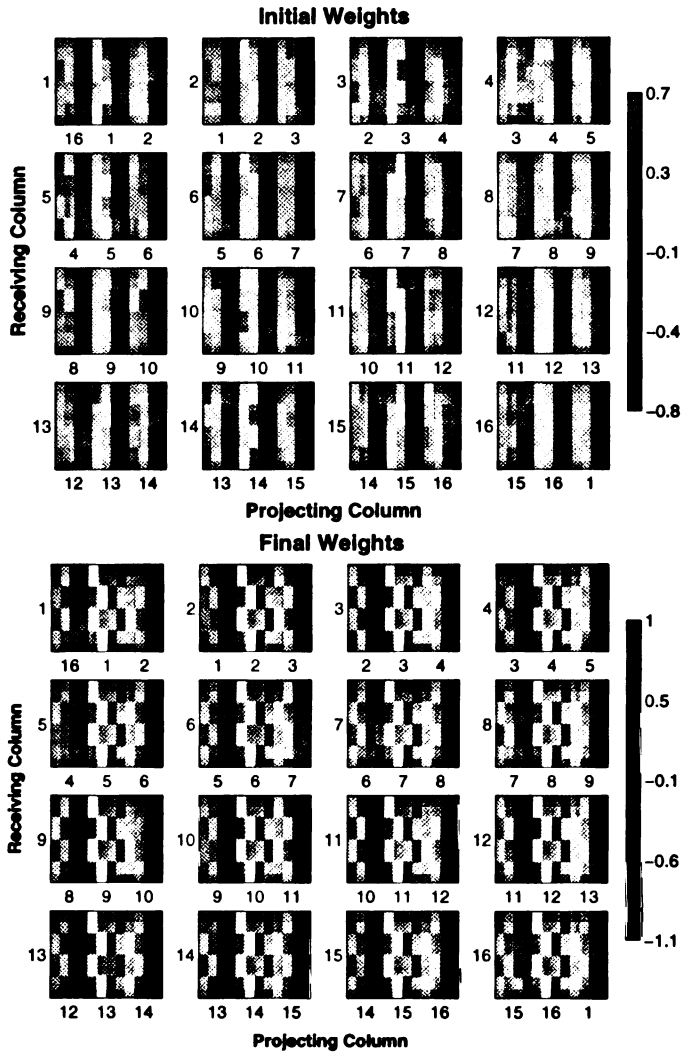


Figure 27. **Top.** Initial random weight matrix for the optimal receptive fields shown in the preceding figure. Each box shows all the weights received by a single column (number on the left) from its neighbors and from within the column. The weights received by cell #1 in column 1, for instance, are represented by the top row of the upper left-hand matrix. The first four columns represent the weights from cells 1-4 of the column to the left (in this case, column 16); the middle four columns represent weights from cells 1-4 within the column; and the right four columns represent weights from cells 1-4 of the column to the right (in this case, column 2). The size of each weight is indicated by the scale to the right of the figure. **Bottom.** Final weight matrix, after 10000 Monte Carlo time steps, for the optimal receptive fields shown in the preceding figure. Note that the weights are now very structured: the middle 4 of the 12 columns for each matrix, for instance, show the appropriate phase-specific connectivity for the within-column connections, and are stereotyped across columns.

Initially the weights have no obvious structure (figure 27, top), except for the fact that the arbor function constrains weights within a column (intracortical arbor value of 1) to be stronger than weights between columns (intracortical arbor value of 0.5), and the identities of excitatory and inhibitory weights are fixed. After the simulation develops, the weights have a stereotyped connectivity (figure 27, bottom): within-column connections are phase-appropriate, and between-columns connections are certainly structured. If we now fix these connections and allow the RFs to vary in both orientation and phase, will we get the periodic map back again?

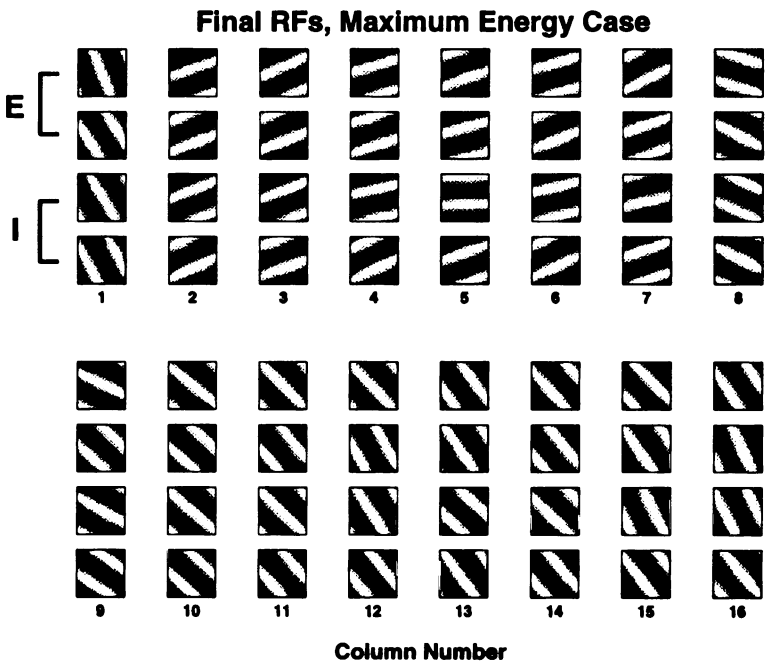


Figure 28. Maximum energy network, developed from the final weight matrix shown in the preceding figure. Given the weight matrix well-fit for the optimal receptive fields, we wanted to see whether the system could find those receptive fields again. 50 Monte Carlo simulations were run, for which the network shown in the figure was optimal. Orientations tend to remain close to identical between columns, with one jump greater than 45°.

The answer is no. Figure 28 shows the network, out of the 50 generated from the weight matrix in figure 27, with the maximum-energy final state. Within each column, cells have developed a common orientation and two opposite phases as expected. However, the progression of preferred orientation between columns, rather than showing a smoothly-changing periodic structure, tends to remain unchanging (col 2-7, 8-16) with mildly (columns 7-8) or strongly (columns 1-2) discontinuous jumps between iso-orientation blocks. In fact, the network appears to tend toward a single orientation – if the RF orientations are all set to 0°, for instance, the energy of the resulting phase-optimized network is very much larger (data not shown – energy of 22336 versus 18316).

In the minimum-energy network shown in figure 29, the tendency to regions of identically-oriented columns, separated by fractures, is even more apparent. Despite its discontinuities, however, this network is still more energetically-favorable than the "ideal" network shown in figure 26. Thus,

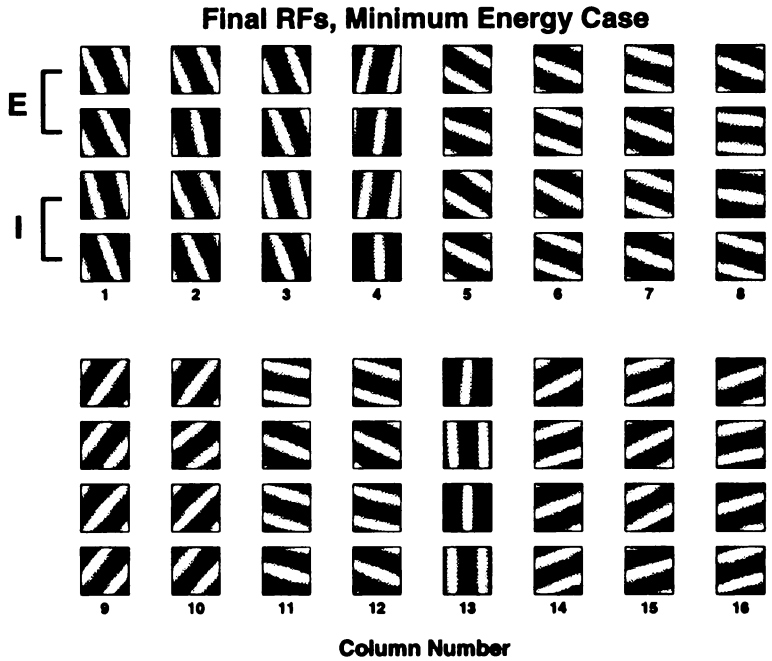


Figure 29. Minimum energy network from among the 50 Monte Carlo simulations run using the weight matrix shown in figure 27, bottom. Orientations tend to be either the same, or widely different, between adjacent columns. Even this outcome, however, is more favorable (*i.e.* has higher energy) than the desired, slowly-varying configuration.

the smoothly-varying case, under this particular energy function, is very unlikely to be reached; as defined by the energy, which represents the total correlation of the network, 50 out of 50 different networks generated from random initial conditions are more favorable outcomes.

These ideas are quantified in figure 30. Figure 30A shows the development of the energy over time for the desired RFs of figure 26 (in light gray, with plastic *weights*) and for the mean of 50 networks developed from the weights shown at the bottom of figure 27 (in dark gray, with plastic *RFs*, with dashed lines representing standard deviations). At no time is the desired network more energetically favorable. Figure 30B shows a histogram of the difference in orientation between adjacent columns in all 50 simulations. The histogram is peaked around zero – as expected since unchanging orientation is more energetically favorable – and not at $\pm 22.5°$, as expected from the desired network and represented by the white bars. In figure 30C, any orientation jump larger than 45° is arbitrarily defined as a fracture, and the number of fractures is plotted against each final state. The fact that increasing energy is associated with fewer fractures might be considered promising, but the truth of the matter is shown in figure 30D. Here, the orientation gradients are shown for the four networks with only one fracture. As can be seen in the figure, the orientations are not smoothly changing; rather, as suggested by figure 30B, and as found for the full developmental model (figure 25), they remain constant (orientation gradient of 0) with occasional jumps.

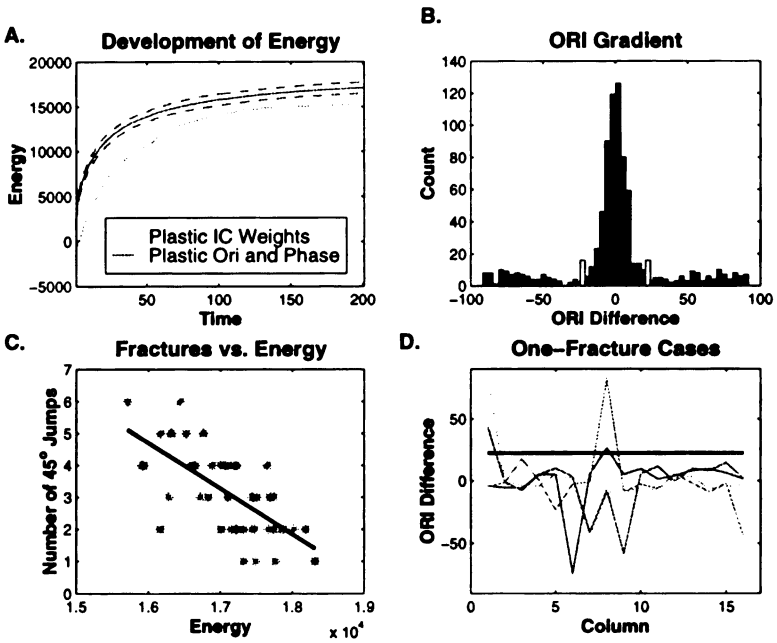


Figure 30. **A**. Development of different networks over time. The wavy, light gray line represents the energy of the network with the desired RFs fixed (figure 26) and the intracortical weights evolving. The smooth solid black line represents the mean evolution of 50 different instantiations of a network in which these developed intracortical weights in turn are fixed, and the RFs are changing. Dashed black lines show one standard deviation. All cases developed with the weights fixed have higher energy than the case with the desired RFs. **B**. The distribution of orientation changes between adjacent columns (the orientation gradients) from all 50 fixed-weight simulations. The changes appear Gaussian around zero, but show a long tail. The white bars at $\pm 22.5^o$ plot the distribution for the ideal RF case. **C**. In this plot, fractures are arbitrarily defined as an orientation jump between adjacent columns of greater than 45^o. There is a strong negative correlation between energy of the network and number of fractures ($p = 7.88 \times 10^{-8}$). **D**. Orientation gradients are shown for four networks in which only one fracture occurs. As implied by **B**, most orientation changes are near zero. The solid black line represents the gradient expected for the desired outcome.

The explanation for this map behavior lies in the relative strengths in the correlations between different receptive fields. To evaluate these correlations, we compare RFs of 8 different orientations and a single arbitrary phase to a roster of RFs with each combination of 8 orientations and 8 phases (figure 31). The results of this comparison are shown in figure 32; receptive fields of a given orientation are clearly best-correlated with other RFs of the same orientation. For example, for the reference orientation of 0^o, the best-correlated comparison RFs also have an orientation of 0^o, and a phase of either 0^o (well-correlated) or 180^o (well-anticorrelated). The best-correlated RF of the neighboring orientation (22.5^o) has an absolute correlation of approximately half, for either correlation or anti-correlation. Thus, if a network seeks to maximize its total energy, networks with long iso-orientation (high-energy) runs and occasional zero-correlation "fractures" should be far more favorable than networks in which the correlations between columns and the corresponding energy are only moderate, but constant. For example, a transition of 90° traversing 4 columns is

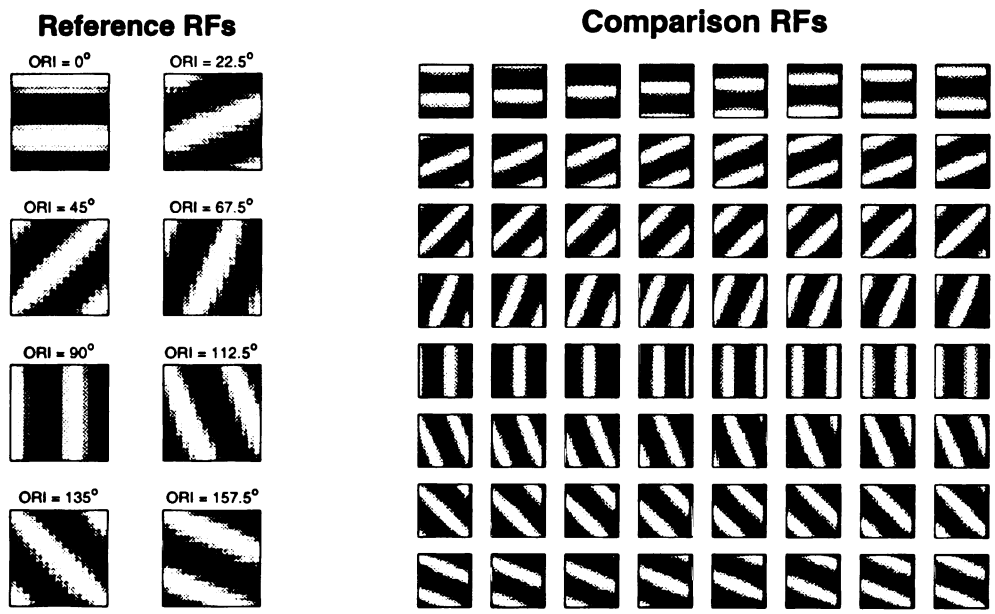


Figure 31. To understand why the networks summarized in figure 30 tend to show little change in orientation between columns, we examine correlations between receptive fields. Reference RFs, shown on the left, differ in orientation; each of these RFs is compared with all of the comparison RFs, which differ in both orientation (top-to-bottom) and phase (left-to-right). Note that adjacent columns should be slightly offset from each other; thus, the comparison RFs have a vertical line of zeros on the right side (which appears as a vertical gray bar) to indicate an area of non-overlap.

roughly 1.5 times as energetically favorable if it occurs through three 0° changes followed by a 90° jump as if it occurs through four 22.5° changes. If a network is to develop consistently changing orientations, then, this difference in correlation must be reduced. In other words, we require a different energy function, one that finds moderate (22.5°) changes in orientation to be only slightly less energetically favorable than 0° changes, so that changes in orientation will be best accomplished by smooth transitions.

Three aspects of these Monte Carlo simulations bear discussing: (1) the desired number of orientation cycles shown in figure 26 is arbitrary; (2) the amount of RF overlap between columns is not varied; (3) the arbor is arbitrarily set to a value of 0.5 between columns. These potential problems are overwhelmed by the results described above: the weights defined by the desired map seek to align all the orientations, regardless of the initial map. The amount of overlap and the arbor value may alter the strength of the correlations between columns, but as shown in figure 25, the transition between smooth and independent orientations is an abrupt one; and neither (2) nor (3) addresses the issue of periodicity. In short, how to get a smooth, periodic orientation map is not answered by simply linking columns together, and understanding map formation will require additional model constraints (see Discussion).

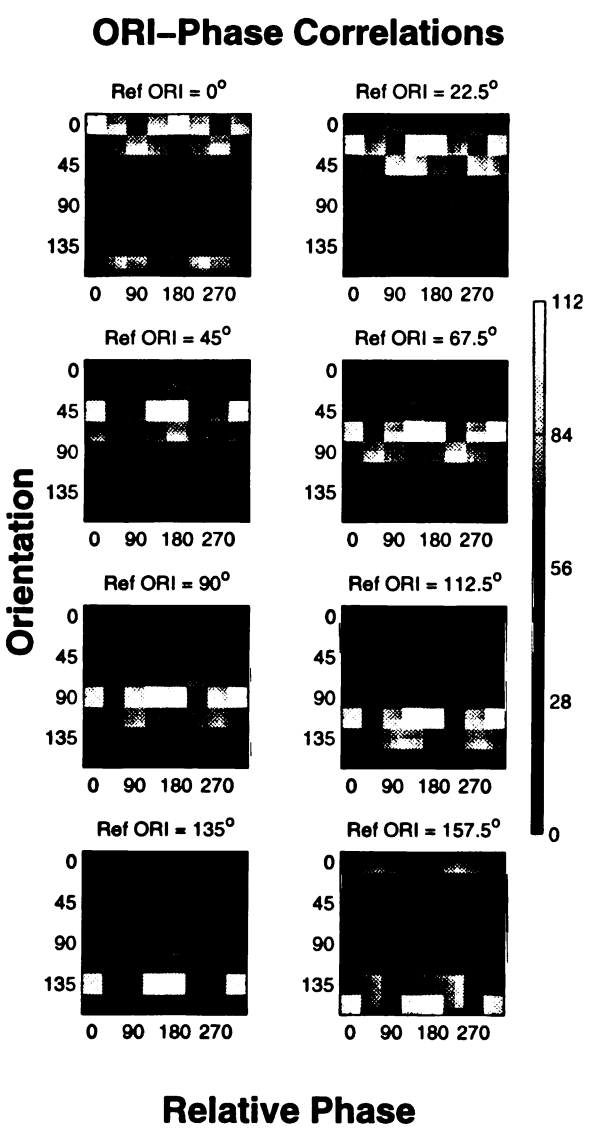


Figure 32. Absolute value of the correlation between each of the 8 reference RFs and all of the comparison RFs, plotted in 8 figures in positions corresponding to those of figure 31. RFs are well- or anti-correlated with adjacent RFs of the same orientation, as shown by the white boxes in "equal orientation" rows for which the row orientation and the reference orientation (noted above each plot) are identical. Note that because of the position offset of the reference and comparison RFs (figure 31), the best-correlated and best anti-correlated RFs are not always at 0° and 180° relative phase. Reference RFs are only about half as strongly correlated with RFs of neighboring orientation, as shown by the rows above and below the equal orientation rows. The absolute values of the correlations, as indicated by the scale bar, are arbitrary.

1.6 Discussion

Recent theoretical work has demonstrated that appropriate LGN activity correlations are sufficient to give rise to simple cell receptive fields when those correlations are coupled with activity-dependent synaptic changes and synaptic competition. Here we advance this finding to account for the generation of a functional layer 4 circuit: that is, these correlations and learning mechanisms suffice to generate not only the appropriate geniculocortical connections, but also the appropriate intracortical weights, and the appropriate local invariance of RF properties. These results follow robustly from a conceptual framework in which neurons in the network seek to maximize the correlations in their firing. Maximally correlated excitatory cells share the same orientation and absolute spatial phase, and are linked by strong excitatory weights; maximally correlated excitatory and inhibitory cells share the same orientation, but opposite absolute spatial phase, and are connected by strong inhibitory weights. From this underlying motivation follows the development of orientation-invariant, but phase-varying, columns whose neurons have properties expected of the mature cortex, including contrast-invariance of orientation tuning.

The same guiding principle, however – maximization of the total correlation of the network – fails to generate orientation maps within the current framework. As demonstrated by one-dimensional Monte Carlo simulations (eg. figure 28), the model reaches a well-correlated state by matching column orientations as well as possible for as long a cortical distance as possible – where the scale of "as long as possible" is determined by the random orientations that initially develop in different parts of the map – and by developing "fractures," where orientation changes very rapidly, between runs of the same orientation. In fact, as the model demonstrates, a fundamental question is not why orientations stay the same within this framework, but why they differ. As currently implemented, the model does not answer this question, but a number of possibilities suggest themselves (see below).

1.6.1 Comparison to Experiment

Independently of the question of maps, the model does match several experimental findings beyond those that formed the original goals – *i.e.* beyond the development of simple cell RFs, correlation-based intracortical circuitry, and orientation-invariant columns.

Many are a direct consequence of the goals themselves. The development of correlation-based intracortical circuitry, for instance, also implies that both inhibition and excitation onto a given cell arise from other neurons of similar preferred orientation (Ferster, 1988; Hirsch et al., 1998, and figure 22), and that the geniculocortical inputs and the full circuit's inputs to a cell share the same orientation tuning (Ferster et al., 1996). Moreover, the presence of inhibitory neurons within the orientation-invariant columns of the model necessitates a variation in both relative and absolute spatial phase. With the addition of retinotopic scatter to the current model (Hubel and Wiesel, 1962), a large number of relative phases can be robustly reproduced, consistent with the work of DeAngelis et al. (1999). Reexamination of their data, as well as those of other work that records neurons in the same location in cortex (eg. Hetherington and Swindale, 1999), should be sufficient to address the question of variation in absolute phase (see "Experimental Predictions").

Additionally, some of the ancillary results raise potential correspondences with existing data – some stronger than others. One such result is the finding that a ramp of the inhibitory strength can be

necessary for robust orientation development if (1) the learning rule depends on post-synaptic spiking (rather than on voltage, or on the presence of other inhibitory inputs), and (2) inhibition is required to be dominant. The difference between post-synaptic voltage and post-synaptic spiking devolves to the question of whether afferents can segregate based on subthreshold post-synaptic membrane potentials. *In vitro*, long-term depression in the presence of post-synaptic hyperpolarization (Bear and Malenka, 1994) suggests such a scenario is possible; whether this scenario operates *in vivo* normally is unknown, but it is known that chronic suppression of intracortical activity nonetheless permits ocular dominance segregation (Hata and Stryker, 1994). That inhibition can be dominant seems to be clearer. Experiments of Ferster and Jagadeesh (1992) suggest that inhibition can be powerful – nonspecific stimulation of LGN afferents induces a quick EPSP followed by a strong IPSP – as does recent work by Hirsch et al. (1998). Hirsch and coworkers found that the excitatory response evoked by a dark stimulus displayed in the OFF subregion of simple cell could be completely suppressed if that stimulus was extended only slightly into an adjacent ON subregion (their figure 5; see also discussion in Ferster and Miller (2000)). If both conditions (1) and (2) hold, and if further experiments suggest that our implementation of the Komatsu findings is incomplete (because in no case does our implementation of the Komatsu learning rule require a ramp of the inhibition), the model could provide an explanation for the progressive increase in inhibitory strength over time observed by Luhmann and Prince (1991) in rat cortical slice: well-formed receptive fields, and well-aligned columns, require time to organize. If inhibition becomes strong too soon, *i.e.* when excitatory connections are not strong enough to effectively reinforce each other, columns misalign. However, the current weight of the evidence would seem to argue against this explanation.

We also find a number of simple cells with null orientation responses. This result appears to be not a product of the form of the intracortical circuitry, but of the form of the receptive field or the amount of inhibition. A number of explanations were discussed, among them that our model of the afferent RF is oversimplified, in that its aspect ratio is too small, for instance (Jones and Palmer, 1987); or that inhibition should be yet stronger. One could also argue that inputs from other cortical cells not included here may improve tuning, and that the inclusion of longer-acting inhibition (*e.g.* $GABA_B$) may effectively increase the response threshold, and thereby remove null responses. On the other hand, examples in the literature from both cat (Azouz et al., 1997, see discussion) and monkey (figure 1 of Celebrini et al., 1993) suggest that some (up to 70% in the former study) simple cells will respond to the null orientation. Of course, these responses could depend on the factors just mentioned, perhaps via the type or level of anesthetic and the concomitant strength of inhibition. It will be interesting to determine how prevalent this property is, and to see whether the null response can be suppressed: studies have reported responses to the null orientation upon application of GABA antagonists (Sillito, 1975; Sato et al., 1996), but none have examined whether GABA agonists can remove already-present null responses. Of note, the model tends to show second response peaks at the null orientation, rather than a single broad peak at the preferred orientation that extends to the null; however, the experimental data do not describe the extents of their null responses in orientation (but again see Celebrini et al. (1993), figure 1).

Another result is the behavior of orientation columns where poor orientation selectivity occurs in multiple-column simulations. In the anesthetized animal, areas in an optical map can show weak responses to oriented stimuli for many reasons, but two interesting possibilities are (1) that neurons

in a particular cortical position do not have strong preferred orientations, or (2) that those neurons are strongly oriented, but they prefer quite different orientations. Recently Maldonado et al. (1997) showed that cells in pinwheel centers of cat striate cortex are well-oriented, but tuned to different orientations. Some caveats are important – the seeing distance of their tetrode could allow them to record more frequently from well-oriented cells in adjacent columns, for example – but their result can also be seen in the model here. In figure 21, the model tends to develop strong ON and OFF subregions. For the most part, relatively unoriented columns within a mature orientation map are composed of well-oriented cells with differing orientation preferences, though there are cases where the model will develop "checkerboard" RFs that have strong preferences for orthogonal stimuli. In either case, the model is more consistent with a well-developed subregion structure; the geniculocortical correlations appear to favor this outcome.

