

Path Integration of Head Direction: Updating a Packet of Neural Activity at the Correct Speed Using Axonal Conduction Delays

Daniel Walters^{1*}, Simon Stringer¹, Edmund Rolls²

1 Oxford Centre for Theoretical Neuroscience and Artificial Intelligence, Department of Experimental Psychology, University of Oxford, United Kingdom, **2** Oxford Centre for Computational Neuroscience, Oxford, United Kingdom

Abstract

The head direction cell system is capable of accurately updating its current representation of head direction in the absence of visual input. This is known as the path integration of head direction. An important question is how the head direction cell system learns to perform accurate path integration of head direction. In this paper we propose a model of velocity path integration of head direction in which the natural time delay of axonal transmission between a linked continuous attractor network and competitive network acts as a timing mechanism to facilitate the correct speed of path integration. The model effectively learns a “look-up” table for the correct speed of path integration. In simulation, we show that the model is able to successfully learn two different speeds of path integration across two different axonal conduction delays, and without the need to alter any other model parameters. An implication of this model is that, by learning look-up tables for each speed of path integration, the model should exhibit a degree of robustness to damage. In simulations, we show that the speed of path integration is not significantly affected by degrading the network through removing a proportion of the cells that signal rotational velocity.

Citation: Walters D, Stringer S, Rolls E (2013) Path Integration of Head Direction: Updating a Packet of Neural Activity at the Correct Speed Using Axonal Conduction Delays. PLoS ONE 8(3): e58330. doi:10.1371/journal.pone.0058330

Editor: William W. Lytton, SUNY Downstate MC, United States of America

Received: September 7, 2012; **Accepted:** February 2, 2013; **Published:** March 19, 2013

Copyright: © 2013 Walters et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Funding: This research was supported by the Economic and Social Research Council (ESRC) and the Wellcome Trust. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

* E-mail: daniel.walters@psy.ox.ac.uk

Introduction

Head direction cells signal the orientation of the animal's head in the horizontal plane [1–3]. In the absence of guiding visual input, a network of head direction cells will accurately represent the current head direction of the animal [3–5]. This is the path integration of head direction, where an animal integrates idiothetic (self-motion) signals to track the current orientation of its head within an environment [6,7].

In many neural network models of the head direction cell system, the head direction cells conceptually form a ring representing the spatial continuum of head directions within the one-dimensional head-direction space. The position of the peak of a single, often Gaussian, packet of neural activity within this ring of head direction cells reflects the current head direction of the animal. By integrating a continuous angular head velocity signal it is possible to shift the position of the packet of neural activity within the head direction cell ring. The changing position of the neural activity packet reflects the changing head direction of the animal. These types of neural network models are thus capable of achieving the path integration of head direction [8–14].

An important computational question is how the head direction cell system is able to accurately perform the path integration of head direction. That is, how the packet of neural activity representing head direction can be updated to accurately reflect the true current head direction of the animal.

The neural network models of [10] and [12] can integrate real rat angular head velocity data to update the neural network activity packet representing head direction and thus perform the path integration of head direction. There is minimal error between the instantaneous network representation of head direction and the instantaneous true head direction of the rat. These neural network models, however, are “hard-wired”: the vector \mathbf{w}_i of the strengths of the synaptic connections w_{ij} between a particular set of presynaptic cells $j = 1, \dots, N$ and a particular postsynaptic cell i is pre-specified before the neural network simulation commences, and no learning takes place at any individual synaptic connection w_{ij} that is a component of this synaptic weight vector \mathbf{w}_i .

It is highly unlikely that the real head direction cell system is hard-wired. Accurate path integration of head direction requires precise control over the current position of a neural activity packet in a neural network representing the continuous head-direction space. That is, the neural activity packet should remain in its current position when the head of the animal is not rotating, and should accurately track the head direction of the animal when the animal's head is rotating. However, the behaviour of a packet of neural activity in a neural network representing a continuous space is highly sensitive to asymmetries in the driving inputs to that packet [15,16]. When the driving inputs are symmetric, i.e. of equal magnitude in all directions, then the activity packet will remain in its current position in the continuous space. Asymmetric

inputs to the packet will result in the packet shifting its position towards the input with greatest magnitude.

Thus, in order to ensure that the packet of neural activity representing head direction is stationary when the animal's head is stationary, and moves accurately to a new position when the animal's head is rotating, a set of very precise synaptic weight matrices is required. Each synaptic weight matrix \mathbf{W} specifies the synaptic connectivity and distribution of synaptic weights between a particular set of presynaptic cells $j=1, \dots, N$ and a particular set of postsynaptic cells $i=1, \dots, M$. As axonal growth and neural migration during brain development are highly stochastic, it is implausible that the required synaptic weight matrices could be entirely genetically pre-specified [14,17,18]. It is therefore far more plausible that the synaptic weight matrices are set up through learning.

[11] proposed a neural network model of the head direction cell system that can learn to perform accurate path integration of head direction. However, this network employed an error correction learning rule in order to converge upon the correct solution. This type of learning rule requires the computation of the difference ($t_i - r_i$) between the desired firing rate t_i of cell i and that cell's current firing rate r_i . This difference, or error, is then used to update the synaptic weight vector \mathbf{w}_j between presynaptic cells $j=1, \dots, N$ and the postsynaptic cell i . In the real brain it is doubtful how such an error term could be calculated and then used as a teaching signal to influence the updating of the synaptic weight vectors in order to converge upon the correct solution [19,20]. Thus, it is unlikely that the real head direction cell system employs an error correction learning scheme in order to achieve accurate path integration of head direction.

In this paper, we present a biologically plausible computational mechanism through which a neural network model can learn to accurately perform the path integration of head direction. This new mechanism incorporates the natural time delay of axonal transmission [21] in order to provide specific time intervals over which associations between packets of neural activity can be learned.

In simulations, we show that a neural network model operating with the proposed timing mechanism is able to learn to perform the path integration of head direction at approximately the same rotational speed as was experienced during training in the presence of visual input. We show that this mechanism can learn to perform the path integration of head direction when trained at different rotational speeds, and when implemented with a distribution of axonal conduction delays. We also show that the model can learn to perform path integration when implemented with axonal conduction delays in either the w^2 synapses or the w^3 synapses but not both, i.e. there are delays in one set of synapses only.

We also explore the implications of the model we present in relation to other neural network models of the path integration of head direction. We demonstrate the vital role played by the axonal conduction delays in enabling the neural network to learn to perform the path integration of head direction. We discuss the role of generalization in models of the path integration of head direction: that is, the ability of the model to perform the path integration of head direction at different speeds by altering the magnitude of a driving rotational velocity signal. The model we propose in this paper does not generalize to different speeds of path integration. This lack of generalization is not, however, a flaw in our model. Instead, we show through simulation results that our model exhibits fault tolerance and graceful degradation after the loss of a significant proportion of cells that signal rotational velocity. This fault tolerance is a direct result of the fact that our

model cannot generalize to different speeds of path integration. We also highlight the differences between the model proposed in this paper and the path integration model of [22].

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 [23], or of spatial view [24]. We also note that the principle of using neural mechanisms to produce natural time intervals over which associations between packets of neural activity can be learned is a general principle. For instance, neuronal time constants can also provide an effective natural time interval over which the path integration of head direction can be learned at approximately the correct speed [22].

The Model

Description and Operation of the Model

Experimental observations of head direction cells. Head direction cells have been discovered in several areas of the rat brain. These include the postsubiculum [1,2], the anterodorsal thalamic nucleus [25], and the lateral mammillary nuclei [26,27]. The head direction cells from these different brain areas exhibit the same general firing characteristics. Each head direction cell responds maximally to a single preferred direction of the head. The directional tuning curve of a single head direction cell is Gaussian or triangular in nature, and has a width in the approximate range of $50^\circ - 110^\circ$ [28]. In this paper we set the tuning width as 90° . The firing rate of an individual head direction cell decreases in a symmetric manner from the centre of the directional tuning curve, with zero, or near-zero, firing rates outside the width of the tuning curve [3,28,29]. Within a population of head direction cells, the preferred head directions of the individual cells will be uniformly distributed such that the population will collectively represent head directions from $0^\circ - 360^\circ$.

Many of the anatomical regions containing head direction cells are interconnected, and it has been suggested that the head direction signal is, in part, generated by an ascending processing stream incorporating the following areas: dorsal tegmental nucleus of Gudden \rightarrow lateral mammillary nuclei \rightarrow anterodorsal thalamic nucleus \rightarrow postsubiculum [30]. This is supported by evidence from lesion studies, where bilateral lesions of the anterodorsal thalamic nucleus have been shown to abolish the head direction signal in the postsubiculum [31]. It has also been shown that bilateral lesions of the lateral mammillary nuclei abolish the directional specificity of head direction cells in the anterodorsal thalamic nucleus [26]. In this paper we focus upon the computational mechanism through which our model achieves the path integration of head direction, and as such we do not attempt to place the different layers of cells in our model in a direct one-to-one relationship with the brain areas known to contain head direction cells. However, we propose that the underlying neural mechanisms instantiated in our model would be present in some form in the ascending processing stream described above.

Model details: Rate-coded neurons. In this paper, each neuron is modelled using 'rate-coding' [19,32]. This means that the model represents only the instantaneous average firing rate r_i of each cell i in the network, and does not represent the exact times of the individual action potentials emitted by the cells. Differential rate-coded neural network models are usually formulated in the following way [33].

Firstly, for each cell i we define a quantity called the cell activation h_i , which reflects the total amount of current that has recently been injected into the cell by all of the presynaptic cells

$j=1, \dots, N$. We assume that the rate of change of the cell activation h_i is proportional to a linear combination of the firing rates r_j of the presynaptic cells $j=1, \dots, N$. The rate of change of cell activation may depend on the firing rates of the presynaptic cells a short time interval Δt in the past due to axonal transmission delays that occur between different layers of cells. The firing rate r_j of each presynaptic cell j is weighted by the synaptic weight w_{ij} from presynaptic cell j to the postsynaptic cell i . Hence, the following general form of 'leaky-integrator' equation is used to model the change in activation h_i of each cell i

$$\tau \frac{dh_i(t)}{dt} = -h_i(t) + \frac{1}{N} \sum_j w_{ij}(t)r_j(t - \Delta t). \tag{1}$$

The term $-h_i(t)$ on the right hand side of equation (1) effects an exponential decay of the cell activation $h_i(t)$ in the absence of any driving input from the presynaptic cells $j=1, \dots, N$, for example when the presynaptic cells are not firing.

Next, the output firing rate r_i of each cell i can be calculated directly from the cell activation h_i , using a variety of different transfer functions $r_i(t)=f(h_i(t))$. For example, the transfer function f may be either an identity function $r_i(t)=h_i(t)$, or a linear function $r_i(t)=\alpha h_i(t)$, or a threshold linear function [19,33]. In this paper, we use a sigmoid transfer function

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

where α and β are the sigmoid threshold and slope respectively. The advantage of the sigmoid transfer function f is that it bounds the firing rate r_i for each cell i between 0 and 1.

Model details: Neural network architecture. The model presented in this paper, shown in Figure 1, comprises two connected networks: a network of head direction cells, and a network of combination cells. The model is proposed as a minimal synaptic architecture required in order to learn to update a packet of head direction cell activity at the same speed as was imposed during training in the light. Thus, the model is not intended to directly represent any one specific area of the brain known to contain head direction cells. Instead, the model is the simplest needed to demonstrate the underlying computational principles used in the path integration of head direction across several areas of the rat brain.

The network of head direction cells (with firing rate r_i^{HD} for 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 firing rate-based models as described above. Within the layer of head direction cells there is a single Gaussian packet of firing activity. The centre of the activity packet represents the current head direction of the simulated agent.

The network of combination cells (with firing rate r_i^{COMB} for 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 above. The combination cells operate as a competitive network and develop their firing properties during training, with individual combination cells learning to represent a combination of a particular rotational velocity with a particular head direction.

The combination cells in our model may be related to head direction cells in the lateral mammillary nucleus, which exhibit

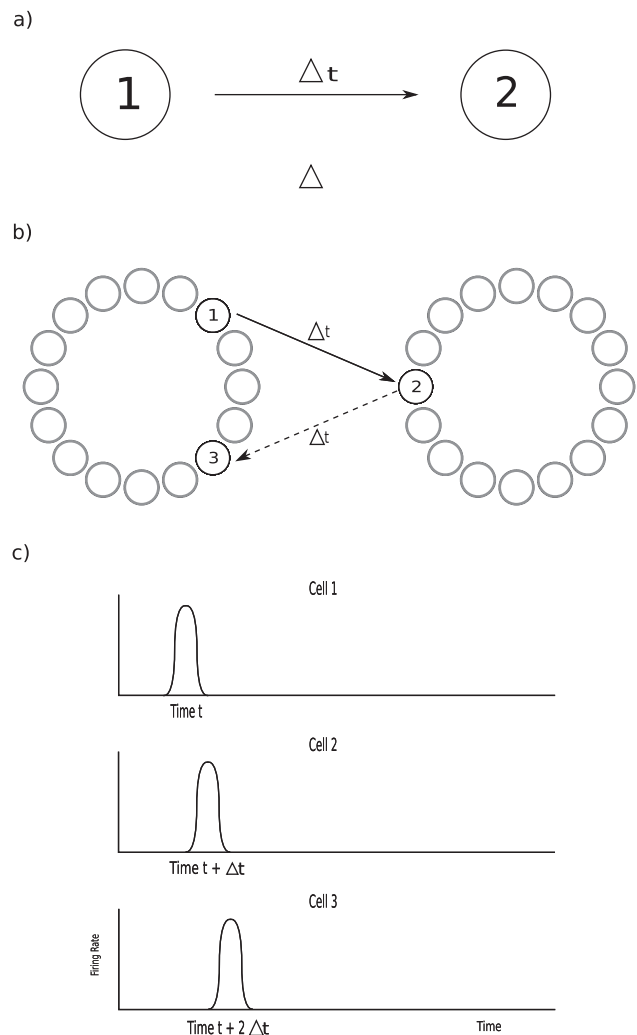


Figure 1. Network architecture for two-layer self-organizing neural network model of the head direction system. The network architecture contains a layer of head direction (HD) cells representing the current head direction of the agent; a layer of combination (COMB) cells representing a combination of 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 connection in the network, which operate as follows. The w_{ij}^1 synapses are Hebb-modifiable recurrent connections between head direction cells. These connections help to support stable packets 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. Consequently, the combination cells only become active when the agent is rotating. The head direction cells receive inputs from the combination cells through the w_{ij}^2 synapses. The w_{ij}^2 and w_{ij}^3 synapses are trained using time-delayed Hebbian associative learning rules, which incorporate a temporal delay Δt in the presynaptic firing rates. These rules introduce asymmetries into the w_{ij}^2 and w_{ij}^3 weight profiles, which play an important role in shifting the packet of head direction cell activity through the head direction cell network at the correct speed on the basis of idiothetic signals alone. doi:10.1371/journal.pone.0058330.g001

modulation of their firing rate by angular head velocity, i.e. the firing increases as the angular head velocity increases [27]. In addition, a small number of angular head velocity cells in the dorsal tegmental nucleus exhibit modulation of their firing rate by head direction, with an increase in firing rate when the head of the rat is within a broad contiguous range of head directions [34,35].

The rotational velocity cells in our model have binarised firing rates, i.e. 0 or 1, and represent idiothetic signals conveying whether the animal is rotating its head in a clockwise or a counter-clockwise direction. The idiothetic signals in the brain are highly likely to have a strong vestibular component, as lesions of the vestibular system can abolish the head direction signal [36,37]. It is also possible that the idiothetic signal contains a motor efference copy component [38], as tight restraint and passive rotation of an animal can significantly reduce the responses of head direction cells [3,39]. In principle, the idiothetic signals in our model could be either vestibular or motor efference copy.

How the path integration mechanism works in a simplified network with two neurons. To introduce the operation of the model consider a simple two-cell system, in which a single presynaptic cell j with a prescribed firing rate $r_j(t)$ is driving activity in a postsynaptic cell i . Cells 1 and 2 in Figure 2a represent this two-cell model. Our model assumes that there is a short axonal delay Δt in signal transmission from the presynaptic cell j to the postsynaptic cell i . This delay means that the chemical signals representing firing of the presynaptic cell j take Δt to reach the postsynaptic cell i . In this case, the rate of change of activation of the postsynaptic cell at time t will depend on the firing rate of the presynaptic cell Δt in the past, that is $r_j(t - \Delta t)$. The cell firing profiles through time, as depicted in Figure 2c, illustrate the effect that a delay of Δt in the transmission of a signal representing the firing of presynaptic cell 1 has upon the firing of postsynaptic cell 2. That is, if the centre of mass of the firing profile of cell 1 occurs at a time t , then the centre of mass of the firing profile of cell 2 will, due to the transmission delay of Δt , occur at time $t + \Delta t$. To model this, an axonal delay Δt is incorporated into the presynaptic firing rate term $r_j(t - \Delta t)$ in the governing equation for the activation $h_i(t)$ of the postsynaptic cell i as follows

$$\tau \frac{dh_i(t)}{dt} = -h_i(t) + w_{ij}(t)r_j(t - \Delta t). \quad (3)$$

We assume that the synaptic weight w_{ij} from the presynaptic cell j to the postsynaptic cell i is modified by a Hebbian-like associative learning rule, which depends multiplicatively on the arrival of chemical signals representing the firing of the presynaptic and postsynaptic cells.

