Integration of sleep drive and navigation in *Drosophila*

Andres Flores Valle[1], **Pedro J. Gonçalves**[2,3,*], **and Johannes D. Seelig**[1,*]

[1]Center of Advanced European Studies and Research (caesar), Bonn, Germany
[2]Max Planck Research Group Neural Systems Analysis, Center of Advanced European Studies and Research (caesar), Bonn, Germany
[3]Computational Neuroengineering, Department of Electrical and Computer Engineering, Technical University of Munich, Germany
[*]pedro.goncalves@caesar.de
[*]johannes.seelig@caesar.de

ABSTRACT

During sleep, the brain undergoes dynamic and structural changes. In *Drosophila*, such changes have been observed in the central complex, a brain area important for sleep control and navigation. For navigation, the structure and function of the central complex suggest a relationship to ring attractors, networks that model head direction. This raises the question about a similar structure-function relationship for sleep control.
Motivated by experimental findings, we develop a ring attractor model that maintains activity at a setpoint in the face of plasticity. In this model, an integrator of sleep drive emerges with a functional role in stabilizing the network. Nevertheless, the network becomes unstable over extended time spent awake. Therefore, a sleep phase is introduced where autonomous attractor dynamics, related to its activity during wakefulness, reset the network. The proposed integration of sleep and head direction circuits captures features of their neural dynamics observed in flies and mice during wakefulness and sleep.

1 Introduction

Sleep affects many different brain functions such as cognition[1] or working memory[2] and sleep dysfunction is related to a range of diseases[3]. Sleep is observed across species and different hypotheses have been put forward to explain the function of sleep[4], among others for example reverse learning of spurious network states (that were created as a byproduct of intended memories)[5–7] or weakening of synapses (synaptic homeostasis hypothesis)[8]. The function of sleep is linked to the dynamic and structural changes that it induces in the brain[9], which in turn are monitored by sleep control circuits[10,11]. The circuits that control sleep are distributed over many different brain areas and cell types[10]. Thanks to the genetic tools[12,13] that allow dissecting neural circuits into small populations of genetically identified cell types, as well as more recently the fly connectome[14], *Drosophila* has emerged as a valuable model for sleep control[11,15–17]. For example, neural correlates of a subsystem of the sleep control circuit that measures sleep drive have been identified[17,18]. Specifically, sleep drive has been linked to so called ring neurons which arborize in concentric rings in the ellipsoid body, a substructure of the central complex[18] of the fly brain.

The central complex has also been shown to be important for navigation. In particular, ring neurons similar to the ones mentioned above are involved in visually guided navigational behaviors[19,20]. In this context, the structure and function of circuits in the central complex are reminiscent of ring attractor networks[21]. Such networks, which are well suited to encode a circular variable, have been suggested to underlie the encoding of head direction, originally in mammals[22] and more recently in flies[21]. In particular, in flies head direction has been found to be encoded in the activities of so-called wedge neurons.

It is currently unknown why the circuits for sleep drive and navigation are integrated in the same brain structure. The morphological similarity of the involved ring neurons and the spatial proximity of the circuits as well as the connectome[14], suggest however that they interact. Given the observed structural changes in the sleep-related ring neurons[18,23,24], this suggests that the head direction system needs to operate in the face of substantial synaptic and

functional changes in connected circuits.

Motivated by these experimental observations of plasticity[18,23,24], we here use theoretical modeling to discuss how sleep homeostasis and navigation can be understood as a combined system (see Figure 1A for a schematic summary). We show that ring neurons can be assigned the role of integrating homeostatic sleep drive (referred to in the following as 'homeostatic integrator', see Figure 1B), while at the same time maintaining an operational head direction system (wedge neurons) in the face of plasticity. To avoid saturation of the homeostatic integrator and to maintain the ring attractor network over extended periods of time, we introduce a sleep phase where plasticity is inverted compared to wakefulness and the resulting autonomous dynamics reset the ring attractor (see Figure 1). Overall, a homeostatic integrator emerges as a functional element of the ring attractor network.

We first implement this approach on a model with two neural populations, one excitatory and one inhibitory, and then expand it into a complete ring attractor network. We discuss how these models can integrate several experimental observations on the navigation and sleep homeostasis systems reported in the literature, such as autonomous dynamics of the head direction system during sleep. We further discuss several predictions of the model that potentially could be tested in experiments. This analysis contributes to an understanding of the generation and dynamics of sleep drive and links the control of sleep to sleep function.

2 Results

2.1 Correlates of ring attractor models in *Drosophila*

We briefly summarize the anatomy of the central complex ring and wedge (also called EPG[25]) neurons that underlies the models developed in the following. Figure S1A shows a ring neuron (red, downloaded from[14]) that projects in a concentric ring in the ellipsoid body, a substructure of the central complex. There are different classes of ring neurons: the sleep homeostatic integrator was identified in[18] as R2 neurons or as R5 neurons in[23]. Whether these are the same as R2 neurons with orientation-tuned receptive fields described in[20] is unclear, although sleep-related R2 neurons also respond to visual input[26]. Wedge neurons (green) also send projections into the ellipsoid body and all wedge neurons together tile the entire neuropil (grey wedges in Figure S1B)[27].

The connectivity between wedge and ring neurons assumed in the model is schematically illustrated in Figure S1B (similar to the connectivity reported in[28]). We assume that the connections in the model are direct (for example the recurrent connectivity of wedge neurons), although this assumption could be relaxed by including indirect connections[29].

Synaptic plasticity in ring and wedge neurons has been discussed in several papers[18,23,24,30,31]. Note that we here focus on plasticity in the connections from wedge to ring neurons, while leaving the connectivity in the opposite direction constant. This is opposite from[30,31], where plasticity from ring to wedge neurons is considered, while the other direction is left constant. The data in[18] is consistent with the plasticity occurring in either direction, while in our view the increase in firing rate of the inhibitory ring neurons with time spent awake seems to speak rather in favor of increasing plasticity in wedge to ring neurons.

As in previous work[25], we identify wedge neurons as the excitatory component in a ring attractor network with recurrent excitation, and ring neurons as the global inhibitory component (Figure S1C).

2.2 Two-population model with plasticity for ring and wedge neurons

The goal of our analysis is to construct a homeostatic sleep integrator from the dynamics of the head direction system. This is motivated by the experimental observation of increasing and decreasing structural and functional changes in homeostatic ring neurons during wakefulness and sleep[18], respectively, and their connection to the head direction system (see Figure S2). While the exact time course of this signal has not been determined in the fly, a homeostatic integrator is generally assumed to increase with time spent awake and decrease during sleep to a baseline (Figure 1B).

We first discuss a population model of wedge and ring neurons based on an excitatory-inhibitory network[32] (Figure 2A). The dynamics of (excitatory) wedge and (inhibitory) ring neurons are described by the following system

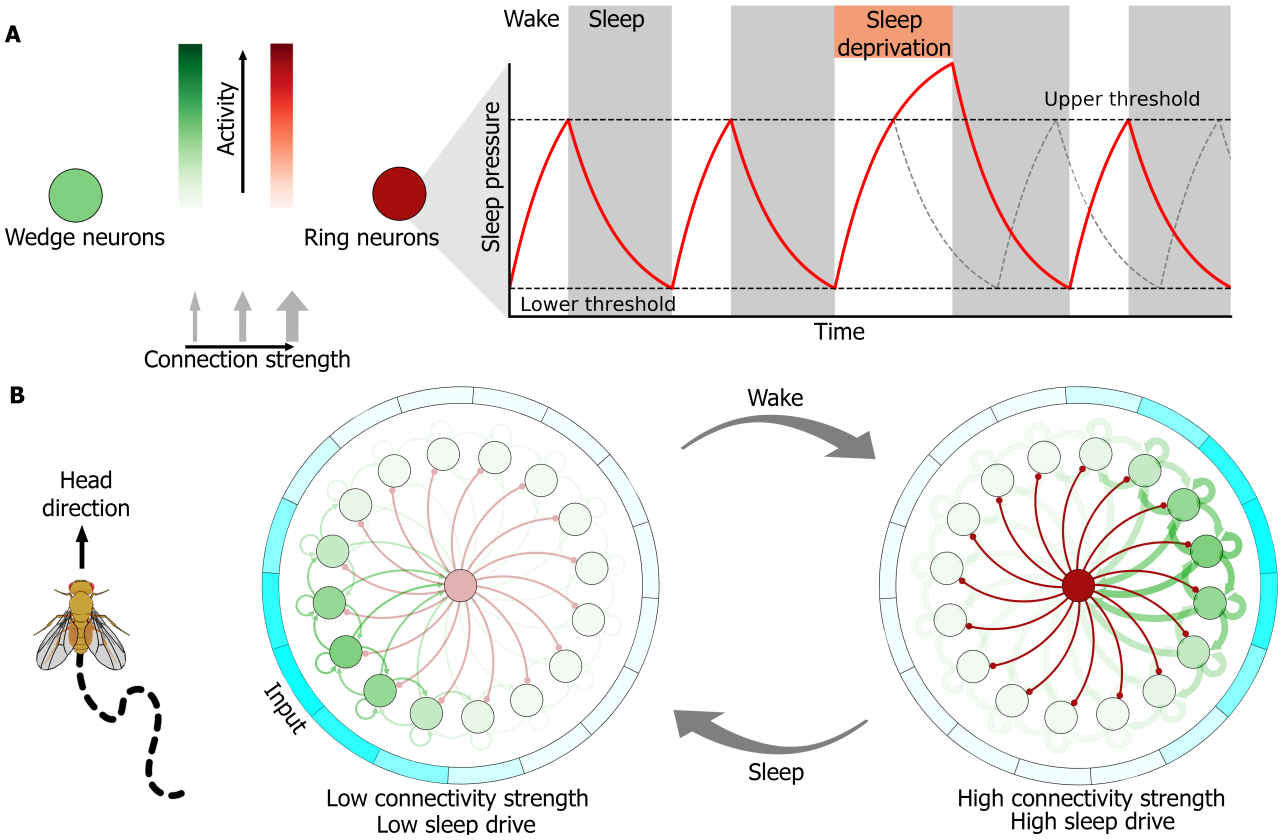

Figure 1. Schematic of the model. **A** Ring neurons encode sleep drive which increases with time spent awake and induces sleep when reaching an upper threshold. Sleep decreases sleep pressure to a lower threshold and the wake phase restarts. Increased sleep drive due to sleep deprivation leads to sleep rebound. **B** Ring attractor network between populations of wedge neurons (green) and populations of ring neurons (red). Wedge neurons encode the head direction of the fly through visual and idiothetic inputs (blue) in a bump-like activity profile. Ring neurons promote low sleep drive after sleep time with correspondingly low connectivity strength (left). Ring neurons promote high sleep drive when connectivity in the network is strengthened after wakefulness (right). Legends for colors and arrow width (connectivity strength) are shown in **A** (left)

of differential equations:

$$\begin{cases} \tau \dot{r}_E(t) = -r_E(t) + \Big[w_{EE}(t) r_E(t) - w_{EI} r_I(t) + \theta \Big]_+ \\ \tau \dot{r}_I(t) = -r_I(t) + \Big[w_{IE}(t) r_E(t) \Big]_+ , \end{cases} \tag{1}$$

where $r_E(t)$ and $r_I(t)$ are the firing rates at time t of wedge neurons and ring neurons, respectively, w_{AB} is the synaptic weight from population B to population A, θ is a constant background input onto wedge neurons, $[\cdot]_+$ represents a threshold-linear function to ensure positive-valued firing rates, and τ is the effective population time constant. We additionally introduce plasticity rules for the excitatory weights w_{EE} and w_{IE} (which we motivate below):

$$\begin{cases} \tau_{EE} \dot{w}_{EE}(t) = r_E(t) r_E(t) \\ \tau_{IE} \dot{w}_{IE}(t) = r_I(t) r_E(t) \Big(r_E(t) - r_0 \Big), \end{cases} \tag{2}$$

where τ_{EE} and τ_{IE} are time constants, and r_0 is a positive presynaptic threshold. While the first plasticity rule is a linear Hebbian rule, the second is a triplet rule with presynaptic threshold; its behavior is similar to the linear

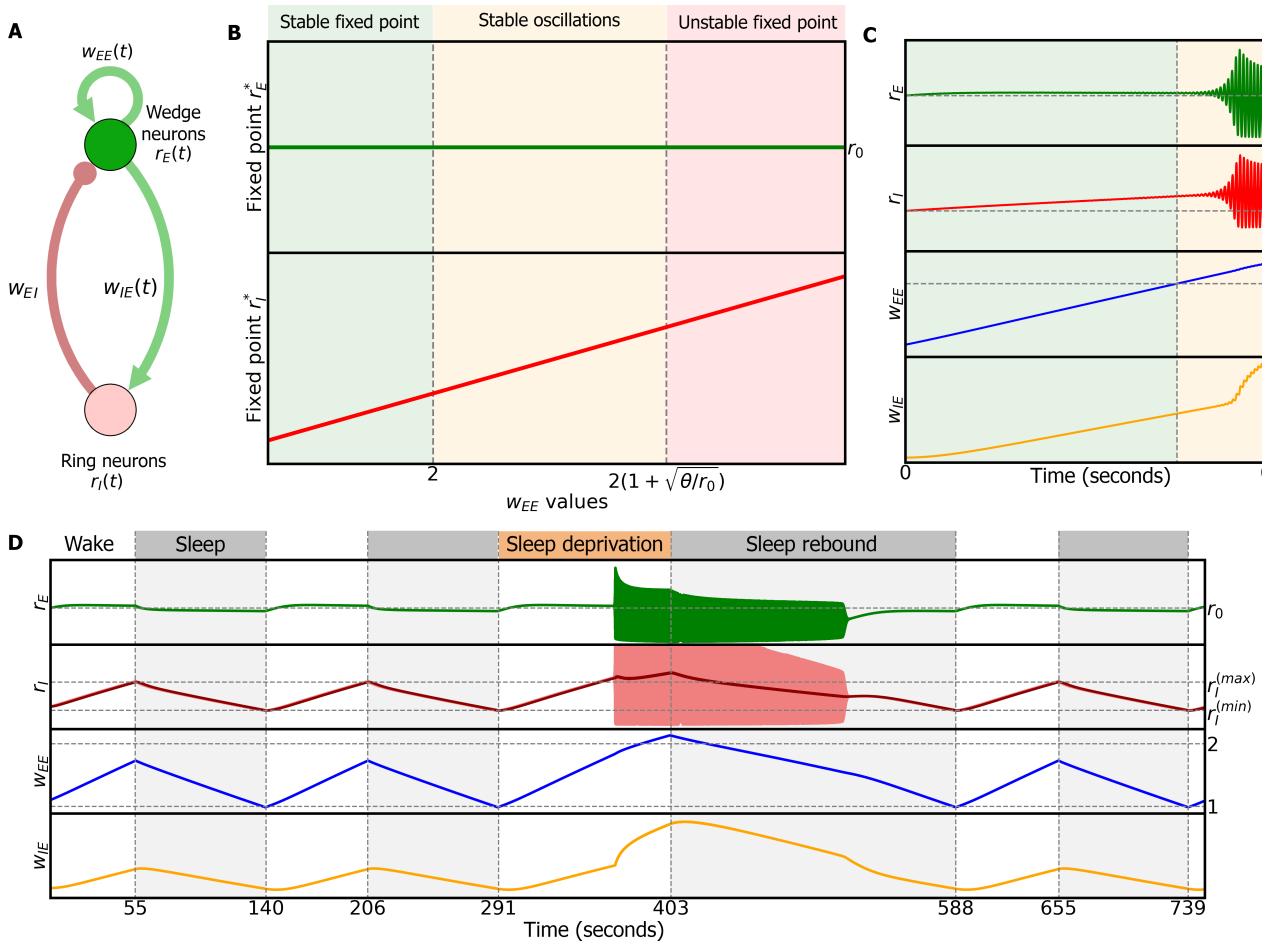

Figure 2. Two-population model. **A** Model describing the dynamics of the activities of excitatory wedge neurons and inhibitory ring neurons, r_E and r_I respectively, and the plasticity of the synaptic weights within the wedge population w_{EE} and from the wedge population to the ring population w_{IE}. Green connections are excitatory, red connections inhibitory. **B** Stability conditions with respect to w_{EE} (fast-timescale limit, see Methods). **C** Dynamics during wake phase. As w_{EE} grows, the system undergoes a Hopf bifurcation (a critical point where the system starts to oscillate) and both populations start to oscillate around a fixed point. **D** Dynamics with alternating sleep and wake phases. Wake phase produces LTP in w_{EE} and w_{IE} and increases the activity of ring neurons. Sleep produces LTD and reduces the activity of ring neurons. Extending the wake period produces sleep deprivation and results in stable oscillations. The subsequently required sleep period for resetting is longer (sleep rebound).

