Supplemental Materials for "Equalization of ocular dominance columns induced by an activity-dependent learning rule and the maturation of inhibition"

Taro Toyoizumi and Kenneth D. Miller

February 5, 2009

Contents

1 Analysis of the semi-linear Hebbian model

In this section, we mathematically analyze the properties of the Hebbian learning rule of Eq. 6, and illustrate, with a simple semi-linear model (linear except for hard upper and lower bounds on weights) the stability or instability of contra-dominated, equalized, and open-eye-dominated patterns under normal conditions and under MD.

1.1 Analytical Framework

Let $\vec{\mathbf{w}}(x) = \begin{pmatrix} w_I(x) \\ w_C(x) \end{pmatrix}$ be the vector of synaptic weights and $\vec{\mathbf{h}}(x) = \begin{pmatrix} h_I \\ h_C \end{pmatrix}$ be the input vector. To begin with, the output firing rate of a neuron at cortical position x is described by

$$\begin{aligned} r(x) &= \sum_{a=C,I} w_a(x) h_a + \int dx' M(x-x') r(x') \\ &= \int dx' K(x-x') \vec{\mathbf{h}}^T \vec{\mathbf{w}}(x), \end{aligned} \tag{S1}$$

where $K(x) = \frac{1}{2\pi} \int dk e^{ikx} \frac{1}{1-\tilde{M}(k)}$ and $\tilde{M}$ is the Fourier transformation of the function describing intracortical connectivity. Note that the half rectification function in Eq. 1 is omitted here so that the equation will be linear and amenable to analysis. We assume that the input mean has the form $\langle \vec{\mathbf{h}} \rangle = \begin{pmatrix} 1 \\ \epsilon \end{pmatrix}$ and the input covariance matrix has the form

$$\langle (\vec{\mathbf{h}} - \langle \vec{\mathbf{h}} \rangle)(\vec{\mathbf{h}} - \langle \vec{\mathbf{h}} \rangle)^T \rangle = q \begin{pmatrix} 1 & c\epsilon \\ c\epsilon & \epsilon \end{pmatrix} \tag{S2}$$

with covariance strength $q > 0$, MD factor $0 \leq \epsilon \leq 1$ and between-eye correlation $0 \leq c \leq 1$. Here, $\epsilon = 1$ corresponds to normal development while smaller values of ϵ model monocular deprivation.

Setting $\mathbf{S} = \frac{1}{2}\begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}$, the average changes of synaptic weights according to the Hebbian learning rule with subtractive normalization is described by

$$\begin{aligned} \Delta\vec{\mathbf{w}}(x) &= \alpha\langle(r(x) - \rho\bar{r}(x))\mathbf{S}\vec{\mathbf{h}}\rangle \\ &= \alpha\langle(\delta r(x) + (1-\rho)\bar{r}(x))\mathbf{S}\vec{\mathbf{h}}\rangle \\ &= \alpha\int dx' K(x-x')\mathbf{S}\left[\langle(\vec{\mathbf{h}} - \langle\vec{\mathbf{h}}\rangle)(\vec{\mathbf{h}} - \langle\vec{\mathbf{h}}\rangle)^T\rangle + (1-\rho)\langle\vec{\mathbf{h}}\rangle\langle\vec{\mathbf{h}}\rangle^T\right]\vec{\mathbf{w}}(x') \end{aligned} \tag{S3}$$

along with the saturation condition of synapses, $0 \leq w_a(x) \leq w_{\max}$. Equation S3 becomes

$$\Delta\vec{\mathbf{w}}(x) = \alpha\int dx' K(x-x')\mathbf{A}\vec{\mathbf{w}}(x), \tag{S4}$$

