Available online at www.sciencedirect.com
ScienceDirect
Neural Networks 19 (2006) 1347–1370
Neural Networks
www.elsevier.com/locate/neunet

2006 Special Issue

The micro-structure of attention

Neill R. Taylor*, Matthew Hartley, John G. Taylor

Department of Mathematics, King's College London, Strand, London, United Kingdom

Received 16 June 2006; accepted 1 August 2006

Abstract

We investigate three possible methods of specifying the microstructure of attention feedback: contrast gain, additive and output gain, using simple single node and 3-layer cortical models composed of graded or spiking neurons. Contrast gain and additive attention are also tested in a spiking network which is simplified by mean field methods. The simulation task uses two stimuli, probe and reference, presented singly or together within the neuronal receptive fields whilst attention is directed towards or away from the receptive field. Model neurons are differentially activated in the different stimuli and attention and equilibrium potentials or average firing rates recorded depending on neuron type are recorded. We compare results for the different modes of attention and architectures with experimental single cell recordings which show how neuronal firing rates change in response to attention, with a bias towards neurons that respond more effectively to the attended stimulus, to investigate which attentional method best fits the experimental data. The simulation results are also mathematically analysed. We conclude that there is most experimental support for contrast gain, although some additional feedback gain would be possible. We propose a tentative method by which attention as contrast gain may occur in the primate brain using acetylcholine and nicotinic receptors.

Keywords: Visual attention; Contrast gain; Output gain; Additive attention; Regression lines; Acetylcholine; Nicotinic receptors

1. Introduction

Attention is a selective filter by which neural processing is limited to important information that is pertinent to current behaviour, by controlling which information passes to higher cortical levels. There has been considerable interest in general features of attention since the time of the ancient Greeks, but with the advent of brain imaging techniques and advanced single cell recordings the nature of attention has been increasingly better analysed. In this paper we wish in particular to consider how to understand various experimental results arising from single cell measurements in behaving monkeys so as to explore the more detailed micro-structure of attention, and thereby to help in building more precise models based on realistic neural activity and its modulation.

Visual attention has been shown to increase the output responses of neurons for single inputs, whilst for multiple inputs the neuronal responses are dependent on the neuron's sensitivity to the individual stimuli, as arises in the well known 'biased competition' model of Desimone and Duncan (1995); recent results (Williford & Maunsell, 2006) are inconclusive as to whether the response change due to attention is non-proportional or proportional. However the manner in which the output response of neurons is increased by attention is still controversial. Our purpose in this paper is to attempt to clarify some aspects of this controversy by means of suitable experimental data and its simulation.

There are essentially three different methods by which attention can be postulated to act: by additive feedback, by contrast gain and by output gain. The first of these needs little further explanation, and has been used by a variety of modellers to describe how attention acts at the micro-level (Deco & Rolls, 2005; Grossberg & Raizada, 2000). In contrast gain the inputs from an attended stimulus to a given cell under attention control are all boosted by some multiplicative constant, arising possibly from some higher-level feedback signal. In output gain the whole output of an attended cell, from whatever inputs, is boosted by the multiplicative feedback attention signal. Both have been investigated in particular psychological paradigms (Carrasco, Ling, & Read, 2004), as well as considerable modelling of a range of psychological

* Corresponding address: Department of Mathematics, King's College London, Strand, London WC2R 2LS, United Kingdom. Tel.: +44 20 7848 1026; fax: +44 20 7848 2017.
E-mail address: neill.taylor@kcl.ac.uk (N.R. Taylor).

0893-6080/$ - see front matter © 2006 Elsevier Ltd. All rights reserved.
doi:10.1016/j.neunet.2006.08.002

Nomenclature

SE	selectivity
SI	sensory interactivity
P_i	normalised average firing rate/potential of node i due to probe stimulus
R_i	normalised average firing rate/potential of node i due to reference stimulus
Pair_i	normalised average firing rate/potential of node i due to probe and reference stimulus
$V(t)$	neuron potential
A	decay
I	total inhibitory input
B	maximum activity
C	shunting inhibition value
E	total excitatory input
τ	time constant
V_s	shunting inhibition
V_m	maximum potential
V_{leak}	the resting potential
I_{excit}	is the total excitatory input
I_{GABA}	is the total inhibitory input
S_i	is the NMDA gating variable for pool I
τ_s	is the NMDA time constant
r_i	is the firing rate for pool I
Φ	input–output function
$I_{\text{syn},i}$	total synaptic input to pool i
$J_{N,ij}$	represent the effective synaptic coupling from pool j to pool i mediated by NMDA
$J_{A,ij}$	represent the effective synaptic coupling from pool j to pool i mediated by AMPA
I_0	mean external input current
I_1	input current to pool 1
I_2	input current to pool 2
$I_{\text{noise},i}$	noise to pool I
τ_{AMPA}	AMPA time constant
η	is Gaussian white noise
σ^2_{noise}	is the variance of the noise
μ_i	input firing rate to pool I
M^{BC}	modulation index
x_j	output of neuron j
w_j^+ and w_j^-	are the excitatory and inhibitory weights
w'_j	is the strength of the connection from the kth higher-order attention control neuron to the ith node
w_j	is the connection from the jth lower-order node to the ith node
$w_{i,j,k}$	sigma–pi weight
attn	attention feedback signal
$V1IN$	input IN from any $V1$ feedback input
$F(IN, \text{PY56})$	the feedback from layer 2/3 to PY56

Abbreviations

PYR5/6	pyramidal layer 5/6 node
IN4	inhibitory layer 4 cell
ST4	excitatory stellate layer 4 node
PYR2/3	pyramidal layer 2/3 cell
IN5/6	layer 5/6 inhibitory node
ACh	acetylcholine
nAChR	nicotinic ACh receptors
NBM	nucleus basalis of Meynert

paradigms also being done using contrast gain (Fragopanagos, Kockelkoren, & Taylor, 2005; Mozer & Sitton, 1998; Taylor & Rogers, 2002).

Here we investigate which attentional mode best fits experimental results obtained by Reynolds, Chelazzi, and Desimone (1999) by investigating simulations based on two types of neurons: graded and spiking within two different models of cortical structure: single nodes and 3-layer cortical models, as well as mean field equations. These results were originally modelled in (Reynolds et al., 1999) using single neurons with a set of inputs from attended and unattended stimuli using contrast gain. The particular results involved regression curves obtained from careful analysis of an ensemble of hundreds of single cells in monkey $V2$ or $V4$, with considerable success. This analysis was extended in (Taylor & Rogers, 2002) to provide a more mathematical underpinning, including a universal approximation formula.

We start the paper with a description of the experimental paradigm we will study by simulation.

2. The Reynolds et al. results

Reynolds et al. (1999) investigated two assumptions of 'biased competition' (1) that there is automatic competition between populations of neurons activated by multiple stimuli, and (2) that under attention the competition is biased to neurons that respond to the attended stimulus. We are particularly interested in the second experiment reported in Reynolds et al. (1999) which recorded neurons in $V2$ and $V4$. For this task two stimuli were used, termed: probe and reference. Stimuli could appear at 4 locations, 2 in the receptive field and 2 outside. In the attend away condition the monkey attended to a location outside the receptive field whilst in the receptive field the probe stimulus, or reference stimulus or both stimuli could appear. In the case where attention was directed towards the receptive field both probe and reference stimuli were presented within the receptive field and attention was directed to either the probe stimulus or reference stimulus. These conditions with attention directed towards or away from the receptive field leads to the following 5 different combinations of stimuli:

(1) probe input alone, attention directed away;
(2) reference stimulus alone, attention directed away;
(3) pair (probe and reference inputs), attention directed away from receptive field;
(4) pair with attention directed to probe stimulus;
(5) pair with attention directed to reference stimulus.

For each input condition each neuron was recorded over 10 trials, and the firing rates averaged. For each neuron the firing

rates were normalised across the 5 input conditions using the largest average firing rate. These values are then used to calculate 2 measures: selectivity (SE) and sensory interactivity (SI) for each neuron. These are defined by:

$$SE_i = P_i - R_i \quad (1)$$

$$SI_i = \text{Pair}_i - R_i \quad (2)$$

where P_i indicates the normalised average firing rate for probe stimulus of neuron i, R_i is the same value for the reference stimulus of neuron i, and Pair_i is the normalised average firing rate for probe and reference inputs, which can be calculated for attend away, attend probe and attend reference. The t test is used with $p < 0.05$ to detect for which neurons the activity has changed significantly between the attend away and attend probe or attend reference cases. Graphs of SE versus SI are then plotted for:

(1) attend away (probe) where these are the points for which activity changes significantly between attend away and attend probe;
(2) attend probe;
(3) attend away (reference) where these are the points for which activity changes significantly between attend away and attend reference;
(4) attend reference.

Regression analysis was performed to find the best fit for each graph for the line:

$$SI = \text{slope} SE + \text{constant}.$$

For $V2$ and $V4$ neurons the results were comparable with the slope being near 0.5 for the different attend away plots, whilst the attend probe slope was greater than the attend away slopes, and the attend reference slopes being smaller than the attend away slopes. The regression constant was near 0 for all $V2$ graphs, whilst there was an increase in $V4$ for attend probe and attend reference as against their respective constant values for attend away, up to 0.15. The slopes were the same for cells from the two modules. These results confirm the second assumption about 'biased competition': that under attention there is a bias of the competition towards neurons that respond more effectively to the attended stimulus.

In a simulation using 100 simple graded neurons, Reynolds et al. (1999) showed that similar regression slopes could be obtained by multiplying the connection strengths from probe or reference inputs to the neurons by a factor of 5 in the attend probe or attend reference conditions.

3. Method and architecture

We perform similar simulations as in Reynolds et al. (1999), but now extended to the range of graded and spiking neurons for both single neurons and a 3-layer cortical model, where attention feedback is also modelled as 3 forms, additive, output gain and contrast gain, and for a system of mean field equations where attention is modelled as contrast gain and additive. Fig. 1 indicates the method by which the attentional modes

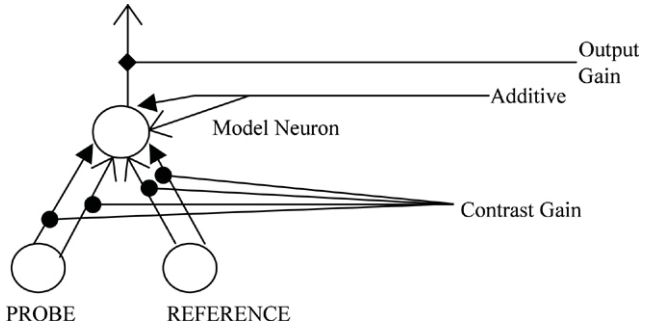


Fig. 1. The different forms of attention. The output neuron is only necessary for the output gain simulation. Closed arrow-heads indicate inhibitory connections and open arrow-heads indicate excitatory weights. Circular and diamond heads indicate multiplicative affects.

are implemented and the organisation of reference and probe inputs.

As in Reynolds et al. (1999) the probe and reference inputs have excitatory and inhibitory efferents to the model neurons. For contrast gain, attention is implemented by multiplying the excitatory and inhibitory connections by a factor >1. In the case of output gain it is the output of the model neurons that is multiplied by this factor. It can be seen that if the outputs of the model neurons feedforward to a higher order region that output gain at the current level would be the same as contrast gain implementation at the higher level.

For both the graded and spiking cases for single neurons the architecture of Fig. 1 is used; for the 3-layer cortex graded case we use the architecture of Grossberg and Raizada (2000) simplified to remove the lateral connectivities between loop structures which form the on-centre, off-surround structure of their simulations. A column or loop is composed of a single pyramidal layer 5/6 node (PYR5/6), an inhibitory layer 4 cell (IN4), an excitatory stellate layer 4 node (ST4) and a pyramidal layer 2/3 cell (PYR2/3). It is known that inputs to cortex enter at layer 4 and layer 6, hence, for the graded model, probe and reference inputs (excitatory and inhibitory) are incident on PYR 5/6 and ST4. In the spiking model we add a layer 5/6 inhibitory node (IN5/6), which means we can remove the inhibitory inputs for the probe and reference to the layer 4 and 5/6 nodes. In graded and spiking models attention is implemented at layer 5/6, for contrast gain this involves the multiplication of the weights from probe and reference to the layer 5/6 nodes, for additive attention inputs are to the layer 5/6 nodes, and for output gain it is the outputs of the layer 5/6 nodes that are multiplied, hence its outputs to the layer 4 nodes are stronger in the attend probe or attend reference case. This is shown in Fig. 2.

3.1. Single node graded case

The single neuron case re-creates the simulation results (Reynolds et al., 1999). We use 100 model neurons with a passive decay of 0.2 and an attentional factor of 3, with noise of ±10%. The activity of a model neuron is described by:

$$\frac{dV}{dt} = -AV - VI + (B - V)E \quad (3)$$

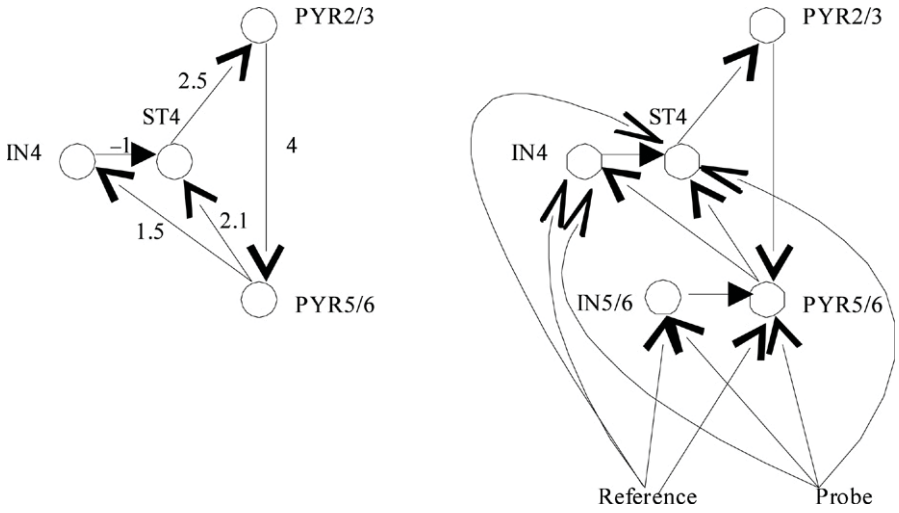


Fig. 2. Architectures for the 3-layer cortex. Left used for graded model, right used for the spiking model.

where, as in Reynolds et al. (1999) $-AV$ is a passive decay term with $A = 0.2$, I is the total inhibitory input (positive value) and the second term on the right-hand side indicates that the neuron has a shunting inhibition value of 0, E is the total excitatory input, and B is the maximum activity (set to 1).

The equilibrium response of this neuron is:

$$\lim_{t\to\infty} V = \frac{BE}{E + I + A}. \tag{4}$$

The connection strengths are randomly set with magnitudes in the range {0, 1}. For all 5 stimulus inputs the stable activation level for the model neurons are found from Eq. (4), and a noise term is added of ±10% from an uniform distribution to the response of the node.

