NX-414: Brain-like computation and intelligence

Alexander Mathis
alexander.mathis@epfl.ch

Lecture 3, March 5

Bayesian perception

One may even say, strictly speaking, that almost all our knowledge is only probable; and in the small number of things that we are able to know with certainty, the principles means of arriving at the truth -induction and analogy- are based on probabilities.

Pierre Simon Laplace. Theorie analytique des probabilities. 1825.

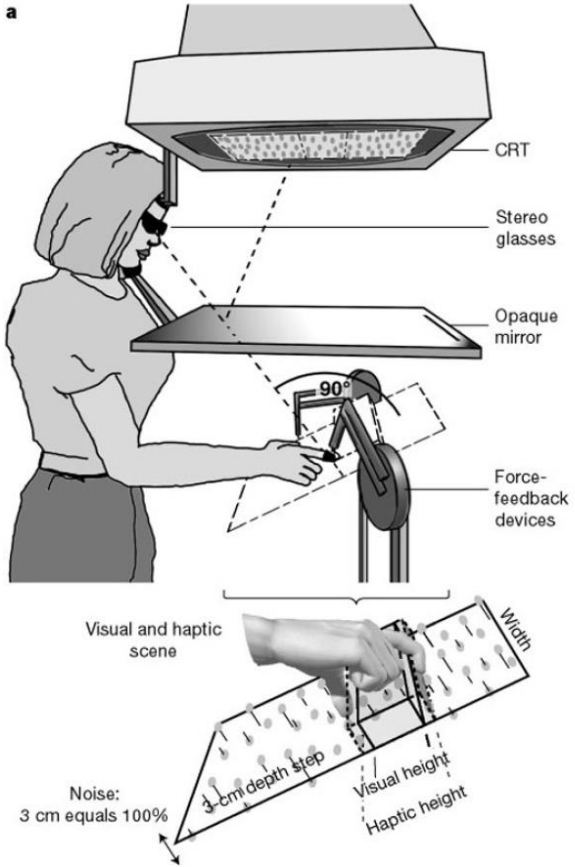
Perception is a probabilistic process

- $P(k|x)$ the conditional probability of evoking spike rate k given that stimulus x was presented (called likelihood for observed k)
- $P(x|k)$ the conditional probability of observing stimulus x given k was recorded

So far we have seen this for k spikes, but the representation can also be modeled as an “abstract variable”.

How does the brain create a coherent and unambiguous percept based on multiple sensory streams?

Visual-haptic cue integration



Paradigm: 2-interval forced-choice (2-IFC) task:

Is h taller than h0?

Description: In the visual-haptic setup, observers view the reflection of the visual stimulus binocularly in a mirror using stereo goggles. The haptic stimulus is presented using two force-feedback devices, one each for the index finger and thumb. With this setup the visual and the haptic virtual scenes can be independently manipulated.

Ernst, M. O. (2006). A Bayesian View on Multimodal Cue Integration Oxford University Press.

Bayesian formulation

joint estimate

tactile estimate

visual estimate

$$\hat{w} = \text{argmax}_w \, P(\text{w}|\text{t}, \text{v})$$

Generative model

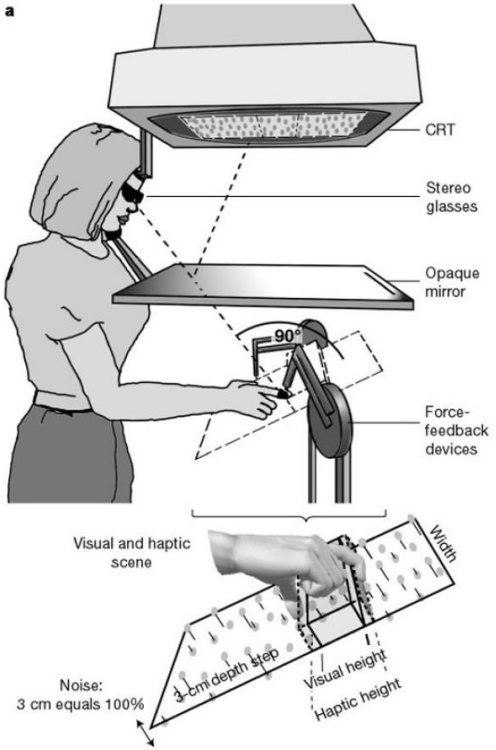
$$P(\text{w}|\text{t}, \text{v}) = \frac{P(t, v \,|w)\, P(w)}{P(t, v)}$$

