Analysis of Linsker's Simulations of Hebbian Rules

David J. C. MacKay
Computation and Neural Systems, Caltech 164–30 CNS,
Pasadena, CA 91125 USA

Kenneth D. Miller
Department of Physiology, University of California,
San Francisco, CA 94143–0444 USA

Linsker has reported the development of center-surround receptive fields and oriented receptive fields in simulations of a Hebb-type equation in a linear network. The dynamics of the learning rule are analyzed in terms of the eigenvectors of the covariance matrix of cell activities. Analytic and computational results for Linsker's covariance matrices, and some general theorems, lead to an explanation of the emergence of center-surround and certain oriented structures. We estimate criteria for the parameter regime in which center-surround structures emerge.

Linsker (1986, 1988) has studied by simulation the evolution of weight vectors under a Hebb-type teacherless learning rule in a feedforward linear network. The equation for the evolution of the weight vector **w** of a single neuron, derived by ensemble averaging the Hebbian rule over the statistics of the input patterns, is[1]

$$\frac{\partial}{\partial t} w_i = k_1 + \sum_j (Q_{ij} + k_2) w_j \quad \text{subject to } -w_{max} \leq w_i \leq w_{max} \qquad (1.1)$$

---

[1]Our definition of equation 1.1 differs from Linsker's by the omission of a factor of $1/N$ before the sum term, where N is the number of synapses. Also, Linsker allowed more general hard limits, $n_E - 1 \leq w_i \leq n_E, 0 < n_E < 1$, which he implemented either directly or by allowing a fraction n_E of synapses to be excitatory $(0 \leq w_i^+ \leq 1)$ and the remaining fraction $1 - n_E$ to be inhibitory $(-1 \leq w_i^- \leq 0)$. These two formulations are essentially mathematically equivalent; this equivalence depends on the fact that the spatial distributions of inputs and correlations in activity among inputs were taken to be independent of whether the inputs were excitatory or inhibitory. Linsker summarized results for $0.35 \leq n_E \leq 0.65$ for his layer $\mathcal{B} \rightarrow \mathcal{C}$, but did not report any dependence of results on n_E within this range and focused discussion on $n_E = 0.5$. At higher layers only $n_E = 0.5$ was discussed. Equation 1.1 is equivalent to $n_E = 0.5$. Our analysis does not depend critically on this choice; what is critical is that the origin be well within the interior of the hypercube of allowed synaptic weights, so that initial development is linear.

where $\mathbf{Q}$ is the covariance matrix of activities of the inputs to the neuron. The covariance matrix depends on the covariance function, which describes the dependence of the covariance of two input cells' activities on their separation in the input field, and on the location of the synapses, which is determined by a synaptic density function. Linsker used a gaussian synaptic density function. Similar equations have been developed and studied by others (Miller et al. 1986, 1989).

Depending on the covariance function and the two parameters k_1 and k_2, different weight structures emerge. Using a gaussian covariance function (his layer $\mathcal{B} \to \mathcal{C}$), Linsker reported the emergence of nontrivial weight structures, ranging from saturated structures through center-surround structures to bilobed-oriented structures.

The analysis in this paper examines the properties of equation 1.1. We concentrate on the gaussian covariances in Linsker's layer $\mathcal{B} \to \mathcal{C}$. We give an explanation of the structures reported by Linsker and discuss criteria for the emergence of center-surround weight structures. Several of the results are more general, applying to any covariance matrix $\mathbf{Q}$. Space constrains us to postpone general discussion, technical details, and discussion of other model networks, to a future publication (MacKay and Miller 1990).

2 Analysis in Terms of Eigenvectors _____

We write equation 1.1 as a first-order differential equation for the weight vector $\mathbf{w}$:

$$\dot{\mathbf{w}} = (\mathbf{Q} + k_2 \mathbf{J})\mathbf{w} + k_1 \mathbf{n} \quad \text{subject to} \quad -w_{\max} \leq w_i \leq w_{\max} \tag{2.1}$$

where $\mathbf{J}$ is the matrix $J_{ij} = 1 \ \forall i, j$, and $\mathbf{n}$ is the DC vector $n_i = 1 \ \forall i$. This equation is linear, up to the hard limits on w_i. These hard limits define a hypercube in weight space within which the dynamics are confined. We make the following assumption:

Assumption 1. The principal features of the dynamics are established before the hard limits are reached. When the hypercube is reached, it captures and preserves the existing weight structure with little subsequent change.

