The insect compass system: from theory to circuitry

Pau Vilimelis Aceituno, Dominic Dall'Osto, Ioannis Pisokas

Abstract

To navigate their environment, insects need to keep track of their heading direction, with some species such as the monarch butterfly able to maintain a stable direction encoding even over long migratory travels. Previous work has shown that insects encode their heading direction as a sinusoidal pattern of activity in a ring of neurons. However, it is unclear whether this sinusoidal encoding is just an evolutionary coincidence, or if it offers some particular advantage. We address this problem by establishing the basic mathematical requirements for heading integration and show that several circuits with different activity patterns can perform the same function. In this family of potential circuits, a sinusoidal activity pattern stands out as the most noise-resilient, but only when coupled with a specific connectivity pattern between neurons. Using network analysis, we compare this optimal connectivity pattern with experimental data from the locust and the fruit fly, finding a remarkably good agreement. Finally, we demonstrate that the circuit we propose can emerge naturally from a Hebbian plasticity rule, showing that the synaptic structure of our proposed network does not need to be explicitly encoded in the genetic program of the insect, but can be acquired during development.

1 Introduction

Insects exhibit an impressive ability to navigate the world, travelling long distances to find food or reach places of interest before returning to their nests (Müller and Wehner, 1988; Menzel and Muller, 1996; Collett, 2019), a feat that requires them to keep track of their orientation across long travels (Homberg, 2015). A basic requirement for these navigation abilities is heading integration – integrating angular velocity signals over time to maintain an estimate of one's current direction relative to a starting point and angle in space (Darwin, 1873; von Frisch, 1967; Mittelstaedt, 1985; Müller and Wehner, 1988).

Electrophysiology and calcium imaging studies in insects have shown that the neural population encoding heading direction has a sinusoidal shaped activation pattern (Labhart, 1988, 2000; Loesel and Homberg, 2001; Pfeiffer et al., 2005; Pfeiffer and Homberg, 2007; Kinoshita et al., 2007; Heinze et al., 2009; Homberg et al., 2011; el Jundi et al., 2014; ElJundi et al., 2019). Furthermore, the neurons that project into that population encode velocity signals as sinusoidal activations (Lyu et al., 2022; Lu et al., 2022).

Theoretical work has speculated that this recurring motif of a sinusoidal activity pattern might be so prevalent because it enables easy elementwise vector addition, where vectors encoded as sinusoidal activity profiles can be added together to give a sinusoidal profile encoding the sum of the vectors. This allows the heading direction to be used by downstream circuitry to track the insect's position (Mittelstaedt, 1985; Wittmann and Schwegler, 1995; Vickerstaff and Paolo, 2005; Haferlach et al., 2007; Wessnitzer et al., 2008; Sakura et al., 2008; Stone et al., 2017). Further work has shown that models closely aligned with biological data can indeed implement heading integration using sinusoidal activity patterns (Hulse et al., 2021; Pfeiffer, 2022; Lyu et al., 2022).

In this work we show that enabling easy vector addition cannot be the unique driving factor for the presence of sinusoidal encodings, as many other circuits with different activity patterns can perform vector addition in the same way. This finding leads us to question whether the sinusoidal activation patterns seen in insect navigation circuits are a coincidental evolutionary artefact, or if they might offer some particular advantage.

To address this question we consider the basic principles necessary for a circuit encoding direction. Of the circuits fulfilling these requirements, the sinusoidal activity pattern offers the best resilience to noise for the encoded information. However, implementing this activity requires a circuit with a specific connectivity pattern between neurons. Thus, our theory provides a concrete prediction for the most noise resilient heading integration circuit. We compare our predicted circuit with connectivity

data for the locust and fruit fly using network analysis tools, showing a strong agreement. Finally, we ask the question how an insect brain might develop such a circuit, finding that a simple Hebbian learning rule is sufficient.

2 A theoretical circuit for head integration

We consider a population of N "compass neurons" with activity that encodes the direction of the insect as an angular variable θ. We represent the activity of this population by a vector where each entry corresponds to the activity of one neuron $\mathbf{a}(\theta) = [\mathbf{a}_1(\theta), \mathbf{a}_2(\theta), ..., \mathbf{a}_N(\theta)]$. We take $N = 8$ neurons, consistent with data from many insect species which possess an eight column organisation, with each one encoding a different direction (Pisokas et al., 2020).

Each neuron's activity is updated depending on its current firing rate and the inputs it receives, both from other neurons in the circuit and externally. We formulate this update rule as follows,

$$\dot{\mathbf{a}}(\theta) = -\mathbf{a}(\theta) + \phi\left[\mathbf{W}\mathbf{a}(\theta) + \mathbf{u}(t)\right], \tag{1}$$

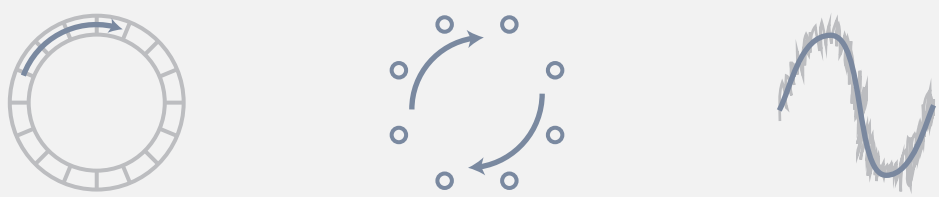
where $\mathbf{W}$ is the circuit's weight matrix, representing the connections between neurons; ϕ is the neural activation function, that converts the total neural input into an output firing rate; and $\mathbf{u}(t)$ is the external input that encodes the insect's velocity, coming from the population of speed encoding neurons (Stone et al., 2017; Hulse et al., 2021; Sayre et al., 2021; Lu et al., 2022; Lyu et al., 2022).

To simplify our derivations we allow the neural activity values to be both positive and negative, interpreting these values as being relative to a baseline neural firing rate. Similarly, we allow the weights to be both positive and negative, a common simplification in computational models (Kadmon and Sompolinsky, 2015; Kriegeskorte and Golan, 2019). This simplification will be addressed in Section 3 where we compare our model with experimental data.

2.1 Mathematical principles for neural heading integration

A circuit capable of performing heading integration must fulfil the requirements outlined in Box 1. The first two requirements allow us to establish the family of possible path integration circuits. We then use the principle of noise minimisation to determine which circuits perform best.

Box 1: Requirements for a heading integration circuit



Ring topology:
The activity should have the same topology as the variable it is encoding to prevent discontinuities. To encode heading, the activity should have the topology of a 1D circle.

Rotational Symmetry:
The heading integration circuit should work similarly, irrespective of the direction in which the insect travels. There shouldn't be a bias for any direction.

Noise minimisation:
The circuit should minimise the noise of the neural representation so the insect can navigate as precisely as possible.

2.2 Constraints on the neural activity

The neural activity should encode the insect's heading direction with the matching topology. Because the heading is a singular angular variable, the topology of the activity space should be a 1D circle, or ring.

Furthermore, the symmetry requirement implies that rotating the heading of the insect should rotate the neural activity vector without changing its shape. Concretely, if the insect is facing north or east, the activity of the neural population as a whole should be the same, but the identity of the neuron with each activity value will be different.

We can formalise the symmetry requirement by considering a heading direction, θ, and a rotation by an integer multiple, k, of the angular spacing between neurons, $\Delta\theta = \frac{2\pi k}{N}$. In this case, individual neuron activities follow the relation

$$\mathbf{a}_n(\theta + \Delta\theta) = \mathbf{a}_{n-k \mod N}(\theta), \tag{2}$$

which enforces that the neural activity vector is circularly rotated as the heading direction changes. This relation can be expressed in the Fourier domain where, by the shift property, the circular rotation becomes multiplication by a complex exponential:

$$\mathcal{F}_f\left[\mathbf{a}(\theta + \Delta\theta)\right] = \mathcal{F}_f\left[\mathbf{a}(\theta)\right] e^{\frac{i2\pi k f}{N}}, \tag{3}$$

where $i = \sqrt{-1}$ and $f \in [0, ..., N-1]$ is the index of the spatial frequency or harmonic – activity with spatial frequency f has f "bumps" around the ring.

Since the complex exponential has unit norm, the magnitude of the Fourier components remain the same for any rotation, $\|\mathcal{F}[\mathbf{a}(\theta + \Delta\theta)]\| = \|\mathcal{F}[\mathbf{a}(\theta)]\| \, \forall \Delta\theta$. Therefore, the shape of the activity pattern, $\mathbf{a} \equiv \mathbf{a}(0)$, can also be fully specified by its Fourier domain representation, $\mathcal{F}[\mathbf{a}]$, and the phase of this activity profile around the network encodes the heading of the insect:

$$\mathcal{F}_f\left[\mathbf{a}(\theta)\right] = \mathcal{F}_f\left[\mathbf{a}\right] e^{i\theta f}. \tag{4}$$

Taking the inverse Fourier transform with the constraint that the neural activities must be real, we get the following form for the neural activity,

$$\mathbf{a}_n(\theta) = \sum_{f=0}^{N-1} \|\mathcal{F}_f\left[\mathbf{a}\right]\| \cos\left(\frac{2\pi f n}{N} + \theta f\right). \tag{5}$$

It should be noted that the phase offset of each cosine waveform scales with the harmonic. This is because higher frequency waveforms have shorter wavelengths (in terms of number of neurons), so to move at the same speed higher frequency waveforms need their phases to rotate more quickly.

This is shown in Fig. 1, bottom row. For a network with $N = 8$ neurons, a 180° rotation shifts the activity waveform by 4 neurons. This 4 neuron shift corresponds to a $\frac{4}{8}$ or 180° phase offset for the $f = 1$ waveform (with a wavelength of 8 neurons), but a $\frac{4}{4}$ or 360° phase offset for the $f = 2$ waveform (with a wavelength of 4 neurons). In this case the $f = 2$ waveform is the same whether the heading angle 0° or 180° is encoded.

This means that higher harmonic waveforms are not on their own sufficient to uniquely determine the encoded angle. As such, a hierarchical decoding scheme is required to resolve these ambiguities by considering activity in all harmonics simultaneously, as detailed in Appendix II.

An alternative activity formulation is to relax the requirement in Eq. 2 that activity should rotate around the network at the same speed as changes in the insect's heading. For example, if activity around the network were to rotate at half the speed of the insect's heading, the $f = 2$ waveform could uniquely encode all angles while preserving the rotational symmetry requirement. Eq. 5 can be reformulated as:

$$\mathbf{a}_n(\theta) = \sum_{f=f_{\text{base}}}^{N-1} \|\mathcal{F}_f\left[\mathbf{a}\right]\| \cos\left(\frac{2\pi f n}{N} + \theta \frac{f}{f_{\text{base}}}\right), \tag{6}$$

where f_{base} is the lowest spatial frequency present in the activity, such that $\frac{N}{f_{\text{base}}}$ is the wavelength of the activity profile.

