Effect of Correlated LGN Firing Rates on Predictions for Monocular Eye Closure vs Monocular Retinal Inactivation

Brian S. Blais

Department of Science and Technology, Bryant University, Smithfield RI and

Institute for Brain and Neural System s, Brown University, Providence RI

Leon N Cooper

Department of Physics, Brown University, Providence RI and

Institute for Brain and Neural Systems, Brown University, Providence RI

Harel Z. Shouval

Department of Neurobiology and Anatomy,

University of Texas Medical School at Houston, Houston TX and

Institute for Brain and Neural Systems, Brown University, Providence RI

(Dated: November 13, 2009)

Monocular deprivation experiments can be used to distinguish between different ideas concerning properties of cortical synaptic plasticity. Monocular deprivation by lid suture causes a rapid disconnection of the deprived eye connected to cortical neurons whereas total inactivation of the deprived eye produces much less of an ocular dominance shift. In order to understand these results one needs to know how lid suture and retinal inactivation affect neurons in the lateral geniculate nucleus (LGN), that provide the cortical input. Recent experimental results by Linden et al. (2009) show that monocular lid suture and monocular inactivation do not change the mean firing rates of LGN neurons but that lid suture reduces correlations between adjacent neurons whereas monocular inactivation leads to correlated firing. These, somewhat surprising, results contradict assumptions that have been made to explain the outcomes of different monocular deprivation protocols. Based on these experimental results we modify our assumptions about inputs to cortex during different deprivation protocols and show their implications when combined with different cortical plasticity rules.

Using theoretical analysis, random matrix theory, and simulations we show that high levels of correlations reduce the ocular dominance shift in learning rules that depend on homosynaptic depression (i.e. Bienenstock Cooper and Munro (BCM) type rules), consistent with experimental results, but have the opposite effect in rules that depend on heterosynaptic depression (i.e. Hebbian/Principal Component Analysis (PCA) type rules).

PACS numbers: 87.19.L-, 87.19.ll, 87.19.lw,

I. INTRODUCTION

Receptive fields in visual cortex are modifiable in the early period of an animals postnatal development. This is thought to depend on synaptic plasticity [1, 2]; the detailed dynamics of such receptive field modifiability has been used to infer the precise form of synaptic plasticity[3, 4]. In a classical paradigm, called monocular deprivation (MD), vision through one eye is deprived in early development. In this paradigm cells in visual cortex tend to disconnect from the deprived eye [5]. We have previously shown how variants of deprivation can be used to distinguish between different classes of learning rules: rules that depend on homosynaptic modification, such as the rule proposed by Bienenstock Cooper and Munro (BCM)[4, 6] and rules that depend on heterosynaptic modification [7] such as the Hebbian-based Oja rule, or PCA rule [4, 8]. Experiments have shown that if monocular deprivation is produced by monocular lid closure (MC) then a rapid loss of response to the deprived eye occurs, while if the retina is inactivated by an injection of TTX (MI), significantly less loss is observed [9, 10]. These results are consistent with homosynaptic BCM-like learning rules. However, the theoretical analysis relies on the, seemingly reasonable, assumption that in the inactivated case (MI) activity in the lateral geniculate nucleus (LGN), which is the cortical input, is reduced compared to the lid closure case (MC). This assumption has been questioned by new experimental results.

In a recent study[11] the activity of neurons in LGN has been recorded during normal vision, when the eye-lid of the experimental animals was sutured and when TTX was injected into the eye. The recordings were conducted on awake animals while they watched movie clips and sinusoidal gratings. The surprising result of these experiments is that MI did not reduce mean activity in LGN when compared to MC; however MI caused an increase in correlations between different cells within the LGN. Previous experimental results in ferret LGN [12, 13], and recent results in mouse LGN [11] indicate that the activity of nearby cells in LGN are correlated, that this activity falls off as a function of the distance between the receptive fields of the cells, and that these correlations exist even in the absence of retinal activity.

In this paper we examine the impact of input correlations during deprivation experiments on two different iconic examples of homosynaptic and heterosynaptic learning rules; BCM and PCA. We find that the consequences of the PCA rule are inconsistent with experimental results but that large correlations in LGN can significantly slow down the loss of response to the deprived eye for BCM neurons in agreement with experiment. Further experimental work to quantitatively determine the correlations within LGN in MI would permit more detailed comparison of theory

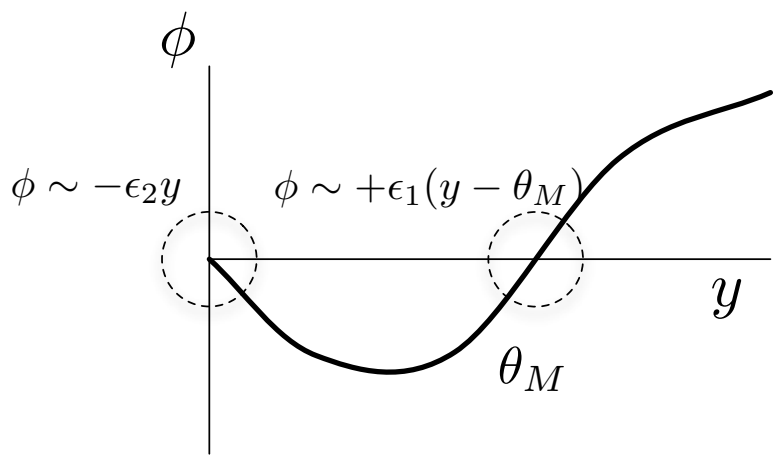

FIG. 1: The modification function, ϕ, for BCM versus the cell output, y. The function has a zero crossing at $y = 0$ and one at $y = \theta_M$, the modification threshold. When the cell is selective its responses to inputs are close to these zero-crossings, so we expand the ϕ function around these and keep only the linear terms, $\phi \sim -\epsilon_2 y$ and $\phi \sim +\epsilon_1(y - \theta_M)$, respectively.

with experiment.

A. Monocular Deprivation: The Original BCM Argument

For completeness, we summarize the original argument from the BCM theory on the consequences of monocular deprivation (MD).[4] It is important to note that, while the analysis provided in this paper assumes a linear neuron, the conclusions are not limited to linear neurons. The numerical simulations presented (see Section IV B) use a non-linear output sigmoid[3] with *quantitatively* similar results.

