Biol Cybern (2010) 103:21–41
DOI 10.1007/s00422-009-0355-0

ORIGINAL PAPER

Biological Cybernetics

Path integration of head direction: updating a packet of neural activity at the correct speed using neuronal time constants

D. M. Walters · S. M. Stringer

Received: 6 July 2009 / Accepted: 27 November 2009 / Published online: 26 May 2010

Abstract A key question in understanding the neural basis of path integration is how individual, spatially responsive, neurons may self-organize into networks that can, through learning, integrate velocity signals to update a continuous representation of location within an environment. It is of vital importance that this internal representation of position is updated at the correct speed, and in real time, to accurately re ect the motion of the animal. In this article, we present a biologically plausible model of velocity path integration of head direction that can solve this problem using neuronal time constants to effect natural time delays, over which associations can be learned through associative Hebbian learning rules. The model comprises a linked continuous attractor network and competitive network. In simulation, we show that the same model is able to learn two different speeds of rotation when implemented with two different values for the time constant, and without the need to alter any other model parameters. The proposed model could be extended to path integration of place in the environment, and path integration of spatial view.

Keywords Path integration · Head direction cells · Continuous attractor network · Competitive network · Neuronal time constant

D. M. Walters (✉) · S. M. Stringer
Department of Experimental Psychology, Oxford Centre for Theoretical Neuroscience and Arti cial Intelligence, South Parks Road, Oxford OX1 3UD, UK
e-mail: daniel.walters@psy.ox.ac.uk
URL: www.oftnai.org

1 Introduction

Classes of cells exist in the brain that primarily signal spatial information. Rat hippocampal place cells respond maximally when the rat is in a particular location in its environment (O'Keefe and Dostrovsky 1971; McNaughton et al. 1983; Muller et al. 1991). Rat entorhinal cortex grid cells are similar to place cells insofar as they re ect the location of the rat—but with each cell responding to several locations in the same environment in a hexagonal grid-like pattern (Hafting et al. 2005; Sargolini et al. 2006). Primate hippocampal spatial view cells code for the particular location at which the primate is looking (Rolls et al. 1997; Georges-François et al. 1999; Rolls and Xiang 2006). Also, head direction cells, which signal when the head of the animal is facing in a particular preferred direction, have been shown in both rats (Ranck 1985; Taube et al. 1990a; Muller et al. 1996) and primates (Robertson et al. 1999).

A key problem to be addressed is how, through learning, these cells may self-organize to form networks that can perform the necessary calculations for spatial navigation. In this article, we consider the process of path integration, where an animal integrates idiothetic (self-motion) cues, such as forward motion or rotation, to update its internal representation of its position or orientation within an environment (Mittelstaedt and Mittelstaedt 1980, 1982; Collett and Zeil 1998; Redish 1999).

It has been shown in previous research that when presented with a continuous input space, a layer of recurrently connected neurons can form a continuous attractor network. Within this network, a local packet of persistent activity (a stable state) can represent a particular directional heading or location (Amari 1977; Taylor 1999). A continuous attractor can support a continuous space of stable memory states, and can, thus, learn to represent any given location or directional

heading. With training on a continuous sensory space, such as location or head direction, after associative Hebbian learning, the strength of the synapses between any two neurons will be a function of the degree of overlap of their receptive fields (the larger the overlap, the stronger the synapse), and, thus, symmetric. Thus, the strengths of the synapses between any given individual neuron and all other neurons in the network are described by an approximately Gaussian distribution, with the synaptic strengths again determined by the degree of receptive-field overlap. This symmetry ensures that a packet of neural activity representing a directional heading or a location can remain stable (unlikely to drift) and persistent within the continuous attractor network when the animal is not moving.

An asymmetric external input is required to drive the activity packet around the continuous attractor network and, thus, update the location of the activity packet in the attractor space. In this manner, the activity packet will track the exact position and orientation of the animal as it moves within its environment. Various methods have been proposed for incorporating both symmetric and asymmetric weights in a continuous attractor, including the introduction of cells that signal rotational velocity to the continuous attractor network (Skaggs et al. 1995; Redish et al. 1996; Samsonovich and McNaughton 1997; Stringer et al. 2002) and imposing the asymmetric weight profile upon the network (Zhang 1996; Song and Wang 2005).

An important question is how the network can learn to update the location of the activity packet at the *same speed* as the animal is moving in its environment during path integration. Even small (but significant) errors in the speed of the packet, compared to the true speed of the animal, could be compounded through time to produce a large discrepancy between the internal representation and the actual location of the animal. In order to achieve successful path integration in a continuous attractor network, there is, thus, a requirement for some biologically plausible mechanism through which the network can self-organize to produce an asymmetry in the synaptic weights that can move the packet of activity in a manner that is calibrated with the speed the animal is moving. We have previously shown how this may be achieved in a biologically plausible manner with Hebbian associative learning, incorporating explicit axonal conduction delays (Walters et al. 2009). In this article, we demonstrate another way in which this may be achieved, using biologically realistic neuronal time constants to effect natural time delays which are then combined with Hebbian associative learning.

We, thus, propose a model which can learn to update the packet of neural activity at the same speed at which the simulated agent is moving. We address the issue of path integration of head direction, and, therefore, our model is of a generic one-dimensional system. We do, however, note that the general principles of this model could be extended to path integration in higher dimensions, i.e. path integration of place in the environment, or of spatial view.

The present model is a development of a model proposed by Stringer and Rolls (2006), in which a continuous attractor is reciprocally linked with a competitive network. In the previous model, a *trace* learning rule was used to introduce a general asymmetry in the synaptic weights during learning. The asymmetry in the weights was in the correct direction, which allowed the network to perform path integration in the correct direction (i.e. clockwise or counter-clockwise). However, the asymmetric weights were not of the correct magnitude to effect the correct speed during replay. In other words, the model of Stringer and Rolls could not automatically learn to perform path integration at the same speed on which the network had been trained in the light.

In the new model presented here, we use a purely Hebbian associative learning rule rather than the trace rule. The use of the Hebb rule allows our model to learn precise associations between successive activity states within the network, which represent particular head directions, over specific natural time intervals present in the network. In the model described in this article, the natural time intervals are effected by the neuronal time constants governing the activations of the cells. This new approach allows the model to learn to perform path integration at approximately the same speed on which the model is trained in the light. In the simulations described later, we show that the model can operate with different values for the neuronal time constants, and, in each case, can learn two different speeds of path integration without the need to alter any of the other parameters governing the behaviour of the model.

In this article, we have used the neuronal time constants for the cell activations to effect a natural time delay over which to learn associations between successive network states, i.e. head directions. We propose, however, that the principle of using some kind of natural time interval in the network to provide the correct timing for updating a packet of neural activity is quite general, and other natural time intervals, such as explicit axonal conduction delays (Walters et al. 2009) or natural oscillations in network activity, might be used by the model in a similar manner.

2 The model

2.1 Conceptual overview of the operation of the model

Our model, shown in Fig. 1, consists of two connected networks. A network of head direction cells (with firing rate r_i^{HD} for postsynaptic head direction cell i) represents the current head direction of the agent, and operates as a continuous attractor performing velocity path integration. Individual head direction cells are simulated using leaky-integrator

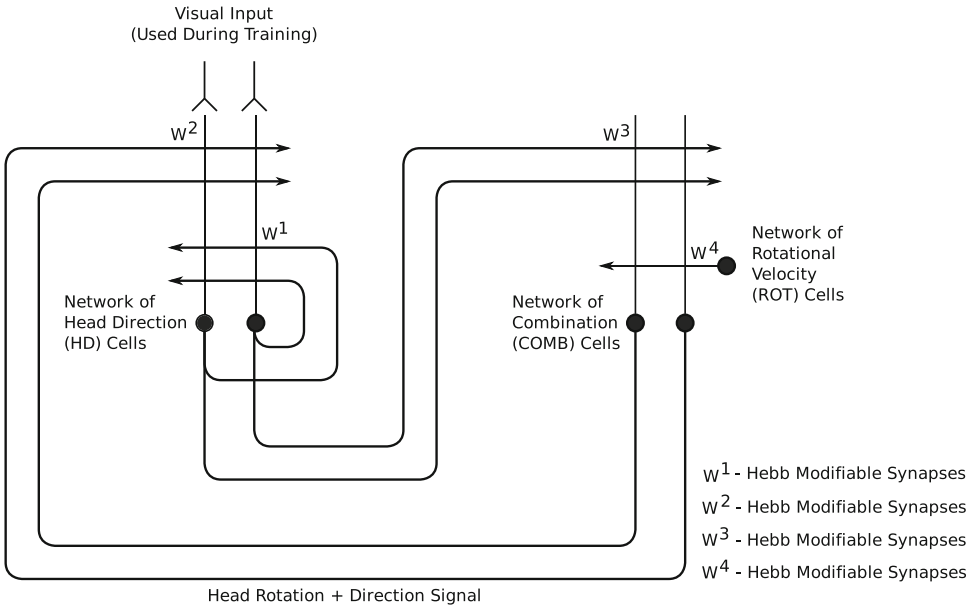

Fig. 1 Network architecture for a two-layer self-organizing neural network model of the head direction system. The model architecture contains a network of head direction (HD) cells representing the current head direction of the agent; a network of combination (COMB) cells representing a combination of head direction and rotational velocity, and a layer of rotational velocity (ROT) cells that become active when the agent rotates. There are four types of synaptic connectivity in the model, which operate as follows. The w_{ij}^1 synapses are Hebb-modifiable recurrent connections between head direction cells. These connections help to support a stable packet of activity within the continuous attractor network of head direction cells in the absence of visual input. The combination cells receive inputs from the head direction cells through the Hebb-modifiable w_{ij}^3 synapses, and inputs from the rotational velocity cells through the Hebb-modifiable w_{ij}^4 synapses. These synaptic inputs encourage combination cells to respond, by competitive learning, to combinations of a particular head direction and rotational velocity. In consequence, the combination cells only become active when the agent is rotating. The head direction cells receive input from the combination cells through the Hebb-modifiable w_{ij}^2 synapses. The neuronal time constants introduce effective transmission delays between the head direction cells and the combination cells. As the agent rotates during training, this in turn introduces, through associative learning, an asymmetry in the w_{ij}^2 and w_{ij}^3 synaptic weight profiles, which results in each combination cell learning to stimulate a different postsynaptic head direction cell to the presynaptic head direction cell that has learned to preferentially stimulate it. These asymmetries are important in shifting the packet of head direction cell activity through the head direction cell network at the correct speed, and on the basis of idiothetic signals alone

firing rate-based models, as described below. That is, our model does not simulate the individual action potentials emitted by cells. Instead, only the instantaneous average firing rate for each cell is simulated evolving through time. Within the layer of head direction cells, there is a single Gaussian packet of firing activity. The centre of this activity packet represents the current head direction of the simulated agent.

A network of combination cells (with firing rate e_i^{COMB} for postsynaptic combination cell i) receives inputs from both the network of head direction cells and a layer of rotational velocity cells (with separate sub-populations of cells signalling clockwise and counter-clockwise rotation). The combination cells are also simulated using leaky-integrator firing rate-based models, as described shortly. The combination cells operate as a competitive network and develop their response profiles during training, with individual combination cells learning to represent a combination of a particular head direction and a particular rotational velocity. Such cells have been found in the brain, having both a Gaussian tuning to head direction and a tuning to the direction of angular head velocity (Taube et al. 1990b; Sharp 1996; Bassett and Taube 2005).