The family of possible neural activity forms that satisfy the heading integration requirements detailed in Section 2.1 can be divided based on which spatial frequencies are active in the encoding. We denote the set of active encoding frequencies as $\mathbf{F}$, defined by:

$$\mathbf{F} = \{f \in [1..N-1] : \|\mathcal{F}_f[\mathbf{a}]\| \neq 0\}. \tag{7}$$

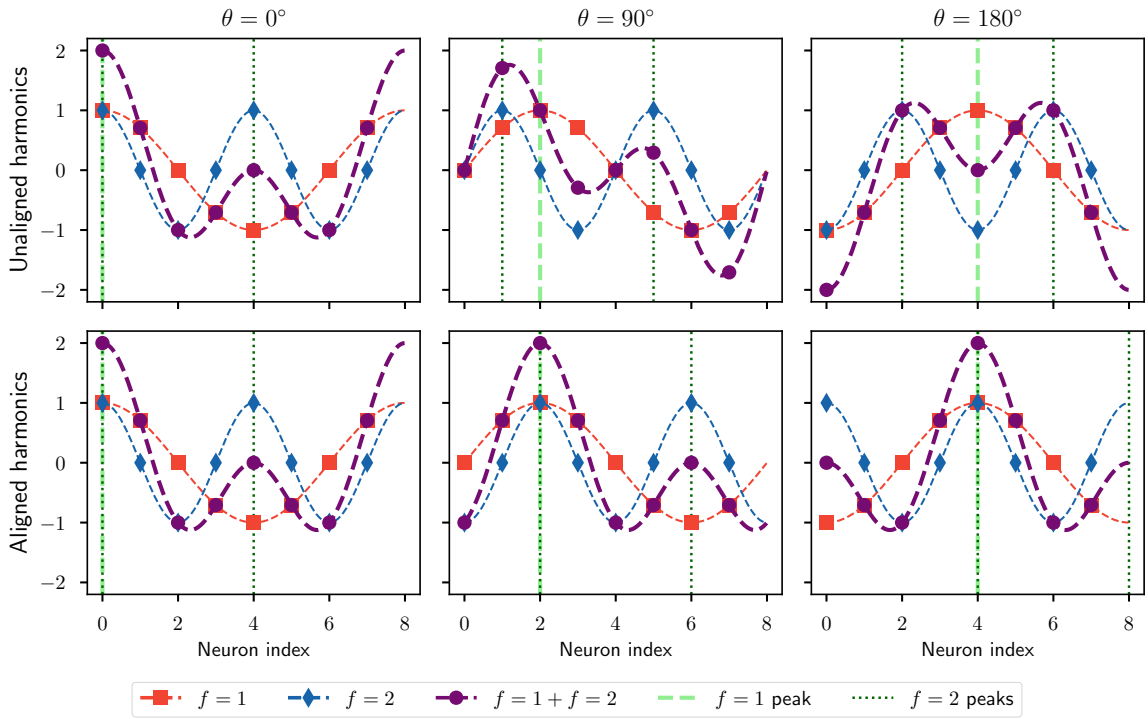


Figure 1: **Encoding heading with multiple harmonics.** Each panel shows the activity profiles encoding a particular heading value. Curves $f = 1$, $f = 2$, and $f = 1 + f = 2$ respectively denote the waveform of the first harmonic, second harmonic, and the sum of the two. The vertical dashed ($f = 1$ peak) and dotted lines ($f = 2$ peaks) indicate the neurons which respond maximally for the first and second harmonic, respectively. **Top row:** Encoding the heading as the sum of two independent harmonics causes the waveform to change shape as the insect rotates, because the waveform for each harmonic can only rotate a distance of $\frac{360°}{f}$ as the insect rotates a full revolution to ensure an unambiguous representation. **Bottom row:** If all harmonics are aligned to rotate at the same speed, the combined waveform shape does not change. However, this alignment implies that higher harmonic waveforms cannot be uniquely mapped back to a heading: here the $f = 2$ encoding is the same for $\theta = 0°$ and $180°$.

2.3 Constraints on heading integration circuits

The basic assumptions outlined earlier also constrain the possible heading integration circuits – such a circuit should allow activity of the required topology and rotational symmetry to stably exist and propagate.

The circular topology requirement constrains the activity in the circuit to have a constant total magnitude. We consider that this constraint is enforced by the nonlinear neural activation function, ϕ, in Eq. 1, as detailed in Appendix I.

If the network activity is at the desired level, and we consider the external input $u(t)$ to be projected onto the ring attractor such that it does not alter the total network activity, then we can consider the network dynamics to be linear at this operating point. Therefore, our circuit dynamics are effectively described as

$$\dot{\mathbf{a}}(\theta) = -\mathbf{a}(\theta) + \mathbf{W}\mathbf{a}(\theta) + \mathbf{u}(t). \tag{8}$$

The rotational symmetry principle also applies to the network. For the same shaped activity waveform to be able to stably exist at any position around the network, the network connectivity should also be rotationally symmetric. For example, the connection strength between the neurons encoding north and north-east directions should be the same as between those encoding south and south-east. Mathematically, this imposes that the weight matrix, $\mathbf{W}$, is circulant, specifically that $\mathbf{W}_{n,m} = \mathbf{W}_{n+k,m+k}$. This matrix is fully specified by its first row, called the *connectivity profile* and

denoted ω, meaning that we can express the product of the matrix with the neural activity as

$$(\mathbf{W}\mathbf{a}(\theta))_k = \mathbf{W}_k\mathbf{a}(\theta) = \sum_{m=0}^{N-1} \mathbf{W}_{k,m}\mathbf{a}_m(\theta) = \sum_{m=0}^{N-1} \mathbf{W}_{0,m-k}\mathbf{a}_m(\theta) = \sum_{m=0}^{N-1} \omega_{m-k}\mathbf{a}_m(\theta), \tag{9}$$

which can be simplified in terms of the convolution operation,

$$\mathbf{W}\mathbf{a}(\theta) = \omega * \mathbf{a}(\theta). \tag{10}$$

Considering the case where the insect is not moving, $\mathbf{u}(t) = 0$, and the network activity is stable, $\dot{\mathbf{a}}(\theta) = 0$, we can combine Eq. 8 and Eq. 10 to get a relation for the stable network activity

$$\mathbf{a}(\theta) = \omega * \mathbf{a}. \tag{11}$$

In the Fourier domain, this simplifies into

$$\mathcal{F}_f\left[\mathbf{a}(\theta)\right] = \mathcal{F}_f\left[\omega\right]\mathcal{F}_f\left[\mathbf{a}(\theta)\right]. \tag{12}$$

As for the activity waveform, we note here that the Fourier transform is taken on the neural indices, not on the temporal domain.

Here we have N equations that have to be satisfied, since the Fourier transform of an N-dimensional vector is also N-dimensional. Each harmonic frequency f, Eq. 12 has two solutions:

- $\mathcal{F}_f\left[\omega\right] = 1$, which implies that activity with this spatial frequency is stable in the network, and so can encode the insect's heading.
- $\mathcal{F}_f\left[\mathbf{a}(\theta)\right] = 0$, meaning that the activity in this harmonic is zero, and thus nothing can be encoded in this frequency.

2.3.1 Minimising noise propagation in the circuit

The only constraint on the connectivity weights given by Eq. 12 is that, if a frequency is used for encoding then $\mathcal{F}_f\left[\omega\right] = 1$. There is no restriction on the weights for the inactive harmonics – they remain free parameters.

But non-encoding channels can still propagate noise, which would be prevented by setting $\mathcal{F}_f\left[\omega\right] = 0$. To illustrate this, we consider white noise denoted by ϵ that is added to the neural activity. When this noisy activity evolves as the dynamics from Eq. 10 dictate,

$$\omega * (\mathbf{a}(\theta) + \epsilon) = \omega * \mathbf{a}(\theta) + \omega * \epsilon, \tag{13}$$

where the term $\omega * \epsilon$ corresponds to noise and should therefore be dampened. Specifically, we want to minimise the variance of that noise

$$\mathbf{Var}\left[\omega * \epsilon\right] = \mathbf{E}\left[\|\omega * \epsilon\|^2\right] = \|\omega\|^2\mathbf{Var}\left[\epsilon\right]. \tag{14}$$

Hence the magnitude of the noise that passes from one time interval to the next is modulated by the magnitude of the weight vector. By Parseval's theorem,

$$\sum_f |\mathcal{F}_f\left[\omega\right]|^2 = \sum_n |\omega_n|^2 = \|\omega\|^2, \tag{15}$$

which implies that to minimise noise propagation in the network we should impose $\mathcal{F}_f\left[\omega\right] = 0$ for all harmonics f where $\mathcal{F}_f\left[\mathbf{a}(\theta)\right] = 0$. We show this in simulation in Fig. 2, where only the first harmonic encodes information and we vary the weights of the other harmonics. The noise is indeed minimised if the weights for all non-encoding harmonics are set to zero.

This result establishes that all non-encoding harmonics in the network should be set to 0 to minimise noise propagation, and allows us to recover the circuit

$$\omega_n = \sum_{f\in\mathbf{F}} \cos\left(\frac{2\pi n}{N} f\right) \tag{16}$$

where $\mathbf{F}$ is the set of harmonics used to encode the heading direction as defined in in Eq. 7.

The choice of $\mathbf{F}$ therefore gives us the harmonics that will be used but also the connectivity of the circuit that encodes the heading direction. We can choose any combination of frequencies and we must pick at least one encoding frequency, thus the number of possible heading integration circuits is $2^N - 1$.

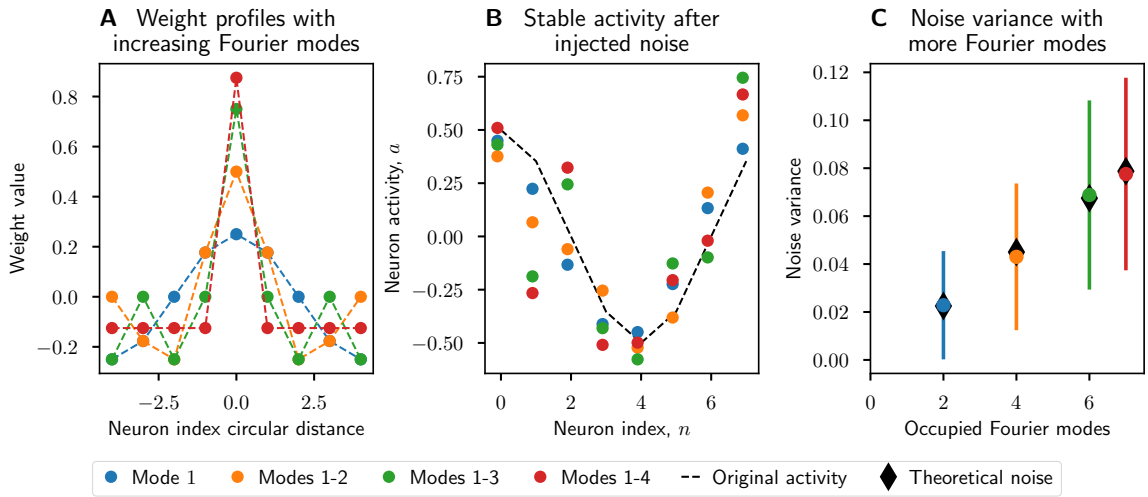


Figure 2: **Increased harmonics number introduce more noise.** Increasing the number of active harmonic frequencies (Fourier modes) increases the effect of errors in the network. **A:** Weight matrix profiles, ω, for networks with increasing numbers of Fourier modes. **B:** Normally distributed noise with zero mean and standard deviation 0.3 was added to the network activity, then the network state was updated until it had settled. Networks with fewer Fourier modes better filtered out noise – reducing the overall effect of added noise. **C:** Noise variance increases linearly as the number of Fourier modes in the weight profile increases, as predicted by Eq. 14. Occupied modes increase in steps of 2 because of the $+f, -f$ symmetry of the Fourier transform.

2.4 Determining the optimal circuit

2.4.1 Multiple harmonic circuits

We broadly separate the family of possible circuits into those that use multiple harmonics, and those that only use a single harmonic.

A limitation of circuits using multiple harmonic is that they do not achieve the best signal-to-noise ratio if the total network activity is limited (by, for example, the energy consumption of the circuit). For a signal with total magnitude $\|\mathbf{a}\|^2$ and noise given by Eq. 14, the signal to noise ratio is:

$$\text{SNR} = \frac{\|\mathbf{a}\|^2}{\|\omega\|^2 \mathbf{Var}\left[\epsilon\right]} = \frac{\|\mathbf{a}\|^2}{|\mathbf{F}| \mathbf{Var}\left[\epsilon\right]} \tag{17}$$

where $|\mathbf{F}|$ is the number of encoding frequencies. Because white noise independently affects all harmonics equally, increasing the number of encoding frequencies increases the amount of noise in the network.

An additional limitation of multiple harmonic circuits is that adding their activity profiles does not necessarily result in another compatible activity profile, because the harmonic waveforms have different phases. A visual example is shown in Fig. 1, bottom row. If the $f = 1 + f = 2$ waveforms at 0° and 180° are added together, the $f = 1$ components will cancel each other out while the $f = 2$ components will reinforce each other, resulting in a waveform with only $f = 2$ activity. This violates the desired property of easy elementwise vector addition, in contrast to single harmonic circuits.

2.4.2 Single harmonic circuits

This leaves us with the conclusion that a single harmonic should be used, reducing the number of possible optimal circuits to N. As we consider networks with $N = 8$ neurons, the possible harmonics are $f = \{1, 2, 3, 4, 5, 6, 7\}$, where the zeroth harmonic is discarded because it just represents the baseline neural activity. This gives the following activities and weights, derived from Eq. 6 and Eq. 16

$$\mathbf{a}_n(\theta) = \cos\left(\frac{2\pi n f}{N} + \theta\right) \quad \omega_n = \cos\left(\frac{2\pi f n}{N}\right). \tag{18}$$

The circuits for the first 4 harmonics are plotted in Fig. 3. But not all of these circuits are valid. In particular, circuits with even frequencies have multiple neurons with identical activity values because they share a common divisor with $N = 8$, explained in detail in Appendix III.2.

For $f = 4$ there are only 2 unique activity values, $\mathbf{a}_n = \pm\cos(\theta)$, so this circuit can only encode one dimension, not a circular topology. For $f = 2$ and $f = 6$ (not plotted), there are 4 unique activity values, allowing the angle to be properly encoded. However, these circuits are degenerate because they are actually composed of two independent subcircuits, each encoding one direction (see Fig. 3 B). Because the subcircuits are independent, they can't constrain the activity to have the required circular topology.