Hebbian rule with presynaptic threshold, but it has a quadratic dependency on the presynaptic activity while ensuring no change in w_{IE} if neural activity is zero (see derivation in Methods(4.3) and in[33]).

These plasticity rules are motivated by the fact that ring neurons monitor sleep drive by increasing their firing rate and by undergoing LTP over time spent awake[18]. We include these observations in the model with the following conditions: (i) The synaptic connection between wedge and ring neurons, w_{IE}, gets stronger during the wake phase. (ii) The firing rate of ring neurons r_I increases during the wake phase. Additionally, since wedge neurons encode the head direction of the fly with a persistent bump of activity that moves between subsets of wedge neurons[21], the activity of the entire population of wedge neurons is constant over time. This leads to the following condition: (iii) The firing rate of the entire population of wedge neurons evolves always to a constant setpoint, $r_E \longrightarrow r_0$. Note that activity is not constrained to the setpoint, but evolves towards it over time, and will deviate from the setpoint due to external (for example visual or self-motion related) input (Figure S4), consistent with experimentally observed

behavior-related changes in bump amplitude.[34]. Condition (iii) is in conflict with (ii) because as w_{IE} increases, ring neurons provide stronger inhibition to wedge neurons so that r_E (the head direction signal) decays as w_{IE} increases. Therefore, to compensate for the effect of LTP in w_{IE}, we introduce the following additional condition: (iv) The recurrent synaptic weight between wedge neurons, w_{EE}, gets stronger during the wake phase. LTP in the connections w_{EE} and w_{IE} must be coupled such that condition (iii) holds. All these conditions are fulfilled with the introduced plasticity rules, as discussed in section 2.3. We note that the dynamics of the plasticity rules are much slower than the dynamics of neural populations, so that $\tau_{EE}, \tau_{IE} >> \tau$.

2.3 Two-population model dynamics during wakefulness

We start by characterizing the dynamics of the model for fixed values of the weights w_{EE} and w_{IE} (fast-timescale limit; see Methods (4.4)). For the sake of the analysis, let us assume that w_{IE} is coupled to w_{EE} so that the dynamics of the model are characterized only by the value of w_{EE}. The fixed point for wedge neurons r_E^* remains constant, and the fixed point for ring neurons r_I^* changes with w_{EE} (Figure 2B). If $w_{EE} \leq 2$ (light green region), the fixed point (r_E^*, r_I^*) is stable, and both neural populations evolve towards these values. This makes wedge neurons to evolve towards a constant activity, r_0, in agreement with assumption (iii) . With increasing w_{EE}, the fixed point of ring neurons r_I^* also increases while the fixed point of wedge neurons r_E^* remains constant, as required by the above conditions (ii) and (iii). If w_{EE} increases further to $2 < w_{EE} \leq 2(1+\sqrt{\theta/r_0})$, the model enters a regime of stable oscillations (light orange region). In this regime, both neural populations oscillate around the fixed point with a frequency that changes with w_{EE} (see Figure S3A and Methods (4.4)). In addition, ring neurons increase their amplitude of oscillations as w_{EE} increases. Finally, when $w_{EE} \geq 2(1+\sqrt{\theta/r_0})$ (light red region), the fixed point is unstable and the activity of both populations diverges.

Figure 2C illustrates the dynamics of the full system, i.e. the activity of wedge and ring neurons, $r_E(t)$ and $r_I(t)$, as well as synaptic weights, $w_{EE}(t)$ and $w_{IE}(t)$. In the beginning (light green region), the fixed point of ring and wedge neurons is stable because $w_{EE} < 2$. When this boundary is crossed, the system enters the regime of stable oscillations (light orange region). Also, in the stable region of the simulation in Figure 2C, w_{EE}, w_{IE} and ring neuron activity r_I increase, while the activity of wedge neurons r_E remains constant as imposed by conditions (i)-(iii). w_{EE}, w_{IE} and r_I constitute therefore a measure of how far the network has moved from its initial state.

2.4 Two-population model dynamics during sleep

In order to reset the system back to its stable state ($w_{EE} < 2$) after prolonged activity (wakefulness), we introduce a sleep phase with inverted plasticity[5–7]. For this we assume that during sleep the recurrent connection between wedge neurons, w_{EE}, gets weaker (condition (vi)) through LTD[4,7] according to the following plasticity rule:

$$\tau_{EE}\dot{w}_{EE}(t) = -r_E(t)r_E(t). \tag{3}$$

The plasticity rule for w_{IE} is the same as in the wake phase. The impact of LTD on the model can be understood by inspecting Figure 2B: with decreasing value of w_{EE}, the fixed points become stable (light green) and the activity of ring neurons decreases (see Methods (4.4)). The switch between the wake and sleep phases is triggered by the activity of ring neurons, which have been identified as the integrator of homeostatic sleep drive in the fly[18,26]. We additionally apply a low-pass filter to the activities of ring neurons (see equation (9) in Methods) to remove possible oscillations and refer to this filtered signal in the following as 'switching signal'.

Figure 2D shows a simulation of the model combining subsequent wake (white regions) and sleep phases (grey). During wakefulness, w_{EE} and w_{IE} undergo LTP and the activity of ring neurons increases (light red line in second row in Figure 2D) and r_E is constant. When the switching signal (dark red line in second row) crosses an upper threshold, $r_I^{(max)}$, the model switches to sleep. During sleep, w_{EE} undergoes LTD due to the switch in plasticity, while the activity of ring neurons decreases. w_{IE} also undergoes LTD (note that the plasticity rule does not change) since it follows the trend of w_{EE} to impose the set-point r_0 to the wedge neurons.

Therefore, sleep resets synaptic plasticity and activity of ring neurons. Once the switching signal reaches a lower threshold, $r_I^{(min)}$, the model is switched back to the wake phase. In the third wake phase, we extend the phase by allowing the switching signal to increase beyond the upper threshold, simulating sleep deprivation. Here, w_{EE}

crosses the bifurcation boundary, $w_{EE} > 2$, and the model enters the domain of stable oscillations. During the following sleep phase, the system needs more time to fully reset and reach the lower threshold. Such sleep rebound after sleep deprivation is an experimentally described feature of sleep homeostasis circuits[17,26].

The time that the system spends in the sleep and wake phases is determined by the time constants of the plasticity rules, τ_{EE} and τ_{IE}, and the upper and lower thresholds of the switching signal, $r_I^{(min)}$ and $r_I^{(max)}$. In our simulations, we set τ_{EE} and τ_{IE} to yield dynamics in the timescale of seconds (for ease of visualization), but larger values lead to similar behavior on longer timescales (minutes or hours, see Figure S8).

2.5 Ring attractor with plasticity for ring and wedge neurons

We now extend the two-population model into a ring attractor network. The network is shown schematically in Figure S1C, implemented with $N = 32$ wedge neurons $r_E^{(i)}$ (based on anatomy[27]). For simplicity, we model the population of ring neurons with a single variable $r_I(t)$. The activities are described by the following system of differential equations:

$$\begin{cases} \tau \dot{r}_E^{(i)}(t) = -r_E^{(i)}(t) + \left[\sum_j^N w_{EE}^{(ij)}(t) r_E^{(j)}(t) - w_{EI} r_I(t) + \theta + I^{(i)}(m)\right]_+ & \text{for} \quad i = 1,\dots,N \\ \tau \dot{r}_I(t) = -r_I(t) + \left[\sum_j^N w_{IE}^{(j)}(t) r_E^{(j)}(t)\right]_+. \end{cases} \tag{4}$$

The plasticity rules are extended from the two-population model:

$$\begin{cases} \tau_{EE} \dot{w}_{EE}^{(ij)}(t) = r_E^{(i)}(t) r_E^{(j)}(t) & \text{for} \quad i,j = 1,\dots,N \\ \tau_{IE} \dot{w}_{IE}^{(i)}(t) = r_I(t) r_E^{(i)}(t) \left(r_E^{(i)}(t) - r_0\right) & \text{for} \quad i = 1,\dots,N. \end{cases} \tag{5}$$

The synaptic weight between a presynaptic wedge neuron j and a postsynaptic wedge neuron i is represented by the time dependent matrix element $w_{EE}^{(ij)}$(t). The matrix is initialized with a Gaussian function that depends on the distance between wedge neurons along the ring. The Gaussian has two parameters, the maximum amplitude, $w_{EE}^{(max)}$, and the standard deviation σ (equation (34) in Methods). Figure 3A illustrates the connectivity from all the wedge neurons to wedge neuron 16. Additionally, $I^{(i)}(m)$ is an input to each wedge neuron i, and represents visual or idiothetic (potentially time-varying) input into the ring attractor. This input is defined as a Gaussian function where the peak is located in a given wedge neuron m (equation (35) in Methods).

2.6 Ring attractor dynamics during wakefulness

The resulting network shows overall similar dynamics as described for the two-population model. As shown below, the stability of a bump of activity, for instance localized around wedge neuron 16, is determined by its total excitatory input, $w_{EE}^{(16,sum)}(t)$ (Figure 3A):

$$w_{EE}^{(16,sum)}(t) = \sum_j^N w_{EE}^{(16,j)}(t). \tag{6}$$

Figure 3B shows the different dynamic regimes of the bump as a function of the parameters $w_{EE}^{(max)}$ and σ that determine the values of $w_{EE}^{(ij)}$ (see Methods (4.7) and Figure S6). The colored lines in Figure 3B are isolines of constant $w_{EE}^{(16,sum)}$, and correspond to the boundaries of distinct dynamics.

The boundaries are similar to the ones found in the two-population model (with w_{EE} instead of $w_{EE}^{(16,sum)}$). As in the two-population model, the bump is stable if $1 < w_{EE}^{(16,sum)} < 2$. As the recurrent weights $w_{EE}^{(ij)}$ increase due to LTP during the wake phase, so do σ and $w_{EE}^{(max)}$; and the bump oscillates if $w_{EE}^{(16,sum)} > 2$. In contrast to the two-population model, the ring attractor does not become unstable when $w_{EE}^{(16,sum)} > 2(1+\sqrt{\theta/r_0})$ but keeps oscillating with very low activity in wedge neurons (see Methods 4.7).

Figure 3C illustrates the dynamics of the system: as $w_{EE}^{(16,sum)}$ increases, the weights $w_{IE}^{(i)}$ (represented by the mean) also increase, leading to increasing activity of ring neurons, in turn decreasing the bump width over time

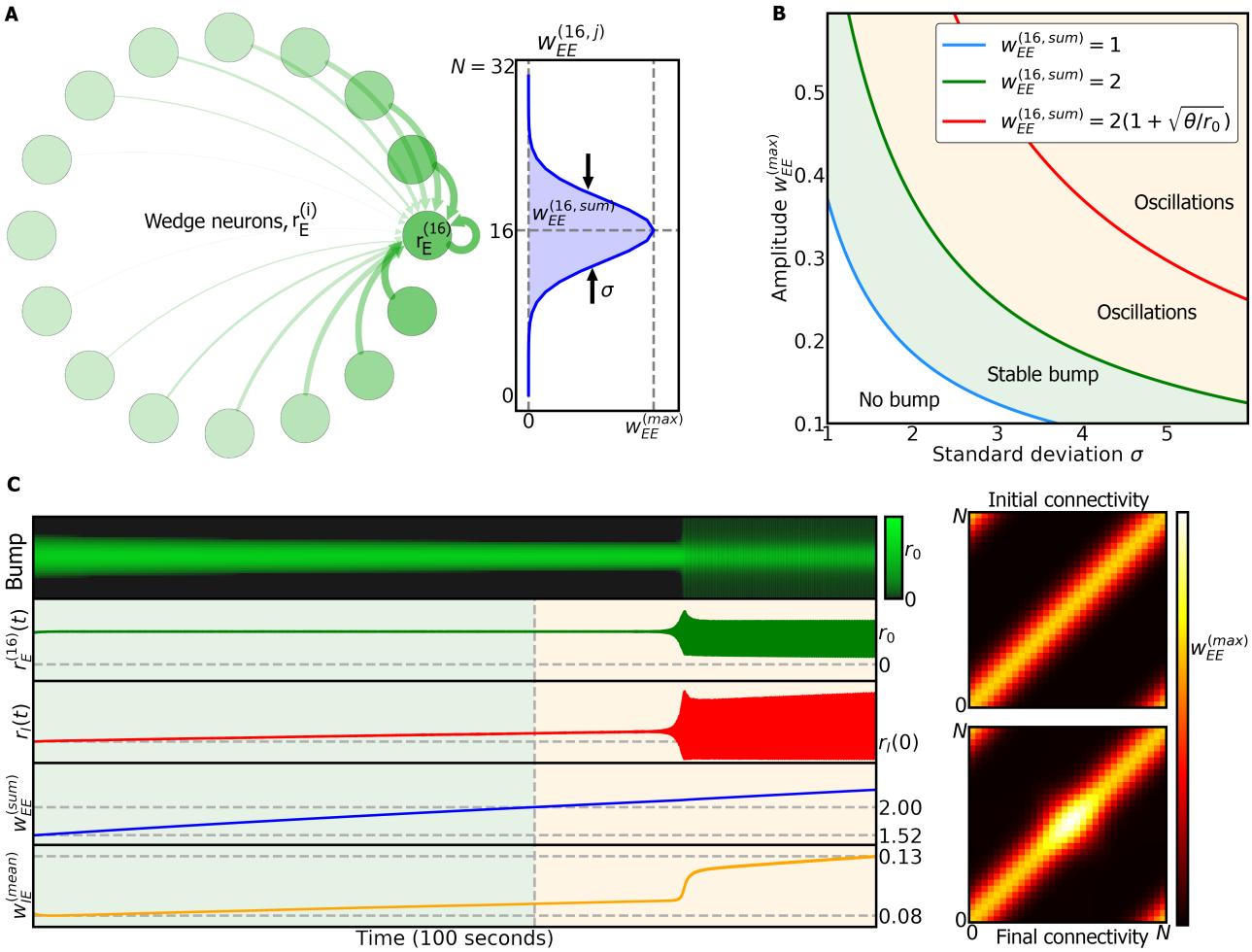

Figure 3. Dynamics of ring attractor network during wake phase. **A** Left side: representation of excitatory connections from all wedge neurons to wedge neuron $r_E^{(16)}$. Right side: Gaussian connectivity from all wedge neurons to wedge neuron 16, with maximum amplitude $w_{EE}^{(max)}$ and standard deviation σ. **B** Model dynamics obtained in the fast-timescale limit (see Methods(4.7)) depend on the parameters of the excitatory connectivity, $w_{EE}^{(max)}$ and σ. **C** Left side: dynamics during wake phase of the ring attractor model where the bump is located around wedge neuron 16. When the total excitatory weight, $w_{EE}^{(16,sum)}$, crosses a threshold, the bump starts to oscillate. Right side, top: initial excitatory connections between wedge neurons. Right side, bottom: final excitatory connections between wedge neurons, changes are a result of synaptic plasticity.

and maintaining activity of wedge neuron 16 constant. When $w_{EE}^{(16,sum)} > 2$, the bump starts to oscillate. In addition, the plasticity rule in the recurrent connections $w_{EE}^{(ij)}$ leads to an increase of the synaptic weights around the bump position (Figure 3C, right side, top *vs.* bottom).

