NECO/Miller/Ms. 1232

Equivalence of a Sprouting–and–Retraction Model and Correlation–Based Plasticity Models of Neural Development

Kenneth D. Miller
Depts. of Physiology and Otolaryngology and Neuroscience Graduate Program
W.M. Keck Center for Integrative Neuroscience
Sloan Center for Theoretical Neurobiology at UCSF
University of California
San Francisco, CA 94143-0444
ken@phy.ucsf.edu

June 25, 1997

In press, *Neural Computation*

Send Correspondence to:
Kenneth D. Miller
Dept. of Physiology, UCSF
513 Parnassus
SF, CA 94143-0444

Phone: 415-476-8217
Fax: 415-476-4929

Abstract

A simple model of correlation-based synaptic plasticity via axonal sprouting and retraction (Elliott *et al.*, 1996a) is shown to be equivalent to the class of correlation-based models (Miller *et al.*, 1989), although these were formulated in terms of weight modification of anatomically fixed synapses. Both models maximize the same measure of synaptic correlation, subject to certain constraints on connectivity. Thus, the analyses of the correlation-based models suffice to characterize the behavior of the sprouting-and-retraction model. More detailed models are needed for theoretical distinctions to be drawn between plasticity via sprouting and retraction, weight modification, or a combination.

The model of Elliott *et al.* involves stochastic search through allowed weight patterns for those that improve correlations. That of Miller *et al.* instead follows dynamical equations that determine continuous changes of the weights that improve correlations. The identity of these two approaches is shown to depend on the use of subtractive constraint enforcement in the models of Miller *et al.*. More generally, to model the idea that neural development acts to maximize some measure of correlation subject to a constraint on the summed synaptic weight, the constraint must be enforced subtractively in a dynamical model.

Models of activity-dependent, correlation-based mechanisms of neural development (Miller 1990a, 1996) have, for simplicity, typically utilized weight modifications of anatomically fixed synapses (but see Malsburg, 1979; Fraser and Perkel, 1990; Montague *et al.*, 1991; Colbert *et al.*, 1994; Elliott *et al.*, 1996a,b). However, anatomical changes in connectivity, for example via synaptic sprouting and retraction guided or stabilized by correlation-based rules, also may play important roles in activity-dependent development and learning.

For example, retraction of axons and dendrites plays a role in many forms of activity-dependent development. In ocular dominance plasticity, thalamocortical afferents withdraw from regions of cortex that come to be dominated by afferents serving the opposite eye (LeVay et al., 1978). Dendrites of postsynaptic neurons also have been shown to avoid inappropriate ocular dominance columns (Katz and Constantine-Paton, 1988; Katz et al., 1989; Kossel et al., 1995). At the neuromuscular junction, activity-dependent competition normally leads to anatomical retraction of the axons of all but one input to each muscle fiber (Purves and Lichtman, 1985). However, retraction of an axon at the neuromuscular junction is preceded by significant loss of its physiological synaptic strength (Colman et al., 1997). Thus, these regressive anatomical changes might simply follow physiological change, and not play a leading or guiding role in activity-dependent development.

Stronger evidence for a central role for anatomical changes is found in optic tectum, where there is a continuous retinotopic reorganization of the visual map throughout life. This implies that both retinotopic and, when they occur, ocular dominance maps in this system must be continually maintained amidst constant anatomical rearrangement of inputs (Constantine-Paton et al., 1990; Debski et al., 1990). Striking anatomical changes also occur during visual cortical development. During the normal development of ocular dominance columns in cat visual cortex, there is a huge increase in the overall number of synapses in visual cortex (Cragg, 1975). At the same time, the terminal arbors of thalamic afferents to visual cortex increase in density and in branching complexity, though not in overall extent (Antonini and Stryker, 1993a). However, in both optic tectum and visual cortex, control of synapse number during development appears to be activity-independent (Hayes and Meyer, 1989; Bourgeois et al., 1989), suggesting that activity-dependent processes help determine *which* connections survive but do not influence the overall number of survivors.

Recent results suggest that activity-dependent processes can directly influence axonal sprouting in visual cortex, at least under abnormal conditions causing denervation of a cortical region. In kitten visual cortex, closure of one eye (monocular deprivation) for short times leads to dramatic loss of arborizations of the thalamocortical afferents corresponding to the closed eye (Antonini and Stryker, 1993b). Reverse deprivation (opening the closed eye and closing the open one) following similar periods of deprivation leads to a strong anatomical shift in favor of the originally closed eye (Movshon and Van Sluyters, 1981). Thus, thalamocortical afferents from the originally closed eye appear to show significant axonal sprouting after reverse deprivation, although this sprouting might occur only into regions that have already been denervated by the newly closed eye (Mioche and Singer, 1989). Similarly, following retinal lesions in adult cats, intracortical axonal sprouting appears to occur into the visual cortical region that formerly responded to the lesioned area (Darian-Smith and Gilbert, 1994).

A rapidly-accumulating set of evidence suggests a possible role for neurotrophins in activity-dependent synaptic plasticity in visual cortex and hippocampus (*e.g.* (Cabelli et al., 1995; Fiorentini et al., 1995; Kang and Schuman, 1995; Korte et al., 1995; McAllister et al., 1995); reviewed in (Thoenen, 1995; Bonhoeffer, 1996)). This also suggests a role for anatomical synaptic rearrangements, because in many systems neurotrophins play a strong role in influencing sprouting and retraction of axonal and dendritic branches (*e.g.* Purves and Lichtman, 1985, McAllister *et al.*, 1995). However, it remains unclear whether neurotrophins play a specific instructional role in activity-dependent plasticity, as opposed to, *e.g.*, a non-specific role in regulating overall levels of sprouting. Furthermore, neurotrophins are also involved in weight modification of anatomically fixed synapses (Figurov et al., 1996; Kang and Schuman, 1995; Korte et al., 1995).

Finally, many other experiential modifications can lead to anatomical changes in synaptic connectivity in a variety of other neural structures (Weiler et al., 1995).

Mechanisms of activity-dependent synaptic sprouting and retraction have not been explicitly included in previous correlation-based models for at least two reasons. First, formulation of the dynamics of such mechanisms has seemed forbidding in the absence of improved experimental characterization. Second, the study of anatomically fixed synapses seems potentially adequate to understand the behavior of more general correlation-based models. For example, if the range of retinotopically allowed axonal exploration in a correlation-based sprouting and retraction model is equivalent to the range of initial axonal connections in a similar model using fixed synapses, then both models would explore the same space of possible connections, and both should converge to the same "most-correlated" set of connections within that space.

The purpose of the present paper is twofold. First, I show that there is a precise such equivalence between one very simple sprouting–and–retraction model – that of Elliott *et al.*, 1996a – and previously formulated models using fixed synapses (Miller, 1990a). Second, I show that the use of resource limitations or competitive constraints (Miller and MacKay,

1994) that are linear in the synaptic weights, such as weight normalization, in the sprouting model naturally correspond to subtractive implementation of such constraints in the fixed-synapse models. More generally, subtractive implementation of linear constraints emerges as a natural result of an energy-minimization viewpoint, whereas it appeared quite arbitrary from the viewpoint of dynamical models.

