Supplementary Math Note
Principles governing the integration of landmark and self-motion cues in entorhinal cortical codes for navigation

Malcolm G. Campbell, Samuel A. Ocko, Caitlin S. Mallory, Isabel I.C. Low, Surya Ganguli, Lisa M. Giocomo

A One-dimensional model

One dimensional representation of an attractor network Consider a large population of neurons living on a one-dimensional neural sheet. For analytical simplicity, we consider the sheet to be continuous, so that position on the sheet is described by a continuous coordinate u, with the firing rate of a neuron at position u given by $s(u)$. Each neuron interacts with neighboring neurons through a translation-invariant connectivity, yielding the dynamics

$$\frac{ds\,(u)}{dt} = -s(u) + \mathcal{F}\left[\int_{u'} \mathsf{J}(u-u')s(u')\right]. \tag{1}$$

Here $\mathsf{J}(u-u')$ defines the synaptic weight from a cell at position u' to one at u, and $\mathcal{F}$ is a nonlinearity.

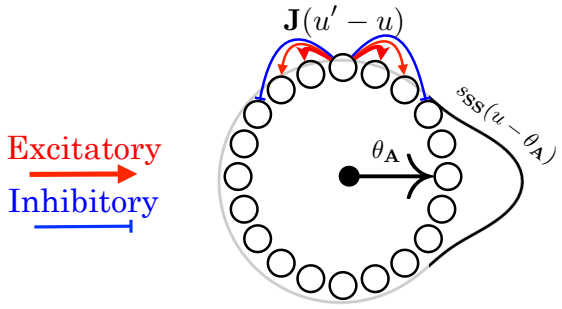

Figure 1: Schematic of a one-dimensional periodic neural sheet with short-range excitation (red arrows) and longer range inhibition (blue arrows). This yields a 1D family of bump-attractor states $s_{\mathbf{SS}}\,(u-\theta_{\mathsf{A}})$, which are mapped onto a single periodic variable θ_{A} representing the translation of the bump pattern.

Many appropriate choices of J and $\mathcal{F}$, corresponding for example to short range excitation and long range inhibition, will yield a family of stable, or steady state, localized bump activity patterns $s_{\mathbf{SS}}\,(u-\theta_{\mathsf{A}})$, parameterized by the position of their peak θ_{A}. Furthermore, for simplicity, we assume periodic boundary conditions on the neural sheet; both the coordinate u along the sheet, and the coordinate θ_{A} specifying a point on the manifold of stable attractor patterns, are angles defined modulo 2π (Fig. 1).

One-dimensional representation of path integration We can mathematically show (details in [1]) that in this representation, adding velocity-conjuctive cells to the network yields dynamics for path integration (Fig. 2a, b):

$$\frac{d\theta_{\mathsf{A}}}{dt} = \mathsf{v}\cdot\mathsf{k}_{\mathsf{A}}. \tag{2}$$

Here k_A is a constant of proportionality that relates animal velocity to the rate of phase advance in the attractor network (2π/Grid Spacing). Solving Eq. 2 allows us to recover path integration, where the resulting integrated attractor phase is *only* a function of the current position $x(t)$:

$$\theta_A(t) = \theta_A(x(t)) = k_A \cdot x(t). \tag{3}$$

The observed firing rate of a cell at neural sheet position u is, likewise, simply a function of animal position given by:

$$s(u, x) = s_{SS}(u - \theta_A(x)) = s_{SS}(u - k_A \cdot x). \tag{4}$$

Because s_{SS} has a neural sheet periodicity of 2π, the spatial firing pattern will have a periodicity of $2\pi/k_A$.

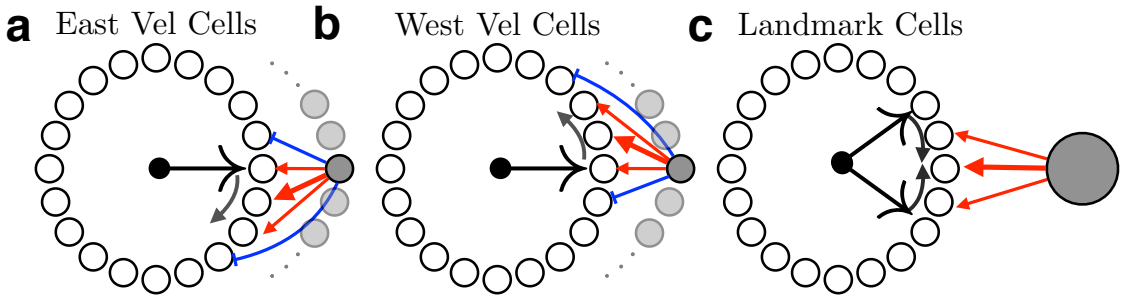

Figure 2: Path integration achieved with a 1D neural sheet and velocity-conjunctive cells. **a)** When the animal moves east, east-conjunctive cells with biased outgoing connections move the attractor pattern in the positive u direction. **b)** When the animal moves west, the attractor pattern is moved in the negative u direction. **c)** Schematic of a landmark cell correcting the attractor bump pattern. A single landmark cell will, given enough time, pull the peak of the bump pattern towards the strongest landmark-attractor synapse, regardless of the direction of the initial offset.

