

# Self-organizing continuous attractor networks and path integration: one-dimensional models of head direction cells

S M Stringer, T P Trappenberg, E T Rolls<sup>1</sup> and I E T de Araujo

Oxford University, Department of Experimental Psychology, South Parks Road,  
Oxford OX1 3UD, UK

E-mail: Edmund.Rolls@psy.ox.ac.uk (Web page: [www.cns.ox.ac.uk](http://www.cns.ox.ac.uk))

Received 7 November 2001, in final form 1 February 2002

Published 30 April 2002

Online at [stacks.iop.org/Network/13/217](http://stacks.iop.org/Network/13/217)

## Abstract

Some neurons encode information about the orientation or position of an animal, and can maintain their response properties in the absence of visual input. Examples include head direction cells in rats and primates, place cells in rats and spatial view cells in primates. ‘Continuous attractor’ neural networks model these continuous physical spaces by using recurrent collateral connections between the neurons which reflect the distance between the neurons in the state space (e.g. head direction space) of the animal. These networks maintain a localized packet of neuronal activity representing the current state of the animal. We show how the synaptic connections in a one-dimensional continuous attractor network (of for example head direction cells) could be self-organized by associative learning. We also show how the activity packet could be moved from one location to another by idiothetic (self-motion) inputs, for example vestibular or proprioceptive, and how the synaptic connections could self-organize to implement this. The models described use ‘trace’ associative synaptic learning rules that utilize a form of temporal average of recent cell activity to associate the firing of rotation cells with the recent change in the representation of the head direction in the continuous attractor. We also show how a nonlinear neuronal activation function that could be implemented by NMDA receptors could contribute to the stability of the activity packet that represents the current state of the animal.

## 1. Introduction

Single-cell recording studies have revealed a number of classes of neurons which appear to encode the orientation or position of an animal with respect to its environment. Examples of

<sup>1</sup> Author to whom any correspondence should be addressed.

such classes of cells include head direction cells in rats (Ranck 1985, Taube *et al* 1990, 1996, Muller *et al* 1996) and primates (Robertson *et al* 1999), which respond maximally when the animal's head is facing in a particular preferred direction; place cells in rats (O'Keefe and Dostrovsky 1971, McNaughton *et al* 1983, O'Keefe 1984, Muller *et al* 1991, Markus *et al* 1995), that fire maximally when the animal is in a particular location, and spatial view cells in primates, that respond when the monkey is looking towards a particular location in space (Rolls *et al* 1997, Georges-François *et al* 1999, Robertson *et al* 1998). An important property of such classes of cells is that they can maintain their response properties when the animal is in darkness, with no visual input available to guide and update the firing of the cells. Moreover, when the animal moves in darkness, the spatial representation is updated by self-motion, that is idiothetic, cues. These properties hold for cells which indicate head direction in rats (Taube *et al* 1996) and macaques (Robertson *et al* 1999), for place cells in the rat hippocampus (O'Keefe 1976, McNaughton *et al* 1989, Quirk *et al* 1990, Markus *et al* 1994) and for spatial view cells in the macaque hippocampus (Robertson *et al* 1998). In this paper we consider how these properties, of idiothetic update in the dark, and of stability of firing in the dark, could arise. The particular model developed in this paper is for the head direction cell system, as this can be treated as a one-dimensional system. We extend this model to place cells in rats elsewhere (Stringer *et al* 2002).

An established approach to modelling the underlying neural mechanisms of head direction cells and place fields is 'continuous attractor' neural networks (CANNs) (see for example Skaggs *et al* (1995), Redish *et al* (1996), Zhang (1996), Redish and Touretzky (1998), Samsonovich and McNaughton (1997)). This class of network can maintain the firing of its neurons to represent any location along a continuous physical dimension such as head direction. These models use excitatory recurrent collateral connections between the neurons to reflect the distance between the neurons in the state space of the animal (e.g. head direction space). Global inhibition is used to keep the number of neurons in a bubble of activity relatively constant, and to help to ensure that there is only one activity packet. The properties of these CANNs have been extensively studied, for example by Amari (1977) and Taylor (1999). They can maintain the packet of neural activity (i.e. the set of active neurons that represent a spatial location) constant for long periods.

A key challenge in these CANN models is how the bubble of neuronal firing representing one location in the continuous state space can be updated based on non-visual, idiothetic, cues to represent a new location in state space. This is essentially the problem of path integration: how a system that represents a memory of where the animal (referred to more generically when modelled as an agent) is in physical space could be updated based on idiothetic cues such as vestibular cues (which might represent a head velocity signal), or proprioceptive cues (which might update a representation of place based on movements being made in the space, during for example walking in the dark). An approach introduced in recent models utilizes the concept of modulating the strength of the synaptic connection weights in one direction in the CANN to 'push' the activity packet in one direction, as the support for a packet is then asymmetric (for review see Zhang (1996)). Such an approach is closely related to dynamic remapping (Dominey and Arbib 1992, Pouget and Sejnowski 1995, Arbib 1997, Guazzelli *et al* 2001), and is the principal mechanism used to shift egocentric spatial representations of neural activity using velocity signals in the model of Droulez and Berthoz (1991).

A major limitation of the models studied so far is that the connections are pre-specified by the modeller, with for example two types of connection to deal with left and right spatial shifts. In this paper we introduce a way to learn by self-organization the correct connection strengths for the CANN to be updated by idiothetic inputs. We note that no simple shift across

a topographically organized map will work for many of the systems to which our model applies, such as the hippocampus, as these systems do not display topographical organization. (That is, nearby cells in the hippocampus do not represent nearby locations in space.)

We also consider the connection strengths between the neurons in the continuous attractor (which is a different issue to how the idiothetic inputs are learned). The idea that Hebbian learning can produce symmetric weight matrices within CANN models that enable stable representations to be formed has been stated and investigated by a number of authors. However, previous simulations of stable representations within CANN models have been where the recurrent synaptic weights are ideal, corresponding to regular and complete training in a completely connected network, thus ensuring that the synaptic weights are perfectly symmetrical in opposite directions in the state space. In this paper we introduce two methods to help with the problem of irregular training. One is the use of a short term memory trace in the Hebbian learning rule used to form the connections between the principal neurons in the continuous attractor (section 2). The memory trace helps to smooth out irregularities. The second method is to introduce a stabilizing influence on activity packets, using a nonlinear neuronal activation function (which could be realized by NMDA receptors) (section 4).

## 2. Self-organization of recurrent synaptic connections within a one-dimensional continuous attractor network of head direction cells

In this section we describe formally the model of a continuous attractor network that we investigate, and show how the weights could be trained, both by a Hebbian method previously discussed by Zhang (1996) and Redish and Touretzky (1998), and with Hebbian learning with a short term memory trace to help with the issue of irregular training, and diluted connectivity in the continuous attractor.

The generic model of a continuous attractor, which has been investigated extensively previously (Amari 1977, Zhang 1996), is as follows. The model is a recurrent attractor network with global inhibition. It is different from a Hopfield attractor network in that there are no discrete attractors formed by associative learning of discrete patterns. Instead there is a set of neurons that are connected to each other by synaptic weights  $w_{ij}^{\text{RC}}$  that are a simple function, for example Gaussian, of the distance between the states of the agent in the physical world (e.g. head directions) represented by the neurons. Neurons that represent similar states of the agent in the physical world have strong connections. The network updates its firing rates by the following ‘leaky-integrator’ dynamical equations. The continuously changing activation  $h_i^{\text{HD}}$  of each head direction cell  $i$  is governed by the equation

$$\tau \frac{dh_i^{\text{HD}}(t)}{dt} = -h_i^{\text{HD}}(t) + \frac{\phi_0}{C^{\text{HD}}} \sum_j (w_{ij}^{\text{RC}} - w^{\text{INH}}) r_j^{\text{HD}}(t) + I_i^{\text{V}}, \quad (1)$$

where  $r_j^{\text{HD}}$  is the firing rate of head direction cell  $j$ ,  $w_{ij}^{\text{RC}}$  is the excitatory (positive) synaptic weight from head direction cell  $j$  to cell  $i$ ,  $w^{\text{INH}}$  is a global constant describing the effect of inhibitory interneurons and  $\tau$  is the time constant of the system. The term  $I_i^{\text{V}}$  represents a visual input to head direction cell  $i$ . In the light, each term  $I_i^{\text{V}}$  is set to have a Gaussian response profile. The Gaussian assumption is not crucial, but is convenient. It is known that the firing rates of head direction cells in both rats (Taube *et al* 1996, Muller *et al* 1996) and macaques (Robertson *et al* 1999) are approximately Gaussian. In this paper we do not present a theory of how the visual inputs become connected to the head direction cells, but simply note that it is a property of head direction cells that they respond in the way described to visual inputs.

When the agent is in the dark, then the term  $I_i^V$  is set to zero<sup>2</sup>. The firing rate  $r_i^{\text{HD}}$  of cell  $i$  is determined from the activation  $h_i^{\text{HD}}$  and the sigmoid function

$$r_i^{\text{HD}}(t) = \frac{1}{1 + e^{-2\beta(h_i^{\text{HD}}(t) - \alpha)}}, \quad (2)$$

where  $\alpha$  and  $\beta$  are the sigmoid threshold and slope, respectively.

One way to set up the weights between the neurons in the continuous attractor network is to use an associative (Hebb-like) synaptic modification rule. In relation to biology, the idea here is that neurons close together in the state space (the space being represented) would tend to be co-active due to the large width of the firing fields. This type of rule has been used by for example Redish and Touretzky (1998). We describe this rule first, and then introduce an associative synaptic modification rule with a short term memory trace which has some advantages. The head direction cells are forced to respond during the initial learning phase by visual cues in the environment, which effectively force each cell to be tuned to respond best to a particular head direction, with less firing as the head direction moves away from the preferred direction. The simple associative learning rule is that the weights  $w_{ij}^{\text{RC}}$  from head direction cell  $j$  with firing rate  $r_j^{\text{HD}}$  to head direction cell  $i$  with firing rate  $r_i^{\text{HD}}$  are updated according to the following (Hebb-like) rule

$$\delta w_{ij}^{\text{RC}} = k r_i^{\text{HD}} r_j^{\text{HD}} \quad (3)$$

where  $\delta w_{ij}^{\text{RC}}$  is the change of synaptic weight and  $k$  is the learning rate constant.

