January 17, 2000

GCNU TR 2000–005

Multiplicative Modulation of Bump Attractors

Maneesh Sahani   **Peter Dayan**

Gatsby Unit

Multiplicative Modulation of
Bump Attractors

Maneesh Sahani   **Peter Dayan**

Gatsby Unit

1 Introduction

A number of electrophysiological studies in visual and visuo-motor cortices have shown that the tuning curves of cells to visual stimulus parameters may be multiplicatively modulated by extra-retinal factors. One example is found in the gain fields of neurons in the posterior parietal cortex. Many of these cells exhibit tuning for the retinotopic location of a visually presented motor target; these tuning curves can be scaled in response to changes in body configuration (Andersen *et al.* 1985; Brotchie *et al.* 1995). More recently, shifts in attention have also been found to modulate neural responses multiplicatively in a number of visual cortical areas; for example, the orientation tuning curves of cells in area V4 (McAdams and Maunsell 1999), or direction tuning curves of cells in area MT (Treue and Trujillo 1999).

Salinas and Abbott (1996) showed in simulations that a form of multiplicative scaling can be reproduced by a simple network model. They considered a recurrent network of continuous output neurons, with lateral connection weights in a centre-surround ("mexican hat") configuration; that is, nearby neurons excite each other, while distant neurons are mutually inhibitory. Such networks have been previously shown to account for a number of the properties observed in cortical responses (Ben-Yishai *et al.* 1995; Somers *et al.* 1995; Carandini and Ringach 1997). Salinas and Abbott showed that if an additional, modulatory, input is provided equally to all cells in such a network, changes in the level of modulation result in apparently multiplicative changes in the output.

In this note, we examine this behaviour of the network more closely, presenting a theoretical analysis of the apparent multiplicative scaling. We show, both through this analysis and through simulations, that the multiplicative behaviour arises when the output of the network achieves a critical width, which is a characteristic of the recurrent connectivity. Thus, the shape of the tuning curves within the multiplicative scaling regime is determined by the recurrent weights of the network, rather than by the stimulus-driven input.

2 The Network

We consider a recurrently connected network of analog neurons. The neurons are tuned to a one-dimensional feature of the stimulus (such as its location in azimuth, or its orientation) and, for convenience, are arranged in a line or ring according to their preferred values, forming a topographic map of the stimulus feature. The recurrent connection strength between two neurons depends solely on their relative stimulus preferences.

The membrane activation level of the ith neuron at a time t, $u_i(t)$, is governed by the differential equation

$$\tau \frac{d}{dt} u_i(t) = -u_i(t) + h_i(t) \tag{1}$$

where τ is a membrane time constant and $h_i(t)$ is the time-varying total input to the ith neuron. This input has three parts: two external sources — both of which are stepped to fixed values initially, and then held constant in time — and a recurrent feedback component which evolves with the network dynamics:

$$h_i(t) = s_i + r + \sum_j J_{ij} m_j(t). \tag{2}$$

The first external source is a stimulus-tuned component, s_i, which depends on the stimulus presented to the network and on the tuning of the cell. The second external source is a modulatory component r, which is taken to be the same for all cells. Finally, the recurrent component of the input is the sum, as j ranges over all the cells in network, of the output of the jth cell $m_j(t)$, weighted by the connection strength from the jth cell to the ith, J_{ij}. In all of the networks we consider, the recurrent connections are symmetric, and depend only on the separation between the two neurons in the topographic map, that is, $J_{ij} = J(|i-j|)$. The connection strength may be positive or negative; the model does not include a separate class of inhibitory cells.

The output of the ith neuron $m_i(t)$ is related to the activation $u_i(t)$ by a nonlinear transfer function, g. We take this to be the threshold-linear function

$$g(u) = \begin{cases} \beta(u - T) & \text{for } u \geq T \\ 0 & \text{for } u < T \end{cases} \tag{3}$$

for some threshold T and slope β.

Since the external inputs are held fixed, and the recurrent weights are symmetric, the network activity evolves from its initial value to to a stable fixed-point attractor. Furthermore, provided the connection strengths are not too large, this attractor will be finite. We write μ_i for the output of the ith neuron at the fixed point; its value is obtained by setting the derivative du_i/dt to 0, resulting in the condition

$$\mu_i = g(\sum_j J_{ij}\mu_j + s_i + r). \tag{4}$$

In this note we will be interested in the properties of this stable solution, rather than in the dynamics of settling.

We will focus on two versions of the basic network.