1.6.2 Experimental Predictions

Certain results of the model provide a set of experimental predictions. First, local layer 4 connections should be predominantly phase-specific: cross-correlation analysis should demonstrate that excitatory connections received by a cell arise primarily from neurons with receptive fields of the same orientation and absolute phase, while inhibitory connections arise primarily from neurons with receptive fields of the same orientation but opposite absolute phase. This specificity, of course, is primarily functional: more promiscuous connectivity may be seen anatomically, but physiological connectivity should be mostly phase-appropriate. (One exception might be $I \rightarrow I$ connections; as discussed in Results, because *all* inhibitory cells respond at the column's null orientation, irrespective of spatial phase, $I \rightarrow I$ weights learn on patterns in which phase specificity is not well-respected, and can consequently show weaker phase selectivity than other weight types in the final state.) Some hints of this type of relationship can be found in DeAngelis et al. (1999, figure 6a). Their "similarity index" $SI = \sum_S (U(S) * V(S)) / \sqrt{\sum_S U(S)^2 \sum_S V(S)^2}$, where U is one receptive field, V is the other, and S represents the input space, is almost a measure of the absolute phase: it is just the point-by-point multiplication of each RF. They find that where the similarity index is high, the degree of cross-correlation is also high, consistent with stronger connections between cells of like RF. Unfortunately, this measure is not capable of addressing absolute phase for intermediate similarity values: RFs that are offset from each other, but match very well where they do overlap, can have the same similarity value as neurons that are centered at the same point, but have moderately different orientations or spatial phases. Future work might look explicitly at the absolute phase of the RFs, as well as considering visual (eg. full-field noise) or pharmacological (eg. glutamate) stimuli that raise cortical firing rates and thereby permit examination of inhibitory troughs in the correlogram. These data will also shed light on whether the finding of two absolute phases per column in the model is a realistic one. The existence of only two such phases may seem simplistic, but it is not yet ruled out by the published experimental data.

Second, RF properties that show local invariance should be those that are shared by anti-correlated, as well as by correlated, RFs. This prediction would not be limited to the visual system, of course; the presence of a correlation-based learning rule, in conjunction with intracortical inhibition, implies that well-correlated and well-anticorrelated neuronal receptive fields should be found within the same column, regardless of modality. In the somatosensory system, for instance, receptive

fields in primary somatosensory cortex (area 3b) resemble visual cortical neurons in that they possess excitatory and inhibitory subregion structure (DiCarlo et al., 1998). A number of the findings presented here for the visual system would be expected to follow through: excitatory connections should form primarily between neurons of similar absolute phase, and inhibitory connections should primarily link neurons of opposite spatial phase, for example.

Third, columnar structure should result very robustly if vertical connections within the cortex are strong. In the model, "columns" develop between groups of cells that are strongly linked; as demonstrated in the multi-column simulations, for instance, increasing the strength of between-column weights can drive the orientations of adjacent columns together, and lead the relative phase of the excitatory cells to sometimes segregate *between*, rather than within, columns. Removing all vertical connections prevents these columns from forming. Thus, one would expect that any manipulation that affected vertical connectivity during development might alter columnar formation. Animal models of cortical dysgenesis that display defective vertical connections suffer from the possibility that other major problems – such as in single neuron processing – might exist. Other approaches, such as those involving pharmacological modification of vertical connections, are not yet known, but a suggestive finding was noted by Xiang et al. (1998). They found that application of a nicotinic agonist to rat visual cortical slice could preferentially activate layer 5 low-threshold spiking (LTS) inhibitory cells, which ramify vertically within cortex. Chronic infusion of a nicotinic agonist during development (a technique previously performed with metabotropic glutamate receptor antagonists in cortex (Hensch and Stryker, 1996)) might preferentially inhibit responses outside of layer 5. If such specificity is possible, and if the activation of these neurons can decorrelate activity between the layers, one might then expect the alignment of orientation tuning to differ between layer 5 and the rest of primary visual cortex.

Lastly, it is unclear how well these predictions apply to the primate visual system. Most of the LGN-recipient cells in layer 4C are unoriented (Blasdel and Fitzpatrick, 1984; Hawken and Parker, 1984), unlike those in layer 4 of the cat, suggesting that if the circuit exists, it is present in the transformation from layer 4C to simple cells. Even if the circuit is absent, however, we would expect to see a bias for excitatory connections between similar RFs and for inhibitory connections between dissimilar RFs, and a dependence of columnar development on the strength of vertical connections.

1.6.3 Model and Assumptions

In the current model, a number of assumptions have been made (see Methods). Following the work of Miller (1994) and Erwin and Miller (1998), one of the most critical is the existence of appropriate correlated LGN activity: disruption of the postulated early LGN activity correlations should disrupt development of all of the circuit elements – receptive fields, intracortical connectivity, and columnar structure. Two recent experiments bear on this issue. Gödecke and Chapman (1998) injected APB into both retinae of developing ferrets to selectively block all ON-responses, leaving OFF-responses intact. If the APB is not infused too early (which leads to silencing of the cortex, and which may have non-specific effects on OFF-responses), cortical activity is retained, but development of orientation selectivity is not. Weliky and Katz (1999) used electrode arrays to directly determine correlations between LGN neurons in ferrets. Unfortunately, they were unable to simultaneously isolate both

retinotopic and center-type specific correlations, and they looked from P24-P27, before the major onset of orientation selectivity development (P30-32; Chapman and Stryker, 1993). Thus, the former experiments furnish consistent, but by no means conclusive, results; the latter, unfortunately, do not further constrain the model.

The learning mechanisms are taken relatively directly from experimental results. An excitatory Hebbian covariance rule is broadly consistent with much of the LTD/LTP literature: stimulation frequencies less than some threshold value lead to LTD, while frequencies greater than a threshold value lead to LTP (Bliss and Collingridge, 1993), though there is no strong evidence that this threshold slides with mean activity. The inhibitory learning rule is derived from the work of Komatsu (Komatsu and Iwakiri, 1993; Komatsu, 1994; Komatsu, 1996), though some simplifications have been made (see Methods). By far the least well-constrained portion of the model, however, is the nature of the competition between synapses. Experimental evidence bearing on this point has just begun to accumulate (Scanziani et al., 1996; Turrigiano et al., 1998), suggesting in the first case that a form of heterosynaptic long-term depression exists in hippocampus *between* cells, and in the second, that pyramidal neurons may be able to multiplicatively rescale their inputs to compensate for activity changes. Neither of these phenomena is analogous to the subtractive constraint used in this work, which operates globally and uniformly on all the synapses to a given cell, and which we choose as a relatively well-studied computational means (Miller and MacKay, 1994) of enforcing competition. Recently, theoretical work has suggested that the experimental dependence of LTP and LTD on spike timing (Bi and Poo, 1998) may be sufficient to give rise to subtractive-style constraints (Song, Miller and Abbott, 1999). Unless this work advances, the question of the proper constraint remains open (Miller, 1996); while a subtractive constraint is not ruled out by existing data, other constraints, similarly plausible, could potentially lead to different model behavior.

1.6.4 Comparison to Other Models

Few other models examine either (1) plastic intracortical connections or (2) the existence of multiple cells at a single cortical location, and none examine both.

The models of Sirosh and Miikkulainen (Sirosh and Miikkulainen, 1997; Sirosh and Miikkulainen, 1994) are among the first, if not the first, to examine the effects of plastic intracortical connections. They take the first step beyond the traditional 2-dimensional model by allowing the individual weights of the fixed intracortical interaction – generally of a radially-symmetric Gaussian or Mexican-hat form (eg. Obermayer et al., 1992; Miller, 1994) – to vary in an activity-dependent fashion. To permit this scheme to work, they assume that (1) inhibition is strengthened between coactive cells, rather than weakened as observed experimentally (Komatsu and Iwakiri, 1993; Komatsu, 1994; Komatsu, 1996); (2) inhibition extends much further than does excitation, thereby retaining an initial Mexican-hat form; and (3) single cells can extend both excitatory and inhibitory weights. The questions with this model concern the level of abstraction. If we view the model very literally, the inhibitory learning rule appears to disagree with the work of Komatsu, and the fact that single cells, for instance, project both excitatory and inhibitory weights is certainly not biological. On the other hand, if we view each "cell" in their model as representing a cortical column, then the inhibitory learning rule *might* represent a columnar version of the increase in inhibition seen onto single cells in response to increased cortical excitability (Rutherford et al., 1997), and the projection

of both excitatory and inhibitory weights might be an approximation to column-column interactions. Even so, the model does not examine the intracolumnar connections discussed here.

Other models also examine either one or the other of these issues. In a similar vein, Rao and Ballard (1999) seek to develop both feedforward and intracortical weights in a model of primary visual cortex. However, their model attempts to explain the influence of cortical feedback on such properties as endstopping; they do not explicitly consider development of columns, but rather the development of appropriate simple cell receptive fields as a process embodied by a neural implementation of a Kalman filter. The developmental model of Olson and Grossberg (1998) does examine the structure of single columns, but the intracolumnar weights are fixed. They establish local "dipole units" for each column in which two pools of cells, each of which receives plastic geniculocortical weights, are constrained to develop in a mutually inhibitory fashion. Thus, although the cells are free to develop their afferent receptive fields, the intracortical interaction "function" is unchanging.

1.6.5 Future Directions

The obvious next steps for the model concern (1) orientation maps and (2) layer 2/3 responses. For goal (1), smoothness constraints seem plausible to implement; creating a learning rule in which the contributions of pre- and post-synaptic elements saturate can deemphasize the difference between smaller and larger correlations, and should improve the ability of adjacent columns to develop slightly different orientations without compromising the overall energy of the network. However, such a network still has no innate tendency towards a periodic solution. Perhaps the most intuitive next step – a simple Mexican hat interaction within layer 4 – will not work; once more than one absolute spatial phase is present at a given location, the least-correlated orientation at the scale determined by the inhibitory portion of the Mexican hat is no longer the orthogonal orientation, as in the 2-dimensional model of Miller (1994), but the *same* orientation, of opposite spatial phase. Consequently, tight excitation coupled with more broadly-ranging inhibition may only encourage long runs of similar orientations, rather than smoothly varying orientation.

The investigation of goal (2) might answer the question of periodicity. Layer 2/3 responses should primarily be those of complex cells. In the literature are a number of possible methods for developing such receptive fields – the feed-forward model of Hubel and Wiesel (1962), the simplex model of Debanne et al. (1998), and the recurrent-gain model of Chance et al. (1999), for example – but assuming that complex cells do develop, perhaps under one of these rubrics, the interaction of these cells could provide a more interesting horizontal structure. Cells insensitive to phase, such as complex cells, should respond to a Mexican hat cortical interaction in the desired fashion: the least correlated cells now differ in orientation. Long-range connections within layer 2/3, which preferentially connect columns of like orientation, could contribute to long-range periodicity (Shouval et al., 1998). One of the primary difficulties with implementing such a scheme is the way in which the feedback from layer 2/3 returns to layer 4. Anatomically, this feedback would have to occur via the loop 2/3→5→6→4. But even if we model this loop in a simplified way as a direct connection from 2/3 to 4, how these projections should be constrained anatomically, let alone physiologically, would certainly require a number of assumptions.

Nonetheless, in its current state the model demonstrates that the appropriate input correlations are both consistent with, and can account for, a number of developmental and mature visual cortical

properties. More generally, these simulations emphasize for the first time that neural activity is a plausible means of establishing both thalamocortical and intracortical cortical circuitry, as well as columnar structure, within layer 4 without strict genetic prespecification. With experiments, many of which have already been started, to test the predictions of the model and to better identify some of the mechanisms included within it, further studies might be better able to address formation of circuitry within other cortical layers, and in higher cortical areas.

Chapter 2

Operation of the Mature Circuit

2.1 Abstract

We study a recently proposed "correlation-based", push-pull model of the circuitry of layer 4 of cat visual cortex (Troyer et al., 1998). The model circuit was previously shown to explain the contrast-invariance of cortical orientation tuning: given the correlation-based connectivity between cells, the circuit's dominant inhibition suppresses the untuned component and sharpens the tuned component of the LGN input, and produces contrast-invariant orientation tuning. Here we show that the model can also account for several contrast-dependent (c-d) "nonlinearities" in cortical responses. These include an advance with increasing contrast in the temporal phase of response to a sinusoidally modulated stimulus; a relative enhancement with increasing contrast of responses to higher temporal frequencies; and contrast saturation that occurs at lower contrasts in cortex than in LGN. Within the model circuit, these properties arise from a mixture of nonlinear cellular and synaptic mechanisms: short-term synaptic depression, spike-rate adaptation, contrast-induced changes in cellular conductance, and the nonzero spike threshold. The former three mechanisms are sufficient to explain the experimentally observed increase in c-d phase advance in cortex relative to LGN. The c-d changes in temporal frequency tuning arise as a threshold effect: voltage modulations in response to higher-frequency inputs are only slightly above threshold at lower contrast, but become robustly suprathreshold at higher contrast. The other three nonlinear mechanisms are also crucial, allowing contrast-dependent temporal frequency tuning to coexist with contrast-invariant orientation tuning. Contrast saturation, and the observation that responses to stimuli of increasing temporal frequency saturate at increasingly high contrasts, can be induced by both synaptic depression and the model's push-pull inhibition. A previous proposal (Carandini and Heeger, 1994) explained these nonlinear response properties by assuming (1) an underlying linear receptive field, produced by a balanced push-pull arrangement of LGN inputs, that guarantees contrast-invariant orientation tuning; and (2) normalization by shunting inhibition, derived equally from cells of all preferred orientations, to account for the nonlinear response properties without disrupting (1). The present proposal extends the model of Troyer et al. (1998) to simultaneously explain both contrast-invariant orientation tuning and these c-d nonlinearities, and requires only processing that is local in orientation, in agreement with intracellular measurements (Ferster, 1986; Anderson et al., 1999).

2.2 Introduction

The response properties of simple cells in layer 4 of cat primary visual cortex (V1) serve as a model system for studying the mechanisms underlying cerebral cortical processing. These cells are perhaps the best-studied cortical cells, and are the site of emergence of the strong selectivity for stimulus orientation seen throughout visual cortex (Hubel and Wiesel, 1962).

One of the defining characteristics of simple cells is the largely linear nature of their responses. Their responses to arbitrary stimuli can be reasonably well predicted from a weighted sum of the luminance stimulus, where the weighting is given by the cell's receptive field, the luminance stimulus is represented as local contrast (deviation from the mean luminance in units of the mean), and negative values of the weighted sum are taken to yield zero response (Hubel and Wiesel, 1962; Jones and Palmer, 1987; DeAngelis et al., 1993). As predicted by a linear response model, the shape of a simple cell's orientation tuning curve is invariant to changes in stimulus contrast (Sclar and Freeman, 1982; Skottun et al., 1987): a change in contrast multiplies all responses by a constant, rather than changing the form of the response tuning curve. However, other aspects of simple cell responses show a nonlinear dependence on stimulus contrast (reviewed in Carandini et al., 1998).

In this paper, we demonstrate for the first time a unified mechanistic account of both the linear and nonlinear aspects of simple cell responses. We have recently demonstrated (Troyer et al., 1998) that a simple intracortical circuit with sufficiently strong inhibition and correlation-based connectivity motivated by numerous intracellular studies (*e.g.*, Ferster, 1986; Ferster, 1988; Nelson et al., 1994; Ferster et al., 1996; Chung and Ferster, 1998; Hirsch et al., 1998; Anderson et al., 1999) can, along with the arrangement of lateral geniculate nucleus (LGN) inputs to simple cells, account for the contrast-invariance of orientation tuning. We now show that this model, without disrupting contrast-invariant orientation tuning, also provides a mechanistic explanation for three nonlinear simple cell response properties: (1) Contrast-dependent phase advance: as the contrast of a sinusoidal grating stimulus increases, the response of a cortical cell occurs earlier in the stimulus cycle (Dean and Tolhurst, 1986; Albrecht, 1995). (2) Contrast-dependent temporal frequency tuning: with increasing stimulus contrast, there is a relative increase in response to high- vs. low-temporal-frequency stimulus gratings (Holub and Morton-Gibson, 1981; Albrecht, 1995). (3) Contrast saturation: cortical responses may plateau with increasing contrast (Albrecht, 1995) at lower contrasts than do LGN responses (Sclar, 1987; Cheng et al., 1995).

These response properties have previously been argued to arise from "normalizing" inhibition: inhibition dependent only on stimulus contrast, and thus derived from a pool of cells of all preferred orientations (Heeger, 1992; reviewed in Carandini et al., 1997, 1998; see also Albrecht and Geisler, 1991). That proposal arose from considering how to "correct" a linear model: simple cells were assumed to receive input that depends linearly on the stimulus, and it was then asked what must be added to also explain response nonlinearities. However, modeling even these linear aspects of simple cell responses can be problematic, as is illustrated by the problem of contrast-invariance of orientation tuning in response to drifting sinusoidal luminance gratings. Because their response rates cannot decrease below zero, LGN mean firing rates increase with contrast (whereas under a linear LGN response model, an increase in stimulus contrast would increase the amplitude of temporal modulation of firing rates without affecting mean rates). Furthermore, cortical cells integrate this input through the nonlinearity of a nonzero spike threshold. Thus, the orientation tuning of the

LGN input to a simple cell widens with increasing stimulus contrast: due to the increase both in modulations and means of LGN firing rates, a broader range of stimulus orientations produce suprathreshold LGN input at higher contrasts.

Our previous work (Troyer et al., 1998) demonstrated that addition of inhibition like that observed experimentally to the LGN input can explain contrast-invariant orientation tuning. The inhibition suppresses an untuned component and sharpens a tuned component of the LGN input, but operates within model circuitry that is local in orientation: the inhibition and excitation received by a cell have orientation tuning peaked around the cell's preferred orientation and falling to small values at orthogonal orientations, as observed experimentally (Ferster, 1986; Anderson et al., 1999). The model contained numerous nonlinearities found in the cortical circuit, including nonlinear LGN response properties, nonzero spike thresholds, spike rate adaptation currents in excitatory cells, and stimulus-induced changes in integration time constants; in the present work we also add short-term synaptic depression (Abbott et al., 1997; Tsodyks and Markram, 1997). Here we show that, due to these nonlinear mechanisms, the model circuit also naturally gives rise to the studied nonlinear response properties, without sacrificing the contrast invariance of orientation tuning.

Specifically, we and others have previously demonstrated that synaptic depression can contribute to contrast-dependent phase advance (Priebe et al., 1997; Chance et al., 1997; Chance et al., 1998); so too can contrast-dependent reductions in time constant of cells in the circuit (Carandini and Heeger, 1994). Now we show that these mechanisms and spike-rate adaptation together can quantitatively account for cat V1 phase advance data. Contrast-dependence of temporal frequency tuning arises as a consequence of the spike threshold nonlinearity in combination with depression, adaptation, and changes in time constant. Finally, contrast saturation is shown to result from the push-pull interaction between inhibitory and excitatory cells as well as from short-term synaptic depression.

Some of these results have appeared in abstract form (Priebe et al., 1997).

2.3 Methods

In this section we present the full details of our methods necessary to replicate our work. To ensure accessibility, the Results section begins with a brief summary of the modeling framework providing the essential information needed for understanding the Results. The reader may wish to start there.

Our approach combined two methods. We used rate-coded neurons to gain a basic understanding of the circuit effects of synaptic depression, and we then used more biophysically realistic neurons to understand these effects in more detail. The more realistic model is similar to that of (Troyer et al., 1998), except that synaptic depression was not included there.

2.3.1 Elements in common to both rate and spiking models

Geniculate responses. Neurons in the lateral geniculate nucleus (LGN) can be classified into ON cells (response to light onset) and OFF cells (response to light offset). For the purposes of this paper, we only considered ON and OFF cells of the X type, which dominate central V1 physiology in the cat (Ferster 1990a,b); and we only consider non-lagged cells, ignoring lagged cells for simplicity (Saul and Humphrey, 1990). Our model was based as much as possible on experimental data obtained at

approximately 5° eccentricity.

Because X cells behave quite linearly (Shapley and Hochstein, 1975), we were able to plausibly approximate the LGN grid as a linear filter. Thus, geniculate firing rates in response to sinusoidal moving gratings were assumed to be the rectified sum of two factors: (1) a sinusoidal modulation with the same temporal frequency as the stimulus, and with amplitude and phase determined as described below; and (2) background firing rates of 15 Hz and 10 Hz for ON and OFF cells, respectively (Levine and Troy, 1986; Kaplan et al., 1987). By rectified sum, we mean that negative values of the sum were set equal to zero. A difference between ON and OFF background firing rates is not important to the results, and was included simply to ensure that results do not depend on precise equality of these rates.

The size of the sinusoidal modulation can be described in terms of the first harmonic (F1) of the LGN firing rates either before or after rectification (the F1 is the amplitude of the response component at the same temporal frequency as the stimulus).[1] For most of the simulations presented in this paper, the pre-rectification F1 – the amplitude of the sinusoidal modulation described in the previous paragraph – in response to a given stimulus of the preferred spatial frequency was chosen separately for ON and OFF cells (because of their different background firing rates) so that the post-rectification F1 – the F1 of the geniculate response after rectification – matched the F1 values determined by Sclar (1987, figure 1) for LGN responses to stimuli of different contrasts and temporal frequencies. However, for some of the simulations (denoted as "flat" in Results), we set the pre-rectification F1 amplitudes to as many as four different arbitrary values (15, 30, 60, and 90 Hz) that were held constant across temporal frequencies in order to remove this aspect of the input variability. The F1s calculated from the LGN *outputs* were significantly lower than these pre-rectification values due to the rectification step and to spatial-frequency filtering by the LGN (see below). To equate the "flat" amplitudes with contrasts (and, more generally, to fit all other contrast-saturation curves), we used matlab's "curvefit" function to fit the pre-rectification F1 values at each temporal frequency for ON cells (from the Sclar data) with Naka-Rushton curves (Albrecht, 1995):

$$R(C) = R_{max} C^n / (C^n + C_{50}^n) \tag{2.1}$$

where R is the cortical response as a function of the contrast C, R_{max} is the maximum saturated response, n is an exponent, and C_{50} is the contrast that evokes a half-maximal response. From the fit curves we found the corresponding contrasts for each pre-rectification F1 at each temporal frequency. We then combined the data derived from "flat" inputs with those from Sclar inputs to generate contrast saturation curves (figure 37).

The pre-rectification F1's as chosen above were further modified for stimuli of nonoptimal spatial frequencies by use of the following center-surround LGN spatial filter (Peichl and Wassle, 1979;

[1]Throughout, we normalize the F1 to equal the amplitude of the sinusoidal component at the frequency of the grating stimulus. If the LGN input has temporal frequency ω, this normalized F1 is given by the sum of the amplitudes of the ω and $-\omega$ frequency components of the Fourier transform, when that transform is normalized so that the F0 or DC is the mean rate; this normalization is standard in neurophysiology(Skottun, Grosof and de Valois, 1991). We have previously (Troyer et al., 1998) incorrectly stated that this normalization of the F1 requires that the Fourier transform have an extra factor of two relative to the normalization that makes the F0 equal to the mean rate. This mistake was due to our ignoring the $-\omega$ component, which has equal amplitude to the ω component; the factor of two is accounted for by including the negative as well as positive frequency components.

Linsenmeier et al., 1982):

$$17e^{-k^2\sigma^2_{\text{center}}/4} - 16e^{-k^2\sigma^2_{\text{surround}}/4};\quad \sigma_{\text{center}} = 15', \sigma_{\text{surround}} = 1^o. \tag{2.2}$$

Here, k is the spatial frequency of the grating; its wavelength λ is given by $\lambda = 2\pi/|k|$. Because all gratings were shown at the preferred spatial frequency of the model cortical cells (0.635 deg/cycle, Troyer et al. (1998)), the modulation amplitude was reduced by the amount predicted by this filter relative to its value at the preferred spatial frequency of the LGN (0.54 cyc/deg, the k for which the filter value is maximized).

The phase of the response was simply taken to be equal to the stimulus phase for ON cells, and 180° out of phase with the stimulus phase for OFF cells. There is no need for an overall time lag (a constant phase lag) relative to the stimulus, since such a lag has no effect on results (one can simply imagine the stimulus to have been presented earlier in time). However, this choice of phase neglects all *changes* in LGN response phase with stimulus contrast. Thus, all changes in cortical response phase in the model resulted solely from geniculocortical depression and intracortical processing (discussed below), allowing us to focus on the question of whether these factors are sufficient to account for the experimentally-observed *difference* between LGN and cortical contrast-dependent phase advance.

Our LGN model is simplified in a number of respects. LGN firing modulations before rectification may not be linear (*e.g.*, Reich et al., 1997), and the precise structure is likely to be anaesthesia-dependent. The background rates also most likely vary with depth and type of anesthesia, as well as with background luminance. For urethane anesthesia and photopic-range luminances, estimates of background firing in the range of 10 to 15 Hertz appear reasonable (Levine and Troy, 1986; Kaplan et al., 1987); however, our data on amplitude of geniculate firing rate modulations come from pentobarbital-anesthetized animals (Sclar, 1987). More accurate models will require such data to be derived under the same experimental conditions. In addition, we have neglected between-cell variability in response amplitudes and phases (*e.g.*, Saul and Humphrey, 1990).

Cortical receptive fields (RFs). Simple cell afferent synaptic weight distributions were each described by a Gabor function. Functions of this form have been shown to provide a good quantitative fit to many simple cell receptive fields (Jones and Palmer, 1987). If x is distance from the RF center along a cell's axis of preferred orientation, and y is distance from the RF center along the perpendicular axis, then the form of the function is:

$$G(x,y) = e^{(-x^2/\sigma_x^2)}e^{(-y^2/\sigma_y^2)}cos(2\pi fx + \phi). \tag{2.3}$$

Here, x and y have units of degrees, σ_x and σ_y are the standard deviations of the respective Gaussian envelopes, and f and ϕ are the spatial frequency (in cycles/degree) and the phase offset (in radians), respectively, of the cosine modulation. We used the parameters $\sigma_y = 0.8204°$, $\sigma_x = 0.4767°$, as described below; the phase offset ϕ was systematically varied. The spatial frequency of the cosine modulation was chosen to be 0.8 cycles/degree; when the Gabor was convolved with the center-surround structure of the LGN cells' RFs, the peak response occurred at a spatial frequency of approximately 0.635 deg/cycle (Troyer et al., 1998, figure 11), consistent with data for 5° eccentricity from Movshon et al. (1978). The values of σ_x and σ_y were chosen to achieve the following: each RF had 2.64 subfields, defined as the ratio of the width of the gaussian envelope at 5% of peak (1.65°)

to the width of a single subregion (0.625°), and an aspect ratio of 4.54, defined as the ratio of the length of the gaussian evelope at 5% of peak (2.84°) to the width of a single subregion. These values were the mean spatial values for simple cell physiologic RFs reported in Jones and Palmer (1987).

