
Characterization of grid-cell phase dynamics during conflict between path integration and sensory landmark cues

Jingxuan Fan

Neuroscience Graduate Program
Harvard University
Cambridge, MA 02138
jfan@g.harvard.edu

Wenyi Zhang

Neuroscience Graduate Program
Harvard University
Cambridge, MA 02138
wenyi_zhang@g.harvard.edu

Albert Chen

Biophysics Graduate Program
Harvard University
Cambridge, MA 02138
albertchen@g.harvard.edu

Abstract

To navigate natural environments, animals leverage both idiothetic path integration and allocentric sensory cues to estimate their spatial position. Previous experimental work has shown that introducing a mismatch between an animal’s locomotion and visual landmarks can lead to either coherent phase shift or complete grid-cell remapping in a context-dependent manner[1,2]. The activity of grid cells is well-captured by a ring attractor network, the dynamics of which can be reduced to its phase dynamics[3]. Here, we verified and extended the analytical results of [1] via numerical simulations. We showed that under mismatches between continuous visual inputs and animal velocity, the phase dynamics of 1D and 2D attractor models exhibit a phase transition between two regimes that correspond to complete and coherent remapping. We also modified the visual inputs such that they are consistent with the experimental setup in [2]. We observed a similar phase transition only when visual landmarks occurred more frequently than an empirically measured threshold. Moreover, we directly simulated the dynamics of the full ring attractor model to recapitulate the process of learning a consistent mapping between visual landmarks and internal representation. Finally, we demonstrated that this learned network architecture exhibits a similar phase transition behavior as analytically predicted in [1] under the scenarios of velocity-visual cue mismatch.

19

1 Introduction

20 Navigation is critical to goal-directed behaviors and relies on information distributed across multiple
21 neural populations in the brain. In order to estimate spatial position accurately, an animal’s internal
22 representation of its environment needs to be consistent with the external world. During navigation,
23 animals combine their self-motion cues, which are shaped by proprioception and motion feedback,
24 with sensory inputs, such as visual and olfactory cues, to construct neural representations of their
25 environments.

26 In the mammalian brain, grid cells in the medial entorhinal cortex (MEC), which have characteristic
27 periodic firing fields, are thought to encode spatial information that supports navigation. Recent

works have explored how grid cells in mice integrate self-motion and visual landmark information. Ocko, Hardcastle, *et al.* (2018) [3] proposed a 1D attractor model in which a grid-cell network is coupled to a velocity-conjunctive attractor that supports path integration. The grid-cell network also receives sensory information from landmark cells with synaptic weights learned via Hebbian plasticity. This neural attractor has the topology of a ring and shares features with other ring attractor models that have been used to model grid cells and other cell types with periodic firing patterns, such as head-direction cells[4-6]. [3] further demonstrated that the dynamics of their attractor network's activity can be reduced to its phase dynamics. This reduced model predicts that within the regime of small mismatch between internal representation and visual landmarks, the phase difference between internal and visual representation stabilizes into a constant phase shift; within the regime of large mismatch, the phase difference enters a periodic steady cycle. These predictions were also recently corroborated by experimental evidence. Using a virtual reality-based navigation task on a linear track, Campbell, Ocko, *et al.* (2018) [1] and Campbell, Attinger, *et al.* (2020) [2] showed that MEC grid cells remap coherently under small velocity-visual cue conflicts while their firing patterns degenerate under large conflicts.

In this investigation, we confirmed the results of the analytic calculations in [1] using numerical simulations and further explored the phase and neural dynamics of the network with more complex visual inputs. We started with the reduced model derived in [3] that summarizes the 1D ring attractor network's activity by its phase. Following [1], we first implemented the reduced model with continuous, constant visual input, which is akin to optic flow information. We then simulated the phase dynamics with only discrete visual landmarks that mimic the experimental setup presented in [2] and extended the reduced model to 2D scenarios. Finally, we returned to the neural dynamics presented in [3] and examined how a ring attractor network learns sensory cues from landmark cells with Hebbian plasticity of synaptic weights, and moreover, how the learned model network behaves following velocity-visual cue mismatch, again using numerical simulations.

1.1 Linear path integration task

We first will describe the general experimental setup that our simulation studies aim to recapitulate. Particularly, we followed the setup in [1] and [2]: mice navigated a linear track in virtual reality for a water reward at the end of the track in a closed-loop experiment. The VR track was unidirectional and 400 cm in length. In [1], the walls of the linear track are patterned and therefore the mice receive continuous optic flow as they navigate. In [2], visual landmark "towers" were placed at every 80 cm; thus, there were 5 equally spaced towers in total. In between landmarks, the walls of the linear track were of a uniform gray color. When a mouse reached the end of the VR track, it was teleported back to the beginning.