In the first instance, we treat a network of finite size in general terms. For this analysis it is useful to collect the input and output variables into vectors representing the entire network. The fixed-point output vector will be written $\boldsymbol{\mu}$. The stimulus-related inputs are collected into the vector $\mathbf{s}$. The modulatory input is the same for all cells, and can thus be written $r\mathbf{1}$, where $\mathbf{1}$ is a vector of ones (of dimensionality appropriate to the context), and r is a scalar as before. The connection weights are collected into a symmetric Toeplitz matrix, J. Taking the transfer function g to act element-by-element on a vector argument, the condition on the attractor output can be written as

$$\boldsymbol{\mu} = g(\mathsf{J}\boldsymbol{\mu} + \mathbf{s} + r\mathbf{1}) \tag{5}$$

In the second case, we treat a special instance of the network in which the connectivity matrix and tuned input are trigonometric functions of the stimulus, which is taken to be an angle θ representing, for example, the direction of motion of a visual stimulus. In this case, it is convenient to take the number of neurons to approach the continuum limit. The stimulus-driven input and network output can then both be written as functions of stimulus angle, $s(\theta)$ and $\mu(\theta)$ respectively, through reference to the topographic map. The sum in (2) becomes an integral in the limit, and the translation invariance of the recurrent connectivity allows it to be written as a convolution with a symmetric kernel, $J(\theta)$. This kernel represents both the density and strength of connections between neurons whose preferred stimuli differ by an angle θ. The fixed-point condition for the continuous network is

$$\mu(\theta) = g\left(\int_{-\pi}^{\pi} \frac{d\theta'}{2\pi} J(\theta - \theta')\mu(\theta') + s(\theta) + r\right). \tag{6}$$

3 Finite Network

Experimental reports of multiplicative scaling are based on the modulation of a single neuron's tuning curve by an extra-retinal signal. Multiple tuning curves are obtained for the same cell, for different settings of the modulatory parameter. These tuning curves are observed to all be centred on the same stimulus, and to be scaled versions of each other.

In simulations, and the subsequent analysis, it is convenient to replace the tuning curve taken for a single cell by the profile of fixed-point activation across the entire network in response to a single stimulus. The profile obtained in this way may be treated as a tuning curve by plotting the output of each cell against its preferred stimulus value. For networks such as those we examine here, which are translationally invariant, symmetric and have stationary parameters, this curve will be identical to that obtained by plotting the fixed-point output of a single neuron in response to various stimuli.
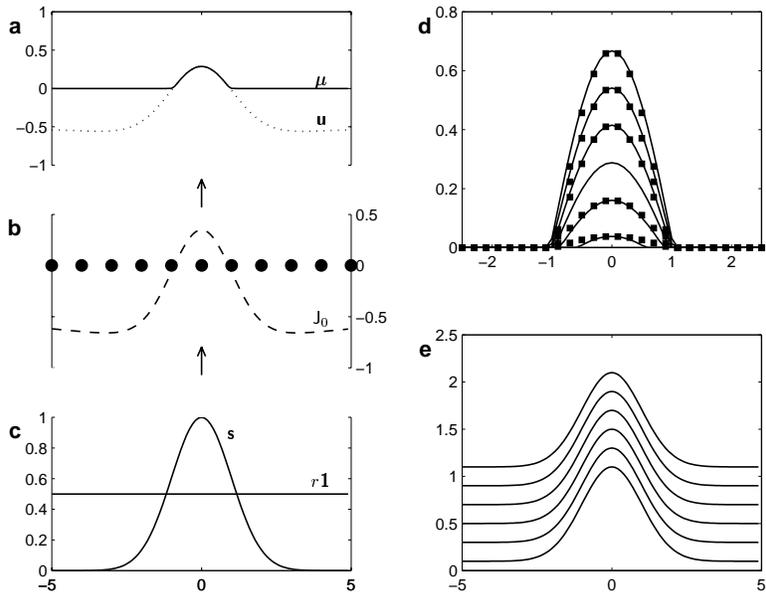
Figure 1: The finite network. Panel **b** shows every tenth cell (along with the 100th cell) in the network of 100 neurons, with units arranged according to their stimulus preference. The superimposed dashed curve ($\mathbf{J}_0$) shows the connection strength from each of the units to the central neuron, which has preferred stimulus 0. Typical inputs to the network are shown in panel **c** (**s** and $r\mathbf{1}$). Panel **a** shows the resulting output; the solid line ($\boldsymbol{\mu}$) shows the fixed-point output, the dotted line (**u**) shows the unrectified output at the fixed point, given by $\beta(\mathbf{J}\boldsymbol{\mu} + \mathbf{s} + r\mathbf{1} - T)$. The solid lines in panel **d** show the central part of the fixed-point output of the network at a variety of modulatory input levels; the summed external inputs ($\mathbf{s} + r\mathbf{1}$) in each case are shown in panel **e**. No squares appear over the output corresponding to a modulation level of 0.5. The squares plotted over the remaining outputs are obtained by multiplication of this output curve. The density of neurons is twice that of the squares. Note the restricted horizontal scale in panel **d**, showing only the central segment of the network. Panels **d** and **e** were adapted from Salinas and Abbott (1996).