In the BCM theory, the left and right weights between LGN and cortical cells, $\mathbf{w}^l$ and $\mathbf{w}^r$ respectively, modify depending on the input to those weights, $\mathbf{x}^l$ and $\mathbf{x}^r$, and a function, ϕ, depending on the output of the entire cell, y, and a modification threshold, θ_M as shown in Figure 1.

$$\mathbf{w}^{l(r)} = \phi(y, \theta_M)\mathbf{x}^{l(r)} \tag{1}$$

In monocular deprivation the deprived eye is presented with noise, $\mathbf{n}$, and the open eye receives patterned input, $\mathbf{x}$. If the left eye is the open eye, the inputs become

$$\left\{\mathbf{x}^l, \mathbf{x}^r\right\} \Rightarrow \left\{\mathbf{x}^o, \mathbf{x}^d\right\} \equiv \{\mathbf{x}, \mathbf{n}\} \tag{2}$$

where we use the superscripts o and d for open- and deprived-eyes, respectively.

If MD begins when the receptive fields have already reached their selective fixed points, and we further assume (without loss of generality) that the threshold, θ_M, adjusts very quickly, then we have for the open eye

$$\begin{aligned} \mathbf{w}^o \cdot \mathbf{x}_i^o &= \theta_M \quad i = 1 \text{ (preferred input)} \\ \mathbf{w}^o \cdot \mathbf{x}_i^o &= 0 \quad i > 1 \text{ (non-preferred input)} \end{aligned} \tag{3}$$

where the preferred inputs are input patterns to which the neuron is selective, and the non-preferred inputs are input patterns to which the neuron is not selective. (In a realistic environment, the actual responses fall with some distribution peaking near zero or θ_M.)

Expanding ϕ around zero and θ_M, we get

$$\begin{aligned} \phi &\simeq +\epsilon_1(y - \theta_M) \quad y \text{ near } \theta_M \\ \phi &\simeq -\epsilon_2 y \quad y \text{ near zero} \end{aligned} \tag{4}$$

The output of the cell for the different patterns entering the open eye is

$$\begin{aligned} y &\simeq \theta_M + \mathbf{w^d} \cdot \mathbf{n} \quad i = 1 \text{ (preferred input)} \\ y &\simeq \mathbf{w^d} \cdot \mathbf{n} \quad i > 1 \text{ (non-preferred input)} \end{aligned} \tag{5}$$

so that the deprived-eye weights modify as

$$\dot{\mathbf{w}}^d \simeq \begin{cases} +\epsilon_1(\theta_M + \mathbf{w}^d \cdot \mathbf{n} - \theta_M)\mathbf{n} & \text{preferred patterns} \\ -\epsilon_2(\mathbf{w}^d \cdot \mathbf{n})\mathbf{n} & \text{non-preferred patterns} \end{cases} \tag{6}$$

For the deprived eye, averaged over the environment and time, we have

$$\begin{aligned} \left\langle \dot{w}_i^d \right\rangle &\simeq -\sum_j \epsilon < w_j^d n_j n_i > = -\sum_j \epsilon < w_j^d >< n_j n_i > \\ &= -\sum_j \epsilon < w_j^d >< n^2 > Q_{ji} \end{aligned} \tag{7}$$

where

$$\begin{aligned} \epsilon &\equiv -(\epsilon_2 N_{\text{non-pref}} - \epsilon_1 N_{\text{pref}})/(N_{\text{non-pref}} + N_{\text{pref}}) \\ < n^2 > &= < n_i^2 > \text{ for all } i \\ < n^2 > Q_{ij} &\equiv < n_i n_j > \end{aligned}$$

and N_{pref} and $N_{\text{non-pref}}$ are the number of preferred and non-preferred patterns, respectively. Note that the diagonal elements of $\mathbf{Q}$ are 1.

If n_i and n_j are uncorrelated in space, both with mean zero

$$\langle n_i n_j \rangle = \langle n_i \rangle \langle n_j \rangle = 0 \text{ when } i \neq j \tag{8}$$

yielding the usual result,

$$\dot{w}_i^d \simeq -\epsilon n^2 w_i^d \tag{9}$$

Thus the deprived-eye weights decrease, and decrease faster if there is more noise (i.e. higher n^2) from the deprived-eye. This deprivation effect is stronger for more selective neurons.

B. Monocular Deprivation: The Original Hebbian (PCA) Argument

The dynamics of Hebbian learning rules are determined by the eigenvectors of the correlation matrix with the largest eigenvalues, or the principal components (PCA) [4, 14]. In one learning rule proposed by Oja in 1982 the weight vector converges to the largest eigenvector and the length of the vector is normalized to 1 [8]. This rule takes the form:

$$\dot{\mathbf{w}} = \mathbf{x}y - y^2\mathbf{w} \tag{10}$$

We now show that the monocular deprivation results cannot be obtained by a PCA learning rule such as that of Oja, given the same assumptions about the inputs. This PCA analysis is not restricted to this one learning rule, but is representative of a general class of Hebb-based learning rules.[4, 14]

An exact solution to Oja's learning rule, which is a form of PCA or stabilized Hebbian learning, is shown by Wyatt and Elfeldel [15] to be

$$\mathbf{w}(t) = \frac{e^{\mathbf{Q}_{\text{FULL}}t}\mathbf{w}_0}{(||e^{\mathbf{Q}_{\text{FULL}}t}\mathbf{w}_0||^2 + 1 - ||\mathbf{w}_0||)^{\frac{1}{2}}}, \tag{11}$$

where $\mathbf{w}_0 = \mathbf{w}(t = 0)$ is the initial state of the weight vector and $\mathbf{Q}_{\text{FULL}}$ is the 2-eye correlation function. We need to include the correlation for the open eye when considering PCA because both the initial development of selectivity with this rule, and the dynamics of deprivation, depend on the open-eye correlations.

1. Normal Rearing

If both eyes have exactly the same input, the 2-eye correlation function, $\mathbf{Q}_{\text{FULL}}$, has the form

$$\mathbf{Q}_{\text{FULL}} = \begin{pmatrix} \mathbf{Q}_{\text{open}} & \mathbf{Q}_{\text{open}} \\ \mathbf{Q}_{\text{open}} & \mathbf{Q}_{\text{open}} \end{pmatrix} \tag{12}$$

where $\mathbf{Q}_{\text{open}}$ is the open-eye correlation function.