This method works well with head direction cells that have Gaussian receptive fields which overlap partly with each other. In the simulations to be described, during the learning phase we set the firing rate  $r_i^{\text{HD}}$  of each head direction cell  $i$  to be the following Gaussian function of the displacement of the head from the preferred firing direction of the cell:

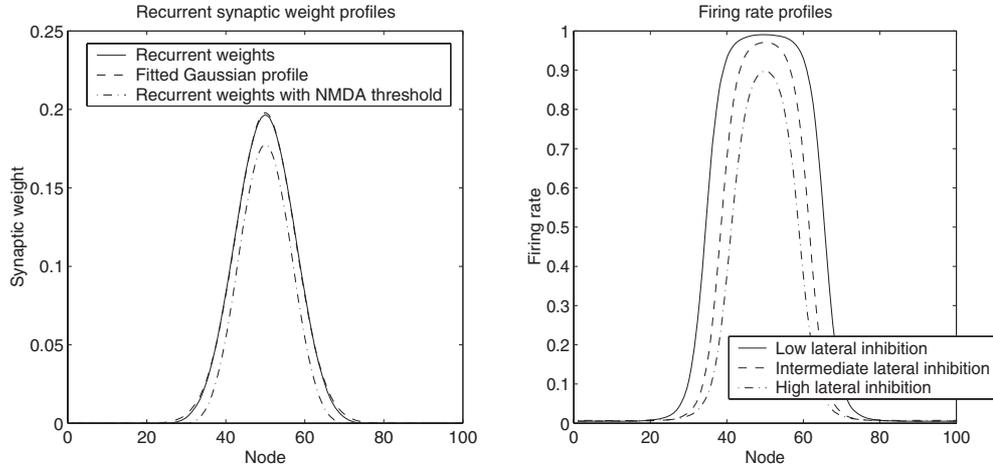
$$r_i^{\text{HD}} = e^{-(s_i^{\text{HD}})^2 / 2(\sigma^{\text{HD}})^2}, \quad (4)$$

where  $s_i^{\text{HD}}$  is the difference between the actual head direction  $x$  (in degrees) of the agent and the preferred head direction  $x_i$  for head direction cell  $i$ , and  $\sigma^{\text{HD}}$  is the standard deviation.  $s_i^{\text{HD}}$  is given by

$$s_i^{\text{HD}} = \text{MIN}(|x_i - x|, 360 - |x_i - x|). \quad (5)$$

To train the network, the agent is rotated through the full range of head directions, and at each head direction the weights are updated by equation (3). This results in nearby cells in head direction space, which need not be at all close to each other in the brain, developing stronger synaptic connections than cells that are more distant in head direction space. In fact, in the case described, the synaptic connections develop strengths which are almost, but not exactly, a Gaussian function of the distance between the cells in head direction space, as shown in figure 1 (left). Interestingly if a nonlinearity is introduced into the learning rule which mimics the properties of NMDA receptors by allowing the synapses to modify only after strong postsynaptic firing is present, then the synaptic strengths are still close to a Gaussian function of the distance between the connected cells in head direction space (see figure 1) (left). We show in simulations presented below that the network can support stable activity packets in the absence of visual inputs after such training.

<sup>2</sup> The scaling factor  $\phi_0 / C^{\text{HD}}$  controls the overall strength of the recurrent inputs to the continuous attractor network, where  $\phi_0$  is a constant and  $C^{\text{HD}}$  is the number of synaptic connections received by each head direction cell from other head direction cells. Scaling the recurrent inputs  $\sum_j (w_{ij}^{\text{RC}} - w_{ij}^{\text{NH}}) r_j^{\text{HD}}(t)$  by  $(C^{\text{HD}})^{-1}$  ensures that the overall magnitude of the recurrent input to each head direction cell remains approximately the same when the number of recurrent connections received by each head direction cell is varied. For a fully recurrently connected continuous attractor network,  $C^{\text{HD}}$  is equal to the total number of head direction cells,  $N^{\text{HD}}$ .



**Figure 1.** Numerical results of the regular training of the one-dimensional continuous attractor network of head direction cells with the Hebb rule equation (3) without weight normalization. Left: the learned recurrent synaptic weights from head direction cell 50 to the other head direction cells in the network arranged in head direction space, as follows. The first graph (solid curve) shows the recurrent synaptic weights learned with the standard Hebb rule equation (3). The second graph (dashed) shows a Gaussian curve fitted to the first graph; these two graphs are almost coincident, although the weights are not strictly Gaussian. The third graph (dash-dot) shows the recurrent synaptic weights learned with the standard Hebb rule equation (3), but with a nonlinearity introduced into the learning rule which mimics the properties of NMDA receptors by allowing the synapses to modify only after strong postsynaptic firing is present. Right: the stable firing rate profiles forming an activity packet in the continuous attractor network during the testing phase in the dark. The firing rates are shown after the network has been initially stimulated by visual input to initialize an activity packet, and then allowed to settle to a stable activity profile in the dark. The three graphs show the firing rates for low, intermediate and high values of the lateral inhibition parameter  $w^{\text{INH}}$ . For both left and right plots, the head direction cells are arranged according to where they fire maximally in the head direction space of the agent when visual cues are available.

A new hypothesis is now proposed for how the appropriate synaptic weights could be set up to deal with irregularities introduced into the synaptic weight connections by irregular training or by randomly diluted connectivity of the synaptic weights (as is present in the brain, Rolls and Treves (1998), Rolls and Deco (2002)). This hypothesis takes advantage of temporal probability distributions of firing when they happen to reflect spatial proximity. If we again consider the case of head direction cells, then the agent will necessarily move through similar head directions before reaching quite different head directions, and so the temporal proximity with which the cells fire can be used to set up the appropriate synaptic weights. The learning rule to utilize such temporal properties is then a trace learning rule which strengthens synaptic connections between neurons based on the temporal probability distribution of the firing. There are many versions of such rules (Rolls and Milward 2000, Rolls and Stringer 2001), but a simple one which works adequately is

$$\delta w_{ij}^{\text{RC}} = k \bar{r}_i^{\text{HD}} \bar{r}_j^{\text{HD}} \quad (6)$$

where  $\delta w_{ij}^{\text{RC}}$  is the change of synaptic weight, and  $\bar{r}^{\text{HD}}$  is a local temporal average or trace value of the firing rate of a head direction cell given by

$$\bar{r}^{\text{HD}}(t + \delta t) = (1 - \eta) r^{\text{HD}}(t + \delta t) + \eta \bar{r}^{\text{HD}}(t) \quad (7)$$

**Table 1.** Typical parameter values for models 1A and 1B.

$\sigma^{\text{HD}}$	$20^\circ$
Learning rate $k$	0.01
Learning rate $\tilde{k}$	0.01
Trace parameter $\eta$	0.9
$\tau$	1.0
$\phi_0$	400
$\phi_1$	400
$\phi_2$	400
$\gamma$	0.5
$\alpha^{\text{HIGH}}$	0.0
$\alpha^{\text{LOW}}$	-0.5
$\beta$	0.1

where  $\eta$  is a parameter set in the interval  $[0,1]$  which determines the contribution of the current firing and the previous trace. For  $\eta = 0$  the trace rule (6) becomes the standard Hebb rule (3), while for  $\eta > 0$  learning rule (6) operates to associate together patterns of activity that occur close together in time. In the simulations described later, it is shown that the main advantage of use of the trace rule (6) in the continuous attractor (with  $\eta > 0$ ) is that it produces a broader profile for the recurrent synaptic weights in the continuous attractor network than would be obtained with the standard Hebb rule (3), and thus broader firing fields (compare figures 1 and 2). Use of the trace rule thus allows broadly tuned head direction cells to emerge (in the light as well as in the dark) even if the visual cues used for initial training produce narrow tuning fields. Use of this rule helps also in cases of diluted connectivity and irregular training, as shown below in figure 3.

With irregular training, the magnitudes of the weights may also be uneven (as well as unequal in opposite directions). To bound the synaptic weights, weight decay can be used in the learning rule (Redish and Touretzky 1998, Zhang 1996). In the simulations described here with irregular training (figures 3, 10 and 12), weight normalization was used<sup>3</sup>.

Simulations to demonstrate the efficacy of both learning procedures are now described<sup>4</sup>. In the simulations, unless otherwise stated, there were 100 neurons in the fully connected network ( $C^{\text{HD}} = N^{\text{HD}} = 100$ ). (We did establish that similar results were obtained with different numbers of neurons in the network.) Typical model parameters for the simulations performed in this paper are given in table 1. Each head direction cell  $i$  was assigned a unique favoured head direction  $x_i$  from  $0^\circ$  to  $360^\circ$ , at which the cell was stimulated maximally by the available visual

<sup>3</sup> To implement weight normalization when it was used, after each time step of the learning phase, the recurrent synaptic weights between neurons within the continuous attractor network were rescaled to ensure that for each head direction cell  $i$  we have

$$\sqrt{\sum_j (w_{ij}^{\text{RC}})^2} = 1, \quad (8)$$

where the sum is over all head direction cells  $j$ . Such a renormalization process may be achieved in biological systems through synaptic weight decay (Oja 1982, Rolls and Treves 1998). The renormalization (8) helps to ensure that the learning rules are convergent in the sense that the recurrent synaptic weights between neurons within the continuous attractor network settle down over time to steady values.

<sup>4</sup> For the numerical integration of the differential equations (1) during the testing phase, we employ the ‘forward Euler’ time-stepping scheme

$$h_i^{\text{HD}}(t + \delta t) = \left(1 - \frac{\delta t}{\tau}\right) h_i^{\text{HD}}(t) + \frac{\delta t}{\tau} \frac{\phi_0}{C^{\text{HD}}} \sum_j (w_{ij}^{\text{RC}} - w^{\text{INH}}) r_j^{\text{HD}}(t) + \frac{\delta t}{\tau} I_i^{\text{V}}, \quad (9)$$

where the time step  $\delta t$  is set sufficiently small to provide a good approximation of the continuous dynamics.

cues during training<sup>5</sup>. We note that although a continuous attractor network is implemented by the synaptic connection strengths, this does not imply that there has to be any regular, maplike, topography in the physical arrangement of the neurons. During the testing phase, the dynamical equations (1) and (2) were implemented with  $\tau = 1$  and  $\phi_0/C^{\text{HD}} = 4$ . In addition, the lateral inhibition constant,  $w^{\text{INH}}$ , was set to half of the value of the maximum excitatory weight  $w_{ij}^{\text{RC}}$ .

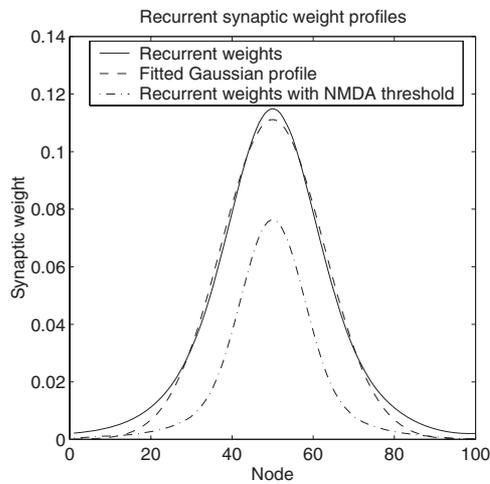
The results of regular training with the associative rule equation (3) are shown in figure 1. During regular training of the weights (without weight normalization, and with  $k = 0.01$ ), the agent was rotated twice through all 100 head directions  $x_i$  associated with the 100 head direction cells (once clockwise, and once anticlockwise). In figure 1 (right) the firing rates of the cells are shown when, after training, the network was activated by visual cues corresponding to a head direction of  $180^\circ$ . The visual cues were transiently applied for  $t = 0-25$ , and then the network was allowed to settle. The stable firing rates that were reached are shown at  $t = 600$  in the figure for three levels of the global inhibition. The values were  $w^{\text{INH}} = 0.3, 0.4$  and  $0.5 \times$  the maximum value of the recurrent synaptic weights, as indicated. The weights that were produced in the network are those illustrated in figure 1 (left).

The results of regular training with the trace associative rule equation (6) are shown in figure 2 (without weight normalization, with  $\eta = 0.9$  and with  $k = 0.01$ ). The magnitudes of the strengths of the connections from node (or neuron) 50 to the other neurons in the 100-neuron continuous attractor model are shown. The weights are broader (and thus also is the resulting width of the head direction cell tuning curve) than after training with the associative rule equation (3). This is the main advantage of using the trace rule in the continuous attractor, given that the tuning of head direction cells is broad (Taube *et al* 1996, Muller *et al* 1996, Robertson *et al* 1999).