3.1   Simulation

The network we use for simulation is identical to that of Salinas and Abbott (1996) and is illustrated in figure 1**a**–**c**. We use 100 cells, with preferred stimulus values ranging from -5 to 4.9 in arbitrary units. The recurrent connection strengths are given by the difference of two Gaussians:

$$J_{ij} = \left[A_E \exp\left(-\frac{(x_i - x_j)^2}{2\sigma_E^2} \right) - A_I \exp\left(-\frac{(x_i - x_j)^2}{2\sigma_I^2} \right) \right] \delta x \tag{7}$$

Here A_E (A_I) sets the strength, and σ_E (σ_I) the extent, of the excitatory (inhibitory) connections. The values x_i and x_j are the preferred stimuli for the ith and jth neurons respectively: thus the extents of the connections are given in the dimensions of the stimulus. The term δx gives the spacing between adjacent neurons; this allows the connection strengths to be specified in a neural-density-independent way. The values of the parameters are those of Salinas and Abbott (the δx term is implicit in their paper): $A_E = 10.5$, $A_I = 7$, $\sigma_E = 1$, $\sigma_I = 10$ and $\delta x = 0.1$. The transfer function parameters, in the notation of (3), are $T = 1$ and $\beta = 0.2$.

Tuned input to the network is Gaussian in shape. Since the network is translation invariant, we lose no generality in taking the presented stimulus position to be 0 in all cases. In this case, the stimulus-related input to the ith neuron is

$$s_i = A_s \exp\left(-\frac{x_i^2}{2\sigma_s^2} \right). \tag{8}$$

We set the strength of the input, A_s, to 1, and the width, σ_s, to 1. This last parameter is smaller than the value (1.5) used by Salinas and Abbott; this reduction will allow us to bring out the behaviour of the network more clearly.

3

The parameters described above are held fixed in all the simulations. Panel **c** of figure 1 shows typical values of the stimulus-related and modulatory inputs, with the resulting fixed-point output (equivalent to the tuning curve of the central neuron) shown in panel **a**. In panel **d**, the attractor outputs of the network for various different levels of modulatory input are shown. Despite an additive change in the total external input to the network (panel **e**), the output (and thus the tuning curves) appears to be multiplicatively scaled by the modulation.

3.2 Analysis

The equivalence of tuning curves and fixed-point output in the symmetric, translation-invariant network, implies that, for the network to show multiplicative scaling, the fixed-point output must change in proportion to its initial value as the modulation is changed. In other words exact multiplicative scaling requires the condition

$$\frac{d\boldsymbol{\mu}}{dr} \propto \boldsymbol{\mu}. \tag{9}$$

A feature of the centre-surround or "mexican hat" recurrent network architecture, evident in figure 1, is that the attractors exhibit "bump" output geometries; a group of cells whose preferred stimuli fall close to the current input are active (forming a bump in the topographic representation), while the surrounding neurons are all below threshold (Amari 1977; Ben-Yishai *et al.* 1995). In the networks we consider here, the inhibitory effect of a single neuron reaches all the other cells in the network. This will limit the dynamics of the network so that attractors will contain only a single bump.

Clearly, one necessary condition for the relationship of (9) to hold, is that the extent of this bump must remain constant as the modulatory input is changed. In particular, those neurons that fall outside the bump should remain below threshold, thus exerting no effect on the level of activation of those cells that fall within it. As a result, if the network does indeed behave multiplicatively, these neurons within the bump may be viewed as comprising a small sub-network, effectively isolated from the remaining units. By definition, all of the neurons within this restricted sub-network have non-zero output. Thus, their activation levels are above threshold, and fall within the linear portion of the transfer function. This feature makes the restricted network amenable to direct solution.

Suppose a central group of n neurons is active in the bump. We write the vector variables associated with the restricted network as follows: $\boldsymbol{\mu}_{\overline{n}}$ contains the non-zero elements of the steady-state output vector $\boldsymbol{\mu}$; $\mathbf{s}_{\overline{n}}$ contains the corresponding elements of the stimulus-related input $\mathbf{s}$; and the square matrix $\mathsf{J}_{\overline{n}}$ represents the corresponding rows and columns of J.