We expand the initial weight vector $\mathbf{w}(0)$ in terms of the eigenvectors, $\mathbf{u}_j$, of the open-eye correlation matrix.

$$\mathbf{w}(0) = \begin{pmatrix} \mathbf{w}^l(0) \\ \mathbf{w}^r(0) \end{pmatrix} = \sum_j \begin{pmatrix} a_j^l \mathbf{u}_j \\ a_j^r \mathbf{u}_j \end{pmatrix} \tag{13}$$

where $\mathbf{u}_j$, the j'th eigenvector of the one-open-eye correlation function $\mathbf{Q}_{\text{open}}$, has the eigenvalue ξ_j and a_j^l, a_j^r are the expansion coefficients for left and right eye respectively. We assume that eigenvectors and eigenvalues are arranged in a descending order, that is $\xi_1 > \xi_2 > \cdots \xi_N$. Inserting this into the Wyatt formula, and taking the limit, we get

$$\mathbf{w}(t \to \infty) = \sqrt{\frac{1}{2}} \begin{pmatrix} \mathbf{u}_1 \\ \mathbf{u}_1 \end{pmatrix} \tag{14}$$

as long as the largest eigenvalue is non-degenerate.

Thus the solution converges to a state in which both eye receptive fields are eigenvectors of the one-eye correlation function $\mathbf{Q}_{\text{open}}$. The higher the ratio between ξ_1 and the smaller eigenvalues the faster it will converge.

2. Monocular Deprivation (MD)

As for BCM, we assume that one eye is open and generates a structured input to the cortical cell, whereas the other eye is deprived and the activity of the deprived LGN inputs is uncorrelated with open-eye inputs.

Thus the full correlation function has the form

$$Q_{\text{FULL}} = \begin{pmatrix} \mathbf{Q}_{\text{open}} & 0 \\ 0 & n^2\mathbf{Q} \end{pmatrix} \tag{15}$$

where n^2 is the variance of the deprived-eye inputs. The eigenvectors and eigenvalues of the open and deprived-eye correlation function are defined as

$$\text{(open eye) } \mathbf{Q}_{\text{open}}\mathbf{u}_i = \xi_i \mathbf{u}_i \tag{16}$$

$$\text{(deprived eye) } \mathbf{Q}\mathbf{v}_j = \lambda_j \mathbf{v}_j \tag{17}$$

MD is started after the neuron has converged to the binocular fixed-point,

$$\mathbf{w}(0) = \sqrt{\frac{1}{2}} \begin{pmatrix} \mathbf{u}_1 \\ \mathbf{u}_1 \end{pmatrix} \tag{18}$$

We expand the initial condition for the deprived-eye in terms of the eigenvectors of the deprived-eye correlation function, $\mathbf{Q}$.

$$\mathbf{u}_1 = \sum_j b_j \mathbf{v}_j \tag{19}$$

The deprived-eye term in the numerator of the Wyatt solution (Equation 11) using the correlation function in Equation 15 is

$$e^{n^2 \mathbf{Q} t}\mathbf{u}_1 = \sum_j e^{n^2 \lambda_j t} b_j \mathbf{v}_j \tag{20}$$

$$\approx e^{n^2 \lambda_1 t} b_1 \mathbf{v}_1 \tag{21}$$

$$= e^{n^2 \lambda_1 t} (\mathbf{u_1} \cdot \mathbf{v_1}) \mathbf{v}_1 \tag{22}$$

where the approximation assumes that the largest eigenvalue of the deprived-eye correlation function is larger then all of the others. In the case of a degenerate largest eigenvalue, a constant N would multiply the term in Equation 22 and none of the conclusions that follow would change.

After this approximation, we arrive at the solution

$$\mathbf{w}(t) = \frac{\begin{pmatrix} e^{\xi_1 t}\mathbf{u}_1 \\ e^{n^2 \lambda_1 t}(\mathbf{u}_1 \cdot \mathbf{v}_1)\mathbf{v}_1 \end{pmatrix}}{(e^{2\xi_1 t} + e^{2n^2 \lambda_1 t}(\mathbf{u}_1 \cdot \mathbf{v}_1)^2)^{1/2}} \tag{23}$$

If the magnitude of the deprived-eye maximum eigenvalue (scaled by the variance) is smaller than the largest eigenvalue of the visual inputs, that is $n^2\lambda_1 < \xi_1$, then the $t \to \infty$ limiting case becomes

$$\mathbf{w}(t \to \infty) = \begin{pmatrix} \mathbf{u}_1 \\ \mathbf{0} \end{pmatrix} \tag{24}$$

and the deprived-eye weights decay.

The rate of decay of the deprived eye depend on the *difference* between the open-eye maximum eigenvalue, ξ_1, and the deprived-eye maximum eigenvalue scaled by the variance, $n^2\lambda_1$. Thus, the *larger* the variance of the noise the *slower* the decay of the deprived-eye weights, which is the opposite behavior to BCM. It is in fact possible,in the presence of very large noise, to get a PCA neuron to have an *increased* response to the deprived eye and a corresponding *decrease* in the open-eye responses.

II. RESULTS

A. Monocular Deprivation with BCM and Correlation of LGN Activity

We now generalize the BCM argument to deal with complex correlations in LGN activity. We explore the possibility that lid closure (MC) leads to LGN noise that is uncorrelated, and that retinal inactivation with TTX (MI), leads to LGN noise that is *correlated* resulting in a possible decrease in the synaptic modification for inactivated inputs.

Denote

$$< w_i^d > \ \rightarrow \ w_i$$
$$< n^2 > \ \rightarrow \ n^2$$

to obtain

$$\dot{w}_i = -\sum_j \epsilon n^2 Q_{ij} w_j \tag{25}$$

In matrix form

$$\dot{\mathbf{w}} = -\epsilon n^2 \mathbf{Q}\mathbf{w} \tag{26}$$

where $\mathbf{Q}$ is a square, N dimensional correlation matrix. The eigenvectors of $\mathbf{Q}$ are

$$\mathbf{Q}\mathbf{v}_i = \lambda_i \mathbf{v}_i \tag{27}$$

We expand the weight vector in the complete set of eigenvectors,

$$\mathbf{w}(t) = \sum_{i=1}^{N} a_i(t)\mathbf{v}_i$$

so that

$$\begin{aligned} \dot{\mathbf{w}(t)} &= \sum_{i=1}^{N} \dot{a}_i(t)\mathbf{v}_i = -\epsilon n^2 \mathbf{Q} \sum_{i=1}^{N} a_i(t)\mathbf{v}_i \\ &= -\epsilon n^2 \sum_{i=1}^{N} \lambda_i a_i(t)\mathbf{v}_i \end{aligned} \tag{28}$$

This gives

$$a_i(t) = a_i(0)e^{-\epsilon n^2 \lambda_i t} \tag{29}$$

and

$$\begin{aligned} \mathbf{w}(t) &= \sum_{i=1}^{N} a_i(0)e^{-\epsilon n^2 \lambda_i t}\mathbf{v}_i \\ &= \sum_{i=1}^{N} a_i(0)e^{-t/\tau_i}\mathbf{v}_i \end{aligned} \tag{30}$$

with the time constants for the decay defined as $\tau_i = (\epsilon n^2 \lambda_i)^{-1}$.