All circuits with odd frequencies are equivalent. For example, as shown in Fig. 3, the connections in the circuit for $f = 3$ are the same as those for the $f = 1$ circuit after the neuron identities are permuted. As detailed in Appendix III.1, the frequencies $f = \{1, 3, 5, 7\}$ always give the same circuit because the odd frequencies are coprime with the number of neurons $N = 8$.

Since the activities and weights of all the non-degenerate circuits $f = \{1, 3, 5, 7\}$ are the same as the base harmonic $f = 1$ up to a permutation, we choose the lowest harmonic $f = 1$, which gives us the following activity and weights:

$$\begin{aligned} \mathbf{a}_n(\theta) &= \cos\left(\frac{2\pi n}{N} + \theta\right) \\ \omega_n &= \cos\left(\frac{2\pi n}{N}\right). \end{aligned} \tag{19}$$

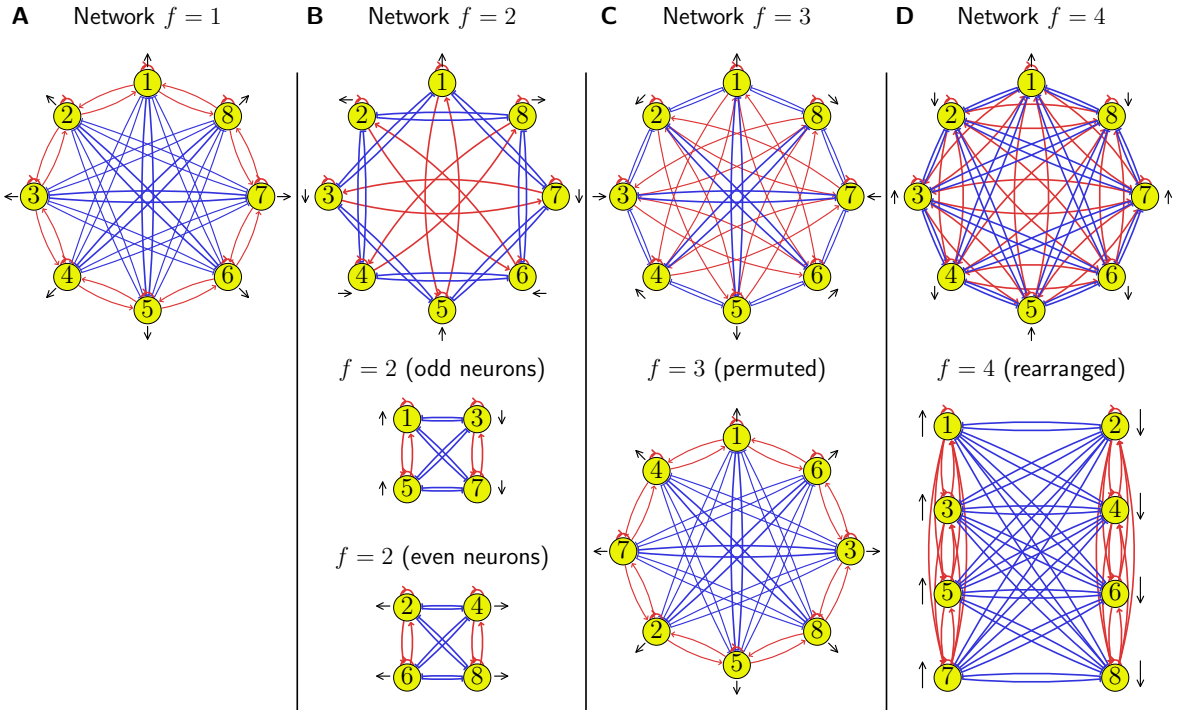


Figure 3: **Circuits for encoding with different harmonics.** We plot four circuits corresponding to each of the individual harmonics $f = 1, 2, 3, 4$. Excitatory synapses are marked in red and inhibitory in blue. Neurons are shown in yellow with each having a black arrow that marks the direction to which it is tuned from Eq. 18. The $f = 1$ circuit is the simplest and constitutes our baseline. For the other cases we plot the original connectivity in the upper row and the rearranged network in the lower row. We find that the $f = 2$ circuit consists of two independent subnetworks encoding orthogonal directions, the $f = 3$ case is identical to $f = 1$ after permuting neuron indices, and the $f = 4$ case results in two connected groups of neurons inhibiting each other, hence it can only encode one direction. As such, all cases either have a degenerate ring structure ($f = 2, 4$) or are equivalent to $f = 1$ after permutation.

3 Comparing the predicted circuit with biological data

Our theory proposes an activity and a circuit for heading integration. A sinusoidal activity profile encoding the heading via its phase has been observed experimentally in species such as the locust and the fruit fly (Zittrell et al., 2023; Turner-Evans et al., 2017), and was part of the motivation for this study. Therefore, we seek to validate our novel prediction – that the circuit weights should be sinusoidal. We test this prediction against connectivity data from the locust and fruit fly, both of which perform heading integration (Kim and Dickinson, 2017; Turner-Evans et al., 2017; Zittrell et al., 2023).

Our model contains a number of simplifications that do not allow a direct comparison to the biological circuit:

- In the model we considered a population of eight neurons, but in insects there are eight neural columns, each with several neuron types (EPG, PEG, PEN, $\Delta 7$). Of these we model only the EPG (known as the compass neurons) which encode the integrated heading direction.
- The synaptic connections between neurons in our model can be both positive and negative, while biological neurons follow Dale's law, meaning that a single neuron can either have only positive or only negative synapses.
- The neural activity in the model was centred around zero and could have both positive and negative activity (firing rates), while real neurons only have positive firing rates.

Therefore, we simplified the biological connectivity to produce an equivalent circuit that could be directly compared with our model prediction. The neural population in our theoretical model corresponds to the biological EPG neurons, as these encode the integrated heading. We considered the other three neurons types that are part of the compass circuit (PEG, PEN, $\Delta 7$, see Pisokas et al. (2020)) as just implementing connections between EPG neurons in accordance with biological constraints. We counted the number of different paths between EPG neurons, accounting for the sign of the connections (whether the path passed through an inhibitory $\Delta 7$ neuron), and used the net path count as a proxy for connectivity strength. This process is explained in detail in Appendix IV.

We then computed the average connectivity profile, how each neuron connected to its neighbours around the ring, and compared this profile to the closest fitting sinusoid. Because the neuron gains, absolute synaptic strength and membrane properties of the neurons are unknown, the units of the net path count are not necessarily equivalent to our abstract connection strength. We therefore fit an arbitrarily scaled and shifted sinusoid:

$$\omega_{m-n} = \beta \cos\left(\frac{2\pi(m-n)}{N}\right) + \gamma, \tag{20}$$

where $m-n$ is the circular distance between two neurons and β, γ are constants that are fit to minimise the precision weighted mean squared error compared with the experimental connectivity profile (see Appendix VI).

Analysing the data from Pisokas et al. (2020) we consider the shortest excitatory and inhibitory pathways between EPG neurons in the locust, which have lengths of 2 and 3 respectively. There are no direct connections of length 1, and all the paths of length ≥ 4 must pass through the same neuron type multiple times. This path counting analysis is shown in Fig. 4 A, and the procedure is detailed in Appendix IV. We find that the connectivity profile between neurons for the locust is very close to sinusoidal in shape, supporting our theoretical prediction.

For the fruit fly, the connectivity is very broad, with neurons that connect almost to every column. To test our theory in this scenario we need to consider not only whether neurons are connected, but also how strongly. We use data from Hulse et al. (2021), which provides synapse counts between pairs of neurons in the fruit fly, and we use these synapse counts as a proxy for connectivity strength, a view that has been validated in previous experiments (Liu et al., 2022; Barnes et al., 2022). After identifying the neurons and connections of interest, we grouped the neurons in eight columns following the logic presented in Pisokas et al. (2020), with detailed methodology explained in Appendix V.

We repeated the path counting analysis for the fruit fly with synapse count data (Fig. 4, B), and found that while the data is noisy, the connectivity profile is consistent with our prediction of a sinusoid. We find that the sinusoid is the best match compared to other possible circuits (see Appendix VI).

We therefore confirm our prediction of a sinusoidal shaped weight pattern in the navigation circuits of both insect species – using connectivity-level data for the desert locust and synapse-count data for the fruit fly.

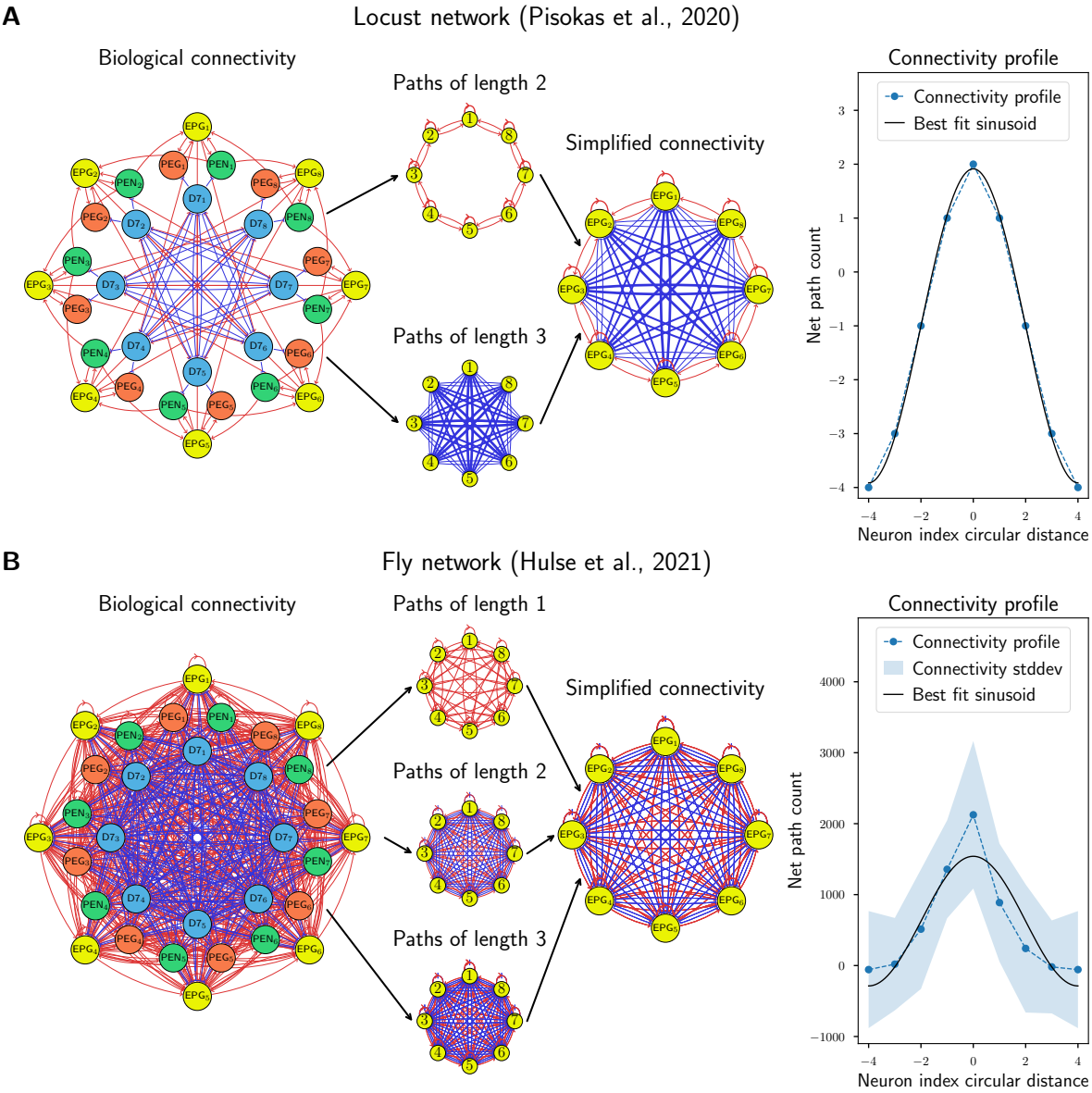


Figure 4: **Comparison to biological data.** Biological network models with 4 distinct neural populations were simplified to equivalent networks with 1 population by counting paths of length 1, 2, and 3 between EPG neurons (which encode the integrated heading direction) and using the net signed path count as a proxy for connectivity strength. The average connectivity profile for each neuron to its neighbours around the ring was compared to the sinusoidal connectivity predicted by our theory. The locust network has no standard deviation because in the original data from Pisokas et al. (2020) the connections were found to be same for neurons in each column. Excitatory connections are shown in red and inhibitory connections in blue, while the strength of a connection is indicated by its width.