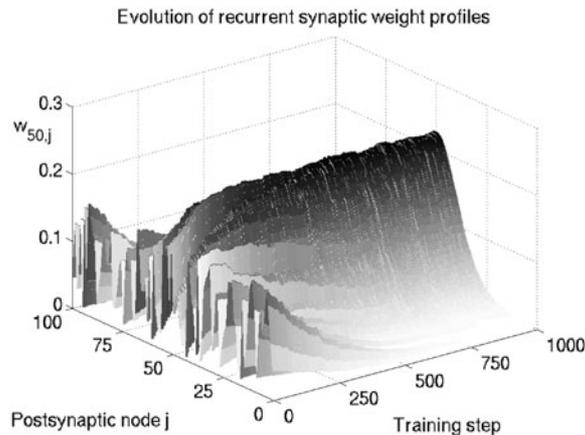
We also investigated the effects of training with a much less regular series of head directions than used so far. In the irregular training procedure investigated, at every training step, a new head direction was chosen relative to the current head direction from a normal distribution with mean zero and standard deviation  $90^\circ$ , and the agent stepped one node towards that direction at every time step. This procedure was repeated for 1000 training steps. The effect of this training on the recurrent weights, which initially had random values, is shown in figure 3. The connection strengths from presynaptic node (or neuron) 50 to other nodes are shown. The figure illustrates the effects of training with the trace rule equation (6) and  $\eta$  set to 0.9, and with weight renormalization using equation (8). The figure shows that training with the trace rule (6) combined with weight normalization (8) can lead over time to quite a smooth synaptic weight profile within the continuous attractor network, which is approximately Gaussian in shape. Furthermore, the weight profile obtained with the trace rule (6) is smoother than that obtained with the standard Hebb rule (3). This improves the stability of the activity packet within the continuous attractor network. However, the synaptic weights were insufficiently symmetrical to produce a completely stable bubble of neuronal activity, and instead the activity packet showed a slow drift as shown in figure 9 (left). As the system is expected to operate without drift, in section 4 we investigate ways in which the activity packet of neuronal activity can be stabilized in a continuous attractor network which does not have completely symmetrical synaptic weights in both directions round or along the attractor.

The conclusion from this section is that we have shown two self-organizing learning procedures that can set up the recurrent synaptic weights in a continuous attractor network of head direction cells to produce a stable bubble of firing which is maintained by a memory process when the visual cues are removed. (With the trace rule, care may be needed to allow

<sup>5</sup> Of course, in real nervous systems the directions for which individual head direction cells fire maximally would be randomly determined by processes of lateral inhibition and competition between neurons within the network of head direction cells.



**Figure 2.** Numerical results of the regular training of the one-dimensional continuous attractor network of head direction cells with the trace rule (6) without weight normalization. The plot shows the learned recurrent synaptic weights from head direction cell 50 to the other head direction cells in the network, as follows. The first graph (solid curve) shows the recurrent synaptic weights learned with the trace rule (6). The second graph (dashed) shows a Gaussian curve fitted to the first graph. The third graph (dash-dot) shows the recurrent synaptic weights learned with the trace rule (6), but with a nonlinearity introduced into the learning rule which mimics the properties of NMDA receptors by allowing the synapses to modify only after strong postsynaptic firing is present. The head direction cells are arranged in the graphs according to where they fire maximally in the head direction space of the agent when visual cues are available.



**Figure 3.** Numerical results of the irregular training of the one-dimensional continuous attractor network of head direction cells with the trace rule (6) and with weight normalization (8). The training consists of 1000 training steps, where for each training step a new head direction is chosen at random, and the agent is then rotated over a number of time steps to that new head direction with learning updates performed at each time step. The plot shows the time evolution of the recurrent synaptic weights from head direction cell 50 to the other head direction cells in the network. The head direction cells are arranged in the plot according to where they fire maximally in the head direction space of the agent when visual cues are available.

the trace value to build up from an initial value of zero during learning.) The trace rule can help to form broader synaptic weight distributions (and thus broader activity packets), and can operate well with irregular training.

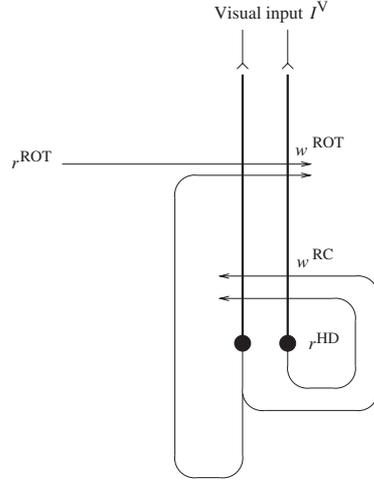
### 3. Continuous attractor models of head direction cells with idiothetic inputs: moving the activity packet by using idiothetic inputs

A key property of head direction cells is their ability to maintain their responses and update their firing based on motion cues when the animal is in complete darkness. One approach to simulating the movement of an activity packet produced by idiothetic cues (which is a form of path integration) is to employ a look-up table that stores for every possible head direction and rotational velocity the corresponding new direction in space (Samsonovich and McNaughton 1997). Another approach involves modulating the recurrent synaptic weights between head direction cells by idiothetic cues as suggested by Zhang (1996), though no possible biological implementation was proposed of how the appropriate dynamic synaptic weight changes might be achieved or how the connectivity could become self-organized. Another mechanism (Skaggs *et al* 1995) relies on a set of cells, termed rotation cells, which are co-activated by head direction cells and vestibular cells and drive the activity of the attractor network by anatomically distinct connections for clockwise and anticlockwise rotation cells. However, no proposal was made about how this could be achieved by a biologically plausible learning process. In order to achieve biological plausibility, the appropriate synaptic connections need to be self-organized by a learning process, and the aim of this section is to propose what the appropriate connections might be, and how they could self-organize.

Path integration is the ability of a system to continuously track and faithfully represent the time-varying head direction and position of a moving agent in the absence of visual input, using idiothetic inputs. We now present a continuous attractor model of head direction cells, model 1A, that is able to solve the problem of path integration in the dark through the incorporation of idiothetic inputs from sets of clockwise and anticlockwise rotation cells, and which develops its synaptic connectivity through self-organization. A closely related model, model 1B, is described in section 5. Although described in the context of head direction cells, the procedures are generic.

There is an initial learning phase with both visual and idiothetic inputs available. During this phase, the visual and idiothetic inputs work together to guide the self-organization of the network connectivity. An underlying assumption of the models is that, when visual cues are available to the agent, the visual inputs dominate other excitatory inputs to the head direction cells. In this case the head direction cells are stimulated by a particular arrangement of visual features in the environment, and hence by particular head directions. After the learning phase, the agent is then able to perform effective path integration in the absence of visual cues, with only idiothetic inputs available. The model (1A) consists of a recurrent continuous attractor network of head direction cells, which receives inputs from the visual system and from a population of head rotation cells. We denote the firing rate of each head rotation cell  $k$  by  $r_k^{\text{ROT}}$ . The population of head rotation cells includes both clockwise and anticlockwise rotation cells, which fire when the agent rotates either clockwise or anticlockwise respectively. The simplest network involves only a single clockwise rotation cell with  $k = 1$ , and a single anticlockwise rotation cell with  $k = 2$ . This is the network simulated below. The architecture of model 1A is shown in figure 4. There are two types of modifiable synaptic connection: (i) purely recurrent connections within the continuous attractor network of head direction cells, and (ii) idiothetic inputs to the head direction cells from the head rotation cells. The firing rates of the rotation cells are assumed to increase monotonically with the angular speed of the agent in the relevant direction (clockwise or anticlockwise for different head rotation cells), as described later.

The hypothesis that underlies model 1A is that the rotation cell firing interacts in a multiplicative way with connections in the continuous attractor using sigma-pi neurons, in such a way that clockwise rotation cells influence connections between neurons in the clockwise



**Figure 4.** General network architecture for one-dimensional continuous attractor models of head direction cells.

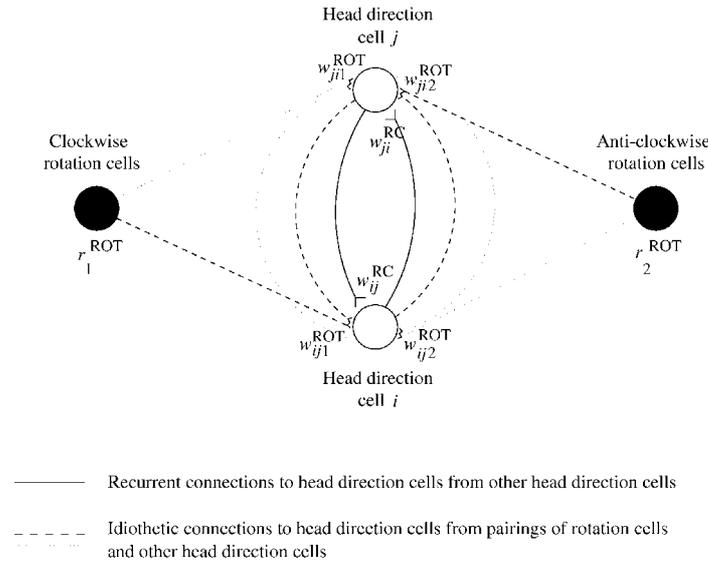
direction in the continuous attractor, and vice versa. Sigma-pi neurons sum the products of the contributions from two or more sources (see Rolls and Deco (2002), Koch (1999), section 21.1.1). A neural architecture that might implement such a system is shown in figure 5. In this figure there is a single clockwise rotation cell with firing rate  $r_1^{\text{ROT}}$ , and a single anticlockwise rotation cell with firing rate  $r_2^{\text{ROT}}$ . In addition, the idiothetic synaptic weights from the clockwise and anticlockwise rotation cells are denoted by  $w_{ij1}^{\text{ROT}}$  and  $w_{ij2}^{\text{ROT}}$  respectively. First, the connections shown by solid lines are the recurrent connections already described for the continuous attractor that can maintain its packet of activity. Second, there are sigma-pi synapses (or more accurately pi synapses) onto neurons in the continuous attractor which multiply together inputs from the rotation cells and from other head direction cells. Different head rotation cells (indexed by  $k$ ) could signal either clockwise or anticlockwise head rotation. The equations that follow describe how the system operates dynamically. How the appropriate connection strengths are learned is addressed in section 3.1. We note that the system could be realized in a number of different ways. For example, the connections that implement the sigma-pi synapses could also be used as the main recurrent connections in the continuous attractor. In this scenario, the recurrent synaptic connections in the continuous attractor would operate as described in section 2, but would have in addition a sigma-pi term that operates with respect to rotation cell inputs. With this regime, the solid connections shown in figure 5 would not be needed.

More formally, for each idiothetic synapse on a head direction cell, the synaptic input is generated by the product of the input from another cell in the continuous attractor, and the input from a rotation cell. Such effects might be produced in neurons by presynaptic terminals (see Koch 1999). The dynamical equation (1) governing the activations of the head direction cells is now extended to include inputs from the rotation cells in the following way. For model 1A, the activation of a head direction cell  $i$  is governed by the equation

$$\tau \frac{dh_i^{\text{HD}}(t)}{dt} = -h_i^{\text{HD}}(t) + \frac{\phi_0}{C^{\text{HD}}} \sum_j (w_{ij}^{\text{RC}} - w^{\text{INH}}) r_j^{\text{HD}}(t) + I_i^{\text{V}} + \frac{\phi_1}{C^{\text{HD} \times \text{ROT}}} \sum_{jk} w_{ijk}^{\text{ROT}} r_j^{\text{HD}} r_k^{\text{ROT}}, \quad (10)$$

where  $r_j^{\text{HD}}$  is the firing rate of head direction cell  $j$ ,  $r_k^{\text{ROT}}$  is the firing rate of rotation cell  $k$  and  $w_{ijk}^{\text{ROT}}$  is the corresponding overall effective connection strength. The first term on the right of

Synaptic connections for Sigma-Pi Model 1A



**Figure 5.** Recurrent and idiothetic synaptic connections to head direction cells in the sigma-pi model 1A. In this figure there is a single clockwise rotation cell with firing rate  $r_1^{\text{ROT}}$ , and a single anticlockwise rotation cell with firing rate  $r_2^{\text{ROT}}$ . In addition, the idiothetic synaptic weights from the clockwise and anticlockwise rotation cells are denoted by  $w_{ij1}^{\text{ROT}}$  and  $w_{ij2}^{\text{ROT}}$  respectively.

equation (10) is a decay term, the second describes the effects of the recurrent connections in the continuous attractor, the third is the visual input (if present) and the fourth represents the effects of the idiothetic connections implemented by sigma-pi synapses<sup>6</sup>. (We note that  $\phi_1$  would need to be set in the brain to have a magnitude which allows the actual head rotation cell firing to move the activity packet at the correct speed, and that this gain control has some similarity to the type of gain control that the cerebellum is believed to implement for the vestibulo-ocular reflex, see Rolls and Treves (1998).) Thus, there are two types of synaptic connection to head direction cells: (i) recurrent synapses from head direction cells to other head direction cells within the recurrent network, whose effective strength is governed by the terms  $w_{ij}^{\text{RC}}$ , and (ii) idiothetic sigma-pi synapses dependent upon the interaction between an input from another head direction cell and a rotation cell, whose effective strength is governed by the terms  $w_{ijk}^{\text{ROT}}$ . At each time step, once the head direction cell activations  $h_i^{\text{HD}}$  have been updated, the head direction cell firing rates  $r_i^{\text{HD}}$  are calculated according to the sigmoid transfer function (2). Therefore, the initial learning phase involves the setting up of the synaptic weights  $w_{ij}^{\text{RC}}$  and  $w_{ijk}^{\text{ROT}}$ .