Synaptic depression. Synaptic depression is a use-dependent decrease in the strength of a neuronal connection; as the firing rate of a presynaptic neuron increases, the effect of its synapses on the postsynaptic cell decreases. As described recently by Abbott et al. (1997) and Tsodyks and Markram (1997), synaptic depression in cortical synapses can be characterized by two parameters: f, the fractional decline in synaptic strength after a spike; and τ, the time constant of recovery of synaptic strength. Following a spike, the synaptic weight is multiplied by the fraction f, where $0 \leq f \leq 1$, after which the weight recovers with time constant τ toward its undepressed value.

The form of equation used to describe depression depended on whether we use rate-coded neurons or spiking neurons. The equation for a spiking model is

$$\tau \frac{dw}{dt} = -w(t) + w_{\max} - \tau c_1 \rho(t) w(t) \tag{2.4}$$

where $\rho(t)$ describes the spike train of the input whose synaptic weight is being considered, and c_1 is a constant to be determined. Demanding that the weight is multiplied by f ($0 \leq f \leq 1$) after a spike determines that $c_1 = -\ln f$ (see Appendix). The equation for a rate model is:

$$\tau \frac{dw}{dt} = -w(t) + w_{\max} - \tau c_2 r(t) w(t) \tag{2.5}$$

where $r(t)$ is the spike rate of the input whose weight is being considered, and c_2 is a constant. We assume the spike train is drawn from a Poisson process with mean rate $r(t)$, and demand that the average of Eq. 2.4 over Poisson realizations yield Eq. 2.5; this constraint determines the constant $c_2 = (1 - f)$ (see Appendix).

The structure of both rate and spiking models is as follows. There are synaptic weights connecting the LGN to the cortex – the geniculocortical (G) weights – and weights connecting the cortical cells to each other – both excitatory-to-excitatory (E) and inhibitory-to-excitatory (I) intracortical weights (figure 33). The intracortical connections instantiate the cat layer 4 circuit model proposed in (Troyer et al., 1998) and developed in chapter 1, as further described below. However, excitatory-to-inhibitory weights were shown to have little influence on the behavior of the mature circuit, and are omitted here. We also have not yet examined the influence of inhibitory-to-inhibitory weights in the mature model, but see the discussion in Troyer et al. (1998).

We examined synaptic depression in each of the three types of weights in the rate model, but only in the geniculocortical weights in the spiking model. In both models, weight values must be changed when depression parameters are changed (to maintain the network in a stable range, Troyer et al. (1998, figure 13)). Exploration of such parameter dependence is computationally expensive in the spiking model, so we have not yet explored intracortical depression in the spiking model.

Values of f and τ for each of the above weights classes (G, E, I) are beginning to be described physiologically; however, the experimental data do not yet converge to unique values of f and τ for any of these classes in the mammalian cortex. Two classes of experiment seem to give self-consistent but differing results: one class in which f and τ are determined by paired-pulse experiments, and another in which f and τ are determined by trains of stimuli (S. Nelson, personal communication). We examined each class of data (the "pulse" and "train" parameter sets) as separately as possible.

Parameters (and parameter sources) can be found in Table 2.

Table 2:	Paired-pulse data			Train data		
	f	τ, in ms	location	f	τ, in ms	location
G	0.563 [1]	99 [1]	LGN→IV	0.465 [4]	371 [4]	LGN→III
E	0.875 [2]	57 [2]	IV→IV	0.8 [5]	472 [5]	II/III→II/III
I	0.8 [3]	179 [3]	IV→IV	0.95 [5]	1017 [5]	II/III→II/III

1. Stratford et al. (1996, Fig. 1g) cat primary visual cortex
2. Tarczy-Hornoch (1996, Fig. 4.5) cat primary visual cortex
3. Tarczy-Hornoch et al. (1998, Fig. 4 (0.2Hz curve)) cat primary visual cortex
4. Gil et al. (1997, Fig. 3) mouse, rat somatosensory cortex
5. Song, Varela, Turrigiano, Abbott and Nelson (1999) rat primary visual cortex

Table 2: Depression Parameters Used. Parameters were derived by simple least-squares fits to data in the figures indicated; note that the τ value from the figure legend in Gil et al. (1997) refers to their exponential fit, *not* to the time constant of a fit to a synaptic depression description. Parameters from the random stimulus train experiments of Song, Varela, Turrigiano, Abbott and Nelson (1999) were taken directly as reported. Geniculocortical data divided readily into pulse (Stratford et al., 1996) and train (Gil et al., 1997) parameter sets. The corresponding intracortical data were then chosen, in the case of the pulse data, from work from the same laboratory (Tarczy-Hornoch, 1996; Tarczy-Hornoch et al., 1998); and in the case of the train data, from other work in the rodent that recorded both E and I depression curves (Song, Varela, Turrigiano, Abbott and Nelson, 1999). Note also that many of the f and τ values in this table do not describe connections within layer IV, the cortical layer we model in this paper. Where possible, when compiling this table we selected values determined (1) in layer IV (2) within primary visual cortex (3) in the cat.

2.3.2 Rate model

In the rate model, the LGN was structured as a 31 × 31, 6.8° × 6.8° retinotopic grid of cells, with retinotopic position varying linearly across the grid. ON cells were positioned at the vertices of the grid, while OFF cells lay at the center of each square within the grid; this offset is motivated by results from Wassle et al. (1981). Each geniculate cell can be thought to represent multiple LGN cells at each retinotopic location (this approximation will be discarded in the spiking model, in which 4 ON cells and 4 OFF cells, offset as above, will be positioned at 30 × 30 grid points). The choice of a 31 × 31 grid in the rate model, as opposed to the 30 × 30 grid in the spiking model, was arbitrary, and was originally made in the rate model solely to allow Gabor RFs to be centered over the single ON-cell at the center of the grid.

We examined 192 model cortical simple cells located at the single retinotopic position defined by the central LGN ON neuron. Of the 192 cortical cells, half were excitatory and half were inhibitory, with both sets of 96 representing each combination of 12 evenly-spaced orientations (at 8° to 173°,

to minimize grid discretization error) and 8 evenly-spaced spatial phases (0^o to 315^o). In this model there was no randomness to the geniculocortical weights; each was set to the value of the Gabor at the corresponding retinal position, where positive (negative) values of the Gabor correspond to weights from ON (OFF) inputs. Connections between cortical cells were correlation-based, as in Troyer et al. (1998): weights between a cell pair were determined by (1) computing the normalized cross-correlation between the cells' geniculocortical RFs, as defined in Troyer et al. (1998); (2) for an inhibitory connection, multiplying this value by -1; (3) setting negative values to zero; (4) raising the resulting value to some power (5, in this case), and assigning this new value as the weight; and (5) after all weights were assigned, normalizing excitatory or inhibitory weights to a given cell by the rms excitatory or inhibitory weight, respectively, to that cell (the weights were subsequently multiplied by an overall gain, described below). Thus, an excitatory cell made strong positive/excitatory weights onto excitatory cells that had RFs well-correlated with its own RF; an inhibitory cell made strong negative/inhibitory weights onto excitatory cells that had RFs well-anticorrelated with its own RF; and all other cell pairs shared weak or no connections. Combined with appropriate physiological constraints on excitatory cell outputs (see below), this connectivity scheme gives rise to a simple circuit (figure 33) that achieves contrast-invariant orientation tuning (Troyer et al., 1998) while obeying known experimental constraints such as the distribution of excitatory and inhibitory connections received by cortical simple cells (Ferster 1986, 1988) and the relative contribution of geniculocortical vs. intracortical weights (Ferster et al., 1996). The connections studied in this circuit can also be shown to arise naturally through activity-dependent development (Chapter 1).

Dynamically, neurons in the rate model obeyed the following equations. Let

$r_k^{I/E}(t)$ = the firing rate of inhibitory/excitatory cell k at time t
$a_k^{I/E}(t)$ = the activity (voltage) of inhibitory/excitatory cell k at time t
$\tau_m^{I/E}$ = the time constant of the inhibitory/excitatory cell membrane
$G_k(t)$ = the geniculocortical input to cell k at time t
$w_{kj}^{e \leftarrow i/e}(t)$ = the synaptic weight of the connection from inhibitory/excitatory cell j to excitatory cell k at time t
$\theta^{I/E}$ = the firing threshold for inhibitory/excitatory cells
$floor$ = a floor on the membrane voltage of the cells (see below)

The firing rate for excitatory or inhibitory cell k is equal to the thresholded membrane potential:

$$\begin{aligned} r_k^E(t) &= \left[a_k^E(t) - \theta^E\right]^+ \\ r_k^I(t) &= \left[a_k^I(t) - \theta^I\right]^+ \end{aligned}$$

where $[x]^+ = x,\ x > 0;\ = 0$, otherwise. An inhibitory cell k then updates its activity by integrating its geniculocortical input. (As described above, excitatory-to-inhibitory intracortical weights have little influence on the behavior of the model, and inhibitory-to-inhibitory weights are not included here - see Discussion, Troyer et al. (1998) - so all input to inhibitory cells is of geniculocortical origin.):

$$\tau_m^I \frac{da_k^I}{dt} = -a_k^I(t) + G_k(t)$$

For an excitatory cell k, the input arises from geniculocortical, inhibitory, and intracortical excitatory

sources; it is simply the sum of these respective contributions:

$$n_k(t) = G_k(t) - \sum_{j=1}^{N} w_{kj}^{e \leftarrow i}(t) r_j^I(t) + \sum_{j=1}^{N} w_{kj}^{e \leftarrow e}(t) r_j^E(t)$$

where N is the number of excitatory or inhibitory neurons. Finally, the activity update for the excitatory cell k is:

$$\tau_m^E \frac{da_k^E}{dt} = -a_k^E(t) + n_k(t), \; a_k^E > floor; \tag{2.6}$$

$$= \left[-a_k^E(t) + n_k(t)\right]^+, \; a_k^E = floor; \tag{2.7}$$

where "floor" is akin to the potassium reversal potential of the cell, as described in the next paragraph.

Outside of the (fixed) f and τ values for depression, the rate model consisted of 8 parameters: the membrane time constants, the firing thresholds, and the gains of geniculocortical (G), inhibitory ($I \rightarrow E$) and excitatory ($E \rightarrow E$) weights, as well as the voltage floor. The geniculocortical, inhibitory, and excitatory gains were scalars that provided an overall multiplicative scaling for each type (G, I, E) of weight. The voltage floor was included merely to represent the lower bound on the membrane voltage imposed in real neurons by the potassium reversal potential (*i.e.* the floor was the value below which any neuron's membrane potential was not allowed to go; if the membrane potential attempted to drop below the floor, it was clamped to the floor potential). This floor was somewhat arbitrarily set to -30, but the value of -30 was not critical; the behavior of the model was quantitatively similar for a floor value of -75, and only marginally different for a very "depolarized" floor value of -5.

Outputs of the model (excitatory cells only) were determined on a cell-by-cell basis and averaged across the appropriate cells. The seven parameters other than the floor were then determined by searches through the parameter space for cases that satisfied the following criteria:

1. $\tau_m^E > \tau_m^I$ (McCormick et al., 1985)
2. $\theta^E > \theta^I$ (McCormick et al., 1985)
3. Standard deviation of a Gaussian fit to the orientation tuning curve < 20 degrees at all contrasts
4. Invariance of orientation tuning width with contrast (Sclar and Freeman, 1982), defined as follows: a ratio of the standard deviation of a Gaussian fit to the orientation tuning curve at low (10%) and high (80%) contrast between 4:5 and 5:4.
5. "Amplification ratio" > 1 and < 5 for both 10% and 80% contrast preferred orientation sinusoidal gratings (defined as ratio of F1 of voltage response with full cortical circuitry intact to F1 of voltage response induced by geniculocortical inputs alone); these values are comparable to the limits suggested in (Ferster et al., 1996) for responses to 2Hz, 64% contrast drifting sinusoidal gratings at the preferred orientation. Responses to gratings of 12 Hz and 16 Hz temporal frequency were not required to match this criterion because of the cortical response attenuation expected from filtering by the membrane time constant at these frequencies.
6. Mean cortical firing rates between 10 and 30 Hz for preferred orientation stimulus at 80% contrast. As for the previous constraint, responses to gratings of 12 Hz and 16 Hz temporal frequency were not required to match this criterion.

Parameter searches were performed separately for each temporal frequency of stimulation; consequently, parameter sets that satisfied the experimental criteria for one temporal frequency did not necessarily satisfy those criteria for other temporal frequencies. For the c-d phase advance plot (figure 40), we show all parameter sets, but we also examined the effects of limiting our analysis to those parameter sets that gave constraint-satisfying responses across all three temporal frequencies tested. Temporal frequency and contrast saturation plots, on the other hand, show results for the same parameter sets across all temporal frequencies. For these plots, the constraints were enforced at each temporal frequency, except that F1 ratios, and mean cortical firing rates at high contrast, were not enforced for temporal frequencies greater than 8 Hz; and none of the constraints were enforced for the "flat" F1 value of 15 Hz. For the "no depression" case, the low bound on mean firing rate at high contrast was also relaxed slightly (to > 9.5 Hz) in order to allow generation of a contrast saturation curve (figure 44).

A number of different combinations of the seven parameters were examined. The range of permissible combinations of four of these parameters varied with the location(s) of depressing synapses; the relative strength of inhibition required to prevent cortical runaway, for instance, was much less when intracortical excitatory depression (E depression) was present. For cases in which E depression was present, we searched through all combinations of the following values for these four parameters:

1. $\theta^E = 2, 4, 6$
2. G gain $= 1.0, 2.0, 4.0, 8.0$
3. $I \rightarrow E$ gain $= 0.15, 0.25, 0.35, 0.45$
4. $E \rightarrow E$ gain $= 0.06, 0.09, 0.12, 0.15$

When E depression was absent, we instead searched through all combinations of the following values for these four parameters:

1. $\theta^E = 3, 6, 9$
2. G gain $= 0.5, 1.0, 2.0, 4.0$
3. $I \rightarrow E$ gain $= 0.25, 0.35, 0.45, 0.55$
4. $E \rightarrow E$ gain $= 0.02, 0.04, 0.06, 0.08$

In all cases, we searched through all combinations of the following values for the remaining three parameters:

5. $\tau_m^E = 8, 12, 16ms$
6. $\tau_m^I = \tau_m^E/2$
7. $\theta^I = 1, 2, 3$ with $\theta^I < \theta^E$

When E depression was present, the number of combinations amounted to 1344: 1728 ($4^3 \times 3^3$) $-$ 384 ($4^3 \times 3 \times 2$, the number of possibilities eliminated by the constraint $\theta^I < \theta^E$). When E depression was absent, only half as many combinations were eliminated by the constraint, and the number of combinations was 1536. Note that we biased our selection towards smaller membrane time constants. Reported *in vitro* values of 20ms for regular-spiking (presumably excitatory) and 12ms

for fast-spiking (presumably inhibitory) neurons (McCormick et al., 1985), and *in vivo* values in the absence of a stimulus of 15-24 msec for regular-spiking cells (Hirsch et al., 1998) almost certainly overestimate the *in vivo* values during stimulation; our selections thus reflected our assumption that the *in vitro* data do not capture significant conductances open in the stimulated *in vivo* state that would shorten the membrane integration time.

To obtain phase advance data, we subtracted the phase of the F1 of the cortical response to 10% contrast gratings from the phase of the F1 of the cortical response to 80% contrast gratings (all of which were shown at the optimal spatial frequency, and at an angle of 38°); we will refer to this phase difference as "contrast-dependent phase advance" (c-d phase advance). As most simulations were of 2s duration, phase analysis was performed on the last 500ms, when the geniculocortical and intracortical excitatory depressing synapses would have reached steady-state. [Intracortical inhibitory synapses fit to the train data (τ = 1017ms) would not have reached steady-state, but the influence of the inhibitory depression is weak. For similar parameters, 6-second runs in which the last 1000ms were analyzed showed small differences in phase advance relative to 2-second runs: 0.014° ± 0.25° (mean ± std deviation) for 2Hz inputs; 0.090° ± 1.00° for 4Hz inputs; 0.14° ± 1.57° for 8Hz inputs.]

The activity and depression equations were discretized using simple first-order Euler methods and 2 ms bins. This bin size did not significantly affect our results. For a case with similar parameters in which gratings at a temporal frequency of 2Hz were presented to a network with "pulse" depression in all weight types (G, E, and I), phase advance for cases satisfying the biological criteria at 2ms resolution differed by only $-0.0622 \pm 0.0447°$ from the same runs repeated at 0.25ms resolution.

2.3.3 Spiking model

The spiking model was implemented as in Troyer et al. (1998); we review the main points here. Retinotopic (retinal cell) positions were modeled as a 30 × 30 grid covering 6.8° × 6.8° of the visual field, with the positions of OFF-center cells offset by 0.11° in each direction relative to the positions of ON-center cells. This density follows the work of Peichl and Wassle (1979) for retinal cells at 5° eccentricity. There were 4 ON and OFF LGN X-cells at each retinotopic position; this value lies between empirical data from Sherman (1985) and Peters and Yilmaz (1993) and follows the assumption of Wörgötter and Koch (1991). Firing rates for each subtype of X-cell were determined by converting the rates used for the rate model into 25%-correlated Poisson spike trains (Alonso et al., 1996).

Cortical neurons were modeled as single-compartment, conductance-based integrate-and-fire neurons, using the NEURON simulation program (Hines and Carnevale, 1997). (NEURON was not used in Troyer et al. (1998), but the implementations are identical). The model excitatory and inhibitory cells were fit to data from regular-spiking cells and fast-spiking cells, respectively, by matching the model cells' spike-frequency vs. injected current (f-I) curves to the examples shown in McCormick et al. (1985). The parameters used to match the model cells to the data were C, the capacitance of a single compartment; V_{thresh} $(-52.5mV)$, the voltage which, when crossed, produced a spike; V_{reset}, the voltage to which the cell returned after spiking; $t_{refract}$, the refractory time after a spike has been elicited; and parameters describing two currents: a leak current (g_{leak} and V_{leak}), and a spike adaptation current (g_{adapt} and V_{adapt}) used only in excitatory neurons (Troyer and Miller, 1997a,b). For ex-

citatory cells, $C = 500pF$, $g_{leak} = 25nS$, $V_{leak} = -73.6mV$, $V_{reset} = -56.5mV$, $t_{refract} = 1.5msec$, $\bar{g}_{\text{adapt}} = 3nS$, $V_{adapt} = -90mV$; for inhibitory cells: $C = 214pF$, $g_{leak} = 18.0nS$, $V_{leak} = -81.6mV$, $V_{reset} = -57.8mV$, $t_{refract} = 1.0msec$. Each time-varying conductance was modeled as a difference of exponentials, $g(t) = \sum_{t_j<t} \bar{g} \left(e^{-(t-t_j)/\tau^{\text{fall}}} - e^{-(t-t_j)/\tau^{\text{rise}}} \right)$ where t_j corresponds to the times of each spike (presynaptic spikes for synaptic conductances, postsynaptic spikes for spike-rate adaptation current). The time constants were, for excitatory synapses, $\tau_{\text{ex}}^{\text{rise}} = 0.25\ msec$, $\tau_{\text{ex}}^{\text{fall}} = 1.75\ msec$; for inhibitory synapses, $\tau_{\text{in}}^{\text{rise}} = 0.75\ msec$, $\tau_{\text{in}}^{\text{fall}} = 5.25\ msec$; for spike-rate adaptation, $\tau_{\text{adapt}}^{\text{rise}} = 1\ msec$, $\tau_{\text{adapt}}^{\text{fall}} = 83.3\ msec$. Finally, all cells received background excitatory input at a rate of 5800 Hz with a conductance $\bar{g}_{\text{ex}}^{\text{bg}} = 0.89\ nS$. The conductances of intracortical ($\bar{g}_{\text{ex}}^{\text{ctx}}$, $\bar{g}_{\text{in}}$) and geniculocortical synapses ($\bar{g}_{\text{ex}}^{\text{gc}}$) are described below.

We examined 1600 excitatory and 400 inhibitory cells in a $2/3mm \times 2/3mm$ patch of layer IV of cat V1, representing $6.8° \times 6.8°$, 5° from the fovea. To determine geniculocortical receptive fields, we no longer used deterministic Gabor functions as in the rate model; rather, we chose synaptic strengths stochastically via repeated sampling from a Gabor function (Troyer et al., 1998). The shape and position of the Gabor function were determined by three parameters: preferred orientation, preferred spatial phase and retinotopic position. Preferred orientation was taken directly from an equal-sized optical map of cat V1 (Troyer et al., 1998, figure 36A; map provided by M.C. Crair and M.P. Stryker) while preferred spatial phase was assigned to cells randomly; retinotopic position was assumed to change linearly across the grid. The mean number of LGN inputs to a single cortical cell was 125 ± 8 (mean $\pm$ std).

The connections between cortical cells were determined as in the rate model, except that they were assigned stochastically rather than deterministically: the weights assigned in the rate model were instead taken as a probability distribution, and weights were assigned by repeated sampling from this distribution, followed by normalization of each synapse type on each cell as in the rate model; full details are as in the computational model of Troyer et al. (1998). Again, excitatory cortical connections occurred between cells with correlated geniculocortical receptive fields, whereas connections between inhibitory and excitatory cells occurred between cells with anti-correlated geniculocortical receptive fields. The mean number of intracortical excitatory and inhibitory synapses onto a single cortical cell was 107.4 ± 29.6 and 26.5 ± 8.6, respectively (80% from excitatory cells, 20% from inhibitory cells).

A major determinant of the function of the model was the relative synaptic strength of geniculocortical and intracortical synapses. These values were determined, after sampling and normalizing the synaptic weights as just described, by multiplying all synaptic weights of a given type (geniculocortical, intracortical excitatory, intracortical inhibitory) by a single constant to set the total strength of such synapses. These values were chosen to constrain the standard deviation of the orientation tuning curve to be < 20 degrees at all contrasts, as in the rate model, and to ensure contrast invariance at all temporal frequencies. Synaptic strength is defined in terms of the integrated current response induced when voltage clamping the cell at V_{thresh} and giving one simultaneous stimulus to all the synapses of a given type. Each excitatory cell received a total inhibitory synaptic strength of 14.726 nA msec, and a total intracortical excitatory synaptic strength of 3.112 nA msec, yielding mean unitary conductance values of $\bar{g}_{\text{in}} = 7.59nS$ and $\bar{g}_{\text{ex}}^{\text{ctx}} = 0.37$ nS. We used three separate values for the geniculocortical synaptic strengths, depending upon the parameters used for geniculocortical

depression. In the "no depression" case, the geniculocortical excitation onto each cortical cell was 3.112 nA msec, with $\bar{g}_{ex}^{gc} = 0.32$ nS. When the "pulse" parameters were used, the geniculocortical synaptic strength increased to 8.86 nA msec, with a mean unitary conductance of $\bar{g}_{ex}^{gc} = 0.92$ nS. When "train" parameters were used, geniculocortical synaptic strength was 26.45 nA msec, with a mean unitary conductance of $\bar{g}_{ex}^{gc} = 2.7$ nS. Note that we held total inhibition fixed, although we could have reduced this value when depression was present (because depression attenuates the untuned (DC) component of the geniculocortical input). Since total inhibition is a free parameter, and reducing (increasing) inhibition broadens (tightens) both orientation and temporal frequency tuning, we have some freedom to control these values, and yet remain within the experimental constraints.

The results presented here for the spiking model show model responses to drifting gratings at 105°. After a 500 msec "blank stimulus", during which time the cortical and LGN cells fired at background rates, a moving grating stimulus was presented for one second. Phase advances were then calculated as in the rate model, but using histograms built from the response of neurons to ten repeated stimulus presentations. Because each cell fired with a distinct phase in response to the gratings, we compared the difference in the phase of the response to 80% and 10% contrast gratings on a cell-by-cell basis, for all excitatory neurons with preferred orientation in the 5°-wide bin around 105° (preferred orientations 102.5° through 107.4°); there were 29 such excitatory neurons for the orientation map used. F1s for cortical responses were determined via Fourier transform of the last 500ms of each 1-second grating presentation.

2.4 Results

We have established in previous work (Troyer et al., 1998) that a cortical circuit with correlation-based intracortical connections, including spatially opponent, push-pull inhibition (figure 33), can account for a number of experimental findings in layer 4 of primary visual cortex in the cat, including contrast-invariant orientation tuning. In this paper we add synaptic depression to the circuit, and explore the effects of depression, spike-rate adaptation, and our model circuit on contrast-dependent nonlinear response properties in primary visual cortex.

2.4.1 Modeling Framework

We begin by summarizing the essential information about our model needed to understand our results.

Intracortical Circuit

We study a circuit (Troyer et al., 1998) in which (1) geniculocortical synaptic weights to a cell are described by Gabor functions, with ON-center (OFF-center) inputs corresponding to positive (negative) portions of the Gabor; and (2) intracortical connections are made between cortical cells based on the correlations between their RFs, *i.e.* between the geniculocortical synaptic weights they receive. An excitatory cell makes strong connections onto other excitatory cells with which it is strongly correlated; an inhibitory cell makes strong connections onto excitatory cells with which it is strongly anticorrelated. The dominant resulting connections follow a "push-pull" scheme, and are illustrated in figure 33. Two excitatory neurons whose ON and OFF subregions are aligned in visual space ("same-phase RFs", with similar preferred orientations) have strong positive, reciprocal connections; while an excitatory neuron whose ON (OFF) subregions overlap an inhibitory neuron's OFF (ON) subregions in visual space ("anti-phase RFs", with similar preferred orientations) receives a strong negative synaptic weight from that inhibitory neuron. Cell pairs with weaker RF alignment or anti-alignment have weaker or no connections. A crucial requirement is that inhibition be dominant: the pathway LGN→I→E must have stronger overall gain than the pathway LGN→E (where E and I indicate excitatory and inhibitory cortical cells, respectively), as assessed by mean feedforward inhibition exceeding mean feedforward excitation for a stimulus with orientation orthogonal to a cell's preferred orientation.

