Network: Computation in Neural Systems **7** (1996) 671–685. Printed in the UK

A coupled attractor model of the rodent head direction system

A David Redish†§, Adam N Elga‡‖ and David S Touretzky†¶

† Computer Science Department and Center for the Neural Basis of Cognition, Carnegie Mellon University, Pittsburgh, PA 15213-3891, USA

‡ 523 Laughlin Hall, Princeton University, Princeton, NJ 08544, USA

Received 28 March 1996, in final form 26 July 1996

Abstract. Head direction (HD) cells, abundant in the rat postsubiculum and anterior thalamic nuclei, fire maximally when the rat's head is facing a particular direction. The activity of a population of these cells forms a distributed representation of the animal's current heading. We describe a neural network model that creates a stable, distributed representation of head direction and updates that representation in response to angular velocity information. In contrast to earlier models, our model of the head direction system accurately tracks a series of actual rat head rotations, and, using biologically plausible neurons, it fits the single-cell tuning curves of real HD cells recorded from rats executing those same rotations. The model makes neurophysiological predictions that can be tested using current technologies.

Introduction

Head direction cells in the postsubiculum (PoS, also known as dorsal presubiculum) were first described by Ranck *et al* [16]. In subsequent work, Taube *et al* [22] characterized these cells as having triangular tuning curves: the firing rate drops off linearly from a peak at the *preferred direction* until it reaches a baseline value. Taube *et al* [23] report that PoS cells typically have baseline-to-baseline tuning curve widths of 100°. Similar cells have been found in the anterior thalamic nuclei (ATN) [4, 11, 21], most of which seem to be in the anterior dorsal (AD) nucleus. See figures 7 and 8 later for sample PoS and ATN tuning curves. These curves can also be modelled very closely by Gaussians with an average standard deviation of approximately 66° [4, 29].

One way of interpreting the activity of these cells is as a distributed representation of the rat's current head direction. A population of HD cells with preferred directions ϕ_i evenly distributed through 360° represents the direction of the weighted vector sum $\sum_i F_i \boldsymbol{v}_i$, where F_i is the normalized firing rate and $\boldsymbol{v}_i$ is a unit vector pointing in direction ϕ_i. This is the weighted circular mean [12], and is also known as a *population vector* encoding [10].

In order for the activity of such a population to represent the rat's head direction while its head is rotating, motion information must update the firing rates of head direction cells. Head direction cells in PoS and ATN are known to be sensitive to visual motion cues [4], and to the positions of landmarks such as a white card on the wall of a cylindrical arena

§ E-mail: dredish@cs.cmu.edu

‖ Present address: Department of Linguistics and Philosophy, Massachusetts Institute of Technology, Bldg 20D-213, Cambridge, MA 02139, USA.

¶ E-mail: dst@cs.cmu.edu

0954-898X/96/040671+15$19.50

[11, 21, 23]. Visual motion cues allow the system to update its current representation based on observed angular velocity, while taking bearings off familiar landmarks provides the absolute heading information needed to correct for velocity integration errors.

The rat's head direction system remains active even in total darkness [13, 21], so other sources of input must be sufficient to update the heading estimate. Blair and Sharp [4] report that PoS and ATN cells are sensitive to vestibular sensation, while Chen *et al* [5, 6] report that some cells in retrosplenial cortex (also known as posterior cingulate cortex) are correlated with both head direction and self-motion. Retrosplenial cortex is tightly coupled to both ATN [26] and postsubiculum [27]. Tactile sensations, perhaps via the vibrissae, might supply additional information about head movements.

Focusing on vestibular input, the head direction representation could be updated by integrating angular head velocity over time. Data showing that some cells in PoS code for angular head velocity and others code for head direction modulated by angular velocity [19] support such a possibility, as does the observation of cells in retrosplenial cortex that are correlated with head direction and modulated by self-motion information [5, 6].

Although the two populations in PoS and ATN seem similar, Blair and Sharp [4] and Taube and Muller [24] have recently shown a difference between them: ATN cell activity is best correlated not with current head direction, but with head direction approximately 20–40 ms in the future. PoS head direction cells, on the other hand, are best correlated with the animal's current (or recent) head direction†.

1. Previous models of the head direction system

McNaughton *et al* [13] proposed an associative mapping model for updating head direction based on angular velocity. In this model, a head direction cell population H and an angular velocity population H' jointly produce activity in a position $\times$ velocity population HH'. Each HH' cell's efferent connections encode a prediction of the animal's future heading. Various anatomical areas, such as parietal and retrosplenial cortex, are discussed as sources of the H and H' signals, and PoS is suggested as a possible site for the HH' population. The model did not address the shape of head direction cell tuning curves or the presence of head direction cells in the anterior thalamic nuclei.

Another approach to modelling the head direction system is the circular shift register. Skaggs *et al* [20] sketched such a solution, and were also the first to introduce attractor dynamics in a head direction model. In their proposal, head direction cells project to corresponding left and right rotation cells, whose activity is also controlled by vestibular cells that fire when the animal is making a left or right turn. The rotation cells project back to either the left or right neighbours of the head direction cell that drives them. Thus, during a turn, the hill of activity over the head direction cell population gradually shifts. Intrinsic connections in the head direction cell population give the system attractor dynamics, which serves to maintain the shape of the hill at all times. Another population, the 'visual cells', makes direct connections to the head direction cells and when active, move the hill of activation to specific locations. This model was an abstract proposal and did not include specific equations or simulation results.

Blair [3] proposed a somewhat different shift register model in which clockwise and anticlockwise angular velocity modulated head direction cells (AVHD cells) in the reticular

† Both Blair and Sharp [4] and Taube and Muller [24] report an optimal correlation of ATN activity with future head direction (about 37 ms in the future) and PoS activity with current head direction, but Blair and Sharp (personal communication) have recently revised their estimates, suggesting that although ATN activity still anticipates future head direction (by 24 ms), PoS activity may lag the current head direction (by 13 ms).

thalamic nucleus selectively inhibit cells in ATN that are offset from the corresponding AVHD cells. The activity of ATN cells is also governed by a modulatory input from angular speed cells hypothesized to exist in the mammilary bodies; this varies the rate of shifting because the ATN cells provide input that helps drive the AVHD cells. ATN also drives HD cells in PoS and retrosplenial cortex. Blair was the first to demonstrate a model in which the ATN representation leads the PoS representation. However, the model did not produce realistic tuning curves, and Blair did not report tracking performance on realistic data sets. He did show that his model could accomplish a single turn at a reasonable speed.