Since all neurons within the restricted network are above threshold, the fixed-point satisfies the linear equation:

$$\boldsymbol{\mu}_{\overline{n}} = \beta(\mathsf{J}_{\overline{n}}\boldsymbol{\mu}_{\overline{n}} + \mathbf{s}_{\overline{n}} + (r - T)\mathbf{1}). \tag{10}$$

Stability requires that all the eigenvalues of $\mathsf{J}_{\overline{n}}$ be less than β^{-1}, and thus we can solve this equation to obtain

$$\boldsymbol{\mu}_{\overline{n}} = (\beta^{-1}\mathbb{I} - \mathsf{J}_{\overline{n}})^{-1}(\mathbf{s}_{\overline{n}} + (r - T)\mathbf{1}), \tag{11}$$

where $\mathbb{I}$ is the identity matrix.

Thus, within the restricted network, the derivative of the attractor with respect to the modulation can be computed:

$$\frac{d\boldsymbol{\mu}_{\overline{n}}}{dr} = (\beta^{-1}\mathbb{I} - \mathsf{J}_{\overline{n}})^{-1}\mathbf{1}. \tag{12}$$

This result is crucial to our understanding of the multiplicative behaviour. It shows that the change in output due to a change in modulation is a function only of the network connectivity and the bump attractor width, not of the exact shape of the current attractor $\boldsymbol{\mu}$, or of the stimulus-related input $\mathbf{s}$. Thus, even if the bump width does remain constant, the condition (9) will only hold if, in addition, the non-zero segment of the output, $\boldsymbol{\mu}$, is proportional to $(\beta^{-1}\mathbb{I} - \mathsf{J}_{\overline{n}})^{-1}\mathbf{1}$. Inspection of (11) suggests that this may occur in two cases. The first case is when $(r - T) \gg s_i$; that is, the modulatory input dominates the tuned input in terms of the current flowing into the cells. The second case is when $(\beta^{-1}\mathbb{I} - \mathsf{J}_{\overline{n}})^{-1}\mathbf{s}_{\overline{n}}$ is similar in shape to $(\beta^{-1}\mathbb{I} - \mathsf{J}_{\overline{n}})^{-1}\mathbf{1}$; that is, within the restricted network the tuned input is similar to a constant input level (with the possible addition of eigenmodes suppressed by the matrix $(\beta^{-1}\mathbb{I} - \mathsf{J}_{\overline{n}})^{-1}$).
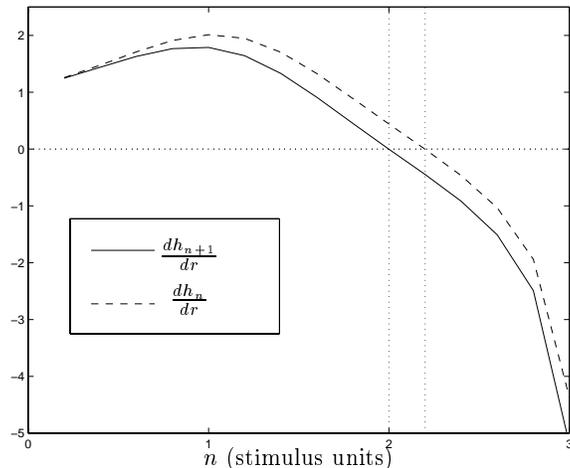
Figure 2: Sensitivity of marginal neuron input to change in modulation, plotted as a function of the attractor width, n (measured in terms of the range of preferred stimuli spanned by cells in the bump). The network is exactly as in figure 1. The solid line shows the derivative of the total input for the neuron just outside the bump, the dashed line for the neuron just inside. The difference between the zero-crossings of the two lines corresponds to a widening of the bump by a single unit on each side.

This latter case holds for stimulus-dependent input which is broadly tuned relative to the scale of the connectivity matrix and the attractor bump, so that it is effectively flat within the activated region of the network. Salinas and Abbott (1996) use precisely such broadly tuned input.

The analysis to this point has assumed that the width of the stable output bump remains constant as the modulation changes. This is a necessary condition for multiplicative modulation, but does it actually hold true in networks of the type discussed here?

Consider the case in which the network inputs have been chosen to yield an output vector $\boldsymbol{\mu}$ that exhibits a bump of width n. We can assess the stability of this width by finding the change in the total input to a neuron just outside the bump (identified by the subscript $n+1$) as a result of a change in modulation input. The total input to this marginal neuron is

$$h_{n+1} = \mathbf{j}_{n+1}^T \boldsymbol{\mu} + s_{n+1} + r, \tag{13}$$

where $\mathbf{j}_{n+1}$ is the vector formed from the $(n+1)$th row of the connection matrix. Since only neurons within the bump are active, we can restrict the vectors in this expression to these central n neurons:

$$h_{n+1} = \mathbf{j}_{n+1,\overline{n}}^T \boldsymbol{\mu}_{\overline{n}} + s_{n+1} + r. \tag{14}$$

Using (12), we thus find the sensitivity of this marginal input to changes in modulation:

$$\frac{dh_{n+1}}{dr} = \mathbf{j}_{n+1,\overline{n}}^T (\beta^{-1}\mathbb{I} - \mathsf{J}_{\overline{n}})^{-1} \mathbf{1} + 1. \tag{15}$$

Again, this result is remarkable in that, as we saw in the case of the change in output, the result is independent of the stimulus-tuned input, as well as of the detailed shape of the steady-state output. It depends only on the network architecture and on the current output width.