1 Equivalence of a sprouting/retraction model and a fixed-synapse model

Elliott *et al.* (1996a) consider a sprouting-and-retraction model of ocular dominance development (Fig. 1A). A two-dimensional grid of cortical cells receives synapses from two two-dimensional input layers, one layer serving each eye. The input layers represent cells in the lateral geniculate nucleus (LGN). Let each synapse from LGN to cortex be labeled by an index $i, j, \ldots$. Assume that all LGN inputs have two activity states, "on" or "off". Let σ_i be a variable with value 1 if the input to synapse i is on, and value -1 if that input is off. Note that multiple synapses come from the same input cell; the σ's for all synapses from a single input are always identical. Let N_j be the cortical neighborhood of synapse j, defined as the set of synapses on the same cortical cell as synapse j or on any of the four adjacent cortical cells.

Elliott *et al.* posit that synaptic rearrangements occur so as to minimize the energy $E = -\frac{1}{2}\sum_{j,i\in N_j}\sigma_i\sigma_j$; that is, so as to maximize the correlations between the activities of synapses on the same and neighboring cortical cells. Arborizations are taken to be localized, so that synapses may only be made within a certain arborization radius r_A of the input cell. Rearrangements occur subject to this and other constraints on connectivity, to be discussed. Input activity patterns are selected from some ensemble. For each input pattern, a small number of randomly chosen allowed rearrangements are tried. Most of the model results were obtained at "zero temperature", meaning that only E-reducing rearrangements are accepted; a few results were obtained at "finite temperature", meaning that E-increasing rearrangements were accepted probabilistically. The present discussion focuses on the zero-temperature case.

To see the equivalence to previous fixed-synapse models, let Greek letters $\boldsymbol{\alpha}, \boldsymbol{\beta}, \ldots$ represent retinotopic position in the input layers, and Roman letters $\mathbf{x}, \mathbf{y}, \ldots$ position in the cortical layer. Let $S^I(\mathbf{x}, \boldsymbol{\alpha})$ represent the number of synapses from the LGN neuron at position $\boldsymbol{\alpha}$ in input layer I to the cortical cell at $\mathbf{x}$. Let $\sigma^I(\boldsymbol{\alpha})$ be the activity (± 1) of the neuron at position $\boldsymbol{\alpha}$ in input layer I. Define an intracortical function $I(\mathbf{x},\mathbf{y})$ to be 1 if $\mathbf{x}=\mathbf{y}$ or if $\mathbf{x}$ and $\mathbf{y}$ are cortical neighbors, and 0 otherwise. Then the energy $E=-\frac{1}{2}\sum_{j,i\in N_j}\sigma_i\sigma_j$ can be rewritten

$$E=-\frac{1}{2}\sum_{\mathbf{x},\mathbf{y},\boldsymbol{\alpha},\boldsymbol{\beta},I,J} I(\mathbf{x},\mathbf{y})\left[S^I(\mathbf{x},\boldsymbol{\alpha})\sigma^I(\boldsymbol{\alpha})\right]\left[S^J(\mathbf{y},\boldsymbol{\beta})\sigma^J(\boldsymbol{\beta})\right] \tag{1}$$

Assuming that the number of synaptic rearrangements per input pattern is small, so that LGN activity patterns may be averaged over, this energy can be rewritten as (Fig. 1C)

$$E=-\frac{1}{2}\sum_{\mathbf{x},\mathbf{y},\boldsymbol{\alpha},\boldsymbol{\beta},I,J} S^I(\mathbf{x},\boldsymbol{\alpha})I(\mathbf{x},\mathbf{y})C^{IJ}(\boldsymbol{\alpha},\boldsymbol{\beta})S^J(\mathbf{y},\boldsymbol{\beta}) \tag{2}$$

where $C^{IJ}(\boldsymbol{\alpha},\boldsymbol{\beta})=\left\langle\sigma^I(\boldsymbol{\alpha})\sigma^J(\boldsymbol{\beta})\right\rangle$, the angle brackets signify average over input patterns, and the notation E has been retained for $\langle E\rangle$. Define the arbor function by $A(\mathbf{x},\boldsymbol{\alpha})=1$, $|\mathbf{x}-\boldsymbol{\alpha}|\le r_A$; $A(\mathbf{x},\boldsymbol{\alpha})=0$, otherwise. Then the localization of arborizations is taken into account by minimizing E subject to $S^I(\mathbf{x},\boldsymbol{\alpha})=0$ whenever $A(\mathbf{x},\boldsymbol{\alpha})=0$.

The correlation-based models studied in (Miller *et al.*, 1989; Miller, 1990a,b, 1994) begin from equations that describe (i) simple Hebbian or other correlation-based plasticity of anatomically fixed synapses, and (ii) cortical activity as a function of the input activity pattern. These equations are combined, and averaged over the ensemble of input activation patterns, to arrive at equations for the development of synaptic connectivity of the form

$$\frac{d}{dt}S^I(\mathbf{x},\boldsymbol{\alpha})=A(\mathbf{x},\boldsymbol{\alpha})\sum_{\mathbf{y},\boldsymbol{\beta},I,J} I(\mathbf{x},\mathbf{y})C^{IJ}(\boldsymbol{\alpha},\boldsymbol{\beta})S^J(\mathbf{y},\boldsymbol{\beta}) \tag{3}$$

Here, $S^I(\mathbf{x},\boldsymbol{\alpha})$ represents the total synaptic strength — the average synaptic strength times the number of synapses — from $\boldsymbol{\alpha}$ in layer I to $\mathbf{x}$. $A(\mathbf{x},\boldsymbol{\alpha})$, the arbor function, describes the retinotopically allowed anatomical strength of connections from an input at $\boldsymbol{\alpha}$ to a cortical cell at $\mathbf{x}$. $I(x,y)$, the intracortical interaction function, describes the influence of synapses at cortical position $\mathbf{y}$ on the growth of simultaneously active synapses at position $\mathbf{x}$, via intracortical connectivity and/or diffusion of trophic or modulatory factors. The correlation functions $C^{IJ}(\boldsymbol{\alpha},\boldsymbol{\beta})$ describe the correlation in activity between an input at position $\boldsymbol{\beta}$ in LGN layer J, and one at $\boldsymbol{\alpha}$ in LGN layer I. Assuming that $I(\mathbf{x},\mathbf{y})=I(\mathbf{y},\mathbf{x})$ and $C^{IJ}(\boldsymbol{\alpha},\boldsymbol{\beta})=C^{IJ}(\boldsymbol{\beta},\boldsymbol{\alpha})$, Eq. 3 can be easily shown to represent gradient descent minimization of the energy E of Eq. 2 in the variables $T^I(\mathbf{x},\boldsymbol{\alpha})=S^I(\mathbf{x},\boldsymbol{\alpha})/\sqrt{A(\mathbf{x},\boldsymbol{\alpha})}$.[1] Again, this energy is minimized subject to $S^I(\mathbf{x},\boldsymbol{\alpha})=0$ whenever $A(\mathbf{x},\boldsymbol{\alpha})=0$.