The matrix $\mathbf{Q} + k_2 \mathbf{J}$ is symmetric, so it has a complete orthonormal set of eigenvectors[2] $\mathbf{e}^{(a)}$ with real eigenvalues λ_a. The linear dynamics within the hypercube can be characterized in terms of these eigenvectors, each of

---

[2]The indices a and b will be used to denote the eigenvector basis for $\mathbf{w}$, while the indices i and j will be used for the synaptic basis.

which represents an independently evolving weight configuration. First, equation 2.1 has a fixed point at

$$\mathbf{w}^{FP} = -k_1(\mathbf{Q} + k_2 \mathbf{J})^{-1}\mathbf{n} = -k_1 \sum_a \frac{\mathbf{e}^{(a)} \cdot \mathbf{n}}{\lambda_a} \mathbf{e}^{(a)} \tag{2.2}$$

Second, relative to the fixed point, the component of $\mathbf{w}$ in the direction of an eigenvector grows or decays exponentially at a rate proportional to the corresponding eigenvalue. Writing $\mathbf{w}(t) = \sum_a w_a(t)\mathbf{e}^{(a)}$, equation 2.1 yields

$$w_a(t) - w_a^{FP} = [w_a(0) - w_a^{FP}]e^{\lambda_a t} \tag{2.3}$$

Thus, the principal emergent features of the dynamics are determined by the following three factors:

1. The principal eigenvectors of $\mathbf{Q} + k_2 \mathbf{J}$, that is, the eigenvectors with largest positive eigenvalues. These are the fastest growing weight configurations.

2. Eigenvectors of $\mathbf{Q} + k_2 \mathbf{J}$ with negative eigenvalue. Each is associated with an attracting constraint surface, the hyperplane defined by $w_a = w_a^{FP}$.

3. The location of the fixed point of equation 1.1. This is important for two reasons: (a) it determines the location of the constraint surfaces and (b) the fixed point gives a "head start" to the growth rate of eigenvectors $\mathbf{e}^{(a)}$ for which $|w_a^{FP}|$ is large compared to $|w_a(0)|$ (see Fig. 3).

3 Eigenvectors of Q

We first examine the eigenvectors and eigenvalues of $\mathbf{Q}$. The principal eigenvector of $\mathbf{Q}$ dominates the dynamics of equation 2.1 for $k_1 = 0$, $k_2 = 0$. The subsequent eigenvectors of $\mathbf{Q}$ become important as k_1 and k_2 are varied. Some numerical results on the spectrum of $\mathbf{Q}$ have appeared in Linsker (1987, 1990) and Miller (1990). Analyses of the spectrum when output cells are laterally interconnected appear in Miller et al. (1986, 1989).

3.1 Properties of Circularly Symmetric Systems. If an operator commutes with the rotation operator, its eigenfunctions can be written as eigenfunctions of the rotation operator. For Linsker's system, in the continuum limit, the operator $\mathbf{Q} + k_2 \mathbf{J}$ is unchanged under rotation of the system. So the eigenfunctions of $\mathbf{Q} + k_2 \mathbf{J}$ can be written as the product of a radial function and one of the angular functions $\cos l\theta$, $\sin l\theta$, $l = 0, 1, 2 \ldots$. To describe these eigenfunctions we borrow from quantum mechanics the notation $n = 1, 2, 3 \ldots$ and $l =$ s, p, d$\ldots$ to denote the function's total number of nodes $= 0, 1, 2 \ldots$ and number of angular

Name	Eigenfunction	λ/N
1s	$e^{-r^2/2R}$	lC/A
2p	$r\cos\theta e^{-r^2/2R}$	l^2C/A
2s	$(1 - r^2/r_0^2)e^{-r^2/2R}$	l^3C/A

$$R = \frac{C}{2}\left(1 + \sqrt{1 + 4A/C}\right)$$
$$l = \frac{R-C}{R} \quad 0 < l < 1)$$
$$r_0^2 = \frac{2A}{\sqrt{1 + 4A/C}}$$

Table 1: The First Three Eigenfunctions of the Operator $Q(\mathbf{r}, \mathbf{r}')$. $Q(\mathbf{r}, \mathbf{r}') = e^{-|\mathbf{r}-\mathbf{r}'|^2/2C}e^{-r'^2/2A}$, where C and A denote the characteristic sizes of the co-variance function and synaptic density function. $\mathbf{r}$ denotes two-dimensional spatial position relative to the center of the synaptic arbor, and $r = |\mathbf{r}|$. The eigenvalues λ are normalized by the effective number of synapses $N = 2\pi A$.

nodes $= 0, 1, 2 \ldots$, respectively. For example, "2s" and "2p" both denote eigenfunctions with one node, which is radial in 2s and angular in 2p (see Fig. 1).

