4.7.2 Stripe width in Miller's model

An important aspect of any model of ocular dominance stripe formation is the dependence of stripe width on the parameters of the model. It is argued in [Miller 1990(a)] that, if we now consider a complete layer of output units in the model, there are several different cases with respect to stripe width. Firstly, stripes only form if the fastest-growing eigenvector of S^D is monocular, and this depends on the parameters. If we assume this is true, then the width depends on the form of the cortical interaction function I, which we have not so far considered. If I is of Mexican hat form (Miller usually used a difference of gaussians), then the periodicity of the resulting ocular dominance stripes is related to the position of the peak of the fourier transform of the cortical interaction function. However, if I is all-excitatory, then Miller argues that the addition of constraints on input units is required to force segregation:

> In the case of a purely excitatory cortical interaction function, the peak of the fourier transform of I corresponds to an infinite wavelength, that is dominance by a single eye everywhere. Addition of constraints on the total synaptic strength of input cells modifies the equation, forcing the left- and right-eye inputs from each retinotopic location to divide between them the cortical surface over which they initially arborize. The result is to suppress the growth of monocular patterns with wavelength longer than an arbor diameter, so that the fastest-growing monocular pattern has a wavelength of about an arbor diameter [Miller 1990(a)].

We will discuss further the issue of including normalization constraints for presynaptic as well as postsynaptic units in chapter 6. In the meantime, we make the following observation with regard to the implications of this for Miller's monocular deprivation results. Miller presents simulations (e.g. [Miller et al 1989]) where monocular deprivation is modelled by reducing the strength of the correlations in one eye. Stripes from the deprived eye are narrower, as in the natural system. However, this only occurs because the normalization constraints on presynaptic units are now relaxed: in particular Miller allowed the sum of synaptic strengths for each presynaptic unit to vary between 0.5 and 1.5 times its initial value. Without this, presynaptic normalization compels each presynaptic unit to take over the same amount of cortex regardless of the strength of the correlations.

4.8 Conclusions

Miller's model provides an explanation for the formation of ocular dominance stripes that has the virtues of both rich possibilities for mathematical analysis, and a clear biological interpretation for some of its parameters. However the model has limitations. Firstly, it is unstable to various types of perturbative correlations, and monocularity breaks down in the presence of certain perturbative correlations either within or between eyes. Secondly, it has not yet been shown that the same mechanism can satisfactorily account for the development of topography.

Thus binocular solutions can be favoured over monocular solutions. However, the following two points need to be taken into consideration before these results can be used to make definite predictions about the weight structure that will finally dominate:

- The absolute size of the eigenvalues needs to be known. Even though monocular solutions are relatively punished, it might be that this effect is not enough for any particular perturbation to make a qualitative difference to the final outcome.
- The form of the perturbing correlation matrices needs to be known. Although our analysis is general, the concrete cases we concentrated on assumed that all components of each perturbing matrix have the same sign. More complicated perturbation matrices will lead to results that are impossible to predict without performing the precise calculations.

These results are significant because perturbations are inevitable in the natural system. It is implausible to think that the correlation functions are identical in the two eyes, or that correlations between the eyes are exactly zero. In the former case there is a particular form of instability in Miller's model, and in the latter case monocularity becomes less likely as cross-eye correlations increase.

It is important to make a distinction between different types of result. Firstly there is the question, are the eigenvectors stable to perturbations? We have shown that for some perturbations they are not. The other type of result is using perturbation theory to take a small (size ϵ) step in analysing the effect of more realistic correlations on model results. One step of this type is the analysis of cross-eye correlations. For animals such as cats and monkeys it seems highly likely (although not yet shown experimentally) that there should be substantial correlations between activities in the two eyes for the case of evoked activity. Perturbation theory gives a hint as to what would happen in Miller's model in this case. Another result of this type is the analysis of adding negative correlations within an eye: again, perturbation theory goes some way to analysing this case.

4.7 Other aspects of Miller's model

We now briefly discuss two other aspects of Miller's model that have relevance later in the thesis.

4.7.1 Constraints

An important point about Miller's system is the type of constraints used to limit synaptic strengths. Besides saturation limits he also used a type of normalization which we refer to as *subtractive*. This is to contrast it with the more common *divisive* type of normalization (as used for instance in the Neural Activity model [Willshaw & von der Malsburg 1976]). These are two different ways of keeping the sum of components or sum of squares of components of the weight vector constant for a particular unit, depending on the way these constraints are enforced. In divisive normalization, after each weight update, every component is divided by an appropriate number to enforce the constraint. In subtractive normalization, a constant amount is subtracted from each component, the constant vector being chosen so as to enforce the constraint. Usually we consider the case where all components of the constraint vector are the same. A detailed analysis of the difference between these cases in this context is presented in [Miller & MacKay unpublished]. For our purposes, there are two points of significance:

- Normalizing in either case by the sum of components is equivalent to normalizing according to the (z_{1S}, z_{1S}) eigenvector. Thus the amount of (z_{1S}, z_{1S}) in the evolving weight structure is kept constant, and perturbations to the eigenvalue of this eigenvector will not affect the nature of the competition.
- If the constraint is enforced divisively, the fastest-growing $(y_i, -y_i)$ solutions have to grow faster than (z_i, z_i) in order for monocularity to develop. However, if the constraint is enforced subtractively, the rate of growth of $(y_i, -y_i)$ only has to be positive to have a chance of winning the race to saturation.