It is also assumed that the chemical signals reflecting the firing of the postsynaptic cell i arrive instantaneously at the synapse w_{ij} . Thus, the rate of change of the synaptic weight w_{ij} at time t will be dependent on the current value of the postsynaptic cell firing rate $r_i(t)$.

In contrast, the chemical signals reflecting the firing of the presynaptic cell j will take Δt to propagate along the axon from the presynaptic cell j to the synapse w_{ij} . Therefore, we assume that the rate of change of the synaptic weight w_{ij} will depend on the firing rate of the presynaptic cell j at a time Δt in the past, that is $r_j(t - \Delta t)$.

Consequently, the synaptic weight w_{ij} is updated according to the product of the terms $r_i(t)$ and $r_j(t - \Delta t)$ as follows

$$\frac{dw_{ij}(t)}{dt} = kr_i(t)r_j(t - \Delta t) \quad (4)$$

where k is a constant learning rate, and $r(t)$ is a sigmoid function of the activation $h(t)$ given above in equation (2).

Equations (3) and (4) arise together naturally by assuming the existence of an axonal delay Δt in the transmission of chemical signals that reflect the firing of the presynaptic cell j . These equations allow the presynaptic and postsynaptic cells to learn a temporally delayed association. That is, if, during learning, the postsynaptic cell was active Δt after the presynaptic cell was active, i.e. so that the product $r_i(t)r_j(t - \Delta t)$ is large, then the presynaptic cell learns to stimulate the postsynaptic cell at a time Δt in the future. (Given that the signals still take Δt to travel from the presynaptic cell to the postsynaptic cell after learning, due to the Δt time delay in the governing equation for the postsynaptic cell activation h_i .) In this way, the correct time intervals between cell activities can 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. Conversely, if the presynaptic cell is active after the postsynaptic cell, then there will be no potentiation of the synaptic connection. Furthermore, if the weight vector of the postsynaptic cell is normalized as described for the full model below, then this will tend to reduce the connection strength. We note that the synaptic learning rule governing weight adaptation is a standard Hebbian rule normally written without the $(t - \Delta t)$ shown in the preceding equation, which merely draws attention to the fact that the presynaptic firing rate takes time Δt to travel to the postsynaptic cell.

The operation of the full network model. Next, we consider the simple two-cell model integrated into the full two-layer path integration neural network architecture shown in Figure 2(b). The left layer of neurons may be considered to correspond to a ring of head direction cells, while the right neuronal layer corresponds to a ring of combination cells representing specific combinations of head direction and angular velocity. There are feedforward synaptic connections from the head direction cells to the combination cells, and feedback connections from the combination cells to the head direction cells.

In the full model, during learning in the light, the activity of the head direction cells is driven strongly by visual input. That is, a packet of activity, corresponding to visual input representing the current head direction of the animal, is imposed on the network of head direction cells. The location of this packet of visually stimulated activity within the network of head direction cells is shifted to match the rotation of the animal. In contrast, the firing of the combination cells is driven purely by feedforward inputs from the head direction cells (as well as rotation cells considered later), with competitive lateral inhibition between the combination cells mediated by inhibitory interneurons.

Temporally-delayed associative learning is achieved over fixed time delays Δt for both the feedforward and feedback connections between the network of head direction cells and the network of combination cells.

During training in the light, the network will learn an association in the feedforward connections over a period Δt from the subset of head direction cells active at time t to the subset of combination cells that are activated at time $t + \Delta t$ by the feedforward inputs from the head direction cells. The layer of combination cells operates as a competitive learning network, with different random subsets of combination cells learning to respond to different clusters of head direction cells. The learning in the

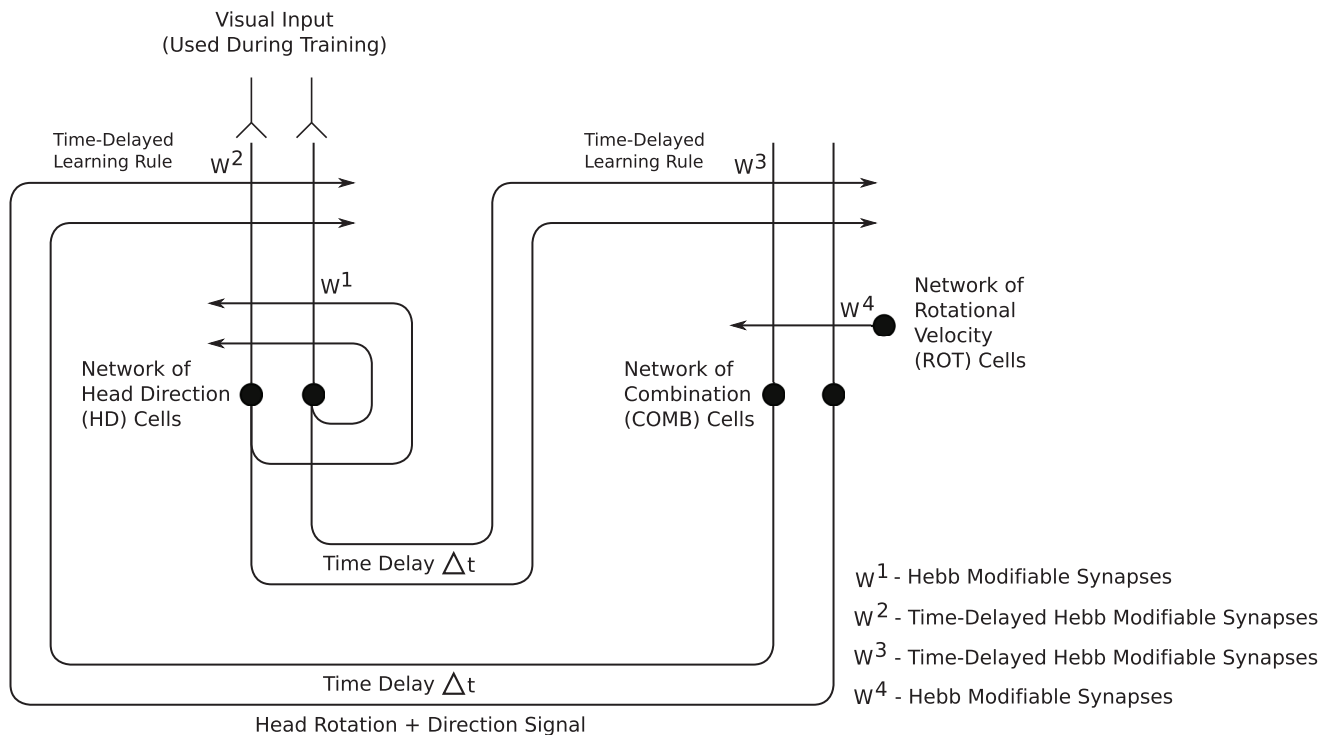


Figure 2. Architecture and operation of a simple two-cell model and its relationship to a larger many-cell model. a) In the simple two-cell model, a presynaptic cell 1 is connected to a postsynaptic cell 2. The propagation of signals from the presynaptic cell 1 to the postsynaptic cell 2 is subject to a transmission delay of duration Δt . b) The simple two-cell model integrated into the full two-layer path integration neural network with reciprocal synaptic connections between the two layers. The left layer of neurons may be considered to correspond to a ring of head direction cells, while the right layer corresponds to a ring of combination cells that represent specific combinations of head direction and angular velocity. Propagation of signals from the head direction cell layer to the combination cell layer, e.g. from cell 1 to cell 2, is subject to a transmission delay of duration Δt . Propagation of signals from the combination cell layer to the head direction cell layer, e.g. from cell 2 to cell 3, is also subject to a transmission delay of duration Δt . c) The time course of the firing profiles of cells 1, 2 and 3 during path integration after training has taken place. Consider a three-cell synaptic pathway, cell 1 \rightarrow cell 2 \rightarrow cell 3. The transmission delay of Δt in the propagation of signals from cell 1 to cell 2 ensures that if the centre of mass of the firing profile of cell 1 occurs at time t , then the centre of mass of the firing profile of cell 2 will occur at time $t + \Delta t$. Similarly, the transmission delay Δt in the propagation of signals from cell 2 to cell 3 ensures that the centre of mass of the firing profile of cell 3 will occur at time $t + 2\Delta t$. doi:10.1371/journal.pone.0058330.g002

feedforward connections to the combination cell layer is thus a form of unsupervised competitive learning [19,32].

Similarly, during training in the light, the network will learn an association in the feedback connections over a period Δt from the subset of combination cells that are activated at time $t + \Delta t$ to the subset of head direction cells that are driven by the external visual input at time $t + 2\Delta t$. Because the firing of the head direction cells is imposed by strong visual input during training, the learning in the feedback connections to the head direction cell layer is a form of supervised pattern association learning [19,32].

If the agent is moving with angular velocity $\frac{d\theta}{dt}$, then over the time $2\Delta t$ it takes for the signal from the head direction cells to be propagated to the combination cells and back to a new set of head direction cells, the agent will have rotated $2\Delta t \frac{d\theta}{dt}$. Therefore, the overall effect of learning in the feedforward and feedback connections is as follows. 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 $t_2 = t_1 + 2\Delta t$, representing a later head direction $\theta_2 = \theta_1 + 2\Delta t \frac{d\theta}{dt}$. This kind of associative learning over

fixed 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.

In Figure 2, cells 1, 2 and 3 represent a particular three-cell synaptic pathway after training has occurred: cell 1 \rightarrow cell 2 \rightarrow cell 3. This three-cell pathway illustrates the effect, depicted in Figure 2c, that signal transmission delays of Δt between cells 1 and 2, and cells 2 and 3, have in terms of enabling a precise association to be learned between a cell (cell 1) representing the head direction of the animal at time t , and a cell (cell 3) representing the head direction of the animal at time $t + 2\Delta t$. That is, after training, the full network performs path integration by propagating time-delayed signals from the head direction cell layer to the combination cell layer, and then back again to the head direction cell layer. Previous neural network models of the path integration of head direction [8–11,13] require an asymmetric driving input to the ring of head direction cells in order to shift the packet of head direction cell activity through the head direction cell network and thus update the internal representation of current head direction. The new model proposed in this paper also functions by self-organizing an asymmetric driving input to the ring of head direction cells, but the novel functionality of this new model is that, due to the transmission delays Δt , this asymmetric

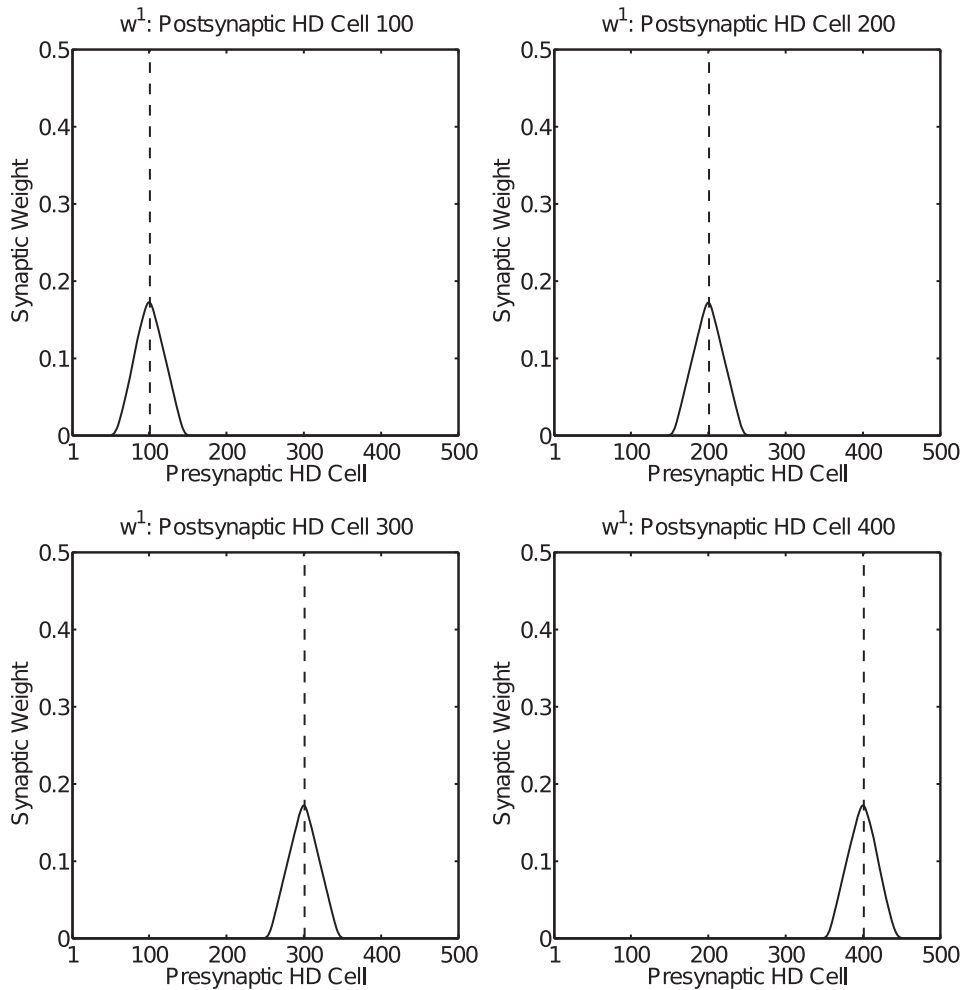


Figure 3. The recurrent synaptic weights w^1 within the network of Head Direction (HD) cells after training with the Hebbian associative learning rule (11) and weight normalization (13). These results are from a simulation with an axonal conduction delay of 100 ms, and a rotational velocity during training of $180^\circ/\text{sec}$ (all other parameters are as given in Table 1). 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 fire 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 plots, the synaptic weight profile is symmetric about the presynaptic HD cell with maximal synaptic strength, and this symmetry helps to support a stable packet of HD cell activity during testing in the absence of visual input.
doi:10.1371/journal.pone.0058330.g003

driving input to the head direction cell ring is now temporally precise. That is, the transmission delays Δt allow the learning of an association between head directions occurring at different points in time, and this asymmetry allows for the accurate replay of the correct sequence of head directions in the absence of visual input, i.e. the path integration of head direction.

Implementation of the Model for Simulations

The activation $h_i^{\text{HD}}(t)$ of head direction cell i in the model at time t is governed by

$$\tau \frac{dh_i^{\text{HD}}(t)}{dt} = -h_i^{\text{HD}}(t) + e_i(t) - \frac{1}{N_{\text{HD}}} \sum_j \tilde{w}^{\text{HD}} r_j^{\text{HD}}(t)$$

$$\begin{aligned} & + \frac{\phi_1}{C_{\text{HD} \rightarrow \text{HD}}} \sum_j w_{ij}^1(t) r_j^{\text{HD}}(t) \\ & + \frac{\phi_2}{C_{\text{COMB} \rightarrow \text{HD}}} \sum_j w_{ij}^2(t) r_j^{\text{COMB}}(t - \Delta t) \\ & - I^{\text{EXTERN}} \end{aligned} \quad (5)$$

where the activation $h_i^{\text{HD}}(t)$ is driven by the following terms.

The term $-h_i^{\text{HD}}(t)$ 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 .

Table 1. Simulation parameter values.

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}	250
\tilde{w}^{COMB}	50
I^{EXTERN}	90
σ^{HD}	20°
Learning Rates k^1, k^2, k^3, k^4	0.1
τ	1.0 ms
λ	100.0
ϑ_1	3.75×10^3
ϑ_2	2.5×10^3
ϑ_3	1.25×10^3
ϑ_4	4×10^2
HD Sigmoid Transfer Function Parameters	
α	0.0
β	0.1
COMB Sigmoid Transfer Function Parameters	
α	10.0
β	0.3
Training Parameters	
No. Training Epochs	50
Speed of Rotation (Experiments 1 & 3)	180°/s
Speed of Rotation (Experiments 2 & 4)	360°/s

HD = Head Direction. COMB = Combination. ROT = Rotational Velocity. Values are constant across all experiments except where noted. doi:10.1371/journal.pone.0058330.t001

The term $\sum_j \tilde{w}^{\text{HD}} r_j^{\text{HD}}(t)$ represents inhibitory feedback within the head direction cell network, where the summation is performed over all presynaptic head direction cells j , \tilde{w}^{HD} is a global constant describing the effect of inhibitory interneurons within the network of head direction cells, and N^{HD} is the total number of head direction cells in the model.