One Dimensional Representation of Landmark Cells We model landmark cells as sensory driven cells i whose firing rates are entirely a function of the animal's immediate position: $\text{Firing}_i(t) = H_i(x(t))$. We assume every landmark cell forms feed-forward connections to all cells in the attractor network with position dependent synaptic weights $W_i(u)$. This yields a complex coupled dynamics between neurons and synapses, where the distribution of attractor network activity patterns, or phases, drives plasticity in synapses from landmark cells to the attractor network (Fig. 2c). In turn, these synaptic weights modify the evolution of the attractor network phase. Despite this complexity, we will see that these dynamics reduce to a simple set of effective dynamics for short-term orientation:

$$\frac{d}{dt}\theta_A = \underbrace{k_A \cdot dx/dt}_{\text{Path Integration}} + \sum_i \underbrace{H_i(x(t))\text{Force}\left(\theta_L^i - \theta_A\right)}_{\text{Landmark Cells}}, \tag{5}$$

where all the synaptic weights of a landmark cell $W_i(u)$ can be represented by a single variable θ_L^i. We note that the dynamics for orientation of the attractor state to landmarks are *first-order*, in that the combination of path-integration and landmark input determine the *velocity* of position self-estimate, not the acceleration. This force function has the same qualitative form as $\sin\left(\theta_L^i - \theta_A\right)$, so we approximate it as such.
We next turn to understanding the outcome of this coupled dynamics and how it relates to mapping the environment.

Linear Track Consider an animal running along a linear track at constant running velocity, having a position $x = vt$. The animal has a 1D attractor network with state θ_A which evolves with a path integration gain of k_A. Without landmarks, the equation for the state of the attractor network is:

$$\frac{d\theta_A}{dt} = k_A v. \tag{6}$$

Consider a scenario in which landmark cells that have uniformly distributed landmark fields are added. After the landmark positions have been learned, each landmark cell has a learned state of $\theta_L^i = k_A(G=1)x_i$, where x_i is the firing field center of that landmark cell. Instead of considering individual landmark cells, it is simpler to consider *average* phase $\theta_L(x)$ of landmark cells firing at a particular location x. After learning, the landmark phases match up with the path integration of the animal, and $\theta_L(x) = k_L(G=1)x$, where $k_L(G=1) = k_A(G=1) = k_0$, the path-integration constant the animal had when it was learning the environment. We note that k_0 is the baseline spatial frequency (2π/Grid Spacing). Eq. 5 then becomes:

$$\begin{aligned}\frac{d\theta_A}{dt} &= k_A v + \omega \cdot \sin\left(\theta_L(x(t)) - \theta_A(t)\right) \\ &= k_A v + \omega \cdot \sin\left(k_L v t - \theta_A(t)\right).\end{aligned} \tag{7}$$

Gain Manipulation When the gain value is manipulated, the velocity in virtual reality (VR) space becomes Gv, $x(t) = Gvt$, and the attractor dynamics become:

$$\begin{aligned}\frac{d\theta_A}{dt} &= k_A v + \omega \cdot \sin\left(\theta_L(x(t)) - \theta_A(t)\right) \\ &= k_A v + \omega \cdot \sin\left(k_0 G v t - \theta_A(t)\right).\end{aligned} \tag{8}$$

This is equivalent to Eq. 7 where we make the scaling $k_L(G) = k_0 \cdot G$.

$$k_A(G)v + \omega \cdot \sin\left(k_L(G)vt - \theta_A(t)\right). \tag{9}$$

If self-motion input is purely *visual*, then $k_A(G)$ will also be proportional to gain, and there is no mismatch between $k_A(G), k_L(G)$; if self-motion input is purely *locomotor*, then $k_A(G)$ is independent of gain. The general relation is:

$$k_A(G) = k_0 \cdot (1 + \Delta G[1 - \%\text{Loc}(G)]) \tag{10}$$

Where $\%$Loc is the percent of velocity input that is locomotor, and $\Delta G = G - 1$. We examine the case where $k_A \neq k_L$.