where we have defined the matrix $\mathbf{A} = \mathbf{S}\left[\langle(\vec{\mathbf{h}} - \langle\vec{\mathbf{h}}\rangle)(\vec{\mathbf{h}} - \langle\vec{\mathbf{h}}\rangle)^T\rangle + (1-\rho)\langle\vec{\mathbf{h}}\rangle\langle\vec{\mathbf{h}}\rangle^T\right] = \begin{pmatrix} u & v \\ -u & -v \end{pmatrix}$ with $u = q(1-c\epsilon)/2 + (1-\rho)(1-\epsilon)/2$ and $v = -q\epsilon(1-c)/2 + (1-\rho)\epsilon(1-\epsilon)/2$. The eigenvalue decomposition of this matrix is written as $\mathbf{A} = \mathbf{P}\mathbf{\Lambda}\mathbf{P}^{-1}$ where the columns of $\mathbf{P}$ are the eigenvectors of $\mathbf{A}$ and $\mathbf{\Lambda}$ is a diagonal matrix whose diagonal entries are the corresponding eigenvalues of $\mathbf{A}$. We find that

$$\mathbf{P} = \frac{1}{\lambda_1}\begin{pmatrix} -v & 1 \\ u & -1 \end{pmatrix} \tag{S5}$$

and

$$\mathbf{\Lambda} = \begin{pmatrix} 0 & 0 \\ 0 & \lambda_1 \end{pmatrix}, \tag{S6}$$

where $\lambda_1 = u - v$. Then, if $\vec{\mathbf{e}}_0$ and $\vec{\mathbf{e}}_1$ are the two eigenvectors of $\mathbf{P}$, we can decompose $\vec{\mathbf{w}}(x) = z_0(x)\vec{\mathbf{e}}_0 + z_1(x)\vec{\mathbf{e}}_1$ where $z_0(x)$ and $z_1(x)$ are given by

$$\begin{pmatrix} z_0(x) \\ z_1(x) \end{pmatrix} = \mathbf{P}^{-1}\vec{\mathbf{w}}(x) = \begin{pmatrix} 1 & 1 \\ u & v \end{pmatrix}\begin{pmatrix} w_I(x) \\ w_C(x) \end{pmatrix}. \tag{S7}$$

We can then reexpress Eq. S4 in terms of the dynamics of z_0 and z_1:

$$\begin{aligned} \Delta z_0(x) &= 0, \\ \Delta z_1(x) &= \alpha\lambda_1\int dx' K(x-x')z_1(x). \end{aligned} \tag{S8}$$

Note that $z_0(x) = w_I(x) + w_C(x)$ does not change in time because of the imposed subtractive normalization constraint, i.e, the sum of contra-eye and ispi-eye synapses should be constant. The saturation condition of synapses is now described by

$$\begin{aligned} 0 \leq -vz_0(x) + z_1(x) \leq \lambda_1 w_{\max}, \\ 0 \leq uz_0(x) - z_1(x) \leq \lambda_1 w_{\max}. \end{aligned} \tag{S9}$$

The learning dynamics of Eq. S8 has an energy function

$$E = -\frac{1}{2}\int\int dxdx' z_1(x)K(x-x')z_1(x'). \tag{S10}$$

The direction of flow at $z = (z_0(x), z_1(x))$ is described by the negative gradient of E.

We assume $K(x)$ is symmetric about zero, $K(x) = K(-x)$, so that its Fourier transform is real. Because $\tilde{K}(k) = 1/(1 - \tilde{M}(k))$ and we assume we are working in a regime in which cortical dynamics are stable, meaning $\tilde{M}(k) < 1$ for all k, the Fourier spectrum of K is always positive, i.e., $\tilde{K}(k) > 0$.

From the positivity of the Fourier spectrum of K, we can show that all the stable fixed points of the dynamics satisfy $z_1(x) = z_{\max}$ or $z_1(x) = z_{\min}$, where all the components of $w_a(x)$ are saturated. Note that $z_{\max} = uw_{\max}$, $z_{\min} = vw_{\max}$ if $v < 0$ and $z_{\min} = 0$ if $v > 0$. To show this, we show there cannot be a stable fixed point $z^* = (z_0^*(x), z_1^*(x))$ of the dynamics that has a non-saturated component. First, if we perturb z around that stable fixed stable point by $\delta z = (\delta z_0(x), \delta z_1(x))$, the energy function changes by

