Below are descriptions of the data files used in the study "Doering et. al. *Noise resistant synchronization and collective rhythm switching in a model of animal group locomotion*":

1. **refractive.csv** – File containing the times (converted to minutes relative to when a focal ant became inactive) of every physical stimulation received by the focal ants and whether a stimulation resulted in a focal ant becoming active (activate = 1). The column *interaction_no*. records the order in which each focal ant's interactions occurred. For example, if an ant received 3 interactions before becoming active, the row with the time of her first interaction would be labeled 1, the row with time of her second interaction would be labeled 2, and the row with her final interaction would be labeled 3.

2. **refractiverepeat.csv –** File containing the times that 11 separate ants (11 in 1 colony of *L.* sp. W and 11 in 1 colony of *L. crassipilis*) became spontaneously active (event = 3), ignored a nestmate stimulation while inactive (event = 0), activated in response to a nestmate stimulation (event = 1), left the nest (event = 5), or became inactive (event = 4; event = s if it was the first instance that the focal ant was observed becoming inactive). We did not assign any behavior to event = 2. The column *rtime* measures the time of certain events (all activations and inactivations) relative to when the focal ant last switched between either the active or inactive state. The column *time* measures the times of all events relative to the start of the colony's video recording.

3. **refractivesample.csv** – File containing the durations that separate individuals spent inactive before either moving spontaneously (spon = 1) or being triggered into movement by a nestmate (spon = 0).

4. **wisc_empirical_parameters.csv** – 100 simulated activity time series generated by the agent-based model using parameters based on *L.* sp. W data. Each column of the file represents a separate time series.

5. **L.crassiplis_35hr_colony.csv -** This file contains all empirical 35-hour time series from *L. crassipilis* prior to the application of rescaling and smoothing. The columns (expect for the first column) are separate time series, and the rows of the first column denote the time points of each data value.

6. **L.crassiplis_9hr_colony.csv -** This file contains all empirical 9-hour time series from *L. crassipilis* prior to the application of rescaling and smoothing. The columns (expect for the first column) are separate time series, and the rows of the first column denote the time points of each data value.

7. **L.crassiplis_individuals.csv -** This file contains all empirical time series from isolated individuals of *L. crassipilis* prior to the application of rescaling, smoothing, or the correction of NaN values to 0 values. The columns (expect for the first column) are separate time series, and the rows of the first column denote the time points of each data value.

8. **L.spW_35hr_colony.csv -** This file contains all empirical 35-hour time series from *L.* sp W prior to the application of rescaling and smoothing. The columns (expect for the first column) are separate time series, and the rows of the first column denote the time points of each data value.

9. **L.spW_9hr_colony.csv -** This file contains all empirical 9-hour time series from *L.* sp W prior to the application of rescaling and smoothing. The columns (expect for the first column) are separate time series, and the rows of the first column denote the time points of each data value.

10. **L.spW_individuals.csv -** This file contains all empirical time series from isolated individuals of *L.* sp W prior to the application of rescaling, smoothing, or the correction of NaN values to 0 values. The columns (expect for the first column) are separate time series, and the rows of the first column denote the time points of each data value.

11. **Ant_ABM_cycles.nlogo** – code for agent-based model.

12. **cras_empirical_parameters.csv** – 100 simulated activity time series generated by the agent-based model using parameters based on *L. crassipilis* data. Each column of the file represents a separate time series.

13. **CV_agg.csv** – The dominant collective period (*waveperiod* column), Coefficient of variation (CV – *fanopeak* column), and mean Inter-beat interval (IBI – *peakmean* column) associated with each empirical time series used in this study. The *name* column denotes the name of the colony (for colony-level time series) or the colony of origin (for time series of isolated individual ants).

14. **Lcomp.R** – R code used for statistical analyses.

15. **noise_uniform_1100_50.csv** – 350 simulated activity time series generated by the agent-based model. Each column of the file represents a separate time series. Columns 1-50 are time series where $\Omega = 0$, in columns 51-100 $\Omega = 50$, in columns 101-150 $\Omega = 100$, in columns 151 - 200 $\Omega = 150$, in columns 201-250 $\Omega = 200$, in columns 251 - 300 $\Omega = 250$, and in columns 301-350 $\Omega = 300$. The model's other parameters were set to: $\langle S \rangle = 4050$ seconds, $A = 216$ seconds, and $\langle R \rangle = 1100$ seconds.