Cond. independence assumption

$$= \frac{P(t|w)P(v|w)\, P(w)}{P(t, v)}$$

$$\propto P(t|w)P(v|w)\, P(w)$$

Bayesian formulation



Generative model: $P(t, v \mid w)$

$$P(v|w) = \frac{1}{\sqrt{2\pi\sigma_v^2}} e^{-\frac{(w-v)^2}{2\sigma_v^2}}$$

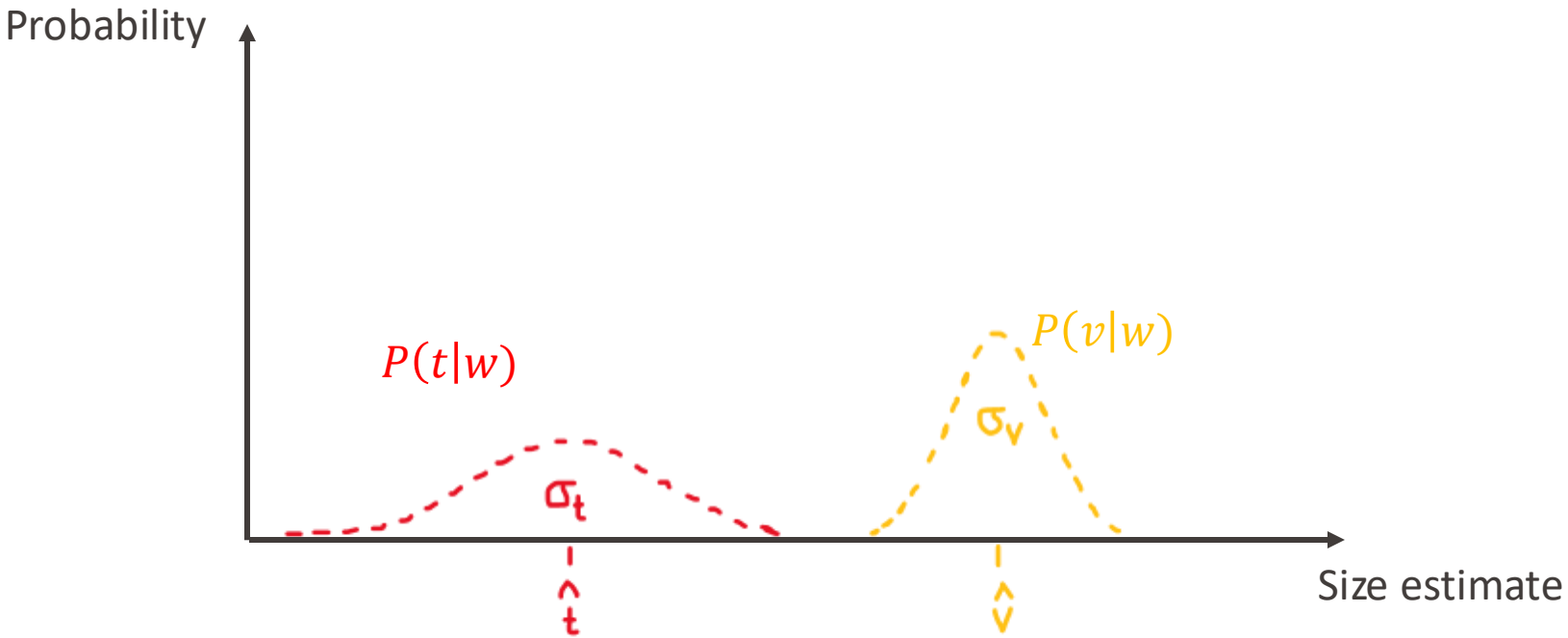
$P(v|w)$ $\quad v = w + \eta$

w

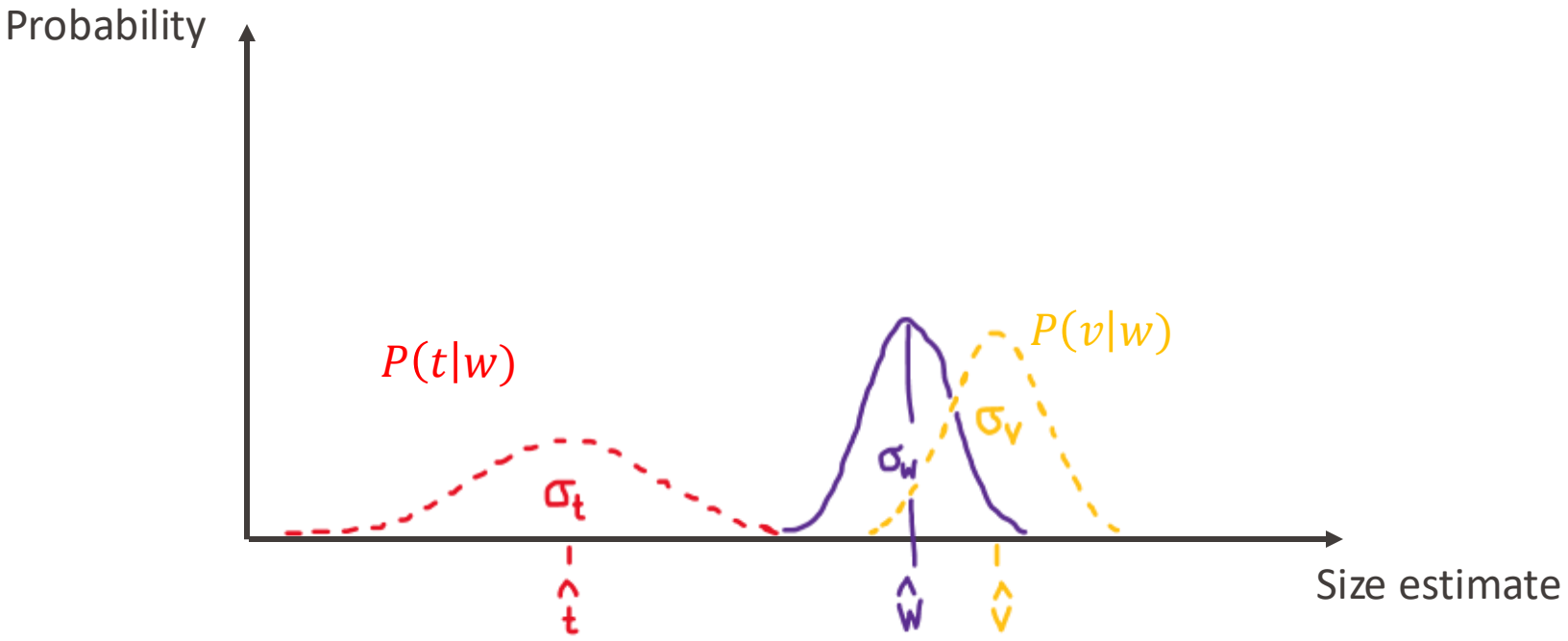
$P(t|w)$ $\quad t = w + \eta$

$$P(t|w) = \frac{1}{\sqrt{2\pi\sigma_t^2}} e^{-\frac{(w-t)^2}{2\sigma_t^2}}$$

Bayesian cue integration



Bayesian cue integration



$$P(\mathrm{w}|\mathrm{t},\mathrm{v}) \propto P(t|w)P(v|w)$$

Mean and variance of the joint estimate

$$P(\mathrm{w}|\mathrm{t},\mathrm{v}) \propto P(t|w)P(v|w)$$

$$\mu_w = \frac{\sigma_t^2}{\sigma_t^2 + \sigma_v^2} v + \frac{\sigma_v^2}{\sigma_t^2 + \sigma_v^2} t$$

$$\sigma_w^2 = \frac{\sigma_t^2 \sigma_v^2}{\sigma_t^2 + \sigma_v^2}$$

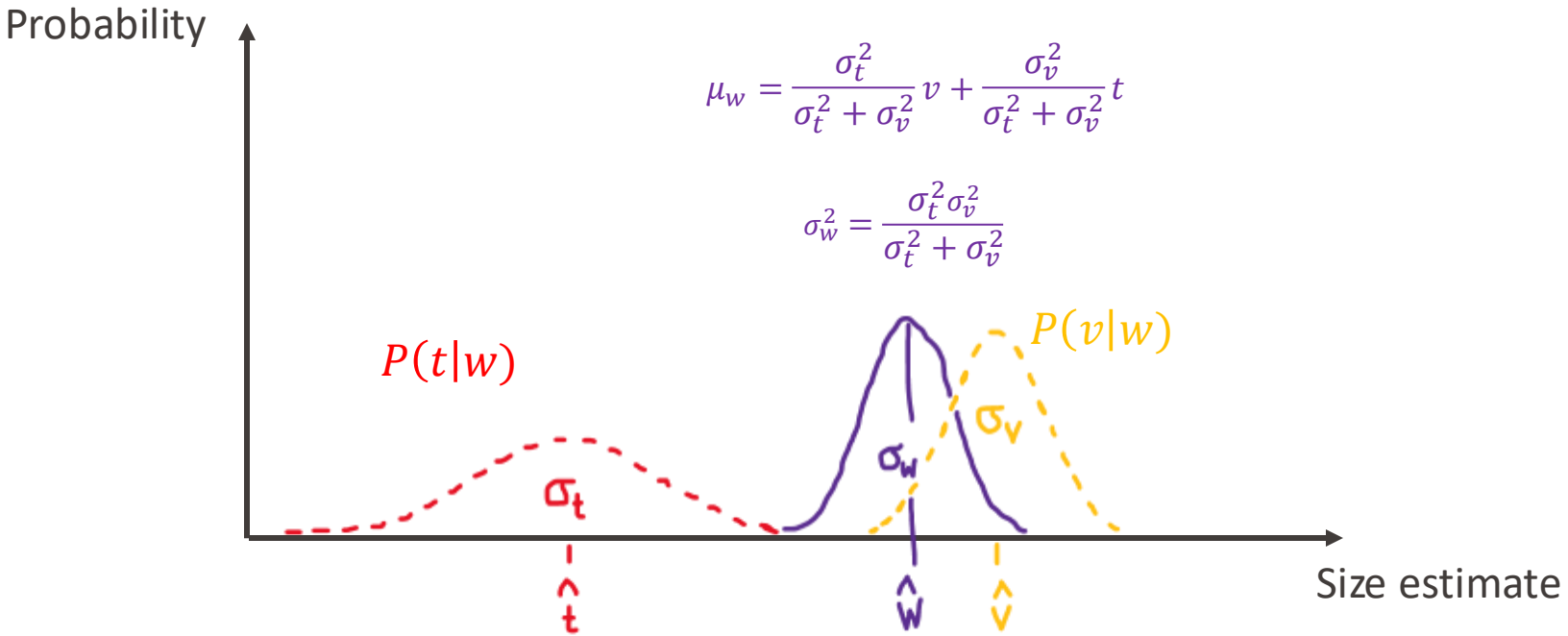
Or equivalently:

$$\mu_w = \frac{\sigma_t^2 \sigma_v^2}{\sigma_t^2 + \sigma_v^2} \left(\frac{1}{\sigma_v^2} v + \frac{1}{\sigma_t^2} t \right)$$

$$\frac{1}{\sigma_w^2} = \frac{1}{\sigma_v^2} + \frac{1}{\sigma_t^2}$$

You'll demonstrate this in the exercises!

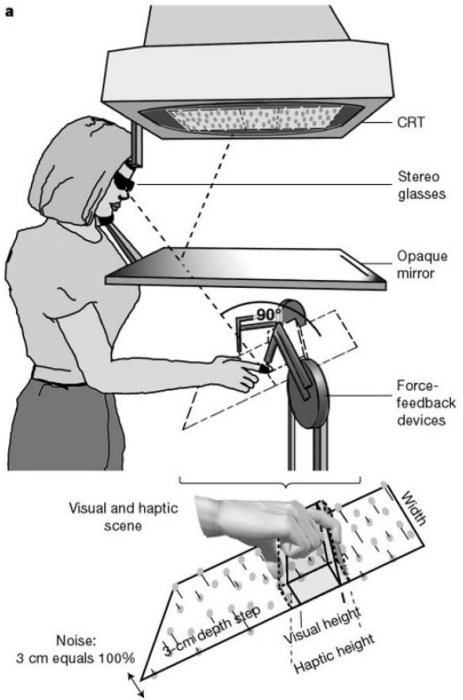
Bayesian cue integration

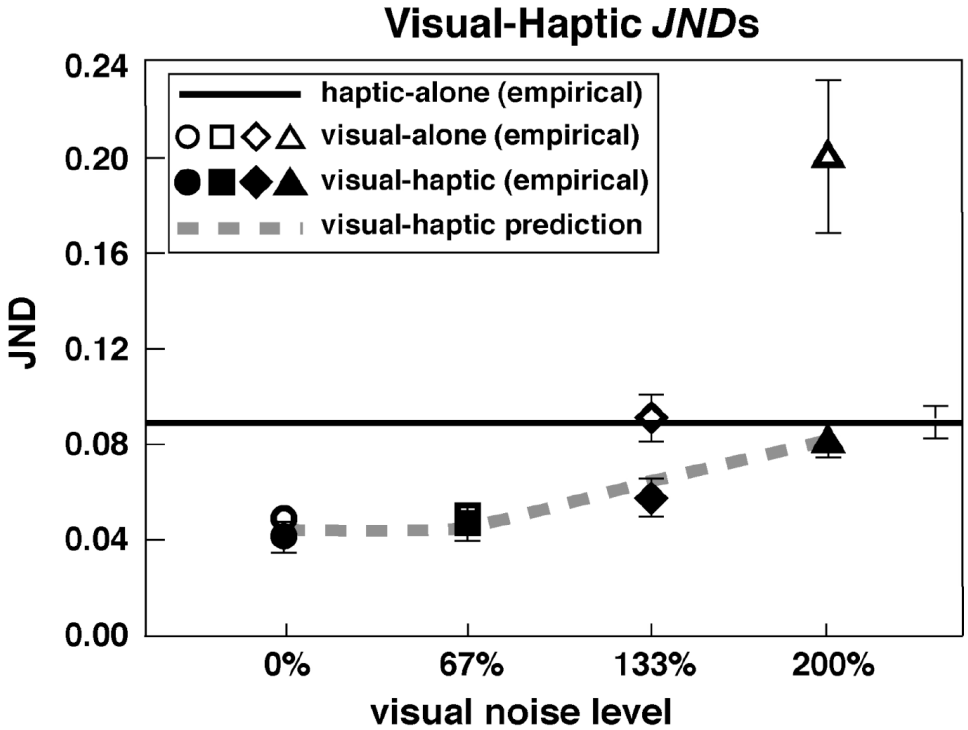


$$P(\mathrm{w}|\mathrm{t},\mathrm{v}) \propto P(t|w)P(v|w)$$

Do humans behave according to this theory?