For monotonic and nonnegative covariance functions, we conjecture that the leading eigenfunctions of Q are ordered in eigenvalue by their numbers of nodes such that the eigenfunction $[nl]$ has larger eigenvalue than both $[(n+1)l]$ and $[n(l+1)]$. This conjecture is obeyed in analytical and numerical results we have obtained for Linsker's and similar systems. The general validity of this conjecture is under investigation.

3.2 Analytic Calculations for $k_2 = 0$. We have solved analytically for the first three eigenfunctions and eigenvalues of the covariance matrix for layer $B \rightarrow C$ of Linsker's network, in the continuum limit (Table 1). 1s, the function with no changes of sign, is the principal eigenfunction of Q; 2p, the bilobed-oriented function, is the second eigenfunction; and 2s, the center-surround eigenfunction, is third.[3]

Figure 1a shows the first six eigenfunctions for layer $B \rightarrow C$ of Linsker (1986).
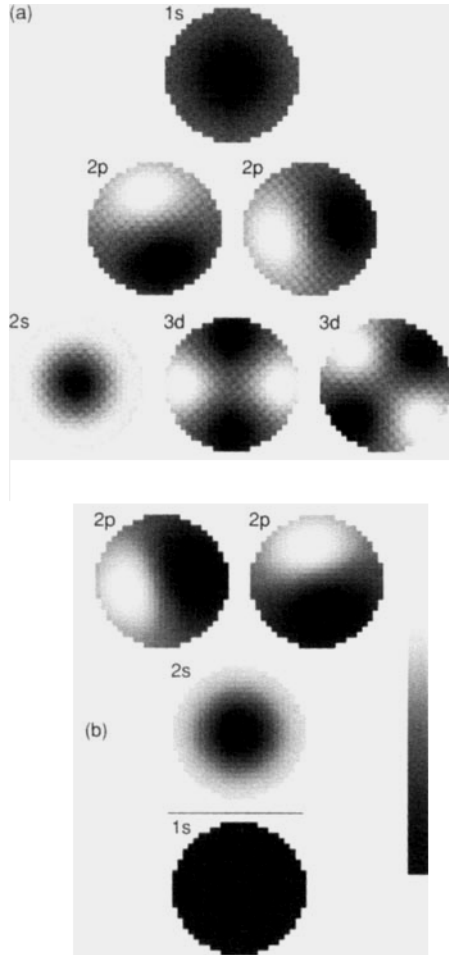
---

[3] 2s is degenerate with 3d at $k_2 = 0$.

Figure 1: Eigenfunctions of the operator $Q + k_2 J$. In each row the eigenfunctions have the same eigenvalue, with largest eigenvalue at the top. Eigenvalues (in arbitrary units): (a) $k_2 = 0$: 1s, 2.26; 2p, 1.0; 2s and 3d, 0.41. (b) $k_2 = -3$: 2p, 1.0; 2s, 0.66; 1s, -17.8. The gray scale indicates the range from maximum negative to maximum positive synaptic weight within each eigenfunction. Eigenfunctions of the operator $(e^{-|\mathbf{r}-\mathbf{r}'|^2/2C} + k_2)e^{-r'^2/2A}$ were computed for $C/A = 2/3$ (as used by Linsker for most layer $\mathcal{B} \rightarrow \mathcal{C}$ simulations) on a circle of radius 12.5 grid intervals, with $\sqrt{A} = 6.15$ grid intervals.

4 The Effects of the Parameters k_1 and k_2 _____

Varying k_2 changes the eigenvectors and eigenvalues of the matrix $Q+k_2J$. Varying k_1 moves the fixed point of the dynamics with respect to the origin. We now analyze these two changes, and their effects on the dynamics.

Definition. Let $\hat{n}$ be the unit vector in the direction of the DC vector n. We refer to $(w \cdot \hat{n})$ as the *DC component* of w. The DC component is proportional to the sum of the synaptic strengths in a weight vector. For example, 2p and all the other eigenfunctions with angular nodes have zero DC component. Only the s-modes have a nonzero DC component.

4.1 General Theorem: The Effect of k_2. We now characterize the effect of adding k_2J to *any* covariance matrix Q.

Theorem 1. *For any covariance matrix Q, the spectrum of eigenvectors and eigenvalues of $Q + k_2J$ obeys the following:*

1. *Eigenvectors of Q with no DC component, and their eigenvalues, are unaffected by k_2.*

2. *The other eigenvectors, with nonzero DC component, vary with k_2. Their eigenvalues increase continuously and monotonically with k_2 between asymptotic limits such that the upper limit of one eigenvalue is the lower limit of the eigenvalue above.*

3. *There is at most one negative eigenvalue.*

4. *All but one of the eigenvalues remain finite. In the limits $k_2 \to \pm\infty$ there is a DC eigenvector $\hat{n}$ with eigenvalue $\to k_2N$, where N is the dimensionality of Q, that is, the number of synapses.*

The properties stated in this theorem, whose proof is in MacKay and Miller (1990), are summarized pictorially by the spectral structure shown in Figure 2.