Solving To solve Eq. 9, it is helpful to move to the rotating reference frame of the landmarks, to examine how the *difference* between the phase of the attractor network and the average phase of landmark cells which are firing, $\Delta\theta(t) = [\theta_A(t) - \theta_L(x(t))]$, evolves as:

$$\frac{d\Delta\theta}{dt} = \frac{d[\theta_A(t) - \theta_L(x(t))]}{dt} = (k_A - k_L)\, v + \omega \cdot \sin\left(\theta_L\left(x(t)\right) - \theta_A(t)\right) = (k_A - k_L)\, v - \omega \sin(\Delta\theta). \tag{11}$$

To understand which ratios are important, we make the above equations dimensionless by defining a dimensionless time:

$$\tau = t/\omega \tag{12}$$

giving dimensionless equations

$$\frac{d\Delta\theta}{d\tau} = \frac{(\mathsf{k_A} - \mathsf{k_L})\,\mathsf{v}}{\omega} - \sin(\Delta\theta) = \mathsf{D} - \sin(\Delta\theta) \tag{13}$$

where we have defined a dimensionless "Decoherence number"

$$\mathsf{D}(\mathsf{G}) = \frac{(\mathsf{k_A}(\mathsf{G}) - \mathsf{k_L}(\mathsf{G}))\,\mathsf{v}}{\omega} = -\mathsf{k}_0 \cdot \frac{\mathsf{v} \cdot \%\mathsf{Loc}(\mathsf{G}) \cdot \Delta\mathsf{G}}{\omega}. \tag{14}$$

The behavior of Eq. 13, and thus the measured firing rate maps, falls into three regimes defined by the decoherence number.

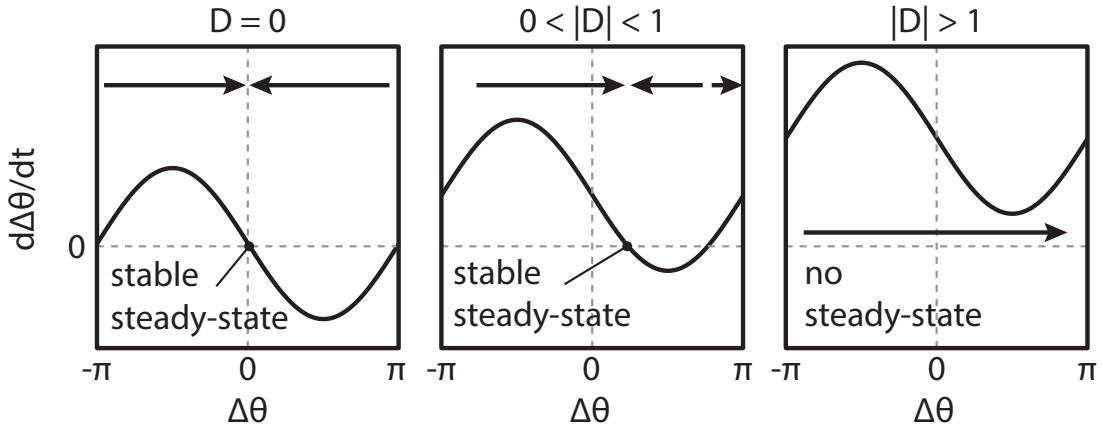

Figure 3: The three regimes of Eq. 13. In the first regime, where $\mathsf{D} = 0$, the steady state will be at $\Delta\theta = 0$. In the second, subcritical, regime, where $0 < |\mathsf{D}| < 1$, there will *still* be a steady state, but with a shifted phase $\Delta\theta$ having the same sign as D. In the third, critical regime, there is no steady state, and $\Delta\theta(t)$ instead reaches a steady cycle. Positive decoherence numbers are shown here; negative decoherence numbers will give rise to negative shifts of equal magnitude. See also Fig. 5 in the main text.

No gain manipulation (zero decoherence number) When the linear track has not been manipulated, the internal representations of path integration and landmarks are in register, i.e. $\mathsf{k_L} = \mathsf{k_A}$. This means $\mathsf{D} = 0$ (Fig. 3 Left). Therefore, $\Delta\theta = 0$, and the attractor network phase is *exactly* the phase of surrounding landmarks.

$$(\theta_{\mathsf{A}})_{\mathsf{G}=1}(\mathsf{x}) = \mathsf{k_L}\mathsf{x} \tag{15}$$

Small gain manipulation (small decoherence number) When $|\mathsf{D}| < 1$, Eq. 11 is solved with a constant shift between the the attractor network state and the landmarks, which saturates at a fixed value of $\Delta\theta(t) = \sin^{-1}(\mathsf{D})$ (Fig. 3 Middle). This gives us a gain-manipulated solution which is simply a phase-shifted version of the non gain-manipulated solution:

$$(\theta_{\mathsf{A}})_{\mathsf{GM}}(\mathsf{x}) = \mathsf{k}_0 \cdot \mathsf{x} + \sin^{-1}(\mathsf{D}), \tag{16}$$

where the phase shift depends on a combination of the animal speed, gain manipulation, and landmark strength. We can express this phase shift as a *spatial* shift in observed firing patterns as a function of virtual position x:

$$(\theta_{\mathsf{A}})_{\mathsf{GM}}(\mathsf{x}) = \mathsf{k}_0 \cdot \left[\mathsf{x} + \frac{\sin^{-1}(\mathsf{D})}{\mathsf{k}_0}\right] = (\theta_{\mathsf{A}})_{\mathsf{G}=1}(\mathsf{x} + \Delta\mathsf{x}), \tag{17}$$

where the spatial shift is given by:

$$\Delta\mathsf{x} = \frac{\sin^{-1}(\mathsf{D})}{\mathsf{k}_0}. \tag{18}$$

Large gain manipulation (large decoherence number) When $|\mathsf{D}| > 1$, there is no *constant* $\Delta\theta$ which can solve Eq. 11; instead, $\Delta\theta(t)$ follows a cyclic pattern (Fig. 3 Right). Despite this remapping, the effect of landmarks on the attractor network does not disappear entirely. This leads to two experimental signatures that we can see.

Cross-correlation due to uneven precession While all $\Delta\theta$ will be seen in the critical regime, they will not be observed at equal frequencies. For example, when $\mathsf{D} > 1$ (gain decrease condition), when the advancing attractor phase is behind the landmark phase, the pull from landmarks accelerates the relative precession rate. When the attractor phase is ahead of the landmark phase, the pull from landmarks slows the relative precession (Fig. 4a) This waxing and waning of the precession rate results in the network spending more time in an advanced attractor phase:

$$\mathsf{Prob}(\Delta\theta) \propto [d\Delta\theta/dt]^{-1} \propto [\mathsf{D} - \sin(\Delta\theta)]^{-1} \tag{19}$$

leading to a small peak in the cross-correlation (Main Figure 5g.)

Mean Precession Rate The new firing field spacing will be in between the original spacing and the spacing given by the new velocity input; this is because the effect of landmarks on the attractor network does not disappear entirely. Instead, the new number of fields will be a weighted average between the original number and the number given by the new gain condition:

$$\text{Fractional Change in Number of Fields} = \frac{\mathsf{k}_{\mathsf{A}} - \mathsf{k}_{\mathsf{L}}}{\mathsf{k}_{\mathsf{L}}} \cdot \mathsf{Precess}(\mathsf{D}), \tag{20}$$

where the mean procession rate $\mathsf{Precess}(\mathsf{D})$ is a function of the decoherence number that varies between 0 and 1 (Fig. 4b).

Proof of mean precession rate We can calculate the dimensionless rate of precession by calculating the change in *time* with respect to *precession*.

$$\frac{d\tau}{d\Delta\theta} = [\mathsf{D} - \sin(\Delta\theta)]^{-1} \Rightarrow \text{Precession Time} = \tau(\Delta\theta = 2\pi) - \tau(\Delta\theta = 0) = \int_{\Delta\phi=0}^{2\pi} [\mathsf{D} - \sin(\Delta\theta)]^{-1} \tag{21}$$

This gives a mean precession rate of

$$\text{Precession Rate} = \frac{2\pi}{\int_0^{2\pi} [\mathsf{D} - \sin(\Delta\theta)]^{-1}} = \frac{\mathsf{D} \cdot 2\pi}{\int_0^{2\pi} \left[1 - \frac{\sin(\Delta\theta)}{\mathsf{D}}\right]^{-1}} = \mathsf{D} \cdot \mathsf{Precess}(\mathsf{D}) \tag{22}$$