A fundamental limitation of neural shift registers is that they can integrate accurately over only a limited range of angular velocities. In order to accomodate the effects of slow rotations, the system must have a fine-scale representation of direction. However, the greater the number of distinct directions that can be represented, the greater the number of intermediate states the system must go through *in the same interval of real time* to accomodate a rapid rotation. Synaptic delays and the relatively low firing rates of HD cells bound the number of successive shift operations that can take place in a given time interval. One solution might be to use a parallel shifter circuit that can shift by any amount in one step, as in [1] or [25]. However, this requires a large number of connections which becomes prohibitive for a fine-scale representation of direction; it also seems an awkward approach to dealing with continuously variable angular velocity inputs.

Zhang [29] presents the first simulation results of a pure attractor model of the head direction system. He defines a Gaussian-shaped hill of activation and derives weights and an activation function that produce self-sustaining activity patterns of this form. He then derives another weight function to dynamically translate the activity pattern to the left or right. However, Zhang does not use a biologically plausible neural model: his units have both excitatory and inhibitory effects and have unbounded firing rates, and he does not separate the roles of ATN and PoS. Zhang also does not measure tracking ability on realistic data sets, leaving this for future work. Zhang shows how a representation of head direction in his attractor network can be made to lead *true* head direction during a turn, but gives no computational reason why this should occur, nor does he explain why ATN should lead true head direction while PoS remains locked to true (or recent) head direction.

In this paper we present an attractor model that accurately tracks head direction with biologically plausible units that display realistic tuning curves. We compare the head direction representation in our postsubicular representation with data recorded by Blair and Sharp from freely moving rats and show that the representation accurately tracks head direction. We also show that our postsubicular and anterior thalamic cells reproduce tuning curves seen in neurophysiological recordings from that same movement sequence. Finally, we show that in our model, the anterior thalamic head direction representation must lead the postsubicular representation.

2. The coupled attractor model

2.1. Neuronal model

We use a neuronal model that is more realistic than standard integrate-and-fire models, but more abstract than compartmental models. We assume for simplicity that action potentials are Boolean events with infinitesimal duration. Given an action potental at time s in a neuron j that synapses onto neuron i, we assume (again for simplicity) that the postsynaptic potential (PSP) in neuron i has an instantaneous rise and exponential fall-off with time constant τ. We model this PSP by the product of a synaptic weight w_{ij} and an α function

$\alpha_j(t-s) = \mathrm{e}^{-(t-s)/\tau_j}$. The voltage $V_i(t)$ of neuron i is then taken to be the linear sum of the effect at time t of all PSPs that have ever occured there. Let $\mathcal{F}_j(s)$ be 1 if cell j fired a spike at time s and 0 otherwise. We write $V_i(t)$ as a tonic inhibition† term γ_i plus a sum over all synapses j of the integral of all PSPs induced by that synapse:

$$V_i(t) = \gamma_i + \sum_j w_{ij} \int_0^t \alpha_j(t-s)\mathcal{F}_j(s)\ \mathrm{d}s. \tag{1}$$

In order to work with a continuous formulation, we replace the spike record $\mathcal{F}_j(s)$ with a probability of firing $F_j(s)$, defined as a sigmoidal function of the voltage $V_j(s)$. This probabilistic approximation of spiking behaviour can also be understood as a model of a neuronal population [15]; $F_j(s)$ is then the fraction of neurons in population j firing at time s. Following [15], we rewrite these equations in a form similar to the Wilson–Cowan equations [28] by defining the *synaptic drive* $S_j(t)$ supplied by neuron j:

$$S_j(t) = \int_0^t \alpha_j(t-s)F_j(s)\ \mathrm{d}s \tag{2}$$

which can be written as a differential equation:

$$\tau_j \frac{\mathrm{d}S_j(t)}{\mathrm{d}t} = -S_j(t) + F_j(t). \tag{3}$$

The synaptic drive $S_j(t)$ of neuron j thus approaches the neuron's firing rate $F_j(t)$ with time constant τ_j. Note that this time constant is the same as the time constant of the decay of the PSP α function, and is thus a function of the presynaptic neuron, not the postsynaptic one.

Synaptic drive can be interpreted as a coarse-grained time-average of the actual firing rate [15, 28] and is not necessarily a measurable neuronal property. It can be understood as the effect neuron j has on neuron i divided by the synaptic weight between them [15].

Finally, we rewrite the equation for the voltage $V_i(t)$ as a function of the impinging synaptic drives of the neurons that synapse on neuron i. This gives us a neuronal model consisting of three equations. Spatial summation is described in (4) and temporal summation in (6):

$$V_i(t) = \gamma_i + \sum_j w_{ij} S_j(t) \tag{4}$$

$$F_i(t) = \frac{1 + \tanh(V_i(t))}{2} \tag{5}$$

$$\tau_i \frac{\mathrm{d}S_i(t)}{\mathrm{d}t} = -S_i(t) + F_i(t). \tag{6}$$

As has been shown in [15], equations (4)–(6) form a consistent neuronal model that can be understood as describing either a continuous approximation to a single neuron, or a population of neurons with $V_i(t)$ being the average voltage and $F_i(t)$ the fraction of neurons in the population firing a spike at time t.

† In practice, for all cases in which our model requires γ to be non-zero, it is inhibitory, but nothing in these equations prohibits γ from being excitatory.

2.2. Maintaining a stable representation

To create a population with a triangular attractor state (actually a Gaussian with standard deviation of roughly 66°) we follow [7] and create an excitatory pool E and an inhibitory pool I, each composed of units governed by equations (4)–(6). For simplicity, we assume that the units in each pool have evenly distributed preferred directions. A unit in the excitatory pool strongly excites those units in both pools whose preferred directions are close to it. A unit in the inhibitory pool weakly inhibits practically all units in both pools, but units close in preferred direction are inhibited slightly more. See figure 1.