4.2 Implications for Linsker's System. For Linsker's circularly symmetric systems, all the eigenfunctions with angular nodes have zero DC component and are thus independent of k_2. The eigenvalues that vary with k_2 are those of the s-modes. The leading s-modes at $k_2 = 0$ are 1s, 2s; as k_2 is decreased to $-\infty$, these modes transform continuously into 2s, 3s respectively (Fig. 2).[4] 1s becomes an eigenvector with negative eigenvalue, and it approaches the DC vector $\hat{n}$. This eigenvector enforces a constraint $w \cdot \hat{n} = w^{FP} \cdot \hat{n}$, and thus determines that the final average synaptic strength is equal to $w^{FP} \cdot n/N$.

_____

[4]The 2s eigenfunctions at $k_2 = 0$ and $k_2 = -\infty$ both have one radial node, but are not identical functions.
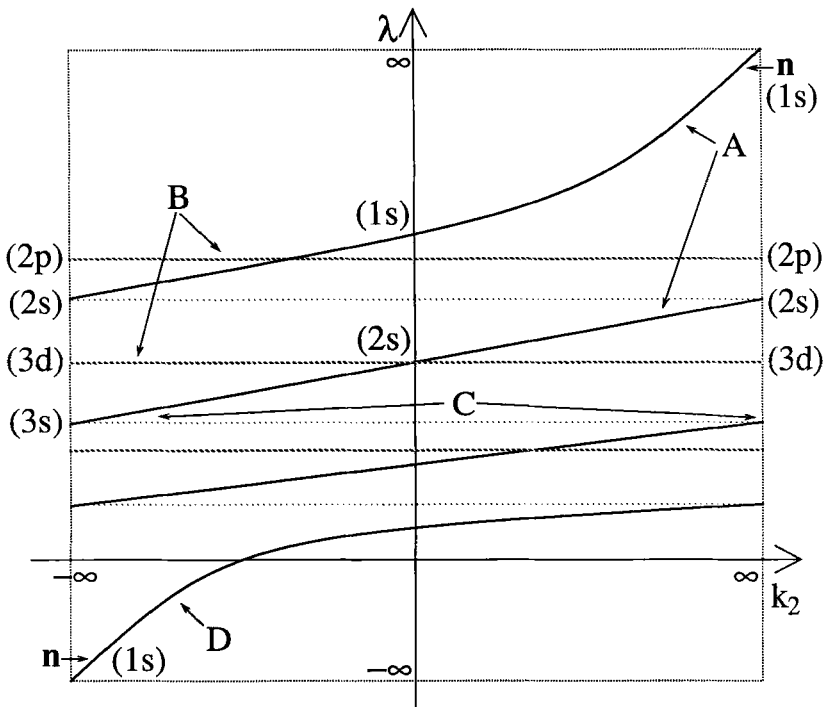
Figure 2: General spectrum of eigenvalues of $Q + k_2 J$ as a function of k_2.
A: Eigenvectors with DC component. B: Eigenvectors with zero DC component. C: Adjacent DC eigenvalues share a common asymptote. D: There is only one negative eigenvalue. The annotations in parentheses refer to the eigenvectors of Linsker's system.

Linsker (1986) used $k_2 = -3$. This value of k_2 is sufficiently large that the properties of the $k_2 \rightarrow -\infty$ limit hold (MacKay and Miller 1990), and in the following we concentrate interchangeably on $k_2 = -3$ and $k_2 \rightarrow -\infty$. The computed eigenfunctions for Linsker's system at layer $\mathcal{B} \rightarrow \mathcal{C}$ are shown in Figure 1b for $k_2 = -3$. The principal eigenfunction is 2p. The center-surround eigenfunction 2s is the principal *symmetric* eigenfunction, but it still has smaller eigenvalue than 2p.