Humans (can) optimally integrate vision and haptic information





Ernst, M. O. (2006). A Bayesian View on Multimodal Cue Integration Oxford University Press.
Ernst & Banks, Nature 2002

Path integration and attractor models

Animal navigation

as captured from 11 March 2021 to 3 November 2021 via GPS

Trends in Ecology & Evolution

https://animallives.org

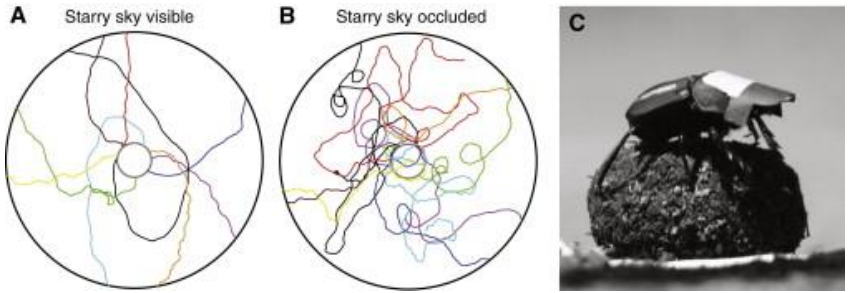
How do animals navigate?

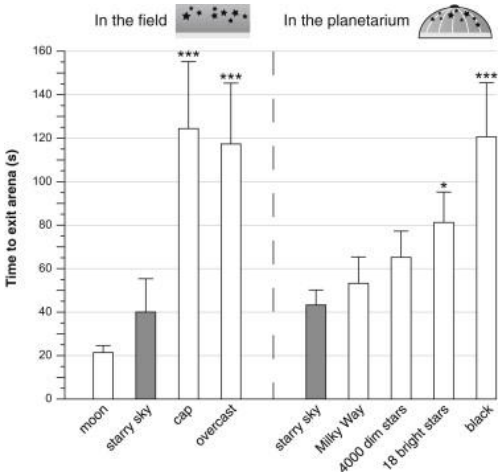
Using various external cues:

- Remembering landmarks
- Magnetoreception
- Olfaction
- Polarized light (e.g., bees)
- Sun/moon/stars
- Way-marking (e.g., wood-mice)

Alternative:

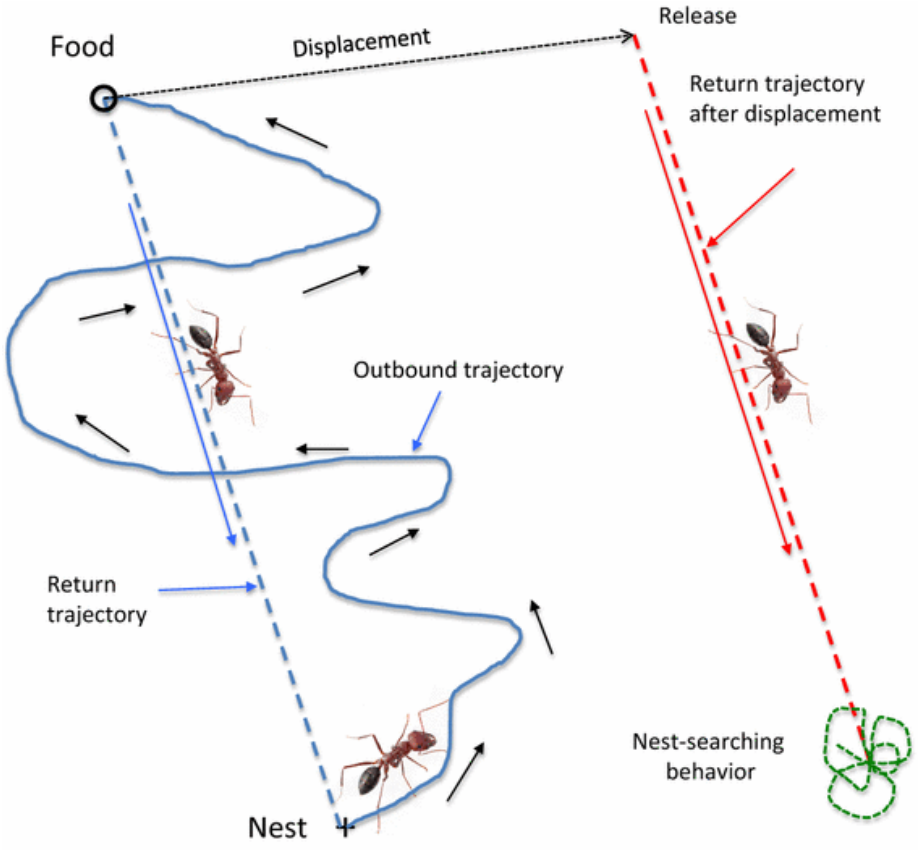
By path integration/dead reckoning…





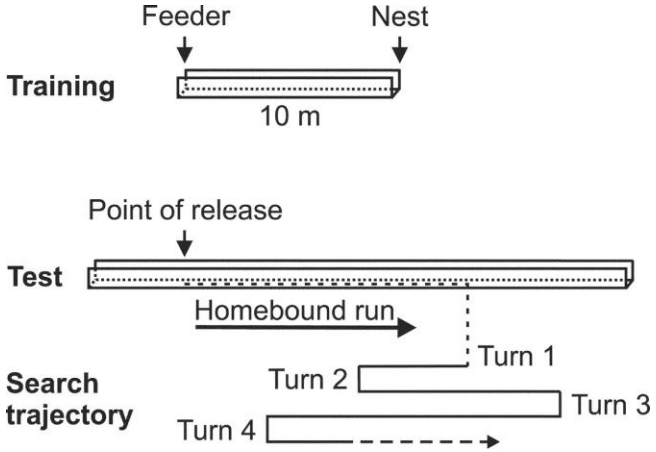
Dung Beetles Use the Milky Way for Orientation, Dacke et al. Current Biology 2013

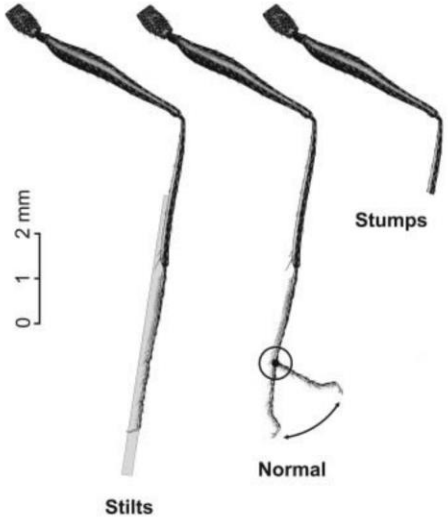
Path integration

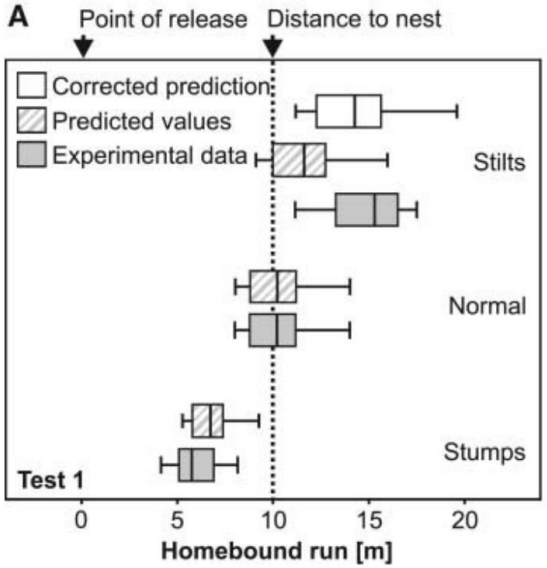


Srinivasan, Journal of Comparative Physiology A, 2015

How do ants estimate the distance?

