

Ring attractor dynamics in the *Drosophila* central brain

Sung Soo Kim,* Hervé Rouault,* Shaul Druckmann,† Vivek Jayaraman†

Ring attractors are a class of recurrent networks hypothesized to underlie the representation of heading direction. Such network structures, schematized as a ring of neurons whose connectivity depends on their heading preferences, can sustain a bump-like activity pattern whose location can be updated by continuous shifts along either turn direction. We recently reported that a population of fly neurons represents the animal's heading via bump-like activity dynamics. We combined two-photon calcium imaging in head-fixed flying flies with optogenetics to overwrite the existing population representation with an artificial one, which was then maintained by the circuit with naturalistic dynamics. A network with local excitation and global inhibition enforces this unique and persistent heading representation. Ring attractor networks have long been invoked in theoretical work; our study provides physiological evidence of their existence and functional architecture.

Studies of neural circuits near the sensory periphery have produced deep mechanistic insights into circuit functions (1, 2). However, it has been more challenging to understand circuit functions in central brain regions dominated by recurrent networks, which often produce complex neural activity patterns. These dynamics play a major role in shaping cognitive functions (3–7), such as the maintenance of heading information during navigation (8–10). A heading representation must be unique (because an animal can face only one direction at a given time) and persistent (to allow an animal to keep its bearings in darkness), yet must allow updating that matches the magnitude and speed of heading changes expected from the animal's movements. Theoretically, this can be accomplished by ring attractor networks (11–14), wherein the position of a localized subset of active neurons in a topological ring represents the animal's heading direction. However, whether the brain uses these hypothesized networks is still unknown (8, 15). A recent study reported that a population of neurons, called E-PG neurons (Fig. 1, C and D; see supplementary materials for nomenclature), in the *Drosophila melanogaster* ellipsoid body (EB) appears to use bump-like neural activity dynamics to represent the animal's heading in visual environments and in darkness (16, 17). Here, we establish essential properties of the network that enables this representation.

We first determined whether the E-PG population activity bump tracks the fly's heading direction relative to its visual surroundings during tethered flight (Fig. 1 and fig. S1). We used two-photon imaging with the genetically encoded calcium indicator GCaMP6f (18) to record dendritic calcium activity of the entire E-PG population in the EB while the fly was flying in a virtual-reality LED arena. The azimuthal velocity of the visual

scene was proportional to the fly's yaw velocity (Fig. 1, A and B). As with walking flies (16), E-PG population activity during flight was organized into a single bump, whether the visual scene contained a single bar (fig. S1B) or a more complex pattern (Fig. 1G). The activity bump closely tracked the fly's heading in flight (Fig. 1K) and persisted in darkness (Fig. 1H). However, unlike in walking, the activity bump seldom tracked the fly's motor actions in darkness (Fig. 1, H and K, and fig. S1C), potentially because tethering deprives the fly of normal sensory feedback about its rotational movements from its halteres (19). Although the location of the activity bump eventually drifted in some flies, the bump's movement was, on average, uncorrelated to the animal's turning movements in darkness (Fig. 1K). These findings suggest that the representation of heading in the E-PG population has intact, visually driven dynamics as well as persistence, but is largely uncoupled from updating by self-motion cues during tethered flight.

To test whether the fly's compass network enforces a unique bump within the EB, we took advantage of the relative persistence of the visually evoked activity bump in darkness, and asked whether this bump could coexist with an "artificial" bump of activity. We used localized optogenetic stimulation to create artificial activity bumps in different locations within the E-PG population. Using a transgenic fly line in which E-PG neurons co-expressed CsChrimson (20) and GCaMP6f, we used alternating two-photon laser scan lines of excitation (higher laser intensity) and imaging (normal laser intensity) to monitor changes in E-PG population dynamics in response to an optogenetically created spot of local activity (Fig. 2, A and B, and fig. S2, A and B). By varying the intensity of stimulation light delivered to the target location, we could create bumps of increased calcium activity (Fig. 2, C to F, and movie S1). As the new bump formed, activity at the previous location began to decline and eventually disappeared (Fig. 2D) without significantly perturbing the fly's behavior (but see fig. S2E). When the optogenetic excitation

was terminated, the amplitude of the artificially created bump settled at levels typically evoked by sensory stimuli and did not disappear; it either stayed in the induced location for several seconds (fig. S2F) or slowly drifted away (see below) (Fig. 3).

The bump's uniqueness may arise through either recurrent mutual suppression or an indirect mechanism whereby strong bump activity in the EB functionally inhibits feedforward sensory inputs to other E-PG neurons. To discriminate between these alternatives, we simultaneously excited two locations on the EB ring. A reference location was excited at a fixed laser power, and a second, spatially offset location was excited at increasing levels of laser power (fig. S2G and movies S3 to S5). We could always suppress the reference bump by increasing laser power at the second location above a certain threshold, consistent with mutual suppression.

Recurrent suppression can ensure a unique activity bump through a simple winner-take-all (WTA) circuit (fig. S3A). However, an animal's representation of its angular orientation should favor more continuous updates based on turning actions. Such gradual, ordered drift to nearby locations would be more consistent with continuous, or ring, attractor models (fig. S3, B to D). We therefore examined changes in the location of an artificially created bump after the stabilization of its peak activity at the "natural" level. The experiments were performed in darkness to untether the bump from any potentially lingering visual input (Fig. 3). If EB dynamics were driven by a WTA network, bumps would be expected to disappear at times and to jump to random distant locations (fig. S3E). In contrast, the bump drifted gradually around the EB (Fig. 3, B and D, and movie S6); this finding suggests that the fly's heading representation is updated through functionally excitatory interactions between neighboring E-PG neurons, consistent with a ring attractor model. These observations together rule out the possibility that network dynamics in darkness result purely from cell-intrinsic mechanisms (21, 22) or slowly decaying visual input. Most important, direct manipulation of E-PG neuron activity changed the network state, which implies that E-PG neurons do not merely mirror dynamics occurring in a different circuit, but are themselves an important component of the ring attractor (23).

We next aimed to dissect the effective connectivity pattern underlying ring attractor dynamics in the E-PG population. A wide range of network structures can, in principle, implement ring attractors (11, 13, 14, 24, 25). We focused our efforts to a model space between two extreme network architectures that are analytically solvable: (i) a "global model" based on global cosine-shaped interactions (fig. S3B) (11, 13, 26) and (ii) a "local model" based on relatively local excitatory interactions (fig. S3D and supplementary text) (24, 27). Under constraints of a fixed bump width of 90° to match physiological observations (Fig. 1J) and an assumption of effectively excitatory visual input without any negative bias, both models could

Howard Hughes Medical Institute, Janelia Research Campus, Ashburn, VA 20147, USA.

*These authors contributed equally to this work.

†Corresponding author. Email: druckmanns@janelia.hhmi.org (S.D.); vivek@janelia.hhmi.org (V.J.)

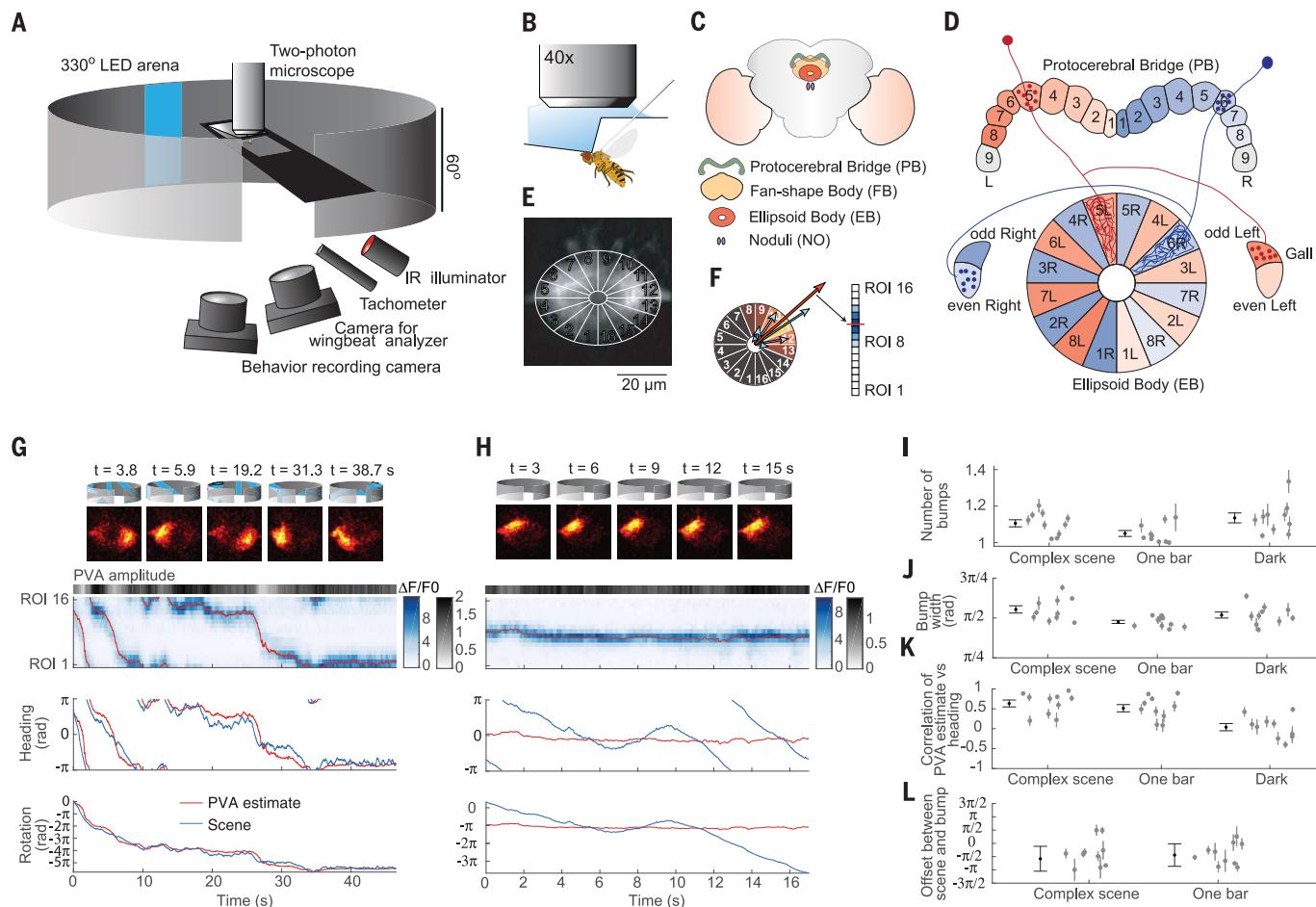


Fig. 1. E-PG neurons encode body orientation relative to the visual world during tethered closed-loop flight.

(**A**) Setup schematic. (**B**) Close-up of tethered flying fly. (**C**) Central complex. (**D**) Dendrites of each E-PG neuron innervate wedge-shaped segment of EB; axons project to corresponding glomeruli in PB and Gall. (**E**) Averaged calcium image of dendritic processes of entire E-PG population segmented into 16 regions of interest (ROIs). (**F**) Position (PVA direction) and strength (PVA amplitude) of bump obtained by summation of 16 vectors whose lengths represent magnitude of fluorescence transients ($\Delta F/F_0$). (**G**) GCaMP6f fluorescence transients in E-PG dendrites during tethered flight in complex visual scene. Top: Visual pattern at sample

time points. Second row: Sample frames of calcium imaging. Third row: $\Delta F/F_0$ of 16 ROIs. Grayscale band denotes PVA amplitude; red line is PVA estimate. Fourth row: PVA estimate and heading (blue). Bottom: Same as fourth row, but unwrapped. (**H**) Fluorescence transients in darkness. (**I**) Number of activity bumps in E-PG population across flies ($n = 10$) for three visual conditions. Each dot with vertical line indicates mean \pm SEM for each fly. Population mean \pm SEM is shown at left of each scatterplot. (**J**) Bump width measured by full width at half maximum. (**K**) Correlation between estimated bump position and heading. (**L**) Angular offset between PVA estimate and scene orientation. Whisker plots, mean \pm circular SD.

explain the basic properties of bump dynamics, including its uniqueness and its persistence in darkness. We therefore probed the network's response to more artificial conditions, such as abrupt visual stimulus shifts.

We first examined experimentally how the E-PG population responded to unnatural, abrupt visual shifts. Depending on the distance of the shift, the E-PG bump either “flowed” continuously (shorter shift distances; Fig. 4, A and C, and movies S7 and S8) or “jumped” to the new location (longer shift distances; Fig. 4, B and C, and movie S9) (16). In simulations, both models predicted a mixture of jump and flow responses, depending on the strength and width of the abruptly shifting visual input (Fig. 4D, fig. S4A, and supplementary text). For example, weak wide input induced flows and strong narrow input evoked jumps (Fig. 4D). However, the jump-flow balance predicted by the two

models differed and was more consistent with the local model in several aspects (Fig. 4D and fig. S4A). First, the visual input strength we inferred from normal conditions was much weaker than required by the global model for bump jumps (fig. S1D). Second, the global model required a much-wider-than-normal range of visual input strengths to explain jumps at multiple distances (Fig. 4D, fig. S1D, and fig. S4A). Third, using parameters consistent with the rest of our findings, we could reproduce the jump-flow ratio observed in Fig. 4C with the local model but not with the global model (fig. S4B).

To obtain more concrete evidence, we compared model predictions to experimentally observed bump dynamics, under conditions in which input strength, polarity, and shift distance were controlled through optogenetic stimulation. To simulate moderate and large input shift distances, we se-

quentially stimulated two small regions in the EB—each with an angular width of 22.5° —separated by either 90° or 180° (Fig. 4, E to G, movie S10, and fig. S4, C to E). We then varied the stimulation laser power to detect the threshold required for the bump to jump (Fig. 4E). The laser power required to elicit a jump was not significantly different between the two different shift distances, favoring the local model (Fig. 4F). We then inferred the strength of input to the network by comparing the amplitude of the optogenetically evoked bump to natural bump amplitudes in darkness. The optogenetic input strength required to induce jumps was smaller than the global model's prediction but matched that of the local model (Fig. 4G) and the range of the inferred visual input strength under normal conditions (fig. S1D, fig. S4, D and E, and movie S11). Finally, when we tested intermediate models that lie between the extremes of