<sup>6</sup> The scaling factor  $\phi_1/C^{\text{HD}\times\text{ROT}}$  controls the overall strength of the idiothetic inputs from the rotation cells, where  $\phi_1$  is a constant, and the term  $C^{\text{HD}\times\text{ROT}}$  is defined as follows.  $C^{\text{HD}\times\text{ROT}}$  is the number of idiothetic connections received by each head direction cell from couplings of individual head direction cells and rotation cells. Scaling the idiothetic inputs  $\sum_{jk} w_{ijk}^{\text{ROT}} r_j^{\text{HD}} r_k^{\text{ROT}}$  by the term  $(C^{\text{HD}\times\text{ROT}})^{-1}$  ensures that the overall magnitude of the idiothetic input to each head direction cell remains approximately the same as the number of idiothetic connections received by each head direction cell is varied. For a fully connected network,  $C^{\text{HD}\times\text{ROT}}$  is equal to the number of head direction cells,  $N^{\text{HD}}$ , times the number of rotation cells,  $N^{\text{ROT}}$ .

### 3.1. Self-organization by learning of synaptic connectivity from idiothetic inputs to the continuous attractor network of head direction cells

In this section we describe how the connections from the idiothetic inputs to the head direction cells can self-organize such that, after the initial learning phase, the idiothetic inputs can correctly shift activity packets from one location to another in the continuous attractor network of head direction cells in the absence of visual cues. The problem to be solved is how, in the absence of visual cues, the clockwise vestibular inputs might move the activity packet in one direction, and by the correct amount, in the continuous attractor, and how the anticlockwise vestibular inputs might move the activity packet the correct amount in the opposite direction. We know of no previous suggestions of how this could be achieved in a biologically plausible way.

The overall hypothesis of how this is achieved in the first model, 1A, is as follows. In the learning phase, when the agent is turning clockwise, the firing of the clockwise rotation cells ( $r_1^{\text{ROT}}$ ) is high (see figure 5). The idiothetic sigma-pi connections ( $w_{ij1}^{\text{ROT}}$ ) to the head direction cells in the clockwise direction in the continuous attractor network are strengthened by an associative learning rule that associates the product of the firing of the clockwise rotation cell ( $r_1^{\text{ROT}}$ ), and the trace ( $\bar{r}_i^{\text{HD}}$ ) of the recent activity of presynaptic head direction cells which is accumulated in the idiothetic synaptic connection, with the current postsynaptic head direction cell firing ( $r_i^{\text{HD}}$ ). The trace enables the idiothetic synapses in the correct direction between cells in the continuous attractor to be selected.

In the models presented here it is assumed that when visual cues are available during the learning phase, the visual inputs to head direction cells dominate the other excitatory inputs, and individual head direction cells fire according to the Gaussian response profile (4). However, even with visual information available, the idiothetic inputs still have a critical role to play in the setting up of the synaptic connections from the rotation cells to the continuous attractor network. In models 1A and 1B, during the initial learning phase with visual input available, the inputs from the visual and vestibular systems are able to work together to guide the self-organization of the network synaptic connectivity using simple biologically plausible learning rules. A common feature of both models 1A and 1B is their reliance on temporal ‘trace’ learning rules for the connections from the rotation cells, that utilize a form of temporal average of recent cell activity. A key property of these types of learning rule is their ability to build associations between different patterns of neural activities that tend to occur in temporal proximity. In the models presented here, trace learning is able to associate, for example, the co-firing of a set of rotation cells (which correspond to vestibular cells and have firing rates that depend on head angular velocity) and an earlier activity pattern in the recurrent network reflecting the previous head direction of the agent, with cells within the recurrent network reflecting the current head direction of the agent.

During the initial learning phase, the response properties of the head direction and rotation cells are set as follows. While the agent is rotated both clockwise and anticlockwise during learning, the visual input drives the cells in the recurrent network of head direction cells as described above. That is, as the head direction of the agent moves away from the preferred direction for the cell, the firing rate  $r_i^{\text{HD}}$  of cell  $i$  is set according to the Gaussian profile (4). While the agent is undergoing rotation, the rotation cells fire according to whether the agent is rotating in the appropriate direction, and with a firing rate that increases monotonically with respect to speed of rotation. Specifically, in the simulations performed later the firing rates of the rotation cells are set in the following simple manner. During the learning phase the agent is rotated alternately in clockwise and anticlockwise directions at a constant speed. Therefore, during this phase we set  $r_1^{\text{ROT}}$  to 1 when the agent is rotating in the clockwise direction, and 0 otherwise. Similarly, we set  $r_2^{\text{ROT}}$  to 1 when the agent is rotating in the anticlockwise

direction, and 0 otherwise. Then, during the subsequent testing phase, the firing rates of the rotation cells  $r_1^{\text{ROT}}$  and  $r_2^{\text{ROT}}$  are varied to simulate different rotation speeds of the agent, and hence produce different translation speeds of the activity packet within the continuous attractor network of head direction cells.

For model 1A, the initial learning phase involves the setting up of the synaptic weights  $w_{ijk}^{\text{ROT}}$ . At the start of the learning phase the synaptic weights  $w_{ijk}^{\text{ROT}}$  may be initialized to either zero or some random positive values. Next, the learning phase proceeds with the agent rotating in the light, with the firing rates of the head direction and rotation cells set as described above, with the synaptic weights  $w_{ijk}^{\text{ROT}}$  updated at each time step according to

$$\delta w_{ijk}^{\text{ROT}} = \tilde{\kappa} r_i^{\text{HD}} \bar{r}_j^{\text{HD}} r_k^{\text{ROT}} \quad (11)$$

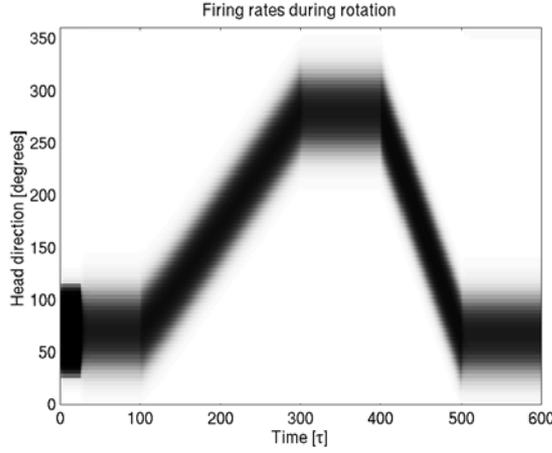
where  $\delta w_{ijk}^{\text{ROT}}$  are the changes in the synaptic weights, and where  $r_i^{\text{HD}}$  is the instantaneous firing rate of the postsynaptic head direction cell  $i$ ,  $\bar{r}_j^{\text{HD}}$  is the trace value of the presynaptic head direction cell  $j$  given by equation (7),  $r_k^{\text{ROT}}$  is the firing rate of rotation cell  $k$  and  $\tilde{\kappa}$  is the learning rate associated with this type of synaptic connection. The essence of this learning process is that when the activity packet has moved, say, in a clockwise direction in the recurrent attractor network of head direction cells, the trace term ‘remembers’ the direction in which the head direction cells have been activated, and the result of  $r_i^{\text{HD}} \bar{r}_j^{\text{HD}}$  is used in combination with the firing  $r_1^{\text{ROT}}$  of the clockwise rotation cell to modify the synaptic weights  $w_{ij1}^{\text{ROT}}$ . This should typically lead to one of the two types of weight  $w_{ij1}^{\text{ROT}}$  or  $w_{ij2}^{\text{ROT}}$  becoming significantly larger than the other for any particular pair of head direction cells  $i$  and  $j$ . For example, if we consider two head direction cells  $i$  and  $j$  that fire maximally for similar head directions, then during the learning phase cell  $i$  may often fire a short time after cell  $j$  depending on which direction the agent is rotating in as it passes through the head direction associated with head direction cell  $j$ . In this situation the effect of the above learning rule would be to ensure that the size of the weights  $w_{ij1}^{\text{ROT}}$  and  $w_{ij2}^{\text{ROT}}$  for the connection  $i, j$  would be largest in the rotational direction (clockwise or anticlockwise) in which cell  $i$  is a short distance from cell  $j$ . The effect of the above learning rule for the synaptic weights  $w_{ij1}^{\text{ROT}}$  and  $w_{ij2}^{\text{ROT}}$  is to generate a synaptic connectivity such that the firing of one of the two classes of rotation cells (clockwise or anticlockwise) should increase the activations  $h_i^{\text{HD}}$  of head direction cells  $i$ , where neurons  $i$  represent head directions that are a small rotation in the appropriate clockwise or anticlockwise direction from that represented by currently active neurons  $j$ . Thus, the co-firing of a particular type of rotation cell (clockwise or anticlockwise) and set of head direction cells representing a particular head direction, should stimulate the firing of further head direction cells such that the pattern of activity within the network of head direction cells evolves continuously to faithfully reflect and track the changing state of the agent<sup>7</sup>.

*3.1.1. Simulation results demonstrating self-organization of idiothetic inputs to head direction cells in continuous attractor network.* The results of the simulation of model 1A with

<sup>7</sup> In order to achieve a convergent learning scheme for model 1A, after each time step the idiothetic synaptic weights  $w_{ijk}^{\text{ROT}}$  may be renormalized by ensuring that for each head direction cell  $i$  we have

$$\sqrt{\sum_j (w_{ijk}^{\text{ROT}})^2} = 1, \quad (12)$$

where such a renormalization is performed separately for each rotation cell  $k$ . The effect of such a renormalization procedure is to ensure that the learning rules are convergent in the sense that the synaptic weights  $w_{ijk}^{\text{ROT}}$  settle to steady values over time.



**Figure 6.** Numerical results for model 1A with sigma-pi neurons after regular training with the Hebb rule (3) for the recurrent connections within the continuous attractor, trace rule (11) for the idiothetic connections and without weight normalization. The plot shows the shift in the activity packet in the continuous attractor network of head direction cells as the agent rotates clockwise and anticlockwise in the dark. The shift is effected by idiothetic inputs to the continuous attractor network from the clockwise and anticlockwise rotation cells. The plot shows the firing rates in the continuous attractor network of head direction cells through time, with the head direction cells arranged in the plot according to where they fire maximally in the head direction space of the agent when visual cues are available.