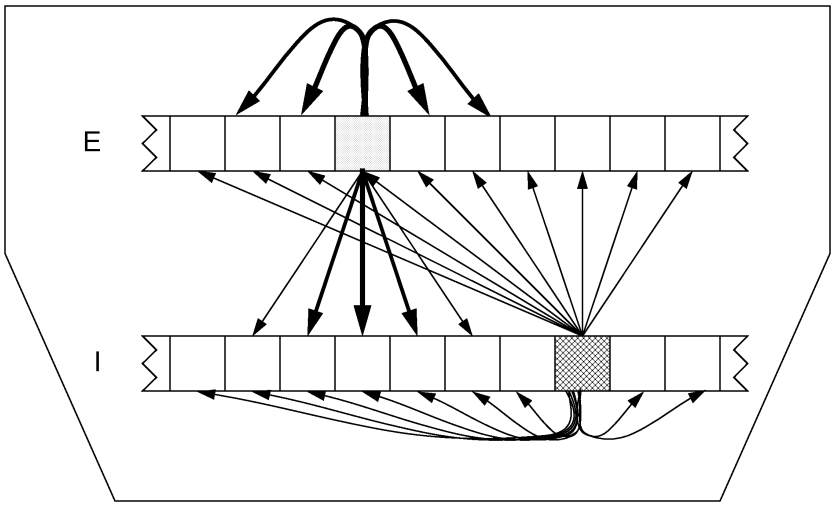

Figure 1. Connection structure of the excitatory (E) and inhibitory (I) pools in one attractor module. Shown are the strong, narrowly focused excitatory connections from an excitatory unit and the weak, diffuse inhibitory connections from an inhibitory unit.

Thus the voltages of the kth elements of pools E and I are respectively given by

$$V_k^{\mathrm{E}}(t) = \gamma_{\mathrm{E}} + \sum_j w_{\mathrm{EE}}(\phi_k - \phi_j) S_j^{\mathrm{E}}(t) + \sum_j w_{\mathrm{EI}}(\phi_k - \phi_j) S_j^{\mathrm{I}}(t) \tag{7}$$

$$V_k^{\mathrm{I}}(t) = \gamma_{\mathrm{I}} + \sum_j w_{\mathrm{IE}}(\phi_k - \phi_j) S_j^{\mathrm{E}}(t) + \sum_j w_{\mathrm{II}}(\phi_k - \phi_j) S_j^{\mathrm{I}}(t) \tag{8}$$

where ϕ_k is the preferred direction of unit k, and $w_{\mathrm{XY}}(\cdot)$ are periodized normalized Gaussian functions that determine the weights on the connections from pool Y to pool X. Inhibitory neurons have an inhibitory effect because w_{EI} and w_{II} are negative (see equation (12) and table 1 later). Since we want inhibitory units to have longer range effects than excitatory ones, we start with the Gaussians

$$g_{\mathrm{E}}(x) = \exp\left(\frac{-x^2}{\sigma_{\mathrm{E}}^2}\right) \qquad \text{and} \qquad g_{\mathrm{I}}(x) = \exp\left(\frac{-x^2}{\sigma_{\mathrm{I}}^2}\right) \tag{9}$$

where $\sigma_{\mathrm{I}} > \sigma_{\mathrm{E}}$. Following [8], we call

$$\sum_{j=-\infty}^{\infty} f(x + P_j) \tag{10}$$

the *periodized version* of $f(x)$ with period P. We define g_{E}^* and g_{I}^* to be the periodized versions of g_{E} and g_{I} with period 360° that have been normalized so that

$$\sum_{k=0}^{N_{\mathrm{E}}-1} g_{\mathrm{E}}^*\left(360\frac{k}{N_{\mathrm{E}}}\right) = 1 = \sum_{k=0}^{N_{\mathrm{I}}-1} g_{\mathrm{I}}^*\left(360\frac{k}{N_{\mathrm{I}}}\right) \tag{11}$$

where N_{X} is the number of units in pool X. w_{EE}, w_{IE}, w_{II}, and w_{EI} are defined to be g_{E}^* and g_{I}^* respectively multipled by appropriate constants:

$$\begin{aligned} w_{\mathrm{EE}} &= \kappa_{\mathrm{EE}}\, g_{\mathrm{E}}^* \\ w_{\mathrm{IE}} &= \kappa_{\mathrm{IE}}\, g_{\mathrm{E}}^* \\ w_{\mathrm{II}} &= \kappa_{\mathrm{II}}\, g_{\mathrm{I}}^* \\ w_{\mathrm{EI}} &= \kappa_{\mathrm{EI}}\, g_{\mathrm{I}}^*. \end{aligned} \tag{12}$$

With appropriately chosen parameters (see [14] for the derivation, and our table 1 for the parameter values used), the excitatory and inhibitory pools will have single regions of excitation in each pool as stable attractor states.

2.3. Updating the representation

Our model of PoS–ATN interactions consists of two coupled attractor modules and an additional inhibitory gain control population which we, like Blair [3], will identify with the mammillary bodies (see section 5). All connections between the two attractor modules are between their excitatory pools.

In addition to the intrinsic connections within each pool and between the two pools making up an attractor, which are necessary for the maintenance of the representation, there are two types of connections between the excitatory PoS:E and ATN:E pools: *matching* and *offset*. Matching connections are reciprocal connections between units with corresponding preferred directions. In the absence of any head rotation, matching connections dominate and serve to synchronize the locations of the peaks in the PoS and ATN pools.

Offset connections are responsible for changing the represented head direction. They come in two forms: left and right. Each element in the excitatory pool of the PoS structure with preferred direction ϕ has a left-offset connection to the unit in the excitatory pool of the ATN structure with preferred direction $\phi - \delta$ and a right-offset connection to the unit with preferred direction $\phi + \delta$, where δ is the amount of the offset and is the same for all units. Although adding in the true tangent vector to accomplish rotation would imply that $\delta = 90^\circ$, the attractor nature of the ATN structure will move the stable state towards any off-peak excitation. In practice, we use $\delta = 10^\circ$ because this deforms the tuning curves less and produces a smoother progression of represented head direction values.