The activity in LGN neurons from the deprived eye can now be characterized by the correlation matrix, $\mathbf{Q}$.

1. Uncorrelated Noise

In the original MD uncorrelated noise argument of BCM, the matrix $\mathbf{Q}$ is

$$\mathbf{Q} \rightarrow \begin{pmatrix} 1 & 0 & 0 & 0 & \cdots \\ 0 & 1 & 0 & 0 & \cdots \\ 0 & 0 & 1 & 0 & \cdots \\ 0 & 0 & 0 & 1 & \cdots \\ & & \vdots & & \ddots \end{pmatrix} = \mathbf{I} \tag{31}$$

This leads to $\lambda_i = 1$ for all $\mathbf{w}$, therefore

$$\mathbf{w}(t) = \mathbf{w}(0)e^{-\epsilon n^2 t}$$

which approaches 0 as $t \to \infty$.

This is the 'normal' BCM uncorrelated noise result and gives the reference time of decay, $\tau = (\epsilon n^2)^{-1}$. Note that the weights all decay in time, with a faster decay for a larger noise variance, n^2.

2. *Fully Correlated Noise*

If the inputs to the deprived eye are *completely* correlated, $\mathbf{Q}$ becomes

$$\mathbf{Q} \to \begin{pmatrix} 1 & 1 & 1 & 1 & \cdots \\ 1 & 1 & 1 & 1 & \cdots \\ 1 & 1 & 1 & 1 & \cdots \\ & & \vdots & & \ddots \end{pmatrix} \tag{32}$$

This results in

$$\lambda_1 = N \text{ for the DC eigenvector } \mathbf{v}_1 = \frac{1}{\sqrt{N}}(1, 1, \cdots, 1)^T \tag{33}$$

$$\lambda_2, \cdots, \lambda_N = 0 \tag{34}$$

so that

$$\mathbf{w}(t) = a_1 \frac{1}{\sqrt{N}} \begin{pmatrix} 1 \\ 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix} e^{-\epsilon n^2 N t} + \sum_{i=2}^{N} a_i(0)\mathbf{v}_i \tag{35}$$

The first term is non-selective and decays rapidly. All of the others do not decay, $\tau_{i \neq 1} = \infty$. Thus, if the initial state is selective there is *no* decay.

3. *Partial Constant Correlation*

Let the inputs to the deprived eye be *partially* correlated so that $\mathbf{Q}$ is[21]

$$\mathbf{Q} \to \begin{pmatrix} 1 & q & q & q & \cdots \\ q & 1 & q & q & \cdots \\ q & q & 1 & q & \cdots \\ & & \vdots & & \ddots \end{pmatrix} \tag{36}$$

We can write $\mathbf{Q}$ as

$$\mathbf{Q} \;=\; (1-q)\mathbf{I} + q\mathbf{J} \tag{37}$$

where $\mathbf{I}$ is the identity and $\mathbf{J}$ is defined as

$$\mathbf{J} \;\equiv\; \begin{pmatrix} 1 & 1 & 1 & 1 & \cdots \\ 1 & 1 & 1 & 1 & \cdots \\ 1 & 1 & 1 & 1 & \cdots \\ & & \vdots & & \ddots \end{pmatrix} \tag{38}$$

The eigenvectors and eigenvalues of $\mathbf{Q}$ are

$$\begin{aligned} \mathbf{v}_1 = \tfrac{1}{\sqrt{N}}(1,1,1,\cdots,1)^T \quad & \lambda_1 = 1 + (N-1)q \\ \mathbf{v}_i \text{ orthogonal to } \mathbf{v}_1 \quad & \lambda_i = 1 - q \text{ for } i > 1 \end{aligned} \tag{39}$$

This gives

$$\mathbf{w}(t) \;=\; a_1(0)e^{-\epsilon n^2(1+(N-1)q)t}\mathbf{v}_1 + \sum_{i=2}^{N} a_i(0)e^{-\epsilon n^2(1-q)t}\mathbf{v}_i \tag{40}$$

The first term is non-selective and decays rapidly. The other terms decay with characteristic time $\tau_q = (\epsilon n^2(1-q))^{-1}$. Note that since $\tau_q^{-1} \leq \epsilon n^2$, the weights decay *more slowly* than the uncorrelated case, and τ_q reaches its maximum at $q = 1$, where there is no decay as in equation 35.

4. Partial Non-Uniform Correlation

We now relax the assumption that the correlation between any two neurons is a constant, q. Let the correlation matrix have the form

$$\mathbf{Q} \;=\; \mathbf{Q}_0 + \mathbf{Q}_1 \tag{41}$$

where $\mathbf{Q}_0$ is the correlation function above,

$$\mathbf{Q}_0 = (1-q)\mathbf{I} + q\mathbf{J} \tag{42}$$

and $\mathbf{Q}_1$ is a symmetric, N-dimensional random-valued matrix whose off-diagonal elements have mean zero with variance, m^2

$$\mathbf{Q}_1 \equiv \begin{pmatrix} 0 & \Delta_{ij} & & \cdots \\ \Delta_{ij} & 0 & 0 & \cdots \\ & & \vdots & \ddots \end{pmatrix} \tag{43}$$

where

$$\begin{aligned} \int p_{ij}(\Delta_{ij})\Delta_{ij}^2 d(\Delta_{ij}) &= m^2 \\ \Delta_{ij} &= \Delta_{ji} \\ \int p_{ij}(\Delta_{ij}) d(\Delta_{ij}) &= 1 \end{aligned}$$

uncorrelated independent random variables of the $N \times N$ symmetric matrix

Matrices such as $\mathbf{Q}_1$ have been analyzed by Wigner[16] and found to have the following eigenvalue distribution in the large N limit

$$\sigma(\lambda) = \begin{cases} \frac{(4Nm^2-\lambda^2)^{1/2}}{2\pi N m^2} & \text{for } \lambda^2 < 4Nm^2 \\ 0 & \text{for } \lambda^2 > 4Nm^2 \end{cases} \tag{44}$$

