

Attentional Recruitment of Inter-Areal Recurrent Networks for Selective Gain Control

Richard H. R. Hahnloser

rhahnloser@mit.edu

Howard Hughes Medical Institute, Department of Brain and Cognitive Sciences, MIT, Cambridge, MA 02139 U.S.A.

Rodney J. Douglas

rjd@ini.phys.ethz.ch

Institute of Neuroinformatics ETHZ/UNIZ, CH-8057 Zürich, Switzerland

Klaus Hepp

klaus.hepp@itp.phys.ethz.ch

Institute for Theoretical Physics, ETHZ, Hönggerberg, CH-8093 Zürich, Switzerland

There is strong anatomical and physiological evidence that neurons with large receptive fields located in higher visual areas are recurrently connected to neurons with smaller receptive fields in lower areas. We have previously described a minimal neuronal network architecture in which top-down attentional signals to large receptive field neurons can bias and selectively read out the bottom-up sensory information to small receptive field neurons (Hahnloser, Douglas, Mahowald, & Hepp, 1999). Here we study an enhanced model, where the role of attention is to recruit specific inter-areal feedback loops (e.g., drive neurons above firing threshold). We first illustrate the operation of recruitment on a simple example of visual stimulus selection. In the subsequent analysis, we find that attentional recruitment operates by dynamical modulation of signal amplification and response multistability. In particular, we find that attentional stimulus selection necessitates increased recruitment when the stimulus to be selected is of small contrast and of small distance away from distractor stimuli. The selectability of a low-contrast stimulus is dependent on the gain of attentional effects; for example, low-contrast stimuli can be selected only when attention enhances neural responses. However, the dependence of attentional selection on stimulus-distractor distance is not contingent on whether attention enhances or suppresses responses. The computational implications of attentional recruitment are that cortical circuits can behave as winner-take-all mechanisms of variable strength and can achieve close to optimal signal discrimination in the presence of external noise.

1 Introduction

The primate visual cortex is divided into many distinct areas that are organized hierarchically by recurrent inter-areal connections. The various areas represent visual space nearly topographically, but the sensory receptive fields of their neurons are quite different. It appears that the role of inter-areal ascending projections is to form progressively higher representations of the visual world, for example, from neurons tuned to moving edges in V1 (Hubel & Wiesel, 1962) up to neurons tuned to the view of particular objects in area IT (Logothetis, Pauls, Bulthoff, & Poggio, 1994; Riesenhuber & Poggio, 1999). The size of neuronal receptive fields grows with hierarchical level. For example, in the macaque monkey, classical receptive field diameters grow from about 1 degree in V1 to 5 degrees in MT, 30 degrees in MST, and in IT they can be even larger than the entire contralateral visual field.

There is evidence that the descending inter-areal connections affect attentional selection and memorization of behaviorally relevant information (Motter, 1993; Tomita, Ohbayashi, Nakahara, & Miyashita, 1999). Attention-related neural activity has been observed in many cortical visual areas of the macaque monkey. The strength of attention decreases with level from MST, MT, V4, V2 down to V1 (Motter, 1993, 1994; Moran & Desimone, 1985; Luck, Chelazzi, Hillyard, & Desimone, 1997; Colby, Duhamel, & Goldberg, 1996; Treue & Maunsell, 1996, 1999; Fuster, 1990; Desimone, 1996). There is evidence that the origin of attentional signals is in prefrontal cortex (Tomita et al., 1999). The generally low firing rates in higher areas suggest that the attentional control circuitry recruits or derecruits feedback networks with neurons in lower areas by driving neurons above or below firing threshold.

Several experiments provide evidence of the involvement of inter-areal feedback in the attentional selection of low-contrast stimuli. In V4, responses of neurons to very low-contrast stimuli are not increased when they are attended in the presence of high-contrast distractors: Response enhancement is possible only above a critical contrast of about 5% (Reynolds, Pasternak, & Desimone, 2000). Also, in the experiments of De Weerd, Peralta, Desimone, and Ungerleider (1999), it was found that restricted lesions of areas V4 and TEO result in monkeys being unable to report the orientation of low-contrast gratings in the presence of high-contrast distractors. However, the monkeys' perceptual performance for low-contrast stimuli was almost unchanged when the distractors were not present. This suggests that the circuits in TEO-V4 are highly involved in attentional selection of low-contrast stimuli rather than in reading out of stimulus orientation.

Previously, we have noted that if attentional stimulus selection is induced by an excitatory top-down bias, then this bias necessarily also leads to a bias in the readout of the selected stimulus (Hahnloser et al., 1999). Here, we resolve this selection/readout dilemma by considering attentional inputs that are just strong enough to recruit neurons in the higher area and their

feedback loops with neurons in the lower area, but without providing for an additional input bias. Stimulus selection is achieved by recruiting more feedback loops for stimuli that are to be attended than for distractor stimuli.

We cast our network as a simple model of two recurrently connected areas, such as MST–MT, V5–V2, or TEO–V4. We explore the computational principles that could underlie the recruitment of feedback between large and small receptive field neurons by simulating physiological experiments in which multiple stimuli appear inside the receptive field of a large field neuron. We study the limits within which attentional selection is possible when the attended stimulus is of low contrast and at a small distance from distractor stimuli. By varying the amount of recruitment and the size of the attended stimulus, we explore the accuracy of readout in a noisy environment.

2 Network Equations

The firing rates of E excitatory neurons in the lower area are denoted by M_x ($x = 1, \dots, E$), those of N excitatory neurons in the higher area by P_i ($i = 1, \dots, N$) and those of I inhibitory neurons in the lower area by I_y ($y = 1, \dots, I$) (see Figure 1a). The indices x and y stand for a one-dimensional topography of the lower area, and the index i stands for a not necessarily topographic labeling of neurons in the higher area. The equations describing the evolution of firing rates are given by:

$$\dot{P}_i = -P_i + \left[p_i + \alpha_F \sum_{x=1}^E M_x \cos(\delta_x - \chi_i)_+ - t \right]_+ \quad (2.1)$$

$$\dot{M}_x = -M_x + \left[m_x + \alpha_B \sum_{i=1}^N P_i \cos(\delta_x - \chi_i)_+ - \beta \sum_{y=1}^I I_y \right]_+ \quad (2.2)$$

$$\dot{I}_x = -I_x + \left[\alpha_I \sum_{i=1}^N P_i \cos(\psi_x - \chi_i)_+ - \beta_I \sum_{y=1}^I I_y \right]_+ \quad (2.3)$$

Here $[f]_+ = \max(0, f)$ denotes rectification and ensures positivity of firing rates.

The receptive field centers δ_x of neurons M_x are regularly spaced, $\delta_x = \delta_{\max} \frac{x-1}{E-1}$. Similarly, the receptive field centers of neurons I_y are given by $\psi_y = \psi_{\max} \frac{y-1}{I-1}$. The inter-areal connections are purely excitatory. Their strength decays as a function of receptive field separation z , according to $\cos(z)_+ = [\cos(z)]_+$. There is uniform inhibition in the lower area (the assumption of uniformity is a simplification that could be relaxed; see Wersing, Beyn, & Ritter, 2001). The parameters α_F , α_B , and α_I determine the strength of excitation, and β and β_I the strength of inhibition.