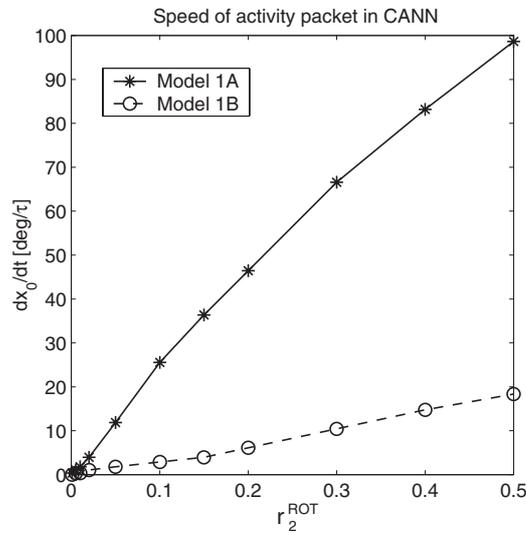
sigma-pi neurons are shown in figures 6–8<sup>8</sup>. (The regular learning regime was used with the associative learning rule equation (3), and the weights were not normalized.) Figure 6 shows the head direction represented in the continuous attractor network in response to rotation cell inputs from the vestibular system. That is, the results shown are for idiothetic inputs in the absence of visual cues. The activity packet was initialized as described previously to a head direction of 75°, and the packet was allowed to settle without visual input. For  $t = 0–100$  there was no rotation cell input, and the activity packet in the continuous attractor remained stable at 75°. For  $t = 100–300$  the clockwise rotation cells were active with a firing rate of 0.15 to represent a moderate angular velocity, and the activity packet moved clockwise. For  $t = 300–400$  there was no rotation cell firing, and the activity packet immediately stopped, and remained still. For  $t = 400–500$  the anticlockwise rotation cells had a high firing rate of 0.3 to represent a high velocity, and the activity packet moved anticlockwise with a greater velocity. For  $t = 500–600$  there was no rotation cell firing, and the activity packet immediately stopped.

<sup>8</sup> For the numerical integration of the differential equations (10) of model 1A we employ the ‘forward Euler’ time-stepping scheme

$$h_i^{\text{HD}}(t + \delta t) = \left(1 - \frac{\delta t}{\tau}\right) h_i^{\text{HD}}(t) + \frac{\delta t}{\tau} \frac{\phi_0}{C^{\text{HD}}} \sum_j (w_{ij}^{\text{RC}} - w^{\text{INH}}) r_j^{\text{HD}}(t) + \frac{\delta t}{\tau} I_i^{\text{V}} + \frac{\delta t}{\tau} \frac{\phi_1}{C^{\text{HD} \times \text{ROT}}} \sum_{jk} w_{ijk}^{\text{ROT}} r_j^{\text{HD}} r_k^{\text{ROT}}, \quad (13)$$

where the time step  $\delta t$  is set sufficiently small to provide a good approximation of the continuous dynamics. Similarly, for the numerical integration of the differential equations (18) of model 1B, we employ the time-stepping scheme

$$h_i^{\text{HD}}(t + \delta t) = \left(1 - \frac{\delta t}{\tau}\right) h_i^{\text{HD}}(t) + \frac{\delta t}{\tau} \frac{\phi_0}{C^{\text{HD}}} \sum_j (\tilde{w}_{ij}^{\text{RC}} - w^{\text{INH}}) r_j^{\text{HD}}(t) + \frac{\delta t}{\tau} I_i^{\text{V}}. \quad (14)$$

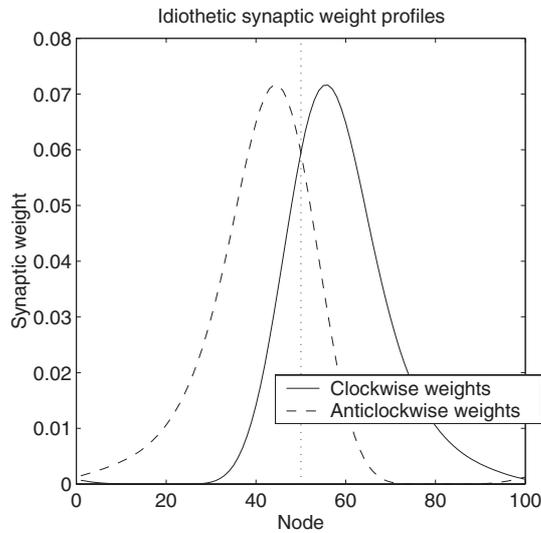


**Figure 7.** Numerical results for models 1A and 1B after regular training with the Hebb rule (3) for the recurrent connections in the continuous attractor, trace rule (11) for the idiothetic connections and without weight normalization. The plot shows the speed of the activity packet in the continuous attractor network of head direction cells for different strengths of the idiothetic input  $r_2^{\text{ROT}}$ , the firing rate of the anticlockwise rotation cell. The speed plotted is the rate of change of the position (in degrees) of the activity packet in the head direction space of the agent with time. The first graph is for model 1A with sigma-pi neurons, and the second graph is for model 1B which relies on modulation of the recurrent weights within the continuous attractor network.

Figure 7 shows that the velocity of the head direction cell activity packet is a nearly linear function of the rotation cell firing, at least within the region shown. Separate curves are shown for the sigma-pi model (1A) and the synaptic modulation model (1B). The results shown in this figure show that the rate of change (velocity) of the represented head direction in the continuous attractor generalizes well to velocities at which the system was not trained. The results show that both models 1A and 1B operate as expected.

Figure 8 shows the synaptic weights from the clockwise and anticlockwise rotation cells to the continuous attractor nodes in model 1A. The left graph is for the connections from the pairing of the anticlockwise rotation cell and head direction node 50 to the other head direction cells  $i$  in the network, that is  $w_{i,50,2}^{\text{ROT}}$ . The right graph is for the connections from the pairing of the clockwise rotation cell and head direction node 50 to the other head direction cells  $i$  in the network, that is  $w_{i,50,1}^{\text{ROT}}$ . The graphs show that, *per hypothesem*, the connections from the anticlockwise rotation cell have self-organized in such a way that they influence more the head direction cells in the anticlockwise direction, as shown by the offset of the graph from node 50. In model 1B, the learned modulation factors  $\lambda_{ijk}^{\text{ROT}}$  are identical to the corresponding weights  $w_{ijk}^{\text{ROT}}$  learned in the sigma-pi model.

Figures 6–8 thus show that self-organization by learning in the ways proposed in models 1A and 1B does produce the correct synaptic connections to enable idiothetic inputs from for example head rotation cell firing to correctly move the activity packet in the continuous attractor. The correct connections are learned during a period of training in the light, when both visual and idiothetic inputs are present. Thus the two new models proposed in this paper provide plausible models for how path integration could be learned in the brain.

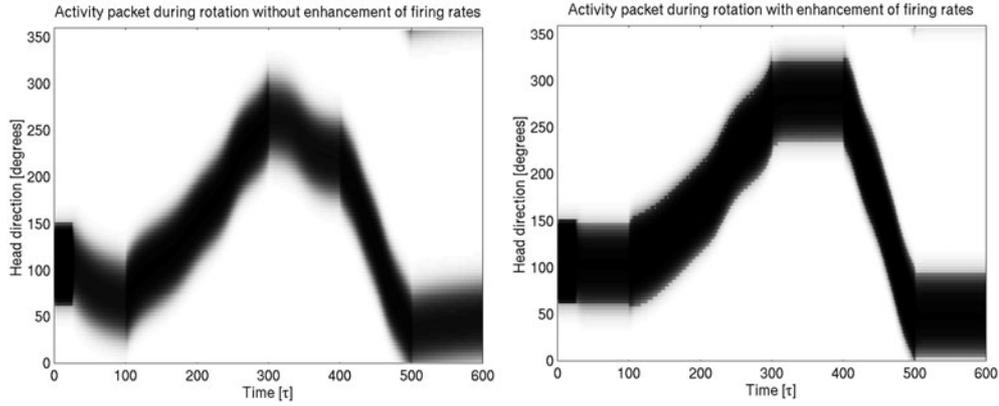


**Figure 8.** Numerical results for model 1A with sigma-pi neurons after regular training with the Hebb rule (3) for the recurrent connections within the continuous attractor, trace rule (11) for the idiopathic connections and without weight normalization. The plot shows the learned idiopathic synaptic weights from the clockwise and anticlockwise rotation cells to the continuous attractor network of head direction cells. The first graph shows the learned idiopathic synaptic weights  $w_{i,50,1}^{\text{ROT}}$  from the coupling of the clockwise rotation cell and head direction cell 50, to the other head direction cells  $i$  in the network. The second graph shows the learned idiopathic synaptic weights  $w_{i,50,2}^{\text{ROT}}$  from the coupling of the anticlockwise rotation cell and head direction cell 50, to the other head direction cells  $i$  in the network. The head direction cells are arranged in the graphs according to where they fire maximally in the head direction space of the agent when visual cues are available.

#### 4. Stabilization of the activity packet within the continuous attractor network when the agent is stationary

With irregular learning conditions (in which identical training with high precision of every node cannot be guaranteed), the recurrent synaptic weights between nodes in the continuous attractor will not be of the perfectly regular form normally required in a CANN. This can lead to drift of the activity packet within the continuous attractor network of head direction cells when no visual cues are present, even when the agent is not moving. This is evident in figure 9. In this section we discuss two alternative approaches to stabilizing the activity packet when it should not be drifting in real nervous systems.

A first way in which the activity packet may be stabilized within the continuous attractor network of head direction cells when the agent is stationary is by enhancing the firing of those cells that are already firing. In biological systems this may be achieved through mechanisms for short term synaptic enhancement (Koch 1999). Another way is to take advantage of the nonlinearity of the activation function of neurons with NMDA receptors, which only contribute to neuronal firing once the neuron is sufficiently depolarized (Wang 1999). The effect is to enhance the firing of neurons that are already reasonably well activated. The effect has been utilized in a model of a network with recurrent excitatory synapses which can maintain active an arbitrary set of neurons that are initially sufficiently strongly activated by an external stimulus (see Lisman *et al* (1998), but see also Kesner and Rolls (2001)). In the head direction cell models, we simulate such biophysical processes by adjusting the sigmoid threshold  $\alpha_i$  for each head direction cell  $i$  as follows. If the head direction cell firing rate  $r_i^{\text{HD}}$  is lower than a threshold value,  $\gamma$ , then the sigmoid threshold  $\alpha_i$  is set to a relatively high value  $\alpha^{\text{HIGH}}$ .



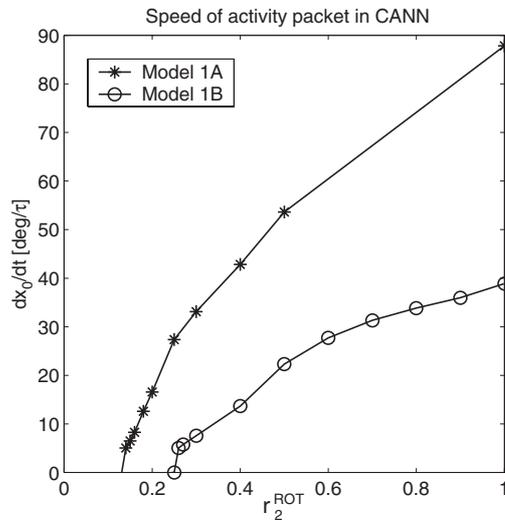
**Figure 9.** Left: numerical results for model 1A with the irregular learning regime using the trace rule (6) for the recurrent connections within the continuous attractor, trace rule (11) for the idiothetic connections and with the weight normalization (8) and (12). The simulation was performed without enhancement of firing of head direction cells that are already highly activated according to equation (15). This plot is similar to figure 6, where the shift in the activity packet in the continuous attractor network of head direction cells is shown as the agent rotates clockwise and anticlockwise in the dark. The shift is effected by idiothetic inputs to the continuous attractor network from the clockwise and anticlockwise rotation cells. The plot shows the firing rates in the continuous attractor network of head direction cells through time, with the head direction cells arranged in the plot according to where they fire maximally in the head direction space of the agent when visual cues are available. The simulation was performed without NMDA-like nonlinearity in the neuronal activation function, and some drift of the activity packet when it should be stable is evident. Right: a similar simulation to that shown on the left, except that it was with enhancement of firing of head direction cells that are already highly activated according to equation (15). No drift of the activity packet was present when it should be stationary.