3.2. Three-layer cortex graded case

We use the neuronal definitions from Grossberg and Raizada (2000), where the potential is defined as:

$$\tau \frac{\mathrm{d}V}{\mathrm{d}t} = -V - (V_i - V)I + (V - V_m)E \tag{5}$$

where τ is the time constant, V_s is the shunting inhibition value set to -1, V_m is the maximum activity level set to 1, and E and I are as before the total synaptic excitatory and inhibitory input, respectively, where the output of PYR2/3 nodes are defined as:

$$\text{Out (PYR2/3)} = \max\{\text{activity (PYR2/3)} - 0.2, 0\} \tag{6}$$

whilst all other output functions are:

$$\text{Out}\,(x) = \max\{\text{activity}\,(x), 0\}. \tag{7}$$

The time constants are 2.5 for ST4 and PYR5/6 nodes, 1.25 for PYR2/3 nodes and 0.1875 for the IN4 nodes. Noise is again randomly added. We record the outputs of the PYR5/6 nodes which are given by Eq. (7), and then perform the same analysis as Reynolds et al. (1999).

3.3. Spiking models

Neuronal potential is defined by the equation:

$$\tau \frac{\mathrm{d}V}{\mathrm{d}t} = (g_{\text{leak}}(V - V_{\text{leak}}) + (V_m - V)I_{\text{excit}} + (V - V_{\text{shunt}})I_{\text{GABA}}) + \text{noise} \tag{8}$$

where $V(t)$ is the potential, V_{leak} is the resting potential -70 mV, $g_{\text{leak}} = -2.5 * 10^{-8}$, I_{excit} is the total excitatory input, I_{GABA} is the total inhibitory input, V_{shunt} is the shunting potential of -80 mV, and V_m is set to 0 mV. The spiking threshold is set to -52 mV.

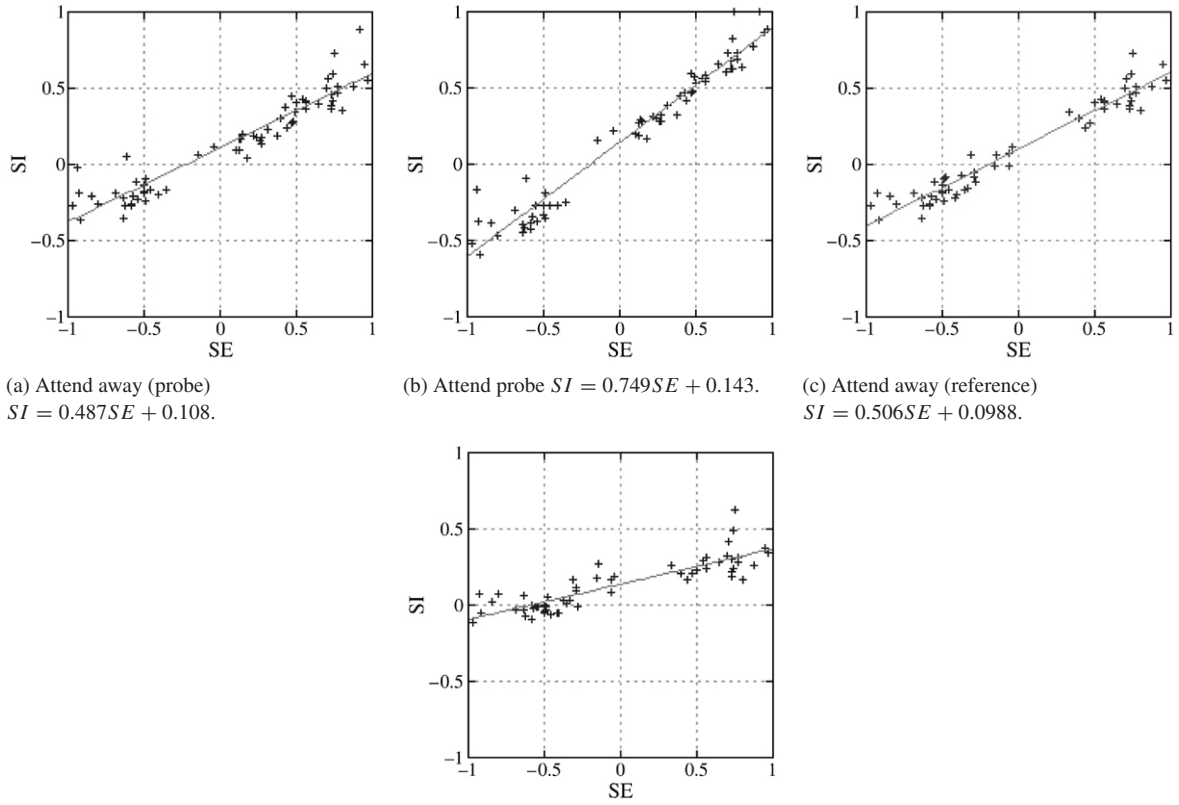
Noise is added to the membrane potential of the neurons which leads to spontaneous firing rates for the single node case in the range 0–11 Hz, with an average firing rate of 5.5 Hz; for the 3-layer cortical model noise causes a spontaneous rate of ∼3 Hz for excitatories and 9 Hz for inhibitories. The probe and reference inputs are each represented by a population of 100 spiking neurons.

3.4. Mean field reduction of the spiking neuron equations

We investigate a mean field reduction of the spiking neuron equations using the analysis of Wong and Wang (2006) which is based on the earlier structure of Brunel and Wang (2001). For the current simulations the architecture is composed of 2 selective pools of excitatory neurons—one preferentially activated by the probe stimulus and the other preferentially activated by the reference stimulus—a pool of non-selective excitatory neurons, and a pool of non-selective inhibitory nodes. The analysis performed by Wong and Wang (2006) reduces this structure to 2 selective pools. The mean field equations are then:

$$\frac{\mathrm{d}S_1}{\mathrm{d}t} = -\frac{S_1}{\tau_s} + (1 - S_1)\gamma r_1 \tag{9}$$

$$\frac{\mathrm{d}S_2}{\mathrm{d}t} = -\frac{S_2}{\tau_s} + (1 - S_1)\gamma r_2 \tag{10}$$



(a) Attend away (probe) $SI = 0.487SE + 0.108$.

(b) Attend probe $SI = 0.749SE + 0.143$.

(c) Attend away (reference) $SI = 0.506SE + 0.0988$.

(d) Attend reference $SI = 0.232SE + 0.137$.

Fig. 3. Contrast gain in single graded neuron model, as in Reynolds et al. (1999). Attentional modulation is *3.

where S_i is the NMDA gating variable for pool i, τ_s is the time constant set to 100 ms, $\gamma = 0.641$, and r_i is the firing rate for pool i. The firing rates are given by:

$$r_1 = \Phi(I_{\text{syn},1}) \tag{11}$$

$$r_2 = \Phi(I_{\text{syn},2}) \tag{12}$$

$$\begin{aligned} I_{\text{syn},1} &= J_{N,11}S_1 - J_{N,12}S_2 + J_{A,11}r_1 - J_{A,12}r_2 \\ &\quad + I_0 + w_{11}I_1 + w_{12}I_2 + I_{\text{noise},1} \end{aligned} \tag{13}$$

$$\begin{aligned} I_{\text{syn},2} &= J_{N,22}S_2 - J_{N,21}S_1 + J_{A,22}r_2 - J_{A,21}r_1 \\ &\quad + I_0 + w_{22}I_2 + w_{21}I_1 + I_{\text{noise},2} \end{aligned} \tag{14}$$

where $I_{\text{syn},i}$ is the total synaptic input to pool i, and $J_{N,ij}$ represent the effective synaptic coupling from pool j to pool i mediated by NMDA, similarly $J_{A,ij}$ represent the effective synaptic coupling mediated by AMPA. The w_{ij} represent the connection strength from stimulus j to pool i. The values used are $J_{N,11} = J_{N,22} = 0.1561$ nA, $J_{N,12} = J_{N,21} = 0.0264$ nA, $J_{A,11} = J_{A,22} = 9.9026 * 10^{-4}$ nA Hz^{-1}, and $J_{A,21} = J_{A,12} = 6.5177 * 10^{-5}$ nA Hz^{-1}. $I_0 = 0.2346$ nA and is the mean effective external input, which along with a specific noise term ($I_{\text{noise},i}$) produces a spontaneous firing rate ~3 Hz. The noise term is governed by:

$$\tau_{\text{AMPA}} \frac{\mathrm{d}I_{\text{noise}}}{\mathrm{d}t} = -I_{\text{noise}} + \eta\sqrt{\sigma^2_{\text{noise}}\tau_{\text{AMPA}}} \tag{15}$$

with $\tau_{\text{AMPA}} = 2$ ms, η is Gaussian white noise, and $\sigma_{\text{noise}} = 0.007$ nA, and σ^2_{noise} is the variance of the noise. The input–output function Φ is also from Wong and Wang (2006) adapted from Abbott and Chance (2005):

$$\Phi(I_{\text{syn},i}) = \frac{270I_{\text{syn},i} - 108}{1 - \exp(-0.1540(270I_{\text{syn},i} - 108))}. \tag{16}$$

The two other external inputs (I_1 and I_2) are defined by:

$$I_i = 0.2243 * 10^{-3}\mu_i \tag{17}$$

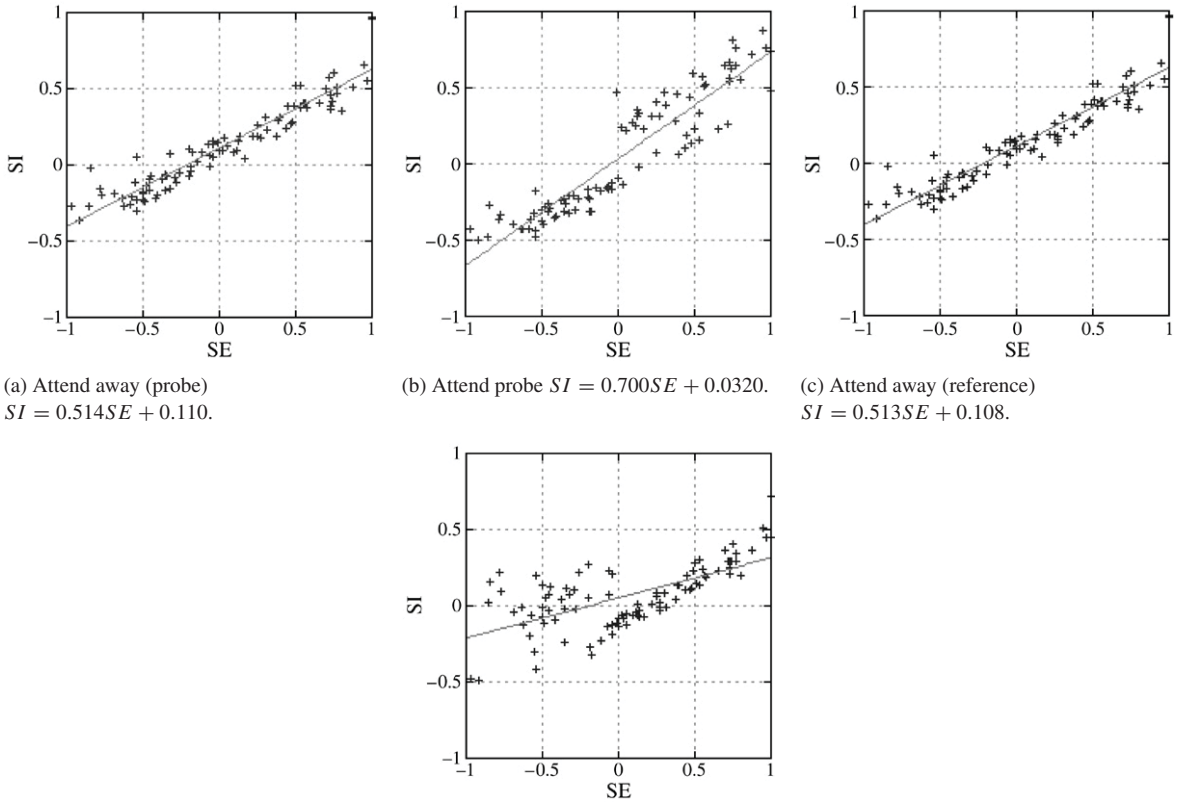
where μ_i is a firing rate and for the simulations performed below takes the value 250 Hz for both stimuli. When testing contrast gain we multiply the appropriate set of input weights w_{ij} by some factor that is larger than 1. In the additive case there is an extra additive input term to the synaptic currents Eqs. (13) and (14); this attention value also uses Eq. (17) where μ here is determined experimentally.

4. Results

4.1. Graded single-layer cortex

The connections from probe and reference to the model neurons are randomly set in the range 0 to 1 from an uniform distribution.

We show the results for contrast gain attention in Fig. 3, where the modulatory factor is 3. These results are similar to that of Reynolds et al. (1999) with the slopes being: 0.749, 0.487, 0.232 and 0.506 for the attend probe, attend away



(a) Attend away (probe) $SI = 0.514SE + 0.110$.

(b) Attend probe $SI = 0.700SE + 0.0320$.

(c) Attend away (reference) $SI = 0.513SE + 0.108$.

(d) Attend reference $SI = 0.263SE + 0.0500$.

Fig. 4. Additive attention in single graded neuron model, with attention as ±0.6 depending on neurons' preferences.

(probe), attend reference and attend away (reference) cases, respectively, albeit for a smaller attentional value. Similar slopes are found if regression lines are calculated for all neurons, i.e. when the t test is not used.

For additive and output gain attention we class the model neurons as probe preferred or reference preferred, where the former has a total of probe weights greater than the total of reference weights, and vice versa for the latter.

Then for additive attention in the attend probe case only the probe preferred model neurons have a positive additive term inputted to the neuron whilst all other model neurons had a negative additive term input, similarly for the attend reference condition only neurons that are reference preferred were affected by positive additive attention, the other neurons all had a negative additive term applied. Using an additive attentional value of 0.2 the results are shown in Fig. 4.

In the case of attention as output gain, we modulate the output by 2 factors $1 + x$ if attending to input that is preferred by neuron, and $\frac{1}{1+x}$ if another input is preferred by the neuron, where $x > 0$. Using a value of 0.3 the results shown in Fig. 5 are found.

Whilst in all cases values for the attentional bias can be found such that the regression slopes match the experimental results (Table 1), it is the results for contrast gain that have the closest match to the experimental results. For output gain and additive attention there are clear neuronal groupings forming in the attend probe and attend reference plots; such splits are not seen in the experimental results (Reynolds et al., 1999).

The neuronal potential equation in Eq. (3) uses a linear output function for probe and reference as well as the model neurons; replacing with an output function that is sigmoidal generates similar distribution of points in the plots and attentional values can be found that give the required regression slopes within experimental ranges.