Figure 2 shows the dependence of the marginal input derivative (dh_{n+1}/dr) on the bump width for the network with parameters taken from Salinas and Abbott (1996), along with a similar input derivative calculated for the neuron just inside the bump attractor (dh_n/dr). These curves demonstrate that for a range of bump widths, an increase in the modulation strength will drive the bump towards a characteristic, stable, width.

If the output starts relatively narrow, an increase in modulation drives up the input to the neuron that lies just outside the bump. A sufficient increase in modulation raises the activation level of this unit above threshold, and thus recruits it into the central bump. This process repeats until the bump reaches a width of 2.0 stimulus units, at which point the net change in input to the marginal unit is almost zero (it is actually very slightly negative). Subsequent increases in modulation leave the bump-width unchanged. The complementary behaviour is observed if the initial output is larger than the

5
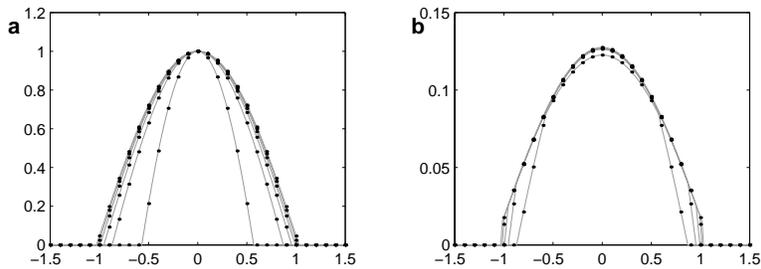
Figure 3: A closer look at the output of the finite network. Panel **a** shows the output curves of figure 1**b**, each having been rescaled to achieve a unit maximum. To focus on properties of the bump, only the central 31 neurons are shown. The dots represent the outputs of the cells in the finite network, the grey lines indicate the corresponding curves for a continuous network with the same connectivity. The inner-most curve corresponds to the smallest modulatory input, with the widths increasing as the modulation rises. Panel **b** shows the differences between the successive outputs shown in figure 1**b**. Again, the inner-most curve corresponds to the two smallest modulation values, and the curves grow wider with increasing modulation.

characteristic width. In this case, an increase in modulation drives down the input to the neuron that lies just inside the bump (as well as to the one just outside). Thus the bump narrows until, once again, a stable point is reached. In this case this stable width appears to be slightly larger, at 2.2; the difference is a result of the finite network size. In a continuous model, a single stable size emerges regardless of direction of approach.

This behaviour can be observed directly in the outputs of the simulation described above. If the fixed-point output curves shown in figure 1 are rescaled to have equal heights, as in figure 3**a**, the progressive widening indicated by the analysis becomes obvious. Furthermore, it is clear that once the limiting width has been achieved, the overall output shape does indeed remain constant (provided that inhibition is broad enough to prevent additional bumps forming as the modulation drives cells far from the center above threshold). Figure 3**b** shows the differences between successive output curves, thus approximating the derivatives of (12). We see that even for narrow outputs (and, therefore, smaller restricted networks) the shape of the derivative resembles a truncated version of the eventual stable output shape. Thus the output is being pushed further towards the characteristic output shape of the network with each increase in modulation. By the time the limiting width is achieved, the output is being driven in large part by the modulatory input.

4   Continuous Network

We turn now to a continuum model in which the neurons are taken to form a dense array and the recurrent weights and tuned input are both given by trigonometric functions. This network was investigated as a model of orientation tuning in V1 by (Ben-Yishai *et al.* 1995), and possesses all the features needed to exhibit the multiplicative scaling that is the subject of this note. It has the advantage that in many cases we can arrive at algebraic solutions to various equations. To avoid complications which arise from the π-periodicity of orientation, we choose to construct a model analogous to that of Ben-Yishai *et al.*, but on a 2π-periodic feature such as stimulus motion direction.