Otherwise, if the head direction cell firing rate  $r_i^{\text{HD}}$  is greater than or equal to the threshold value,  $\gamma$ , then the sigmoid threshold  $\alpha_i$  is set to a relatively low value  $\alpha^{\text{LOW}}$ . This is achieved in the numerical simulations by resetting the sigmoid threshold  $\alpha_i$  at each time step depending on the firing rate of head direction cell  $i$  at the previous time step. That is, at each time step  $t + \delta t$  we set

$$\alpha_i = \begin{cases} \alpha^{\text{HIGH}} & \text{if } r_i^{\text{HD}}(t) < \gamma \\ \alpha^{\text{LOW}} & \text{if } r_i^{\text{HD}}(t) \geq \gamma \end{cases} \quad (15)$$

where  $\gamma$  is a firing rate threshold. The sigmoid slopes were set to a constant value,  $\beta$ , for all cells  $i$ . This procedure has the effect of enhancing the current position of the activity packet within the continuous attractor network, and so prevents the activity packet drifting erratically due to the noise in the recurrent synaptic weights, as illustrated in figure 9 (left). The effect can also be seen in the threshold nonlinearity introduced into the relation between the velocity of the activity packet and the idiothetic input signal (figure 10).

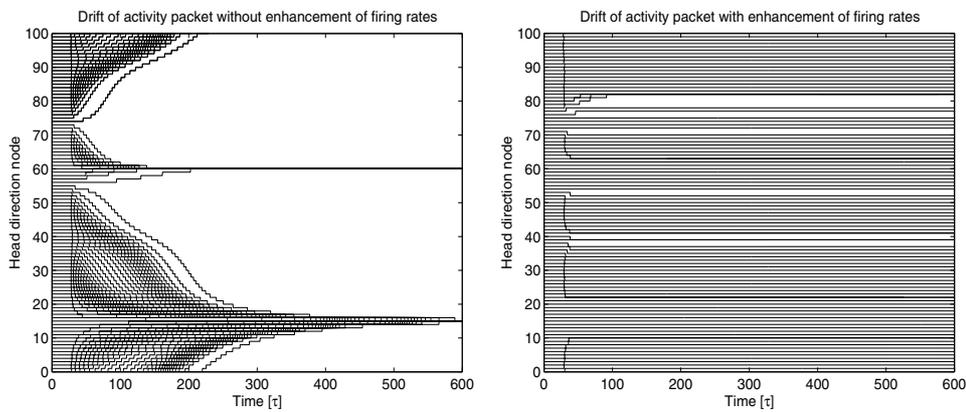
The effects of using the additional nonlinearity in the activation function of the neurons in the continuous attractor models 1A and 1B is illustrated in figures 9 and 10. For the simulations shown in figure 9, irregular learning was used with the trace rule equation (6) with  $\eta = 0.9$ , and with weight normalization. Figure 9 (left) shows the results when running without the NMDA nonlinearity (i.e. with  $\alpha^{\text{HIGH}} = \alpha^{\text{LOW}} = 0$ ), and figure 9 (right) shows the results with the NMDA nonlinearity for which  $\gamma$  was 0.5,  $\alpha^{\text{HIGH}}$  was 0.0 and  $\alpha^{\text{LOW}}$  was  $-5.0$ . In figure 9 (left) the activity packet was initialized as described previously to a head direction of  $108^\circ$ , and the packet was allowed to settle without visual input. For  $t = 0-100$  there was no rotation cell input, and the activity packet in the continuous attractor showed some drift. For  $t = 100-300$



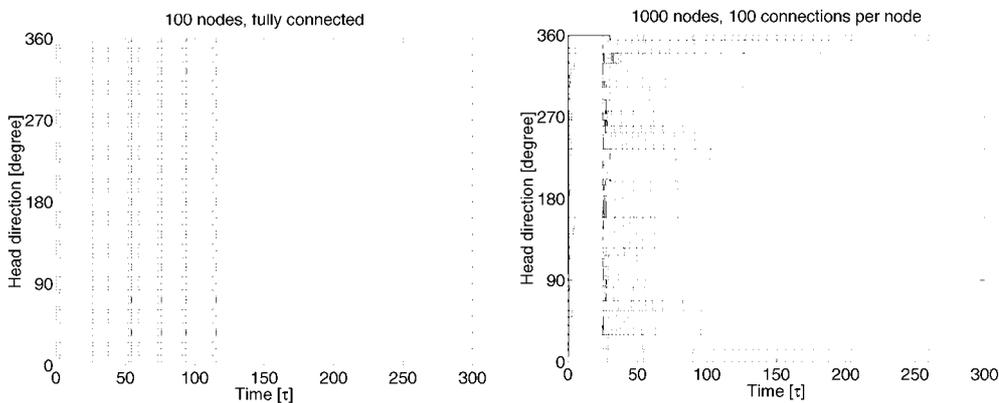
**Figure 10.** Numerical results for models 1A and 1B with the regular learning regime using the Hebb rule (3) for the recurrent connections within the continuous attractor, trace rule (11) for the idiothetic connections and without weight normalization. The plot shows the speed of the activity packet in the continuous attractor network of head direction cells for different strengths of the idiothetic input  $r_2^{\text{ROT}}$ , the firing rate of the anticlockwise rotation cell. This plot is similar to figure 7, except that in the simulations the firing of head direction cells that are already highly activated is enhanced according to equation (15). The speed plotted is the rate of change of the position (in degrees) of the activity packet in the head direction space of the agent with time. The first graph is for model 1A with sigma-pi neurons, and the second graph is for model 1B which relies on modulation of the recurrent weights within the continuous attractor network.

the clockwise rotation cells were active with a firing rate of 0.095 to represent a moderate angular velocity, and the activity packet moved clockwise. For  $t = 300\text{--}400$  there was no rotation cell firing, and the activity packet showed some drift. For  $t = 400\text{--}500$  the anticlockwise rotation cells had a firing rate of 0.08, and the activity packet moved anticlockwise. For  $t = 500\text{--}600$  there was no rotation cell firing, and the activity packet showed a little drift. In figure 9 (right) it is shown that the drift is eliminated when the NMDA nonlinearity is used, while the effects of the rotation cell firing still operate. (The testing conditions used for the right part of the figure were the same as those on the left, except that the clockwise firing rate was 0.135 and the anticlockwise firing rate was 0.16.) When we inspect figure 10 we see that a threshold nonlinearity is introduced into the relation between the velocity with which the activity packet moves along the continuous attractor, and the rotation cell firing rate (cf figure 7). The nonlinearity reflects the fact that the head direction cell firing tends to be maintained in a fixed population of continuous attractor neurons by their nonlinear activation function.

An advantage of using the nonlinearity in the activation function of a neuron (produced for example by the operation of NMDA receptors) is that this tends to enable packets of activity to be kept active without drift even when the packet is not in one of the energy minima that can result from irregular learning (or from diluted connectivity in the continuous attractor as described below). This is illustrated by the fact that after the irregular learning regime described above, there tends to be a small number of stable locations for the continuous attractor activity packet, as shown in figure 11 (left). When the NMDA nonlinearity is used with the parameters described above, we see from figure 11 (right) that the number of stable states in the continuous attractor is higher. Thus, use of this nonlinearity increases the number of locations in the continuous physical state space at which a stable activity packet can be maintained.



**Figure 11.** Numerical results for model 1A with the irregular learning regime using the trace rule (6) for the recurrent connections within the continuous attractor, trace rule (11) for the idiotic connections and with the weight normalization (8) and (12). The plot shows the drift of the activity packet within the continuous attractor network of head direction cells when the visual input is removed, and the agent itself remains stationary (i.e. not rotating). Both left and right plots show many time courses of the position of the activity packet within the continuous attractor network in the head direction space of the agent for different initial locations. The left plot shows results without enhancement of the firing of head direction cells that are already highly active, while the right plot shows results with enhancement of the firing of head direction cells that are already highly active according to equation (15).



**Figure 12.** Effects of undertraining. Left: a sigma-pi network (model 1A) with 100 recurrent connections per neuron in the fully connected continuous attractor ( $C^{\text{HD}} = 100$ ) was trained with ten different equispaced head directions. The simulations were run without the NMDA nonlinearity in the activation functions in the neurons of the continuous attractor. It was found that the network settled in the correct head direction state from among the ten head direction states originally trained when stimulated (from time  $t = 0-50$ ) with a visual input corresponding to any head direction. Right: a similar experiment, but with 10% connectivity in the continuous attractor network with 100 recurrent connections per neuron. A number of stable states were found, although there were fewer than the number trained.

A second way in which real nervous systems might overcome the drift that may be produced in continuous attractor networks by irregular learning, or by diluted connectivity in the recurrent attractor connections (which might make the recurrent connections in different directions from a given neuron asymmetric thus leading to drift), is by using training for only a limited number of locations in the state space. One way to introduce the concept is to recall that the number

of stable states that can be trained in a discrete attractor (with fully distributed random binary patterns) is  $0.14C$  where  $C$  is the number of recurrent synapses onto each neuron (Hopfield 1982). With sparse patterns, the capacity is

$$p \approx \frac{C}{a \ln(1/a)} k \quad (16)$$

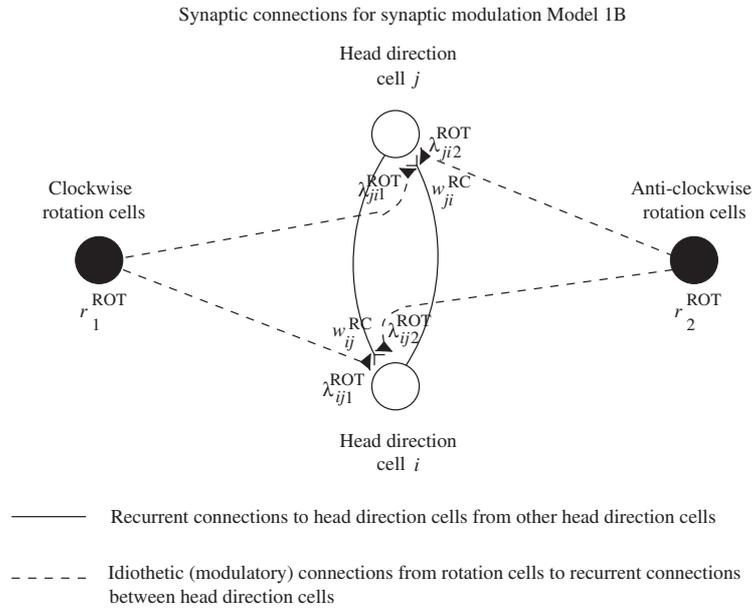
where  $k$  is a factor that is roughly of the order of 0.2–0.3 and depends weakly on the detailed structure of the firing rate distribution, on the connectivity pattern etc, and where  $a$  is the sparseness defined by

$$a = \frac{(\sum_i r_i/N)^2}{\sum_i (r_i^2/N)} \quad (17)$$

where  $r_i$  is the firing rate of the  $i$ th neuron in a set of  $N$  neurons (Treves and Rolls 1991). If discrete attractor networks are trained with more patterns than set by this critical capacity, then the system will undergo a phase transition and become disordered (i.e. it will be in a spin-glass phase), and will not operate correctly (Hopfield 1982). The concept we propose now for continuous attractors is to train the continuous attractor recurrent network with a limited number of Gaussian patterns which will not exceed some critical capacity (see Battaglia and Treves (1998)). The resulting system should support an activity packet that moves correctly in the state space when pushed by the idiothetic inputs, while at the same time having a fixed number of stable states which will prevent the system from drifting when there are no visual or idiothetic inputs. Effectively the energy distribution will have minima, but there is still a continuous mapping from the state of the continuous attractor to the state space of the agent. We name these systems semi-continuous attractor neural networks (S-CANNs), noting that there is continuity in the underlying representation, and at the same time a discrete number of stable states.