4 Learning rules and development

Having validated the connectivity of our theoretical circuit by comparing it to experimental data, we ask whether our circuit lends itself to biological development. Specifically, we show that even though our circuit requires precise connection strengths, this connectivity can be developed naturally by a Hebbian learning rule.

Because the weight matrix is circulant, its eigenvalues are equal to the Fourier spectrum of its first row, which we derived to only have one nonzero value, for $f = 1$. The network therefore only has a single eigenvalue and so projects activity into a single dimension. This operation is similar to classical dimensionality reduction methods such as principal component analysis, which can be implemented by Hebbian-like learning rules (Dayan et al., 2003).

We thus analyse the effects of incorporating Oja's rule into our model, a classical variant of Hebbian learning where the synaptic strength between two neurons grows when both neurons are active simultaneously, and the total synaptic strength is regularised to prevent exploding weight growth,

$$\frac{d\mathbf{W}_{n,m}}{dt} = \eta \left(a_m(t) a_n(t) - a_n^2(t) \mathbf{W}_{n,m} \right), \tag{21}$$

where $\mathbf{W}_{n,m}$ is the synaptic connection strength from neuron m to neuron n, a_m and a_n are the pre- and postsynaptic activities, and η sets the speed of the weight update dynamics relative to the activity dynamics.

For our analysis we integrate the weight updates over some long period of time, after which the insect has visited many directions,

$$\Delta\mathbf{W}_{n,m} = \eta \int_\theta \left[a_m(\theta) a_n(\theta) - a_n(\theta)^2 \mathbf{W}_{n,m} \right] d\theta, \tag{22}$$

where θ is the integration space. We will assume that the insect moves uniformly over the full circle. Applying the activity from Eq. 19, we can find the fixed point of this update rule, when $\Delta\mathbf{W}_{n,m} = 0$:

$$\mathbf{W}_{n,m} = \frac{\eta \int_\theta \cos\left(\theta + \frac{2\pi n}{N}\right) \cos\left(\theta + \frac{2\pi m}{N}\right) d\theta}{\eta \int_\theta \cos^2\left(\theta + \frac{2\pi n}{N}\right) d\theta} = \cos\left(\frac{2\pi(n-m)}{N}\right). \tag{23}$$

Combined with Eq. 9 and Eq. 19, this result means that if there is sinusoidal activity in the network, the weights will naturally converge to the optimal sinusoidal values by way of Oja's rule. These weights will then enforce that only sinusoidal activity can propagate in the network, increasing robustness to noise.

Note that this rule or related ones do not work for combinations of harmonics, because the two harmonics would compete in the learning, as shown in Appendix VII. This supports the choice of a single encoding harmonic.

This finding has two interesting consequences from a biological standpoint: First, our circuit can emerge even when starting with only very coarse initial weights, without the need for high-precision initial connectivity. Second, this simple plasticity rule allows the system to repair or recover from perturbations in its synapses as shown by simulations in Fig. 5.

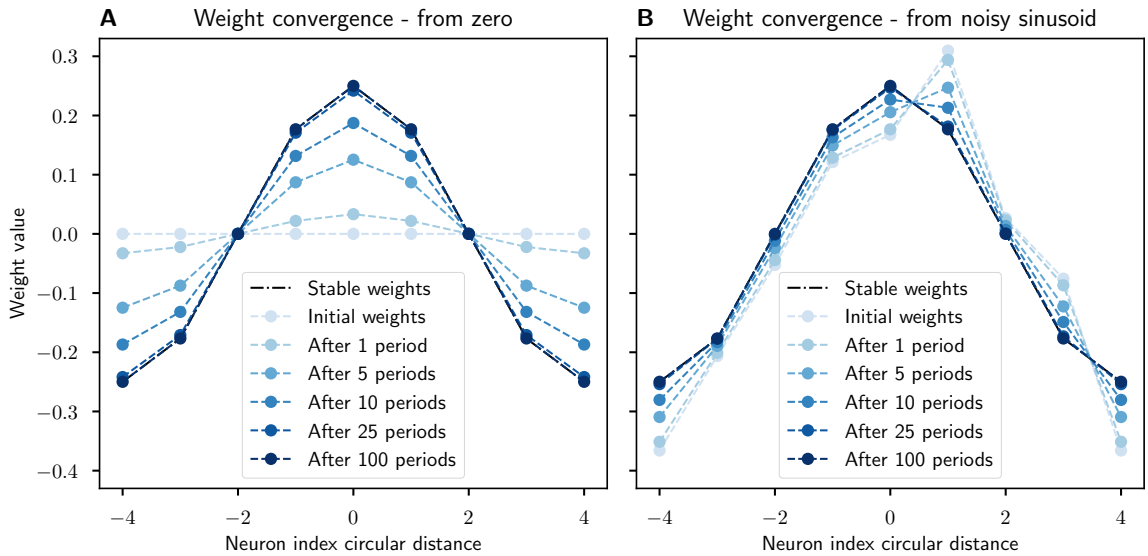


Figure 5: **Recovery of synaptic weights with a Hebbian learning rule.** The synaptic weights converge to a sinusoidal pattern under Oja's rule when the network activity is dynamic and a sinusoidal input is provided. On the left, the weights start at zero and slowly converge to the prescribed sinusoidal profile, showing that this connectivity can emerge from scratch. On the right, the sinusoidal weights are perturbed by noise but learning ensures that the weight profile is corrected. In both cases the network's initial activity is corrupted with zero mean Gaussian noise. Noisy sinusoidal input is provided to rotate the activity bump around the network at a speed of 1/8 neurons per timestep. The simulation runs for 100 periods. Parameters: $N = 8$, integration timestep $\Delta t = 0.01, \eta = 0.1, \|\mathbf{a}\| = 1, \sigma_W = 0.2, \sigma_a = 0.2, \sigma_u = 0.2$.

5 Evolution of the eight column circuit

We now discuss whether there might be a reason that insect head direction circuits typically have an eight column architecture (Pisokas et al., 2020). The derivations leading to Eq. 19 are valid for other values of N, so there is no reason *a priori* to expect $N = 8$.

Recent studies in genetics (Johnston et al., 2022; Dingle et al., 2018), argue that an observed organism is more likely to have resulted from a simple than a more complex genome – evolution favours simplicity. We note that powers of two are easier to generate with replication dynamics than other numbers, because they just require each cell to divide a set number of times. Other numbers require that at some point, two cells resulting from a division must behave differently, necessitating more complex signalling mechanisms and making this possibility less likely to be observed.

As we show in Appendix III.3, not all numbers of neurons support a working circuit. the circuits for $N = 2$ and $N = 4$ are degenerate – either resulting in a single dimensional encoding, or two disconnected circuits that do not enforce the required ring topology. $N = 8$ is therefore the smallest power of two that allows for a non-degenerate circuit. This hints at the possibility that $N = 8$ is not arbitrary, but rather that it is the evolutionarily simplest circuit capable of performing heading integration.

6 Discussion

In this work we derived an optimal noise-minimising circuit that encodes heading direction, then showed that the proposed circuit matches experimental data from insects. Finally, we showed that such a circuit can be developed and maintained by a biological learning rule, and proposed a mathematical argument for the $N = 8$ column structure often found in the compass circuits of insects. In this section we discuss the implications and limitations of these contributions, and outline potential follow-up work.

Heading integration circuits in insects has been extensively studied in previous literature, with models ranging in complexity from simplified conceptual networks similar to our own (Wittmann and Schwegler, 1995) to sophisticated models agreeing with biological data and featuring multiple neuron types (Pisokas et al., 2020; Lyu et al., 2022). Previous theoretical work has argued that a sinusoidal activity encoding is such a common motif in insect navigation because it facilitates elementwise vector addition (Lyu et al., 2022; Wittmann and Schwegler, 1995). However, this cannot be the full story, because we derive a whole family of circuits that have the same property. By showing that the sinusoidal activity emerges as the theoretically most noise resilient heading integration circuit, and verifying that the corresponding circuit matches experimental data, we close this explanatory gap.

We also show that our proposed circuit can be developed by a simple Hebbian based learning rule, and speculate that it might be the simplest possible heading integration circuit (in terms of evolutionary complexity). This agrees with predictions by recent theories, that evolution should tend to favour simplicity where possible (Johnston et al., 2022).

Our work still has some unaddressed limitations, notably about the topology of the activity. The use of a circular topology to encode the heading direction of the insect is valid only in 2D environments. But our model species, the fruit fly, actually lives in a 3D environment. We argue that even if flies live in a 3D environment, the third dimension (up-down) is different from the other two. In particular, flies typically only perform long distance navigation in the other two dimensions, as if in 2D. Further studies are however necessary to examine the full effects of 3D motion.

Similarly, we could investigate circuits that integrate position, not only heading, which would require us to replace the ring topology with that of a 2D plane. This would be particularly interesting for foraging insects like bees or ants, whose ability to remember their position with respect to a nest is the subject of many experimental and computational studies. Position integrating neurons in these insects also have sinusoidal activity patterns (Haferlach et al., 2007; Vickerstaff and Paolo, 2005).

Finally, another interesting avenue for future work is to compare the encoding of direction in insects with that of mammals, which encode heading in a fundamentally different way that uses many more neurons, which are located in the CA1 region of the hippocampus. This raises a critical question – why would the circuitry and encoding be different if navigation should follow similar principles across all species? We speculate that the difference might lie in the type of navigation that the two classes use. Insects often rely on a reference system that is globally anchored to a certain point or phenomenon, whether it is their nests for ants and bees, the polarisation of sunlight for locusts, or the pattern of the milky way in dung beetles. On the other hand, mammals such as mice typically do not use global clues, but rely on local landmarks that are context-dependent and only occur in specific locations. Mammals must therefore have a flexible encoding that can be updated as different environments are explored. This would require a different set of principles than those selected here.

References

Christopher L Barnes, Daniel Bonnéry, and Albert Cardona. Synaptic counts approximate synaptic contact area in drosophila. *Plos one*, 17(4):e0266064, 2022.

Kenneth P Burnham and David R Anderson. Model selection and multimodel inference. *A practical information-theoretic approach*, 2, 2004.

T. S. Collett. Path integration: how details of the honeybee waggle dance and the foraging strategies of desert ants might help in understanding its mechanisms. *Journal of Experimental Biology*, 222 (11):jeb205187, May 2019. ISSN 1477-9145. doi: 10.1242/jeb.205187.

Charles Darwin. Origin of certain instincts. *Nature*, 7(179):417–418, April 1873. ISSN 0028-0836. doi: 10.1038/007417a0.

Peter Dayan, Laurence F Abbott, et al. Theoretical neuroscience: computational and mathematical modeling of neural systems. *Journal of Cognitive Neuroscience*, 15(1):154–155, 2003.

Kamaludin Dingle, Chico Q Camargo, and Ard A Louis. Input–output maps are strongly biased towards simple outputs. *Nature communications*, 9(1):761, 2018.

Basil el Jundi, Keram Pfeiffer, Stanley Heinze, and Uwe Homberg. Integration of polarization and chromatic cues in the insect sky compass. *Journal of Comparative Physiology A: Neuroethology, Sensory, Neural, and Behavioral Physiology*, 200(6):575–589, mar 2014. ISSN 14321351. doi: 10.1007/s00359-014-0890-6.

B. ElJundi, E. Baird, M.J. Byrne, and M. Dacke. The brain behind straight-line orientation in dung beetles. *Journal of Experimental Biology*, 222, 2019. doi: 10.1242/jeb.192450.

Thomas Haferlach, Jan Wessnitzer, Michael Mangan, and Barbara Webb. Evolving a neural model of insect path integration. *Adaptive Behavior*, 15(3):273–287, 2007. ISSN 1059-7123. doi: 10.1177/1059712307082080.

Stanley Heinze, Sascha Gotthardt, and Uwe Homberg. Transformation of polarized light information in the central complex of the locust. *The Journal of neuroscience: the official journal of the Society for Neuroscience*, 29(38):11783–93, sep 2009. ISSN 1529-2401. doi: 10.1523/JNEUROSCI.1870-09.2009.

Uwe Homberg. Sky compass orientation in desert locusts—evidence from field and laboratory studies. *Frontiers in behavioral neuroscience*, 9:346, 2015.

Uwe Homberg, Stanley Heinze, Keram Pfeiffer, Michiyo Kinoshita, and Basil El Jundi. Central neural coding of sky polarization in insects. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 366(1565):680–687, 2011. ISSN 14712970. doi: 10.1098/rstb.2010.0199.