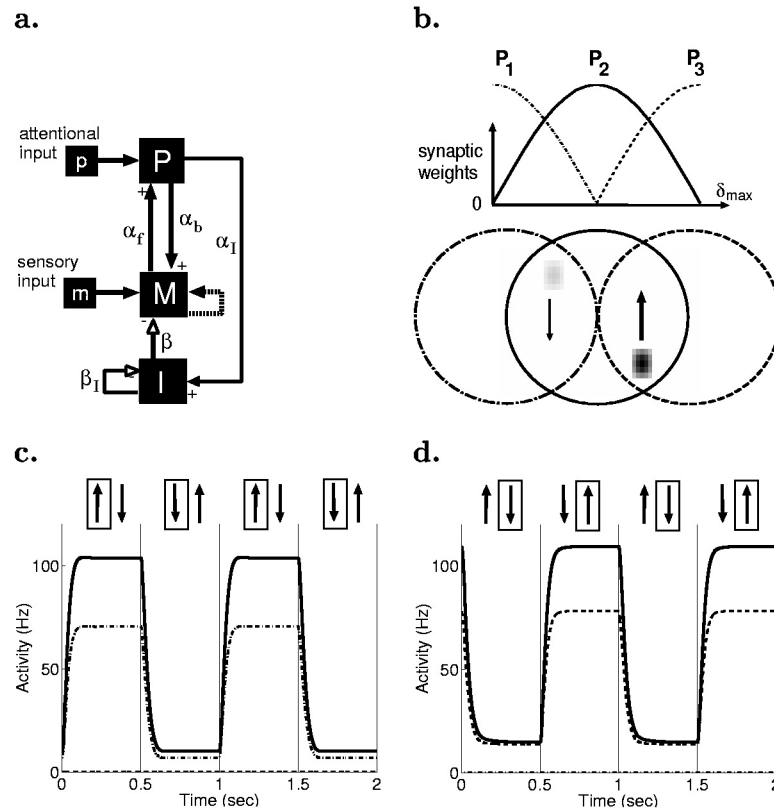


Figure 1: Attending to one of two stimuli moving in antiphase. (a) Schematic of network architecture of excitatory neurons in higher area (P) and excitatory and inhibitory neurons in lower area (M and I). The Greek letters denote the synaptic coupling strengths. (b) Two moving visual stimuli are placed in the receptive field of neuron P_2 in area MST (thick circle, schematic). Synaptic weights of the three MST neurons with MT neurons are shown on top. (c) The attentional selection of the left stimulus (indicated by the rectangle) is modeled by recruiting pointer neurons P_1 and P_2 . Their responses are shown by the solid and the dash-dotted lines, correlating with the upward movement of the left stimulus. $p_1 = p_2 = 1, p_3 = 0$. (d) This time, pointer neurons P_2 and P_3 are recruited. Their responses (solid and dashed lines) correlate with the upward movement of the right stimulus. $E = 320, \delta_{max} = \pi, N = 3, I = 32, \psi_{max} = \pi, \chi_1 = 0, \chi_2 = \pi/2, \chi_3 = \pi, \alpha_F = .5, \alpha_B = 2, \alpha_l = 2, \beta = 60.07, \beta_l = 60, a = 6, t = 1, h_{up} = 1, h_{down} = .1$.

For the simulations, the visual input m_x contains either one or two localized stimuli of the form $m_x = h \cos(\frac{180^\circ}{a}(\delta_x - r))_+$ if $|\delta_x - r| \leq a/2$ and $m_x = 0$ otherwise. Here, h corresponds to the contrast of a stimulus (or its luminosity), a to the stimulus width (in degrees), and r to its retinal location, $0^\circ < r < \delta_{\max}$. Neurons P_i have nonzero firing thresholds $t > 0$ that express the difficulty of driving neurons in higher areas by visual stimulation alone. The attentional inputs p_i to neurons in the higher area are set to either zero (no recruitment) or t (recruitment). Appendix A describes a simple method for selecting appropriate values for the five coupling parameters α_F , α_B , α_i , β , and β_i .

3 Example of Recruitment

Typically, neurons in various higher visual areas are able to respond to just the attended stimulus in their receptive field, filtering out nonattended stimuli (Moran & Desimone, 1985; Desimone, 1998; Reynolds, Chelazzi, & Desimone, 1999). For example, Treue and Maunsell recorded from neurons in areas MT and MST of alert monkeys during a visual attention task involving two moving stimuli (Treue & Maunsell, 1996, 1999). The monkeys were instructed to attend to one of them and to respond quickly to a change of speed. Both stimuli fell inside the receptive field of a recorded neuron, and alternately, one stimulus moved in the neuron's preferred direction while the other moved in the antipreferred direction. They found that most of the time, the neuronal response correlated strongly with the direction of motion of the attended stimulus, not the distractor stimulus. And when monkeys attended to a stimulus outside the receptive field, neuronal response was suppressed, not correlating with the preferred movement direction of either of the two stimuli in the receptive field.

We chose these experiments by Treue and Maunsell to illustrate the operation of recruitment (we do not provide a complete model for the MT-MST interactions). We simulated the response behavior of $N = 3$ motion-selective neurons P_1 , P_2 , and P_3 in area MST (the higher area) and $E = 320$ motion-direction selective neurons M_x in area MT (the lower area). Receptive field centers in MST are $\chi_1 = 0^\circ$, $\chi_2 = \delta_{\max}/2$, and $\chi_3 = \delta_{\max} = 180^\circ$. There are two vertically moving dots. Neuron P_2 sees both dots in its receptive field, whereas neurons P_1 and P_3 each see only one dot. The two dots oscillate in antiparallel directions to each other (see Figure 1b). Because we assume that the one-dimensional map in MT encodes only the horizontal dimension, "vertical movement" was simulated by contrast changes. We assume that all MT and MST neurons have an equal upward direction preference, simply modeled by setting the contrast of the "downward-moving dot" to one-tenth of the contrast for the "upward-moving dot" (in other words, the contrasts of the two dots flipped back and forth between two values).

The common threshold t of the neurons in MST is large. In the model, MST neurons are activated by visual stimulation only when they are also

recruited by attentional input, $p_i = t$. But only those neurons are recruited whose receptive fields overlap with the current focus of attention. For example, when the monkey attends to the left stimulus, neurons P_1 and P_2 are recruited, in which case neuron P_2 responds mainly during the upward movement of the dot on the left (see Figure 1c). Similarly, the response of P_2 is bound to the upward movement of the right dot when the monkey attends to the right and neurons P_2 and P_3 are recruited (see Figure 1d).

Neurons P_1 and P_3 are active only when the attended stimulus is the one in their receptive field, in which case they respond to its upward movement. When the monkey attends to the other stimulus outside their receptive field, they remain silent.

In analogy with the Treue and Maunsell experiments, attention causes stimulus competition beyond receptive field boundaries (neurons P_1 and P_3), as well as within receptive field boundaries (neuron P_2).

