From the fly connectome to exact ring attractor dynamics

Tirthabir Biswas[1, 2, a, *] Angel Stanoev[2, b, *] Sandro Romani[2, c, †] and James E. Fitzgerald[1, 2, 3, 4, 5, d, †]

[1] *Department of Neurobiology, Northwestern University, Evanston, IL 60208, USA.*

[2] *Janelia Research Campus, Howard Hughes Medical Institute, Ashburn, VA 20147, USA.*

[3] *Department of Physics and Astronomy, Northwestern University, Evanston, IL 60208, USA.*

[4] *Department of Engineering Sciences and Applied Mathematics, Northwestern University, Evanston, IL 60208, USA.*

[5] *NSF-Simons National Institute for Theory and Mathematics in Biology, Chicago, IL 60611, USA.*

(Dated: October 30, 2024)

Abstract

A cognitive compass enabling spatial navigation requires neural representation of heading direction (HD), yet the neural circuit architecture enabling this representation remains unclear. While various network models have been proposed to explain HD systems, these models rely on simplified circuit architectures that are incompatible with empirical observations from connectomes. Here we construct a novel network model for the fruit fly HD system that satisfies both connectome-derived architectural constraints and the functional requirement of continuous heading representation. We characterize an ensemble of continuous attractor networks where compass neurons providing local mutual excitation are coupled to inhibitory neurons. We discover a new mechanism where continuous heading representation emerges from combining symmetric and anti-symmetric activity patterns. Our analysis reveals three distinct realizations of these networks that all match observed compass neuron activity but differ in their predictions for inhibitory neuron activation patterns. Further, we found that deviations from these realizations can be compensated by cell-type-specific rescaling of synaptic weights, which could be potentially achieved through neuromodulation. This framework can be extended to incorporate the complete fly central complex connectome and could reveal principles of neural circuits representing other continuous quantities, such as spatial location, across insects and vertebrates.

I. Introduction

Animals navigating through their environment must maintain an accurate sense of direction, requiring their brains to represent heading direction (HD) through neural activity patterns [1, 2]. However, the precise neural circuit mechanisms enabling this representation remain unclear. Several theoretical models based on ring attractor networks [3–6], a type of continuous attractor network, have been proposed to explain HD systems [7–16], stimulating the search for identifying such networks and dynamics in the brain. Core feature of ring attractors is a localized pattern, a bell-shaped bump of activity within a ring-like neural structure, which is stably-maintained through recurrent excitations between similarly tuned neurons and distal inhibitions, and can be readily moved to different locations around the ring. Such a bump of activity is thus suitable for encoding circular variables, and it has been used to characterize mammalian HD networks whose neurons exhibit reliable orientation tuning and whose population activity is limited to a ring-like manifold [17] and updated by self-motion [1, 18]. Lack of knowledge and data about the connectivity patterns within these networks, however, prohibit the understanding of the relationship between their structure and function. Most of the HD models, therefore, assume idealized architectures with large number of neurons [19, 20] and perfect circular symmetry to qualitatively describe the HD dynamics, that however, do not fully capture the biological complexity and multi-population interactions revealed by connectomic data. A recently proposed such model was also able to achieve continuous angular representations with only a small number of neurons [21].

To bridge this gap between theory and biology, we set out to study the HD of the insect brain where there is connectivity and cell type identity data with engineering access for functional imaging and manipulability of specific neural populations, and theory that can interface both of these facets. Recent work in the fruit fly, *Drosophila*

*: shared contribution
†: joint correspondence

[a] Electronic address: tirthabir.biswas@northwestern.edu
[b] Electronic address: stanoeva@janelia.hhmi.org
[c] Electronic address: romanis@janelia.hhmi.org
[d] Electronic address: james.fitzgerald@northwestern.edu

melanogaster, has uncovered its HD network in the central complex (CX) of the brain, that exhibits ring attractor-like dynamics and is capable of angular velocity integration [22–25]. Key compass neurons in this system are the excitatory EPG neurons located in a toroidal structure of the CX, the ellipsoid body [22, 26] (Figs. 1a and 1b). Their population is comprised of less than 50 compass neurons, organized in 8 computational units anatomically arranged in a circle (Fig. 1a) [26]. This seemingly simple system nevertheless manages to solve a difficult task – experiments show that these compass neurons can encode heading angles with a precision much higher than the resolution limit of 45°, [21], suggested by their 8-unit architecture. This suggests an existence of a novel encoding strategy that emerges from the interaction between excitatory compass neurons and inhibitory populations.

The recent availability of detailed connectomic data [27–30] from the fly central complex inspired us to find biologically-compatible network models that satisfy both architectural constraints and functional requirements, see [31] for a recent study applying connectomic constraint to fly visual system. We therefore developed a framework to identify the theoretical conditions that comprise a ring attractor in general, which can then be applied in concert with any connectomic data to identify the neural mechanisms that enable continuous orientation encoding in the fly brain. To this end, we studied the interactions between the EPG neurons and its hypothesized inhibitory population of $\Delta 7$ neurons in the protocerebral bridge [32], shown to also encode and propagate the HD bump of activity [26, 33, 34]. We discovered two distinct classes of models that can generally arise from interactions between multiple populations of neurons. Networks with all active inhibitory neurons simplify to a fully-symmetric effective configuration, whereas regimes with inactive inhibitory neurons generate asymmetric feedback to the EPG neurons and give rise to a *mirror-symmetric* effective system, that nevertheless can support continuous heading representation. In the following sections we summarize our approach for linking the theoretical ring attractor ensemble framework with the fly connectome.

II. Towards constructing a continuous ring attractor consistent with connectomic data

The angular location inferred from the vector averaged activity (Fig. 1c) tracks the heading direction of a moving fly to within a few degrees. To model the biology of the heading direction system we will consider a threshold-linear recurrent network with eight compass neurons or computational units (Fig. 1d) that each correspond to activity patterns observed in distinct neuronal compartments (Figs. 1e and 1f, also see Methods). In light of the experimental observation that approximately half of the neurons are typically active (Fig. 1b), here we will focus on configurations with only four active neurons (Fig. 1e), although most of our analysis can be generalized to other configurations. The choice of a threshold-linear input-output function ensures that once the input drives to the neurons are negative, they can remain inactive and generate activity patterns (Fig. 1f) that closely resemble what is observed. Finally, since the bump-like activity can be sustained for a few seconds even in the dark, we imposed that the steady-state activity profiles be self-sustainable.

Our first goal will be to find the ensemble of all two-population networks that can self-sustain a continuum of steady state activity profiles smoothly spanning the entire 360 degrees of angular locations. Taking inspiration from the approximate 8-fold circular symmetry observed in the connectomic data involving the excitatory compass neurons and the secondary population of inhibitory neurons (Fig. 1g), we will consider *effective* networks involving just the compass neurons that can be mapped from a two population network with exact 8-fold circular symmetry. For instance, a class of effective networks we will consider will obey 8-fold circular symmetry as well as left-right symmetry leading to only 4 independent effective weights as depicted by four different colors in Fig. 1i. The synapse strengths between the computational units in Fig. 1i should be thought of as representing *effective* connectivity which not only include all the individual synapses between neurons in the two compartments of compass neurons, but also indirect connections via neurons in the inhibitory population (Fig. 1h). For a given active set of neurons, requiring a continuum of steady-state solutions with differing average angular locations will impose two relationships among the weight parameters. Further imposing that the configurations be stable, have a single peak and that the inactive neurons receive (net) negative input drives for consistency will constrain the networks through inequalities. However, as we shall see, this will still leave us with a two dimensional space of *continuous self-sustaining ring attractor networks*, Appendix B provides the mathematical details.

Somewhat surprisingly we will find out that due to the nonlinearity in the input-output function it also becomes possible to map synapse counts which obey 8-fold circular symmetry to effective networks whose weights are neither symmetric nor circularly symmetric, but only preserve a residual symmetry. However, we will be able to apply the methods we just discussed to identify continuous ring attractor networks with these residual symmetries as well,

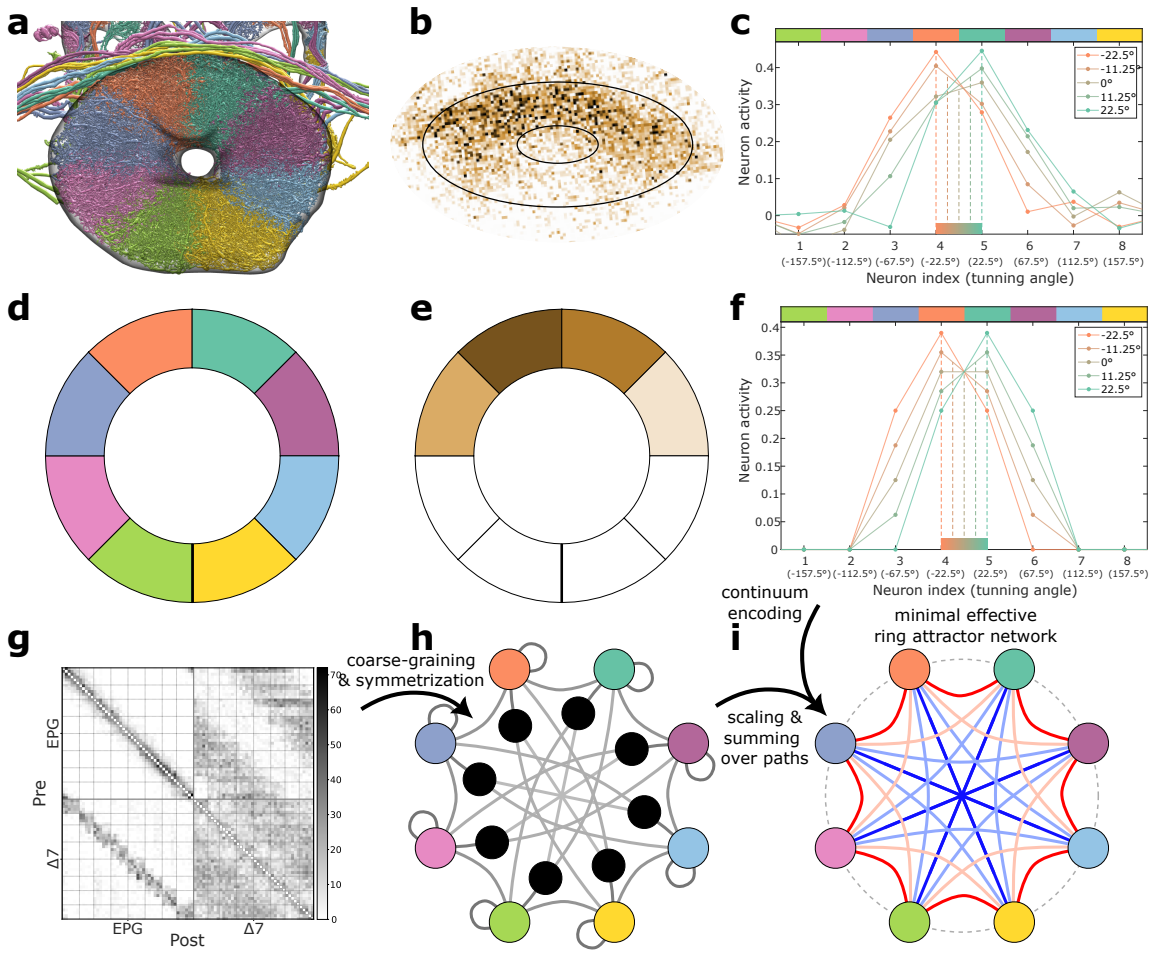

FIG. 1: **Constructing effective ring attractors from connectomic data.** (a) EPG neurons from the hemibrain dataset in the ellipsoid body (EB). Neuron processes are color-coded according to the EB wedges that they arborize. (b) Example calcium activity imaging data of the EPG neuron population in the EB. (c) 5 example activity profiles (solid lines) whose mean values (dashed lines) are distributed uniformly across the interval between the 4th and 5th wedges. (d) and (e) Schematics of the neural units' identities and activities, respectively, according to the color codes in (a) and (b). (f) Example activity profiles with mean values matching the ones in (c), sampled from a continuum of solutions generated by a ring attractor model. (g) Connectivity matrix between the EPG and Δ7 neurons from the hemibrain dataset. Individual estimated wedges are separated by thin gray lines. (h) Schematic of a coarse-grained two-population model of a readout population (EPG, color-coded as in (d)) and a secondary population (black). (i) Schematic of an effective single-population ring attractor model derived by aggregating the two-population model (h) and parameterized to satisfy the continuum encoding conditions.

the details are contained in Appendix C.

Our next goal will be to investigate whether one can find connectome-constrained ring attractors that are consistent with the observed fly connectome. Since the EPG-Δ7 network has been previously implicated in the encoding of heading direction, here we will specifically consider the EPG-Δ7 connectome (Fig. 1h) coarse-grained and circularly symmetrized (Fig. 1g) to align it with our modeling assumptions. To convert the synapse counts to synaptic weights we will assume that the former is multiplied with four different scaling parameters, depending upon the pre and post-synaptic partners, EPG → EPG, EPG → Δ7, Δ7 → EPG, and Δ7 → Δ7. We thus have a four dimensional parameter space of *realizable* networks consistent with the connectome. Next, we will derive the effective networks that one obtains from the EPG-Δ7 network by summing over all relevant paths (including via Δ7 neurons) between the EPG compartments. We can now ascertain whether we can find a realizable effective network that is also a continuous ring attractor (Fig. 1i). In other words, are there scaling parameters such that the weights of the effective network that we obtain from the connectome, obey all the equality and inequalities we theoretically derived.

III. Ensembles of continuous self-sustaining ring attractors

Our goal in this section is to provide the underlying biological principles and theoretical insights that led us to the ensemble of self-sustaining ring attractors that can uniquely encode the continuum of 360^0 angles, and can be obtained from two-population networks preserving 8-fold circular symmetry. The first step in determining our ensemble is to reduce the two population model into an *effective* single population model of compass neurons as outlined above and detailed in Appendix D. What is novel about this procedure is that one not only obtains effective networks which are fully symmetric, *i.e.* , preserves 8-fold circular symmetry and left-right symmetry, but also *mirror-symmetric* networks which only preserve part of these symmetries. This occurs when some of the inhibitory neurons are inactive as is possible in a threshold-linear network. In this case, for every path from one compass neuron to another compass neuron, the reverse path may not be available because some of the inhibitory neurons in the path may have become inactive. This makes the effective weights asymmetric and thereby reduces the symmetries of the network. We emphasize that the nonlinear input-output transfer function is a key aspect of our model that ensures half of the compass neurons can remain inactive in any given bump profile. Fig. 2a, depicts a network with 8 compass neurons, $\{y_i,\ i = 1, \ldots, 8\}$, where the neurons, y_3, y_4, y_5, y_6, are active. It illustrates how the effective connection between y_4 and y_5 sums over different pathways that can either lead to symmetric or mirror-symmetric networks.

In Appendix E, we also show that without any loss of generality we can assume that the self couplings are absent in the effective network that makes the analysis technically simpler. Accordingly, as shown in Figs. 2b and 2c, there are three and six independent weight-parameters that parametrizes all effective synaptic weights between the four active neurons in the symmetric and mirror-symmetric networks, respectively.

Now, the requirement that the compass neurons can self-sustain a bump of activity through their recurrent connections essentially means that the steady state activities must be able to equal the total input drives the neurons receive from all the other neurons in a self consistent manner. Mathematically this means that if $\bar{W}$ represents the effective weight-matrix for the active neurons (Figs. 2b and 2c), then $\bar{W}$ must be degenerate, *i.e.* , have an eigenvector (to be identified with the neuronal activity profile) with eigenvalue one. Given the symmetries of our model we can claim (Appendix F) that if there is only one such linearly independent eigenvector, then the activity profile must be symmetric; one can, of course, scale the amplitude of the activity (Fig. 2d), but such a scenario would lead to only eight discrete average angular locations, not a continuum (Fig. 2e). What if an asymmetric eigenvector with eigenvalue one did exist? For the symmetric model, it is easy to see why this would imply the existence of another eigenvector which is a mirror copy of the first leading to a *doubly degenerate* weight matrix, $\bar{W}$ (Fig. 2f). In this case though, one *would* have a two dimensional eigenspace of steady state solutions, not only whose amplitudes can vary, but also the average angular locations across a continuum, precisely what we required (Fig. 2g). In Appendix C we show that similar arguments follow for the mirror-symmetric weight matrix.

There is a different way to interpret the double degeneracy condition. In a linear theory one can check that a single eigenvector with eigenvalue one could either be symmetric or anti-symmetric, see Appendix F, but the latter is not a valid activity profile in a threshold-linear network because activities cannot be negative. The doubly degenerate case, however, can be interpreted as having a symmetric ($\vec{y}_s$) and an anti-symmetric ($\vec{y}_a$) eigenvector with eigenvalues one, the requirements restricting the weights to reside on the scarlet and violet hypersurfaces respectively, in the weight space as depicted for the symmetric model in Fig. 2h. The doubly degenerate ensemble thus lives along their intersection indicated by the brown and grey curves. As illustrated in Fig. 2i, the anti-symmetric eigenvector can add to the components of a symmetric eigenvector on one side while cancel on the other side producing an asymmetric profile. The amount of asymmetry will depend on the relative strengths along the two eigenvectors and determine the average angular location, while their overall strength controls the amplitude of the neuronal activity, see Fig. 2j. It turns out that for the symmetric model, the continuum of angles for a given active set precisely spans 45 degrees as needed to uniquely encode the entire angular range by virtue of the 8-fold symmetry. However, for the mirror-symmetric model this range could be smaller or larger than 45 degrees. The unique encoding requirement imposes an additional condition on the weight parameters. Thus while the symmetric model is characterized by a one-parameter (three weight-parameters minus two eigenvalue conditions) family of ring attractors, the mirror symmetric model is three dimensional (six weight-parameters minus two eigenvalue and one uniqueness condition).

Although the networks we have found can support a continuum of steady states, unless these activity profiles are stable, we will not have a viable ring-attractor. Now, W has four eigenvalues, two of which are one by construction. Perturbative stability (*i.e.* stable under small fluctuations around steady states) thus requires the two other eigenvalues to be less than one, leading to decaying dynamical modes, see Appendices B and C.