On a subset of trials, conflicts between self-motion and visual cues were introduced by manipulating the gain of the transformation from the rotation of the running ball to the mouse's progression on the VR track. On a gain increase trial, for example, a mouse would encounter a landmark tower sooner than it would otherwise expect, whereas on a gain decrease trial it would encounter a landmark later than expected.

2 Numerical simulation of 1D ring attractor phase dynamics

We begin by considering the 1D ring attractor model grid-cell activity presented in [3].

$$\begin{aligned}
 \frac{ds(u)}{dt} = & -\frac{s(u)}{\tau_m} + \mathcal{G}\left(\int_{u'} J(u-u')s(u')\right) \\
 & + \epsilon_{PI} s_{EC}(u - \Delta\phi_{PI}) + \epsilon_{PI} s_{WC}(u + \Delta\phi_{PI}) + \epsilon_{LM} \sum_i W_i(u) s_i^L(t).
 \end{aligned} \quad (1)$$

On the right hand side, the first two terms represent the intrinsic dynamics of the network. The third and fourth terms represent inputs from east and west velocity-conjunctive cells, respectively, while the last term represents inputs from landmark cells. Ocko and colleagues showed that under strong and stable intrinsic dynamics, we can reduce the description of the neural dynamics to a one-dimensional variable θ_A , which represents the phase of the network, by treating all other inputs

74 as a small perturbation. Thus, we have that

$$\frac{d\theta_A}{dt} = k_A v + \omega \sum_i H_i(x(t)) \mathcal{F}(\theta_L^i - \theta_A). \quad (2)$$

75 Here, the first term contains the contributions from velocity-conjunctive cells, where path integration
76 gain k_A relates the animal velocity v to the rate of internal phase advance. The second term contains
77 contributions from landmark cells indexed by i and modulated by ω , where $H_i(x)$ is the firing field
78 of the i^{th} cell and \mathcal{F} is a force function that pulls the internal phase θ_A towards the landmark phase θ_L^i
79 of the i^{th} cell. It can be shown that \mathcal{F} is the negative derivative of the spatial autocorrelation function
80 of the stable (cosine) bump profile. When there is continuous and spatially uniform visual input, we
81 have that

$$\frac{d\theta_A}{dt} = k_A v + \omega \sin(\theta_L(x(t)) - \theta_A). \quad (3)$$

82 2.1 Continuous visual input

83 To better understand the phase dynamics of the attractor network, we numerically simulated Eq.(3)
84 under various gain factors and reproduced the analytical behavior described in [1] (Figure 1). When
85 the gain is manipulated, the rate of landmark phase advance deviates from that of internal phase
86 advance such that Eq.(3) becomes

$$\frac{d\theta_A}{dt} = k_A v + \omega \sin(k_L v t - \theta_A) \quad (4)$$

$$= k_A v + \omega \sin(k_0 G v t - \theta_A). \quad (5)$$

87 When the gain G is equal to one, we have that $k_A = k_L = k_0$, which is inversely proportional to the
88 grid spacing. To characterize the behavior of the network, a dimensionless decoherence number was
89 introduced in [1], given by

$$D = \frac{v(k_A - k_L)}{\omega}. \quad (6)$$

90 [1] describes 3 different regimes of the network dynamics determined solely by the decoherence num-
91 ber D . Here we present three example figures for each regime. When there is no gain manipulation,
92 i.e. $G = 1$, we have that $D = 0$ and the internal and landmark representations are in phase (Figure
93 1a). In the subcritical regime where $0 < |D| < 1$, the landmark cues pull the internal phase θ_A
94 towards the landmark phase θ_L that is consistent with the external environment and therefore the phase
95 difference between θ_A and θ_L reaches a constant (Figure 1b). The subcritical regime corresponds to
96 gain manipulations under which grid cells coherently remap. On the other hand, in the supercritical
97 regime where $|D| > 1$, the internal phase θ_A becomes untethered from the landmark phase θ_L and
98 the phase difference $\Delta\theta$ enters a steady cycle (Figure 1c). In this regime of gain manipulations, the
99 spatial firing patterns of the grid cells degenerate, leading to complete remapping.