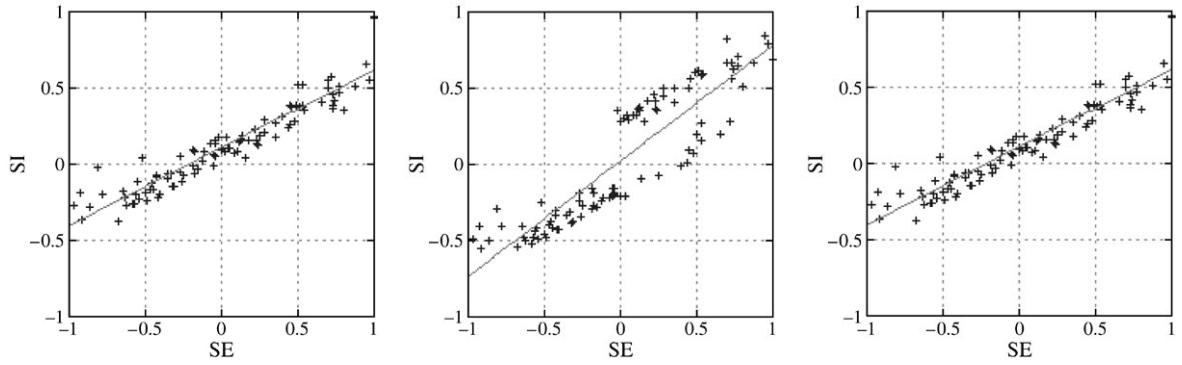
4.2. Graded 3-layer cortex

For the 3-layer cortical model the probe and reference efferents target PYR5/6 and ST4 nodes; here we consider three possible forms the efferents can take:

(1) the weights to ST4 and PYR5/6 in a given column are the same;
(2) the weights to ST4 are close to the weights to PYR5/6 in the same column (say differ by ±10%);
(3) or both sets of weights are completely random.

In each case connections are initialised in the range {0, 1}.

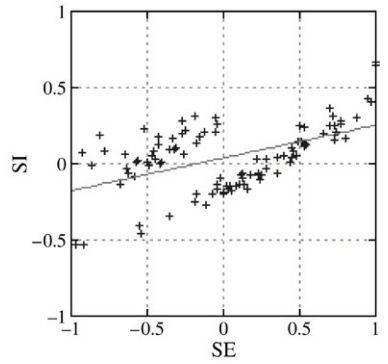
Where the initialisation is totally random, the nature of the thresholded output function for PYR2/3 nodes can result in many PYR2/3 nodes having no output and hence the 3-layer cortical model reduces for the most part to the single node model; unsurprisingly similar results for regression lines are found as for the single node case.



(a) Attend away (probe) $SI = 0.513SE + 0.105$.

(b) Attend probe $SI = 0.757SE + 0.0222$.

(c) Attend away (reference) $SI = 0.513SE + 0.105$.



(d) Reference $SI = 0.215SE + 0.0375$.

Fig. 5. Output gain in single graded neuron model selectively applied of form 1.3 and 0.76. Using the two tailed t test with $p < 0.005$ leads to all results being plotted.

Table 1
Regression slopes for contrast gain, additive with specificity, and output gain with specificity using two-tailed t test, and experimental results for $V2$ and $V4$

	Contrast gain	Additive	Output gain	$V2$	$V4$
Away probe	0.487	0.514	0.513	0.47	0.49
Probe	0.749	0.700	0.757	0.69	0.83
Away reference	0.506	0.513	0.513	0.55	0.6
Reference	0.232	0.263	0.215	0.24	0.21

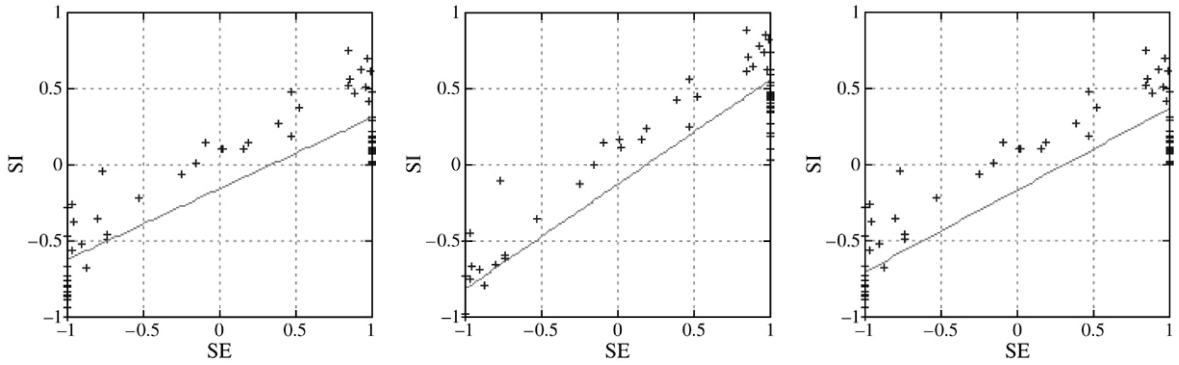
Both weight initialisations 1 and 2 produce PYR2/3 outputs in more columns than the random case in at least one of the five tasks. For each attentional form few differences occur between the different weight organisations 1 and 2, so here we present results only using type 2 from above.

Fig. 6 shows the results for contrast gain attention, using modulation during the attend probe and reference tasks of *2. The regression slopes fit with the experimental results, though there are a number of plotted points that lie on the $SE = \pm 1$ lines. This is due to the output function and shunting inhibition value of -1; the combination of which can lead to zero output; if PYR5/6 have maximum output for probe input alone, and output of zero for reference stimulus alone (or vice versa) we generate points that will lie on $SE = \pm 1$. This distribution of points also causes the regression constants to be negative, as can be seen by inspection of Fig. 6.

Additive attention using a component of magnitude 0.25 and applied differentially as in the single node case produces the results shown in Fig. 7. Inspecting the results for attend probe in Fig. 7 it can be seen that there are again two neuronal groupings; the division in attend reference plot also exists but is less clear.

The attentional values used in the single neuron case for differential output gain of 1.5 and 0.76 are not sufficient to cause the size of changes expected from the experimental results although the slopes do have the correct direction of change. The results plotted in Fig. 8 use $x = 1$ to give attentional values of 2 and 0.5. Slope changes are still smaller than required. In both the attend probe and attend reference conditions there are splits into 2 groups; such a split is not seen in the experimental results, and if all simulated neurons are plotted the 2 distinct groupings remain.

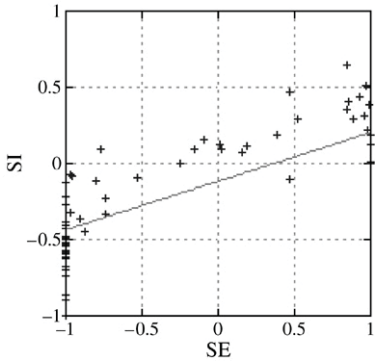
As with the previous results for the single neuron model, we can find values for the three attentional forms such that the regression slopes lie near the experimental results (summarised in Table 2). However, additive and output gain both show two



(a) Away away (probe) $SI = 0.466SE - 0.158$.

(b) Attend probe $SI = 0.686SE - 0.129$.

(c) Away attend (reference) $SI = 0.537SE - 0.169$.



(d) Attend reference $SI = 0.319SE - 0.118$.

Fig. 6. Contrast gain in a 3-layer cortex using shunting inhibition of −1 and a threshold of 0.2 for PYR2/3 output. Weights to ST4 are ±10% of those to PYR56 nodes.

Table 2
Regression slopes for contest gain, additive with specificity, and output gain with specificity using two-tailed t test, and experimental results for $V2$ and $V4$

	Contrast gain	Additive	Output gain	$V2$	$V4$
Away probe	0.466	0.477	0.480	0.47	0.49
Probe	0.686	0.714	0.702	0.69	0.83
Away reference	0.537	0.534	0.489	0.55	0.6
Reference	0.319	0.290	0.262	0.24	0.21

groupings in each of the attend probe and attend reference plots.

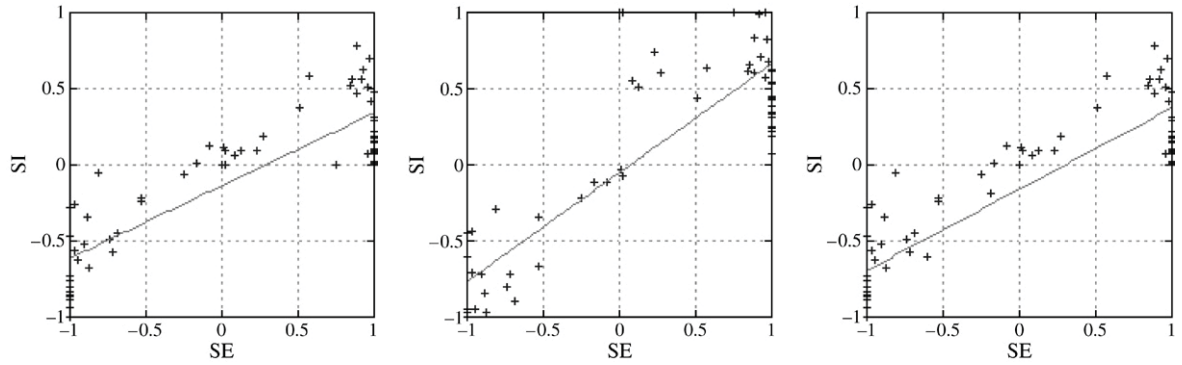
4.3. Spiking single-layer cortex

As described previously the probe and reference stimuli are both represented by a population of 100 nodes; this provides a more even spike train input to the model neurons. Weights are again randomly set with a particular excitatory or inhibitory population totalling up to 1.5.

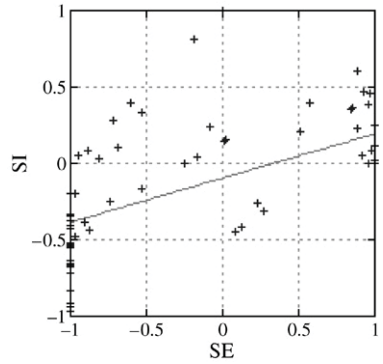
In the contrast gain model an attentional factor of *1.5 gives the results shown in Fig. 9. The regression slopes all show the correct trends (attend probe > attend away (probe), and attend reference < attend away (reference)), although in this case the regression slope for attend reference is rather lower than the experimental results. The regression constants are larger than the experimental results.

For additive attention we add two further neurons, one active in the attend probe condition with excitatory weights to the model neurons classed as probe preferred and inhibitory weights to all other nodes; the second new node is active in the attend reference case and has excitatory weights to reference preferred nodes and inhibitory otherwise. These controllers of additive attention fire at the rate of 15 Hz during their respective attend conditions. The results are shown in Fig. 10, where divisions can be seen for the attend probe and attend reference plots.

The output gain version has x set to 0.2 giving differential attentional values of 1.2 and 0.833; here we plot the graphs using t test with $p < 0.05$ (Fig. 11) and without t test where all neurons are plotted (Fig. 12). These values give, when all neurons are plotted, a slightly larger value for the attend probe case and a slightly smaller value for the attend reference case than the experimental results, but a smaller value for x of say 0.15 which does not produce enough points when the t test is used to define a regression line from a sample of 100 neurons (only 2 points for the attend probe case). The results when all



(a) Away probe $SI = 0.477SE - 0.139$. (b) Probe $SI = 0.714SE - 0.0518$. (c) Away probe $SI = 0.534SE - 0.162$.



(d) Reference $SI = 0.290SE - 0.0971$.

Fig. 7. Additive attention in graded 3-layer model with preference of ±0.25.

neurons are plotted show the correct trends and relationships but when the t test is used (Fig. 11) the trends are correct for the regression slopes with: attend probe > attend away (probe) and attend reference < attend away (reference) but the overall regression slope relationship has attend reference > attend probe conflicting with experimental results. From Fig. 12 the separation into 2 clusters of points can be seen once again for the attend probe and attend reference cases. The regression constants show a large increase in the attend probe and attend reference cases as against the respective attend away regression constants.

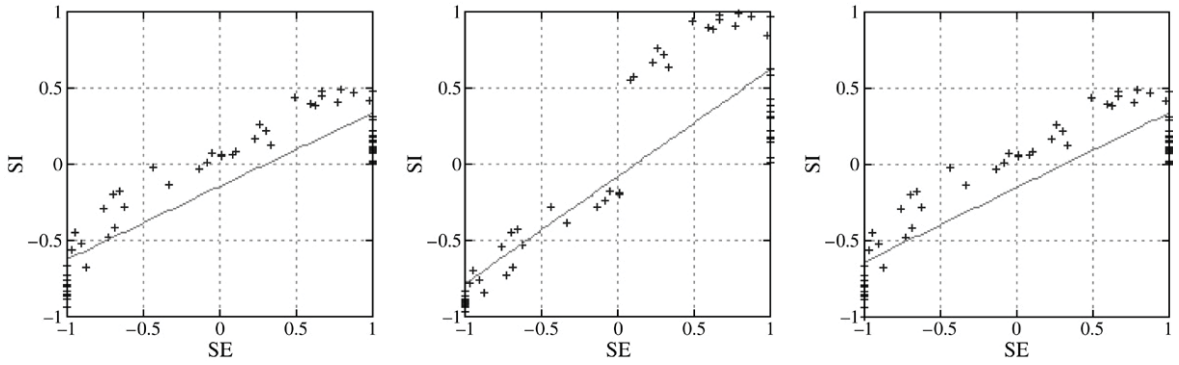
We consider that the attentional affects might be the combination of two of the attentional modes we are considering. Since contrast gain has the best distribution of points along the regression line we combine this with additive attention which has the better results for the regression constant values. Reducing the contrast gain affects to *1.2 and reducing the firing rates of the additive inputs to 7.5 Hz gives the plots shown in Fig. 13. The additive component could be additive attention or simply an additive feedback between cortical regions.

The best fit to the experimental results is the combination of contrast gain and additive; there are no different groupings within these plots as seen for the additive attention only case. Similar affects are seen for output gain though only when all neurons are plotted. The contrast gain case has the good regression line slopes without different groupings but has higher regression constant values than the experimental results; in the combination model the regression constants are reduced into the range of the experimental results for $V4$. Table 3 summarises the regression slope values for the simulations in this section.

4.4. Spiking 3-layer cortex

Contrast gain is modelled as previously by multiplying the reference or probe weights by a factor >1 in the attend reference and attend probe conditions, respectively. Weights to PYR5/6 and IN5/6 nodes are multiplied but not those to IN4 and ST4 (see Section 6). Using an attentional factor of 1.8 gives the results shown in Fig. 14, in all cases it is the firing rates of PYR5/6 nodes that are recorded. The regression slopes show the correct trends and have values that fit with the experimental results but the regression constants are higher than the single-cell results.