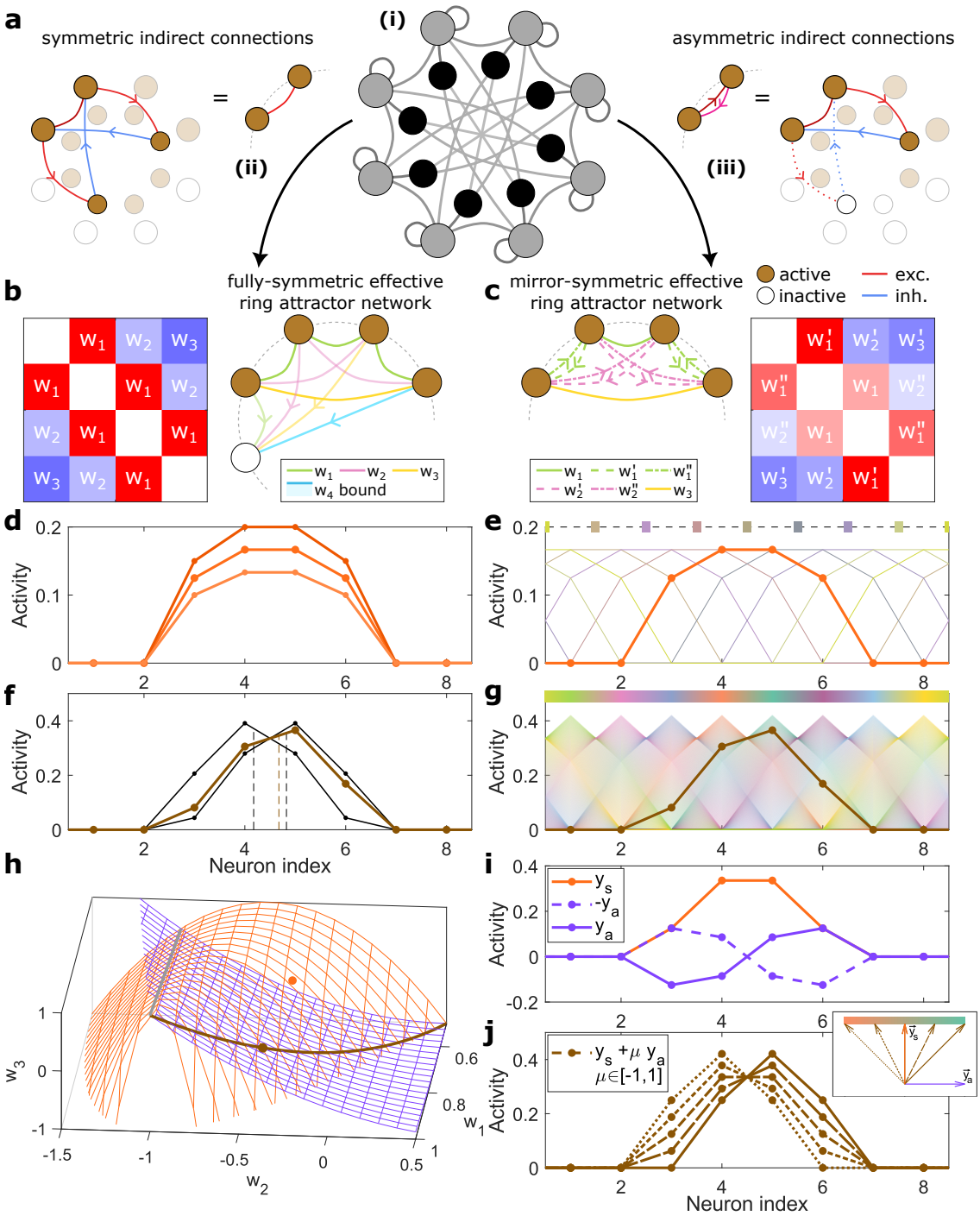

FIG. 2: **Constructing an effective ring attractor model with a continuum-encoding property.** (a) The two-population model (i) can be aggregated into two distinct effective ring attractor networks: (ii) fully-symmetric network, by assuming all neurons in the second population (inner circle) are active, or (iii) mirror-symmetric network, arising from asymmetric indirect connections when there are inactive neurons in the second population (inner circle) in steady state. In (iii) the dotted path is absent leading to the asymmetric weights. (b) and (c) Effective matrices for the 4 active neurons of the primary population for the fully- and mirror-symmetric models, respectively. (d) Example symmetric profile for the single-degeneracy case. Darker/lighter scarlet profiles depict the flexibility in amplitude increase/decrease. (e) 8 total profiles and their respective mean values at discrete uniformly-distributed locations (top) for the single-degeneracy case. (f) Example profile (brown) (as in (d)) for the double-degeneracy case, which enables asymmetric mirrored pairs of basis profiles (black). (g) Continuum of profiles and their mean values (top) span the whole neuron space. (h) Intersection points (brown) between the weight subspaces satisfying the symmetric conditions (scarlet) and the antisymmetric conditions (violet) yield valid ring attractors. (i) Symmetric (scarlet) and antisymmetric (violet) profiles (eigenvectors, inset) define the basis vectors for the activity profiles (j). Different values of $\mu \in [-1, 1]$ define different steady state profiles (or vectors, inset) in the specific angular interval (between neurons 4 and 5 in the figure).

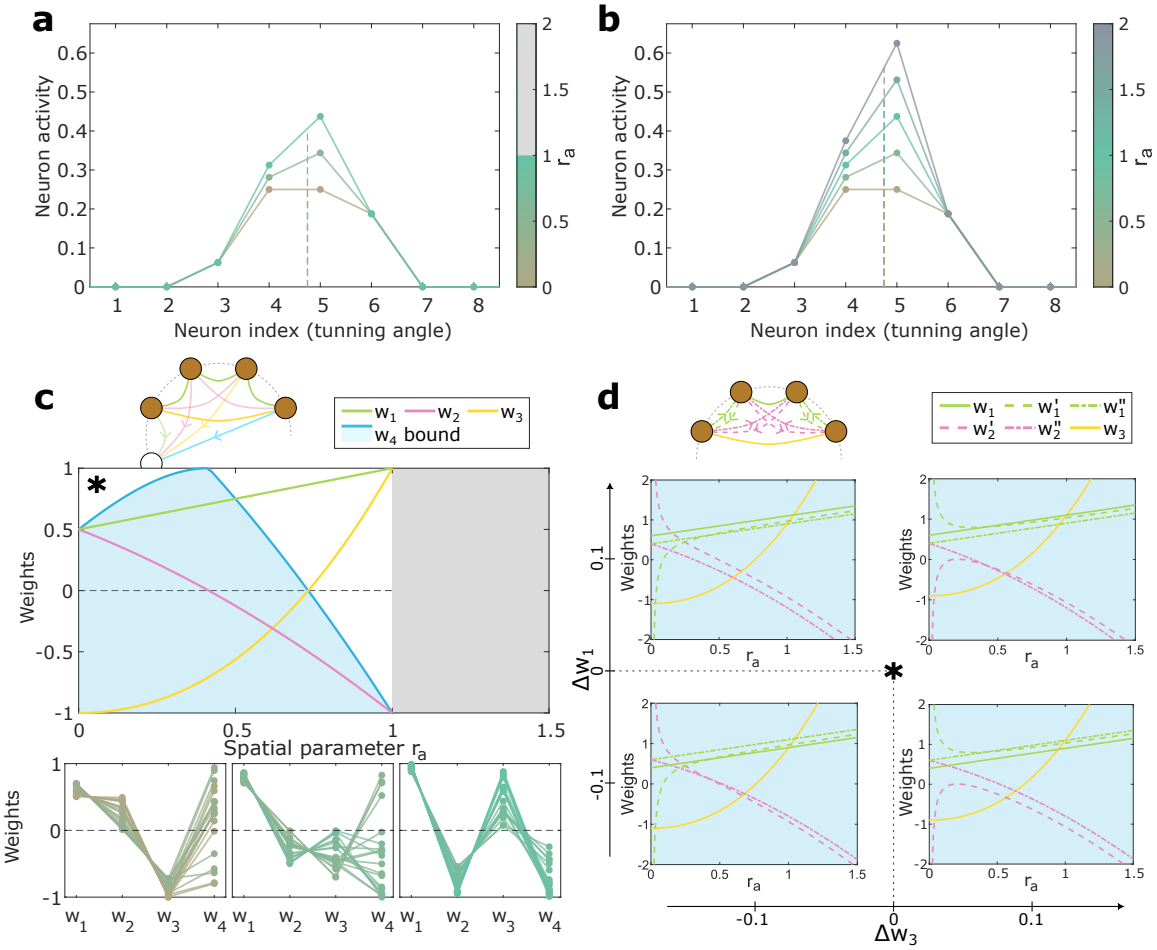

FIG. 3: **Spatial parameter r_a ties the shapes of the weight and activity profiles.** (a) and (b) Dependence of the shape of the neuron activity profiles on $r_a \in [0, 1]$ for the fully-symmetric case (a) and $r_a \in [0, 2]$ range for the mirror-symmetric case (a) which allows $r_a > 1$ unlike the symmetric model. (c) and (d) Dependence of the weights on the r_a parameter values for the fully- and mirror-symmetric cases, respectively. Sample weight profiles are shown for the fully-symmetric case (c, bottom) for three segments of r_a defined by the sign-changing points of w_2 ($r_a = \sqrt{2} - 1$) and w_3 ($r_a = \sqrt{3} - 1$). Values for w_4 were randomly sampled between -1 and its upper bound dependent on r_a (blue region). Weight profile shapes for the mirror-symmetric case (d) are dependent on two additional parameters $\Delta w_1 \equiv w_1 - w_{1,\mathtt{sym}}(r_a)$ and $\Delta w_3 \equiv w_3 - w_{3,\mathtt{sym}}(r_a)$.

This condition imposes inequalities on the weight space and, for instance, eliminates the grey branch in Fig. 2h. We also explore non-perturbative stability through a combination of analytical and numerical techniques where we allow the network to start with a random active set and activity profile of neurons and check that there are no unstable or growing modes, see Methods.

For both symmetric and mirror-symmetric networks, the steady state activity profiles are characterized by two free parameters, the amplitude denoted by $\sigma > 0$, and the shape parameter, $\mu \in [-1, 1]$, that linearly encodes the average angular location spanning 45^0. For a given σ, μ, the activity profiles that one generates in these models, interestingly only depend on one more key parameter combination involving the synaptic weights, $r_a \equiv y_{a4}/y_{a3} = y_{a5}/y_{a6}$, the ratio between the activities in the anti-symmetric eigenstate. This quantity is completely determined in terms of the weights and therefore fixed for a given ring-attractor, see Appendices B and C. It also is simply related to the ratio of the neuronal activities in the symmetric profile: $r_s \equiv y_{s4}/y_{s3} = y_{s5}/y_{s6} = r_a + 2$, and therefore controls the overall shape of the activity profiles. For a given angle, as one increases r_a, the profiles have sharper peaks as depicted in Figs. 3a and 3b. While in the symmetric model (Fig. 3a) this parameter is restricted to be in the range, $r_a \in (0, 1)$, the mirror symmetric profiles can have sharper peaks with r_a exceeding one (Fig. 3b), and therefore, in principle, offers a possibility of distinguishing the two different ring-attractor models using experimental activity measurements.

So far we have discussed the equality constraints on the recurrent weights between the active compass neurons coming from requiring a continuum of stable steady state profiles. However, self-consistency of the active set imposes several inequalities that involve effective *feedforward* weights from the active compass neurons to the inactive ones where the effective weights, as before, include indirect paths via the inhibitory population. Specifically, we must ensure that the total drives that the active neurons provide to all the inactive neurons are negative for all the steady state activity profiles. Due to the left-right symmetry of the models, one need only consider the drives to y_7 and y_8 (or equivalently y_1 and y_2). Since the space of steady states is two dimensional, for each inactive neuron we end up with two inequalities, and therefore four inequalities altogether. For the symmetric model, it turns out that two of these are automatically satisfied leading to two remaining inequalities. Fig. 3c depicts how these inequalities restrict w_4 as we vary r_a. Since the equality constraints discussed previously can be solved to yield unique functions, $w_{i,\mathtt{sym}}(r_a)$, $i = 1, \ldots, 3$, Fig. 3c comprehensively characterizes the two parameter ensemble of symmetric continuous ring attractors, details are included in Appendix B. In particular we notice that w_1 has to always be excitatory, $1/2 \leq w_1 < 1$, while at least one of w_2 or w_3 has to be inhibitory. Thus our model captures the intuitive requirement of local excitation and distal inhibition, but clearly the inhibition has significant flexibility. We plot some sample weight profiles in Fig. 3c (bottom, see Methods for details on how we generated them) that are categorized into three groups: $w_2 > 0, w_3 < 0$ (left), $w_2, w_3 < 0$ (middle), and $w_2 < 0, w_3 > 0$ (right). w_4 remains more flexible.

The mirror-symmetric model is much more flexible because the feedforward weights onto y_7 and y_8 are completely independent and they are also not related to the recurrent couplings. Thus we have an eleven dimensional space of ring attractors (three specifying the recurrent couplings, and four feedforward weights for each of the inactive neurons, y_7, y_8) that have to satisfy four inequalities stemming from the self-consistency arguments, essentially making the feedforward weights unconstrained at a single parameter level. Fig. 3d attempts to capture the various parameters characterizing the mirror symmetric ensemble, see Appendix C for more details. In Appendix G we also provide a partial characterization of externally driven fully symmetric ring attractor networks that require only a singly-degenerate effective weight matrix similar to what was considered in [21] and also point out the difficulty they may have in decoding the angular location from the bump profile when the external drive varies.

IV. Ring attractors consistent with EPG-Δ7 connectome

In the previous sections we have articulated how to obtain the ensemble of all two population threshold-linear networks that can produce continuous, stable and unique angular encoding. In this section we will test whether the fruit fly connectomics data involving the EPG-Δ7 system (Fig. 4a) can realize such a network. The EPG neurons are approximately divided into 8 compartments with 4 or 5 neurons in each and their recurrent connectivity obeys an approximate 8-fold circular symmetry: If C_{ij} represents the total number of synapses from neurons in the ith compartment to neurons in the jth compartment, then

$$C_{ij} \approx C_{i+1,j+1} \ , \tag{IV.1}$$

where from here onwards $(i + 8)$ will be identified with i. The Δ7 neurons can also be divided into 8 compartments (see Methods for details) which preserve similar connectivity patterns, so that Eq. (IV.1) is valid whether i, j represents an EPG index or a Δ7 index. In many ways it is the presence of the 8-fold circular symmetry that motivates the segmentation of the EPGs and Δ7's into 8 compartments. However, this symmetry is approximate, and thus to be consistent with our theoretical framework we symmetrized synapse counts (Fig. 4b) so that, $C_{ij} = C_{i+1,j+1}$, and their values are the average over the 8 instances of a given *type* of pre- and post-synaptic pair, see Methods.

While the synapse counts are expected to be proportional to the synaptic strengths, the proportionality constant can vary depending upon the pre- and post-synaptic neuron type. Accordingly, we choose four different scale factors, γ^{ab}, $a, b \in \{E, I\}$, to proportionately relate synapse counts and weights,

$$W_{ij} = \gamma^{ab} C^{ij} \ , \tag{IV.2}$$

where a, b labels the post and pre synaptic neuronal type, and E, I indicates the excitatory EPGs and inhibitory Δ7s respectively (Fig. 4c). Thus from the connectomic data we obtain a four dimensional space of realizable EPG-Δ7 networks. However, since the reduction to the effective network always involves $E \to I$ and $I \to E$ connections in conjunction, only the product $\widetilde{\gamma}^2 \equiv -\gamma^{EI}\gamma^{IE}$, appears in the effective network. Thus the connectome leads us to a three dimensional subspace of *realizable* effective networks. Simple dimensional arguments involving the recurrent weight space involving the active compass neurons can now tell us whether it is reasonable to expect to find connectome-constrained continuous ring attractors.

We first note that the intersection of an A and B dimensional linear subspace in an N dimensional

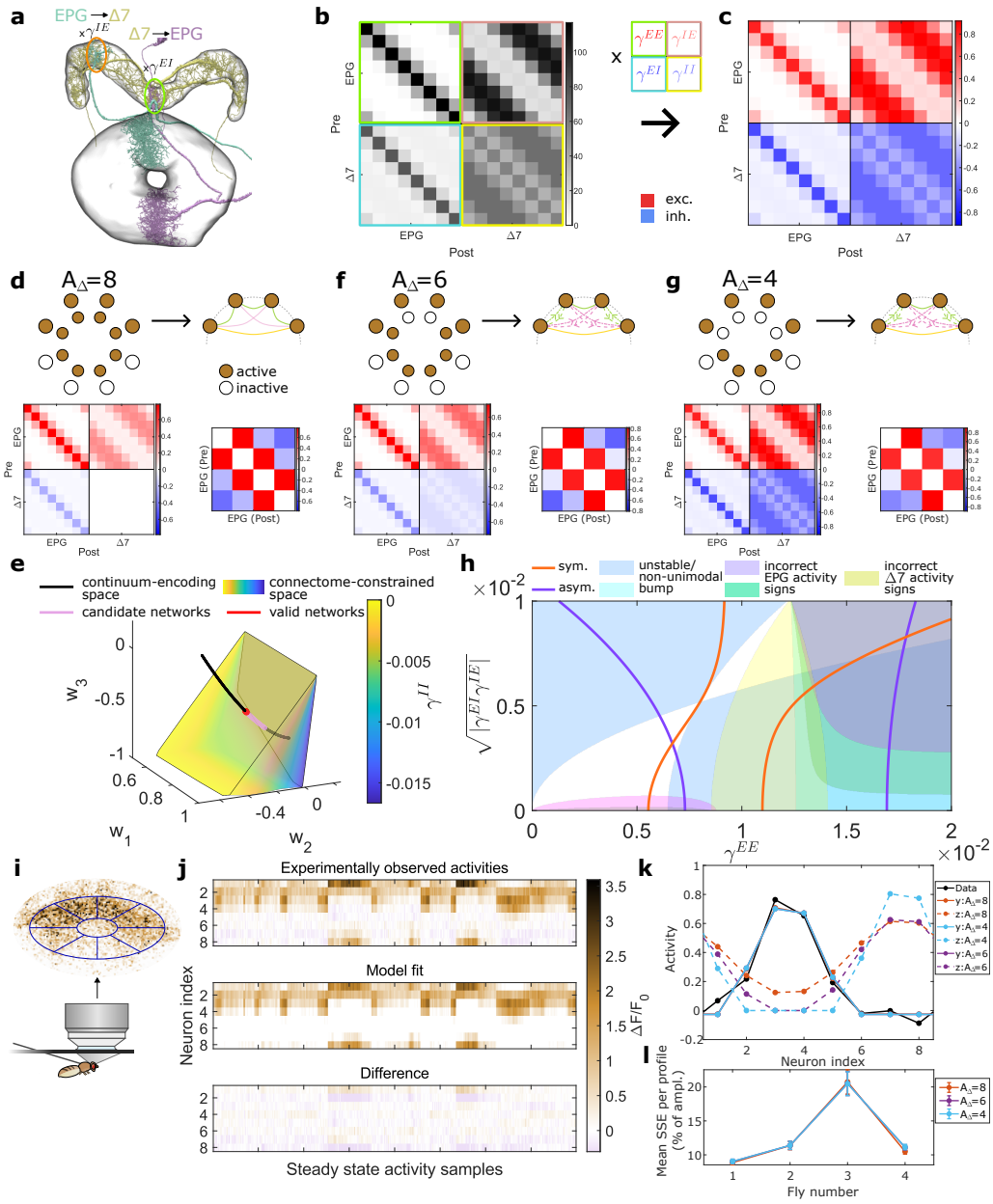

FIG. 4: **Constructing ring attractors from the EPG-Δ7 connectomic data.** (a) EPG (from the EB) and Δ7 neurons link between each other in the protocerebral bridge (PB), depicted using the hemibrain dataset. (b) Coarse-grained synaptic count matrix between the EPG and Δ7 neurons. Each block submatrix is parameterized with a single γ-parameter that scales the values within that submatrix, to generate the matrix in (c). Red - excitatory, blue - inhibitory weights. (d) Top: Two-population network (left; outer layer - EPG, inner layer - Δ7 population) with $A_\Delta = 8$ active Δ7 neurons aggregates to an effective fully-symmetric network of active EPG neurons (right). Bottom: Corresponding scaled and effective matrices. Value of $\gamma^{II} = 0$ was used. (e) Intersection points (pink line) between the theoretical conditions of the continuum-encoding space (black) and connectome-constrained space (color-coded volume) yield candidate weight profile solutions of ring attractors. Further constraints on the activity profiles narrow down the solution space of valid ring attractors (red). (f) and (g) Same as in (d), but for the cases with $A_\Delta = 6$ and $A_\Delta = 4$ active Δ7 neurons - mirror-symmetric effective networks. (h) γ-parameter view of the solution space (analogous to (e)), additionally showing regions with invalid ring attractors, color-coded according to the different inequality constraints they do not satisfy. (i) Schematic of the setup for two-photon calcium imaging of the EPG neuron population in the EB for tethered immobile flies in darkness. Top: Mean $\Delta F/F$ values of the calcium signal was computed in 8 regions of interest (ROIs) around the EB for each time point. (j) Top: Unwrapped activity profiles for different time points, during which the EPG bump position was relatively stationary. Middle and bottom: Corresponding fits with the fully-symmetric model and their differences to the data, respectively. (k) Example model fits for a single EPG activity profile (solid lines) and the respective predictions for the activity of the Δ7 population (dashed) for the three different models (d,f,g). (l) Summary statistics (mean±s.e.m.) of the fitting data with the three aforementioned models for multiple flies. The SSE per profile is normalized to the bump amplitude.

space is, $(A + B - N)$ dimensional. While the subspaces we encounter are not linear, this basic intuition mostly carries forward. Let us first consider the symmetric model that requires all $\Delta 7$ neurons to remain active (Fig. 4d). The ensemble of symmetric continuous attractors span a one dimensional subspace (A=1), while the realizable networks span a three dimensional volume (B=3) in the three dimensional weight-space (N=3) of active neurons. Thus we can expect to find connectome-constrained ring attractors in the form of a finite one-dimensional curve in the weight-space (Fig. 4e). However, these solutions also has to satisfy various inequalities that has been previously discussed. Additionally, one has to ensure that the drives to the $\Delta 7$ neurons are consistent with their assumed activity-state (all active for the symmetric networks). So, for the symmetric model we need to ensure that all the $\Delta 7$ drives are positive. It turns out that for the EPG-$\Delta 7$ connectome, these inequalities severely constrain the solutions so that the recurrent connectivity among the $\Delta 7$'s, $\gamma^{II} \approx 0$, and we effectively end up with a single symmetric connectome-constrained ring attractor network (Fig. 4e). Now the ensemble of ring attractors in the mirror symmetric model (Figs. 4f and 4g) curves out a three dimensional subspace (A=3). Thus we expect a zero dimensional intersection with the three dimensional subspace (B=3) of realizable networks in the six dimensional weight space (N=6). Indeed we were able to find unique connectome constrained ring attractors assuming either six $\Delta 7$s (Fig. 4f), or four $\Delta 7$s (Fig. 4g) are active, Appendix D contains more details. No consistent solutions compatible with all the inequalities are possible when only two $\Delta 7$s are active.