16. **refractive_period_sim.csv** – 350 simulated activity time series generated by the agent-based model. Each column of the file represents a separate time series. Columns 1-50 are time series where $R = 500$, in columns 51-100 $R = 600$, in columns 101-150 $R = 700$, in columns 151 - 200 R

= 800, in columns 201-250 R = 900, in columns 251 - 300 R = 1000, and in columns 301-350 R = 1100. The model's other parameters were set to: <S> = 4050 seconds, A = 216 seconds, and Ω = 0.

17. **refractive_period_sim_exp.csv** – 350 simulated activity time series generated by the agent-based model where the refractory parameter R was sampled from an exponential distribution. Each column of the file represents a separate time series. Columns 1-50 are time series where R = 500, in columns 51-100 R = 600, in columns 101-150 R = 700, in columns 151 - 200 R = 800, in columns 201-250 R = 900, in columns 251 - 300 R = 1000, and in columns 301-350 R = 1100. The model's other parameters were set to: <S> = 4050 seconds, A = 216 seconds, and Ω = 0.

18. **Time_series_analysis_Leptothorax.m** – MATLAB file that calculates the dominant oscillation period (according to wavelet analysis), CV (coefficient of variation of the time between activity peaks), and IBI (inter-beat interval aka. mean time between activity peaks) of empirical time series from this study.

19. **multirhythm_sim.m** – MATLAB file that performs Lomb–Scargle periodogram analysis on time series generated by the agent-based model. This script outputs some of the plots from figure 6 in the main text that show the appearance of multiple rhythms in the simulated time series.

20. **multirhythm_empirical.m** – MATLAB file that performs Lomb–Scargle periodogram analysis on the empirical 35-hour time series. This script outputs some of the plots from figure 2 and figure 3 in the main text that show the appearance of multiple rhythms in the empirical time series.

21. **Leptothorax_sim_exp_collective_time_series_all_boxplot.m** – MATLAB file that calculates the dominant oscillation period (according to wavelet analysis) on every collective-level empirical and simulated time series in this study. This script outputs figure S3, which is the full version of Figure 5C in the main text.

22. **wisc_5deg.csv** – 100 simulated activity time series generated by the agent-based model using parameters based on *L.* sp. W data. For each active ant in these simulations, the heading of the agent's next step was determined by adding ± 5 degrees to its prior heading. This is in contrast to the ± 45 degrees used in the simulations contained in the file "**wisc_empirical_parameters.csv**". Using ± 5 degrees causes agents to have straighter walking paths compared to agents that adjusted their headings by ± 45 degrees. Each column of the file represents a separate time series.

23. **cras_5deg.csv** – 100 simulated activity time series generated by the agent-based model using parameters based on *L. crassipilis* data. For each active ant in these simulations, the heading of the agent's next step was determined by adding ± 5 degrees to its prior heading. This is in contrast to the ± 45 degrees used in the simulations contained in the file "**wisc_empirical_parameters.csv**". Using ± 5 degrees causes agents to have straighter walking

paths compared to agents that adjusted their headings by ± 45 degrees. Each column of the file represents a separate time series.

24. **wisc_360deg.csv** – 100 simulated activity time series generated by the agent-based model using parameters based on *L.* sp. W data. For each active ant in these simulations, the heading of the agent's next step was determined by adding ± 360 degrees to its prior heading. This is in contrast to the ± 45 degrees used in the simulations contained in the file "**wisc_empirical_parameters.csv**". Using ± 360 degrees causes agents to have more random walking paths compared to agents that adjusted their headings by ± 45 degrees. Each column of the file represents a separate time series.

25. **cras_360deg.csv** – 100 simulated activity time series generated by the agent-based model using parameters based on *L. crassipilis* data. For each active ant in these simulations, the heading of the agent's next step was determined by adding ± 360 degrees to its prior heading. This is in contrast to the ± 45 degrees used in the simulations contained in the file "**wisc_empirical_parameters.csv**". Using ± 360 degrees causes agents to have more random walking paths compared to agents that adjusted their headings by ± 45 degrees. Each column of the file represents a separate time series.