The model assumes a continuum of neurons tuned to the direction of a moving visual stimulus. The preferred directions of the neurons will be indicated by the continuous variable $\theta \in [-\pi, \pi)$. The density of connections between neurons tuned to the orientations θ_1 and θ_2 is given by

$$J(\theta_1, \theta_2) = \frac{1}{\pi} \left(J_0 + J_1 \cos(\theta_1 - \theta_2) \right) \tag{16}$$

with $J_0 < 0$ and $J_1 > 0$. We assume without loss of generality that the visual stimulus is presented at 0 orientation. The input tuning curves of the cells are also taken to be cosine functions, resulting in stimulus-related input given by

$$s(\theta) = s_1 \cos \theta. \tag{17}$$

6
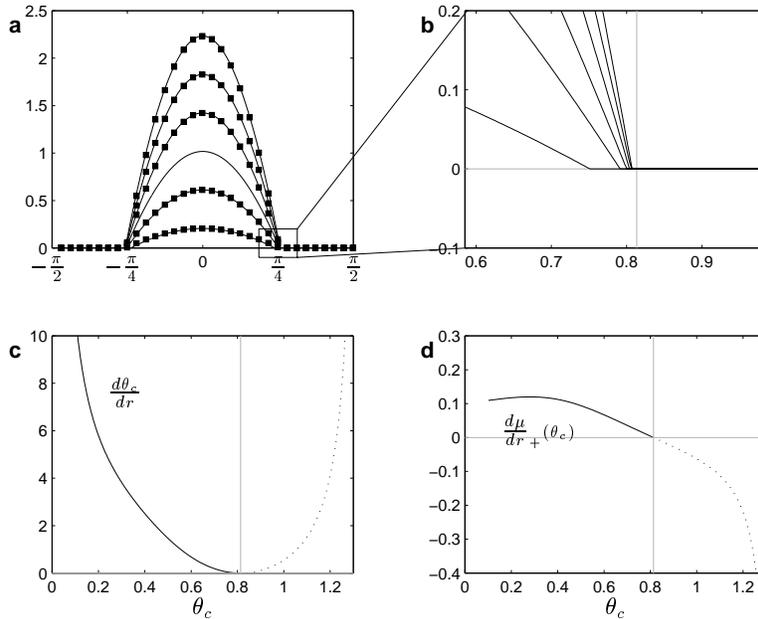
Figure 4: Output of the continuous network. Panel **a** shows the output of the network at various modulation levels, $r = 1, 4, 7, 10, 13, 16$. See figure 1 for an explanation of the squares. Panel **b** shows the same outputs on a larger scale, focusing on the change in θ_c. Panel **c** shows the modulation-induced change in attractor width as a function of θ_c. The zero corresponds to the observed limit width in the simulations. Panel **d** shows the change in output of the marginal unit as the modulation is increased (compare to figure 2). The dotted vertical lines in panels **b-d** all correspond to the limiting $\theta_c^* = 0.8134$. Only values of θ_c less than, or equal to this are achievable.

The modulatory input, r, is once again a constant independent of the neurons' preferred orientations.

The steady state equation for this model is given by

$$\mu(\theta) = g\left(\int_{-\pi}^{\pi} \frac{d\theta'}{\pi} \left(J_0 + J_1 \cos(\theta - \theta') \right) \mu(\theta') + r + s_1 \cos\theta \right) \tag{18}$$

If we write μ_0 for the zeroth Fourier series coefficient of the output, $\int \frac{d\theta}{\pi} \mu(\theta)$, and μ_1 for the first coefficient, $\int \frac{d\theta}{\pi} \cos(\theta) \mu(\theta)$, we obtain a solution for the steady state:

$$\mu(\theta) = \begin{cases} \beta(\mu_1 J_1 + s_1)(\cos\theta - \cos\theta_c) & |\theta| \leq \theta_c \\ 0 & |\theta| > \theta_c \end{cases} \tag{19}$$

Here, θ_c, the critical angle, defines the width of the central bump of the attractor. It is related to the inputs and the Fourier components of the output by the following system of equations:

$$\cos\theta_c = \frac{T - (r + \mu_0 J_0)}{s_1 + \mu_1 J_1} \tag{20}$$

$$\mu_0 = \frac{2\beta}{\pi}(\mu_1 J_1 + s_1)(\sin\theta_c - \theta_c \cos\theta_c) \tag{21}$$

$$\mu_1 = \frac{\beta}{\pi}(\mu_1 J_1 + s_1)(\theta_c - \tfrac{1}{2}\sin 2\theta_c). \tag{22}$$

The fixed-point output curves for a number of different modulation levels are shown in figure 4**a**. These plots were made using the following parameter settings: $J0 = -86$, $J1 = 100$, $\beta = 1$, $T = 1$, $s_1 = 1.5$. The enlargement in panel **b** reveals a slow change in the attractor width.