This architecture can account for cortical orientation tuning and its contrast invariance (Troyer et al., 1998). How does this circuitry account for orientation tuning? For a stimulus at a cell's preferred orientation and spatial phase, other neurons with similar preferred orientation and spatial phase – both excitatory and inhibitory – are strongly activated. However, the inhibition is directed onto cells with similar preferred orientation but antiphase (opposite spatial phase) RFs. In the case of a drifting sinusoidal grating of the preferred orientation, the resulting inhibition received by a cell comes *out-of-phase* with its excitation, permitting excitatory cells to respond during the temporal phase in which more excitation is received than inhibition (figure 33A). As the orientation is shifted away from preferred, temporal modulation of both feedforward excitation and feedforward inhibition decreases. Since inhibition is dominant in the mean, at some orientation the modulation is small enough that inhibition is dominant at all times, and the cell cannot fire. In particular, for a stimulus

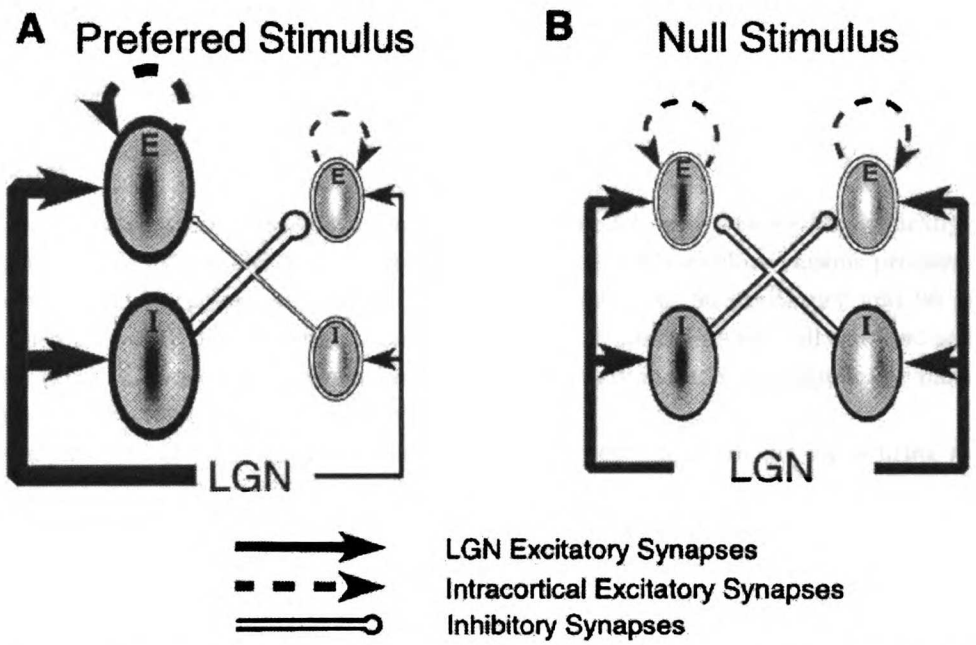


Figure 33. Cartoon of the cortical circuit studied. All neurons receive excitatory geniculocortical connections from the LGN as determined by Gabor functions (illustrated by modulations on a gray/uniform background): ON inputs at a given retinal location within the RF are represented by black, OFF inputs by white. All illustrated RFs are centered at a common retinotopic position. Neurons with RFs of similar preferred orientation but opposite spatial phase are connected by inhibitory synaptic weights (white, with black outline), while neurons with similar preferred orientations and similar spatial phases are connected by excitatory synaptic weights (black arrows). **A**. Response to a full-field sinusoidal grating of the preferred orientation. When the stimulus maximally overlaps the RFs on the left, the geniculocortical input to those cells is maximal (large solid-black arrows), while the input to the RFs of opposite spatial phase (those on the right of **A**) is minimal (small solid-black arrows). Neurons of the well-stimulated phase will fire robustly, and the strongly-activated inhibitory cells send inhibition only to the weakly-stimulated anti-phase excitatory neurons, which do not fire. As a result, as the grating moves across the neurons' RFs, the excitatory cortical neurons will produce a strongly time-varying response at the same temporal frequency as that of the input. **B**. Response to a full-field sinusoidal grating of the null orientation. Because LGN cells respond to all orientations, the geniculocortical input is still present, but the input to each phase is approximately equal. Inhibition is equally strong from neurons of each phase to their anti-phase excitatory-cell partners. Since inhibition is dominant, none of the excitatory cells fire. The actual circuit studied included cells of many preferred orientations and spatial phases and, for the spiking model, many retinotopic positions. Connections were based on correlations between RFs. Cartoon illustrates dominant connections; resulting circuit behavior can be well understood from this simplified version of the circuit (Troyer et al., 1998).

at a cell's null orientation (perpendicular to the preferred), inhibitory neurons of both spatial phases are simultaneously activated. In this case, the fact that inhibition is both strong and anti-phase leads excitatory cells of both phases to be *simultaneously* inhibited (figure 33B).

The contrast-invariance of orientation tuning arises because an increase in contrast equally increases the geniculocortical drive to a given cell and to the anti-phase cells from which it receives inhibition. Thus, the basic argument just given for orientation tuning applies invariantly at all contrasts. A more detailed analysis is given in Troyer et al. (1998).

Rate Model

We studied two forms of model: a conceptual rate model, and a more biophysically-accurate spiking model. The rate model was developed in order to quickly explore various properties of the cortical circuit and its circuit elements. The model consisted of 96 excitatory and 96 inhibitory neurons, with RFs of 12 different orientations and 8 different spatial phases, all centered at the same retinotopic point. Connections between cortical neurons were made deterministically based on the correlation between their RFs, as described above.

Model neuron firing rates were calculated as the weighted sum of all the input firing rates from geniculocortical, intracortical excitatory, and intracortical inhibitory sources, rectified at a threshold – hence the term, "rate model". The model was described by eight parameters: the thresholds and membrane time constants of excitatory and inhibitory cells, the gains of geniculocortical (G), intracortical inhibitory-to-excitatory (I), and excitatory-to-excitatory (E) cell connections, and a lower bound on the membrane voltage. Appropriate values for these eight variables were obtained by constraining the output of the circuit to match a set of experimental findings, including the width and contrast-invariance of orientation tuning; this set did not include the nonlinear response properties studied here (see Methods).

Spiking Model

To expand upon the insights obtained from the rate model in a more biophysically realistic framework, we used the spiking model of Troyer et al. (1998). 1600 excitatory and 400 inhibitory neurons were laid out in a $2/3mm \times 2/3mm$ cortical grid, with retinotopic position constrained to move smoothly across the grid, and with orientations determined by an experimentally measured map from cat V1. The spatial phase of each RF (which determines the location of its ON and OFF subregions) was chosen randomly. Connections between cortical cells were then made probabilistically based on the correlation between the RFs. All neurons were conductance-based integrate-and-fire cells, matched to data from McCormick et al. (1985) as explained in (Troyer and Miller, 1997a,b). Excitatory neurons had spike-rate adaptation currents. We included only fast (AMPA and GABA-A) synaptic currents, deferring examination of slow currents (*e.g.*, NMDA and GABA-B) to future work (*e.g.* Krukowski et al., 1998). Again, parameters were chosen to achieve appropriately narrow, contrast-invariant orientation tuning, and nonlinear response properties were then studied (see Methods).

Visual Stimuli and LGN Inputs

Visual inputs to the models were drifting full-field sinusoidal gratings. LGN responses were assumed to arise from a spike rate that was the sum of a linear stimulus-induced temporal modulation and a constant background rate, with rates rectified at zero. Amplitudes of the stimulus modulation were matched to LGN data on responses across contrast and temporal and spatial frequency (Sclar, 1987), as described in the Methods. The rate model used this rate directly as the LGN response, while the spiking model used Poisson spike trains sampled from these rates.

The geniculocortical synaptic weights to the simple cells in the model layer 4 were described by Gabor functions (see Methods), with parameters matched to experimental measurements of simple cell RFs. In the rate model, the geniculocortical (G) weights were defined deterministically by the Gabor distribution, with negative Gabor values indicating OFF weights; the spiking model RFs were established probabilistically by sampling from the Gabor distribution.

Synaptic Depression

Synaptic depression is a use-dependent decrease in synaptic efficacy (Abbott et al., 1997; Markram and Tsodyks, 1996); as the firing rate of a presynaptic neuron increases, the influence of single synapses from that cell onto the postsynaptic neuron declines (figure 34). Intuitively, this relationship holds because higher firing frequencies prevent recovery from depression between input spikes, as discussed below.

One can characterize synaptic depression by two parameters: f, the fractional synaptic weight change after a spike ($0 \leq f \leq 1$), and τ, the time constant of recovery from depression. Smaller values for f lead to a greater loss of synaptic strength after every spike; smaller values of τ cause faster recovery from this depression. In both the rate and spiking models, like forms of depression are used: the rate-model depression equation is equal to the average, over Poisson-sampled spike trains, of the spiking-model depression equation (see Appendix); and their behavior in simulations is qualitatively and quantitatively quite similar (figure 39).

In the experimental literature, two classes of data appear to be present: one in which synaptic depression is studied through the use of paired-pulse stimuli, and one in which depression is characterized by probing with trains of stimuli (S. Nelson, personal communication). These two types of experiment result in different measured values for f and τ, which we call the "pulse" and "train" parameters, respectively (see Table 2). Given this experimental uncertainty in parameter values, we examine all results under both choices of parameters.

2.4.2 Experimental Findings Addressed

Contrast-Dependent Phase Advance

Simple cells respond earlier in time to drifting gratings as the contrast of those gratings increases, as quantified by the difference in the phase of the first harmonic (F1) of the cortical spiking responses at each contrast (Dean and Tolhurst, 1986; Albrecht, 1995). We reviewed the literature to determine the size of this contrast-dependent (c-d) phase advance (figure 35). We examined both V1 and LGN c-d phase advance, because only the difference between these values needs to be accounted for by cortical mechanisms. In all cases we report the advance over three octaves of contrast (*e.g.* the

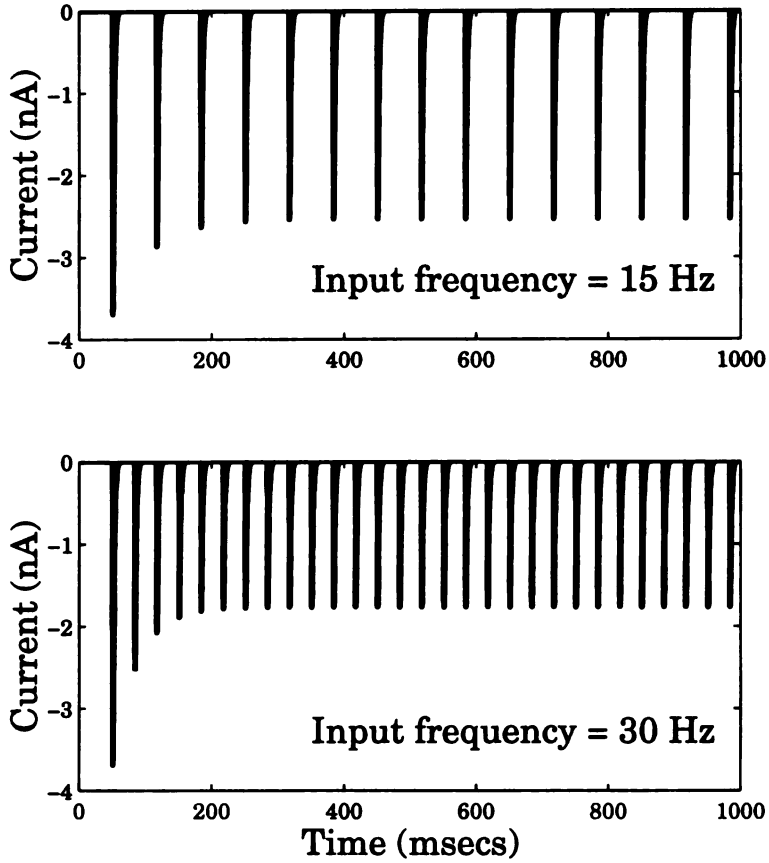


Figure 34. Amount of depression varies with input rate Geniculocortical current received by a cortical neuron in the spiking model; in both cases, $f = 0.563$ and $\tau = 99msec$ ("pulse" parameters). The cortical neuron is voltage-clamped at the spike-threshold, and all 120 LGN inputs to the cell are activated synchronously at the indicated rates; each current "spike" represents the arrival of input spikes. The initial current response is $-3.6875nA$ in both cases. In **A** (input rate of 15 Hz), steady state current per cycle is $-2.5259nA$; in **B** (input rate of 30 Hz), steady state current per cycle is $-1.7655nA$.

relative advance between 10% and 80% contrast).

For V1 simple cells in the cat, c-d phase advance has been measured for approximately 30 cells (Dean and Tolhurst, 1986) in one study, and for over 100 cells in another (Albrecht, 1995). Mean c-d phase advances were comparable: 42° for a 2 Hz grating in the former study, 47° and 49° for 2Hz and 8Hz gratings, respectively, in the latter. In the LGN, X cells show 25° mean c-d phase advance in response to 8Hz (Sclar, 1987) and 3Hz (Saul and Humphrey, 1990) gratings, while Y cells demonstrate as much or more c-d phase advance as cortical simple cells. Both the LGN and cortical measurements are characterized by large standard deviations. Without a knowledge of the X or Y nature of the geniculocortical inputs to the cortical cells studied previously, it is difficult to know how much c-d phase advance the cortex must add, or even whether it adds any at all. An additional uncertainty is raised by the fact that we are modeling layer 4, where the first transformation of LGN inputs occurs. Further cortical transformations could add more c-d phase advance, so layer 4 might show less c-d phase advance than the cortical mean; however, the data on cortical cells were not broken down by layers.

We make perhaps the simplest assumption: that cortical layer 4 should account for the mean difference in c-d phase advance between X cells and V1 simple cells. Thus, we assume that layer 4 must account for roughly 20 degrees of c-d phase advance over 3 octaves of contrast. Note that we

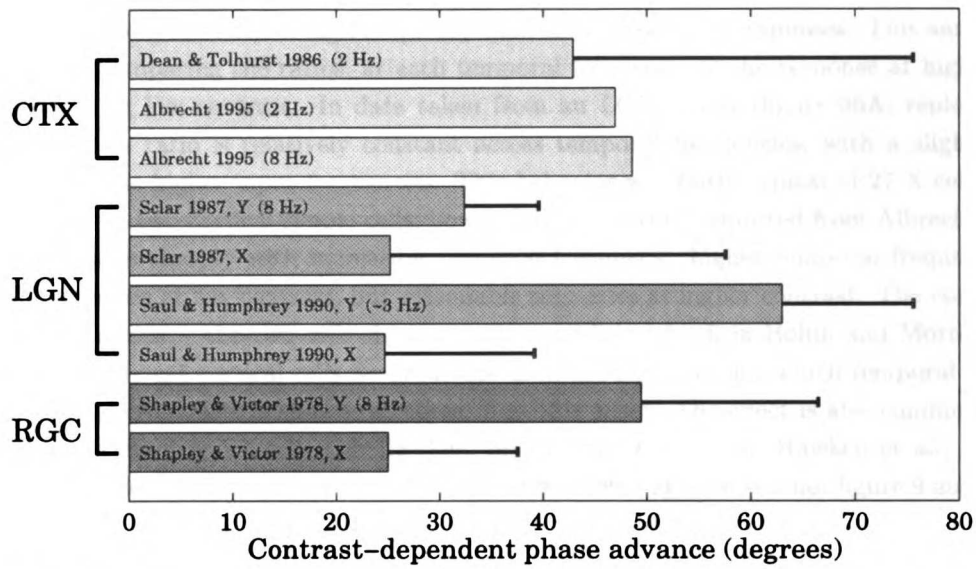


Figure 35. Experimentally-determined values for c-d phase advance in three parts of the visual pathway: retinal ganglion cells (RGC, bottom, in dark gray), lateral geniculate nucleus (LGN, middle, in gray), and striate cortex (CTX, top, in light gray). C-d phase advance is quantified here as the relative phase difference between responses to stimuli differing by 3 octaves of contrast. The contrasts below are Michelson contrasts ($(I_{max}-I_{min})/2*I_{mean}$). Data represented are all from cat and are as follows: (1) Dean and Tolhurst (1986): responses to 5% and 25% contrast drifting gratings for 29 V1 simple cells. We linearly extrapolated, from 2.3 octaves of contrast to 3, the reported mean and standard deviation of c-d phase advance. (2) Albrecht (1995) (taken from Discussion of that paper): 2 Hz: responses to 5% and 25% contrast drifting gratings, linearly extrapolated to 3 octaves. 8Hz: responses to drifting gratings at 10% and 80% contrast (note, phase advance at 8Hz between 3.5% and 28.3% contrast was 33% larger). All data for V1 simple cells; standard deviations and number of cells were not reported. (3) Sclar (1987): responses to 10% and 80% contrast drifting gratings for 27 X and 51 Y cells. (4) Saul and Humphrey (1990): responses to drifting gratings of optimal temporal frequencies for 19 non-lagged X and 8 non-lagged Y cells over a range of contrasts (0.0025% to 96%). Their linear fits to phases of suprathreshold responses provided slopes with accompanying standard deviations (both in cycles of phase per octave of contrast), which we multiplied by 3 (converted to degrees) to obtain changes over 3 octaves. (5) Shapley and Victor (1978): 3.5% and 28.3% contrast (2.5% and 20% RMS contrast) for 8 X and 18 Y cells. Responses to counterphase gratings including 6-8 different temporal frequencies with total contrast as indicated; phase advance of 8 Hz component was determined. In those papers in which phase advance was determined for both X and Y cells, the same temporal frequency was used for each data set; in the figure, this frequency is indicated in the "Y-cell" bar only. We were guided through this data by the lucid discussion of Albrecht (1995).

do not include LGN c-d phase advance in our simulations, so the simulations should be compared only to this *difference* between experimentally observed LGN and V1 c-d phase advance.

Contrast-Dependent Changes in Temporal Frequency Tuning

In response to an increase in stimulus contrast, cortical temporal-frequency tuning curves can show a relatively greater amplification of high frequency than of low frequency responses. This amplification is measured by comparing the ratios, at each temporal frequency, of the response at high contrast to the response at low contrast. In data taken from an LGN X cell (figure 36A; replotted from Sclar (1987)), this ratio is relatively constant across temporal frequencies, with a slightly larger amplification of the higher input frequencies. This behavior was fairly typical of 27 X cells studied in Sclar (1987). In two cortical simple cells, however (Figs. 36B,C; replotted from Albrecht (1995)), this ratio increases sharply with increasing temporal frequency: higher temporal frequencies give very small responses at low contrast, but reasonable responses at higher contrast. The cortical data for cats is very sparse – the two cells shown, and one additional cell in Holub and Morton-Gibson (1981), are the only cat cortical cells we are aware of in the literature for which temporal frequency tuning at multiple contrasts is shown; all three show this effect. The effect is also common, though not universal, in monkey V1 cells (M. Hawken, private communication; Hawken et al., 1992; of 3 published tuning curves, effect is seen in Carandini et al., 1997, figure 6 but not figure 9 and not seen in Albrecht, 1995, figure 11), suggesting that preferential amplification of high temporal frequencies may be a common V1 property. However, there is no data as to whether, or how strongly, this effect is seen in layer 4 neurons. Moreover, Y cells show a more pronounced amplification of high frequency responses (Sclar, 1987) than do X cells. Just as for c-d phase advance, without knowledge of the relative X and Y cell input to studied simple cells, it is unclear how much amplification, if any, is accomplished by the cortex. We again make the assumption that the cortex must account for the difference in response between X cells and V1 simple cells. Lastly, these data also suggest, as does one published cell in monkeys (Carandini et al., 1997, figure 6), that increases in contrast might shift the peak of the temporal frequency response curve to higher frequencies, in addition to relatively amplifying such frequencies.

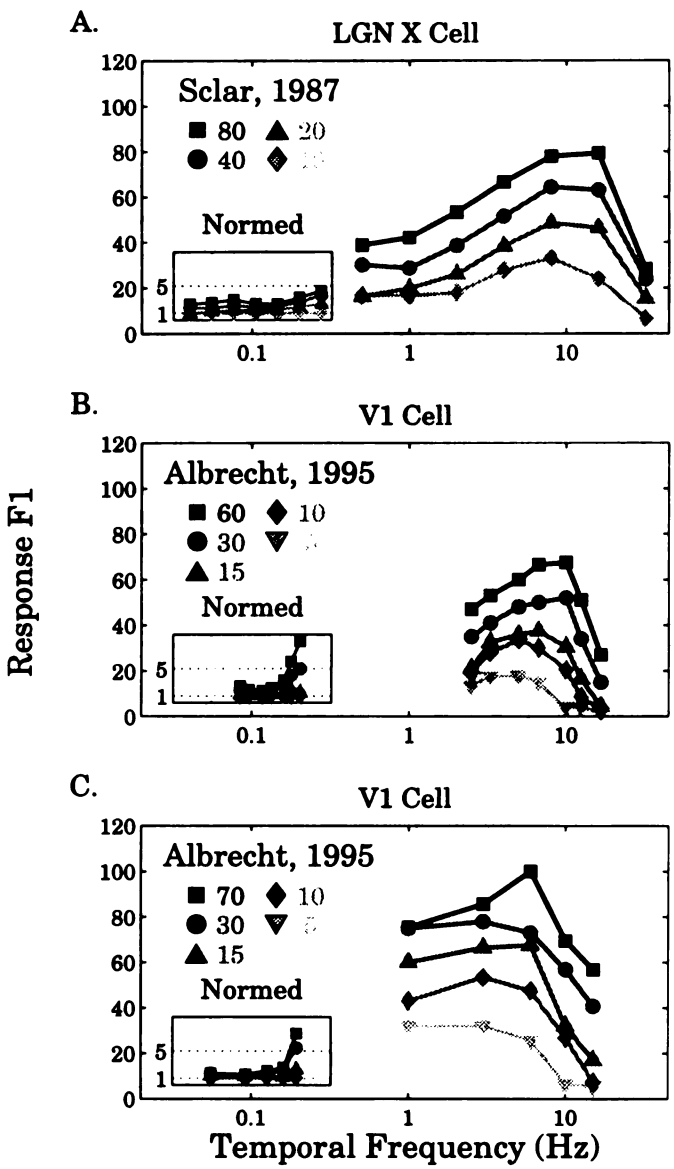


Figure 36. Contrast enhances responses to higher temporal frequencies in V1, but not LGN. Experimentally-determined F1 responses at different temporal frequencies and contrasts for one LGN X-cell and two V1 cells. Lighter grays to darker grays, with corresponding symbols: increasing contrast, with values noted in legend for each figure. To show the relative increase in high-frequency responses with increasing contrast, we plot normalized data in the insets: each response is divided by the response at the corresponding frequency and 10% contrast (so that all 10% responses are normalized to 1). The 5% V1-cell curves are omitted from the normalized data; we normalize by 10% to better compare to our model curves in figure 42 (for which the low contrast is 10%). Dashed lines within the inset indicate ratios of 1 (lower line) and 5 (upper line). **A.** LGN X cell temporal frequency response; raw data replotted from Sclar (1987). **B.** V1 cell's temporal frequency response; raw data replotted from Albrecht (1995). **C.** A second V1 cell's temporal frequency response; raw data replotted from Albrecht (1995).

Saturation of Responses With Increasing Contrast

Simple-cell responses tend to reach a plateau with increasing stimulus contrast (figure 37B,C); this is known as contrast saturation. This cannot be explained by intrinsic saturation of the cell's ability to fire. As evidenced, for example, by the contrast-invariance of orientation tuning, saturation does not occur at a fixed response level, but rather at different response levels for different stimuli (*e.g.*, so that orientation tuning curves are similar in shape at saturating contrasts and at low contrasts). LGN inputs show contrast saturation as well (figure 37A). If LGN input firing does not change with increasing contrast, neither will cortical firing; thus, a cortical explanation for contrast saturation is needed only to the extent that cortical cells saturate at lower contrasts than their LGN inputs.

While the LGN X cell in figure 37 indeed saturates at higher contrasts than the cortical cells in that figure, it is not clear whether this is a general phenomenon. Contrast saturation can be measured by a parameter C_{50}: the contrast at which response is half of the maximal, saturating response (determined from a fit of the Naka-Rushton equation, Eq. 2.1, to the contrast-response curve). In Table 3, we show the value of C_{50} for the cells of figure 37 and for 5 additional LGN X cells for which we found contrast response curves in the literature, along with the mean value reported for over 100 cat cortical simple cells in Albrecht (1995). From these values, it is not obvious whether additional mechanisms are required to explain cortical contrast saturation in cats beyond LGN contrast saturation. The same uncertainty applies in monkey, where V1 cells saturate over a range of contrasts similar to the combined saturation ranges of magnocellular and parvocellular LGN cells (Sclar et al., 1990).