[1]Symmetrize Eq. 3 by the transformation $T^I(\mathbf{x},\boldsymbol{\alpha})=S^I(\mathbf{x},\boldsymbol{\alpha})/\sqrt{A(\mathbf{x},\boldsymbol{\alpha})}$, $A(\mathbf{x},\boldsymbol{\alpha})\neq 0$; $T^I(\mathbf{x},\boldsymbol{\alpha})=0$,

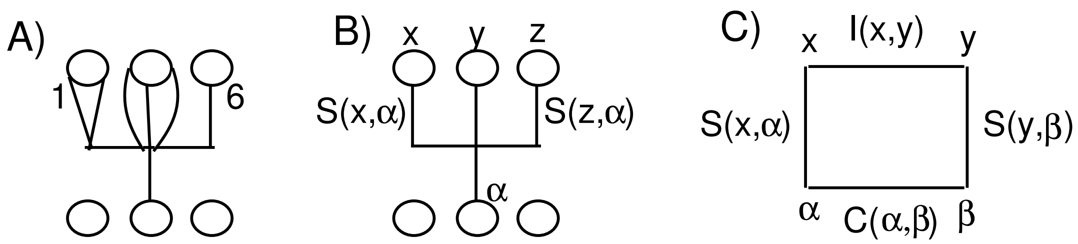

Figure 1: Equivalence of the Models of Elliott *et al.* and of Miller *et al.*.
For simplicity, the models are illustrated for the case of only a single input type (*e.g.*, a single eye).
A. Model of Elliott et al. Each synapse is given an individual label. Here, six synapses are shown from the central presynaptic cell to the three postsynaptic cells; these synapses are labelled from 1 to 6 (only the labels '1' and '6' are shown). The activity at each synapse is $\sigma_i \in \{-1, 1\}$, where i is the synapse label. All synapses from the same input cell have the same activity, hence in this case σ_i is identical for $i = 1, \ldots, 6$. Synaptic connections are made and broken, within retinotopic limits described by an arbor function A, to maximize $\sum_{j,i \in N_j} \sigma_i \sigma_j$, where N_j is the cortical neighborhood of synapse j, defined in the text.
B. Model of Miller et al. The presynaptic cells are labeled by an index $\boldsymbol{\alpha}, \boldsymbol{\beta}, \ldots$, and the postsynaptic cells are labeled by an index $x, y, \ldots$. There is a single variable, $S(\mathbf{x}, \boldsymbol{\alpha})$, describing the total synaptic strength of connection between cells $\boldsymbol{\alpha}$ and $\mathbf{x}$; this is equivalent to the *number* of synapses connecting the two cells in the model of Elliott *et al.*. Thus, given the configuration shown in (A), the equivalent strengths in (B) in the model of Miller *et al.* would be $S(\mathbf{x}, \boldsymbol{\alpha}) = 2$, $S(\mathbf{y}, \boldsymbol{\alpha}) = 3$, $S(\mathbf{z}, \boldsymbol{\alpha}) = 1$.
C. The energy maximized by both models. The graphic shows the factors underlying the interaction energy between two weights, $S(\mathbf{x}, \boldsymbol{\alpha})$ and $S(\mathbf{y}, \boldsymbol{\beta})$. This energy is the *product* of the four factors shown: the two weights, the correlation between the two presynaptic cells $C(\boldsymbol{\alpha}, \boldsymbol{\beta})$, and the intracortical influence between the two postsynaptic locations $I(\mathbf{x}, \mathbf{y})$. Both models maximize an energy (negative of Eq. 2, which is minimized) given by the sum over all weight pairs of this interaction energy.

Thus, an identical energy (Eq. 2; Fig. 1C) is minimized by both the sprouting and retraction model of Elliott *et al.* (1996a) and the correlation-based models of Miller *et al.* (1989), which assume plasticity of anatomically fixed synapses. However, the energies are minimized subject to differing constraints, discussed in the appendix. The constraints imposed in the two models have a similar character, involving upper and/or lower limits on the strengths of individual connections and on the summed presynaptic and/or postsynaptic strength associated with each cell. The choices of constraints do not represent intrinsic differences between sprouting-and-retraction and fixed-synapse models. Either type of model might be run with either set of constraints.

However, a given set of constraints may be dynamically *enforced* in different, inequivalent ways (Miller and MacKay, 1994). I now address the form of enforcement needed to render a dynamical model equivalent to the model of Elliott *et al.*.

2 Constraint Enforcement and Energy Minimization

Constraints on connectivity are utilized in developmental models to incorporate several biological facts. These include: that biological development is competitive, so that differences between the inputs' activities, rather than the amounts of their activities, determines their fate in synaptic competition (Wiesel and Hubel, 1965; Guillery, 1972; Lo and Poo, 1991; Balice-Gordon and Lichtman, 1995); that resources are limited; and that weights are single-signed and finite. Because we know virtually nothing of the mechanisms by which biological competition is achieved, there is little biological guidance as to the form such constraints should take or the manner in which they should be implemented (see discussion in Miller, 1990a, 1996b; Miller and MacKay, 1994).

In the statistical model of Elliott *et al.*, connection patterns are searched for those that decrease the energy. The search consists of a sequence of small random perturbations of synaptic connectivity. Constraints on connectivity are enforced by limiting the perturbations to those that obey the constraints. For example, to enforce constraints preserving total

otherwise (Miller, 1990a; Miller and Stryker, 1990; MacKay and Miller, 1990). Note that $S^I(\mathbf{x},\boldsymbol{\alpha}) = 0$ whenever $A(\mathbf{x},\boldsymbol{\alpha}) = 0$, so no information is lost by this transformation. With the stated assumptions on C and I, the resulting equation for $T^I(\mathbf{x},\boldsymbol{\alpha})$ is symmetric under $(\mathbf{x},\boldsymbol{\alpha}) \leftrightarrow (\mathbf{y},\boldsymbol{\beta})$, and performs gradient descent in the energy

$$E = -\frac{1}{2}\sum_{\mathbf{x},\mathbf{y},\alpha,\boldsymbol{\beta},I,J} T^I(\mathbf{x},\boldsymbol{\alpha})\sqrt{A(\mathbf{x},\boldsymbol{\alpha})}I(\mathbf{x},\mathbf{y})C^{IJ}(\boldsymbol{\alpha},\boldsymbol{\beta})\sqrt{A(\mathbf{y},\boldsymbol{\beta})}T^J(\mathbf{y},\boldsymbol{\beta}) \tag{4}$$

Since $\sqrt{A(\mathbf{x},\boldsymbol{\alpha})}T^I(\mathbf{x},\boldsymbol{\alpha}) = S^I(\mathbf{x},\boldsymbol{\alpha})$, this E is identical to that of Eq. 2.

synaptic strength projected by each afferent, and retinotopically restricting the cortical locations accessible to each afferent, perturbations consist of movement of a randomly chosen synapse to a new cortical location that is chosen randomly from among those retinotopically accessible to the corresponding afferent.

In the dynamical model of Miller *et al.*, a deterministic time derivative of the weights is computed at each time step. The equations must be constrained to ensure that this time derivative respects the constraints on connectivity. There are many ways to achieve this, which are not equivalent either mathematically or in developmental outcome (Miller and MacKay, 1994). Intuitively: the space of allowed connectivity patterns forms a *constraint surface* in the space of weights. The unconstrained time derivative may point off of this surface, in which case it must be corrected by return to the constraint surface. But, to which point on the constraint surface should the weight pattern be moved? This freedom corresponds to the multiple, inequivalent ways of dynamically constraining the equation.