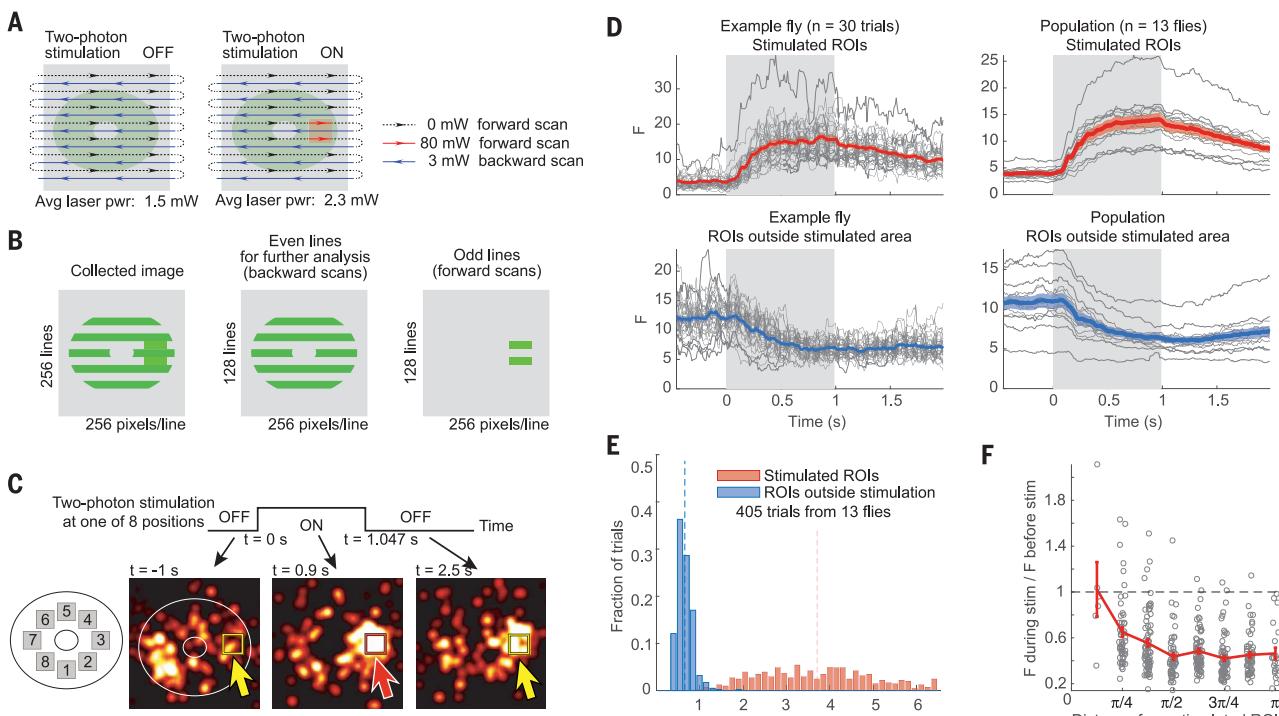


Fig. 2. E-PG neurons compete by mutually suppressing each other through recurrent connections. (A) Schematic of simultaneous calcium imaging and localized optogenetic stimulation. (B) Analysis procedure for collected images. (C) Top: Temporal profile of two-photon optogenetic stimulation. Bottom: Three sample frames (smoothed with Gaussian filter). Yellow rectangle with arrow, stimulus OFF; red rectangle with arrow, stimulus ON. (D) Time course of calcium dynamics from example fly (left) and population (right). Gray background, optogenetic stimulation period; gray lines, individual trials (left) or flies (right). Top: Mean F of stimulated

ROIs. Bottom: Mean of the four most active ROIs outside optogenetically stimulated area before stimulation. Thick colored lines and colored shaded area denote mean and SEM, respectively. (See fig. S2C for control experiment.) (E) Distribution of fluorescence ratio during and before stimulation. $P < 0.001$, Wilcoxon rank sum test between stimulated (red) and outside stimulation (blue) areas. (See fig. S2D for control experiment.) (F) Suppression by optogenetic stimulation. The x axis indicates distance from stimulation position to existing bump; $P < 0.001$, t test for each distance. Limited sample size prevented a statistical test for $\pi/8$.

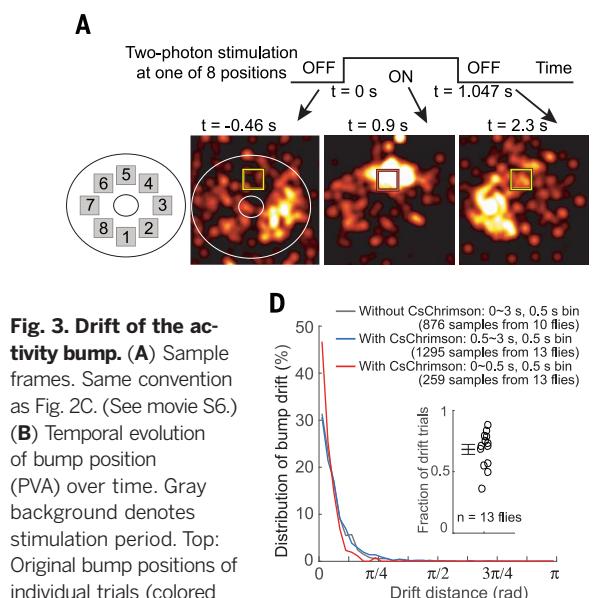
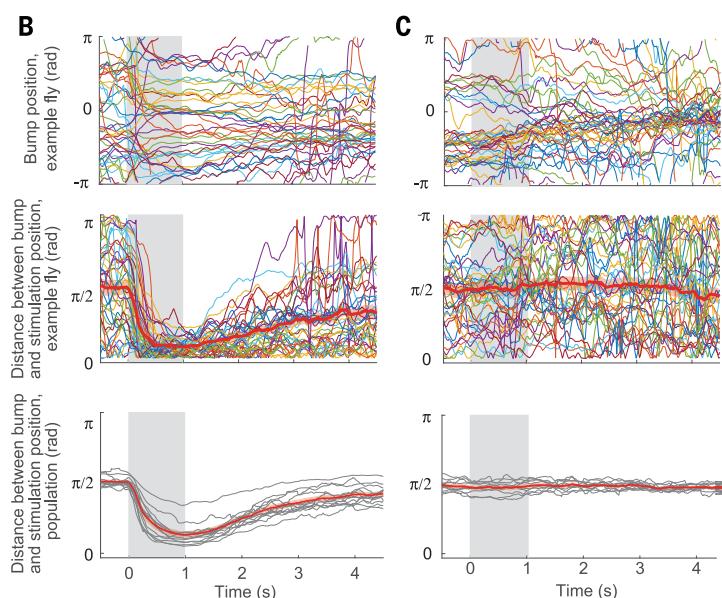


Fig. 3. Drift of the activity bump. (A) Sample frames. Same convention as Fig. 2C. (See movie S6.) (B) Temporal evolution of bump position (PVA) over time. Gray background denotes stimulation period. Top: Original bump positions of individual trials (colored thin lines are PVA estimates). Second row: Distance between bump and stimulation position. Red line and shade denote mean \pm SEM. Bottom: Population mean \pm SEM (red) across flies (gray lines). (C) Same as (B), without CsChrimson. (D) Distribution of bump drift distances after the end of optogenetic stimulation. Colored lines represent different conditions. $P = 0.324$ between gray and



blue, $P < 0.0001$ between blue and red, $P < 0.0001$ between gray and red; two-sample Kolmogorov-Smirnov tests without multiple-comparisons correction. Distributions are skewed toward short drift distances. Inset shows fraction of trials with drifting bump in each fly ($P = 0.0008$, t test compared to 0.5).

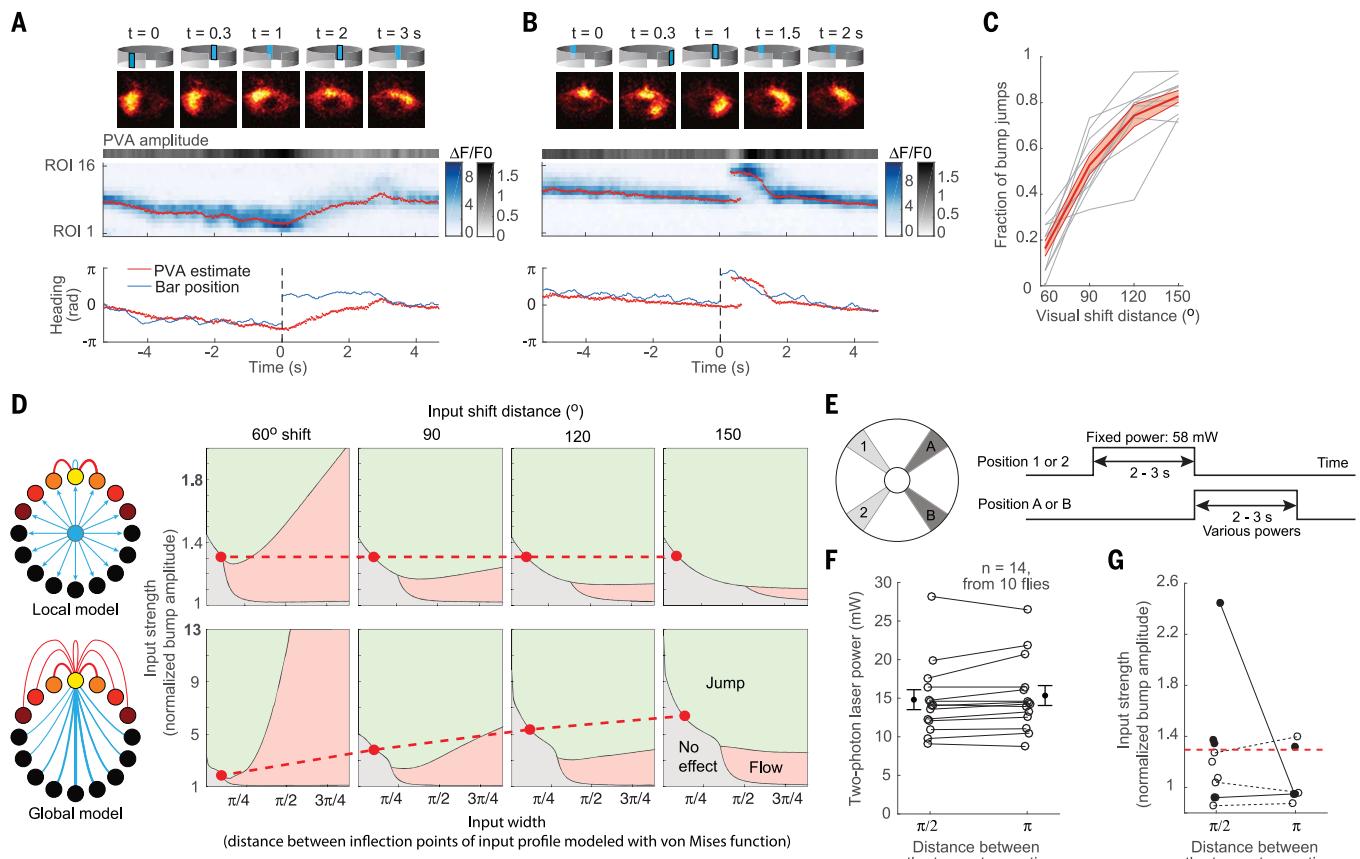


Fig. 4. Probing the connectivity profile of the ring attractor network.

(**A**) Example of bump “flow” in response to abrupt shift of vertical bar. Same convention as Fig. 1G. Red dots are bump positions estimated from Bayesian sampling method. (**B**) Bump “jump.” (**C**) Jump probability increases with distance of visual input shift. Red line and shading denote mean \pm SEM. (**D**) Input-response phase diagrams. Top: Response of local model (fig. S3D) to various input widths, strengths, and abrupt shift distances. Bottom: Global model (fig. S3B). Note that the y axis increments are different between the two models. Red lines denote input strength for bump jump with narrow input, which is constant for the local model and

increases with shift distances for the global model. (**E**) Schematics of stimulation protocol to detect the threshold input strength for bump jump in response to narrow (22.5°) input. Two 22.5° areas were sequentially stimulated. (**F**) Laser power required to make bump jump from the first stimulation position (1 or 2) to a fixed second stimulation position (A or B). $P = 0.102$, paired t test. (**G**) Input strength, estimated by normalized bump amplitude, required for bump jump from fixed first stimulation position to second stimulation position. Red dashed line denotes simulated threshold of the local model. Solid dots are trials with first stimulation at position 1; open dots are trials with first stimulation at position 2.

the local and global models (fig. S4, H and I, and supplementary text), we found that any model that exhibited the observed jumps in response to a weak 22.5° -wide input had narrow connectivity profiles (fig. S4I). All these observations were once again consistent with the local model.

In mammals, heading representations are thought to be distributed across multiple neural populations and multiple brain areas (8). In *Drosophila* as well, the compass system likely involves multiple cell types, including neurons in the protocerebral bridge (PB) (7, 23). Further, occasional changes observed in the dynamics suggest network modulation by other factors not yet known. For example, we sometimes observed sudden changes in E-PG dynamics, as when the amplitude of the sensory-evoked activity bump changed depending on whether or not the tethered fly was flying (see supplementary materials) and, occasionally, during flight [population vector average (PVA) amplitude plots in Fig. 1, G and H, Fig. 4, A and B, and fig. S1B]. Nonetheless, the E-PG population provides a

powerful physiological handle on the internal representation of heading (16): a single activity bump moving through topographically arranged neurons. The experimental approach this enabled provides one avenue for investigating which of multiple populations are key circuit components of a computation and which simply read out the results of that computation. We found that the artificial bump created by directly manipulating E-PG population activity displays natural dynamics, which indicates that these neurons are a key component of the heading circuit.

Our finding that the uniqueness of the E-PG activity bump is ensured via global competition strengthens the conclusion that this population encodes an abstract internal representation of the fly’s heading direction (16). Such abstract representations permit an animal to untether its actions from the grasp of its immediate sensory environment and thereby confer flexibility in both time and behavioral use. Combining an analysis of artificially induced bump dynamics with theoretical

modeling allowed us to interrogate this recurrent circuit architecture. We found that the effective network connectivity profile was consistent with ring attractor models characterized by narrow local excitation and flat long-range inhibition. This neural circuit motif of local excitation and long-range inhibition is ubiquitous across many brain areas and across animal taxa (28–31). Such observations support the idea that common circuit motifs might be evolutionarily adapted to serve as crucial building blocks of cognitive function.

REFERENCES AND NOTES

- R. I. Wilson, *Annu. Rev. Neurosci.* **36**, 217–241 (2013).
- A. Borst, M. Helmstaedter, *Nat. Neurosci.* **18**, 1067–1076 (2015).
- J. J. Knierim, K. Zhang, *Annu. Rev. Neurosci.* **35**, 267–285 (2012).
- E. Akopyan et al., *Nat. Neurosci.* **10**, 494–504 (2007).
- H. S. Seung, *Proc. Natl. Acad. Sci. U.S.A.* **93**, 13339–13344 (1996).
- R. Chaudhuri, I. Fiete, *Nat. Neurosci.* **19**, 394–403 (2016).
- K. Wimmer, D. Q. Nykamp, C. Constantinidis, A. Compte, *Nat. Neurosci.* **17**, 431–439 (2014).

8. J. S. Taube, *Annu. Rev. Neurosci.* **30**, 181–207 (2007).
9. J. S. Taube, R. U. Muller, J. B. Ranck Jr., *J. Neurosci.* **10**, 420–435 (1990).
10. B. L. McNaughton, F. P. Battaglia, O. Jensen, E. I. Moser, M. B. Moser, *Nat. Rev. Neurosci.* **7**, 663–678 (2006).
11. R. Ben-Yishai, R. L. Bar-Or, H. Sompolinsky, *Proc. Natl. Acad. Sci. U.S.A.* **92**, 3844–3848 (1995).
12. W. E. Skaggs, J. J. Knierim, H. S. Kudrimoti, B. L. McNaughton, *Adv. Neural Inf. Process. Syst.* **7**, 173–180 (1995).
13. K. Zhang, *J. Neurosci.* **16**, 2112–2126 (1996).
14. X. Xie, R. H. Hahnloser, H. S. Seung, *Phys. Rev. E* **66**, 041902 (2002).
15. A. Peyrache, M. M. Lacroix, P. C. Petersen, G. Buzsáki, *Nat. Neurosci.* **18**, 569–575 (2015).
16. J. D. Seelig, V. Jayaraman, *Nature* **521**, 186–191 (2015).
17. T. Wolff, N. A. Iyer, G. M. Rubin, *J. Comp. Neurol.* **523**, 997–1037 (2015).
18. T. W. Chen *et al.*, *Nature* **499**, 295–300 (2013).
19. M. H. Dickinson, *Philos. Trans. R. Soc. London Ser. B* **354**, 903–916 (1999).
20. N. C. Klapoetke *et al.*, *Nat. Methods* **11**, 338–346 (2014).
21. M. Yoshida, M. E. Hasselmo, *J. Neurosci.* **29**, 4945–4952 (2009).
22. G. Major, D. Tank, *Curr. Opin. Neurobiol.* **14**, 675–684 (2004).
23. K. S. Kakaria, B. L. de Bivort, *Front. Behav. Neurosci.* **11**, 8 (2017).
24. S. Amari, *Biol. Cybern.* **27**, 77–87 (1977).
25. P. Song, X. J. Wang, *J. Neurosci.* **25**, 1002–1014 (2005).
26. D. Hansel, H. Sompolinsky, in *Methods in Neuronal Modeling: From Ions to Networks*, C. Koch, I. Segev, Eds. (MIT Press, 1998), pp. 499–567.
27. A. M. Turing, *Philos. Trans. R. Soc. London Ser. B* **237**, 37–72 (1952).
28. H. Ko *et al.*, *Nature* **473**, 87–91 (2011).
29. M. Weliky, K. Kandler, D. Fitzpatrick, L. C. Katz, *Neuron* **15**, 541–552 (1995).
30. H. Ozeki, I. M. Finn, E. S. Schaffer, K. D. Miller, D. Ferster, *Neuron* **62**, 578–592 (2009).
31. S. P. Mysore, A. Asadollahi, E. I. Knudsen, *J. Neurosci.* **30**, 1727–1738 (2010).