All offset connections have strengths modulated by angular head velocity in the following way. While the head is rotating to the right, right-offset connections have strength proportional to the speed of rotation and left-offset connections have strength zero. The opposite holds true for rotations to the left. During periods when the animal is not turning, both sets of offset connections have zero strength; only the matching connections remain effective, synchronizing the PoS and ATN representations.

The angular velocity modulation of the offset connections to ATN is an empirically determined function which we denote $\xi(\dot{\phi})$. As shown later in figure 4, the relationship is nonlinear but monotonic. We implement the compensatory gain control signal from the mammillary bodies by setting the tonic inhibition on all ATN:E cells to $\gamma_{\mathrm{E}} = -\frac{1}{2}\xi(\dot{\phi})$. This non-specific modulation is applied to all units in the ATN:E population, independent of preferred direction.

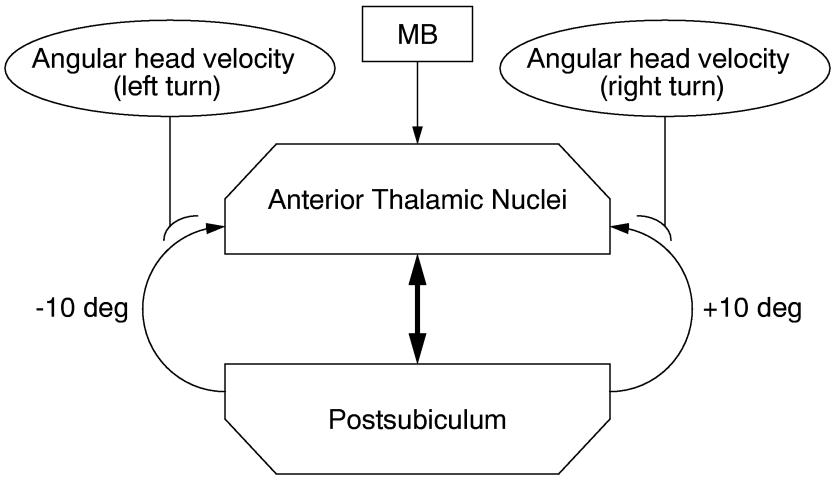

Figure 2. Connections between postsubiculum (PoS), anterior thalamic nuclei (ATN), and mammillary bodies (MB). Matching connections are drawn as a straight arrow; left- and right-offset connections as curved arrows. Activity of the *angular head velocity (left turn)* units is proportional to angular speed during left turns, and zero otherwise. They modulate the strength of the left-offset connections. The *angular head velocity (right turn)* behaves analogously.

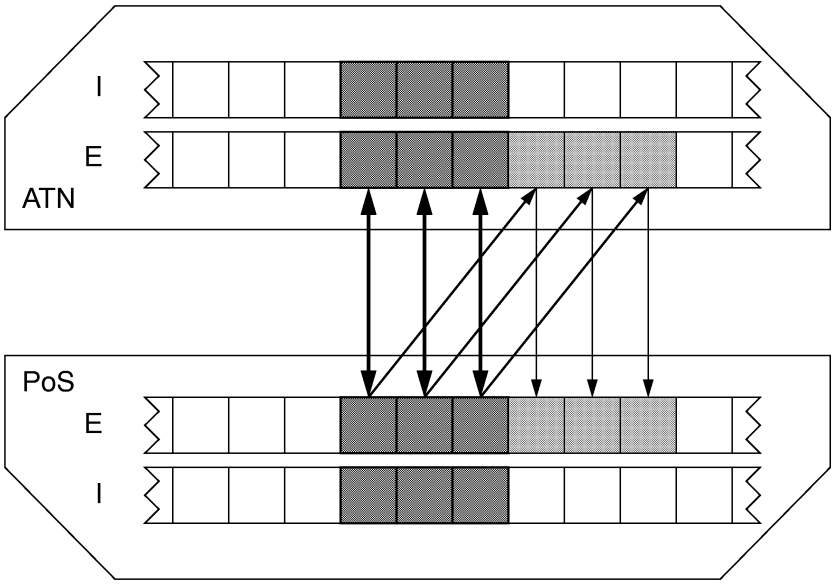

Figure 3. Active connections between PoS and ATN during a head rotation to the right. Shown are matching connections (vertical arrows) and right-offset connections (slanted arrows) from the shaded units. The left-offset connections, inactive during a right turn, are omitted. Mammillary body input (providing tonic inhibition to the ATN:E population) is also not shown.

This matching-plus-offset connections architecture entails that the locations of Gaussians in the two attractor modules will be synchronized during periods of no rotation, but during rotations ATN will lead PoS. Furthermore, the amount of lead will depend on the angular

velocity of the rotation, but due to the gain control mechanism, the shape of the hill in ATN will remain largely unchanged.

Our model is thus compatible with the data presented in the introduction: that PoS and ATN are reciprocally connected, and that cells in PoS are better correlated with current or recent heading, while ATN cells are better correlated with future heading. The model is summarized in figures 2 and 3.

Table 1. Simulation parameters.

Time step	0.1 ms
Number of units in a pool [N_E, N_I]	100
Excitatory unit time constant [τ_E]	1 ms
Excitatory unit tonic inhibition [γ_E]	−1.5
Inhibitory unit time constant [τ_I]	0.2 ms
Inhibitory unit tonic inhibition [γ_I]	−7.5
Standard deviation of g_E (σ_E)	30°
E → E connection weight (κ_{EE})	5.0
E → I connection weight (κ_{IE})	16.0
Standard deviation of g_I (σ_I)	360°
I → I connection weight (κ_{II})	−8.0
I → E connection weight (κ_{EI})	−12.0
PoS → ATN matching connection strength	1.0
ATN → PoS matching connection strength	0.6
Offset [δ, section 2.3]	10°
PoS → ATN offset connection strength [$\xi(\dot{\phi})$]	See figure 4
MB → ATN connection strength	$-\frac{1}{2}\xi(\dot{\phi})$