In the model described here, we do not incorporate explicit axonal conduction delays as used by Walters et al. (2009). Instead, we show how the neuronal time constants in the current model can introduce effective natural time delays in the propagation of signals from the layer of head directions cells to the layer of combination cells, and then back again from the combination cells to the head direction cells. The effect of the neuronal time constants is that when a presynaptic cell is driving a postsynaptic cell, the activity profile of the postsynaptic cell is slightly delayed in time from the activity profile of the presynaptic cell. One useful statistic that we have considered is the time delay between the centres of mass of the presynaptic and postsynaptic temporal activity profiles. In particular, the centre of mass of the activity profile of the postsynaptic cell will be delayed by a small fixed time interval Δt after the centre of mass of the activity profile

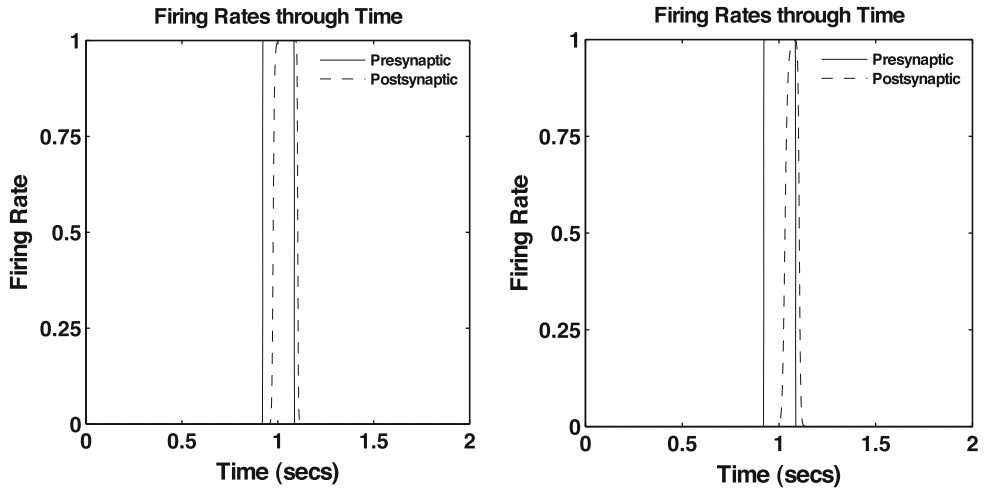

Fig. 2 Firing rates through time of two simulations of a simple two-cell system, with the presynaptic cell firing rate indicated by a *solid line* and the postsynaptic cell firing rate indicated by a *dashed line*. *Left* Simulation in which the time constant τ of the governing equation for postsynaptic cell activation was set to 50 ms. As can be seen, there is a delay between the centres of mass of the temporal firing profiles of the presynaptic cell and the postsynaptic cell. *Right* Simulation in which the time constant τ of the governing equation for postsynaptic cell activation was set to 100 ms. There is a clear delay between the centres of mass of the temporal firing profiles of the presynaptic cell and the postsynaptic cell. The delay is more pronounced than in the *left plot*, which indicates that a longer time constant τ will produce a longer delay Δt between the presynaptic and postsynaptic activity profiles

of the presynaptic cell. Although there is a small delay Δt between the centres of mass, the synaptic connection from the presynaptic cell to the postsynaptic cell is still strengthened by the firing rate-based associative Hebbian rule. This learning rule strengthens the connections between the two cells in proportion to the product of their current instantaneous firing rates, and, hence, strengthens the connection during learning because the periods of activity of the two cells still partially overlap through time. This effectively allows the network to learn associations between neural activity in the head direction cell and combination cell networks over specific fixed time intervals Δt, and, thus, the network learns to shift a packet of head direction cell activity at the correct speed.

In order to explain the operation of the model, consider a simple two-cell system. The activation h_i of the postsynaptic cell i is governed by

$$\tau \frac{\mathrm{d}h_i(t)}{\mathrm{d}t} = -h_i(t) + w_{ij}(t) r_j(t) \tag{1}$$

where τ is the neuronal time constant, $r_j(t)$ is the firing rate of the presynaptic cell j, and w_{ij} is the synaptic weight from the presynaptic cell j to the postsynaptic cell i. The firing rate of the postsynaptic cell is given by the sigmoid transfer function

$$r_i(t) = \frac{1}{1 + \mathrm{e}^{-2\beta(h_i(t)-\alpha)}} \tag{2}$$

where α is the sigmoid threshold and β is the slope. Although the presynaptic cell j and postsynaptic cell i may fire together, a relatively large time constant τ, say 50–100 ms, will ensure that the centre of mass of the temporal profile of the postsynaptic cell activity is delayed by a small time interval Δt from the centre of mass of the temporal profile of the presynaptic cell activity. A larger neuronal time constant τ will lead to a longer effective delay Δt. This is shown in Fig. 2 and Table 1.

Table 1 Simulation results for the simple two-cell system

Results from two-cell system	
τ (ms)	Δt (ms)
50	36.8
60	43.2
70	49.2
80	54.7
90	59.7
100	64.2

For each of six simulations, we varied the neuronal time constant τ and measured the effective time delay Δt between the centres of mass of the temporal profile of the presynaptic cell firing and the temporal profile of the postsynaptic cell firing. In relation to the simple model simulated, an increase of 10 ms in the time constant τ produces an increase of approximately 5–6 ms in the time delay Δt between the centres of mass of the presynaptic and postsynaptic cell temporal firing profiles

During learning, the synaptic weight w_{ij} from presynaptic cell j to postsynaptic cell i is strengthened according to the Hebbian associative learning rule

$$\frac{\mathrm{d}w_{ij}(t)}{\mathrm{d}t} = k r_i(t) r_j(t) \tag{3}$$

where k is the learning rate, r_j is the firing rate of the presynaptic cell, and r_i is the firing rate of the postsynaptic cell. This

learning rule depends upon the instantaneous ring rates of the presynaptic and postsynaptic cells, and does not depend upon the precise timing of the action potentials emitted by these cells (which are not explicitly simulated in this model). The learning rule will strengthen the weight between the presynaptic and postsynaptic cells during the short periods in which both of these cells have high ring rates. However, due to the neuronal nite time constant of the postsynaptic cell, the network learns how to associate a temporal activity pro le in the presynaptic cell with a slightly delayed temporal activity pro le in the postsynaptic cell. In particular, the centre of mass of the temporal activity pro le of the postsynaptic cell is always delayed Δt from the centre of mass of the temporal activity pro le of the presynaptic cell. This will hold true both during and after learning. The consequence of this is that the presynaptic cell learns to drive activity in the postsynaptic cell with the same effective time delay Δt that was present during learning. Moreover, a longer postsynaptic time constant τ ensures that the presynaptic cell learns to stimulate the postsynaptic cell a greater time interval Δt after the presynaptic cell ring. It is this delay Δt between presynaptic and postsynaptic cell ring, which allows the correct time intervals between cell activities to be learned and replayed by cells in the model. This is essential for the model to be able to replay the temporal sequence of cell activity at the speed that was experienced during learning.

In order to further illustrate how the neuronal time constant leads to effective signal delays, we simulated a two-cell system according to Eqs. 1 and 2 with a non-modi able synaptic weight $w_{ij} = 1$ between the presynaptic and postsynaptic cells. Figure 2 displays the presynaptic and postsynaptic cell ring rates through time for two simulations of the two-cell system. In the left plot, the neuronal time constant τ for the postsynaptic cell was set to 50 ms. It is evident that there is a delay between the temporal ring pro les of the presynaptic cell and the postsynaptic cell. In the right plot, the neuronal time constant was set to 100 ms. The delay between the presynaptic and postsynaptic cell ring is more pronounced, which illustrates that a longer time constant will produce a longer delay Δt between the presynaptic and postsynaptic temporal activity pro les. In order to further quantify this mechanism, we conducted six simulations of the two-cell system, each with different values for the time constant τ. For each simulation, we calculated the time delays Δt between the centres of mass of the presynaptic cell temporal ring pro le and the postsynaptic cell temporal ring pro le. Table 1 shows the calculated delay Δt for the six simulations conducted with different values of the postsynaptic neuronal time constant τ. Each increase of 10 ms in the time constant τ leads to an approximate increase of 5–6 ms in the effective delay Δt. These results show that the time delay is approximately proportional to the neuronal time constant τ. The time constant can, thus, serve as a reliable timing mechanism over which associations can be made between presynaptic and postsynaptic cell activities.

Now let us reconsider the full model, shown in Fig. 1, that is simulated in this paper. In the simulations presented below we implemented relatively large neuronal time constants $\tau^{\mathrm{COMB}} = 100\&150$ ms for the combination cells, and a relatively small time constant $\tau^{\mathrm{HD}} = 1$ ms for the head direction cells. This meant that the main effective delay in neurotransmission was in the head direction cell to combination cell w^3 synaptic connections. During training, visual input is available to guide the ring of the head direction cells so that they are forced to re when the agent is facing their preferred head direction. At any moment in time the individual combination cells will respond to an earlier pattern of head direction cell ring at time $t - \Delta t$ through the synaptic connections w^3. At the same time, the combination cells will strengthen their w^2 synaptic connections onto the head direction cells that are currently active at time t. If the agent is rotating with velocity $\frac{\mathrm{d}\theta}{\mathrm{d}t}$, then over this time interval Δt, the agent will have rotated $\Delta t \frac{\mathrm{d}\theta}{\mathrm{d}t}$. This is the overall association that the network learns to perform. That is, the network learns that an activity pattern in the head direction cell network at time t_1 representing a particular head direction θ_1 should stimulate (via the network of combination cells) a new pattern of activity in the head direction cell network at time $t_2 = t_1 + \Delta t$, representing a later head direction $\theta_2 = \theta_1 + \Delta t \frac{\mathrm{d}\theta}{\mathrm{d}t}$. This kind of associative learning over xed time intervals enables the model to learn the correct velocity for updating the packet of neural activity in the head direction cell network, and thus allows path integration to occur at the correct speed.

2.2 Model equations and implementation

The two-layer model architecture is shown in Fig. 1. During training, each head direction cell receives an external visual input e_i, which carries information about the head direction of the agent. When visual cues are available, these external inputs dominate other excitatory inputs to the head direction cells, and force each head direction cell to respond best to a particular head direction of the agent. The Gaussian tuning of the head direction cell response pro le leads, for any given head direction cell, to a decrease in cell ring as the head direction of the agent moves away from the preferred head direction of that cell.

The activation h_i^{HD} of head direction i in the model is governed by

$$\tau^{\mathrm{HD}} \frac{\mathrm{d}h_i^{\mathrm{HD}}(t)}{\mathrm{d}t} = -h_i^{\mathrm{HD}}(t) + e_i(t) - \frac{1}{N^{\mathrm{HD}}} \sum_j \tilde{w}^{\mathrm{HD}} r_j^{\mathrm{HD}}(t)$$

$$+\frac{\phi_1}{C^{\mathrm{HD}\to\mathrm{HD}}}\sum_j w_{ij}^1(t)r_j^{\mathrm{HD}}(t) + \frac{\phi_2}{C^{\mathrm{COMB}\to\mathrm{HD}}}\sum_j w_{ij}^2(t)r_j^{\mathrm{COMB}}(t) - I^{\mathrm{EXTERN}} \quad (4)$$

where the activation $h_i^{\mathrm{HD}}(t)$ is driven by the following terms: $\sum_j \tilde{w}^{\mathrm{HD}} r_j^{\mathrm{HD}}(t)$ represents inhibitory feedback within the head direction cell network, where the summation is performed over all the presynaptic head direction cells j, $\tilde{w}^{\mathrm{HD}}$ is a global constant describing the effect of inhibitory interneurons within the network of head direction cells, and $n^{\mathrm{HD}} = 500$ is the total number of head direction cells in the model. The term $\sum_j w_{ij}^1(t)r_j^{\mathrm{HD}}(t)$ represents excitatory feedback within the layer of head direction cells. The term $r_j^{\mathrm{HD}}(t)$ is the presynaptic firing rate of head direction cell j, and w_{ij}^1 is the excitatory (positive) synaptic weight from presynaptic head direction cell j to postsynaptic head direction cell i.[1] Further terms in Eq. 4 are as follows. The term $-h_i^{\mathrm{HD}}(t)$ is a decay term such that, in the absence of further presynaptic input, the activation level of the postsynaptic head direction cell will decay to zero according to the time constant τ^{HD}. The term $e_i(t)$ represents an external visual input to postsynaptic head direction cell i. When there is no visual input, the term e_i is set to zero. Thus, in the absence of visual input, the key term driving the head direction cell activations in Eq. 4 is a sum of inputs from the presynaptic combination cells, $\sum_j w_{ij}^2(t)r_j^{\mathrm{COMB}}(t)$. The term $r_j^{\mathrm{COMB}}(t)$ is the firing rate of the presynaptic combination cell j, and w_{ij}^2 is the strength of the corresponding synapse between that presynaptic combination cell and the postsynaptic head direction cell i.[2] The term I^{EXTERN} represents external feedforward inhibition to the head direction cell network: this is necessary during the learning phase to ensure that, in the presence of visual input, only a small subset of head direction cells (those representing head directions nearby in the head-direction space) are active at any one point in time, i.e. the standard deviation of the head direction cell activity packet remains small. In the absence of external visual input (during the testing phase), the term I^{EXTERN} is set to zero.

The firing rate $r_i^{\mathrm{HD}}(t)$ of postsynaptic head direction cell i is determined from the activation $h_i^{\mathrm{HD}}(t)$ of that cell and the sigmoid activation function

$$r_i^{\mathrm{HD}}(t) = \frac{1}{1+\mathrm{e}^{-2\beta(h_i^{\mathrm{HD}}(t)-\alpha)}} \quad (5)$$

where α and β are the sigmoid threshold and slope respectively.