Equivalence to the model of Elliott *et al.* is achieved if constraints achieve energy minimization within the constrained space of possible weight patterns. Thus, given an unconstrained energy and a corresponding unconstrained gradient descent dynamics, the constraint implementation must achieve gradient descent in the energy *on the constraint surface.* Mathematically, this is achieved by *perpendicular projection* of the unconstrained gradient descent dynamics onto the constraint surface.[2]

In previous work (Miller and MacKay, 1994), two methods were studied of enforcing constraints limiting the sum of weights over a cell, subtractive and multiplicative. At each timestep, after adding the unconstrained derivative to the weight vector, subtractive enforcement involves subtracting a fixed amount from each weight to return the weight vector to the constraint surface. Multiplicative enforcement involves multiplying each weight by a fixed amount, and thus subtracting an amount proportional to the weight itself. In the case of a constraint on the sum of the synaptic weights, such as $\sum_{\boldsymbol{\alpha},I} S^I(\mathbf{x},\boldsymbol{\alpha}) = S_{\text{post}}$ — or more generally, of a constraint on a linear combination of the synaptic weights, which yields a hyperplane constraint surface — perpendicular projection of the gradient vector corresponds to subtractive constraint enforcement (Miller and MacKay, 1994).[3] For a hypersphere surface,

[2]Beginning from a point $\mathbf{x}$, the local change in energy per unit movement in direction $\hat{\mathbf{v}}$, $\hat{\mathbf{v}}\cdot\hat{\mathbf{v}} = 1$, is given by $\nabla E(\mathbf{x}) \cdot \hat{\mathbf{v}}$ where $\nabla E(\mathbf{x})$ is the gradient of the energy at $\mathbf{x}$. Thus, the direction of maximum decrease of the energy along the constraint surface is the direction with maximum dot product with the negative of the gradient vector. This is the direction found by perpendicular projection of the negative of the gradient vector — the derivative vector under gradient descent dynamics — onto the constraint surface.

[3]There are some technicalities involving the use of the arbor function, discussed in more detail in (Miller and MacKay, 1992, App. B). Briefly: one must work in the symmetric T representation discussed in foot-

multiplicative enforcement is perpendicular projection. For other surfaces, perpendicular projection cannot be characterized as either multiplicative or subtractive.

If there are multiple constraints, one on each postsynaptic cell (or, one on each presynaptic cell), perpendicular projection onto the full, multi-cell constraint surface is accomplished by perpendicular enforcement of each single-cell constraint.[4] These arguments apply also to restrictions to a constraint region bordered by hyperplanes, such as $S_{\mathrm{post}}^{\min} \leq \sum_{\boldsymbol{\alpha},I} S^I(\mathbf{x},\boldsymbol{\alpha}) \leq S_{\mathrm{post}}^{\max}$. Such a constraint can be imposed dynamically by allowing unconstrained development on the interior of the region, and subtractive enforcement of the constraint at each border. This will result in gradient descent minimization of the energy over the allowed region.

The case of simultaneous constraints on both pre- and postsynaptic cells is similar. Perpendicular constraint enforcement of the full set of constraints minimizes energy. Assuming the single-cell constraint surfaces are hyperplanes, perpendicular enforcement corresponds to subtractive enforcement. However, the details of formulating perpendicular constraints in this case are somewhat more complicated. These are described briefly in (Miller, 1997).

Constraints on the strength of individual weights are also linear constraints. Thus, perpendicular projection onto the hyperplane bounding the constrained volume of allowed synaptic weights is achieved by subtracting off the non-allowed weight change from the gradient descent vector. This corresponds to a "hard nonlinearity" — dynamics are linear within the constrained volume, but no weight changes are allowed that take a weight beyond its allowed limits. Single-synapse constraints fail to commute with constraints over summed pre- and/or postsynaptic weights. To avoid this conflict, the methods for formulating perpendicular constraints described in (Miller, 1997) can be used. In practice, more ad hoc methods are generally used to resolve the conflict; while this seems unlikely to significantly alter the developmental outcome from constrained energy minimization, this has not been

note 1, in which the dynamics are gradient descent. There, the subtractive constraint is a perpendicular projection onto the constraint surface, although it generally does not appear perpendicular in the S representation. In the S representation, the constraint conserving $\sum_{\boldsymbol{\alpha},I} S^I(\mathbf{x},\boldsymbol{\alpha})$ is enforced through subtraction of a multiple of $A(\mathbf{x}-\boldsymbol{\alpha})$ from $\frac{d}{dt}S^I(\mathbf{x},\boldsymbol{\alpha})$. Transforming to the T representation, the constraint conserves $\sum_{\boldsymbol{\alpha},I} \sqrt{A(\mathbf{x}-\boldsymbol{\alpha})}T^I(\mathbf{x},\boldsymbol{\alpha})$, and is enforced through subtraction of a multiple of $\sqrt{A(\mathbf{x}-\boldsymbol{\alpha})}$. Thus, the constraint is perpendicular in the T representation: it conserves $\mathbf{n}(\mathbf{x})\cdot\mathbf{T}(\mathbf{x})$, and is enforced through subtraction of a multiple of $\mathbf{n}(\mathbf{x})$, where $\mathbf{n}(\mathbf{x})$ has elements $\sqrt{A(\mathbf{x}-\boldsymbol{\alpha})}$.

[4]The full constraint surface is the intersection of the constraint surfaces associated with each cell. The normals of the single-cell constraint surfaces are all normal to the full constraint surface and to one another. The sequence of perpendicular projections onto each single-cell constraint surface results in a projection onto the full constraint surface along a linear combination of the single-cell normals, and any such linear combination is perpendicular to the full constraint surface. See (Miller and MacKay, 1992, App. C) for mathematical formulation of the projection operators that implement constraints on a network of cells.

studied.

In summary, perpendicular constraint enforcement achieves gradient descent on the multi-cell constraint surface. If constraints are linear in the synaptic strengths, this is achieved by subtractive enforcement. Subtractive enforcement thus realizes the intuitive picture that the biological dynamics are searching, via a correlation-based mechanism, for a "most correlated" set of inputs subject to such linear constraints. Such a search might occur by sprouting and retraction, by synaptic strength modification, or by some combination. While previous work (Miller and MacKay, 1994) demonstrated the differences in outcome resulting from, and in overall energies minimized by, subtractive vs. multiplicative enforcement, the relationship to energy minimization or correlation maximization *over the constraint surface* was not previously noted.

3 Application of the Equivalence to Understand the Results of Elliott *et al.*.

The main results of the model of Elliott *et al.* are to show that ocular dominance segregation will occur under their model; that the periodicity of the resulting columns corresponds to the arbor diameter; and that the degree of segregation increases with the distance over which inputs within an eye are correlated.[5] All of these results follow directly from prior quantitative analyses of the model of Miller *et al.* (which used subtractive constraint enforcement) (Miller et al., 1989; Miller, 1990a).