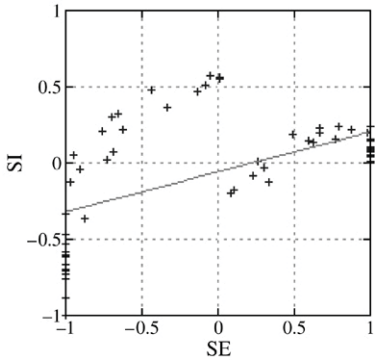
Attention as additive is modelled as previously by testing the model loops and by responses defining the columns as reference preferred or probe preferred. The additive term is again applied differentially as a positive or negative term, using the structure described for the spiking single node model with the same spike rate of 15 Hz. For the current architecture all neurons are plotted when the t test is applied with $p < 0.05$ and the results are shown in Fig. 15. The division into 2 distinct groupings in the attend reference and attend probe cases is again present.



(a) Attend away (probe) $SI = 0.480SE - 0.147$.

(b) Attend probe $SI = 0.702SE - 0.0826$.

(c) Attend away (reference) $SI = 0.489SE - 0.153$.



(d) Attend reference $SI = 0.262SE - 0.0590$.

Fig. 8. Output gain in graded 3-layer model applied selectively with attention = *2 and *0.5.

Table 3
Regression slopes for attend away (probe and reference), attend probe, and attend reference

	Contrast	Additive	Output	Contrast + Additive	$V2$	$V4$
Away (probe)	0.433	0.492	0.310	0.487	0.47	0.49
Probe	0.707	0.836	0.404	0.841	0.69	0.83
Away (ref)	0.526	0.567	0.652	0.556	0.55	0.6
Reference	0.100	0.179	0.464	0.245	0.24	0.21

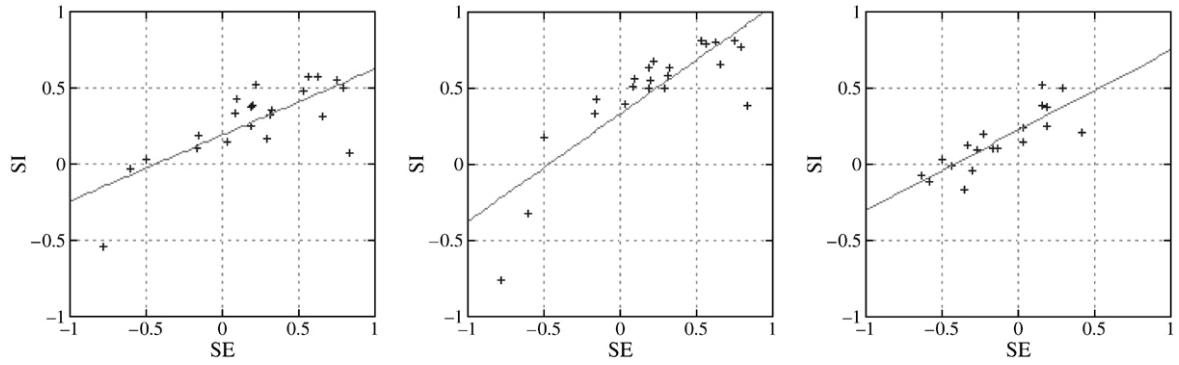
Output gain attention is applied differentially as previously using $x = 0.5$. Figs. 16 and 17 show the results for this form of attention with and without using the t test, respectively. We plot the non-t test results since for this model it is difficult to find an attentional modulation value that produces regression slopes that are as low as the experimental results.

As in the single node model we consider the combination of contrast gain and additive attention (Fig. 18). We reduce the attentional affects to *1.3 for contrast gain and an additive input of 7.5 Hz.

The combination of contrast gain and additive attention give the best fit to the experimental results in terms of regression slopes and distribution of points, though the regression constants are higher here than for the experimental results. The results of contrast gain only are similar to the combined but with higher regression constants, whilst for additive and output gain we can find regression slopes in the experimental ranges for which the distribution of points in the plots are different to the experimental results with different groupings forming in the attend probe and attend reference conditions. Table 4 summarises the regression slope values.

4.5. Mean-field approach

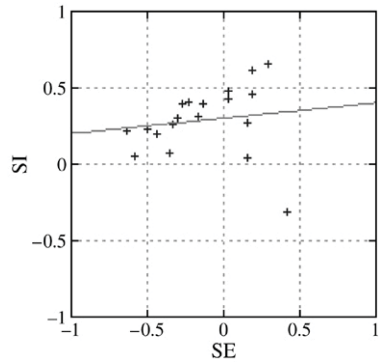
For the mean field results we use 50 copies of the mean field equations with the inputs I_1 and I_2 effected by different values of weights, which are random in the range $\{0, 1\}$. We test only 2 forms of attention: contrast gain and additive. Two input architectures are modelled: one has each pool receiving inputs from only one stimulus; the second has some crossover between inputs with the non-preferred stimulus having an input weight to a pool of maximum 10% of the weight value of the preferred stimulus to the same pool. The latter mimics to some extent the mixing of pool activations between $V2$ and $V4$ in the model of Deco and Rolls (2005). In the model of Deco and Rolls (2005) the biased competition effects from additive attention depend not only on the relationship between the feedforward and feedback weights between $V2$ and $V4$,



(a) Attend away (probe) $SI = 0.433SE + 0.189$.

(b) Attend probe $SI = 0.707SE + 0.331$.

(c) Attend away (reference) $SI = 0.526SE + 0.220$.



(d) Attend reference $SI = 0.100SE + 0.301$.

Fig. 9. Spiking single node, attention modelled as contrast gain. Attention = *1.5.

Table 4
Regression slopes for attend away (probe and reference), attend probe, and attend reference

	Contrast	Additive	Output	Contrast + Additive	$V2$	$V4$
Away (probe)	0.488	0.531	0.483	0.519	0.47	0.49
Probe	0.741	0.758	0.978	0.760	0.69	0.83
Away (ref)	0.586	0.531	0.567	0.542	0.55	0.6
Reference	0.281	0.385	0.483	0.261	0.24	0.21

but also the effective stimuli firing rates and additive attention value. The affective stimuli firing rates are, in these simulations, modified by the input connection strengths, hence we would need for each combination of probe and reference weights to find the region of biased competition to set any feedforward and feedback weights in a multi-layer model. Since we are working with a single cortical region we used a modified version of Deco and Rolls (2005) modulation index M^{BC} reduced to either $V2$ or $V4$, such that:

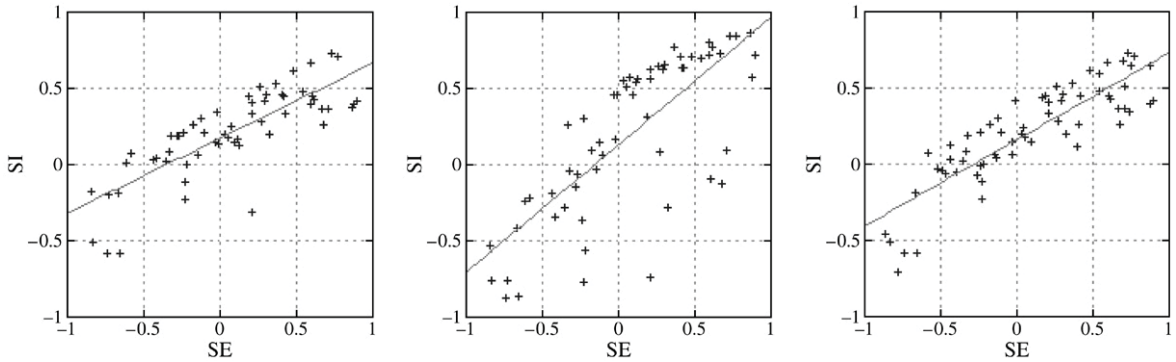
$$M^{BC} = 1 - \frac{\left(\left|\frac{M^P - a}{a}\right| + \left|\frac{M^R - b}{b}\right|\right)}{2}$$

where for $V2$ $a = 0.1$ and $b = 0.08$, and for $V4$ $a = 0.3$ and $b = 0.25$, and P indicates probe preferring pool and R reference preferring pool, with:
M^P = (P firing rate (attend probe) − P firing rate (attend away))/P firing rate (attend away) and
M^R = (R firing rate (attend away) − R firing rate (attend probe))/R firing rate (attend away)
All firing rates are calculated for the presentation of both stimuli.

As we do not have a connected $V2$ and $V4$ M^{BC} was calculated for different additive attention values. The highest M^{BC} values we calculated were in the range 0.5–0.6 which lies within (Deco & Rolls, 2005) the weak biased competition region. With unit input weight strengths for probe and reference the peak M^{BC} values was caused by an additive component of 109 Hz; weaker input weights required smaller additive components.

The results for the 2 different stimuli weight models using contrast gain are shown in Figs. 19 and 20, and those for additive attention are shown in Figs. 21 and 22

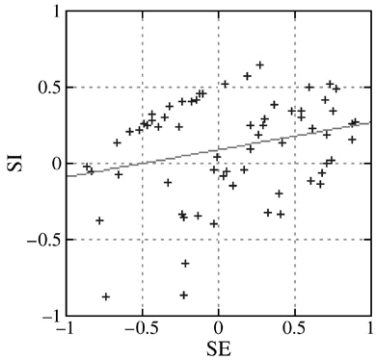
The regression slopes for the contrast gain method are close to the experimental results, whilst those for the additive form have not altered enough from the attend away regression slopes



(a) Attend away (probe) $SI = 0.492SE + 0.171$.

(b) Attend probe $SI = 0.836SE + 0.130$.

(c) Attend away (reference) $SI = 0.567SE + 0.160$.



(d) Attend reference $SI = 0.179SE + 0.0888$.

Fig. 10. Additive attention in spiking single node model.

to be in the experimental range. We can find additive values that give regression slopes close to the experimental results, but this attention needs to take a value of 120 Hz which is close to half the maximum input firing rate, and near the 109 Hz calculated earlier when determining the peak M^{BC} values. In most attend probe or attend reference plots there are some groupings of points, whether attention is additive or contrast gain; this suggests that the architecture itself is unsuitable for investigating the mode of attention in terms of the regression line results of Reynolds et al. (1999). The regression slopes for the different attentional feedback forms in the mean-field approach are shown in Table 5.

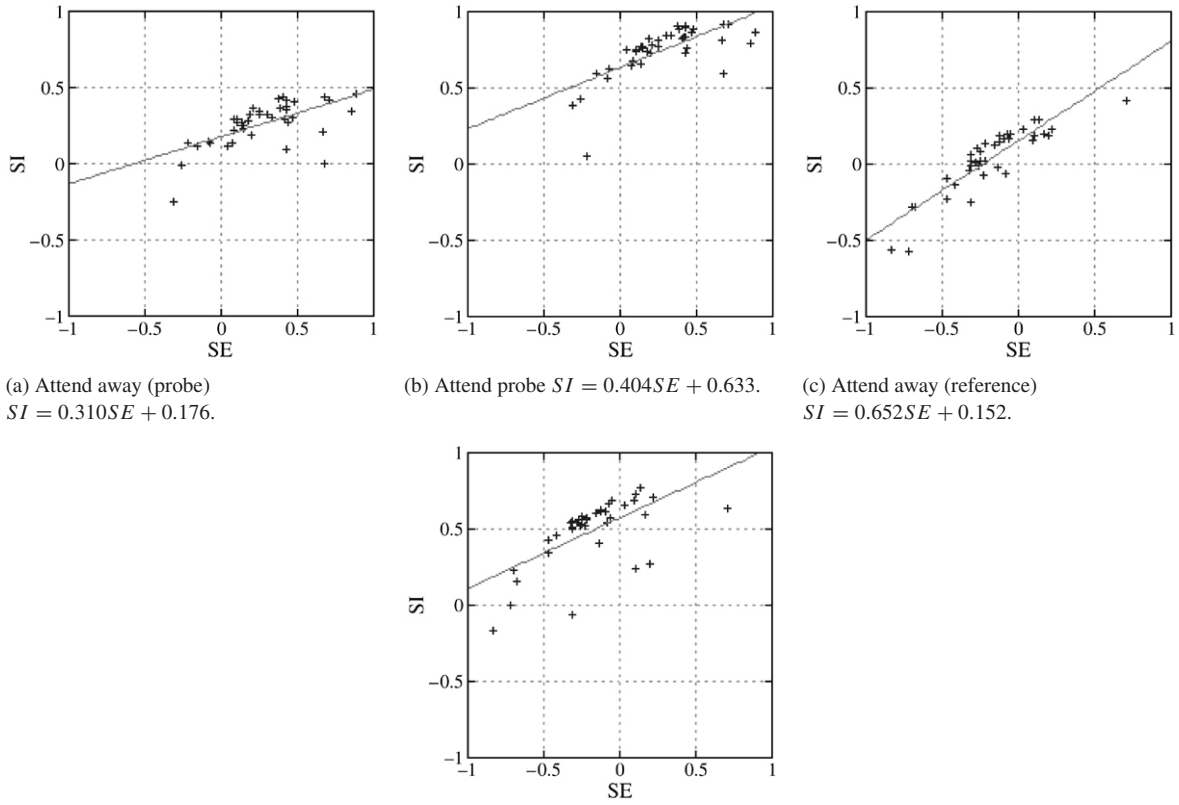
4.6. Conclusions on the simulations

We have investigated the action of three different forms of attention (contrast gain, additive and output gain) in various single cell and cortical models. The task involves five different stimuli conditions: presentation of probe alone, presentation of reference alone, presentation of both with attention directed away, presentation of stimuli with attention directed towards probe, and presentation of stimuli with attention directed towards reference. From the resulting firing rates or equilibrium potentials depending if the model used spiking neurons or graded-response neurons we calculated the SE and SI values as defined in Eqs. (1) and (2) for each neuron. A regression line analysis was performed finding the best fit for the line $SI = mSE + c$, where m is the regression line slope and c the regression constant. The task and analysis follows that performed in Reynolds et al. (1999) using single cell recordings from nodes in $V2$ and $V4$.

The three methods of attention modelled produced varying results. The output gain version was generally the poorest; it being difficult to reliably find modulation factors that gave regression slopes within the experimental range for the 3-layer cortical models. Whilst the slopes for the single node models did fit the experimental results, all results (single and multi-layer models) had distinct neuronal groupings within the attend

Table 5
Regression slope values for the mean field simulations and single cell experimental results

	Contrast gain (Fig. 19)	Additive (Fig. 21)	Contrast gain (Fig. 20)	Additive (Fig. 22)	$V2$	$V4$
Away probe	0.517	0.495	0.500	0.447	0.47	0.49
Probe	0.732	0.608	0.697	0.578	0.69	0.83
Away reference	0.516	0.500	0.492	0.447	0.55	0.6
Reference	0.275	0.360	0.274	0.332	0.24	0.21



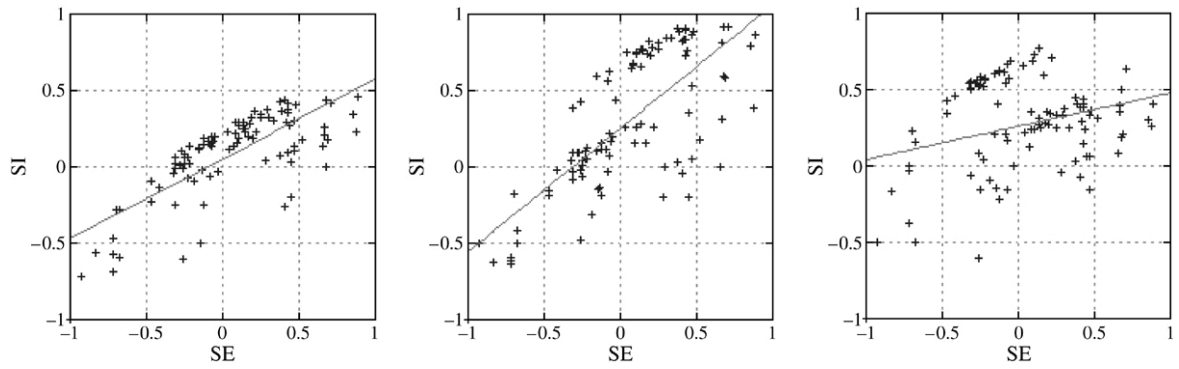
(a) Attend away (probe) $SI = 0.310SE + 0.176$.