Experimental C_{50} Values

Cell	Temporal Frequency (Hz)							
	1	2/2.5	3/3.3	4/5	6/6.7	8/10	12/12.5	15/16/16.7
LGN (1)	> 100	> 100		> 100	35	23.0	21.9	21.8
LGN (2)			12.7					
LGN (3)			14.9					
LGN (4)			6.7					
LGN (5)			8.2					
LGN (6)				5.7				
V1(1)		> 100	30.0	16.7	21.6	18.5	28.9	39.1
V1(2)	7.7		7.4		13.1	16.2		22.7
V1 mean	15.5							

Table 3. C_{50} values from Naka-Rushton curves (Eq. 2.1) fit to experimental data. LGN (1), V1(1), V1(2): cells of figure 37A,B,C respectively. LGN (2–4): cells from Cheng et al. (1995); LGN (5): cell from Chino et al. (1994). LGN (6): cell from Kaplan et al. (1987). V1 mean: mean from over 100 cat simple cells, each at or near its optimal temporal frequency, reported in Albrecht (1995). LGN (1): temporal frequencies (TF's) 1, 2, 4, 8, 16 Hz. LGN (2–5): TF 3.1 Hz. LGN (6): TF 4 Hz. V1(1): TF's 2.5, 3.3, 5, 6.7, 10, 12.5, 16.7 Hz. V1(2): TF's 1, 3, 6, 10, 15 Hz. For individual cells, we performed least-squares fits of Naka-Rushton curves to published contrast-response data. Best-fit value greater than 100 indicate that the response did not show saturation over the measured contrasts.

Although the need for a cortical explanation of saturation is unclear, we will address here the degree to which the mechanisms we study can contribute to such saturation. We address this for two reasons: (1) Such a cortical explanation may be needed elsewhere if not in V1. For example, in monkey MT, cells clearly saturate at much lower contrasts than either LGN or V1 cells (Sclar et al., 1990); and (2) Contrast saturation has provided a central motivation for formulating other cortical models (Albrecht and Geisler, 1991; Heeger, 1992; Carandini et al., 1998), and so it seems important to address alternative mechanisms even while the issue remains unsettled experimentally.

Lastly, data on contrast saturation suggest an additional point: simple cell responses saturate at higher contrasts as temporal frequency increases. This effect was noted by Albrecht (1995) in discussing the two cells shown in figure 37, and is shown particularly prominently by the cell of figure 37C. Similar findings have been noted in monkeys (Carandini et al., 1997).

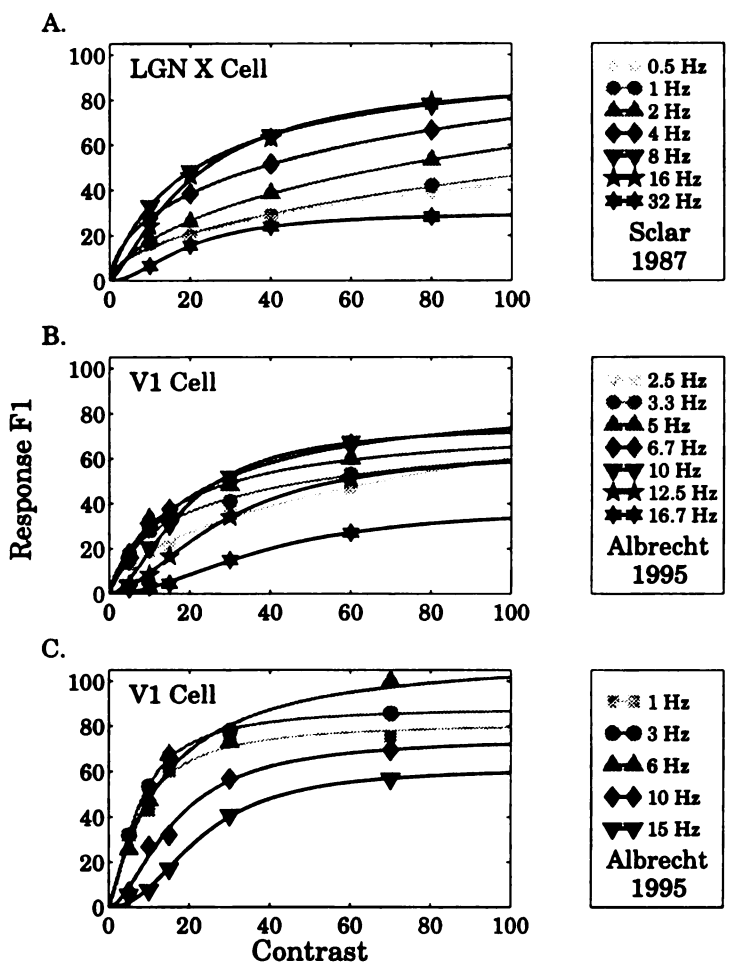


Figure 37. Experimental data for the temporal frequency dependence of contrast saturation. Experimentally-determined F1 responses at different temporal frequencies and contrasts for one LGN X-cell and two V1 cells. Figures are replotted from the data shown in figure 36, and fitted to Naka-Rushton curves (see Methods). Lighter grays to darker grays: increasing contrast. **A.** LGN X-cell contrast saturation response; raw data replotted from Sclar (1987). **B.** V1 cell's contrast saturation response; raw data replotted from Albrecht (1995). **C.** A second V1 cell's contrast saturation response; raw data replotted from Albrecht (1995). Note that (for the most part – see text and Table 3) the responses for both V1 cells saturate at lower contrasts than do the LGN responses, and that cortical responses to higher temporal frequency inputs saturate at higher contrasts.

2.4.3 Model Findings

Contrast-Dependent Phase Advance

At least three mechanisms can contribute to additional cortical c-d phase advance beyond that of the LGN inputs: synaptic depression, spike-rate adaptation, and contrast-dependent increases in conductance. Synaptic depression is evoked by the presynaptic spiking response to the grating stimulus, and differentially suppresses the later portions of the input, and thus of the postsynaptic response, over each stimulus cycle. As illustrated in figure 38, this shifts the response peak forward in time. Because the effect of synaptic depression grows with presynaptic firing rate, and thus with contrast, this shift increases with stimulus contrast, yielding a c-d phase advance. Spike-rate adaptation is evoked by postsynaptic rather than presynaptic spiking response, but otherwise it causes c-d phase advance for the same reasons as synaptic depression. Finally, as emphasized in studies of the normalization model (*e.g.*, Carandini et al., 1998), increases in postsynaptic conductance cause a decrease in membrane time constant, and this decrease in integration time causes the phase of responses to advance. If conductance grows with stimulus contrast, this also yields a c-d phase advance.

We first examined the role of synaptic depression. We began by studying the effects of the depression parameters, f (the loss of synaptic strength after each presynaptic action potential) and τ (the time constant of depression) (figure 39). We restricted study for these parameter searches to the geniculocortical synapses: we examined the c-d phase advance of the total geniculocortical input to simple cells in response to optimally oriented spatial gratings drifting at three temporal frequencies. Synaptic depression yields c-d phase advances of 5-10 degrees across a broad range of parameters.

The parameter dependence of this effect can be understood as follows. A smaller f, representing stronger depression, induces stronger c-d phase advance, up to a point. Once f becomes small enough that the synaptic weights are close to zero within the stimulus cycle at some contrast, further increases in constrast have less and less additional effect, so too great a reduction in f can decrease the c-d phase advance (figure 39a). Smaller τ yields greater recovery from depression between spikes, hence less depression and less c-d phase advance. As τ increases, the depression becomes stronger and the phase advance increases, until τ becomes comparable to the period of the stimulus cycle. At this point, τ is preventing synaptic weight recovery between response cycles. Further increases in τ have little effect on c-d phase advance: such increases change the dynamic range over a cycle, lowering the mean synaptic weight and mean response, but do not seem to appreciably alter the time course of depression and recovery within that dynamic range (of course, as $\tau \to \infty$, the steady-state response level will go to zero, and c-d phase advance will become undefined). Finally, an increase in temporal frequency is roughly equivalent to moving the graphs down and to the left: at higher temporal frequency, there is less time in each cycle for depression to occur, so a larger f is needed to get an equivalent amount of depression; and, there is less time in each cycle to recover from depression, so a smaller τ gives an equivalent amount of recovery.

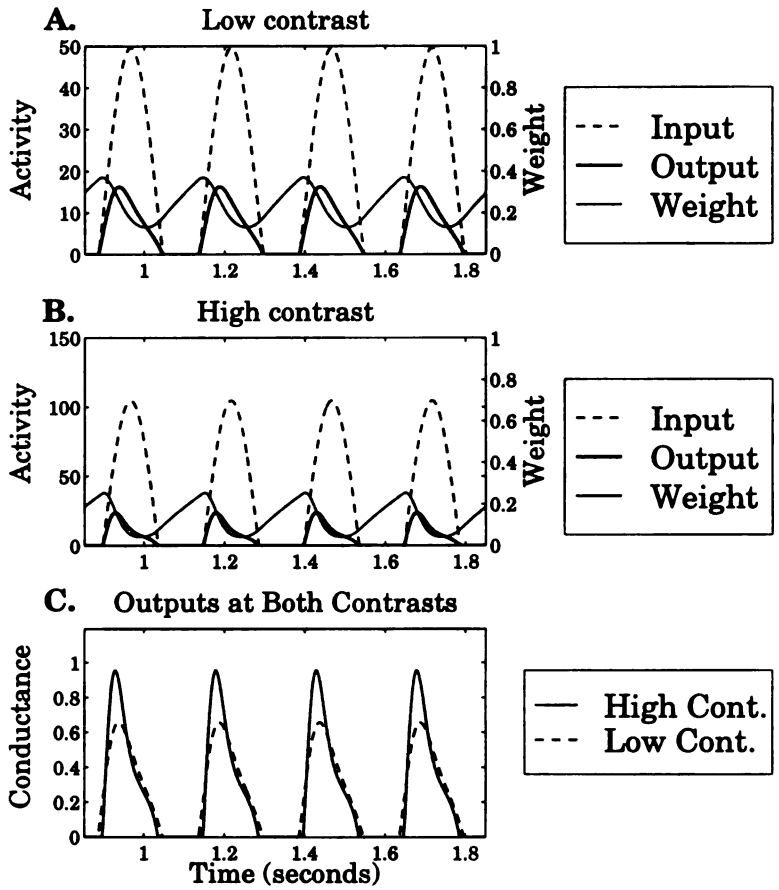


Figure 38. Geniculocortical synaptic depression induces both an absolute and a relative phase advance. **A, B**: Steady-state responses to a drifting sinusoidal grating at a cortical cell's preferred orientation and spatial frequency. Dashed lines show firing rate, noted on the ordinate, of a single LGN ON-cell input to the cell; thin solid lines, strength of the synaptic weight from that LGN input to the cortical cell, normalized by the weight's maximum value; thick solid lines, conductance contributed by that LGN ON input to the cortical cell, scaled by an arbitrary factor for display purposes (but maintaining the relative difference between low and high contrast conductances across figures). The weight depresses as input rate increases, and recovers after input rate declines. Consequently, the peak of the conductance curve, the product of the weight times the rate, shifts forward in time relative to the input. This shift in the peak correlates well with the absolute phase advance. The cell's output, which in the absence of intracortical connections is just the LGN conductance temporally filtered by the cell's time constant (and rectified), will show a phase advance similar to that of this single conductance. **A.** Responses to a low contrast (10%) stimulus. **B.** Responses to a high contrast (80%) stimulus. **C.** Comparison of the conductances (now shown unscaled, and measured in Hz) induced by this particular connection at low and high contrast. The steeper and stronger synaptic depression at higher contrast leads to an earlier peak of cortical response in each cycle and thus to a greater phase advance. Depression parameters: $f = 0.465$, $\tau = 371$ ("train" parameters).

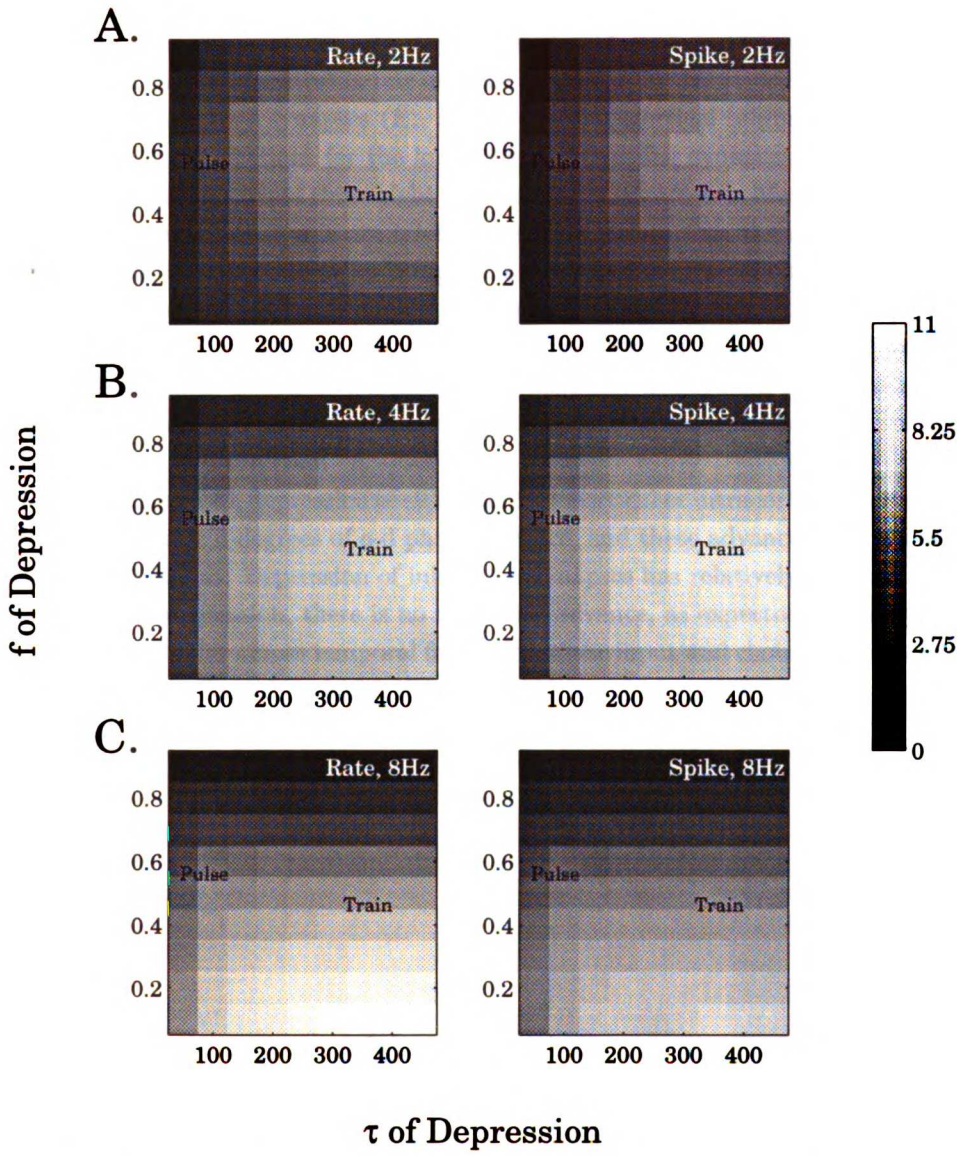


Figure 39. The dependence of contrast-dependent geniculocortical phase advance on f and τ in the rate and spiking models, shown for drifting gratings of (**A**) 2Hz, (**B**) 4Hz, and (**C**) 8Hz temporal frequency. In this figure, response is simply the summed geniculocortical input to simple cells, ignoring cortical integration. Black indicates less phase advance, white indicates more phase advance. The experimental values of the geniculocortical f and τ parameters for the pulse and train data sets (Table 2) are marked by the words "Pulse" and "Train." Gratings were of optimal orientation and spatial frequency; mean c-d phase advance across simple cells of multiple spatial phases is shown. For the rate model (left panels; label in upper right-hand corner of plot), c-d phase advance is measured for the sum over a cell's geniculocortical inputs of synaptic weight times firing rate. For the spiking model (right panels), c-d phase advance is measured for the net current due to geniculocortical inputs, which spike according to a Poisson sampling of firing rates. The amount of geniculocortical c-d phase advance is similar, but not exactly equal, for the two models. Differences presumably are due to the particular Poisson spike trains sampled in the spiking model, as the rate model should give the average over Poisson samples (see Appendix).

We next examined the relative contributions of synaptic depression at different synaptic loci in the full model circuit, using the rate model. This model has no spike-rate adaptation and has a fixed membrane time constant, so only depression should contribute to the c-d phase advance. Synaptic depression can be operative in one of three locations: in the geniculocortical synapses (G), in the intracortical excitatory synapses (E), and in the intracortical inhibitory synapses (I). This yields eight possible configurations for the locations of depressing synapses. We compared the c-d phase advance produced in each of these eight configurations (figure 40). Matching these data across the different depression conditions is not trivial; one must ensure that the data are comparable by matching firing rates, for example, or by using the same set of parameters in all cases. We chose to show the distribution of results for all model parameter sets that satisfied the known experimental constraints (see Methods) at a given temporal frequency. Similar plots in which we include only model parameter sets that fit the constraints at *all* temporal frequencies give similar results with less variability, but there are no such parameter sets within our search range for some cases (the "I+E" and "G" cases, and the train "E" case).

As evidenced by figure 40, depression of either geniculocortical or intracortical excitatory synapses can induce approximately 5 degrees of c-d phase advance, and these advances sum when depression is present in both locations. Depression of inhibitory synapses has relatively little additional effect. In the absence of any depression, there is no c-d phase advance, as expected. These general results are for the most part similar across temporal frequency of the input and choice of synaptic depression parameters ("pulse" vs. "train"), except that the train parameter set tends to produce somewhat larger phase shifts than the pulse set, as is also evident in figure 39.

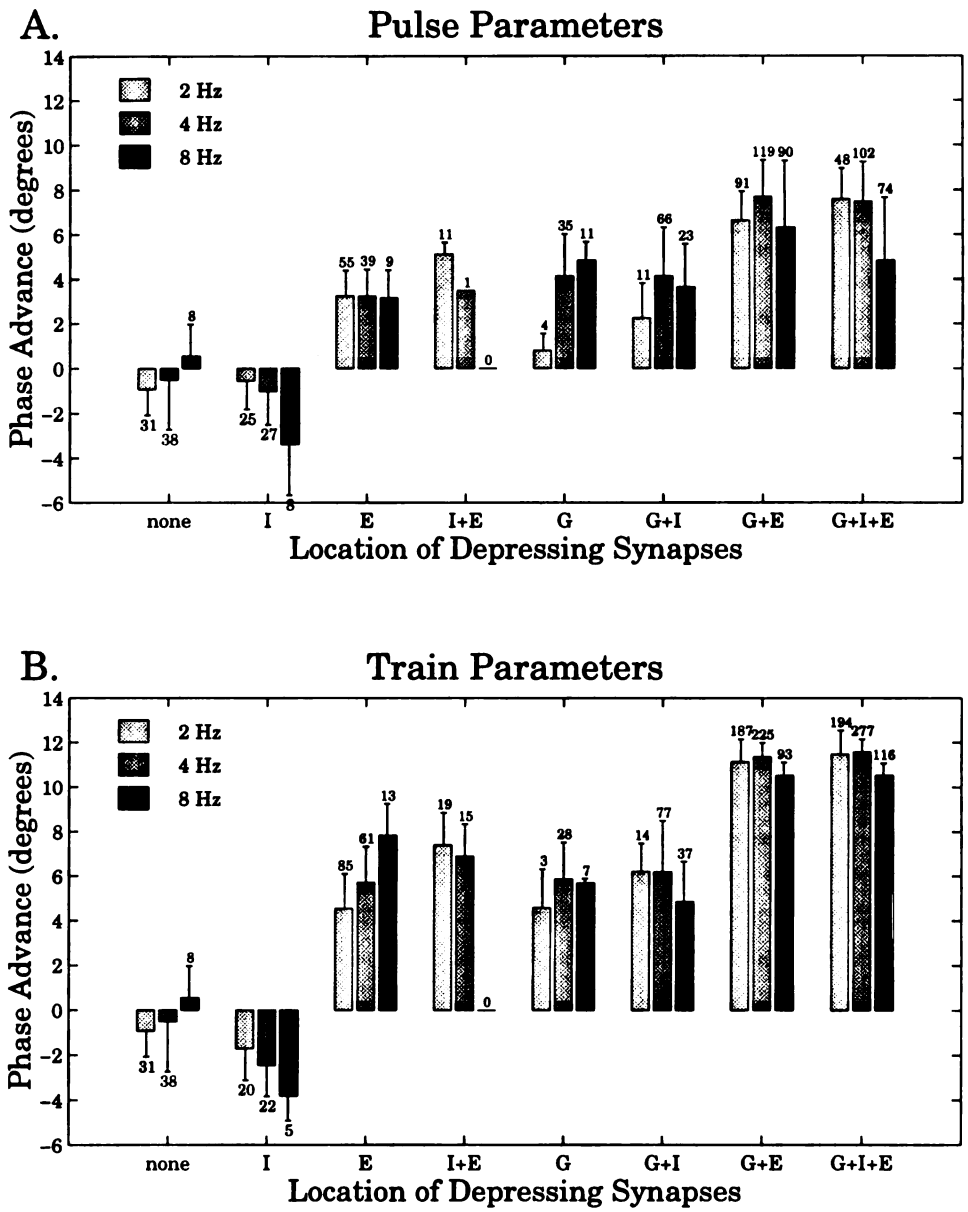


Figure 40. Please see figure legend, next page.

Figure 40. Dependence of c-d phase advance on the location of depressing synapses. Graphs show c-d phase advance ± standard deviation for those rate model parameter sets that produce constraint-satisfying outputs (see Methods) at a given temporal frequency. The number of parameter sets contributing to each data point is noted above each error bar. C-d phase advance is shown for grating inputs with temporal frequencies of 2 Hz, 4 Hz, and 8 Hz, represented by the light gray, gray, and dark gray bars, respectively. The location of the depressing synapses, if any, in each of the cases is indicated by the letter(s) on the abscissa (G = geniculocortical, I = inhibitory, E = excitatory intracortical). The 3 bars above "I + E", for example, indicate the c-d phase advance ± standard deviation for 2 Hz, 4 Hz, and 8 Hz grating inputs when the rate model includes depression in both the I and the E synapses, but not in the G synapses. **A.** C-d phase advance for "pulse" depression parameters. **B.** C-d phase advance for "train" depression parameters. Note that the rate model is completely deterministic; the standard deviations arise from the averaging of all constraint-satisfying parameter sets for the given temporal frequency and location(s) of depressing synapses. If we restricted ourselves to parameter sets that satisfied constraints across all three temporal frequencies, results were very similar where such parameters were found (except that standard deviations were much smaller); but no such sets were found for some locations of depressing synapses (see text).

To consider the additional effects of spike-rate adaptation and of contrast-dependent changes in membrane time constant, we turn to the spiking model. In this model, depression was included only at geniculocortical synapses, for reasons described in Methods. In the absence of depression ("D") or adaptation ("A"), a c-d phase shift of 3-4 degrees appears (figure 41, "No A, No D"), consistent with a contrast-induced decrease in membrane time constant (which predicts changes of 2-5 degrees). Adding either adaptation alone ("A, no D") or geniculocortical depression alone adds roughly another 5 degrees, and the effects of these two mechanisms together are additive.

With all three mechanisms present, the spiking model shows shows mean c-d phase advance of $13 - 15°$, relative to LGN, for either set of depression parameters (figure 41). Depression in intracortical excitatory synapses can easily add another 5° (figure 40). This suggests that these mechanisms may be sufficient to account for the differences between LGN X-cell and V1 c-d phase advances that have been observed in cats (figure 35). However, while we have found that the effects of geniculocortical depression add with those of intracortical E depression (figure 40) and with those of adaptation (figure 41), we have not studied the three together. We tried modeling adaptation in the rate model, but did not see an effect on c-d phase advance. In our simple rate model, adaptation was proportional to the rate, and therefore was active even at low rates. In reality and in our spiking model, the net effect of adaptation increases faster than linearly with firing rate: the mean adaptation current increases proportionally to the rate, but the effect of this current on spiking increases with rate, because at higher rates (smaller inter-spike-intervals), there is less time for the spike-induced current to decay between spikes. This difference appears to be critical to the c-d phase advance induced by adaptation. Rather than include a more complicated (and underconstrained) dependence of adaptation on rate, we elected to study only the effects of synaptic depression in the rate model, and to study adaptation only in the spiking model. Conversely, as discussed in Methods, for reasons of computational complexity, we did not study depression of intracortical synapses in the spiking model.

We also examined the dependence of phase shift on stimulus orientation in the rate model (data not shown). C-d phase advance remains essentially constant across orientations that give reasonable

response.

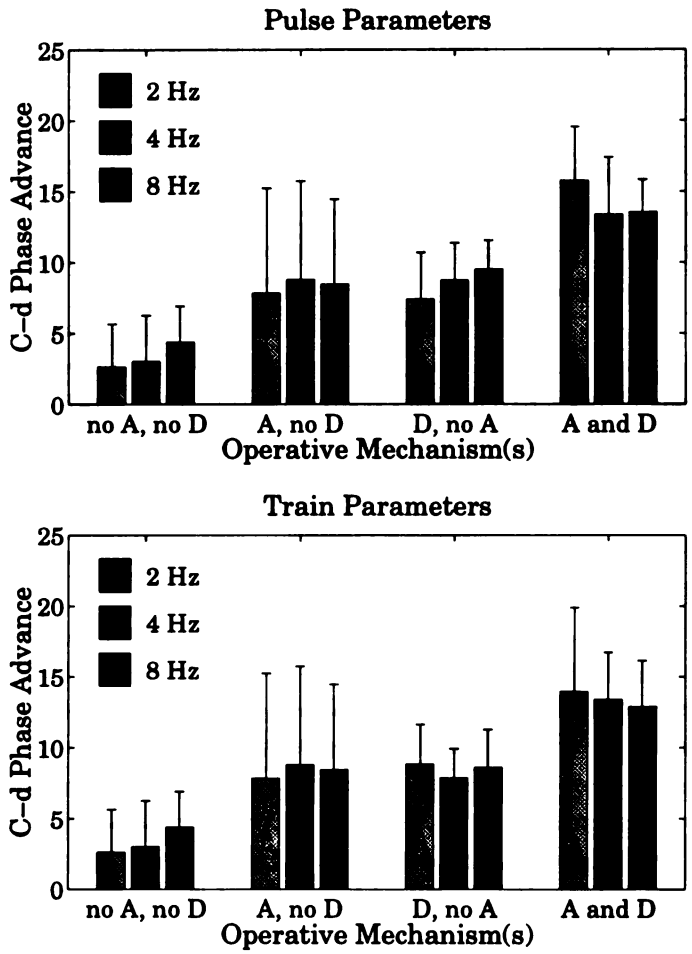


Figure 41. C-d phase advance of spiking model in the presence of different temporal nonlinearities for different temporal frequencies of input. Light gray: input gratings at a temporal frequency of 2Hz; gray: input gratings at 4Hz; dark gray: input gratings at 8Hz. Four different types of simulations were run: from left to right, simulations with neither spike-rate adaptation nor synaptic depression ("no A, no D"), simulations with only spike-rate adaptation ("A, no D"), simulations with only synaptic depression ("D, no A"), and simulations with both mechanisms present ("A and D"). We calculated a c-d phase advance from the peri-stimulus time histograms for each of the 29 cells examined, then computed the mean and standard deviation across cells for each condition. Both spike-rate adaptation and geniculocortical synaptic depression induce a c-d phase advance, and the advance increases when both are present simultaneously.