100 To better characterize the behavior of phase difference across the three regimes, we varied D across
101 $-4 < D < 4$ with 100 sampling points and simulated the phase dynamics (Figure 2). We used
102 two parameters to describe the time progression of internal versus landmark phase difference: 1)
103 major frequency of the periodic steady cycle, which we defined as the frequency with the highest
104 power peak in fast Fourier transform analysis and 2) steady phase shift value $\Delta\theta$, which we defined
105 as the average over the 100 time steps towards the end of the simulation. Steady phase shift value
106 $\Delta\theta$ was only calculated for the cases that could achieve a stable dynamics. We found that within
107 the $-1 < D < 1$ regime (shaded by grey), the phase difference was able to reach steady state as
108 predicted and the steady state value scale roughly linearly with D within the range $-\pi$ to π . We also
109 observed that within the $|D| > 1$ regime, the major frequency of the phase difference scale almost
110 linearly with $|D|$ with some exceptions close to the transition boundary $|D| = 1$ (Figure 2 left).

111 2.2 Discrete visual input

112 We next interrogated whether the attractor network exhibited similar behaviors when sensory informa-
113 tion is no longer available at every moment in time. Although continuous visual input enables clean
114 analytical solutions, visual information may not always be readily accessible in real environment.
115 In [2], for instance, the animal encountered a salient landmark—a visual "tower"—only once every
116 80 cm. In between landmarks, there is minimal visual cues that inform the animal about its current
117 location.

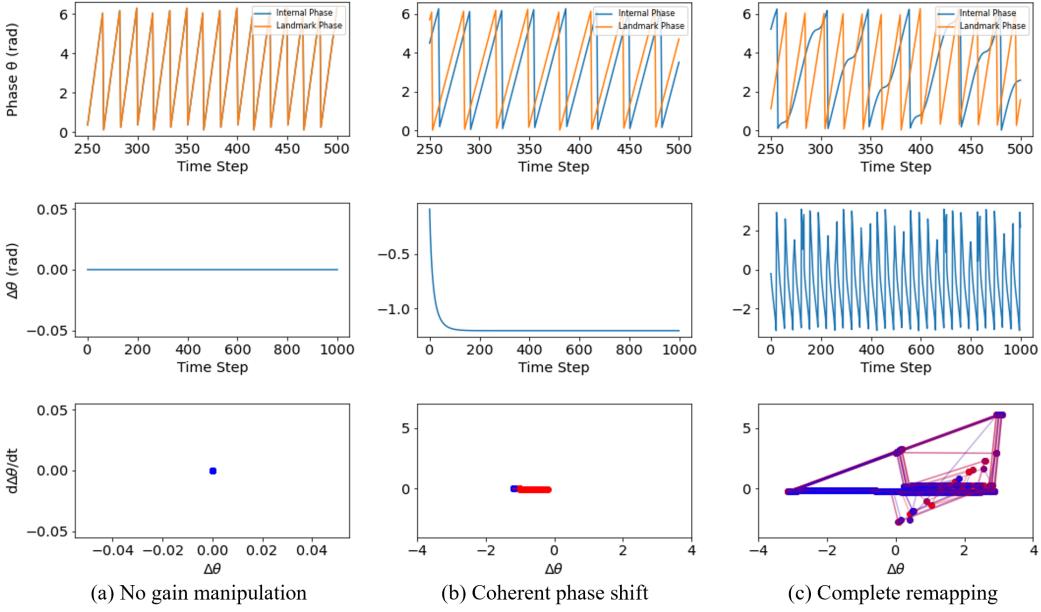


Figure 1: Phase dynamics with continuous landmark input. (a) Phase dynamics when there is no gain manipulation, i.e. decoherence number $D = 0$. Top: Internal phase θ_A and landmark phase θ_L across time. Middle: Phase difference $\Delta\theta$ across time. Bottom: Phase space of $\Delta\theta$. Red indicates the beginning of the simulation while blue the end of the simulation. (b) Similar to (a), but for $D = 0.94$. (c) Similar to (a), but for $D = 2.4$.

Motivated by the experiments of [2], we simulated the network phase dynamics with five discrete visual inputs equally spaced along the 400-cm track (Figure 3). Using Eq.(2), we modeled the firing field H of each landmark cells as a boxcar function centered on the corresponding landmark position x_{LM} . When the animal position x was within a certain radius of x_{LM} , H took on the value of one; everywhere else, H was zero. For simplicity, we set the radius of the firing field for all landmark cells to be 20 cm.