The term $\sum_j w_{ij}^1(t) r_j^{\text{HD}}(t)$ represents excitatory feedback within the layer of head direction cells, where the summation is over those presynaptic head direction cells that have excitatory synapses onto the postsynaptic head direction cell i . The term $r_j^{\text{HD}}(t)$ 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 . The scaling factor $\frac{\phi_1}{C^{\text{HD} \rightarrow \text{HD}}}$ controls the overall strength of the recurrent inputs to the network of head direction cells, where ϕ_1 is a constant, and $C^{\text{HD} \rightarrow \text{HD}}$ is the number of synapses onto each postsynaptic head direction cell from the presynaptic head direction cells.

In the absence of visual input, the key term driving the head direction cell activations is a sum of inputs from the presynaptic combination cells $\sum_j w_{ij}^2(t) r_j^{\text{COMB}}(t - \Delta t)$, where the summation is performed only over the subset of combination cells that have excitatory synapses onto the postsynaptic head direction cell i . The firing rate $r_j^{\text{COMB}}(t - \Delta t)$ is subject to a delay of Δt in transmission from presynaptic combination cell j to postsynaptic head direction cell i , and $w_{ij}^2(t)$ is the strength of the synapse between presynaptic combination cell j and postsynaptic head direction cell i . The scaling factor $\frac{\phi_2}{C^{\text{COMB} \rightarrow \text{HD}}}$ controls the overall strength of the combination cell inputs, where ϕ_2 is a constant, and $C^{\text{COMB} \rightarrow \text{HD}}$ is the number of synapses onto each postsynaptic head direction cell from the presynaptic combination cells.

The term I^{EXTERN} is a constant that represents external feedforward inhibition to the head direction cell network: this is necessary during the learning phase to ensure that, during 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 activity packet remains small. In the absence of external visual input, during the testing phase, the term I^{EXTERN} is set to zero.

The visual input to the postsynaptic head direction cell i is represented by the term $e_i(t)$. This visual input carries information about the current head direction of the agent, and when visual cues are available, the term $e_i(t)$ dominates other excitatory inputs to postsynaptic head direction cell i and forces this head direction cell to respond best to a particular head direction of the agent. Each head direction cell is assigned a unique preferred head direction in the range $0^\circ - 360^\circ$, and the current visual input to postsynaptic head direction cell i is set to the following Gaussian response profile

$$e_i^{\text{HD}} = \lambda e^{-(s_i^{\text{HD}})^2 / 2(\sigma^{\text{HD}})^2} \quad (6)$$

where s_i^{HD} is the difference between the actual head direction x of the agent and the preferred head direction x_i for head direction cell i , λ is a scaling factor expressing 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

$$s_i^{\text{HD}} = \text{MIN}(|x_i - x|, 360 - |x_i - x|). \quad (7)$$

The combination cells self-organize their firing responses through competitive learning [32,40]. The layer of combination cells thus operates as a competitive network. The activation $h_i^{\text{COMB}}(t)$ of postsynaptic combination cell i at time t is governed by

$$\begin{aligned} \tau \frac{dh_i^{\text{COMB}}(t)}{dt} &= -h_i^{\text{COMB}}(t) \\ &- \frac{1}{N^{\text{COMB}}} \sum_j \tilde{w}^{\text{COMB}} r_j^{\text{COMB}}(t) \end{aligned}$$

$$\begin{aligned}
& + \frac{\phi_3}{C^{\text{HD} \rightarrow \text{COMB}}} \sum_j w_{ij}^3(t) r_j^{\text{HD}}(t - \Delta t) \\
& + \frac{\phi_4}{C^{\text{ROT} \rightarrow \text{COMB}}} \sum_j w_{ij}^4(t) r_j^{\text{ROT}}(t) \quad (8)
\end{aligned}$$

with the terms defined as follows.

The term $\sum_j \tilde{w}_{ij}^{\text{COMB}} r_j^{\text{COMB}}(t)$ represents inhibitory feedback within the combination cell network, where the summation is performed over all presynaptic combination cells j , \tilde{w}^{COMB} is the global lateral inhibition constant describing the effect of inhibitory interneurons within the combination cell network, and N^{COMB} is the total number of combination cells in the model.

The term $\sum_j w_{ij}^3(t) r_j^{\text{HD}}(t - \Delta t)$ is the input from the head direction cells, where the summation is performed only over the subset of presynaptic head direction cells that have excitatory synapses onto the postsynaptic combination cell i . The firing rate $r_j^{\text{HD}}(t - \Delta t)$ is delayed by Δt in transmission from the presynaptic head direction cell j to the postsynaptic combination cell i , and $w_{ij}^3(t)$ is the strength of the synapse between presynaptic head direction cell j and postsynaptic combination cell i . The scaling factor $\frac{\phi_3}{C^{\text{HD} \rightarrow \text{COMB}}}$ controls the overall strength of the inputs from the head direction cells, where ϕ_3 is a constant, and $C^{\text{HD} \rightarrow \text{COMB}}$ is the number of synapses onto each postsynaptic combination cell from the presynaptic head direction cells.

The term $\sum_j w_{ij}^4(t) r_j^{\text{ROT}}(t)$ is the input from the rotational velocity cells, where the summation is performed over the subset of presynaptic rotational velocity cells j that have excitatory synapses onto the postsynaptic combination cell i . The firing rate of presynaptic rotational velocity cell j is given by $r_j^{\text{ROT}}(t)$, and $w_{ij}^4(t)$ is the corresponding strength of the synapse from this cell. The scaling factor $\frac{\phi_4}{C^{\text{ROT} \rightarrow \text{COMB}}}$ controls the overall strength of the inputs from the rotational velocity cells, where ϕ_4 is a constant, and $C^{\text{ROT} \rightarrow \text{COMB}}$ is the number of synapses onto each postsynaptic combination cell from the presynaptic rotational velocity cells. 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 firing, 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 .

The firing rates $r_i^{\text{HD}}(t)$ and $r_i^{\text{COMB}}(t)$ of postsynaptic head direction cell i and postsynaptic combination cell i respectively are determined from the activations $h_i^{\text{HD}}(t)$ and $h_i^{\text{COMB}}(t)$ of these cells and the sigmoid activation function given in equation (2). For a postsynaptic combination cell i , the threshold α is set to a high value to ensure that, after self-organization through competitive learning, each individual postsynaptic combination cell i will function like a logical AND gate. That is to say that temporally conjunctive inputs from the presynaptic head direction cells and the presynaptic rotational velocity cells are required in order for the postsynaptic combination cells to fire. Because the postsynaptic head direction cells are already selective for a particular head direction, there is no requirement to set the threshold α to a high value in the sigmoid activation function for the head direction cells.

The synaptic weights $w_{ij}^3(t)$ from the presynaptic head direction cells to the postsynaptic combination cells are subject to a delay Δt

in the transmission of the signal from presynaptic head direction cell j to postsynaptic combination cell i , and are updated by a local associative Hebb rule as follows

$$\frac{dw_{ij}^3(t)}{dt} = k^3 r_i^{\text{COMB}}(t) r_j^{\text{HD}}(t - \Delta t) \quad (9)$$

where k^3 is the learning rate, $r_i^{\text{COMB}}(t)$ is the instantaneous firing rate of postsynaptic combination cell i , and $r_j^{\text{HD}}(t - \Delta t)$ is the time-delayed firing rate of presynaptic head direction cell j .

Similarly, the synaptic weights $w_{ij}^2(t)$ from the presynaptic combination cells to the postsynaptic head direction cells are subject to a delay Δt in the transmission of the signal from presynaptic combination cell j to postsynaptic head direction cell i , and are updated as follows

$$\frac{dw_{ij}^2(t)}{dt} = k^2 r_i^{\text{HD}}(t) r_j^{\text{COMB}}(t - \Delta t) \quad (10)$$

where k^2 is the learning rate, $r_i^{\text{HD}}(t)$ is the instantaneous firing rate of postsynaptic head direction cell i , and $r_j^{\text{COMB}}(t - \Delta t)$ is the time-delayed firing rate of the presynaptic combination cell j .

The excitatory recurrent synaptic weights $w_{ij}^1(t)$ within the layer of head direction cells incorporate instantaneous signal transmission between presynaptic head direction cell j and postsynaptic head direction cell i , and are therefore not subject to a time delay Δt in signal transmission. The weights are updated as follows

$$\frac{dw_{ij}^1(t)}{dt} = k^1 r_i^{\text{HD}}(t) r_j^{\text{HD}}(t) \quad (11)$$

where k^1 is the learning rate, $r_i^{\text{HD}}(t)$ is the instantaneous firing rate of the postsynaptic cell i , and $r_j^{\text{HD}}(t)$ is the instantaneous firing rate of the presynaptic cell j .

The synaptic weights $w_{ij}^4(t)$ from the presynaptic rotational velocity cells to the postsynaptic combination cells incorporate instantaneous signal transmission between presynaptic rotational velocity cell j and postsynaptic combination cell i . The weights are updated as follows

$$\frac{dw_{ij}^4(t)}{dt} = k^4 r_i^{\text{COMB}}(t) r_j^{\text{ROT}}(t) \quad (12)$$

where k^4 is the learning rate, $r_i^{\text{COMB}}(t)$ is the instantaneous firing rate of postsynaptic combination cell i , and $r_j^{\text{ROT}}(t)$ is the instantaneous firing rate of presynaptic rotational velocity cell j .

All synaptic weights $w_{ij}^1(t)$, $w_{ij}^2(t)$, $w_{ij}^3(t)$, and $w_{ij}^4(t)$ are renormalized by rescaling after updating to ensure that

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

where the sum is over all presynaptic cells j . Such a renormalization process may be achieved in biological systems through synaptic weight decay [19,41]. The renormalization helps to ensure that the learning rules are convergent in the sense that the synaptic weights settle down over time to steady values, i.e. the weights do not grow unbounded.

During training in the presence of visual input, the agent rotates on the spot and the synaptic connections are established as follows.

Visual input drives the network of head direction cells according to Gaussian head direction related tuning profiles. The combination cell network is driven by inputs from the head direction cells and the rotational velocity cells. The synaptic weights w_{ij}^1 , w_{ij}^2 , w_{ij}^3 , and w_{ij}^4 are updated according to the simple and local learning rules discussed above. During training, the 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. That is, the combination cells learn to reflect the current head direction of the agent, but are most active when the agent is rotating. At the same time, the model is able to learn to perform velocity path integration of head direction by associating presynaptic head direction cell activity Δt in the past with current postsynaptic combination cell activity, and associating presynaptic combination cell activity Δt in the past with current postsynaptic head direction cell activity. Thus, considering the synaptic pathway between a head direction cell (HD1), a combination cell (COMB), and a second head direction cell (HD2), so that we have HD1 \rightarrow COMB \rightarrow HD2, the firing of HD1 at time t will produce firing in COMB at time $t + \Delta t$, which in turn will produce firing in HD2 at time $t + 2\Delta t$. The network therefore learns to associate a head direction θ_1 occurring at time t with a head direction θ_2 occurring at time $t + 2\Delta t$. In this manner, the model learns to update the packet of head direction cell activity using inputs from the combination cell network.

Training and Testing Protocol

In this paper, we do not try to explain how head direction cell firing properties develop in the presence of visual input: we only seek to explain how the head direction cell system learns to accurately perform the path integration of head direction in the dark. Therefore, during the initial learning phase in the light, we simply impose head direction cell-like firing properties on the head direction cells via hard-wired visual inputs $e_i(t)$. During training, the agent is simulated with visual input available and the visual input changes as the agent rotates through 360° , reflecting the changing head direction of the agent. The rotational velocity of the agent is simulated by means of updating the external visual input $e_i(t)$ to the network of head direction cells, such that the angular position of the external visual input rotates through 360° in constant increments. In order to keep the computational expense of the simulations low, we only train and test the model on one rotational speed at a time in both clockwise and counter-clockwise directions. We believe that, in principle, the model should be able to cope with learning multiple rotational speeds in both directions of rotation in the same set of synaptic weight matrices.

One clockwise rotation of the agent through consecutive positions $0^\circ - 360^\circ$, followed by one counter-clockwise rotation of the agent through consecutive positions $360^\circ - 0^\circ$, constitutes an epoch of training. During the clockwise rotation, half of the rotational velocity cells are set to have a firing rate of 1, with the remaining half of the rotational velocity cells set to a firing rate of 0. The rotational velocity cells with non-zero firing rates thus signal clockwise rotation. During the counter-clockwise rotation, the rotational velocity cells that signal clockwise rotation are set to have a firing rate of 0, and the remaining half of the rotational velocity cells are set to have a firing rate of 1. The rotational velocity cells with non-zero firing rates thus signal counter-clockwise rotation. In total, 50 training epochs are performed.

At the start of the training phase, the synaptic weights w_{ij}^1 , w_{ij}^2 , w_{ij}^3 , w_{ij}^4 are initialized to random positive values. As the training

phase progresses, the changing activation levels and firing rates of the head direction cells and combination cells are simulated according to equations (5), (8), and (2). The synaptic weights w_{ij}^3 and w_{ij}^2 are updated according to equations (9) and (10) respectively, and the synaptic weights w_{ij}^1 and w_{ij}^4 are updated according to equations (11) and (12). All synaptic weights undergo renormalization according to equation (13). In this way, the model learns, through self-organization, to perform accurate path integration of head direction.

Upon completion of the training phase, the simulation continues with the testing phase. The activation levels and firing rates of all cells in the model are set to zero. The agent is then orientated to an initial head direction and simulated with visual input available, but no rotational velocity cells active, for a period of 1 second. The visual input is then removed by setting all e_i terms to zero, and the agent then remains at the initial head direction for a period of a further 1 second. The purpose of this first part of the testing phase is to allow a stable packet of activity, representing the initial head direction, to develop in the head direction cell continuous attractor network.

After this, the activations and firing rates of all cells in the model are recorded. The agent remains at the initial head direction for a further period of 1 second. The firing rates of the rotational velocity cells that signal clockwise rotation are then set to 1 for a period of 1 second. The firing rates of the rotational velocity cells that signal counter-clockwise rotation remain set to 0 during this time period. The clockwise rotational velocity cells are turned on to verify that the model has learned to perform the path integration of head direction in a clockwise direction of rotation. That is, that the model can use idiothetic signals, i.e. rotational velocity signals, in the absence of visual input in order to update the internal representation of head direction.

The firing rates of all the rotational velocity cells are then set to 0 and remain in this state for a period of 1 second. This is in order to establish that the head direction cell continuous attractor network is functioning correctly, i.e. can maintain a stable and persistent packet of activity at any head direction within the continuum of head directions.

The firing rates of the rotational velocity cells that signal counter-clockwise rotation are then set to 1 for a period of 1 second. The firing rates of the rotational velocity cells that signal clockwise rotation remain set to 0 during this period. This part of the testing phase verifies that the model can perform the path integration of head direction in a counter-clockwise direction of rotation.

The firing rates of all the rotational velocity cells are then set to 0 for a final period of 1 second. This final period of the testing phase again verifies that the head direction cell continuous attractor network can maintain a stable and persistent packet of activity at any head direction, i.e. the continuous attractor network possesses a continuum of stable states. Theoretically, because the model should learn to perform the path integration of head direction at the same speed in both the clockwise and counter-clockwise directions of rotation, the position of the head direction cell continuous attractor activity packet at the end of the testing phase should be the same as the position of the head direction cell continuous attractor activity packet at the start of the testing phase.

Throughout the testing phase no learning is permitted, i.e. equations (9), (10), (11), (12), and (13) are not simulated and the synaptic weights w_{ij}^1 , w_{ij}^2 , w_{ij}^3 , and w_{ij}^4 are not updated.

For the duration of both the training and the testing phases, all differential equations that are currently being simulated are

approximated by Forward Euler finite difference schemes with a timestep of 0.0001 seconds.

Results

Demonstration of Core Model Performance

In this experiment, the model was simulated with the agent rotating at a velocity of $180^\circ/\text{s}$ during training in the presence of visual input. The axonal conduction delays Δt were set to 100ms. All other parameters for the model are given in Table 1. The model was then tested in the absence of visual input to determine if, through learning, the model could update a packet of head direction cell activity at the same speed as the rotational velocity during training. The results from the testing phase are shown in Figures 3, 4, 5, 6, 7 and Table 2.