The occurrence of ocular dominance column formation was shown in the previous analysis to depend on two factors. First, the input activities must be such that, after averaging over activity patterns, two inputs of the same eye are always better correlated in their activities than two inputs of the opposite eye at the same retinotopic separation, at least for small retinotopic separations; and two same-eye inputs must be no worse correlated than two opposite-eye inputs for larger retinotopic separations up to an arbor radius.[6] This condition

[5] Again, this describes the zero-temperature version of the model. By studying finite temperature, Elliott *et al.* found evidence for a phase transition that separates a high-temperature (high noise in the weight modification process), disordered regime from a low-temperature, ordered regime. Finite temperature results of course also apply to either model; one can construct a stochastic dynamics corresponding to the deterministic dynamics at finite temperature (*e.g.* (van Kampen, 1992, Chap. 9)).

[6] Mathematically: the assumption is made of symmetry between the eyes, so that $C^{\mathrm{LL}} = C^{\mathrm{RR}}$ and $C^{\mathrm{LR}} = C^{\mathrm{RL}}$. Then the condition is that $C^{\mathrm{D}}(\alpha) \equiv C^{\mathrm{LL}}(\alpha) - C^{\mathrm{LR}}(\alpha)$ has the peak of its Fourier transform at frequency 0, or at a frequency that corresponds to a wavelength long relative to the arbor diameter.

is easily met by the activity patterns used by Elliott *et al.*: a circle of activated cells within a single eye, centered at a random position, with all other cells inactive. Second, there must be some locally excitatory interactions between cortical cells, so that the ocular dominance of neighboring cells is coupled. This is achieved by the neighborhood function of Elliott *et al.*, which favors the development of correlated receptive fields on neighboring cortical cells.

The spatial period of the resulting ocular dominance columns was shown previously to depend on the intracortical interaction function and the constraints, if any, on the total strength of projection of each input cell. The intracortical interaction used by Elliott *et al.* is purely excitatory, and one of their constraints is that the total projection from each presynaptic cell is held constant. For this case, the period of ocular dominance columns was previously shown to correspond to an arbor diameter, as found empirically by Elliott *et al.*.

The degree of segregation was shown in the previous analysis to depend on the correlations in input activities. Two factors are involved. First, when the same-eye correlations are larger than the opposite-eye correlations only over distances small relative to an arbor radius, segregation is weak; but as the distance over which the same-eye correlations are larger than the opposite-eye correlations becomes larger, segregation becomes stronger. Second, anti-correlations between the two eyes enhance the degree of segregation, while positive correlations between the two eyes reduce it. Elliott *et al.* find that increasing the radius of the circle of activated cells used in their activity patterns leads to sharper segregation, and that for larger arbor radius larger activity circles are needed to get a similar degree of segregation. Furthermore they consider two models, one of which has anticorrelations between the eyes, the other of which has either reduced anticorrelations or else positive correlations between the eyes, and find that the latter model requires much wider-ranging within-eye correlations (larger activity discs for a given arbor width) to achieve a similar degree of segregation, relative to the former model.[7] These results are all as predicted by the previous analysis. A quantitative account could be made by computing the precise within-

[7] In the "relocation" model of Elliott *et al.*, all synaptic changes occur by choosing an active synapse and relocating it; the change is accepted if it decreases the energy (in the zero-temperature version). Because all changes are conditional on a synapse being active, the correlations $C^{IJ}(\boldsymbol{\alpha},\boldsymbol{\beta}) = \langle\sigma(\boldsymbol{\alpha})\sigma(\boldsymbol{\beta})\rangle$ must also be computed conditional on this. Since the opposite eye is always uniformly inactive, the two eyes are completely anticorrelated, $C^{LR}(\boldsymbol{\alpha},\boldsymbol{\beta}) \equiv -1$ for all $\boldsymbol{\alpha},\boldsymbol{\beta}$. In their "interchange" model, synaptic changes occur by interchanging two synapses, one active, one inactive; again, the change is accepted if it decreases the energy. Because the changes involve an inactive as well as an active synapse, the correlations of an inactive synapse with the inactive neurons in the opposite eye must also be considered in computing the correlations, so the anticorrelations between the eyes will be reduced and/or positive correlations between the eyes may be induced, depending on the sizes of the activity circle and the arbors.

and between-eye correlations determined by their activity patterns and sampling procedures.

Elliott *et al.* also model monocular deprivation. Because they conserve the total strength projected by each presynaptic arbor, they must model deprivation by hand-setting the strength of each eye's arbors. Thus, their model deals only with the arrangements of these inputs, given a fixed difference in input strength between the eyes. As found previously in the model of Miller *et al.*, the period of the ocular dominance stripes is unchanged, but one eye's stripes become thinner and the other's thicker. In the studies of Miller *et al.*, the final strength of each eye's projection, as well as the stripe layout, emerged dynamically.

4 Discussion

I have shown that a simple, stochastic, correlation-based sprouting-and-retraction model of synaptic development can be understood within the framework of correlation-based developmental models (Miller, 1990a,1996), although those models were formulated in terms of modification of anatomically fixed synapses. The various models all descend in an identical energy, corresponding to maximizing a certain measure of synaptic correlations. The analyses of the correlation-based framework accurately predict the outcomes of the sprouting-and-retraction model. The sprouting-and-retraction model emphasizes energy minimization or, equivalently, correlation maximization. This perspective in turn provides a principled reason for the use of subtractive enforcement of linear constraints on connectivity (or for multiplicative enforcement of quadratic constraints).

Thus, the hypothesis in the introduction has been confirmed and sharpened. Models of sprouting and retraction and of modification of anatomically fixed synapses can both explore the same space of possible connections and converge to the same "most-correlated" set of connections within that space. This occurs provided both maximize the same measure of correlation over the constrained set of allowed weight patterns.

The correlation-based framework was developed to allow analysis of the outcome of synaptic competition under a wide group of mechanisms. Development in this framework depends on biologically measurable functions that describe correlations in the thalamic inputs, retinotopic restrictions on arborizations, and intracortical interactions. The dependence on these functions has been characterized through both analysis and simulations (Miller *et al.*, 1989; Miller, 1990a, 1994, 1996). Different biological mechanisms are distinguished by the different predictions they make for the shape of these functions. For example, the model of Elliott *et al.* is restricted to positive intracortical interaction functions.

Previously, such correlation-based models were shown to embrace very simple models of a variety of underlying mechanisms, including activity-dependent competition for diffusible modification factors (*e.g.* neurotrophins) by anatomically fixed synapses (Miller et al., 1989; Miller, 1990a). Sprouting and retraction mechanisms were not previously included because it seemed likely that density-dependent and diffusive terms may arise in simplified models of their dynamics, *e.g.* if there is a tendency for sprouting to occur from more-occupied areas into less-occupied areas; and no such terms were included in the framework studied.[8] Such terms do not arise in the model of Elliott *et al.* because of their assumption that all synaptic changes allowed by the constraints (see appendix) and by the relocation or interchange rule (the effects of which can be incorporated in the correlation function, see footnote 7) are tried with equal probability, and are accepted or rejected according to their contribution to the energy. That is, synaptic densities and gradients do not affect sprouting and retraction probabilities other than through the energy and the constraints.