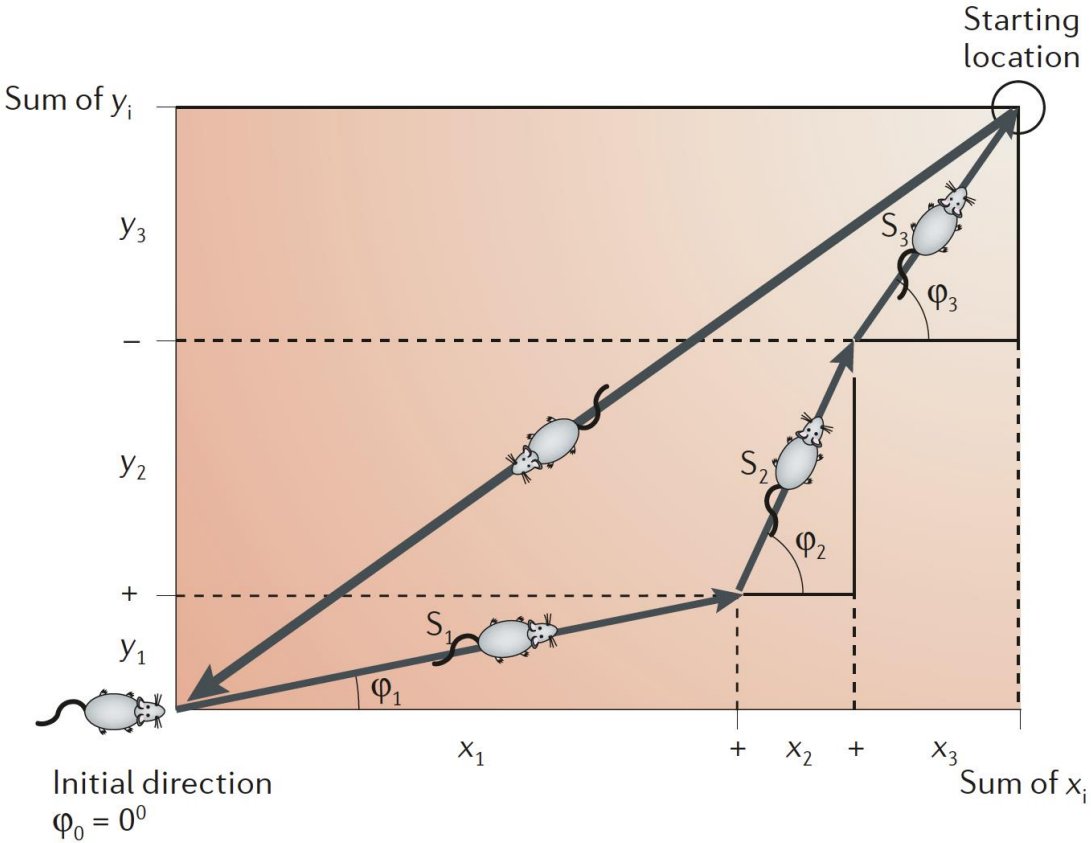






Wittlinger, Wehner & Wolf, Science 2006

Path integration



McNaugthon et al., Nature Review Neuroscience 2006

Head directions (angular orientation)

Head-direction cells (in rats)

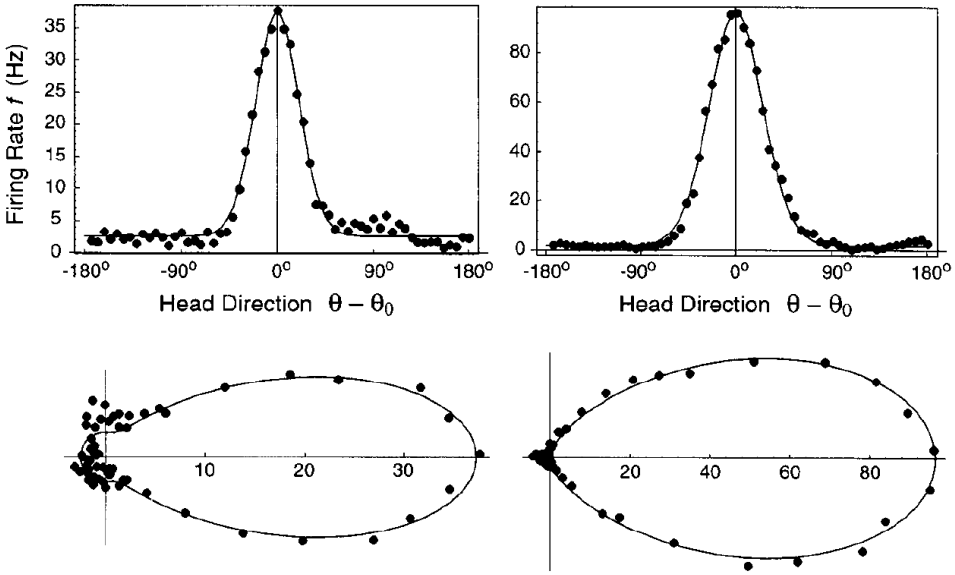


Figure 2. A typical HD cell has stereotyped directional tuning curve. Its firing rate reaches maximum when the head direction θ is aligned with the preferred direction θ_0. Data points and fitting curves are shown in both Cartesian and polar coordinates. *A*, An anterior thalamic HD cell with medium peak rate. Data from Figure 4*B* in Taube (1995). Parameters: $K = 8.08$, $A = 2.53$ Hz, and $Be^K = 34.8$ Hz. *B*, A postsubicular HD cell with high peak rate. Data from Figure 3*C* in Taube et al. (1990a). Parameters: $K = 5.29$, $A = 1.72$ Hz, and $Be^K = 94.8$ Hz. (Digital data courtesy of J. S. Taube.)

Fit with von Mises tuning curve (circular Gaussian)

$$f = A + Be^{K\cos(\theta-\theta_0)}$$

Cells found by Ranck/Taube …
Figure adapted from Zhang 1996, J. Neurosci

Head-direction properties

- HD cells only fire when head of rat is pointing in the preferred direction in the horizontal plane, regardless of location
- HD preferred direction is stable across environments (unless animal is disoriented)
- If linked to landmarks, rotating landmarks rotates HD cells; HD cells can be associated with a landmark within minutes (and remember the association)
- HD fire in total darkness (presumably by integration self-motion cues!)
- However, HD direction may drift during long recordings in the dark (see next slide)
- Preferred directions of different HDs are tightly coupled, so that a landmark rotation induces a consistent rotation (stable differences)!

Zhang 1996, J. Neurosci

A drifting multi-unit recording

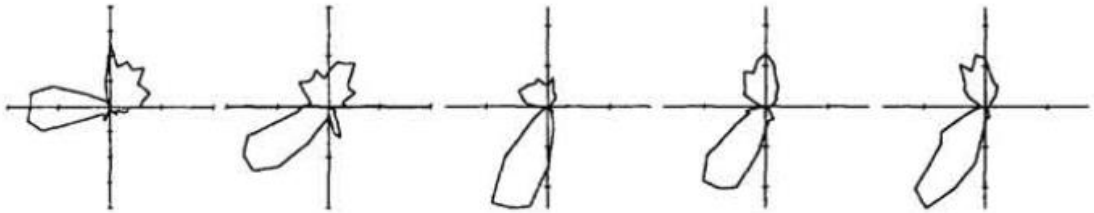
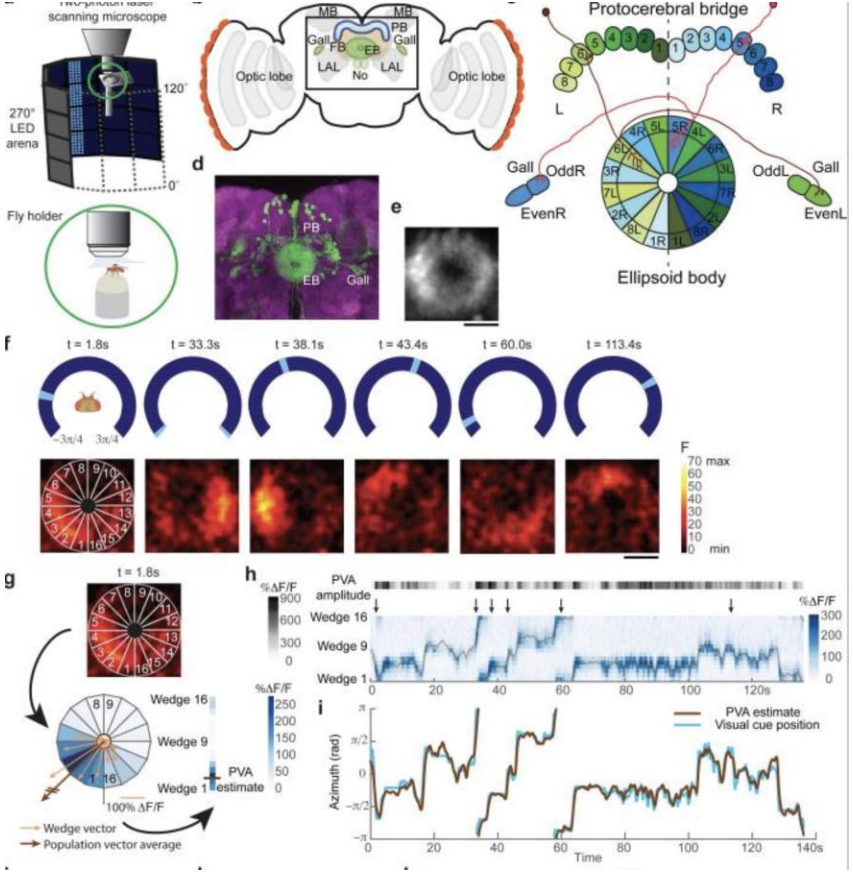


Figure 2: Shifts in alignment of a head direction cell over the course of a single recording session (one minute intervals).