(b) Attend probe $SI = 0.404SE + 0.633$.

(c) Attend away (reference) $SI = 0.652SE + 0.152$.

(d) Attend reference $SI = 0.464SE + 0.572$.

Fig. 11. Output gain in single spiking node model. Attention is 1.2 and 0.833 and t test is used.



(a) Attend away $SI = 0.517SE + 0.0517$.

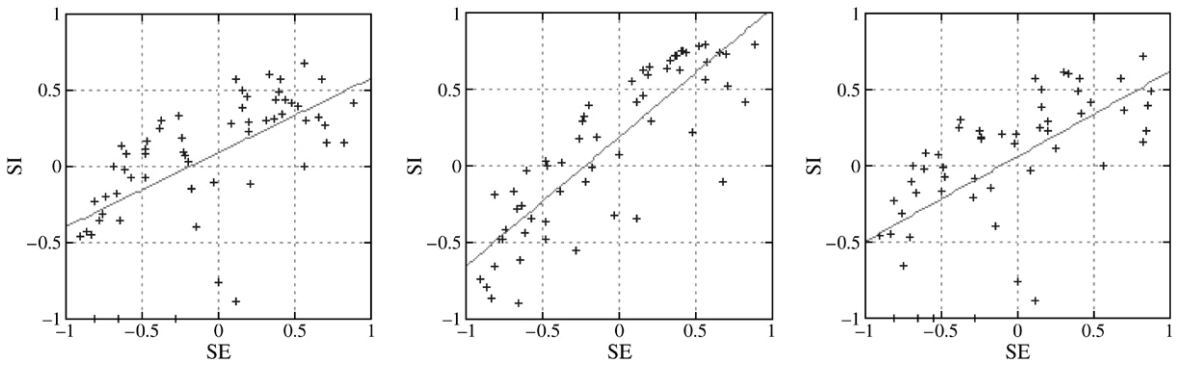
(b) Attend probe $SI = 0.808SE + 0.252$.

(c) Attend reference $SI = 0.218SE + 0.260$.

Fig. 12. Output gain in single spiking node model. Attention as above. All neurons plotted.

probe and attend reference plots; such distributions were not reported for the single-cell experimental recordings (Reynolds et al., 1999). The regression constant values were, overall, the most dissimilar to the experimental results showing large increases from the attend away cases to the attend probe and attend reference plots.

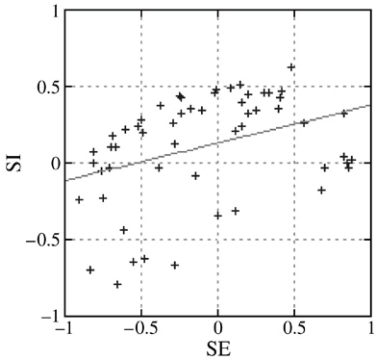
Where attention was modelled as additive, except in the mean field results, we could find a suitable additive graded or spike train input that would generate regression slopes in the expected range. However in these results the attend probe and attend reference results show two different distributions within the plots, each of which have quite different regression slopes if each population was treated separately; the calculated regression slope plotted on each graph lies somewhere in between the regression slope for each population. The additive form of attention did perform well in the regression constant values. This to some extent was caused by the differential application of the additive attention that led to the excitation of the population that preferred the attended stimulus whilst, at the same time, inhibiting the population that preferred the other stimulus. This resulted in a similar increase and decrease either side of the SI axis, which leads to a small change in the attend



(a) Attend away (probe) $SI = 0.487SE + 0.0890$.

(b) Attend probe $SI = 0.841SE + 0.187$.

(c) Attend away (reference) $SI = 0.556SE + 0.0595$.



(d) Attend reference $y = 0.245SE + 0.130$.

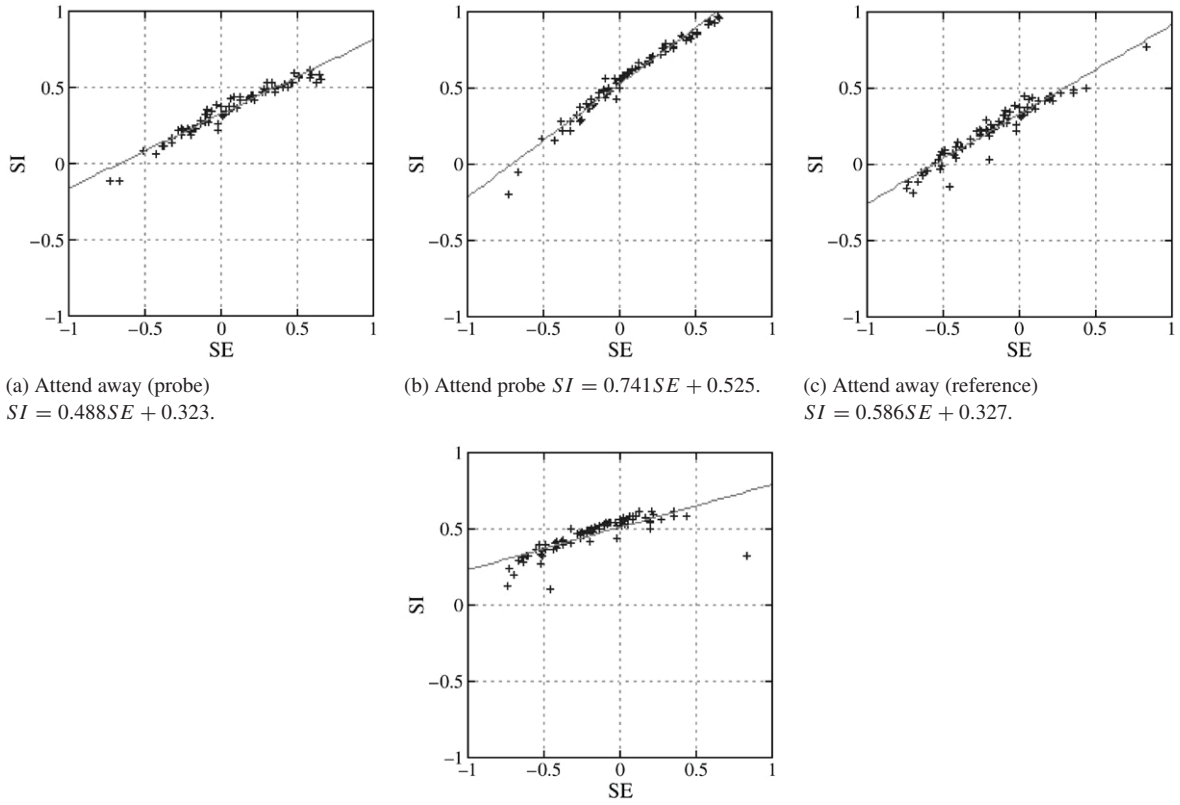
Fig. 13. Using a combination of contrast gain (*1.2) and additive attention (± weights from nodes at 8 Hz) in single node spiking model.

probe and attend reference regression constant values from the values obtained in the attend conditions.

The contrast gain results gave the best fits to the experimental regression slope results, where a single form of attention was used. It was quite easy to find values for the different models—single and multi-layer, graded and spiking—that generated good regression slope values. Where the contrast gain models struggled was in the regression constant values, which were less than those calculated for the output gain models, but still had values, in the spiking models, that were too large, and showed a large increase moving from attend away to attend probe and attend reference.

The combination of additive attention and contrast gain produced the best results for the single layer spiking models. By reducing the modulatory factor of contrast gain and the additive firing rate we managed to keep the best parts of each form; the good matching of the slopes by contrast gain and the good matching of the constants by additive attention but without generating the splits seen in additive attention alone. The results though for the spiking 3-layer model with both attentional forms were not as good.

The mean field results were disappointing for both contrast gain and additive attention. A modulatory value for the contrast gain form could be found that fitted the experimental data for the regression slope values, whilst the regression constants were again too large. Attention as additive required a large attentional component to generate regression slopes comparable to the single-cell results. In all simulations the distributions in the SE–SI plots showed different populations. The distinct groupings seen throughout Figs. 19–22 are due to the architecture which results in the almost total annihilation of firing in the non-preferred pools for the input of a single stimulus. Hence, when only the probe is presented the firing rates of the reference preferred pools is below spontaneous, whilst probe rates are much greater than spontaneous; the reverse results when only the reference stimulus is presented, which then leads to a polarisation into 2 groups when SE (which is probe – reference) is calculated. The validity of using the mean field equations to model the experimental results must also be questioned. The experimental results come from single-cell recordings of up to 100 neurons whilst each set of mean field equations represents a ‘collapsed’ population of nodes (hundreds or even thousands of spiking neurons). Hence, the range of experimental results is already contained within the results of a single set of mean field equations. When we solve 50 mean field equations for different input weights we are not looking at results comparable to the single-cell recordings. To investigate the (Deco & Rolls, 2005) model SE–SI plots we need then to use the full spiking architecture. The interaction between neurons that prefer the same stimulus but have different response levels due to varying input strengths combined with the small region for which feedforward and feedback connection strengths lead to biased competition from an additive attention component seem excessively complex.



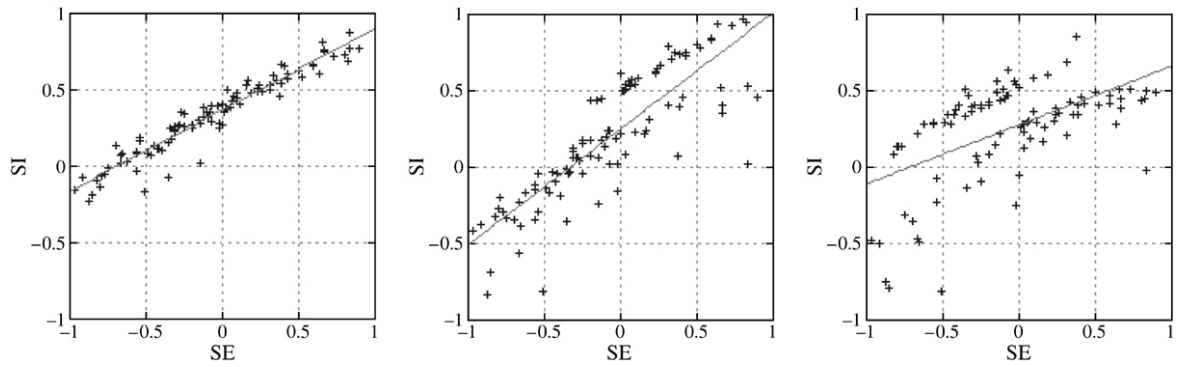
(a) Attend away (probe) $SI = 0.488SE + 0.323$.

(b) Attend probe $SI = 0.741SE + 0.525$.

(c) Attend away (reference) $SI = 0.586SE + 0.327$.

(d) Attend reference $SI = 0.235SE + 0.493$.

Fig. 14. 3-layer spiking network, attention as contrast gain *1.8.



(a) Attend away $y = 0.531x + 0.367$.

(b) Attend probe $y = 0.758x + 0.251$.

(c) Attend reference $y = 0.385x + 0.276$.

Fig. 15. Attention as additive in 3-layer spiking model. The results using the t test and without t test are exactly the same.

Indeed it is not clear for such a structure that feedback and feedforward weights could be calculated to produce biased competition from additive attention and which also shows multiplicative gain.

5. Theoretical analysis of the simulation results

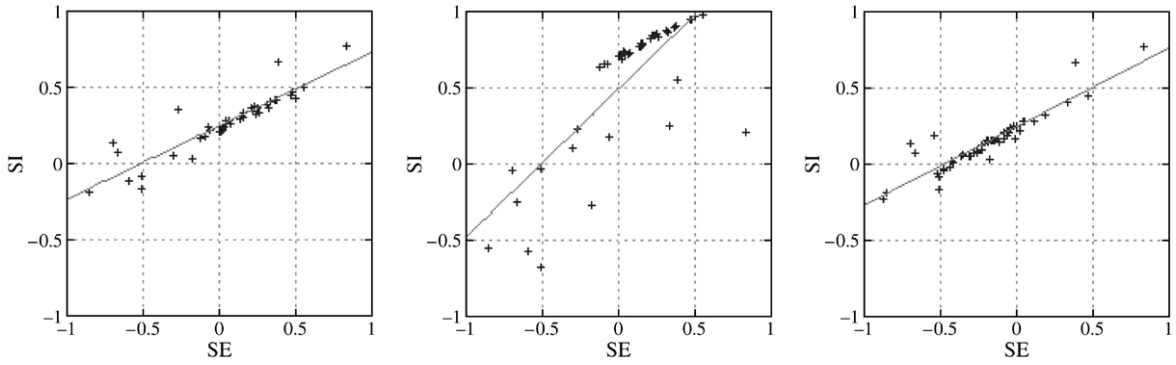
We analyse mathematically here the detailed simulation results presented in the previous Section 4 in order to gain a further understanding of the various results. We start by considering the shunting inhibition neuron, whose equation is as used in Reynolds et al. (1999). The neuron membrane potential V satisfies the first-order differential equation:

$$\frac{\mathrm{d}V}{\mathrm{d}t} = -AV + (B - V)E + (C - V)I \tag{18}$$

where E and I are the excitatory and inhibitory inputs to the neuron:

$$E = \sum_j x_j w_j^+ \tag{19a}$$

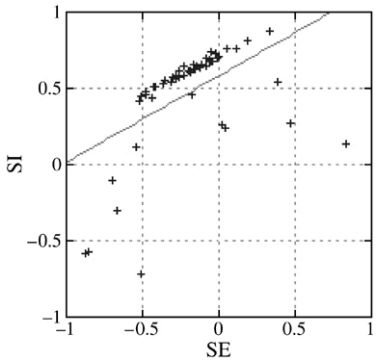
$$I = \sum_j x_j w_j^- \tag{19b}$$



(a) Attend away (probe) $SI = 0.483SE + 0.246$.