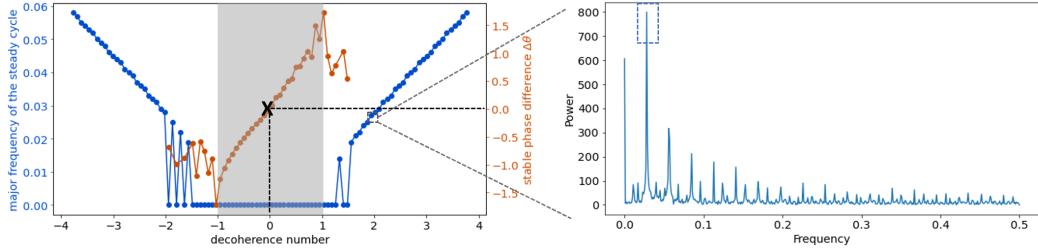


Figure 2: Characterization of time progression of phase difference under continuous optic flow visual input. Left: Time progression of phase difference characterized by major frequency and steady phase difference $\Delta\theta$ over 3 regimes of decoherence. Right: An example of fast Fourier transform analysis performed for phase difference within the $|D| > 1$ regime. The major frequency is labeled by a blue square.

The behavior of the attractor network with discrete visual inputs is analogous to the behavior under continuous visual input. We found that when there was no gain manipulation, the internal and landmark phases were essentially identical (Figure 3a). When the gain was set to 1.5, the landmarks anchored the internal phase: The phase shift $\Delta\theta$ oscillated within a narrow range about a nonzero mean (Figure 3b). The small oscillations reflect the visual landmarks pulling the internal phase θ_A towards the landmark phase θ_L^i . This is similar to the subcritical regime in the continuous-input scenario in which grid cells coherently remap. When the gain was increased to 1.7, the phase shift $\Delta\theta$

131 showed not only small oscillations but also large variations, suggesting that the internal phase breaks away from the landmark frame of reference (Figure 3c). Note that the decoherence number alone is

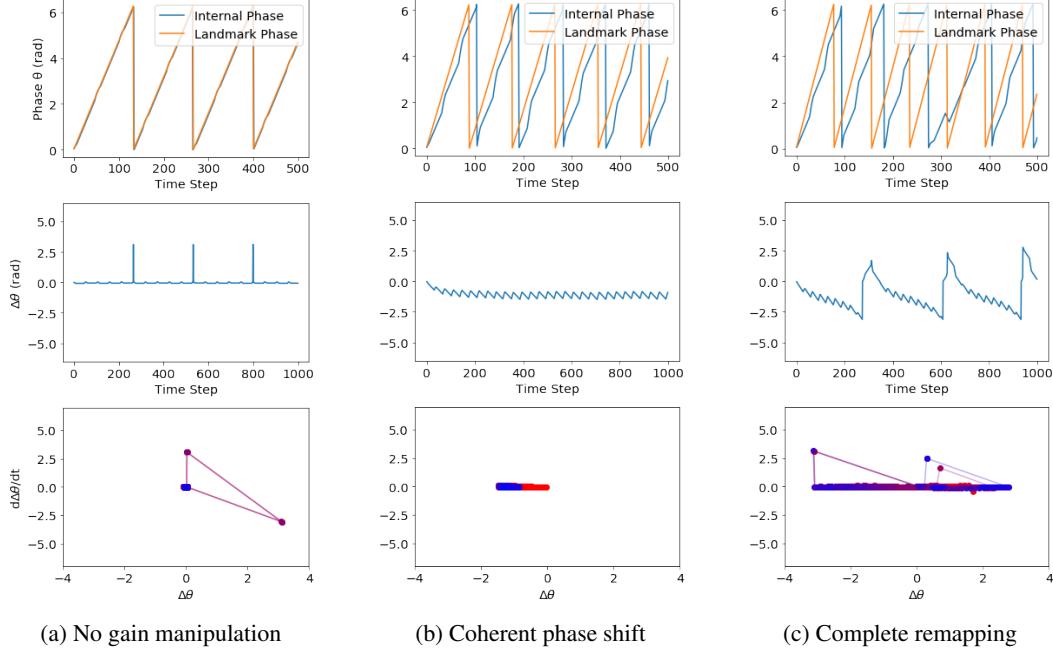


Figure 3: Phase dynamics with discrete landmark input. (a) Phase dynamics when there is no gain manipulation, i.e. $G = 1$. Top: Internal phase θ_A and landmark phase θ_L across time. Middle: Phase difference $\Delta\theta$ across time. Large deviations from zero are likely due to phase switching from 2π from 0 at periodic boundary. Bottom: Phase space of $\Delta\theta$. Red indicates the beginning of the simulation while blue the end of the simulation. (b) Similar to (a), but for $G = 1.5$. (c) Similar to (a), but for $G = 1.7$.