Contrast-Dependent Changes in Temporal Frequency Tuning

We next studied the contrast dependence of temporal frequency tuning. As in our studies of contrast-dependent phase advance, we wanted to isolate the cortical contribution to temporal frequency tuning – in this case, to understand the cortical response in the absence of any incoming temporal information beyond the stimulus-driven temporal modulation of the input rates. Experimentally, the LGN inputs show temporal-frequency dependence in the amplitude of their rate modulations (their response F1; figure 36A). Thus, we found it convenient to consider an even simpler model of LGN responses, in which the LGN response F1 was constant across temporal frequencies at a given contrast, with larger F1's representing higher contrast. We refer to such an LGN response profile as "flat", in distinction to the experimental tuning of figure 36A, which we refer to as "Sclar" tuning (because the experimental data is from Sclar (1987)). We present data in all cases for both flat (dashed lines in figure 42) and Sclar (solid lines in figure 42) inputs. Flat data allow us to examine cortical contributions to temporal tuning; Sclar data permit us to view cortical responses to more

realistic LGN inputs.

We first considered cortical responses in the absence of depression or spike-rate adaptation. With flat LGN inputs, the LGN input conductances of course show a flat tuning profile (figure 42A, top row, dashed lines), and the growth in LGN input conductance with increasing contrast is constant across temporal frequencies (see "normed" responses in figure 42A, which show responses at high input level divided by responses at low input level). However, the output of the cortical cells is not flat: higher-frequency responses are much attenuated (figure 42A, middle row), both for flat and for Sclar inputs. Moreover, these smaller, higher-frequency cortical responses grow more rapidly with increasing contrast than lower-frequency responses, especially in the spiking model and especially with flat LGN inputs. That is, as in the experimental data, increased contrast enhances cortical responses to higher temporal frequencies.

We next added spike-rate adaptation (figure 42A, bottom right) and synaptic depression (figure 42B,C). Both mechanisms suppress responses to lower-frequency stimuli much more strongly than responses to higher-frequency stimuli. By so doing, adaptation and depression convert the low-pass cortical response to flat LGN inputs into a band-pass one, and more generally strongly enhance the band-pass nature of cortical responses to the more realistic Sclar LGN inputs. Both mechanisms can contribute to the relative enhancement of higher-temporal-frequency responses with increasing contrast. This is because both are more strongly activated by lower-frequency than by higher-frequency stimuli. In the rate model, for example, the relative enhancement of higher-temporal-frequency responses can be seen much more robustly in the presence of depression (figure 42B) than in its absence (figure 42A), while in the spiking model, adaptation (bottom rows of figure 42A,B,C) tends to eliminate any relative enhancement of lower frequencies while preserving such enhancement at higher frequencies. If too strong, however, depression also suppresses the contrast-dependent differences between LGN input conductances, making different contrasts appear more alike to the cortical cell. This reduces the strength of contrast-dependent response enhancement at all temporal frequencies, and can largely eliminate the relative enhancement of higher frequencies (*e.g.*, figure 42C, rate model).

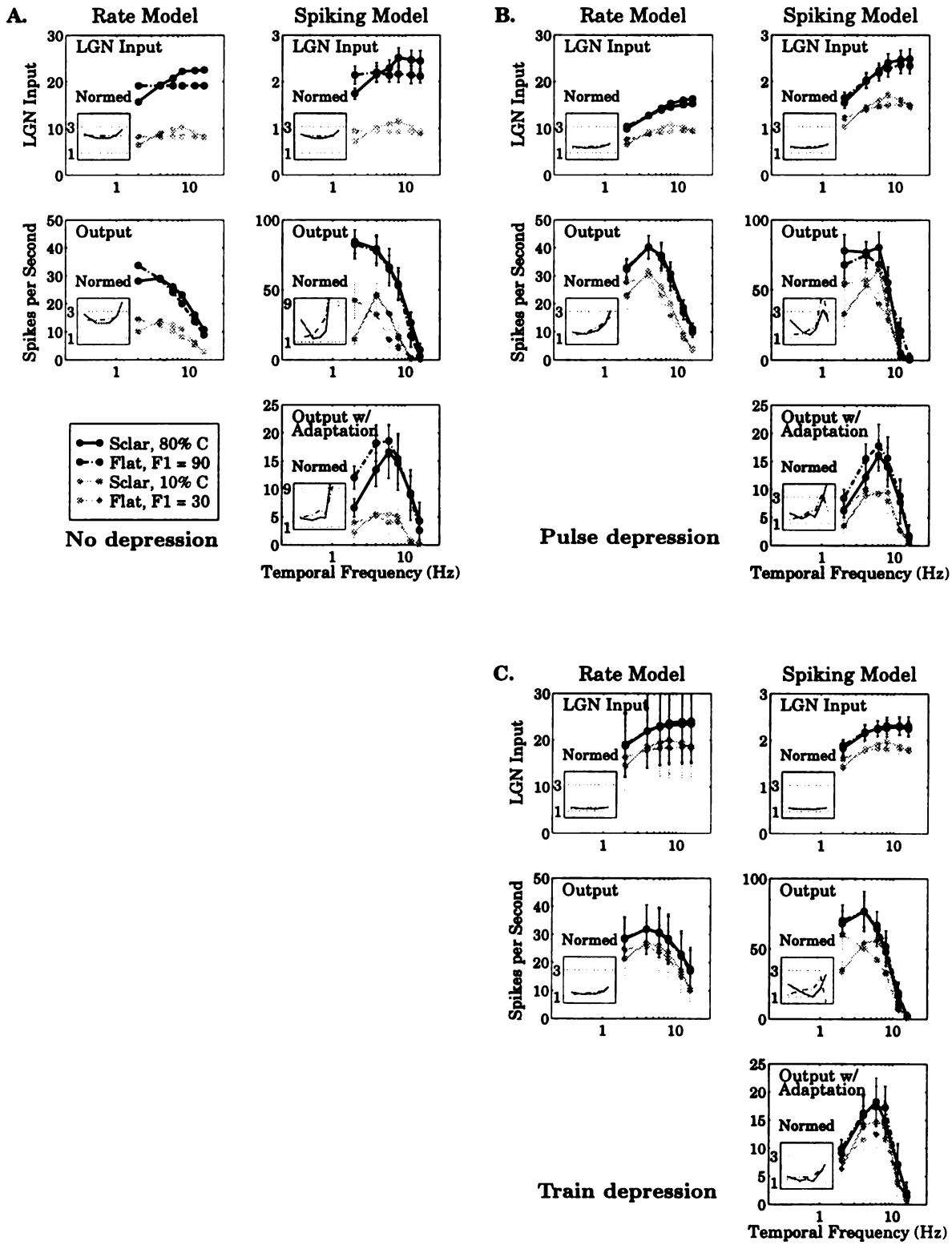


Figure 42. Please see figure legend, next page.

Figure 42: The contrast dependence of temporal frequency tuning for different outputs of the rate and spiking models. X-axes: temporal frequency. Y-axes: for top rows, F1 of summed LGN input to a simple cell – measured in Hertz for the rate model and in nano-mhos for the spiking model; for all other plots, F1 of excitatory simple cell firing response, measured in spikes per second. Two types of LGN inputs were used. Model responses to "flat" LGN inputs are indicated by dashed lines; responses to experimentally measured "Sclar" LGN inputs, figure 36A (Sclar, 1987), are indicated by solid lines. Gray lines show responses to 10% contrast ("Sclar") or LGN response F1 of 30 ("flat"), while black lines indicate responses to 80% contrast or F1 of 90. Note that the F1 values of the flat inputs are set before LGN outputs are calculated – *i.e.* they are "pre-rectification" values (see Methods). Insets: Responses at high input level divided, frequency-by-frequency, by low input level responses. For each of **A, B, C**: left column shows rate model, right column spiking model; top row shows LGN input to simple cell, middle row shows simple-cell firing responses without spike-rate adaptation currents, bottom row shows simple-cell spiking responses with spike-rate adaptation currents in excitatory cells (spiking model only). **A**. No synaptic depression. *Top row*: Because depression is absent, conductances very closely follow the temporal frequency dependence of LGN response amplitudes. *Middle row*: Because of filtering by the membrane time constant at higher, but not lower, temporal frequencies (see text), as well as inhibition in the model circuit, both types of model show low-pass behavior, as well as a relative amplification of high temporal frequency responses with contrast (insets). *Bottom row*: Note the band-pass nature of the response induced by spike-rate adaptation, in addition to the relative amplification of high temporal frequencies (inset). **B**. "Pulse" depression. *Top row*: With depression present, LGN input conductances no longer closely follow the temporal frequency dependence of LGN response amplitudes; low frequency responses show a relative attenuation, even for "flat" inputs. *Middle, Bottom rows*: Cortical outputs are correspondingly band-pass, and show relative amplification of high temporal frequencies. **C**. "Train" depression. *Top row*: As with pulse depression, conductances show high-pass behavior. Difference between conductances at low vs. high contrast is strongly attenuated by the strong depression. *Middle, Bottom rows*: Cortical outputs: difference between contrasts, and relative emplification of high temporal frequencies, is attenuated relative to pulse depression. The large standard deviations for the rate model result from the fact that two classes of parameter sets, with two distinct values for the feedforward gain (and thus different response amplitudes), satisfied the constraints. Rate model plots include results only for parameter sets that satisfied experimental constraints (see Methods) at every temporal frequency. There was 1 parameter set for no depression, and 3 and 33 parameter sets for both geniculocortical and intracortical pulse and train depression, respectively (plots show average over parameter sets). In all cases, spiking model results show averages over 29 cells. All error bars indicate standard deviations.

A third nonlinearity that contributes to the differential enhancement of higher temporal frequency responses is the change with contrast in the degree of low-pass cellular filtering. Cellular and synaptic time constants act as low-pass filters, causing the modulation of the simple cell's voltage response (the first harmonic or F1 of the voltage response) to decrease with increasing temporal frequency, and accordingly causing a similar decrease in the peak voltage response.[2] The membrane time

[2]A linear RC model of a cell with time constant τ produces modulated first harmonic responses to temporal frequencies f proportional to $1/\sqrt{1+(2\pi f\tau)^2}$; $f = 1/\tau$ diminishes the maximum response by 84%. Membrane time constants of 8-16 msec, as used in the rate model, would produce corresponding attenuations of 14%-35% at 16 Hz, and 22%-46% at 12 Hz, relative to responses at 2 Hz. The time constant in the spiking model covers a similar range; it depends on synaptic input, including stimulus-independent background firing, and is approximately 15 msec in the

constant of a cortical cell shrinks as the amount of synaptic input to the cell, and thus the membrane conductance, increases. Thus, the membrane time constant is smaller at higher contrasts, yielding less attenuation of voltage responses to higher temporal frequencies than at lower contrasts. This effect is captured in the spiking model, but not in the rate model. The effect is relatively small: the mean time constant in the spiking model shrinks from 12.5 msec to 8 msec between the low (F1=30) and high (F1=90) flat input levels. Assuming a linear model of voltage response, this yields about an 18% increase in the high-contrast voltage F1 at 12 Hz relative to that expected from the low-contrast time constant.

Finally, the nonlinearity of a nonzero spiking threshold plays an important role in the contrast-dependent enhancement of higher-temporal-frequency responses. Figure 43A shows responses in the spiking model to four levels of flat LGN input, when adaptation but not depression is present (same data as figure 43A, bottom row, dashed lines, but with two additional LGN input levels included). Note the responses at 12 Hz input frequency: for input F1's of 15 or 30 spikes per second, the response is close to zero, whereas for input F1's of 60 or 90 spikes per second, the response is significantly higher. The basis for this increase can be seen in figure 43B, which shows the corresponding intracellular voltage traces in response to 12 Hz inputs for a randomly chosen cell with spiking turned off; the spike threshold of -52.5 mV is indicated as a dashed line. Only the responses for the two higher input levels cross spike threshold for this cell during the time period shown; the modest attenuation of voltage modulation due to membrane filtering is, on average, sufficient to quench most spiking responses at the lower input levels. Higher input modulation levels, however, yield higher voltage modulations that consistently cross threshold and give spiking responses that are much more amplified, relative to low input responses, than the corresponding voltage modulations might indicate. Of course, this threshold effect requires the dominant inhibition in our circuit model, which ensures that the mean response to a sinusoidal grating is always subthreshold and spiking occurs only on voltage modulations (Troyer et al., 1998).

absence of a stimulus, 12.5 (DC) $\pm$ 1.5 (F1) msec for F1=30 flat LGN inputs, and 8 (DC) $\pm$ 2.5 (F1) msec for F1=90 flat LGN inputs.

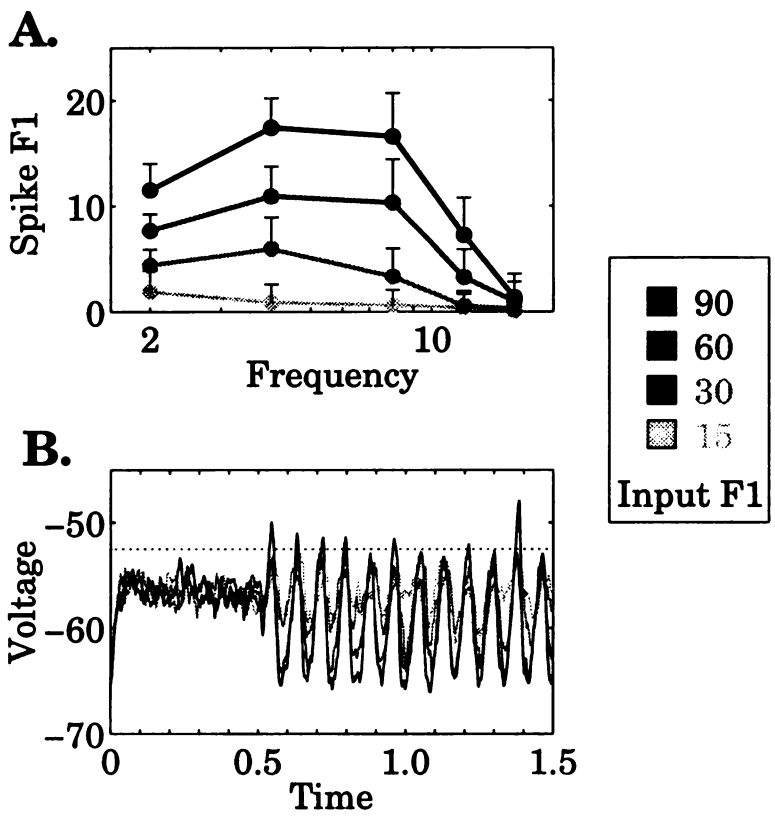


Figure 43. The iceberg effect: the appearance of higher temporal frequency responses at higher contrasts **A.** Temporal frequency tuning curves for "flat" F1 inputs of 15, 30, 60, and 90 Hz (pre-rectification values), color-coded from light gray to black, respectively, for a spiking model simulation in which adaptation was present, but synaptic depression was not (same as figure 42A, bottom right, but with two additional F1 values). Error bars indicate standard deviations of the means across 29 cells. Note that responses to 12 Hz input gratings are present for input F1's of 60 and 90 Hz, but essentially absent for input F1's of 15 and 30 Hz. **B.** Intracellular voltage traces, for a randomly chosen cortical cell, in response to a single presentation of a 12 Hz temporal frequency grating at each of the 4 F1 input levels used in A (corresponding, at 12 Hz, to contrasts of 3.9%, 7.8%, 18.7%, and 41.2%). Spiking responses in the cell have been turned off; spiking threshold is indicated by the dotted line. Synaptic conductances for LGN inputs and nonspecific *in vivo* "background" inputs (see Methods) were turned on at time 0. A blank stimulus was presented for the first 0.5 seconds of the trace, after which the grating stimulus appeared. Note that, for input F1 values of 15 or 30 Hz, the membrane voltage never crossed spike threshold. For higher input F1 values (60 or 90 Hz), the membrane potential did reach threshold, as corroborated by the increase in the spiking response indicated in A. Traces were achieved as follows: all conductances onto a cell, including spike-rate adaptation conductances, were recorded during simulations of A. These conductances were then "played back" to the cell with spiking turned off.

This differential amplification of spiking responses, relative to voltage responses, depends critically on the three other nonlinearities discussed above – depression, adaptation, and conductance-induced decreases in membrane time constant. This can be seen by comparing the contrast-dependence of temporal frequency tuning to the contrast-invariance of orientation tuning. In the absence of these nonlinearities, a shift to higher temporal frequencies would be very much like a shift toward non-preferred orientations: both shifts yield smaller voltage modulations (for increasing temporal frequency, modulation shrinks due to cellular and synaptic filtering; for shifts to non-preferred orientations, modulation shrinks because the firing-rate modulations of the different LGN inputs to a cell become increasingly desynchronized (Troyer et al., 1998)). Yet orientation tuning is kept contrast-invariant in our model by the dominant antiphase inhibition. For temporal frequency tuning to vary with contrast, it must be the case that the relationship between the modulation and the inhibition changes with temporal frequency, so that responses can grow with contrast by greater factors at higher temporal frequencies, yet grow by the same factor for all orientations at a given temporal frequency. This appears to be due largely to the three nonlinearities discussed above, which all act in the appropriate way: each tends to cause a greater contrast-dependent boost in higher-frequency than lower-frequency responses, but to more equally affect the growth with contrast of responses to different orientations at a given temporal frequency. In addition, the LGN input firing rates show a slightly greater contrast-dependent increase at high than at low temporal frequencies. When all of these effects are eliminated, little relative enhancement of higher temporal-frequency responses is seen in our model (figure 42A, rate model, flat inputs).

Note also that we see at best only a weak shift in the peak of the temporal frequency tuning curve with increasing contrast(Figs. 2.4.1B,C, bottom right). At present, there is no experimental data as to whether LGN-recipient cells in cat layer 4 show such a shift in peak. If they do not, but instead show only a relative increase in responses to higher temporal frequencies at higher contrast, this could be sufficient to induce shifts in the tuning peaks of downstream cells.

Saturation of Responses With Increasing Contrast

Lastly we examined the saturation of cortical responses with increasing contrast (figure 44). Even in the absence of depression or spike-rate adaptation, model cortical responses tend to saturate somewhat earlier than their LGN inputs, particularly at lower temporal frequencies (figure 44B, Table 4). If either pulse or train depression is active (figure 44C,D), saturation occurs significantly earlier than in either the LGN inputs or the models without depression. (The one exception is at the highest temporal frequency of the spiking model with adaptation, for which responses are small and the measure of saturation probably inaccurate.) Moreover, clearly in the depression cases, and also somewhat in the examples lacking depression, there is a tendency for responses to higher temporal frequencies to saturate later than responses to lower temporal frequencies: for cases with depression, C_{50} values increase monotonically with temporal frequency if the lowest temporal frequency is excluded. The same pattern is seen in the V1 cell of figure 37C (less so in the cell of 11B), though the model C_{50} values are somewhat lower than those measured by Albrecht.

Model C_{50} Values

Cell	Temporal Frequency (Hz)					
	2	4	6	8	12	16
LGN input	> 100	> 100	35	23.0	21.9	21.8
Rate: No dep.	29.6	23.3	18.0	14.6	16.8	19.6
Rate: Pulse	7.3	3.9	5.5	6.7	10.6	14.4
Rate: Train	5.7	2.2	2.9	3.4	5.9	8.6
Spiking-NoA: No dep.	16.6	8.5	10.0	10.9	29.7	90.1
Spiking-NoA: Pulse	10.2	4.2	4.8	6.3	12.1	11.0
Spiking-NoA: Train	9.2	4.8	4.7	4.7	9.7	*
Spiking-A: No dep.	24.7	15.9	20.0	14.1	25.1	37.8
Spiking-A: Pulse	6.6	5.7	7.0	6.1	12.7	14.9
Spiking-A: Train	5.2	3.0	3.9	3.7	8.0	22.3

Table 4. C_{50} values from Naka-Rushton curves (Eq. 2.1) fit to model data of figure 43. Model used as input the LGN curves of figure 37A ("LGN (1)" in Table 3); curves for 6 and 12 Hz were obtained by interpolation. NoA, A: no adaptation, adaptation, respectively. *: Best fit had negative n, *i.e.* Naka-Rushton curve does not provide good fit.

The contrast saturation effects induced by synaptic depression can be readily understood. As demonstrated by Abbott et al. (1997) and Tsodyks and Markram (1997), in the presence of depression, as a presynaptic neuron's firing rate increases to values much larger than $1/\tau$ (where τ is the time constant of recovery from depression), the overall postsynaptic effect of its synapses – proportional to rate times efficacy – saturates at a plateau value. The postsynaptic cell cannot "see" further increases in rate. Thus, the impact on the cortical cells of contrast-induced increases in LGN firing rates will saturate earlier than it would without depression – cortical cells will not "see" further increases in contrast. This saturation will occur at higher contrasts for higher temporal frequencies, because depression more strongly suppresses lower than higher-frequency inputs.

As shown in the table, however, cortical responses can saturate at lower contrasts than LGN, even when depression is absent. These effects result from the inhibition in our circuit model. Because the cortical response is determined by a thresholded version of the membrane voltage, for a sinusoidal input grating the response of the cortex can be largely understood from the peak membrane voltage. We estimate this peak as the sum of the mean voltage and the modulation amplitude (first harmonic) of the voltage. In the absence of inhibition, this peak voltage closely follows the modulation of the LGN input: tuning curves of peak voltage and of LGN modulation show very similar C_{50}'s under various conditions (data not shown). However, when inhibition is added, the peak voltage can show C_{50} values that are lower than the corresponding LGN values, because the inhibition in the model both decreases the slope of, and adds a constant negative DC offset to, the curve of peak voltage vs. contrast. The DC offset originates from the background firing of the LGN, which, because the cortex is inhibition-dominated, is net inhibitory. By both flattening and shifting the cortical response curve closer to zero, inhibition effectively causes cortical neurons to saturate sooner than their inputs.

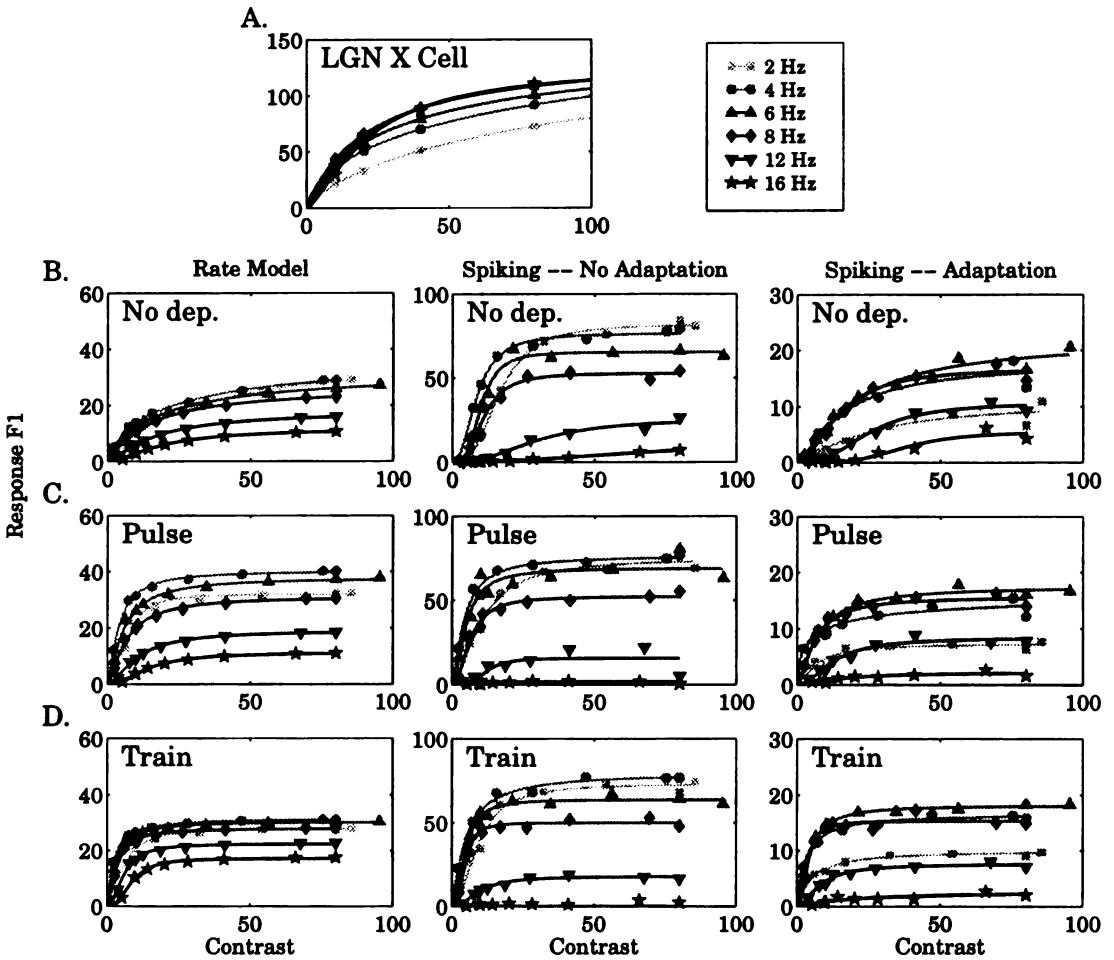


Figure 44. Model data for the temporal frequency dependence of contrast saturation. Model contrast saturation curves. Light gray to dark gray: increasing temporal frequency. Rate model plots show results for 1 parameter set for no depression, and averages over 3 and 32 parameter sets for pulse and train depression, respectively. In all cases, spiking model results show averages over 29 cells. **A.** Sclar input data (pre-rectification F1 for ON cells versus contrast). **B-D.** Left column: rate model; middle column: spiking model, without adaptation; right column: spiking model, with adaptation. Data points for "flat" inputs are combined with those for Sclar inputs (see Methods). **B.** No depression **C.** Pulse Depression **D.** Train depression Results for pulse and train depression are qualitatively similar. As described in the text, cortical responses tend to saturate at lower contrasts than do their LGN inputs, and responses to higher temporal frequencies saturate at higher contrasts.