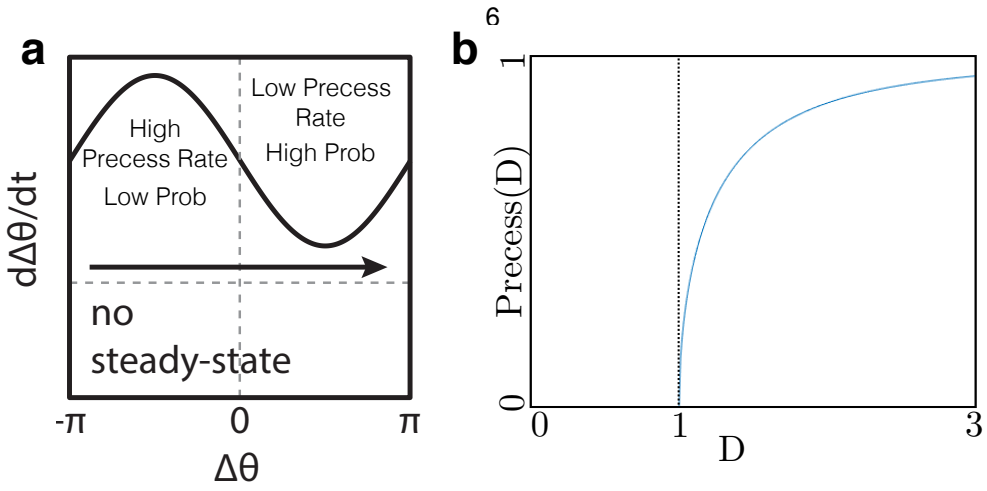

Figure 4: **a)** Schematic of waxing and waning of precession rates. At $\Delta\theta > 0$ the precession rate is slowed by the landmarks (High Probability of Observing), while at $\Delta\theta < 0$ the precession is accelerated (Low Probability of Observing). **b)** Plot of the precession number as a function of decoherence number. For $|\mathrm{D}| < 1$, $\mathrm{Precess}(\mathrm{D}) = 0$. In the limit of $\mathrm{D} \to 1^+$, the system is barely critical and precesses very slowly as $\mathrm{Precess}(\mathrm{D}) \to 0^+$. In the limit of $\mathrm{D} \to \infty$, $\mathrm{Precess}(\mathrm{D}) \to 1$. Note that the same behavior applies for negative decoherence numbers, as $\mathrm{Precess}(\mathrm{D}) = \mathrm{Precess}(-\mathrm{D})$.

Where we have defined the dimensionless "precession number" which is a function of the decoherence number

$$\mathrm{Precess}(\mathrm{D}) = \frac{2\pi}{\int_0^{2\pi}\left[1 - \frac{\sin(\Delta\theta)}{\mathrm{D}}\right]^{-1}}. \tag{23}$$

Moving back to our original units gives the average rate of phase advancement, which we can simplify

$$\left(\frac{\bar{d\theta}}{dt}\right) = \frac{d\theta_{\mathrm{L}}}{dt} + \frac{\bar{\Delta\theta}}{dt} = \mathrm{v}k_{\mathrm{L}} + \mathrm{D}\cdot\mathrm{Precess}(\mathrm{D})\cdot\omega = \mathrm{v}k_{\mathrm{L}} + \frac{(k_{\mathrm{A}} - k_{\mathrm{L}})\,\mathrm{v}}{\omega}\cdot\mathrm{Precess}(\mathrm{D})\cdot\omega = \tag{24}$$

$$\mathrm{v}\cdot[k_{\mathrm{L}}\,(1 - \mathrm{Precess}(\mathrm{D})) + k_{\mathrm{A}}\cdot\mathrm{Precess}(\mathrm{D})], \tag{25}$$

and thus the number of fields observed will be given by a *weighted average* of k_{L} and k_{A}, with the weighting given by the decoherence number:

$$\text{Fractional Change in Number of Fields} = \frac{k_{\mathrm{A}} - k_{\mathrm{L}}}{k_{\mathrm{L}}}\cdot\mathrm{Precess}(\mathrm{D}) \tag{26}$$

B Effect of grid spacing

Here, we consider the dependence of our results on grid spacing. In the model, grid spacing is inversely proportional to $k_{\mathrm{A}}(\mathrm{G} = 1)$, which we call k_0. The decoherence number can then be written as

$$\mathrm{D}(\mathrm{G}) = \frac{\mathrm{v}\cdot\Delta k}{\omega} = -\frac{\mathrm{v}\cdot k_0\cdot\%\mathrm{Loc}(\mathrm{G})\cdot\Delta\mathrm{G}}{\omega} \tag{27}$$

Where $\%$Loc is the percent of velocity input that is locomotor, and $\Delta\mathrm{G} = \mathrm{G} - 1$. Note that the decoherence

number will thus be inversely proportional to grid spacing ($2\pi/\mathrm{k}_0$). [1] One prediction of Eq. 27 is that the decoherence transition will occur at less extreme gain values for grid cells with small spacing, as $\mathrm{D} \propto \mathrm{k}_0 \cdot \Delta\mathrm{G}$ (Supplementary Fig. 15a).

Since we only recorded a small number of gain values, we cannot observe a grid-spacing-dependent decoherence transition. However, we are able to examine the effect of grid spacing on firing pattern shifts in the sub-critical regime in our dataset. Counterintuitively, the predicted shift in firing patterns at fixed gain will *not* have a strong dependence of the grid spacing. According to Eq. 16, the *phase* shift will be larger for small grid spacing:

$$\Delta\theta = \sin^{-1}(\mathrm{D}) = -\sin^{-1}\left(\mathrm{k}_0 \cdot \frac{\mathrm{v} \cdot \%\mathrm{Loc} \cdot \Delta\mathrm{G}}{\omega}\right) \underset{\sim}{\propto} \mathrm{k}_0 \tag{28}$$

Where the last step comes from the small-angle approximation $\sin^{-1}(\mathrm{D}) \approx \mathrm{D}$. However, the *spatial* shift is equal to the phase shift *times* the grid spacing, and thus will yield only a slight dependence on grid spacing (Supplementary Fig. 15b):

$$\Delta\mathrm{x} = \sin^{-1}(\mathrm{D})/\mathrm{k}_0 = -\frac{\sin^{-1}\left(\mathrm{k}_0 \cdot \frac{\mathrm{v}\cdot\%\mathrm{Loc}\cdot\Delta\mathrm{G}}{\omega}\right)}{\mathrm{k}_0} \approx \frac{\mathrm{v} \cdot \%\mathrm{Loc} \cdot \Delta\mathrm{G}}{\omega}. \tag{29}$$

This is what we observe in our data (Supplementary Fig. 15c-f).