4.3 Effect of k_1. Varying k_1 changes the location of the fixed point of equation 2.1. From equation 2.2, the fixed point is displaced from the origin only in the direction of eigenvectors that have nonzero DC

component, that is, only in the direction of the s-modes. This has two important effects, as discussed in Section 2: (1) The s-modes are given a head start in growth rate that increases as k_1 is increased. In particular, the principal s-mode, the center-surround eigenvector 2s, may outgrow the principal eigenvector 2p. (2) The constraint surface is moved when k_1 is changed. For large negative k_2, the constraint surface fixes the average synaptic strength in the final weight vector. To leading order in $1/k_2$, Linsker showed that the constraint is $\sum w_j = k_1/|k_2|$.[5]

4.4 Summary of the Effects of k_1 and k_2. We can now anticipate the explanation for the emergence of center-surround cells: For $k_1 = 0$, $k_2 = 0$, the dynamics are dominated by 1s. The center-surround eigenfunction 2s is third in line behind 2p, the bilobed function. Making k_2 large and negative removes 1s from the lead. 2p becomes the principal eigenfunction and dominates the dynamics for $k_1 \simeq 0$, so that the circular symmetry is broken. Finally, increasing $k_1/|k_2|$ gives a head start to the center-surround function 2s. Increasing $k_1/|k_2|$ also increases the final average synaptic strength, so large $k_1/|k_2|$ also produces a large DC bias. The center-surround regime therefore lies sandwiched between a 2p-dominated regime and an all-excitatory regime. $k_1/|k_2|$ has to be large enough that 2s dominates over 2p, and small enough that the DC bias does not obscure the center-surround structure. We now estimate this parameter regime.

5 Criteria for the Center-Surround Regime

We use two approaches to determine the DC bias at which 2s and 2p are equally favored. This DC bias gives an estimate for the boundary between the regimes dominated by 2s and 2p.

1. **Energy Criterion:** We first estimate the level of DC bias at which the weight vector composed of (2s plus DC bias) and the weight vector composed of (2p plus DC bias) are energetically equally favored. This gives an estimate of the level of DC bias above which 2s will dominate under simulated annealing, which explores the entire space of possible weight configurations.

2. **Time Development Criterion:** Second, we estimate the level of DC bias above which 2s will dominate under simulations of time development of equation 1.1. We estimate the relationship between the parameters such that, starting from a typical random distribution of initial weights, the 2s mode reaches the saturating hypercube at the same time as the 2p mode.

---

[5]To next order, this expression becomes $\sum w_j = k_1/|k_2 + \bar{q}|$, where $\bar{q} = \langle Q_{ij} \rangle$, the average covariance (averaged over i and j). The additional term largely resolves the discrepancy between Linsker's g and $k_1/|k_2|$ in Linsker (1986).

Both criteria will depend on an estimate of the complex effect of the weight limits $-w_{max} \le w_i \le w_{max}$. (Without this hypercube of saturation constraints, 2p will always dominate the dynamics of equation 1.1 after a sufficiently long time.) We introduce $g = k_1/(|k_2|Nw_{max})$ as a measure of the average synaptic strength induced by the DC constraint, such that $g = 1$ means all synapses equal w_{max}.[6] Noting that a vector of amplitude $\sqrt{N}w_{max}$ has rms synaptic strength w_{max}, we make the following estimate of the constraint imposed by the hypercube (discussed further in MacKay and Miller 1990):

Assumption 2. When the DC level is constrained to be g, the component $h(g)$ in the direction of a typical unit AC vector at which the hypercube constraint is "reached" is $h(g) = \sqrt{N}w_{max}(1 - g)$.

Assumptions 1 and 2 may not adequately characterize the effects of the hypercube on the dynamics, so the numerical estimates of the precise locations of the boundaries between the regions may be in error. However, the qualitative picture presented by these boundaries is informative.

5.1 Energy Criterion. Linsker suggested analysis of equation 1.1 in terms of the energy function on which the dynamics perform constrained gradient descent. The energy of a configuration $\mathbf{w} = \sum w_a \mathbf{e}^{(a)}$ is

$$E = -\frac{1}{2}\sum_a \lambda_a w_a^2 - \sqrt{N}k_1 \sum_a w_a n_a \tag{5.1}$$

where n_a is the DC component of eigenvector $\mathbf{e}^{(a)}$. We consider two configurations, one with w_{2p} equal to its maximum value $h(g)$ and $w_{2s} = 0$, and one with $w_{2p} = 0$ and $w_{2s} = \text{sign}(n_{2s})h(g)$. The component w_{1s} is the same in both cases. All the other components are assumed to be small and to contribute no bias in energy between the two configurations. The energies of these configurations will be our estimates of the energies of saturated configurations obtained by saturating 2p and 2s, respectively, subject to the constraints. We compare these two energies and find the DC level $g = g^E$ at which they are equal:[7]

$$g^E = \frac{1}{1 + 2\frac{|n_{2s}k_2|}{(\lambda_{2p} - \lambda_{2s})/N}}$$

For Linsker's layer $\mathcal{B} \rightarrow \mathcal{C}$ connections, our estimate of g^E is 0.16.