$$\delta E \;=\; -\int dx \delta z_1(x) v_1^*(x) - \frac{1}{2}\int\int dx dx' \delta z_1(x) K(x - x') \delta z_1(x'), \tag{S11}$$

where the negative gradient of E at z^* is described by $v_1^*(x) = \int dx' K(x - x') z_1^*(x')$. Because the fixed point is stable, any allowed perturbation (meaning any perturbation that does not carry synapses past their weight limits) must yield a positive δE. The second term of Eq. S11 is always negative because the Fourier spectrum of K is all positive (analogous to a positive definite matrix in a discrete case). Therefore, the first term must be positive for any allowed perturbation. However, if at least one component of $z_1(x)$ has not saturated, we can always choose a non-zero $\delta z_1(x)$ to be perpendicular to $v_1^*(x)$, i.e., $\int \delta z_1(x) v_1^*(x) = 0$, by perturbing only the non-saturated component of $z_1(x)$. The gradient of E at z^* must be zero in the direction of any nonsaturated component of z^*, because otherwise z^* would not be a fixed point. This is contradictory because, then, the energy function is decreased by the perturbation δz, so z^* is not stable. Therefore, all the components of $z_1(x)$ are either $z_{\max}$ or $z_{\min}$ at stable fixed points.

1.2 Stability of contra-dominated, equalized, and open-eye-dominated patterns under the covariance learning rule

In the main text, we explained that the initially contralaterally-dominated pattern stayed intact under the covariance learning rule with subtractive normalization until cortical inhibition matured (Fig. 2C). This contralaterally-dominated pattern was destabilized with the maturation of inhibition that initiates the CP and the equalized pattern became stable (Fig. 2D). Subsequently, if the contra-eye was closed, an MD shift to the open-eye-dominated pattern occurred in the presence of large noise (Fig. 2E), but the equalized pattern remained stable against MD without the noise. Here, we explain these results using the linear analysis described in section 1.1. In particular, the maturation of inhibition changes K from a more low-pass to a more band-pass filter (as shown in Fig. S1A), and we show that this destabilizes the contralaterally-dominated pattern. We also show that closing the contra-eye after the equalization of OD columns does not cause an MD shift without noise, because the equalized pattern remains stable.

From the analysis in the section 1.1, the dynamics of the covariance learning under subtractive normalization are described by Eq. S8, and the $z_1(x)$ are equal to either $z_{\max}$ or $z_{\min}$ at stable fixed points. Hence, in order for $z_1(x)$ to be a stable fixed point, its flow, $\Delta z_1(x) \propto \int K(x - x') z_1(x') dx'$, should be positive at x where $z_1(x) = z_{\max}$ and negative at x where $z_1(x) = -z_{\min}$. We simulated Eq. (S8), starting from the same initial condition described in Fig. 2B. In Figure S1B, each panel shows z_1 (red lines) and the corresponding $\int K(x-x')z_1(x')dx'$ (blue lines) under a different condition. The top-left panel shows the contralaterally-dominated pattern of z_1 before the CP and the corresponding flow, $\int K(x-x')z_1(x')dx'$ (blue lines). Before the CP, K is essentially a low-pass filter but with a delta-peak at the center. Hence, $\Delta z_1(x)$ is a superposition of $z_1(x)$ itself and its low-passed or smoothed version. We can see from the figure that this z_1 is stable, having two peaks inherited from $z_1(x)$; the stability of $z_1(x)$ is essentially due to the delta-peak of K. The top-right panel shows the same pattern z_1, along with the new flow $\int K(x-x')z_1(x')dx'$ induced after the maturation of inhibition that initiates the CP. Now, K is more a band-pass filter, and this creates another positive peak of Δz_1 in between the two peaks of z_1, making this pattern unstable. This shows how the maturation of inhibition destabilizes the initial, contralaterally-dominated condition.