One potential problem with the fixed-synapse formalism is that it contains an initial bias in favor of weight patterns that involve uniform connectivity: if weights are initialized as small perturbations around some anatomical arbor function, then a weight pattern similar to the arbor function initially has a much larger size than weight patterns orthogonal to it. This can lead to a bias in favor of the development of such patterns, even if such patterns have less favorable energy than others (*e.g.* see discussion of the effect of the DC mode's "head start" in (MacKay and Miller, 1990)). Although the zero-temperature model of Elliott *et al.* does not necessarily avoid this problem, a sprouting-and-retraction framework more generally may provide a natural means of avoiding such bias. The addition of noise, as in the high-temperature version of the Elliott *et al.* model, may also be of benefit.

It is often stated that Hebbian rules result in a set of weights corresponding to the "principal component" of the input data – the principal eigenvector of the input covariance matrix. This is based on the simple case of a linear covariance plasticity rule for a single, linearly activated postsynaptic cell. For this case, multiplicative constraint enforcement does lead to the principal component (*e.g.* (Miller and MacKay, 1994)), which in this case minimizes the energy of Eq. 2 over any *hypersphere* in weight space. However, even for this

---

[8] The probability for an input from $\boldsymbol{\alpha}$ of type I to sprout to $\mathbf{x}$ may increase with its local innervation density, $\sum_{\mathbf{x}'} f(|\mathbf{x}-\mathbf{x}'|)S^I(\mathbf{x}',\boldsymbol{\alpha})$ (here, $f(\mathbf{x})$ is some weighting function); it may decrease with total innervation density at $\mathbf{x}$, $\sum_{J,\boldsymbol{\alpha}'} S^J(\mathbf{x},\boldsymbol{\alpha}')$, or with the gradient of this density $\nabla_{\mathbf{x}} \sum_{J,\boldsymbol{\alpha}'} S^J(\mathbf{x},\boldsymbol{\alpha}')$ (if areas that are less innervated, either absolutely or relative to neighboring areas, are more likely to receive sprouts; note, however, that if total synaptic strength is conserved on each postsynaptic cell, this density is constant and its gradient is zero). The probability might also depend on the input's total projection strength, $\sum_{\mathbf{x}'} S^I(\mathbf{x}',\boldsymbol{\alpha})$, or on the gradient of this total, $\nabla_{\boldsymbol{\alpha}} \sum_{\mathbf{x}'} S^I(\mathbf{x}',\boldsymbol{\alpha})$ (*e.g.* smaller arbors, either in absolute terms or relative to neighboring arbors, might have a better chance of making new sprouts).

case, on the *hyperplane* corresponding to a linear weight constraint, the principal eigenvector is not an energy minimum, and in fact is not in general a "special" point of any kind for the energy. That is, if Hebbian development with linear weight constraints acts to maximize correlations, then even in this simple case the dynamics do not evolve to the principal component of the data. More complex plasticity or activity rules or network connectivity also ensure that the dynamics do not evolve to the principal component. It is therefore quite incorrect to equate Hebbian development with learning of the principal component.

In previous work (Miller et al., 1989; Miller and MacKay, 1994), it was noted that ocular dominance segregation can emerge under subtractive constraint enforcement even in the presence of partial correlation of the activities of the two eyes (as is presumably induced by vision). Under multiplicative enforcement, in contrast, ocular dominance segregation cannot emerge unless there is anti-correlation between the two eyes. Thus, subtractive enforcement seems to give a better match to the biology, at least in this respect. By connecting subtractive enforcement to correlation maximization over the allowed weight patterns, the present work gives a natural biological grounding to a constraint enforcement method that also seems favored (at least relative to multiplicative enforcement) by the match of developmental outcome to biology.

Because subtractive enforcement can lead weights to saturate at their most extreme allowed values, it has been suggested (Miller and MacKay, 1994) that it may be a poor choice for models of adult plasticity, where continuous plasticity occurs in response to changing input activity distributions (Kaas, 1991). However, we demonstrate in (Miller, 1997) that this tendency to saturation can be eliminated by combining pre- and postsynaptic constraints. Elliott *et al.* (1996b) have found that their model, which does not show such saturation (discussed in appendix), can account for some aspects of adult plasticity. Thus, these aspects can be accounted for by subtractively constrained dynamics. However, in their model, the number of synapses projected by different input classes in response to a given input activity regime is set by hand, and only the arrangement of these synapses is determined by the learning rule. Thus, it remains an open question under what conditions the energy-minimizing models with linear weight constraints discussed here can account for the reversible changes in projection strength that occur with changes in input activities in adult cortical plasticity. Presumably, the key requirement is that the perturbations of activity sufficiently shift the energy landscape on the constraint surface so that the locations of the minima are shifted.

Do biological correlation-based mechanisms rely more on sprouting and stabilization/retraction of synapses of fixed strength, or on modification of synaptic strengths? There is thus far little evidence to guide us. The two processes are obviously not mutually exclusive: for example, the decision whether to stabilize or retract newly sprouted synapses may be guided by the strengthening or weakening of their physiological strengths. In kitten visual cortex, the rapidity of monocular deprivation effects, which emerge physiologically within hours (Mioche and Singer, 1989), weakly suggests that physiological changes in synaptic efficacy may precede (and thus perhaps guide) anatomical restructuring of connections. This is also suggested by recent results showing that anatomical changes in thalamocortical afferent arbors are incomplete after four days of monocular deprivation (Antonini and Stryker, 1996), whereas physiological effects of monocular deprivation appear complete after just two days (Hensch et al., 1995). Similarly, at the neuromuscular junction, anatomical elimination of a synapse is preceded by loss of its synaptic efficacy (Colman et al., 1997), including loss of postsynaptic receptors (Balice-Gordon and Lichtman, 1993, 1995). That weight modification and axonal sprouting and retraction may be coupled is suggested by studies in several systems showing that decrease in the number of synapses onto a postsynaptic cell is associated with increase in strength of the remaining synapses (Herrera and Grinnell, 1980, 1981; Jackson and Parks, 1982; Pockett and Slack, 1982; Liu and Tsien, 1995).

Can modeling help us to distinguish between mechanisms relying on sprouting and retraction, and those relying on modification of synaptic strengths? To the extent to which the two are mathematically equivalent, of course, no theoretical distinctions can be drawn. It is conceivable that practical experimental distinctions could be drawn from simple models to which density-dependent and diffusive terms are added, as described above. However, to draw firm theoretical distinctions, more realistic knowledge of the rules governing synaptic sprouting, retraction, and modification is needed. It remains to be determined how such knowledge will modify the basic understandings of correlation-based development already achieved through study of the simple models discussed here.

Appendix: Constraints in the Two Models

The energy E of Eq. 2 is minimized in the two models subject to the following constraints:

1. **Localization of connectivity:** $S^I(\mathbf{x}, \boldsymbol{\alpha}) = 0$ whenever $A(\mathbf{x}, \boldsymbol{\alpha}) = 0$, in both models.
2. **Upper and lower limits on connection strength:**
   - Elliott *et al.*: $0 \leq S^I(\mathbf{x}, \boldsymbol{\alpha})$ for all $I, \mathbf{x}, \boldsymbol{\alpha}$. No explicit upper limit.
   - Miller *et al.*: $0 \leq S^I(\mathbf{x}, \boldsymbol{\alpha}) \leq s_{\max} A(\mathbf{x}, \boldsymbol{\alpha})$ for all $I, \mathbf{x}, \boldsymbol{\alpha}$, where $s_{\max}$ represents the maximum value of a synaptic weight.