ACKNOWLEDGMENTS

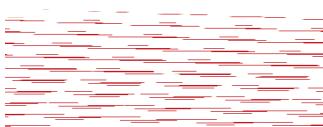
We thank A. Jenett, T. Wolff, and G. Rubin for their generous sharing of the split line SS00096 before publication, and, in particular, T. Wolff for her scientific insights; B. Pfeiffer, A. Wong, D. Anderson, and G. Rubin for generously sharing codon-optimized GCaMP6f DNA before publication; C. Dan for codon-optimized GCaMP6f flies; S. Romani for contributing to insightful discussions regarding bump jumps in the global model and the effect of

input width; Janelia Fly Core and, in particular, K. Hibbard and S. Coffman, for support; J. Liu for virtual reality support; D. Flickinger and V. Goncharov for microscope design support; J. Arnold for fly holder design; Vidrio for ScanImage support; S. Rowell and E. Nielson for operational support; and A. Karpoval, A. Hermundstad, M. Ahrens, S. Romani, Y. Sun, and members of V.J.’s lab for useful discussions and comments on the manuscript. Supported by the Howard Hughes Medical Institute. Code for physiological data and analyses is available at http://research.janelia.org/jayaraman/KimRouaultScience2017_Downloads.

SUPPLEMENTARY MATERIALS

www.sciencemag.org/content/356/6340/849/suppl/DC1
Materials and Methods
Supplementary Text
Figs. S1 to S4
Movies S1 to S11
References (32–46)
Code for ring attractor simulation and Bayesian sampling

29 November 2016; accepted 20 April 2017
Published online 4 May 2017
10.1126/science.aal4835



Ring attractor dynamics in the *Drosophila* central brain
Sung Soo Kim, Hervé Rouault, Shaul Druckmann and Vivek Jayaraman (May 4, 2017)
Science **356** (6340), 849–853. [doi: 10.1126/science.aal4835]
originally published online May 4, 2017

Editor's Summary

Representing direction in the fly

A population of cells called compass neurons represents a fruitfly's heading direction. Kim *et al.* used imaging and optogenetics in behaving flies to elucidate the functional architecture of the underlying neuronal network. They observed local excitation and global inhibition between the compass neurons. The features of the network were best explained by a ring attractor network model. Until now, this hypothesized network structure has been difficult to demonstrate in a real brain.

Science, this issue p. 849

This copy is for your personal, non-commercial use only.

Article Tools Visit the online version of this article to access the personalization and article tools:
<http://science.scienmag.org/content/356/6340/849>

Permissions Obtain information about reproducing this article:
<http://www.sciencemag.org/about/permissions.dtl>

Supplementary Materials for

Ring attractor dynamics in the *Drosophila* central brain

Sung Soo Kim, Hervé Rouault, Shaul Druckmann,* Vivek Jayaraman*

*Corresponding author. Email: druckmanns@janelia.hhmi.org (S.D.); vivek@janelia.hhmi.org (V.J.)

Published 4 May 2017 on *Science* First Release

DOI: 10.1126/science.aal4835

This PDF file includes:

Materials and Methods

Figs. S1 to S4

Captions for movies S1 to S11

Supplementary Text

References

Other supplementary material for this manuscript includes:

Movies S1 to S11

Source codes for ring attractor simulation and bump sampling

MATERIALS AND METHODS

Nomenclature

E-PG neurons are variously referred to as eb-pb-vbo (32), EIP (33), PBg.b-EBw.s-gall.b (17), and EBw.s (16) neurons in the fly literature. Here, we follow a new convention agreed upon by a few research groups working in the central complex: E-PG, where E (Ellipsoid Body) before ‘-’ represents predominantly spiny and putatively postsynaptic processes, and P (Protocerebral Bridge) and G (Gall) after ‘-’ represent predominantly bouton-like and likely presynaptic processes.

Fly stocks

An optogenetic reagent, CsChrimson and the calcium indicator, GCaMP6f, were driven by split-GAL4 (34, 35) SS00096, which was made by A. Jenett and anatomically characterized by T. Wolff in the Rubin lab and generously shared pre-publication. All calcium imaging experiments were performed on 4-6 day old female flies expressing fly codon-optimized UAS-GCaMP6f (18), the DNA for which was generously shared pre-publication by B. Pfeiffer, A. Wong, D. Anderson and G. Rubin. All optogenetic experiments were performed with 4-6 day old female flies generated by crossing split-GAL4 SS00096 with a recombinant of UAS-CsChrimson-mCherry-tag (20) and UAS-GCaMP6f (18). Flies were randomly picked from their housing vials for all experiments. All flies were raised from the egg stage on standard cornmeal and soybean-based medium (36) with 0.2 mM all-*trans*-retinal (20).

Fly preparation for imaging during head-fixed flight

Flies were anaesthetized on a cold plate at 4°C before individual tethering. The front legs were removed to prevent the fly from pushing the fly-holder during imaging. The fly’s proboscis was pressed into its head capsule and immobilized with wax to minimize brain movement. The fly was placed in a brass sarcophagus and a UV curable adhesive was used to glue a pin on the back of the thorax at an angle perpendicular to the body axis using a micromanipulator. The fly was then positioned under a custom-designed stainless steel shim with a pyramidal shape similar to those previously used for tethered flying and walking fly experiments (37-39). For purely visual experiments, the back of the head capsule was kept nearly vertical to maximize exposure of the fly’s eyes to the surrounding LED arena (see below). For optogenetic experiments, the fly’s head was tilted upwards to compensate for depth-dependent light scattering by orienting the EB optimally with respect to the microscope’s focal plane. UV curable adhesive was used to fix the head under the shim. The cuticle at the top of the head was carefully removed using a dissection needle and forceps. Fat cells and trachea blocking the light path were removed. The fly holder (including the micromanipulator) was then transferred to the two-photon microscope and an LED arena was place around the fly (Fig. 1A).

Visual stimulation

Visual arena

Visual stimuli were presented on a cylindrical LED display (40) spanning 330° in azimuth and 60° in elevation that was vertically placed. The display was covered with a color filter and a diffuser as previously described (16, 37, 38). In all experiments, the horizontal rotational velocity of a visual scene was controlled using the difference between the left and right wingbeat

amplitudes (Δ WBA) (41). The wingbeat amplitude of each wing was computed online by analyzing images acquired with a camera, using custom built image analysis software written in MATLAB, similar to a previously described method (39). The image acquisition rate of the camera was 119.2 Hz. The flies' wingbeat frequency was between 170-230 Hz. Thus, the camera's exposure time allowed it to capture the moving shadow of 1 or 2 wingbeats. For closed-loop experiments, the gain was 7.63 % for each degree of Δ WBA, i.e. a Δ WBA of 43.26° would produce a full pattern rotation in 1 s. Air was puffed at the fly if it stopped flying. The fly's behavior was recorded using a separate camera with a dedicated software. Acquisition of each frame was triggered by a TTL signal from the microscope frame trigger.

Visual stimuli for closed-loop flying experiments

We used three different visual stimuli (fig. S1A); condition 1, a bright vertical bar spanning 60° in elevation and 15° in azimuth; condition 2, a pattern containing several objects with unique features; condition 3, a dark scene with no bright pattern. Each of 10 flies was placed at the center of the arena and each visual stimulus was presented for 20 s in closed-loop, with 10 repetitions, for a total of 30 trials. To allow neural activity to adapt to each new pattern, the first 4 s were discarded and only the last 15 s of each trial were used for further analysis. For each fly, condition 2 was presented for 60 s to allow the fly to adjust to a new environment. Then, ten 20 s trials of condition 2 were followed with 2 s of darkness between consecutive trials. Then, 60 s of condition 1 was again presented to allow the fly adjust to a new scene. Finally, ten 20 s trials of condition 1 and ten 20 s trials of condition 3 were alternated.

In order for the position of the dark "pattern" to be consistent with the other conditions, the accumulated wingbeat amplitude difference was measured and used to estimate the body orientation of the fly (condition 3), meaning that the dark "pattern" was under closed-loop control in the same way as for condition 1 and 2, but without any actual visual stimulus displayed. The angular position of the darkness scene was then recorded and used for further analysis.

In instantaneous visual shift experiments (visual jump experiments, Fig. 4, A and B), only condition 1 (one bar) and 3 (darkness) were used. In each trial, a bar (condition 1) was presented in closed-loop for 5 s, before its position was suddenly changed by either -150°, -120°, -90°, -60°, 60°, 90°, 120°, or 150°. Closed-loop control of the bar position was then enabled for another 5 s (for a total of 10 s for each trial). The eight jump conditions were randomly permuted for each repeat, with a total of 8 repeats per fly. Condition 3 (darkness) was inserted after every 4 jumping conditions. In total, there were 80 trials for each of 10 flies.

The position of the pattern, wingbeat amplitudes, wingbeat frequency, air-puffing signal, and two-photon frame trigger were all simultaneously collected using a custom software written in MATLAB that utilized data acquisition hardware.

Two-photon calcium imaging

Calcium imaging was performed using a custom built two-photon microscope (42). We used a 40x objective (NA 1.0, 2.8mm WD) and a GaAsP photomultiplier tube. A chameleon Ultra II

laser tuned to 930 nm was used as the excitation source with a maximum power of 15 mW at the sample. We used the same saline as in previous studies (16) with adjusted calcium concentration at 2.0 mM. We imaged the EB from 6-plane volumes (with an additional fly-back blank plane) at a rate of 7.8 Hz (256x256 resolution) with an equal spacing of 5 μ m between individual scanning planes. For the high-powered anatomical scan (Fig. 1D), we used 1 μ m spacing over 150 μ m depth with 80mW.

Two-photon optogenetic stimulation

Scanning pattern during imaging-only periods

A single two-photon laser source was used for both imaging and optogenetic stimulation, by temporally modulating the laser power (Fig. 2A). Custom MATLAB software was written to modulate the beam profile, in which the laser power was set to 0 mW during forward scanning and to its normal power during backward scanning. Compared to recent developments of simultaneous two-photon imaging and stimulation (43, 44), this method does not require new hardware. The red-shifted channelrhodopsin, CsChrimson, expressed in E-PG neurons with the SS00096 driver line, was extremely sensitive to two-photon stimulation. With moderate scanning laser power (typically >4.5mW), the difference in dwell time (fig. S2A) from the sinusoidal scanning pattern of the resonant galvo mirrors and the different exposure frequencies from volume scanning were sufficient to bias activity in the recurrent EB network. To avoid such bias, an extremely low laser power (3 mW) was used for calcium imaging during the backward scanning phase, resulting in a relatively low signal to noise ratio due to the low two-photon laser intensity used for imaging. Without optogenetic stimulation, the mean laser power was 1.5 mW (even lines only at 3 mW).

Complications of CsChrimson excitation

The instantaneous laser power required to shift the bump to a desired EB position by optogenetic CsChrimson activation was different across the structure. Potential sources of non-uniform efficiency may include inconsistent CsChrimson expression across neurons, depth-dependent laser scattering, different laser dwell times from the resonant galvo mirror, different exposure frequencies from volume scanning, and slightly tilted brain position. 80 mW was strong enough to create a bump anywhere in the EB, but as little as 15 mW was enough to activate the most superficial part of EB.

Scanning pattern during imaging and simultaneous optogenetic stimulation

To activate CsChrimson, the powerbox feature of ScanImage was applied only to the forward scanning phase (Fig. 2A). For a single spot stimulus protocol, the instantaneous laser power was increased to 80 mW when the laser passed a small defined area of the imaging field. The length of each side of the rectangle-shape stimulated area was 14.6% of the x- and y-axis of the imaging area, covering 2.12% of the whole imaging area. The mean laser power for each focal plane was 0.848 mW (80 mW x 0.0212 x 0.5 (odd lines only)), resulting in a mean scanning power of approximately 2.3 mW (1.5 mW + 0.8 mW).

Activation of one position in the EB

Before the start of optogenetic stimulation, a visual bar was presented to engage the fly in closed-loop control and to enhance bump activity. The visual bar was removed at the onset of

the optogenetic stimulation of one of eight positions in the EB (Fig. 2C). The stimulation (instantaneous power 80 mW) was maintained for 1.047 s (100 frames) and then stopped. There was no additional visual stimulus to avoid disturbing the optogenetically induced activity bump. Imaging was performed at a frame rate of 95.5 Hz, a volume rate of 13.64 Hz (6 imaging planes and a fly-back frame), and a resolution of 256x128.

Measuring the behavioral effect of optogenetic bump shifting (fig. S2E)

The stimulation protocol was the same as Fig. 2C, but with the following differences. The visual bar was not removed during stimulation but always presented under the fly's closed-loop control. The temporal gap between consecutive trials were always larger than 7 seconds. The stimulation duration was 2.5s.

Laser power vs fluorescence level

The calcium response to various two-photon (TP) power levels was measured (fig. S2B). For calibration, trials in which flies were not flying were used, because responses were more stationary. The most superficial part of the EB was used for calibration. The same stimulation protocol as described in "Activation of one position" was used.

Simultaneous activation of two positions

The powerbox feature was used to modulate the laser power at two different positions in the EB (fig. S2G). These two positions were selected before the start of each session and fixed for each fly, with the criterion that comparable laser powers lead to optogenetic activation in either position. Fixed laser power (18.4 mW) was used for one position, and 18 different laser powers (evenly divided from 0 to 83 mW, 3 repetitions for each) were used for the other. Imaging was performed at a 54.29 Hz frame rate, a 256x256 resolution, a 7.76 Hz volume rate (6 imaging planes plus a fly-back frame), and optogenetic stimulation applied for 100 frames (1.842 s).