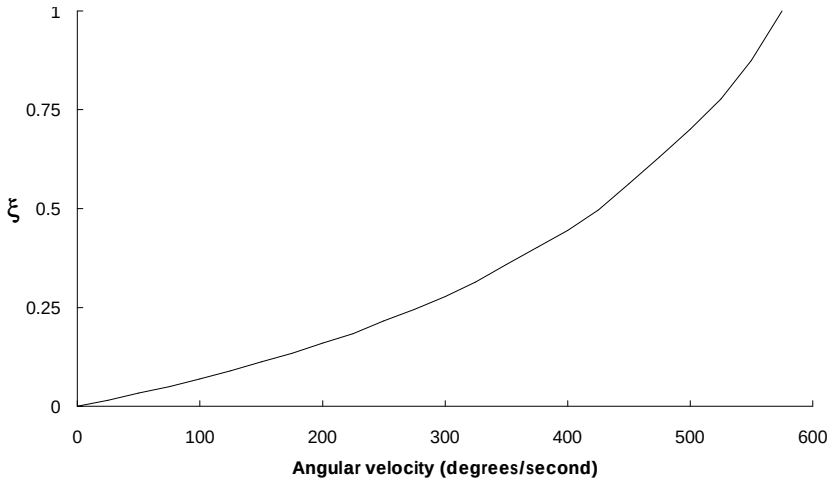

Figure 4. Connection strength of the right-offset connection between PoS:E and AD as a function of angular velocity. This function $\xi(\dot{\phi})$ was determined experimentally.

3. Results

We show that our model tracks head direction accurately, using cells with tuning curves similar to real cells in postsubiculum and the anterior thalamic nuclei. For our simulations, we used units with evenly spaced preferred directions and parameters shown in table 1.

We begin by comparing the model to data recorded by Blair and Sharp from freely moving rats (for details on recording methods, see [4]). These data included the rat's head direction, sampled at 60 Hz. Missing data points were linearly interpolated. To counteract quantization error, we calculated the head direction at every time as the average direction over a 133 ms window centred at that time. This filtering smoothed out fine fluctuations without removing important detail from the angular velocity trace. We then estimated the rat's angular head velocity over the 16 ms time period between two samples ϕ_t and ϕ_{t+1} as the change in head direction divided by 16 ms. These head velocity estimates served as the vestibular input $\dot{\phi}$ for our simulations.

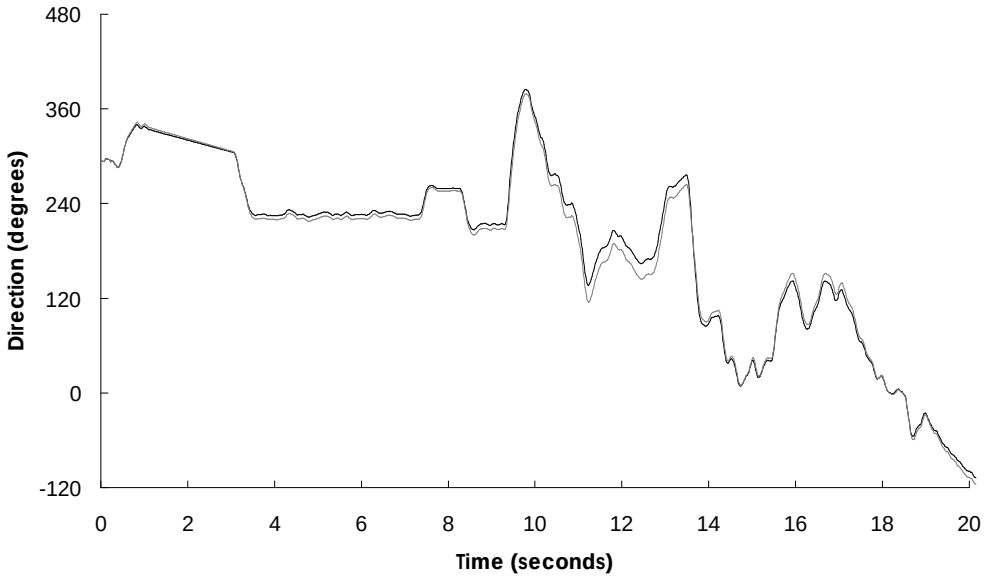

Figure 5. Performance of the model tracking 20 s of rodent head rotations. Head direction data courtesy of Blair and Sharp. The full curve is original data, and the dotted curve is the direction represented by PoS:E population.

Simulations thus consisted of (i) initializing the units to random states, (ii) allowing the two modules to settle to stable attractor states (approx 20–30 ms), (iii) identifying the direction represented in PoS:E with the initial head direction sample, and (iv) allowing the system to run using the $\dot{\phi}$ sequence calculated as per above, and at each step comparing the direction represented in PoS:E with the measured head direction of the animal.

Figure 5 shows the model's ability to integrate actual rodent head movements. The cumulative HD tracking error fluctuated, but typically did not exceed 20° for simulations shorter than 3 min. Figure 6 shows this for the simulation in figure 5.

Tracking accuracy was largely dependent on four parameters: how strongly vestibular input modulated offset connections $\xi(\dot{\phi})$ (see figure 4), the offset amount δ, and the time constants τ_{E} and τ_{I}. The offset connection modulation $\xi(\dot{\phi})$ controlled the angular velocity of the Gaussian for a given input head velocity, while the amount of offset ($\delta = 10°$)

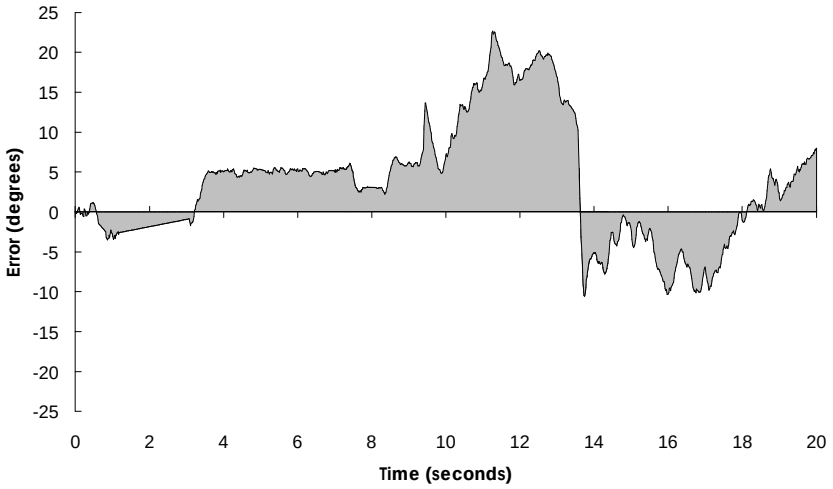

Figure 6. Cumulative tracking error over the course of tracking 20 s of rodent head rotations. Same simulation as figure 5.

controlled the lead of the ATN population. The τ parameters controlled the resistance of units to changing activation, determining the inertia of each pool.