One can take a different approach to understand how the theoretical ensemble of ring attractors can be combined with connectomic information to produce connectome-constrained models. This approach is computationally easier to implement and to scale up to complex networks with more neuronal populations. Rather than working in the weight space, here one works in the γ-space, whose dimensionality grows more slowly with complexity as compared to weight-space. We first note that starting from the synapse count matrix and applying the methods described above yields all the effective weights as functions of three parameters: $w_i = w_i(\gamma^{EE}, \gamma^{II}, \widetilde{\gamma})$. All the equality and inequality constraints that were derived in terms of the various effective weight parameters then can be recast as equalities and inequalities in the γ -space. In Fig. 4h, we show the equality constraints as the scarlet (symmetric eigenvector condition) and violet (anti-symmetric eigenvector condition) curves in γ-space where we have set $\gamma^{II} = 0$. The colored regions are excluded by various inequalities, but we are left with allowed white regions where the scarlet and violet curves are seen to have an intersection. This then corresponds to the connnectome-constrained ring attractor network. By varying $0 < \widetilde{\gamma} \lesssim 7 \times 10^{-4}$, we obtain a small continuous family of viable ring attractors (short red curve). Since the mirror-symmetric model has to satisfy three equations, with three flexible γ parameters we only get unique solutions once the inequalities are imposed.

Figs. 4d, 4f and 4g depict the three connectome-constrained EPG-$\Delta 7$ weight matrices that emerges as self-sustaining stable continuous ring attractor networks where $A_\Delta = 8$, 6, and 4, $\Delta 7$ neurons are active, respectively, We note the striking difference between the symmetric and mirror-symmetric models: the former allows only tiny recurrent couplings between the $\Delta 7$ neurons, $\gamma^{II}/\gamma^{EE} \sim \mathcal{O}(10^{-2})$, and $\widetilde{\gamma} \sim \gamma^{EE}$, whereas in both the mirror-symmetric models, $\widetilde{\gamma} \sim \gamma^{EE} \sim \gamma^{II}$. Somewhat surprisingly, all these different weight-matrices give rise to similar activity profiles for the EPG's that are also consistent with the observed neuronal patterns (Figs. 4i to 4l). The difference between experimentally observed activities and the model fits that were obtained by varying the amplitude, σ, and the shape parameter, μ, were very similar for all the three networks and for all the four flies that we imaged (Fig. 4l). These differences were small as can be seen from the comparison of the steady state experimental activity profiles and theoretical fits from the symmetric network (Fig. 4j). Not surprisingly, the three different connectome-constrained ring attractors lead to discernibly different $\Delta 7$ activity profiles (Fig. 4k) that can, in principle, be used to test and discriminate between them.

V. Connectome properties that lead to a viable ring attractor

We have seen that there exists a two dimensional subspace of symmetric networks, and an eleven dimensional subspace of mirror-symmetric networks that are continuous ring attractors. Nevertheless the subspace of ring attractors occupy a subspace of zero-measure among the space of all (mirror)symmetric networks because the synaptic weights have to be fine tuned to satisfy specific relationships. On the positive side though, the scale factors that relate the synapse counts to synaptic weights offer flexibility. So how difficult is it for a connectome to be able to realize ring attractor dynamics? To test this we generated $\mathcal{O}(1000)$ connectomes by randomly varying synapse count matrices (see Methods) while respecting the 8-fold circular symmetry. About 15% of the time it was possible to adjust the scale factors to realize a symmetric ring attractor network starting with these connectomes. This intriguing result suggests to us that while the fly connectome had to be approximately

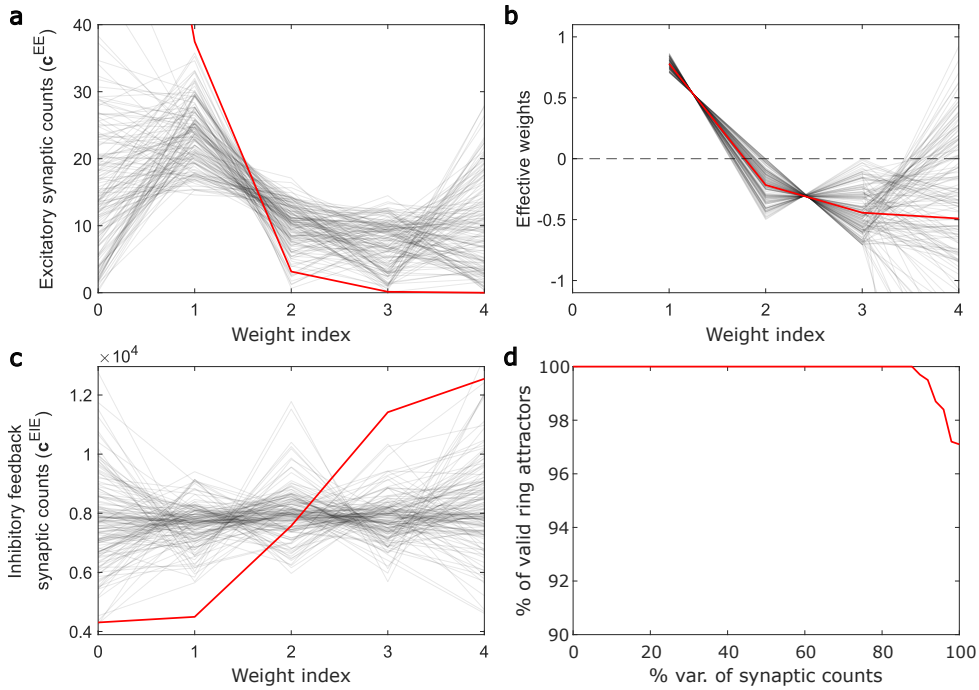

FIG. 5: **Conditions for generating valid continuous ring attractors.** (a-c) Randomly sampled synaptic weight profiles that resulted in valid continuous ring attractors. Example profiles (∼15% of the successfully sampled from the 2000 initially sampled profiles) are shown in black for the excitatory $\mathbf{c}^{EE}$ synapse-counts (a), the corresponding $\mathbf{c}^{EIE}$ inhibitory feedback synapse counts (c), and the resulting effective weights (b). The $\mathbf{c}^{EE}$ and $\mathbf{c}^{EIE}$ profiles were normalized to match the connectomic profiles (red) at indices 1-3 on average. Profiles from the second group of weight profiles (Fig. 3c, $w_2 < 0$, $w_3 < 0$) are only shown. (d) Ring attractors persist upon increasing perturbations around the connectome-derived parameter set. Each parameter value was uniformly randomly sampled in the ±(% var.) range around its connectome-derived value. The mean success rate of 1000 samples per % var. value was calculated.

in the right regions (and it is) within the space of synapse connectivity for our proposed mechanism to work, the ensemble of synapse count matrices that can lead to a viable ring attractor is actually not a zero-measure subspace! We also noticed that while the synapse counts between the compass neurons (Fig. 5a) mirrored the pattern of the effective weights (Fig. 5b), the inhibition necessary for supporting the activity bump could be much more flexible (Fig. 5c). However, interestingly the connectomic inhibitory weights exhibit a very specific trend the reason for which is unclear to us.

While evolution could have led the fruit fly connectome to arrive at a network state that can support continuous angular encoding, such an encoding strategy must be robust against small fluctuations in changes to individual synapse counts and synapse strengths. To explore possible ways to accomplish this issue we decided to vary synapse counts around the fly connectome. Somewhat surprisingly, even when the synapse counts varied by about 90%, it was possible to adjust the scale factors to, almost always, yield a viable ring attractor. Fig. 5d shows how even a relatively large variation of the connectome often supported a ring attractor network. We hypothesize that there may exist a homeostatic mechanism which can compensate for small changes in individual syanpse strengths by appropriately adjusting the scale factors via global feedback mechanism that track stability of the activity profiles.

VI. Discussion

In this paper we have presented a mechanism to obtain a threshold-linear ring-attractor network that can self-sustain stable neuronal activities encoding angular locations continuously in the process. Moreover, our framework is able to obtain the entire ensemble of such continuous ring attractors given the specific number of neurons (active and inactive) involved. Although the synaptic weights have to satisfy certain relationships to realize the continuum of steady state patterns, we show how the flexibility in how synapse counts scale to synapse weights may allow us to find ring attractors amidst our ensemble that are consistent with known connectome. Specifically, we found three connectome-constrained ring attractors involving the EPG and Δ7 neurons in fruitfly heading direction system, each with their own distinctive Δ7 activity patterns.

As the compass EPG neurons in fruit flies are excitatory, inhibitory pathways involving other neuronal populations becomes a key ingredient of the ring attractor network. While here we have focused our attention on the inhibitory $\Delta 7$ neurons, our framework can consider alternatives such as ER6 ring neurons that provide the same broad inhibitory signal to all the EPG neurons in contrast with the $\Delta 7$ neurons that target EPG neurons opposite to the activity bump. Preliminary analysis suggests that it is possible to realize a connectome-constrained ring attractor network with EPGs and ER6s. Our formalism can also incorporate more than one columnar cell types: every additional cell type adds new scale factors associated with new pre-and post-synaptic pairs, allowing an expansion of *realizable* networks. Therefore we will find more ring attractor networks that are consistent with the connectome but these will come with their distinctive activity predictions for all the neuronal populations involved.

Another key biological requirement is for ring attractor networks to be able to integrate velocity inputs. While a connectome based mechanistic implementation of the same is out of the scope of the present paper, in Appendix H we present an illustration of how appropriate velocity signals to the compass neurons can smoothly change the shape of the activity profile involving a given set of active neurons, so that the angular location can integrate the velocity information.

While the exquisite connectomic knowledge motivated us to consider the fruit fly heading direction system, we believe our framework can be generalized to describe other model animals. For instance the central complex is remarkably conserved among insects, [35] and [36] provides a detailed comparison of the fruit fly heading direction system with that of cockroach and locusts respectively. We note that the key requirement for our continuous enconding mechanism to work is the double degeneracy of the effective weight matrix, the number of active and inactive neurons (neuronal compartments) and the precise synaptic connectivity, that can all differ from one insect species to another, only changes the details of the implementation. Unlike insects, in larger animals such as rodents, the heading direction system contains several hundreds of cells that do not exhibit any obvious anatomical alignment of symmetry. Nevertheless, could our network provide an *effective* coarse-grained description of the network dynamics?

A critical component of ring attractor networks, including ours, is that the synaptic weights have to satisfy specific relationships. While one can imagine years of evolution leading to such a fine tuned network state, it is imperative to have a mechanism in place to address spontaneous changes to connectivity. Intriguingly, our analysis shows that small changes in synapse counts may be compensated by small changes in scale factor that don't act at the individual synapse level but only at the network level. Thus, one can conjecture, a homeostatic mechanism where a detection of increase or decrease in the average neuronal activity triggers compensatory changes in the scale factors that can be mediated via neuromodulators or other global signaling pathways.

Symmetries are also popular considerations for constructing ring attractor networks but we believe this is not a fundamental requirement. We conjecture that as long as we ensure the effective active weight matrix is doubly degenerate, we will have two encoding coordinates that may be able to represent amplitude and angular location, see also [37] for considerations related to asymmetric ring attractors. There may be additional requirements related to smooth transitions between the active sets as was the case in our mirror-symmetric example. Indeed, our mirror-symmetric network which violates circular symmetry and left-right symmetry, is able to encode angles continuously and stably. We also point out that [38, 39] has shown how to generate networks that can reproduce a prescribed set of steady state patterns, and these networks need not have any symmetries. This may allow us to generate new asymmetric ring attractor networks by enforcing that they reproduce the activity patterns we have derived. To summarize, we believe the mechanism we have proposed may provide a general and robust way to encode a continuous variable and provides promising new research directions.

VII. Methods

Experimental procedure

Fly preparation

Imaging experiments were performed on 6-8 days old female UAS-jGCaMP7f;SS00096-GAL4 flies, reared at 23°C in 60% relative humidity with a 16:8 light:dark cycle on standard cornmeal fly food. To express the genetically encoded calcium indicator jGCaMP7f [40] in EPG neurons specifically, jGCaMP7f flies (20XUAS-IVS-Syn21-jGCaMP7f-p10 in VK00005; RRID:BDSC_79031) were crossed with the stable split EPG GAL4 driver line SS00096 [41]. The flies were prepared for imaging as previously described [22, 42]. Briefly, flies were anesthetized at 4°C, their proboscis was fixed with wax to reduce brain movements, their thorax was UV-glued to a tether pin for stability and manipulation, and they had their legs removed and the stumps and dorsal abdomen glued to reduce spontaneous motor activity. The

fly's head was positioned with a micromanipulator in a holder with a recording chamber and UV-glued to it to immobilize the head for dissection and brain imaging. To gain optical access to the brain, a section of cuticle between the ocelli and antennae was removed, along with the underlying fat tissue and air sacs. Throughout the experiment, the head in the imaging holder was submerged in saline containing (in micromolar): NaCl (103), KCl (3), TES (5), trehalose (8), glucose (10), NaHCO3 (26), NaH2PO4 (1), CaCl2 (2.5) and MgCl2 (4), with a pH of 7.3 and an osmolarity of 280mOsm.

Two-photon calcium imaging

For each fly we collected imaging data during trials in darkness. Calcium imaging was performed with a custom-built two-photon microscope controlled with ScanImage (version 2022, Vidrio Technologies) [43]. Excitation of jGCaMP7f was generated with an infrared laser (920nm; Chameleon Ultra II, Coherent) with approximately $15mW$ of power, as measured after the objective (×60 Olympus LUMPlanFL/IR, 0.9 numerical aperture). Fast Z-stacks (eight planes with $8\mu m$ spacing and three fly-back frames) were collected at 10Hz by raster scanning (128×128 pixels, $\sim 75 \times 75\mu m^2$) using an 8kHz resonant-galvo system and piezo-controlled Z positioning. Focal planes were selected to cover the full extent of EPG processes in the EB. Emitted light was directed (primary dichroic: 735, secondary dichroic: 594), filtered (filter A: 680 SP, filter B: 514/44) and detected with a GaAsP photomultiplier tube (H10770PB-40, Hamamatsu).

Data analysis and model fitting

The recorded XYZ stacks of EPG activity in the EB were first averaged over the Z dimension, and then drift-corrected over time for alignment. The aligned XYZ stacks were then averaged over time to get a mean EPG activity profile in 3D. To account for the 3D orientation of the EB within the brain, relative to the objective angle, a 3D mask of the EB was constructed from the masks in each individual slice separating the EB signal pixels from background. PCA was then performed on the 3D signal intensity points within the EB mask to find the optimal plane on which the EB signal should be projected. An ellipse was fitted through the projected data, to account for the XY rotation of the EB, and automatically segmented into 8 wedges, similarly as in Fig. 4i, top. Average fluorescence signal $F^i(t)$ was calculated for every wedge $i = \overline{1,8}$ and time point $t = \overline{1,T}$, and the baseline fluorescence signal $F^i{}_0$ per wedge was calculated as the 10th percentile from the $F^i(t)$ values. Finally, the $\Delta F/F_0 = (F - F_0)/F_0$ values were calculated for every wedge and time point, giving an $8 \times T$, $\Delta F/F_0$ data matrix.

We calculated the population vector average (PVA) at each time point t to estimate the EPG bump amplitude and orientation, as the circular mean of the 8-dimensional $\Delta F/F_0$ vector at t and the corresponding wedge angles (CircStat toolbox [44]). Time points where there is a lack of bump strength, as measured by the likelihood of uniformity of the circular data (p-value$\leqslant$0.5 of Rayleigh's test; CircStat), were excluded from the analysis. A moving average with 11 frames of the PVA angle was used to estimate the instantaneous bump velocity. Time points where the bump velocity was greater than $20°/s$ were excluded from the analysis to filter out the non-stationary states.

To estimate the model fit of each of the three models (with $A_\Delta = 8$, $A_\Delta = 6$ and $A_\Delta = 4$) on the EPG activity data, we used their estimated values for the spatial parameter r_a. For each time point we then estimated the σ amplitude and μ position of the bump using Eq. B.29 (Appendix), as well as an additional offset parameter to accommodate the baseline. Non-linear least squares (*lsqnonlin*, Matlab) was performed over the $\Delta F/F_0$ data circularly shifted for each of the 8 possible consecutive activity sets of EPG wedges, and the best fit was chosen for that time point and the active EPG set stored. Predictions for the $\Delta 7$ activities were subsequently generated using Eq D.74 (Appendix).

Constructing connectome-derived ring attractors

To estimate the average synaptic strength in the connectivity between and within the computational units of EPG and $\Delta 7$ cell types from the hemibrain dataset [26, 27], similarly as in [33], we estimate the average synaptic strength at the level of wedges in the EB (between EPGs) and glomeruli in the PB (EPG-$\Delta 7$ and between $\Delta 7$). Each $\Delta 7$ was assigned to a glomeruli (wedge) based on the maximal cumulative output synaptic counts to EPG neurons grouped by wedges. Due to the circular and mirror symmetry, for 8 computational units (wedges) the resulting vectors are 5-dimensional. For example, for EPG-EPG connectivity we obtain $\mathbf{c}^{EE} = ({c^{EE}}_0, {c^{EE}}_1, {c^{EE}}_2, {c^{EE}}_3, {c^{EE}}_4)$, where ${c^{EE}}_0$ is the average strength within a unit (self-loops), ${c^{EE}}_1$ is synaptic strength between neighboring units, while ${c^{EE}}_4$ is the synaptic strength between units on opposite sides of the EB.