Figure 3 shows the recurrent w^1 synaptic weights within the head direction cell network after training with the Hebbian associative learning rule (11) and weight normalization (13). Each of the plots shows the learned synaptic weights to a different postsynaptic head direction cell from the 500 presynaptic head direction cells, with the presynaptic head direction cells arranged in the plots according to where they fire maximally in the head-direction space of the agent when visual input is available. In each plot, a dashed vertical line marks the presynaptic head direction cell with which the postsynaptic head direction cell has maximal w^1 synaptic weight. For all plots, the synaptic 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 a stable packet of head direction cell activity can be maintained in the head direction cell network when the agent is stationary in the absence of visual input.

Figure 4 displays the synaptic weights w^3 from the head direction cell network to the combination cell network after competitive learning with the time-delayed Hebbian associative learning rule (9) and weight normalization (13). The individual plots show 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 fire maximally in the head-direction space of the agent in the presence of visual input. A dashed vertical line indicates the presynaptic head direction cell with which the postsynaptic combination cell has maximal w^3 synaptic strength. For the w^3 synapses, each postsynaptic combination cell received synapses from only 5% of the presynaptic head direction cells. This diluted connectivity was implemented to preserve competitive learning in the combination cell network and ensure that individual combination cells learn to respond to a combination of a particular head direction and rotational velocity. If full w^3 connectivity is implemented then Continuous Transformation (CT) learning occurs [42]. Under a CT learning paradigm, individual postsynaptic combination cells learn to respond to all possible head directions due to the continuity of the head-direction space and the overlapping nature of the head direction cell receptive fields. Previous research by the authors has shown a diluted w^3 connectivity of 5% is sufficient to avoid any CT learning effect in the w^3 synapses [13]. 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 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. Because the model parameters ϕ_3 , ϕ_4 , and the threshold α of the combination cell sigmoid transfer function are tuned to ensure that a strong rotational velocity cell input through the w^4 synapses is also necessary to fire individual combination cells, these cells can be said to learn to respond to combinations of particular head directions and clockwise or counter-clockwise rotational velocity.

In Figure 5, the plots display the synaptic weights w^2 from the combination cell network to the head direction network after learning with the time-delayed Hebbian associative learning rule (10) and weight normalization (13). In each plot, the learned synaptic weights are shown from a different presynaptic combination cell to the 500 postsynaptic head direction cells. As in Figures 3 and 4, the postsynaptic head direction cells are arranged according to where they fire maximally in the head-direction space of the agent in the presence of visual input. The dashed vertical lines indicate the postsynaptic head direction cell with which the presynaptic combination cell has maximal w^3 weight as shown in Figure 4. In 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 representing 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 current head direction of the agent will have changed in the time it takes for a signal to travel from the head direction cell network along the w^3 synapses to the combination cell network, and back to the head direction cell network through the w^2 synapses. The w^2 and w^3 axonal conduction delays therefore act as a timing mechanism that enable individual combination cells, through learning, to update the packet of activity in the head direction cell network and reflect the changing head direction of the agent. The w^2 and w^3 synapses are thus both necessary and sufficient to allow the packet of head direction cell activity to update at the same speed as the agent is rotating.

The plots in Figure 6 display the w^4 synaptic weights from the layer of rotational velocity cells to the combination cell network after learning with the Hebbian associative learning rule (12) and weight normalization (13). Each of the plots shows the learned synaptic weights to a different postsynaptic combination cell from the 500 presynaptic rotational velocity cells. Because the firing profile of the presynaptic rotational velocity cells is binary, with each cell signalling that the agent is either rotating in a given direction or is not, the w^4 synaptic weight profiles can be described as a step function of the presynaptic rotational velocity cell firing rate. As can be seen in the plots, each postsynaptic combination cell receives positive synaptic weights from the subset of exactly 250 presynaptic rotational velocity cells that signal either clockwise or counterclockwise rotation (but not both subsets). Thus, the postsynaptic combination cells have learned to be maximally stimulated by a particular head direction and a particular rotational velocity and, in conjunction with the axonal conduction delays, learn to stimulate a different postsynaptic head direction cell to the presynaptic head direction cell from which they receive maximal w^3 stimulation; but only when the presynaptic rotational velocity cells are co-firing with the head direction cells.

The firing rates of the head direction, combination, and rotational velocity cells are shown in the plots in Figure 7. The top left plot displays the firing rates of the head direction cells during training of the model. Throughout training, the activity in the head direction cells is driven by the presence of external visual

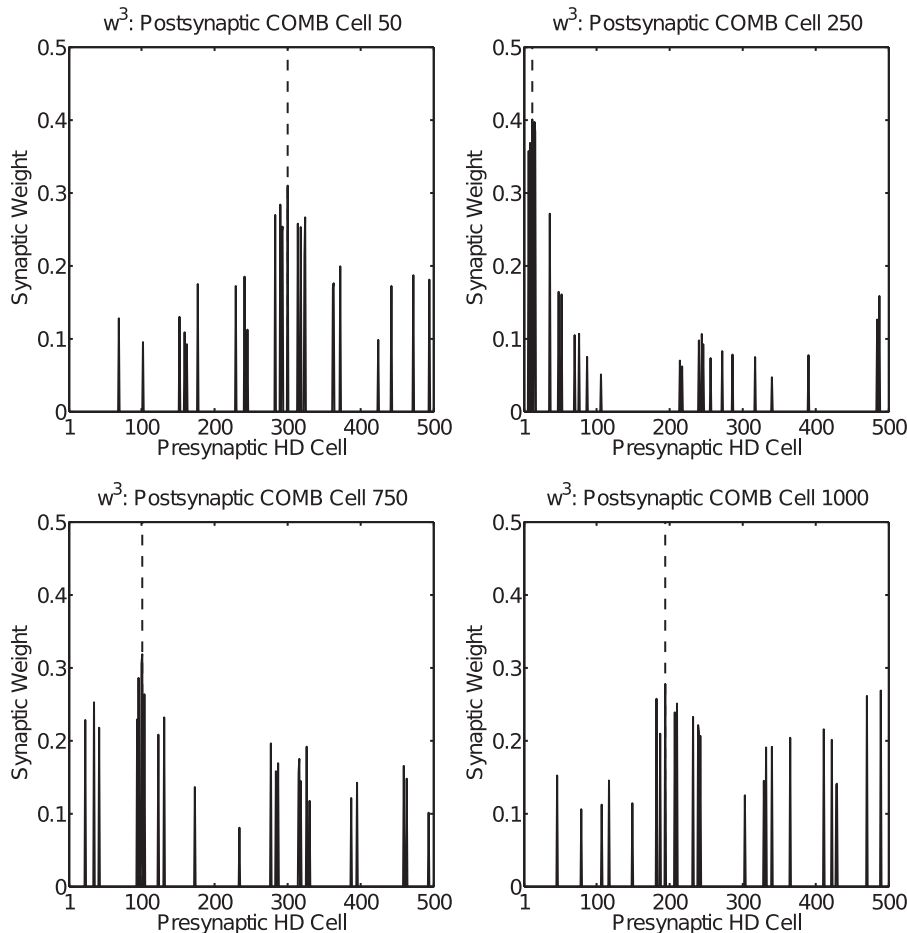


Figure 4. The synaptic weights w^3 from the Head Direction (HD) cell network to the Combination (COMB) cell network after competitive learning with the time-delayed Hebbian associative learning rule (9) and weight normalization (13). These results are from a simulation with an axonal conduction delay of 100 ms, and a rotational velocity during training of $180^\circ/\text{sec}$ (all other parameters are as given in Table 1). 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 fire 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 synaptic 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 head direction cells. Given that the model parameters ϕ_3 , ϕ_4 , and the threshold α of the COMB cell sigmoid transfer function are tuned to ensure that a strong rotational velocity cell input through the w^4 synapses is also needed in order 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.
doi:10.1371/journal.pone.0058330.g004

input. During the time interval 0.0–2.25 seconds, the agent rotated in a clockwise direction. During the time interval 2.25–4.5 seconds, the agent rotated in a counter-clockwise direction. The top right plot displays the firing rates of the head direction cells during testing in the absence of visual input. During the time interval 0.0–1.0 seconds, there was no firing in the rotational velocity cells (bottom left plot) and consequently no firing in the 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 seconds, the 250 rotational velocity cells representing clockwise rotation became active (bottom left), which in turn stimulated an activity packet in the network of combination cells through the w^4 synapses in conjunction with the head direction cell input through the w^3 synapses (bottom right). Due to the axonal conduction delays Δt and the asymmetry in the weight profiles of the w^2 synapses

(compared to the w^3 synapses), the activity packet in the combination cell network 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 seconds, the rotational velocity cells, and thus the combination cells, were quiescent in their firing and a stable packet of activity remained in the head direction cell network. During the time interval 3.0–4.0 seconds, the 250 rotational velocity cells representing counter-clockwise rotation became active (bottom left) and thus stimulated an activity packet 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 of rotation. During the time interval

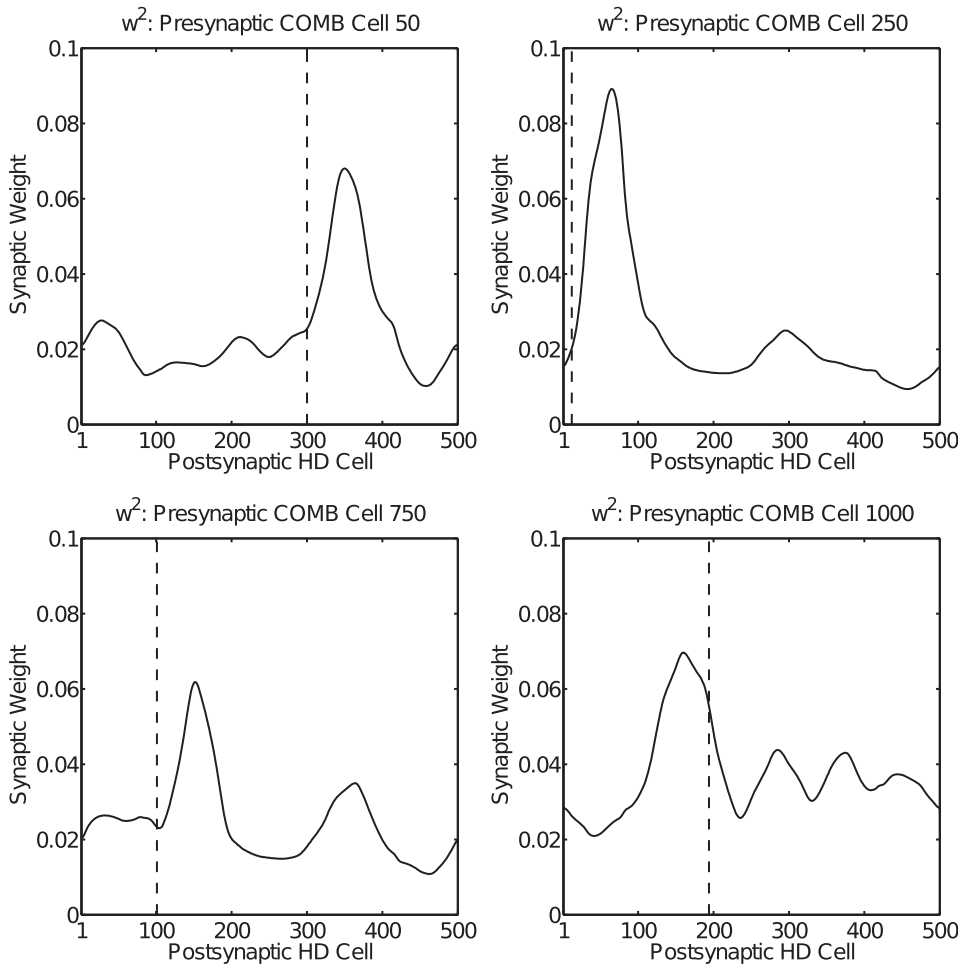


Figure 5. The synaptic weights w^2 from the Combination (COMB) cell network to the Head Direction (HD) cell network after learning with the time-delayed Hebbian associative learning rule (10) and weight normalization (13). These results are from a simulation with an axonal conduction delay of 100 ms, and a rotational velocity during training of $180^\circ/\text{sec}$ (all other parameters are as given in Table 1). 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 fire 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 synaptic weight as shown in Figure 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 it takes for a signal to travel from the HD cells through the w^3 synapses to the COMB cells, and back again to the HD cells through the w^2 synapses. Thus, the axonal conduction delays in the w^2 and w^3 synapses act, through learning, as a timing mechanism that enables the update of the packet of head-direction cell activity at the same speed as the agent is rotating.
doi:10.1371/journal.pone.0058330.g005

4.0–5.0 seconds, the rotational velocity and combination cells ceased firing 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 compared to training, the speed of update of the head direction cell activity packet was recorded. The measurement was taken according to

$$\text{speed} = \left| \frac{p_2 - p_1}{t_2 - t_1} \right| \tag{14}$$

where p_1 and p_2 represent the start and end positions (in degrees) of the packet of head direction cell activity respectively; and t_1 and t_2 represent the time (in seconds) at which the start and end packet

positions were obtained. The packet positions were calculated as follows

$$p = \frac{\sum_i r_i \theta_i}{\sum_i r_i} \tag{15}$$

where r_i is the firing 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 seconds during both the clockwise (1.25–1.75 seconds) and counter-clockwise (3.25–3.75 seconds) periods of rotation in the testing phase. (We started recording 0.25 seconds after the rotational velocity cells were turned on because it takes time for signals to travel through the w^2 synapses to update the head direction cells.)

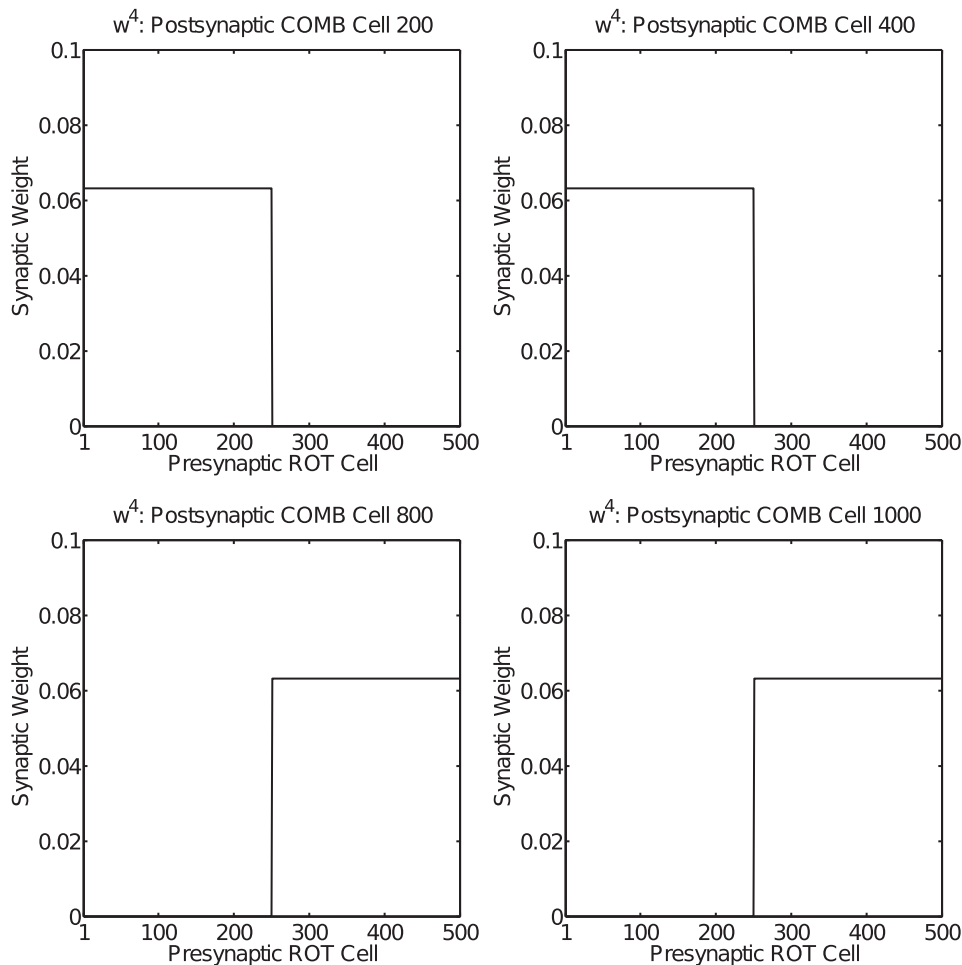


Figure 6. The synaptic weights w^4 from the layer of Rotational Velocity (ROT) cells to the Combination (COMB) cell network after learning with the Hebbian associative learning rule (12) and weight normalization (13). These results are from a simulation with an axonal conduction delay of 100 ms, and a rotational velocity during training of $180^\circ/\text{sec}$ (all other parameters are as given in Table 1). Each of the four plots shows the learned synaptic weights to a different postsynaptic COMB cell from the 500 presynaptic ROT cells. As the firing profile of the presynaptic ROT cells is binary, the w^4 synaptic weight profiles in the plots above can be described as a step function of the presynaptic ROT cell firing rate. doi:10.1371/journal.pone.0058330.g006

Five simulations were conducted with the same model parameters in each case except for different random synaptic connectivity, and different random synaptic weight initialization. Table 2 summarizes the statistics calculated on the results. The mean speed of rotation across the five simulations was $144.1^\circ/\text{s}$ (S.D. = $18.9^\circ/\text{s}$) for the period of clockwise rotation, and $132.3^\circ/\text{s}$ (S.D. = $19.0^\circ/\text{s}$) for the period of counter-clockwise rotation. This is compared to a true speed of $180^\circ/\text{s}$ during training in the presence of visual input. The standard deviations for both the clockwise rotation and the counter-clockwise rotation indicate that the measured speeds of rotation for the simulations showed a low enough dispersion around the means to conclude that the model is robust to different initial random synaptic weights and random synaptic connectivity. To further compare the recorded speed during testing to the speed enforced during training, a calculation was carried out of the mean speed during testing as a percentage of the speed during training. During the period of clockwise rotation, the model updated the packet of head direction at a mean speed that was 80.1% of the speed used during training. For the period of counter-clockwise rotation, the model updated the activity packet at a mean speed that was 73.5% of the speed during training. It can be seen that the speeds recorded during testing are similar to

those experienced by the model during training. The network has learned to achieve this automatically without careful hand-tuning of a parameter specifically governing the speed, as was used by [13]. However, while the speed at testing is a reasonable approximation to the speed during training, it is also interesting to note that the speeds during testing are regularly below those imposed during training. In future work we will investigate what architectural features contribute to this ($\sim 20\%$) underestimation of the speed during path integration.