In order to find the eigenvalues of $\mathbf{Q}$ we use a perturbation theory argument. In the non-perturbed case:

$$\begin{aligned} \mathbf{Q}_0 \mathbf{v}_i &= \lambda_i^0 \mathbf{v}_i \\ \lambda_1^0 &= 1 + (N-1)q \\ \lambda_2^0, \cdots, \lambda_N^0 &= 1 - q \end{aligned}$$

If we treat $\mathbf{Q}_1$ as a perturbation, in the lowest order

$$\lambda_{\alpha>1} = \lambda_\alpha^0 + \langle \alpha | \mathbf{Q}_1 | \alpha \rangle \tag{45}$$

Because the zero-order eigenvalues are largely degenerate we must diagonalize $\mathbf{Q}_1$ over the degenerate states; but this is just the problem solved by Wigner. We can therefore add the Wigner distribution to the unperturbed eigenvalues, $\lambda_2, \cdots, \lambda_N = 1 - q$, to get

$$\lambda_2, \cdots, \lambda_N = 1 - q + \text{ (Wigner distribution)} \tag{46}$$

This result is close to exact (beyond the perturbation argument). Since

$$\mathbf{Q} \;=\; (1-q)\mathbf{I} + q\mathbf{J} + \mathbf{Q}_1 \tag{47}$$

and a transformation that diagonalizes $\mathbf{Q}_1$

$$U\mathbf{Q}_1 U^T = (\text{diagonal}) \tag{48}$$

does not change $\mathbf{I}$

$$U\mathbf{I}U^T = \mathbf{I} \tag{49}$$

For $\mathbf{J}$, excluding the DC eigenvalue, gives us the invariant subspace that corresponds to the degenerate eigenvalue $1-q$. The matrix corresponding to this subspace is spherical (all the eigenvalues are the same) and the eigenvalue distribution is unchanged in any orthogonal (rotational) transformation.

Finally, we have the result that the inverse of the time constant for the decay, τ_i^{-1}, is distributed between $\epsilon n^2(1-q-2m\sqrt{N})$ and $\epsilon n^2(1-q+2m\sqrt{N})$.

An obvious problem of marrying the Wigner random matrix with the correlation matrix of Equation 41, is that not all of the matrices, $\mathbf{Q}$, that result are correlation matrices (some of the resulting eigenvalues may be negative). This restricts the possible values of m to be smaller than $(1-q)/2\sqrt{N}$.

This problem can be solved by introducing a Wishart distribution[17]. Let

$$\mathbf{Q} \simeq \left(\sqrt{\mathbf{Q}_0} + m\mathbf{W}\right)^2 \tag{50}$$

where $\mathbf{W}$ is a symmetric random matrix with independent $\mathcal{N}(0,1)$ distributed elements. We introduce

$$m = m_o\sqrt{\frac{1-q}{N}} \tag{51}$$

and obtain a distribution for the eigenvalues

$$\rho(\lambda) = \frac{\sqrt{4m_o^2 - (\sqrt{\lambda/(1-q)}-1)^2}}{4\sqrt{(1-q)\lambda m_o^2\pi}} \tag{52}$$

This distribution is shown in Figure 2.

To determine the effect of this distribution on the decay time we have simulated the response of the deprived eye neurons in an environment that corresponds to the appropriate correlation matrix (as described in methods). The results are shown in Figure 3A.

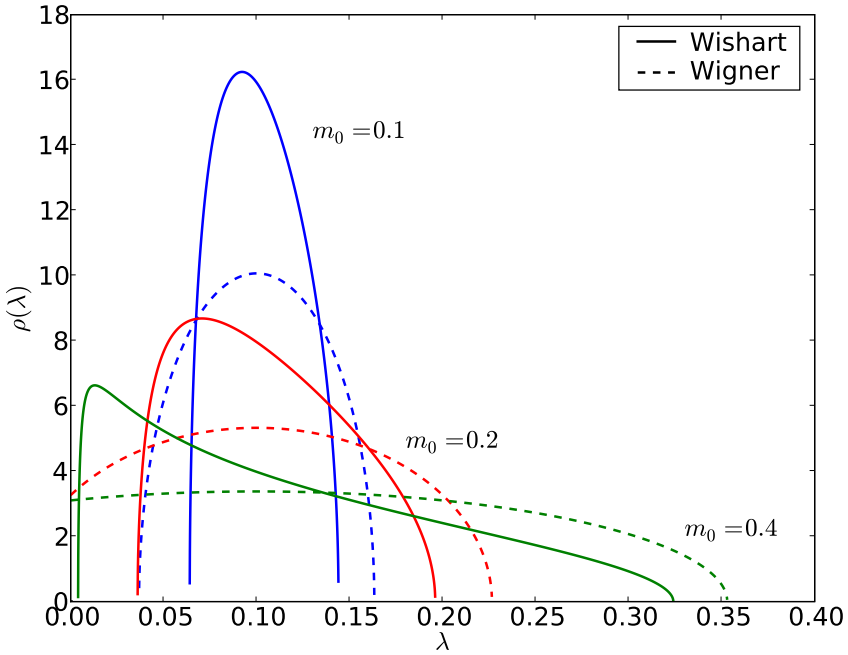

FIG. 2: Eigenvalue distributions with correlation $q = 0.9$. Shown are the Wishart distribution (solid) and the Wigner distribution cut off at $\lambda = 0$ (dashed) for several values of the variance of the noise.

We conclude that the greater the correlation between LGN afferents, the longer the lifetime for the decay of the deprived eye and that this result is insensitive to noise added to the precise shape of the correlation function. It remains to be determined experimentally how great this correlation is.

B. Monocular Deprivation with PCA and Correlated of LGN Activity

We next explore the consequences of correlation in LGN activity on the predictions of the PCA rule in monocular deprivation.

1. *Uncorrelated Noise*

For completeness, we present the result for uncorrelated noise. If the deprived-eye inputs are completely uncorrelated, the correlation is given by Equation 31, and the largest eigenvalue is $\lambda_1^{\text{deprived}} = 1$. As stated above, this leads to the prediction that the *larger* the variance of the noise the *slower* the decay of the deprived-eye weights.