Consecutive activation of two positions

Four narrow, wedge shaped stimulus profiles, each covering 22.5°, were designed to limit optogenetic stimulation to one angular wedge-width (17). Protocol 1 (Fig. 4, E to G): The position of the four stimulation profiles, labeled as 1, 2, A and B, are shown in Fig. 4E. Profile 1 or 2 was first used for ~2-3 s to position the bump. It was then turned off while profile A or B was activated for the next ~2-3 s. The transition from the first to the second profile was immediate, but the timing of the transition could not be precisely controlled due to technical limitations, resulting in varying stimulation durations. The stimulation power for profiles 1 and 2 was pre-determined before a session at a setting high enough to capture the bump, and was kept constant throughout the session. The power for the second profile was manually adjusted for every trial to keep it near the threshold required to cause the bump to jump (fig. S4C). For example, if a given trial showed a successful jump at a given position, the power was slightly lowered for the next trial for that position. If the bump didn't jump, the power was slightly increased for the next trial. Whether the bump jumped or not was subjectively determined during the experiment, but quantitatively assessed during subsequent analysis (see below). The number of repetitions was not predetermined. Sessions were terminated if flies stopped flying despite occasional air puffs, or if 40 or more trials were collected for each pair of profiles, whichever came first. To avoid bias, profiles 1 and 2 were always paired for a given 2nd stimulation power (fig. S4C). For example, if profile 1 was paired with profile A of 20 mW, then the next trial was

profile 2 paired with the same profile A using the same power, 20 mW. We noticed that there was sometimes a slow fluctuation of the threshold power required for a jump, something that was controlled by this pairing strategy. Imaging was performed at a 54.29 Hz frame rate, a 256x256 resolution, a 7.76 Hz volume rate (6 imaging planes plus a fly-back frame), and optogenetic stimulation was applied for 100 frames (1.842 s). Protocol 2 (fig. S4, D and E): The position of four profiles, labeled as 1, A, B, and C, are shown in Fig. 4D. The position of profile 1 was first stimulated for ~2-3 s to capture the bump. The profile was then turned off while one of three remaining profiles was stimulated for the next ~2-3 s.

Data analysis

We used MATLAB for data analysis. All errors and error bars shown are standard error of the mean (s.e.m.) unless indicated otherwise. To avoid bias, no statistical methods were used to predetermine the power and the sample size. All experiments (except the single spot optogenetic experiment: 13 flies) were performed until data from 10 flies was collected. All standard statistical tests were two-tailed tests unless indicated otherwise.

Calculation of fluorescence changes

The background noise level was predetermined by measuring the oscillatory noise from PMT. This level was then subtracted from all imaging data, and the data was half-rectified before further analysis. A running maximum intensity projection (MIP) of a volume (7 planes including a fly-back frame) at a given time was computed for each pixel. Then, 16 ROIs were manually assigned (Fig. 1E), as previously described (16). The number of wedges was selected based on the anatomically-characterized number of EB wedges. Next, time series for each ROI were obtained by taking the average of the fluorescence signal within the ROI at each point in time. For calcium imaging experiments without optogenetics, $\Delta F/F_0$ was computed using F_0 as the mean of the lowest 10% of signals in each ROI. No further temporal smoothing was applied unless indicated otherwise. For optogenetic stimulation experiments, the F was extremely low and noisy. Thus, we took the raw F (after MIP), but the signal was temporally smoothed using boxcar averaging over 21 planes (equivalent to 3 volumes) to increase SNR.

Population vector average (PVA) and its amplitude for Fig. 1

As a simple measure of the bump position and strength, the PVA was computed as the weighted average across EB wedges, with the weight determined by the fluorescence level ($\Delta F/F_0$), and the vector determined by the position of each ROI in the EB (Fig. 1F). The amplitude of the PVA was determined as the length of the average vector. We used brewermap (S. Cobeldick, MathWorks file exchange) with color schemes from <http://colorbrewer2.org/> to generate color maps for all PVA plots except for PVA amplitude, which we display in grayscale (Fig. 1G). For display, either the visual cue position (for trials with a complex visual scene and trials with a single visual bar) or the virtual visual scene position (for trials in darkness) was offset by its median difference (circular distance) from the PVA. The offset occasionally changed between trials for the same fly as reported in previous work (16). Note that the PVA amplitude varied over time. This variability was quantified by first normalizing PVA amplitude by dividing it by

the mean PVA amplitude for each fly. The standard deviation of the normalized PVA amplitude across all trials from all 10 flies was then computed. The standard deviation was 0.4155 for the complex scene, 0.3390 for the one-bar scene, and 0.4482 for darkness. Input fluctuation may contribute to this variance, but the fluctuation was observed even if the fly was not flying in darkness ($\text{std} = 0.3694$), suggesting that input fluctuation alone may not fully explain PVA variability. EB network dynamics may be constantly modulated by other factors, including the fly's internal state, perhaps through signals carried by neuromodulators.

Calculation of the number of bumps

For each frame, a bump was defined as any contiguous set of ROIs with $\Delta F/F_0$ greater than a threshold value (defined to be the mean + 1 s.d.; Fig. 1I). We considered two additional methods of identifying bumps, including (i) a mean threshold method, in which the mean $\Delta F/F_0$ over the entire trial and across ROIs was used as a threshold (mean method, fig. S1C, top), and (ii) a Bayesian sampling method (Bayesian, fig. S1C second row). Under the Bayesian sampling method, the fluorescence level ($\Delta F/F_0$) was fit at each time point using a mixture of von Mises functions. Two assumptions were imposed during fitting: 1) Temporal coupling, which ensures that consecutive time points are likely to have similar fits, thereby reducing abnormal discontinuities in bump position, and 2) Sparseness, which imposes a cost for larger numbers of bumps, thereby reducing overfitting. We never found more than 2 bumps after fitting. The number of bumps for each trial was calculated as the mean of the number of bumps over time. See the Supplementary Text for more details of the Bayesian sampling method.

Calculation of the width of a bump

For the calculation of the width of the bump (full width at half-maximum, or FWHM), we first excluded every time point with either zero bumps or more than one bump. Then, for each time point, the bump width was calculated as the angular distance between two points at which the amplitude was half of its peak value. This was calculated from either raw data (Fig. 1J), or from the fit obtained from the Bayesian sampling method (fig. S1C third row).

Standard deviation of difference

PVA was offset to best match the orientation of the visual scene. Then, the standard deviation of the difference between PVA and the visual scene was calculated (fig. S1C bottom).

Correlation analysis

'Unwrapped' time series were first computed as a cumulative sum of all angular displacements (Fig. 1K). Pearson's correlation coefficients were then computed between two entire 'unwrapped' time series.

Offset between the estimated bump position and the pattern position

For a given trial, the offset between the absolute (to the experimenter) scene position and the PVA estimate was calculated as the median circular distance between the two positions. We then computed the angular average of this offset over all trials for each fly. Note that the visual arena, covering 330° , was mapped to 360° , as was the position of the scene. In Fig. 1L, the error

bar represents the circular s.d. (45) of offsets across trials. Data for the darkness condition was excluded as there was no significant correlation between the PVA and the estimated scene position.

Analysis of bump jump vs. flow

The dynamics of the activity bump in response to discontinuous movement of the visual stimulus (i.e., an abrupt visual input shift) were categorized as either jump or flow. Trials in which the fly did not fly during the visual stimulus displacement (defined from 0.5s before to 2s after the visual stimulus displacement) were excluded from further analysis. Then, each trial was analyzed using a Bayesian sampling method to fit data (see the Supplementary Text). Using this fit, the initial bump position was calculated as the circular mean of the bump positions between 0.5s before and at the moment of the visual jump. If the bump positions of the next 4 consecutive time points were within 45° of the current one, the bump was considered to be continuous, meaning that the activity in the current and the following 4 time points was considered to be part of the same, continuous bump. If this condition was met for the 2.5s after the visual stimulus jump, the trial was determined to exhibit a ‘bump flow’. If this condition was not met, it was determined that a new bump was detected. However, to conclude that this newly detected bump had exhibited a jump, the new bump was traced backward in time toward the moment of the visual stimulus jump using the same continuity checking method (within 45° threshold for 4 consecutive time points). If the new bump was created after the visual stimulus jump, then the trial was determined to exhibit a ‘bump jump’. Otherwise, the trial was discarded as a noisy trial.

Analysis of single spot optogenetic stimulation

To precisely determine the ROIs that were optogenetically stimulated, the same ROIs defined for images from even-line scans (imaging lines, Fig. 2C) were applied to images from odd-line scans (stimulation lines). Because the baseline fluorescence of odd line scans is nearly zero, any ROIs with mean F (fluorescence level) higher than mean + 3 s.d of baseline were determined to be affected by optogenetic stimulation. The number of ROIs affected by optogenetic stimulation varied from 2 to 4 depending on the position of the stimulation. The mean F (raw fluorescence level, not $\Delta F/F_0$ because of the extremely low baseline F_0) from even-line scans, i.e. imaging lines, of these ROIs was calculated over time. The F of an existing bump before stimulation was estimated as the average of the maximum four ROIs outside the stimulated region immediately before the stimulation onset. The number 4 was selected because the bump width defined by FWHM is between 3 (67.5°) and 4 (90°) ROIs (Fig.1). Trials in which the existing bump overlapped with stimulated ROIs were excluded from further analyses. For Fig. 2F, we computed the ratio of Fs between the last 0.475 s (50 frames) before and the last 0.475 s during optogenetic stimulation. We separately computed this ratio using ROIs of the existing bump, and using ROIs of the stimulated area. Because the distribution of the ratio is naturally skewed, we used a nonparametric statistical test (Wilcoxon rank-sum test). The same analyses were performed for control flies (fig. S2D), whose genotype was the same as those flies used in Fig. 1. Finally, all trials were sorted by the absolute distance (Fig. 2G) between the center of an existing bump (the position of the ROI with maximum F before stimulation) and the center of the optogenetic stimulation (the position of the ROI with max F in images from odd-line scans). Then the same ratio calculated for Fig. 2F was used to plot Fig. 2G.

Analysis of behavioral effect of optogenetic bump shifts (fig. S2E)

To detect behavioral effects in optogenetic bump shifting experiments, trials with extensive turning before stimulation were excluded — specifically, trials with circular variance of scene orientation greater than 0.2 for 5 s before the initiation of optogenetic stimulation. Trials were also excluded if the fly stopped flying during this 5 s period or the 2.5 s period of stimulation. Trials with an unsuccessful bump jump (as determined by the Bayesian sampling method, see above) were excluded. 47 trials were collected from 10 flies. Control trials were collected from the same flies. At the end of each optogenetic stimulation, the first 8-second segment with circular variance less than 0.2 was selected as a control trial, if it existed. If a new optogenetic stimulation began in the next 3s, it was excluded. The end of this 8-second segment was aligned to 0 s in fig. S2E. 98 control trials were collected from 10 flies. The mean turning amounts between 1 and 2.5 s after the initiation of optogenetic stimulation were used for statistical tests. For the x-axis of the bottom right plot in fig. S2E, the difference of mean bump positions (-1.5 – 0 s vs 1 – 1.5 s) was calculated from an application of the Bayesian sampling method. For the y-axis of the same plot, the visual bar position was used instead.

Analysis of two-spot simultaneous stimulation in Fig. 2

The ROI time courses in two simultaneously stimulated areas were compared across different power levels. The mean F profiles were normalized by the maximum F for each fly and combined to produce fig. S2G.

Analysis of drift in Fig. 3

To determine whether the bump stays at (or equivalently drifts from) or jumps away from the optogenetically stimulated position after the end of stimulation, the Bayesian sampling method (Supplementary Text) was applied to each trial in single spot optogenetic stimulation experiments (Fig. 2C). If the number of bumps in a given trial was larger than 1 for the first 10 frames after the end of stimulation, then it was determined either to be too noisy due to low SNR, or to be an unsuccessful bump shift by optogenetic stimulation. Both cases were excluded from further analysis. Trials with zero bumps at any time point between 0 and 1.5s after the end of stimulation were also determined to be noisy and excluded from analysis. If there was only one bump for 1.5 seconds after the end of stimulation, the bump was determined to have stayed at or drifted from the stimulated position in that trial. If the number of bumps was greater than 1, then the bump was determined to have changed its position abruptly. We calculated the proportion of bumps that drifted (Fig. 3D inset). The time course of bump position over time (Fig. 3B) was calculated as follows. The time course of each ROI was smoothed using a boxcar moving average method (21-frames or 3-volume span) to suppress the noise. Then, the PVA estimate was calculated at each time point (Fig. 3B top left). The PVA was shifted relative to the stimulation site, which was the mean PVA of the stimulated site from odd-line scans during the second half of the stimulation period (Fig. 3B bottom left). Finally, we computed the absolute value of this shifted PVA, and its mean value across trials (Fig. 3B top right). To generate the

drift distribution in Fig. 3C, we discretized the 3 s duration after the end of two-photon stimulation into 0.5 s bins, and we calculated the mean PVA for each bin. The difference of PVAs between consecutive bins was collected across all trials and flies, and histograms were generated. For flies with CsChrimson, the slow off-kinetics of CsChrimson held the bump position closer to the stimulated area for approximately 0.5 s after the cease of stimulation. Thus, a separate histogram was generated by comparing the PVA of this period to the PVA for the last 0.5 s of the stimulation period.

Estimation of laser power threshold necessary to make the bump jump

The goal of this analysis was to measure the power needed to make the bump jump from the initial position to a new position (Fig. 4E). Thus, we excluded all trials with non-jumping bumps between two consecutive optogenetic stimuli. To this end, the Bayesian sampling method (Supplementary Text) was used for the period between 0.5s before the start and 0.5s before the end of the second stimulus. If the bump flowed or stayed at the 1st stimulation position, the trials were excluded from further analysis. Then, the Bayesian sampling method was applied for the period between 0.5s before and 1.5s after the end of the second stimulus. If the number of bumps was not 1 (noisy) or the fly did not fly during that period (the bump may not be stable), the trial was excluded from further analysis. Trials with a clear jumping bump were used to estimate the threshold laser power for a jump. For a given pair of stimulation positions (e.g., profiles 1 and A), the number of trials with a successful bump jump had to be greater than 2 to be included in the data. The mean of the minimum 2 laser powers was determined as the estimate of the threshold power necessary to make the bump jump (Fig. 4F).

Estimation of input power using normalized bump amplitude

Greater input power or laser power induces a larger bump amplitude during stimulation. Thus, input power can be estimated by comparing the amplitude of the stimulated bump to the amplitude of the normal bump in the absence of input (Fig. 4, D and G, fig. S4, B, E, G and I). The bump amplitude during stimulation was determined as the average bump amplitude during the last 0.5 s of the 2nd stimulation. To normalize this amplitude, the reference bump amplitude without the stimulus (i.e., in darkness) seems sufficient. However, GCaMP expression level may be different across ROIs, meaning the ‘observed’ bump amplitude may be different across positions in EB even if the ‘effective’ bump amplitude is the same. Thus, the reference bump amplitude was obtained from the same ROI of the 2nd stimulation in darkness. To this end, the mean bump amplitude during the period between 1s and 1.5 s after the end of 2nd stimulation was used to normalize the bump amplitude during the stimulation. Trials in which the bump drifted away more than 30° after the end of the 2nd stimulus were excluded from this analysis. Finally, the normalized bump amplitudes from trials with minimum power were averaged for each profile pair (1-A, 2-A, 1-B, 2-B for protocol 1 in Fig. 4G, and 1-A, 1-B, and 1-C for protocol 2 in fig. S4E).

Simulation of models

For the models with extreme connectivity profiles (Fig. 4D), the parameters were first selected to generate a bump of zero-to-zero width of 90°. For the globally-connected model, which had two

parameters, one parameter was determined to fix the width of the bump, and the other parameter was swept to try and fit the optogenetic results. This sweep failed to fit the physiological results for any value. The model parameters for Fig. 4D were $J_0 = -8$, $J_1=6$ (see Supplementary Text for parameters). In the local model, the model parameters for Fig. 4D were $\alpha= 3$, $\beta=20$, $D=0.1$ (see Supplementary Text for parameters). The input to the model was assumed to have the shape of a von Mises (VM) function. We then observed the model's behavior while varying the width, power (estimated from the normalized bump amplitude), and instantaneous visual shift. Tracking the activity maxima, responses were grouped into three categories: No response (the maxima remained at the initial position), flow (the maxima varied continuously), and jump (discontinuous positional switch of maxima over time) (Fig. 4D, gray, red, green respectively). Input phase diagrams were obtained for various instantaneous visual shift distances.