The bottom-left panel shows the equalized pattern of z_1 that forms during the CP and the corresponding flow $\int K(x-x')z_1(x')dx'$. This equalized pattern is stable given the maturation of inhibition. Under the subsequent MD condition, $z_{\max}$ and $z_{\min}$ change their values because of the change in the closed-eye input strength, ϵ. The bottom-right panel shows the equalized pattern of z_1 after MD – the same pattern as in the lower left panel, except that the values of $z_{\max}$ and $z_{\min}$ are changed to reflect the MD – along with the corresponding flow. The flow is barely changed, so that the equalized pattern remains stable under MD. This small change of $\Delta z_1(x)$ after MD can be understood from the fact that the change in z_1 is an almost equal increase across all x, that is, it represents the addition of a DC (spatially invariant) component; and K has a very small DC component, $\tilde{K}(0)$, due to the strong cortical inhibition. Even though z_1 is uniformly pushed toward more positive or less negative values, almost half of cortical areas still have negative $\Delta z_1(x)$, preventing the MD from destabilizing the equalized condition.

2 Parameter Dependence of Model Results

As described in the main text, both the Hebbian learning rule with subtractive weight normalization ("Hebbian rule") and the Hebbian learning rule with homeostatic regulation ("homeostatic rule") explained the equalization of OD columns by the maturation of inhibition and the OD shift under subsequent MD. However, the outcome of the Hebbian rule was more sensitive to intracortical connectivity and noise levels than the homeostatic rule. While the homeostatic rule robustly shows equalization and an MD shift with parameter set 1 as well as with parameter set 2, the Hebbian rule does not equalize with parameter set 2 (Fig. S2).

As shown in Fig.7, the depression of closed-eye synapses is conditioned on the monocular deprivation factor f_{MD}. The precise f_{MD} value at which closed-eye depression behavior switches to closed-eye homeostatic potentiation depends on the decay coefficient γ for active synapses; small γ makes the closed-eye-potentiating domain larger (Fig. S3)

3 Comparison of our Model to That of Swindale (1980)

As we discuss briefly in the main text, Swindale (1980, Appendix A) also addressed the problem of equalization and MD. Here we further discuss how these models differ.