The recurrent synapses $w_{ij}^1(t)$ in the head direction cell network are trained by a local associative Hebb rule

$$\frac{\mathrm{d}w_{ij}^1(t)}{\mathrm{d}t} = k^1 r_i^{\mathrm{HD}}(t)r_j^{\mathrm{HD}}(t) \quad (6)$$

which increases the strength of the synapses between those head direction cells that represent nearby directions in the head-direction space, and which tend to be co-active due to broadly tuned, overlapping receptive fields. In order to bound the synaptic weights, weight normalization was used. In order to implement the normalization, the recurrent synaptic weights $w_{ij}^1(t)$ were rescaled after each timestep of the learning phase to ensure that for each postsynaptic head direction cell i we have

$$\sqrt{\sum_j (w_{ij}^1(t))^2} = 1 \quad (7)$$

where the sum is over all the presynaptic 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 helps in ensuring that the learning rules are convergent in the sense that the recurrent synaptic weights within the continuous attractor network settle down over time to steady values, i.e. the weights do not grow unbounded.

During learning, associative synaptic modification in the recurrent synapses w_{ij}^1, in conjunction with the continuity of the head-direction space, allows the strength of the synapses between any two head direction cells to reflect the distance between the head directions represented by those two cells. The recurrent connectivity implemented by the w_{ij}^1 synapses allows the network of head direction cells to operate as a continuous attractor network and support stable packets of neural activity in the absence of external visual input; thus, the agent can operate in the dark.

The learning rule to update the $w_{ij}^2(t)$ synapses from presynaptic combination cell j to postsynaptic head direction cell i is expressed by

$$\frac{\mathrm{d}w_{ij}^2(t)}{\mathrm{d}t} = k^2 r_i^{\mathrm{HD}}(t)r_j^{\mathrm{COMB}}(t) \quad (8)$$

which increases the strength of the synapses between co-active postsynaptic head direction cells and presynaptic combination cells. In order to bound the $w_{ij}^2(t)$ synaptic weights, rescaling occurred after each timestep of the learning phase, as in Eq. 7, to ensure that for each postsynaptic head direction cell i, we have

[1] The scaling factor $\frac{\phi_1}{C^{\mathrm{HD}\to\mathrm{HD}}}$ controls the overall strength of the recurrent inputs to the network of head direction cells, where ϕ_1 is a constant, and $C^{\mathrm{HD}\to\mathrm{HD}}$ is the number of synapses onto each postsynaptic head direction cell from the presynaptic head direction cells.

[2] The scaling factor $\frac{\phi_2}{C^{\mathrm{COMB}\to\mathrm{HD}}}$ controls the overall strength of the combination cell inputs, where ϕ_2 is a constant, and $C^{\mathrm{COMB}\to\mathrm{HD}}$ is the number of synapses onto each postsynaptic head direction cell from the presynaptic combination cells.

$$\sqrt{\sum_j (w_{ij}^2(t))^2} = 1 \tag{9}$$

where the sum is over all the presynaptic combination cells j.

After learning has occurred, and during testing in the absence of visual input, the w_{ij}^2 synapses are responsible for stimulating new head direction cells in the direction the agent is rotating, so as to shift the location of the head direction cell activity packet and, thus, perform path integration.

The combination cells are driven by synaptic inputs $w_{ij}^3(t)$ from the head direction cell network, and synaptic inputs $w_{ij}^4(t)$ from the layer of rotational velocity cells (which are external inputs to the network that are only active when the agent is rotating). During the training phase, the w_{ij}^3 and w_{ij}^4 synapses onto the combination cells self-organize using Hebbian competitive learning rules, which enable the combination cell network to operate as a competitive network that learns to represent different combinations of a particular head direction and rotational velocity. The activation of the postsynaptic combination cell i is governed by

$$\begin{aligned}\tau^{\text{COMB}} \frac{\mathrm{d}h_i^{\text{COMB}}(t)}{\mathrm{d}t} &= -h_i^{\text{COMB}}(t) \\ &\quad - \frac{1}{N^{\text{COMB}}} \sum_j \tilde{w}^{\text{COMB}} r_j^{\text{COMB}}(t) \\ &\quad + \frac{\phi_3}{C^{\text{HD}\to\text{COMB}}} \sum_j w_{ij}^3(t) r_j^{\text{HD}}(t) \\ &\quad + \frac{\phi_4}{C^{\text{ROT}\to\text{COMB}}} \sum_j w_{ij}^4(t) r_j^{\text{ROT}}(t)\end{aligned} \tag{10}$$

with the terms de ned as follows. The term $\sum_j \tilde{w}^{\text{COMB}} r_j^{\text{COMB}}(t)$ represents inhibitory feedback within the combination cell network, where the summation is performed over all postsynaptic combination cells j, $\tilde{w}^{\text{COMB}}$ is the global lateral inhibition constant describing the effect of inhibitory interneurons within the combination cell network, and $N^{\text{COMB}} = 1000$ is the total number of combination cells in the model. The term $\sum_j w_{ij}^3(t) r_j^{\text{HD}}(t)$ is the input from the head direction cells, where $r_j^{\text{HD}}(t)$ is the ring rate of presynaptic head direction cell j, and $w_{ij}^3(t)$ is the corresponding strength of the synapse from this cell.[3] The term $\sum_j w_{ij}^4(t) r_j^{\text{ROT}}(t)$ is the input from the rotational velocity cells, where the ring rate of presynaptic rotational velocity cell j is given by $r_j^{\text{ROT}}(t)$, and $w_{ij}^4(t)$ is the strength of the corresponding synapse from this cell.[4] Activity within the combination cell network is driven by the head direction cell network if, and only if, the rotational velocity cells are also active. If the rotational velocity cells cease ring, i.e. the agent is stationary, then the activity in the combination cell network decays to zero according to the term $-h_i^{\text{COMB}}(t)$ and the time constant τ^{COMB}.

The ring rate $r_i^{\text{COMB}}(t)$ of postsynaptic combination cell i is determined from the activation $h_i^{\text{COMB}}(t)$ and the sigmoid activation function

$$r_i^{\text{COMB}}(t) = \frac{1}{1 + \mathrm{e}^{-2\beta(h_i^{\text{COMB}}(t) - \alpha)}} \tag{11}$$

where α and β are the sigmoid threshold and slope, respectively. The threshold α is set to a high value to ensure that each individual postsynaptic combination cell will function similar to a logical AND gate. That is to say that temporally conjunctive inputs from the presynaptic head direction cells and presynaptic rotational velocity cells are required in order for the postsynaptic combination cells to re. In this manner, individual combination cells become selective to combinations of a particular head direction occurring a small time interval Δt in the past and a rotational velocity, and, thus, a competitive network emerges.

The synaptic weights $w_{ij}^3(t)$ from the head direction cells to the combination cells are updated during learning according to

$$\frac{\mathrm{d}w_{ij}^3(t)}{\mathrm{d}t} = k^3 r_i^{\text{COMB}}(t) r_j^{\text{HD}}(t) \tag{12}$$

which increases the strength of the synapses between co-active presynaptic head direction cells and postsynaptic combination cells. In order to bound the synaptic weights, rescaling was employed after each timestep of the learning phase to ensure that for each postsynaptic combination cell i, we have

$$\sqrt{\sum_j (w_{ij}^3(t))^2} = 1 \tag{13}$$

where the sum is over all presynaptic head direction cells j.

The synaptic weights $w_{ij}^4(t)$ from the rotational velocity cells to the combination cells are updated during learning according to

$$\frac{\mathrm{d}w_{ij}^4(t)}{\mathrm{d}t} = k^4 r_i^{\text{COMB}}(t) r_j^{\text{ROT}}(t) \tag{14}$$

and the weights are bound by rescaling after each timestep of learning to ensure that for each postsynaptic combination

[3] The scaling factor $\frac{\phi_3}{C^{\text{HD}\to\text{COMB}}}$ controls the overall strength of the inputs from the head direction cells, where ϕ_3 is a constant, and $C^{\text{HD}\to\text{COMB}}$ is the number of synapses onto each postsynaptic combination cell from the presynaptic head direction cells.

[4] The scaling factor $\frac{\phi_4}{C^{\text{ROT}\to\text{COMB}}}$ controls the overall strength of the inputs from the rotational velocity cells, where ϕ_4 is a constant, and $C^{\text{ROT}\to\text{COMB}}$ is the number of synapses onto each postsynaptic combination cell from the presynaptic rotational velocity cells.

Table 2 Simulation parameter values (values are constant across all experiments except where specified)

Network parameters	
No. HD cells	500
No. COMB cells	1000
No. ROT cells	500
No. w^1 synapses onto each HD cell	500
No. w^2 synapses onto each HD cell	1000
No. w^3 synapses onto each COMB cell	25
No. w^4 synapses onto each COMB cell	500
$\tilde{w}^{HD}$	375
$\tilde{w}^{COMB}$	50
I^{EXTERN}	150
σ^{HD}	20°
Learning rates k^1, k^2, k^3, k^4	0.1
λ	200.0
τ^{HD}	1.0 ms
τ^{COMB} (Experiments 1 and 2)	150.0 ms
τ^{COMB} (Experiments 3 and 4)	100.0 ms
ϕ_1	3.75×10^3
ϕ_2	2.5×10^3
ϕ_3	5×10^3
ϕ_4	4×10^2
HD sigmoid transfer function parameters	
α	0.0
β	1.5
COMB sigmoid transfer function parameters	
α	10.0
β	1.5
Training parameters	
No. Training epochs	50
Speed of rotation (Experiments 1 and 3)	360°/s
Speed of rotation (Experiments 2 and 4)	180°/s

HD Head direction, *COMB* Combination, *ROT* Rotational velocity

cell i we have

$$\sqrt{\sum_j (w_{ij}^4(t))^2} = 1 \tag{15}$$

where the sum is over all presynaptic rotational velocity cells j.

During training, the w_{ij}^3 and w_{ij}^4 synapses onto the postsynaptic combination cells self-organize using competitive learning to enable the combination cells to learn to represent combinations of particular head directions and rotational velocities.

In the models simulated, we set the neuronal time constant τ^{COMB} for the combination cells to a large value as given in Table 2. This was due to the results from some preliminary simulations, which indicated that having a relatively large time constant τ^{COMB} compared to the neuronal time constant τ^{HD} for the head direction cells would produce the best operation of the model. The large value of the time constant τ^{COMB} means that, through learning, the current activity in the combination cell network will be associated with activity that occurred Δt in the past in the head direction cell network due to the effective time delay Δt between the activity packet in the head direction cell network and the activity in the combination cell network. Thus, after unsupervised learning, different cells in the combination cell network will respond to different combinations of a particular head direction that occurred Δt in the past and a rotational velocity (i.e. direction of movement through the head-direction space).

2.3 Training and testing

In order to keep the simulation run-time reasonable, the models were simulated with a relatively small architecture containing 500 head direction cells and 1000 combination cells. There were also 500 rotational velocity cells—with 250 of the cells (1–250) responding to clockwise rotation of the agent, and the remaining 250 cells (251–500) responding to counter-clockwise rotation. The head direction cells were mapped onto a regular grid of different head directions, such that each postsynaptic head direction cell i had a preferred head direction x_i of the agent at which the cell would be maximally stimulated by the visual input.

Throughout the training phase, the agent moved round a 1D circular space. A full clockwise rotation of the agent through consecutive positions 0–360°, followed by a full counter-clockwise rotation of the agent through consecutive positions 360–0°, constituted 1 epoch of training. The training phase was complete after 50 epochs had been performed. As the agent rotated during training, the activations and firing rates of the head direction cells were simulated according to the dynamical Eqs. 4 and 5.

During training in the light, the external visual inputs I_i^{EXTERN} included in Eq. 4, were the dominant influence upon the firing of the head direction cells. Since each head direction cell is tuned to fire maximally to visual input from a particular head direction, for each postsynaptic cell i the visual input e_i was set to the following Gaussian response profile

$$e_i^{HD} = \lambda e^{-(s_i^{HD})^2/2(\sigma^{HD})^2} \tag{16}$$

where s_i^{HD} is the difference between the actual head direction x of the agent and the preferred head direction x_i for postsynaptic head direction cell i, λ is a scaling factor that expresses the strength of the non-modifiable visual input synapses onto the postsynaptic head direction cells, and σ^{HD} is the standard deviation. For each postsynaptic head direction

cell i, the difference s_i^{HD} is given by

$$s_i^{HD} = \text{MIN}\,(|x_i - x|, 360 - |x_i - x|). \quad (17)$$

In the training phase of the simulations, the activity of the combination cells is driven by the w_{ij}^3 synaptic inputs from the presynaptic head direction cells, and the w_{ij}^4 synaptic inputs from the rotational velocity cells. One of the purposes of the training phase was to determine that the model can self-organize, through competitive learning, to correctly set up the combination cells to respond to combinations of a particular head direction occurring Δt in the past and clockwise or counter-clockwise rotation. Throughout the training, the activations and the ring rates of the combination cells were simulated according to the dynamical Eqs. 10 and 11.