For the fully symmetric case we typically set $\gamma^{II} = 0$ (except for Fig. 4d), as its range was very narrow ($[-7.15 \times 10^{-4}, 0]$). A given connectome data (from the hemibrain data or from the sampled connectomes), typically of the form $\left(\begin{smallmatrix} C^{EE} & C^{EI} \\ C^{IE} & C^{II} \end{smallmatrix}\right)$, is constructed from the corresponding 5-dimensional synaptic weight vectors $\mathbf{c}^{EE}$, $\mathbf{c}^{EI}$, $\mathbf{c}^{IE}$ and $\mathbf{c}^{II}$. An effective inhibitory feedback matrix can also be computed as $C^{EIE} = C^{EI}\,(I - \gamma^{II} C^{II})^{-1}\, C^{IE}$, with its weight vector $\mathbf{c}^{EIE}$ in the first column. From the excitatory $\mathbf{c}^{EE}$ and inhibitory feedback

$\mathbf{c}^{EIE}$ synaptic weight vectors a 3×2 effective connectome matrix is constructed: $\hat{C} = \begin{pmatrix} c^{EE}{}_1 & c^{EIE}{}_1 \\ c^{EE}{}_2 & c^{EIE}{}_2 \\ c^{EE}{}_3 & c^{EIE}{}_3 \end{pmatrix}$. A linear combination of these two vectors should yield the effective weights $\mathbf{w} = (w_1, w_2, w_3)$, thus a two-dimensional mapping parameter vector $\hat{\gamma} = (\hat{\gamma}^{EE}, \hat{\gamma}^{EI}\hat{\gamma}^{IE})$ from the effective connectome matrix to the effective weights: $\hat{C}\,\hat{\gamma} = \mathbf{w}(r_a)$ would yield the connectome-derived ring attractor, with the effective weights satisfying the analytically derived conditions (cited in Table 1, Appendix), expressed as a function of r_a. Here $\hat{\gamma}^{EI}\hat{\gamma}^{IE}$ is a lumped parameter. This translates into finding the value r_a in the range $[0, 1]$ that minimizes $||(\hat{C}\,\hat{C}^{\dagger} - I)\,\mathbf{w}(r_a)||_2$, which was computed numerically with precision of 10^{-10}. The correct γ values were then calculated using $\gamma = (\gamma^{EE}, \gamma^{EI}\gamma^{IE}) = \hat{\gamma}./(1 + (c^{EE}{}_0, c^{EIE}{}_0)^T\,\hat{\gamma})$, accounting for the normalization step due to the subtraction of the w_0 self-loop weights. The lumped $\gamma^{EI}\gamma^{IE}$ parameter was further distributed to $\gamma^{EI} = -\gamma^{IE} = -\sqrt{|\gamma^{EI}\gamma^{IE}|}$, to account for the correct signs. The final weights were then tested to check if they satisfy the inequalities cited in Table 2 (Appendix; whether the presumed active and inactive EPGs have the correct signs and whether w_4 is below its upper bound), and γ values were tested for $\gamma^{EE} > 0$, $\gamma^{EI} < 0$ and $(1 + (c^{EE}{}_0, c^{EIE}{}_0)^T\,\gamma^{map}) > 0$. Furthermore, for the cases where $\gamma^{II} < 0$ (Fig. 4d) the correct signs of the active and inactive $\Delta 7$ neurons were also checked in steady state for $\mu = 1$ and $\sigma = 1$ (Eq. B.29, Appendix).

In the mirror-symmetric case, there are effectively three γ parameters that would map the connectome data to the effective weights. There are also three equations that the effective weights should satisfy for generating a ring attractor (Table 1, Appendix), thus solving a system of nonlinear equations would yield the fixed-point solutions. For this, we used Matlab's *fsolve* solver with the trust-region-dogleg algorithm initiated from 1000 different γ values, a combinatorial set of 10 values per single γ parameter, sampled uniformly from $\gamma^{EE} \in [0, 0.025]$, $\gamma^{EI}\gamma^{IE} \in [-0.002, 0]$ and $\gamma^{II} \in [-0.02, 0]$, respectively. For each candidate γ vector the effective matrix was computed as

$$\bar{W}_{ij}^{EE} \equiv \frac{\gamma^{EE} C_{ij}^{EE} + \gamma^{EI}\gamma^{IE} C_{ij}^{EIE}}{1 - \gamma^{EE} C_{ii}^{EE} - \gamma^{EI}\gamma^{IE} C_{ii}^{EIE}}\,, \tag{VII.3}$$

for all $i \neq j$, and the diagonal terms were set to zero, $\bar{W}_{ij}^{EE} = 0$, see Appendix D for details. From the effective matrix the 6 effective weights are extracted and left-hand sides of the equations cited in Table 1 (Appendix) calculated. Through iterative procedure the sum of squares of the left-hand side values is minimized by *fsolve* to calculate the fixed point. Similarly with the fully-symmetric case, the corresponding additional inequalities are checked to filter out invalid solutions.

For both, the fully-symmetric and the mirror-symmetric ring attractors the marginal stability of the system was verified numerically by perturbing the continuum states with additive noise of 10% and 50% from the maximal state component and checking whether the system converges back to the continuum states using the *ode23* solver in Matlab. This happened 100% of the time for the three computed connectome-derived ring attractors.

Generating randomly sampled connectomes

To generate a random connectome sample, each value of the $\mathbf{c}^{EE}$, $\mathbf{c}^{EI}$, $\mathbf{c}^{IE}$ and $\mathbf{c}^{II}$ vectors was randomly sampled and then compiled into a matrix form. The resulting system was then tested whether it fulfills the stability and continuum conditions of the fully-symmetric model, as well as if the assumed active states of EPG and $\Delta 7$ neurons was correct. Value of $\gamma^{II} = 0$ was used. For Figs. 5a to 5c the random samples were uniform on a 0-40 range. As the γ parameters allow arbitrary scaling, the resulting profiles were normalized to match the hemibrain connectome model values. For Fig. 5d samples for each parameter c_x and % variation value were drawn uniformly from the $c_x \pm (\%var.)$ range. 2000 and 1000 number of samples per condition were drawn for the two cases, respectively.

VIII. Acknowledgments

We would like to thank Vivek Jayaraman, Daniel Turner-Evans, Marcella Noorman, and Ann M. Hermundstad for insightful discussions. We are grateful to Vivek Jayaraman and Brad K. Hulse for experimental support. This work was supported by the Howard Hughes Medical Institute. JEF acknowledges institutional support to the National Institute for Theory and Mathematics in Biology from the National Science Foundation (grant number DMS-2235451) and the Simons Foundation (grant number MPTMPS-00005320).

[1] Jeffrey S Taube. The head direction signal: origins and sensory-motor integration. *Annual Review of Neuroscience*, 30:181–207, 2007.

[2] Brad K Hulse and Vivek Jayaraman. Mechanisms underlying the neural computation of head direction. *Annual Review of Neuroscience*, 43:31–54, 2020.

[3] Mikail Khona and Ila R Fiete. Attractor and in-

tegrator networks in the brain. *Nature Reviews Neuroscience*, 23(12):744–766, 2022.

[4] Rishidev Chaudhuri and Ila Fiete. Computational principles of memory. *Nature Neuroscience*, 19(3):394–403, 2016.

[5] Carlos D Brody, Ranulfo Romo, and Adam Kepecs. Basic mechanisms for graded persistent activity: discrete attractors, continuous attractors, and dynamic representations. *Current Opinion in Neurobiology*, 13(2):204–211, 2003.

[6] James J. Knierim and Kechen Zhang. Attractor dynamics of spatially correlated neural activity in the limbic system. *Annual Review of Neuroscience*, 35(1):267–285, 2012. PMID: 22462545.

[7] W.E. Skaggs, J.J. Knierim, H.S. Kudrimoti, and B.L. McNaughton. A model of the neural basis of the rat's sense of direction. *Advances in Neural Information Processing Systems*, 7:173–180, 1995.

[8] Shun-ichi Amari. Dynamics of pattern formation in lateral-inhibition type neural fields. *Biological Cybernetics*, 27(2):77–87, 1977.

[9] R Ben-Yishai, R L Bar-Or, and H Sompolinsky. Theory of orientation tuning in visual cortex. *Proceedings of the National Academy of Sciences*, 92(9):3844–3848, 1995.

[10] K Zhang. Representation of spatial orientation by the intrinsic dynamics of the head-direction cell ensemble: a theory. *Journal of Neuroscience*, 16(6):2112–2126, 1996.

[11] DA Redish, AN Elga, and DS Touretzky. A coupled attractor model of the rodent head direction system. *Network: Computation in Neural Systems*, 7(4):671–685, 1996.

[12] Marcelo Camperi and Xiao-Jing Wang. A model of visuospatial working memory in prefrontal cortex: recurrent network and cellular bistability. *Journal of Computational Neuroscience*, 5(4):383–405, 1998.

[13] Albert Compte, Nicolas Brunel, Patricia S Goldman-Rakic, and Xiao-Jing Wang. Synaptic mechanisms and network dynamics underlying spatial working memory in a cortical network model. *Cerebral Cortex*, 10(9):910–923, 2000.

[14] Patricia E Sharp, Hugh T Blair, and Jeiwon Cho. The anatomical and computational basis of the rat head-direction cell signal. *Trends in Neurosciences*, 24(5):289–294, 2001.

[15] Kyobi S. Kakaria and Benjamin L. de Bivort. Ring attractor dynamics emerge from a spiking model of the entire protocerebral bridge. *Frontiers in Behavioral Neuroscience*, 11, 2017.

[16] Ran Darshan and Alexander Rivkind. Learning to represent continuous variables in heterogeneous neural networks. *Cell Reports*, 39(1):110612, 2022.

[17] Rishidev Chaudhuri, Berk Gerçek, Biraj Pandey, Adrien Peyrache, and Ila Fiete. The intrinsic attractor manifold and population dynamics of a canonical cognitive circuit across waking and sleep. *Nature Neuroscience*, 22(9):1512–1520, 2019.

[18] JS Taube, RU Muller, and JB Ranck. Head-direction cells recorded from the postsubiculum in freely moving rats. i. description and quantitative analysis. *Journal of Neuroscience*, 10(2):420–435, 1990.

[19] Klaus Wimmer, Duane Q Nykamp, Christos Constantinidis, and Albert Compte. Bump attractor dynamics in prefrontal cortex explains behavioral precision in spatial working memory. *Nature Neuroscience*, 17(3):431–439, 2014.

[20] Yoram Burak and Ila R Fiete. Accurate path integration in continuous attractor network models of grid cells. *PLOS Computational Biology*, 5(2):e1000291, 2009.

[21] Marcella Noorman, Brad K Hulse, Vivek Jayaraman, Sandro Romani, and Ann M Hermundstad. Maintaining and updating accurate internal representations of continuous variables with a handful of neurons. *Nature Neuroscience*, pages 1–11, 2024.

[22] Johannes D Seelig and Vivek Jayaraman. Neural dynamics for landmark orientation and angular path integration. *Nature*, 521(7551):186–191, 2015.

[23] Sung Soo Kim, Hervé Rouault, Shaul Druckmann, and Vivek Jayaraman. Ring attractor dynamics in the drosophila central brain. *Science*, 356(6340):849–853, 2017.

[24] J. Green, A. Adachi, K. K. Shah, J. D. Hirokawa, P. S. Magani, and G. Maimon. A neural circuit architecture for angular integration in drosophila. *Nature*, 546(7656):101–106, 2017.

[25] Daniel B. Turner-Evans, Stephanie Wegener, Hervé Rouault, Romain Franconville, Tanya Wolff, Johannes D. Seelig, Shaul Druckmann, and Vivek Jayaraman. Angular velocity integration in a fly heading circuit. *eLife*, 6:e23496, 2017.

[26] Brad K Hulse, Hannah Haberkern, Romain Franconville, Daniel B Turner-Evans, Shin-ya Takemura, Tanya Wolff, Marcella Noorman, Marisa Dreher, Chuntao Dan, Ruchi Parekh, et al. A connectome of the drosophila central complex reveals network motifs suitable for flexible navigation and context-dependent action selection. *Elife*, 10, 2021.

[27] Louis K Scheffer, C Shan Xu, Michal Januszewski, Zhiyuan Lu, Shin-ya Takemura, Kenneth J Hayworth, Gary B Huang, Kazunori Shinomiya, Jeremy Maitlin-Shepard, Stuart Berg, et al. A connectome and analysis of the adult drosophila central brain. *elife*, 9:e57443, 2020.

[28] Zhihao Zheng, J Scott Lauritzen, Eric Perlman, Camenzind G Robinson, Matthew Nichols, Daniel Milkie, Omar Torrens, John Price, Corey B Fisher, Nadiya Sharifi, et al. A complete electron microscopy volume of the brain of adult drosophila melanogaster. *Cell*, 174(3):730–743, 2018.

[29] Sven Dorkenwald, Arie Matsliah, Amy R Sterling, Philipp Schlegel, Szi-Chieh Yu, Claire E McKellar, Albert Lin, Marta Costa, Katharina Eichler, Yijie Yin, et al. Neuronal wiring diagram of an adult brain. *Nature*, 634(8032):124–138, 2024.

[30] Elizabeth C Marin, Billy J Morris, Tomke Stürner, Andrew S Champion, Dominik Krzeminski, Griffin Badalamente, Marina Gkantia, Christopher R Dunne, Katharina Eichler, Shin-ya Takemura, et al. Systematic annotation of a complete adult male drosophila nerve cord connectome reveals principles of functional organisation. *BioRxiv*, pages 2023–06, 2023.

[31] Janne K Lappalainen, Fabian D Tschopp, Sridhama Prakhya, Mason McGill, Aljoscha Nern, Kazunori Shinomiya, Shin-ya Takemura, Eyal Gruntman, Jakob H Macke, and Srinivas C Turaga. Connectome-constrained networks predict neural activity across the fly visual system.

Nature, pages 1–9, 2024.

[32] Tanya Wolff, Nirmala A Iyer, and Gerald M Rubin. Neuroarchitecture and neuroanatomy of the drosophila central complex: A gal4-based dissection of protocerebral bridge neurons and circuits. *Journal of Comparative Neurology*, 523(7):997–1037, 2015.

[33] Daniel B Turner-Evans, Kristopher T Jensen, Saba Ali, Tyler Paterson, Arlo Sheridan, Robert P Ray, Tanya Wolff, J Scott Lauritzen, Gerald M Rubin, Davi D Bock, et al. The neuroanatomical ultrastructure and function of a biological ring attractor. *Neuron*, 108(1):145–163, 2020.

[34] Cheng Lyu, LF Abbott, and Gaby Maimon. Building an allocentric travelling direction signal via vector computation. *Nature*, 601(7891):92–97, 2022.

[35] Adrienn G Varga and Roy E Ritzmann. Cellular basis of head direction and contextual cues in the insect brain. *Current Biology*, 26(14):1816–1828, 2016.

[36] Ioannis Pisokas, Stanley Heinze, and Barbara Webb. The head direction circuit of two insect species. *Elife*, 9:e53985, 2020.

[37] Ran Darshan and Alexander Rivkind. Learning to represent continuous variables in heterogeneous neural networks. *Cell Reports*, 39(1), 2022.

[38] Tirthabir Biswas and James E Fitzgerald. Geometric framework to predict structure from function in neural networks. *Physical review research*, 4(2):023255, 2022.

[39] Tirthabir Biswas, Tianzhi Lambus Li, and James E Fitzgerald. Tensor formalism for predicting synaptic connections with ensemble modeling or optimization. *arXiv preprint arXiv:2310.20309*, 2023.

[40] Hod Dana, Yi Sun, Boaz Mohar, Brad K Hulse, Aaron M Kerlin, Jeremy P Hasseman, Getahun Tsegaye, Arthur Tsang, Allan Wong, Ronak Patel, et al. High-performance calcium sensors for imaging activity in neuronal populations and microcompartments. *Nature methods*, 16(7):649–657, 2019.

[41] Heather Dionne, Karen L Hibbard, Amanda Cavallaro, Jui-Chun Kao, and Gerald M Rubin. Genetic reagents for making split-gal4 lines in drosophila. *Genetics*, 209(1):31–35, 2018.

[42] Johannes D Seelig, M Eugenia Chiappe, Gus K Lott, Anirban Dutta, Jason E Osborne, Michael B Reiser, and Vivek Jayaraman. Two-photon calcium imaging from head-fixed drosophila during optomotor walking behavior. *Nature methods*, 7(7):535–540, 2010.

[43] Thomas A Pologruto, Bernardo L Sabatini, and Karel Svoboda. Scanimage: flexible software for operating laser scanning microscopes. *Biomedical engineering online*, 2:1–9, 2003.

[44] Philipp Berens. Circstat: a matlab toolbox for circular statistics. *Journal of statistical software*, 31:1–21, 2009.

Appendix

Contents

A Table **2**

B Continuous symmetric ring attractor networks **5**

B.1 The network & symmetry considerations 5

B.2 Eigenvalue constraints on the weights 7

B.3 Constraints from requiring single-bump activity profiles 9

B.4 Constraints from the non-positivity of inactive neurons 11

C Continuous *mirror-symmetric* ring attractor networks **12**

C.1 Constrains on the weights from eigenvalue requirements 12

C.2 Constraints from neuronal activity 14

C.3 Solution Space 15

C.4 Constraints from inactive neurons 16

D From the fly connectome to a ring attractor **17**

D.1 Effective neuronal network that includes indirect neuronal pathways . 17

D.1.1 Effective EPG network 17

D.1.2 Effective feedforward weights from EPG to other neuronal populations 19

D.2 Ring attractor constraints on effective weights 19

D.2.1 Constraints on effective EPG network 19

D.2.2 Constraints on effective feedforward weights from active EPG to other neurons 20

D.3 Scaling parameters connecting synapse counts to synapse weights . . 21

D.4 Finding scale factors that lead to ring attractors 22

E Equivalence between ring attractors with or without self couplings 23

F Symmetries of eigenvectors for (mirror) symmetric weight matrices 24

F.1 Single Eigenvector 25

F.2 Symmetry properties of double degenerate eigenspaces 25

F.3 Decomposition in terms of symmetric and antisymmetric vectors . . . 26

G Driven continuous attractors 27

G.1 General considerations for obtaining a continuum of steady-state profiles 27

G.2 Symmetric driven ring attractor networks relevant for fruit fly 29

G.3 External drive and decoding angular coordinate 30

H Velocity Integration 31

A Table

Appendix notations related to Activity	
Activity of compass neurons:	y^i, where $i = 1, \ldots, \mathcal{N}$
Activity of inhibitory neurons:	z^i, where $i = 1, \ldots, \mathcal{N}$
Activity of all, $\mathcal{A}_y$, active compass neurons:	$\vec{y}$, an $\mathcal{A}_y \leq \mathcal{N}$ dimensional vector
Activity of all, $\mathcal{A}_z$, active inhibitory neurons:	$\vec{z}$, an $\mathcal{A}_z \leq \mathcal{N}$ dimensional vector
Mirror partners:	i and $i^\star$, are mirror partners if $i + i^\star = \mathcal{N} + 1$, e.g., $\{3, 4, 5, 6\} = \{3, 4, 4^\star, 3^\star\}$
Encoded angle by the ith neuron:	$\theta_i = \frac{\pi}{\mathcal{N}} \left(i - \frac{\mathcal{N}+1}{2} \right)$, $i = 1, \ldots, \mathcal{N}$
Symmetric and antisymmetric eigenvectors, respectively, spanning the active compass neurons, $\{y_3, y_4, y_5, y_6\}$:	$\vec{y}_s = \begin{pmatrix} \psi_s \\ \chi_s \\ \chi_s \\ \psi_s \end{pmatrix}$, and $\vec{y}_a = \begin{pmatrix} \psi_a \\ \chi_a \\ -\chi_a \\ -\psi_a \end{pmatrix}$
(Anti-)symmetric eigenvector with eigenvalue one, and general eigenvectors	$\vec{\epsilon}_s, \vec{\epsilon}_a$, and $\vec{\epsilon}_\nu$, $\nu = 1, \ldots, 4$, respectively.
Ratios characterizing symmetric and antisymmetric eigenvectors.	$r_s = \frac{\chi_s}{\psi_s}$ and $r_a = \frac{\chi_a}{\psi_a}$ respectively.
Amplitude of activity profile:	$\sigma = \frac{\langle y_i \rangle}{4(1+r_s)} = \frac{\sum_{i=1}^{\mathcal{N}} y_i}{4(1+r_s)}$
Shape of activity profile:	$\mu = \frac{8\langle \theta_i \rangle}{\pi} = \frac{8 \sum_{i=1}^{\mathcal{N}} \theta_i y_i}{\pi}$
Steady state profiles:	$\vec{y} = \sigma(\vec{y}_s + \mu \vec{y}_a) = \sigma \begin{pmatrix} 1 - \mu \\ r_s - \mu r_a \\ r_s + \mu r_a \\ 1 + \mu \end{pmatrix}$
Two dimensional (anti-)symmetric vectors:	$\vec{y}_s^{2D} = \begin{pmatrix} \psi_s \\ \chi_s \end{pmatrix}$, and $\vec{y}_a^{2D} = \begin{pmatrix} \psi_a \\ \chi_a \end{pmatrix}$.