C Generalization to Two-Dimensional Attractor Networks

For generalization to two dimensional attractor networks, we now have grid cells on a *two-dimensional* neural sheet (Fig. 5 a). The state of the attractor network is represented by a vector $\vec{\theta}_{\mathbf{A}}$. This vector is periodic, and unlike the 1D attractor state θ_{A}, it is periodic on a rhombus (Fig. 5 b), such that:

$$\vec{\theta}_{\mathbf{A}} \equiv \vec{\theta}_{\mathbf{A}} + (2\pi, 0) \equiv \vec{\theta}_{\mathbf{A}} + (\pi, \sqrt{3}\pi), \tag{30}$$

where the vector $(\pi, \sqrt{3}\pi) = 2\pi(\cos(60°), \sin(60°))$ corresponds to the $60°$ periodicity of the network. Likewise, the multiple peak synaptic strengths of each landmark cell can be represented as a two-dimensional vector $\vec{\theta}_{\mathbf{L}}$(Fig. 5 f1) which lives on the same periodic rhombus as $\vec{\theta}_{\mathbf{A}}$:

$$\vec{\theta}_{\mathbf{L}} \equiv \vec{\theta}_{\mathbf{L}} + (2\pi, 0) \equiv \vec{\theta}_{\mathbf{L}} + (\pi, \sqrt{3}\pi) \tag{31}$$

The full dynamics are:

$$\frac{d\vec{\theta}_{\mathbf{A}}}{dt} = \vec{\mathrm{k}}_{\mathrm{A}}\mathrm{v} + \omega \cdot \mathrm{Force}\left(\vec{\theta}_{\mathbf{A}}(t) - \vec{\theta}_{\mathbf{L}}(t)\right), \qquad \frac{d\vec{\theta}_{\mathbf{L}}}{dt} = \vec{\mathrm{k}}_{\mathrm{L}}\mathrm{v}. \tag{32}$$

After learning, the landmark phases match up with the path integration of the animal, and $\vec{\theta}_{\mathbf{A}}(\mathrm{x}) = \vec{\mathrm{k}}_{\mathrm{L}}(\mathrm{G}=1)\mathrm{x}$, where $\vec{\mathrm{k}}_{\mathrm{A}}(\mathrm{G}=1) = \vec{\mathrm{k}}_{\mathrm{L}}(\mathrm{G}=1) = \vec{\mathrm{k}}_0$, the path-integration constant the animal had when it was learning the environment. The magnitude of $\vec{\mathrm{k}}_0$ is inversely proportional to the grid spacing; its direction gives the angle of the observed slice through the 2D grid pattern [2]. The force law (Fig. 5 d)

[1] It is possible that ω and $\%$Loc are functions of k_0. Here we assume they are both held constant.

chosen is a truncated sin function:

$$\text{Force}(\vec{\Delta\theta}) = \begin{cases} -\sin\left(\left|\vec{\Delta\theta}\right|\right) \cdot \frac{\vec{\Delta\theta}}{\left|\vec{\Delta\theta}\right|} & \left|\vec{\Delta\theta}\right| < \pi \\ 0 & \left|\vec{\Delta\theta}\right| \geq \pi. \end{cases} \tag{33}$$

We can, likewise, get a set of dimensionless equations similar to in the 1D case

$$\frac{d\vec{\Delta\theta}}{d\tau} = \text{v}\frac{\left(\vec{\text{k}}_\text{A} - \vec{\text{k}}_\text{L}\right)\text{v}}{\omega} - \text{Force}(\vec{\Delta\theta}) = \vec{\text{D}} - \text{Force}(\vec{\Delta\theta}) \tag{34}$$

Where we have defined a dimensionless decoherence *vector*

$$\vec{\text{D}}(\text{G}) = \frac{\left(\vec{\text{k}}_\text{A}(\text{G}) - \vec{\text{k}}_\text{L}(\text{G})\right)\text{v}}{\omega}. \tag{35}$$

For convenience, we decompose this decoherence vector into a magnitude and direction, $\vec{\text{D}} = \text{D} \cdot \hat{\text{D}}$. We can likewise characterize this into three regimes, analogous to the three regimes of Eq. 13.