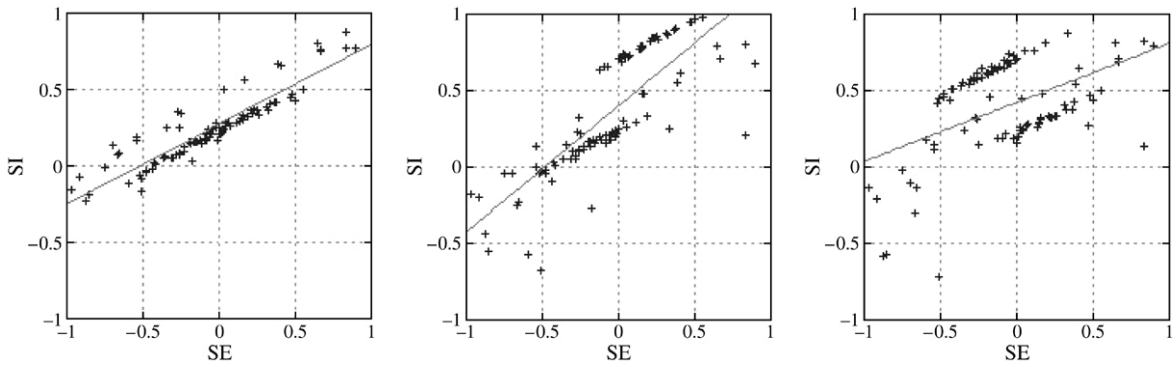
(b) Attend probe $SI = 0.978SE + 0.494$.

(c) Attend away (reference) $SI = 0.567SE + 0.263$.



(d) Attend reference $SI = 0.483SE + 0.580$.

Fig. 16. Output gain attention in spiking 3-layer model, t test used.



(a) Attend away $SI = 0.518SE + 0.271$.

(b) Attend probe $SI = 0.824SE + 0.400$.

(c) Attend reference $SI = 0.387SE + 0.423$.

Fig. 17. Output gain attention in 3-layer spiking model, all neurons plotted.

where w_j^+ and w_j^- are the excitatory and inhibitory weights respectively onto the given neuron from the jth neuron with input strength x_j (so violating Dale's law, but allowing for mimicking the effects of inhibitory interneurons). The temporally stable solution to Eq. (18) is the value:

$$V(\infty) = \left(\frac{BE + CI}{E + I + A}\right). \tag{20}$$

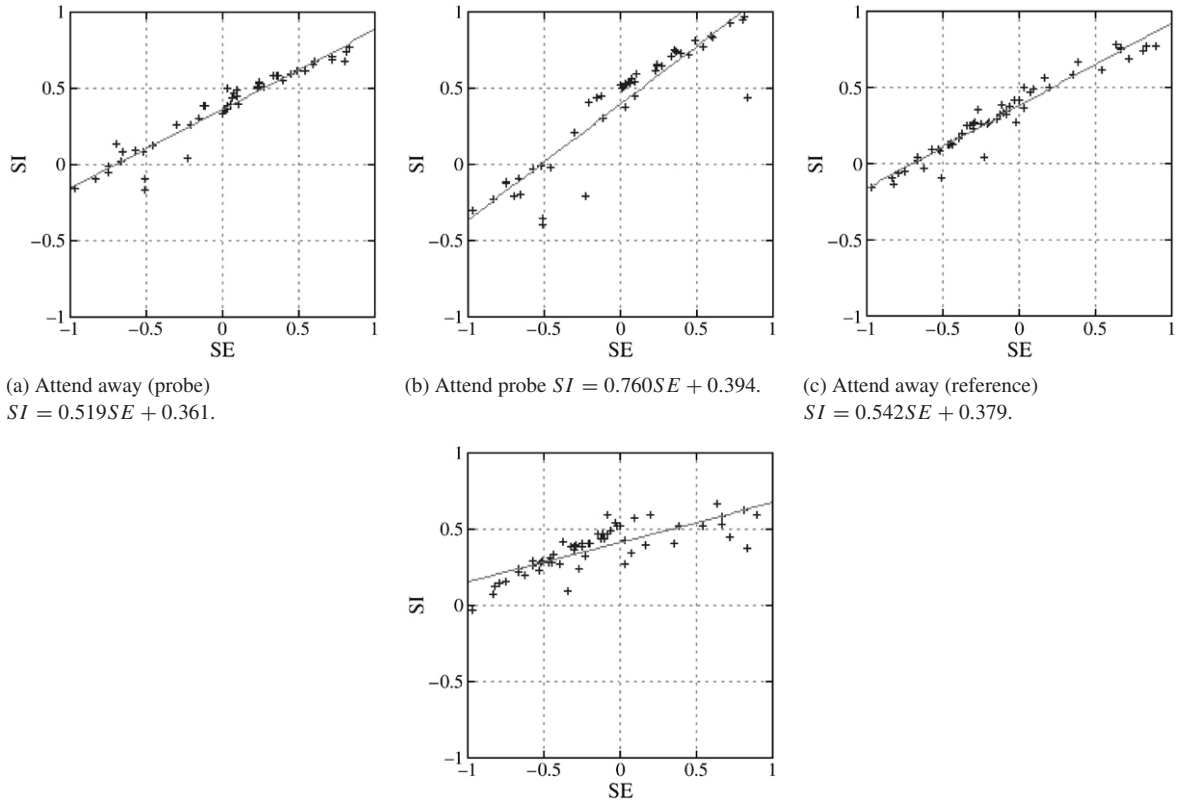
The resulting formula for the values of the sensitivity indices already introduced, SE and SI, are then, using (18) and (19) and extending the analysis in Taylor and Rogers (2002),

Appendix B

$$SI = (P + R) - R = \frac{B(E_1 + E_2) + C(I_1 + I_2)}{E_1 + E_2 + I_1 + I_2 + A} - \frac{BE_1 + CI_1}{E_1 + I_1 + A} \tag{21a}$$

$$SE = P - R = \frac{BE_2 + CI_2}{E_2 + I_2 + A} - \frac{BE_1 + CI_1}{E_1 + I_1 + A} \tag{21b}$$

where the summations in (19) are reduced to being only over the probe and reference stimuli, denoted by the indices $j = 1$ for R and 2 for P.



(a) Attend away (probe) $SI = 0.519SE + 0.361$.

(b) Attend probe $SI = 0.760SE + 0.394$.

(c) Attend away (reference) $SI = 0.542SE + 0.379$.

(d) Attend reference $SI = 0.261SE + 0.412$.

Fig. 18. Mixture of contrast gain and additive in 3-layer spiking model, t test used.

In Taylor and Rogers (2002) the mean of the formulae (21), over a set of random choices of the excitatory and inhibitory weights onto the attended neuron were approximated by assuming independence of the various factors entering on the right hand side of Eq. (21). We obtained there the simplified 'universal' formula

$$\langle SI \rangle = \frac{\langle SE \rangle}{1+u} + \text{threshold} \tag{22}$$

where

$$\text{threshold} = A(\text{constant}) \tag{23a}$$

(so being proportional to the excitatory membrane resting potential, the constant being determined by properties of the random ensemble of weights), and

$$u = \left\langle \frac{X_1}{X_2} \right\rangle = \left\langle \frac{R}{P} \right\rangle \tag{23b}$$

is the relative attention factor to the reference as compared to the probe. In the attend away case, $u = 1$, in the attend probe case $u = 1/5$ and in the attend reference case $u = 5$ (taking the multiplicative gain contrast from attention to be the value of 5 as used by Reynolds et al. (1999)). This leads to the slopes of 0.5 (attend away), 0.83 (attend probe), and 0.17 (attend reference). These are very close to the values determined for the collection of neurons recorded from $V2$ and $V4$ in Reynolds et al. (1999).

We now extend the linear regression formula (22) to the case of a non-zero inhibitory resting membrane potential C in Eq. (18), leading to the same formula as in (22) but now with the threshold term modified to be

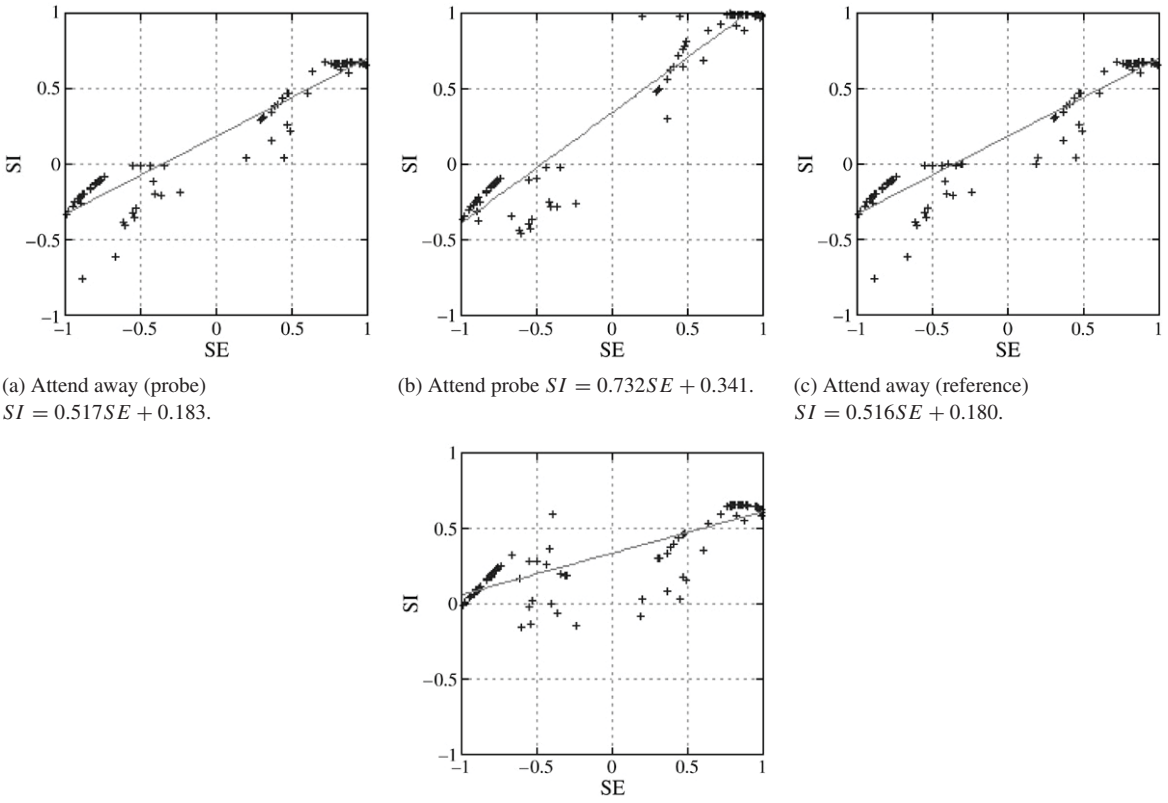
$$\text{threshold} = A(\text{constant}) + C(\text{constant}') \tag{23c}$$

(where the constant$'$ in (23c) arises from the various mean values of the statistical ensemble).

Turning to the details of the simulations of Sections 3 and 4, we already explained above the results of Fig. 3, which shows the slope dependence of the regression line for the SI–SE plot on the position of the focus of attention in the case of a single graded node. The effects arising in this plot when attention is additive, as shown in Fig. 4, may be explained from the fact that such an attention signal is seen as a purely additive shift in the excitatory component:

$$E = \sum_j x_j w_j^+ + \text{attn} \tag{24}$$

(where the added term 'attn' on the right hand side of Eq. (24) is the additive attention feedback term, assumed excitatory). On repeating the earlier analysis leading to Eq. (22) we obtain the same formula, but now with A replaced by $(A + \text{attn})$. Thus all that changes due to the additive attention feedback term is that the SE–SI regression line is shifted upwards, with the threshold increasing by a term proportional to the attn component; there is no change in slope of the regression SE–SI



(a) Attend away (probe) $SI = 0.517SE + 0.183$.

(b) Attend probe $SI = 0.732SE + 0.341$.

(c) Attend away (reference) $SI = 0.516SE + 0.180$.

(d) Attend reference $SI = 0.275SE + 0.335$.

Fig. 19. Single layer model, 50 mean field equations using different weight values from the probe to pool 1 and reference input to pool 2, no crossover between inputs and pools. Contrast gain case where the weights from probe and reference are multiplied by a factor of 1.8 during the appropriate attend condition.

plot. This may differ between the attend away, attend probe and attend reference cases, leading to different values for the SI-axis crossing in the three different attend conditions. However the slope will be very close to 0.5 in all three cases.

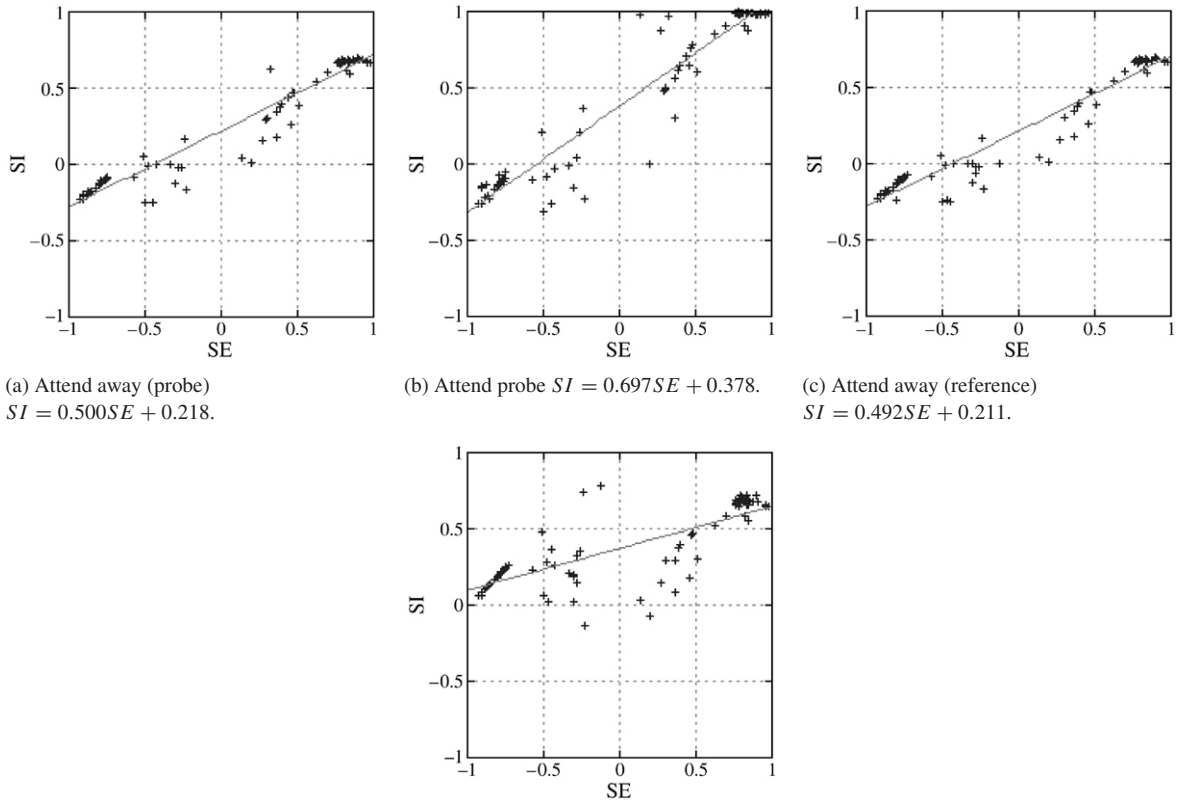
We note that the results of Fig. 4 do show differences of slope between the attend away, attend probe and attend reference cases, in agreement with experiment. This can be explained as arising by the selection of neurons by choosing those which are probe preferred or reference preferred, across the ensemble of neurons, the two sets being defined by probe weights > reference weights or vice versa. However we see that there is almost identity in slope between those components of the points in Fig. 4 which arise as a linearly-distributed population. This is the lower left-sided population in Fig. 4(b) and the left-sided population in Fig. 4(d). A similar feature also arises in Fig. 10(b) and 10(d) in the spiking one-layer cortex case, where the linear populations in those two figures again have the slope of 0.5, with the other points arising, it can be conjectured, due to the selection process.