The explanation of why in Figure 1 the activity of neuron P_2 correlates with the movement direction of the attended stimulus is quite simple: the recruitment of either neuron P_1 or P_3 contributes feedback amplification, enhancing responses in MT to the left or right stimulus. And because responses are enhanced in MT, responses will be enhanced in MST as well. This explains why neuron P_2 —receiving the same attentional input $p_2 = t$ in both Figures 1c and 1d—can have a response that is biased according to which fellow MST neuron is recruited.

4 Loss of Attentional Selection for Low-Contrast and Nearby Stimuli —

In this section, we analyze the conditions under which attentional recruitment enables persistent selection of a behaviorally relevant stimulus in the presence of distractor stimuli. By “persistent selection,” we mean that neural responses are locked to the selected stimulus and persist when distractor stimuli change (in other words, neural responses are as if the distractors were not present). We explore the sensitivity of persistent selection to various parameters such as stimulus contrast (relative to distractor contrast) and stimulus location (relative to distractor location).

We will consider only a one-dimensional network and nearby stimuli that fall between $r = 0^\circ$ and $r = 90^\circ$. Consequently, we restrict the map in the lower area to $\delta_{\max} = 90^\circ$ and $\psi_{\max} = 90^\circ$. Because we are mainly interested in the effects of recruitment, half of the neurons in the higher area have receptive field centers at $\chi_1^i = 0^\circ$ and the other half at $\chi_2^i = 90^\circ$ ($i = 1, \dots, N/2$). By discretizing the receptive field centers to just two values separated by 90 degrees, the population activity in the higher area gets a simple interpretation: the activity vector formed by pairs $\mathbf{P}^i = (P_1^i, P_2^i)$ of neurons forms a vector whose direction indicates the center of activity in the lower area (and thus the location of the selected stimulus). This population vector property stems from the geometrical fact that there are sine and cosine synaptic connection profiles between the two areas (which in turn is based

on the equality $\cos(\alpha - 90^\circ) = \sin(\alpha)$. As in our previous work, the activity vectors \mathbf{P}^i in the higher area shall be referred to as pointers (Hahnloser et al., 1999).

Recruiting pointers that share receptive field centers is mathematically equivalent to changing the synaptic weights α_F , α_B , and α_I made by a single pointer. In Figure 2a, a stimulus and a distractor of equal contrast are presented to the network at three different separations. The steady response in the higher area is read out as the pointer angle,

$$\gamma = \arctan\left(\frac{\sum_i P_2^i}{\sum_i P_1^i}\right),$$

and is plotted as a function of the number N_+ of recruited pointers (N_+ is defined by $p_1^k = p_2^k = t$ for $k \leq N_+$ and $p_1^k = p_2^k = 0$ for $k > N_+$). All recruited pointers are initialized so as to express a preattentive bias to the left, $\mathbf{P}^i(0) = (2, 0)$. This initialization tends to induce attentional selection of the stimulus on the left.

When stimulus and distractor are directly adjacent to each other, then persistent selection arises only for about $N_+ > 20$. Persistent selection requires fewer pointers the farther apart the stimuli are.

In a similar way as for stimulus separation, the selection of a low-contrast stimulus is possible only within limits. Figure 2b shows a diagram in which we plot (as a function of N_+) the minimal relative contrast of a stimulus that still permits its selection by attentional recruitment. Relative contrast is defined as the contrast of the stimulus divided by the contrast of the distractor.

When the stimulus is close to the distractor and 32 pointers are recruited, then persistent selection is possible only if its contrast is at least 20% of the distractor contrast. However, when the stimulus is farther away, then many fewer pointers are required for persistent selection at the same relative contrast. In other words, for a fixed number of recruited pointers, the minimal stimulus contrast allowing for persistent selection increases as the stimulus and the distractor move closer together.

We have analyzed the sensitivity of these results to the strength of synaptic weights. Interestingly, we found that altering the strength of the excitatory feedback (e.g., decreasing α_B by 4%) does not have a noticeable influence on the distance sensitivity in Figure 2a. However, this decrease of excitatory feedback has a dramatic effect on the contrast sensitivity in Figure 2b. The reduced strength of excitatory feedback for the intermediate stimulus-distractor separation results in a highly reduced performance for selecting low-contrast stimuli. In the next section, we show that the reason for the decreased selectability is that a small reduction in feedback can cause the net effect of recruitment on neurons in the lower area to be inhibitory rather than excitatory, without affecting the strength of competition.

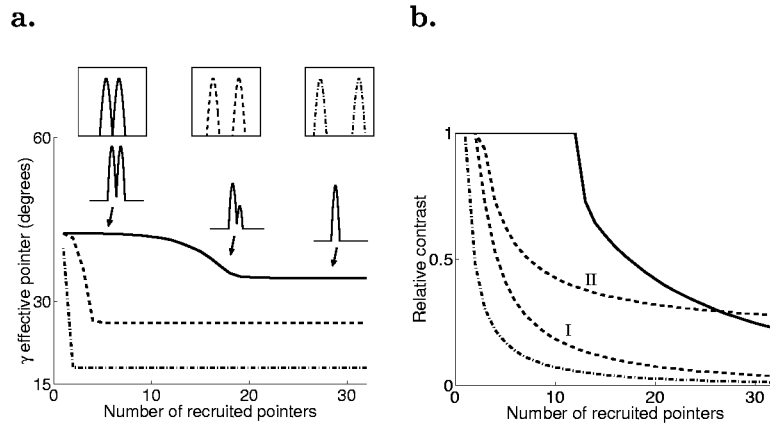


Figure 2: Distance and contrast dependence of attentional selection. (a) The effective pointer angle γ is plotted as a function of the number of recruited pointers. Two identical stimuli are presented at three different interstimulus separations (insets above, 17, 34, and 51 degrees). The initial conditions of pointer neurons $\mathbf{P}^i(0) = (2, 0)$ tend to cause selection of the left stimulus, the right stimulus representing a distractor. The closer the two stimuli are, the more pointers are required for persistent selection. For the first case, where the stimuli are directly adjacent to each other (solid line), three snapshots of steady map activity are shown, corresponding to interpolation (left), partial selection (middle), and persistent selection (right). (b) The minimal relative contrast allowing for persistent selection is shown as a function of the number N_+ of recruited pointers. The curves were determined by slowly decreasing the contrast of the selected stimulus until γ starts to deviate from the center of the selected stimulus. Again, curves are plotted for three stimulus-distractor separations. The solid, dashed, and dash-dotted curves correspond to $\alpha_B = 0.625$. The dashed curve labeled II corresponds to $\alpha_B = 0.6$ —to be compared with the dashed curve labeled I. $\beta = 3.755$, $\beta_I = 60$, $\alpha_F = 0.1$, $\alpha_I = 10$, $I = 32$, $N = 64$, $E = 320$.

5 Winner-Take-All and Attentional Enhancement of Responses

Here we compare inter-areal feedback and winner-take-all (WTA) mechanisms. We show that recruitment has the effect of changing the strength of the WTA mechanism.