No gain manipulation (zero decoherence vector) When the linear track has not been manipulated, the internal representations of path integration and landmarks are in register, meaning $\vec{\text{D}} = 0$. Therefore, $\vec{\Delta\theta} = 0$, and the attractor network phase as function of position x is *exactly* the phase of surrounding landmarks.

$$\left(\vec{\theta}_\mathbf{A}\right)_{\text{G}=1}(\text{x}) = \vec{\text{k}}_\text{L}\text{x} \tag{36}$$

Small gain manipulation (small decoherence vector) When $|\vec{\text{D}}| < 1$, Eq. 34 is solved with a constant shift between the the attractor network state and the landmarks(Fig. 5 e1), which saturates at a fixed value of

$$\vec{\Delta\theta}(t) = \sin^{-1}(\text{D}) \cdot \hat{\text{D}} \tag{37}$$

This gives us a gain-manipulated solution which is simply a phase-shifted version of the non gain-manipulated solution:

$$\left(\vec{\theta}_\mathbf{A}\right)_{\text{GM}}(\text{x}) = \vec{\text{k}}_0 \cdot \text{x} + \sin^{-1}(\text{D}) \cdot \hat{\text{D}} \tag{38}$$

where the phase shift depends on a combination of the animal speed, gain manipulation, and landmark strength. Because the decoherence vector points in the same direction as $\vec{\text{k}}_\text{L}$, we can express this phase shift as a *spatial* shift in observed firing patterns as a function of virtual position x:

$$\left(\vec{\theta}_\mathbf{A}\right)_{\text{GM}}(\text{x}) = \vec{\text{k}}_0 \cdot \left[\text{x} + \frac{\sin^{-1}(\text{D})}{\left|\vec{\text{k}}_0\right|}\right] = (\theta_\text{A})_{\text{G}=1}(\text{x} + \Delta\text{x}) \tag{39}$$

Where the spatial shift is given by:

$$\Delta\text{x} = \frac{\sin^{-1}(\text{D})}{\left|\vec{\text{k}}_0\right|}. \tag{40}$$

Large gain manipulation (large decoherence vector) When $|\vec{\mathsf{D}}| > 1$, there is no constant $\vec{\Delta\theta}$ which can solve Eq. 34; instead, $\vec{\Delta\theta}(t)$ follows a complicated trajectory. Unlike the 1D case, the decohered regime has many sub-regimes. When $\vec{\mathsf{k}}_0$, and thus $\vec{\mathsf{D}}$ point in the $(1, 0)$ or the $(1/2, \sqrt{3}/2)$ directions, the behavior is effectively one-dimensional-this corresponds to the linear track being perfectly oriented along a particular axis of the animal's grid, and we recover the solution of Eq. 13, yielding rescaling of patterns (Fig. 5 f1).

However, when $\vec{\mathsf{k}}_0$, and thus $\vec{\mathsf{D}}$ are not entirely aligned with a periodic direction, we see a steady cycle that involves shifts in $\vec{\Delta\theta}$ *orthogonal* to the decoherence vector (Fig. 5 e2). The resulting slice through a 2D grid will have some combination of an orthogonal offset (Fig. 5 f2), a mean angular offset (Fig. 5 f3), and curvature (Fig. 5 f4). This will not only change the observed number of firing fields, but will also cause some fields to appear and disappear entirely, yielding not only rescaling but also *remapping*. We observed these effects in our data (Supplementary Fig. 5 and 6). There are other decohered sub-regimes which involve simultaneous precession in multiple directions; these yield the same effects of orthogonal shift, angular offset, and curvature of the 2D slice.

References

[1] Ocko S, Hardcastle K, Giocomo LM, Ganguli S (2018) Emergent elasticity in the neural code for space. *bioRxiv*.

[2] Yoon K, Lewallen S, Kinkhabwala AA, Tank DW, Fiete IR (2016) Grid cell responses in 1d environments assessed as slices through a 2d lattice. *Neuron* 89(5):1086–1099.