A very similar architecture to the 3-layered cortical model of Grossberg and Raizada (2000) was created for use with graded neurons, and was shown in Fig. 2. It consists of a pyramidal neuron in layers 2/3, a pair of inhibitory and stellate neurons in layer 4, and a pyramidal neuron in layers 5/6 (as in Grossberg and Raizada (2000)). We consider the effect of attention modulation on the pyramidal neuron PY56 in the layers 5/6, which is taking inputs from the lower cortical site ($V1$), from attention feedback, and from feedback from layers 2/3. Then the extension of the stable layer 5/6 neuron potential $V(56)$, extending Eq. (20) to the 3-layered case, and with $C = 0$, is:

$$V(56) = \frac{B(56)E(V1IN, \text{attn}, F(IN, \text{PY56}))}{A(56) + (E + I)(fn)} \tag{25}$$

and where E and I, the excitatory and inhibitory inputs to the PYR56 neuron, depend explicitly and additively on the input IN from any $V1$ feedback input, denoted as $V1\ IN$ in Eq. (25), the attention feedback signal, denoted attn there, and the feedback from layer 2/3 to PY56 denoted as $F(IN, \text{PY56})$, we denote these variables by fn. The latter feedback leads to a recurrent equation for the value of $V(56)$ from (25). We work in the parameter region where there is no bifurcation to non-zero continuing excitation of PYR56, since there is no observation of such a bifurcation in the temporal data for the $V2$ or $V4$ neurons, nor did Grossberg and Raizada (2000) appeal to such a bifurcation effect to fit the temporal data of Reynolds et al. (1999).

The results of Fig. 6, arising from the architecture, and with shunting inhibition set at 0, are to be expected to be close to those of Fig. 3, with attention dependency of the slope of the SE–SI line as given by the universal $\frac{1}{1+u}$ (which can also be deduced by the similarity of (25) to (20)); the shifts of the crossing of the SI-axis are also explained as in the graded case. Also the effect of attention in the simulation of Fig. 7



(a) Attend away (probe) $SI = 0.500SE + 0.218$.

(b) Attend probe $SI = 0.697SE + 0.378$.

(c) Attend away (reference) $SI = 0.492SE + 0.211$.

(d) Attend reference $SI = 0.274SE + 0.371$.

Fig. 20. Contrast gain results with attentional modulation of 1.8. All nodes get inputs from probe and reference stimuli.

mirrors closely that of contrast gain: the inputs only from the attended input are boosted by attention. However the slopes are not expected to satisfy the general formulae (22) and (23) since the contrast gain itself arises only from the larger weights involved in the relevant set of neurons.

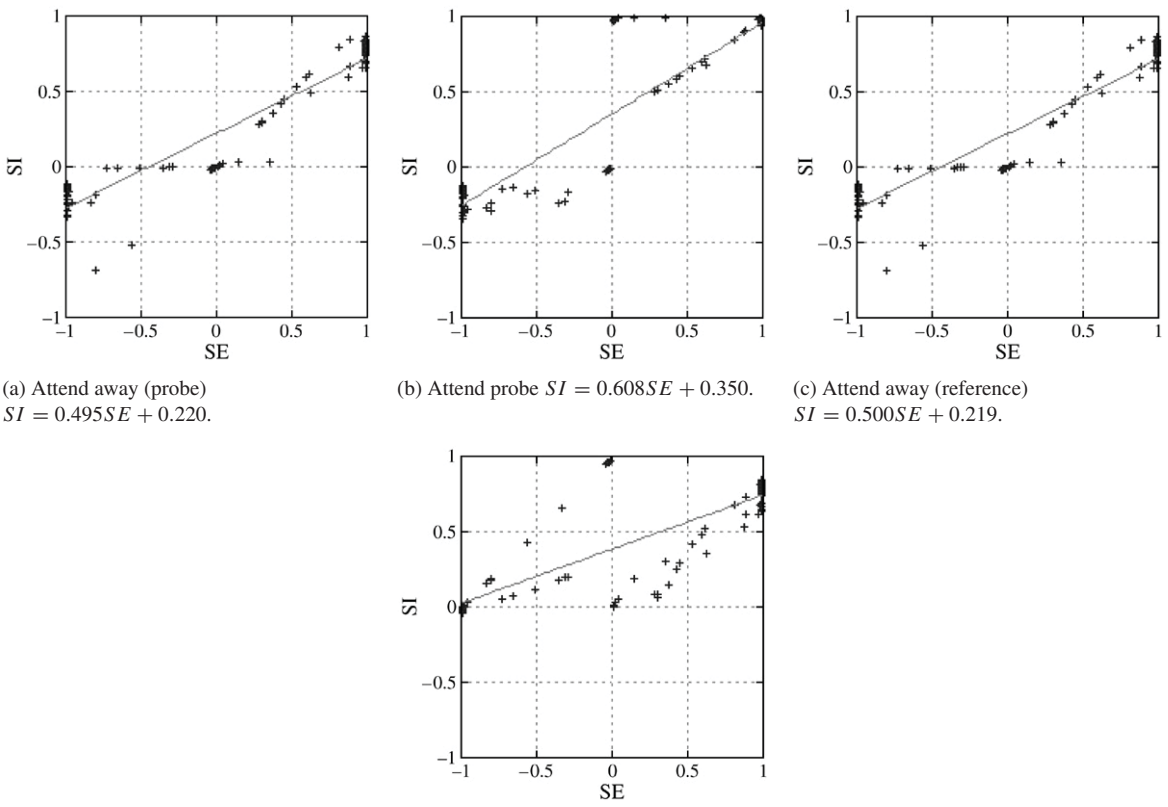
When the inhibitory shunting membrane potential is changed from 0 (as used in the Reynolds et al. (1999) simulation) to -1, then the term C in Eqs. (18) and (20) is now present. For the contrast gain model of attention feedback, the SI–SE curves in the three different attention conditions will still have the slopes they did in the case of $C = 0$, but now the crossing of the SI axis will be altered by a term proportional to C, so will be negative. This is precisely what is observed in a separate simulation (not reported on here) in the 3-layered case; a similar results arises in the 1-layered case. Let us consider the case that the attention feedback is additive in the three-layer case. If we assume that there is no recurrence from the layer 2/3 to layer 5, so that $F(IN, \mathrm{PY}56)$ now becomes solely a function of IN, then we may analyse the solution (25) (with its additive structure) as in Taylor and Rogers (2002) so as to lead to a similar linear regression equation for SI against SE as observed experimentally, but now with slope not completely independent of the attention feedback signal (as in the one-layer case). The bias is due to the selection process involved in choosing neurons with probe weights (for preferred probe) or reference weights (in the other case) being greater than the other set of weights. There are now two changes brought about by attention:

(1) to change the crossing of the regression line on the SI axis;
(2) modification of the regression slopes by the extra amplification or reduction in taking account of the selection of weights in the ratio u of Eq. (23b). For the probe preferred case then the constant $u < 1$, whilst for the reference preferred case we will have $u > 1$, in each case the degree of amplification or reduction about 1 for the value of u being expected to depend on the mean weights assigned to the two populations. Thus we expect, as shown in Fig. 10 (and 7), that the total slopes will be able to be adjusted close to the experimental data by the selection process of the relevant neuron sub-populations.

We have not considered in detail the spiking simulations either in the 1-layer or 3-layer cases. The two-population structures of various plots arising from the spiking neuron simulations were discussed in Section 4; the other aspects of the slopes, we suggest, are already covered by our analysis in the light of an expected averaging process turning the spiking neuron equations into those involving graded neurons.

6. The neurochemistry of attention and its implications for micro-attention

Of the four primary neuromodulators: acetylcholine (ACh), norepinephrine/noradrenaline, serotonin and dopamine, the most important seem to be ACh and noradrenaline. Serotonin's



(a) Attend away (probe) $SI = 0.495SE + 0.220$.

(b) Attend probe $SI = 0.608SE + 0.350$.

(c) Attend away (reference) $SI = 0.500SE + 0.219$.

(d) Attend reference $SI = 0.360SE + 0.382$.

Fig. 21. Attention as additive, only positive component of +60 Hz used. No crossover with inputs.

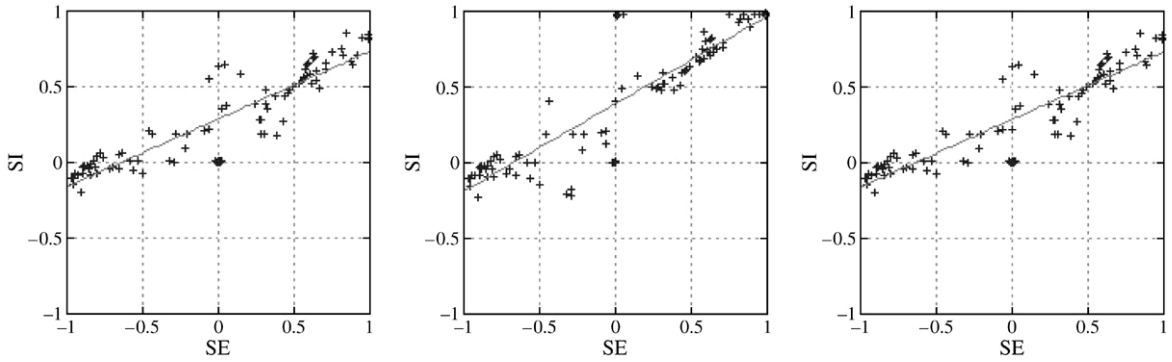
role in visual attention seems to be limited, while dopamine affects attention (Ye, Xi, Peng, Wang, & Guo, 2004), but in a manner that is not yet entirely clear. The result of noradrenaline action is alerting to new stimuli, not the focusing of attention (Foote, Berridge, Adams, & Pineda, 1991). Whilst ACh has been shown to have a number of attentional actions—see (Davidson & Marrocco, 2000; Witte, Davidson, & Marrocco, 1997) for example—it is the most important for the focusing of attention (Witte et al., 1997). ACh action is mediated by two receptors: muscarinic and nicotinic. When utilising muscarinic receptors ACh has a number of affects: alerting (Witte et al., 1997), orienting (Davidson & Marrocco, 2000), and modulation of visual attribute processing (Mentis et al., 2001). However, it is nicotinic ACh receptors (nAChR) that are the most important for visual attention (Witte et al., 1997).

The question then is: what is the mechanism by which nAChR allow ACh to focus attention? We have discussed this elsewhere (IJCNN, 2005), so summarise the arguments provided there. The source of cortical ACh is the nucleus basalis of Meynert (NBM), and the spread of axons from NBM to cortex is in general diffuse (Everitt & Robbins, 1997; Kimura, 2000; Lucas-Meunier, Fossier, Baux, & Amar, 2003). Studies of the distribution of ACh varicosities in the cortex (Turrini et al., 2001; Umbriaco, Watkins, Descarries, Cozzari, & Hartman, 1994) have shown differing results for the number of varicosities that are synaptic: 15% in the rat, 45% in the macaque. Which leads to the conclusion that ACh undergoes volume distribution and hence is unable to support focusing of attention on a specific stimulus or goal. The proportion of synaptic ACh varicosities rises to 67% in humans (Smiley, Morrell, & Mesulam, 1997) which suggests that a more specific action of ACh could be a possibility. Indeed many of the components needed for ACh to act in a focusing manner are present:

(1) high proportion of synaptic varicosities (this at least is true in humans (Smiley et al., 1997));
(2) that cortical nicotinic receptors act in an amplificatory manner (to justify the two multiplicative attention methods we are modelling) on nearby synaptic weights. At least for cortical interneurons there is evidence for the effect of nAChR on nearby synaptic sites (Alkondon, Pereira, Eisenberg, & Alburquerque, 2000); the amplificatory action is also supported (Kimura, 2000);
(3) that attention feedback control signals axon boutons arrive near to the nicotinic receptors on the cortical neurons, this can occur since the higher level feedback can go to layer 5 which has relatively high densities of synaptic varicosities.

We suggest that focusing of attention can act in a multiplicative manner via nAChR at layer 5 cortical neurons due to local coincidence of ACh varicosities and feedback axon boutons.

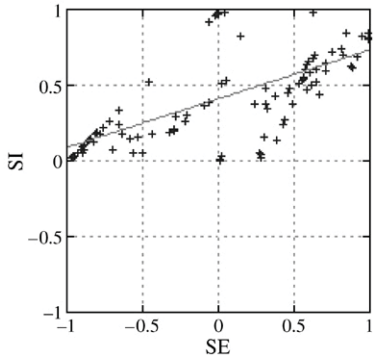
There is still open the question of how learning of attention in feedback style is achieved. We speculated in Taylor, Hartley, and Taylor (2005) that with this neurochemistry and



(a) Attend away (probe) $SI = 0.447SE + 0.285$.

(b) Attend probe $SI = 0.578SE + 0.388$.

(c) Attend away (reference) $SI = 0.447SE + 0.284$.



(d) Attend reference $SI = 0.322SE + 0.409$.

Fig. 22. Additive attention component of 60 Hz. All neurons have inputs from probe and reference.

architecture the multiplicative attention model would have the sigma–pi form:

$$w_{i,j,k} = w_j w'_k \tag{26}$$

where w'_k is the strength of the connection from the kth higher-order attention control neuron to the ith node, and w_j is the connection from the jth lower-order node. The learning of this product form is expected through the separate learning of the separate factors in the product term on the right-hand side of Eq. (25), leading to a learning rule of:

$$\Delta w_{i,j,k} = \mathrm{OUT}_i[\mathrm{OUT}_j w'_k + \mathrm{OUT}_k w_j]. \tag{27}$$

We have used the sigma–pi form of (26) to explore the more global effects of attention in Taylor et al. (2005), Taylor, Panchev, Hartley, Kasderidis, and Taylor (2006), and could have used such a structure here in the spiking models, rather than simply multiplying the weight values when attention was present.