Model Performance with Different Conduction Delays and Rotational Velocities

This experiment was conducted to investigate the effect of simulating the model with values of either 100ms or 50ms for the axonal conduction delays in the w_{ij}^2 and w_{ij}^3 synapses, and training the model at rotational velocities of either $180^\circ/\text{s}$ or $360^\circ/\text{s}$. All of the other model parameters were as given in Table 1.

Figure 8 displays the firing rates of the head direction, combination, and rotational velocity cells recorded from a model implemented with axonal conduction delays of 100ms, and trained at a rotational velocity of $360^\circ/\text{s}$. The conventions are the same as

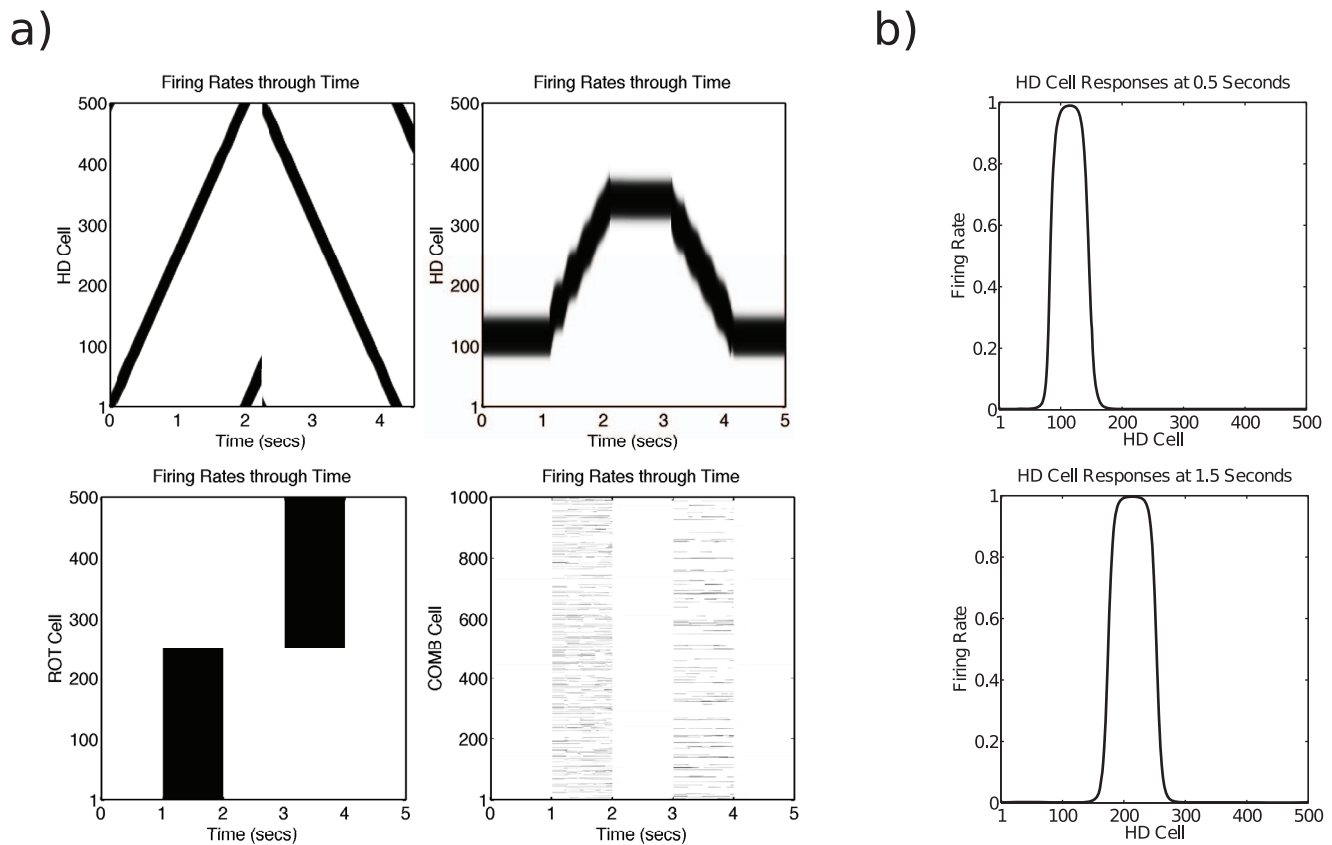


Figure 7. Firing rates of Head Direction (HD), Combination (COMB) and Rotational Velocity (ROT) cells during training and testing.

These results are from a simulation with an axonal conduction delay of 100 ms, and a rotational velocity during training of $180^\circ/\text{sec}$ (all other parameters are as given in Table 1). a) **Top Left:** Firing rates in the network of 500 HD cells during the 4.5 seconds of training, with the HD cells driven by visual input (0.0–2.25 seconds: agent rotating clockwise; 2.25–4.5 seconds: agent rotating counter-clockwise). **Top Right:** Firing rates in the network of 500 HD cells during the 5 seconds of testing in the absence of visual input. During the intervals 1.0–2.0 seconds, and 3.0–4.0 seconds, the packet of HD cell activity moves through the head-direction space of the agent driven by presynaptic COMB cell firing. During other periods, the packet of HD cell activity remains stationary and persistent due to quiescence in the presynaptic COMB cell firing. **Bottom Left:** Firing rates in the layer of 500 ROT cells during the 5 seconds of testing in the absence of visual input (1.0–2.0 seconds: ROT cells representing clockwise rotation are active; 3.0–4.0 seconds: ROT cells representing counter-clockwise rotation are active). **Bottom Right:** Firing rates in the network of 1000 COMB cells during the 5 seconds of testing in the absence of visual input. In the interval 1.0–2.0 seconds, 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 seconds, the COMB cells become active due to the firing of the 250 ROT cells representing counter-clockwise rotation of the agent. In all plots, regions of high firing are represented by darker shading. It can be seen that the firing rates are rather binarized. That is, individual cells are either quiescent, or are maximally active with a firing rate of 1.0. b) The firing rates of the HD cells recorded at two different points in time. The top plot shows the firing rates of the HD cells at 0.5 seconds, while the HD cell network is maintaining a stable and persistent packet of activity in one location. The bottom plot shows the firing rates of the HD cells at 1.5 seconds, while the packet of neural activity is being shifted through the HD cell network due to the driving influence of the active ROT cells and the presynaptic COMB cells, i.e. during path integration. In both plots, the HD cell network firing profile is approximately Gaussian, reflecting the Gaussian profile of the individual HD cell response profiles.

doi:10.1371/journal.pone.0058330.g007

for Figure 7, and the interpretation is also the same. When the same model is trained at twice the rotational velocity, the same mechanism of a time-delayed w^2 synaptic input to the postsynaptic head direction cells from the presynaptic combination cells stimulates head direction cells representing head directions further along the current direction of rotation, and thus updates the packet of head direction cell activity i.e. the model has learned to perform velocity path integration of head direction (top right plot). As the bottom left and bottom right plots display, the current model also operates in the same way as the model reported in the previous section by requiring that the head direction cell input through the w^3 synapses be temporally conjunctive with the rotational velocity cell input through the w^4 synapses, in order to stimulate an activity packet in the combination cell network. When there is no rotational velocity cell input, the firing in the

combination cell network decays to zero, and a stable packet of activity is maintained in the head direction cell network (top right plot).

Measurements of the speed of update of the packet were taken according to equations (14) and (15). Measurements were taken across a 0.5 second interval during periods of both clockwise and counter-clockwise rotation. Five simulations were conducted, with identical model parameters except for different random synaptic connectivity, and different random synaptic weight initialization. The results from these simulations are summarized in Table 2. The mean speed of rotation during the period of clockwise rotation was $290.0^\circ/\text{s}$ (S.D. = $2.5^\circ/\text{s}$), and for the period of counter-clockwise rotation was $291.6^\circ/\text{s}$ (S.D. = $4.4^\circ/\text{s}$). This is compared to a true speed during training of $360^\circ/\text{s}$. It can be concluded that the model is robust to changes in different random

Table 2. Speed of movement of the head direction cell activity packet during testing in the absence of visual input.

100 ms Delay; 180°/s Rotational Velocity		
	Clockwise	Counter-Clockwise
Mean Speed	144.1°/s	132.3°/s
Standard Deviation	18.9°/s	19.0°/s
Percentage	80.1%	73.5%
100 ms Delay; 360°/s Rotational Velocity		
	Clockwise	Counter-Clockwise
Mean Speed	290.0°/s	291.6°/s
Standard Deviation	2.5°/s	4.4°/s
Percentage	80.5%	81.0%
50 ms Delay; 180°/s Rotational Velocity		
	Clockwise	Counter-Clockwise
Mean Speed	131.7°/s	111.3°/s
Standard Deviation	7.2°/s	5.9°/s
Percentage	73.1%	61.8%
50 ms Delay; 360°/s Rotational Velocity		
	Clockwise	Counter-Clockwise
Mean Speed	231.8°/s	231.5°/s
Standard Deviation	7.9°/s	7.7°/s
Percentage	64.4%	64.3%

For all of the four experiments the results are taken from five simulation runs, each with different random synaptic connectivity and different random synaptic weight initialization. The table inlines the mean speed of the activity packet across the five simulations during testing as calculated according to equations (14) and (15), the standard deviation, and the mean speed of the packet as a percentage of the speed of rotation of the agent during training in the presence of visual input. In each case, results are inlined for both clockwise and counter-clockwise rotation of the activity packet during testing.

doi:10.1371/journal.pone.0058330.t002

synaptic weights and different random synaptic connectivity. The mean speed of the activity packet during testing as a percentage of the rotational velocity during training was also calculated. During the period of clockwise rotation, the mean speed was 80.5% of the rotational velocity during training. For the period of counter-clockwise rotation, the mean speed was 81.0% of the rotational velocity during training. Similar to the model reported in the previous section, it can be seen that the speeds recorded during testing are a reasonable approximation to those experienced by the model during training, although there is a (~20%) underestimation. Moreover, the same model with the same parameter set is able to learn two completely different speeds of rotational velocity during training, and reproduce (within a small margin of error) those speeds during testing.

The left plot of Figure 9 displays the firing rates of the head direction cells recorded during the testing phase of a model implemented with axonal conduction delays of 50ms, and trained at a rotational velocity of 180°/s. The conventions are the same as for the top right plots in Figures 7 and 8. The results demonstrate that with a different axonal conduction delay, the model is still able to maintain a stable packet of head direction cell activity when the rotational velocity cells (and thus, the combination cells) are not firing. Furthermore, when the subset of 250 rotational velocity cells representing either clockwise or counter-clockwise rotation

are firing, then an activity packet in the combination cell network is stimulated and drives the head direction cell activity packet in the correct direction. The model has thus learned to perform velocity path integration of head direction.

Five simulations with identical model parameters except for different random synaptic connectivity, and different synaptic weight initialization, were conducted and measurements of the speed of update of the head direction cell activity packet were taken according to equations (14) and (15). Table 2 summarizes the results. The mean speed during the period of clockwise rotation was 131.7°/s (S.D. = 7.2°/s), and was 111.3°/s (S.D. = 5.9°/s) during the period of counter-clockwise rotation. This is compared to a true speed during training of 180°/s. The mean speed of rotation as a percentage of the rotational velocity during training was calculated to be 73.1% for the period of clockwise rotation, and 61.8% during the period of counter-clockwise rotation. The speeds recorded during testing are approximately those imposed on the model during training. However, a comparison of these simulations and the simulations conducted at the same rotational velocity but with an axonal conduction delay of 100ms reveals no overlap in the mean values or standard errors, which suggests a significant difference between the results. A two-tailed Wilcoxon Rank-Sum test ($W_S = 21; p = 0.01$) revealed a statistically significant difference

between the two sets of results and led us to reject the null hypothesis that the results from the two experiments were drawn from identical populations. Thus, the duration of the axonal delay appears to be an important parameter affecting the accuracy of path integration in the network.

The right plot of Figure 9 shows the firing rates of the head direction cells recorded during the testing phase of a model implemented with axonal conduction delays of 50ms, and trained at a rotational velocity of $360^\circ/\text{s}$. It is clear from this figure that the model has learned to perform velocity path integration of head direction.

Measurements of the speed of update of the head direction cell activity packet were taken according to equations (14) and (15), across five simulations with identical model parameters except for different random synaptic connectivity, and different random synaptic weight initialization. The results are summarized in Table 2.

The mean speed of rotation, calculated across five simulations with identical model parameters but with different random synaptic connectivity and different random synaptic weight initialization, was $231.8^\circ/\text{s}$ (S.D. = $7.9^\circ/\text{s}$) during the clockwise rotation, and $231.5^\circ/\text{s}$ (S.D. = $7.7^\circ/\text{s}$) during the counter-clockwise rotation. This is compared to a true speed during training of $360^\circ/\text{s}$. For the period of clockwise rotation the mean speed was calculated to be 64.4% of the speed imposed during training, and for counter-clockwise rotation it was 64.3%. Thus, the model rotated at just under two-thirds ($\sim 64\%$) of the trained speed. Comparing these simulations to the simulations conducted at the same rotational velocity, but with an axonal conduction delay of 100 ms, reveals no overlap in the mean values, which suggests a significant difference between the results. A two-tailed Wilcoxon Rank-Sum test ($W_S = 23; p = 0.01$) revealed a statistically significant difference between the two sets of results and led us to reject the null hypothesis that the results from the two experiments were drawn from identical populations.

A Distribution of Axonal Conduction Delays

The model described so far learns to perform path integration by learning associations between changes in head direction over fixed axonal transmission delays Δt on individual axons. All that is required is that the same time delay Δt occurs in both the activation equations (5) and (8), and the learning rules (9) and (10). Because the association is learned across individual axons, in theory, it should be possible for the model to operate successfully with different transmission delays Δt on different axons. This hypothesis is tested next by running the model with a uniform distribution of different axonal delays, within the interval [1ms,100ms], across different axons. The model was trained with a rotational velocity of $360^\circ/\text{s}$.

Figure 10 displays the firing rates of the head direction cells recorded during testing in the dark. It is clear from the plot that a model implemented with a uniform distribution of axonal conduction delays in the interval [1ms,100ms], and trained with a rotational velocity of $360^\circ/\text{s}$, can learn to perform path integration of head direction.

The speed of update of the head direction cell activity is displayed in Table 3. For the clockwise rotation, the mean speed of the head direction cell activity packet update is $215.0^\circ/\text{s}$ (S.D. = $6.5^\circ/\text{s}$). For the counter-clockwise rotation, the mean speed of packet update is $213.1^\circ/\text{s}$ (S.D. = $9.1^\circ/\text{s}$). When the agent rotates in the clockwise direction during testing, it achieves a mean rotational speed that is 59.7% of the speed imposed during training. For the counter-clockwise rotation, the mean speed is 59.2% of the speed imposed during training.