Reproduction of the probability function for a jump

The visual input to the network is not likely to be perfectly identical over trials, suggesting the variance or noise in the input may underlie the non-step-like bump ratio plot across abrupt visual shift distances (Fig. 4C). Using the input-response phase diagrams, we investigated the effect of this potential variance of input. We explored two dimensions: width and strength. First, we assumed a simple Gaussian distribution along these two dimensions. Combinations of different means and standard deviations of a two-dimensional Gaussian function were tested. Two areas (jump and flow) of each abrupt shift distance were weighted-summed with a given Gaussian distribution function. Inputs with no effect (gray area of Fig. 4D) were excluded from further analysis. The weighted sum of values that resulted in jump and flow were compared to reproduce the jump probability plot. If the difference of the jump probability at each input shift distance between the model and the physiology was less than 0.1, it was determined to be included as a viable parameter set. There were many viable parameter sets: The particular example shown in the fig. S4B had the following parameters: input width of mean 86 and s.d. 50, input strength of mean 1.1 and s.d. 0.13. We repeated the same procedure for the global model and we were not able to find a parameter set that reproduced the plot in Fig. 4C. Note that, under different assumptions, such as slow quasistatic input dynamics, more generous criteria for viable parameter sets, or a log-normal rather than Gaussian distribution, we were able to reproduce the jump probability plot with the global model as well.

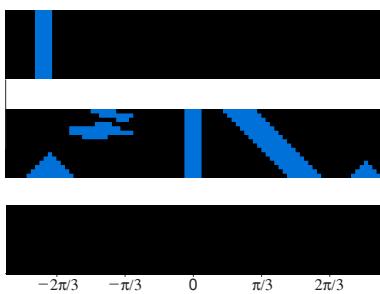
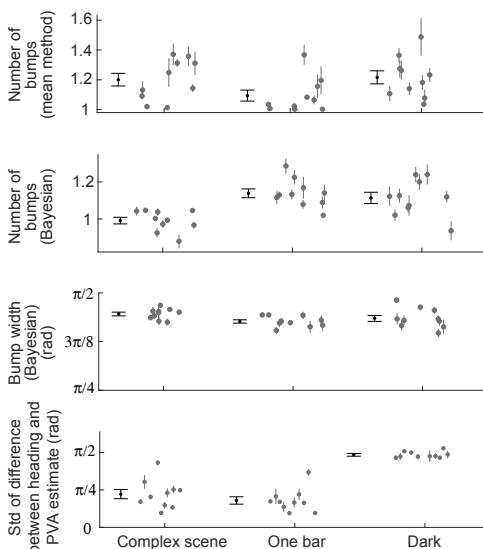
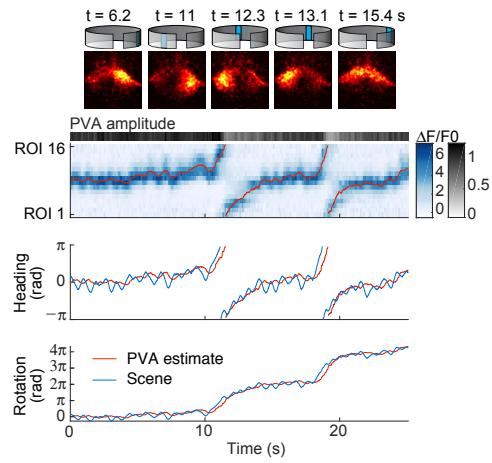
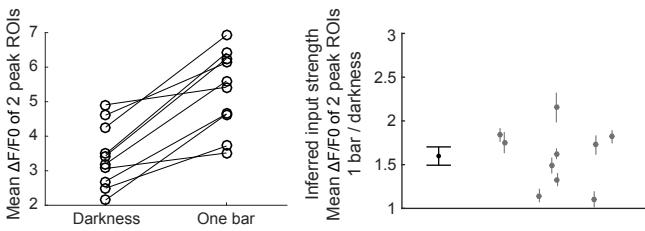
Fig. S1**A****C****B****D**

Fig. S1

(A) Three visual patterns used in the study. (B) Neural response to a single bar. The convention is the same as in Fig. 1G. (C) Population statistics obtained using different methods. Top row, number of bumps (mean method, see Materials and methods). Second row, number of bumps (Bayesian sampling method, see Materials and methods). Third row, FWHM (Bayesian sampling method, see Materials and methods). Bottom row: standard deviation of difference between PVA and heading (or scene orientation) (D) Left panel, average bump amplitude in different environments. At each time point, the $\Delta F/F_0$ of the two ROIs with the maximum $\Delta F/F_0$ were averaged. This value was then averaged across all time points of all trials for each of the dark and one-bar visual conditions for each fly. Right panel, ratio of bump amplitude between dark and one-bar conditions shown at left. The bump amplitude for the one-bar condition of each trial was compared to the mean bump amplitudes from all dark trials. Vertical bars, s.e.m for each fly from 10 trials. Overall population mean = 1.6, s.e.m = 0.105 (n = 10 flies).

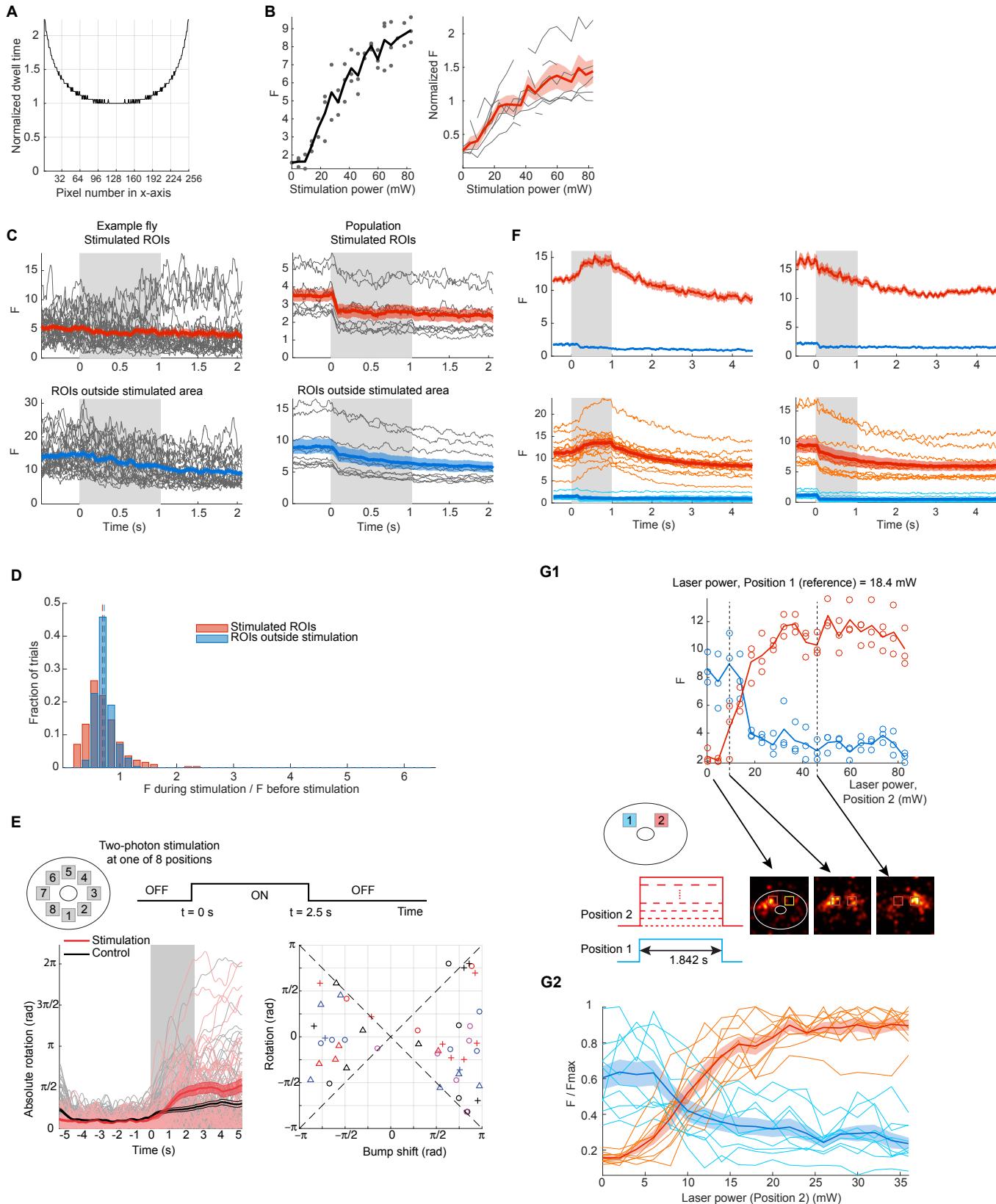
Fig. S2

Fig. S2

(A) Because the resonant galvo mirrors follow a sinusoidal trajectory, the laser exposure to the sample is spatially inhomogeneous. Shown is the two-photon laser dwell time at each pixel relative to the minimum. (B) Dose-response curve. Left panel, an example curve from a fly. 19 levels of laser power evenly divided in a 0 - 83 mW range were used to stimulate the most superficial part of the EB. Each power level was repeated three times. Black dots, single trial. Data points were excluded from the analysis if the fly was flying, or if the stimulation overlapped with an existing activity bump. Black line, average. Right panel, population curve. Grey lines, dose-response curve of each fly, normalized by natural F, which is defined as the mean of the highest 1% of F in an ROI, which, in turn, had the maximum mean F in darkness for 50 seconds performed before the experiment. Discontinuous lines represent rejected data points.

(C) The same as Fig. 2D, but CsChrimson was not expressed in E-PG neurons. See Movie S2. The small drop after initiation of the optogenetic stimulation ($t = 0$ s) is because of noise induced by the blue light from the vertical stripe in the LED arena, which leaked into the photomultiplier tube (PMT) but was turned off at $t = 0$ s. (D) The same as Fig. 2E, but CsChrimson was not expressed in E-PG neurons. Wilcoxon ranksum test, $p = 0.112$. (E) Effect of optogenetic stimulation on the fly's stripe fixation behavior. Left, absolute value of fly's turning amount when optogenetic stimulation was applied (red, $n = 47$ trials from 10 flies) versus not applied (black, $n = 98$ trials from 10 flies). Gray box, period of optogenetic stimulation. Only trials with stable fixation for 5 s before optogenetic stimulation were analyzed. Flies showed significant turning compared to control conditions. Thick lines with shade: Mean \pm sem, (Wilcoxon ranksum test: $p=0.00018052$ for $1\text{ s} < t < 2.5\text{ s}$). Right, relationship between the amount of bump shift (PVA) and fly's turning. Markers with the same shape and color are from the same fly. Note that the variance of turning increases with the distance of the bump shift. Dashed lines: identity and inverse identity lines. (F) Extended time plot of mean F of 4 ROIs with maximum $\Delta F/F_0$ (red), or minimum $\Delta F/F_0$ (blue). ROIs at each time point may vary as the bump may drift. Top left, with CsChrimson (40 trials from a fly). Top right, without CsChrimson (40 trials from a fly). Grey boxes, stimulation period. Thick colored lines, mean. Shaded area, s.e.m. Bottom row, population data. Bottom left, from 13 flies with CsChrimson. Orange lines, mean F of maximum 4 ROIs across all trials from each fly. Thick red line and shaded area, population average and s.e.m. Cyan lines, mean F of minimum 4 ROIs. Thick blue line and shaded area, population average and s.e.m. Bottom right, from 10 flies without CsChrimson. Same convention. The higher fluorescence before stimulation ($t < 0$) compared to the fluorescence after stimulation ($t > 1$) in all 4 panels is because of noise induced by the blue light from the vertical stripe in the LED arena, which leaked into the PMT but was turned off at $t = 0$ s. (G1) Competition between two groups of optogenetically stimulated neurons. Top, a single fly example. Circles, individual trials. One red and one blue circle are paired, though the pairs are not indicated for visual clarity. Blue dots and line, fluorescence level of ROIs stimulated by a fixed laser power (18.4 mW). Red dots and line, fluorescence level of the other ROIs stimulated by various laser powers (x-axis). Bottom, snapshots of competition with different power levels. Left red rectangle in each snapshot, position of fixed power stimulation. Right yellow or red rectangle, position of stimulation with varying laser powers. (G2) Population average. Thin cyan lines, fluorescence level from individual flies measured from the position with fixed laser power. Thick blue line and shade, mean and s.e.m. of population. Thin orange lines, fluorescence level from individual flies measured from the position with varying laser power. Thick red line and shade, mean \pm s.e.m. of population.

Fig. S3

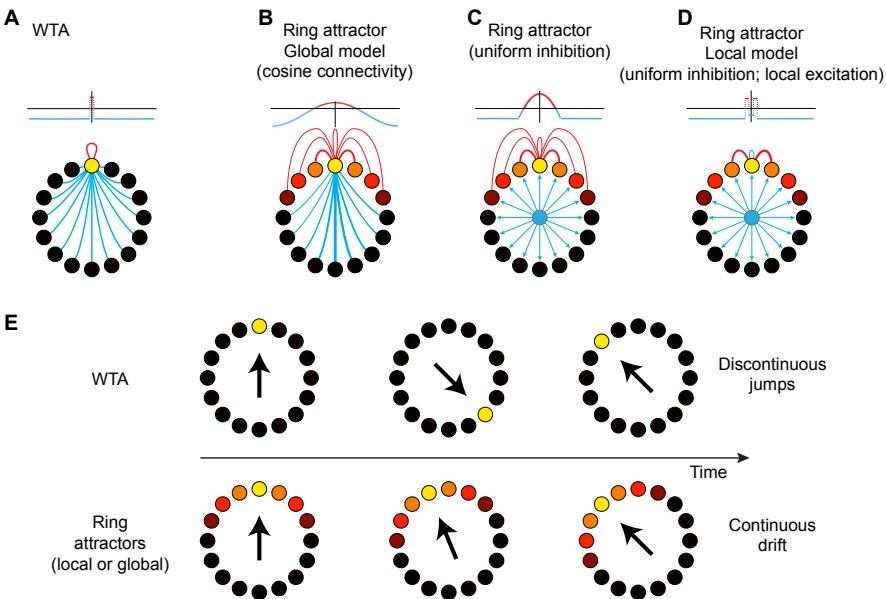


Fig. S3

Here we describe simple models of persistent localized activity. In order to simplify the modeling, we consider rate models for the neurons composing the neuronal populations and assume first order ODEs for their temporal evolution. We also assume rotational symmetry along a circle for the neurons: $\tau \partial_t f_n(t) = -f_n(t) + \phi \left(\sum_{m=0}^{N-1} K_m f_{n-m}(t) \right)$, where f_n are the firing rates, n runs over the neurons, the functions ϕ_n describe the interaction between these neurons, and the indices have implicit periodic boundary conditions. The non-linearity ϕ , and the convolution kernel K further constrain the dynamics. These terms correspond to cell autonomous neural operation and effective recurrent connectivity respectively. We explored three different connectivity structures, limit cases that each have distinct dynamical properties. **(A)** Winner take all (WTA): $K_1 = \dots = K_{N-1}$. This network produces a persistent winner, but without any special relationship between nearest neighbors. Thus, spontaneous state transitions induced by noise produce spatially discontinuous jumps of activity (e.g., top row). Circles in the schematized networks represent active neurons. Yellow, high activity. Black, sub-threshold potential. Red lines, excitatory connections. Blue lines, inhibitory connections. Line thickness, synaptic weight. **(B)** Ring attractor with global cosine interaction: K_n varies in a smooth fashion with n . Conventions for the schematized networks are the same as in **A**. We call this model a ‘global model’ for simplicity. **(C)** Ring attractor with uniform inhibition. Conventions for the schematized networks are the same as in **A**. The blue-colored neuron at the center of the diagram depicts a postulated inhibitory neuron, whose activity is proportional to the summated activity of all neurons. **(D)** Ring attractor with strictly local excitation and uniform inhibition: $K_2 = \dots = K_{N-2}$. We focus on this nearest neighbor model rather than the model in **B** for mathematical tractability. See Supplementary Text for mathematical formulation. In this study, we refer to this model as a ‘local model’ for simplicity. Conventions for the schematized networks are the same as in **C**. **(E)** Dynamics predicted by the different models. Unlike WTA networks (top, see **A**), ring attractor networks (bottom, see **B** to **D**) feature interactions between neighbors, which lead to smooth noise-induced transitions of activity (i.e., drift) between neighboring units.