2.5 Discussion

We have established that a simple circuit model of cat layer 4 that achieves contrast-invariant orientation tuning can also account for three contrast-dependent (c-d) "non-linearities" in simple cell responses to sinusoidal stimuli: c-d phase advance, c-d increase in responsiveness to higher temporal frequencies, and contrast saturation. This work demonstrates that these response nonlinearities can arise locally through the many nonlinear elements present in the LGN responses and cortical circuitry. The observed c-d phase advance can be largely or entirely accounted for by the combined effects of geniculocortical and intracortical synaptic depression, spike-rate adaptation currents in cortical cells, and c-d changes in cortical cell conductance. The c-d increase in responsiveness to higher temporal frequencies arises from the interaction of these nonlinearities with the spike threshold. Finally, the inhibition in our model circuit causes cortical cell responses to saturate at slightly lower contrasts than do the LGN inputs, while synaptic depression causes a much stronger decrease in cortical saturating contrast relative to LGN.

The circuit model used here has previously been shown to account for a wide variety of intracellular and extracellular observations related to orientation tuning in cat layer 4, including the aforementioned locality in orientation of the excitation and inhibition received by simple cells (Troyer et al., 1998). These results are inherited by the present model, which simply adds a new synaptic mechanism, synaptic depression, while keeping the same circuitry. Further nonlinearities beyond those discussed here can also be understood from this model. For example, in Krukowski et al. (1998), we showed that the anti-phase inhibition in the model circuit naturally explains "cross-orientation inhibition" (suppression of response to a grating of the preferred orientation by simultaneous display of an orthogonal grating). While other response properties such as direction selectivity remain to be addressed, the model circuit appears to provide an excellent candidate framework for understanding processing within the classical receptive field in cat layer 4.

2.5.1 Theoretical and Experimental Limitations of the Present Work

As we emphasized in the Results, the data on response nonlinearities remains quite sparse. None of the data in cats are known to be from layer 4 (though most are from identified simple cells), so we do not know the degree to which layer 4 cells exhibit these nonlinearities. In addition, LGN Y cells show stronger response nonlinearities than X cells, but the dependence of nonlinear cortical response properties on the X or Y nature of the input received has not been studied.

Furthermore, none of the data for LGN and cortical responses have been recorded under the same conditions from the same animal in cats (in monkeys, there is one such study on the issue of contrast saturation (Sclar et al., 1990)). The difficulty of such experiments is clear, but such data are necessary for determining the contribution of the cortex to c-d phenomena. Temporal response properties in particular might be quite mutable by different types of anesthesia: increases in inhibition, as induced by barbiturates, can cause a lower temporal frequency cutoff in responses at a given contrast in our circuit model, while blockade of NMDA receptors, *e.g.* by ketamine, can have the reverse effect (Krukowski et al., 1998).

Further data are needed to test the dependence of c-d phase advance on temporal frequency and on stimulus orientation, particularly in cat layer 4, and thereby to limit potential parameters and

mechanisms. Albrecht (1995) reported a weak positive correlation between c-d phase advance and temporal frequency across cat and monkey simple cells. While our average results on phase advance tend to show no dependence on temporal frequency, individual parameter sets can show such dependence (*e.g.*, figure 39). Similarly, data for a few cells in monkey V1 (Carandini et al., 1997) demonstrate little dependence of c-d phase advance on stimulus orientation. While our average rate-model phase advance results show no dependence on stimulus orientation (and thus on response strength) for orientations that give appreciable response, we have not carefully examined the parameter dependence of this result; and a dependence may exist for components of c-d phase advance due to adaptation or conductance changes, which were not included in the rate model.

Several of our explanations depend on the existence of sufficient synaptic depression *in vivo*. One study reported that cortical depression appears weaker *in vivo* than *in vitro* (Sanchez-Vives et al., 1998), but speculated that this may result simply from the greater baseline rate of depression *in vivo* due to background activity, an effect included in our modeling. Support for a functional depression-like mechanism *in vivo* was reported by Nelson (1991a,b): responses in cat V1 were suppressed by repetition of visual stimuli in a manner consistent with both synaptic depression and a presynaptic origin. We attempted to control for the uncertainty in the strength of depression by studying two different *in vitro* parameter sets; they showed little difference in behavior except in the degree of contrast-dependence of temporal frequency tuning.

The model weakly suggests that geniculocortical depression may be less strong than in either of these parameter sets. Geniculocortical synaptic depression with these parameters seems to lead model cells to saturate too early, relative to cortical cells (Tables 3, 4). However, it is possible that nonlinearities in LGN temporal response profiles beyond the simple rectification considered here may alter the strength of depression effects on LGN inputs. For example, LGN responses tend to occur over significantly less than a half-cycle of a sinusoidal stimulus (*e.g.*, Reich et al., 1997); this would be likely to affect response saturation in a similar manner as a shift to a higher temporal frequency, for which saturation occurs at higher contrasts.

The similarity of results in both the simpler rate model and the more elaborate spiking model, and the ability to understand their differences in terms of the specific additional nonlinear mechanisms present in the spiking model, give confidence that the understandings achieved here of the contribution of each nonlinear mechanism to each nonlinear response property are fairly robust. Further mechanisms not considered here may also play a role, such as further nonlinearities in LGN responses, other active membrane conductances beyond spike-rate adaptation (McCormick, 1990), nonlinearities of dendritic integration (*e.g.*, Larkum et al., 1999), synaptic facilitation, which is seen at many excitatory synapses onto inhibitory interneurons (Thomson et al., 1993), or the presence of NMDA receptors, which can alter temporal frequency tuning in our model (Krukowski et al., 1998). We believe the present results establish the viability of a local explanation of contrast-response nonlinearities, although further experimental and theoretical work will be needed to quantitatively test this explanation. We discuss some such tests below.

2.5.2 Experimental Tests

The present explanations of c-d phase advance can be directly tested by blocking spike-rate adaptation and/or synaptic depression and determining whether this decreases c-d phase advance. Spike-rate adaptation can be blocked by several pharmacological agents (Nicoll, 1988; Baskys, 1992). If applied iontophoretically to individual cells, these should reduce c-d phase advance (although spike-rate adaptation may not be as strong *in vivo* as *in vitro*, Tang et al. (1997)). Selective intervention against synaptic depression is more difficult, however (see discussion in Chance et al. (1998)).

The combined role of LGN response nonlinearities and geniculocortical synaptic depression in both c-d phase advance and contrast saturation could be assayed in intracellular recordings from simple cells, by using electrically evoked cortical suppression (Chung and Ferster, 1998) to isolate geniculocortically-driven currents during presentation of sinusoidal grating stimuli. By comparing c-d response properties of these input currents to those of the cell's voltage response with the cortical circuit intact, the degree of involvement of cortical mechanisms could be assessed. Comparisons to average LGN firing properties might be used to assay the role of geniculocortical synaptic depression; we would predict that these input currents would show greater c-d phase advance and earlier contrast saturation than LGN firing rates.

The explanation of c-d changes in temporal frequency tuning could be tested by measurements of the membrane potential in response to high-temporal-frequency gratings of increasing contrast. In cells showing c-d enhancement of higher-temporal-frequency spiking responses, we predict a threshold effect: as contrast increases, the spiking response should increase faster than the voltage response.

2.5.3 Applicability of the Model to Other Species

Contrast nonlinearities have also been studied in monkeys. Data there, though also limited, seem qualitatively consistent with those in cats (Hawken et al., 1992; Carandini and Heeger, 1994; Albrecht, 1995; Carandini et al., 1997). However, response properties in the LGN-input-recipient portions of monkey layer 4 are quite different from those in cat layer 4: while cat layer 4 consists very largely of classical simple cells (cells with aligned and oriented, segregated ON and OFF subregions) with strong orientation tuning (Gilbert, 1977; Bullier and Henry, 1979), monkey layer 4C has few such cells (Blasdel and Fitzpatrick, 1984; Hawken and Parker, 1984). Thus, our model circuit is unlikely to apply directly to monkeys. Nonetheless, the general explanations of the effects of synaptic depression, spike-rate adaptation, and c-d changes in conductance on c-d temporal phase advance and contrast saturation are likely to apply. Our remaining explanations, of the c-d shift in temporal frequency tuning and of contrast saturation effects beyond those due to depression, involve the role of inhibition in our model circuit and the low-pass nature of cellular and synaptic filtering. The filtering should be similar across systems. The major features of inhibition involved arise from the spatial opponency and dominance of the inhibition; we can imagine that these may also be general principles of cortical layer 4 (discussed in Troyer et al., 1998).

2.5.4 Comparison to Other Models

The importance of understanding the nonlinearities studied here has been emphasized by studies of the normalization model (Albrecht and Geisler, 1991; Heeger, 1992; Carandini et al., 1997, 1998).

These studies have strongly influenced the field's thinking: as a phenomenological description of cortical processing, the normalization model integrates a wealth of data in a simple way.

However, as a mechanistic explanation, this model is problematic. It incorrectly assumes that inputs are linear in stimulus contrast (see Introduction). Perhaps more seriously, the model's explanations require unrealistically high membrane time constants. The model explains the temporal nonlinearities addressed here through decreases in membrane time constant τ with increasing contrast, induced by the increase in membrane conductance from the normalizing inhibition. The phase advance and the high-temporal-frequency cutoff F at a given contrast are determined by τ in the model; c-d changes in τ yield c-d changes in these effects. However, V1 cells often show low-contrast (or even high-contrast - Saul and Humphrey (1992)) cutoffs at $F = 10 - 15$ Hz (Albrecht, 1995; Carandini et al., 1997, figure 40). For such a cutoff to be simply due to τ, one must have $\tau > 1/F$, *i.e.* greater than 66-100 msec (see footnote 2). Yet time constants of cortical cells ***in vivo*** are only 15-24 msec (Hirsch et al., 1998) at rest, and will only decrease under visual stimulation. Similarly, a 20° c-d phase shift in response to a 2 Hz stimulus - a temporal advance of 28 msec - would require a c-d decrease in τ of 28 msec.[3] Such a large decrease between 10% and 80% contrast seems unlikely.

The normalization model also requires shunting inhibition that depends only on contrast, independent of orientation. This is necessary, for example, to explain contrast saturation or c-d phase shifts of responses to non-preferred stimuli. Experimental data now show that there is a contrast-dependent conductance increase which, at preferred orientations, can be as high as two- or three-fold, but which is tuned for orientation (Borg-Graham et al., 1998; Hirsch et al., 1998; Anderson et al., 1999). The corresponding reduction in time constant could certainly ***contribute*** to phase advances; a small such contribution is seen in figure 41. It could also contribute to the threshold effect that we argue explains contrast-dependent changes in temporal frequency tuning. However, any such contribution would show orientation tuning like that of the conductances.

Another model - that of Chance et al. (1998) - independently arrived at some of the same qualitative ideas that we have developed here (see Chance et al., 1997; Priebe et al., 1997). In particular, they also pointed out that synaptic depression of feedforward synapses could yield a c-d phase advance, although curiously, they did not find such a shift for temporal frequencies of 8 Hz or higher. They did not address the temporal frequency or saturation nonlinearities addressed here, study nonlinear mechanisms other than synaptic depression, or work in the context of a circuit model that also achieved linear response properties.

2.5.5 Conclusion: Nonlinear and Linear Response Properties

As the circuit model presented here has emphasized, many aspects of cortical processing are inherently nonlinear, including spike thresholds, adaptation, synaptic depression, and conductance effects. On the other hand, many spiking responses of cat simple cells can be understood roughly in terms of linear summation of inputs (*e.g.*, DeAngelis et al., 1995; Sclar and Freeman, 1982; Skottun et al., 1987; Skottun, De Valois, Grosof, Movshon, Albrecht and Bonds, 1991). Based on these

[3]Conductance varies across a stimulus cycle, but a simple analysis can be obtained by regarding τ as fixed for a given contrast. Then the formula for phase advance, in units of time, due to a change from low-contrast time constant τ_0 to high-contrast time constant τ_1, at temporal frequency F, is $[\arctan(2\pi F\tau_0) - \arctan(2\pi F\tau_1)]/2\pi F$ (*e.g.*, Carandini et al., 1997). For $2\pi F\tau \ll 1$, *e.g.* $F \ll 8Hz$ for typical cortical resting time constants in vivo of $\tau = 20msec$ (Hirsch et al., 1998), $\arctan(2\pi F\tau)/2\pi F \approx \tau$, hence the phase advance is simply $\tau_0 - \tau_1$.

findings, one theoretical approach is to consider simple cells as a rectified linear filter, and to seek nonlinear corrections that can give a more complete account of spiking responses (*e.g.*, Albrecht and Geisler, 1991; Carandini et al., 1997, 1998).

While this approach is useful in describing spiking behavior, we suggest that when mechanistic explanations are sought, the problem should be turned on its head. Simple cell responses must be understood in terms of cortical cells and circuits, which are inherently nonlinear. The difficulty is explaining why the behavior of a circuit model appears linear in key respects. For example, even if the input to a simple cell is modeled by a simple linear filter, inclusion of a neuron's nonzero spike threshold yields an "iceberg effect": an increase in contrast renders suprathreshold some inputs that previously gave no response, so tuning widens. We have shown how antiphase inhibition, if dominant over feedforward excitation, can eliminate this iceberg effect for orientation and yield contrast-invariant orientation tuning (Troyer et al., 1998), even with realistic, nonlinear inputs. However, such inhibition need not provide a similar correction for temporal frequency tuning in response to orientations that drive the cell, particularly given temporal-frequency-dependent nonlinear mechanisms. In fact, we have found that responses to high temporal frequencies can show such an iceberg effect.

Thus, the key mechanistic question is not why simple cell properties are nonlinear, but rather how they come to appear linear. Once the latter has been explained in a circuit model, one can see to what extent other, nonlinear behavior may emerge naturally from such biological nonlinearities as thresholds, synaptic depression, adaptation, and conductance changes.

Appendix A

Appendix: A Rate Model of Synaptic Depression

We wish to derive an expression for a depressing synapse in a spiking model that correctly reproduces the following behavior: after each presynaptic spike, the corresponding synaptic weight is multiplied by f; in between presynaptic spikes, the synaptic weight decays exponentially back to its maximal value, w_{max}, with time constant τ. We also wish to derive a rate model that is equivalent to an appropriate average of this spiking model.

We begin by deriving an equation for depression in a spiking model. Let $w(t)$ be the weight at time t. Let the presynaptic spike train be denoted by $\rho(t) = \sum_i \delta(t - t_i)$, where presynaptic spike times are denoted as t_i.[1] Our desired equation will be of the form

$$\tau \frac{dw}{dt} = -w(t) + w_{max} - \tau c \rho(t) w(t), . \tag{A.1}$$

where c is a yet-to-be-defined constant. In the absence of a presynaptic spike ($\rho = 0$), this equation simply says that w decays exponentially toward w_{max} with time constant τ, as desired. The form of the last term is set from the fact that this term, representing τ times the change in weight after a spike: (1) must be proportional to the current value of the weight, $w(t)$; (2) must be proportional to $\rho(t)$ so that it will be zero in the absence of a spike, and will be infinite (an infinite value of $\frac{dw}{dt}$, and thus a discontinuous change in w) in the presence of a spike; and (3) must have the same dimensions as w, which is achieved by multiplying by τ, leaving c as a dimensionless constant.

The value of c is determined as follows. Let the times infinitesimally before and infinitesimally after t_i be denoted t_i^- and t_i^+, respectively. Then the spike-induced depression is represented by the equation $w(t_i^+) = f w(t_i^-)$. We would like to compare this equation to the result of integrating Eq. A.1 from t_i^- to t_i^+. Because we will be integrating over an infinitesimal interval, only integrands that are infinite during that interval will give a nonzero result. Because the value of w changes discontinuously, $\frac{dw}{dt}$ is infinite in the interval; so too is $\rho(t)$. The other two terms will integrate to zero, so we need only consider the last term on the right side. However, this term, $\tau(c\rho(t)w(t))$,

[1] $\delta(x)$ is the Dirac delta function, defined by $\delta(x) = 0$, $x \neq 0$; $\int_{-\epsilon}^{\epsilon} dx \delta(x) = 1$ for any $\epsilon > 0$. This function can be realized, for example, as $\lim_{\Delta \to 0} \left\{ \frac{1}{2\Delta},\ |x| \leq \Delta;\ 0,\ |x| > \Delta \right\}$.

presents an additional problem: we cannot integrate $\int w(t)dt$ because we don't know how $w(t)$ itself is changing over the interval - *e.g.* should $w(t)$ be $w(t_i^+)$ or $w(t_i^-)$? To solve this problem, we divide Eq. A.1 by $w(t)$ and multiply by dt/τ before integrating, yielding [2]

$$\int_{w(t_i^-)}^{w(t_i^+)} \frac{dw}{w} = -c \int_{t_i^-}^{t_i^+} \sum_j \delta(t - t_j)dt \tag{A.2}$$

or

$$c = -\ln \frac{w(t_i^+)}{w(t_i^-)} = -\ln f \tag{A.3}$$

Thus, our equation for synaptic depression is

$$\tau \frac{dw}{dt} = -w(t) + w_{max} + \tau(\ln f)\rho(t)w(t) \tag{A.4}$$

This equation can be integrated to yield

$$w(t) = w(0) \exp\left(-\frac{t}{\tau} + N(t,0)\ln f\right) + \frac{w_{max}}{\tau} \int_0^t dt_1 \exp\left(-\frac{(t-t_1)}{\tau} + N(t,t_1)\ln f\right) \tag{A.5}$$

where $N(t_2,t_1) \equiv \int_{t_1}^{t_2} \rho(s)ds$ is the spike count in the interval (t_1,t_2).

We now use this equation to derive an equation for the mean weight, $\bar{w}(t) = E[w(t)]$, in terms of the mean rate, $r(t) = E[\rho(t)]$. Here $E[\cdot]$ means an expectation or mean value over some set of stochastic realizations. In this case, we assume the spike train $\rho(t)$ is a Poisson process with mean rate $r(t)$, so the expectation value is over Poisson realizations of spike trains with this time-varying mean rate. The spike count, $N(t_2,t_1)$, is therefore poisson-distributed with mean $\int_{t_1}^{t_2} r(s)ds$. The equation for $\bar{w}(t)$ is found by taking the expectation value of both sides of Eq. A.5, where non-stochastic (deterministic) quantities can be brought outside the expectation values:

$$\bar{w}(t) = w(0) \exp\left(-\frac{t}{\tau}\right) E\left[\exp\left(N(t,0)\ln f\right)\right] + \frac{w_{max}}{\tau} \int_0^t dt_1 \exp\left(-\frac{(t-t_1)}{\tau}\right) E\left[\exp\left(N(t,t_1)\ln f\right)\right] \tag{A.6}$$

Thus, to compute $\bar{w}(t)$, we must compute expectation values of the form $E[\exp(c\eta)]$, where η is poisson-distributed with mean λ:

$$E[\exp(c\eta)] = \sum_{k=0}^{\infty} P(\eta = k)\exp(ck) = \sum_{k=0}^{\infty} \exp(-\lambda)\left(\frac{\lambda^k}{k!}\right)\exp(ck) \tag{A.7}$$

$$= \exp(-\lambda) \sum_{k=0}^{\infty} \frac{(\lambda \exp(c))^k}{k!} = \exp(-\lambda)\exp(\lambda\exp(c)) \tag{A.8}$$

$$= \exp(-\lambda(1 - e^c)) \tag{A.9}$$

Applying this result to Eq. A.6 yields

$$\bar{w}(t) = w(0) \exp\left(-\frac{t}{\tau} - (1-f)\bar{N}(t,0)\right) + \frac{w_{max}}{\tau} \int_0^t dt_1 \exp\left(-\frac{(t-t_1)}{\tau} - (1-f)\bar{N}(t,t_1)\right) \tag{A.10}$$

[2] Note that these operations yield a term, $\frac{1}{\tau}\int_{t_i^-}^{t_i^+} \frac{w_{max}}{w(t)}dt$, which could also in principle be nonzero, if $w(t) = 0$. However, $w(t)$ can never reach zero for nonzero f and finite τ.

where $\bar{N}(t_2, t_1)$ is the mean number of spikes resulting between times t_1 and t_2 from a Poisson process with mean rate $r(t)$.

Finally, we want to find the differential equation for $d\bar{w}(t)/dt$ that produces Eq. A.10 as a solution. After some work, this equation can be found to be

$$\tau \frac{d\bar{w}}{dt} = -\bar{w}(t) + w_{max} - \tau(1-f)r(t)\bar{w}(t) \tag{A.11}$$

where we have noted that the mean rate $r(t) = E[\rho(t)]$ is given by $r(t) = \bar{N}'(t, 0)$. There are only two differences between our original Eq. A.4 for w , and Eq. A.11 for $\bar{w}$:

$$\begin{aligned} \rho(t) &\longrightarrow r(t) \\ -lnf &\longrightarrow (1-f) \end{aligned}$$

For purposes of simulation of the rate model, we need a discrete version, which is given by:

$$\bar{w}(t + \Delta t) = \bar{w}(t)\left[1 - \frac{\Delta t}{\tau} - \Delta t(1-f)r(t)\right] + \frac{\Delta t}{\tau} w_{max} \tag{A.12}$$

Note that for $f = 1$ (i.e. no depression), the depression term disappears, as it should. Equation A.12 serves as the update rule in the rate model.

Bibliography

Abbott LF, Varela JA, Sen K, Nelson SB (1997) Synaptic depression and cortical gain control. Science 275:220–224.

Ahmed B, Anderson J, Martin K, Nelson J (1997) Map of the synapses onto layer 4 basket cells of the primary visual cortex of the cat. J Comp Neurol 380:230–242.

Albrecht DG (1995) Visual cortex neurons in monkey and cat: Effect of contrast on the spatial and temporal phase transfer functions. Vis Neurosci 12:1191–1210.

Albrecht DG, Geisler WS (1991) Motion selectivity and the contrast-response function of simple cells in the visual cortex. Vis Neurosci 7:531–546.

Albus K, Wolf W (1984) Early post-natal development of neuronal function in the kitten's visual cortex: A laminar analysis. J Physiol 348:153–185.

Alonso JM, Usrey WM, Reid RC (1996) Precisely correlated firing in cells of the lateral geniculate nucleus. Nature 383:815–819.

Anderson J, Carandini M, Ferster D (1999) Orientation tuning of input conductance in cat primary visual cortex. J Neurosci (submitted).

Antonini A, Stryker MP (1996) Plasticity of geniculocortical afferents following brief or prolonged monocular occlusion in the cat. J Comp Neurol 369:64–82.

Azouz R, Gray CM, Nowak LG, McCormick DA (1997) Physiological properties of inhibitory interneurons in cat striate cortex. Cerebral Cortex 7:534–545.

Baskys A (1992) Metabotropic receptors and 'slow' excitatory actions of glutamate agonists in the hippocampus. Trends Neurosci 15:92–96.

Bear M, Singer W (1986) Modulation of visual cortical plasticity by acetylcholine and noradrenaline. Nature 320:172–176.

Bear MF, Malenka RC (1994) Synaptic plasticity: LTP and LTD. Curr Opin Neurobiol 4:389–399.

Bi G, Poo M (1998) Synaptic modifications in cultured hippocampal neurons: Dependence on spike timing, synaptic strength, and postsynaptic cell type. J Neurosci 18:10464–10472.

Bienenstock EL, Cooper LN, Munro PW (1982) Theory for the development of neuron selectivity: Orientation specificity and binocular interaction in visual cortex. J Neurosci 2:32–48.

Bishop C (1995) Neural Networks for Pattern Recognition. Oxford University Press.

Blasdel GG, Fitzpatrick D (1984) Physiological organization of layer 4 in macaque striate cortex. J Neurosci 4:880–895.

Bliss TVP, Collingridge GL (1993) A synaptic model of memory: Long-term potentiation in the hippocampus. Nature 361:31–39.

Borg Graham LJ, Monier C, Fregnac Y (1998) Visual input evokes transient and strong shunting inhibition in visual cortical neurons. Nature 393:369–373.

Braastad BO, Heggelund P (1985) Development of spatial receptive-field organization and orientation selectivity in kitten striate cortex. J Neurophysiol 53:1158–1178.

Bullier J, Henry GH (1979) Laminar distribution of first-order neurons and afferent terminals in cat striate cortex. J Neurophysiol 42:1271–1281.

Cabelli RJ, Hohn A, Shatz CJ (1995) Inhibition of ocular dominance column formation by infusion of NT-4/5 or BDNF. Science 267:1662–1666.

Callaway EM, Katz LC (1992) Development of axonal arbors of layer 4 spiny neurons in cat striate cortex. J Neurosci 12:570–582.

Carandini M, Heeger DJ (1994) Summation and division by neurons in visual cortex. Science 264:1333–1336.

Carandini M, Heeger DJ, Movshon JA (1997) Linearity and normalization in simple cells of the macaque primary visual cortex. J Neurosci 17:8621–8644.

Carandini M, Heeger DJ, Movshon JA (1998) Linearity and gain control in V1 simple cells. In: Cerebral cortex, vol. 13 (Jones EG, Peters A, eds). New York: Plenum.

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

Chance F, Nelson S, Abbott L (1999) Complex cells as cortically amplified simple cells. Nat Neurosci 2:277–282.

Chance FS, Nelson SB, Abbott LF (1997) Effects of synaptic depression of temporal nonlinearities in responses of model simple cells. Soc Neurosci Abstr 23:1266.