2.7 Ring attractor dynamics during sleep

As for the two-population model, we introduce a sleep phase where LTP in recurrent weights $w_{EE}^{(ij)}$ is switched to LTD by changing the sign of the plasticity rule:

$$\tau_{EE}\dot{w}_{EE}^{(ij)}(t) = -r_E^{(i)}(t)r_E^{(j)}(t) \qquad \text{for} \qquad i,j = 1,\ldots,N. \tag{7}$$

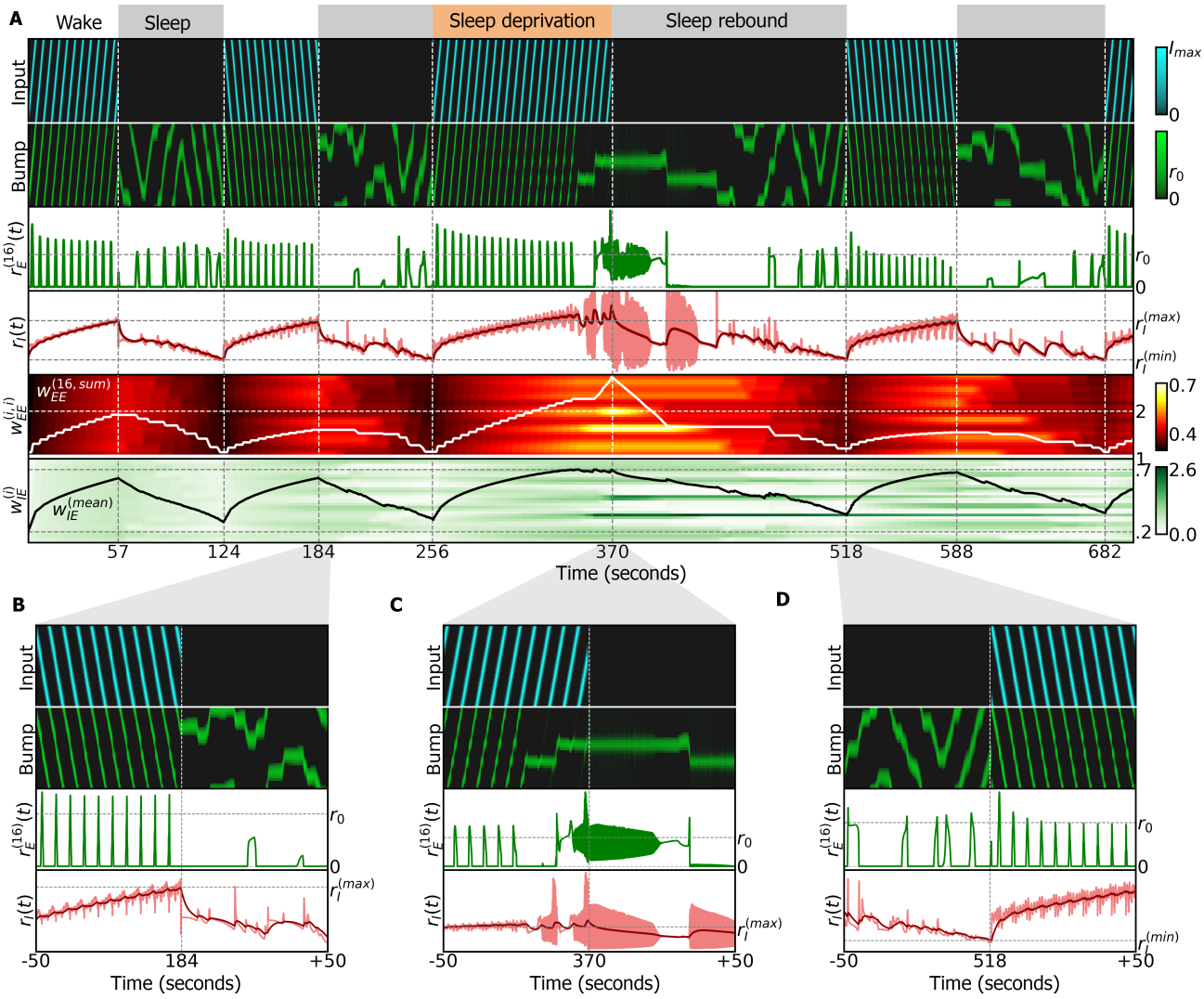

Figure 4. Wake and sleep phases in the ring attractor. **A** Entire simulation over a period of 450 seconds. Top row: input (only on during wake phase), alternating between clockwise and counter-clockwise rotations at $0.5Hz$. Second row: ring attractor bump activity. Third row: activity of wedge neuron 16. Fourth row, light red: activity of ring neurons. Dark red: filtered activity. Switching between sleep and wake depends on filtered activity crossing thresholds $r_I^{(min)}$ and $r_I^{(max)}$. In the third wake epoch, sleep deprivation is produced by extending the wake phase through a higher upper threshold $r_I^{(max)}$. Fifth row: diagonal elements of the connectivity matrix $w_{EE}^{(ij)}$. The white line is the sum of all excitatory connections to wedge neuron 16. It passes threshold 2 at around 240 seconds leading to oscillations. The full connectivity matrix $w_{EE}^{(ij)}$ at the switch times is shown in Figure S7. Sixth row: connectivity $w_{IE}^{(i)}$; white line is mean value. **B** Zoom-in around 113 seconds: switch from wake to sleep phase. **C** Zoom-in around 240 seconds: extended wake phase leads to oscillatory behavior. Circuit switches to sleep. **C** Zoom-in around 341 seconds: switch from sleep to wake phase.

Additionally, the ring attractor network does not receive input during sleep phase, so that $I^{(i)} = 0$. This condition is consistent with the fact that flies do not move during sleep and visual input to ring neurons is inhibited[26].

Figure 4A shows a simulation of the ring attractor during wake (top, white region) and sleep phases (grey region). In the wake phase, a rotating input with a constant frequency of $0.5Hz$ is provided; the input reverses direction between consecutive wake phases (top row). During the wake phase, the bump in the ring attractor closely follows

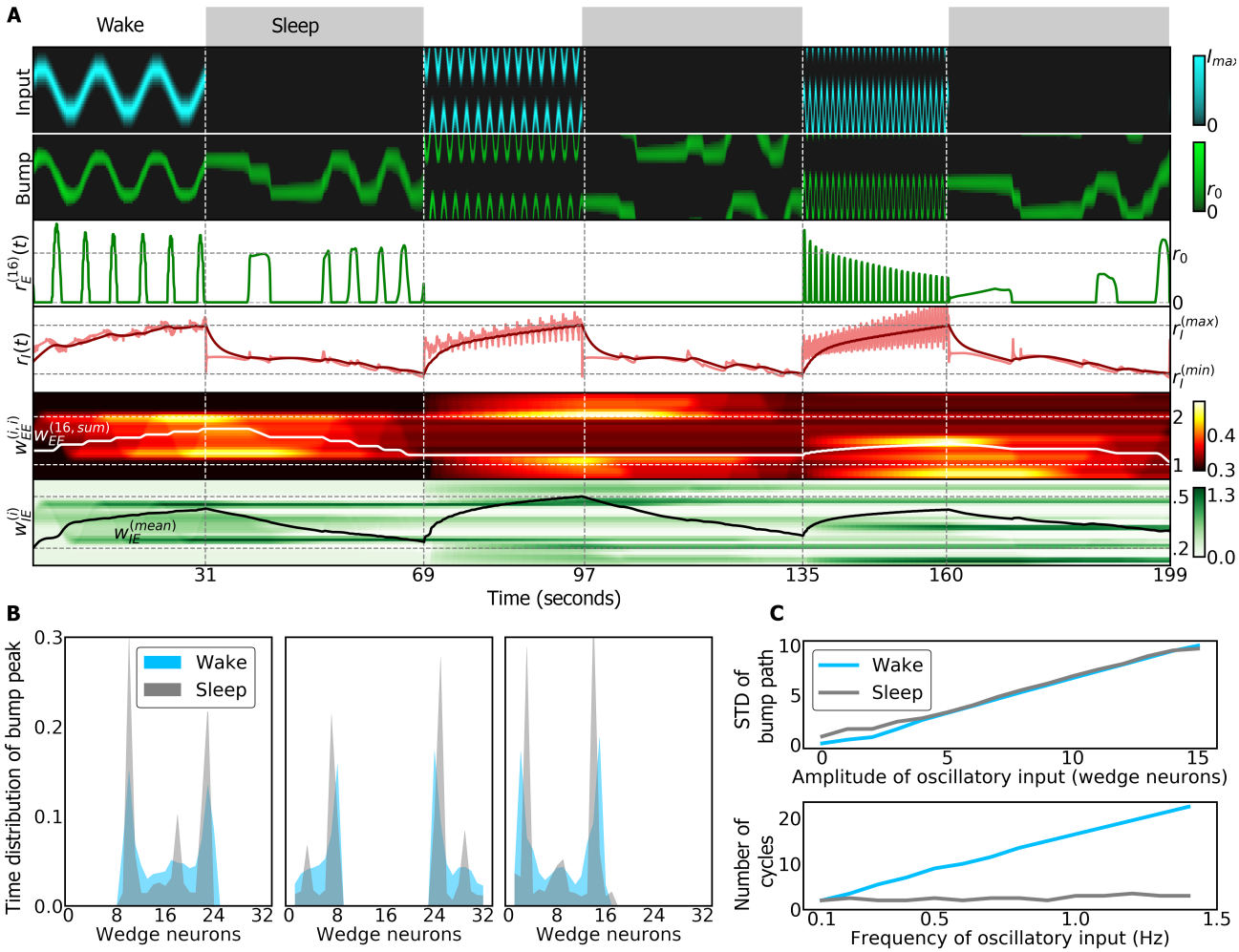

Figure 5. Simulation of the ring attractor network with sinusoidal inputs during wake phase. **A** In each wake phase, a sinusoidal input is provided to the ring attractor (top row) with increasing frequencies in consecutive wake phases. During sleep, an autonomous bump of activity revisits wedge neurons active in the previous wake phase (second row). The third row shows the activity of wedge neuron 16 and the forth row shows the activity of the ring neuron (light red) and its filtered activity (dark red) used to switch between wake and sleep phases. The last two rows represent the synaptic weights that increase during wkefulness and decrease during sleep. **B** Normalized distribution of time that the bump peak is localized in each wedge neuron during wake (blue) and sleep (grey) phases. The first, second and third plots show the distribution for the first, second and third wake and sleep phases, respectively. **C** Top: standard deviation (STD) of the bump path during a total of 15 simulations where different amplitudes of a sinusoidal inputs are provided during the wake phase. During the wake phase, the STD of the bump path is proportional to the amplitude (grey). During sleep, the autonomous bump path has a correlated STD (Pearson correlation coefficient: 0.99 ($p = 1.5 \cdot 10^{-17}$). Bottom: number of cycles of the bump path during a total of 15 simulations with different frequencies of a sinusoidal input during wake phase. During the wake phase, the number of cycles is proportional to the input frequency. However, during sleep, the number of cycles does not correlate with the input frequency: Pearson correlation coefficient: 0.76 ($p = 0.0015$).

the input (second row), while the activity of ring neurons increases (light red line in the forth row). The second to last row shows the diagonal elements of $w_{EE}^{(ij)}$ while the last row shows $w_{IE}^{(i)}$. We again use the filtered activity of ring neurons (switching signal, dark red line in the forth row) for switching between wake and sleep. When the switching

signal reaches the upper threshold, $r_I^{(max)}$, the system switches to sleep and visual input is turned off (Figure 4B). The autonomous bump movement resets the excitatory connections and the activity of ring neurons decreases.

As with the two-population model, the timescale over which the integration takes place depends on the time constants of the plasticity rules. Heterogeneity in the weights $w_{EE}^{(ij)}$ (due to synaptic plasticity) makes the bump drift across wedge neurons. During sleep, such autonomous dynamics reset the network connectivity in the network[7]. On the other hand, wake phases without visual input also show drift (see Figure S8). This wake drift is however different from sleep drift, since it ends once the bump reaches the location of strongest recurrent excitation, making the synaptic weights grow in this location until sleep is initiated. Such wake drift can be reduced by slowing the plasticity rules (see Figure S9 and Methods(4.9))

Once the switching signal reaches the lower threshold, the system switches back to wake phase and the input is turned on again (Figure 4B). Furthermore, when extending the wake period (sleep deprivation, upper orange region), $w_{EE}^{(16,sum)}$ for wedge neuron 16 crosses the boundary for stability, $w_{EE}^{(16,sum)} > 2$. In this case, the bump in the ring attractor starts to oscillate (Figure 4C; see section 2.6). Towards the end of the extended wake phase, the bump stops tracking the input. In the subsequent sleep phase, more time is required to reset the excitatory weights and to reach the lower threshold $r_I^{(max)}$, resulting in sleep rebound (Figure 4D).

To investigate how ring attractor dynamics during sleep are linked to its dynamics during the preceding wake phase, we provided a sinusoidal input with amplitude A and frequency f. In Figure 5A, we used a fixed amplitude, $A = N/4$ and varied the frequency in each wake phase $(0.1, 0.5,$ and 1 Hz). During sleep, the bump revisits wedge neurons that were active in the preceding wake phase, as seen in the distributions of the time spent around each wedge neuron during the first (left), second (center) and third (right) wake phase (in blue) and during the following sleep phase (in grey) (Figure 5B). To further probe the amplitude dependence, we simulated a wake phase and the subsequent sleep phase, and in each simulation varied the amplitude A of a sinusoidal input in the range $[0, 16]$, with fixed frequency of $f = 1Hz$, during the wake phase. The standard deviation (STD) of the paths of the bump during the wake (blue) and sleep (grey) phase closely match (Figure 5C). Similarly, we investigated the frequency dependence with a stimulus with fixed amplitude, $A = N/4$, and varying frequency in the range $[0.1, 1.5]Hz$. The number of oscillatory cycles grows linearly with the input frequency in the wake phase (blue), but remains constant during sleep. Therefore, the dynamics during sleep do not depend on the input frequency during the wake phase (Figure 5C). Additionally, the frequency of oscillations of the bump around the ring can increase as the bump approaches the lower threshold of the switching signal (before waking up, Figure S9A). Such autonomous dynamics are reminiscent of activity observed during sleep in mice[35].

3 Discussion

In the brain of *Drosophila*, structurally very similar neurons in the center of the brain have been assigned functionally very different roles. On the one hand, navigation related ring neurons encode spatial memory[19] or visual features[20]. For these ring neurons, ring attractor networks offer a compelling structure-function relationship which can provide a rationale for their ring-shaped structure. On the other hand, sleep related ring neurons serve as homeostatic sleep integrator, encoding sleep drive through structural[18,24] and activity changes[18,23]. To elucidate the relationship between these suggested navigation and sleep functionalities of ring neurons, we therefore asked what role the homeostatic integrator could play in the ring attractor framework.

To address this question, we developed two models, a two-population model (Figure 2), as well as a ring-attractor network with ring neurons and head direction-encoding wedge neurons (Figure 3). In these models, the function of ring neurons is to maintain head direction activity at a setpoint by balancing proposed wedge neuron plasticity. Activity evolves towards this setpoint over time, but allows deviations due to external (for example visual or self-motion related) input (Figure S4), consistent with experimentally observed behavior-related changes in bump amplitude[34]. In this way, ring neurons are assigned a functional role in the head direction system while at the same time showing features of experimentally observed plasticity and homeostatic sleep integration.

Interestingly, both the two-population model as well as the ring attractor model, show similar dynamics, with related boundaries between different regimes (Figures 2B and 3B). In both models, the introduced competing

plasticity rules aim to maintain head direction (bump) activity at a setpoint (Figure 2C and Figure 3C) at the cost of changing synaptic weights (Figures 2B, 2C and 3C). These weight changes result in ring neurons representing sleep drive with the characteristics of a homeostatic integrator (Figures 2D and 4). Prolonged activity during wakefulness ultimately leads to unstable behavior (Figures 2B, 3B and 3C). Therefore, to restore the head direction system to baseline, we introduced a sleep phase where the synaptic plasticity between wedge neurons is changed from LTP to LTD[4,7] (Figures 2D and 4). This results in autonomous dynamics of the ring attractor which, however, are linked to its dynamics during wakefulness through its spatial (Figures 5A and 5B) but not through its frequency distribution (Figures 5A and 5C).