Appendix notations related to Weights	
8 dimensional weight matrix, 4 dimensional effective weight matrix, and 4 dimensional effective matrix without self-couplings	W, $\widetilde{W}$ and $\bar{W}$ respectively.
Weight matrix elements:	W_{ij}, where $i, j = 1, \ldots, \mathcal{N}$
Symmetric weight matrix without self-couplings:	$\bar{W} = \begin{pmatrix} 0 & w_1 & w_2 & w_3 \\ w_1 & 0 & w_1 & w_2 \\ w_2 & w_1 & 0 & w_1 \\ w_3 & w_2 & w_1 & 0 \end{pmatrix}.$
Mirror-symmetric weight matrix without self-couplings:	$\bar{W} = \begin{pmatrix} 0 & w_1' & w_2' & w_3 \\ w_1'' & 0 & w_1 & w_2'' \\ w_2'' & w_1 & 0 & w_1'' \\ w_3 & w_2' & w_1' & 0 \end{pmatrix}.$
Weight submatrices involving different (excitatory or inhibitory) pre/post-synaptic pairs:	W^{EE}, W^{EI}, W^{IE} and W^{II}.
Synapse-count submatrices involving different (excitatory or inhibitory) pre/post-synaptic pairs:	C^{EE}, C^{EI}, C^{IE} and C^{II}.
Scale-factors for different pre/post-synaptic pairs:	γ^{EE}, γ^{EI}, γ^{IE}, γ^{II}.
Scale-factor combination characterizing excitatory-inhibitory-excitatory loop:	$\widetilde{\gamma} = \sqrt{-\gamma^{EI}\gamma^{IE}}$.
Weight submatrices from active to active/inactive neurons:	$W_{\bullet\bullet}$, $W_{\circ\bullet}$, respectively.
Effective Weight matrices from active EPGs to (in)active EPGs and (in)active Δ7s:	$\widetilde{W}^{EE}_{\bullet\bullet}, \widetilde{W}^{EE}_{\circ\bullet}, \widetilde{W}^{IE}_{\bullet\bullet}$, and $\widetilde{W}^{IE}_{\circ\bullet}$.
8×4 dimensional effective Weight matrices without self-couplings from active EPGs to all EPGs and all Δ7s:	$\bar{W}^{EE}$ and $\widetilde{W}^{IE}$.
Effective weights from active EPGs to other neurons:	$\{u_3 = u_6^\star, u_4 = u_5^\star, u_5, u_6\}$
Eigenvalues for (anti) symmetric and eigenvectors:	$\lambda_s,\ \lambda_a$.

B Continuous symmetric ring attractor networks

In this appendix, we will consider symmetric recurrent networks and work out when and how they can behave as continuous ring attractors. Specifically, we will focus on a threshold-linear network with $\mathcal{N}$ neurons with all-to-all cross couplings but no self-coupling.

B.1 The network & symmetry considerations

Let us consider a recurrent threshold-linear neuronal network of an even number of $\mathcal{N}$ neurons (or computational units) whose firing rates or activities are denoted by y_i, with $i = 1, \ldots, \mathcal{N}$. We are going to choose the convention that y_i encodes the angle

$$\theta_i = \frac{\pi}{\mathcal{N}} \left(i - \frac{\mathcal{N}+1}{2} \right) \ . \tag{B.1}$$

This convention ensures that the encoded angles are distributed evenly around zero. Thus, for $\mathcal{N} = 8$, the eight neurons encode the angles, $\pm\pi/8, \pm 3\pi/8, \pm 5\pi/8, \pm 7\pi/8$. If the neurons attain a steady state, We will use a rate network model,

$$\dot{y}_i = -y_i + \Phi \left(\sum_{j=1}^{\mathcal{N}} W_{ij} y_j + s_i \right) \ , \tag{B.2}$$

to describe the evolution of the firing rate or neuronal activity where W_{ij} are the synaptic weights[1], s_i are drives that in principle could vary from neuron to neuron, and $\Phi(x) = \max(0, x)$ is the threshold-linear transfer function that transforms the inputs from the presynaptic neurons into the given neuron's input drive. We are primarily going to consider steady state activity profiles which must now satisfy

$$y_i = \Phi \left(\sum_{j=1}^{\mathcal{N}} W_{ij} y_j + s_i \right) \ , \tag{B.3}$$

In this appendix, we will consider self-sustaining networks that can support bump-like excitation in the absence of any external stimuli or positive bias. In Appendix G we consider the case with non-zero positive bias. Furthermore, we will impose discrete rotational symmetry, $W_{i+1,j+1} = W_{ij}$, and assume symmetric weights, $W_{ij} = W_{ji}$. In the subsequent appendix, we will consider more general matrices with reduced symmetry properties. Finally, to keep the algebra simple we will assume the self couplings to be zero, $W_{ii} = 0$. In Appendix E, we will show how a network with non-zero self-couplings can be mapped to a network without them, so this assumption does not constitute a loss of generality.

[1] We will later identify these weights to be effective weights that include indirect paths between the neurons via other neurons not included in this effective network.

In the main paper we argued that in order to have a ring attractor that self-supports (doesn't require an external drive) a continuum of steady-state profiles, the $\mathcal{A}_y \times \mathcal{A}_y$ dimenstional $\bar{W}$ submatrix involving the active neurons, must be doubly degenerate[2]. To understand the reason, let us focus on the active neurons, which we will label by indices, $\bar{i}, \bar{j}$, Since the drive to the active neurons must be positive, $\Phi \to 1$, and the activity of the active neurons must satisfy linear steady-state equations,

$$\sum_{\bar{j} \neq \bar{i}} W_{\bar{i}\bar{j}} y_{\bar{j}} = y_{\bar{i}}, \text{ or } \bar{W}\vec{y} = \vec{y} \ . \tag{B.4}$$

In other words, the $\mathcal{A}_y$-dimensional steady-state solution vector involving the active neurons, $\vec{y}$, is an eigenvector of $\bar{W}$ with eigenvalue one. For simplicity, we will assume that both $\mathcal{N}$ and $\mathcal{A}_y$ are even numbers and for notational convenience we will, henceforth, consider two indices, $i, i^\star$, mirror partners if

$$i + i^\star = \mathcal{N} + 1 \ . \tag{B.5}$$

According to this definition, the the weights obey the symmetry

$$\bar{W}_{i^\star j^\star} = \bar{W}_{ij} \ . \tag{B.6}$$

Now suppose that there exists only one such linearly independent eigenvector. Then, in Appendix F we prove that due to the symmetry (B.6), $\vec{y}$ must be either *symmetric*, $y_{i^\star} = y_i$, or *anti-symmetric*, $y_{i^\star} = -y_i$. Since the anti-symmetric pattern implies negative responses that are not allowed in a threshold-linear network, we are only left with a symmetric eigenvector. However, any symmetric profile will have an average angular value,

$$\langle \theta \rangle \equiv \sum_{i=1}^{\mathcal{N}} \theta_i y_i \ , \tag{B.7}$$

that corresponds to the central angle.The only flexibility the steady state profiles therefore have is in being able to change the bump amplitude, but this doesn't change the encoded angle. Accordingly, for every active set, and there are $\mathcal{N}$ of them, we can only encode a single angle. If $\mathcal{N}$ is large, this discrete encoding strategy may be able to provide adequate angular resolution, but it is insufficient to explain the accuracy of fruit fly's heading direction system with only eight computational units.

In order to have a continuous encoding of angles we need an eigenvector that is neither symmetric nor anti-symmetric, so that its mirror pair, defined by

$$y_i^\star \equiv y_{i^\star} \ , \tag{B.8}$$

is also another linearly independent eigenvector due to the symmetry properties of $\bar{W}$. The two eigenvectors together then will span a two-dimensional eigenspace of steady-state solutions. As we will see explicitly, different linear combinations of these eigenvectors can now indeed encode a continuum of average angular location.

[2]Strictly speaking $\bar{W} - \mathtt{I}$ is doubly degenrate.

B.2 Eigenvalue constraints on the weights

In Appendix F, we will show that the two dimensional eigenspace can also be described as a vector space spanned by a symmetric and anti-symmetric eigenvector. Thus to obtain all the steady state activity profiles, we will find the symmetric and antisymmetric eigenvectors separately, and obtain the constraints they impose on the weight-space along the way.

Henceforth, in this section we will specialize to the case, $\mathcal{N} = 8$ and $\mathcal{A}_y = 4$, which is relevant for the fruit fly ring attractor, although many of our techniques can be generalized to other network sizes and active set sizes. In our special case, the symmetric activity eigenvector can be represented as a four dimensional column vector,

$$\vec{y}_s = \begin{pmatrix} \psi_s \\ \chi_s \\ \chi_s \\ \psi_s \end{pmatrix} , \tag{B.9}$$

where for later convenience we will consider $i = 3, 4, 5, 6$ to be the active neurons. The sub-matrix, $\bar{W}$, relevant for the active set that respects all the symmetries must be of the form,

$$\bar{W} = \begin{pmatrix} 0 & w_1 & w_2 & w_3 \\ w_1 & 0 & w_1 & w_2 \\ w_2 & w_1 & 0 & w_1 \\ w_3 & w_2 & w_1 & 0 \end{pmatrix} , \tag{B.10}$$

containing three weight parameters, w_1, w_2m and w_3. Substituting the symmetric eigenvector ansatz (B.9) into the eigenvalue equation, $\bar{W}\vec{y}_s = \lambda_s \vec{y}_s$, we obtain a reduced eigenvalue equation (see Appendix F),

$$\begin{pmatrix} w_3 & w_1 + w_2 \\ w_1 + w_2 & w_1 \end{pmatrix} \begin{pmatrix} \psi_s \\ \chi_s \end{pmatrix} = \lambda_s \begin{pmatrix} \psi_s \\ \chi_s \end{pmatrix} , \tag{B.11}$$

because two of the four original equations are redundant. Imposing the *characteristic* equation, $\texttt{Det}(\bar{W} - \lambda \texttt{I}) = 0$, be satisfied for eigenvalue, $\lambda = 1$, we obtain a constraint equation on the weights:

$$(1 - w_3)(1 - w_1) = (w_1 + w_2)^2 . \tag{B.12}$$

To get a stable continuous attractor, one of the eigenvalues must be one and the other less than one. This implies that their sum[3] must be less than two, or

$$w_1 + w_3 < 2 . \tag{B.13}$$

[3]The sum of eigenvalues equals the trace of a matrix.

Thus in conjunction with (B.13), (B.12) ensures that we have stable symmetric steady-state solutions.

Now, one could also obtain the eigenvalues straightforwardly as

$$\lambda_{s\pm} = \frac{1}{2}\left[(w_3 + w_1) \pm \sqrt{(w_3 - w_1)^2 + 4(w_1 + w_2)^2}\right] \ . \tag{B.14}$$

and imposing $\lambda_{s+} = 1$, we can also obtain the required stable symmetric eigenvector. Technically though, this equation is more difficult to implement as compared to (B.13) and (B.12).

Let us next find the anti-symmetric steady-state solution. The procedure is very similar to the symmetric case, except we have to work with the anti-symmetric ansatz,

$$\vec{y}_a = \begin{pmatrix} \psi_a \\ \chi_a \\ -\chi_a \\ -\psi_a \end{pmatrix} \ . \tag{B.15}$$

The eigenvalue equation now reduces to (Appendix F),

$$\begin{pmatrix} -w_3 & w_1 - w_2 \\ w_1 - w_2 & -w_1 \end{pmatrix} \begin{pmatrix} \psi_a \\ \chi_a \end{pmatrix} = \lambda_a \begin{pmatrix} \psi_a \\ \chi_a \end{pmatrix} \ . \tag{B.16}$$

The two eigenvalues are

$$\lambda_{a\pm} = \frac{1}{2}\left[-(w_3 + w_1) \pm \sqrt{(w_3 - w_1)^2 + 4(w_1 - w_2)^2}\right] \ , \tag{B.17}$$

and to get a stable continuous attractor, we need $\lambda_{a+} = 1$ and $\lambda_{a+} + \lambda_{a-} < 2$. The first condition implies that the weights must satisfy

$$(1 + w_3)(1 + w_1) = (w_1 - w_2)^2 \ . \tag{B.18}$$

The second condition implies the inequality

$$w_1 + w_3 > -2 \ . \tag{B.19}$$

We note that the two inequalities, Eqns.(B.13) and (B.19) readily combines into

$$|w_1 + w_3| < 2 \ . \tag{B.20}$$

Finally, we must put the symmetric and anti-symmetric eigenvector conditions together. Eliminating w_3 from Eqs. (B.12, B.18), after some algebra we obtain a relation between w_1 and w_2:

$$(1 + w_2)(2w_1^2 + w_2 - 1) = 0 \ . \tag{B.21}$$

Thus either $w_2 = -1$, which as we shall see in the next subsection, leads to inconsistent activity profiles, or

$$w_2 = 1 - 2w_1^2 \ . \tag{B.22}$$

For the former solution branch, substituting $w_2 = -1$ in Eqs. (B.12 and B.18), we get

$$(1 + w_1)(w_1 - w_3) = 0 \ , \text{ and } (1 - w_1)(w_1 - w_3) = 0 \ . \tag{B.23}$$

The only consistent solution therefore implies $w_1 = w_3$.

For the other solution branch, substituting w_2 in either Eqs. (B.12 or B.18), one can obtain w_3 in terms of w_1:

$$w_3 \ = \ w_1(4w_1^2 - 3) \ . \tag{B.24}$$

We thus have a one-parameter family of symmetric connectivity matrices that give rise to continuous ring attractors. In this latter case, substituting w_3 in the inequalities (B.20) one can succinctly express the bounds w_1 must satisfy as

$$|4w_1^3 - 2w_1| < 2 \Rightarrow |w_1||2w_1^2 - 1| < 1 \ . \tag{B.25}$$

One can check that the above inequality is satisfied if and only if

$$|w_1| < 1 \ . \tag{B.26}$$

For the branch $w_2 = -1$, and $w_1 = w_3$, it is easy to see that we, in fact, end with the same inequality, (B.26).

B.3 Constraints from requiring single-bump activity profiles

By substituting $\lambda_s = 1$ and $\lambda_a = 1$ in the equations (B.11) and (B.16) respectively, we can explicitly obtain the corresponding symmetric and anti-symmetric eigenvectors. For instance, from second row of the eigenvalue equation, (B.11), We get,

$$(w_1 + w_2)\psi_s + w_1\chi_s = \chi_s \Rightarrow \ r_s \equiv \frac{\chi_s}{\psi_s} = \frac{w_1 + w_2}{1 - w_1} \ , \tag{B.27}$$

where we note that $w_1 \neq 1$ according to (B.26). Very similar manipulations involving the antisymmetric eigenvalue equation, (B.16), yields

$$r_a \equiv \frac{\chi_a}{\psi_a} = \frac{w_1 - w_2}{1 + w_1} \ , \tag{B.28}$$

Now, the two-dimensional space of steady-state solutions span the linear combinations of eigenvectors with eigenvalue one. Thus, by redefining $\psi_s = \sigma$ and $\psi_a = \mu\sigma$, one

can use σ and μ to parametrize the two-dimensional space of steady-state attractor profiles as

$$\vec{y} = \sigma \left[\begin{pmatrix} 1 \\ r_s \\ r_s \\ 1 \end{pmatrix} + \mu \begin{pmatrix} -1 \\ -r_a \\ r_a \\ 1 \end{pmatrix} \right] = \sigma(\vec{\epsilon}_s + \mu \vec{\epsilon}_a) \ , \tag{B.29}$$

where $\vec{\epsilon}_s, \vec{\epsilon}_a$ are symmetric and antisymmetric eigenvectors of $(W - \mathtt{I})$ with eigenvalue zero. Immediately we observe that to ensure non-negativity of all the responses,

$$\sigma > 0 \ , \text{ and } -1 \leq \mu \leq 1 \ . \tag{B.30}$$

We must also satisfy $r_s > 0$. Further, if we want to have a profile that has a single peak, we must have $r_a > 0$, otherwise one develops alternating peaks and troughs. In particular, we now notice that for the solution branch with $w_2 = -1$, $r_s = -1$ making the activity profile unphysical. This is the reason we did not consider this solution branch any further.

Continuing with the viable solutions given by (B.22) and (B.24), we find that the expressions for r_s and r_a simplify considerably:

$$r_s \equiv= 2w_1 + 1 \ , \text{ and } r_a \equiv \frac{\chi_a}{\psi_a} = 2w_1 - 1 \ , \tag{B.31}$$

and therefore,

$$r_s - r_a = 2 \ \Rightarrow r_s \geq r_a \ . \tag{B.32}$$

Thus one really just has to satisfy

$$r_a \geq 0 \Rightarrow w_1 \geq \frac{1}{2} \ . \tag{B.33}$$

To summarize, combining (B.26) and (B.33) we find that w_1 must be constrained to lie within the range

$$\frac{1}{2} \leq w_1 < 1 \ , \tag{B.34}$$

while w_2 and w_3 are given by (B.22) and (B.24).

It is instructive to compute $\langle \theta \rangle$ as a natural choice for decoding the angular coordinate, as one changes the steady state profiles. Using our conventions, for the active set, $i = 3, 4, 5, 6$, encoding angles $\{-\frac{3\pi}{8}, -\frac{\pi}{8}, \frac{\pi}{8}, \frac{3\pi}{8}\}$, respectively, we have

$$\begin{aligned} \langle \theta \rangle = \frac{\sum \theta_i y_i}{\sum y_i} &= \frac{\pi}{8} \left[\frac{-3(1-\mu) - (r_s - \mu r_a) + (r_s + \mu r_a) + 3(1+\mu)}{(1-\mu) + (r_s - \mu r_a) + (r_s + \mu r_a) + (1+\mu)} \right] \\ &= \mu \left(\frac{\pi}{8} \right) \left(\frac{3 + r_a}{1 + r_s} \right) = \mu \left(\frac{\pi}{8} \right) \ , \end{aligned} \tag{B.35}$$

where y_i values were substituted from (B.29), and in the last step we used the relation between the ratios (B.32). Thus, as μ varies in the interval, $[-1, 1]$, $\langle\theta\rangle$ covers an angle of $\pi/4$, as required to encode the entire angular range of 2π with 8 consecutive active sets.