3. **Presynaptic total connection strength:**

   - Elliott *et al.*: The total number of synapses per presynaptic neuron is constant and conserved: $\sum_{\mathbf{x}} S^I(\mathbf{x}, \boldsymbol{\alpha}) = S_{\text{pre}}$ for all $I, \boldsymbol{\alpha}$ and for some constant S_{pre}. To model monocular deprivation, this constant is decreased for inputs from one eye.
   - Miller *et al.*: The effects of imposing a limitation on the sum over presynaptic weights, such as $\sum_{\mathbf{x}} S^I(\mathbf{x}, \boldsymbol{\alpha}) = S_{\text{pre}}$ or $0.5 S_{\text{pre}} \leq \sum_{\mathbf{x}} S^I(\mathbf{x}, \boldsymbol{\alpha}) \leq 1.5 S_{\text{pre}}$, are considered, but such a limitation is usually not used.

4. **Postsynaptic total connection strength:**

   - Elliott *et al.*: The total number of synapses per postsynaptic neuron is constrained to remain within some range: $S_{\text{post}}^{\text{min}} \leq \sum_{\boldsymbol{\alpha}, I} S^I(\mathbf{x}, \boldsymbol{\alpha}) \leq S_{\text{post}}^{\text{max}}$ for all $\mathbf{x}$, where $S_{\text{post}}^{\text{min}}$ (generally 1) and $S_{\text{post}}^{\text{max}}$ are constants. A model in which the number of synapses per cortical cell is constant and conserved — that is, in which $\sum_{\boldsymbol{\alpha}, I} S^I(\mathbf{x}, \boldsymbol{\alpha}) = S_{\text{post}}$ for all $\mathbf{x}$ — is also considered.
   - Miller *et al.*: The total synaptic strength per postsynaptic neuron is constant and conserved: $\sum_{\boldsymbol{\alpha}, I} S^I(\mathbf{x}, \boldsymbol{\alpha}) = S_{\text{post}}$ for all $\mathbf{x}$ and for some constant S_{post}.

The constraints imposed in the two models have a similar character. One notable difference is the presence of an upper limit on the strength of individual connections in the model of Miller *et al.*, and the absence of such a limit in the model of Elliott *et al.*. In Miller and MacKay (1994), it was shown that an upper limit on the strength of individual connections is generally needed under subtractive constraints to attain distributed receptive and/or projective fields (see discussion in (Miller, 1997)). Yet Elliott *et al.*, in the absence of an explicit upper weight limit, find apparently stable weight configurations with distributed receptive and projective fields.

There seem to be at least two explanations. First, as discussed in (Miller, 1997), the combination of pre- and postsynaptic constraints can eliminate the need for an upper weight limit. Second, the particular intracortical interaction used by Elliott *et al.* may play a role. The self-interaction L_{ii} of a connection $T_i \equiv T^I(\mathbf{x}, \boldsymbol{\alpha})$ (see footnote 1) with itself under the gradient descent dynamics is $L_{ii} = \sqrt{A(\mathbf{x}, \boldsymbol{\alpha})} I(\mathbf{x}, \mathbf{x}) C^{II}(\boldsymbol{\alpha}, \boldsymbol{\alpha}) \sqrt{A(\mathbf{x}, \boldsymbol{\alpha})}$. The interaction L_{ij} between two different connections from the same input cell, $T_i \equiv T^I(\mathbf{x}, \boldsymbol{\alpha})$ and $T_j \equiv T^I(\mathbf{y}, \boldsymbol{\alpha})$ is $L_{ij} = \sqrt{A(\mathbf{x}, \boldsymbol{\alpha})} I(\mathbf{x}, \mathbf{y}) C^{II}(\boldsymbol{\alpha}, \boldsymbol{\alpha}) \sqrt{A(\mathbf{y}, \boldsymbol{\alpha})}$. Given constraints only on presynaptic cells, an upper weight limit is required if $L_{ii} > |L_{ij}|$ for all $j \neq i$, where j and i are labels for two weights from a single presynaptic cell (Miller and MacKay, 1994). This condition is not satisfied in the model of Elliott *et al.*, because $I(\mathbf{x}, \mathbf{y}) = 1$ both for $\mathbf{y} = \mathbf{x}$ and for $\mathbf{y}$ the neighbor of $\mathbf{x}$, and A is constant where it is nonzero. Thus, a given input may equally well distribute its weights among a cell and its four nearest neighbors, or concentrate its inputs onto the central cell, without altering the interactions between its connections. Thus, even if only presynaptic constraints are applied, the model of Elliott *et al.* may not require an upper weight limit to achieve distributed receptive and projective fields.

Acknowledgements: I thank Terry Elliott for stimulating me to think about these issues, and for useful discussions, and Michael Crair, Virginia De Sa, Ed Erwin, David MacKay, Michael Silver, and Todd Troyer for helpful comments on the manuscript. This work is supported by NIH grant EY11001 and grants from the Searle Scholar's Program and the Lucille P. Markey Charitable Trust.

References

Antonini A. and Stryker M.P. (1993a). Development of individual geniculocortical arbors in cat striate cortex and effects of binocular impulse blockade. *J. Neurosci.*, 13:3549–3573.

Antonini A. and Stryker M.P. (1993b). Rapid remodeling of axonal arbors in the visual cortex. *Science*, 260:1819–1821.

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

Balice-Gordon R.J. and Lichtman J.W. (1993). In vivo observations of pre- and postsynaptic changes during the transition from multiple to single innervation at developing neuromuscular junctions. *J. Neurosci.*, 13:834–855.

Balice-Gordon R.J. and Lichtman J.W. (1995). Long-term synapse loss induced by focal blockade of postsynaptic receptors. *Nature*, 372:519–524.

Bonhoeffer T. (1996). Neurotrophins and activity-dependent development of the neocortex. *Curr. Opin. Neurobiol.*, 6:119–126.

Bourgeois J.-P., Jastreboff P.J., and Rakic P. (1989). Synaptogenesis in visual cortex of normal and preterm monkeys: Evidence for intrinsic regulation of synaptic overproduction. *Proc. Natl. Acad. Sci. USA*, 86:4297–4301.

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

Colbert C.M., Fall C.P., and Levy W.B. (1994). Using adaptive synaptogenesis to model the development of ocular dominance in kitten visual cortex. In Eeckman F. H., editor, *Computation in Neurons and Neural Systems*, pages 139–144. Kluwer, Boston.

Colman H., Nabekura J., and Lichtman J.W. (1997). Alterations in synaptic strength preceding axon withdrawal. *Science*, 275:356–361.

Constantine-Paton M., Cline H.T., and Debski E. (1990). Patterned activity, synaptic convergence and the NMDA receptor in developing visual pathways. *Ann. Rev. Neurosci.*, 13:129–154.

Cragg B.G. (1975). The development of synapses in the visual system of the cat. *J. Comp. Neurol.*, 160:147–166.

Darian-Smith C. and Gilbert C.D. (1994). Axonal sprouting accompanies functional reorganization in adult cat striate cortex. *Nature*, 368:737–740.

Debski E.A., Cline H.T., and Constantine-Paton M. (1990). Activity-dependent tuning and the NMDA receptor. *J. Neurobiol.*, 21:18–32.

