A mechanism for persistent information in oscillatory networks

M. Rule

September 9th, 2014

Abtract: This is a PDF transcript of the slides for the Methods in Computational Neuroscience (MCN) final presentation talk, fall of 2014. It builds on Rule et al. (2011) "The origin and properties of flicker-induced geometric phosphenes", which showed that oscillatory drive can excite patterned instabilities in visual cortex. These patterns constitute attractor states that are stabilized by an external oscillatory drive. This project explored the idea that similar oscillation-stabilized attractors might serve as a flexible working memory. Unfortunately, these notes are rather brief as I never had the time to write this up properly.

1 Hypothesis and model

Hypothesis Attractor networks are hypothesized to store information in the brain, as a form of working memory. However, stationary rate dynamics are seldom observed in the brain: real networks have a tendency to oscillate due to the latency of inhibitory feedback. Can we generalize the idea of information storage in attractor states to oscillatory networks? In this work, we will show how shared oscillatory drive can activate a mode with distinct stable trajectorie, rather than fixed points.

A rate model E-I oscillator We begin by elaborating on the model of Rule et al. (2011), which consisted of a two-population Wilson-Cowan model with excitatory (E) and inhibitory (I) cells (Figuse 1). The euqations for this model are:
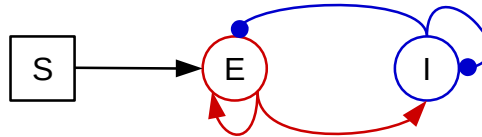

Figure 1: Network configuration for a neural mass model of a network containing excitatory (E) and inhibitory (I) neurons, driven by an external stimulus (S). Depending on the network parameters, the excitatory and inhibitory cells can exhibit a steady state, or they can oscillate spontaneously. This network can be configured so that it is normally quiescent, but can be driven in resonance by an external oscillatory stimulus. See Rule et. al (2011) for further details.

$$\tau_e \dot{E} = -E + f(A_{ee}E - A_{ei}I + \theta_e + S(t) + \eta_e(t)),$$
$$\tau_i \dot{I} = -I + f(A_{ie}E - A_{ii}I + \theta_i + \eta_i(t)),$$
$$f(x) = \frac{1}{1 + e^{-x}},$$
$$\eta \sim \mathcal{N}(0, \sigma^2),$$

where f is a logistic sigmoidal firing-rate nonlinearity, A are synaptic coupling constants, θ are (negative) thresholds, and τ the response time-constants. We define two inputs to the system: an external

stimulus $S(t)$, and Gaussian white noise $\eta^{\dagger}$. We set the network parameters as in Rule et al. (2011) so that this system exhibits a damped oscillatory mode, which is easily excited by an external stimulus at the same frequency.

Weakly coupled oscillators   The potential for such a system to serve as a short-term memory is highlighted by the two-population model explored in Rule et al. (2011) (See "Results: Global dynamics for a highly reduced model"). In this result, two neural oscillators are joined together with weak inhibitory coupling (Figure 2). With weak inhibitory coupling, the system exhibits instability when receiving symmetric drive ($S_1 = S_2$, Figure 2).
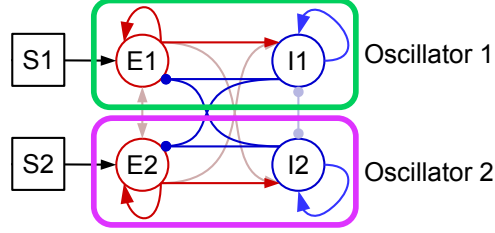

Figure 2: Two neural oscillators coupled with weak inhibitory coupling. This corresponds to the "highly-reduced" two-population model explored in Rule et al. (2011).

As shown in Rule et al. (2011) Figure 7. "Global picture for the simple 4-dimensional model", this instability can take the form of either a symmetric oscillation (in which the oscillators take turns firing during the peak of stimulation), or an asymmetric mode in which a single oscillator dominates, depending on the driving frequency.

The equations for a pair of neural oscillators coupled with weak inhibition are:

$$\tau_e \dot{E}_1 = -E_1 + f(A_{ee}^{self} E_1 - A_{ei}^{self} I_1 - A_{ei}^{other} I_2 + \theta_e + S_1(t) + \eta_{e,1}(t))$$

$$\tau_e \dot{E}_2 = -E_2 + f(A_{ee}^{self} E_2 - A_{ei}^{self} I_2 - A_{ei}^{other} I_1 + \theta_e + S_2(t) + \eta_{e,2}(t))$$

$$\tau_i \dot{I}_1 = -I_1 + f(A_{ii}^{self} I_1 - A_{ie}^{self} E_1 - A_{ie}^{other} E_2 + \theta_i + \eta_{i,1}(t))$$

$$\tau_i \dot{I}_2 = -I_2 + f(A_{ii}^{self} I_2 - A_{ie}^{self} E_2 - A_{ie}^{other} E_1 + \theta_i + \eta_{i,2}(t)),$$

where the synaptic coefficients are now broken down into A^{self} and A^{other} contributions.