The introduction of plasticity was motivated by the observation of structural, synaptic and functional changes in homeostatic ring neurons[18,23,24] as well as their likely connectivity to the head direction system[14,28]. In past work, plasticity has been introduced in ring attractor models to circumvent the problem of fine tuning[36–38], in contrast to the role of plasticity proposed here, which is to maintain bump activity at a setpoint. The proposed models can capture the increase in activity in ring neurons during wakefulness[18,24] (Figures 2 and 4), although the exact time course of this integrator signal has not been measured. Additionally, in the proposed models, sleep deprivation leads to a qualitative change in the behavior of ring neurons towards oscillatory dynamics which is reminiscent of the experimentally observed transition to bursting dynamics[18] or increase in oscillatory dynamics[23].

How plasticity in sleep-related ring neurons[18,24] can drive the observed increase in ring neuron activity with sleep pressure[18,23] is currently unknown. Since it is the activity of ring neurons, and not the plasticity, which is hypothesised to trigger sleep[18,23,26], our models assume plasticity that directly modifies the activity of ring neurons linearly (Fig2D and 4A) (a possibility that is consistent with the data in[18]). A different plasticity mechanism from a different class of visual ring neurons to wedge neurons that flexibly modifies the alignment of the head direction signal with respect to a visual scene was discussed in[30,31]. Whether such plasticity could be combined with the plasticity proposed here needs to be explored in future work.

Our models make several predictions that potentially could be tested experimentally. First, we assume LTP in the recurrent connections of wedge neurons during wakefulness (which could also be achieved through an intermediate population[29]). This plasticity in w_{EE} could be a result of correlated activity between neighboring wedge neurons. Second, as a result of this plasticity, the bump width changes with w_{EE}, decreasing over time spent awake due to LTP in wedge neurons and therefore increasing spatial resolution (and *vice versa* during sleep, Figure S6C). Generally, there is a range of bump widths that can be sustained by the ring attractor (Figure 3). A third model prediction is that extended wakefuness can disrupt the head direction system by producing oscillatory or bursting behavior and will lock the bump position in place (independent of visual input, Figure 4A, sleep deprivation). Additionally, sleep results in autonomous dynamics in the ring attractor model (Figure 4), with the network transitioning towards faster dynamics towards the end of the sleep phase (Figure S9). Such autonomous dynamics are reminiscent of activity of the head direction system observed in mice during sleep[35].

We further assume that the plasticity rule in the head direction system changes during sleep from LTP to LTD. This change of plasticity has been proposed in several models of sleep (for example[39]; see[4] for review). A potential mechanism could for example be neuromodulation of an STDP gate[40,41], which has been observed in insects[42,43] and could involve the strong innervation of the central complex by neuromodulatory neurons[44]. The resulting weakening of synaptic strength during sleep underlies several hypotheses about sleep function[4–6,8]. The approach implemented here is additionally based on the idea of reverse learning[5–7]: during sleep, attractors within the ring attractor network generated during wakefulness are removed and the corresponding increased weights are weakened.

While synaptic changes during sleep and wakefulness are observed across the fly brain (for example[45]), one could hypothesize that such activity related changes are stronger in areas where activity is persistent, such as the head direction system[21]. Therefore, ring neurons might require resetting through sleep sooner than other transiently active neurons and might therefore be responsible for signaling sleep drive. We additionally did not differentiate between different ring attractor inputs (for example visual or idiothetic) and such different signals could also be integrated in different ring neurons or homeostats[11] (taking for example into account that visual experience increases sleep need[46]).

The developed model focuses only on a part of the fly's sleep control circuit, the homeostat, and could be

integrated with other parts of this circuit, similar to models that describe the full sleep-wake cycle (see[47,48] for review). For example, the dorsal fan-shaped body neurons (dFB) in the central complex have been reported to promote sleep when active[17]. These dFB neurons receive excitatory input from ring neurons through indirect pathways[18]. A conceptual model[26] proposed a recurrent circuit for sleep homeostasis where ring neurons integrate visual input and therefore time spent awake, and dFB neurons act as a switch for sleep when the input from ring neurons crosses a threshold. During sleep, dFB neurons inhibit visual input, allowing ring neurons to reset their activity. While this conceptual model does not explain sleep homeostasis in darkness, it has been alternatively suggested that ring neurons could integrate locomotor activity[49]. The model developed here is compatible with both these suggestions and can integrate any input signal, visual or self-motion related.

Overall, the interaction of the homeostatic integrator and the head direction systems together with mounting evidence for a close structure-function relationship in these circuits, suggest that a relationship between the control and function of sleep could be established in this network using theoretical modeling and experiments.

References

1. Killgore, W. D. Effects of sleep deprivation on cognition. In *Progress in brain research*, vol. 185, 105–129 (Elsevier, 2010).
2. Krause, A. J. *et al.* The sleep-deprived human brain. *Nat. Rev. Neurosci.* **18**, 404 (2017).
3. Tobaldini, E. *et al.* Sleep, sleep deprivation, autonomic nervous system and cardiovascular diseases. *Neurosci. & Biobehav. Rev.* **74**, 321–329 (2017).
4. Rennó-Costa, C., da Silva, A. C. C., Blanco, W. & Ribeiro, S. Computational models of memory consolidation and long-term synaptic plasticity during sleep. *Neurobiol. learning memory* **160**, 32–47 (2019).
5. Crick, F. & Mitchison, v. The function of dream sleep. *Nature* **304**, 111–114 (1983).
6. Hopfield, J. J., Feinstein, D. & Palmer, R. 'unlearning'has a stabilizing effect in collective memories. *Nature* **304**, 158–159 (1983).
7. Thiele, J., Diehl, P. & Cook, M. A wake-sleep algorithm for recurrent, spiking neural networks. *arXiv preprint arXiv:1703.06290* (2017).
8. Tononi, G. & Cirelli, C. Sleep and the price of plasticity: from synaptic and cellular homeostasis to memory consolidation and integration. *Neuron* **81**, 12–34 (2014).
9. Cirelli, C. & Tononi, G. Linking the need to sleep with synaptic function. *Science* **366**, 189–190 (2019).
10. Liu, D. & Dan, Y. A motor theory of sleep-wake control: arousal-action circuit. *Annu. review neuroscience* (2019).
11. Kirszenblat, L. & van Swinderen, B. Sleep in drosophila. In *Handbook of Behavioral Neuroscience*, vol. 30, 333–347 (Elsevier, 2019).
12. Jenett, A. *et al.* A gal4-driver line resource for drosophila neurobiology. *Cell reports* **2**, 991–1001 (2012).
13. Simpson, J. H. & Looger, L. L. Functional imaging and optogenetics in drosophila. *Genetics* **208**, 1291–1309 (2018).
14. Xu, C. S. *et al.* A connectome of the adult drosophila central brain. *BioRxiv* (2020).
15. Beckwith, E. J. & French, A. S. Sleep in drosophila and its context. *Front. physiology* **10**, 1167 (2019).
16. Dubowy, C. & Sehgal, A. Circadian rhythms and sleep in drosophila melanogaster. *Genetics* **205**, 1373–1397 (2017).
17. Donlea, J. M., Alam, M. N. & Szymusiak, R. Neuronal substrates of sleep homeostasis; lessons from flies, rats and mice. *Curr. opinion neurobiology* **44**, 228–235 (2017).

18. Liu, S., Liu, Q., Tabuchi, M. & Wu, M. N. Sleep drive is encoded by neural plastic changes in a dedicated circuit. *Cell* **165**, 1347–1360 (2016).

19. Ofstad, T. A., Zuker, C. S. & Reiser, M. B. Visual place learning in drosophila melanogaster. *Nature* **474**, 204–207 (2011).

20. Seelig, J. D. & Jayaraman, V. Feature detection and orientation tuning in the drosophila central complex. *Nature* **503**, 262 (2013).

21. Seelig, J. D. & Jayaraman, V. Neural dynamics for landmark orientation and angular path integration. *Nature* **521**, 186 (2015).

22. Skaggs, W. E., Knierim, J. J., Kudrimoti, H. S. & McNaughton, B. L. A model of the neural basis of the rat's sense of direction. In *Advances in neural information processing systems*, 173–180 (1995).

23. Raccuglia, D. *et al.* Network-specific synchronization of electrical slow-wave oscillations regulates sleep drive in drosophila. *Curr. Biol.* **29**, 3611–3621 (2019).

24. Huang, S., Piao, C., Beuschel, C. B., Götz, T. & Sigrist, S. J. Presynaptic active zone plasticity encodes sleep need in drosophila. *Curr. Biol.* (2020).

25. Hulse, B. K. & Jayaraman, V. Mechanisms underlying the neural computation of head direction. *Annu. Rev. Neurosci.* **43** (2019).

26. Donlea, J. M. *et al.* Recurrent circuitry for balancing sleep need and sleep. *Neuron* **97**, 378–389 (2018).

27. Wolff, T., Iyer, N. A. & Rubin, G. M. 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).

28. Turner-Evans, D. B. *et al.* The neuroanatomical ultrastructure and function of a biological ring attractor. *bioRxiv* 847152 (2019).

29. Green, J. & Maimon, G. Building a heading signal from anatomically defined neuron types in the drosophila central complex. *Curr. opinion neurobiology* **52**, 156–164 (2018).

30. Kim, S. S., Hermundstad, A. M., Romani, S., Abbott, L. & Jayaraman, V. Generation of stable heading representations in diverse visual scenes. *Nature* **576**, 126–131 (2019).

31. Fisher, Y. E., Lu, J., D'Alessandro, I. & Wilson, R. I. Sensorimotor experience remaps visual input to a heading-direction network. *Nature* **576**, 121–125 (2019).

32. Dayan, P. & Abbott, L. F. Theoretical neuroscience: computational and mathematical modeling of neural systems. (2001).

33. Gerstner, W. & Kistler, W. M. Mathematical formulations of hebbian learning. *Biol. cybernetics* **87**, 404–415 (2002).

34. Dana, H. *et al.* High-performance calcium sensors for imaging activity in neuronal populations and microcompartments. *Nat. Methods* **16**, 649–657 (2019).

35. Peyrache, A., Lacroix, M. M., Petersen, P. C. & Buzsáki, G. Internally organized mechanisms of the head direction sense. *Nat. neuroscience* **18**, 569–575 (2015).

36. Renart, A., Song, P. & Wang, X.-J. Robust spatial working memory through homeostatic synaptic scaling in heterogeneous cortical networks. *Neuron* **38**, 473–485 (2003).

37. Itskov, V., Hansel, D. & Tsodyks, M. Short-term facilitation may stabilize parametric working memory trace. *Front. computational neuroscience* **5**, 40 (2011).

38. Seeholzer, A., Deger, M. & Gerstner, W. Stability of working memory in continuous attractor networks under the control of short-term plasticity. *PLoS computational biology* **15**, e1006928 (2019).

39. Nere, A. T., Hashmi, A., Cirelli, C. & Tononi, G. Sleep-dependent synaptic down-selection (i): modeling the benefits of sleep on memory consolidation and integration. *Front. neurology* **4**, 143 (2013).

40. Brzosko, Z., Mierau, S. B. & Paulsen, O. Neuromodulation of spike-timing-dependent plasticity: past, present, and future. *Neuron* **103**, 563–581 (2019).

41. Pawlak, V., Wickens, J. R., Kirkwood, A. & Kerr, J. N. Timing is not everything: neuromodulation opens the stdp gate. *Front. synaptic neuroscience* **2**, 146 (2010).

42. Cassenaer, S. & Laurent, G. Conditional modulation of spike-timing-dependent plasticity for olfactory learning. *Nature* **482**, 47–52 (2012).

43. Feldman, D. E. The spike-timing dependence of plasticity. *Neuron* **75**, 556–571 (2012).

44. Kahsai, L. & Winther, Å. M. Chemical neuroanatomy of the drosophila central complex: distribution of multiple neuropeptides in relation to neurotransmitters. *J. Comp. Neurol.* **519**, 290–315 (2011).

45. Bushey, D., Tononi, G. & Cirelli, C. Sleep and synaptic homeostasis: structural evidence in drosophila. *Science* **332**, 1576–1581 (2011).

46. Kirszenblat, L., Yaun, R. & van Swinderen, B. Visual experience drives sleep need in drosophila. *Sleep* (2019).

47. Postnova, S. Sleep modelling across physiological levels. *Clocks & Sleep* **1**, 166–184 (2019).

48. Abel, J. H., Lecamwasam, K., St Hilaire, M. A. & Klerman, E. B. Recent advances in modeling sleep: from the clinic to society and disease. *Curr. Opin. Physiol.* **15**, 37–46 (2020).

49. Liang, X. *et al.* Morning and evening circadian pacemakers independently drive premotor centers via a specific dopamine relay. *Neuron* **102**, 843–857 (2019).

Funding

Max Planck Society, Center of Advanced European Studies and Research. P.J.G was supported by the German Research Foundation (DFG) through SFB 1089 'Synaptic Microcircuits'.

Acknowledgements

We would like to thank Raoul-Martin Memmesheimer, Gerry Rubin, Julijana Gjorgjieva, Felipe Kalle Kossio, and Marina Elaine Wosniack for helpful discussions.

Disclosures

The authors declare that there are no conflicts of interest related to this article.

4 Methods

4.1 Numerical simulation of models

We numerically solved both the two-population model and the ring attractor model with forward Euler with a time step of $dt = 0.0001$ seconds. Our code is implemented in Python, and we will make it available upon publication.

4.2 Firing rate models

We model the dynamics of the populations of wedge and ring neurons with rate-based models described by the following differential equations:

$$\begin{cases} \tau \dot{r}_E(t) = -r_E(t) + \Big[w_{EE}(t) r_E(t) - w_{EI} r_I(t) + \theta \Big]_+ \\ \tau \dot{r}_I(t) = -r_I(t) + \Big[w_{IE}(t) r_E(t) \Big]_+, \end{cases} \tag{8}$$

where, τ is the neural time constant, and $[\cdot]_+$ is the neural input-output relationship.

Switching between sleep and wake phases is dependent on the firing rate of ring neurons. We filter ring neuron activity with a low-pass filter using the following differential equation:

$$\tau_f \dot{r}_I^{(f)}(t) = -r_I^{(f)}(t) + r_I(t), \tag{9}$$

where τ_f is the time constant of the filter. We switch the wake and sleep phases in the following way:

$$\begin{cases} \text{if} \quad r_I^{(f)}(t) > r_I^{(max)} & \longrightarrow \quad \text{switch to sleep phase (plasticity rules in (13))} \\ \text{if} \quad r_I^{(f)}(t) < r_I^{(min)} & \longrightarrow \quad \text{switch to wake phase (plasticity rules in (14))} \end{cases} \tag{10}$$

4.3 Plasticity rules

For rate-based models, a general Hebbian plasticity rule for a synaptic weight w_{ij} can be defined as follows:

$$\tau_{ij}\dot{w}_{ij} = F(w_{ij}, r_i, r_j), \tag{11}$$

where τ_{ij} is the time constant of the rule, and $F(\cdot)$ is a function that depends on the synaptic weight, w_{ij}, and on pre- and postsynaptic activities, r_j and r_i, respectively[33]. The function $F(\cdot)$ needs to fulfill Hebb's condition: to produce a change in the synaptic weight w_{ij}, the pre- and postsynaptic neurons must be active: $r_i > 0$, $r_j > 0$. In principle, this function is unknown, but we can expand it in a Taylor series[33] around $r_I = r_E = 0$:

$$\tau_{ij}\dot{w}_{ij} \approx c_{00} + c_{10}r_i + c_{11}r_j + c_{20}r_i^2 + c_{21}r_ir_j + c_{22}r_j^2 + c_{30}r_i^3 + c_{31}r_i^2r_j + c_{32}r_ir_j^2 + c_{33}r_j^3 + O(r^4), \tag{12}$$

where each coefficient depends on the connection strength $c_{mn} = c_{mn}(w_{ij})$. The values of these coefficients determine the plasticity rule. For instance, Hebbian plasticity rules that are linear in the neural activities can be obtained by setting second or higher order coefficients to zero[33]. Keeping higher order coefficients leads to rules with non-linearities.