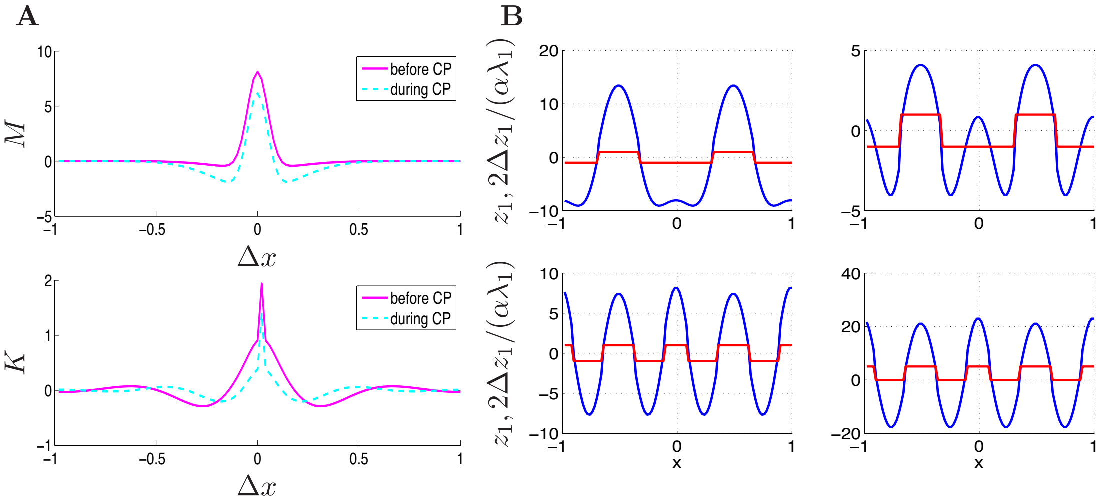

Figure S1: (A) The function describing intracortical connections, M (*Top*), and the resulting linear filter K (*Bottom*). The amount of cortical inhibition is increased af the onset of the CP. (B) Each panel shows the non-trivial eigen mode z_1 (red) and $\int K(x - x')z_1(x')dx'$ (blue), which up to a proportionality constant is the derivative Δz_1. Stability requires that regions where z_1 is positive have positive derivative, and negative regions have negative derivative. *Top-left*: The contralaterally-dominated pattern is stable before the maturation of inhibition. *Top-right*: The contralaterally-dominated pattern becomes unstable after the maturation of inhibition. *Bottom-left*: The equalized pattern becomes stable after the maturation of inhibition. *Bottom-right*: The equalized pattern remains stable under MD.

We define a variable $w_D(x) = w_C(x) - w_I(x)$ representing the difference between contralateral-eye (w_C) and ipsilateral-eye (w_I) input to cortical position x. Equalization means that $\tilde{w}_D(0) \propto \int_{-\infty}^{\infty} dx\, w_D(x) = 0$ where $\tilde{w}_D(0)$ is the Fourier transform of $w_D(x)$ at frequency 0, or the "DC" pattern of w_D. An ocular dominance shift toward the ipsi eye due to MD of the contra eye will manifest as $\tilde{w}_D(0) < 0$. We also define a variable $w_S(x) = w_C(x) + w_I(x)$. Due to our normalization constraint, this is equal to a constant, $w_S(x) = c_s$ for all x and all times. Swindale (1980, Appendix A) made a similar assumption.

The analysis of equalization and MD in Swindale (1980, Appendix A) considered a linear model of the dynamics of w_D. Such a model applies when the dynamics have a stable fixed point, and deviations from that fixed point are sufficiently small. A linear version of our model means that we are ignoring the nonlinearity represented by the $[]_+$ in Eq. 1, and ignoring nonlinearities associated with saturation of weights at maximal or minimal allowed values. When treated as linear equations, both our model and the model of Swindale (1980, Appendix A) yield an equation for $\tilde{w}_D(0)$ of the form

$$\frac{d}{dt}\tilde{w}_D(0) = \tilde{P}^D(0)\tilde{w}_D(0) + Z \tag{S12}$$

In (Swindale 1980, Appendix A), $\tilde{P}^D(0)$ is the Fourier transform at zero frequency of $P^D(x) = P^{CC}(x) + P^{II}(x)$ where $p^{XY}(x)$ is a phenomonological interaction between a synapse from eye X (contra or ipsi) and a synapse of eye Y separated across cortex by a distance x, and Z is c_s times the Fourier transform at zero frequency of $P^{CC}(x) - P^{II}(x)$ ((which is proportional to the integral of this quantity). In our model, $\tilde{P}^D(0) = \left(C^{CC} + C^{II} - (C^{IC} + C^{CI})\right)\tilde{K}(0)$,

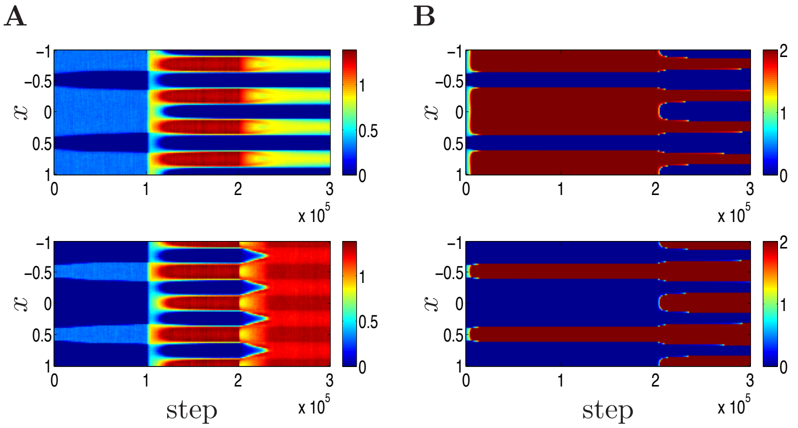

Figure S2: Development of OD columns under the homeostatic plasticity rule with the parameter set 1 (A) and under the pure Hebbian plasticity rule with the parameter set 2 (B). Panels plot synaptic strengths in color as a function of cortical position, x, and simulated timestep as in Fig. 2A or Fig. 4A. While the homeostatic learning rule show similar behavior for the two parameter sets, the pure Hebbian plasticity rule is sensitive to parameter choices and does not equalizes with the parameter set 2.