Elliott T., Howarth C.I., and Shadbolt N.R. (1996a). Axonal processes and neural plasticity. I: Ocular dominance columns. *Cerebral Cortex*, 6:781–788.

Elliott T., Howarth C.I., and Shadbolt N.R. (1996b). Axonal processes and neural plasticity. II: Adult somatosensory maps. *Cerebral Cortex*, 6:789–793.

Figurov A., Pozzo-Miller L.D., Olafsson P., Wang T., and Lu B. (1996). Regulation of synaptic responses to high-frequency stimulation and LTP by neurotrophins in the hippocampus. *Nature*, 381:706–709.

Fiorentini A., Berardi N., and Maffei L. (1995). Nerve growth factor preserves behavioral visual acuity in monocularly deprived kittens. *Vis. Neurosci.*, 12:51–55.

Fraser S.E. and Perkel D.H. (1990). Competitive and positional cues in the patterning of nerve connections. *J. Neurobiol.*, 21:51–72.

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

Hayes W.P. and Meyer R.L. (1989). Impulse blockade by intraocular tetrodoxin during optic regeneration in goldfish: HRP-EM evidence that the formation of normal numbers of optic synapses and the elimination of exuberant optic fibers is activity independent. *J. Neurosci.*, 9:1414–1423.

Hensch T.K., Crair M.C., Ruthazer E.S., Fagiolini M., Gillespie D.C., and Stryker M.P. (1995). Robust two-day ocular dominance plasticity revealed by single-unit recording and intrinsic signal imaging of kitten area 17. *Soc. Neuro. Abs.*, 21:2023.

Herrera A. and Grinnell A. (1980). Transmitter release from frog motor nerve terminals depends on motor unit size. *Nature*, 287:649–651.

Herrera A. and Grinnell A. (1981). Contralateral denervation causes enhanced transmitter release from frog motor nerve terminals. *Nature*, 291:495–497.

Jackson H. and Parks T. (1982). Functional synapse elimination in the developing avian cochlear nucleus with simultaneous reduction in cochlear nerve axon branching. *J. Neurosci.*, 2:1736–1743.

Kaas J.H. (1991). Plasticity of sensory and motor maps in adult mammals. *Ann. Rev. Neurosci.*, 14:137–167.

Kang H. and Schuman E.M. (1995). Long-lasting neurotrophin-induced enhancement of synaptic transmission in the adult hippocampus. *Science*, 267:1658–1662.

Katz L.C. and Constantine-Paton M. (1988). Relationships between segregated afferents and postsynaptic neurons in the optic tectum of three-eyed frogs. *Journal of Neuroscience*, 8:3160–3180.

Katz L.C., Gilbert C.D., and Wiesel T.N. (1989). Local circuits and ocular dominance columns in monkey striate cortex. *Journal of Neuroscience*, 9:1389–1399.

Korte M., Carroll P., Wolf E., Brem G., Thoenen H., and Bonhoeffer T. (1995). Hippocampal long-term potentiation is impaired in mice lacking brain-derived neurotrophic factor. *Proc. Natl. Acad. Sci. USA*, 92:8856–8860.

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

LeVay S., Stryker M.P., and Shatz C.J. (1978). Ocular dominance columns and their development in layer IV of the cat's visual cortex: a quantitative study. *J. Comp. Neurol.*, 179:223–244.

Liu G. and Tsien R.W. (1995). Properties of synaptic transmission at single hippocampal synaptic boutons. *Nature*, 375:404–408.

Lo Y.-J. and Poo M.-M. (1991). Activity-dependent synaptic competition in vitro: Heterosynaptic suppression of developing synapses. *Science*, 254:1019–1022.

MacKay D.J.C. and Miller K.D. (1990). Analysis of Linsker's applications of Hebbian rules to linear networks. *Network*, 1:257–298.

Malsburg C. von der (1979). Development of ocularity domains and growth behavior of axon terminals. *Biol. Cyber.*, 32:49–62.

McAllister A.K., Lo D.C., and Katz L.C. (1995). Neurotrophins regulate dendritic growth in developing visual cortex. *Neuron*, 15:791–803.

Miller K.D. (1990a). Correlation-based models of neural development. In Gluck M.A. and Rumelhart D.E., editors, *Neuroscience and Connectionist Theory*, pages 267–353. Lawrence Erlbaum Ass., Hillsdale, NJ.

Miller K.D. (1990b). Derivation of linear Hebbian equations from a nonlinear Hebbian model of synaptic plasticity. *Neural Comput.*, 2:321–333.

Miller K.D. (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 K.D. (1996a). Receptive fields and maps in the visual cortex: Models of ocular dominance and orientation columns. In Domany E., van Hemmen J.L., and Schulten K., editors, *Models of Neural Networks III*, pages 55–78. Springer-Verlag, NY. Available as ftp://ftp.keck.ucsf.edu/pub/ken/miller95.ps.

Miller K.D. (1996b). Synaptic economics: Competition and cooperation in synaptic plasticity. *Neuron*, 17:371–374.

Miller K.D. (1997). Simultaneous constraints on pre- and post-synaptic cells: Effects on synaptic saturation. *(Unpublished).* Available as ftp://ftp.keck.ucsf.edu/pub/ken/pre_post.ps.gz.

Miller K.D., Keller J.B., and Stryker M.P. (1989). Ocular dominance column development: Analysis and simulation. *Science*, 245:605–615.

Miller K.D. and MacKay D.J.C. (1992). The role of constraints in Hebbian learning. Technical Report Memo 19, Caltech, Pasadena, CA. Available as ftp://ftp.keck.ucsf.edu/pub/ken/constraints-tr.ps.

Miller K.D. and MacKay D.J.C. (1994). The role of constraints in Hebbian learning. *Neural Comput.*, 6:100–126.

Miller K.D. and Stryker M.P. (1990). The development of ocular dominance columns: Mechanisms and models. In Hanson S.J. and Olson C.R., editors, *Connectionist Modeling and Brain Function: The Developing Interface*, pages 255–350. MIT Press/Bradford, Cambridge, MA.

Mioche L. and Singer W. (1989). Chronic recordings from single sites of kitten striate cortex during experience-dependent modifications of receptive-field properties. *Journal of Neurophysiology*, 62:185–197.

Montague P.R., Gally J.A., and Edelman G.M. (1991). Spatial signaling in the development and function of neural connections. *Cereb. Cortex*, 1:199–220.

Movshon J.A. and Van Sluyters R.C. (1981). Visual neural development. *Ann. Rev. Psych.*, 32:477–522.

Pockett S. and Slack J. (1982). Pruning of axonal trees results in increased efficacy of surviving nerve terminals. *Brain Res.*, 243:350–353.

Purves D. and Lichtman J.W. (1985). *Principles of Neural Development.* Sinauer Associates Inc., Sunderland, MA.

Thoenen H. (1995). Neurotrophins and neuronal plasticity. *Science*, 270:593–598.

van Kampen N.G. (1992). *Stochastic Processes in Physics and Chemistry, 2nd Ed.* North-Holland, Amsterdam.

Weiler I.J., Hawrylak N., and Greenough W.T. (1995). Morphogenesis in memory formation: synaptic and cellular mechanisms. *Behav. Brain Res.*, 66:1–6.

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