2. Partial Constant Correlation

If the input correlations to the deprived-eye are of the form in Equation 36 then we know the eigenvectors and eigenvalues, given in Equation 39. If the initial (selective) one-eye vector, $\mathbf{u}_1$ has no DC component (i.e. $\mathbf{u}_1 \cdot (1, 1, 1, \cdots) = 0$) then the numerator term of the deprived-eye Wyatt solution (20) becomes

$$\sum_j e^{n^2 \lambda_j t} (\mathbf{u}_1 \cdot \mathbf{v}_j) \mathbf{v}_j = e^{n^2(1-q)t} \sum_j (\mathbf{u}_1 \cdot \mathbf{v}_j) \mathbf{v}_j \tag{53}$$

$$= e^{n^2(1-q)t} \mathbf{u}_1 \tag{54}$$

Thus the Wyatt solution yields

$$\mathbf{w}_{\mathrm{MI}}(t) = \begin{pmatrix} \mathbf{w}^{\mathrm{open}}(t) \\ \mathbf{w}^{\mathrm{deprived}}(t) \end{pmatrix} = \frac{\begin{pmatrix} e^{\xi_1 t} \mathbf{u}_1 \\ e^{n^2(1-q)t} \mathbf{u}_1 \end{pmatrix}}{(e^{2\xi_1 t} + e^{2n^2(1-q)t})^{1/2}} \tag{55}$$

which will result in the decay of the deprived-eye weights if $\xi_1 > n^2(1-q)$, where ξ_1 comes from the natural images. In the special case that $\xi_1 \gg n^2$, the deprived-eye weights decay exponentially

$$||\mathbf{w}^{\mathrm{deprived}}|| \sim \exp(-\xi_1 t) \tag{56}$$

The largest eigenvalue, ξ_1, coming from the natural images tends to be large compared to $n^2(1-q)$ (on the order of 10), thus there is very little dependence on q, as seen in Figure 3. In practice there is a small DC component due to the finite size of the receptive field, which would grow exponentially, but the test stimulus does not register any changes due to the DC component of the receptive field.

3. Partial Correlation with Variance

It follows from Wigner's work on stochastic matrices, explored in Section II A 4, if the correlation function from Equation 36 is perturbed with random components, the distribution of eigenvalues has a maximum value of $\lambda < 2 - 2q$ before the distribution becomes partly negative and we no longer have a correlation matrix. Generalizing this to the Wishart distribution, we can obtain eigenvectors with larger eigenvalues. In practice, however, this has little effect.(Figure 3)

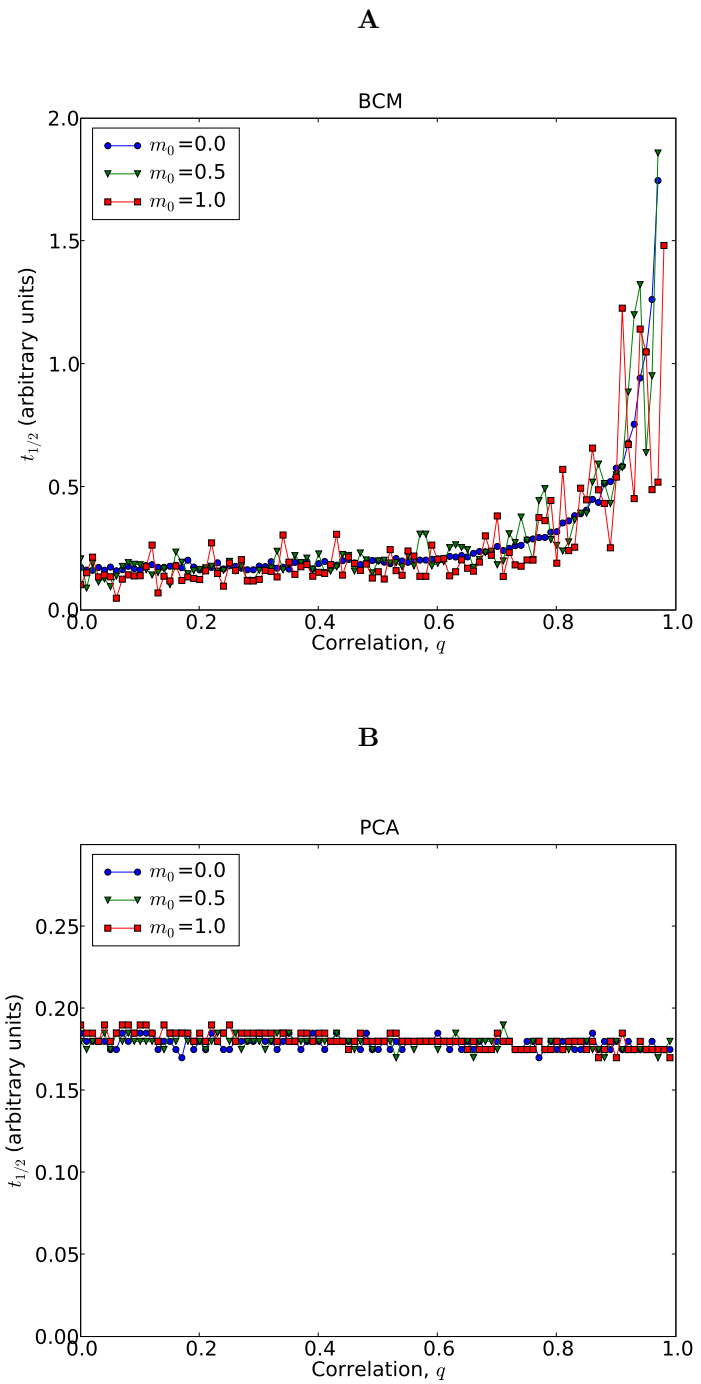

FIG. 3: Simulations of Monocular Deprivation versus the correlation value, q. Both learning rules are generally insensitive to the deviation from pure correlation, as specified by the variance parameter m_0 in the Wishart distribution. (A) The BCM learning rule exhibits a *slow down* of the effects of deprivation as the correlation increases. (B) Hebbian learning shows very little dependence on the correlation, q.

III. DISCUSSION

The BCM theory [4, 6] makes the testable prediction that the ocular dominance shift due to lid closure (MC) is more rapid than that due to retinal inactivation with TTX (MI). This prediction motivated the experiment of Rittenhouse et al (1999) that attempted to test the noise-dependence of loss of deprived-eye response by comparing MC with MI (produced by injecting TTX into the retina of the deprived eye shutting down retinal neuronal activity). It was implicitly assumed in this study that shutting down retinal activity would reduce LGN noise to cortical cells. The results of Rittenhouse et al (1999) were in accord with the BCM prediction: MI showed a much slower loss of response to the deprived eye than MC. These results were confirmed and expanded in studies in mice[10]. But this prediction of BCM was based on the assumption that retinal inactivation would reduce the activity level in the layers of LGN corresponding to the inactivated eye.