In agreement with [7], we observed that large values of τ_{I} can cause oscillatory behaviour. τ_{E} controlled the qualitative behaviour of the system: small values (< 2 ms) allowed accurate tracking of turns at up to 600° s^{-1} with little distortion of the hill shapes and small differences between PoS and ATN hill locations. Larger values of τ_{E} produced

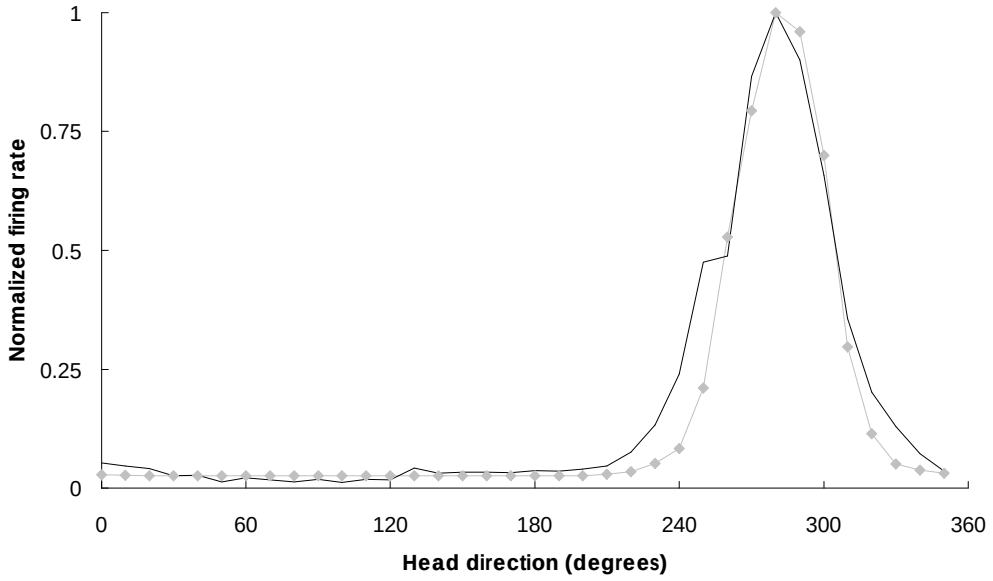

Figure 7. Comparison of tuning curves for an actual postsubiculum cell (full curve) and a simulated PoS:E cell (dotted curve with diamonds) during tracking of a series of rat head rotations. Single-cell recording data courtesy of Blair and Sharp.

distorted hill shapes when tracking fast turns, but yielded larger differences between PoS and ATN hill locations.

Tuning curves of model PoS:E and ATN:E cells (say cell i) were determined by recording $F_i(t)$ at each time t and storing $F_i(t)$ and the actual head direction (not the represented direction) at that point in the simulation. A histogram of F_i was then generated, binned by head direction in 10° bins. Tuning curves of real PoS and ATN cells were generated from spike timing data supplied by Blair and Sharp. And as can be seen in figures 7 and 8, the model shows an excellent fit to the real PoS and ATN cells.

Given the parameters in table 1, the direction represented by the ATN:E population leads the direction represented by the PoS:E population by approximately 10 ms over a wide range of turning speeds. Although this is smaller than the 20–40 ms discrepancy reported in neurophysiological experiments [4, 24], our model at least replicates the qualitative observation that ATN leads PoS. With further refinement, we hope to come closer to replicating the actual lead of ATN relative to the PoS population.

4. Predictions

4.1. Existence of attractors

The attractor nature of the postsubiculum and the anterior thalamic nuclei, an assumption common to several HD models, can be tested by injecting noise into the PoS and/or ATN populations. Because PoS cells are not arranged topographically (i.e. nearby neurons have unrelated preferred directions [18]), noise can be injected into the postsubiculum by microstimulation. The representation of head direction should be disrupted by this procedure, but if the postsubiculum is really an attractor system, the disruption should only be transient. Even in darkness, without external or self-motion cues, the PoS population's

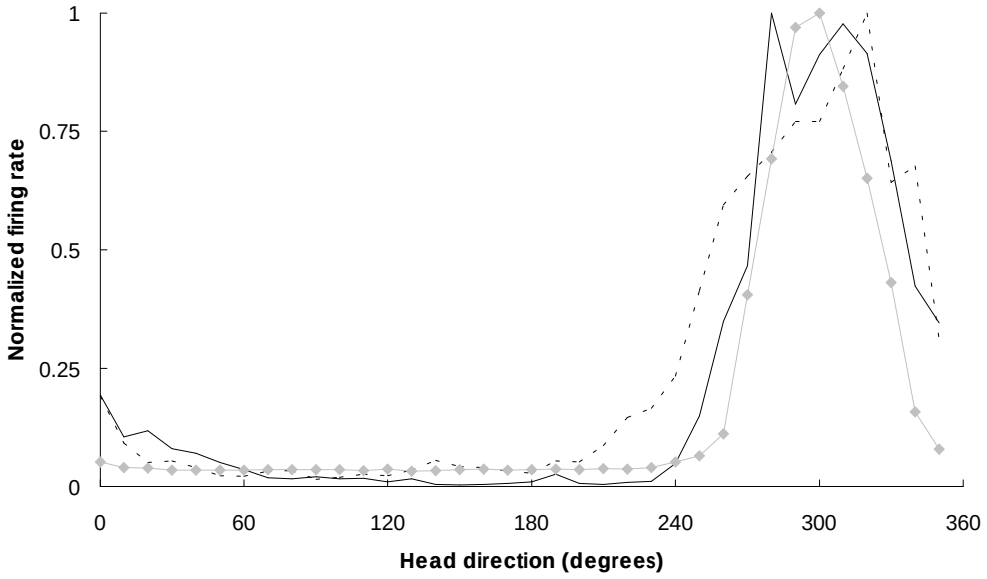

Figure 8. Comparison of tuning curves from two actual ATN cells (full and dashed curves) and a simulated ATN cell (dotted curve with diamonds) during tracking of a series of rat head rotations. Single-cell recording data courtesy of Blair and Sharp.

firing rates should return to a well formed representation of head direction. A similar experiment could test the attractor nature of the ATN population.