We note that for the neuronal activity packets used in the simulations, the sparseness was approximately 0.3, leading to an estimated number of stable states of the semi-continuous attractor a little above the  $0.14C$  expected for a sparseness of 0.5. Given that typical cortical cells might receive 3000–5000 recurrent collateral synapses (see Rolls and Treves (1998), ch 10), the number of such stable states might be in the order of 1000 in the brain. If head directions were to be maintained with a resolution of for example  $3^\circ$ , then only 120 discrete stable states might be necessary.

We tested this concept by running the self-organizing training procedure on the sigma-pi network (model 1A) with 100 connections per neuron ( $C^{\text{HD}} = 100$ ) and different extents of diluted connectivity down to 10% connectivity. A difference to the training regime used previously is that to ensure that the capacity of a recurrent network operating with discrete attractor states was not exceeded, the number of different head directions in which the network was trained was reduced to ten (so that the loading of the network, defined as  $\alpha = p/C^{\text{HD}}$  where  $p$  is the number of patterns that can be retrieved, was 0.1). Moreover, to ensure that stability was not produced by other means, the simulations were run without the NMDA nonlinearity in the activation functions in the neurons of the continuous attractor. It was found that the network settled into the correct head direction state from among the ten head direction states originally trained when stimulated with a visual input corresponding to any head direction (figure 12, left). (A correct state for the network to settle into was taken as one in which the maximum value of the activity packet was closer to the correct trained head direction than to any other trained head direction.) It was also found that this type of training could make the activity packet in the continuous attractor stable in a reasonable number of the trained locations even when the continuous attractor had diluted connectivity (figure 12, right), although with the diluted connectivity there appeared to be fewer than the trained number of stable states.



**Figure 13.** Recurrent and idiopathic synaptic connections to head direction cells in the synaptic modulation model 1B.

## 5. Model 1B

In model 1A described above, there were two separate sets of synapses: the recurrent synapses  $w_{ij}^{\text{RC}}$  between the head direction cells in the continuous attractor network, and the idiopathic synapses  $w_{ijk}^{\text{ROT}}$  from the head rotation cells to the head direction cells. These two sets of synapses operated in a completely independent manner. However, an alternative way of formulating the mechanism is have the firing rates of rotation cells modulate the strength of the recurrent connections between the cells within the continuous attractor. More specifically, in model 1B, rotation cell firing modulates in a multiplicative way the strength of the recurrent connections in the continuous attractor in such a way that clockwise rotation cells modulate the strength of the synaptic connections in the clockwise direction in the continuous attractor. In a similar way, anticlockwise rotation cells modulate the connections between cells in the anticlockwise direction in the continuous attractor. The appropriate connection strengths can be learned in the same way as in the implementation above. Figure 13 shows a possible neural implementation, where the modulation factors  $\lambda^{\text{ROT}}$  from the rotation cells represent the modulatory influences that the rotation cells exert on the recurrent synapses. The concept of synaptic modulation was used by Zhang (1996), though no possible biological implementation was proposed of how the appropriate dynamic synaptic weight changes might be achieved.

More formally, for model 1B, the dynamical equation (1) governing the activations of the head direction cells is now extended to include inputs from the rotation cells in the following way. The activation of a head direction cell  $i$  is governed by the equation

$$\tau \frac{dh_i^{\text{HD}}(t)}{dt} = -h_i^{\text{HD}}(t) + \frac{\phi_0}{C^{\text{HD}}} \sum_j (\tilde{w}_{ij}^{\text{RC}} - w^{\text{INH}}) r_j^{\text{HD}}(t) + I_i^{\text{V}}, \quad (18)$$

where  $r_j^{\text{HD}}$  is the firing rate of head direction cell  $j$ , and where  $\tilde{w}_{ij}^{\text{RC}}$  is the modulated strength of the synapse (effective weighting function) from head direction cell  $j$  to head direction cell  $i$ .

The modulated synaptic weight  $\tilde{w}_{ij}^{\text{RC}}$  is given by

$$\tilde{w}_{ij}^{\text{RC}} = w_{ij}^{\text{RC}} \left( 1 + \phi_2 \sum_k \lambda_{ijk}^{\text{ROT}} r_k^{\text{ROT}} \right) \quad (19)$$

where  $r_k^{\text{ROT}}$  is the firing rate of rotation cell  $k$ , and  $\lambda_{ijk}^{\text{ROT}}$  is the corresponding modulation factor. In addition, the parameter  $\phi_2$  governs the overall strength of the idiothetic inputs. Thus, there are two types of synaptic connection to head direction cells: (i) recurrent connections from head direction cells to other head direction cells within the recurrent network, whose strength is governed by the terms  $\tilde{w}_{ij}^{\text{RC}}$ , and (ii) idiothetic connections from rotation cells to the head direction cell network, which now have a modulating effect on the synapses between the head direction cells, and whose strength is governed by the modulation factors  $\lambda_{ijk}^{\text{ROT}}$ . As for model 1A, once the head direction cell activations  $h_i^{\text{HD}}$  have been updated at the current time step, the head direction cell firing rates  $r_i^{\text{HD}}$  are calculated according to the sigmoid transfer function (2).

The initial learning phase involves the setting up of the synaptic weights  $w_{ij}^{\text{RC}}$  and the modulation factors  $\lambda_{ijk}^{\text{ROT}}$ . The recurrent synaptic weights  $w_{ij}^{\text{RC}}$  and the modulation factors  $\lambda_{ijk}^{\text{ROT}}$  are set up during an initial learning phase similar to that described for model 1A above, where the recurrent synaptic weights  $w_{ij}^{\text{RC}}$  are updated according to equation (3), and the modulation factors  $\lambda_{ijk}^{\text{ROT}}$  are updated according to

$$\delta \lambda_{ijk}^{\text{ROT}} = \tilde{k} r_i^{\text{HD}} \bar{r}_j^{\text{HD}} r_k^{\text{ROT}} \quad (20)$$

where  $\delta \lambda_{ijk}^{\text{ROT}}$  are the changes in the modulation factors, and where  $r_i^{\text{HD}}$  is the instantaneous firing rate of the postsynaptic head direction cell  $i$ ,  $\bar{r}_j^{\text{HD}}$  is the trace value of the presynaptic head direction cell  $j$  given by equation (7),  $r_k^{\text{ROT}}$  is the firing rate of rotation cell  $k$  and  $\tilde{k}$  is the learning rate associated with this type of synaptic connection.

We note that models 1A and 1B are closely related. Indeed, for the special case  $w_{ijk}^{\text{ROT}} = w_{ij}^{\text{RC}} \lambda_{ijk}^{\text{ROT}}$ , the two models become mathematically identical (except for constants). However, with the learning regime used, this is not the case, and the general relation is as follows. If model 1A is expressed in a similar form to model 1B, equation (18) of model 1B applies to both models, and equation (19) of model 1B becomes for model 1A

$$\tilde{w}_{ij}^{\text{RC}} = w_{ij}^{\text{RC}} + \phi_1 \sum_k w_{ijk}^{\text{ROT}} r_k^{\text{ROT}}. \quad (21)$$

## 6. Discussion

In this paper we first demonstrated that continuous attractor networks can be trained by associative learning rules (equations (3) and (6)) which learn the correct recurrent connections based on the overlap of the firing fields of neurons. We also showed that the resulting activity packets were stable (in that they did not drift when the network maintained its activity without the initiating input stimulus) if all locations in the state space were trained, or more generally, if the training was sufficiently regular. We went on to show that the activity packet will not generally be stable (without drift) if the continuous attractor has diluted connectivity, or if irregular training is used. We then proposed two methods for maintaining the stability of the activity packet under these conditions, involving nonlinearity in the activation function of the neurons, and training at only a limited number of locations in the state space of the agent. The two methods are discussed below.

We also proposed and tested two methods for training the system to use idiothetic (self-motion) inputs to move the activity packets in the continuous attractor correctly. The methods

are generic, and the particular case simulated as an example was how rotation cell firing might move the activity packet in a continuous attractor network representing head direction. Our aim was to produce methods that would show how the appropriate connections might self-organize. We sought also to discover methods that used local learning rules for the synaptic modification, because local learning rules are more biologically plausible than alternatives. (A local synaptic learning rule is one in which the information to modify the synapse is available in the pre-synaptic and post-synaptic elements, and does not need to be transported in from elsewhere, see Rolls and Treves (1998)). We discovered two such methods for enabling the appropriate connections to self-organize, described in this paper as models 1A and 1B. Model 1A used sigma-pi neurons and the architecture shown in figure 5. Model 1B used the concept of modulation of synaptic strengths in the appropriate direction in the continuous attractor network depending on the direction of rotation as suggested by Zhang (1996), but actually provided suggestions about how this might be achieved, and a demonstration that the model (1B) worked. Both models made novel use of a trace synaptic modification rule to enable the recent change of the head direction being represented in the continuous attractor to be correctly associated with the current idiothetic rotation signal.

The actual biophysical mechanisms that are needed to implement the self-organization by learning of the idiothetic connections in both models must, necessarily given the computational structure of the problem to be solved, include three terms. In the models these terms are the idiothetic cell firing, and two signals in the head direction continuous attractor network to specify the direction of change of the activity packet representing the current head direction. In both models the multiplicative interactions required, namely sigma-pi operation or synaptic strength modulation, could be performed by presynaptic contacts. However, multiplicative interactions of the type needed in these models might be achieved in a number of other biophysically plausible ways described by Koch (1999, section 21.1.1) and Jonas and Kaczmarek (1999).

The models described in this paper show how path integration could be achieved in a system that self-organizes by associative learning. The path integration is performed in the sense that the representation in a continuous attractor network of the current location of the agent in the state space can be continuously updated based on idiothetic (self-motion) cues, in the absence of visual inputs. The path integration described in this paper refers to updating the representation of head direction using head rotation velocity inputs, and is extended by Stringer *et al* (2002) to an agent performing path integration in a two-dimensional space (such as the floor of a room, or open terrain) based on idiothetic inputs from head direction cell firing and from linear whole body velocity cues. We note that whole body motion cells are present in the primate hippocampus (O'Mara *et al* 1994) and that head direction cells are present in the primate presubiculum (Robertson *et al* 1999).

Previous models of path integration suffer from the problem that they tend to operate as look-up tables, with no suggested process by which the required connections might be formed. For example, a continuous attractor network of head direction cells might have its pattern of firing rates updated by two sets of idiothetic inputs conveying information about the agent rotating either clockwise or anticlockwise. However, a hard-wired model of this process must effectively rely on a form of 'look-up' table to be able to move the activity packet in the correct way, so as to properly track and represent the state of the agent. That is, with the continuous attractor network in a particular state, and with incoming idiothetic inputs representing, say, a rotation or translation, which new cluster of neurons is to be activated in the attractor network is determined through a pre-set matrix that encodes which neurons in the attractor network are stimulated given the current state of the attractor network and the idiothetic inputs. In Samsonovich and McNaughton (1997), this look-up table takes the form

of an intermediate layer of neurons, with pre-set synaptic connections such that an incoming combination of signals from the continuous attractor and the idiothetic input fires a set of cells in the intermediate layer that will in turn stimulate the appropriate cells in the continuous attractor network such that the new attractor state properly reflects the agent's new state. However, the apparent absence of any spatial regularity in the cell response properties of the continuous attractor networks makes such innate hard-wiring unlikely. This is because the 'look-up' table instantiated by the intermediate layer of neurons must give an output that is consistent in the following sense. Consider a continuous attractor network of head direction cells. If a  $30^\circ$  clockwise rotation shifts the attractor network from state  $S_1$  to state  $S_2$ , and if a further  $40^\circ$  clockwise rotation shifts the attractor network from state  $S_2$  to state  $S_3$ , then a  $70^\circ$  clockwise rotation should be able to shift the network from state  $S_1$  to state  $S_3$ . This form of consistency places severe constraints on the synaptic weights to and from the intermediate layer, and it is not clear how an innate connectivity could be hard-wired that met these constraints, especially without some form of spatial organization of response properties in the continuous attractor network. In addition, if the cell response properties of the continuous attractor network do need to develop through learning and self-organization as suggested above, then this might create further difficulties for such a hard-wired 'look-up' table instantiated in an intermediate layer. Hence, such a hard-wired 'look-up' table mechanism for moving the localized activity packet within the continuous attractor networks appears unlikely, and a more biologically plausible mechanism for idiothetic updating of the continuous attractor network may need to develop through learning and self-organization. Two such models have been described in this paper.