A uniform input to the lower area can be viewed as a setting in which each neuron has the same chances of being activated at a steady state (this is due to translational invariance of feedback, $\sin^2 \alpha + \cos^2 \alpha = 1$; see also Hahnloser et al., 1999). The winning neurons (the ones that are activated) are determined by the initial conditions of the dynamics. In Figure 3a, for fixed initial conditions, we see that a localized response to uniform stimulation emerges. The recruitment of many pointers leads to a substantial

narrowing of the activity profile. In Figure 3b, the response width w of the steady response profile is plotted as a function of the number of recruited pointers N_+ . The circles are simulation results, and the full line corresponds to an analytical calculation done in appendix B. It can be seen that w is a monotonically decreasing function of N_+ . This behavior can be interpreted as a WTA mechanism whose softness or hardness is modulated by the number of recruited pointers. The more of them are recruited, the harder is the WTA mechanism.

As an interesting limit to this recruitment-induced strengthening of WTA mechanisms, in appendix C, we calculate the hard WTA limit, in which only one neuron M_x can be active at a steady state. This limit has similarities to a maximum operation that has been suggested to be of relevance for object recognition (Riesenhuber & Poggio, 1999). We find that for a hard WTA, the number N_+^{hard} of pointers that have to be recruited grows quadratically in E . This scaling law suggests that there are not enough neurons to achieve an exact maximum operation by recruitment, but that at best, an approximation to the maximum operation is possible.

As in the previous section, we have analyzed the effect of reducing the strength of excitatory feedback. In Figure 3c, we reduced α_B by 4%. In this case, recruiting pointers does not enhance responses as it did in Figure 3a, but it suppresses responses. In appendix A, we show that the polarity of signal gain depends on the balance between excitatory and inhibitory feedback gain. Signal enhancement occurs if the gain of excitatory feedback set by α_F and α_B is larger than the gain of inhibitory feedback set by α_F , α_I , β_I , and β .

Interestingly, although the signal gains in Figures 3a and 3c are different, the response widths are not. Hence, the hardness of the WTA is insensitive to the exact tuning of synaptic strength, as was the stimulus-separation sensitivity of attentional selection in Figure 2a.

6 Attentionally Controlled Noise Suppression

Psychophysical studies and electrophysiological recordings show that visual attention can enhance the discriminative performance of macaque monkeys (Spitzer, Desimone, & Moran, 1988; Lu & Doshier, 1998) and the discriminative responses of single neurons (Spitzer et al., 1988; McAdams & Maunsell, 1999). Signal discrimination is in many ways equivalent to signal estimation in the presence of external noise.

Population vectors (low-dimensional representations of the activity of many neurons) are possible signal estimators. They can achieve an unbiased readout of sensory input signals if the input noise is uncorrelated between neurons (Seung & Sompolinsky, 1993). In other words, the mean readout of a population vector over many stimulus repetitions is equal to the value of the sensory signal. This is a highly desirable feature of any readout method. However, population vectors do not always have a good performance in averaging out noise. For example, if the neurons supporting the population

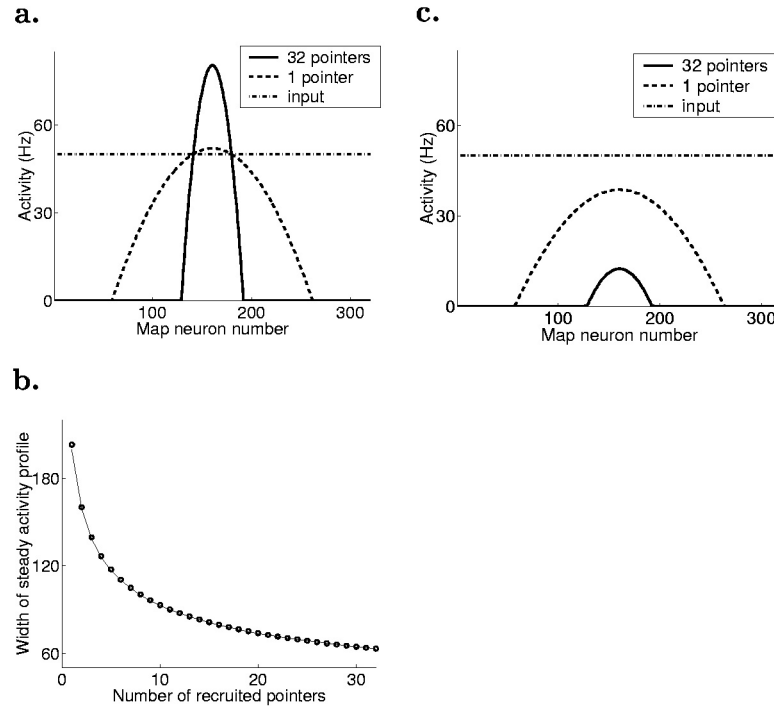


Figure 3: Transition from soft to hard winner-take-all. (a) Steady map response (full and dashed line) to uniform input (dash-dotted line). Attentional recruitment (from 1 to 32 pointers) leads to a sharpened response profile of similar total activity, but with enhanced peak response. The gain of excitatory feedback is approximately equal to that of the inhibitory feedback, $\alpha_B = 0.625$ (see appendix A). (c) The gain of excitatory feedback is smaller than in a, $\alpha_B = 0.6$. Recruitment leads to a strong down modulation of the population response. (b) The response width decays monotonically with the number of recruited pointers (e.g., the strength of the WTA competition increases). For a given number of recruited pointers, the response width is invariant to small changes in α_B (not shown). The circles represent simulations of equations 2.1 to 2.3 and the full line corresponds to a plot of equation B.2. $E = 320$, $I = 32$, $N = 64$. $\beta = 3.755$, $\beta_I = 60$, $\alpha_F = 0.1$, $\alpha_I = 10$.

vector have very narrow tuning curves, then the mean squared error of the readout tends to become very large. In general, any unbiased estimator should have the smallest possible variance, because the variance determines how well two similar stimuli can be discriminated.

Pouget, Zhang, Deneve, and Latham (1998) have examined the advantage of recurrence in the problem of large variability of population vectors.

They found that lateral excitatory connections in cortex can substantially reduce the uncorrelated noise between neurons. In this way, the noisy input to a map of recurrently connected neurons is restored to a steady-state activity that can then be read out by a population vector with near-optimal accuracy. However, the near-optimal readout is achieved only if the stimulus width closely matches the intrinsic tuning width of synaptic connections. Here we show that a much broader range of near-optimality can be achieved when inter-areal feedback loops are recruited according to some prior knowledge of stimulus width.

In our network, the inter-areal feedback combines desirable features of both of the above readout methods. That is, pointers can read out the activity of a map by an unbiased population vector and provide the necessary feedback to cancel uncorrelated noise. Thus, it is possible to have the best of both worlds.

We set the task of the network to extract the location of a stimulus f_x of variable width a , where $f_x(r) = \cos(\frac{\pi}{a}(\delta_x - r))$ if $|\delta_x - r| \leq a/2$ and $f_x = 0$ otherwise (r is the location of the stimulus). We assume that there is prior knowledge available about the stimulus width a . The noise is modeled by adding to f_x some random number η taken from a gaussian distribution with zero mean and fixed variance σ^2 . In this way, the input probability density m_x is given by

$$P(m_x|r) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(m_x - f_x(r))^2}{2\sigma^2}}. \quad (6.1)$$

Figure 4a shows the response of the network to this noisy stimulus under conditions of both few and many recruited pointers. Similar results hold if the noise is Poissonian rather than gaussian.