The coupled attractor architecture of our model further predicts that the two representations will always be internally consistent and aligned with each other. When two or more HD cells have been recorded simultaneously, the difference between their preferred directions never changes [21, 23]. The attractor nature of our model accounts for this observation and suggests that even after injecting noise into postsubiculum, the difference in preferred direction between two simultaneously recorded cells will return to its original value. In our model, the matching connections align the representations in ATN and PoS, suggesting that the difference in preferred directions between cells in both these areas will also not change from session to session or environment to environment.

4.2. Symmetry of ATN tuning curves

Our model updates the representated head direction by adding extra activation into the ATN, offset by an angle of $\pm\delta$. During rotations, this causes the tuning curves of ATN cells to be stretched to the rotating side. This should be easily testable using the *asymmetry* measurement of Taube *et al* [22], defined as the absolute value of the ratio of the slopes of the left and right legs of the triangular tuning curve. The rotation-side leg (i.e. the right leg for clockwise rotations, the left for anticlockwise rotations) should have a flatter slope than the opposite leg, and faster rotations should produce stronger asymmetries.

5. Discussion

In order to get our model to accurately track the head trajectory of a real animal while maintaining realistic tuning curves in all four populations (excitatory and inhibitory pools of PoS and ATN), we had to make a number of biologically questionable assumptions.

All neurons in both pools of an attractor have tuning curves with similar widths and heights, and preferred directions are evenly distributed around the circle. These simplifying assumptions were made for the purpose of keeping the simulation tractable. Real head direction cells show a wide range of tuning widths and heights [22, 23, 21]. All three assumptions could be relaxed by assuming that there is a distribution around each of these parameters (tuning width, tuning height, preferred direction) and that the values used are average values. We have shown previously how similar assumptions could be relaxed in a model of motor control [17].

The integration time constants τ_E and τ_I are in the millisecond and tenth-of-millisecond range. In order to track high-speed turns, the model requires small integration time constants. We do not see this as a major problem, for two reasons. First, a population of neurons with staggered firing times can produce temporal discriminations much finer than the integration constants of individual cells. As Pinto *et al* [15] have shown, the single-neuron model presented in subsection 2.1 can also be interpreted as a representation of a neuronal population. Second, the fact that a rat can move at high angular velocities does not imply that the head direction system can track them†. Visual cues (if available)

† Etienne [9] has shown that a hamster can be disoriented by moderate rotation (on the order of five full revolutions), which implies that the head direction tracking system accumulates error even at normal turning speeds.

can help realign the head direction representation after the turn. If further investigation of the rodent head direction system shows that it cannot track high velocity turns without error, then the model can use more realistic time constants. With values of $\tau_E = 10\,\text{ms}$ and $\tau_I = 2\,\text{ms}$, and an offset value $\delta = 50^\circ$, we have determined a $\xi(\dot{\phi})$ function that will track head direction accurately for turning speeds up to $100^\circ\ \text{s}^{-1}$.

The model includes intrinsic connections within ATN. There are few or no intrinsic connections within the anterior thalamic nuclei, and no evidence for inhibitory interneurons within the anterior dorsal (AD) nucleus† (see [2] for a review). We have found that if ATN is not organized as an attractor but merely sums the inputs from matching and offset connections, ATN tuning curves deform drastically during rotations. Since this has not been observed in the rat, we suggest that ATN does have attractor dynamics. However, ATN does not appear to contain the necessary machinery for an attractor module. We do not know which anatomical structures interconnected with ATN might contribute to its attractor nature, but we offer this as a prediction.

The model requires multiplicative connections between the PoS and ATN populations, because the strength of the offset connections is a function of angular velocity. No such connections have been seen, however, the crucial aspect of the theory is that the loop from PoS to ATN and back must be modulated by vestibular input. Both Taube *et al* [22] and Sharp [18] report that some cells in postsubiculum are sensitive to both head direction and angular velocity. If these PoS cells project to the ATN:E population, they can serve as the origin of the offset connections, and then the offset connections would not have to be modulated by angular velocity since the presynaptic neurons already are. Another possible source of the offset connections is retrosplenial cortex, which includes head direction cells modulated by self-motion [5, 6] and is interconnected with the anterior thalamic nuclei [2, 26].

Cells in the mammillary bodies are assumed to fire at a rate proportional to angular velocity. Blair [3] makes a similar proposal, except that he assumes firing rate *decreases* with angular velocity. Our model requires an inhibitory gain control input to the ATN:E population because without it, the combined input from the offset and matching connections distorts the ATN tuning curves. In the cat, the effect of stimulating the mammillothalamic tract is inhibitory, but the actual synapses seen in rat would seem to be excitatory; see [2] for a review. We identify this gain control population with the mammillary bodies, but more work needs to be done to determine whether the mammillary bodies could actually play this role. As yet there are no published data on neural recordings from the mammillary bodies.

In summary, we have described a coupled attractor model of the rodent head direction system whose components closely match the tuning curves of cells in PoS and ATN. The model tracks actual head rotations with good accuracy, and the ATN representation of head direction leads the PoS representation during turns, in qualitative agreement with neurophysiological observations. The model makes explicit predictions about the attractor nature of PoS and ATN representations, and the shape of ATN tuning curves during rotations, which we hope to see tested in future experiments.

† Most head direction neurons in ATN probably come from the anterior dorsal (AD) nucleus [4, 21].