In our model, the plasticity rule in $w_{EE}(t)$ during the wake phase is linear, obtained by setting $c_{21} = 1$ and all other coefficients to zero, while the plasticity rule in w_{IE} is non linear on the presynaptic neural activity, obtained by setting $c_{21} = -r_0$, $c_{32} = 1$ and the other coefficients to zero. This leads to the following Hebbian plasticity rules:

$$\begin{cases} \tau_{EE}\dot{w}_{EE}(t) = r_E(t)r_E(t) \\ \tau_{IE}\dot{w}_{IE}(t) = r_I(t)r_E(t)\Big(r_E(t) - r_0\Big). \end{cases} \tag{13}$$

During sleep, we can obtain the changed sign of the plasticity rule in w_{EE} by setting $c_{21} = -1$ and all other coefficients to zero:

$$\begin{cases} \tau_{EE}\dot{w}_{EE}(t) = -r_E(t)r_E(t) \\ \tau_{IE}\dot{w}_{IE}(t) = r_I(t)r_E(t)\Big(r_E(t) - r_0\Big). \end{cases} \tag{14}$$

To ensure that synapses remain excitatory or inhibitory throughout the system's dynamics, the plasticity rules are threshold-rectified at zero if the synaptic weights are zero:

$$\begin{cases} \dot{w}_{EE}(t) = \Big[\dot{w}_{EE}(t)\Big]_+ & \text{if} \quad w_{EE}(t) = 0 \\ \dot{w}_{IE}(t) = \Big[\dot{w}_{IE}(t)\Big]_+ & \text{if} \quad w_{IE}(t) = 0. \end{cases} \tag{15}$$

The parameters for the two-population model are shown in Table S1.

To analyze this model, we assume that the dynamics of the synaptic plasticity in equations (13) and (14) are much slower than the dynamics of the neural populations in equations (8):

$$\frac{\tau}{\tau_{EE}} \longrightarrow 0 \quad \text{and} \quad \frac{\tau}{\tau_{IE}} \longrightarrow 0. \tag{16}$$

Below, we discuss the stability and global behavior of the two population model in two different limits. First, the fast-timescale limit where $\tau_{EE}, \tau_{IE} \longrightarrow \infty$ and the evolution of the model therefore only depends on the dynamics of the neural populations. Secondly, the slow-timescale limit where $\tau \longrightarrow 0$ and therefore the dynamics of the plasticity rules are governing the evolution of the model.

τ	τ_f	τ_{EE}	τ_{IE}	w_{EI}	θ	r_0	$r_I^{(min)}$	$r_I^{(max)}$
0.01	2	10000	10000	0.5	10	10	20	35

Table S1. Parameter values used in the two-population model. Time constants τ, τ_f, τ_{EE} and τ_{IE} are measured in seconds. We consider the rest dimensionless quantities.

4.4 Two population model: fast-timescale limit

Assuming $\tau_{EE}, \tau_{IE} \longrightarrow \infty$ means that synaptic plasticity is sufficiently slow compared to the dynamics of the neural populations so that it can be assumed to be constant. This offers the advantage of treating the synaptic weight w_{EE} as a free parameter with a fixed value, assuming that w_{IE} has already evolved through its plasticity rule to its equilibrium value, $r_E \longrightarrow r_0$. Therefore, for a given value of w_{EE}, we set the value of w_{EI} such that the fixed point for the wedge neurons is r_0. In this way, the value of w_{IE} is coupled to the value of w_{EE}. The stability of the 2-dimensional system given by equations in (8) is then analyzed with respect to the value of w_{EE}. Since the system is piecewise linear due to the threshold function $[\cdot]_+$, we perform a linear analysis assuming that the inputs to the neurons are positive. Under these conditions, the fixed point of the system, (r_E^*, r_I^*), is given by the following expressions:

$$\begin{cases} r_E^* = w_{EE} r_E^* - w_{EI} r_I^* + \theta = r_0 \\ r_I^* = w_{IE} r_0. \end{cases} \tag{17}$$

Since we force the fixed point of wedge neurons to be r_0, we can extract the equilibrium value of w_{IE} as a function of w_{EE}:

$$w_{IE} = \frac{\theta}{w_{EI} r_0} + \frac{w_{EE} - 1}{w_{EI}}. \tag{18}$$

The fixed point of the system can be described with respect to w_{EE} as:

$$\begin{cases} r_E^* = r_0 \\ r_I^* = \frac{\theta}{w_{EI}} + r_0 \frac{w_{EE}-1}{w_{EI}} \end{cases} \tag{19}$$

Both the fixed point of ring neuron activity, r_I^*, and the equilibrium value of the connectivity, w_{IE}, depend linearly on w_{EE}, implying that if w_{EE} increases, both w_{IE} and r_I^* increase as well as long as the fixed point is stable. We analyze the stability of the system by calculating the eigenvalues:

$$\lambda_\pm = \frac{1}{2\tau}\left((w_{EE} - 2) \pm \sqrt{(w_{EE} - 2)^2 - \frac{4\theta}{r_0}} \right). \tag{20}$$

Both eigenvalues are shown in Fig. S3A with respect to different values of w_{EE}. This leads to four different cases:

$$\begin{cases} (1) \quad w_{EE} \in \left[1, \quad 2 - 2\sqrt{\theta/r_0}\right] \\ (2) \quad w_{EE} \in \left(2 - 2\sqrt{\theta/r_0}, \quad 2\right] \\ (3) \quad w_{EE} \in \left(2, \quad 2 + 2\sqrt{\theta/r_0}\right] \\ (4) \quad w_{EE} > 2 + 2\sqrt{\theta/r_0} \end{cases} \tag{21}$$

In case (1), both eigenvalues are real and negative: the fixed point is stable (equation (19)). In case (2), the real part of the eigenvalues is negative and the imaginary part is non-zero: the system evolves towards the fixed point

with damped oscillations. In case (3), the real part of the eigenvalues is positive and the imaginary part is non-zero: the system diverges towards infinity, oscillating with amplitudes which increase exponentially. In case (4), the eigenvalues are real and positive: the fixed point is unstable. This analysis predicts a bifurcation in the stability of the fixed point when $w_{EE} = 2$. This behavior is shown in figure 2C.

The non-linearity of the linear threshold function changes the behavior of the model slightly. While the behavior stays the same for the cases (1), (2) and (4), because the model is mostly in the linear regime, case (3) differs and the non-linearity produces stable cycles around the fixed point. This behavior is found empirically from simulating the non-linear model, and is summarized in Figure 2B.

4.5 Two population model: slow-timescale limit

In this section, we analyze the conditions under which the synaptic rules in equations (13) and (14) stabilize the model. We therefore focus on the slow-timescale limit where the firing rates change sufficiently fast compared with the synaptic weights that these changes can be considered instantaneous ($\tau \longrightarrow 0$). We again first consider the linear range of the function $[\cdot]_+$ where the inputs to the neurons are positive. We approximate the instantaneous value of the firing rates in equation (8) as follows:

$$\begin{cases} r_E = w_{EE}(t) r_E - w_{EI} r_I + \theta \\ r_I = w_{IE}(t) r_E \end{cases} \tag{22}$$

This linear system allows extracting the values of r_E and r_I in terms of the synaptic weights as:

$$\begin{cases} r_E = \frac{\theta}{1 - w_{EE} + w_{EI} w_{IE}} \\ r_I = \frac{w_{IE}\theta}{1 - w_{EE} + w_{EI} w_{IE}} \end{cases} \tag{23}$$

We can now compute the vector field for wedge and ring neuron activity as a consequence of the slow dynamics of synaptic plasticity:

$$\begin{cases} \dot{r_E} = \frac{\partial r_E}{\partial w_{EE}} \dot{w}_{EE} + \frac{\partial r_E}{\partial w_{IE}} \dot{w}_{IE} \\ \dot{r_I} = \frac{\partial r_I}{\partial w_{EE}} \dot{w}_{EE} + \frac{\partial r_I}{\partial w_{IE}} \dot{w}_{IE} \end{cases} \tag{24}$$

4.5.1 Slow-timescale limit in the wake phase

Considering the plasticity rules during the wake phase (13), equation (24) leads to the following system of differential equations:

$$\begin{cases} \dot{r_E} = \frac{w_{EI} r_I \tau_{EE} - \tau_{IE}}{\tau_{EE} \tau_{IE}} \frac{r_E^3}{\theta} \left(- r_E + \left(1 + \frac{\tau_{IE}}{w_{EI} r_I \tau_{EE} - \tau_{IE}} \right) r_0 \right) \\ \dot{r_I} = \frac{r_I^3 r_E}{\theta} \left(\frac{1}{\tau_{EE}} + \frac{1 - w_{EE}}{\tau_{IE}} (r_E - r_0) \right) \end{cases} \tag{25}$$

The second equation gives the dynamics of ring neurons, which increase activity with $1/\tau_{EE}$. The first equation gives the dynamics of the population of wedge neurons approaching a setpoint only when the effective decay time constant (the first factor in the right hand side) is positive, otherwise the equation diverges to infinity and the system is unstable. This gives the following criterion for τ_{EE} and τ_{IE}:

$$\frac{\tau_{EE}}{\tau_{IE}} > \frac{1}{w_{EI} r_I} \tag{26}$$

This relationship supports the idea that plasticity in w_{IE} has to be fast enough with respect to the LTP in w_{EE}; otherwise, if w_{EE} increases faster than w_{IE}, the model diverges. Let us compute the upper limit of inequality (26), which corresponds to the minimum of r_I. For that, we approximate the firing rate of wedge neurons by $r_E = r_0$, and $r_I = w_{IE} r_0$. Then, we can write w_{IE} as a function of w_{EE} as in equation (18). The upper limit of inequality (26) will therefore happen at the minimum of w_{EE}. As the minimum is $w_{EE} = 1$, the upper limit of the stability condition is:

$$\frac{\tau_{EE}}{\tau_{IE}} > \frac{1}{\theta}. \tag{27}$$

When this condition holds, the setpoint of wedge neurons in equation (25), which is not r_0 as approximated previously, is given by

$$r_E^* = \left(1 + \frac{1}{w_{EI} r_I \frac{\tau_{EE}}{\tau_{IE}} - 1}\right) r_0, \tag{28}$$

which is generally different from r_0 due to the inertia of the dynamics of w_{EE}. In the limit $\tau_{EE} \longrightarrow \infty$ (no plasticity in w_{EE}), the fixed point is r_0, as expected from the fast-timescale analysis.

Figure S3B shows the vector field for the system of equations (25). The green line shows the trajectory of the setpoint in wedge neurons as the activity of ring neurons increase. As r_I increases due to increasing w_{EE}, the setpoint in r_E approaches r_0.

4.5.2 Slow-timescale limit in the sleep phase

We can perform the same analysis during sleep by considering the plasticity rules in equation (14), so that the equations (24) lead to the following system of differential equations:

$$\begin{cases} \dot{r_E} = \frac{w_{EI} r_I \tau_{EE} + \tau_{IE}}{\tau_{EE} \tau_{IE}} \frac{r_E^3}{\theta} \left(-r_E + \left(1 - \frac{\tau_{IE}}{w_{EI} r_I \tau_{EE} + \tau_{IE}}\right) r_0 \right) \\ \dot{r_I} = \frac{r_I^3 r_E}{\theta} \left(-\frac{1}{\tau_{EE}} + \frac{1 - w_{EE}}{\tau_{IE}} (r_E - r_0) \right). \end{cases} \tag{29}$$

The second equation shows how ring neurons decrease their activity with $1/\tau_{EE}$, at the same rate as in the wake phase. The first equation shows a fixed point for wedge neurons that is lower than r_0, due to the inertia of a decreasing w_{EE} during sleep, given by:

$$r_E^* = \left(1 - \frac{1}{w_{EI} r_I \frac{\tau_{EE}}{\tau_{IE}} + 1}\right) r_0. \tag{30}$$

Figure S3C shows the vector field given by equation (29) during the sleep phase, where the trajectory of the setpoint of wedge neurons (the green line) diverges from r_0 as the activity of ring neurons decreases.

4.6 Ring attractor network with plasticity

We expand the two-population model to a ring attractor network. A total of $N = 32$ individual wedge neurons are modeled by $r_E^{(i)}(t)$. For simplicity, ring neurons are modeled as a population, $r_I(t)$. The model is schematically shown in Figure S1C. The dynamics of the ring attractor network are given by the following equations:

$$\begin{cases} \tau \dot{r}_E^{(i)}(t) = -r_E^{(i)}(t) + \left[\sum_j^N w_{EE}^{(ij)}(t) r_E^{(j)}(t) - w_{EI} r_I(t) + \theta + I^{(i)}(m) \right]_+ \quad \text{for} \quad i = 1, \ldots, N \\ \tau \dot{r}_I(t) = -r_I(t) + \left[\sum_j^N w_{IE}^{(j)}(t) r_E^{(j)}(t) \right]_+. \end{cases} \tag{31}$$

The synaptic plasticity rules are also extended from the two-population model during the wake phase:

$$\begin{cases} \tau_{EE} \dot{w}_{EE}^{(ij)}(t) = r_E^{(i)}(t) r_E^{(j)}(t) \quad \text{for} \quad i, j = 1, \ldots, N \\ \tau_{IE} \dot{w}_{IE}^{(i)}(t) = r_I(t) r_E^{(i)}(t) \left(r_E^{(i)}(t) - r_0 \right) \quad \text{for} \quad i = 1, \ldots, N, \end{cases} \tag{32}$$

and during the sleep phase:

$$\begin{cases} \tau_{EE} \dot{w}_{EE}^{(ij)}(t) = -r_E^{(i)}(t) r_E^{(j)}(t) \quad \text{for} \quad i, j = 1, \ldots, N \\ \tau_{IE} \dot{w}_{IE}^{(i)}(t) = r_I(t) r_E^{(i)}(t) \left(r_E^{(i)}(t) - r_0 \right) \quad \text{for} \quad i = 1, \ldots, N. \end{cases} \tag{33}$$

We initialize the synaptic weights $w_{EE}^{(ij)}$ with a Gaussian function with amplitude $w_{EE}^{(max)}$ and standard deviation σ:

$$w_{EE}^{(ij)} = w_{EE}^{(max)} \exp\left(-\frac{\left(\min\left[|i-j|, \quad N - |i-j| \right] \right)^2}{2\sigma^2} \right). \tag{34}$$

τ	τ_f	τ_{EE}	τ_{IE}	w_{EI}	θ	r_0	$r_I^{(min)}$	$r_I^{(max)}$	$w_{EE}^{(max)}$	σ	I_{max}	I_σ
0.01	2	10000	10000	0.5	10	10	22	30	0.3	1.5	5	2

Table S2. Parameter values used in the ring attractor network. As in the previous model, time constants τ, τ_f, τ_{EE} and τ_{IE} are measured in seconds. We consider the rest dimensionless quantities.

Additionally we provide a Gaussian input to the ring attractor around a given wedge neuron m with amplitude I_{max} and standard deviation I_σ:

$$I^{(i)}(m) = I_{max}\exp\left(-\frac{\left(\min\left[|i-m|, \quad N-|i-m|\right]\right)^2}{2I_\sigma^2}\right). \tag{35}$$

This input allows changing the position of the bump in the simulations, and can represent visual or idiothetic input to update the head direction of the animal.

We again use the low-pass filtered activity of ring neurons to switch between sleep and wake phases, as in equation (9). The values of the parameters in Table S2 are used in all simulations unless stated otherwise.

4.7 Ring attractor network: bump stability analysis

To analyze to stability of the ring attractor model, we use an approach similar to the one in the fast-timescales analysis of the two-population model. First, we assume no plasticity in the recurrent connections $w_{EE}^{(ij)}$ but only in $w_{IE}^{(ij)}$,

$$\begin{cases} \tau_{EE}\dot{w}_{EE}^{(ij)}(t) = 0 \\ \tau_{IE}\dot{w}_{IE}^{(i)}(t) = r_I(t)r_E^{(i)}(t)\left(r_E^{(i)}(t) - r_0\right) \qquad \text{for} \qquad i = 1,\ldots,N. \end{cases} \tag{36}$$

We analyze the stability and behavior of the network while gradually changing $w_{EE}^{(ij)}$. We nitialize the ring attractor network with a bump profile, where neuron number 16 has maximum activity $r_E^{(16)} > r_E^{(j)}, \forall j \neq 16$. Given that only the connections $w_{IE}^{(i)}$ are plastic, the activity of ring neurons converges to a stable value given by

$$r_I \longrightarrow w_{IE}^{(16)} r_0. \tag{37}$$

This can be understood by looking at the plasticity rule for $w_{IE}^{(i)}$ (36). First, all the synaptic weights evolve so that wedge neuron 16 approaches the activity r_0. As all wedge neurons receive the same global inhibition, and wedge neuron 16 has maximum activity, the activity of the other wedge neurons is lower than r_0. At this point, the weights $w_{IE}^{(k)}$ for wedge neurons $k \neq 16$ with non-zero activity, decrease over time until reaching zero. On the other hand, if a wedge neuron $q \neq 16$ has zero activity, it does not provide any input to the ring neurons. As for wedge neuron 16, the synaptic plasticity rule changes the value of $w_{IE}^{(16)}$ such that its activity approaches r_0. An example of this behavior can be seen in Figure S5A, where we initialize the system with an input such that the peak of the bump is at wedge neuron 16. The synaptic weights $w_{IE}^{(k)}$ for wedge neurons $k \neq 16$ which have non-zero activity evolve towards zero.