132 insufficient in predicting which regime the attractor network is in when the visual inputs are discrete.
 133 The decoherence transition depends also on the radius of the landmark cell firing fields as well as
 134 the number of landmarks per unit length. If we fixed the parameters of the network ($D = -0.424$,
 135 Number of landmarks = 5) and varied the radius of the landmark cell firing field, we found that the
 136 attractor network underwent a phase transition when the radius was at around 18.1 cm (Figure 4a-b).
 137 On the other hand, when we fixed the radius of the firing field to be at 10 cm and varied the number
 138 of landmarks (again $D = -0.424$), the transition between the two regimes occurred when there were
 139 9 landmarks placed along the 400-cm track, meaning the inter-landmark distance was around 45 cm
 140 (Figure 4c-d). Together, these results suggest that there can be coherent remapping only if there are
 141 enough landmark cells and their firing fields are wide enough such that they approximately tile the
 142 linear track.
 143

144 3 Numerical simulation of 2D attractor phase dynamics

145 Based on the framework from [1], here we extended our attractor network to two dimensions with
 146 grid cells on a neural sheet. In this case, the attractor phase $\vec{\theta}_A$ is periodic on a rhombus, and satisfies

$$\vec{\theta}_A = \vec{\theta}_A + (2\pi, 0) = \vec{\theta}_A + (\pi, \sqrt{3}\pi). \quad (7)$$

147 The full dynamics is similar to the one-dimension scenario, as shown in Eq.(8) and Eq.(9). The phase
 148 behavior could be similarly sub-divided into three regimes defined by the norm of the decoherence
 149 factor, with $|\vec{D}| \leq 1$ generating coherent phase shift, and with $|\vec{D}| > 1$ generating complete
 150 remapping (Figure 5). Note that here we only consider continuous optic flow visual input.

$$\frac{d\vec{\theta}_A}{dt} = \vec{k}_A v + \omega \cdot \text{Force}(\vec{\theta}_A(t) - \vec{\theta}_L(t)). \quad (8)$$

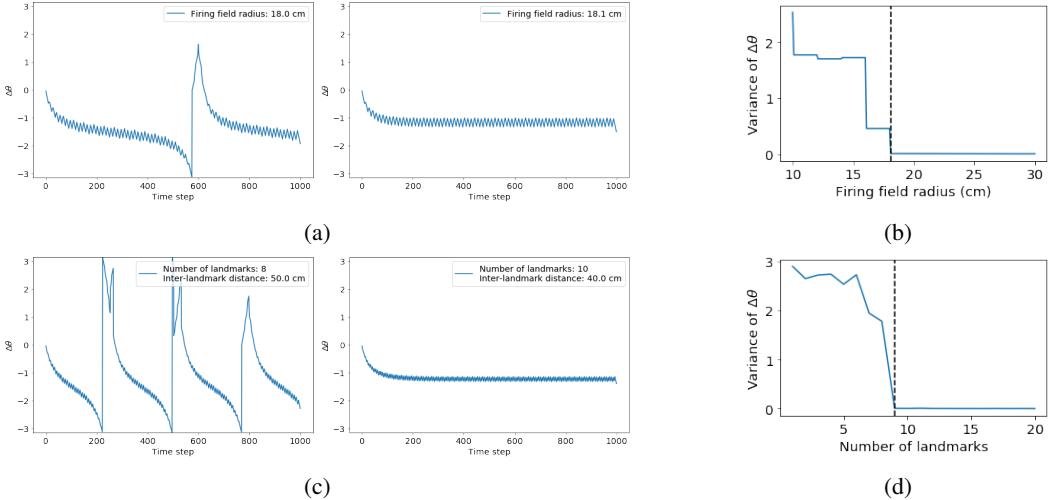


Figure 4: Decoherence transition depends on both firing field radius and inter-landmark distance. For this set of simulations, $G = 2$ and $D = -0.424$. (a) Example plots of phase shift $\Delta\theta$ across time showing complete (left) and coherent (right) remapping for two different firing field radii. (b) Variance of phase shift $\Delta\theta$ for different firing field radii. Vertical broken line indicates critical radius at which attractor network transitions from the supercritical regime to the subcritical regime. (c) Similar to (a) but for different number of landmarks. (d) Variance of phase shift $\Delta\theta$ for different numbers of landmark. Vertical broken line indicates minimum number of landmarks needed for coherent remapping.