At the beginning of the training phase, the synaptic weights w_{ij}^1, w_{ij}^2, w_{ij}^3 and w_{ij}^4 were all initialized to random positive values. These weights were updated at every timestep of the training phase according to Eqs. 6, 8, 12, and 14, respectively, and thus allowed to self-organize.

After the training phase was over, the simulations continued with the testing phase, during which the packet of head direction cell activity was required to update on the basis of idiothetic signals alone (the visual inputs t_i were set to zero). During the testing phase, the dynamical Eqs. 4, 5, 10 and 11 were simulated, but without any learning in any of the synapses i.e. Eqs. 6, 8, 12 and 14 were not simulated.

At the start of the testing phase, all of the ring rates r_i^{HD}, r_i^{COMB} and r_i^{ROT} were set to zero. The agent was oriented to an initial head direction and simulated with visual input available, but with no rotational velocity cells being active, for a period of 1 s. For the period that the agent maintained this head direction, the visual input term t_i for each postsynaptic head direction cell i was set to a Gaussian response pro le identical to that used by the head direction cell during the training phase as given by Eq. 16. The visual input was then removed by setting all the t_i terms to zero, and the agent was allowed to rest in the same initial head direction for a further 1 s. The purpose of this initial part of the testing was to allow the development within the head direction cell network of a stable packet of activity (representing the initial head direction).

The agent was allowed to remain stationary for a further 1 s, then the ring rates r_i^{ROT} of the 250 clockwise rotational velocity cells were set to 1 (fully active) for 1 s. The co- ring of the rotational velocity cells with the head direction cells stimulated a subset of the cells within the combination cell network. The ring of these combination cells then stimulated cells in the head direction cell network representing head directions in the clockwise direction of rotation, and the packet of head direction cell activity moved through the continuous attractor. This part of the testing phase was to ascertain that the representation of current head direction can be updated on the basis of idiothetic signals alone.

The ring rates r_i^{ROT} of all the rotational velocity cells were then set to 0 for 1 s. As the 250 clockwise rotational velocity cells ceased ring, the driving input to the combination cells disappeared, and the combination cells also ceased ring. In consequence, the combination cell network no longer provided a driving input to the head direction cell network, and the packet of head direction cell activity remained stationary at the last-visited head direction. This was to demonstrate that the head direction cell network could maintain a stable packet of activity representing any given head direction.

The ring rates r_i^{ROT} of the 250 counter-clockwise rotational velocity cells were then set to 1 for 1 s. In an identical manner to the period of clockwise rotation, the head direction cell activity packet then updated in a counter-clockwise direction, to demonstrate that counter-clockwise rotation could also be achieved on the basis of idiothetic signals alone.

The ring rates r_i^{ROT} of all the rotational velocity cells were then set to 0 for a nal 1 s. Again, this was to demonstrate that the head direction cell network could maintain a stable packet of activity at the last-visited head direction.

In the simulations, all the differential Eqs. 4, 6, 8, 10, 12, 14 were approximated by Forward Euler nite difference schemes with a timestep of 0.0001 s.

3 Simulations

3.1 Experiment 1 τ^{COMB} 150 ms; 360°/s rotational velocity

In this experiment, the model was simulated with the agent rotating at a velocity of 360°/s during training in the presence of visual input. The model parameters used for this experiment are given in Table 2. Testing in the absence of visual input was then carried out to determine whether the model had learned to update a packet of head direction cell activity at the same speed as the rotational velocity imposed during training. The results from the testing phase are shown in Figs. 3 to 7, and also in Table 3.

Figure 3 displays the recurrent synaptic weights w_{ij}^1 within the head direction cell network after training with the corresponding Hebbian associative learning rule (Eq. 6) and weight normalization (Eq. 7). Each of the four plots shows the learned synaptic weights to a different postsynaptic head direction cell from the 500 presynaptic head direction cells. Within each plot, the presynaptic head direction cells are arranged according to where they re maximally in the head-direction space of the agent when visual input is available. In each plot, a dashed vertical line indicates the presynaptic head direction cell with which the postsynaptic head direction cell has maximal w^1 synaptic weight. For all the plots, the syn-

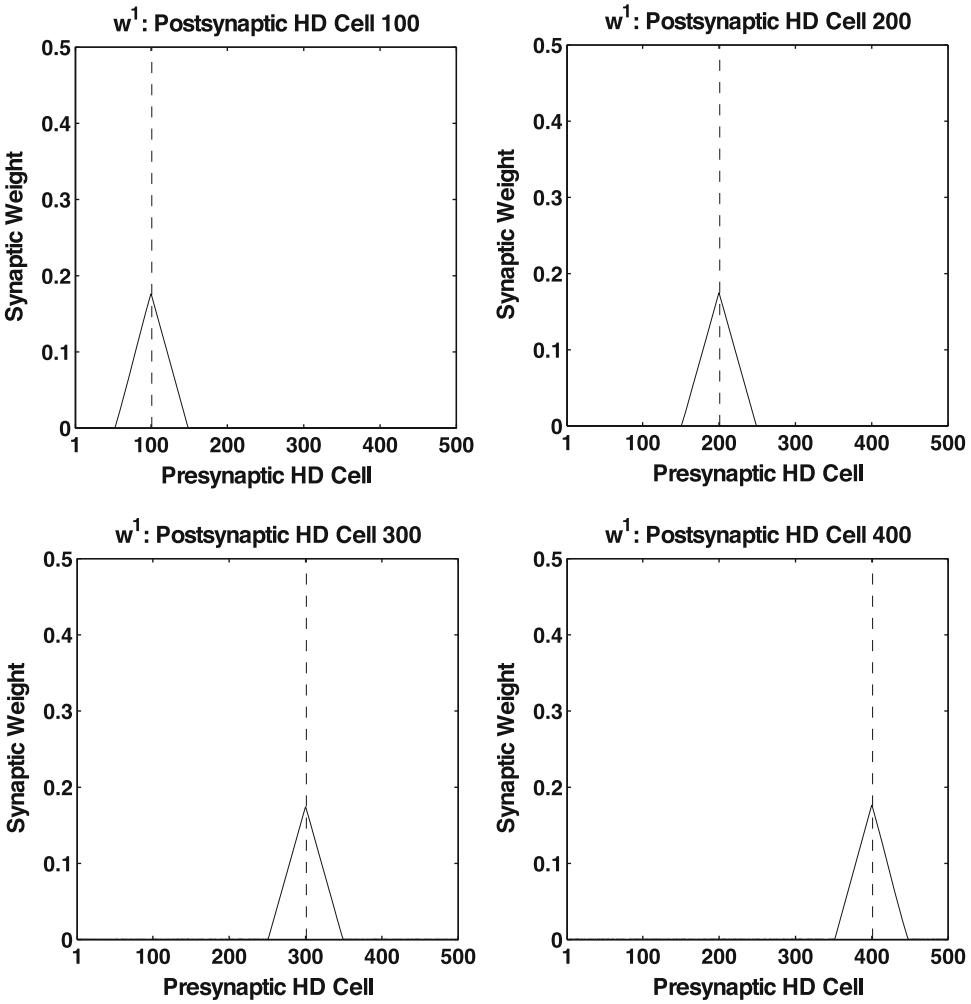

Fig. 3 The recurrent synaptic weights w^1 within the network of head direction (HD) cells after training with the Hebbian associative learning rule (6) and weight normalization (7). These results are from a simulation with a time constant τ^{COMB} of 150 ms, and a rotational velocity during training of 360°/s (all other parameters are as given in Table 2). Each of the four plots shows the learned synaptic weights to a different postsynaptic HD cell from the other 500 presynaptic HD cells in the network. In the plots, the 500 presynaptic HD cells are arranged according to where they are maximally in the head-direction space of the agent when visual input is available. For each plot, a dashed vertical line indicates the presynaptic HD cell with which the postsynaptic HD cell has maximal w^1 synaptic weight. In all the plots, the synaptic weight profile is symmetric about the presynaptic HD cell with maximal synaptic strength, and this symmetry helps in supporting a stable packet of HD cell activity during testing in the absence of visual input

aptic weight profile across the presynaptic head direction cells is clearly symmetric about the individual presynaptic head direction cell with maximal w^1 synaptic strength. This symmetry ensures that the head direction cell network can maintain a stable packet of head direction cell activity when the agent is stationary in the absence of visual input.

In Fig. 4, the plots display the w^3 synaptic weights from the head direction cell network to the combination cell network after competitive learning with the corresponding associative Hebbian learning rule (12) and weight normalization (13). Each individual plot shows the learned synaptic weights to a different postsynaptic combination cell from the 500 presynaptic head direction cells, with the presynaptic head direction cells arranged in the plots according to where they are maximally in the head-direction space of the agent in the presence of visual input. For each plot, a dashed vertical line indicates the presynaptic head direction cell with which the postsynaptic combination cell has maximal synaptic strength. Diluted synaptic connectivity was implemented for the w^3 synapses to ensure that competitive learning was preserved in the combination cell network. Thus, individual postsynaptic combination cells learned to respond to the combination of a particular head direction occurring in the past and a rotational velocity, rather than all possible head directions due to the continuity of the head-direction space and the overlapping receptive fields of the head direction cells (Stringer and Rolls 2006). With the exception of this diluted connectivity, each of the synaptic weight profiles is centred on a region of similarly tuned head direction cells and is approximately symmetric about the presynaptic head direction cell with maximal w^3 synaptic weight. The symmetric profile demonstrates that individual postsynaptic combination cells have learned to be preferentially stimulated by a subset of presynaptic head direction cells representing a preferred head direction Δt in the past. Since the model parameters ϕ_3, ϕ_4 and the threshold α of the combination cell sigmoid transfer function (11) are tuned to ensure that a strong rotational velocity cell input through the w^4 synapses is also required to fire individual postsynaptic combination cells, these cells can be said to learn to respond to combinations of a particular head direction Δt in the past and clockwise or counter-clockwise rotational velocity.

The plots in Fig. 5 display the synaptic weights w^2 from the combination cell network to the head direction cell net-