B.4 Constraints from the non-positivity of inactive neurons

Finally, let us look at the constraints that one obtains while ensuring that the total input drive that the inactive neurons receive is negative, consistent with the ansatz for the steady state profiles. We do need to check that the consistency holds for all the steady state profiles parametrized by μ and σ. Due to the symmetric properties we only need to consider drives, d_i, onto the neurons, $i = 7, 8$. The inactivity of y_7 implies

$$d_7 = w_1 y_6 + w_2 y_5 + w_3 y_4 + w_4 y_3 \leq 0 \ . \tag{B.36}$$

Substituting the activity profile (B.29), one can rewrite the above inequality as

$$\begin{aligned} & (1+\mu)\left[w_1 + w_2\left(\frac{r_s + r_a}{2}\right) + w_3\left(\frac{r_s - r_a}{2}\right)\right] \\ + \quad & (1-\mu)\left[w_4 + w_2\left(\frac{r_s - r_a}{2}\right) + w_3\left(\frac{r_s + r_a}{2}\right)\right] \leq 0 \ . \end{aligned} \tag{B.37}$$

We see that the factors multiplying $(1 \pm \mu)$ must both be negative, otherwise at either $\mu = 1$ or -1, the inequality will be violated. So,

$$w_1 + w_2\left(\frac{r_s + r_a}{2}\right) + w_3\left(\frac{r_s - r_a}{2}\right) \leq 0 \ , \text{ and } w_4 + w_2\left(\frac{r_s - r_a}{2}\right) + w_3\left(\frac{r_s + r_a}{2}\right) \leq 0 \ . \tag{B.38}$$

Further, since $(1 \pm \mu)$ are always positive, the above condition guarantees that the inequality will be satisfied for all values of μ. The inequalities simplify once one uses the expressions for r_s, r_a:

$$C_1 \equiv w_1 + 2w_1 w_2 + w_3 \leq 0 \ , \text{ and } C_2 \equiv w_4 + w_2 + 2w_1 w_3 \leq 0 \ . \tag{B.39}$$

It is easy to check using (B.22, B.24) that the left hand side of the first inequality is identically zero and therefore the inequality doesn't lead to any new constraints. The second inequality can be re-expressed as a bound on w_4 in terms of w_1:

$$w_4 \leq -(8w_1^4 - 8w_1^2 + 1) \ . \tag{B.40}$$

One can perform calculations that are very similar to above to derive constraints from inactivity of y_8:

$$d_8 = w_2 y_6 + w_3 y_5 + w_4 y_4 + w_3 y_3 \leq 0 \ , \tag{B.41}$$

leads to the inequality

$$\begin{aligned} & (1+\mu)\left[w_2 + w_3\left(\frac{r_s+r_a}{2}\right) + w_4\left(\frac{r_s-r_a}{2}\right)\right] \\ + \quad & (1-\mu)\left[w_3 + w_3\left(\frac{r_s-r_a}{2}\right) + w_4\left(\frac{r_s+r_a}{2}\right)\right] \leq 0 \ . \end{aligned}$$

This inequality is satisfied iff,

$$w_2 + w_3\left(\frac{r_s+r_a}{2}\right) + w_4\left(\frac{r_s-r_a}{2}\right) \leq 0 \ , \text{ and } w_3 + w_3\left(\frac{r_s-r_a}{2}\right) + w_4\left(\frac{r_s+r_a}{2}\right) \leq 0 \ , \tag{B.42}$$

which can be simplified to yield

$$w_2 + 2w_1w_3 + w_4 \leq 0 \ , \text{ and } C_3 \equiv w_3 + w_1w_4 \leq 0 \ . \tag{B.43}$$

The first inequality is the same as was obtained earlier, while the second is a new constraint that gives rise to another bound for w_4:

$$w_4 \leq -(4w_1^2 - 3) \ . \tag{B.44}$$

To summarize, the ensemble of self-sustaining continuous ring attractors obeying the symmetries assumed is characterized by two parameters, w_1, and w_4, where the latter must satisfy the bounds (B.40, B.44), w_1 must lie in the range (B.34) and w_2, w_3 are given by (B.22, B.24). It is interesting to note that one could parametrize the weights in terms of an anlge, ϕ according to

$$w_1 = \cos\phi \ , w_2 = -\cos 2\phi \ , w_3 = \cos 3\phi \ , w_4 < \min(-\cos 4\phi, 1 - 2\cos 2\phi) \ , \tag{B.45}$$

where, $0 < \phi < \pi/3$.

C Continuous *mirror-symmetric* ring attractor networks

In this section we will generalize the analysis above to effective weight matrices with reduced symmetry properties that arise naturally from the connectomic features present in fruit flies.

C.1 Constrains on the weights from eigenvalue requirements

As we will discuss in appendix F, once the compass neurons whose activity encodes angular locations can drive each other indirectly via other neuronal populations, the

effective connectivity between the compass neurons need not obey all the symmetries assumed in the previous appendix. For instance, even if the entire connectivity structure preserves circular symmetry, the effective connectivity between the compass neurons need not preserve the symmetry due to the presence of the threshold nonlinearity. In appendix F, we will argue that for the fruit fly heading direction system a reduced *mirror-symmetric* effective weight matrix defined via,

$$W_{i^\star j^\star} = W_{ij} \ , \tag{C.46}$$

is what is appropriate to consider. For the active set we are considering ($i = 3, 4, 5, 6$) the mirror symmetric weight matrix is parameterized by six weight parameters, $\{w_1, w_1', w_1'', w_2', w_2'', w_3\}$:

$$\bar{W} = \begin{pmatrix} 0 & w_1' & w_2' & w_3 \\ w_1'' & 0 & w_1 & w_2'' \\ w_2'' & w_1 & 0 & w_1'' \\ w_3 & w_2' & w_1' & 0 \end{pmatrix} \ . \tag{C.47}$$

The arguments of the previous section, Appendix B, on the eigenvalue requirements of the active weight submatrix, goes through for the mirror-symmetric configuration as well, see appendix B, and thus $\bar{W}$ must still have two degenerate eigenvectors with eigenvalues of one. Moreover, we again need one of the eigenvectors to be symmetric and the other to be anti-symmetric. Substituting the symmetric eigenvector (B.9) in the eigenvalue equation (B.4), we again can obtain a reduced eigenvalue equation:

$$\begin{pmatrix} w_3 & w_1' + w_2' \\ w_1'' + w_2'' & w_1 \end{pmatrix} \begin{pmatrix} \psi_s \\ \chi_s \end{pmatrix} = \lambda_s \begin{pmatrix} \psi_s \\ \chi_s \end{pmatrix} \ . \tag{C.48}$$

The eigenvalues of a 2 by 2 matrix have closed form expressions:

$$\lambda_{s\pm} \equiv T \pm \sqrt{T^2 - D} \ , \text{ where} \tag{C.49}$$

$$D \equiv Det(\bar{W}) = w_3 w_1 - (w_1' + w_2')(w_1'' + w_2'')$$

$$T \equiv \frac{1}{2} Tr(\bar{W}) = \frac{1}{2}(w_3 + w_1) \ . \tag{C.50}$$

In order to have a stable attractor then, we must have

$$\lambda_{s+} = \frac{1}{2}\left[w_3 + w_1 + \sqrt{(w_3 - w_1)^2 + 4(w_1' + w_2')(w_1'' + w_2'')}\right] = 1 \tag{C.51}$$

Just like in the symmetric network, the above equation is equivalent to requiring,

$$(1 - w_3)(1 - w_1) = (w_1' + w_2')(w_1'' + w_2'') \ . \tag{C.52}$$

along with (B.13), $w_1 + w_3 < 2$.

Let us next determine the degenerate antisymmetric eigenvector,

$$\vec{y}_a = \begin{pmatrix} \psi_a \\ \chi_a \\ -\chi_a \\ -\psi_a \end{pmatrix} . \tag{C.53}$$

As with the symmetric case, the eigenvalue problem reduces to a two dimensional problem:

$$\begin{pmatrix} -w_3 & w_1' - w_2' \\ w_1'' - w_2'' & -w_1 \end{pmatrix} \begin{pmatrix} \psi_a \\ \chi_a \end{pmatrix} = \lambda_a \begin{pmatrix} \psi_a \\ \chi_a \end{pmatrix} . \tag{C.54}$$

The eigenvalues now read

$$\lambda_{a\pm} = \frac{1}{2}\left[-(w_3 + w_1) \pm \sqrt{(w_3 - w_1)^2 + 4(w_1' - w_2')(w_1'' - w_2'')}\right] , \tag{C.55}$$

or equivalently, the weights must satisfy the relation,

$$(1 + w_3)(1 + w_1) = (w_1' - w_2')(w_1'' - w_2'') , \tag{C.56}$$

along with the inequality (B.19), $w_1 + w_3 > -2$.

To summarize, to have a stable continuous attractor we must satisfy

$$(1 \pm w_3)(1 \pm w_1) = (w_1' \mp w_2')(w_1'' \mp w_2'') , \tag{C.57}$$

and

$$|w_1 + w_3| < 2 . \tag{C.58}$$

We point out that these equations indeed reproduce the symmetric model if we set, $w_1'' = w_1' = w_1$ and $w_2'' = w_2'$.

C.2 Constraints from neuronal activity

Exactly as in the symmetric model, the existence of two degenerate eigenvectors ensures that we have a two dimensional space of attractor profiles (B.29), The ratios r_s and r_a can be calculated similar to the symmetric case as well:

$$r_s \equiv \frac{\chi_s}{\psi_s} = \frac{1 - w_3}{w_1' + w_2'} = \frac{w_1'' + w_2''}{1 - w_1} , \text{ and} \tag{C.59}$$

$$r_a \equiv \frac{\chi_a}{\psi_a} = \frac{1 + w_3}{w_1' - w_2'} = \frac{w_1'' - w_2''}{1 + w_1} . \tag{C.60}$$

We first note that to ensure positivity of all the responses[4], we must satisfy

$$r_s = \frac{1 - w_3}{w_1' + w_2'} > 0 \ . \tag{C.61}$$

If we further demand that $\mu = 1$ should correspond to a symmetric configuration involving $i = 4, 5, 6$ which can then smoothly transition to the next set of active set, $i = 4, 5, 6, 7$, then we must have $y_4 = y_6$ at $\mu = 1$

$$\Rightarrow r_s - r_a = 2 \Rightarrow \frac{1 - w_3}{w_1' + w_2'} - \frac{1 + w_3}{w_1' - w_2'} = 2 \Rightarrow w_2'^2 - w_1^{2'} - w_2' - w_3 w_1' = 0 \ . \tag{C.62}$$

Finally, to be consistent with the observed bump profiles we will require that the bump has a single maximum in the middle, which implies

$$r_a = \frac{1 + w_3}{w_1' - w_2'} > 0 \ , \tag{C.63}$$

which immediately ensures the previous inequality (C.61), by virtue of (C.62).

C.3 Solution Space

To summarize, we have seen that a mirror-symmetric weight matrix has six independent weight paramaters characterizing the weight matrix involving the active set. There are however three equations, (B.12, B.18, C.62) that these parameters must satisfy to produce a smooth continuous ring attractor. Thus we expect a three dimensional solution space of ring attractors. In this section we will provide a convenient parametrization of the solution space that we frequently use for analysis and results.

The three parameters we will choose to parametrize the solutions are, w_1, w_3, and r_a. We note that the weights, w_1, w_3, are the only two weights that are symmetric, while r_a enables us to characterize the activity profiles in a uniform way that is valid for both the symmetric and mirror-symmetric models. Using the definitions of r_a and $r_s = r_a + 2$, one can straightforwardly obtain w_1', w_2' in terms of w_3 and r_a:

$$w_1' = \frac{1}{2}\left(\frac{1 - w_3}{r_s} + \frac{1 + w_3}{r_a}\right) \ , \text{ and } w_2' = \frac{1}{2}\left(\frac{1 - w_3}{r_s} - \frac{1 + w_3}{r_a}\right) \ . \tag{C.64}$$

The constraints from eigenvalue can then be use to express w_1'', w_2'' in terms of w_1 and r_a:

$$w_1'' = \frac{1}{2}\left[(1 - w_1) r_s + (1 + w_1) r_a\right] \ , \text{ and } w_2'' = \frac{1}{2}\left[(1 - w_1) r_s - (1 + w_1) r_a\right] \ . \tag{C.65}$$

The ranges of the three parameters are then restricted as

$$|w_1 + w_3| < 2 \ , \text{ and } r_a > 0 \ . \tag{C.66}$$

[4]By looking at the $\mu = 0$ case, for instance.

It is useful to point out a couple of key differences between the mirror symmetric and the symmetric case. Unlike the symmetric case, here w_1 and w_3 are free parameters, and thus the allowed range of w_1 can vary depending upon the value of w_3:

$$-2 - w_3 < w_1 < 2 - w_3 \ . \tag{C.67}$$

Unlike the symmetric case, w_1 and r_a are also independent in the general mirror-symmetric ring attractor. Thus while r_a inherrited a maximum bound of one from the bound of w_1, this is no longer true, and r_a, in principle, can have arbitrarily large values, it only needs to be positive (C.63). Thus the profiles in the mirror-symmetric attractor can deviate considerably from the symmetric case.

C.4 Constraints from inactive neurons

Finally, let us look at the constraints on the weights from the active to the inactive neurons as the net drive to each of the inactive neurons must be negative for all activity patterns. Note, that the synapses to the neurons that are inactive are completely independent of the recurrent weights between the active neurons. Thus if $\vec{u}$ denotes the synapse vector onto any of the inactive neurons, we must have

$$\vec{y} \cdot \vec{u} \leq 0 \ \forall \ \mu \ . \tag{C.68}$$

Substituting the activity profile (B.29), then we get

$$u_3(1-\mu) + u_4(r_s - \mu r_a) + u_5(r_s + \mu r_a) + u_6(1+\mu) \leq 0 \ . \tag{C.69}$$

As in the symmetric case, one can re-express the left hand side as a linear combination of two terms involving prefactors $(1 \pm \mu)$, so that the above inequality is satisfied $\forall \ \mu$ provided the coefficients corresponding to $(1 \pm \mu)$ are negative:

$$2u_3 + r_s(u_5 + u_4) + r_a(u_4 - u_5) < 0 \ , \text{ and } 2u_6 + r_s(u_5 + u_4) - r_a(u_4 - u_5) < 0 \ . \tag{C.70}$$

So, there are four independent weights that only have to satisfy two inequalities, which can be further simplified to

$$2u_3 + 2u_5 + (r_s + r_a)u_4 < 0 \ , \text{ and } 2u_6 + 2u_4 + (r_s + r_a)u_5 < 0 \ . \tag{C.71}$$

Thus there is a lot of flexibility, and a given weight can take on arbitrary values as long as other weights are accordingly adjusted.

While (C.71) provides an analytical inequality condition for a viable ring attractor network, there is a simple numerical way to check that these inequalities are satisfied. As we just argued, in order for (C.69) to be satisfied for all the profiles, all we have to check is that these conditions are met for the two extreme configurations, $\mu = \pm 1$.

D From the fly connectome to a ring attractor

In this appendix we will consider a neural network that not only contains the compass neurons, but also other neuronal populations that provide indirect pathways between the compass neurons. Our reason for considering such a scenario is the well known fact that while the EPG's can directly provide local excitatory drives to each other, the broader distal mutual inhibition has to come indirectly via other neuronal populations. Here we will consider a single neuronal population that provides indirect pathways between the EPG neurons. The appendix is organized as follows: First we will discuss how one can map such a network to an effective neuronal network model involving recurrently connected compass neurons that also provide feedforward inputs to the other neuronal populations. It relies on the ability to absorb all the indirect pathways involving inhibitory neuronal populations into effective weights between the active compass neurons and from the active compass neurons to all other neurons. Next, we will derive all the constraints that the effective weights must satisfy in order to have a consistent ring attractor network. Since the effective weights are functions of the original network weights, the various equalities and inequalities we will derive in effect constrains the original network. Next, using the measured synapse count matrices involving the different relevant neuronal populations one can obtain a family of realizable neuronal networks that are parametrized by a few scaling parameters that converts synapse counts to weights. If there exists an intersection between the theoretical ensemble of continuous ring attractors and the realizable networks from the connectome, then would have discovered connectome-constrained continuous ring attractor networks. In the process we will be able to specify how one should adjust the scaling parameters to map the connectome constrained neuronal network into the viable ring attractor networks.

D.1 Effective neuronal network that includes indirect neuronal pathways

D.1.1 Effective EPG network

Recurrent weights between active neurons: Let us consider an additional set of recurrent neurons, z_i's for e.g., that provide indirect inhibitory pathways between the EPG neurons for supporting localized bump of activities. For simplicity we will assume $i = 1, \ldots, \mathcal{N}$. The evolution equations are then modeled as

$$\tau_y^{-1}\dot{y} = -y + \Phi(W^{EE}y + W^{EI}z) \ , \text{ and } \tau_z^{-1}\dot{z} = -z + \Phi(W^{IE}y + W^{II}z) \ , \quad \text{(D.72)}$$

where τ_z, τ_y are the neuronal time constants, and W^{EE}, W^{EI}, W^{IE} and W^{II} are all $\mathcal{N}\times\mathcal{N}$ weight matrices. Now suppose $\mathcal{A}_z$ is the set of active z neurons when a given set of $\mathcal{A}_y$ compass neurons are active. We are going to assume that this set doesn't change depending upon the profile parameters, μ, σ, characterizing the neuronal activity

profile. In this case, the steady state equations for the active neuronal set is given by

$$y = W^{EE}_{\bullet\bullet} y + W^{EI}_{\bullet\bullet} z \ , \text{ and } z = W^{IE}_{\bullet\bullet} y + W^{II}_{\bullet\bullet} z \ , \tag{D.73}$$

where $W^{EE}_{\bullet\bullet}, W^{EI}_{\bullet\bullet}, W^{IE}_{\bullet\bullet}$ and $W^{II}_{\bullet\bullet}$ are respectively $\mathcal{A}_y \times \mathcal{A}_y, \mathcal{A}_y \times \mathcal{A}_z, \mathcal{A}_z \times \mathcal{A}_y$, and $\mathcal{A}_z \times \mathcal{A}_z$ dimensional submatrices of W^{EE}, W^{EI}, W^{IE} and W^{II} respectively relevant for the active neurons. Hence forth we will use $\bullet, \circ$, to indicate the activity state, on or off respectively, of the pre/post synaptic neurons.

The second equation can be solved[5] to yield z in terms of y:

$$z = \left(\mathtt{I} - W^{II}_{\bullet\bullet}\right)^{-1} W^{IE}_{\bullet\bullet} y \ . \tag{D.74}$$

Substituting z in the first equation, we get

$$y = \tilde{W}^{EE}_{\bullet\bullet} y \ , \text{ where } \tilde{W}^{EE}_{\bullet\bullet} \equiv W^{EE}_{\bullet\bullet} + W^{EI}_{\bullet\bullet} \left(\mathtt{I} - W^{II}_{\bullet\bullet}\right)^{-1} W^{IE}_{\bullet\bullet} \ . \tag{D.75}$$

The *effective* weight matrix, $\tilde{W}^{EE}_{\bullet\bullet}$, not only captures the recurrent direct connections between the EPG neurons (first term) but also indirect pathways between the EPG neurons via z neurons[6].

Feedforward weights from active to inactive EPG's: While the requirement to be a consistent continuous ring attractor is most stringent on the effective recurrent couplings between the active EPG neurons (see Appendices B and C), one also needs to check that the assumed neuronal states (active or inactive) of all other neurons in the network that are consistent with the input drives that these neurons receive. For e.g., for the network under study we need to check that the inactive EPG's, the active z's, and the inactive z's, receive negative, positive and negative input drives respectively. Such consistency checks can be performed most easily by deriving effective feedforward couplings between the active EPG's and the other neuronal populations in question. Let us first derive the effective couplings between the active and inactive EPG's.

The total drive, D^e's that the inactive EPG neurons receive can be written as

$$D^y = W^{EE}_{\circ\bullet} y + W^{EI}_{\circ\bullet} z = [W^{EE}_{\circ\bullet} + W^{EI}_{\circ\bullet} (\mathtt{I} - W^{II}_{\bullet\bullet})^{-1} W^{IE}_{\bullet\bullet}] y \ , \tag{D.76}$$

where $W^{EE}_{\circ\bullet}$ and $W^{EI}_{\circ\bullet}$ are $(\mathcal{N} - \mathcal{A}_y) \times \mathcal{A}_y$ and $(\mathcal{N} - \mathcal{A}_y) \times \mathcal{A}_z$ dimensional submatrices of W^{EE} and W^{EI} that connect the active EPG and z neurons respectively to the inactive EPG's. In other words, the effective feedforward off-diagonal $(\mathcal{N} - \mathcal{A}_y) \times \mathcal{A}_y$ dimensional matrix connecting the active to the inactive sector of EPG's is given by

$$\tilde{W}^{EE}_{\circ\bullet} \equiv W^{EE}_{\circ\bullet} + W^{EI}_{\circ\bullet} (\mathtt{I} - W^{II}_{\bullet\bullet})^{-1} W^{IE}_{\bullet\bullet} \ . \tag{D.77}$$

We note in passing that we will use $\tilde{}$ to denote effective matrices that includes the indirect pathways.