For Miller's model, this means that monocularity can develop under *divisive* enforcement of constraints only if the two eyes are *anticorrelated*, whereas under *subtractive* enforcement of constraints there is the potential for monocularity to develop even when the two eyes are *positively correlated* (although we have shown above that this cannot be guaranteed). We discuss the difference between these two types of enforcement of the constraints in the context of a simple competitive model in chapter 6.

according to other eigenvectors. Thus just considering $(y_i, -y_i)$ we see that as before within-eye anticorrelations favour binocularity.

4.5.4 Perturbations in growth rates

We calculated above the change in the eigenvalues caused by perturbing correlations. However, perturbations can also affect the growth rates of the different eigenvectors in the context of the evolution of a general weight vector. We briefly demonstrate that, for a general weight vector satisfying Miller's initial condition that the difference between w_L and w_R is small, the effect of the perturbation is to bias the competition towards the zero DC (binocular) solutions.

Consider the expansion of (w_L, w_R) in terms of the eigenvectors of Q_*^ϵ after time t:

$$\begin{aligned}
\begin{pmatrix} w_L \\ w_R \end{pmatrix} &= \sum_{x_i} (\lambda_i + \epsilon\xi_i)^t \begin{pmatrix} \alpha_i x_i + \epsilon a_i \\ \beta_i x_i + \epsilon b_i \end{pmatrix} [(\alpha_i x_i + \epsilon a_i).w_L + (\beta_i x_i + \epsilon b_i).w_R] \\
&+ \sum_{x_i} (\lambda_i + \epsilon\zeta_i)^t \begin{pmatrix} -\beta_i x_i + \epsilon c_i \\ \alpha_i x_i + \epsilon d_i \end{pmatrix} [(-\beta_i x_i + \epsilon c_i).w_L + (\alpha_i x_i + \epsilon d_i).w_R] \\
&+ \sum_{y_i} (\mu_i + \epsilon\gamma_i)^t \begin{pmatrix} y_i + \epsilon e_i \\ -y_i + \epsilon f_i \end{pmatrix} [(y_i + \epsilon e_i).w_L + (-y_i + \epsilon f_i).w_R] \\
&+ \sum_{z_i} (\nu_i + \epsilon\delta_i)^t \begin{pmatrix} z_i + \epsilon g_i \\ z_i + \epsilon h_i \end{pmatrix} [(z_i + \epsilon g_i).w_L + (z_i + \epsilon h_i).w_R]
\end{aligned}$$

If we assume that the difference between w_L and w_R is of $\mathcal{O}(\epsilon)$, we can calculate terms of $\mathcal{O}(\epsilon)$ in the expansion, using $(\lambda_i + \epsilon\xi_i)^t \approx \lambda_i(1 + \epsilon t\xi_i/\lambda_i)$, *etc.* We find that $w_L - w_R$ grows as $\lambda_i^t(1 + \xi_i t)$ in the direction of x_i and μ_i^t in the direction of y_i. This extra t term could bias the competition towards the zero DC solutions, depending on the sign of ξ_i.

4.6 Summary and significance of the results

Linsker's correlational learning rule can be generalised to the binocular case. A particular version of this was studied independently by Miller: he looked in detail only at the case of identical correlations in the two eyes and zero correlations between the two eyes. We have presented a more general analysis, and shown what happens to the eigenvalues and eigenvectors when small perturbations to Miller's correlations are considered. The main results are as follows:

- When positive correlations are added between the eyes while correlations within each eye remain identical, the eigenvalues of the monocular eigenvectors decrease. The other non-zero DC eigenvectors have increased or decreased eigenvalues depending on whether they are of the form $(+,+)$ or $(+,-)$ respectively, but to a lesser extent. The previously degenerate zero DC eigenvectors split, one eigenvalue being slightly greater than before and the other being slightly less.

- If there are identical negative correlations within each eye (a case investigated empirically by Miller), all eigenvalues are reduced, but the monocular eigenvectors are reduced the most. In addition, eigenvectors begin to change their form.

- When the correlations within the two eyes are made slightly different, the zero DC eigenvectors change their character: they are not stable to the perturbation. This also breaks the degeneracy. If the perturbing correlations are negative in each eye, then the same conclusions hold as for the case of identical negative correlations as above.

- General perturbations can lead to a slightly increased *growth rate* for components of the evolving weight vector in the direction of binocular solutions as compared to an unchanged growth rate for monocular solutions.