For the continuous network we can differentiate the expression for the fixed-point output $\mu(\theta)$ with respect to the modulatory input level r directly, without the need to assume a restricted sub-network of constant size. For notational compactness, we introduce the terms $\nu_0 = r + \mu_0 J_0$, $\nu_1 = s_1 + \mu_1 J_1$ and $W_i = \beta J_i / \pi$ for $i = 0, 1$. The derivative is well-defined everywhere but at the critical angle.

$$\frac{d\mu}{dr}(\theta) = \begin{cases} \beta(\nu_0' + \nu_1' \cos\theta) & |\theta| < \theta_c \\ 0 & |\theta| > \theta_c \end{cases}. \tag{23}$$

7

The primes denote derivatives with respect to r, which are given by

$$\nu_0' \;=\; \frac{1 - W_1(\theta_c + \frac{1}{2}\sin 2\theta_c)}{D}, \tag{24}$$

$$\nu_1' \;=\; \frac{2W_1 \sin\theta_c}{D}, \tag{25}$$

where

$$D = (1 - 2W_0\theta_c)(1 - W_1(\theta_c + \tfrac{1}{2}\sin 2\theta_c)) - 4W_0 W_1 \sin^2\theta_c. \tag{26}$$

At the critical angle the derivative is not well-defined; however the left and right derivatives (for $dr \to 0^-$ and $dr \to 0^+$ respectively) can be evaluated:

$$\frac{d\mu}{dr}_- (\theta_c) = 0 \tag{27}$$

$$\frac{d\mu}{dr}_+ (\theta_c) = \beta(\nu_0' + \nu_1' \cos\theta_c). \tag{28}$$

The second of these derivatives corresponds (upto a factor of β) to the quantity dh_n/dr calculated in (15) above. It is shown in figure 4**d**.

In this case we can exploit the existence of these derivatives to also directly calculate the change of the attractor output width with changes in modulation. Differentiating (20) and using the expressions for ν_i' above, we find that

$$\frac{d\theta_c}{dr} = \frac{\left(1 - W_1(\theta_c - \frac{1}{2}\sin 2\theta_c)\right)^2}{D s_1 \sin\theta_c}. \tag{29}$$

This derivative is shown as a function of θ_c in figure 4**c**. The limiting width of the network, θ_c^*, is achieved when $d\theta_c/dr = 0$. For the cosine network this implies the simple condition

$$\sin 2\theta_c^* = 2\left(\theta_c^* - \frac{\pi}{\beta J_1}\right) \tag{30}$$

For the values of the parameters that were used in figure 4, numerical solution of this equation yields the value $\theta_c^* = 0.8134$.

5 Discussion

The modification of sensory neuronal responses by changes in modulatory variables, such as body configuration or attention, is a subject of considerable theoretical and experimental interest (Andersen *et al.* 1985; Pouget and Sejnowski 1992; Salinas and Abbott 1995; Andersen and Zipser 1988; Salinas and Abbott 1997). A number of experiments, in a variety of visual and visuo-motor cortical areas, have suggested that the modulatory influence takes the form of a gain-change or multiplicative scaling in the neuronal response curves (Andersen *et al.* 1985; Brotchie *et al.* 1995; McAdams and Maunsell 1999; Treue and Trujillo 1999). Particularly in the attentional data, the support for precisely multiplicative modulation seems strong.

Following Salinas and Abbott (1996), we have seen that a form of multiplicative scaling can arise in a simple recurrent network model with centre-surround lateral connectivity, broadly-tuned stimulus-related input, and modulation that results in a uniform input current in all cells. Where Salinas and Abbott studied the behaviour of the network through simulation alone, we have here added an analytic understanding of some of its properties.

Our analysis indicates that the network exhibits approximately multiplicative scaling within a particular regime. The two central results, (12), (15), show that as the modulatory input is varied, the fixed-point output of the network changes shape in a way that depends only on the network connectivity and the extent of the non-zero region of the current attractor. In particular, it does not depend on the details of the current output shape, nor on the stimulus-related input. Only when the extent of the central bump reaches a particular limiting width, characteristic of the network, and also the fixed-point output has the same shape as the modulation-related change (given by (12)), will exactly multiplicative

scaling be observed. Thus, the shape of the multiplied curve is dictated by the network connectivity, rather than by the shape of the input.

Fortunately, this regime is easy to achieve. We have seen that an increase in modulation drives the output of the network towards this characteristic width and shape, regardless of the output response to the stimulus-related input alone. Furthermore, a broadly tuned stimulus-related input may carry a considerable flat component when restricted to the neurons that fall within this characteristic width. If this is the case, the stimulus-related input itself will drive the network towards the characteristic output. Finally, the network might operate in the "marginal phase" regime (identified by Ben-Yishai *et al.* 1995) where the amplitude of the stimulus-related input is extremely small and serves only to break symmetry in the network, while the network output is driven by a combination of a low threshold and a high "background" un-tuned input.