[5] Assuming $(\mathtt{I} - W^{II}_{\bullet\bullet})$ is nonsingular.

[6] Expanding $\left(\mathtt{I} - W^{II}_{\bullet\bullet}\right)^{-1}$ as a Taylor series explitly illustrates the contribution from different paths involving the inhibitory populations

D.1.2 Effective feedforward weights from EPG to other neuronal populations

Since the derivation of the effective recurrent weights between the EPG neurons required us to posit the states of other neuronal populations, we need to check that the sign of the drives to these neurons are consistent with our assumptions. Here we are going to continue to focus on a single z population, but our results can easily be generalized to incorporate more than one neuronal population. The drives, D^z, that the active z neurons receive are given by

$$D^z = W^{IE}_{\bullet\bullet} y + W^{II}_{\bullet\bullet} z = [W^{IE}_{\bullet\bullet} + W^{II}_{\bullet\bullet}(\mathtt{I} - W^{II}_{\bullet\bullet})^{-1} W^{IE}_{\bullet\bullet}] y \ . \tag{D.78}$$

Accordingly, the $\mathcal{A}_z \times \mathcal{A}_y$ dimensional effective feedforward weight matrix connective the active EPG's with active z's are given by

$$\tilde{W}^{IE}_{\bullet\bullet} \equiv W^{IE}_{\bullet\bullet} + W^{II}_{\bullet\bullet}(\mathtt{I} - W^{II}_{\bullet\bullet})^{-1} W^{IE}_{\bullet\bullet} \ . \tag{D.79}$$

In a very similar manner, one can also derive effective weight matrix connecting the active EPG's to the inactive z neurons as,

$$\tilde{W}^{IE}_{\circ\bullet} \equiv W^{IE}_{\circ\bullet} + W^{II}_{\circ\bullet}(\mathtt{I} - W^{II}_{\bullet\bullet})^{-1} W^{IE}_{\bullet\bullet} \ , \tag{D.80}$$

where $W^{IE}_{\circ\bullet}$ and $W^{II}_{\circ\bullet}$ are $(\mathcal{N} - \mathcal{A}_z) \times \mathcal{A}_y$ and $(\mathcal{N} - \mathcal{A}_z) \times \mathcal{A}_z$ dimensional submatrices of W^{IE} and W^{II} that connect the active EPG and z neurons respectively to the inactive z neurons.

D.2 Ring attractor constraints on effective weights

To summarize, in the last section we provided a prescription to obtain effective weight matrices, $\tilde{W}^{EE}_{\bullet\bullet}$, $\tilde{W}^{EE}_{\circ\bullet}$, $\tilde{W}^{IE}_{\bullet\bullet}$, and $\tilde{W}^{IE}_{\circ\bullet}$, which represent recurrent connectivity between active EPG neurons (y) and feedforward connectivity from the active EPGs to inactive EPGs, active z's and inactive z's, respectively. We are now going to enumerate the constraints these matrix elements must satisfy to have continuous attractor.

D.2.1 Constraints on effective EPG network

We observe that the $\mathcal{A}_y \times \mathcal{A}_y$ and $(\mathcal{N} - \mathcal{A}_y) \times \mathcal{A}_y$ dimensional effective weight matrices, $W^{EE}_{\bullet\bullet}$ and $W^{EE}_{\circ\bullet}$ (D.75 and D.77) respectively can naturally be combined into an $\mathcal{N} \times \mathcal{A}_y$ effective matrix connecting inputs from active EPG's with all the EPG's. As discussed in Appendix E, any recurrent network with self-couplings is equivalent to a recurrent network without self-couplings as far as the steady state solutions are concerned. According to (E.95) the equivalent weights, $\bar{W}^{EE}_{ij}$, for our effective EPG

network without self couplings is given by,

$$\begin{aligned}\bar{W}^{EE}_{ij} &\equiv \frac{\tilde{W}^{EE}_{ij}}{1-\tilde{W}^{EE}_{ii}} \ ,\forall\ i\neq j\ ,\ i\in\mathcal{N}\ \&\ j\in\mathcal{A}_y \\ &= \frac{W^{EE}_{ij}+\sum_{k\in\mathcal{A}_z,l\in\mathcal{A}_y} W^{EI}_{ik}\left(\mathtt{I}-W^{II}\right)^{-1}_{kl} W^{IE}_{lj}}{1-W^{EE}_{ii}-\sum_{k\in\mathcal{A}_z,l\in\mathcal{A}_y} W^{EI}_{ik}\left(\mathtt{I}-W^{II}\right)^{-1}_{kl} W^{IE}_{li}}\ . \end{aligned} \tag{D.81}$$

One can now read-off the effective independent weight-parameters, w, u's, from their definitions in the Appendices B and C and impose the conditions that they must satisfy to realize a continuous ring attractor. These constraints are summarized in Appendix D.

For the symmetric model the constraints are given by

$$\begin{aligned}(1\pm\bar{w}_3)(1\pm\bar{w}_1) = (\bar{w}_1\mp\bar{w}_2)^2 \quad &, \quad \tfrac{1}{2}<\bar{w}_1<1\ , \\ \bar{w}_2+2\bar{w}_1\bar{w}_3+\bar{w}_4\leq 0 \quad &, \quad \text{and } \bar{w}_3+\bar{w}_1\bar{w}_4\leq 0\ . \end{aligned} \tag{D.82}$$

where the "bar" indicates that we are referring to the effective weights derived from (D.81). For the mirror-symmetric model the relevant equations for the effective recurrent weights read,

$$\begin{aligned}(1\pm\bar{w}_3)(1\pm\bar{w}_1) = (\bar{w}'_1\mp\bar{w}'_2)(\bar{w}''_1\mp\bar{w}''_2) \quad &, \quad \bar{r}_s-\bar{r}_a=2 \\ -2-\bar{w}_3<\bar{w}_1<2-\bar{w}_3 \quad &, \quad \text{and } \bar{r}_a>0\ , \end{aligned} \tag{D.83}$$

where we have defined

$$\bar{r}_s = \frac{1-\bar{w}_3}{\bar{w}'_1+\bar{w}'_2}\ ,\ \text{and } \bar{r}_a = \frac{1+\bar{w}_3}{\bar{w}'_1-\bar{w}'_2}\ . \tag{D.84}$$

Unlike the symmetric model, in the mirror symmetric model the feedforward weights from active to inactive EPG's are not related to the above recurrent effective weights. For every inactive neuron, the effective weights onto it must satisfy

$$2\bar{u}_3+2\bar{u}_4+(\bar{r}_s+\bar{r}_a)\bar{u}_5<0\ ,\ \text{and } 2\bar{u}_6+2\bar{u}_4+(\bar{r}_s+\bar{r}_a)\bar{u}_5<0\ , \tag{D.85}$$

where $\bar{u}_3, \bar{u}_4, \bar{u}_5, \bar{u}_6$ refers to the effective weights from the active EPG neurons, y^3, y^4, y^5, y^6 to the inactive EPG neuron given by (D.77).

D.2.2 Constraints on effective feedforward weights from active EPG to other neurons

The inequality constraints on weights from active to inactive EPG neurons were derived by ensuring that the effective feedforward drive the inactive neurons receive from the active EPG neurons, is negative for all the activity profiles ($-1\leq\mu\leq 1$). Since we have derived the effective feedforward weights from the active EPG's to the z neurons, the constraints coming from our assumptions about the activity states

(active or inactive) of z neurons can be derived following the same mathematical path as the derivation for the inactive EPG neurons. Explicitly, in the context of the more general mirror-symmetric model the effective weights onto a given neuron must satisfy (C.71):

$$2\bar{u}_3 + 2\bar{u}_5 + (\bar{r}_s + \bar{r}_a)\bar{u}_4 < 0 \ , \quad \text{and} \quad 2\bar{u}_6 + 2\bar{u}_4 + (\bar{r}_s + \bar{r}_a)\bar{u}_5 < 0, \quad \text{(D.86)}$$

$$\text{or, } 2\bar{u}_3 + 2\bar{u}_5 + (\bar{r}_s + \bar{r}_a)\bar{u}_4 > 0 \ , \quad \text{and} \quad 2\bar{u}_6 + 2\bar{u}_4 + (\bar{r}_s + \bar{r}_a)\bar{u}_5 > 0, \quad \text{(D.87)}$$

depending on whether the z was assumed to be inactive or active respectively. Here $\bar{u}_3, \bar{u}_4, \bar{u}_1, \bar{u}_6$ refers to the effective feedforward weights (D.79 and D.80) from the active EPG neurons, y^3, y^4, y^5, y^6 to the z neuron, respectively.

D.3 Scaling parameters connecting synapse counts to synapse weights

In the previous subsections we have obtained the constraints that the synaptic weights must satisfy to become a continuous ring attractor. The connectome, however, does not directly measure the various weights involved but rather only provides information about the number of synapses between different neurons. We are going to assume that for a given type of pre- and post-synaptic neuronal pair, there exists an independent scaling parameter that converts the synapse counts into synaptic weights proportionately. Since there are four different neuronal pairs, we have four different scaling parameters:

$$W^{EE} = \gamma^{EE} C^{EE}, \ W^{EI} = \gamma^{EI} C^{EI} \ , W^{IE} = \gamma^{IE} C^{IE}, \ W^{II} = \gamma^{II} C^{II} \ , \qquad \text{(D.88)}$$

where C's are synapse count matrices that will be inferred from data, and γ's are the proportionality constants between the synapse count and synaptic weights.

We can now express all the effective matrices as functions of the scaling factors. Explicitly, the effective weight matrices between the active EPG neurons are given by

$$\begin{aligned}
\tilde{W}^{EE}_{\bullet\bullet} &= \gamma^{EE} C^{EE}_{\bullet\bullet} + \gamma^{EI}\gamma^{IE} C^{EI}_{\bullet\bullet} \left(1 - \gamma^{II} C^{II}_{\bullet\bullet}\right)^{-1} C^{IE}_{\bullet\bullet} \ , \\
\tilde{W}^{EE}_{\circ\bullet} &= \gamma^{EE} C^{EE}_{\circ\bullet} + \gamma^{EI}\gamma^{IE} C^{EI}_{\circ\bullet} \left(1 - \gamma^{II} C^{II}_{\bullet\bullet}\right)^{-1} C^{IE}_{\bullet\bullet} \ , \\
\tilde{W}^{IE}_{\bullet\bullet} &= \gamma^{IE} C^{IE}_{\bullet\bullet} + \gamma^{II}\gamma^{IE} C^{II}_{\bullet\bullet} \left(1 - \gamma^{II} C^{II}_{\bullet\bullet}\right)^{-1} C^{IE}_{\bullet\bullet} \ , \\
\tilde{W}^{IE}_{\circ\bullet} &= \gamma^{IE} C^{IE}_{\circ\bullet} + \gamma^{II}\gamma^{IE} C^{II}_{\circ\bullet} \left(1 - \gamma^{II} C^{II}_{\bullet\bullet}\right)^{-1} C^{IE}_{\bullet\bullet} \ ,
\end{aligned} \qquad \text{(D.89)}$$

where the $C_{\bullet\bullet}$ and $C_{\circ\bullet}$ matrices represent the synapse count submatrices involving the active neurons, and active to inactive neurons.

We note that this matrix includes diagonal elements representing self-couplings. According to (D.81) the equivalent weights without self couplings that couple the

EPG neurons are given by

$$\bar{W}^{EE}_{ij} = \frac{\gamma^{EE} C^{EE}_{ij} + \sum_{k\in\mathcal{A}_z, l\in\mathcal{A}_y} (\gamma^{EI}\gamma^{IE}) C^{EI}_{ik} \left(1-\gamma^{II}C^{II}\right)^{-1}_{kl} C^{IE}_{lj}}{1-\gamma^{EE}C^{EE}_{ii} - \sum_{k\in\mathcal{A}_z, l\in\mathcal{A}_y} (\gamma^{EI}\gamma^{IE}) C^{EI}_{ik} \left(1-\gamma^{II}C^{II}\right)^{-1}_{kl} C^{IE}_{li}} \ , \tag{D.90}$$

$\forall\ i \in \mathcal{N} \neq j \in \mathcal{A}_y$. In a similar manner one can define an effective matrix that combines weights from the active EPG to the active and inactive z's, *i.e.* , the matrices $\tilde{W}^{IE}_{\circ\bullet}$ and $\tilde{W}^{IE}_{\bullet\bullet}$ into another $\mathcal{N}\times\mathcal{A}_y$ matrix, $\tilde{W}^{IE}$:

$$\tilde{W}^{IE}_{ij} = \gamma^{IE} C^{EE}_{ij} + \sum_{k\in\mathcal{A}_z, l\in\mathcal{A}_y} \gamma^{II}\gamma^{IE} C^{II}_{ik} \left(1-\gamma^{II}C^{II}\right)^{-1}_{kl} C^{IE}_{lj} \ , \forall\ i\in\mathcal{N}\neq j\in\mathcal{A}_y \ . \tag{D.91}$$

We note that since only the signs of the drives matter, rescaling the matrix elements to eliminate self couplings are unnecessary.

To summarize, all the effective weights, $\bar{w}, \bar{u}$'s become functions of three[7] scaling parameters:

$$\gamma^{EE} \ , \gamma^{II} \ , \text{ and } \gamma' \equiv \sqrt{-\gamma^{EI}\gamma^{IE}} \ , \tag{D.92}$$

according to (D.90). As we have seen, to be a continuous ring attractor the effective synaptic weights must satisfy several equalities and inequalities. This in turn means that the above three scale factors have to satisfy several equalities and inequalities to map the connectome into a plausible ring attractor.

D.4 Finding scale factors that lead to ring attractors

To be a ring attractor the synaptic weights must satisfy several equalities and inequalities. Most of these constraints originate in the requirement of the effective active submatrix to be able to encode a continuous angular coordinate smoothly, consistently and robustly. We will first schematically enumerate these constraints in a tabular form referring to the appropriate equation derived earlier. Next, we will discuss constraints on the effective feedforward weights from the active to all other neurons that comes from requiring the signs of the drives to these neurons be consistent with the assumptions we made about their activity state.

Thus we start by enumerating all the constraints on effective recurrent couplings between active EPG's.

[7]We point out that although in (D.91) γ^{IE} appears individually, it can be factored out and therefore its value doesn't affect the inequality constraints.

Constraint	Symmetric	Mirror-symmetric
Continuous encoding	(B.12) & (B.18)	(C.52) & (C.56)
Smooth transition	None	(C.62)
Stability	(B.26)	(C.58)
Profile requirements	(B.33)	(C.63)

Table 1: References to the equations that define the ring attractor conditions

The inequality constraints that limits the ranges of weights from the active EPG's to all the other neurons are tabulated as

Constraint	Symmetric	Mirror-symmetric
Inactive EPG states	(B.43)	(C.71)
Active z states	(D.87)	(D.87)
Inactive z states	(D.86)	(D.86)

Table 2: References to the inequalities that limit the parameter space of valid solutions

E Equivalence between ring attractors with or without self couplings

In this appendix we will show how a steady state solution in a recurrent network can be mapped to a steady-state solution in a recurrent network with no self-couplings, and vice-versa. This will enable us to work with the slightly simpler networks of the latter kind without loss of generality. To see this let us start with the steady state equation:

$$y_i = \Phi\left(\sum_{j=1}^{\mathcal{N}} W_{ij} y_j + s_i\right) \ . \tag{E.93}$$

If, $i = 1, \ldots, \mathcal{A}_y$, denotes the active neurons, their activity must satisfy linear steady-state equations,

$$y_i = \sum_{j=1}^{\mathcal{N}} W_{ij} y_j + s_i \Rightarrow y_i = \sum_{j \neq i} \bar{W}_{ij} y_j + \bar{s}_i \ , \tag{E.94}$$

where we have defined new *equivalent* weights and drives as

$$\bar{W}_{ij} \equiv \frac{W_{ij}}{1 - W_{ii}} \ , \text{ and } \bar{s}_i \equiv \frac{s_i}{1 - W_{ii}} \ , \tag{E.95}$$

assuming $W_{ii} \neq 1$, $\forall\, i$. This means that finding steady state solutions in the original network is equivalent to finding solutions in a different simplified network that have no self-couplings.

To find solutions to the nonlinear equations (B.3) one also must satisfy the inequalities,

$$y_i = \sum_{j=1}^{\mathcal{N}} W_{ij} y_j + s_i < 0 \ , \tag{E.96}$$

for all y_i that are inactive. As long as $W_{ii} \neq 1$ though, the inequalities can be straightforwardly written as

$$y_i = \frac{1}{1 - W_{ii}} \left(\sum_{j \neq i}^{\mathcal{N}} W_{ij} y_j + s_i \right) = \sum_{j \neq i}^{\mathcal{N}} \bar{W}_{ij} y_j + \bar{s}_i < 0 \ . \tag{E.97}$$

To summarize, finding a steady state solution in an arbitrary network can be reduced to a problem of finding a solution in a network with no self-couplings and all other weights and external drives rescaled according to (E.95).

Let us now characterize the ensemble of ring attractors in a network with self-couplings from the knowledge of ring attractors in *equivalent* networks with no self-couplings. Suppose, we have already found weight matrices without self couplings and associated bump profiles that produces continuous attractors. In other words, we have $\bar{W}_{ij}(\{w_a\})$, and associated $y_i(\mu, \sigma)$ that satisfies the eigenvalue equation. From (E.95) then all networks with self-couplings, W_{ii}, that correspond to weights $\bar{W}_{ij}(\{w_a\})$ given by (E.95) in the equivalent network without self-couplings is a valid ring attractor network as they would support the same continuous steady state solutions. Specifically, the ensemble of weight matrices are given by

$$W_{ij}(\{w_a\}, \{W_{ii}\}) = \begin{cases} (1 - W_{ii}) \bar{W}_{ij}(\{w_a\}) & \text{if } i \neq j \\ W_{ii} & \text{if } i = j \end{cases} \ . \tag{E.98}$$

We now have an extra parameters in W_{ii} that characterizes the weight matrices. The activity profiles remain unchanged.

F Symmetries of eigenvectors for (mirror) symmetric weight matrices

In this appendix we will prove a set of results related to the symmetry of the relevant eigenvectors that will help us characterize the ensemble of continuous ring attractors and the continuum of steady states it supports.

F.1 Single Eigenvector

Consider an eigenvector of W with eigenvalue λ,

$$\sum_j W_{ij} y_j = \lambda y_j \ . \tag{F.99}$$

We claim that if there is only one eigenvector with eigenvalue λ, and W is a mirror symmetric matrix, then the eigenvector must be symmetric or anti-symmetric.

We remind readers that a mirror symmetric matrix satisfies

$$W_{i^\star j^\star} = W_{ij} \ . \tag{F.100}$$

Let us now construct $\vec{y}$'s mirror partner, $\vec{y}^\star$, as

$$y_i^\star = y_{i^\star} \ . \tag{F.101}$$

Then

$$\sum_j W_{ij} y_j^\star = \sum_j W_{ij} y_{j^\star} = \sum_j w_{i^\star j^\star} y_{j^\star} = \lambda y_{i^\star} = \lambda y_i^\star \ , \tag{F.102}$$

where the four equalities follow respectively from (F.101), (F.100), (F.99) and (F.101). In other words, $y^\star$ is also an eigenvector with eigenvalue λ. y and $y^\star$ must therefore be proportional to each other. Further they clearly have the same norm implying either

$$\vec{y}^\star = \vec{y} \text{ or } \vec{y}^\star = -\vec{y} \ , \tag{F.103}$$

i.e. , symmetric or anti-symmetric respectively. Finally, we note that the pure anti-symmetric eigenstate is ruled out because it contains negative activities.