where C^{XY} describes the correlation between the activities of inputs serving eye X and eye Y and $\tilde{K}(0) = (1 - \tilde{M}(0))^{-1}$ where $\tilde{M}(0)$ is proportional to the integral of the function $M(x)$ describing intracortical connectivity, and $Z = c^s\left(C^{CC} - C^{II} + (C^{CI} - C^{IC})\right)\tilde{K}(0)$. In both models $Z = 0$ if the two eyes have equal activities, while if the ipsi (contra) eye is deprived Z becomes negative (positive).

The models differ in the sign of $\tilde{P}^D(0)$. Swindale assumed that this was negative. Then $\tilde{w}_D(0) = Z/|\tilde{P}^D(0)|$ is a stable fixed point. In our model, $\tilde{P}^D(0)$ is always positive: increasing the strength of inhibition makes $\tilde{M}(0)$ more and more negative but this makes $\tilde{P}^D(0)$ smaller and smaller while keeping it positive. (Note that stability of the activity dynamics requires that $\tilde{M}(0) < 1$, and in turn is required for the expression of the growth rate in terms of $\tilde{K}(0)$ to be valid, so $\tilde{K}(0)$ can never be negative). The result is that $\tilde{w}_D(0) = -Z/\tilde{P}^D(0)$ is an unstable fixed point, and a linear model cannot be sufficient to describe the dynamics. Equalization occurs because a periodic pattern of w^D grows sufficiently faster than the DC pattern in the linear regime that, in the nonlinear regime in which synapses are saturating, the DC pattern is suppressed.

In practice, the key differences between the models involve recovery from saturated initial conditions. In our model, both equalization and monocular deprivation begin from a condition in which most synapses are saturated, and as a result it can be difficult for the eye whose representation should expand (the open eye under monocular deprivation; the weaker ipsilateral eye in the case of equalization) to activate the cells that they would have to take over to expand. For example, in the case of monocular deprivation, as we discuss in the text, the saturated condition when MD is initiated means that the open-eye synapses are both near their maximal possible value in open-eye-dominated regions and near their minimal possible value in closed-eye-dominated regions. This, along with mature inhibition, means that stimulation of the open eye largely or only inhibits the cells dominated by the closed eye. As a result, the open eye has difficulty taking over the cells initially dominated by the closed eye without an additional mechanism, *e.g.* homeostatic plasticity, to allow weak synapses to

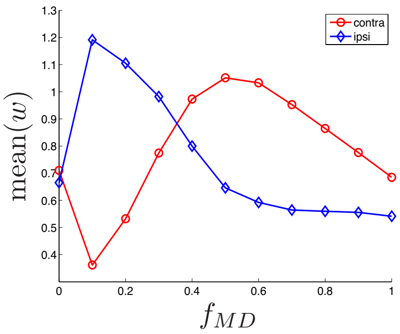

Figure S3: The precise f_{MD} value that yields switching from closed-eye depression to closed-eye homeostatic potentiation during MD depends on the decay coefficient γ for active synapses. With $\gamma = 5$Hz for $h_a > 1$Hz, the switching behavior occurs at around $f_{MD} = 0.3$ (c.f. Fig. 7).

recover and compete. Similar considerations apply to the case of equalization. In contrast, in a linear model, at the stable fixed point preceeding initiation of MD, ocular dominance varies in a sinusoidal rather than square-wave pattern. Thus, the open eye can activate cells in the closed-eye dominated regions near the borders of the regions, allowing the open eye to take over these cells and expand its representation under MD.