We have computed the mean and standard deviation of the readout γ for 5000 presentations of the same stimulus at $r = 45^\circ$, but with different noise η . The mean of γ converged toward r , in agreement with unbiased estimation. In Figure 4b, we have plotted the standard deviation $S(\gamma)$ as a function of the number of recruited pointers. For a broad stimulus (dashed line), the standard deviation is minimal when the number of recruited pointers falls between 3 and 5. On the other hand, when the stimulus is narrow (full line), the readout is optimal when about 7 to 14 pointers are recruited. Thus, we find that whichever number of pointers should be recruited depends critically on the stimulus width.

Figure 4c shows the dependence of the standard deviation $S(\gamma)$ on stimulus width, for 1, 4, and 32 recruited pointers. The stimulus location was held fixed at $r = 45^\circ$, and its width a was varied in steps of 4.5 degrees from 4.5 to 81 degrees. The standard deviation was computed using $n = 1000$ presentations of the stimulus for each width.

To get a sense of how large these standard deviations of the pointer estimates are in absolute terms, we have compared them to the minimal

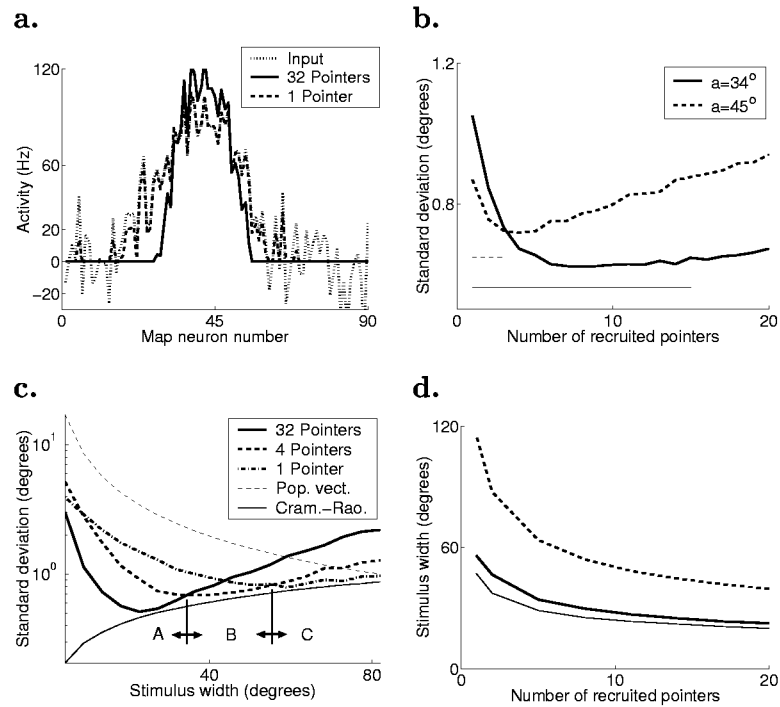


Figure 4: Recruiting feedback for suppressing noise. (a) A noisy stimulus of width $a = 54^\circ$ is indicated by the dotted line. The steady activity in the lower area is shown for the case of 32 recruited pointers (solid line) and for the case of one recruited pointer (dashed line). $\sigma^2 = .04$, $E = 90$. (b) Standard deviations of pointer readout as a function of the number of recruited pointers. A relatively broad stimulus $a = 45^\circ$ results in a minimal standard deviation for about three to five recruited pointers (dashed line). The lower bound as given by equation 6.3 is shown by the fine dashed line. For a slightly narrower stimulus, $a = 34^\circ$, the standard deviation is minimal for 6 to 15 recruited pointers. The fine full line shows the lower bound for this stimulus width. For both stimuli, the optimal pointer estimates deviate by about 10% from the theoretical minimum. $\sigma^2 = .04$. (c) Standard deviations as a function of stimulus width a . Narrow stimuli (in region A) require strong feedback (32 pointers). Broader stimuli (in regions B and C) require weaker feedback (4 and 1 pointer, respectively). The fine dashed line represents the performance of the population vector estimate, equation 6.4. Its performance is similar to the 1-pointer case. $\sigma^2 = .04$. (d) As suggested in b, for every number of recruited pointers, there is a different stimulus width a_{best} for which the standard deviation of readout is minimal (thick full line: $\sigma^2 = .04$ and fine full line: $\sigma^2 = .01$). The dashed line shows the response width w to uniform input. $E = 80$, $N = 40$, $I = 20$. $\alpha_F = 0.4$, $\alpha_B = 0.1$, $\alpha_I = 2.5$, $\beta_I = 24$, $\beta = 0.9656$.

achievable standard deviation $S(\gamma_{opt})$ achievable by any readout method. For a large network ($E \gg 1$), this minimum is given by the Cramér–Rao bound (Cover & Thomas, 1991), defined by the inverse of the square root of the Fisher information:

$$S(\gamma_{opt}) = \frac{1}{\sqrt{\sum_x \langle -\frac{d^2}{dr^2} P(m_x|r) \rangle}}. \quad (6.2)$$

A calculation shows that for large E ,

$$S(\gamma_{opt}) = \sigma \sqrt{\frac{a}{\pi E}}, \quad (6.3)$$

which corresponds to the fine line in Figure 4c. Hence, an optimal readout has an error that is proportional to the square root of the stimulus width and is inversely proportional to the square root of the number of neurons.

It is illustrative to compare the pointer readout to the readout achieved by a population vector defined by $\mathbf{v} = (v_1, v_2) = \sum_x m_x (\cos \delta_x, \sin \delta_x)$. We have calculated the standard deviation for large E , under the approximation that fluctuations are small, in which case only the component $z = 1/\sqrt{2}(v_2 - v_1)$ of the population vector orthogonal to the stimulus direction r matters:

$$S(\gamma_{pop}) \simeq \frac{S(z)}{|\langle \mathbf{v} \rangle|} = \sigma \frac{(\pi^2 - a^2)}{4a \cos(a/2)} \sqrt{\frac{\pi - 2}{2\pi E}}. \quad (6.4)$$

This standard deviation is shown as a fine dashed line in Figure 4c. It tends to diverge for narrow stimuli but approaches the optimal standard deviation as the stimulus width approaches 90 degrees (at this angle, there is a mathematical equivalence between the population vector and the maximum likelihood estimates).

We have found that for any stimulus width, there are a number of recruited pointers N_+ , for which the standard deviation of readout is surprisingly close to the Cramér–Rao bound (see Figure 4c). For small N_+ (dash-dotted line), the readout has a large standard deviation for narrow stimuli and decreases as stimuli become broader. This behavior is similar to that of the population vector and confirms the intuitive notion that when feedback is weak, pointers are nothing but population vectors. As more pointers are recruited (dashed line), the standard deviation is nonmonotonic and has a minimum at a stimulus width of about $a = 45^\circ$. Finally, when many pointers are recruited (full line), then the minimal standard deviation is achieved for even narrower stimuli, at about $a = 25^\circ$.