Brad K Hulse, Hannah Haberkern, Romain Franconville, Daniel B Turner-Evans, Shin-ya Takemura, Tanya Wolff, Marcella Noorman, Marisa Dreher, Chuntao Dan, Ruchi Parekh, Ann M Hermundstad, Gerald M Rubin, and Vivek Jayaraman. A connectome of the drosophila central complex reveals network motifs suitable for flexible navigation and context-dependent action selection. *eLife*, 10, 2021. doi: 10.7554/elife.66039.

Iain G Johnston, Kamaludin Dingle, Sam F Greenbury, Chico Q Camargo, Jonathan PK Doye, Sebastian E Ahnert, and Ard A Louis. Symmetry and simplicity spontaneously emerge from the algorithmic nature of evolution. *Proceedings of the National Academy of Sciences*, 119(11):e2113883119, 2022.

Jonathan Kadmon and Haim Sompolinsky. Transition to chaos in random neuronal networks. *Physical Review X*, 5(4):041030, 2015.

Irene S Kim and Michael H Dickinson. Idiothetic path integration in the fruit fly drosophila melanogaster. *Current Biology*, 27(15):2227–2238, 2017.

Michiyo Kinoshita, Keram Pfeiffer, and Uwe Homberg. Spectral properties of identified polarized-light sensitive interneurons in the brain of the desert locust schistocerca gregaria. *The Journal of experimental biology*, 210(Pt 8):1350–61, 2007. ISSN 0022-0949. doi: 10.1242/jeb.02744.

Nikolaus Kriegeskorte and Tal Golan. Neural network models and deep learning. *Current Biology*, 29 (7):R231–R236, April 2019. ISSN 09609822. doi: 10.1016/j.cub.2019.02.034.

Thomas Labhart. Polarization-opponent interneurons in the insect visual system. *Nature*, 331(6155): 435–437, 1988. ISSN 0028-0836. doi: 10.1038/331435a0.

Thomas Labhart. Polarization-sensitive interneurons in the optic lobe of the desert ant cataglyphis bicolor. *Naturwissenschaften*, 87(3):133–136, 2000. ISSN 0028-1042. doi: 10.1007/s001140050691.

Tony X. Liu, Pasha A. Davoudian, Kristyn M. Lizbinski, and James M. Jeanne. Connectomic features underlying diverse synaptic connection strengths and subcellular computation. *Current Biology*, 32 (3):559–569.e5, February 2022. ISSN 0960-9822. doi: 10.1016/j.cub.2021.11.056.

Rudi Loesel and Uwe Homberg. Anatomy and physiology of neurons with processes in the accessory medulla of the cockroach leucophaea maderae. *Journal of Comparative Neurology*, 439(2):193–207, 2001. ISSN 0021-9967. doi: 10.1002/cne.1342.

Jenny Lu, Amir H. Behbahani, Lydia Hamburg, Elena A. Westeinde, Paul M. Dawson, Cheng Lyu, Gaby Maimon, Michael H. Dickinson, Shaul Druckmann, and Rachel I. Wilson. Transforming representations of movement from body- to world-centric space. *Nature*, 601(7891):98–104, 2022. ISSN 0028-0836. doi: 10.1038/s41586-021-04191-x.

Cheng Lyu, L. F. Abbott, and Gaby Maimon. Building an allocentric travelling direction signal via vector computation. *Nature*, 601(7891):92–97, 2022. ISSN 0028-0836. doi: 10.1038/s41586-021-04067-0.

R. Menzel and U. Muller. Learning and memory in honeybees: From behavior to neural substrates. *Annual Review of Neuroscience*, 19:379–404, 1996. ISSN 0147-006X. doi: 10.1146/annurev.ne.19.030196.002115.

Horst Mittelstaedt. Analytical cybernetics of spider navigation. In Friedrich G. Barth, editor, *Neurobiology of Arachnids*, pages 298–316. Springer Berlin Heidelberg, 1985. ISBN 978-3-642-70348-5. doi: 10.1007/978-3-642-70348-5_15.

M. Müller and Rüdiger Wehner. Path integration in desert ants, *Cataglyphis fortis*. *Proceedings of the National Academy of Sciences*, 85(14):5287–5290, 1988. doi: 10.1073/pnas.85.14.5287.

Keram Pfeiffer. The neuronal building blocks of the navigational toolkit in the central complex of insects. *Current Opinion in Insect Science*, page 100972, 2022.

Keram Pfeiffer and Uwe Homberg. Coding of azimuthal directions via time-compensated combination of celestial compass cues. *Current Biology*, 17(11):960–965, jun 2007. ISSN 0960-9822. doi: 10.1016/J.CUB.2007.04.059.

Keram Pfeiffer, Michiyo Kinoshita, and Uwe Homberg. Polarization-sensitive and light-sensitive neurons in two parallel pathways passing through the anterior optic tubercle in the locust brain. *Journal of neurophysiology*, 94(6):3903–15, 2005. ISSN 0022-3077. doi: 10.1152/jn.00276.2005.

Ioannis Pisokas, Stanley Heinze, and Barbara Webb. The head direction circuit of two insect species. *eLife*, 9:e53985, July 2020. ISSN 2050-084X. doi: 10.7554/eLife.53985.

Midori Sakura, Dimitrios Lambrinos, and Thomas Labhart. Polarized skylight navigation in insects: model and electrophysiology of e-vector coding by neurons in the central complex. *Journal of neurophysiology*, 99(2):667–82, 2008. ISSN 0022-3077. doi: 10.1152/jn.00784.2007.

Marcel Ethan Sayre, Rachel Templin, Johanna Chavez, Julian Kempenaers, and Stanley Heinze. A projectome of the bumblebee central complex. *eLife*, 10:e68911, 2021. doi: 10.7554/elife.68911.

Thomas Stone, Barbara Webb, Andrea Adden, Nicolai Ben Weddig, Anna Honkanen, Rachel Templin, William Wcislo, Luca Scimeca, Eric Warrant, and Stanley Heinze. An anatomically constrained model for path integration in the bee brain. *Current Biology*, 27(20):3069–3085, 2017. ISSN 0960-9822. doi: 10.1016/j.cub.2017.08.052.

Daniel Turner-Evans, Stephanie Wegener, Hervé Rouault, Romain Franconville, Tanya Wolff, Johannes D Seelig, Shaul Druckmann, and Vivek Jayaraman. Angular velocity integration in a fly heading circuit. *eLife*, 6:2112–2126, May 2017. ISSN 2050-084X. doi: 10.7554/elife.23496.

R. J. Vickerstaff and E. A. Di Paolo. Evolving neural models of path integration. *Journal of Experimental Biology*, 208(17):3349–3366, 2005. doi: 10.1242/jeb.01772.

Karl von Frisch. *The dance language and orientation of bees*. Harvard University Press, 1967. ISBN 9780674418776.

Jan Wessnitzer, Thomas Haferlach, Michael Mangan, and Barbara Webb. Path integration using a model of e-vector orientation coding in the insect brain: Reply to vickerstaff and di paolo. *Adaptive Behavior*, 16(4):277–280, 2008. doi: 10.1177/1059712308094947.

Thomas Wittmann and Helmut Schwegler. Path integration - a network model. *Biological Cybernetics*, 73(6):569–575, 1995. ISSN 3401200. doi: 10.1007/bf00199549.

Frederick Zittrell, Kathrin Pabst, Elena Carlomagno, Ronny Rosner, Uta Pegel, Dominik M. Endres, and Uwe Homberg. Integration of optic flow into the sky compass network in the brain of the desert locust. *Frontiers in Neural Circuits*, 17, 2023. ISSN 1662-5110. doi: 10.3389/fncir.2023.1111310.

Code availability

We make the code used for all analysis and figure generation available at https://github.com/Dominic-DallOsto/insect-navigation-sinusoidal-optimality.

Acknowledgements

We would like to thank Benjamin F. Grewe for his support and helpful comments.

I Path integration dynamics for heading and position

Here we make more precise statements about the dynamics of the circuit. The topology of the activity is a circular line attractor, so that any perturbation falls back into a circle and the position of the activity around the circle represents an angle. In our circuit with the dynamics from Eq. 1, this circular line attractor is achieved by setting

$$\begin{aligned} \|\phi'[\mathbf{Wa}]\| &= 1 \quad \forall \|\mathbf{a}\| = r \\ \|\phi'[\mathbf{Wa}]\| &< 1 \quad \forall \|\mathbf{a}\| > r \\ \|\phi'[\mathbf{Wa}]\| &> 1 \quad \forall \|\mathbf{a}\| < r. \end{aligned} \tag{24}$$

where r is the radius of the attractor. Effectively we can write the original dynamics around the line attractor directly in the spatial Fourier basis,

$$\dot{\mathcal{F}_f[\mathbf{a}]} = -\mathcal{F}_f[\mathbf{a}] + (1 + \alpha(r - \|\mathbf{a}\|))\mathcal{F}_f[\omega](\mathcal{F}_f[\mathbf{a}] + \delta\mathcal{F}_f[\mathbf{a}]), \tag{25}$$

where $\alpha(r - \|\mathbf{a}\|)$ is some non-linear function.

As the activity always falls back to the circular line attractor, the heading integration is linear along the circle. This implies that any small movement of the animal is projected onto the circle then linearly integrated. Notice that the linearity in the circular line is not a computational assumption, but rather it emerges from the principle of symmetry.

In more intuitive terms, the neurons have a saturating nonlinearity where they reduce their gain as they fire more, implying that any deviation from the circular line attractor will return back to it. As the activity of the network goes beyond that of the line attractor, the gain reduces and the activity decreases, and when the activity of the network is smaller than that of the line attractor the gain is larger and the activity increases.

II Ambiguities in multiple harmonic decoding with drift

We consider the case where multiple harmonics are used, and their phases have drifted. For example, we consider the harmonics f_1 and f_2. The activity is then

$$\mathbf{a}_n(\theta) = \cos\left(\frac{2\pi f_1 n}{N} + f_1\theta\right) + \cos\left(\frac{2\pi f_2 n + f_2\theta}{N}\right). \tag{26}$$

If the phase of the second harmonic drifts by $\delta\theta$,

$$\mathbf{a}_n^{\delta\theta}(\theta) = cos\left(\frac{2\pi f_1 n}{N} + f_1\theta\right) + \cos\left(\frac{2\pi f_2 n + f_2\theta + \delta\theta}{N}\right). \tag{27}$$

We can calculate the alignment between the activity with drift and the activity without as a dot product of the activity vectors,

$$\frac{\langle \mathbf{a}(\theta), \mathbf{a}^{\delta\theta}(\theta)\rangle}{\|\mathbf{a}(\theta)\|\|\mathbf{a}^{\delta\theta}(\theta)\|} = \frac{1}{\|\mathbf{a}(\theta)\|^2}\sum_n \mathbf{a}_n(\theta)\mathbf{a}_n^{\delta\theta}(\theta) = 1 + \cos(\delta\theta). \tag{28}$$

However, there are other angles where the alignment is better. For example, we can consider an estimate angle $\hat{\theta}$ with its corresponding activity $\mathbf{a}(\hat{\theta})$ gives

$$\frac{\langle \mathbf{a}(\hat{\theta}), \mathbf{a}^{\delta\theta}(\theta)\rangle}{\|\mathbf{a}(\theta)\|\|\mathbf{a}^{\delta\theta}(\theta)\|} = \|\mathbf{a}(\theta)\|^2\sum_n \mathbf{a}_n(\theta + \delta\theta)\mathbf{a}_n^{\delta\theta}(\theta) = \cos(\theta - \hat{\theta}) + \cos(\theta - \hat{\theta} + \delta\theta). \tag{29}$$

For small $\delta\theta$, this is maximised when $\hat{\theta} = \theta + \delta\theta/2$. However, since θ is circular if the drift is π, there are two possible positions given by $\hat{\theta} = \theta \pm \pi/2$.

This implies that we have two possible positions that could be decoded.

III Equivalent circuits and degeneracies

III.1 Equivalence under permutation

The activity of neurons given by Eq. 5 implies that the preferred angle of neuron n is given by

$$\varphi_n = 2\pi \frac{fn \mod N}{N} \tag{30}$$

where in our case $N = 8$. Table S.1 shows $nf \mod (N)$ evaluated for all neurons in the network using different harmonics. We notice that for $f = \{1, 3, 5, 7\}$ all the numbers from zero to seven appear, while for $f = \{2, 6\}$ we only get the even numbers, for $f = 4$ we get only zero and four, and for $f = 8$ there is only zero.

The explanation is based on number theory. If f and N have their greatest common divisor $\gcd(N, f) = d$, then $nf \mod N = 0$ for $n = N/d$. This implies that the preferred angle of neuron $n = 0$ is the same as that of $n = N/d$. When N, f are co-prime $d = 1$, n goes from 0 to $N - 1$ without repeating any value. However, when $d > 1$, the neuron $n = N/d$ has the same tuning as the neuron $n = 0$, the neuron $n = N/d + 1$ has the same tuning as the neuron $n = 1$ and so on. In other words, the angular tuning of the neurons has a period of $\gcd(N, f)$.