The model can thus learn to perform path integration of head direction at approximately the speed imposed during training when the model is implemented with a uniform distribution of axonal conduction delays. This is an important result because it demonstrates that the model has biological validity: it is much more likely that there is a distribution of axonal conduction delays in the brain, rather than a single value of the conduction delay that is implemented across all synapses. The result is also important because it shows that the model can function correctly across a wide range of axonal conduction delays, from 1ms to 100ms.

Conduction Delays in One Direction

In the simulations presented so far in this paper, the models have been implemented with axonal conduction delays in both the w^2 synapses from the presynaptic combination cells to the postsynaptic head direction cells, and in the w^3 synapses from the presynaptic head direction cells to the postsynaptic combination cells. The model should still function correctly if implemented with axonal conduction delays in only one set of the w^2 or w^3 synapses. This experiment explores the results of implementing this model architecture.

Axonal Conduction Delays in the w^2 Synapses Only. In this version of the model, there are axonal conduction delays in the w^2 synapses only. Synaptic transmission through the w^3 synapses is now instantaneous with the firing of the presynaptic head direction cells. The model was simulated with an imposed rotational velocity of $360^\circ/\text{s}$, and an axonal conduction delay of 100ms. The network parameters are given in Table 1.

The left-hand plot of Figure 11 displays the firing rates of the head direction cells recorded during testing in the absence of visual input. It is evident that a model implemented with axonal conduction delays in the set of w^2 synapses only can learn to perform path integration of head direction when trained with a rotational velocity of $360^\circ/\text{s}$.

When rotating in a clockwise direction, the mean speed of update of the head direction cell activity packet is $227.5^\circ/\text{s}$ (S.D. = $5.0^\circ/\text{s}$). The results are given in Table 4. When rotating in a counter-clockwise direction, the mean speed of update of the head direction cell activity packet is $227.9^\circ/\text{s}$ (S.D. = $3.6^\circ/\text{s}$). The mean speed during clockwise rotation is 63.2% of the speed imposed during training in the light. For the counter-clockwise rotation, the mean speed is 63.3% of the speed imposed during training.

Axonal Conduction Delays in the w^3 Synapses Only. This version of the model only has axonal conduction delays in the set of w^3 synapses. The model was simulated with the network parameters given in Table 1, and trained at a rotational velocity of $360^\circ/\text{s}$. The axonal conduction delay was set to 100ms.

The right-hand plot of Figure 11 displays the firing rates of the head direction cells recorded during testing in the dark. The plot clearly shows that a model implemented with axonal conduction delays only in the w^3 synapses between the presynaptic head direction cells and the postsynaptic combination cells can still learn to perform path integration of head direction.

When rotating clockwise in the absence of visual input, the mean speed of the head direction cell activity packet is $237.2^\circ/\text{s}$ (S.D. = $4.8^\circ/\text{s}$). For the counter-clockwise rotation, the mean speed is $247.7^\circ/\text{s}$ (S.D. = $28.2^\circ/\text{s}$). The head direction cell activity packet updates in a clockwise direction at a speed that is 65.9% of the speed that is imposed during training. When rotating counter-clockwise, the head direction cell activity packet updates at a speed that is 68.8% of the speed imposed during training. These results are given in Table 4.

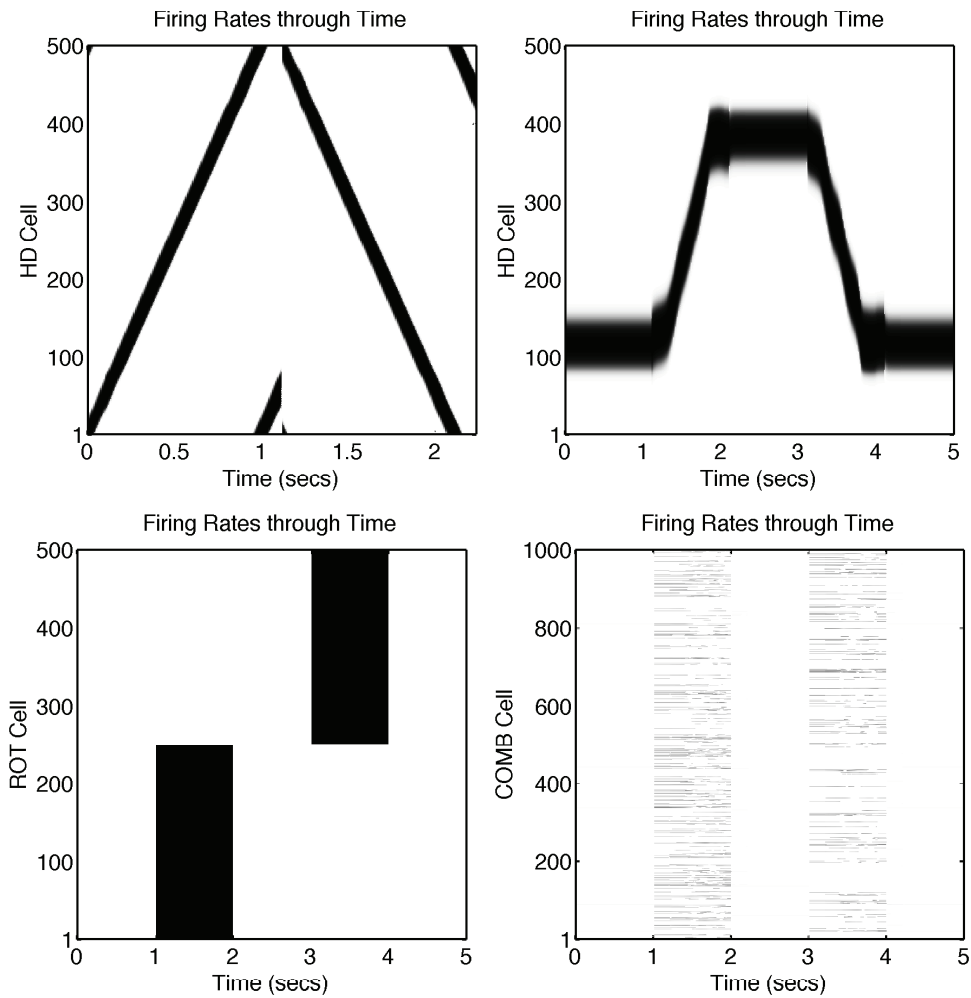


Figure 8. Firing rates of Head Direction (HD), Combination (COMB), and Rotational Velocity (ROT) cells during training and testing. These results are from a simulation with an axonal conduction delay of 100 ms, and a rotational velocity during training of $360^\circ/\text{sec}$ (all other parameters are as given in Table 1). Conventions are as for Figure 7. doi:10.1371/journal.pone.0058330.g008

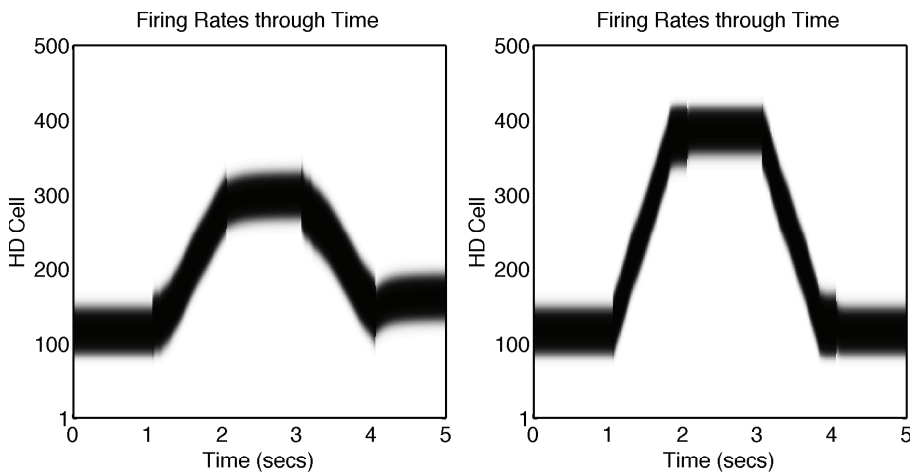


Figure 9. Firing rates of Head Direction (HD) cells during testing. These results are from simulations with an axonal conduction delay of 50 ms, and rotational velocities during training of $180^\circ/\text{sec}$ (Left plot), and $360^\circ/\text{sec}$ (Right plot). Conventions are as for the top-right plots in Figures 7 and 8. doi:10.1371/journal.pone.0058330.g009

The results of this experiment demonstrate that the model can still function correctly when implemented with axonal conduction delays in only one set of the w^2 or w^3 synapses, but not both sets. There is, however, a slight reduction in model performance in comparison to the model simulated with axonal conduction delays in both sets of synapses. The model is able to learn the correct behaviour with a single set of axonal conduction delays because those delays still provide a natural time interval over which associations between different head directions can be learned by the model. Thus, upon replay in the absence of visual input, path integration of head direction is performed.

Conduction Delays are Vital for Path Integration

This experiment was conducted to demonstrate that a non-zero value Δt for the axonal conduction delays is required if the model is to successfully learn path integration of head direction. Δt was set to 0ms, and the model was simulated with a rotational velocity of $360^\circ/\text{s}$ imposed during training.

Figure 12 displays the firing rates of the head direction cells recorded during testing in the absence of visual input. It is clear that the model fails to learn to perform path integration of head direction when the axonal conduction delays Δt are set to 0ms. This is because the activity profile in the postsynaptic combination cell i will not be delayed through time from the activity profile of the presynaptic head direction cell j . Because the axonal conduction delays are set to 0ms in the current model, signal transmission is instantaneous with the presynaptic cell firing. This means that, in the current experiment, the current head direction θ_1 will, after signal transmission to the combination cell layer and back again to the head direction cell layer, become associated with itself. That is, a head direction cell representing head direction θ_1 does not become associated with a head direction cell representing head direction θ_2 occurring at some later time in the future. Consequently, the model will not learn to perform path integration of head direction.

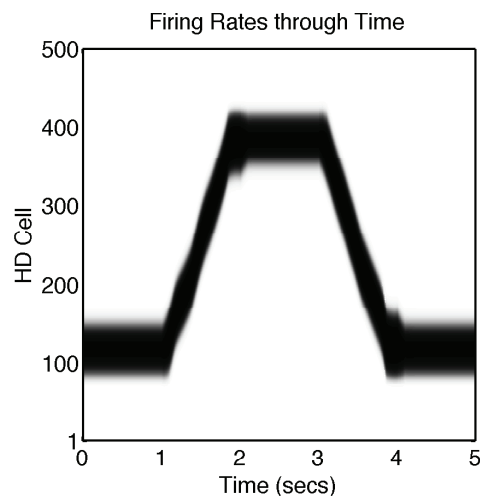


Figure 10. The firing rates of the head direction (HD) cells recorded during the testing phase. The rates are recorded from a model implemented with a uniform distribution of axonal conduction delays for w^2 and w^3 synapses within the interval from 1ms to 100ms. The model was trained with a rotational velocity of $360^\circ/\text{s}$. Conventions are as for the top-right plots in Figures 7 and 8. It is clear from the plot that the model can learn to perform path integration of head direction when simulated with a range of axonal conduction delays. doi:10.1371/journal.pone.0058330.g010

Table 3. Speed of movement of the head direction cell activity packet during testing in the absence of visual input with a distribution of axonal conduction delays.

Speed of Packet Update: A Distribution of Delays		
	Clockwise	Counter-Clockwise
Mean Speed	215.0°/s	213.1°/s
Standard Deviation	6.5°/s	9.1°/s
Percentage	59.7%	59.2%

The results are taken from five simulation runs of a model implemented with a distribution of axonal conduction delays. Each of the five models was implemented with different random synaptic connectivity and different random synaptic weight initialization. The table inlines the mean speed of the activity packet across the five simulations during testing as calculated according to equations (14) and (15), the standard deviation, and the mean speed of the packet as a percentage of the speed of rotation of the agent during training in the presence of visual input. In each case, results are inlined for both clockwise and counter-clockwise rotation of the activity packet during testing. doi:10.1371/journal.pone.0058330.t003

The results of the current experiment also highlight the difference between the current model and an earlier model of velocity path integration of head direction reported by [13]. This previous model does not incorporate any form of natural time interval over which associations between head directions at different times can be learned. In order to achieve accurate path integration, the model of [13] instead requires the manual tuning of a scaling factor controlling the strength of synaptic input to the head direction cell continuous attractor network. Because there are no axonal conduction delays in the previous model of [13], the results of this current experiment serve to show that the axonal conduction delays must be set to a non-zero value (e.g. 50ms) if path integration is to be achieved in the current model without the explicit manual tuning of a scaling factor. Thus, the axonal conduction delays are vitally important to the functioning of the current model.

Generalization and Robustness of the Model

Many previous neural network models that can perform velocity path integration of head direction exhibit the property of generalization [10–13]. That is, a linear increase in the magnitude of the signal within the model that represents the angular velocity of the head leads to a linear increase in the speed at which a packet of head direction cell activity within the model moves through the head-direction space.

The model we present in this paper operates in a different manner, which offers robustness instead of such generalization. After the model learns to perform the path integration of head direction at a particular rotational velocity, increasing or decreasing the firing rates of the rotational velocity cells will not produce a corresponding increase or decrease in the velocity of the head direction cell activity packet through the head-direction space. This is because the model learns a specific association between head direction θ_1 at a time t_1 and a later head direction θ_2 at a time $t_2 = t_1 + 2\Delta t$. In effect, for each learned head rotational velocity, the model learns a “look-up” table: given a particular head direction, the learned synaptic weight matrices dictate what the new head direction should be at a time $t_1 + 2\Delta t$ in the future. Increasing or decreasing the firing rates of the rotational velocity cells does not have any significant effect upon the change in the head direction from t_1 to t_2 .

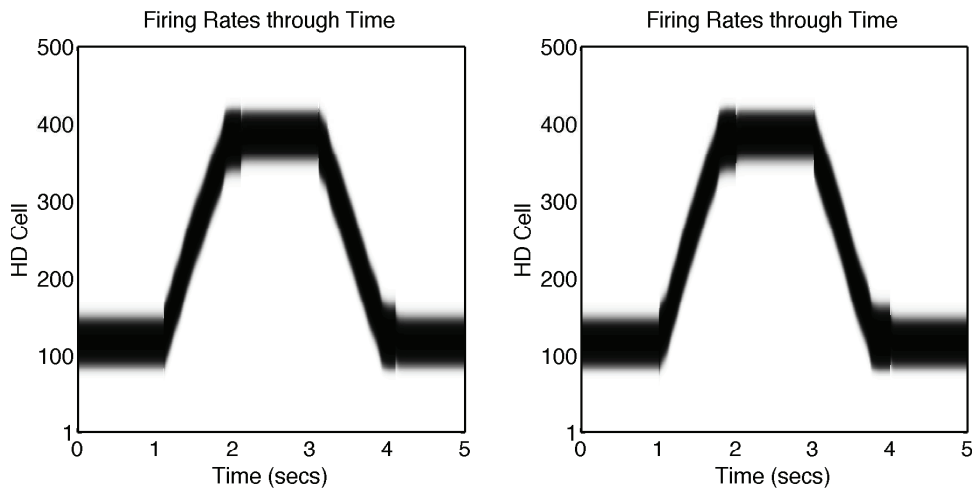


Figure 11. Firing rates of the head direction cells from models with axonal conduction delays in one set of the w^2 or w^3 synapses only. The left-hand plot displays the head direction cell firing rates recorded during the testing, in the absence of visual input, of a model simulated with axonal conduction delays in the w^2 synapses only. It is clear that the model can still learn to perform path integration of head direction. The right-hand plot displays the firing rates of head direction cells recorded during the testing, in the absence of visual input, of a model simulated with axonal conduction delays in the w^3 synapses only. It is also clear that the model can still learn to perform path integration of head direction. doi:10.1371/journal.pone.0058330.g011

Rather than generalization, our model exhibits robustness, or fault tolerance, to the loss of a number of the rotational velocity cells. This is an important property of biological neural networks. In this experiment, we took a fully-trained model (axonal conduction delay of 100ms, rotational velocity of $180^\circ/\text{s}$) and, through random sampling without replacement from a uniform probability distribution, we deleted 25%(125) of the rotational velocity cells. We then simulated the testing phase of the model in order to determine the effect that the loss of a proportion of the rotational velocity cells has upon the speed of path integration. We conducted five simulations, each with identical model parameters except that the random selection of rotational velocity cells to be deleted was different for each simulation.

Figure 13 displays the firing rates of the head direction cells recorded during the testing phase of the degraded network. It is clear that the network can still perform the path integration of head direction even after “losing” 25% of the rotational velocity cells that were present during the initial training of the model. The network thus exhibits robustness to cell loss.