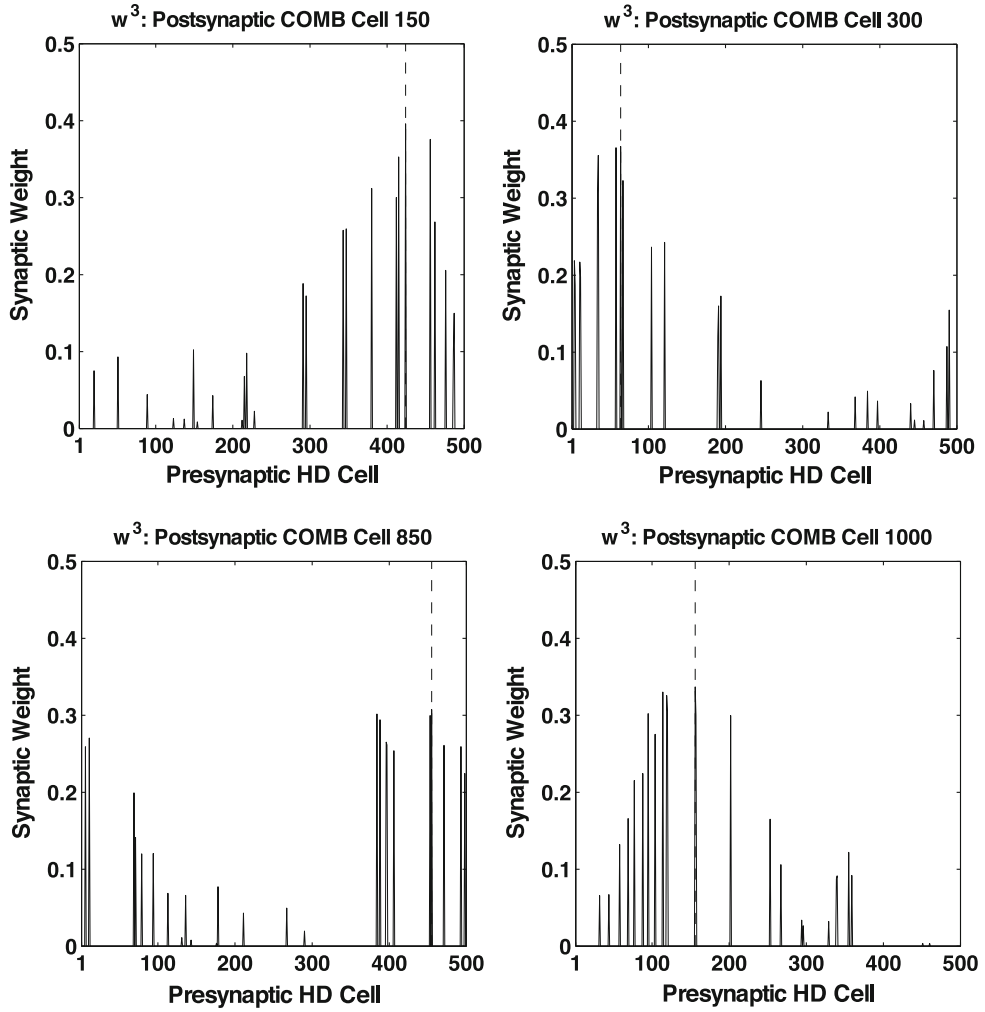

Fig. 4 The synaptic weights w^3 from the head direction (HD) cell network to the combination (COMB) cell network after competitive learning with the Hebbian associative learning rule (12) (and weight normalization (13). These results are from a simulation with a time constant τ^{COMB} of 150 ms, and a rotational velocity of 360°/s (all other parameters are as given in Table 2). Each of the four plots shows the learned synaptic weights to a different postsynaptic COMB cell from the 500 presynaptic HD cells. The presynaptic HD cells are arranged in the plots according to where they are maximally in the head-direction space of the agent when visual input is available. For each plot, a dashed vertical line indicates the presynaptic HD cell with which the postsynaptic COMB cell has maximal w^3 weight. Except for the effects of diluted synaptic connectivity, each of the weight profiles is centred on a region of similarly tuned HD cells, with a profile that is approximately symmetric about the presynaptic HD cell with maximal synaptic strength. Thus, the learned w^3 synaptic weights show that individual COMB cells learn to receive maximal stimulation from particular HD cells. Given that the model parameters ϕ_3, ϕ_4, and the threshold α of the COMB cell sigmoid transfer function (11) are tuned to ensure that a strong rotational velocity cell input through the w^4 synapses is also needed to fire the COMB cells, these cells in fact learn to respond to combinations of a particular head direction and clockwise or counter-clockwise rotational velocity

work after learning with the corresponding Hebbian associative learning rule (8) and weight normalization (9). In each plot, the learned synaptic weights are shown from a different presynaptic combination cell to the 500 postsynaptic head direction cells, with the head direction cells arranged in the plots according to where they are maximally in the head-direction space of the agent when visual input is available. The dashed vertical lines indicate, for each plot, the postsynaptic head direction cell with which the presynaptic combination cell has maximal w^3 weight as per Fig. 4. For all four plots, the w^2 synaptic weight profile is asymmetric about the postsynaptic head direction cell with maximal w^3 weight. This asymmetry shows that the presynaptic combination cell has learned to preferentially stimulate a postsynaptic head direction cell that represents a different head direction to the head direction cell from which the combination cell receives maximal w^3 stimulation. Thus, the asymmetry reflects the fact that, during training in the presence of visual input, the

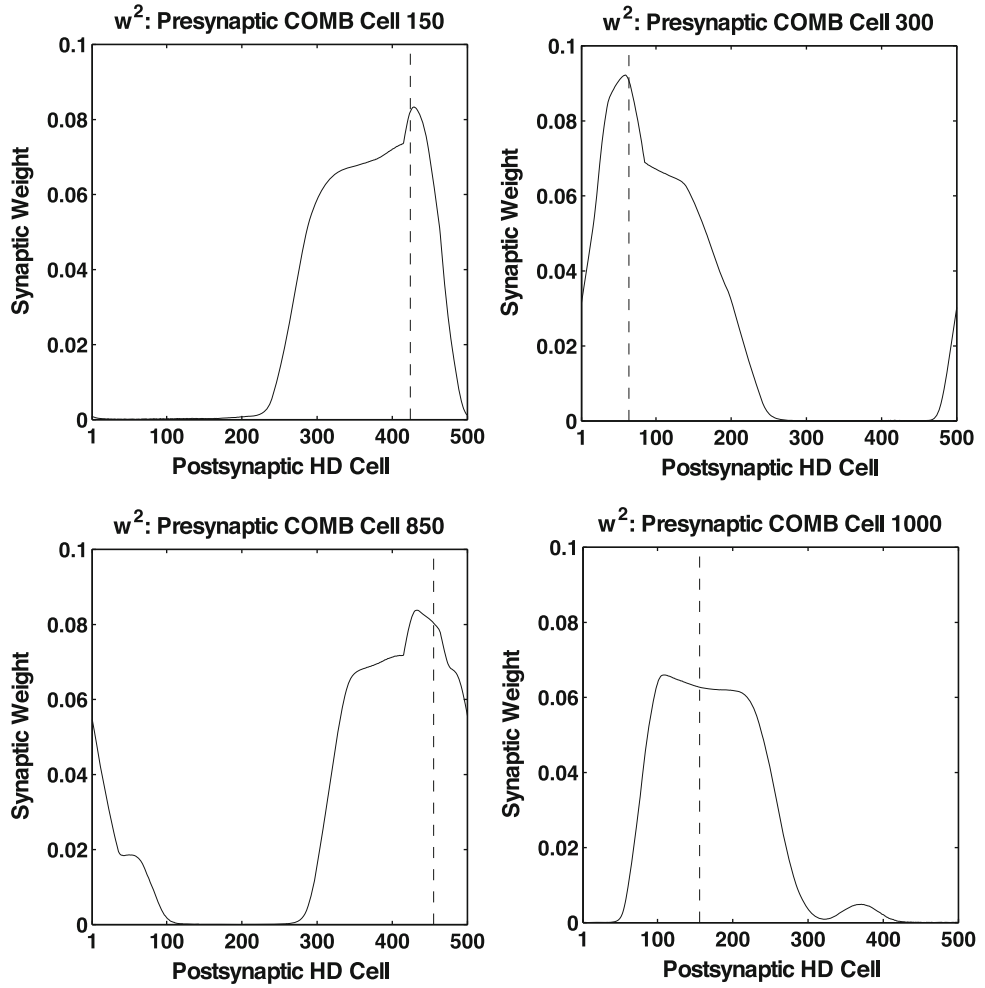

Fig. 5 The synaptic weights w^2 from the combination (COMB) cell network to the head direction (HD) cell network after learning with the Hebbian associative learning rule (8) and weight normalization (9). These results are from a simulation with a time constant τ^{COMB} of 150 ms, and a rotational velocity of 360°/s (all other parameters are as given in Table 2). Each of the four plots shows the learned synaptic weights from a different presynaptic COMB cell to the 500 postsynaptic HD cells. The postsynaptic HD cells are arranged in the plots according to where they are maximally in the head-direction space of the agent when visual input is available. For each plot, a dashed vertical line indicates the postsynaptic HD cell with which the presynaptic COMB cell has maximal w^3 weight as shown in Fig. 4. In each plot, the w^2 synaptic weight profile is asymmetric about the postsynaptic HD cell with maximal w^3 synaptic weight, indicating that the presynaptic COMB cell preferentially stimulates an HD cell representing a different head direction to the HD cell from which the COMB cell receives maximal w^3 synaptic weight. This reflects the fact that the packet of HD cell activity will have moved through the head-direction space of the agent in the time delay Δt between the firing of the presynaptic head direction cells and the resulting firing of the postsynaptic combination cells. Thus, the effective natural time delays produced by the neuronal time constants τ^{COMB} and τ^{HD} act, through learning, as a timing mechanism that enables the update of the packet of HD cell activity at the same speed as the agent is rotating

packet of head direction cell activity (representing current head direction) will have moved through the head-direction space of the agent in the natural time delay Δt between the activity profile in the head direction cell network and the activity profile in the combination cell network. This results, at any given moment in time, in the most active combination cell learning to stimulate a different postsynaptic head direction cell to the presynaptic head direction cell it has learned to be stimulated by, reflecting the changing head direction of the agent. Therefore, the neuronal time constants τ^{COMB} and τ^{HD} act, through learning, as a timing mechanism that is sufficient for the purpose of enabling the update of the packet of head direction cell activity at the same speed as the agent is rotating.

Figure 6 displays the w^4 synaptic weights from the layer of rotational velocity cells to the combination cell network after learning with the corresponding Hebbian associative learning rule (14) and weight normalization (15). Each of the plots shows the learned synaptic weights to a different postsynaptic combination cell from the 500 presynaptic rotational veloc-

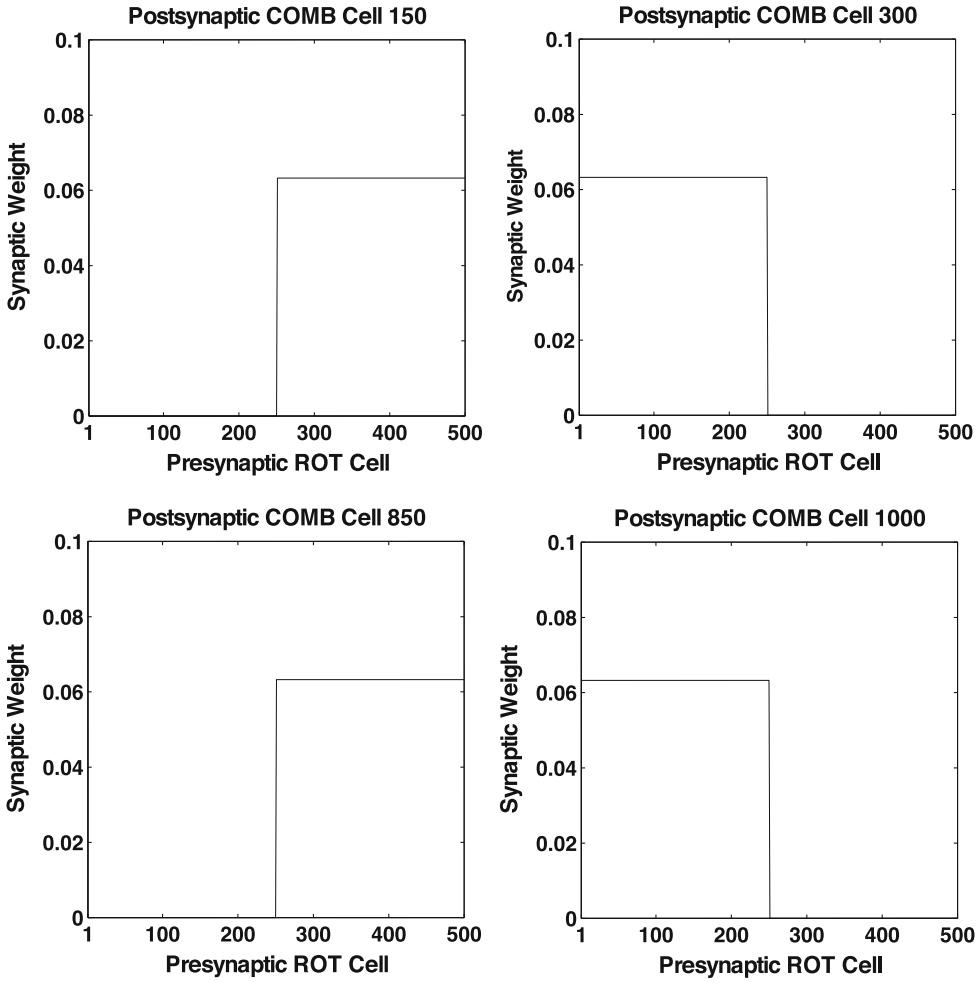

Fig. 6 The synaptic weights w^4 from the layer of rotational velocity (ROT) cells to the combination (COMB) cell network after learning with the Hebbian associative learning rule (14) and weight normalization (15). These results are taken from a simulation with a time constant τ^{COMB} of 150 ms, and a rotational velocity of 360°/s (all other parameters are as given in Table 2). Each of the four plots shows the learned synaptic weights to a different postsynaptic COMB cell from the 500 presynaptic ROT cells. As the ring pro le of the presynaptic ROT cells is binary, the w^4 synaptic weight pro les in the plots above can be described as a step function of the presynaptic ROT cell ring rate $r_j^{\mathrm{ROT}}(t)$. Combination cells 300 and 1000 have learned to respond to the clockwise direction of rotation. Combination cells 150 and 850 have learned to respond to the counter-clockwise direction of rotation

ity cells. Since the ring pro le of the presynaptic rotational velocity cells is binary, with each cell signalling that the agent is either rotating in a particular direction (cell fully ring) or is not (cell not ring), the w^4 synaptic weight pro les can be described as a step function of the presynaptic rotational velocity cell ring rate $r_j^{\mathrm{ROT}}(t)$. As the four plots display, each postsynaptic combination cell receives positive synaptic weights from the subset of exactly 250 presynaptic rotational velocity cells that signal either clockwise or counter-clockwise rotation (but not both subsets). Combination cells 300 and 1000 have learned to respond to the 250 rotational velocity cells representing clockwise rotation. Combination cells 150 and 850 have learned to respond to the 250 rotational velocity cells representing counter-clockwise rotation. Thus, the postsynaptic combination cells have learned to be maximally stimulated by a particular head direction occurring Δt in the past and a particular rotational velocity. The postsynaptic combination cells will, in turn, stimulate a different postsynaptic head direction to the presynaptic head direction cell that preferentially stimulates them; but this will only occur if the rotational velocity cells are temporally co- ring with the head direction cells.