This divides the possible circuits into the following groups:

- $f = \{1, 3, 5, 7\}$ which contain $N = 8$ angular tunings with values $[0, \pi/4, \pi/2, 3\pi/4, \pi, 5\pi/4, 3\pi/2, 7\pi/4]$ because every odd number is coprime with 8 and thus the $\gcd(N, f) = 1$.
- $f = \{2, 6\}$ which cycle through 4 possible directions $[0, \pi/2, \pi, 3\pi/2]$ because $\frac{N}{\gcd(N,f)} = 4$
- $f = 4$ which can only represent 2 possible directions $[0, \pi]$.
- $f = 8$ which can only represent the direction 0

	$n = 0$	$n = 1$	$n = 2$	$n = 3$	$n = 4$	$n = 5$	$n = 6$	$n = 7$	$\gcd(8, f)$
$f = 1$	0	1	2	3	4	5	6	7	1
$f = 2$	0	2	4	6	0	2	4	6	2
$f = 3$	0	3	6	1	4	7	2	5	1
$f = 4$	0	4	0	4	0	4	0	4	4
$f = 5$	0	5	2	7	4	1	6	3	1
$f = 6$	0	6	4	2	0	6	4	2	2
$f = 7$	0	7	6	5	4	3	2	1	1
$f = 8$	0	0	0	0	0	0	0	0	8

Table S.1: Distribution of neural activity phases for different harmonics: We computed $fn \mod 8$ for $n = \{1, 2, 3, 4, 5, 6, 7\}$ and $f = \{1, 2, 3, 4, 5, 6, 7, 8\}$ and found that all possible phases appear for any odd number. This happens because the $\mod 8$ operation imposes an abelian group structure, namely $\mathbb{Z}/8$ and any coprime with 8 will be a generator of the whole group. If we use instead $f = 2$ or $f = 6$ we have that we can divide the $N = 8$ and f by 2 and we get the abelian group $\mathbb{Z}/4$ which has four elements. The same argument applies to $f = 4$, leaving only two elements, and for $f = 8$ we get a single element.

III.2 Degenerate circuits

Given the groupings presented in the previous subsection, we notice that not all of them allow the encoding of a full angle. Notably, $f = 8$ the neurons only encode one angle, and for $f = 4$ only two complementary angles are encoded $0, \pi$. This implies that we cannot represent a circle, because we need two different dimensions to do so, but in both these cases can only encode one.

For $f = 2, 6$ we obtain two groups of neurons that cover four angles, thus it is possible to cover a circle. However, the circuit is degenerate, as shown in Fig. 3. In this case there are no connections between even and odd neurons. Thus, we have two groups of neurons that are disconnected, and each group encodes only one direction (either north-south or east-west). While the encoding could work in principle, the circuit is decoupled, meaning that there is nothing in this circuit that prevents the activity from having activities outside the circular topology: as the two circuits are disconnected, a given value of activity in one group does not restrict the activity in the other.

III.3 Circuits with different neuron counts

We evaluate the viability of circuits with other values of N. From findings in the the previous appendix, we note that choosing N to be a prime number implies that all frequencies will be coprime with N, and thus that all the neurons will have a different tuning.

For $N = 2$, the two angles are $0, \pi$, in whichthe circuit cwn only encode one dimension and thus cannot encode a circle, just as we had for $N = 8, f = 4$ in the subsection III.2.

For $N = 4$, if $f = 2$ the circuit can encode only the angles $0, \pi$ and we have the same case as for $N = 2$. But if $f = 1$ of $f = 3$, the neurons are tuned to the angles $0, \pi/2, \pi, 3\pi/2$ which can encode a circle. However, by looking at the connectivity matrix from Eq. 18, we find that neurons $1, 3$ are connected to one another, but not to $2, 4$, just as we had for $N = 8, f = 4$ in the subsection III.2. This implies that the circuit does not enforce a circular topology, and thus it does not work.

Notice that we could think of another structure where each neuron encodes a different direction. Intuitively, we would have a north-south neuron and an east-west neuron, as we would have in cartesian coordinates. This would have the same problem as $N = 4$, where the two neurons would not interact and thus the topology would be wrong. Additionally, when the insect heads north, the north-south neuron would be very active, while when it heads south it would be very inactive. This means that the circuit as a whole would have a different firing rate depending on the direction, breaking the symmetry assumption.

IV Path counting

We counted the number of different paths between EPG neurons, accounting for the sign of the connections (whether the path passed through an inhibitory $\Delta 7$ neuron), and used the net path count as a proxy for connectivity strength. The results of this analysis are shown in Section 3, but we will detail an example using the locust network from Pisokas et al. (2020), shown in Fig. S.1.

We consider the shortest excitatory and inhibitory pathways between EPG neurons, which have lengths of 2 and 3 respectively. There are no direct paths between EPG neurons. In this case there are two paths of length 2 that implement self-excitatory connections for EPG_1:

- $\mathrm{EPG}_1 \to \mathrm{PEG}_1 \to \mathrm{EPG}_1$
- $\mathrm{EPG}_1 \to \mathrm{PEN}_1 \to \mathrm{EPG}_1$

And there is one connection of path length 2 that connects EPG_1 to each of its nearest neighbours:

- $\mathrm{EPG}_1 \to \mathrm{PEN}_1 \to \mathrm{EPG}_2$
- $\mathrm{EPG}_1 \to \mathrm{PEN}_1 \to \mathrm{EPG}_8$

For inhibitory connections there are four paths of length 3 that connect EPG_1 to the neuron on the opposite side of the ring, EPG_5. These are:

- $\mathrm{EPG}_1 \to \Delta 7_5 \to \mathrm{PEN}_5 \to \mathrm{EPG}_5$
- $\mathrm{EPG}_1 \to \Delta 7_5 \to \mathrm{PEG}_5 \to \mathrm{EPG}_5$
- $\mathrm{EPG}_1 \to \Delta 7_4 \to \mathrm{PEN}_4 \to \mathrm{EPG}_5$
- $\mathrm{EPG}_1 \to \Delta 7_6 \to \mathrm{PEN}_6 \to \mathrm{EPG}_5$

There are three connections of path length 3 connecting EPG_1 to EPG_4:

- $\mathrm{EPG}_1 \to \Delta 7_4 \to \mathrm{PEN}_4 \to \mathrm{EPG}_4$
- $\mathrm{EPG}_1 \to \Delta 7_4 \to \mathrm{PEG}_4 \to \mathrm{EPG}_4$
- $\mathrm{EPG}_1 \to \Delta 7_5 \to \mathrm{PEN}_5 \to \mathrm{EPG}_4$

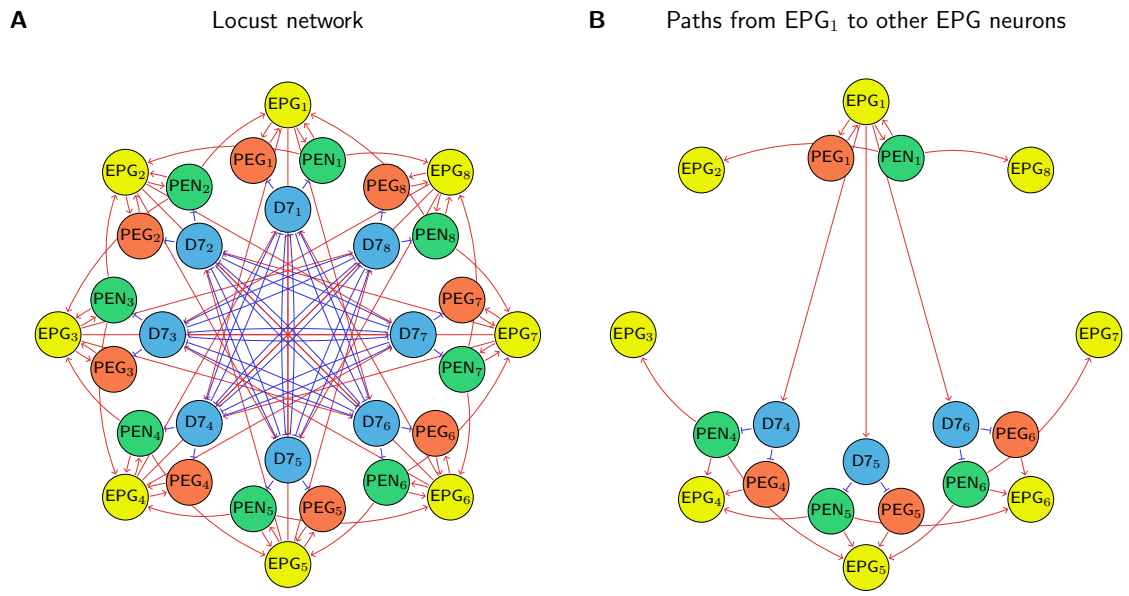
Finally, there is one path of length 3 connecting EPG_1 to the neuron perpendicular to it around the ring, EPG_3:

- $\text{EPG}_1 \rightarrow \Delta 7_4 \rightarrow \text{PEN}_4 \rightarrow \text{EPG}_3$

We then add these connections together to give the net path count profile for the network. Because the network is rotationally symmetric, the connections from EPG_1 are the same as the connections from EPG_2 but with the neurons indices incremented by 1. The generalised profile is shown in Table S.2.

$\text{EPG}_n \rightarrow$	$n-4$	$n-3$	$n-2$	$n-1$	n	$n+1$	$n+2$	$n+3$
Excitatory path count	0	0	0	1	2	1	0	0
Inhibitory path count	4	3	1	0	0	0	1	3
Net path count	-4	-3	-1	1	2	1	-1	-3

Table S.2: Net path count profile between EPG neurons in the locust circuit.



Supplementary Figure S.1: The full locust connectivity network from Pisokas et al. (2020) (left), and the connections with path lengths 2 and 3 from EPG_1 to all other EPG neurons in the network (right). Because the network is rotationally symmetric these path counts generalise to all EPG neurons as shown in Table S.2.

V Data preprocessing

The following section details our data preprocessing to produce a central complex circuit for the fruit fly similar to that in Pisokas et al. (2020) using synaptic counts from the fruit fly connectome data set (Hulse et al., 2021). Full details are shown in our available code.

We first identified the 6 neuron types which corresponded to the neurons of interest in the central complex (Table S.3).

Neuron class	Number
PEN_a(PEN1)	20 neurons
PEN_b(PEN2)	22 neurons
PEG	18 neurons
EPG	46 neurons
EPGt	4 neurons
Delta7	42 neurons

Table S.3: Neuron types and their numbers in the fruit fly connectome data set (Hulse et al., 2021).

We grouped the two PEN and two EPG populations together for further analysis.

From the connectome data we created a connectivity matrix (Fig. S.2) containing the number of synapses between all pairs of neurons of the above mentioned types. This contained 129473 synapses between 152 neurons in total.

Next, neurons in the same glomerulus were grouped together. The EPG neurons were divided between 18 glomeruli, L1-L9 and R1-R9. The PEN neurons were also divided between 16 glomeruli, L2-L9 and R2-R9. The PEG neurons were divided into 18 glomeruli with only 1 neuron in each, L1-L9 and R1-R9. The Delta7 neurons had 10 unique sub-types (Table S.4). which were grouped into 8 glomeruli based on their left glomerulus index – i.e. L4R5_R and L4R6_R were grouped in L4.

Delta7 neuron sub-types	Glomerulus group
Delta7(PB15)_L1L9R8_R	Delta7 L1
Delta7(PB15)_L2R7_R	Delta7 L2
Delta7(PB15)_L3R6_R	Delta7 L3
Delta7(PB15)_L4R5_R	Delta7 L4
Delta7(PB15)_L4R6_R	Delta7 L4
Delta7(PB15)_L5R4_L	Delta7 L5
Delta7(PB15)_L6R3_L	Delta7 L6
Delta7(PB15)_L6R4_L	Delta7 L6
Delta7(PB15)_L7R2_L	Delta7 L7
Delta7(PB15)_L7R3_L	Delta7 L7
Delta7(PB15)_L8R1R9_L	Delta7 L8

Table S.4: List of the Delta7 neuron sub-types in the fruit fly connectome data set (Hulse et al., 2021). These differ in the glomeruli their pre- and post-synaptic terminals innervate. We are referring to those altogether as Delta7 neurons in the present account.

After this grouping we had a connectivity matrix of 60 neurons in total (Fig. S.3).

For the EPG, PEN, and PEG neurons we then grouped glomeruli mirrored in the two hemispheres together – i.e. L1 and R1 were grouped into glomerulus 1. This resulted in a connectivity matrix of 34 neurons (Fig. S.4).