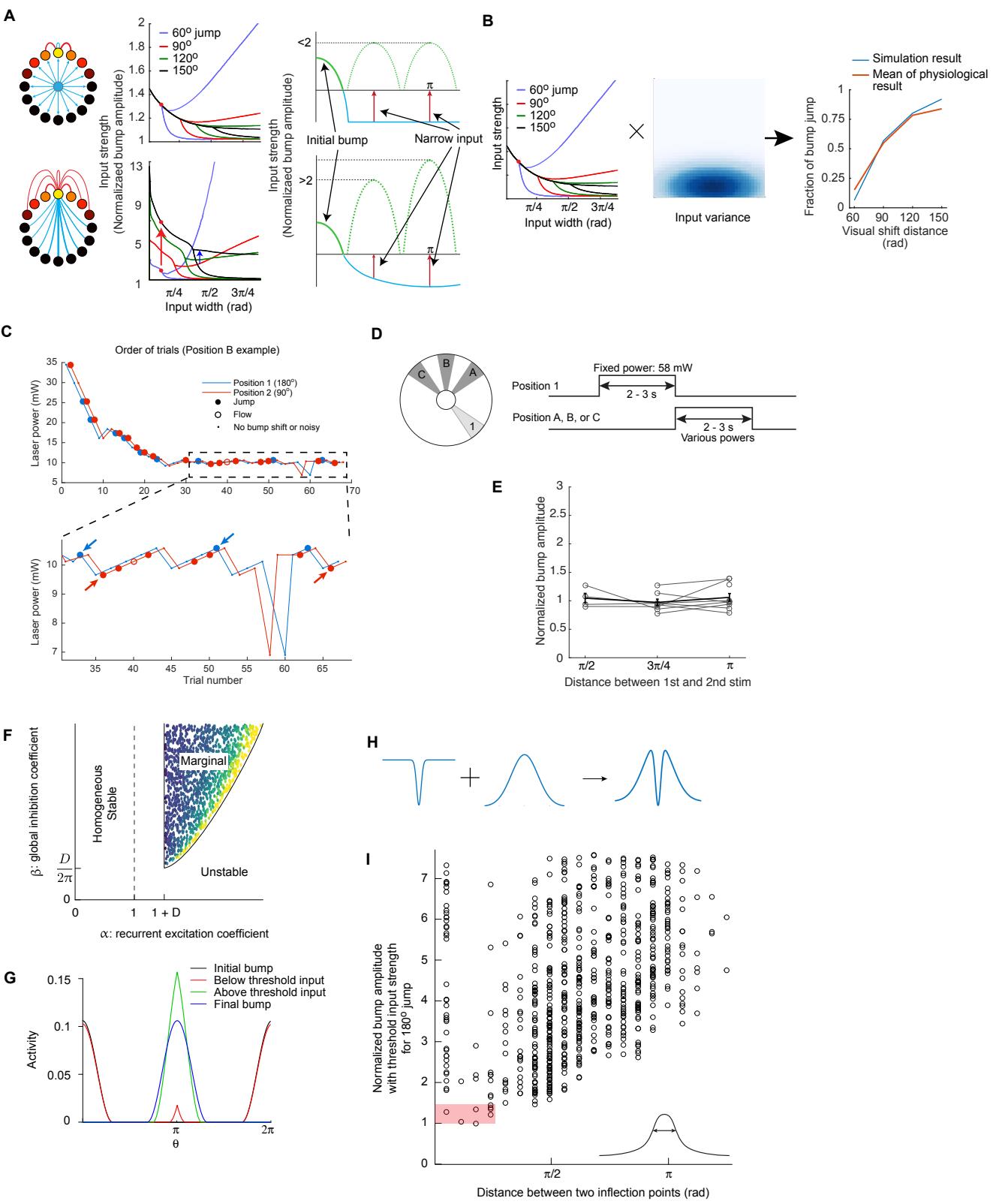
Fig. S4


Fig. S4

(A) Quantitative differences between the predictions of two extreme models. Top row, local model (fig. S3D). Input-response phase diagrams (Fig. 4D) are overlaid. The input strength required to transition from flow to jump decreases over input shift distances at all input widths. Red dot, predicted threshold input strength to induce the bump to jump with narrow input width of 22.5°. Note that this strength did not vary with all input shift distances from 90° to 150°. Top right, the cartoon visualizes the relationship between the threshold input strength, input shift, and bump amplitude for the local model. For example, when a narrow input was applied at 90° distance from the initial bump position with input strength at threshold for bump jump, the bump amplitude will be less than 2 for the local model. The strength can be inferred by taking relative bump amplitude compared to the one without input (Materials and methods). Importantly, for local model, this inferred strength should be the same for all input shift distances. The length of red arrows indicates the input power required to generate spikes. Bottom row, a global model (fig. S3B). Overlaid input-response phase diagram. The blue arrow illustrates clear inversion of the threshold input strength required for the bump to jump with an input shift distance at around 90° input width. Red dots and a red arrow, increasing threshold input strength required to induce a bump in response to different input shift distances from 60° to 150°. Bottom right, the model predicts that the jump of the activity bump requires strong input (>2) and it increases with the input shift distance, unlike the local model. (B) An example of simulated input variance that can reproduce the jump probability observed in physiological experiments (Fig. 4C). Input phase diagrams (left panel) were multiplied point-by-point across four input shift distances with a probability distribution function of input (the center panel). Right panel, reproduction of the jump probability. (C) Example execution of the sequential optogenetic stimulation protocol. To estimate the stimulation strength necessary to make the bump jump from various distances (1, 2) to a fixed position (B), and to avoid bias between two distances (1, 2), laser powers were paired for both distances (Materials and methods). (D) Schematic of an experiment to probe the threshold input strength for a jump at three distances. (E) Input strength, estimated by normalized bump amplitude. Convention is the same as Fig. 4G, but for three distances using the protocol in D. See Materials and methods. Wilcoxon ranksum test between $\pi/2$ and $3\pi/4$, $p = 0.57$. Between $\pi/2$ and π , $p = 0.84$. Between $3\pi/4$ and π , $p = 0.2$. (F) Parameter phase diagram for the local model (fig. S3D). See the Supplementary Text for a mathematical formulation. The “marginal” area represents a region of parameter space that ensures a unique persistent bump even without input. The “homogeneous” area represents states where the steady state profile of E-PG population activity is flat. The “unstable” area represents states in which the population activity diverges in time. Colored dots represent the input strength required for the bump jump at distance. Yellow dots represent very high threshold. Dark dots represent low threshold. Physiological and optogenetic data showing threshold less than 2 suggest that the connectivity of E-PG neurons, if locally connected, is far from the yellow area. Note that for the yellow dots near the boundary, the input strength needed to move the bump can be analytically computed to be approximately 6. α , recurrent excitation coefficient. β , global inhibition coefficient. D , diffusion coefficient. See Supplementary Text for more details. (G) Bump profile in the presence of input in a local model. Black line: a bump profile without input. Red line: a bump profile with a weak narrow input placed at a displacement of 180° from the existing bump. Note that there are two bumps. Green line: a bump profile with a narrow, supra-threshold input strength. Note that the original bump (black) disappeared, and only one bump remains. Blue line: a bump profile after the input is removed after jump. See the Supplementary Text for more details. (H)

Schematic of generating connectivity profiles using two von Mises (VM) functions. Two VM functions have the same center, but different width and amplitude. The summed function will have an additional bias parameter, for a total of 5 free parameters. (I) Input strength (normalized bump amplitude) for a jump of distance π across different connectivity widths. The connectivity width is defined as the distance between the two inflection points of the summed connectivity profile. Red rectangle, experimentally constrained space. See the Supplementary Text for more details.

MOVIES S1 – S11

Movie S1 Activity bump shifts its position in response to two-photon optogenetic stimulation. Top, images from simultaneous optogenetic stimulation and two-photon calcium imaging were divided into an imaging part (even lines, top left, smoothed with a Gaussian filter), and stimulation part (odd lines, top right, raw). Yellow rectangle, the position of optogenetic stimulation. Rectangle becomes red during stimulation (Fig. 2, A to C). Bottom left, behavioral recording of a tethered flying fly during the experiment, showing the vertical bar under closed-loop control. Bottom right, a color plot of F (not $\Delta F/F_0$). Red dashed horizontal lines, start and end of optogenetic stimulation.

Movie S2 Activity bump is not affected by two-photon stimulation without CsChrimson expressed in E-PG neurons.

Same convention as for Movie S1.

Movie S3 to S5 Mutual suppression between E-PG neurons during simultaneous stimulation of E-PG neurons in two parts of EB.

Same convention as for Movie S1. Top left, two powerboxes were simultaneously stimulated. The power of the left box was fixed (18.4mW), while that of the right box varied (3, 0mW; 4, 9.2mW; 5, 46mW). The video is slowed down to highlight mutual suppression.

Movie S6 Activity bump may drift after optogenetic stimulation.

See Fig. 3. Same convention as for Movie S1.

Movie S7 Activity bump may flow in response to a sudden position change of a vertical bar. Top left, raw two-photon imaging of calcium dynamics with GCaMP6f. Bottom left, behavioral recording of a tethered flying fly in closed loop control of a vertical bar. Right, color plot of $\Delta F/F_0$. Red dashed horizontal line, time of jump.

Movie S8 Annotated bump flow.

The same data as Movie S7, but only showing data around the time of the abrupt visual shift. Slowed down 10 times.

Movie S9 Activity bump may jump in response to abrupt shift of a stripe.

Same convention as for Movie S7.

Movie S10 An example trial of sequential optogenetic stimulation spaced 180° apart.

See Fig. 4E and Materials and methods. Same convention as for Movie S1.

Movie S11 An example trial of sequential optogenetic stimulation spaced 135° apart.

See fig. S4, D and E and Materials and methods. Same convention as for Movie S1.

Supplementary Text — Theory

Contents

1 Attractor models	2
2 Stationary states	2
3 Bump diffusion without input	5
4 Dynamics in the presence of localized input	6
4.1 Local model	6
4.2 Global interaction model	8
5 Input phase diagram	8
6 Numerical exploration of the parameters	8
7 Inference of bump configurations	9
7.1 Likelihood function and posterior distribution	9
7.2 Sampling	10
7.3 Neighbor coupling	11

1 Attractor models

The persistent activity observed on the E-PG neurons is the result of the interactions of many neurons within a recurrent circuit. Anatomical studies provide valuable informations on the structure of this circuit (17). But much less is known at the level of functional interactions. We wish to study *in silico*, the dynamical properties of several classes of recurrent circuits in order to understand the key components underlying the activity of the E-PG neurons. In order to simplify the modeling, we consider rate models for the N neurons composing the E-PG neuron population and assume first order ODEs for their temporal evolution:

$$\tau \partial_t f_n(t) = -f_n(t) + \phi_n(f_1(t), \dots, f_N(t)) \quad (1)$$

where the index n runs over the N neurons. The E-PG neurons are positioned around the ring shape structure of the ellipsoid body (EB) and we did not notice any significant property of the dynamics depending on the absolute position on the circuit. Hence, we will assume rotation symmetry along the ring. This means the dynamical equation (1) can be rewritten:

$$\tau \partial_t f_n(t) = -f_n(t) + \phi \left(\sum_{m=0}^{N-1} K_m f_{n-m}(t) \right) \quad (2)$$

where the indices have implicit boundary conditions. Since the dynamics have stationary states and neither clockwise nor anti-clockwise rotations seems to be favored, we can assume that K is mirror symmetric: $K_n = K_{-n}$.

Under these general assumptions, two elements further constraint the dynamics: the non-linearity function or $f - I$ curve ϕ , and the convolution kernel K . These terms corresponding to cell autonomous neural operation and effective recurrent connectivity respectively.

We are going to explore three different connectivity structures:

- **winner take all (WTA)** $K_1 = \dots = K_{N-1}$
- **nearest neighbor interaction (local model)** $K_2 = \dots = K_{N-2}$
- **global interaction (global model)** K_n varies in a smoothly with n .

These three different limit cases have distinct dynamical properties that we shall explore. The goal is to understand the functional connectivity structure of the E-PG neuron network.

2 Stationary states

We first study the shape of the stationary states of the three previous cases. We know that the persistent activity observed in EB takes the form of a single continuous bump of activity. We first see how this activity can be accounted

for by the models. We consider unless otherwise specified that the $f - I$ curve is threshold linear:

$$\phi(I) = (I + 1)\Theta(I + 1) = [I + 1]_+ \quad (3)$$

where Θ is the Heaviside function. This hypothesis mainly assumes neuron activity are not saturating in the normal range of activity (see fig. S1D).

Winner take all stationary states The stationary states f_h^0 verify:

$$f_n^0 = \phi((K_0 - K_1)f_n^0 + K_1S) \quad (4)$$

where $S = \sum_{n=0}^{N-1} f_n^0$. The homogeneous state exists and is stable when

$$(N - 1)K_1 < 1 - K_0 \quad (5)$$

$$K_1 > K_0 - 1 \quad (6)$$

The marginal state whereby only one neuron is active exists and is stable when

$$K_0 < 1 \quad (7)$$

$$K_1 < K_0 - 1 \quad (8)$$

Note that although we know several E-PG neurons are active to form the bump of activity, at this stage it could be that this activity is a result of an underlying WTA that feeds the E-PG neurons in a feed-forward manner.

Local model stationary states It is possible to solve the stationary states and stability conditions for the local model. We will first describe this solution. In a second part, we will take the continuous limit and show that the local model can constitute a continuous attractor displaying finite width bump stationary states.