The ring rates of the head direction, combination and rotational velocity cells are shown in the plots in Fig. 7. The top left plot displays the ring rates of the head direction cells during the training phase of the experiment. Throughout the training, the activity in the head direction cells is driven by the presence of external visual input. In the time interval 0.0–2.25 s, the agent rotated in a clockwise direction, with counter-clockwise rotation occurring in the time interval 2.25–4.5 s. The top right plot displays the ring rates of the head direction cells in the testing phase without external visual input. During the time interval 0.0–1.0 s, there was no ring in the rotational velocity cells (*bottom left plot*), and a moderate amount of baseline activity in the network of combination cells (*bottom right plot*); thus, there was a stable packet of head direction cell activity supported by the w^1 recurrent synapses. During the time interval 1.0–2.0 s, the 250 rotational velocity cells representing clockwise rotation became active (bottom left), stimulating activity in the combination cells (bottom right) through the w^4 synapses in conjunction with the head direction cell input through the w^3 synapses. Owing to the high value of the neuronal time constant τ^{COMB} and the asymmetry in the w^2 and w^3

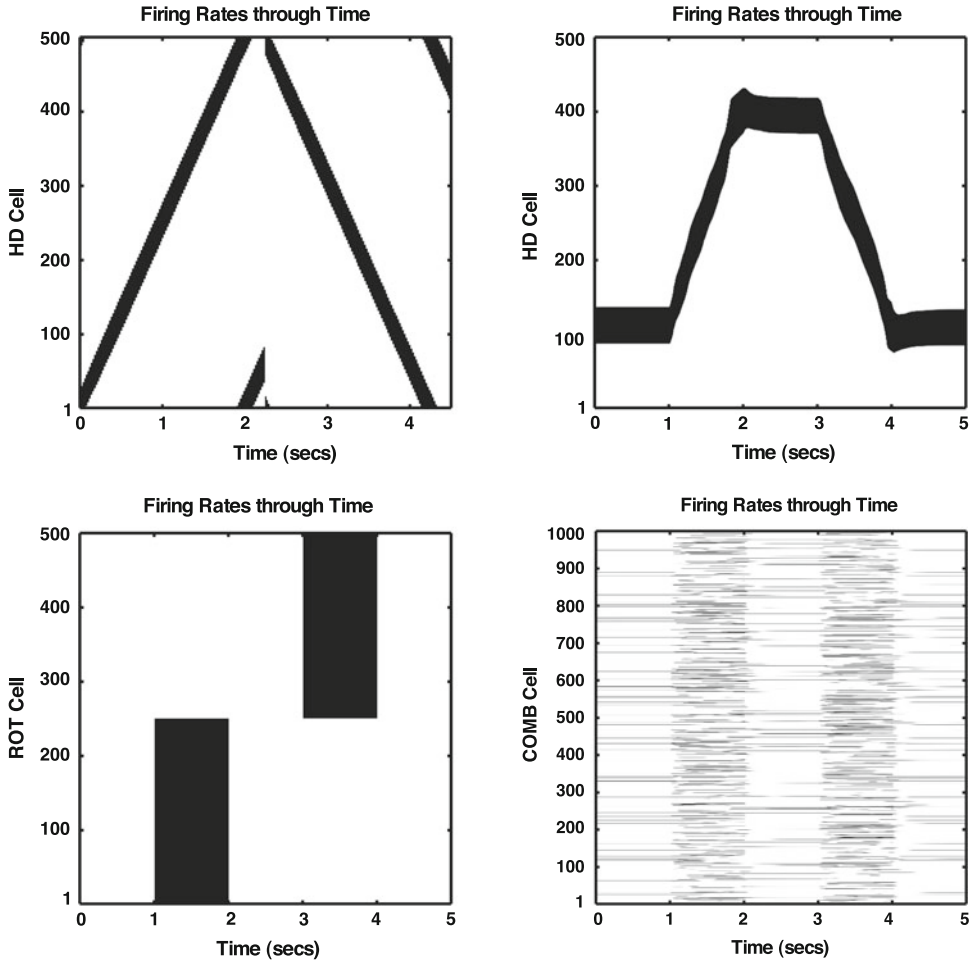

Fig. 7 Firing rates of head direction (HD), combination (COMB), and rotational velocity (ROT) cells during training and testing. These results are taken from a simulation with a time constant τ^{COMB} of 150 ms, and a rotational velocity of 360°/s (all other parameters are as given in Table 2). *Top left* Firing rates in the network of 500 HD cells during the 4.5 s of training, with the HD cells driven by visual input (0.0–2.25 s: agent rotating clockwise; 2.25–4.5 s: agent rotating counter-clockwise). *Top right* Firing rates in the network of 500 HD cells during the 5 s of testing in the absence of visual input. During the interval 0.0–1.0 s, there was a stable packet of activity in the HD cell network. In the interval 1.0–2.0 s, the ROT cells representing clockwise rotation were turned on and the packet of HD cell activity moved through the network in a clockwise direction. During the interval 2.0–3.0 s, the packet of HD cell activity remained stable in the network. In the interval 3.0–4.0 s, the ROT cells representing counter-clockwise rotation were turned on and the packet of HD activity moved through the network in a counter-clockwise direction. During the interval 4.0–5.0 s, the packet of HD cell activity remained stable in the HD cell network. *Bottom left*: Firing rates in the layer of 500 ROT cells during the 5 s of testing in the absence of visual input (1.0–2.0 s: ROT cells representing clockwise rotation are active; 3.0–4.0 s: ROT cells representing counter-clockwise rotation are active). *Bottom right* Firing rates in the network of 1000 COMB cells during the 5 s of testing in the absence of visual input. In the interval 1.0–2.0 s, the COMB cells become active due to the firing of the 250 ROT cells representing clockwise rotation of the agent. In the interval 3.0–4.0 s, the COMB cells become active due to the firing of the 250 ROT cells representing counter-clockwise rotation. At other periods, there is a moderate amount of baseline activity in the COMB cell network. In all plots, regions of high firing are represented by darker shading

synaptic weight profiles produced during learning, the firing of the combination cells during this period of the testing phase stimulated head direction cells representing head directions further along in the clockwise direction of rotation. Thus the head direction cell activity packet moved through the head-direction space of the agent, and the model performed velocity path integration of head direction. During the time interval 2.0–3.0 s, the rotational velocity cells were quiescent in their firing, and there was only a moderate amount of baseline activity in the combination cells; thus, the head direction cell activity packet remained stable in the head direction cell network. During the time interval 3.0–4.0 s, the 250 rotational velocity cells representing counter-clockwise rotation started firing (bottom left), in turn, stimulating activity in the combination cell network (bottom right). In an identical mechanism to that for clockwise rotation, the model thus performed velocity path integration of head direction, but this time in the counter-clockwise direction

Table 3 Speed of movement of the head direction activity packet during testing in the absence of visual input

	Clockwise	Counter-clockwise
τ^{COMB} 150 ms; 360°/s Rotational velocity		
Mean speed	180°/s	201°/s
Standard error	19.4°/s	8.0°/s
Percentage	50%	56.0%
τ^{COMB} 150 ms; 180°/s Rotational velocity		
Mean speed	125.1°/s	125.5°/s
Standard error	.9°/s	7.5°/s
Percentage	69%	69.7%
τ^{COMB} 100 ms; 360°/s Rotational velocity		
Mean speed	208.4°/s	224.1°/s
Standard error	15.0°/s	3.2°/s
Percentage	58%	62.3%
τ^{COMB} 100 ms; 180°/s Rotational velocity		
Mean speed	117.2°/s	104.2°/s
Standard error	.5°/s	4.3°/s
Percentage	65.1%	57.9%

For all of the four experiments, the results are computed over six simulation runs. Within an experiment, each of the six simulations had identical model parameters but different random synaptic connectivity and different random synaptic weight initializations. The table displays the mean speed of the activity packet across the six simulations during testing as calculated according to Eqs. 18 and 19, the standard error of the mean; and the mean speed of the activity packet as a percentage of the speed of rotation of the agent that was imposed during training in the presence of visual input. For each experiment, results are displayed for both clockwise and counter-clockwise rotations of the activity packet during testing

of rotation. During the time interval 4.0–5.0 s, the rotational velocity cells ceased ring, the activity level in the combination cell network returned to a baseline level and again there was a stable packet of activity in the head direction cell network.

In order to determine whether the model could perform velocity path integration of head direction at the same speed during testing as was imposed during training, the speed of update of the head direction cell activity packet was recorded. The measurement was taken according to

$$\text{speed} = \left| \frac{p_2 - p_1}{t_2 - t_1} \right| \tag{18}$$

where p_1 and p_2 represent the start and end positions respectively (in degrees) of the packet of head direction cell activity; and t_1 and t_2 represent the time (in s) at which the start and end packet positions were obtained. The packet positions were calculated as follows:

$$p = \frac{\sum_i r_i \theta_i}{\sum_i r_i} \tag{19}$$

where r_i is the ring rate of postsynaptic head direction cell i, and θ_i is the preferred head direction for postsynaptic head direction cell i in the presence of visual input.

Measurements of speed were taken for 0.5 s during both the clockwise (1.25–1.75 s) and counter-clockwise (3.25–3.75 s) periods of rotation during testing. (The recording was started 0.25 s after the rotational velocity cells started ring to allow time for the combination cells to become active, due to the high value of the neuronal time constant τ^{COMB}, and start driving the head direction cell activity packet through the head direction cell network.)

Six simulations were conducted, and in each simulation the same model parameters were employed, with the exception of different random synaptic connectivity and different random synaptic weight initialization. Table 3 summarizes the statistics we carried out on the results. The mean speed of rotation across the six simulations was 180°/s for clockwise rotation, and 201°/s for counter-clockwise rotation. This is compared to a true speed of 360°/s during training the presence of external visual input. Standard errors of 19.4°/s for clockwise rotation, and 8.0°/s for counter-clockwise rotation indicate a small amount of variation across the simulations, but the standard errors are still suf ciently small to conclude that the simulations produced speeds of rotation consistent enough with one another to produce reliable results. In order to further compare the recorded speed during testing to the speed imposed during training, the mean speed was calculated during testing as a percentage of the speed during training. For the period of clockwise rotation, the model updated the packet of head direction cell activity at a speed that was 50% of that imposed during training. For the period of counter-clockwise rotation, the speed was 56.0% of the speed during training. The model has learned

to achieve this automatically without the need to carefully hand-tune a parameter specifically governing the speed as was used by Stringer and Rolls (2006). We note, however, that while the speed at testing is a reasonable approximation (and within the same order of magnitude) of the speed imposed during training, the speeds during testing are regularly below those during training. In future study, we will investigate what architectural features contribute to this underestimation of the speed during path integration.

3.2 Experiment 2 τ^{COMB} 150 ms; 180°/s rotational velocity

We conducted this experiment to demonstrate that if the model was trained with the same parameter set but at half the rotational velocity compared to Experiment 1, then during the testing phase the model would still perform velocity path integration of head direction at the correct speed. Thus, the same model can learn and perform path integration at different rotational velocities. Except for the difference in rotational velocity, all the model parameters were the same as for Experiment 1 (as given in Table 2).