In Figure 4d, the full line corresponds to the best stimulus width $a_{best}(N_+)$ as a function of the number of recruited pointers (a_{best} is the width at which the standard deviation of the readout is smallest). Again, strong recruitment

is better for narrow stimuli, and weak recruitment is better for broad stimuli. As can be seen, $a_{best}(N_+)$ is not very sensitive to the variance of the noise. Furthermore, its dependence on N_+ is similar to the response width w to a uniform input without noise, shown by the dashed line (see also Figure 3a). w is larger than a_{best} by about 30 degrees but decreases in a similar way. Hence, the strength of feedback (N_+) is optimal when its implicit response width (defined by uniform input) is slightly larger than the stimulus width to be encoded.

To summarize, Figure 4 makes the point that attentional recruitment (based on prior information about stimulus width) can yield a substantial improvement in signal estimation in comparison to locally recurrent networks (corresponding to the case where the number of recruited pointers is fixed). Increased recruitment is needed for small stimuli. Also, increased recruitment is needed for very noisy environments. However, the dependence of optimal recruitment is less sensitive on noise variance than on stimulus size. We expect that a similar improvement also holds for two-dimensional (2D) spatial receptive fields (however, unlike in the one-dimensional case, narrow 2D stimuli have an equal discriminability as broad 2D stimuli; there, the Cramér–Rao bound is independent of stimulus width; Zhang & Sejnowski, 1999).

7 Discussion

Recruitment of neurons and their feedback loops has been studied previously in a different context. In a model of the oculomotor integrator (Seung, Lee, Reis, & Tank, 2000), recruitment was postulated to serve the role of maintaining a precise tuning of feedback amplification, compensating for saturation nonlinearity. Here, recruitment is postulated in the context of inter-areal networks to accentuate multistability. By increasing both the excitatory and inhibitory gain of inter-areal feedback, persistent attentional selection of low-contrast and nearby stimuli becomes possible.

In agreement with experiments, we have found a distractor-dependent limit for the contrast of a stimulus, below which attentional selection is impossible (see Figure 2b). Our results predict that besides the contrast dependence of attentional selection, there should be an additional dependence on stimulus separation (see Figure 2a), for which there is only little experimental evidence so far (De Weerd et al., 1999).

In most electrophysiological experiments, selection of a stimulus enhances visual responsiveness at attended locations (only a few experiments have shown suppressed responses at attended locations; Motter, 1993). The results in Figure 2 suggest that this response enhancement is due to the general dominance of excitatory inter-areal feedback gain over the inhibitory gain. If inter-areal circuits were wired to reduce responses at attended locations, then the selectability of low-contrast stimuli would be impaired in comparison to circuits wired to enhance responses at attended locations.

Thus, in order to be able to attend to low-contrast stimuli, attentional effects should be enhancing rather than suppressing. Support for a dominance of excitatory gain over inhibitory gain comes from anatomical studies (Johnson & Burkhalter, 1997) and a more recent finding, where after cooling of area MT, 33% of the neurons in area V2 showed a significant decrease in response to visual stimulation, whereas only 6% showed an increase (Hupé et al., 1998).

We have shown that attentional recruitment can give cortical processing the ability to adjust signal processing to achieve near-optimal noise reduction for a broad range of stimulus sizes (see Figure 4). This ability generalizes previous results, where recurrent connections have been shown to be near-optimal for only a limited range of stimulus sizes (Deneve, Latham, & Pouget, 1999). However, we point out that the question remains how cortex would be able to recruit the appropriate amount of feedback, given some noisy environment and stimulus to be expected. We do not provide a solution for this problem, but we imagine that the appropriate computation involving the use of some prior information to deduce recruitment level is done in prefrontal cortex.

Our results on noise reduction are comparable to psychophysical studies where attention has been suggested to activate WTA competition between visual filters (Lee, Itti, Koch, & Braun, 1999). Lee et al. have fitted their psychophysical data from orientation discrimination tasks to a model where there is a divisive normalization between simple cortical visual filters. They found best agreement between model and data when the effect of attention was to change the exponents of the divisive normalization between filters rather than any other parameter of the filter interactions. Because the exponents determine the strength of competition between cortical filters, their result is consistent with our finding that attentional recruitment has the effect of hardening WTA competition.

Appendix A: Parameter Selection

In the model equations 2.1 to 2.3, there are five coupling constants: α_F , α_B , β , α_I , and β_I . Here we show how to choose these constants as a function of the feedback gain they result in.

In the following, excitatory gain means the loop gain of the excitatory feedback to the map, via neurons P^i ; and inhibitory gain means the loop gain of the effective inhibitory feedback to the map, via neurons I . The total gain of feedback is the sum of these two loop gains. The relative gain of feedback is important for stability, because in our network, the eigenvectors of the positive and negative feedback loops corresponding to the largest and smallest eigenvalues, respectively, are similar to each other (the excitatory connections are broad, comparable to global inhibitory connections).

First, choose the strength of localizing inhibition β_I according to the number n_I of inhibitory neurons to be active at a steady state (it can be shown

that n_I is independent of the visual input to the map). Assume that the different pointers \mathbf{P}^i get the same attentional input and are thus parallel to each other, forming a single “effective pointer.” Denote the common steady pointer angle by $\gamma = \arctan(P_2^i/P_1^i)$. In this case, the inhibitory neurons I_x receive a feedforward input profile $\alpha_I P \cos(\gamma - \psi_x)$, where $P = \sum_i \|\mathbf{P}^i\|$ is the length of the effective pointer. n_I is determined by the steady state of equation 2.3 (e.g., a symmetric profile centered at γ). The cut-off relation $I_z = 0$ determines the border ψ_z of this profile. In terms of the angular separation between inhibitory neurons, $\gamma - \psi_z = n_I \hat{\psi} / 2$ where $\hat{\psi} = \frac{\pi}{2(I-1)}$, we get

$$\alpha_I P \cos(n_I \hat{\psi} / 2) = \beta_I S, \quad (\text{A.1})$$

where $S = \sum_x I_x$ is the total activity of inhibitory neurons (angles are now measured in radians). In analogy to equation B.2, by integrating the steady state of equation 2.3 over x in the large N (continuum) limit, we can calculate n_I to third-order approximation, using the cut-off relation, equation A.1:

$$n_I = 2 \left(\frac{3}{2\beta_I \psi_2^2} \right)^{1/3}. \quad (\text{A.2})$$

Notice that the width n_I depends on only the strength of recurrent inhibition β_I and not on α_I or the length P of the effective pointer. For an inhibitory map of $I = 32$ neurons, if we choose $\beta_I = 60$ in equation A.2, we obtain $n_I = 4$ active inhibitory neurons at a steady state. If smaller values for β_I are chosen, then the inhibitory activity profile tends to be broader. But a broad inhibitory activity profile is not very desirable for our simulations, since it can result in unwanted boundary effects.