The value of $w_{IE}^{(16)}$ is determined by the bump profile, because $r_E^{(16)}$ is receiving input from any wedge neuron with non-zero activity. Thus $w_{IE}^{(16)}$ has to balance the total excitation to set the activity of the wedge neuron 16 to r_0. Furthermore, the bump profile is determined by the parameters of the recurrent connectivity profile $w_{EE}^{(ij)}$, i.e. $w_{EE}^{(max)}$ and σ (equation (34)). If we fix the amplitude $w_{EE}^{(max)}$ and increase the standard deviation σ, the width of the bump, that is the number of active wedge neurons, decreases. This is because more wedge neurons are providing input to wedge neuron 16 as σ increases, and the value of $w_{IE}^{(16)}$ increases to set $r_E^{(16)} = r_0$, which in turn provides more

inhibition through ring neurons to all the wedge neurons and lowers their activities, reducing at the same time the bump width. This behavior is seen in Figure S5A and S5B.

The bump also shows oscillatory behavior (Fig. S5C) depending on the value of σ. In general, the state of the ring attractor network and its stability can be described in terms of the recurrent connectivity distribution, $w_{EE}^{(ij)}$, consistent with the fast-timescale limit analysis in the two-population model.

To investigate how the behavior and stability of the ring attractor network depend on the recurrent connections $w_{EE}^{(ij)}$, we simulated the ring attractor model for a grid of values for the parameters $w_{EE}^{(max)}$ and σ:

$$\begin{cases} w_{EE}^{(max)} \in [0.1, 0.6) & \text{with step} \quad 0.005 \\ \sigma \in [1, 6) & \text{with step} \quad 0.05. \end{cases} \tag{38}$$

In total, we performed 10000 simulations where we initialize the bump peak in wedge neuron 16 with a predefined input of 0.5 seconds, and let the network evolve for 10 seconds. We then analyzed the stability and behavior of the system in the last second of each simulation, therefore assuming that the state of the network does not change. Figure S5 shows an example of 3 simulations with different σ values; the light orange band across all simulations highlights the region used for analysis.

In this region we computed for each simulation the following:

- **Oscillation frequency**: the frequency at which wedge neuron 16 oscillates. For this, we computed the Discrete Fourier Transform, and the resulting peak value corresponds to the oscillation frequency.
- **Mean bump FWHM**: the mean value over time of the full width at half maximum of the bump, a proxy for the width of the bump.
- **Maximum bump peak**: the maximum value of the wedge neuron 16 over time.
- **Mean bump peak**: the mean value over time of the wedge neuron 16. If the bump does not oscillate, this value is equivalent to the maximum bump peak value.
- **Mean ring neuron activity**: the mean of ring neuron activity over time.

These measures are displayed in Figure S6B,S6C,S6D, S6E and S6F, as a function of the recurrent connectivity parameters $w_{EE}^{(max)}$ and σ. Figure *S6*B shows how the network starts oscillating with increasing $w_{EE}^{(max)}$ and σ. On the other hand, the FWHM in Figure *S6*B and S6C shows how at low $w_{EE}^{(max)}$ and σ values, the bump disappears and all wedge neurons have constant activity at r_0. As the parameter values increase, a bump of activity appears and the FWHM decreases, as observed in Figure3C and S5A and B. Finally, the activity of ring neurons increases as $w_{EE}^{(max)}$ and σ increase.

We further characterise the behavior and stability of the bump in the network with the total excitatory connectivity to the wedge neuron with maximal activity, i.e. wedge neuron 16, $w_{EE}^{(16,sum)}$:

$$w_{EE}^{(16,sum)} = \sum_{j}^{N} w_{EE}^{(16,j)}. \tag{39}$$

Figure S6A highlights the isolines where $w_{EE}^{(16,sum)}$ is constant for different values of $w_{EE}^{(max)}$ and σ. Note how the constant values of $w_{EE}^{(16,sum)} = 2$ and $w_{EE}^{(16,sum)} = 2(1+\sqrt{\theta/r_0})$ coincide with the boundaries of the different dynamic regimes. Therefore, the quantity $w_{EE}^{(16,sum)}$ has in the ring attractor model a similar role as does the recurrent connection w_{EE} in the two-population model. For $w_{EE}^{(16,sum)} < 1$, the bump disappears, similar to the two-population network model when $w_{EE} < 1$. However, unlike in the two-population model, which is unstable for $w_{EE} > 2(1+\sqrt{\theta/r_0})$, in the ring attractor network, for $w_{EE}^{(16,sum)} > 2(1+\sqrt{\theta/r_0})$ the wedge neurons are strongly inhibited by high activity in ring neurons.

From the above analysis we extracted regions of stability that are shown in Figure 3B. Constant lines of $w_{EE}^{(16,sum)}$ are computed as follows: in the continuous limit, i.e. $N \longrightarrow \infty$, the total excitatory connectivity is given by the following integral:

$$w_{EE}^{(16,sum)} = \int w_{EE}^{(max)} e^{\frac{x^2}{2\sigma^2}} dx = w_{EE}^{(max)} \sigma\sqrt{2\pi}. \tag{40}$$

Therefore, the isolines of constant $w_{EE}^{(16,sum)}$ are given by $w_{EE}^{(max)} = w_{EE}^{(16,sum)}/\sigma\sqrt{2\pi}$. For the discrete case, we empirically found $w_{EE}^{(16,sum)}$ to be well approximated by:

$$w_{EE}^{(max)} = \frac{w_{EE}^{(16,sum)}}{C\sigma}, \tag{41}$$

for any constant line $w_{EE}^{(16,sum)}$, where $C \approx 2.697$.

4.8 Ring attractor network: autonomous bump path analysis

To simulate how the bump in the ring attractor changes position to update the head direction during the wake phase, we use a simple clockwise or counter-clockwise rotating input with frequency f defined by:

$$m(t) = \pm fNt. \tag{42}$$

$m(t)$ is the wedge neuron where the ring attractor receives the Gaussian input $I^{(i)}(m(t))$ (equation (35)), and it is a cyclic variable, so that:

$$\begin{cases} \text{if} \quad m(t) = N+1 \quad \longrightarrow \quad m(t) = 1 \\ \text{if} \quad m(t) = 0 \quad \longrightarrow \quad m(t) = N. \end{cases} \tag{43}$$

During the sleep phase, the bump in the ring attractor shows autonomous dynamics (Figure 4A). To investigate the relationship between the path of the bump during sleep and in the preceding wake phase, we use a sinusoidal input during the wake phase, defined by the amplitude A, frequency f, and the center C:

$$m(t) = C + Asin(2\pi ft). \tag{44}$$

An example of this input with different frequencies f and centers C is shown in Figure 5A. We can obtain the position of the bump during wake and sleep phases as:

$$P_{bump}(t) = \text{argmax}_i \left[r_E^{(i)}(t) \right], \tag{45}$$

so that the position of the bump corresponds to the wedge neuron with maximum activity. We can now compute the distribution of times that the bump is localized around each wedge neuron i during sleep and wake phases, respectively, as:

$$\begin{cases} t_{wake}^{(i)} = \int_{wake} \delta\Big(P_{bump}(t) - i\Big) dt \\ t_{sleep}^{(i)} = \int_{sleep} \delta\Big(P_{bump}(t) - i\Big) dt, \end{cases} \tag{46}$$

where $\delta(\cdot)$ is the Dirac delta function and integrals extend over the wake and sleep phases. Figure 5B shows these distributions normalized for the three wake phases in Figure 5A and their following sleep phases. The distributions are very similar, meaning that during the sleep phase, the bump revisits the same wedge neurons that were active during the wake phase.

We further asked how the autonomous bump path changes during sleep with respect to the amplitude of the sine-shaped input, A, and frequency, f, during sleep. We first fixed the frequency of the input at $f = 1Hz$ and the

center at $C = 16$ while varying the amplitude in the range of $[0, 15]$ with an increment of 1. This resulted in 15 simulations where we computed the standard deviation of the bump path during sleep and wake:

$$\begin{cases} \mathrm{STD}_{wake} = \sqrt{\frac{1}{T_{wake}} \int_{wake} \left(P_{bump}(t) - 16\right)^2 dt} \\ \mathrm{STD}_{sleep} = \sqrt{\frac{1}{T_{sleep}} \int_{sleep} \left(P_{bump}(t) - 16\right)^2 dt}, \end{cases} \tag{47}$$

where T_{wake} and T_{sleep} are the duration of wake and sleep phases, respectively. Figure 5C shows the standard deviation in both phases with respect to the amplitude A. Note the similarity between both phases.

Secondly, we fixed the value of the amplitude at $A = 8$ and the center at $C = 16$ while varying the frequency, f, in the range $[0.1, 1.5]$ at increments of $0.1Hz$, resulting in 15 simulations. We quantify the number of cycles during both the sleep and wake phase. During the wake phase, the number of cycles is proportional to the input frequency f. Therefore a linear relationship between the number of cycles during sleep and the frequency of the input would give correlation between the frequency and the autonomously rotating bump path. However, Figure 5C, bottom, shows how the number of cycles during sleep does not change as the input frequency increases. An example of this can be seen in Figure 5A where we increase the frequency in consecutive wake phases and the path of the bump during sleep does not increase the rotation frequency.

4.9 Ring attractor network: bump drift during wake phase

In the simulations and analyses above, we provided input during the wake phase and the ring attractor network closely followed the input with a bump of activity. However, a ring attractor network should be able to sustain the bump of activity in the absence of input. It is known that small changes in the synaptic connections of wedge neurons $w_{EE}^{(ij)}$ can cause drifts of the bump in the absence of input[36–38].

To test for drift during the wake phase, we used a flashing rotating input that turns on and off. The input around a wedge neuron m is on for 0.2 sec (equation (35)), and then is turned off for 0.3 sec. Therefore, for N neurons the rotating input frequency is $1/(0.5N)$. Figure S8A shows such a simulation with three wake and sleep phases and S8B, C and D show zoom-ins around different times. Note how the bump drifts from the provided visual input, due to the synaptic changes in $w_{EE}^{(ij)}$.

This drift depends on the plasticity time constants. For instance, Figure S9 shows a simulation with 100 times larger time constants, τ_{EE}, $\tau_{IE} = 10^6$. The duration of wake and sleep phases are now in the order of hours, compared to the simulation in Figure S9. Note in the zoom-ins of Figures S9B, C and D that the bump of activity is sustained during the off time of the visual input without drifting. Since there are many more rotations during the wake phase and the synaptic changes in $w_{EE}^{(ij)}$ are very small in each rotation, the weights increase all together very homogeneously.

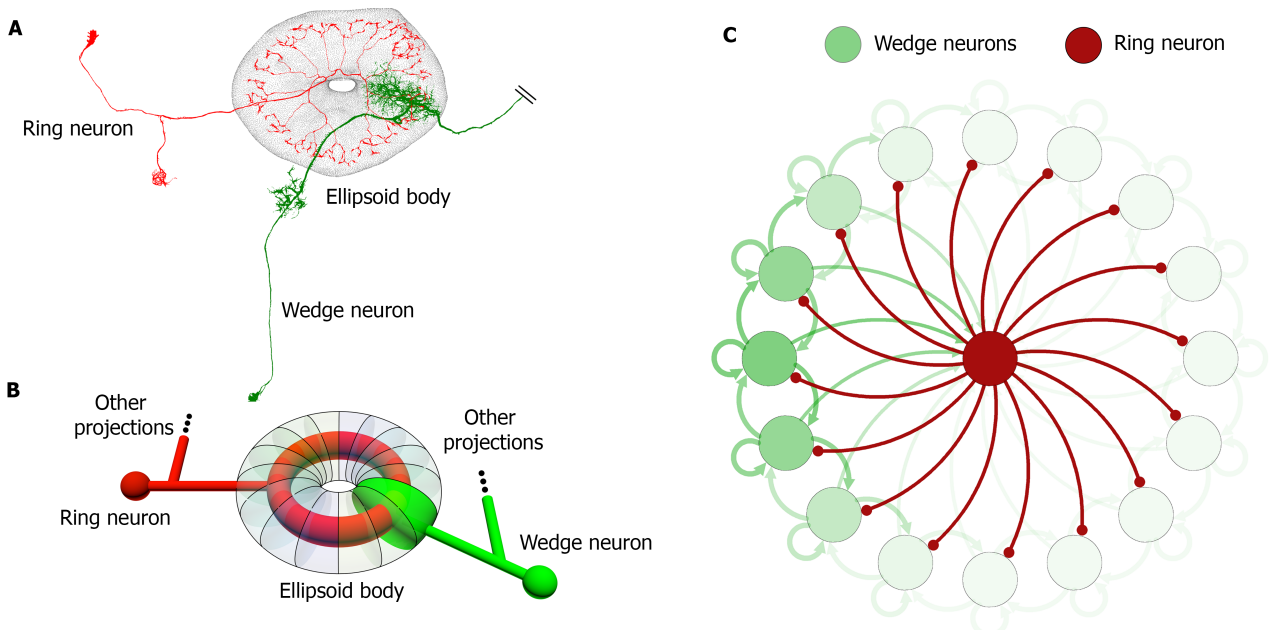

Figure S1. Wedge and ring neuron connectivity. **A** Projections of a ring neuron (R2) (red) and wedge (or EPG) neurons (green) in the ellipsoid body, based on EM data from the fly connectome (top view)[14]. **B** Schematic projections of a ring neuron and a wedge neuron in the ellipsoid body. **C** Ring attractor architecture is linked to wedge neurons and ring neurons. The green color in wedge neurons indicate their activity, illustrating how a bump of activity is located within a set of wedge neurons.