151

$$\text{Force}(\vec{\Delta\theta}) = \begin{cases} -\sin |\vec{\Delta\theta}| \cdot \frac{\vec{\theta}_A}{|\vec{\theta}_A|} & \text{if } |\vec{\theta}_A| < \pi \\ 0 & \text{if } |\vec{\theta}_A| \geq \pi. \end{cases} \quad (9)$$

152 Parallel to the 1D case, here we also examined the major frequencies (second-largest frequencies
153 from Fourier transform) and steady state differences as $|\vec{D}|$ progresses (Figure 6). The general trend
154 is quite similar to the 1D scenario (Figure 2), except that the dynamics at $|\vec{D}| \geq 1$ is generally
155 less periodic, and the stability of the frequency progression is correlated with the velocity direction.
156 When velocity is along the $(\sqrt{3}, 1)$, the dynamics seems to achieve the most stable progression of
157 periodicity. This is possibly due to an even scaling of velocity input along the two grid-cell directions,
158 and in effect identical to the 1-dimensional case.

159 4 Numerical simulation of neural dynamics

160 After we have numerically simulated the phase dynamics of the reduced model in 1D and 2D
161 scenarios, we explored the behavior of the full ring attractor model. To set up a ring attractor network
162 with velocity integration, we adapted and modified the model used in [3] (Eq.(10)). The ring network
163 has recurrent connections that are locally excitatory and long-range inhibitory. It receives input from
164 non-velocity conjunctive cells and velocity-conjunctive cells of west and east directions. The activity
165 of velocity-conjunctive cells are shifted versions of the non-velocity conjunctive cells and they are
166 coupled to the ring network by a small perturbation factor ϵ_{PI} . According to the derivation in [3], the
167 small perturbation strength ensures that the network can linearly integrate the animal velocity.

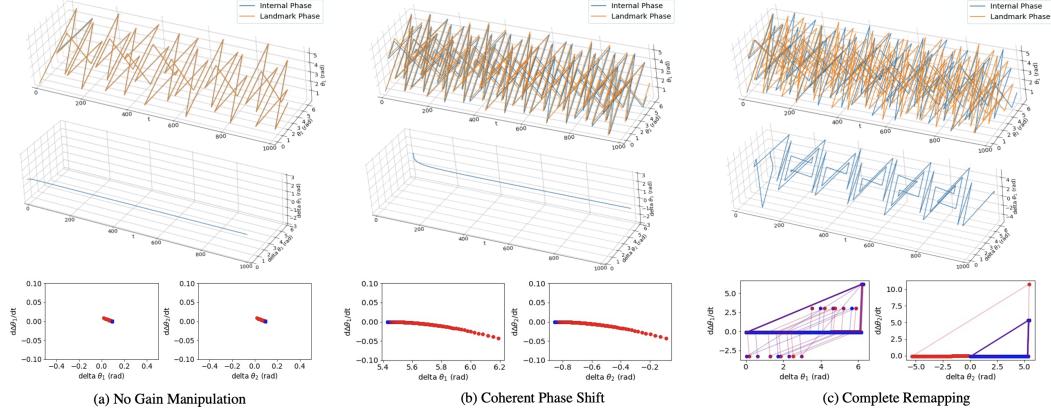


Figure 5: 2D dynamics with continuous visual input. (a) Progression of phase dynamics when $|\vec{D}| = 0$. (b) Phase dynamics when $|\vec{D}| = 0.86$. (c) Phase dynamics when $|\vec{D}| = 1.33$.

$$\begin{aligned} \tau_m \frac{ds(u)}{dt} = & -s(u) + \mathcal{G} \left(\int_{u'} J(u - u') s(u') + b \right. \\ & \left. + \epsilon_{PI} \int_{u'} J(u - u') s_{EC}(u') + \epsilon_{PI} \int_{u'} J(u - u') s_{WC}(u') \right) \quad (10) \end{aligned}$$

168 Simulating using the above setup, we observed that the network integrates velocity: the bump
169 amplitude stabilizes and the bump phase moves at a certain angular velocity that scales linearly with
170 the animal velocity (Figure 7a). The scaling factor depends linearly on the outbound phase shift of
the velocity-conjunctive cell activity (Figure 7b).