The networks described in this paper were able to maintain a stable activity packet if the recurrent weight profiles were identical for each node in the continuous attractor network, were symmetrical in the two directions round the continuous attractor and if nearby nodes were more strongly coupled than more distant nodes. This could be achieved by using an associative learning rule to link nodes with overlapping Gaussian firing rate profiles, complete connectivity in the continuous attractor and regular training. If the recurrent connectivity between the neurons in the continuous attractor network was incomplete, or if there was some irregularity in the training regime, then the activity packet tested without visual input drifted. Two ways to overcome the drift were found. One was to use supra-linearity in the neuronal activation function, to encourage neurons already firing to keep firing. Such a function might be implemented by NMDA receptors (Wang 1999). The second way was to train the network at a limited number of locations in the state space. We conjecture that the number of locations must not exceed what would be the capacity of the networks when operating as discrete attractors. Provided that the recurrent weights between neurons in the continuous attractor are still a simple function of the physical distance to be represented in the environment, the network can still represent continuous space, though in a semi-continuous way. The stability in this scenario is achieved by virtue of the fact that there are effectively discrete energy minima in the semi-continuous energy landscape. Of course, care must be taken not to overtrain these semi-continuous attractor networks with too many locations, or they will enter a disorganized, spin-glass phase (Amit 1989). Prevention of overtraining in the natural world might be assisted by only learning when new or salient visual inputs produced by new places or views that would fire head direction cells are encountered.

An interesting aspect of the operation of continuous attractors with diluted connectivity trained in only a limited number of locations in the state space is that the exact node on which the activity packet is centred is close to one of the training nodes, but that the number of such stable locations in the state space increases approximately as one would expect the capacity of a discrete attractor network to increase. We postulate that this is because the diluted connectivity produces minima in the energy landscape superimposed on the minima introduced

by the limited training, and that these energy minima support stable firing provided that the critical capacity of the network operating as a discrete attractor is not exceeded.

The model makes a number of predictions, which can in principle be tested experimentally. One prediction is that for the idiotic learning rule equation (11) to operate correctly using its temporal trace, the presynaptic term  $\bar{r}_j^{\text{HD}}$  must precede or reflect the history before the post-synaptic term becomes active. This requirement of the model appears to be supported (see Abbott and Nelson 2000). Another prediction is that the idiotic inputs (in the case considered in this paper, from head rotation cells) need to operate with nonlinearity of the type specified by sigma-pi neurons. In this context, we remark that very few idiotic sigma-pi synapses would suffice to implement the mechanism for path integration described in this paper. The reason for this is that the introduction of any asymmetry into the continuous attractor functional connectivity will suffice to move the activity packet. The prediction is thus made that the connectivity of the idiotic inputs could be quite sparse in brain systems that perform path integration.

### Acknowledgments

This research was supported by the Medical Research Council, grant PG9826105, by the Human Frontier Science Programme and by the MRC Interdisciplinary Research Centre for Cognitive Neuroscience.

### References

- Abbott L F and Nelson S B 2000 Synaptic plasticity: taming the beast *Nature Neurosci.* **3** 1178–83
- Amari S 1977 Dynamics of pattern formation in lateral-inhibition type neural fields *Biol. Cybern.* **27** 77–87
- Amit D J 1989 *Modeling Brain Function. The World of Attractor Neural Networks* (Cambridge: Cambridge University Press)
- Arbib M A 1997 From visual affordances in monkey parietal cortex to hippocampo-parietal interactions underlying rat navigation *Phil. Trans. R. Soc. B* **352** 1429–36
- Battaglia F P and Treves A 1998 Attractor neural networks storing multiple space representations: a model for hippocampal place fields *Phys. Rev. E* **58** 7738–53
- Dominey P F and Arbib M A 1992 A cortico-subcortical model for generation of spatially accurate sequential saccades *Cereb. Cortex* **2** 153–75
- Droulez J and Berthoz A 1991 A neural network model of sensoritopic maps with predictive short-term memory properties *Proc. Natl Acad. Sci. USA* **88** 9653–7
- Georges-François P, Rolls E T and Robertson R G 1999 Spatial view cells in the primate hippocampus: allocentric view not head direction or eye position or place *Cereb. Cortex* **9** 197–212
- Guazzelli A, Bota M and Arbib M A 2001 Competitive Hebbian learning and the hippocampal place cell system: modeling the interaction of visual and path integration cues *Hippocampus* **11** 216–39
- Hopfield J J 1982 Neural networks and physical systems with emergent collective computational abilities *Proc. Natl Acad. Sci. USA* **79** 2554–8
- Jonas E A and Kaczmarek L K 1999 The inside story: subcellular mechanisms of neuromodulation *Beyond Neurotransmission* ed P S Katz (New York: Oxford University Press) ch 3, pp 83–120
- Kesner R and Rolls E T 2001 Role of long term synaptic modification in short term memory *Hippocampus* **11** 240–50
- Koch C 1999 *Biophysics of Computation* (Oxford: Oxford University Press)
- Lisman J E, Fellous J M and Wang X J 1998 A role for NMDA-receptor channels in working memory *Nature Neurosci.* **1** 273–5
- Markus E J, Barnes C A, McNaughton B L, Gladden V L and Skaggs W 1994 Spatial information content and reliability of hippocampal CA1 neurons: effects of visual input *Hippocampus* **4** 410–21
- Markus E J, Qin Y L, Leonard B, Skaggs W, McNaughton B L and Barnes C A 1995 Interactions between location and task affect the spatial and directional firing of hippocampal neurons *J. Neurosci.* **15** 7079–94
- McNaughton B L, Barnes C A and O'Keefe J 1983 The contributions of position, direction, and velocity to single unit activity in the hippocampus of freely-moving rats *Exp. Brain Res.* **52** 41–9

- McNaughton B L, Leonard B and Chen L 1989 Cortical–hippocampal interactions and cognitive mapping: a hypothesis based on reintegration of the parietal and inferotemporal pathways for visual processing *Psychobiology* **17** 230–5
- Muller R U, Kubie J L, Bostock E M, Taube J S and Quirk G J 1991 Spatial firing correlates of neurons in the hippocampal formation of freely moving rats *Brain and Space* ed J Paillard (Oxford: Oxford University Press) pp 296–333
- Muller R U, Ranck J B and Taube J S 1996 Head direction cells: properties and functional significance *Curr. Opin. Neurobiol.* **6** 196–206
- Oja E 1982 A simplified neuron model as a principal component analyser *J. Math. Biol.* **15** 267–73
- O’Keefe J 1976 Place units in the hippocampus of the freely moving rat *Exp. Neurol.* **51** 78–109
- O’Keefe J 1984 Spatial memory within and without the hippocampal system *Neurobiology of the Hippocampus* ed W Seifert (London: Academic) pp 375–403
- O’Keefe J and Dostrovsky J 1971 The hippocampus as a spatial map: preliminary evidence from unit activity in the freely moving rat *Brain Res.* **34** 171–5
- O’Mara S M, Rolls E T, Berthoz A and Kesner R P 1994 Neurons responding to whole-body motion in the primate hippocampus *J. Neurosci.* **14** 6511–23
- Pouget A and Sejnowski T J 1995 Spatial representations in the parietal cortex may use basis functions *Advances in Neural Information Processing Systems* vol 7, ed G Tesauro, D S Touretzky and T K Leen (Cambridge, MA: MIT Press) pp 157–64
- Quirk G L, Muller R U and Kubie J L 1990 The firing of hippocampal place cells in the dark depends on the rat’s recent experience *J. Neurosci.* **10** 2008–17
- Ranck J B Jr 1985 Head direction cells in the deep cell layer of dorsolateral presubiculum in freely moving rats *Electrical Activity of the Archicortex* ed G Buzsáki and C H Vanderwolf (Budapest: Akadémiai Kiadó)
- Redish A D, Elga A N and Touretzky D S 1996 A coupled attractor model of the rodent head direction system *Network: Comput. Neural Syst.* **7** 671–85
- Redish A D and Touretzky D S 1998 The role of the hippocampus in solving the Morris water maze *Neural Comput.* **10** 73–111
- Robertson R G, Rolls E T and Georges-François P 1998 Spatial view cells in the primate hippocampus: effects of removal of view details *J. Neurophysiol.* **79** 1145–56
- Robertson R G, Rolls E T, Georges-François P and Panzeri S 1999 Head direction cells in the primate presubiculum *Hippocampus* **9** 206–19
- Rolls E T and Deco G 2002 *Computational Neuroscience of Vision* (Oxford: Oxford University Press)
- Rolls E T and Milward T 2000 A model of invariant object recognition in the visual system: learning rules, activation functions, lateral inhibition and information-based performance measures *Neural Comput.* **12** 2547–72
- Rolls E T, Robertson R G and Georges-François P 1997 Spatial view cells in the primate hippocampus *Eur. J. Neurosci.* **9** 1789–94
- Rolls E T and Stringer S M 2001 Invariant object recognition in the visual system with error correction and temporal difference learning *Network* **12** 111–29
- Rolls E T and Treves A 1998 *Neural Networks and Brain Function* (Oxford: Oxford University Press)
- Samsonovich A and McNaughton B 1997 Path integration and cognitive mapping in a continuous attractor neural network model *J. Neurosci.* **17** 5900–20
- Skaggs W E, Knierim J J, Kudrimoti H S and McNaughton B L 1995 A model of the neural basis of the rat’s sense of direction *Advances in Neural Information Processing Systems* vol 7, ed G Tesauro, D S Touretzky and T K Leen (Cambridge, MA: MIT Press) pp 173–80
- Stringer S M, Rolls E T, Trappenberg T P and de Araujo I E T 2002 Self-organizing continuous attractor networks and path integration: two-dimensional models of place cells *Network: Comput. Neural Syst.* at press
- Taube J S, Goodridge J P, Golob E G, Dudchenko P A and Stackman R W 1996 Processing the head direction signal: a review and commentary *Brain Res. Bull.* **40** 477–86
- Taube J S, Muller R U and Ranck J B Jr 1990 Head-direction cells recorded from the postsubiculum in freely moving rats. I. Description and quantitative analysis *J. Neurosci.* **10** 420–35
- Taylor J G 1999 Neural ‘bubble’ dynamics in two dimensions: foundations *Biol. Cybern.* **80** 393–409
- Treves A and Rolls E T 1991 What determines the capacity of autoassociative memories in the brain? *Network* **2** 371–97
- Wang X J 1999 Synaptic basis of cortical persistent activity: the importance of NMDA receptors to working memory *J. Neurosci.* **19** 9587–603
- Zhang K 1996 Representation of spatial orientation by the intrinsic dynamics of the head-direction cell ensemble: a theory *J. Neurosci.* **16** 2112–26