The study of Linden et al 2009 was done to directly check the assumption that TTX reduction of retinal activity actually reduced LGN activity. Instead it was found that LGN activity is not reduced on average but becomes highly correlated. These results require a reexamination of the analysis and the expected results of different deprivation experiments in the BCM theory as well as in other theories of synaptic plasticity [4].

Our present study was initiated to investigate the consequences on the BCM and the PCA theories for MC vs MI in situations when retinal inactivation leads to LGN noise that is not reduced but is *correlated*. We note that theoreticians had not investigated this previously because no one thought this would be a consequence of TTX retinal inactivation: a wonderful reminder that the real world produces surprises most of us would not expect.

Our present analysis shows that the BCM theory in the situation where the average noise activity is unchanged[11] predicts a reduced loss of response of the inactivated eye (compared to lid closure) that depends on the amount of correlation between firing rates of LGN neurons. For zero correlation and identical noise variance, MI would produce the same loss as MC. But for total correlation, there is no loss at all! In contrast we find that in the PCA theory the magnitude of the noise correlations does not significantly effect the time course of deprivation experiments.

Although the simulations and analysis here have assumed the existence of both positive and negative weights, this detail is not important to the results. It is true, for example, that Equation 26 with the all-ones correlation function (Equation 32) in the case of all-positive weights leads to all weights decaying to zero, not just the DC component. However, a small amount of mean-field inhibition[4, 19, 20] is sufficient to keep this from occurring. It has been shown elsewhere that the

BCM theory is valid for simulations using only positive weights[4].

We observe that the correlations needed here to get the MI effect are fairly high, $q \sim 0.8$ whereas the correlations measured in Linden et. al. 2009 are fairly low, around 0.1. This arises from a difference in the of the definition of the correlation measure with spike-based neurons versus rate-based neurons. Although this comparison requires more research, we show in the appendix (see Section A) that in one simple case a high rate-based correlation leads to a low spike-based correlation.

Given our results, it becomes interesting to test the correlation dependence of the rate of fall-off of response from the deprived eye in MI. If such tests can be made, they would provide a more detailed check of theoretical predictions and, hopefully, continue the dialogue between theory and experiment.

IV. METHODS

A. Correlated Environments

Given the form of a correlation function, such as shown for the Wishart distribution (Equation 50) or more simply the partial constant correlation (Equation 36) we can generate an environment of input vectors which has that correlation.

Assuming a matrix $\mathbf{Q}$ is specified, that is symmetric and real. Such a matrix can be decomposed, with the Cholesky algorithm, into

$$\mathbf{Q} = \mathbf{L}\mathbf{L}^T$$

In order to generate the environment, we start by generating N random vectors, $\mathbf{u}_i$ with i from 0 to N, where the elements are drawn from a normal distribution with zero mean and unit variance. The correlation function of these vectors is the unit matrix. We generate new correlated random vectors ($\mathbf{v}_i$) from the random vectors $\mathbf{u}_i$ using the matrix $\mathbf{L}$ obtained from the Cholesky decomposition.

$$\mathbf{v}_i \equiv \mathbf{L}\mathbf{u}_i$$

These vectors have the appropriate correlation function which is demonstrated by the direct calculation.

$$\langle \mathbf{v}_i \mathbf{v}_i^T \rangle \;=\; \langle (\mathbf{L}\mathbf{u})(\mathbf{L}\mathbf{u})^T \rangle$$

$$= \mathbf{L}\langle \mathbf{u}\mathbf{u}^T \rangle \mathbf{L}^T$$
$$= \mathbf{L}\mathbf{L}^T = \mathbf{Q}$$

One can think of $\mathbf{Lu}$ as a multi-dimensional spherical cloud of points, each dimension scaled by the eigenvalues of $\mathbf{Q}$, and then rotated so that the principal components are pointing in the direction of the eigenvectors of $\mathbf{Q}$.

B. Numerical Simulations

We use 7×7 circular patches from images of natural scenes to represent the normal visual environment[3, 4]. The images are processed by a retinal difference of Gaussians, with the biologically observed approximate 3 to 1 ratio of the surround to the center of the ganglion receptive field[18]. Neurons with a particular learning rule are trained with natural scene stimulus to both eyes until we obtain binocular oriented receptive fields. To model deprivation we continue training but present correlated vectors to the deprived eye, where the vectors are derived from the correlation function in Equation 50 using the method described in Section IV A.

To quantitatively measure the timing of the deprivation experiments, we measure the response of the neurons using oriented stimuli and then estimate the characteristic half-time for the decay of neuronal response. We report a *negative* half-time if the response increases. The results of the simulations are shown in Figure 3.

Acknowledgments

This work was partially supported by the Collaborative Research in Computational Neuroscience (CRCNS) grant (NSF #0515285) from the National Science Foundation and by the National Eye Institute. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Appendix A: On the correspondence between rate and spike correlations

Here we examine one simple case where the correspondence between rate correlations and spike correlations can be calculated in a straightforward fashion. Here we analyze the correlations between the spike trains of two neurons. Spikes are generated by a doubly stochastic process. At each time bin there is a probability that the neurons will spike. These probabilities, which are related

to the spike rate, are updated synchroneously in the two neurons, and these rates are correlated. Switch to a different rate occures randomly, with a memoryless process, resulting in an exponential distribution with a time constant τ.

The rate values of the two neurons change in a correlated manner (as in the rest of the paper) with a mean rate r_0 and a correlation function:

$$\mathbf{Q} = n^2 \begin{pmatrix} 1 & q \\ q & 1 \end{pmatrix} \tag{A1}$$

We introduce a variable $S_i(t)$ such that $S_i(t) = 1$ if there is a spike at time t for neuron i, and 0 otherwise. The central quantity we measure is $< S_i(t)S_j(t+\rho) >$ where i and j are the neuron indexes and ρ is the temporal shift. If ρ is within the same time bin, then the rates are correlated. If ρ is in a different bin, then the rates are uncorrelated. Correlations within the same time bin are denoted by $< \quad >_s$ and in a different time bin by $< \quad >_d$. Consequently

$$< S_i(t)S_j(t+\rho) >=< S_i(t)S_j(t+\rho) >_s P_s(\rho)+ < S_i(t)S_j(t+\rho) >_d (1 - P_s(\rho)) \tag{A2}$$

where $P_s(\rho) = \exp(-\rho/\tau)$ is the probability that two bins a time ρ apart belong to the same time bin.