Recurrent networks with this type of centre-surround connectivity have been studied before in the context of the sharpening of broad thalamic input and generation of contrast-invariant tuning curves in primary visual cortex (Ben-Yishai *et al.* 1995; Somers *et al.* 1995; Carandini and Ringach 1997). In particular, these studies argued that the shapes of the tuning curves of cells may be dominated by the effects of lateral connectivity within the network, rather than by the shape of the input delivered to it. We observe here that it is precisely in those cases where this is true, that the network can also exhibit multiplicative modulation. Indeed, in the model networks, the two types of behaviour arise through virtually identical mechanisms.

Clearly, this is not the only mechanism through which multiplicative scaling effects may arise. However, given the known prevalence of lateral connections within the neocortex, care must be taken to evaluate alternative models in the context of recurrent circuitry. For example, a simple hypothesis might be that the extra-retinal modulation directly affects the gain of the individual cells directly — modifying the slope parameter, β, of the transfer function — perhaps through some neuromodulator pathway. However, in a centre-surround network of the type discussed here, such a manipulation does not result in purely multiplicative scaling. Simulations using the Gaussian network described above (not shown) indicate that as β is increased, with the input, which is now entirely stimulus-related, being held fixed, the fixed-point output grows both taller *and narrower*. This narrowing behaviour can also be seen by inspection of (30). Thus, in the presence of strong lateral connections, a direct neuronal gain modulation does not result in multiplicative scaling of the fixed-point output.

One important discrepancy between the data and the model we have discussed must be noted. In some experiments (notably that of McAdams and Maunsell 1999) the response of a neuron is, on average, elevated above its background level in the presence of any stimulus, even when this stimulus is orthogonal to its preferred one. Furthermore, the amplitude of even this cross-orientation response appears to be multiplicatively scaled by attention. The mechanism for scaling discussed here depends on a region of zero output outside the central bump. Even if this zero idealized analog output were to correspond to a non-zero, but small, firing rate in a real, noisy neuron, we would not expect it to be affected by the modulation. Thus, some extension of this simple model is needed to capture this aspect of attentional scaling.

6   Acknowledgments

References

Amari, S. (1977). Dynamics of pattern formation in lateral-inhibition type neural fields. *Biological Cybernetics 27*(2), 77–87.

Andersen, R. and D. Zipser (1988). The role of the posterior parietal cortex in coordinate transformations for visual-motor integration. *Canadian Journal of Physiology and Pharmacology 66*(4), 488–501.

Andersen, R. A., G. K. Essick, and R. M. Siegel (1985). Encoding of spatial location by posterior parietal neurons. *Science 230*(4724), 456–458.

Ben-Yishai, R., R. L. Bar-Or, and H. Sompolinsky (1995). Theory of orientation tuning in visual cortex. *Proceedings of the National Academy of Sciences of the United States of America 92*, 3844–3848.

Brotchie, P. R., R. A. Andersen, L. H. Snyder, and S. J. Goodman (1995). Head position signals used by parietal neurons to encode locations of visual-stimuli. *Nature 375*(6528), 232–235.

Carandini, M. and D. L. Ringach (1997). Predictions of a recurrent model of orientation selectivity. *Vision Research 37*(21), 3061–3071.

McAdams, C. J. and J. H. R. Maunsell (1999). Effects of attention on orientation-tuning functions of single neurons in macaque cortical area V4. *Journal of Neuroscience 19*(1), 431–441.

Pouget, A. and T. J. Sejnowski (1992). A distributed common reference frame for egocentric space in the posterior parietal cortex. *Behavioural and Brain Sciences 15*(4), 787–788.

Salinas, E. and L. Abbott (1995). Transfer of coded information from sensory to motor networks. *Journal of Neuroscience 15*(10), 6461–6474.

Salinas, E. and L. Abbott (1996). A model of multiplicative neural responses in parietal cortex. *Proceedings of the National Academy of Sciences of the United States of America 93*(21), 11956–11961.

Salinas, E. and L. Abbott (1997). Invariant visual responses from attentional gain fields. *Journal of Neurophysiology 77*(6), 3267–3272.

Somers, D. C., S. B. Nelson, and M. Sur (1995). An emergent model of orientation selectivity in cat visual cortical simple cells. *Journal of Neuroscience 15*(8), 5448–5465.

Treue, S. and J. C. M. Trujillo (1999). Feature-based attention influences motion processing gain in macaque visual cortex. *Nature 399*, 575–579.