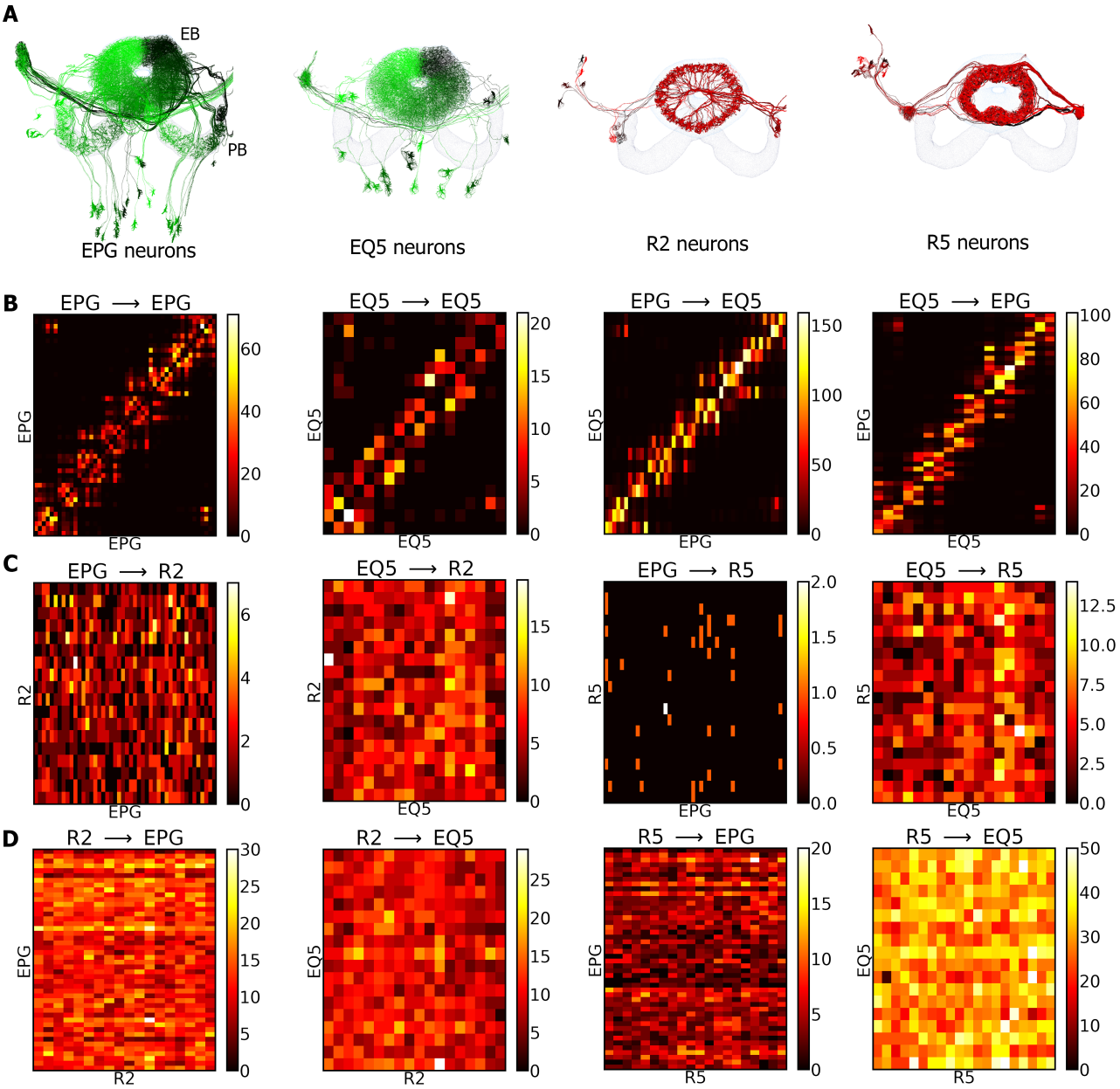

Figure S2. Connectivity between wedge and ring populations based on the fly connectome[14]. **A** In green, two populations of wedge neurons: EPG and EQ5. In red, ring neuron populations that were proposed to encode sleep pressure in the fly: R2[18] and R5[23]. **B** Recurrent connectivity between wedges in the EB between both EPG and EQ5 populations, which are represented as w_{EE} in the proposed models (plastic weights). **C** Connectivity from wedge neurons to ring neurons, represented as w_{IE} (plastic weights) in the proposed models. **D** Connectivity from ring neurons to wedge neurons, modeled as w_{EI} (fixed weights). Connectivity matrices and images are downloaded from[14].

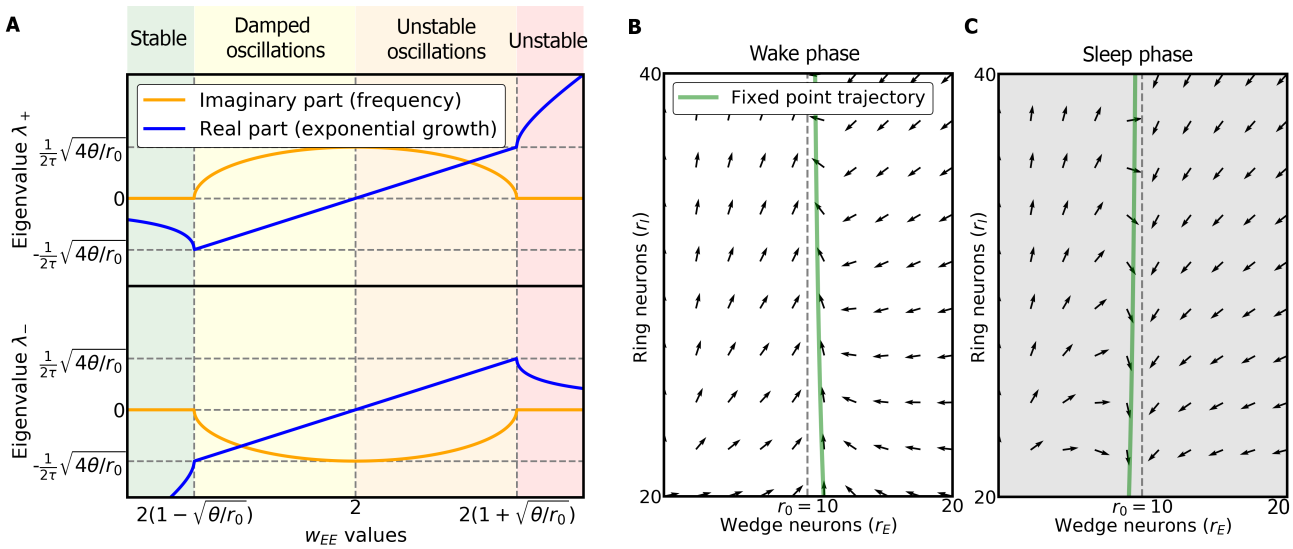

Figure S3. Analysis of the two population model in the fast and slow-timescale limits. **A** Eigenvalues in the two-population model in the fast-timescales limit. The real (in blue) and imaginary (in orange) parts of the eigenvalues are plotted as a function of w_{EE}. **B** Vector field of wedge and ring neurons dependent on synaptic plasticity during the wake phase in the slow-timescale limit. The green line represents the set point trajectory of wedge neurons. **C** Vector field of wedge and ring neurons due to plasticity in the sleep phase in the slow-timescale limit. The green line is the trajectory of the set point in wedge neurons.

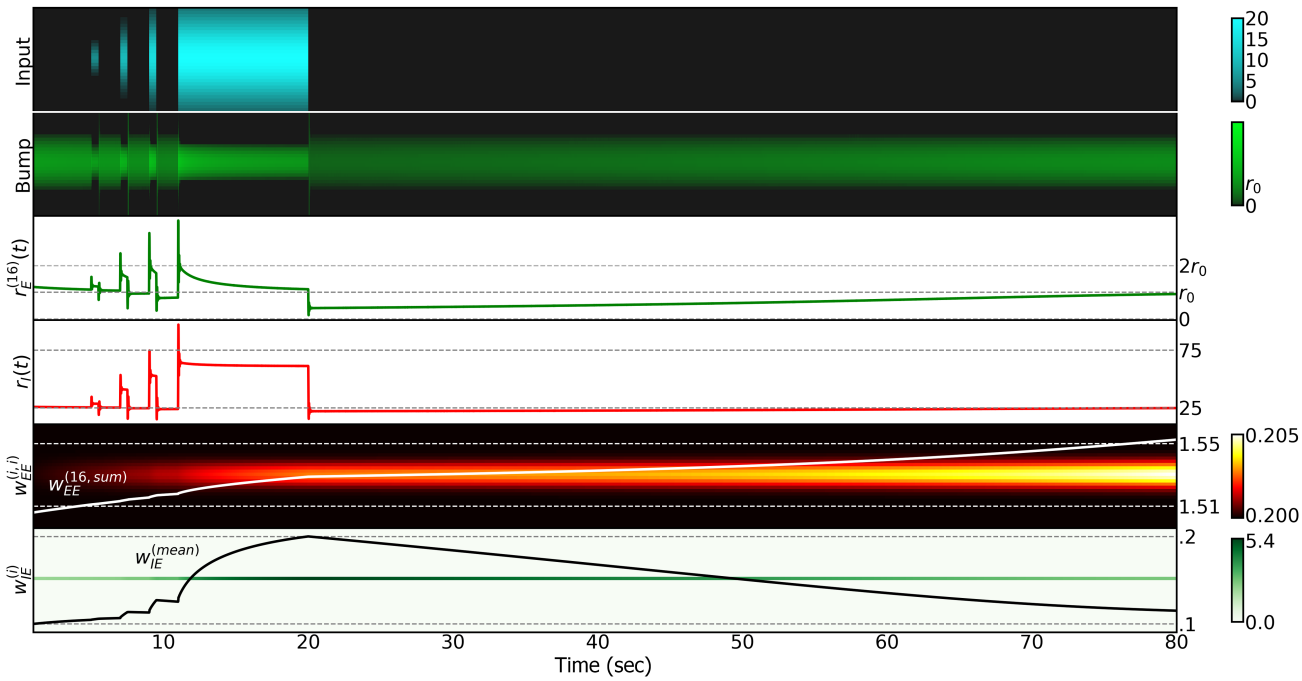

Figure S4. Bump profile in the ring attractor network in response to inputs of different amplitudes I_{max} and standard deviations I_σ. When input (first row) is provided to the ring attractor the plasticity rule for $w_{IE}^{(i)}$ brings the activity of the wedge neuron $r_E^{(16)}$ (where the bump peak is located, second row) back to r_0 (third row). The bump is not constrained to have constant activity (third row), but always relaxes towards r_0 over time in the absence of changing input. For ease of visualization, we used a slower time constant for the plasticity of $w_{EE}^{(ij)}$ ($\tau_{EE} = 1000000$), thus avoiding oscillations in the bump throughout the simulation.

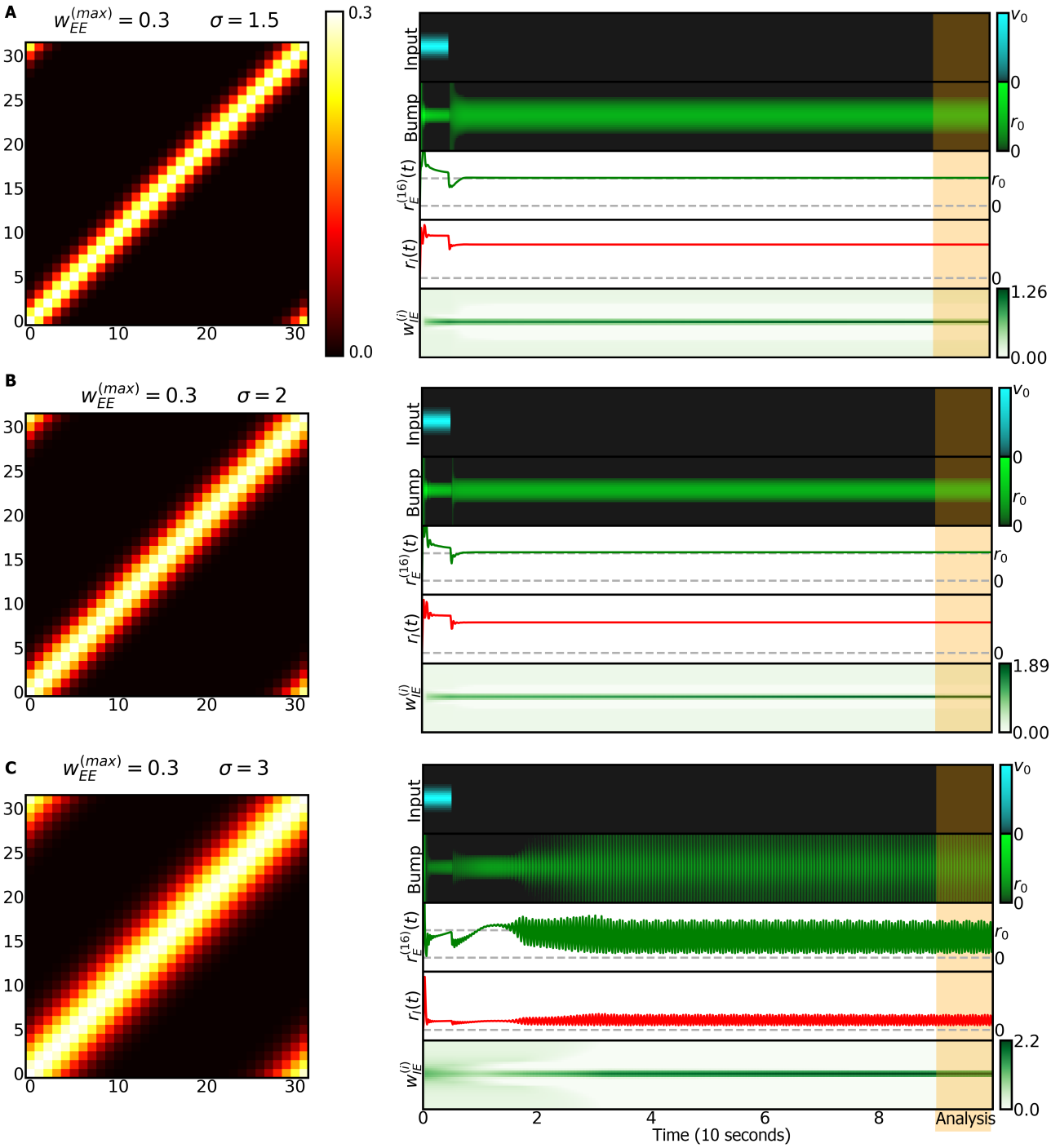

Figure S5. Ring attractor network with no plasticity in $w_{EE}^{(ij)}$ and different initial distributions of weights. **A** Left: initial $w_{EE}^{(ij)}$ values; right: 10 seconds of simulated dynamics. First row: input positioning the bump around wedge neuron 16. Second row: bump profile over time. Third row: activity of wedge neuron 16 evolving towards r_0 due to plasticity in $w_{IE}^{(i)}$. Forth row: activity of ring neurons. Fifth row: the synaptic weights $w_{IE}^{(i)}$. The orange band at the end of the simulations represents the period in which the stability of the bump is analyzed. **B** Same as A but with larger σ. **C** Same as A, B but again increasing σ. In this simulation, the bump shows stable oscillations.