5.2 Time Development Criterion. The energy criterion does not take into account the initial conditions from which equation 1.1 starts. We now derive a second criterion that attempts to do this.

---

[6]This is equal to twice Linsker's g.

[7]λ/N is written as a single entity because $\lambda \propto N$. Also $n_{2s} \sim 1/k_2$, so $n_{2s}k_2$ tends to a constant as $k_2 \rightarrow \infty$.
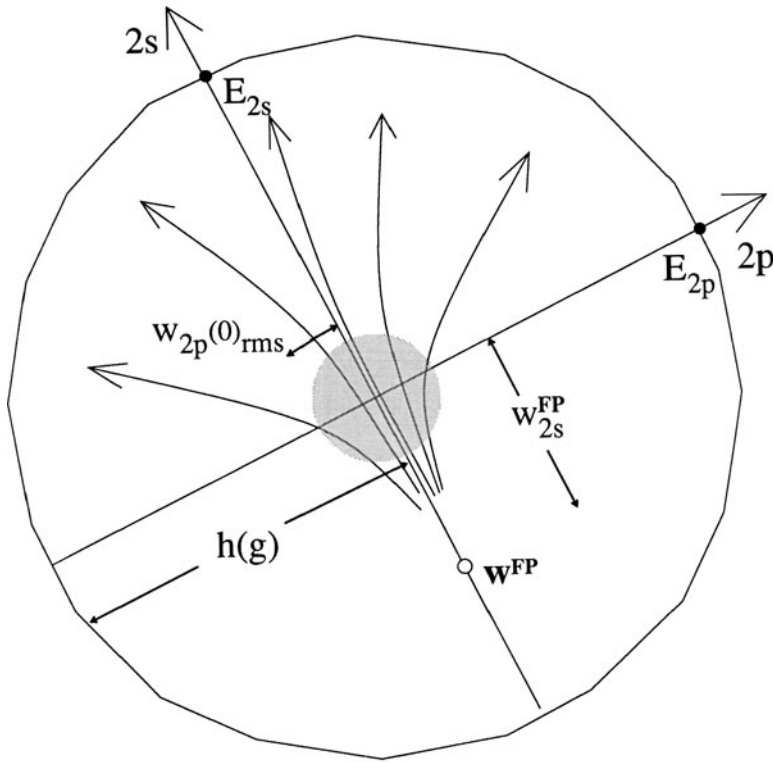
Figure 3: Schematic diagram illustrating the criteria for 2s to dominate. The polygon of size $h(g)$ represents the hypercube. *Energy criterion*: The points marked E_{2p} and E_{2s} show the locations at which the energy estimates were made. *Time development criterion*: The gray cloud surrounding the origin represents the distribution of initial weight vectors. If $w_{2p}(0)$ is sufficiently small compared to w_{2s}^{FP}, and if the hypercube is sufficiently close, then the weight vector reaches the hypercube in the direction of 2s before w_{2p} has grown appreciably.

If the initial random component in the direction of 2p, $w_{2p}(0)$, is sufficiently smallcompared to w_{2s}^{FP}, which provides 2s with a head start, then w_{2p} may never start growing appreciably before the growth of w_{2s} saturates (Fig. 3). The initial component $w_{2p}(0)$ is a random quantity whose

typical magnitude can be estimated statistically from the weight initial-ization parameters. $w_{2p}(0)_{rms}$ scales as $1/\sqrt{N}$ relative to the nonrandom quantity w_{2s}^{FP}. Hence the initial relative magnitude of w_{2p} can be made arbitrarily small by increasing N, and the emergence of center-surround structures may be achieved at any g by using an N sufficiently large to suppress the initial symmetry breaking fluctuations.

We estimate the boundary between the regimes dominated by 2s and 2p by finding the choice of parameters such that $w_{2p}(t)$ and $w_{2s}(t)$ reach the hypercube at the same time. We evaluate the time t_{2s} at which w_{2s} reaches the hypercube.[8] Our estimate of the typical starting component for 2p is $w_{2p}(0)_{rms} = \sqrt{2}\sigma(g)w_{max}$ where $\sigma(g)$ is a dimensionless standard deviation derived in MacKay and Miller (1990). We set $w_{2p}(t_{2s}) = h(g)$, and solve for N^*, the number of synapses above which w_{2s} reaches the hypercube before w_{2p}, in terms of g:

$$\sqrt{N^*} = \frac{\sqrt{2}\sigma(g)}{(1-g)} \left(1 + \frac{1-g}{g}\frac{\lambda_{2s}/N}{|n_{2s}k_2|}\right)^{\lambda_{2p}/\lambda_{2s}} \qquad (5.2)$$