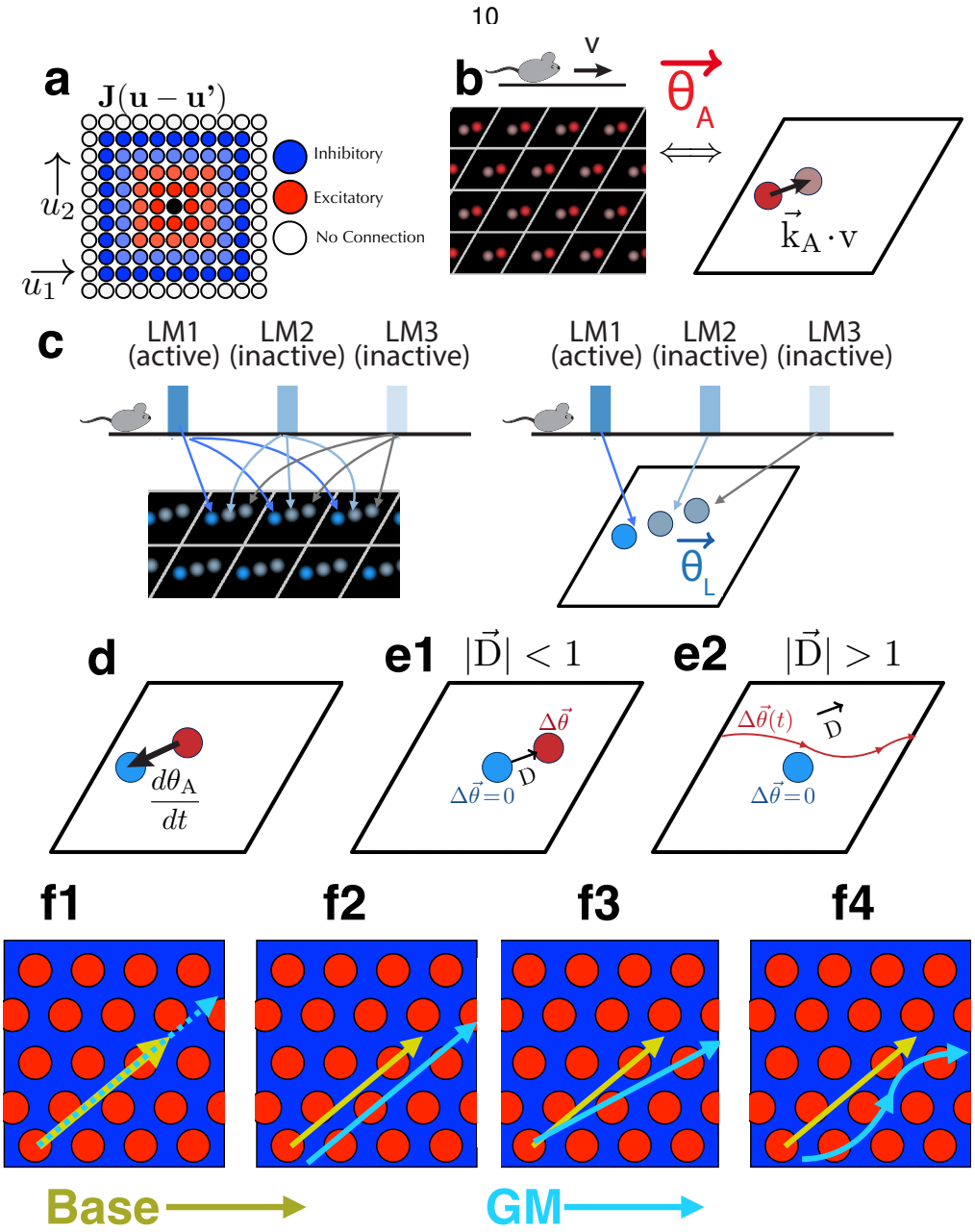

Figure 5: **a)** A 2D neural sheet with short-range excitation and long-range inhibition, analogous to Fig. 1. Each neuron on the continuous sheet now has coordinates $\mathbf{u} = (u_1, u_2)$. **b)** A 2D analogue of the reduced dimension attractor state representation of Fig. 1. A periodic multi-bump state (high firing rate in red) can be represented by a *two-dimensional* variable $\vec{\theta}_\mathbf{A}$ over the periodic rhombus. Lines drawn on top represent the "unit cell" to guide the eye. As the animal moves at a rate of v, the underlying attractor state moves at a rate of $\vec{k}_A \cdot v$. **c)** The landmark cell Hebbian weights have the same profile as 2D attractor states. As the animal travels along the track, the peak synaptic weight of firing landmark cells will also be translated. **d)** Analogously, there is a force law, where the state of an attractor network $\vec{\theta}_\mathbf{A}$ will be pulled towards the landmark state $\vec{\theta}_\mathbf{L}$. **e1)** Schematic of one of many decohered regimes. When $|\vec{D}| < 1$, $\Delta\vec{\theta}_\mathbf{A}$ approaches a steady state, yielding a constant shift in observed firing fields. **e2)** When $|\vec{D}| > 1$, there is no steady state, and $\Delta\vec{\theta}_\mathbf{A}$ will precess around the rhombus. The decoherence vector can be decomposed into two components-one parallel to the periodicity of the rhombus, and another which is orthogonal. There will be precession in the parallel direction (horizontal), and a net (but fluctuating) shift in the orthogonal (vertical) direction. **f1)** If a decohered $\vec{\theta}_\mathbf{A}$ took the same path through the 2D grid as it did in the original gain condition, this would result in a simple rescaling of the observed firing pattern. However, because $\Delta\vec{\theta}_\mathbf{A}$ evolves in 2D, it will not in general precess around $\Delta\vec{\theta}_\mathbf{A} = 0$. The dynamics in general will yield a slice through a 2D grid with some combination of orthogonal offset **(f2)**, mean angular offset **(f3)**, and curvature **(f4)**. Each of these effects will yield new and dropped fields, yielding not only rescaling but also *remapping*.