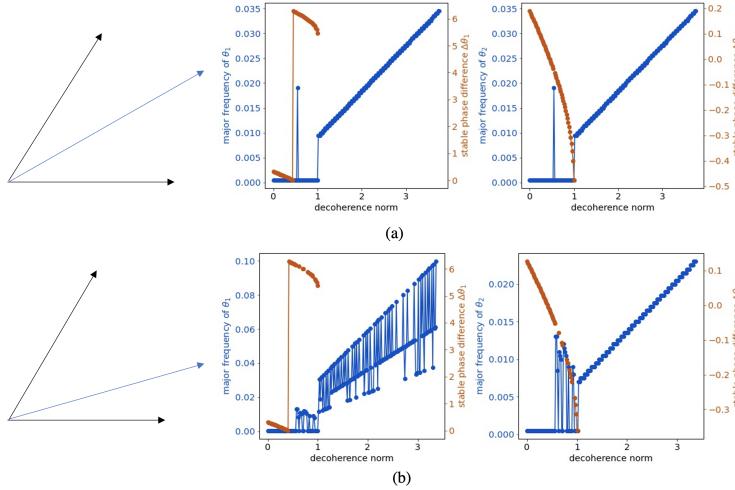


Figure 6: Characterization of 2D phase dynamics under continuous optic flow visual input. Major frequency and steady phase difference $\Delta\theta$ of the phase dynamics vary with the norm of decoherence factor \vec{D} . (a) Schematics of the velocity direction (blue) $\vec{v} = (\sqrt{3}, 1)$ in grid space (Left). Major frequency and steady phase difference of the first dimension θ_1 (Middle) and of the second dimension θ_2 (Right). (b) Schematics of the velocity direction (blue) $\vec{v} = (3, 1)$ in grid space (Left). Major frequency and steady phase difference of θ_1 (Middle) and θ_2 (Right).

171
172 We next took into consideration the visual landmark inputs. To mimic the training process in which the
173 animals learn a consistent mapping between the ring neurons and the visual neurons. We incorporated
174 into the model a group of visual neurons that are coupled to the main ring through feed-forward

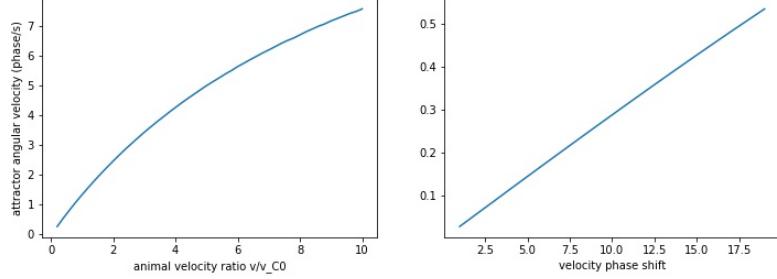


Figure 7: Factors that affect the velocity of internal dynamics. Left: Attractor angular velocity increases with animal velocity. Right: Attractor angular velocity scales linearly with the phase shift of velocity conjunctive cells.

175 excitatory weights (Eq.(11)). These weights undergo Hebbian plasticity with weight decay over the
176 time scale of one training session (Eq.(12)).

$$\begin{aligned} \tau_m \frac{ds(u)}{dt} = & -s(u) + \mathcal{G} \left(\int_{u'} J(u - u') s(u') + b \right. \\ & \left. + \epsilon_{\text{PI}} \int_{u'} J(u - u') s_{\text{EC}}(u') + \epsilon_{\text{PI}} \int_{u'} J(u - u') s_{\text{WC}}(u') + \epsilon_{\text{LM}} \sum_i W_i(u) s_i^{\text{L}}(t) \right) \quad (11) \end{aligned}$$

$$\frac{dW_i(u)}{dT} = \frac{\int_t s(u, t) s_i^{\text{L}}(t)}{\int_t s_i^{\text{L}}(t)} - W_i(u). \quad (12)$$

177 We demonstrated that after 10 sessions of training, the network can learn a set of feed-forward weights
that consistently map the phase from visual neurons to ring neurons (Figure 8a).