5.3 Discussion of the Two Criteria. Figure 4 shows g^E and $N^*(g)$. The two criteria give different boundaries. In regime A, 2p is estimated to both emerge under equation 1.1, and to be energetically favored. Sim-ilarly, in regime C, 2s is estimated to dominate equation 1.1, and to be energetically favored. In regime D, the initial fluctuations are so big that although 2s is energetically favored, symmetry breaking structures can dominate equation 1.1.[9] Lastly, in regime B, although 2p is energetically favored, 2s will reach saturation first because N is sufficiently large that the symmetry breaking fluctuations are suppressed. Whether this satu-rated 2s structure will be stable, or whether it might gradually destabilize into a 2p-like structure, is not predicted by our analysis.[10] The possible difference between simulated annealing and equation 1.1 makes it clear that if initial conditions are important (regimes B and D), the use of sim-ulated annealing on the energy function as a quick way of finding the outcome of equation 1.1 may give erroneous results.

Figure 4 also shows the areas in the parameter space in which Linsker made the simulations he reported. The agreement between experiment and our estimated boundaries is reasonable.

---

[8]We set $w_{2s}(0) = 0$, neglecting its fluctuations, which for large N are negligible compared with w_{2s}^{FP}.

[9]If the initial component of 2s is *toward* the fixed point, the 2s component must first shrink to zero before it can then grow in the opposite direction. Thus, large fluctuations may either hinder or help 2s, while they always help 2p.

[10]In a one-dimensional model system we have found that both cases may be obtained, depending sensitively on the parameters.
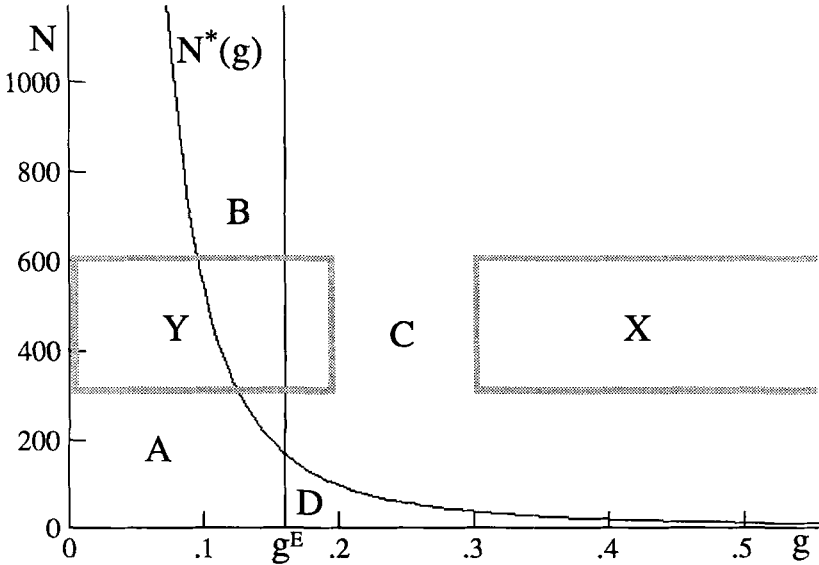
Figure 4: Boundaries estimated by the two criteria for $C/A = 2/3$. To the left of the line labeled g^E, the energy criterion predicts that 2p is favored; to the right, 2s is favored. Above and below the line $N^*(g)$, the time development criterion estimates that 2s and 2p, respectively, will dominate equation 1.1. The regions X, Y, mark the regimes studied by Linsker: (X) $N = 300 - 600$, $g = 0.3$–0.6: the region in which Linsker reported robust center-surround; (Y) $N = 300$–600, $g <\sim 0.2$: asymmetric center-surround structures and (near $g = 0$) bilobed cells.

6 Conclusions and Discussion

For Linsker's $B \rightarrow C$ connections, we predict four main parameter regimes for varying k_1 and k_2.[11] These regimes, shown in Figure 5, are dominated by the following weight structures:

$k_2 = 0$, $k_1 = 0$ The principal eigenvector of Q, 1s.

$k_2 = $large positive   The flat DC weight vector, which leads
    and/or $k_1 = $ large   to the same saturated structures as 1s.