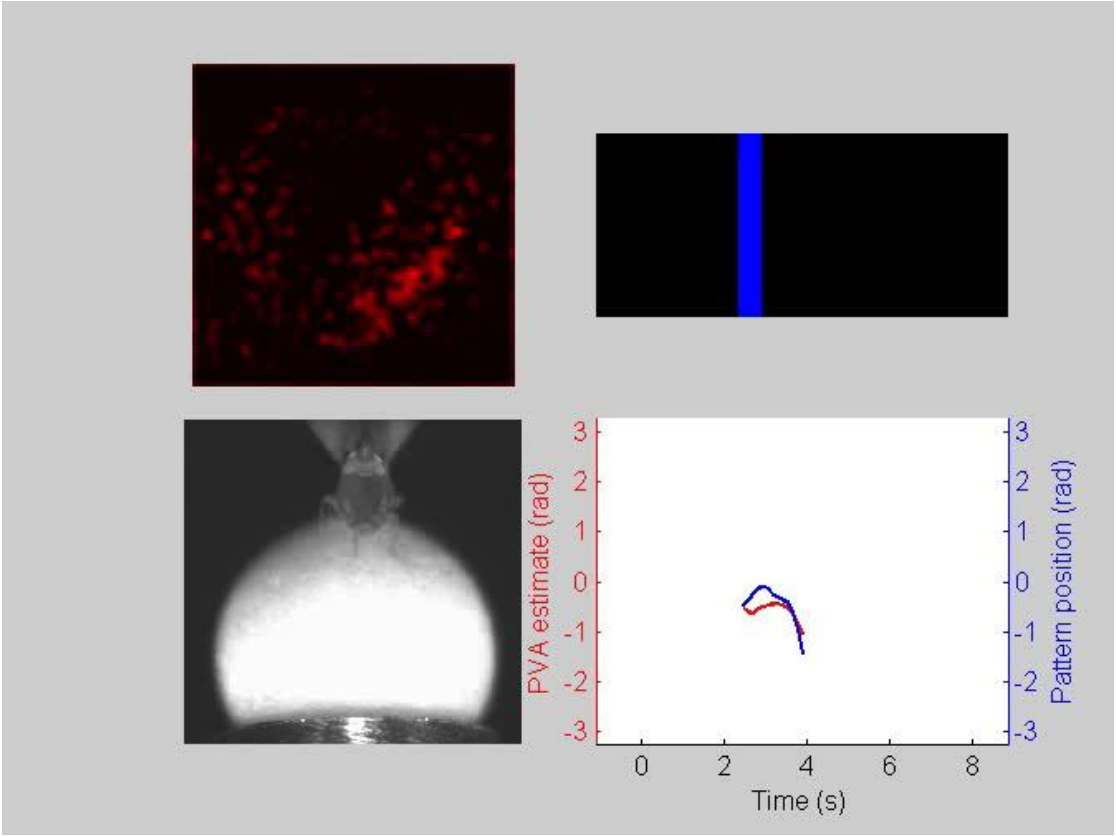
Skaggs et al. 1995, NeurIPS

Neural dynamics for landmark orientation and angular path integration



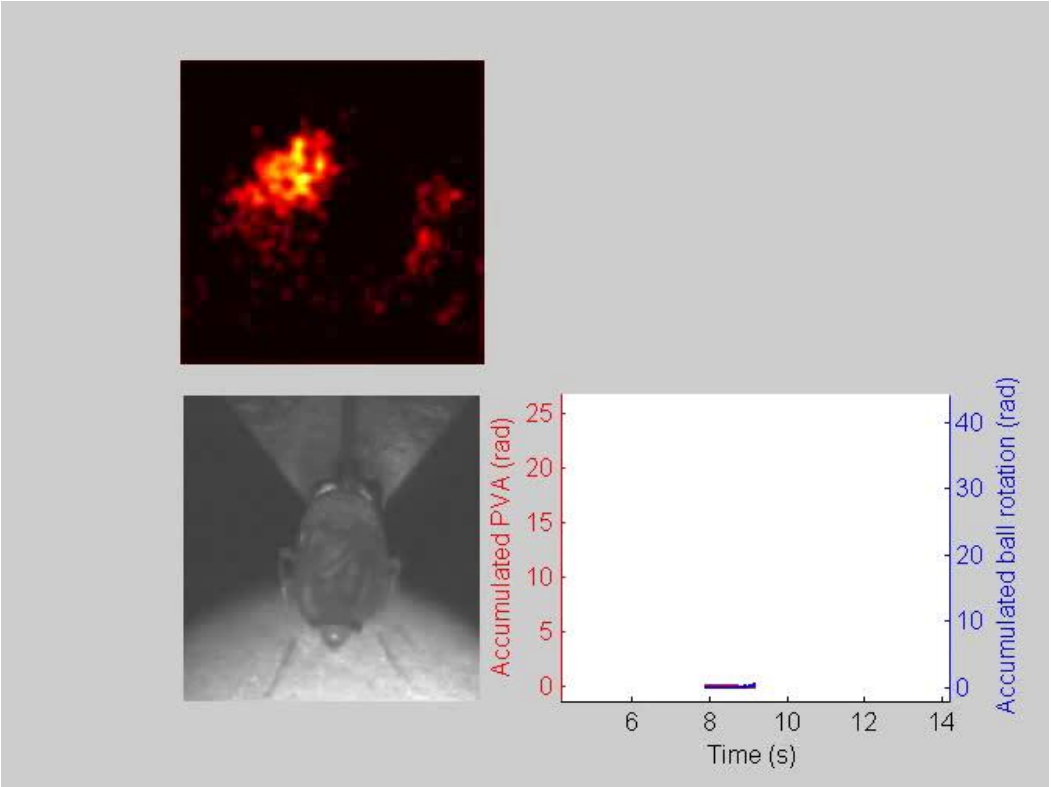
Seelig & Jayaraman Nature 2013

Compass-like representation of landmark orientation



Seelig & Jayaraman Nature 2013

Activity of the central complex in the absence of a visual cue (darkness)



Seelig & Jayaraman Nature 2013

Recurrent neural networks and path integration

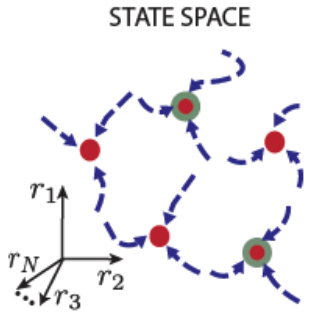
Questions:

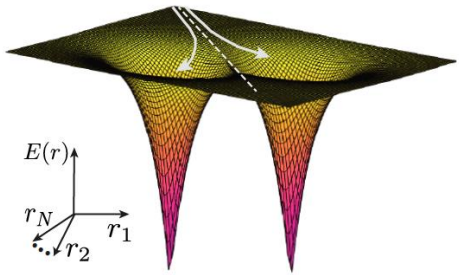
1. How to *engineer* neural systems for path integration? -> ring attractors
2. Next week: How to *learn* path integration from scratch?
3. Next week: Are the solutions related?

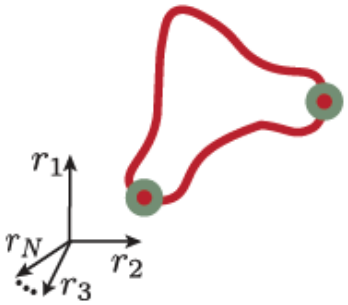
Recap: path integration is a fundamental ability that depends on accumulating *velocity* signals (from the vestibular, proprioceptive …senses) to form a representation where one is in space.

In mammals, head direction, grid and place cells have been implicated

Attractor models





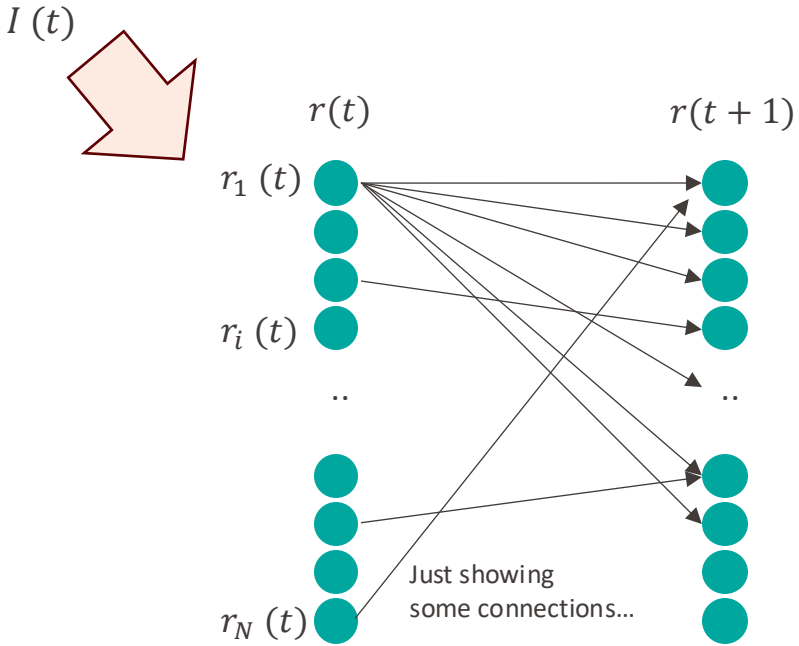


Functions:

- Associative memory
- Decision making
- Path integration
- …

Khona & Fiete, Nat Review Neuro 2021

Recurrent neural network (RNN)