F.2 Symmetry properties of double degenerate eigenspaces

Let us next consider a two dimensional eigenspace associated with an eigenvalue λ. If $\vec{y}$ is an eigenvector which is neither symmetric nor anti-symmetric, then one can define two new eigenvectors,

$$\vec{y}_\pm \equiv \vec{y} \pm \vec{y}^\star \ . \tag{F.104}$$

By construction we now have an eigenbasis where one eigenvector, $\vec{y}_+$ is symmetric, and the other is $\vec{y}_-$ is antisymmetric. In other words, a doubly degenerate eigenspace can be decomposed into a symmetric and an antisymmetric eigenspace.

What if the entire double degenerate eigenspace is symmetric? In this case the average encoded position is always at zero, and such a scenario cannot lead to continuous angular encoding, and therefore will not be considered in this paper. A completely anti-symmetric eigenspace, on the other hand, is disallowed because negative activities are not allowed. Thus the symmetric and antisymmetric mixture is the only viable eigenspace that can lead to continuous angular encoding.

F.3 Decomposition in terms of symmetric and antisymmetric vectors

In this section we show that the dynamic and the steady state equations naturally decompose into two equations one for the symmetric and the other for the anti-symmetric parts of the activity. First we note that one can decompose any activity vector into a symmetric and an anti-symmetric piece:

$$\vec{y} = \vec{y}_s + \vec{y}_a = \begin{pmatrix} \psi_s \\ \chi_s \\ \chi_s \\ \psi_s \end{pmatrix} + \begin{pmatrix} \psi_a \\ \chi_a \\ -\chi_a \\ -\psi_a \end{pmatrix} . \tag{F.105}$$

The entire space is spanned by two symmetric basis vectors, $\vec{\epsilon}_{s1}, \vec{\epsilon}_{s2}$, and two anti-symmetric basis vectors, $\vec{\epsilon}_{a1}, \vec{\epsilon}_{a2}$. Moreover, the symmetric and the antisymmetric spaces are orthogonal to each other, and one can find the symmetric and antisymmetric projection of any arbitrary vector via

$$\vec{y}_s = \hat{P}_s(\vec{y}) = \frac{1}{2}(\vec{y} + \vec{y}^\star) \ , \text{ and } \vec{y}_a = \hat{P}_a\vec{y} = \frac{1}{2}(\vec{y} - \vec{y}^\star) \ , \tag{F.106}$$

where $\hat{P}_s$ and $\hat{P}_a$ are the symmetric and anti-symmetric projection operators respectively.

Let us now look at the time evolution equation

$$\dot{\vec{y}} = -\vec{y} + W\vec{y} + \vec{b} \ , \tag{F.107}$$

where $\vec{b}$ is a constant external drive. In the driven ring attractor network (Appendix G) we will consider a symmetric $\vec{b}$, while for velocity integration (Appendix H) an antisymmetric $\vec{b}$ will become relevant. Using the projection operations we can now get two equations, one for the symmetric and the other for the antisymmetric part:

$$\begin{aligned} \dot{\vec{y}}_s &= -\vec{y}_s + \hat{P}_s(\bar{W}\vec{y}) + \vec{b}_s \ , \\ \dot{\vec{y}}_a &= -\vec{y}_a + \hat{P}_a(\bar{W}\vec{y}) + \vec{b}_a \ , \end{aligned} \tag{F.108}$$

We are now going to show that for a mirror symmetric matrix the evolution of symmetric and the antisymmetric components decouple. For the symmetric part, we have

$$\bar{W}\vec{y}_s = \begin{pmatrix} 0 & w_1' & w_2' & w_3 \\ w_1'' & 0 & w_1 & w_2'' \\ w_2'' & w_1 & 0 & w_1'' \\ w_3 & w_2' & w_1' & 0 \end{pmatrix} \begin{pmatrix} \psi_s \\ \chi_s \\ \chi_s \\ \psi_s \end{pmatrix} = \begin{pmatrix} (w_1' + w_2')\chi_s + w_3\psi_s \\ w_1\chi_s + (w_1'' + w_2'')\psi_s \\ w_1\chi_s + (w_1'' + w_2'')\psi_s \\ (w_1' + w_2')\chi_s + w_3\psi_s \end{pmatrix} , \tag{F.109}$$

which is again a symmetric vector. In fact it become convenient to work with the 2 dimensional vector,

$$\vec{y}_s^{2D} = \begin{pmatrix} \psi_s \\ \chi_s \end{pmatrix} , \tag{F.110}$$

with evolution equation

$$\dot{\vec{y}}_s^{2D} = -\vec{y}_s^{2D} + \bar{W}_s \vec{y}_s^{2D} + \vec{b}_s \ , \text{ where } \bar{W}_s = \begin{pmatrix} w_3 & w_1' + w_2' \\ w_1'' + w_2'' & w_1 \end{pmatrix} . \tag{F.111}$$

Similarly for the antisymmetric part, we have

$$\bar{W}\vec{y}_a = \begin{pmatrix} 0 & w_1' & w_2' & w_3 \\ w_1'' & 0 & w_1 & w_2'' \\ w_2'' & w_1 & 0 & w_1'' \\ w_3 & w_2' & w_1' & 0 \end{pmatrix} \begin{pmatrix} \psi_a \\ \chi_a \\ -\chi_a \\ -\psi_a \end{pmatrix} = \begin{pmatrix} (w_1' - w_2')\chi_a - w_3\psi_a \\ -w_1\chi_a + (w_1'' - w_2'')\psi_a \\ w_1\chi_a - (w_1'' - w_2'')\psi_a \\ -(w_1' - w_2')\chi_a - w_3\psi_a \end{pmatrix} , \tag{F.112}$$

which is an symmetric vector. Again it is convenient to work with the 2 dimensional vector,

$$\vec{y}_a^{2D} = \begin{pmatrix} \psi_a \\ \chi_a \end{pmatrix} , \tag{F.113}$$

with evolution equation

$$\dot{\vec{y}}_a^{2D} = -\vec{y}_a^{2D} + \bar{W}_a \vec{y}_a^{2D} + \vec{b}_s \ , \text{ where } \bar{W}_a = \begin{pmatrix} -w_3 & w_1' - w_2' \\ w_1'' - w_2'' & -w_1 \end{pmatrix} . \tag{F.114}$$

G Driven continuous attractors

G.1 General considerations for obtaining a continuum of steady-state profiles

In this appendix we are going to consider ring attractors that can encode angles continuously but needs an external drive to sustain the localized bump of activity. Specific examples of these ring attractors in the context of linear networks have been found recently, here we are going to provide the entire ensemble of threshold-linear symmetric networks. The potential advantage with these models is that we only need to impose a single degeneracy condition. However, as we will see, the external drive also affects the average angle that could make decoding the angle more difficult.

Since we now have an external drive we are going to consider dynamical equations of the form

$$\dot{\vec{y}} = -\vec{y} + \Phi(W\vec{y} + \vec{s}) \ , \tag{G.115}$$

where $\vec{s}$ is assumed to be a constant drive, and we assume that W has a single eigenvector with eigenvalue one to generate a continuum of steady-state profiles with different average angles. As for the self-sustaining networks, we focus on the steady state solutions that must now satisfy

$$\vec{y} = \bar{W}\vec{y} + \vec{s} \ , \tag{G.116}$$

where $\vec{y}$ and $\vec{s}$ are now representing four dimensional vectors.corresponding to the active neuronal directions. In analogy with solutions to linear differential equations, the general solution to the steady state equation can be expressed as a sum of a homogeneous and particular solution,

$$\vec{y} = \vec{y}_h + \vec{y}_p \ , \tag{G.117}$$

where the homogeneous solutions satisfy the eigenvalue equation,

$$\bar{W}\vec{y}_h = \vec{y}_h \ , \tag{G.118}$$

and, as we shall see momentarily, can provide the continuum of steady state profiles we require, while $\vec{y}_p$ is any solution that satisfies (G.116).

Let us first look at a particular solution satisfying,

$$\bar{W}_{ij}y_j + s_i = y_i \ , \tag{G.119}$$

where we are now going to assume that the external drive do not depend on the neuron, so $s_i = s \ , \forall \ i$. Then, we find that its mirror, $\vec{y}^\star$, also solves the equation:

$$\bar{W}_{ij}y_j^\star + s_i = \bar{W}_{ij}y_{j^\star} + s_i = \bar{W}_{i^\star j^\star}y_{j^\star} + s_{i^\star} = y_{i^\star} = y_i^\star \ , \tag{G.120}$$

where in the first equality we just used the definition of a mirror, $y_i^\star = y_{i^\star}$, and in the second inequality we used the symmetric property of the weight matrix, $\bar{W}_{ij} = \bar{W}_{i^\star j^\star}$, and $s_{i^\star} = s_i = s$. And, therefore one can find a symmetric configuration,

$$\vec{y}_s \equiv \frac{1}{2}(\vec{y} + \vec{y}^\star) \ , \tag{G.121}$$

that also satisfies (G.116) and can serve as the particular solution.

Let us next look at the homogeneous solutions. We have already shown (Appendix F) that as long there is only a single eigenvalue of one, the eigenstate corresponding to the homogeneous equation must be either symmetric or anti-symmetric. Now, if both the homogeneous and particular solution is symmetric, a continuous attractor with changing average location is not possible, the average is going to be always at the center. Thus we need an anti-symmetric eigenstate, $\vec{\epsilon}_a$ that characterizes a continuum of homogeneous solutions, $\vec{y}_h = \eta\vec{\epsilon}_a$, so that the general solution can be written in the form

$$\vec{y} = \vec{y}_s + \eta\vec{\epsilon}_a \ , \tag{G.122}$$

where we point out that $\vec{y}_s$ depends on s linearly which will be important later on.

G.2 Symmetric driven ring attractor networks relevant for fruit fly

We are now going to focus on the special case with $N = 8$ neuronal/computational units with steady state configurations involving $A = 4$ active neurons, as is pertinent for the fly ring attractor network. For simplicity, here we will specialize to the symmetric case, although most of our results should generalize to the mirror-symmetric networks as well. In Appendix B we already derived the condition for having an asymmetric eigenvector for the symmetric model. To summarize the results, to satisfy the eigenvalue condition the weights must satisfy,

$$(1 + w_3)(1 + w_1) = (w_1 - w_2)^2 \ . \tag{G.123}$$

and stability is ensured as long as,

$$-w_1 - w_3 < 2 \Rightarrow w_1 + w_3 > -2 \ . \tag{G.124}$$

The anti-symmetric eigenvector is given by

$$\vec{y}_a = \eta \vec{\epsilon}_a = \eta \begin{pmatrix} 1 \\ r_a \\ -r_a \\ -1 \end{pmatrix} \ , \text{ where } r_a = \frac{1 + w_1}{w_1 - w_2} \ . \tag{G.125}$$

Let us next try to find a symmetric particular solution. Substituting, the ansatz for a symmetric solution,

$$\vec{y}_s = \begin{pmatrix} \psi_s \\ \chi_s \\ \chi_s \\ \psi_s \end{pmatrix} \ , \tag{G.126}$$

we can rewrite (G.116) in matrix form:

$$\begin{pmatrix} 1 - w_1 & -w_1 - w_2 \\ -w_1 - w_2 & 1 - w_3 \end{pmatrix} \begin{pmatrix} \psi_s \\ \chi_s \end{pmatrix} = s \begin{pmatrix} 1 \\ 1 \end{pmatrix} \ . \tag{G.127}$$

Since we have assumed that we have a singly degenerate $\bar{W}$, the above matrix must be invertible and one obtains,

$$\begin{pmatrix} \psi_s \\ \chi_s \end{pmatrix} = \bar{s} \begin{pmatrix} 1 - w_3 & w_1 + w_2 \\ w_1 + w_2 & 1 - w_1 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \bar{s} \begin{pmatrix} 1 + w_1 + w_2 - w_3 \\ 1 + w_2 \end{pmatrix} \tag{G.128}$$

where

$$\bar{s} \equiv \frac{s}{(1-w_1)(1-w_3)-(w_1+w_2)^2} \ . \tag{G.129}$$

Thus the entire bump profile can be re-expressed as

$$\vec{y} = \sigma \begin{pmatrix} 1 \\ r_s \\ r_s \\ 1 \end{pmatrix} + \eta \begin{pmatrix} 1 \\ r_a \\ -r_a \\ -1 \end{pmatrix} = \sigma(s)[\vec{\epsilon}_s + \mu(s)\vec{\epsilon}_a] \ , \tag{G.130}$$

where

$$r_s \equiv \frac{1+w_2}{1+w_1+w_2-w_3} \ , \sigma(s) \equiv \frac{s(1+w_1+w_2-w_3)}{(1-w_1)(1-w_3)-(w_1+w_2)^2} \ , \text{ and } \mu(s) \equiv \frac{\eta}{\sigma(s)} \ . \tag{G.131}$$

G.3 External drive and decoding angular coordinate

By inspection of (G.130) and (B.29), it is clear that the steady state activity profiles look very similar, but importantly the parameters describing the different activity profiles, σ, μ, for a given set of weights become functions of the external drive. As we will now explain, this could lead to difficulty in decoding the heading direction as the average angular location now depends on the external drive. To see this it is more convenient to decompose the evolution equation in a symmetric and asymmetric piece (see Appendix F):

$$\vec{y} = \vec{y}_s + \vec{y}_a = \begin{pmatrix} \psi_s \\ \chi_s \\ \chi_s \\ \psi_s \end{pmatrix} + \begin{pmatrix} \psi_a \\ \chi_a \\ -\chi_a \\ -\psi_a \end{pmatrix} = \begin{pmatrix} y_s^{2D} \\ y_s^{2D} \end{pmatrix} + \begin{pmatrix} y_a^{2D} \\ -y_a^{2D} \end{pmatrix} \ , \tag{G.132}$$

and the evolution equation (F.107) now decomposes into two separate equations,

$$\dot{y}_s^{2D} = -y_s^{2D} + \bar{W}_s y_s^{2D} + s(t)\mathbf{1} \ , \text{ and } \dot{y}_a^{2D} = -y_a^{2D} + \bar{W}_a y_a^{2D} \ , \tag{G.133}$$

where we are now allowing the external drive to change, and we have defined,

$$\bar{W}_s \equiv \begin{bmatrix} w_3 & w_1+w_2 \\ w_1+w_2 & w_1 \end{bmatrix} , \ \bar{W}_a \equiv \begin{bmatrix} -w_3 & w_1-w_2 \\ w_1-w_2 & -w_1 \end{bmatrix} , \text{ and } \mathbf{1} \equiv \begin{pmatrix} 1 \\ 1 \end{pmatrix} \ . \tag{G.134}$$

The key point to note now is that the symmetric and the antisymmetric components evolve independently. Thus if $\vec{y}_a$ is at an initial steady state configuration, it will continue to remain in this state, but with a changing $s(t)$, the symmetric part will start to change with time, $\vec{y}_s = \vec{y}_s(t)$. This in particular will change the activity profile and the average angular location making decoding the heading angle from average neuronal activity difficult. To be specific, let us compute the average angle according to (B.35). For our steady state profiles we have

$$\langle\theta\rangle = \mu\left(\frac{\pi}{8}\right)\left(\frac{3+r_a}{1+r_s}\right) = \left[\frac{\pi\eta}{8s(t)}\right]\left(\frac{3+r_a}{1+r_s}\right)\left[\frac{(1-w_1)(1-w_3)-(w_1+w_2)^2}{1+w_1+w_2-w_3}\right] . \tag{G.135}$$

Imagine now $s(t)$ changing from some initial value to a different value. We know that this will not have any affect on the antisymmetric component of the activity vector and hence η will remain unchanged, but $\langle\theta\rangle$ will change with the changing shape of the activity profile because s has changed. Obviously, if s is a fixed external drive none of these decoding problem will arise. However, the average neuronal activity,

$$\langle y\rangle = \frac{1}{4}\sum_i^4 y_i = \frac{\sigma(1+r_s)}{2} = \frac{s(1+w_1+w_2-w_3)(1+r_s)}{2[(1-w_1)(1-w_3)-(w_1+w_2)^2]} , \tag{G.136}$$

was seen to vary from one steady state to another in the experimental data. This would imply that the external drive must be varying as well.

H Velocity Integration

In this appendix we are going to provide an example of a velocity drive that can continually deform the shape of the activity profile so that the bump of activity can faithfully and smoothly track the angular location. We are only going to focus on activity changes for a given fixed set of active neurons in the context of the effective network. We can therefore consider the linear network dynamics,

$$\dot{\vec{y}} = -\vec{y} + \bar{W}\vec{y} + \vec{v} , \tag{H.137}$$

where $\vec{v}$ is assumed to be a constant velocity drive. Again, the solution can be obtained as a sum of a *particular* solution ($\vec{y}_p$) and solutions ($\vec{y}_h$) to the homogeneous equation,

$$\dot{\vec{y}}_h = -\vec{y}_h + \bar{W}\vec{y}_h . \tag{H.138}$$

The homogeneous equation is a linear differential equation, and it is well known that its solution can be decomposed into modes:

$$\vec{y}_h(t) = \sum_{\nu=1}^4 b_\nu e^{\lambda_\nu t}\vec{\epsilon}_\nu , \tag{H.139}$$

where $\vec{\epsilon}_\nu$'s are constant eigenvectors that satisfy,

$$(\bar{W} - \mathtt{I})\vec{y}_\nu = \lambda_\nu \vec{\epsilon}_\nu \ . \tag{H.140}$$

Now we know that for our models two eigenvalues of $\bar{W} - \mathtt{I}$ are zero, corresponding to the steady state eigen vectors and the other two have negative eigenvalues from stability considerations. Thus the general solution to the homogeneous equation can be written as

$$\vec{y}_h(t) = \sigma_0(\vec{\epsilon}_s + \mu_0 \vec{\epsilon}_a) + b_3 e^{\lambda_3 t}\vec{\epsilon}_3 + b_4 e^{\lambda_4 t}\vec{\epsilon}_4 \ , \tag{H.141}$$

where $\lambda_3, \lambda_4 < 0$ are the two remaining eigenvalues. We note that because these eigenvalues are negative, they give rise to decaying modes that loose significance with time and therefore will be ignored henceforth.

Let us now focus on finding a particular solution. For the purpose of illustration we will consider a velocity drive that is along the "shape-shifting" direction:

$$\vec{v} = v_a \vec{\epsilon}_a \ . \tag{H.142}$$

Then we claim that we can find a particular solution of the form

$$\vec{y}_p = b_a(t)\vec{\epsilon}_a \ . \tag{H.143}$$

Since $(W - \mathtt{I})\vec{\epsilon}_a = 0$, substituting the particular solution in (H.137) we have

$$\dot{b}_a \vec{\epsilon}_a = v_a \vec{\epsilon}_a \Rightarrow \dot{b}_a = v_a \ . \tag{H.144}$$

A particular solution is then given by

$$\vec{y}_p = v_a t \vec{\epsilon}_a \ . \tag{H.145}$$

Thus one can express the entire solution as

$$\vec{y}(t) = \sigma_0[\vec{\epsilon}_s + \mu(t)\vec{\epsilon}_a] \ , \text{ where } \mu(t) = \mu_0 + \left(\frac{v_a}{\sigma_0}\right) t \ . \tag{H.146}$$

In other words, the rate of change of angular location, $\omega \equiv d\langle\theta\rangle/dt$ is proportional to the velocity drive, v_a,

$$\omega = \frac{\pi v_a}{\sigma_0} \ . \tag{H.147}$$

It is interesting to note that fruit flies in the dark are only able to proportionately track the heading direction. Perhaps, it requires an additional cue to adjust the proportionality factor and track heading directions precisely.