------

[11]not counting the symmetric regimes $(k_1, k_2) \leftrightarrow (-k_1, k_2)$ in which all the weight structures are inverted in sign.
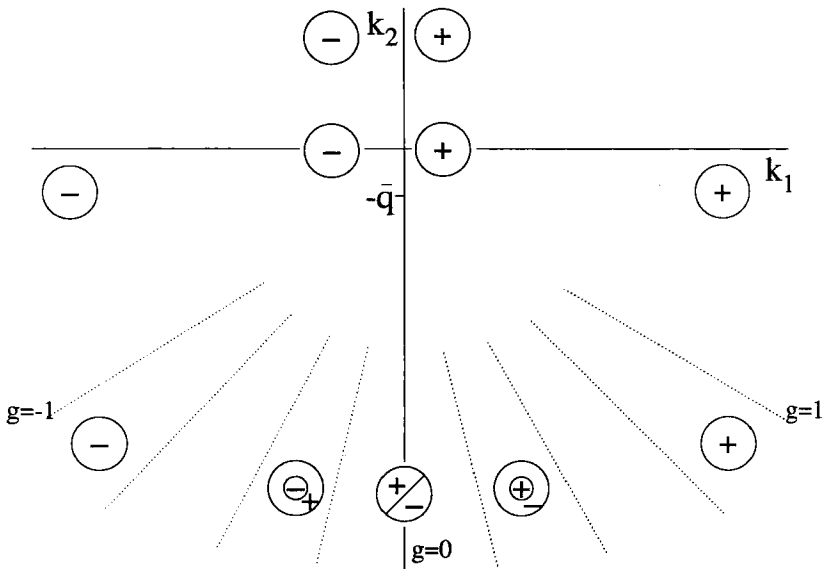
Figure 5: Parameter regimes for Linsker's system. The DC bias is approximately constant along the radial lines, so each of the regimes with large negative k_2 is wedge shaped.

k_2 =large negative, $k_1 \simeq 0$	The principal eigenvector of $Q + k_2 J$ for $k_2 \to -\infty$, 2p.
k_2 =large negative, k_1 = intermediate	The principal eigenvector of $Q + k_2 J$ for $k_2 \to -\infty$ with nonzero DC component, 2s. The size of this regime can depend on the size of the symmetry-breaking fluctuations, and hence on the number of synapses.

Higher layers of Linsker's network can be analyzed in terms of the same four regimes; the principal eigenvectors are altered, so that different structures can emerge (MacKay and Miller 1990).

Linsker suggested that the emergence of center-surround structures may depend on the peaked synaptic density function that he used (Linsker 1986, p. 7512). However, with a flat ("pillbox") density function, the eigenfunctions are qualitatively unchanged, so we expect that center-surround structures may emerge by the same mechanism.

The development of the interesting cells in Linsker's layer $B \to C$ depends on the use of negative synapses and on the use of the terms k_1 and k_2 to enforce a constraint on the final percentages of positive and negative synapses. Both of these may be biologically problematic (Miller 1990; MacKay and Miller 1990). A linear Hebb rule like Linsker's can be derived without the use of negative synapses by examining the *difference* between the innervation strengths of two equivalent excitatory projections, for example, left-eye and right-eye inputs (Miller et al. 1989) or ON-center and OFF-center inputs (Miller 1989). However, in this case the constants k_1 and k_2 disappear from the equation for the development of the difference of synaptic strengths because these constants take on equal values for each of the two equivalent populations. Therefore, there will only be one regime, in which the principal eigenvector of Q dominates. Such a model can nonetheless develop interesting receptive field structures if oscillations exist in the covariance functions of the input layer, and particularly if lateral interactions are introduced in the output layer (Linsker 1987; Miller *et al.* 1989; Miller 1989, 1990).

References

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

Linsker, R. 1987. Towards an organizing principle for perception: Hebbian synapses and the principle of optimal neural encoding. *IBM Research Report RC 12830*.

Linsker, R. 1988. Self-organization in a perceptual network. *Computer* **21**(3), 105–117.

Linsker, R. 1990. Designing a sensory processing system: What can be learned from principal components analysis? *Proc. Int. Joint Conf. on Neural Networks, Jan. 1990*, M. Caudill, ed., pp. II:291–97. Lawrence Erlbaum, Hillsdale, NJ.

MacKay, D. J. C., and Miller, K. D. 1990. Analysis of Linsker's application of Hebbian rules to linear networks. *Network*, to appear.

Miller, K. D. 1989. Orientation-selective cells can emerge from a Hebbian mechanism through interactions between ON- and OFF-center inputs. *Soc. Neurosci. Abstr.* **15**, 794.

Miller, K. D. 1990. Correlation-based mechanisms of neural development. In *Neuroscience and Connectionist Theory*, M.A. Gluck and D.E. Rumelhart, eds., pp. 267–353. Lawrence Erlbaum, Hillsdale, NJ.

Miller, K. D., Keller, J. B., and Stryker, M. P. 1986. Models for the formation of ocular dominance columns solved by linear stability analysis. *Soc. Neurosci. Abstr.* **12**, 1373.

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