The values of the other parameters are determined by separating the inhibitory from the excitatory feedback loops. First, choose some values for α_F and α_B such that their product is small. In this case, the excitatory feedback loop mediated by a single pointer is weak (having weak pointers means that the increments in feedback strength that arise by recruiting pointers are small and can be precisely controlled). The activities of excitatory neurons in the lower and the higher area are typically of equal magnitude if α_F is smaller than α_B . By choosing an inversely proportional connection strength from pointer onto inhibitory neurons, $\alpha_I = 1/\alpha_F$, the amplification from map onto pointers cancels the reduction from pointers onto inhibitory neurons, just as if the map fed onto inhibitory neurons directly and with unitary weights. Because both excitatory and inhibitory feedback are mediated by pointers, it is convenient to express their gains in terms of the length P of the effective pointer. In this view, the excitatory gain G_E onto the map is

$$G_E = \alpha_B P. \quad (\text{A.3})$$

The inhibitory gain G_I onto the map is $G_I = -\beta S$. Using S from equation A.1 we find

$$G_I = -\frac{\alpha_I \beta}{\beta_I} P \cos(n_I/2\delta_I). \quad (\text{A.4})$$

In our simulation, we have chosen β such that the sum of excitatory and inhibitory gain is equal to zero, $G_E + G_I = 0$, from which we get

$$\beta = \frac{\alpha_F \alpha_B \beta_I}{\cos(n_I \delta_I / 2)}. \quad (\text{A.5})$$

In order to obtain numerical values for β , we replace n_I from equation A.2.

As expected, with this choice of parameters, the gain of cortical amplification is about unity in Figures 2 to 4 and is independent of the number of pointers recruited. Notice that the previous calculation of feedback gain is valid only for large pointer map networks, because it is based on a continuum approximation. Nevertheless, these parameters yield stable attractor dynamics for all inputs. But the network is very close to the limit beyond which stability breaks down. For example, for the parameter settings of Figure 3, decreasing β in equation A.5 by only 2% results in unstable dynamics with unbounded amplification. Hence, although we want the excitatory gain to be at least as large as the inhibitory gain, we are limited in the amount by which they can differ from each other. This result is comparable to a previous calculation, where inhibition was instantaneous and we were able to construct a Lyapunov function if the excitatory gain was not larger than the inhibitory gain by more than $1/E$ (Hahnloser et al., 1999).

The fact that the recurrent inhibition in equation 2.2 is not instantaneous but mediated by separate inhibitory neurons can sometimes lead to oscillations (Li & Dayan, 1999). However, it does not lead to oscillations if the time constant of equation 2.3 is small (paradoxically, the time constant is small if β_I is large—if only a small number of inhibitory neurons are active).

Appendix B: Soft Winner-Take-All

As an approximation to the limit, in which the number of neurons E in the lower area becomes infinite, the steady state of equations 2.2 and 2.3 can be transformed into the second-order differential equation $M_x'' = -M_x + c$ (where $'$ denotes derivation with respect to x). For uniform input, the solution is a cosine-shaped profile that can be centered at an (almost) arbitrary angle χ : $M_x = H[\cos(\chi - \delta_x)]_+ + c$, where $\chi - w/2 \leq \delta_x \leq \chi + w/2$. The amplitude H , the width w , and the offset c of the profile are unknowns that can be inferred from $M_{\chi \pm w/2} = 0$, $M(\chi) = H$ and $\int_{\chi-w/2}^{\chi+w/2} M_x dx = 2H \sin(w/2) + cw$. This leads to

$$w - \sin w = \frac{\pi}{N_+ \alpha_F \alpha_B (E - 1)}, \quad (\text{B.1})$$

where N_+ is the number of recruited pointers. To third-order approximation in w , we find

$$w \simeq \left(\frac{6\pi}{N_+ \alpha_F \alpha_B (E - 1)} \right)^{1/3}. \quad (\text{B.2})$$

The excellent match of equation B.2 with the simulation results can be seen in Figure 3b. Surprisingly, based on equation B.1, the width w does not depend on the strength of inhibition given by β and β_I , nor does it depend on the number I of inhibitory neurons. In fact, the width depends on only the strengths of excitation α_F and α_B . Increasing the strength of inhibition does not change the width of the profile, only its amplitude H .

The sharpening of activity with increasing number of recruited pointers can be understood in terms of the mathematical principle of forbidden sets (Hahnloser, Sarpeshkar, Mahowald, Douglas, & Seung, 2000). This principle was derived for symmetric linear threshold networks. The principle says that some sets of map neurons cannot be simultaneously active at a steady state, because their connectivity expresses “forbidden” or unstable differential modes (differential modes are eigenvectors with both negative and positive components). Because all unstable modes are differential, at least one map neuron will eventually fall below the rectification nonlinearity, and so the largest eigenvalue of the feedback will decrease. In this way, map neurons are progressively inactivated by the network dynamics. The process halts when the largest eigenvalue becomes smaller than one, where stability is achieved and a stable activity pattern can be formed (the set of active map neurons becomes “permitted”). By recognizing that the number N_+ of recruited pointers has a multiplicative influence on eigenvalues of inter-areal feedback (recruiting twice as many pointers results in a doubling of feedback gain), we arrive at a simple understanding of what causes the WTA mechanism. The more pointers are recruited, the more map neurons have to be inactivated by the network dynamics in order for the active neurons to form a permitted set.

Appendix C: Hard Winner-Take-All

If the number of attentionally recruited neurons in the higher area increases progressively and all of these neurons participate in similar feedback loops with neurons in the lower area, then at a certain point, the feedback will be so strong that only a single neuron in the lower area can be active at a steady state. Under these conditions, the network implements a hard winner-take-all mechanism. Here we calculate exactly how many pointers are required to achieve this regime. Assume that the parameters are selected according to appendix A. We study the steady states of equations 2.1 to 2.3, to establish that no neuron other than neuron M_s can be active at a steady state (the choice of s is arbitrary). In other words, denoting steady states by underlin-

ing, we require that $\underline{M}_{s'} = 0$, where $s' \neq s$ holds true for all stationary inputs m_x .

We proceed by assuming that there are N_+ recruited pointers and that neurons are at a steady state in which only a single neuron M_s is active in the lower area. In this case, the length of the effective pointer is $P = N_+ \alpha_I \underline{M}_s$. Using this expression with the relationship between parameters as given in appendix A, in particular equation A.1, we find that

$$\underline{M}_s = m_s. \quad (\text{C.1})$$

There is an amplification gain of exactly one (this intermediate result is surprisingly consistent with the continuum assumption made in appendix A).

In order for M_s to be the only activated neuron, the neurons $M_{s\pm 1}$ adjacent to M_s should not be activated, even in the most extreme case where their input is equally large, $m_{s+1} = m_s$ (notice that if m_{s+1} were larger than m_s , then we might as well assume that M_{s+1} is the single active neuron in steady state, which leads us back to the same argument). Thus, neurons that are not nearest neighbors of M_s do not need to be considered; they have a smaller probability of being activated than nearest neighbors, because the excitatory feedback loops via pointers decay with distance in the lower area (Hahnloser et al., 1999).