Chance FS, Nelson SB, Abbott LF (1998) Synaptic depression and the temporal response characteristics of V1 cells. J Neurosci 18:4785–4799.

Chapman B, Stryker MP (1993) Development of orientation selectivity in ferret visual cortex and effects of deprivation. J Neurosci 13:5251–5262.

Chapman B, Stryker MP, Bonhoeffer T (1996) Development of orientation preference maps in ferret primary visual cortex. J Neurosci 16:6443–6453.

Chapman B, Zahs KR, Stryker MP (1991) Relation of cortical cell orientation selectivity to alignment of receptive fields of the geniculocortical afferents that arborize within a single orientation column in ferret visual cortex. J Neurosci 11:1347–1358.

Cheng H, Chino YM, Smith EL, Hamamoto J et al. (1995) Transfer characteristics of lateral geniculate nucleus X neurons in the cat: Effects of spatial frequency and contrast. J Neurophysiol 74:2548–2557.

Chino YM, Cheng H, Smith ELr, Garraghty PE, Roe AW, Sur M (1994) Early discordant binocular vision disrupts signal transfer in the lateral geniculate nucleus. Proc Natl Acad Sci USA 91:6938–6942.

Chung S, Ferster D (1998) Strength and orientation tuning of the thalamic input to simple cells revealed by electrically evoked cortical suppression. Neuron 20:1177–89.

Cohen MA, Grossberg S (1983) Absolute stability of global pattern formation and parallel memory storage by competitive neural networks. IEEE Transactions on Systems Man and Cybernetics 13:815–826.

Crair MC, Gillespie DC, Stryker MP (1998) The role of visual experience in the development of columns in cat visual cortex. Science 279:566–570.

Crair MC, Malenka RC (1995) A critical period for long-term potentiation at thalamocortical synapses. Nature 375.

Crair MC, Ruthazer ES, Gillespie DC, Stryker MP (1997) Ocular dominance peaks at pinwheel center singularities of the orientation map in cat visual cortex. J Neurophysiol 77:3381–3385.

Dean AF, Tolhurst DJ (1986) Factors influencing the temporal phase of response to bar and grating stimuli for simple cells in the cat straite cortex. Exp Brain Res 62:143–151.

DeAngelis G, Ghose G, Ohzawa I, Freeman R (1999) Functional micro-organization of primary visual cortex: receptive field analysis of nearby neurons. J Neurosci 19:4046–4064.

DeAngelis GC, Ohzawa I, Freeman RD (1993) Spatiotemporal organization of simple-cell receptive fields in the cat's striate cortex. II. Linearity of temporal and spatial summation. J Neurophysiol 69:1118–1135.

DeAngelis GC, Ohzawa I, Freeman RD (1995) Receptive-field dynamics in the central visual pathways. Trends Neurosci 18:451–458.

Debąnne D, Shulz D, Fregnac Y (1998) Activity-dependent regulation of 'on' and 'off' responses in cat visual cortical receptive fields. J Physiol 508:523–548.

DiCarlo J, Johnson K, Hsiao S (1998) Structure of receptive fields in area 3b of primary somatosensory cortex in the alert monkey. J Neurosci 18:2626–2645.

Erwin E, Miller KD (1998) Correlation-based development of ocularly-matched orientation maps and ocular dominance maps: Determination of required input activity structures. J Neurosci 18:9870–9895.

Erwin E, Obermayer K, Schulten K (1995) Models of orientation and ocular dominance columns in the visual cortex: A critical comparison. Neural Comput 7:425–468.

Feller M, Wellis D, Stellwagen D, Werblin F, Shatz C (1996) Requirement for cholinergic synaptic transmission in the propagation of spontaneous retinal waves. Science 272:1182–1187.

Ferster D (1986) Orientation selectivity of synaptic potentials in neurons of cat primary visual cortex. J Neurosci 6:1284–1301.

Ferster D (1988) Spatially opponent excitation and inhibition in simple cells of the cat visual cortex. J Neurosci 8:1172–1180.

Ferster D (1990a) X- and Y-mediated current sources in areas 17 and 18 of cat visual cortex. Vis Neurosci 4:135–145.

Ferster D (1990b) X- and Y-mediated synaptic potentials in neurons of areas 17 and 18 of cat visual cortex. Vis Neurosci 4:115–133.

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

Ferster D, Jagadeesh B (1992) EPSP-IPSP interactions in cat visual cortex studied with in vivo whole-cell patch recording. J Neurosci 12(4):1262–1274.

Ferster D, Lindstrom S (1983) An intracellular analysis of geniculo-cortical connectivity in area 17 of the cat. J Physiol 342:181–215.

Ferster D, Miller K (2000) Neural mechanisms of orientation selectivity in the visual cortex. Annu Rev Neurosci . (To appear.).

Fregnac Y, Imbert M (1984) Development of neuronal selectivity in the primary visual cortex of the cat. Physiol Rev 64:325–434.

Ghosh A, Shatz CJ (1992) Involvement of subplate neurons in the formation of ocular dominance columns. Science 255:1441–1443.

Gil Z, Connors B, Amitai Y (1997) Differential regulation of neocortical synapses by neuromodulators and activity. Neuron 19:679–686.

Gilbert CD (1977) Laminar differences in receptive field properties of cells in cat primary visual cortex. J Physiol 268:391–421.

Godecke I, Chapman B (1998) Effects of on-center retinal ganglion cell blockade on development of visual cortical responses. Soc Neurosci Abstr 24:1051.

Goodhill GJ (1993) Topography and ocular dominance: A model exploring positive correlations. Biol Cybern 69:109–118.

Guillery RW (1972) Binocular competition in the control of geniculate cell growth. J Comp Neurol 144:117–130.

Hammond P, Pomfrett CJ (1990) Influence of spatial frequency on tuning and bias for orientation and direction in the cat's striate cortex. Vision Res 30:359–369.

Hata Y, Stryker MP (1994) Control of thalamocortical afferent rearrangement by postsynaptic activity in developing visual cortex. Science 265:1732–35.

Hawken MJ, Parker AJ (1984) Contrast sensitivity and orientation selectivity in lamina IV of the striate cortex of old world monkeys. Exp Brain Res 54:367–372.

Hawken MJ, Shapley RM, Grosof DH (1992) Temporal frequency tuning of neurons in macaque V1: Effects of luminance contrast and chromaticity. Inv Opth and Vis Sci Suppl 33:955.

Hebb DO (1949) The Organization of Behavior. New York: John Wiley and Sons, Inc.

Heeger DJ (1992) Normalization of cell responses in cat striate cortex. Vis Neurosci 9:181–198.

Hendry S, Schwark H, Jones E, Yan J (1987) Numbers and proportions of gaba-immunoreactive neurons in different areas of monkey cerebral cortex. J Neurosci 7:1503–1519.

Hensch T, Stryker M (1996) Ocular dominance plasticity under metabotropic glutamate receptor blockade. Science 272:554–557.

Hetherington P, Swindale N (1999) Receptive field and orientation scatter studied by tetrode recordings in cat area 17. Vis Neurosci 16:637–652.

Hines M, Carnevale T (1997) The NEURON simulation environment. Neural Computation 9:1179–1209.

Hirsch JA, Alonso JM, Reid RC, Martinez L (1998) Synaptic integration in striate cortical simple cells. J Neurosci 18:9517–9528.

Holub RA, Morton Gibson M (1981) Response of visual cortical neurons of the cat to moving sinusoidal gratings: Response-contrast functions and spatiotemporal interactions. J Neurophysiol 46:1244–1259.

Hubel DH, Wiesel TN (1962) Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. J Physiol 160:106–154.

Issa N, Trachtenberg J, Chapman B, Zahs K, Stryker M (1999) The critical period for ocular dominance plasticity in the ferret's visual cortex. J Neurosci 19:6965–6978.

Jones JP, Palmer LA (1987) An evaluation of the two-dimensional Gabor filter model of simple receptive fields in cat striate cortex. J Neurophysiol 58:1233–1258.

Jones JP, Stepnoski A, Palmer LA (1987) The two-dimensional spectral structure of simple receptive fields in cat striate cortex. J Neurophysiol 59:1212–1232.

Kaplan E, Purpura K, Shapley RM (1987) Contrast affects the transmission of visual information trough the mammalian lateral geniculate nucleus. J Physiol 391:267–288.

Kisvarday Z (1992) Gabaergic networks of basket cells in the visual cortex. Progress in Brain Res 90:385–406.

Komatsu Y (1994) Age-dependent long-term potentiation of inhibitory synaptic transmission in rat visual cortex. J Neurosci 14:6488–6499.

Komatsu Y (1996) $GABA_B$ receptors, monoamine receptors, and postsynaptic inositol trisphosphate-induced Ca^{2+} release are involved in the induction of long-term potentiation at visual cortical inhibitory synapses. J Neurosci 16:6342–6352.

Komatsu Y, Iwakiri M (1993) Long-term modification of inhibitory synaptic transmission in developing visual cortex. Neuroreport 4:907–910.

Kossel A, Lowel S, Bolz J (1995) Relationships between dendritic fields and functional architecture in striate cortex of normal and visually deprived cats. J Neurosci 15:3913–3926.

Krukowski AE, Hoffman A, Miller KD (1998) Correlation-based intracortical connectivity in striate cortex can account for temporal frequency tuning and 'cross-orientation' inhibition. Soc Neurosci Abstr 24:261.

Larkum ME, Zhu JJ, Sakmann B (1999) A new cellular mechanism for coupling inputs arriving at different cortical layers. Nature 398:338–341.

Levine MW, Troy JB (1986) The variability of the maintained discharge of cat dorsal lateral geniculate cells. J Physiol 375:339–359.

Linsenmeier RA, Frishman LJ, Jakiela HG, Enroth Cugell C (1982) Receptive field properties of X and Y cells in the cat retina derived from contrast sensitivity measurements. Vision Res 22:1173–1183.

Linsker R (1986) From basic network principles to neural architecture (series). Proc Natl Acad Sci USA 83:7508–7512, 8390–8394, 8779–8783.

Luhmann H, Prince D (1991) Postnatal maturation of the gabaergic system in rat neocortex. J Neurophysiol 65:247–263.

Malach R, Amir Y, Harel M, Grinvald A (1993) Relationship between intrinsic connections and functional architecture revealed by optical imaging and in vivo targeted biocytin injections in primate striate cortex. Proc Natl Acad Sci USA 90:10469–10473.

Maldonado PE, Godecke I, Gray CM, Bonhoeffer T (1997) Orientation selectivity in pinwheel centers in cat striate cortex. Science 276:1551–1555.

Markram H, Tsodyks M (1996) Redistribution of synaptic efficacy between neocortical pyramidal neurons. Nature 382:807–810.

Martin KA, Whitteridge D (1984) Form, function and intracortical projections of spiny neurones in the striate visual cortex of the cat. J Physiol 353:463–504.

Mastronarde DN (1989) Correlated firing of retinal ganglion cells. Trends Neurosci 12:75–80.

McCormick DA (1990) Membrane properties and neurotransmitter actions. In: The synaptic organization of the brain (Shepard G, ed), pp 32–66. Oxford: Oxford University Press.

McCormick DA, Connors BW, Lighthall JW, Prince DA (1985) Comparative electrophysiology of pyramidal and sparsely spiny stellate neurons of the neocortex. J Neurophysiol 54:782–805.

Meister M, Lagnado L, Baylor DA (1995) Concerted signaling by retinal ganglion cells. Science 270:1207–1210.

Miller KD (1990) Correlation-based models of neural development. In: Neuroscience and connectionist theory (Gluck MA, Rumelhart DE, eds), pp 267–353. Hillsdale, NJ: Lawrence Erlbaum Ass.

Miller KD (1994) A model for the development of simple cell receptive fields and the ordered arrangement of orientation columns through activity-dependent competition between ON- and OFF-center inputs. J Neurosci 14:409–441.

Miller KD (1996) Synaptic economics: Competition and cooperation in synaptic plasticity. Neuron 17:371–374.

Miller KD, Erwin E, Kayser A (1999) Is the development of orientation selectivity instructed by activity? J Neurobiol (to appear).

Miller KD, Keller JB, Stryker MP (1989) Ocular dominance column development: Analysis and simulation. Science 245:605–615.

Miller KD, MacKay DJC (1994) The role of constraints in Hebbian learning. Neural Comput 6:100–126.

Miyashita M, Tanaka S (1992) A mathematical model for the self-organization of orientation columns in visual cortex. NeuroReport 3:69–72.

Movshon JA, Thompson ID, Tolhurst DJ (1978) Spatial and temporal contrast sensitivity of neurones in areas 17 and 18 of the cat visual cortex. J Physiol 283:101–120.

Nelson DA, Katz LC (1995) Emergence of functional circuits in ferret visual cortex visualized by optical imaging. Neuron 15:23–34.

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

Nelson SB (1991a) Temporal interactions in the cat visual system. I. Orientation-selective suppression in the visual cortex. J Neurosci 11:344–356.

Nelson SB (1991b) Temporal interactions in the cat visual system. III. Pharmacological studies of cortical suppression suggest a presynaptic mechanism. J Neurosci 11:369–380.

Nicoll RA (1988) The coupling of neurotransmitter receptors to ion channels in the brain. Science 241:545–551.

Niebur E, Worgotter F (1994) Design principles of columnar organization in visual cortex. Neural Comp 6:602–614.

Obermayer K, Blasdel GG, Schulten K (1992) A statistical mechanical analysis of self-organization and pattern formation during the development of visual maps. Phys Rev A 45:7568–7589.

Olson S, Grossberg S (1998) A neural network model for the development of simple and complex cell receptive fields within cortical maps of orientation and ocular dominance. Neural Networks 11:189–208.

Peichl L, Wassle H (1979) Size, scatter and coverage of ganglion cell receptive field centres in the cat retina. J Physiol 291:117–141.

Peters A, Yilmaz E (1993) Neuronal organization in area-17 of cat visual cortex. Cerebral Cortex 3:49–68.

Press WH, Teukolsky SA, Vetterling WT, Flannery BP (1992) Numerical Recipes in C, second edn. Cambridge: Cambridge University Press.

Priebe NJ, Kayser AS, Krukowski AE, Miller KD (1997) A model of simple cell orientation tuning: The role of synaptic depression. Soc Neurosci Abstr 23:2061.

Rao R, Ballard D (1999) Predictive coding in the visual cortex: a functional interpretation of some extra-classical receptive-field effects. Nat Neurosci 2:79–87.

Reich DS, Victor JD, Knight BW, Ozaki T, Kaplan E (1997) Response variability and timing precision of neuronal spike trains in vivo. J Neurophysiol 77:2836–2841.

Reid RC, Alonso JM (1995) Specificity of monosynaptic connections from thalamus to visual cortex. Nature 378:281–284.

Riesenhuber M, Bauer HU, Brockmann D, Geisel T (1998) Breaking rotational symmetry in a self-organizing map model for orientation map development. Neural Comp 10:717–730.

Ritter H, Martinetz T, Schulten K (1992) Neural Computation and Self-Organizing Maps: An Introduction. Reading, MA: Addison-Wesley.

Rutherford LC, Dewan A, Lauer HM, Turrigiano GG (1997) Brain-derived neurotrophic factor mediates the activity-dependent regulation of inhibition in neocortical cultures. J Neurosci 17:4527–4535.

Sanchez Vives MV, McCormick DA, Nowak LG (1998) Is synaptic depression prevalent in vivo and does it contribute to contrast adaptation. Soc Neurosci Abstr 24:??

Sato H, Katsuyama N, Tamura H, Hata Y, Tsumoto T (1996) Mechanisms underlying orientation slectivity of neurons in the primary visual cortex of the macaque. J Physiol 494:757–771.

Saul AB, Humphrey AL (1990) Spatial and temporal response properties of lagged and nonlagged cells in cat lateral geniculate nucleus. J Neurophysiol 64:206–224.

Saul AB, Humphrey AL (1992) Evidence of input from lagged cells in the lateral geniculate nucleus to simple cells in cortical area 17 of the cat. J Neurophysiol 68:1190–1208.

Scanziani M, Malenka RC, Nicoll RA (1996) Role of intercellular interactions in heterosynaptic long-term depression. Nature 380:446–450.

Sclar G (1987) Expression of "retinal" contrast gain control by neurons of the cat's lateral geniculate nucleus. Exp Brain Res 66:589–596.

Sclar G, Freeman RD (1982) Orientation selectivity in the cat's striate cortex is invariant with stimulus contrast. Exp Brain Res 46:457–461.

Sclar G, Maunsell JH, Lennie P (1990) Coding of image contrast in central visual pathways of the macaque monkey. Vision Res 30:1–10.

Shapley R, Hochstein S (1975) Visual spatial summation in two classes of geniculate cells. Nature 256:411–413.

Shapley RM, Victor JD (1978) The effect of contrast on the transfer properties of cat retinal ganglion cells. J Physiol London 285:275–298.

Sherman SM (1985) Functional organization of the W-, X-, and Y-cell pathways in the cat: A review and hypothesis. Progr Psychobiol Physiol Psychol 11:233–314.

Shouval HZ, Goldberg DH, Jones JP, Shelton WA, Beckerman M, Cooper LN (1998) Lateral connectivity can provide a scaffold for orientation maps: A computational study. Soc Neuro Abs 24(323.15):813.

Sillito A M (1975) The contribution of inhibitory mechanisms to the receptive field properties of neurones in the striate cortex of the cat. J Physiol 250:305–329.

Sirosh J, Miikkulainen R (1994) Cooperative self-organization of afferent and lateral connections in cortical maps. Biol Cybern 71:66–78.

Sirosh J, Miikkulainen R (1997) Topographic receptive fields and patterned lateral interaction in a self-organizing model of the primary visual cortex. Neural Comput 9:577–594.

Skottun BC, Bradley A, Sclar G, Ohzawa I, Freeman RD (1987) The effects of contrast on visual orientation and spatial frequency discrimination: A comparison of single cells and behavior. J Neurophysiol 57:773–786.

Skottun BC, De Valois RL, Grosof DH, Movshon JA, Albrecht DG, Bonds AB (1991) Classifying simple and complex cells on the basis of response modulation. Vision Res 38:1079–86.

Skottun BC, Grosof DH, de Valois RL (1991) On the responses of simple and complex cells to random dot patterns. Vision Res 31:43–46.

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

Song S, Miller K, Abbott L (1999) Competitive hebbian learning through spike-timing-dependent synaptic plasticity. In preparation .

Song S, Varela J, Turrigiano G, Abbott L, Nelson S (1999) The dynamics of synaptic depression at monosynaptic inhibitory inputs to visual cortical pyramidal neurons. In: Proceedings of the computational neuroscience meeting, CNS98 (Bower JM, ed). NY: Plenum. To appear.

Steele P, Mauk M (1999) Inhibitory control of ltp and ltd: Stability of synapse strength. J Neurophysiol 81:1559–1566.

Stratford KJ, Tarczy Hornoch K, Martin KA, Bannister NJ, Jack JJ (1996) Excitatory synaptic inputs to spiny stellate cells in cat visual cortex. Nature 382:258–261.

Swindale NV (1982) A model for the formation of orientation columns. Proc R Soc Lond B 215:211–230.

Tang AC, Bartels AM, Sejnowski TJ (1997) Effects of cholinergic modulation on responses of neocortical neurons to fluctuating input. Cerebral Cortex 7:502–509.

Tarczy Hornoch K (1996) Physiology of Synaptic Inputs to layer IV of Cat Visual Cortex. PhD thesis, Oxford University, England.

Tarczy Hornoch K, Martin KA, Jack JJ, Stratford KJ (1998) Synaptic interactions between smooth and spiny neurones in layer 4 of cat visual cortex in vitro. J Physiol London 508:351–363.

Thomson AM, Deuchars J, West DC (1993) Single axon excitatory postsynaptic potentials in neocortical interneurons exhibit pronounced paired pulse facilitation. Neurosci 54:347–360.

Troyer TW, Krukowski A, Priebe NJ, Miller KD (1998) Contrast-invariant orientation tuning in cat visual cortex: Feedforward tuning and correlation-based intracortical connectivity. J Neurosci 18:5908–5927.

Troyer TW, Miller KD (1997a) Integrate-and-fire neurons matched to physiological f-I curves yield high input sensitivity and wide dynamic range. In: Computational neuroscience: Trends in research 1997 (Bower JM, ed), pp 197–201. New York: Plenum Press.

Troyer TW, Miller KD (1997b) Physiological gain leads to high ISI variability in a simple model of a cortical regular spiking cell. Neural Comput 9:971–983.

Tsodyks MV, Markram H (1997) The neural code between neocortical pyramidal neurons depends on neurotransmitter release probability. Proc Natl Acad Sci USA 94.

Tsumoto T, Suda K (1982) Laminar differences in development of afferent innervation to striate cortex neurones in kittens. Exp Brain Res 45:433–446.

Turrigiano GG, Leslie KR, Desai NS, Rutherford LC, Nelson SB (1998) Activity-dependent scaling of quantal amplitude in neocortical neurons. Nature 391:892–896.

Vidyasagar TR, Siguenza JA (1985) Relationship between orientation tuning and spatial frequency in neurones of cat area 17. Exp Brain Res 57:628–631.

Wassle H, Boycott B, Illing R (1981) Morphology and mosaic of on- and off-beta cells in the cat retina and some functional considerations. Proc R Soc Lond B 212:177–195.

Webster MA, De Valois RL (1985) Relationship between spatial-frequency and orientation tuning of striate-cortex cells. J Opt Soc Am A 2:1124–1132.

Weliky M, Katz LC (1997) Disruption of orientation tuning in primary visual cortex by artificially correlated neuronal activity. Nature 386:680–685.

Weliky M, Katz LC (1999) Correlational structure of spontaneous neuronal activity in the developing lateral geniculate nucleus *in vivo*. Science 283:to appear.

Wiesel TN, Hubel DH (1965) Comparison of the effects of unilateral and bilateral eye closure on cortical unit responses in kittens. J Neurophysiol 28:1029–1040.

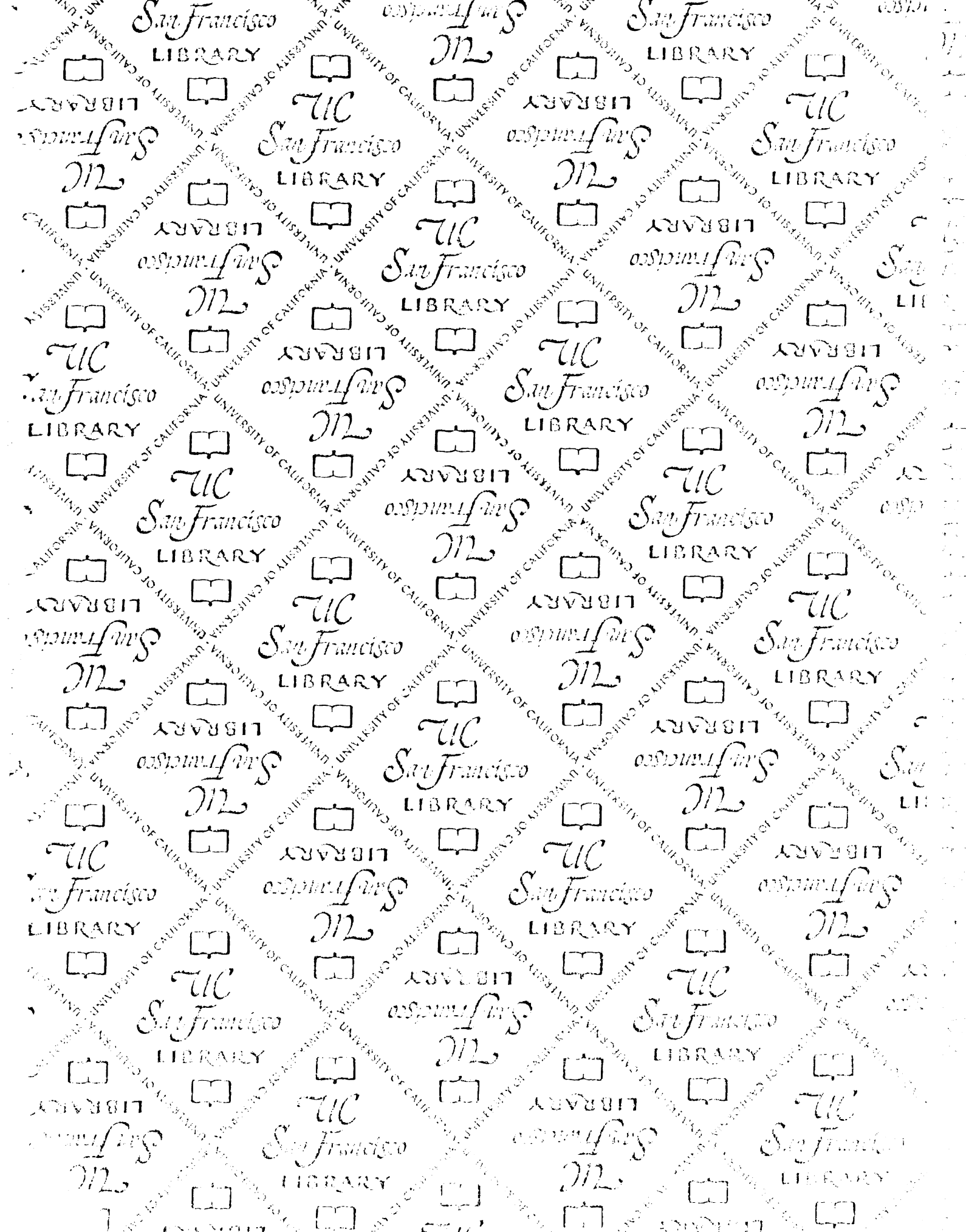
Wong RO (1999) Retinal waves and visual system development. Annu Rev Neurosci 22:29–47.

Wong RO, Oakley DM (1996) Changing patterns of spontaneous bursting activity of On and Off retinal ganglion cells during development. Neuron 16:1087–1095.

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

Xiang Z, Huguenard J, Prince D (1998) Cholinergic switching within neocortical inhibitory networks. Science 281:985–988.

Yuste R, Peinado A, Katz LC (1992) Neuronal domains in developing neocortex. Science 257:666–669.


UC
San Francisco
LIBRARY
UNIVERSITY OF CALIFORNIA