Table 5 displays the speed of update of the head direction cell activity packet. During the period of clockwise rotation, the mean speed of the head direction cell activity packet is $143.1^\circ/\text{s}$ (S.D. = $0.6^\circ/\text{s}$). During the period of counter-clockwise rotation, the mean speed of the activity packet is $134.4^\circ/\text{s}$ (S.D. = $0.3^\circ/\text{s}$). The mean clockwise rotational speed is 79.5% of the speed imposed during training, and the mean counter-clockwise speed is 74.6% of the speed imposed during training. Moreover, the original simulation, before degradation of the rotational velocity cells, rotated at an average speed of $144.1^\circ/\text{s}$ in the clockwise direction and $132.3^\circ/\text{s}$ in the counter-clockwise direction. A comparison of the speeds of rotation of the model in its original and degraded states reveals that deleting 25% of the rotational velocity cells results in a change of only $\sim 2\%$ in the speed of path integration in both the clockwise and counter-clockwise directions.

Further simulations were conducted in order to determine the extent to which the model exhibits fault tolerance to the loss of a number of rotational velocity cells. Figure 14 displays the results of a series of simulations in which the percentage of rotational velocity cells deleted was increased in increments of 10%. For each

percentage of rotational velocity cells abolished, five simulations were conducted each with identical model parameters but with different random synaptic weight initializations. The average speed of rotation was then calculated over both directions of rotation across all simulations. The unbroken black line in Figure 14 represents the average rotational speed of the current model presented in this paper with different percentages of the rotational velocity cells abolished. It can be seen that the model is robust to deletion of up to 40% of the rotational velocity cells. That is, the average rotational speed, as a percentage of the speed imposed during training in the light, does not vary by a large amount until 50% of the rotational velocity cells are deleted, at which point the model can no longer perform path integration.

We predict that neural network models of the path integration of head direction that exhibit generalization to different speeds of path integration will not also exhibit robustness to loss of rotation cells. We hypothesize that a model that displays a linear relationship between the firing rates of the angular head velocity inputs and the speed of path integration will experience a decrease in the speed of path integration of $\sim 25\%$ as a result of “losing” 25% of its rotational velocity cells. Such a decrease in the speed of the head direction cell activity packet would effectively mean that this type of degraded model will no longer be able to accurately perform the path integration of head direction in the absence of visual input.

Comparison between the Current Model and the Model of Walters and Stringer (2010)

There are a number of differences between the current model and the model proposed by [22]. Firstly, the two models exhibit different behaviour as a result of deleting a percentage of the rotational velocity cells. As reported in the previous section, the model presented in this paper exhibits path integration that is robust to the loss of up to 40% of the rotational velocity cells. We conducted a similar set of simulations, in which an increasing percentage of the rotational velocity cells are deleted, using our previous model [22] for the purposes of comparison between the two models. The dashed line in Figure 14 displays the average

Table 4. The speed of movement of the head direction cell activity packet during testing with axonal conduction delays in either the w^2 or w^3 synapses but not both.

Results: w^2 Delays Only		
	Clockwise	Counter-Clockwise
Mean Speed	227.5°/s	227.9°/s
Standard Deviation	5.0°/s	3.6°/s
Percentage	63.2%	63.3%
Results: w^3 Delays Only		
	Clockwise	Counter-Clockwise
Mean Speed	237.2°/s	247.7°/s
Standard Deviation	4.8°/s	28.2°/s
Percentage	65.9%	68.8%

The results inlined are the average of five simulations conducted with identical model parameters, but with different random synaptic weight initializations and different random synaptic connectivities. When the model is implemented with axonal conduction delays either only in the w^2 synapses, or only in the w^3 synapses, then the model can still learn to perform path integration of head direction at approximately the same speed as was experienced during training. doi:10.1371/journal.pone.0058330.t004

speed of rotation achieved by the model of [22] as an increasing percentage of the rotational velocity cells are abolished. It can clearly be seen that the path integration accuracy decreases as the percentage of rotational velocity cells that have been deleted increases. In contrast to the current model reported in this paper, the previous model is able to perform path integration of head direction to a modest degree of accuracy even with 90% of the rotational velocity cells abolished.

A second key difference between the current model and the previous model of [22] lies in the form of the hebbian learning rules that update the w^2 and w^3 synapses. In the current model, the learning rules that update the w^2 and w^3 synapses, equations (10) and (9) respectively, incorporate presynaptic firing rates that occurred at a time $t - \Delta t$ in the past. These learning rules effectively ‘look backward’ through time to the presynaptic firing rate that occurred $t - \Delta t$ in the past. The interval of time over which the two learning rules look backward is set to be equal to the length of the conduction delay Δt that is present in the transmission of signals between the head direction cells and the combination cells (w^3 synapses), and is also present in the transmission of signals between the combination cells and the head direction cells (w^2 synapses).

The key point here is that a learning rule incorporating a time delay Δt in the presynaptic term will force the postsynaptic cell to learn to respond to the particular presynaptic cells that were active at $t - \Delta t$ in the past and were thus responsible for firing the postsynaptic cell at the current time t , given that there is a Δt axonal transmission delay from the presynaptic to postsynaptic cells. Thus, the combination of a Δt delay in the presynaptic term in the learning rule, coupled with a Δt axonal transmission delay, permits a postsynaptic cell to learn to respond to the particular subset of presynaptic cells that are responsible for driving it. This consistency allows asymptotic convergence of the learning process and the synaptic weights.

However, the earlier model of [22] did not incorporate an explicit time delay Δt into the learning rules, even though there was still an effective delay in the response of the postsynaptic cell to presynaptic inputs due to a long neuronal time constant. This meant that when the postsynaptic cell fired due to presynaptic

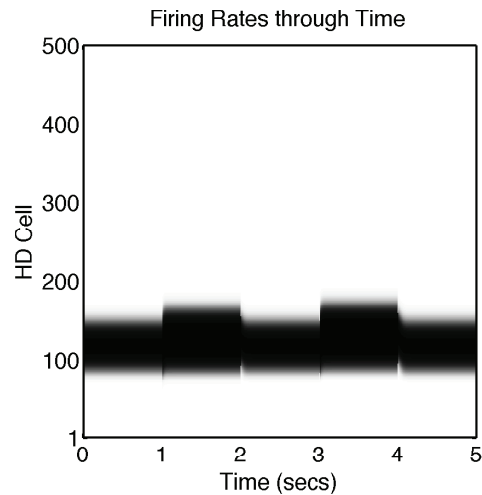


Figure 12. Path integration performance with no axonal conduction delays. Firing rates of the head direction (HD) cells recorded during the testing phase of a model simulated with an axonal conduction delay of 0ms, and trained with rotational velocity of 360°/s. Conventions are as for the top-right plots in Figures 7 and 8. It is evident that the model fails to learn to perform path integration of head direction. doi:10.1371/journal.pone.0058330.g012

activity that had occurred some small time interval in the past (dependent on the postsynaptic cell time constant), the learning rule strengthened the afferent synaptic weights from the currently active presynaptic cells rather than the cells that were actually responsible for driving the postsynaptic cell at the current time. It is easy to see how this inconsistency could retard the convergence of the learning process.

In view of the above differences between the two models, we predicted that the current model presented in this paper would converge faster during training than the model of [22]. Indeed, by comparing simulations with the current model with the earlier model of [22], we did indeed find that the convergence of the synaptic weights was significantly faster in the current model.

Convergence results for the current model presented in this paper are displayed by the unbroken lines in the left and right plots of Figure 15. The left plot displays the root mean square (RMS) change across successive blocks of five epochs for the w^3 synaptic weight matrix. As the learning converges, the RMS change in a synaptic weight matrix across successive blocks of five epochs will decrease. This is clearly the case for the w^3 synaptic weight matrix of the current model as shown in the left plot of Figure 15. The right plot of Figure 15 also demonstrates that there is convergence of the w^2 synaptic weight matrix for the current model.

The dashed line in the left plot of Figure 15 represents the change in the w^3 synaptic matrix in the model of [22] during training. The RMS decreases as the number of training epochs experienced increases, and thus there is convergence in the values of the w^3 synaptic weight matrix through time. However, the synaptic weight matrix w^3 in the model of [22] does not converge as quickly as the current model reported in this paper. Similarly, the dashed line in the right plot of Figure 15 displays the change in the w^2 synaptic weight matrix in the model of [22]. Again, it is clear that the w^2 synaptic weight matrix also shows convergence but does not converge as quickly as the w^2 synaptic weight matrix in the model reported in this paper.

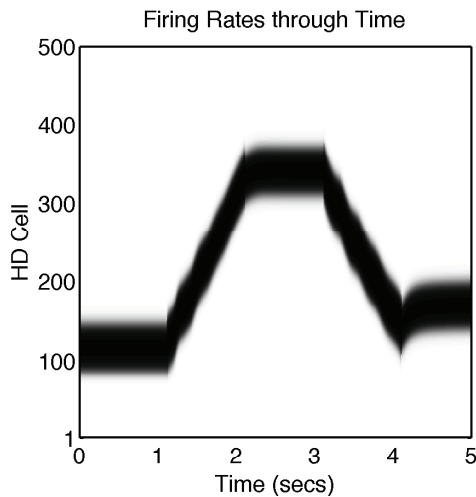


Figure 13. Path integration performance with a reduced number of rotational velocity cells. Firing rates of the head direction (HD) cells recorded during the testing phase of a model simulated with 25%(125) of the rotational velocity cells deleted after training. The original model was trained with axonal conduction delays of 100ms, at a rotational velocity of 180°/s. It is clear that the model can still perform the path integration of head direction even in a degraded state, and thus the model exhibits robustness to damage. doi:10.1371/journal.pone.0058330.g013

We thus conclude that the incorporation of the presynaptic firing at a time $t - \Delta t$ in the past in the learning rules (9) and (10) of the current model, which operates in tandem with the axonal conduction delays Δt , promotes significantly faster convergence of the w^3 and w^2 synaptic weight matrices than in the model of [22].

Discussion

The model we describe in this paper presents a computational method by which a packet of head direction cell activity can be accurately updated given an external signal representing the velocity of self-motion. The results are essentially achieved by using axonal transmission times to provide delays, and learning associations between the input velocity signal combined with the earlier head direction represented in the head direction cell network, and the new head direction.

Our model operates by learning what is effectively a ‘look-up’ table for the path integration of head direction at a particular rotational velocity. This “look-up” table functionality ensures that the model exhibits robustness to loss of a proportion of the rotational velocity cells. We know of no other neural network model of the path integration of head direction that displays such fault tolerance. An implication of our model is that the rat must learn to perform path integration at every angular head velocity it experiences: there will thus be a “look-up” table for every speed of path integration and there is no generalization to new, previously unlearned, speeds of path integration.

In this paper, we did not simulate the model being trained on more than one speed of path integration at a time, as we wished to focus upon the computational properties of the model when trained at a single rotational velocity. In principle, the model should accommodate more than one speed of path integration: this just requires learning and maintaining more than one look-up table mediated by the competitive layer of combination cells. In turn, learning multiple speeds of path integration requires that the rotational velocity cells can signal more than one rotational

Table 5. Speed of movement of the head direction cell activity packet during testing in the absence of visual input with removal of rotational velocity cells.

Speed of Packet Update with Degraded Model		
	Clockwise	Counter-Clockwise
Mean Speed	143.1°/s	134.4°/s
Standard Deviation	0.6°/s	0.3°/s
Percentage	79.5%	74.6%

The results are taken from five simulation runs of a model implemented with 25%(125) of the rotational velocity cells randomly deleted after training had taken place. For each of the five simulations, the exact rotational velocity cells that were deleted were different, i.e. the random selection process was reset with a different seed for each simulation. The table inlines the mean speed of the activity packet across the five simulations during testing as calculated according to equations (14) and (15), the standard deviation, and the mean speed of the packet as a percentage of the speed of rotation of the agent during training in the presence of visual input. In each case, results are inlined for both clockwise and counter-clockwise rotation of the activity packet during testing. doi:10.1371/journal.pone.0058330.t005

velocity. One way of achieving this would be to have many rotational velocity cells, each signalling a unique rotational velocity.

However, the firing properties of real angular head velocity cells in the brain have rather different firing properties than the rotation cells implemented in our model simulations described above. In fact, cells have been reported in the rat brain, particularly in the dorsal tegmental nucleus, with firing rates that are a monotonically increasing function of the angular head velocity of the rat [34,35]. [34] have also reported that the angular head velocity cells can be either symmetric in that they respond similarly to clockwise and counter-clockwise rotation, or asymmetric in that they respond to either clockwise or counter-clockwise rotation but not both. These angular head velocity cells have been found to have different slopes and are linear over different ranges of angular head velocity [34].

Given these more complex firing properties of angular head velocity cells in the brain, how might the network learn to perform multiple different speeds of path integration? One solution might be that the firing rate vector over the entire population of angular head velocity cells would point in different directions for different angular head velocities. A subsequent competitive network could thus learn to form separate representations for different angular head velocities, where the different angular head velocities are represented by different combination cells [19,32]. Each unique direction of the population firing rate vector would then effectively select a different look-up table within the network and thus enable accurate path integration of head direction at different speeds.

In the simulations reported in this paper, the combination cells learn to respond to combinations of a particular head direction and rotational velocity. To date, however, there is no evidence for the existence of cells in the rat brain that respond with such high specificity to a particular head direction *and* a particular rotational velocity. Instead, there are cells that do respond to head directions and angular head velocities, but over a range of velocities (with a monotonic increase in firing rate as angular head velocity increases), and often with a broader preference for head direction than the head direction cells found in the postsubiculum, anterodorsal thalamic nucleus and lateral mammillary nuclei [27,34,35]. Moreover, these angular-head-velocity by head-direction cells are not the majority cell type in the brain areas in which

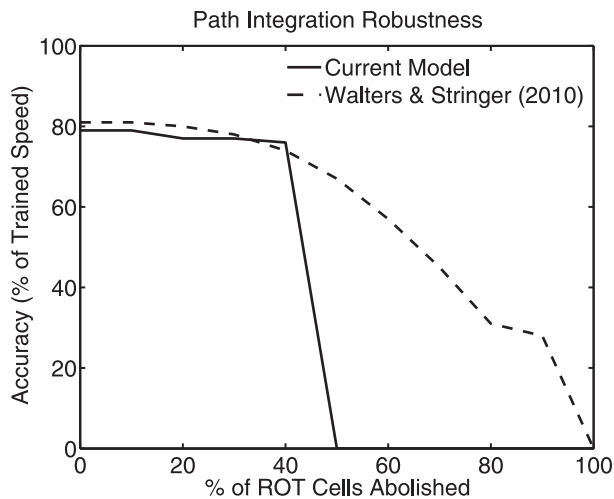


Figure 14. The robustness of the path integration of head direction to loss of a percentage of the rotational velocity cells.

Results are shown for the current model and for a previous model by the authors [22]. Both models perform very well, i.e. are robust, with up to 40% of the rotational velocity cells abolished. The model reported in this paper is able to perform path integration at a relatively constant level of accuracy as the percentage of rotational velocity cells that are abolished increases from 0% to 40%. When the percentage of abolished rotational velocity cells is greater than or equal to 50% then the model reported in this paper can no longer perform the path integration of head direction. The previous model of [22] can perform the path integration of head direction to some modest degree of accuracy with 90% of the rotational velocity cells abolished.

doi:10.1371/journal.pone.0058330.g014

they are found – particularly in the dorsal tegmental nucleus – with [34] reporting that only 5 out of 44 angular head velocity cells (11%) showed modulation of their firing rates by head direction. This is in contrast to our current model, where the combination cells outnumber any other cell type by a ratio of 2:1. The difference between the response profiles of the combination cells in our model and the known response profiles of the head direction by angular head velocity cells in the rat brain is an outstanding issue that we will investigate further in future research. We hypothesize that replacing the simplified rotation cells in the current model with more realistic angular head velocity cells may produce combination cells with more biologically realistic firing profiles.

We emphasize that, with the exception of the model by [22], we know of no previous model that can solve the problem of angular path integration of head direction using self-organizing learning with purely associative Hebbian learning rules. [11] used an error correction learning rule and a special purpose network architecture (with separate head direction subnetworks required for each direction of idiothetic signal) to produce a convergent learning scheme in a one-dimensional head direction cell system. This error correction approach is less biologically plausible than the associative learning model described in this paper.