Acknowledgments

We thank Bard Ermentrout for crucial assistance with attractor dynamics, David Pinto for illuminating discussions, and Tad Blair and Pat Sharp for generously sharing their rat recording data with us. We also acknowledge two anonymous reviewers for their helpful comments. Adam Elga was supported by the Neural Processes in Cognition Summer Program 1995.

References

[1] Anderson C H and Van Essen D C 1987 Shifter circuits: a computational strategy for dynamic aspects of visual processing *Proc. Natl Acad. Sci. USA* **84** 6297–301

[2] Bentivoglio M, Kultas-Ilinsky K and Illinsky I 1993 Limbic thalamus: Structure, intrinsic organisation, and connections *Neurobiology of Cingulate Cortex and Limbic Thalamus: A Comprehensive Handbook* ed B A Vogt and M Gabriel (Basel: Birkhäuser) pp 71–122

[3] Blair H T 1996 A thalamocortical circuit for computing directional heading in the rat *Advances in Neural Information Processing Systems* vol 8, ed D S Touretzky, M C Mozer and M E Hasselmo (Cambridge, MA: MIT Press) pp 152–8

[4] Blair H T and Sharp P E 1995 Anticipatory head direction signals in anterior thalamus: Evidence for a thalamocortical circuit that integrates angular head motion to compute head direction *J. Neurosci.* **15** 6260–70

[5] Chen L L, Lin L H, Barnes C A and McNaughton B L 1994 Head-direction cells in the rat posterior cortex: II. Contributions of visual and ideothetic information to the directional firing *Exp. Brain Res.* **101** 24–34

[6] Chen L L, Lin L H, Green E J, Barnes C A and McNaughton B L 1994 Head-direction cells in the rat posterior cortex: I. Anatomical distribution and behavioral modulation *Exp. Brain Res.* **101** 8–23

[7] Ermentrout B and Cowan J 1979 A mathematical theory of visual hallucination patterns *Biol. Cybern.* **34** 137–50

[8] Ermentrout B 1995 *XTC—a tool for modeling spatial evolution equations* Computer program and documentation available at ftp://ftp.math.pitt.edu/pub/bardware/

[9] Etienne A S 1987 The control of short-distance homing in the golden hamster *Cognitive Processes and Spatial Orientation in Animals and Man* ed P Ellen and C Thinus-Blanc (Dordrecht: Martinus Nijhoff) pp 233–51

[10] Georgopoulos A P, Caminiti R, Kalaska J F and Massey J T 1983 Spatial coding of movement: a hypothesis concerning the coding of movement direction by motor cortical populations *Neural Coding of Motor Performance* ed J Maisson *et al*, *Exp. Brain Res.* Suppl. 7 327–36

[11] Knierim J J, Kudrimoti H S and McNaughton B L 1995 Place cells, head direction cells, and the learning of landmark stability *J. Neurosci.* **15** 1648–59

[12] Mardia K V 1972 *Statistics of Directional Data* (New York: Academic)

[13] McNaughton B L, Chen L L and Markus E J 1991 'Dead reckoning,' landmark learning, and the sense of direction: a neurophysiological and computational hypothesis *J. Cognitive Neurosci.* **3** 190–202

[14] Murray J D 1989 *Biomathematics* vol 19 *Mathematical Biology* (Berlin: Springer)

[15] Pinto D J, Brumberg J C, Simons D J and Ermentrout G B 1996 A quantitative population model of whisker barrels: re-examining the Wilson–Cowan equations *J. Computat. Neurosci.* **3** 247–64

[16] Ranck J B Jr 1984 Head-direction cells in the deep cell layers of dorsal presubiculum in freely moving rats *Soc. Neurosci. Abs.* **10** 599

[17] Redish A D and Touretzky D S 1994 The reaching task: Evidence for vector arithmetic in the motor system? *Biol. Cybern.* **71** 307–17

[18] Sharp P E 1996 Multiple spatial/behavioral correlates for cells in the rat postsubiculum: multiple regression analysis and comparison to other hippocampal areas *Cerebral Cortex* **6** 238–59

[19] Sharp P E, Blair H T, Etkin D and Tzanetos D B 1995 Influences of vestibular and visual motion information on the spatial firing patterns of hippocampal place cells *J. Neurosci.* **15** 173–89

[20] Skaggs W E, Knierim J J, Kudrimoti H S and McNaughton B L 1995 A model of the neural basis of the rat's sense of direction *Advances in Neural Information Processing Systems 7* ed G Tesauro, D S Touretzky and T K Leen (Cambridge, MA: MIT Press) pp 173–80

[21] Taube J S 1995 Head direction cells recorded in the anterior thalamic nuclei of freely moving rats *J. Neurosci.* **15** 1953–71

[22] Taube J S, Muller R I and Ranck J B Jr 1990 Head direction cells recorded from the postsubiculum in freely moving rats. I. Description and quantitative analysis *J. Neurosci.* **10** 420–35
[23] Taube J S, Muller R I and Ranck J B Jr 1990 Head direction cells recorded from the postsubiculum in freely moving rats. II. Effects of environmental manipulations *J. Neurosci.* **10** 436–47
[24] Taube J S and Muller R U 1995 Head direction cell activity in the anterior thalamic nuclei, but not the postsubiculum, predicts the animal's future directional heading *Soc. Neurosci. Abs.* **21** 946
[25] Touretzky D S, Redish A D and Wan H S 1993 Neural representation of space using sinusoidal arrays *Neural Comput.* **5** 869–84
[26] van Groen T, Vogt B A and Wyss J M 1993 Interconnections between the thalamus and retrosplenial cortex in the rodent brain *Neurobiology of Cingulate Cortex and Limbic Thalamus: A Comprehensive Handbook* ed B A Vogt and M Gabriel (Basel: Birkhäuser) pp 123–50
[27] van Groen T and Wyss J M 1990 The postsubicular cortex in the rat: Characterization of the fourth region of the subicular cortex and its connections *Brain Res.* **529** 165–77
[28] Wilson H R and Cowan J D 1972 Excitatory and inhibitory interactions in localized populations of model neurons *Biophys. J.* **12** 1–24
[29] Zhang K 1996 Representation of spatial orientation by the intrinsic dynamics of the head-direction cell ensemble: A theory *J. Neurosci.* **16** 2112–26