The plots in Fig. 8 display the firing rates of the head direction, combination, and rotational velocity cells during training and testing of the model. The conventions are the same as for Fig. 7, and so is the interpretation. When the same model is trained at half the rotational velocity compared to Experiment 1, the same mechanism of a high value for the neuronal time constant τ^{COMB} on the activations of the combination cells produces a natural time delay Δt between the activity profile in the head direction cell network and the activity profile in the combination cell network. During this delay, Δt, the head direction of the agent will have changed. Thus, at any given moment in time, the most active combination cell learns to preferentially stimulate the head direction cell representing the new head direction. This learned association will help to shift a packet of head direction cell activity through the head direction cell network. Much similar to Experiment 1, the current model also requires temporally conjunctive inputs to the combination cells through the w^3 synapses from the head direction cells, and the w^4 synapses from the rotational velocity cells, in order for the combination cells to start firing and, in turn, stimulate an activity packet in the head direction

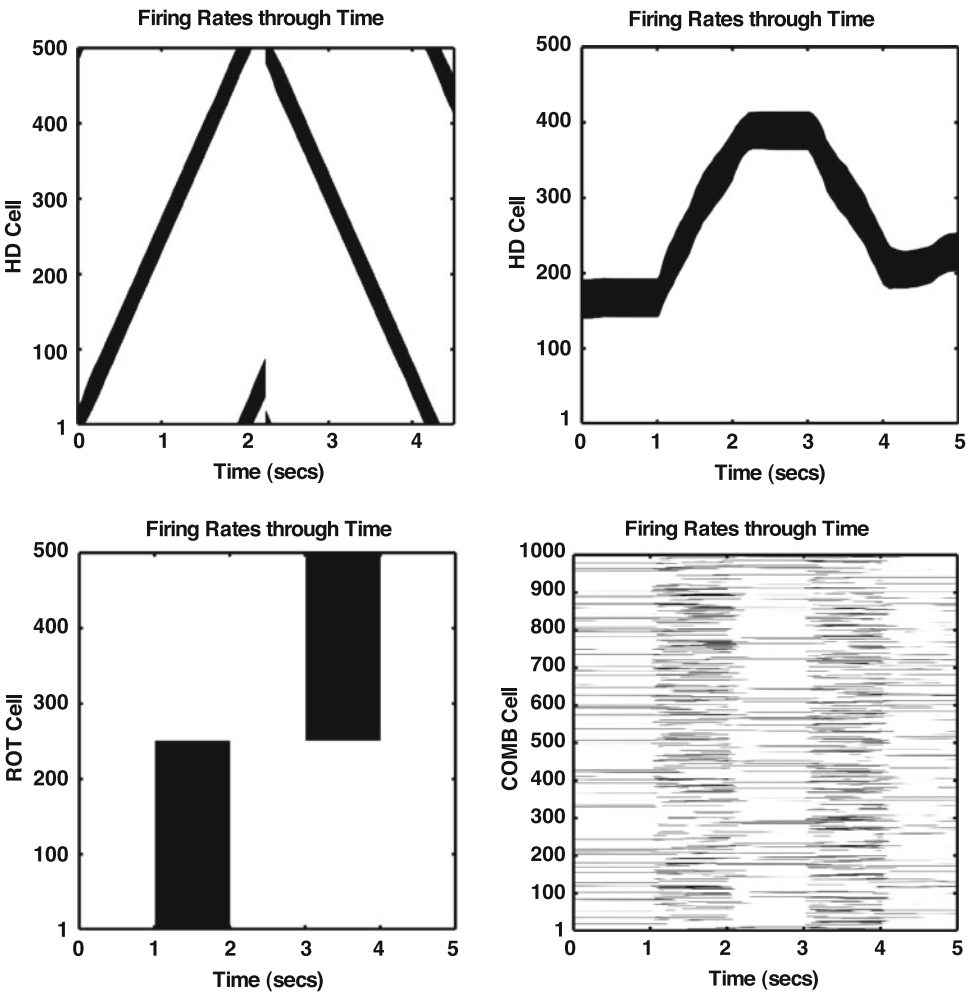

Fig. 8 Firing rates of head direction (HD), combination (COMB) and rotational velocity (ROT) cells during training and testing. These results are taken from a simulation with a time constant τ^{COMB} of 150 ms, and a rotational velocity of 180°/s (all other parameters are as given in Table 2). Conventions are as for Fig. 7

cell network (bottom left and right plots). When there is no ring in the layer of rotational velocity cells, there is only a moderate amount of baseline activity in the combination cell network, and, thus, there is a stable packet of activity in the head direction cell network (top right plot).

In order to determine whether the model had learned to update the packet of head direction cell activity at the same speed during testing as was imposed during training, we measured the speed of the update according to Eqs. 18 and 19. We took measurements across a 0.5 s interval for both clockwise and counter-clockwise rotation, and repeated these measurements across six different simulations that had identical model parameters with the exception of different random synaptic connectivity and different random synaptic weight initialization. The results we obtained are summarized in Table 3.

The mean speed of rotation was 122.5°/s for the period of clockwise rotation, and 125°/s for the period of counter-clockwise rotation. This is compared to the true speed during training of 180°/s. Standard errors of 7.7°/s for the clockwise measurements and 5.7°/s for the counter-clockwise measurements, indicate a small degree of variability across the six simulations, but the standard errors are small enough to conclude that the simulations produced speeds of rotation consistent enough with one another to be considered as reliable results. We also calculated the mean speed of the activity packet during testing as a percentage of the rotational velocity imposed during training. For the period of clockwise rotation, the mean speed during testing was 69.5% of the rotational velocity imposed during training. For the period of counter-clockwise rotation, the mean speed was 69.7% of the rotational velocity imposed during training. Similar to Experiment 1, the speeds recorded during testing of the model are a reasonable approximation to those experienced during training of the model, although there is a small but regular underestimation. Moreover, the same model with the same set of parameters is able to learn two completely different rotational velocities during training, and reproduce (within a small margin of error) those velocities during the testing phase.

3.3 Experiment 3: τ^{COMB} 100 ms; 360°/s rotational velocity

This experiment was conducted to demonstrate that the same model could learn to perform velocity path integration of head direction when implemented with a different value, compared to Experiments 1 and 2, for the neuronal time constant τ^{COMB}. (Thus, the principle of the neuronal time constants for the cell activations producing effective natural time delays Δt over which associations between cell activities can be learned is a general one.) Except for the value of the time constant τ^{COMB}, all the other model parameters are identical to those used for Experiment 1 (as given in Table 2).

The left plot in Fig. 9 displays the firing rates of the head direction cells during the 5 s of testing the model in the absence of visual input. The conventions are the same as for the top right plots in Figs. 7 and 8, and demonstrate that when there is no firing in the layer of rotational velocity cells, the combination cell network does not become active enough to shift an activity packet in the head direction cell network; thus, there is a stable packet of head direction cell activity. When the rotational velocity cells start firing (to signal either clockwise or counter-clockwise rotation), then the combination cells reach an activation level sufficient to shift an activity packet in the head direction cell network, and the model updates its representation of current head direction. The model has, thus, learned to perform velocity path integration of head direction.

We took measurements of the speed of update of the head direction cell activity packet according to equation 18 (and

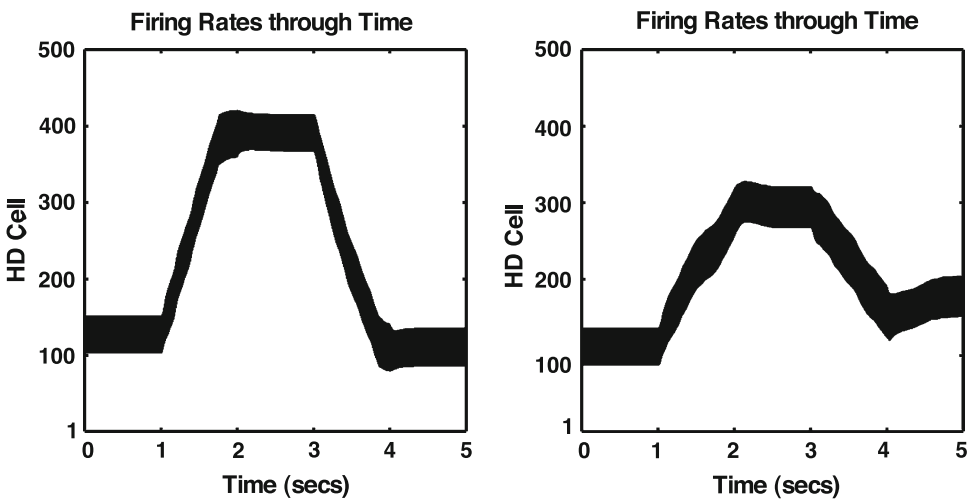

Fig. 9 Firing rates of head direction (HD) cells during testing. These results are taken from simulations with a time constant τ^{COMB} of 100 ms and rotational velocities during training of 360°/s (*Left plot*) and 180°/s (*Right plot*). Conventions are as for the *top-right plots* in Figs. 7 and 8

(19). The measurements were taken across six different simulations with identical model parameters except for different random synaptic connectivities and different random synaptic weight initializations. The results are summarized in Table 3. The mean speed of rotation was 250.4°/s during the period of clockwise rotation, and 212.4°/s during the period of counter-clockwise rotation. This is compared to the true rotational velocity imposed during training of 360°/s. Standard errors of 15°/s for the clockwise measurements and 32°/s for the counter-clockwise measurement, indicate a small degree of variation across the simulations. The standard errors are small enough, however, to conclude the recorded rotational velocities were consistent enough with one another to produce reliable results. The mean speed of rotation as a percentage of the rotational velocity imposed during training was calculated to be 69.8% for the period of clockwise rotation and 62.3% for the period of counter-clockwise rotation. The speeds recorded during testing are approximately the same as those that the model experienced during training, and although there is an underestimation for both clockwise and counter-clockwise rotations, the mean speeds are within the same order of magnitude as the speed imposed during training. The model can thus be implemented with a different value for the time constant τ^{COMB}, and can still learn (to within a small margin of error) to update the packet of head direction cell activity at the same speed it experienced during training.

3.4 Experiment 4: τ^{COMB} 100 ms; 180°/s rotational velocity

By using these values we aimed at demonstrating that the same model could be implemented with different values for the neuronal time constant τ^{COMB}, and could be trained at different rotational velocities when implemented with these different values, and in all the cases could still learn to perform velocity path integration of head direction. With the exception of a different value for the time constant τ^{COMB}, and a different rotational velocity, the model parameters are identical to those used in the previous experiments (as given in Table 2).

The right plot of Fig. 9 displays the firing rates of the head direction cells during the 5 s of testing in the absence of visual input. The conventions are the same as for the top right plot in Figs. 7 and 8, and show that when there is no firing in the layer of rotational velocity cells, the combination cells do not become active enough to stimulate the head direction cell network, and, thus, the packet of head direction cell activity remains stable. When the rotational velocity cells do begin to fire, the combination cells will, through learning, preferentially stimulate head direction cells representing head directions further along in the current direction of rotation. The model, with the same parameters as used in Experiment 3, has, thus, learned to perform velocity path integration of head direction at the new rotational speed of 180°/s.

We calculated the speed of update of the head direction cell activity packet according to Eqs. 18 and 19. The measurements were taken across six different simulations with identical model parameters except for different random synaptic connectivities and different random synaptic weight initializations. The results are summarized in Table 3. The mean speed of rotation was 127°/s during the period of clockwise rotation and 102.4°/s during the period of counter-clockwise rotation. This is compared to the true speed of 180.0°/s imposed upon the model during training. We calculated a standard error of 5.4°/s for the clockwise measurements, and 4.8°/s for the counter-clockwise measurements. The standard error indicate little variation in the speed of update of the activity packet across simulations, and, thus, the six simulations are consistent enough with one another to produce reliable results. The mean speed of clockwise rotation was calculated to be 69.5% of the true speed imposed during training. The mean speed of counter-clockwise rotation was calculated to be 57% of the speed imposed during training. For both directions of rotation, there is an underestimation of the true speed encountered during training. The speeds recorded are still within the same order of magnitude of the speed imposed during training, and, thus, sufficiently close to one another to conclude that the model can be implemented with different values for the time constant τ^{COMB}, and trained at different rotational velocities, and can, in each case, still learn (to within a small margin of error) to perform velocity path integration of head direction at the correct speed.

4 Discussion

The model that we presented in this article can learn to update a packet of head direction cell activity by the correct amount given an external input representing the velocity of self-motion. The principle of operation is that the neuronal time constants for the governing equations of the cell activations produce effective natural time delays between the activity profiles in the presynaptic cells and the postsynaptic cells, and it is this delay Δt over which associations between cell activities can be learned and replayed at the correct speed. In the current model, a high value for the neuronal time constant τ^{COMB} leads to a time delay Δt between the presynaptic head direction cell activity profile and the postsynaptic combination cell activity profile. During training in the presence of visual input, the head direction of the agent will have changed during the delay Δt; thus, at any given moment in time, the most active combination cell will strengthen its synapse with the head direction cell representing the new head direction, and this head direction cell will

be a different cell to the one which preferentially stimulates the combination cell. Essentially, the model learns to make associations between a current head direction and a combination of a particular head direction that occurred in the past and a rotational velocity.