Finally, as in (Pisokas et al., 2020) we grouped the PEG9 and PEG1 neurons and EPG9 and EPG1 neurons together. The former because they both output to the same ellipsoid body segment, and the latter because they both receive common input. This resulted in a connectivity matrix with 32 neurons (Fig. S.5) which we used for the analysis in Fig. 4 as this network had the 8-fold rotational symmetry compatible with our theoretical model.

VI Fitting weights

Having calculated the paths or synaptic strengths between EPG neurons in the network, we next computed the average connectivity profile for the network – how each EPG neuron connected to its neighbours a certain distance around the ring. We then compared this connectivity profile to the closest fitting sinusoid to quantify to what degree our prediction of sinusoidal weights was consistent with biological data.

We use the vector

$$\hat{\omega}_d = \beta \cos\left(\frac{2\pi d}{N}\right) + \gamma, \tag{31}$$

where $d = ((m-n)+4 \mod 8) - 4 \in [-4..3]$ is the signed circular distance from neuron m to neuron n.

In the networks from Pisokas et al. (2020) the weights are rotationally symmetric, so we minimise the mean squared error between the sinusoidal profile and the biological connectivity profile,

$$\beta^*, \gamma^* = \arg\min_{\beta,\gamma} \frac{1}{N} \sum_{d=-N/2}^{N/2-1} \left(\hat{\omega}_d(\beta,\gamma) - \omega_d\right)^2, \tag{32}$$

by using least squares. The network derived from the synaptic count data in Hulse et al. (2021) is not rotationally symmetric so each value in the average connectivity profile, ω_d, has a corresponding standard deviation, σ_d. In this case we minimise the precision-weighted mean squared error, which emphasises fitting connectivity profile values that are more consistently seen in the network:

$$\beta^*, \gamma^* = \arg\min_{\beta,\gamma} \frac{1}{N} \sum_{d=-N/2}^{N/2-1} \left(\frac{\hat{\omega}_d(\beta,\gamma) - \omega_d}{\sigma_d}\right)^2. \tag{33}$$

The result of fitting sinusoidal profiles to the data is shown in Section 3. Since the fit for the fruit fly is not as clear as for the locust, we also compared the quality of the fit with harmonics $f = 2$ and a combination of $f = 1$ and $f = 2$, which gives us worse results as shown in Fig. S.6 and Table S.5.

For completeness we also fit the observed weights in the fruit fly by a Gaussian curve. Since Gaussians also have a width, this requires an extra parameter to fit, which yields

$$\beta_g e^{-\frac{n^2}{\sigma_g}} + \gamma \tag{34}$$

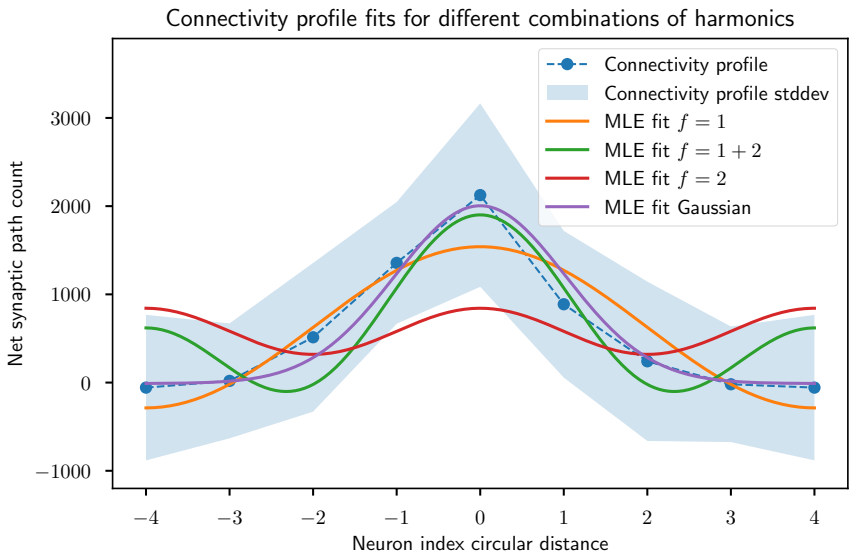
where $\beta_g, \sigma_g, \gamma_g$ are parameters to fit. We find a good agreement with the weights in Fig. S.6, and a lower root mean square error (RMSE) than with any combination of harmonics. However, we note that our sinusoidal model has two parameters instead of three, so to make a fair comparison we use the Akaike Information Criterion (AIC) Burnham and Anderson (2004), which is given by

$$\text{AIC} = 2p - 2\ln(L) = 2p - 16\text{MSE}, \tag{35}$$

where p is the number of parameters and L is the log-likelihood, which in this case is the sum of squared errors, or 8MSE with $N = 8$. Note that it is common to add a correction to the AIC when the number of samples is small (Burnham and Anderson, 2004), which is given by

$$\text{AICc} = \text{AIC} + \frac{2p^2 + 2p}{N - p - 1}, \tag{36}$$

where N is the number of samples (equivalent to the number of neurons). We find that the lowest AIC and AICc (corresponding to the best model) corresponds to the harmonic $f = 1$.



Supplementary Figure S.6: The weight profile with sinusoidal weights of frequency 1 provides the best fit to the experimental data when also accounting for the number of model parameters, compared to the second frequency, the combination of the first two frequencies, or a Gaussian profile. Note that the -4 and 4 neuron indices are the same and just duplicated for visualisation purposes.

Curve	RMSE	AIC	AICc
$f = 1$	0.3213	5.65	8.15
$f = 1, 2$	0.4409	7.11	9.61
$f = 2$	0.8478	15.50	18
Gaussian	0.1941	6.60	12.60

Table S.5: Quality of fit for different possible weight structures. The Root Mean Square Error is weighted by the variance attached to each connection (in units of standard deviations). The Akaike Information Criterion (AIC) and its sample-size correction (AICc) for the weight profiles in Fig. S.6. The best fit after correcting for model size is $f = 1$.

VII Convergence of Oja's rule with multiple harmonics

Integrating Oja's learning rule updates (Eq. 21) over all angles in the position space gives us the following:

$$\Delta\omega = \mathbf{a} * \mathbf{a} - \|\mathbf{a}\|^2\omega, \tag{37}$$

which we can transform into the Fourier domain:

$$\mathcal{F}_f[\Delta\omega] = \mathcal{F}_f[\mathbf{a}]^2 - \|\mathbf{a}\|^2\mathcal{F}_f[\omega]. \tag{38}$$

The steady state solution when the weight update is 0 is:

$$\mathcal{F}_f[\omega] = \frac{\mathcal{F}_f[\mathbf{a}]^2}{\|\mathbf{a}\|^2}. \tag{39}$$

By Parseval's theorem, $\|\mathbf{a}\|^2 = \sum_f \mathcal{F}_f[\mathbf{a}]^2$ so the Fourier spectrum of the steady state weights is just the normalised Fourier spectrum of the input:

$$\mathcal{F}_f[\omega] = \frac{\mathcal{F}_f[\mathbf{a}]^2}{\sum_f \mathcal{F}_f[\mathbf{a}]^2}. \tag{40}$$

From this we can see that the stable L_1 norm of ω is 1,

$$\sum_f |\mathcal{F}_f[\omega]| = \frac{1}{\sum_f \mathcal{F}_f[\mathbf{a}]^2} \sum_f \mathcal{F}_f[\mathbf{a}]^2 = 1. \tag{41}$$

If we combine this result with Eq. 12, we find that in the case of a single encoding frequency, $\mathcal{F}_{f^*}[\mathbf{a}] > 0$, Oja's rule will result in a single stable harmonic with $\mathcal{F}_{f^*}[\omega] = 1$ because $\sum_f \mathcal{F}_f[\mathbf{a}]^2 = \mathcal{F}_{f^*}[\mathbf{a}]^2$.

If Oja's rule has a normalising factor added to account for the number of encoding harmonics used, $|\mathbf{F}|$,

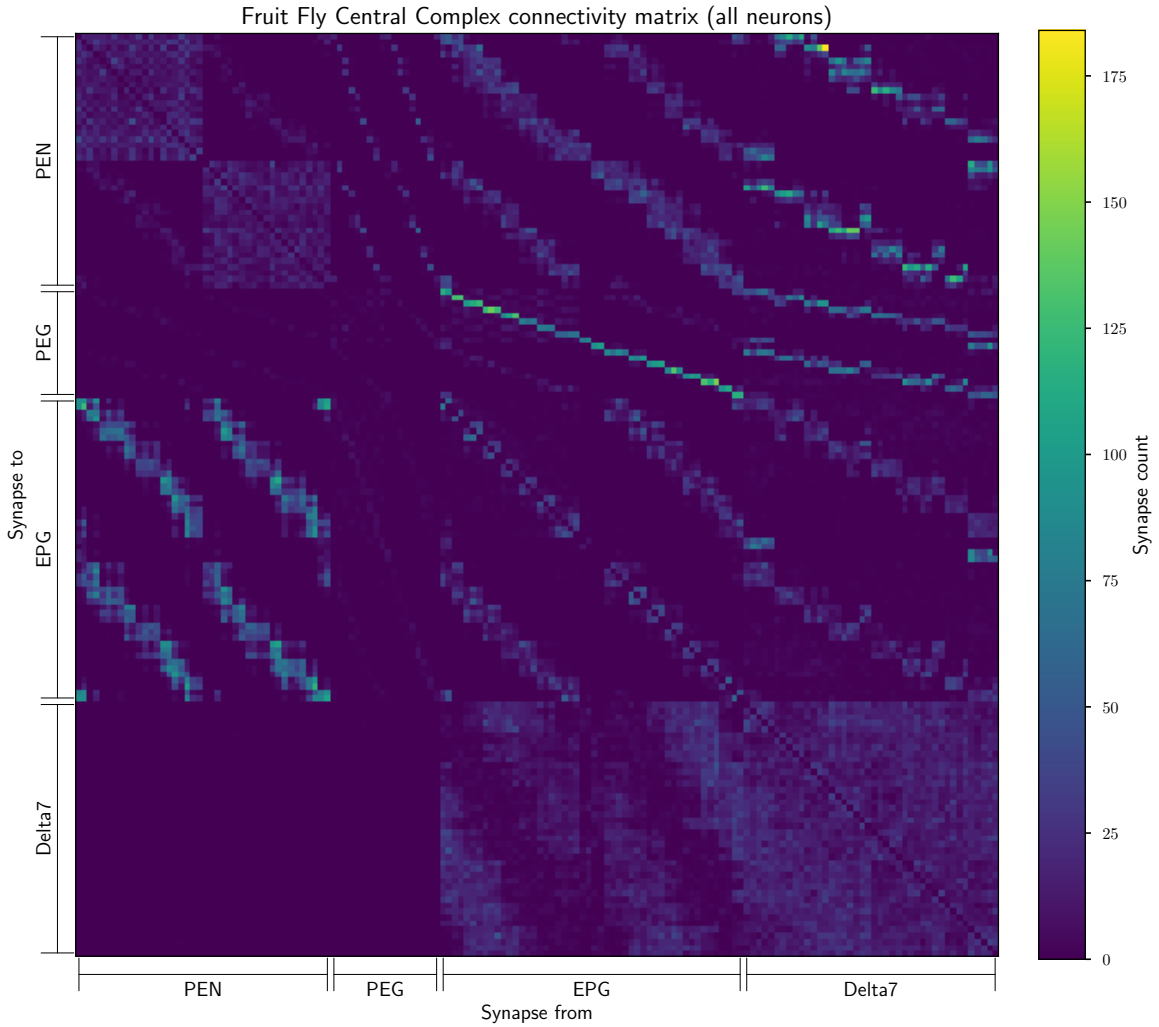
$$\Delta\omega = \mathbf{a} * \mathbf{a} - \frac{1}{|\mathbf{F}|}\|\mathbf{a}\|^2\omega, \tag{42}$$