All of our Experiments did not reproduce the exact speed during testing that they were trained upon. This could be due to a variety of factors. For example, the w^2 and w^3 synaptic connections incorporate axonal conduction delays that allow these connections to learn specific associations over temporally successive head directions. However, the role of the w^1 synapses is to stabilise the activity packet within the head direction cell network. Hence the w^1 connections do not take part in shifting the packet of head direction cell activity, and may in fact retard the movement

of the activity packet during path integration. This would lead to the actual speed of replay underestimating the correct value. A possible solution may be to simulate a model that does not contain w^1 recurrent synapses. [12] demonstrated a model that was able to perform velocity path integration of head direction without requiring recurrent synaptic weights, although they did not allow the synaptic weight profiles to self-organize; instead imposing the required asymmetries upon the network. It would be interesting to see whether it is possible to develop a model that can self-organize to perform path integration without recurrent w^1 synapses, and whether this model would be able to cope with smaller conduction delays without any reduction in the accuracy of update of the head direction cell activity packet. In this case, in the absence of recurrent connections within the head direction cell network, an activity packet could be stabilised in the absence of visual input by the feedforward and feedback connections between the head direction cells and the combination cells.

Another factor that may contribute to the undershoot in the replayed speed of path integration could be that due to computational cost our simulated models were relatively small, with only 2000 neurons in total (in order to keep the simulation run-time reasonable). The significant computational cost of the simulations was due to the need to train the layer of combination cells as a competitive network that self-organized its afferent synaptic connections over many training epochs, and that this had to be simulated using a continuous time differential model with a small numerical integration timestep of 0.1 ms. These model features made extensive exploration of the parameter space of the model infeasible for larger network sizes than what was explored in this paper. The relatively small network architecture simulated in the paper may be more susceptible to noise due to, for example, the limited number of combination cells providing an uneven representation of the continuous space of head directions. This problem would be exacerbated in a small network by the diluted connectivity of the afferent w^3 connections onto the combination cells, whereby each postsynaptic combination cell receives w^3 connections from only a random 5% of the head direction cells. We thus hypothesize that a larger model may show even less difference between the recorded speeds during testing and the actual speed during training, when compared to the results presented here. However, we are not sure why this kind of random noise would lead to a systematic undershoot of the true speed, rather than an overshoot. An alternative cause of the systematic underestimate of the true speed might be a less accurate model of the neuronal dynamics. For example, the present model is rate-coded. That is, the current model does not explicitly represent the exact timings of the individual action potentials or “spikes” emitted by cells. Instead, the model represents only an average firing rate for each neuron. More time-accurate dynamics would be introduced by implementing an integrate and fire model, in which the exact times of the spikes emitted by neurons are represented. For example, attractor neural networks built from “integrate and fire” neurons can perform very fast memory recall [19]. It might be that the implementation of more accurate integrate and fire neurons could reduce the path integration error in the model presented here. It would be straightforward to develop an integrate and fire version of the model presented here using corresponding associative learning rules that might utilise the timings of the spikes from the pre- and post-synaptic neurons [43,44]. This would need to be investigated through future simulation work. Nevertheless, the rate-coded model presented in this paper shows how a network using associative learning rules could learn to perform path integration at nearly the correct speed albeit with a mild underestimate. This work therefore provides a

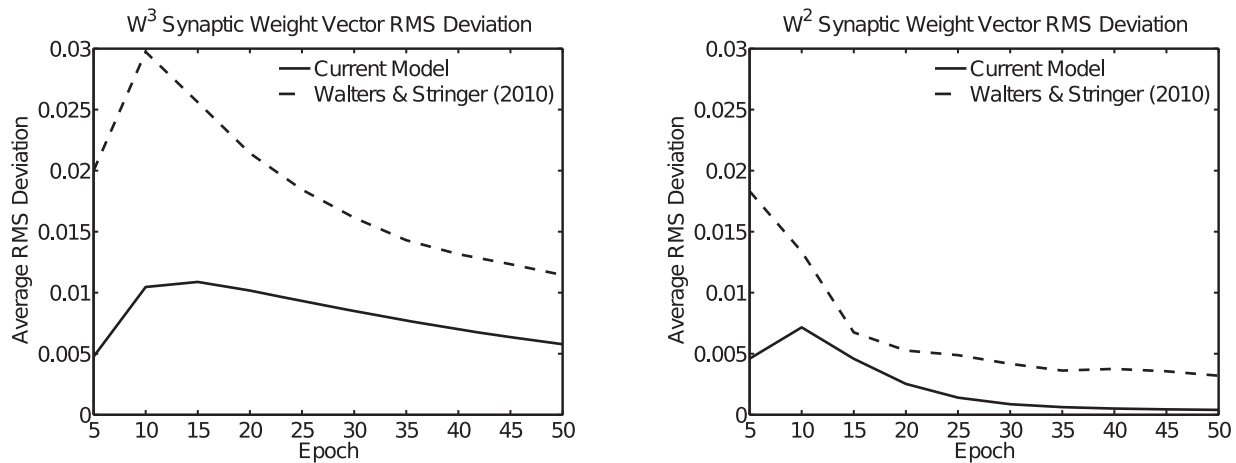


Figure 15. Convergence of the w^2 and w^3 synaptic weight vectors through time. Results are shown for the current model and for a previous model by the authors [22]. The plots display the root mean square (RMS) change in the synaptic weights across successive blocks of five training epochs for either the w^3 connections (left plot) or w^2 connections (right plot). For both models, the RMS change in the weights is monotonically decreasing after about 10–15 epochs. This implies that both the w^3 and w^2 synaptic weights are asymptotically converging to steady values. However, it is evident that convergence of the synaptic weights is much more rapid for the current model than the model of [22]. doi:10.1371/journal.pone.0058330.g015

basis for future investigation of path integration in more biologically realistic integrate and fire networks.

The models we simulated used fixed delays of either 50 ms or 100 ms for the axonal transmission delay. With typical conduction velocities of unmyelinated fibres of 0.1 m/s [21], this would correspond to axonal lengths of 5 mm and 10 mm respectively. We did simulate the network with smaller conduction delays of 10 ms, but found that the network was unable to perform path integration after training. We believe that this was due to the relatively small size of the network architecture used in the current simulations. We therefore used the larger conduction delays reported in this paper. We expect that a larger network should cope well with smaller axonal conduction delays.

The model described in this paper, using axonal transmission delays, offers a significant improvement in the accuracy of path integration over a previous model proposed by [22] that instead relies on long neuronal time constants of the order of 100 ms. This is most noticeable from the observation that the current model learns to perform path integration in the order of $\sim 75\%–80\%$ of the speed imposed during training in the light, whereas the model of [22] learns to perform path integration in the order of $\sim 50\%–60\%$ of the speed imposed during training. Although in Figure 14 in the current paper the two models exhibit the same accuracy of path integration with up to $\sim 20\%$ of the ROT cells abolished, in many previous simulations the model of [22] showed a reduced accuracy of the learned path integration compared to

the current model presented in this paper. In the current model described above we set the neuronal time constants τ to a value of 1.0ms in order to isolate the effects of axonal delays in these simulations and thus avoid confounding these two different timing mechanisms. In this paper we have highlighted the main differences between the current model and the previous model of [22].

We note that the principle of using axonal transmission delays to produce temporally precise neural processing is also argued to subserve other neural functionality. In particular, axonal transmission delays of varying lengths are argued to be vitally important to the process of sound localization in the auditory processing pathways of the owl [45–47].

Finally, we suggest that path integration implemented in the way described here could be performed in other brain systems, including hippocampal place cells [48], entorhinal cortex grid cells [49,50], and the hippocampal spatial view system of neurons that respond when a primate looks at a particular location in space, and which are updated by idiothetic eye movements made in the dark [51–54].

Author Contributions

Conceived and designed the experiments: DW SS ER. Performed the experiments: DW SS ER. Analyzed the data: DW SS ER. Wrote the paper: DW SS ER.

References

- Ranck JB Jr (1985) Head direction cells in the deep cell layer of dorsolateral presubiculum in freely moving rats. In: Buzsáki G, Vanderwolf C, editors, *Electrical Activity of the Archicortex*, Akadémiai Kiadó, Budapest.
- Taube JS, Muller RU, Ranck JB Jr (1990) Head-direction cells recorded from the postsubiculum in freely moving rats. I. Description and quantitative analysis. *Journal of Neuroscience* 10: 420–435.
- Taube JS, Muller RU, Ranck JB Jr (1990) Head-direction cells recorded from the postsubiculum in freely moving rats. II. Effects of environmental manipulations. *Journal of Neuroscience* 10: 436–447.
- Goodridge JP, Taube JS (1995) Preferential use of the landmark navigational system by head direction cells in rats. *Behavioral Neuroscience* 109: 49–61.
- Goodridge JP, Dudchenko PA, Worboys KA, Golob EJ, Taube JS (1998) Cue control and head direction cells. *Behavioral Neuroscience* 112: 749–761.
- Mittelstaedt ML, Mittelstaedt H (1980) Homing by path integration in a mammal. *Naturwissenschaften* 67: 566–567.
- Etienne AS, Jeffery KJ (2004) Path integration in mammals. *Hippocampus* 14: 180–192.
- Skaggs WE, Knierim JJ, Kudrimoti HS, McNaughton BL (1995) A model of the neural basis of the rat's sense of direction. In: Tesouro G, Touretzky D, Leen T, editors, *Advances in Neural Information Processing Systems*, MIT Press, Cambridge, MA, volume 7. pp. 173–180.
- Blair HT, Sharp PE (1995) Anticipatory head-direction cells in anterior thalamus: Evidence for a thalamocortical circuit that integrates angular head motion to compute head direction. *Journal of Neuroscience* 15: 6260–6270.
- Redish AD, Elga AN, Touretzky DS (1996) A coupled attractor model of the rodent head direction system. *Network: Computation in Neural Systems* 7: 671–685.

11. Hahnloser RHR (2003) Emergence of neural integration in the head-direction system by visual supervision. *Neuroscience* 120: 877–891.
12. Song P, Wang XJ (2005) Angular path integration by moving “hill of activity”: A spiking neuron model without recurrent excitation of the head-direction system. *Journal of Neuroscience* 25: 1002–1014.
13. Stringer SM, Rolls ET (2006) Self-organizing path integration using a linked continuous attractor and competitive network: Path integration of head direction. *Network: Computation in Neural Systems* 17: 419–445.
14. Stratton P, Wyeth G, Wiles J (2010) Calibration of the head direction network: A role for symmetric angular head velocity cells. *Journal of Computational Neuroscience* 28: 527–538.
15. Amari S (1977) Dynamics of pattern formation in lateral-inhibition type neural fields. *Biological Cybernetics* 27: 77–87.
16. Taylor JG (1999) Neural “bubble” dynamics in two dimensions: Foundations. *Biological Cybernetics* 80: 393–409.
17. Scott EK, Luo L (2001) How do dendrites take their shape? *Nature Neuroscience* 4: 359–365.
18. Dickson BJ (2002) Molecular mechanisms of axon guidance. *Science* 298: 1959–1964.
19. Rolls ET, Treves A (1998) *Neural Networks and Brain Function*. Oxford University Press, Oxford.
20. Rolls ET, Deco G (2002) *Computational Neuroscience of Vision*. Oxford University Press, Oxford.
21. Girard P, Hupé JM, Bullier J (2001) Feedforward and feedback connections between areas V1 and V2 of the monkey have similar rapid conduction velocities. *Journal of Neurophysiology* 85: 1328–1331.
22. Walters DM, Stringer SM (2010) Path integration of head direction: updating a packet of neural activity at the correct speed using neuronal time constants. *Biological Cybernetics* 103: 21–41.
23. Stringer SM, Rolls ET, Trappenberg TP, De Araujo IET (2002) Self-organizing continuous attractor networks and path integration: Two-dimensional models of place cells. *Network: Computation in Neural Systems* 13: 429–446.
24. Stringer SM, Rolls ET, Trappenberg TP (2005) Self-organizing continuous attractor network models of hippocampal spatial view cells. *Neurobiology of Learning and Memory* 83: 79–92.
25. Taube JS (1995) Head direction cells recorded in the anterior thalamic nuclei of freely moving rats. *Journal of Neuroscience* 15: 70–86.
26. Blair HT, Cho J, Sharp PE (1998) Role of the lateral mammillary nucleus in the rat head direction circuit: a combined single unit recording and lesion study. *Neuron* 21: 1387–1397.
27. Stackman RW, Taube JS (1998) Firing properties of rat lateral mammillary single units: head direction, head pitch, and angular head velocity. *Journal of Neuroscience* 18: 9020–9037.
28. Sharp PE (2005) Regional distribution and variation in the firing properties of head direction cells. In: Wiener SI, Taube JS, editors, *Head Direction Cells and the Neural Mechanisms of Spatial Orientation*, MIT Press, Cambridge, MA.
29. Taube JS (2007) The head direction signal: Origins and sensory-motor integration. *Annual Review of Neuroscience* 30: 181–207.
30. Bassett JP, Taube JS (2005) Head direction signal generation: Ascending and descending information streams. In: Wiener SI, Taube JS, editors, *Head Direction Cells and the Neural Mechanisms of Spatial Orientation*, MIT Press, Cambridge, MA. pp 83–109.
31. Goodridge JP, Taube JS (1997) Interaction between the postsubiculum and anterior thalamus in the generation of head direction cell activity. *Journal of Neuroscience* 17: 9315–9330.
32. Hertz J, Krogh A, Palmer RG (1991) *Introduction to the Theory of Neural Computation*. Addison Wesley, Wokingham, UK.
33. Levine DS (1991) *Introduction to Neural and Cognitive Modeling*. Lawrence Erlbaum Associates, New Jersey.
34. Bassett JP, Taube JS (2001) Neural correlates for angular head velocity in the rat dorsal tegmental nucleus. *Journal of Neuroscience* 21: 5740–5751.
35. Sharp PE, Tinkelman A, Cho J (2001) Angular velocity and head direction signals recorded from the dorsal tegmental nucleus of Gudden in the rat: implications for path integration in the head direction cell circuit. *Behavioral Neuroscience* 115: 571–588.
36. Stackman RW, Taube JS (1997) Firing properties of head direction cells in the rat anterior thalamic neurons: Dependence on vestibular input. *Journal of Neuroscience* 17: 4349–4358.
37. Stackman RW, Clark AS, Taube JS (2002) Hippocampal spatial representations require vestibular input. *Hippocampus* 12: 291–303.
38. Miles FA, Evars EV (1979) Concepts of motor organization. *Annual Review of Psychology* 30: 327–362.
39. Knierim JJ, Kudrimoti HS, McNaughton BL (1995) Place cells, head direction cells, and the learning of landmark stability. *Journal of Neuroscience* 15: 1648–1659.
40. Rumelhart DE, Zipser D (1985) Feature discovery by competitive learning. *Cognitive Science* 9: 75–112.
41. Oja E (1982) A simplified neuron model as a principal component analyser. *Journal of Mathematical Biology* 15: 267–273.
42. Stringer SM, Perry G, Rolls ET, Proske JH (2006) Learning invariant object recognition in the visual system with continuous transformations. *Biological Cybernetics* 94: 128–142.
43. Bi GG, Poo MM (1998) Synaptic modifications in cultured hippocampal neurons: Dependence on spike timing, synaptic strength, and postsynaptic cell type. *Journal of Neuroscience* 18: 10464–10472.
44. Froemke RC, Dan Y (2002) Spike-timing-dependent synaptic modification induced by natural spike trains. *Nature* 416: 433–438.
45. Carr CE, Konishi M (1988) Axonal delay lines for time measurement in the owl's brainstem. *Proceedings of the National Academy of Sciences of the USA* 85: 8311–8315.
46. Carr CE, Konishi M (1990) A circuit for detection of interaural time differences in the brainstem of the barn owl. *Journal of Neuroscience* 10: 3227–32246.
47. Carr CE (1993) Processing of temporal information in the brain. *Annual Review of Neuroscience* 16: 223–243.
48. O'Keefe J, Dostrovsky J (1971) The hippocampus as a spatial map: Preliminary evidence from unit activity in the freely moving rats. *Brain Research* 34: 171–175.
49. Hafting T, Fyhn M, Molden S, Moser MB, Moser EI (2005) Microstructure of a spatial map in the entorhinal cortex. *Nature* 436: 801–806.
50. Sargolini F, Fyhn M, Hafting T, McNaughton BL, Witter MP, et al. (2006) Conjunctive representation of position, direction, and velocity in entorhinal cortex. *Science* 312: 758–762.
51. Robertson RG, Rolls ET, Georges-François P (1999) Head direction cells in the primate pre-subiculum. *Hippocampus* 9: 206–219.
52. Robertson RG, Rolls ET, Georges-François P (1998) Spatial view cells in the primate hippocampus: Effects of removal of view details. *Journal of Neurophysiology* 79: 1145–1156.
53. Rolls ET (1999) Spatial view cells and the representation of place in the primate hippocampus. *Hippocampus* 9: 467–480.
54. Rolls ET, Xiang JZ (2006) Spatial view cells in the primate hippocampus, and memory recall. *Reviews in the Neurosciences* 17: 175–200.