In total, we suggest that the detailed action of ACh supports the contrast gain microstructure model of attention feedback. The presence of some additive feedback component is to be expected as arising from lower-level Hebbian correlation learning. From an information feedback point of view there is clearly greater specificity in the sigma–pi feedback, since greater information is needed to specify the set of three-index symbols $w_{i,j,k}$ defining the sigma–pi feedback weights than solely the $w_{i,j,}$ for the additive feedback weights

7. Discussion

We can summarise our results by stating that only the contrast gain mechanism of attention feedback control can model closely the regression plots obtained by Reynolds et al. (1999), showing the best fits to the regression slope values and distribution of points in the plots. However the architecture and possibly the equations need to be made more complex to account for all the regression line data.

7.1. General comments

Two models of attention feedback (output gain and additive feedback) each lead to clear differences with the Reynolds et al. (1999) regression plots. In the case of output gain we cannot reliably replicate the regression slope values across all models and there exists two populations that form in the attend probe and attend reference conditions. Similar divisions occur if output gain is only applied to one population at a time, rather than our differential application shown here. If an excitatory additive component only is used we need to use a larger modulatory factor to obtain results near to the experimental regression slope values. The addition of an inhibitory pool is unlikely to help the division since overall our total firing rates between the attend away and the attend probe and attend reference results are similar, hence an inhibitory pool would not change its firing rate significantly and is more likely to affect the lower spiking pool actually helping to accentuate the

division. The additive attention model suffers also from two populations in the attend probe and attend reference plots. In this the differential application of differential additive attention (positive to the population preferential to the attended stimulus, negative to the others) mimics to some extent the effects of an inhibitory pool. If the additive attention had a spectrum of response from a strong positive term to a strong negative term as neurons' response changes from strongly preferential to attended stimulus to strongly non-preferential to the attended stimulus, respectively, then we would expect to see a single population of points in the attend stimulus plots. Such a response seems highly specific and difficult to achieve by learning.

In the light of various claims by other authors we need however to assess our results. Thus in Grossberg and Raizada (2000) it is claimed that a three-layered cortical model, with graded neurons, would satisfactorily simulate the results of Reynolds et al. (1999). A careful perusal of the paper of Grossberg and Raizada (2000) shows, however, that there was no simulation of the detailed regression results of Reynolds et al. (1999) that have been investigated here. Indeed one of our reasons to use the 3-layered cortical model of (Grossberg & Raizada, 2000) was to extend their simulation of the results of (Reynolds et al., 1999) to the regression analysis. The results of our investigation, presented above and further commented on above in this section indicate that the model of (Grossberg & Raizada, 2000) does not give results fitting the regression data of (Reynolds et al., 1999). This failure can be put down to a number of deficits possessed by attention as additive feedback, especially the lack of stimulus specificity in amplifying the attended stimulus input to a specific neuron which is itself receiving attention feedback. Across an ensemble of such neurons the stimulus specificity is not sharp enough. There is support for our results from (Hamker, 2004), who only used graded neurons and a single layered cortex; nor did he consider additive or output gain forms of attention feedback.

Our results are also supported somewhat by recent experimental results (Williford & Maunsell, 2006) though they found whilst contrast gain could describe the results, they concluded that response gain and activity gain were marginally better fits. However their experimental paradigm is different to that used by (Reynolds et al., 1999) which has been used here; additionally, as they point out, the attentional amplification was lower for their results than in other studies. Whilst for our results presented above the noise present for probe and reference inputs (in the spiking models) is not sufficient to cause spontaneous firing of these inputs, in simulations where the probe and reference inputs do fire spontaneously at rates in the range 0–10 Hz (mean ∼5 Hz) directing attention to the probe or stimulus does lead to changes in firing rates of the target neurons (the Model neurons in the single-layer architecture, and the PYR5/6 nodes in the 3-layer structure) of up to 3 Hz, though the degree of change is dictated by the attentional modulation level. Indeed changes in neuronal responses caused by the effects of attention on noise were found in the experimental results (Williford & Maunsell, 2006).

A similar lack of sharpness also occurs for additive feedback attention when analysed for spiking neurons. The situation is not as clean here, since there are both non-linear as well as stochastic effects. However we have seen that there are differences in the spiking case between the various forms of attention by direct simulation. We have attempted to explore this further by use of mean field approximations (as in Deco and Rolls (2005), Brunel and Wang (2001), Wong and Wang (2006)) although we have not been able to reproduce the experimental results of Reynolds et al. (1999). Thus we left this approach aside; it is unclear to us that there would be an obvious reason why the additive approach would lead to results similar to the contrast gain case (unless some form of 'bump' activity was formed due to lateral connectivity).

The differences seen here between additive and contrast gain attention feedback in the graded case appear already to arise at the single neuron level, where a universal formula was obtained for the regression slope and constant (Taylor & Rogers, 2002). That for the regression slope was simple for contrast gain, being the value $\frac{1}{1+u}$, where $u = 1/5$, 1 and 5 for attend reference, away and to the probe, respectively. This formula was even simpler for the case of additive feedback, when all the slopes are identical, independent of attention mode. The extension of this universal formula to the 3-layered graded case or to the 1- or 3-layered cortex of spiking neurons is more difficult, but indicates the need for strong non-linearities or stochastic effects to get round the universal formula which found no effect of attention on the regression slope in the simplest case.

We can also ask under what conditions one form of attention feedback at a certain level in a hierarchical system can be modified to appear as another form as one proceeds up the hierarchy. Thus the modification of an additive component to a neural network with Mexican hat type of lateral interaction, with short range excitation and long-range inhibition, can lead to output in which the additive component becomes an output gain effect (Salinas & Abbott, 1997). It has been suggested that such an effect is at the basis of the gain modulation on neurons in parietal lobe by eye position. However for there to be such an effect arising from one layer to the next in either the ventral or dorsal visual hierarchy would require such a Mexican hat type of lateral connectivity. The existence of such connectivity has been proposed by several researchers (Amari, 1989; Ermentrout & Cowan, 1980) to allow for the creation of 'bump' localised persistent solutions for neural activity in the 'neural fields' suggested as approximations to the visual cortical areas such as $V1$, $V2$, $V4$ etc. However more detailed analyses of the responses of cortical cells in these areas have cast doubt on such connectivity and so on the mechanism of turning additive to output gain attention feedback signals.

There are a number of psychophysical results (Carrasco et al., 2004) which indicate that there could well be an output gain component arising at a global level in visual attention. Thus the problem of how this is to be understood still arises.

Also there is considerable visual feedback from one layer to a lower one in the visual hierarchy (Lamme, Super, & Spekreijse, 1998; Salin & Bullier, 1995). Much of this feedback is expected to be additive, but may not be directly

involved in attention feedback at all, but only in lower level visual processing up to semantic level, helping make it more efficient in its own right (Mumford, 1991a, 1991b). This is supported when we consider the neuromodulator most involved in attention, that being acetylcholine.

7.2. Conclusions

We have shown using different forms of neuron (graded and spiking) and various structures (single neurons and multi-layer cortical models) and mean field equations that the form of attention that best fits the $SE–SI$ graphs of Reynolds et al. (1999) from single-cell recordings is contrast gain; if we combine forms then a combination of contrast gain and additive attention give the best fit. A mathematical analysis confirmed the experimental results. We have also by investigating the literature suggested a tentative method by which the multiplicative affects of attention may occur, by utilising ACh and nicotinic receptors. We note that the conclusions of a recent modelling study (Sripati & Johnson, 2006) were that attention either worked by altering the neuronal firing threshold which can be achieved by a constant additive input, i.e. the additive attention feedback modelled throughout this paper, or by adaptation of the neuronal conductance where the number of open channels are modulated by some neuromodulator, i.e. the multiplicative form used for contrast gain. Indeed this second method is how we have previously shown multiplicative attention may act by the dependence of the number of open channels on the ACh concentration (Taylor et al., 2005) as a possible method of implementing the neurochemistry of attention outlined above. In all (Sripati & Johnson, 2006) investigated 6 possible attentional methods but did not look at combinations of 2 or more. We note also that the authors did not consider the data in Reynolds et al. (1999) at all; so the results are somewhat complementary.

More work on contrast gain attention is required since we cannot replicate the regression constant values of the single cell recordings reliably; with the 3-layer spiking cortical model, especially, having regression constant values that were too large, though the results for the single neuron spiking case were within the experimental range. The next stage is to investigate the effects and model of attention in a multiple region model, as in Deco and Rolls (2005).

Acknowledgements

NRT would like to thank the Engineering and Physical Sciences Research Council (EPRSC), UK; one of us (JGT) would like to thank the EC for financial support under the GNOSYS cognitive robot project (FP6-IST-2-003835) to allow him to pursue this work, and MH and JGT would like to thank the EC for financial support under the MATHESIS project.

References

Abbott, L. F., & Chance, E. S. (2005). Drivers and modulators from push–pull and balanced synaptic input. *Progress in Brain Research*, *149*, 147–155.

Alkondon, M., Pereira, E. F., Eisenberg, H. M., & Alburquerque, E. X. (2000). Nicotinic receptor activation in human cerebral cortical interneurons: A mechanism for inhibition and disinhibition of neuronal networks. *Journal of Neuroscience*, *20*, 66–75.

Amari, S. (1989). Dynamical stability of pattern formation of cortical maps. In M. A. Arbib, & S. Amari (Eds.), *Dynamic interactions in neural networks: Models and data* (pp. 15–34). New York: Springer.

Brunel, N., & Wang, X. -J. (2001). Effects of neuromodulation in a cortical network model of object working memory dominated by recurrent inhibition. *J. Computational Neuroscience*, *11*, 63–85.

Carrasco, M., Ling, S., & Read, S. (2004). Attention alters appearance. *Nature Neuroscience*, *7*, 308–313.

Davidson, M. T., & Marrocco, R. T. (2000). Local infusion of scopolamine into intraparietal cortex slows covert orienting in Rhesus monkeys. *Journal of Neurophysiology*, *83*, 1536–1549.

Deco, G., & Rolls, E. T. (2005). Neurodynamics of biased competition and cooperation for attention: A model with spiking neurons. *Journal of Neurophysiology*, *94*, 295–313.

Desimone, R., & Duncan, J. (1995). Neural mechanisms of selective visual attention. *Annual Review of Neuroscience*, *18*, 193–222.

Ermentrout, G. B., & Cowan, J. D. (1980). Large scale spatially organized activity in neural nets. *SIAM Journal of Applied Mathematics*, *38*, 1–21.

Everitt, B. J., & Robbins, T. W. (1997). Central cholinergic systems and cognition. *Annual Review of Psychology*, *48*, 649–684.

Foote, S. L., Berridge, C. W., Adams, L. M., & Pineda, J. A. (1991). Electrophysiological evidence for the involvement of the locus coeruleus in alerting, orienting, and attending. *Progress in Brain Research*, *88*, 531–532.

Fragopanagos, N., Kockelkoren, S., & Taylor, J. G. (2005). A neurodynamic model of the attentional blink. *Cognitive Brain Research*, *24*, 568–586.

Grossberg, S., & Raizada, R. D. (2000). Contrast-sensitive perceptual grouping and object-based attention in the laminar circuits of primary visual cortex. *Vision Research*, *40*(10–12), 1413–1432.

Hamker, F. H. (2004). Predictions of a model of spatial attention using sum- and max-pooling functions. *Neurocomputing*, *56*, 329–343.

Kimura, F. (2000). Cholinergic modulation of cortical function: A hypothetical role in shifting the dynamics in cortical network. *Neuroscience Research*, *38*, 19–26.

Lamme, V. A., Super, H., & Spekreijse, H. (1998). Feedforward, horizontal, and feedback processing in the visual cortex. *Current Opinion in Neurobiology*, *8*, 529–535.

Lucas-Meunier, E., Fossier, P., Baux, G., & Amar, M. (2003). Cholinergic modulation of the cortical neuronal network. *Pflugers Archiv*, *446*, 17–29.

Mentis, M. J., Sunderland, T., Lai, J., Connolly, C., Krasuski, J., Levine, B., et al. (2001). Muscarinic versus nicotinic modulation of a visual task: A PET study using drug probes. *Neuropsychopharmocology*, *25*, 555–563.

Mumford, D. (1991a). On the Computational Architecture of the Neo-Cortex I: The role of thalamo-cortical loops. *Biological Cybernetics*, *65*, 135–145.

Mumford, D. (1991b). On the computational architecture of the neo-cortex II: The role of cortico-cortical loops. *Biological Cybernetics*, *66*, 241–251.

Mozer, M. C., & Sitton, M. (1998). Computational modeling of spatial attention. In H. Pashler (Ed.), *Attention* (pp. 341–393). London: UCL Press.

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.

Salin, P. -A., & Bullier, J. (1995). Corticocortical connections in the visual system: Structure and function. *Physiological Reviews*, *75*(1), 107–154.

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

Smiley, J. F., Morrell, F., & Mesulam, M. M. (1997). Cholinergic synapses in the human cerebral cortex: An ultrastructural study in serial sections. *Experimental Neurology*, *144*, 361–368.

Sripati, A. P., & Johnson, K. O. (2006). Dynamic gain changes during attentional modulation. *Neural Computation*, *18*(6), 1847–1867.

Taylor, J. G., Hartley, M., & Taylor, N. R. (2005). Attention as Sigma-Pi controlled ACh-based feedback. In *Proc. of IJCNN'05*.

Taylor, J. G., & Rogers, M. (2002). A control model of the movement of attention. *Neural Networks*, *15*, 309–326.

Taylor, N. R., Panchev, C., Hartley, M., Kasderidis, S., & Taylor, J. G. (2006). Occlusion, attention and object representations. In S. Kolias, A. Stafylopatis, W. Duch, & E. Oja (Eds.), *Proc. ICANN* (pp. 592–601). Berlin: Springer.

Turrini, P., Casu, M. A., Wong, T. P., De Koninck, Y., Ribeiro-da-Silva, A., & Cuello, A. C. (2001). Cholinergic nerve terminal establish classical synapses in the rate cerebral cortex: Synaptic patterns and age-related atrophy. *Neuroscience*, *105*, 277–285.

Umbriaco, D., Watkins, K. C., Descarries, L., Cozzari, C., & Hartman, B. K. (1994). Ultrastructural and morphometric features of the acetylcholine innervation in adult rat parietal cortex: An electron microscope study. *Journal of Comparative Neurology*, *348*, 351–373.

Williford, T., & Maunsell, J. H. R. (2006). Effects of spatial attention on contrast response functions in macaque area *V*4. *Journal of Neurophysiology*, *96*, 40–54.

Witte, E. A., Davidson, M. C., & Marrocco, R. T. (1997). Effects of altering brain cholinergic activity on covert orienting of attention: Comparison of monkey and human performance. *Psychopharmocology*, *132*, 324–334.

Wong, K. -F., & Wang, X. -J. (2006). A recurrent network mechanism of time integration in perceptual decisions. *Journal of Neuroscience*, *26*(4), 1314–1328.

Ye, Y., Xi, W., Peng, Y., Wang, Y., & Guo, A. (2004). Long-term but not short-term blockade of dopamine release in Drosophila impairs orientation during flight in a visual attention paradigm. *European Journal of Neuroscience*, *20*, 1001–1007.