then the steady state solution for the weights becomes

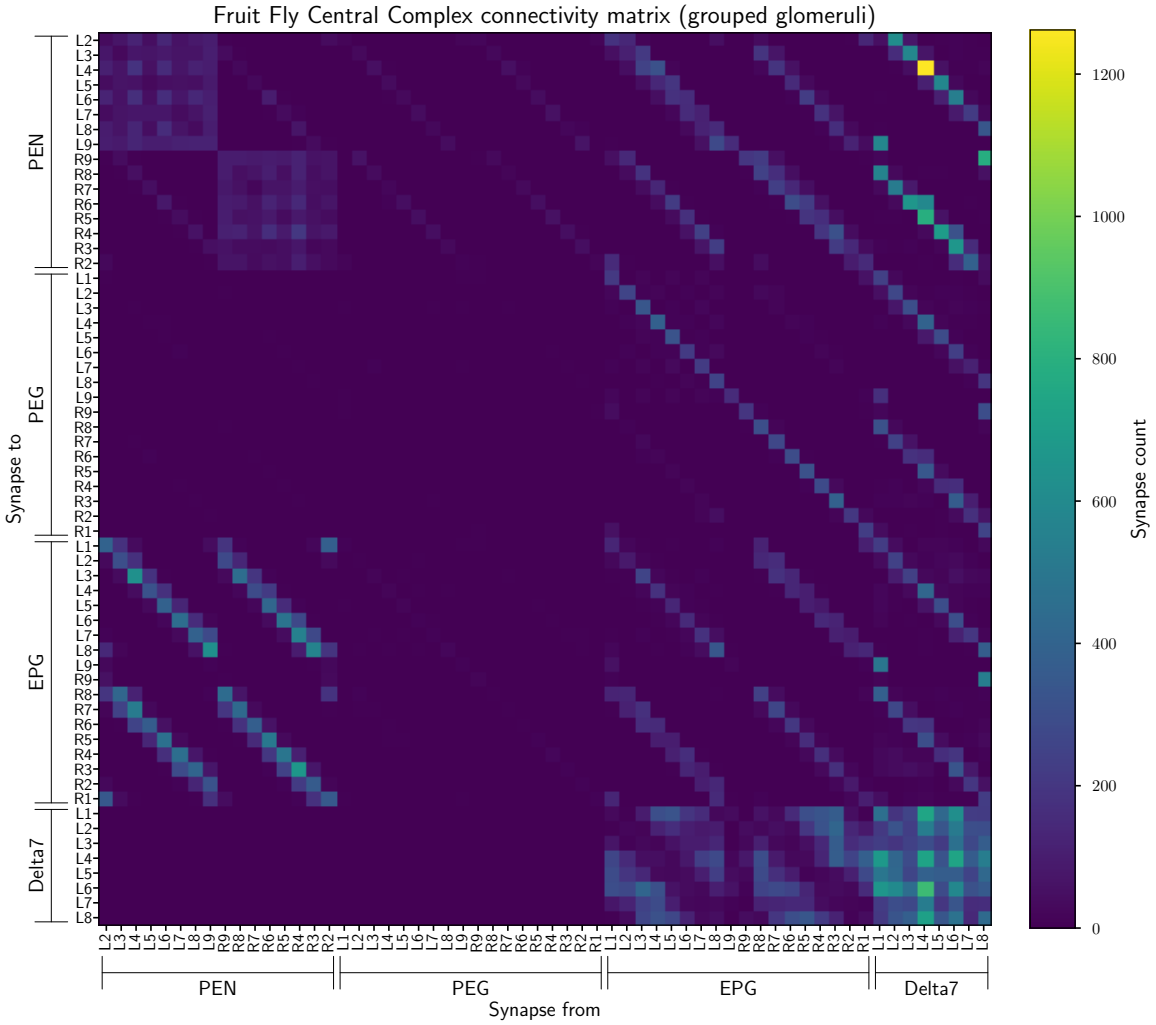
$$\mathcal{F}_f[\omega] = |\mathbf{F}|\frac{\mathcal{F}_f[\mathbf{a}]^2}{\sum_f \mathcal{F}_f[\mathbf{a}]^2}, \tag{43}$$

where the L_1 norm of $\omega = |\mathbf{F}|$. In this case multiple harmonics could develop stable values of $\mathcal{F}_f[\omega] = 1$, but only if the activity magnitudes for these harmonics are identical. If any perturbation affects the activities, the harmonic which happens to have the larger activity will begin to dominate the other by the dynamics of Eq. 43, until only one remains.

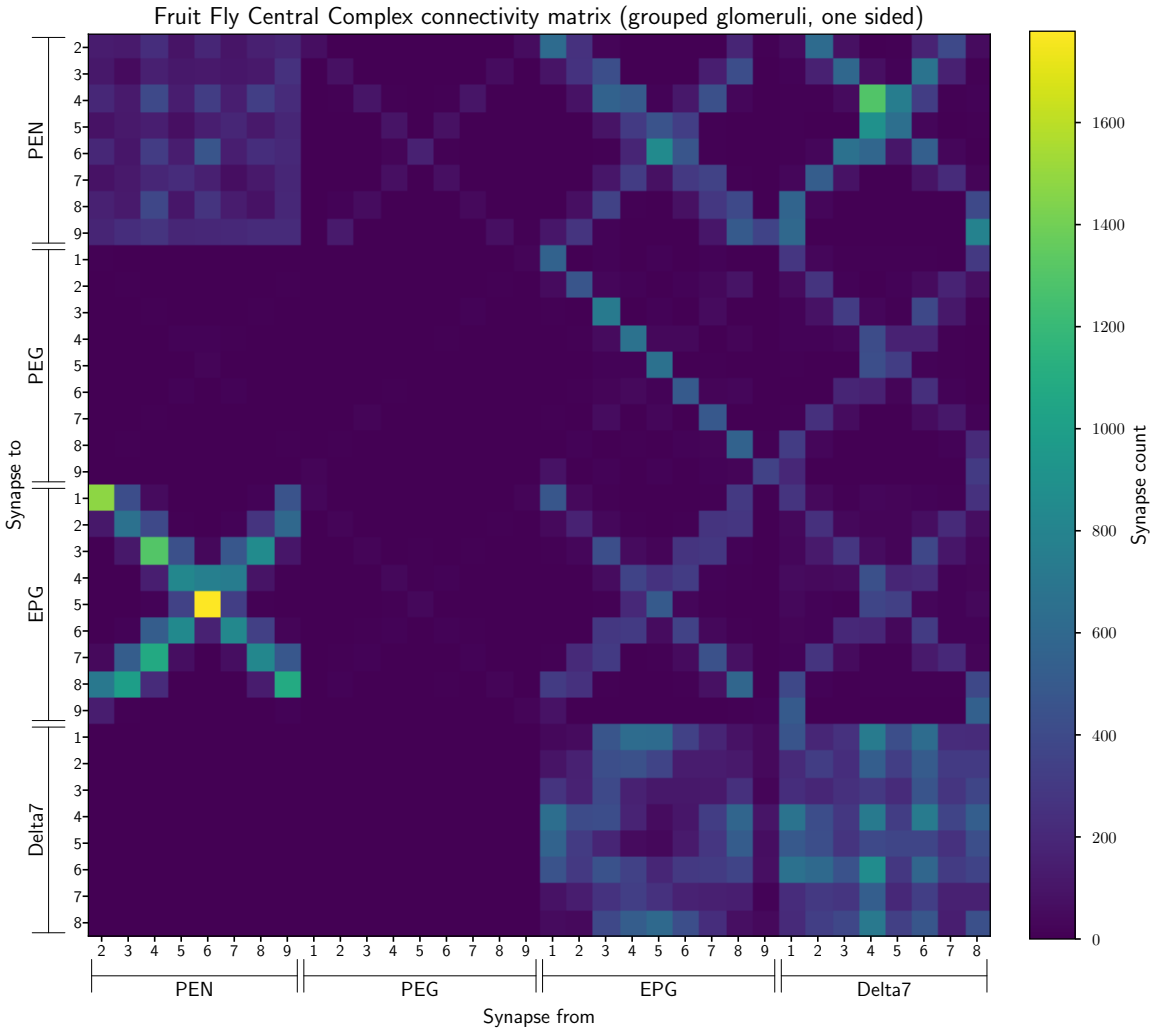
Therefore, the only case where Oja's rule results in a stable solution robust to perturbations is when only one harmonic is used.



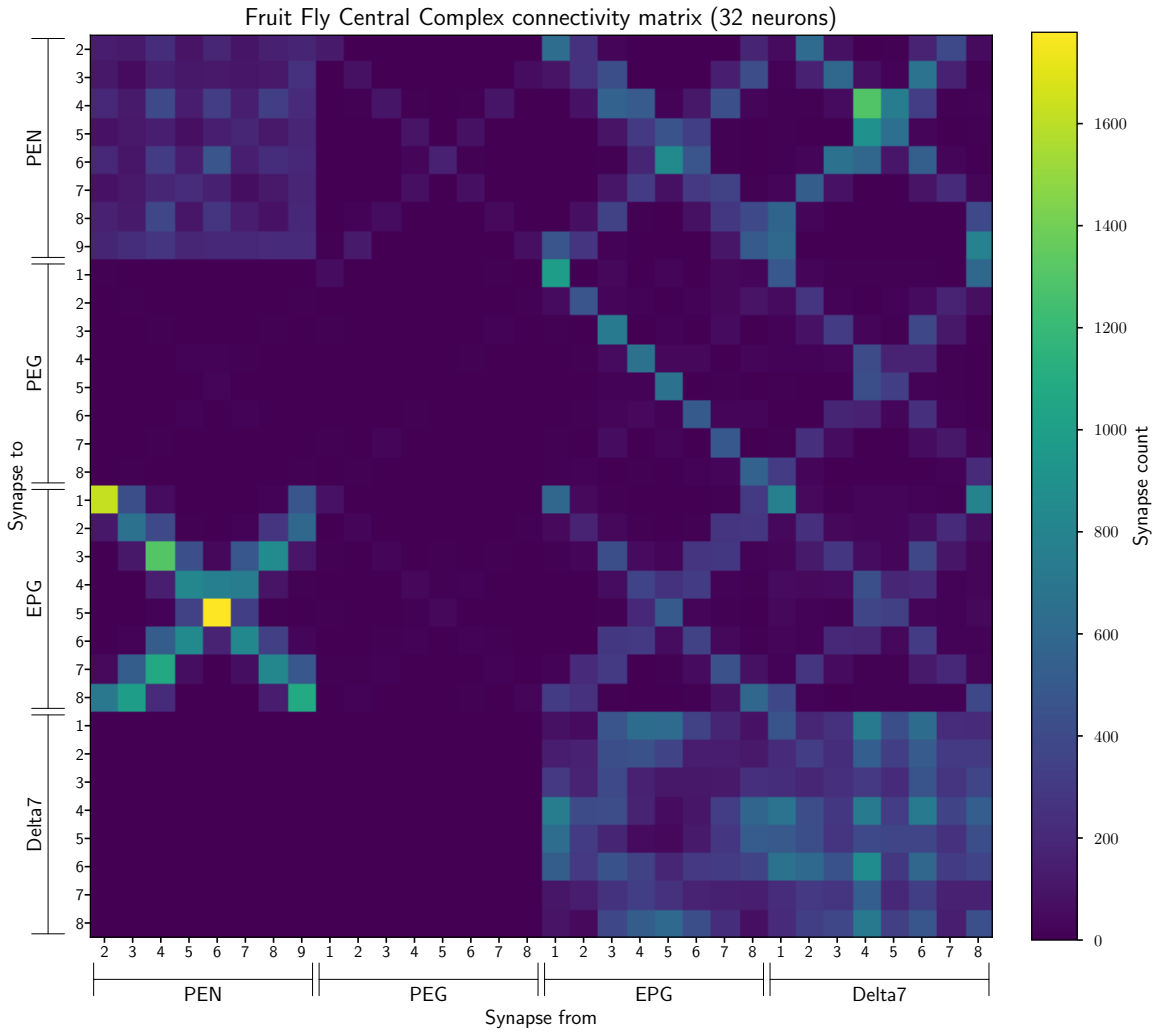
Supplementary Figure S.2: The connectivity matrix for the fruit fly containing all 152 neurons. On the horizontal axis are the names of the pre-synaptic neurons while on the vertical axis the names of the post-synaptic neurons. Neurons within each cell type are ordered by the glomerulus they innervate and arbitrarily within glomerulus.



Supplementary Figure S.3: The connectivity matrix for the fruit fly neurons grouped by innervated glomerulus so containing 60 groups. On the horizontal axis are the names of the pre-synaptic neurons while on the vertical axis the names of the post-synaptic neurons. Neuron groups are ordered by the glomerulus they innervate.



Supplementary Figure S.4: The connectivity matrix for the fruit fly neurons grouped by glomerulus and aggregated from both hemispheres, so containing 34 groups. On the horizontal axis are the names of the pre-synaptic neuron groups while on the vertical axis are the names of the post-synaptic neuron groups. Neuron groups are ordered by the glomerulus they innervate.



Supplementary Figure S.5: The connectivity matrix for the fruit fly neurons grouped by glomerulus and aggregated from both hemispheres, with PEG9 & PEG1 neurons and EPG9 & EGP1 neurons grouped, so containing 32 groups in total. On the horizontal axis are the names of the pre-synaptic neuron groups while on the vertical axis the names of the post-synaptic neuron groups. Neuron groups are ordered by the glomerulus they innervate.