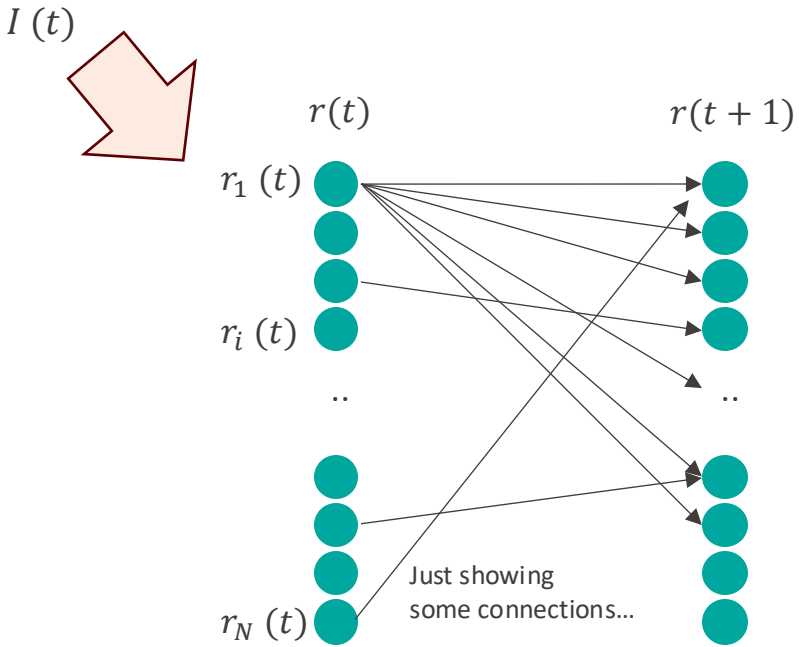


rate of neuron *i* at time *t*

Rate update equation

$$r(t+1) = W\, r(t) + I\ (t)$$

Recurrent neural network (RNN)



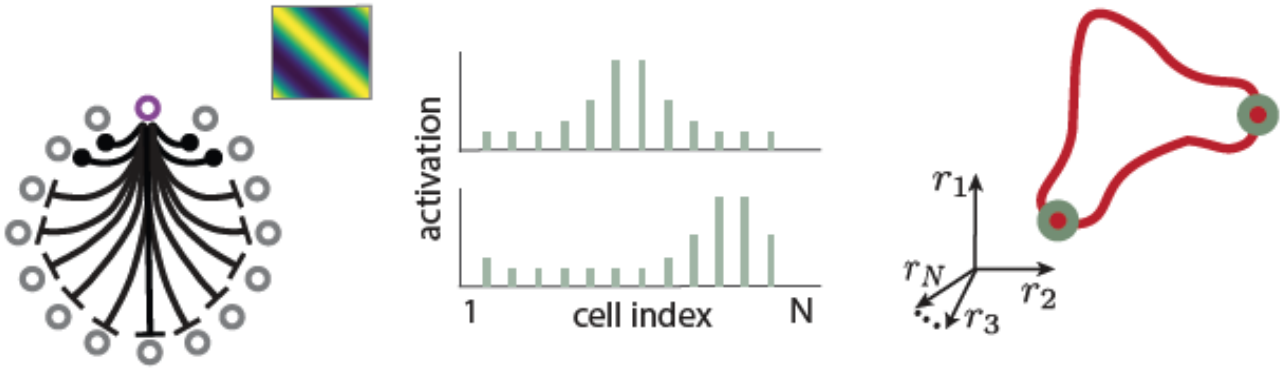
Rate + membrane equation

$$u(t+1) = Wr(t) + I(t)$$

$$\mathrm{r(t+1)} = \sigma(u(t+1))$$

element-wise nonlinearity

Ring attractor models



Khona & Fiete, Nat Review Neuro 2021

Average firing rate for all HD cells with same preferred phase θ:

$$f = f(\theta, t)$$

Similarly,

$$u = u(\theta, t)$$

defines the average net inputs received by these units. The input is linked to the firing rate by some nonlinearity:

$$f = \sigma(u)$$

The dynamics are governed continuous dynamics :

$$\tau \frac{\partial u}{\partial t} = -u + w * f$$

With $w * f = 1/2\pi \int_0^{2\pi} w(\theta - \varphi, t) f(\varphi) d\varphi$

Zhang 1996, J. Neurosci

What do we know & what do we want to know?

$$\tau \frac{\partial u}{\partial t} = -u + w * f$$

Notes

- The dynamics are the dynamics of a typical recurrent neural network (RNN)
- We built-in radial symmetry for the weights

Can we find weights, so that we get head-direction cells?

$$\tau \frac{\partial u}{\partial t} = -u + w * f$$

Finding a weight kernel that solves the equation

$$\tau \frac{\partial u}{\partial t} = -u + w * f$$

(Non-regularized version doesn't converge!)

We can find good approximate solutions, by encouraging "small solutions":

Fredholm-integral equation of first kind:

$$\mathcal{L} = \frac{1}{2\pi} \int_0^{2\pi} (w * f - u)^2 \, d\varphi + \frac{\lambda}{2\pi} \int_0^{2\pi} \mathrm{w}^2 \, d\varphi$$

$$= \sum_{n=-\infty}^{\infty} |\widehat{u_n} - \widehat{w_n}\widehat{f_n}|^2 + \lambda \sum_{n=-\infty}^{\infty} |\widehat{w_n}|^2$$

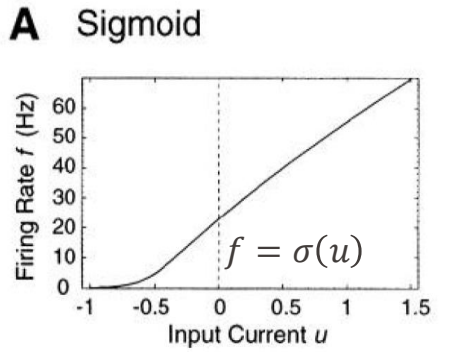
$$\widehat{w_n} = \frac{\widehat{u_n}\,\widehat{f_n}}{\lambda + |\widehat{f_n}|^2}$$

We can get the weights
via the inverse Fourier transform

Note that $\widehat{u_n}$ is given via $f = \sigma(u)$, where is σ invertible!

Zhang 1996, J. Neurosci

Synaptic weights supporting stable states



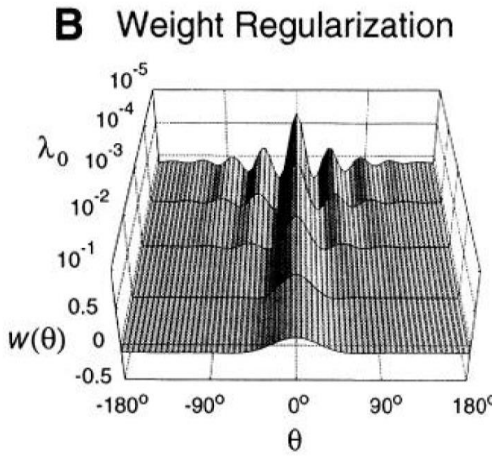


Figure 4. Synaptic weight distribution supporting stable activity profiles assuming sigmoid function of the form in Equation A. In *A*, the input current is in arbitrary units. Large firing rate at zero current is caused by the constant bias c. Parameters: $\beta = 0.8$, $b = 10$, $c = 0.5$, and $a = 6.34$ (determined by the scaling condition $\sigma(1 - c) = f_{max} = 40$ Hz). *B*, Synaptic weight distribution function $w(\theta)$ solved at different levels of regularization for smallness. Function $w(\theta)$ describes the average strength of synaptic weights between units whose preferred directions differ by the angle θ. For appropriate scaling, we use $\lambda_0 \equiv \lambda/\max|\hat{f}_n|^2$ to quantify the regularization. The desired static profile is of the form in Equation 1 with $K = 8$, $A = 1$ Hz, and $f_{max} = A + Be^K = 40$ Hz. *C*, When the weight regularization is too strong, the actual stable firing profile tends to be blunter than the desired one, or even becomes totally flat (not shown). On the other hand, when the regularization is too weak, the stable profile may suddenly become multiple-peaked.