We now calculate both terms separately, starting with the more complicated within-bin correlations.

$$< S_i(t)S_j(t+\rho) >_s= \delta_{ij} \left[\delta(\rho)r_i + (1 - \delta(\rho))\, r_i^2\right] + (1 - \delta_{ij})\, r_i r_j \tag{A3}$$

where r_i and r_j are the instantaneous rates, but for simplicity the dependence on time is not explicitly presented (within-bin rates satisfy $r_i(t) = r_i(t+\rho)$). We now take the average over the joint distribution of r_i and r_j and assume that this is equivalent to taking the temporal average. This average is denoted by $\overline{()}$. We use: $\overline{r_i} = r_0$, $\overline{r_i^2} = n^2 + r_0^2$ and $\overline{r_i r_j} = qn^2 + r_0^2$. We therefore get:

$$\overline{< S_i(t)S_j(t+\rho) >_s} = \delta_{ij} \left[\delta(\rho)r_0 + (1 - \delta(\rho))\,(n^2 + r_0^2)\right] + (1 - \delta_{ij})\,(qn^2 + r_0^2) \tag{A4}$$

The second term has the form:

$$\overline{< S_i(t)S_j(t+\rho) >_d} = \overline{< S_i(t) >}\ \overline{< S_j(t+\rho) >} = r_0^2 \tag{A5}$$

putting these together we get that

$$\begin{aligned} \sigma_{ij}(\rho) &= \overline{< S_i S_j(\rho) >} - \overline{< S_i(t) >}\ \overline{< S_j(t+\rho) >} && \text{(A6)} \\ &= P_s(\rho) \left\{\delta_{ij} \left[\delta(\rho)(r_0 - r_0^2) + (1 - \delta(\rho))n^2\right] + (1 - \delta_{ij})qn^2\right\} && \text{(A7)} \end{aligned}$$

The correlation function calculated in Linden et.al. (2009) is: $C_{ij}(\rho) = \sigma_{ij}(\rho)/(r_0(1-r_0)$ if both neurons have the same mean rates.

Using this we get

$$C_{ij}(\rho) \;=\; P_s(\rho)\left\{\delta_{ij}\left[\delta(\rho) + (1-\delta(\rho))\frac{n^2}{r_0(1-r_0)}\right] + (1-\delta_{ij})\frac{qn^2}{r_0(1-r_0)}\right\} \tag{A8}$$

As an example, suppose that we have a $q = 0.8$, a mean frequency of 20Hz, which means that $r_0 = 0.02\text{s}^{-1}$, and the standard deviation of 10 Hz in choosing the firing rate, which means that $n = 0.01$. This imples that a peak of A8 for $i \neq j$, is 0.001 and and integral over the function from -10ms to 10ms is 0.018, a much lower correlation value than the one implied by $q = 0.8$.

---

[1] M F Bear and C D Rittenhouse. Molecular basis for induction of ocular dominance plasticity. *J Neurobiol*, 41(1):83–91, 1999.

[2] Frank Sengpiel and Peter C Kind. The role of activity in development of the visual system. *Curr Biol*, 12(23):R818–26, December 10 2002.

[3] B. S. Blais, N. Intrator, H. Shouval, and L. N Cooper. Receptive field formation in natural scene environments: comparison of single cell learning rules. *Proceedings of the National Academy of Sciences*, 10(7):1797–1813, 1998.

[4] Leon N Cooper, Nathan Intrator, Brian S. Blais, and Harel Z. Shouval. *Theory of cortical plasticity.* World Scientific, New Jersey, 2004.

[5] T. Wiesel and D. Hubel. Comparison of effect of unilateral and bilateral eye closure on cortical unit response in kittens. *Journal of Physiology*, 180(180):106–154, 1962.

[6] E. L. Bienenstock, L. N Cooper, and P. W. Munro. Theory for the development of neuron selectivity: orientation specificity and binocular interaction in visual cortex. *Journal of Neuroscience*, 2:32–48, 1982.

[7] Brian Blais, Harel Shouval, and Leon N Cooper. The role of presynaptic activity in monocular deprivation: Comparison of homosynaptic and heterosynaptic mechanisms. *Proc. Natl. Acad. Sci.*, 96:1083–1087, 1999.

[8] E Oja. A simplified neuron model as a principal component analyzer. *Journal of Mathematical Biology*, 15:267–273, 1982.

[9] Cynthia D. Rittenhouse, Harel Z. Shouval, Michael A. Paradiso, and Mark F. Bear. Evidence that monocular deprivation induces homosynaptic long-term depression in visual cortex. *Nature*, 397:347–350, 1999.

[10] Mikhail Y. Frenkel and Mark F. Bear. How monocular deprivation shifts ocular dominance in visual cortex of young mice. *Neuron*, 44(6):917–923, 2004.

[11] Monica Linden, Arnold J. Heynen, Robert H. Haslinger, and Mark F. Bear. Thalamic activity that drives visual cortical plasticity. *Nature Neuroscience*, Advance Online Publication(doi:10.1038/nn.2284), 2009.

[12] Michael Weliky and Lawrence C. Katz. Correlational structure of spontaneous neuronal activity in the developing lateral geniculate nucleus in vivo. *Science*, 285:599–604, 1999.

[13] Tomokazu Ohshiro and Michael Weliky. Simple fall-off pattern of correlated neural activity in the developing lateral geniculate nucleus. *Nat Neurosci*, 9(12):1541–1548, December 2006.

[14] K. D. Miller and D. J. C. MacKay. The role of constraints in Hebbian learning. *Neural Computation*, 6:98–124, 1994.

[15] J. L. Wyatt and I. M. Elfadel. Time-domain solutions of Oja's equations. *Neural Computation*, 7(5):915–922, 1995.

[16] Eugene P. Wigner. On the distribution of the roots of certain symmetric matrices. *Annals of Mathematics*, 67(2):325–364, March 1958.

[17] Armando Bazzani, Gastone C. Castellani, and Leon N Cooper. Eigenvalue distribution for a class of covariance matrices. *unpublished*, 2009.

[18] R.A. Linsenmeier, L. J. Frishman, H. G. Jakiela, and C. Enroth-Cugell. Receptive field properties of X and Y cells in the cat retina derived from contrast sensitivity measurments. *Vision Research*, 22:1173–1183, 1982.

[19] Cooper, L. N. and Scofield, C. L. (1988). Mean-field theory of a neural network. *Neural Computation*, 85:1973–1977.

[20] Intrator, N. and Cooper, L. N. (1992). Objective function formulation of the BCM theory of visual cortical plasticity: Statistical connections, stability conditions. *Neural Networks*, 5:3–17.

[21] This becomes totally correlated (previous example) when $q \to 1$.