In the discrete case, stationary states can be written as branches of cosine. We rewrite the dynamical equation as follows:

$$\tau \partial_t f_n = -f_n + \left[\alpha f_n + D(f_{n-1} + f_{n+1} - 2f_n) - \beta \sum_{m=0}^{N-1} f_m + 1 \right]_+ \quad (9)$$

The uniform state verifies:

$$f_k^0 = \frac{1}{1 - \alpha + N\beta} \quad (10)$$

It exists and is stable when:

$$1 - \alpha + \beta N > 0 \quad (11)$$

$$1 - \alpha + 2D(1 - \cos 2\pi/N) > 0 \quad (12)$$

When the local excitation α and wide inhibition β are strong enough, a localized bump of activity forms. This bump has the shape of a branch of cosine. As an example, we will consider a bump vanishing at 0 in which activity spans M neurons. In that case, the bump can be written as:

$$f_n^0 = A(\sin(\omega n - \phi) + \sin \phi) \quad (13)$$

The stationary state condition imposes:

$$2 \sin \omega/2 = \sqrt{\frac{\alpha - 1}{D}} \quad (14)$$

$$A(\alpha - 1) \sin \phi = \beta S - 1 \quad (15)$$

$$DA(\sin(\omega - \phi) + \sin(\phi)) < \beta S - 1 \quad (16)$$

$$\sin((M + 1)\omega - \phi) = -\sin \phi \quad (17)$$

where S is the total activity. The inequality (16) which imposes a negative input current for the neuron $n = 0$ can be shown after some algebra using the other stationary conditions to be equivalent to:

$$2\pi/\omega - 1 < M + 1 < 2\pi/\omega \quad (18)$$

hence M is uniquely defined and the bump has a width as close as possible to the period of the cosine $2\pi/\omega$. Finally, one can compute S and find the amplitude A :

$$\tan \phi = \frac{\sin(M + 1)\omega}{\cos(M + 1)\omega - 1} \quad (19)$$

$$S/A = \frac{\sin \omega \cos \phi}{1 - \cos \omega} + (M + 1) \sin \phi \quad (20)$$

$$1/A = (1 - \alpha) \sin \phi + \beta \left(\frac{\sin \omega \cos \phi}{1 - \cos \omega} + (M + 1) \sin \phi \right) \quad (21)$$

This state exists when:

$$2(1 - \cos 2\pi/N) < \frac{\alpha - 1}{D} < 4 \quad (22)$$

$$1/A > 0 \quad (23)$$

We also performed the linear stability analysis around the bump stationary state. The eigenmodes (except one) are sines and cosines centered on the maximum of the stationary bump. We showed that when the bump exists with the previous conditions, it is linearly stable. Moreover, the stationary states with multiple bumps are unstable.

Finally, it is possible to take the continuous limit $N \rightarrow +\infty$ to that discrete model. In that case, the system constitutes a continuous attractor model. The evolution equation for neurons labeled by a continuous variable θ between 0 and 2π with periodic boundary condition can be written as follows:

$$\partial_t f(\theta, t) = -f(\theta, t) + \left[\alpha f(\theta, t) + D \partial_{\theta\theta} f(\theta, t) - \beta \int_0^{2\pi} f(\theta, t) d\theta + 1 \right]_+ \quad (24)$$

Taking the continuous limit of the discrete model, or solving the above equations, we show that the stationary states can be written:

$$f^0(x) = A(1 + \cos \omega x) \quad (25)$$

with

$$\omega = \sqrt{\frac{\alpha - 1}{D}} \quad (26)$$

$$1/A = 1 - \alpha + 2\pi\beta/\omega \quad (27)$$

and the existence now writes:

$$\alpha > 1 + D \quad (28)$$

$$1/A > 0 \quad (29)$$

The phase diagram for the continuous limit is shown on fig. S4F.

Global model stationary states It is generally difficult to extract stationary states in the general K_i case. But a stereotypical case has been studied in great detail (11, 26). The Ben-Yishai model treats the continuous limit case where $K(\theta) = J_0 + J_1 \cos(\theta)$. The phase diagram has been studied in great detail. The stationary states in the marginal phase are branches of cosine:

$$f^0(\theta) = \Theta(A_1 \cos(\theta - \theta_0) - A_0) \quad (30)$$

We see from the previous paragraphs that the stationary states are not very discriminant regarding the shape of the activity bump. In order to characterize each model, we need to study their dynamical properties and in particular response to time varying inputs.

3 Bump diffusion without input

In the absence of tuned input, the bump of activity evolves under the influence of noise. In order to study the influence of noise, we add a gaussian white noise to the input current of the neurons. The evolution equation can now be written:

$$\tau \partial_t f_n(t) = -f_n(t) + \left[\sum_{m=0}^{N-1} K_m f_{n-m}(t) + \sigma \eta(t) \right]_+ \quad (31)$$

where $\eta(t)$ is a white noise: $\langle \eta(t) \rangle = 0$ and $\langle \eta(t)\eta(t') \rangle = \delta(t - t')$.

For the winner take all model, increasing the amplitude of the noise leads to an abrupt transition for most of the parameters. At low noise, the winner is very stable as expected. However there is a level of noise above which there is not clear winner and all the neurons are noisy. Just below this threshold, it takes again an exponentially long time to observe transitions. This is only

at the boundary of the marginal phase (when β is as low as possible) that the situation changes. In that case, the time to observe transitions is not too high (fig. S3E top).

For continuous attractor (both local and global), weak noise induces a diffusion of the bump position (fig. S3E bottom).

4 Dynamics in the presence of localized input

In this section, we study the activity of the network in the presence of a localized input:

$$\partial_t f(\theta, t) = -f(\theta, t) + [K * f(\theta, t) + 1 + I\delta(\theta)]_+ \quad (32)$$

4.1 Local model

In the continuous limit, the localized input induces a discontinuity in the slope of the activity profile.

In such a situation, a bump of persistent activity can be sustained at a position different from the localized input (fig. S4G, red profile). And at the position of the input, a weaker bump of activity with a discontinuous slope is present. As we will see, this situation arises when the input is weaker than a threshold depending on the parameter of the model. When the input is stronger than this threshold, the first bump disappear and a bump of high amplitude arise at the position of the input (fig. S4G, green profile).

We have analytically studied this phenomenon both for the discrete version and for the continuous limit. It is more easily formulated for the continuous version though, and we detail the analytical treatment below only for that case.

We study the two bump solutions to the following equation:

$$f(\theta, t) = \left[\alpha f(\theta, t) + D\partial_{\theta\theta} f(\theta, t) - \beta \int_0^{2\pi} f(\theta, t) d\theta + 1 + I\delta(\theta) \right]_+ \quad (33)$$

The first bump, which does not overlap with the input can be written:

$$f_1(\theta) = A(\cos(\omega\theta + \phi_1) + 1) \quad (34)$$

and has its activity between $\theta = \phi_1 - \pi/\omega$ and $\theta = \phi_1 + \pi/\omega$

The second bump can be written:

$$f_2(\theta) = A(\cos(|\omega\theta| + \phi) + 1) \quad (35)$$

and has its activity between $\theta = (\pi - \phi)/\omega$ and $\theta = -(\pi - \phi)/\omega$

Notice that the prefactor A is the same for both bumps because the curvatures of the bumps of activity must be equal when the activities reach zero.

The total activity $S = \int_0^{2\pi} f(\theta, t) d\theta$ has the following expression:

$$S = 2A(2\pi - \phi - \sin \phi)/\omega \quad (36)$$

Integrating the equation (33) around $\theta = 0$, one obtains:

$$I = 2AD\omega \sin \phi \quad (37)$$

The stationary equation at the points where the activity reaches zero gives:

$$A(1 - \alpha + 2\beta(2\pi - \phi - \sin \phi)/\omega) = 1 \quad (38)$$

One can then obtain the input I as a function of the phase ϕ :

$$I = \frac{\sin \phi}{(2\pi - \phi_0 - \sin \phi_0)\beta/(\alpha - 1) - \omega/2} \quad (39)$$

This input is maximum for a phase ϕ_0 verifying:

$$2\pi - \phi + \tan \phi = \frac{(\alpha - 1)\omega}{2\beta} \quad (40)$$

For the profile to be stable, one can show that the phase has to verify $\pi/2 < \phi_0 < \phi < \pi$. The maximum input I_0 able to sustain two bumps verifies:

$$I_0 = \frac{\alpha - 1}{\beta(1 - 1/\cos \phi_0)} \quad (41)$$

In order to understand the behavior of the bump amplitude with and without input, we consider the domain of the phase space close to the bifurcation to the marginal phase. In that region, $1 - \alpha + 2\pi\beta/\omega$ is small and we pose: $1 - \alpha + 2\pi\beta/\omega = \epsilon$. We then have:

$$\tilde{\phi}_0^3 \equiv (\pi - \phi_0)^3 = \frac{3\epsilon\omega}{2\beta} \quad (42)$$

$$I_0 = \frac{\alpha - 1}{2\beta}(1 - \tilde{\phi}_0^2/4) \quad (43)$$

If we now consider there is only one bump in the presence of the input I_0 , the activity profile can be written as:

$$f_s(\theta) = A_s(\cos(|\omega\theta| + \phi_s) + 1) \quad (44)$$

The ratio of the maximum activity with and without input has the following expression:

$$R = A_s(\cos \phi_s + 1)\epsilon \quad (45)$$

When the input strength is I_0 , one obtains:

$$R \approx 6 \quad (46)$$

which closely corresponds to yellow dots in fig. S4F.

4.2 Global interaction model

The amplitude of localized input current being infinite in the continuous limit, one can write:

$$\partial_t f(\theta, t) = -f(\theta, t) + [K * f(\theta, t) + 1]_+ + I\delta(\theta) \quad (47)$$

We are effectively in the case of the cosine input studied previously (26). If the input is applied quasi-stationarily, we are in the case of the weak input limit, and according to (26), the original bump should flow toward the position of the input.

5 Input phase diagram

For the local and global model, we explored the dynamics of the bump in present of shifted input position.

We initialize the state of the system so that the center of the bump is positioned at the origin of the angular positions. In a second phase, an input of a given width w and amplitude a is abruptly applied. We then look for the changes in the activity profiles. The situations might arise:

- if the input is too localized and its amplitude is not high enough, neurons do no reach the firing threshold and no change in activity is observed.
- if the input is wide enough, the active neurons receive a gradient of input and the bump moves progressively toward the center of the applied input.
- finally, if the input is strong enough, the activity will be dominated by the input and a new bump will develop at the center of the input. This will generate a jump of the bump from its initial position to the new position.

6 Numerical exploration of the parameters

So far, we compared two limit cases of the connectivity: the so-called local model whereby the neurons outside the bump have all the same potential and the global model whereby this potential varies smoothly and decreasing as a function of the distance from the bump.

We would like to explore a larger set of models encompassing the two previous limit scenarios. In order to do that, we consider that the connectivity is a sum of two von Mises functions (fig. S4H):

$$K(\theta) = a_1 f(\theta|0, \kappa_1) - a_2 f(\theta|0, \kappa_2) + C \quad (48)$$

where $f(\theta|\mu, \kappa)$ is the von Mises probability density function of parameter μ, κ (μ gives the center of the distribution and κ the concentration of the probability around μ) at the position θ . The global model is recovered when $\kappa_1, \kappa_2 \rightarrow 0$ and the local model is obtain when $\kappa_1, \kappa_2 \rightarrow +\infty$.

It means the attractor model is defined by 5 parameters (κ_1 , κ_2 , a_1 , a_2 and C). We chose to sample this space uniformly.

The width is defined as the position of the first inflection point of the connectivity function.

We study the ratio of amplitude of the bump with and without input in more general convolution kernels. We consider kernels of sum of Von Mises functions:

This gives 5 parameters to explore. On top of that, we vary the width of the input:

$$I(\theta) = I_1 V(\theta; \kappa_I) + O_I \quad (49)$$

The simulation result indicates strong dependence of the ratio of the bump amplitude on the connectivity width (fig. S4I).

7 Inference of bump configurations

The population vector average PVA describes the position and amplitude of a single bump of activity on a circle. As soon as several bumps coexist, as it is the case for the jump and flow experiment assays, one need to rely on other quantification means to describe the activity configurations. Since the calcium measurements are inherently noisy, we chose to rely on a Bayesian inference framework. We first describe the likelihood function which describe the probability of bump configurations. Then we present the sampling method which allows us to infer the configurations given the calcium intensities.

7.1 Likelihood function and posterior distribution

We consider a trial dataset composed of intensities for each time point t and ROI r : $I_{t,r}$. We suppose that the signal is a superposition of several bumps modeled as von Mises functions. Moreover, we suppose that the noise is gaussian and uncorrelated both in time and position (see below for an account of time dependence of the calcium signal). This gives the following expression for the intensities:

$$I_{t,r} = \sum_{i=1}^{N_t} af(\theta_r | \mu, \kappa) + \eta_{t,r} \quad (50)$$

where $f(x|\mu, \kappa)$ is the von Mises probability density function of parameter μ, κ (μ gives the center of the distribution and κ the concentration of the probability around μ) at the position x . $\eta_{t,r}$ is an uncorrelated white noise of variance σ . N_t is the number of bumps at time t . Notice that the width of the bumps are approximatively: $1/\sqrt{\kappa}$.

This means that the probability to observe the measured signal $I_{t,r}$ knowing the bump parameters can be written as follows:

$$p(I_{t,r} | N_t, \mu, a) = \frac{1}{N} \prod_{t=1}^T \prod_{r=1}^R \exp \left(-\frac{(I_{t,r} - \sum_{i=1}^{N_t} af(\theta_r | \mu, \kappa))^2}{2\sigma^2} \right) \quad (51)$$

where N is a normalization constant.

Since the signal intensities are known, but not the parameters of the bumps, one needs to "invert" the previous probability. This is done by applying the Bayes rule:

$$p(N_t, \mu, a | I_{t,r}) = \frac{p(I_{t,r} | N_t, \mu, a) p(N_t, \mu, a)}{p(I_{t,r})} \quad (52)$$

The denominator is only a normalization constant since it does not depend on the parameters we are looking for.

We still have to determine the prior: $p(N_t, \mu, a)$. In order to keep a non-informative prior on the bump amplitudes, we are going to take an improper prior (*ie* non normalized probability):

$$p(N_t, \mu, a) \propto e^{-\alpha \sum_{t=1}^T N_t} \prod \Theta(a) \quad (53)$$

where Θ is the Heaviside function. α constraints the number of bumps in the system and introduces a sparsity bias. We used $\alpha = 4$.

Here, we can make a connection with the statistical physics of gas. Formally, the bumps in our system play the role of particles and they interact through the measurements: particles will compete for places where the signal intensities are high. Notice that in our system, the number of bumps, or particles, is not fixed: this is the situation of the grand-canonical ensemble in statistical physics. As we wrote, α is equivalent to the opposite of the chemical potential of the system. This equivalence will be useful for the sampling method (see below).

We can now write the *log* of the posterior, up to an additive constant as:

$$\mathcal{L} = \log p(I_{t,r} | N_t, \mu, a) + \log p(N_t, \mu, a) \quad (54)$$

$$= - \sum_{t,r} \frac{\left(I_{t,r} - \sum_{i=1}^{N_t} af(\theta_r | \mu, \kappa) \right)^2}{2\sigma^2} - \alpha \sum_{t=1}^T N_t \quad (55)$$

The last equation implicitly assumes that the bump amplitudes are positive.

7.2 Sampling

One way to infer the position of the bumps is to look for the most probable configuration. This entails the maximization of log-posterior probability \mathcal{L} from eq. (55). This function is not convex and gradient descent methods would fail to reliably maximize it. We then rely on a sampling method with the help of the Metropolis algorithm.

Making again the connection with the physics of gas, we are trying to simulate the grand-canonical ensemble by a Monte Carlo method.

A state is defined by a set of N_t , μ , κ and a . The state transitions should allow to reach any one of these states from any other (ergodicity). We define three kinds of transitions at any time t :

- **Creation or deletion of bumps** For creation, a position is chosen at random between 0 and 2π . The bump is initialized with an amplitude 2 and $\kappa = 2.5$. This bump is then added to the list of bumps already present at this time t . For deletion, a bump is chosen at random and simply removed from the list of bumps.
- **Change of position** A random number with normal distribution of standard deviation $\sigma_\mu = 0.5$ is chosen and added to the position of one of the bumps.
- **Change of amplitude** A random number with normal distribution of standard deviation 1 is drawn and added to the amplitude of one of the bumps.

Choosing the transition probabilities appropriately is important. They should verify the detailed balance equations in order for the algorithm to converge toward the right distribution. But they should also be carefully chosen to maximize the convergence speed. In fact, the transition probabilities do not exactly verify the detailed balance equations but we have not noticed any significant biases in the inferred bumps.