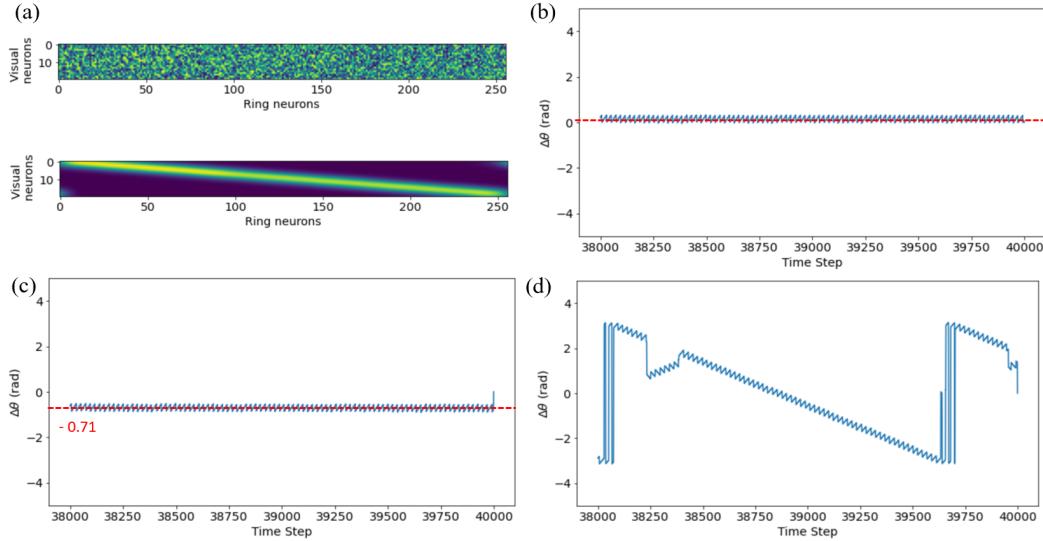


Figure 8: Characterization of visual input learning and dynamics under velocity-visual cue mismatch in full ring attractor network. (a) Ring attractor network learns a consistent visual landmark-to-internal phase mapping after training. (b), (c), and (d) Example behaviors of the learned network undergoing velocity-visual flow mismatch manipulation. (b) indicates no mismatch; (c) indicates small mismatch; and (d) indicates large mismatch.

178

179 After training, we fixed the visual-to-ring neuron weights to mimic the behavior over one testing
180 session. By fixing the visual flow and varying the animal velocity, we could create a mismatch
181 between the internal phase and the visual landmark phase. We demonstrate that the network's
182 behavior towards the end of the session approaches our previous simulation of the reduced model
183 with discrete landmark inputs. Specifically, when the animal velocity remains the same, the difference
184 between internal and landmark phase approaches zero (Figure 8b). When the animal velocity was
185 manipulated to be differed from the visual flow in a small range, the difference between internal
186 and landmark phase approached a constant number between $-\pi$ and π . Figure 8c gives an example
187 of such manipulation. When the animal velocity is manipulated to be differed from the visual flow
188 substantially, the difference between internal and landmark phase appears to be a periodic steady
189 cycle as shown by the example in Figure 8d.

190 However, it is noticeable that in this case the mismatch between the animal velocity and the visual
191 flow cannot be easily described by a single dimensionless decoherence number defined by [1]. The
192 empirically leaned weight matrix between visual and ring neurons renders the ratio between locomotor
193 and visual input strength variable. Therefore, we only provide some examples of the behaviors that
194 the network can achieve under different mismatch regimes without providing a definitive boundary.
195 Further analysis is needed to fully characterize the phase transition behaviors of this network.

196 References

- 197 [1] Campbell, M.G., Ocko, S.A., Mallory, C.S., Low, I.I.C., Ganguli, S., and Giocomo, L.M.. (2018) Principles
198 governing the integration of landmark and self-motion cues in entorhinal cortical codes for navigation. *Nat.*
199 *Neurosci* **21**(8):1096-1106.
- 200 [2] Campbell, M.G., Attinger, A., Ocko, S.A., Ganguli, S., and Giocomo, L.M. (2020) Distinct algorithms for
201 combining landmarks and path integration in medial entorhinal, visual and retrosplenial cortex. bioRxiv doi:
202 10.1101/2020.10.05.327106.
- 203 [3] Ocko, S.A., Hardcastle, K., Giocomo, L.M., and Ganguli, S. (2018) Emergent elasticity in the neural code
204 for space. *Proc. Natl. Acad. Sci. U.S.A.* **115**(50):E11798-E11806.
- 205 [4] Skaggs, W.E., Knierim, J.J., Kudrimoti, H.S., and McNaughton, B.L. (1994) A model of the neural basis of
206 the rat's sense of direction. In *Advances in Neural Information Processing Systems* 7.
- 207 [5] Samsonovich, A. and McNaughton, B.L. (1997) *J. Neurosci.* **17**(15):5900-5920.
- 208 [6] Xie, X., Hahnloser R.H.R., and Seung, H.S. (2002) Double-ring network model of a head-direction system.
209 *Phys. Rev. E Stat. Nonlin. Soft Matter Phys.* **66**:041902.