Zhang 1996, J. Neurosci

Synaptic weights supporting stable states

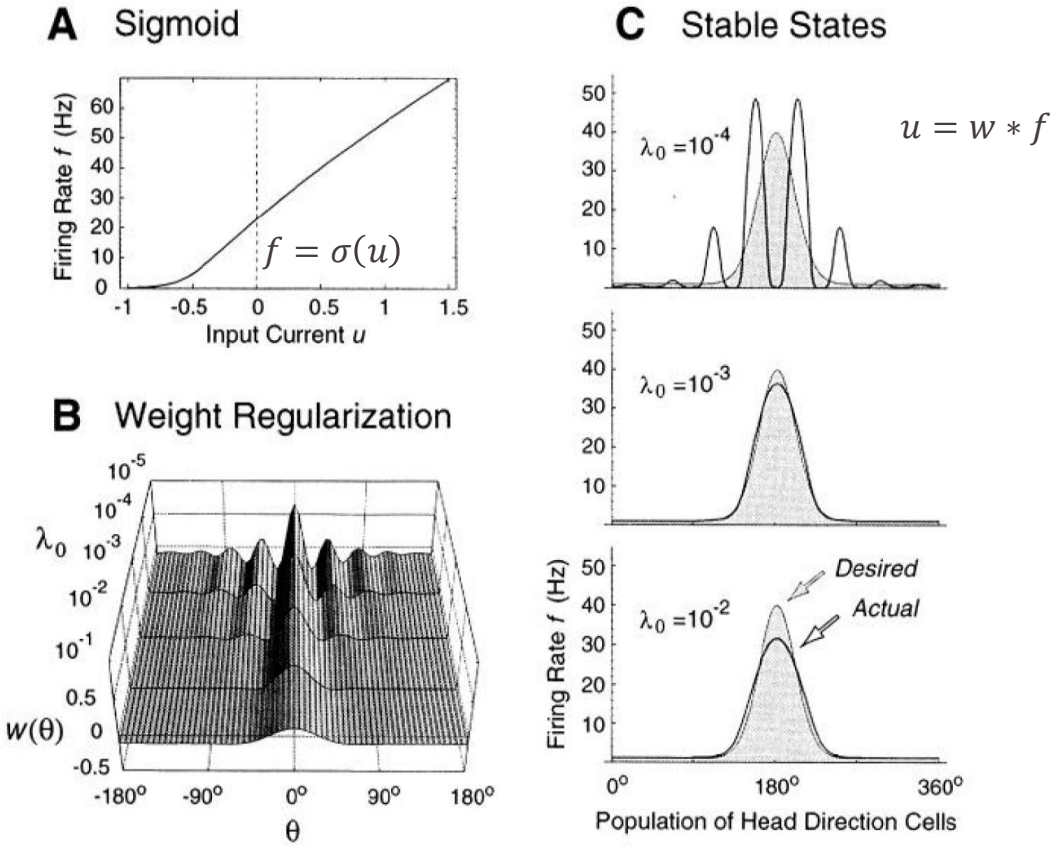
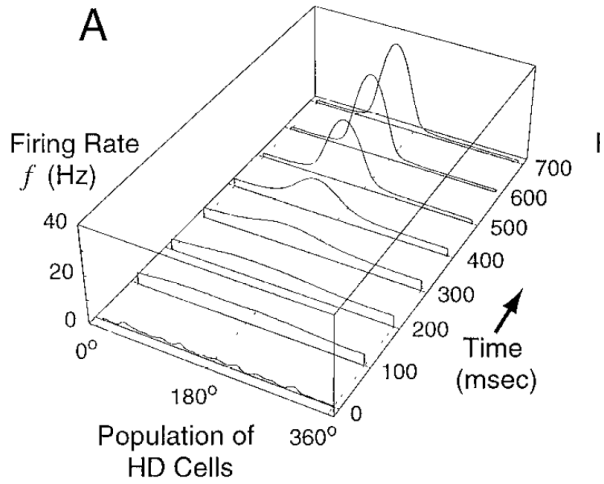


Figure 4. Synaptic weight distribution supporting stable activity profiles assuming sigmoid function of the form in Equation A. In *A*, the input current is in arbitrary units. Large firing rate at zero current is caused by the constant bias c. Parameters: $\beta = 0.8$, $b = 10$, $c = 0.5$, and $a = 6.34$ (determined by the scaling condition $\sigma(1 - c) = f_{max} = 40$ Hz). *B*, Synaptic weight distribution function $w(\theta)$ solved at different levels of regularization for smallness. Function $w(\theta)$ describes the average strength of synaptic weights between units whose preferred directions differ by the angle θ. For appropriate scaling, we use $\lambda_0 \equiv \lambda/\max|\hat{f}_n|^2$ to quantify the regularization. The desired static profile is of the form in Equation 1 with $K = 8$, $A = 1$ Hz, and $f_{max} = A + Be^K = 40$ Hz. *C*, When the weight regularization is too strong, the actual stable firing profile tends to be blunter than the desired one, or even becomes totally flat (not shown). On the other hand, when the regularization is too weak, the stable profile may suddenly become multiple-peaked.

Zhang 1996, J. Neurosci

Network converges to stereotyped, localized activity



Zhang 1996, J. Neurosci

Network converges to stereotyped, localized activity

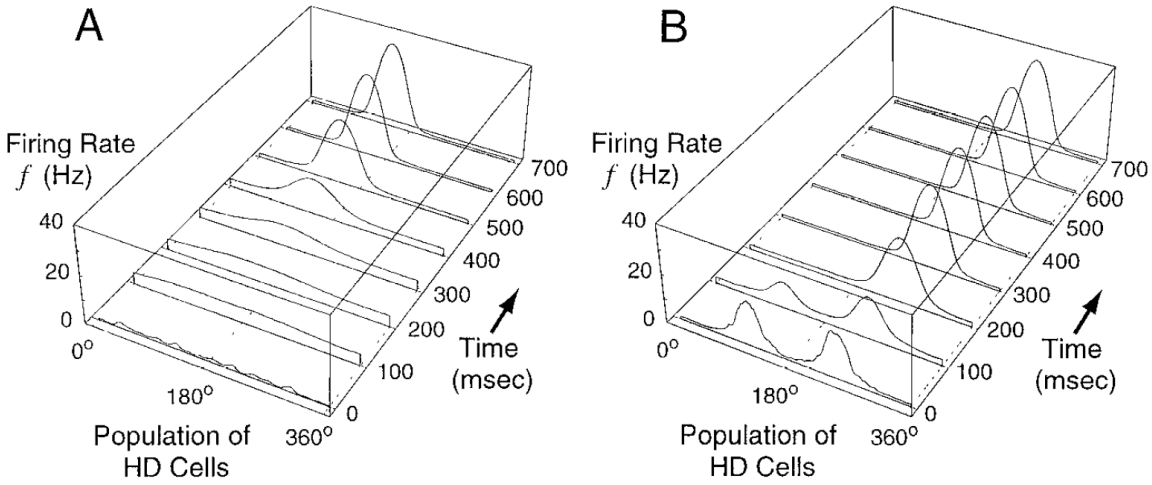


Figure 3. Snapshots of the time evolution of the model obtained by numerical integration of the continuous Equation 2, showing the emergence of the same stable firing profile from two arbitrary initial states. The final profile has stereotyped shape, but its peak potentially can be centered anywhere (neutral equilibrium). The HD cells are indexed by their preferred directions (ranging from 0° to 360°). Free parameter $\tau = 10$ msec. Other parameters are as in Figure 4 under the regularization $\lambda_0 = 10^{-3}$.

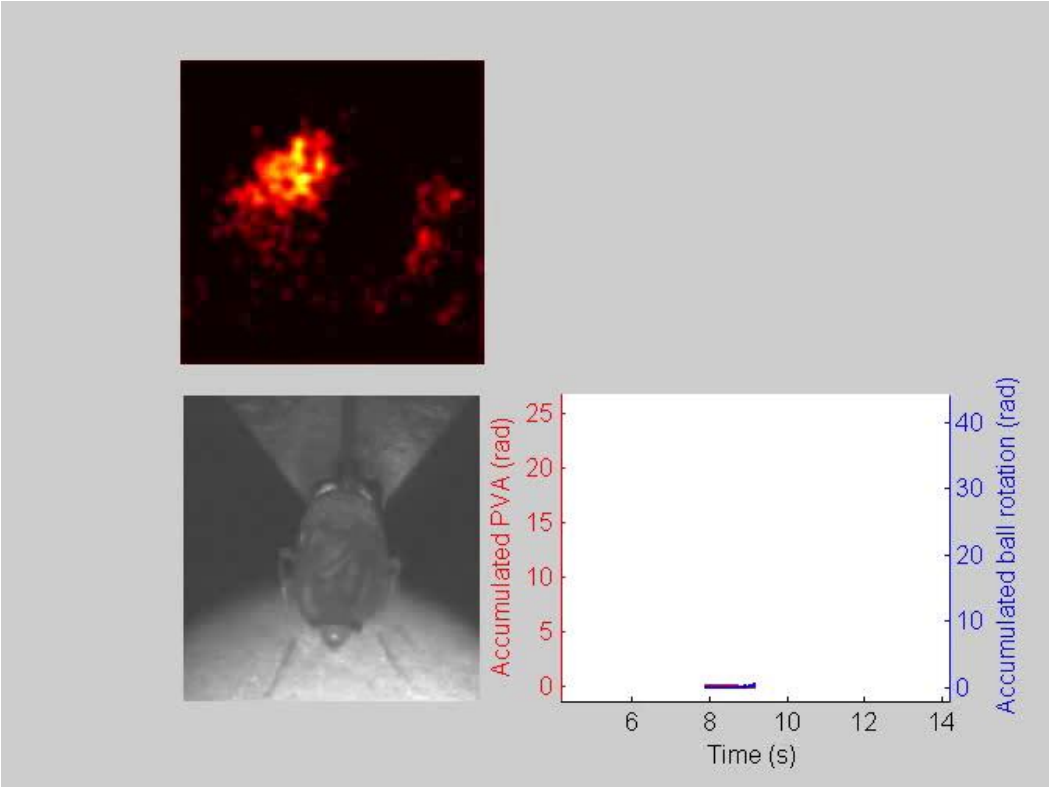
Persistent activity, independent of input (compare to HD activity in darkness)

Zhang 1996, J. Neurosci

How can we update the compass (head-direction)?

Based on self-motion cues?