Creation transitions were chosen with probability 0.01. Deletion were chosen N_t times more often (to compensated for the entropic factor (46)). We create bumps with amplitude from a given distribution but we delete them irrespective of their amplitudes, which is a potential source of bias. However, this should be very negligible since the amplitude will diffuse fast toward a value minimizing the energy of the bump.

Change of positions are chosen with probability 0.5 and amplitude changes are chosen in the other cases.

We then apply the Metropolis algorithm: we choose a move; if the it increases the likelihood, we accept it, if it does not, we accept it only with probability:

$$\exp(\beta \Delta \mathcal{L}) \quad (56)$$

where β is the inverse temperature of the system. β is chosen high ($\beta = 5$), so that the temperature is low and at any time point close to convergence, the state is close to the maximum likelihood.

7.3 Neighbor coupling

The observed calcium signal is correlated in time, so that we expect *a priori* the bump states to be also correlated in time as well. We incorporated this information in the prior.

We add another term to eq. (55):

$$\mathcal{L}' = \mathcal{L} + J_c \sum_t \sum_{i=1}^{N_t} \sum_{j=1}^{N_{t+1}} \exp \left(-\frac{(\mu_{i,t} - \mu_{j,t+1})^2}{2\sigma_c^2} \right) \quad (57)$$

where σ_c is chose to be 0.5 (half a ROI) and J_c is the intensity of coupling (we chose $J_c = 1.8$).

References

1. R. I. Wilson, Early olfactory processing in *Drosophila*: Mechanisms and principles. *Annu. Rev. Neurosci.* **36**, 217–241 (2013). [doi:10.1146/annurev-neuro-062111-150533](https://doi.org/10.1146/annurev-neuro-062111-150533) [Medline](#)
2. A. Borst, M. Helmstaedter, Common circuit design in fly and mammalian motion vision. *Nat. Neurosci.* **18**, 1067–1076 (2015). [doi:10.1038/nn.4050](https://doi.org/10.1038/nn.4050) [Medline](#)
3. J. J. Knierim, K. Zhang, Attractor dynamics of spatially correlated neural activity in the limbic system. *Annu. Rev. Neurosci.* **35**, 267–285 (2012). [doi:10.1146/annurev-neuro-062111-150351](https://doi.org/10.1146/annurev-neuro-062111-150351) [Medline](#)
4. E. Aksay, I. Olasagasti, B. D. Mensh, R. Baker, M. S. Goldman, D. W. Tank, Functional dissection of circuitry in a neural integrator. *Nat. Neurosci.* **10**, 494–504 (2007). [doi:10.1038/nn1877](https://doi.org/10.1038/nn1877) [Medline](#)
5. H. S. Seung, How the brain keeps the eyes still. *Proc. Natl. Acad. Sci. U.S.A.* **93**, 13339–13344 (1996). [doi:10.1073/pnas.93.23.13339](https://doi.org/10.1073/pnas.93.23.13339) [Medline](#)
6. R. Chaudhuri, I. Fiete, Computational principles of memory. *Nat. Neurosci.* **19**, 394–403 (2016). [doi:10.1038/nn.4237](https://doi.org/10.1038/nn.4237) [Medline](#)
7. K. Wimmer, D. Q. Nykamp, C. Constantinidis, A. Compte, Bump attractor dynamics in prefrontal cortex explains behavioral precision in spatial working memory. *Nat. Neurosci.* **17**, 431–439 (2014). [doi:10.1038/nn.3645](https://doi.org/10.1038/nn.3645) [Medline](#)
8. J. S. Taube, The head direction signal: Origins and sensory-motor integration. *Annu. Rev. Neurosci.* **30**, 181–207 (2007). [doi:10.1146/annurev.neuro.29.051605.112854](https://doi.org/10.1146/annurev.neuro.29.051605.112854) [Medline](#)
9. J. S. Taube, R. U. Muller, J. B. Ranck Jr., Head-direction cells recorded from the postsubiculum in freely moving rats. I. Description and quantitative analysis. *J. Neurosci.* **10**, 420–435 (1990). [Medline](#)
10. B. L. McNaughton, F. P. Battaglia, O. Jensen, E. I. Moser, M. B. Moser, Path integration and the neural basis of the ‘cognitive map’. *Nat. Rev. Neurosci.* **7**, 663–678 (2006). [doi:10.1038/nrn1932](https://doi.org/10.1038/nrn1932) [Medline](#)
11. R. Ben-Yishai, R. L. Bar-Or, H. Sompolinsky, Theory of orientation tuning in visual cortex. *Proc. Natl. Acad. Sci. U.S.A.* **92**, 3844–3848 (1995). [doi:10.1073/pnas.92.9.3844](https://doi.org/10.1073/pnas.92.9.3844) [Medline](#)
12. W. E. Skaggs, J. J. Knierim, H. S. Kudrimoti, B. L. McNaughton, A model of the neural basis of the rat’s sense of direction. *Adv. Neural Inf. Process. Syst.* **7**, 173–180 (1995). [Medline](#)
13. K. Zhang, Representation of spatial orientation by the intrinsic dynamics of the head-direction cell ensemble: A theory. *J. Neurosci.* **16**, 2112–2126 (1996). [Medline](#)
14. X. Xie, R. H. Hahnloser, H. S. Seung, Double-ring network model of the head-direction system. *Phys. Rev. E* **66**, 041902 (2002). [doi:10.1103/PhysRevE.66.041902](https://doi.org/10.1103/PhysRevE.66.041902) [Medline](#)

15. A. Peyrache, M. M. Lacroix, P. C. Petersen, G. Buzsáki, Internally organized mechanisms of the head direction sense. *Nat. Neurosci.* **18**, 569–575 (2015). [doi:10.1038/nn.3968](https://doi.org/10.1038/nn.3968) [Medline](#)
16. J. D. Seelig, V. Jayaraman, Neural dynamics for landmark orientation and angular path integration. *Nature* **521**, 186–191 (2015). [doi:10.1038/nature14446](https://doi.org/10.1038/nature14446) [Medline](#)
17. T. Wolff, N. A. Iyer, G. M. Rubin, Neuroarchitecture and neuroanatomy of the *Drosophila* central complex: A GAL4-based dissection of protocerebral bridge neurons and circuits. *J. Comp. Neurol.* **523**, 997–1037 (2015). [doi:10.1002/cne.23705](https://doi.org/10.1002/cne.23705) [Medline](#)
18. T. W. Chen, T. J. Wardill, Y. Sun, S. R. Pulver, S. L. Renninger, A. Baohan, E. R. Schreiter, R. A. Kerr, M. B. Orger, V. Jayaraman, L. L. Looger, K. Svoboda, D. S. Kim, Ultrasensitive fluorescent proteins for imaging neuronal activity. *Nature* **499**, 295–300 (2013). [doi:10.1038/nature12354](https://doi.org/10.1038/nature12354) [Medline](#)
19. M. H. Dickinson, Haltere-mediated equilibrium reflexes of the fruit fly, *Drosophila melanogaster*. *Philos. Trans. R. Soc. London Ser. B* **354**, 903–916 (1999). [doi:10.1098/rstb.1999.0442](https://doi.org/10.1098/rstb.1999.0442) [Medline](#)
20. N. C. Klapoetke, Y. Murata, S. S. Kim, S. R. Pulver, A. Birdsey-Benson, Y. K. Cho, T. K. Morimoto, A. S. Chuong, E. J. Carpenter, Z. Tian, J. Wang, Y. Xie, Z. Yan, Y. Zhang, B. Y. Chow, B. Surek, M. Melkonian, V. Jayaraman, M. Constantine-Paton, G. K.-S. Wong, E. S. Boyden, Independent optical excitation of distinct neural populations. *Nat. Methods* **11**, 338–346 (2014). [doi:10.1038/nmeth.2836](https://doi.org/10.1038/nmeth.2836)
21. M. Yoshida, M. E. Hasselmo, Persistent firing supported by an intrinsic cellular mechanism in a component of the head direction system. *J. Neurosci.* **29**, 4945–4952 (2009). [doi:10.1523/JNEUROSCI.5154-08.2009](https://doi.org/10.1523/JNEUROSCI.5154-08.2009) [Medline](#)
22. G. Major, D. Tank, Persistent neural activity: Prevalence and mechanisms. *Curr. Opin. Neurobiol.* **14**, 675–684 (2004). [doi:10.1016/j.conb.2004.10.017](https://doi.org/10.1016/j.conb.2004.10.017) [Medline](#)
23. K. S. Kakaria, B. L. de Bivort, Ring attractor dynamics emerge from a spiking model of the entire protocerebral bridge. *Front. Behav. Neurosci.* **11**, 8 (2017). [doi:10.3389/fnbeh.2017.00008](https://doi.org/10.3389/fnbeh.2017.00008) [Medline](#)
24. S. Amari, Dynamics of pattern formation in lateral-inhibition type neural fields. *Biol. Cybern.* **27**, 77–87 (1977). [doi:10.1007/BF00337259](https://doi.org/10.1007/BF00337259) [Medline](#)
25. P. Song, X. J. Wang, Angular path integration by moving “hill of activity”: A spiking neuron model without recurrent excitation of the head-direction system. *J. Neurosci.* **25**, 1002–1014 (2005). [doi:10.1523/JNEUROSCI.4172-04.2005](https://doi.org/10.1523/JNEUROSCI.4172-04.2005) [Medline](#)
26. D. Hansel, H. Sompolinsky, in *Methods in Neuronal Modeling: From Ions to Networks*, C. Koch, I. Segev, Eds. (MIT Press, 1998), pp. 499–567.

27. A. M. Turing, The chemical basis of morphogenesis. *Philos. Trans. R. Soc. London Ser. B* **237**, 37–72 (1952). [doi:10.1098/rstb.1952.0012](https://doi.org/10.1098/rstb.1952.0012)
28. H. Ko, S. B. Hofer, B. Pichler, K. A. Buchanan, P. J. Sjöström, T. D. Mrsic-Flogel, Functional specificity of local synaptic connections in neocortical networks. *Nature* **473**, 87–91 (2011). [doi:10.1038/nature09880](https://doi.org/10.1038/nature09880) [Medline](#)
29. M. Weliky, K. Kandler, D. Fitzpatrick, L. C. Katz, Patterns of excitation and inhibition evoked by horizontal connections in visual cortex share a common relationship to orientation columns. *Neuron* **15**, 541–552 (1995). [doi:10.1016/0896-6273\(95\)90143-4](https://doi.org/10.1016/0896-6273(95)90143-4) [Medline](#)
30. H. Ozeki, I. M. Finn, E. S. Schaffer, K. D. Miller, D. Ferster, Inhibitory stabilization of the cortical network underlies visual surround suppression. *Neuron* **62**, 578–592 (2009). [doi:10.1016/j.neuron.2009.03.028](https://doi.org/10.1016/j.neuron.2009.03.028) [Medline](#)
31. S. P. Mysore, A. Asadollahi, E. I. Knudsen, Global inhibition and stimulus competition in the owl optic tectum. *J. Neurosci.* **30**, 1727–1738 (2010). [doi:10.1523/JNEUROSCI.3740-09.2010](https://doi.org/10.1523/JNEUROSCI.3740-09.2010) [Medline](#)
32. U. Hanesch, K. F. Fischbach, M. Heisenberg, Neuronal architecture of the central complex in *Drosophila melanogaster*. *Cell Tissue Res.* **257**, 343–366 (1989). [doi:10.1007/BF00261838](https://doi.org/10.1007/BF00261838)
33. C. Y. Lin, C.-C. Chuang, T.-E. Hua, C.-C. Chen, B. J. Dickson, R. J. Greenspan, A.-S. Chiang, A comprehensive wiring diagram of the protocerebral bridge for visual information processing in the *Drosophila* brain. *Cell Rep.* **3**, 1739–1753 (2013). [doi:10.1016/j.celrep.2013.04.022](https://doi.org/10.1016/j.celrep.2013.04.022) [Medline](#)
34. B. D. Pfeiffer, T.-T. B. Ngo, K. L. Hibbard, C. Murphy, A. Jenett, J. W. Truman, G. M. Rubin, Refinement of tools for targeted gene expression in *Drosophila*. *Genetics* **186**, 735–755 (2010). [doi:10.1534/genetics.110.119917](https://doi.org/10.1534/genetics.110.119917) [Medline](#)
35. H. Luan, N. C. Peabody, C. R. Vinson, B. H. White, Refined spatial manipulation of neuronal function by combinatorial restriction of transgene expression. *Neuron* **52**, 425–436 (2006). [doi:10.1016/j.neuron.2006.08.028](https://doi.org/10.1016/j.neuron.2006.08.028) [Medline](#)
36. A. Guo, L. Li, S. Z. Xia, C. H. Feng, R. Wolf, M. Heisenberg, Conditioned visual flight orientation in *Drosophila*: Dependence on age, practice, and diet. *Learn. Mem.* **3**, 49–59 (1996). [doi:10.1101/lm.3.1.49](https://doi.org/10.1101/lm.3.1.49) [Medline](#)
37. J. D. Seelig, M. E. Chiappe, G. K. Lott, A. Dutta, J. E. Osborne, M. B. Reiser, V. Jayaraman, Two-photon calcium imaging from head-fixed *Drosophila* during optomotor walking behavior. *Nat. Methods* **7**, 535–540 (2010). [doi:10.1038/nmeth.1468](https://doi.org/10.1038/nmeth.1468) [Medline](#)
38. J. D. Seelig, V. Jayaraman, Feature detection and orientation tuning in the *Drosophila* central complex. *Nature* **503**, 262–266 (2013). [Medline](#)

39. G. Maimon, A. D. Straw, M. H. Dickinson, Active flight increases the gain of visual motion processing in *Drosophila*. *Nat. Neurosci.* **13**, 393–399 (2010). [doi:10.1038/nn.2492](https://doi.org/10.1038/nn.2492) [Medline](#)
40. M. B. Reiser, M. H. Dickinson, A modular display system for insect behavioral neuroscience. *J. Neurosci. Methods* **167**, 127–139 (2008). [doi:10.1016/j.jneumeth.2007.07.019](https://doi.org/10.1016/j.jneumeth.2007.07.019) [Medline](#)
41. A. Bahl, G. Ammer, T. Schilling, A. Borst, Object tracking in motion-blind flies. *Nat. Neurosci.* **16**, 730–738 (2013). [doi:10.1038/nn.3386](https://doi.org/10.1038/nn.3386) [Medline](#)
42. T. A. Pologruto, B. L. Sabatini, K. Svoboda, ScanImage: Flexible software for operating laser scanning microscopes. *Biomed. Eng. Online* **2**, 13 (2003). [doi:10.1186/1475-925X-2-13](https://doi.org/10.1186/1475-925X-2-13) [Medline](#)
43. J. P. Rickgauer, K. Deisseroth, D. W. Tank, Simultaneous cellular-resolution optical perturbation and imaging of place cell firing fields. *Nat. Neurosci.* **17**, 1816–1824 (2014). [doi:10.1038/nn.3866](https://doi.org/10.1038/nn.3866) [Medline](#)
44. A. M. Packer, L. E. Russell, H. W. Dalgleish, M. Häusser, Simultaneous all-optical manipulation and recording of neural circuit activity with cellular resolution in vivo. *Nat. Methods* **12**, 140–146 (2015). [doi:10.1038/nmeth.3217](https://doi.org/10.1038/nmeth.3217) [Medline](#)
45. P. Berens, CircStat: A MATLAB toolbox for circular statistics. *J. Stat. Softw.* **31**, 1–21 (2009). [doi:10.18637/jss.v031.i10](https://doi.org/10.18637/jss.v031.i10)
46. D. Frenkel, B. Smit, *Understanding Molecular Simulation: From Algorithms to Applications* (Academic Press, ed. 2, 2001).