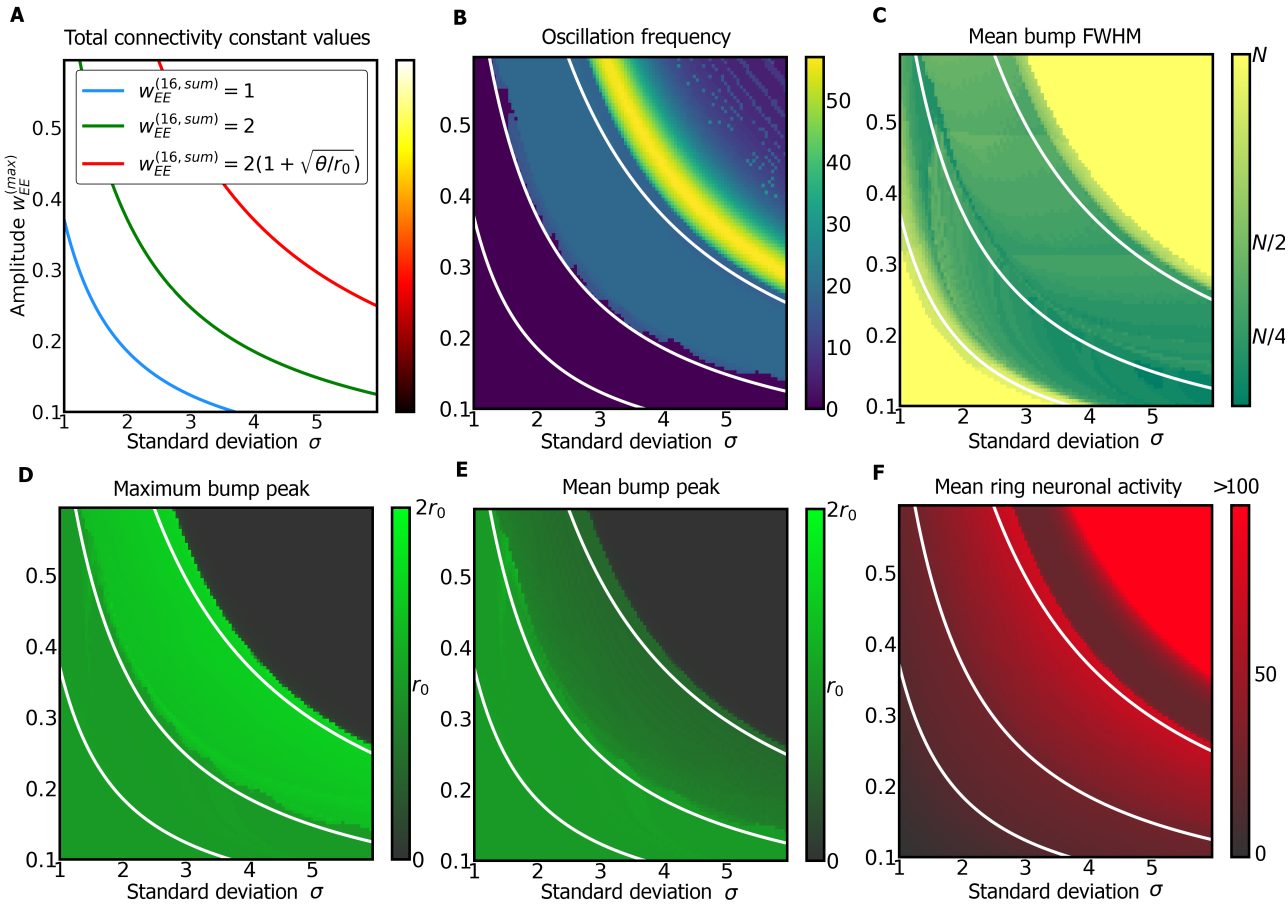

Figure S6. Stability analysis of the bump in the ring attractor network. All subfigures are plots of different measures as a function of the grid of values of σ and $w_{EE}^{(max)}$ that define the fixed synaptic weights $w_{EE}^{(ij)}$ in the bump stability analysis (Methods (4.7)). All variables in B, C, D, E, and F are calculated in a time slot of 1 sec at the end of each simulation (orange area in Figure 2). **A** Isolines of $w_{EE}^{(16,sum)}$. These isolines are overlaid in white in the other subfigures and represent approximate boundaries of stability. **B** Frequency of the oscillations in the bump across the analysis time slot. Above the isoline $w_{EE}^{(16,sum)} = 2$, the bump oscillates. **C** Mean FWHM over the analysis time slot. Below the isoline $w_{EE}^{(16,sum)} = 1$ and above the isoline $w_{EE}^{(16,sum)} = 2(1+\sqrt{\theta/r_0})$ there is no bump because the FWHM is equal to the number of neurons, N. **D** Maximum peak of the bump in the analysis time slot (maximum activity of wedge neuron 16, $max(r_E^{(16)}(t))$). Below the isoline $w_{EE}^{(16,sum)} = 2$, the maximum is r_0 as forced by the plasticity rule in $w_{IE}^{(i)}$. Between the isolines $w_{EE}^{(16,sum)} = 2$ and $w_{EE}^{(16,sum)} = 2(1+\sqrt{\theta/r_0})$, the bump oscillates and the amplitude is given by the maximum, in this case $2r_0$. Above $w_{EE}^{(16,sum)} = 2(1+\sqrt{\theta/r_0})$, there is close-to-zero activity in wedge neurons, since the maximum is near zero. **E** Mean of the bump peak over the analysis slot time (mean activity of the wedge neuron 16, $< r_E^{(16)}(t) >$). In case of no oscillations, this value should be equivalent to the maximum bump peak in panel D. During oscillations, the value closely represents the center of oscillations. Note that between isolines $w_{EE}^{(16,sum)} = 2$ and $w_{EE}^{(16,sum)} = 2(1+\sqrt{\theta/r_0})$ this value is lower than r_0. **F** Mean activity over the analysis time slot of ring neurons, $< r_I(t) >$. Below the isoline $w_{EE}^{(16,sum)} = 2(1+\sqrt{\theta/r_0})$ the activity of ring neurons increases. Above the isoline, the activity of ring neurons first decreases and then rapidly increases. We clipped the values of the mean activity above 100 to facilitate visualization, but the increment of activity in this area reached values over 1000.

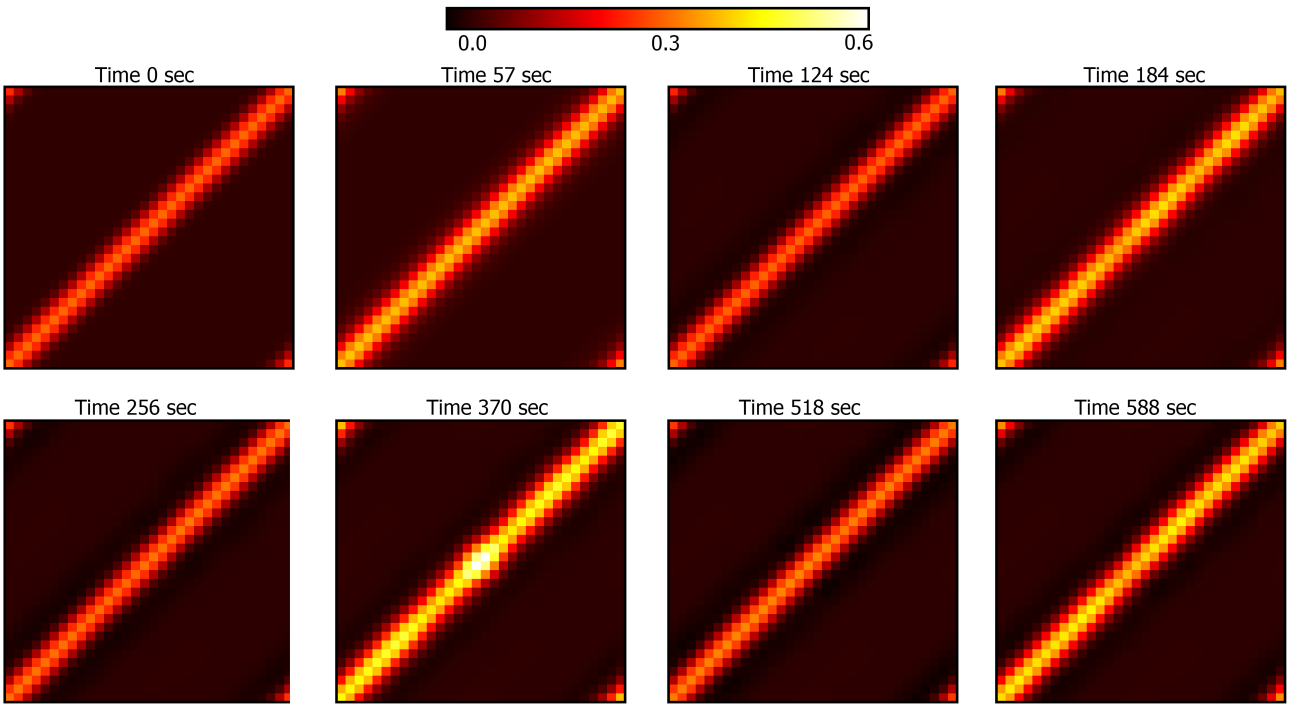

Figure S7. Values of the synaptic weights $w_{EE}^{(ij)}$ in the simulation of Figure 3. From left to right and from top to bottom: each plot shows the values each time there is a switch between wake and sleep phase and *vice versa*. Top, left: initial values of the weights set by the initial values of $w_{EE}^{(max)}$ and σ in Table S2. Top, center left: weights at the end of the first wake phase. Top, center right: after the subsequent sleep phase. Top, right: after the second wake phase. Bottom, left: weights after the second sleep phase. Bottom, center left: after sleep deprivation. Bottom, center right: after sleep rebound. Bottom, right: weights after the third wake phase. Note how the weights after each wake phase are increased (specially after sleep deprivation in the sixth plot) and how the sleep phase resets the weights to close to the initial conditions.

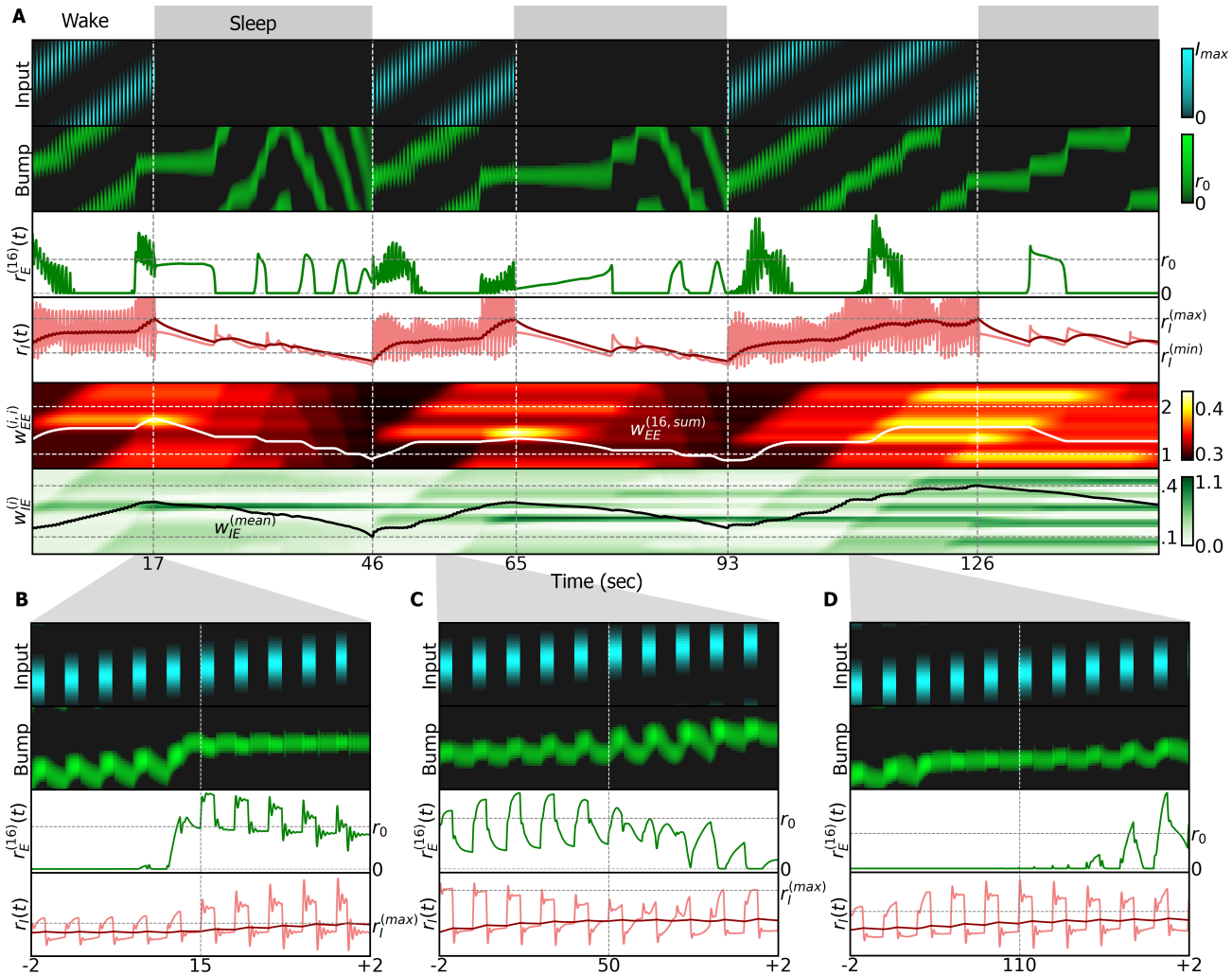

Figure S8. Simulation of the ring attractor network with flashing input during the wake phase (details in Methods(4.9)). **A** Simulation with three wake phases and three sleep phases. A flashing input that turns on and off is provided during the wake phase. During the off period, the ring attractor sustains a bump of activity that drifts due to changes in the synaptic weights $w_{EE}^{(ij)}$. **B** Zoom-in around 15 seconds in the first wake phase of the simulation. **C** Zoom-in around 50 seconds in the second wake phase. **D** Zoom-in around 110 seconds in the third wake phase.

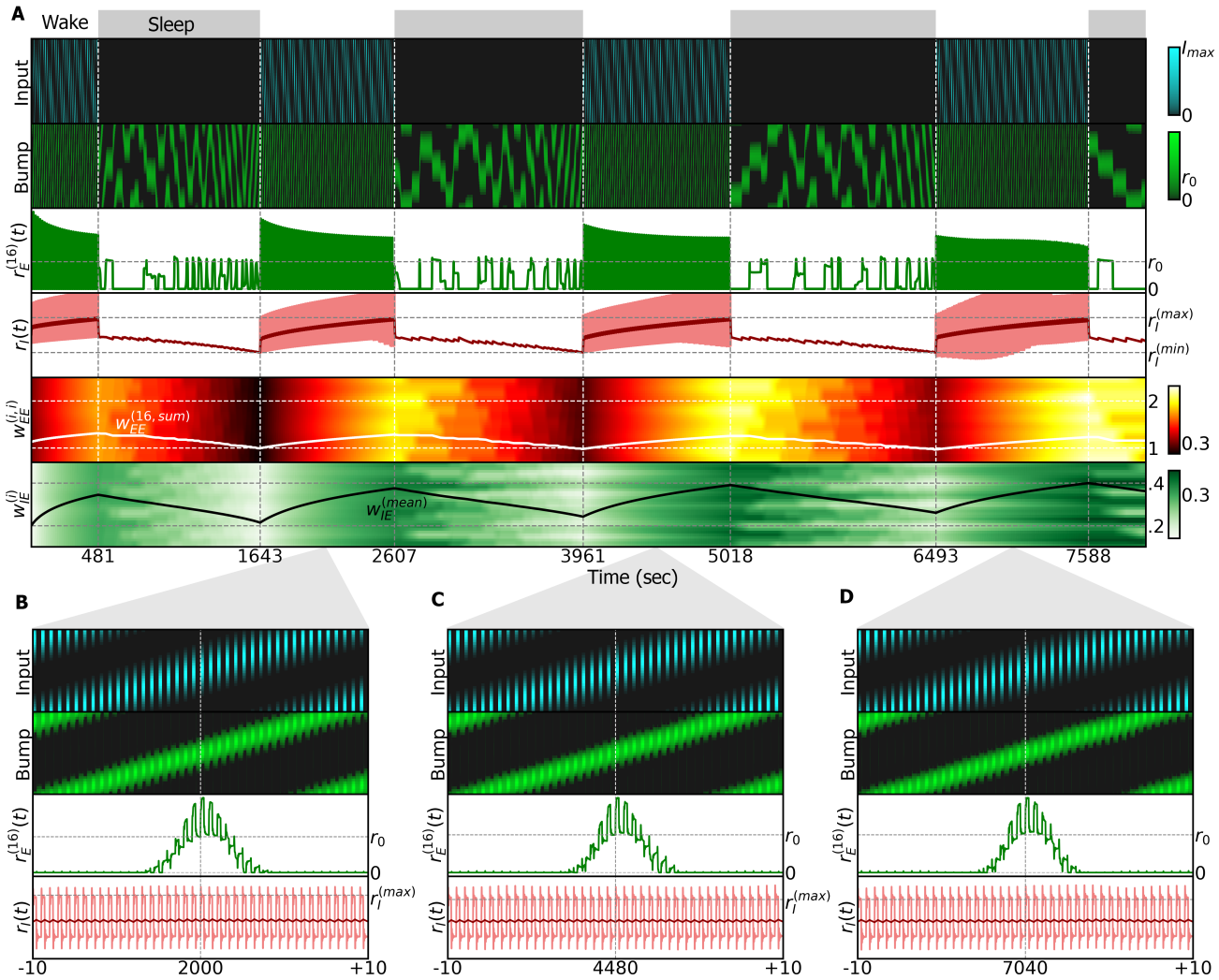

Figure S9. Simulation of the ring attractor network with flashing input during the wake phase with slow plasticity time constants (details in Methods(4.9)). The time constants in this simulation are 100 times larger than in the previous ring attractor simulations. **A** Full simulation with three wake phases and three sleep phases. We used the same input rotation frequency as in Figure S7. As in Figure S7, we provide a flashing input that turns on and off during the wake phase. **B** Zoom-in around 2000 seconds in the first wake phase of the simulation. **C** Zoom-in around 4480 seconds in the second wake phase. **D** Zoom-in around 7040 seconds in the third wake phase. Note how the duration of sleep and wake phases are now in the order of hours due to the increased time constants in the plasticity rules. Note also how the bump of activity is sustained in place without drifting after the input switches off.