In the four experiments we conducted, the model did not reproduce the exact speed during testing that was imposed during training. This could be due to the fact that our simulated models were relatively small, containing only 2000 neurons in total. This small architecture may be more susceptible to noise in the system due to an inability to average the noise out as well as a larger architecture, and the noise may lead to the packet of head direction cell activity updating at a slower speed than the model was trained with. A model with a larger architecture may, thus, show less difference between the speeds recorded during testing and those imposed during training, when compared to the results we presented in this article. Unfortunately, owing to the large simulation run-time, we were unable to investigate the performance in a larger network in this study. We hope to examine this issue further in future research.

We know of no previous model that can update a packet of head direction cell activity at the correct speed through self-organization with biologically plausible *associative* learning rules. Hahnloser (2003) used an *error correction* learning rule in a network with separate head direction sub-networks for each direction of idiothetic signal to produce a convergent learning scheme in a one-dimensional head direction cell system. One disadvantage of that model, however, is that the use of an error correction learning rule makes it less biologically plausible than the model presented in this article.

In order for the current model to operate correctly for different rotational velocities, it is necessary to have different rotational velocity cells tuned to those different velocities. This in turn requires that the model is trained for each different velocity. If a distributed representation of rotational velocity is implemented (with each rotational velocity cell tuned to a range of velocities and overlapping with its neighbours) then the model should generalize well across different velocities. We have shown in this article that the model can, without the need to alter any model parameters, learn to perform path integration when trained with different rotational velocities, and it would be interesting to confirm that the same model can be trained on different rotational velocities at the same time. Training with just a few velocities should suffice to allow the model to generalize over a range of velocities.

In the simulations of the two-layer network described above, the neuronal time constants τ^{COMB} implemented for the combination cells were much larger than the neuronal time constants τ^{HD} used for the head direction cells. In further simulations, not described in this paper, we also varied the values of the neuronal time constants for the head direction cells. In these simulations, we found that the accuracy of path integration during testing was significantly degraded by having larger values for the neuronal time constants for the head direction cells. Thus, we observed an asymmetry between the effects of the neuronal time constants for the combination cells and the head direction cells. The path integration mechanism described in this paper was successful only over the effective delays in neurotransmission from the head direction cells to the combination cells due to large τ^{COMB}, and not over delays in neurotransmission from the combination cells to the head direction cells due to large τ^{HD}. The reason for this is as follows.

The combination cells operate as a competitive network, and simply learn through the w^3 synapses to represent the head direction some time interval Δt in the past (when there is a rotation signal also present). However, the learning rule for the w^2 synapses must operate as a pattern association network, and associates the combination cell firing with the current pattern of activity in the head direction cell network imposed by the visual training signal. During testing, however, if the head direction cells have a large neuronal time constant, then there is a significant delay in neurotransmission from the combination cells to the head direction cells. This means that the combination cells are firing up the head direction cells too late to keep accurate track of the true head direction of the agent. The only way to ameliorate this inaccuracy is to implement a relatively small neuronal time constant for the head direction cells. Alternatively, we hypothesise that one way in which the path integration accuracy of the model might be improved is by introducing explicit time delays into the learning rules for the w^2 and w^3 synaptic weights, which will match the delays in neurotransmission through these synapses. This has already been investigated by Walters et al. 2009). These authors have simulated path integration in a similar network architecture with axonal transmission delays incorporated into the equations governing both the cell activations and the corresponding learning rules. These authors have reported more accurate path integration during testing.

We chose to vary the neuronal time constant τ^{COMB} across the experiments conducted, whilst keeping the neuronal time constant τ^{HD} constant, to clarify the principle that an increase in the value of the time constant τ^{COMB} is approximately proportional to an increase in the effective natural time delay Δt. Thus, the same model can learn to perform path integration when implemented with different time delays Δt. Moreover, the analytical argument given in Sect. 2.1 should apply to any individual three-cell circuit of head direction cell → combination cell → head direction cell over a broad range of possible neuronal time constants. Further, individual three-cell circuits should be able to operate in parallel with other three-cell circuits with different time constants, i.e. where

τ^{HD}, τ^{COMB} and τ^{HD} for the three cells are set to different values. Therefore, we hypothesise that the model should still self-organize and perform path integration at the correct speed if the neuronal time constants τ^{HD} and τ^{COMB} for individual head direction and combination cells were drawn from a distribution of time constants, instead of being homogeneous for each cell type. We plan to test this hypothesis in future simulations.

In the models simulated, we used values of 100 and 150 ms for the neuronal time constant τ^{COMB}, and a constant value of 1 ms across all simulations for the neuronal time constant τ^{HD}. A possible way in which the different time constants—the key principle of operation for the current model—could be implemented in the brain is through the use of AMPA receptor-mediated synaptic transmission for faster time constants (in the case of the current model, this would be the head direction cells), and NMDA receptor-mediated synaptic transmission for slower time constants (the combination cells in the current model). The time constant for AMPA currents is approximately 2 ms, while the time constant for NMDA currents is approximately 100–150 ms (Hestrin et al. 1990; Spruston et al. 1995; Brunel and Wang 2001). Thus, the values we have implemented are within the range of biological plausibility.

One of the key components of the current model is the presence of combination cells that learn to respond to combinations of a particular head direction and a rotational velocity. For the proposed architecture to be implemented in the brain, there would, thus, need to be classes of neurons that respond to combinations of spatial information and velocity. Previous research has revealed that there are neurons that respond to a combination of head direction and angular rotation velocity (Taube et al. 1990b; Sharp 1996; Bassett and Taube 2005). It has also been shown that hippocampal place cells can have their activity modulated by running speed (McNaughton et al. 1983), and can even be tuned to particular velocities (Wiener et al. 1989). The current theory provides an account of the presence of such cells in the hippocampus. The origin of these signals, and indeed a possible brain region for path integration of place to be implemented, is the dorsocaudal medial entorhinal cortex which contains spatial grid cells that are modulated by head direction and forward velocity (Hafting et al. 2005; Sargolini et al. 2006).

Finally, we suggest that path integration implemented in the way described in this article could be performed in other brain systems, including the hippocampal place cell system (O'Keefe and Dostrovsky 1971; McNaughton et al. 1983; Muller et al. 1991) and the hippocampal spatial view system of neurons that respond when a primate looks at a particular location in space, which are updated by idiothetic eye movements made in the dark (Robertson et al. 1998, 1999; Rolls 1999; Rolls and Xiang 2006).

Acknowledgement This research was supported by the Economic and Social Research Council and the Wellcome Trust.

References

Amari S (1977) Dynamics of pattern formation in lateral-inhibition type neural fields. Biol Cybern 27:77–87

Bassett J, Taube JS (2005) Head direction signal generation: ascending and descending information streams. In: Wiener SI, Taube JS (eds) Head direction cells and the neural mechanisms of Spatial orientation, pp 83–109. MIT Press, Cambridge

Brunel N, Wang X-J (2001) Effects of neuromodulation in a cortical network model of object workring memory dominated by recurrent inhibition. J Comput Neurosci 11:63–85

Collett TS, Zeil J (1998) Places and landmarks: an arthropod perspective. In: Healy S (ed) Spatial representation in animals. Oxford University Press, Oxford

Georges-François P, Rolls ET, Robertson RG (1999) Spatial view cells in the primate hippocampus: allocentric view not head direction or eye position or place. Cereb Cortex 9:197–212

Hafting T, Fyhn M, Molden S, Moser MB, Moser EI (2005) Microstructure of a spatial map in the entorhinal cortex. Nature 436:801–806

Hahnloser RHR (2003) Emergence of neural integration in the head-direction system by visual supervision. Neuroscience 120:877–891

Hestrin S, Sah P, Nicoll R (1990) Mechanisms generating the time course of dual component excitatory synaptic currents recorded in hippocampal slices. Neuron 5:247–253

McNaughton BL, Barnes CA, 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 52:41–49

Mittelstaedt H, Mittelstaedt ML (1982) Homing by path integration. In: Papi F, Wallraff HG (eds), Avian navigation, pp 290–297. Springer, Berlin

Mittelstaedt ML, Mittelstaedt H (1980) Homing by path integration in a mammal. Naturwissenschaften 67:566–567

Muller RU, Kubie JL, Bostock EM, Taube JS, Quirk GJ (1991) Spatial firing correlates of neurons in the hippocampal formation of freely moving rats. In: Paillard J (ed) Brain and space, pp 296–333. Oxford University Press, Oxford

Muller RU, Ranck JB Jr, Taube JS (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 15:267–273

O'Keefe J, Dostrovsky J (1971) The hippocampus as a spatial map: preliminary evidence from unit activity in the freely moving rat. Brain Res 34:171–175

Ranck JB Jr (1985) Head direction cells in the deep cell layer of dorsolateral presubiculum in freely moving rats. In: Buzsáki G, Vanderwolf CH (eds) Electrical activity of the archicortex. Akadémiai Kiadó, Budapest

Redish AD (1999) Beyond the cognitive map. MIT Press, Cambridge

Redish AD, Elga AN, Touretzky DS (1996) A coupled attractor model of the rodent head direction system. Netw Comput Neural Syst 7:671–685

Robertson RG, Rolls ET, Georges-François P (1998) Spatial view cells in the primate hippocampus: effects of removal of view details. J Neurophysiol 79:1145–1156

Robertson RG, Rolls ET, Georges-François P (1999) Head direction cells in the primate pre-subiculum. Hippocampus 9:206–219

Rolls ET, Treves A (1998) Neural networks and brain function. Oxford University Press, Oxford

Rolls ET (1999) Spatial view cells and the representation of place in the primate hippocampus. Hippocampus 9:467–480

Rolls ET, Xiang JZ (2006) Spatial view cells in the primate hippocampus, and memory recall. Rev Neurosci 17:175–200

Rolls ET, Robertson RG, Georges-François P (1997) Spatial view cells in the primate hippocampus. Eur J Neurosci 9:1789–1794

Samsonovich A, McNaughton BL (1997) Path integration and cognitive mapping in a continuous attractor neural network model. J Neurosci 17:5900–5920

Sargolini F, Fyhn M, Hafting T, McNaughton BL, Witter MP, Moser MB, Moser EI (2006) Conjunctive representation of position, direction, and velocity in entorhinal cortex. Science 312:758–762

Sharp PE (1996) Multiple spatial-behavioral correlates for cells in the rat postsubiculum: multiple regression analysis and comparison to other hippocampal areas. Cereb Cortex 6:238–259

Skaggs WE, Knierim JJ, Kudrimoti HS, McNaughton BL (1995) A model of the neural basis of the rat's sense of direction. In: Tesauro G, Touretzky DS, Leen TK (eds) Advances in neural information processing Systems, vol 7, pp 173–180. MIT Press, Cambridge

Song P, Wang X-J (2005) Angular path integration by moving "hill of activity": a spiking neuron model without recurrent excitation of the head-direction system. J Neurosci 25:1002–1014

Spruston N, Jonas P, Sakmann B (1995) Dendrite glutamte receptor channels in rat hippocampal CA3 and CA1 pyramidal neurons. J Physiol 482:325–352

Stringer SM, Rolls ET (2006) Self-organizing path integration using a linked continuous attractor and competitive network: path integration of head direction. Netw Comput Neural Syst 17:419–445

Stringer SM, Trappenberg TP, Rolls ET, De Araujo IET (2002) Self-organizing continuous attractor networks and path integration: One dimensional models of head direction cells. Netw Compuat Neural Syst 13:217–242

Taube JS, Muller RU, Ranck JB Jr (1990a) Head-direction cells recorded from the postsubiculum in freely moving rats. I. Description and quantitative analysis. J Neurosci 10:436–447

Taube JS, Muller RU, Ranck JB Jr (1990b) Head-direction cells recorded from the postsubiculum in freely moving rats. II. Effects of environmental manipulations. J Neurosci 10:436–447

Taylor JG (1999) Neural 'bubble' dynamics in two dimensions: foundations. Biol Cybern 80:393–409

Walters DM, Stringer SM, Rolls ET (2009) Path integration of head direction: updating a packet of neural activity at the correct speed using axonal conduction delays (submitted)

Wiener SI, Paul CA, Eichenbaum H (1989) Spatial and behavioural correlates of neuronal activity. J Neurosci 9:2737–2763

Zhang K (1996) Representation of spatial orientation by the intrinsic dynamics of the head-direction cell ensemble: a theory. J Neurosci 16:2112–2126