A simple calculation of the steady-state \underline{M}_{s+1} yields

$$\underline{M}_{s+1} = m_{s+1} + N_+ \alpha_F \alpha_B \underline{M}_s (\cos(\hat{\delta}) - 1), \quad (\text{C.2})$$

where $\hat{\delta} = \frac{\pi}{2(E-1)}$. We determine the number of recruited pointers N_+^{hard} beyond which the WTA is hard by using the constraint $\underline{M}_{s+1} = 0$. We find that for

$$N_+^{hard} \geq \frac{1}{\alpha_F \alpha_B (1 - \cos(\hat{\delta}))} \simeq \frac{(E-1)^2}{\alpha_F \alpha_B}, \quad (\text{C.3})$$

no neuron other than M_s is active at a steady state. We see that the number of neurons N_+^{hard} in the higher area to be recruited increases quadratically with the number of neurons in the lower area, which suggests that this hard WTA limit is an interesting computational limit rather than a possible tool that can be used by inter-areal circuits as studied here.

Acknowledgments

We acknowledge comments on the manuscript by Martin Giese and the support of the Swiss National Science Foundation and the Körber Foundation.

References

- Colby, C., Duhamel, J.-R., & Goldberg, M. (1996). Visual, presaccadic, and cognitive activation of single neurons in monkey lateral intraparietal area. *J. Neurophysiol.*, *76*(5), 2841–2852.
- Cover, T., & Thomas, A. (1991). *Information theory*. New York: Wiley.
- Deneve, S., Latham, P., & Pouget, A. (1999). Reading population codes: A neural implementation of ideal observers. *Nature Neuroscience*, *2*(8), 740–745.
- Desimone, R. (1996). Neural mechanisms for visual memory and their role in attention. *Proc. Natl. Acad. Sci. USA*, *93*(24), 13494–13499.
- Desimone, R. (1998). Visual attention mediated by biased competition in extrastriate visual cortex. *Philosophical Transactions of the Royal Society (London), B Biological Sciences*, *353*, 1245–1255.
- De Weerd, P., Peralta, M., Desimone, R., & Ungerleider, L. (1999). Loss of attentional stimulus selection after extrastriate cortical lesions in macaques. *Nature Neuroscience*, *2*(8), 753–758.
- Fuster, J. (1990). Inferotemporal units in selective visual attention and short-term memory. *J. Neurophysiol.*, *64*(3), 681–697.
- Hahnloser, R. H., Douglas, R. J., Mahowald, M., & Hepp, K. (1999). Feedback interactions between neuronal pointers and maps for attentional processing. *Nature Neuroscience*, *2*(8), 746–752.
- Hahnloser, R. H., Sarpeshkar, R., Mahowald, M., Douglas, R. J., & Seung, S. (2000). Digital selection and analog amplification coexist in a silicon circuit inspired by cortex. *Nature*, *405*, 947–951.
- Hubel, D., & Wiesel, T. (1962). Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. *J. Physiol.*, *160*, 106–154.
- Hupé, J., James, A., Payne, B., Lomber, S., Girard, P., & Bullier, J. (1998). Cortical feedback improves discrimination between figure and background by V1, V2 and V3. *Nature*, *394*, 784–787.
- Johnson, R., & Burkhalter, A. (1997). A polysynaptic feedback circuit in rat visual cortex. *J. Neurosci.*, *17*(18), 7129–7140.
- Lee, D., Itti, L., Koch, C., & Braun, J. (1999). Attention activates winner-take-all competition among visual filters. *Nature Neuroscience*, *2*(4), 375–381.
- Li, Z., & Dayan, P. (1999). Computational differences between asymmetrical and symmetrical networks. *Network*, *10*, 59–77.
- Logothetis, N., Pauls, J., Bulthoff, H., & Poggio, T. (1994). View dependent object recognition by monkeys. *Curr. Biol.*, *4*(5), 401–414.
- Lu, Z.-L., & Doshier, B. (1998). External noise distinguishes attention mechanisms. *Vision Research*, *38*(9), 1183–1198.
- Luck, S. J., Chelazzi, L., Hillyard, S., & Desimone, R. (1997). Neural mechanisms of spatial selective attention in areas V1, V2, and V4 of macaque visual cortex. *J. Neurophysiol.*, *77*(1), 24–42.
- McAdams, C. J., & Maunsell, J. H. (1999). Effects of attention on the reliability of individual neurons in monkey visual cortex. *Neuron*, *23*, 765–773.
- Moran, J., & Desimone, R. (1985). Selective attention gates visual processing in the extrastriate cortex. *Science*, *229*, 782–784.

- Motter, B. (1993). Focal attention produces spatially selective processing in visual cortical areas V1, V2, and V4 in the presence of competing stimuli. *J. Neurophysiol.*, *70*(3), 909–919.
- Motter, B. (1994). Neural correlates of feature selective memory and pop-out in extrastriate area V4. *J. Neurosci.*, *14*(4), 2190–2199.
- Pouget, A., Zhang, K., Deneve, S., & Latham, P. (1998). Statistically efficient estimation using population coding. *Neural Computation*, *10*, 373–401.
- Reynolds, J., Pasternak, T., & Desimone, R. (2000). Attention increases sensitivity of V4 neurons. *Neuron*, *26*, 703–714.
- Reynolds, J. H., Chelazzi, L., & Desimone, R. (1999). Competitive mechanisms subserve attention in macaque areas V2 and V4. *Journal of Neuroscience*, *19*(5), 1736–1753.
- Riesenhuber, M., & Poggio, T. (1999). Hierarchical models of object recognition in cortex. *Nature Neuroscience*, *2*, 1019–1025.
- Seung, H. S., Lee, D. D., Reis, B. Y., & Tank, D. D. (2000). Stability of the memory of eye position in a recurrent network of conductance-based model neurons. *Neuron*, *26*, 259–271.
- Seung, H., & Sompolinsky, H. (1993). Simple models for reading neuronal population codes. *Proc. Natl. Acad. Sci. USA*, *90*, 10749–10753.
- Spitzer, H., Desimone, R., & Moran, J. (1988). Increased attention enhances both behavioral and neuronal performance. *Science*, *240*, 338–340.
- Tomita, H., Ohbayashi, M., Nakahara, K., & Miyashita, Y. (1999). Top-down signal from prefrontal cortex in executive control of memory retrieval. *Nature*, *401*, 699–703.
- Treue, S., & Maunsell, J. (1996). Attentional modulation of visual motion processing in cortical areas MT and MST. *Nature*, *382*, 539–541.
- Treue, S., & Maunsell, J. (1999). Effects of attention on the processing of motion in macaque middle temporal and medial superior temporal visual areas. *Journal of Neuroscience*, *19*(17), 7591–7602.
- Wersing, H., Beyn, W.-J., & Ritter, H. (2001). Dynamical stability conditions for recurrent neural networks with unsaturating piecewise linear transfer functions. *Neural Computation*, *13*, 1811–1825.
- Zhang, K., & Sejnowski, T. (1999). Neuronal tuning: to sharpen or broaden? *Neural Computation*, *11*, 75–84.