Activity of the central complex in the absence of a visual cue (darkness)



Seelig & Jayaraman Nature 2013

Shifting activity in the ring attractor

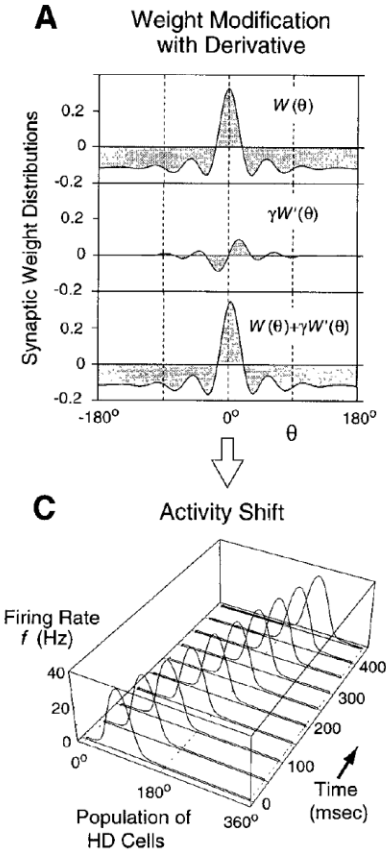
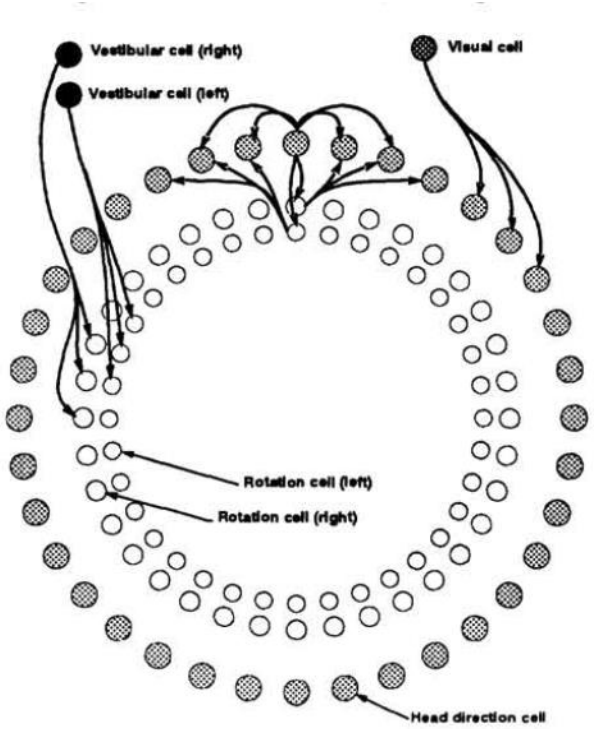


Figure 6. Dynamic activity shift occurs if and only if the synaptic weight distribution function has a nonzero odd component. Intuitively, the odd components in these examples serve to excite the right neighbors so as to move the activation toward the right, and at the same time inhibit the left neighbors so as to erase the trail. *A*, *C*, When the odd component is proportional to the derivative of the static even-weight distribution, the shift does not disturb the shape of the static activity profile. We chose $\gamma = -0.063$, which yields the speed $\gamma/\tau \simeq 360°$/sec. Other parameters including the even-weight distribution $W(\theta)$ are identical to those in Figure 3. *B*, *D*, When the odd component is sinusoidal, the traveling profile has a different shape. We chose $\alpha = 0.00201$ so that the averages of $|\alpha \sin \theta|$ and $|\gamma W'(\theta)|$ are equal.

Skagg's model for the head-direction system



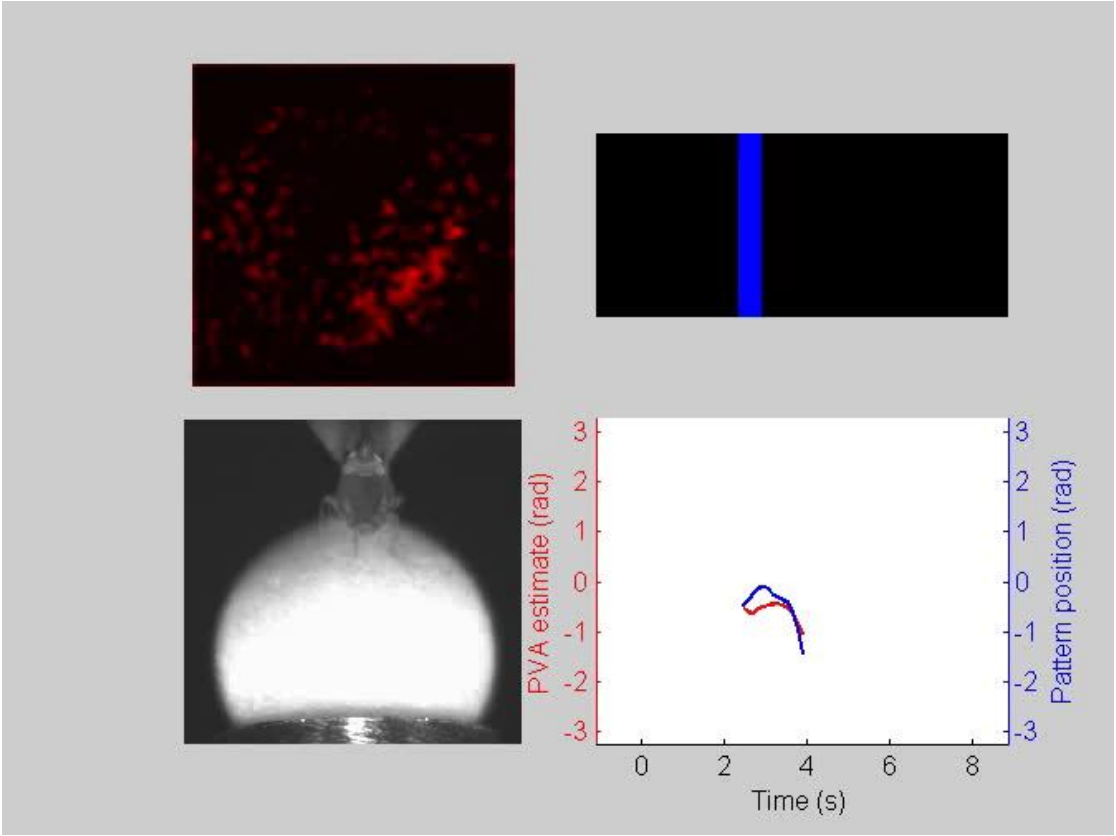
Note: asymmetric connectivity
from
the "shift-circuit"

Skaggs et al. 1995, NeurIPS

How can we update the compass (head-direction)?

Based on landmarks?

Compass-like representation of landmark orientation



Seelig & Jayaraman Nature 2013

Calibration with external view-dep. input

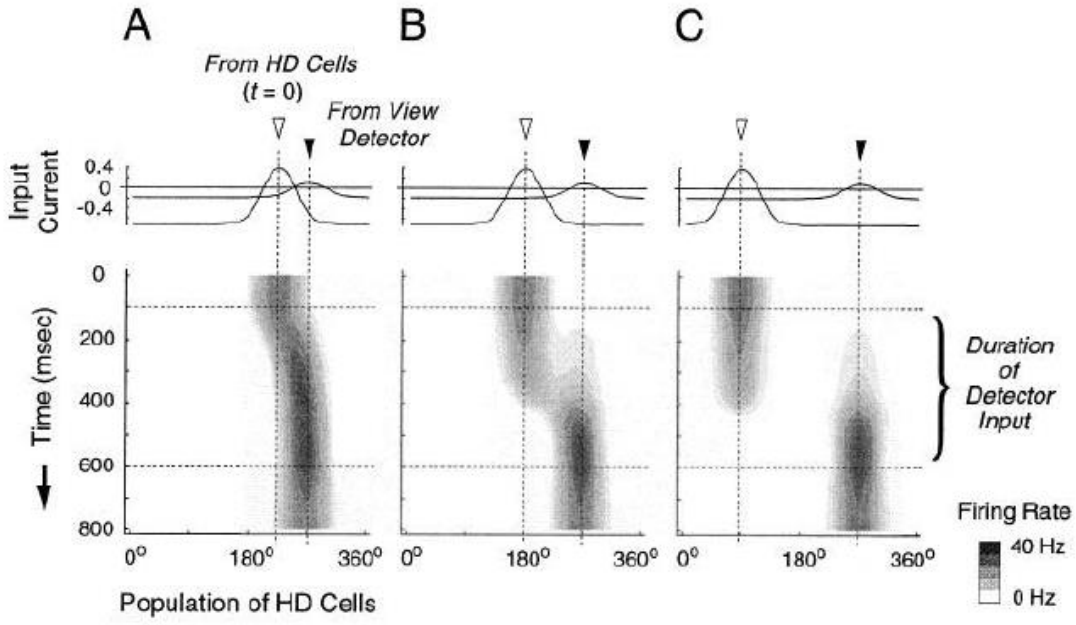


Figure 10. External input from a local-view detector calibrates the internal direction maintained by the HD cell network. In each of the three examples, the external input always starts at time 100 msec and holds constant for 0.5 sec. The distribution of the external input has the same shape as the distribution of the intrinsic input among the HD cells in the static state, but its magnitude is 25% as strong. Other parameters are as in Figure 3.

Zhang 1996, J. Neurosci

Take-home messages

- Discussion of Bayesian multisensory integration (*you will work on this in the problem set*)
- Attractor models are powerful models of brain function (and make several non-trivial predictions)
- Path integration is an important brain function
- We also highlighted classic results in rats & more recent ones in fruit flies
- Attractor models can implement path integration; we focused on ring attractors.
- Next week, you'll see that one can "learn" attractor-like models from normative goals