2   Simulations

We first confirm that the two-population model (Figure 2) can be coaxed to enter one of two oscillatory modes, depending on its initial state. This serves as a proof-of-principle for a simple, 1-bit oscillatory memory. We will then build on this idea to explore memory circuits with higher capacity.

---

†Adding noise inside the nonlinearity is technically wrong; it should be phrased as an Îto stochastic differential equation, and the noise should be added outside the nonliterary so that it may be scaled correctly with the integration time-step, Δt. However, moving the noise inside the firing-rate nonlinearity ensures that firing rates remain bounded on $(0, 1)$. You can approximately map this to the Îto formalism either by accounting for the current slope of $f(\cdot)$ when scaling noise by $\sqrt{dt}$, or by creating a second set of equations for neural *activations*, and incorporating the noise term into these. In this work, making this more formal doesn't really buy us anything, so we will simply say that the results of integrating these equations are specific to the particular choice of time step Δt used for numeric integration.

Weakly coupled oscillators: "experiment" In this simulation, the system begins at steady state, with both populations identical. We then providing a driving input to *one* population, increasing the firing rate of the excitatory cells therein. This sets up the initial conditions of the system. The input is then removed, and the network is driven with *shared* oscillatory drive. We will show that the drive sustains the memory of the initial input, and allows it to be read-out from the firing rates of the E populations at a later time (Figure 3).

Storing information in an ensemble of 30 oscillators This intuition generalizes to $N > 2$ population models. In this next simulation, we consider $N = 30$ oscillatory populations, all identically coupled to each other with weak inhibition. In this case, we have arbitrarily assigned 15 populations into "group 1" and the remaining into "group 2". A stimulus arriving to group one drives the activity of all populations in this group, creating an asymmetry in the network. With oscillatory drive in the "hold" period, this asymmetry can still be observed several cycles later, confirming that the system retains the memory trace (Figure 4).

Two islands of encoded assembly stability All simulations presented so far were with weak noise ($\sigma^2 < 0.032$), and primarily reflect the stability of the mean-field model to small perturbations. We next evaluated how this system behaves in the presence of larger noise levels. Figure 5 illustrates the stability of the encoded memory for a range of noise values. It also evaluates different frequencies and amplitudes of the synchronous oscillatory drive used to stabilize the memory. As in Rule et al. (2011), we find two islands of stability, which correspond to the symmetric and asymmetric modes of population oscillation.
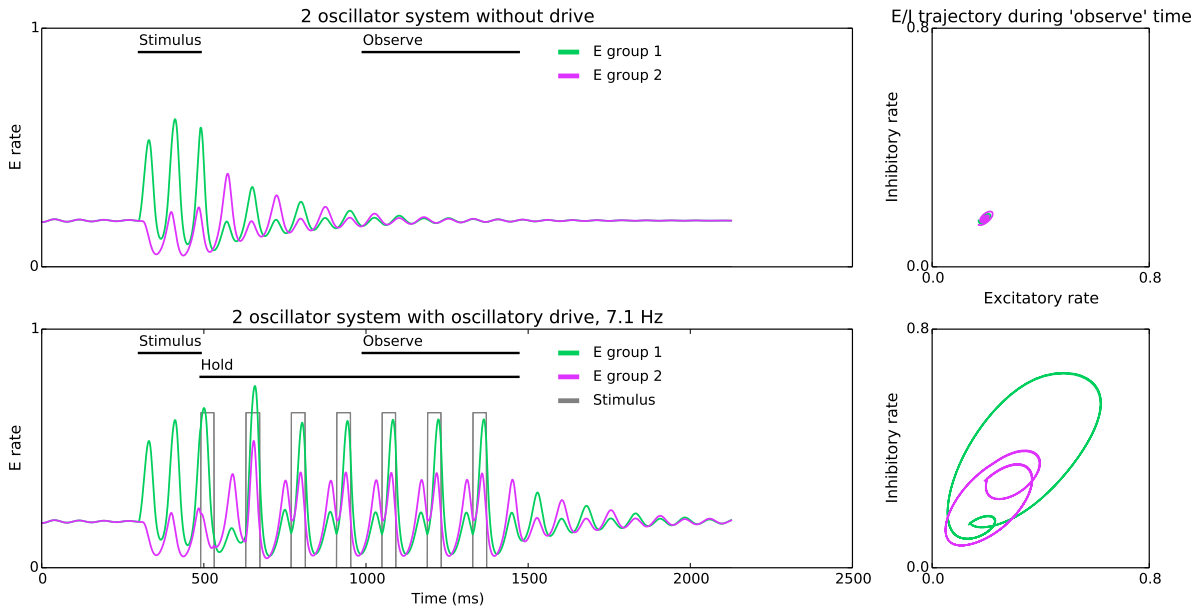

Figure 3: *Memory retention in the two-population model of Rule et al. (2011).* Starting from rest, a stimulus is delivered to the first oscillator (green) for $t \in [300, 500)$ ms. We test the stability of this stimulus-driven network response during a hold period $t \in [500, 1500)$ ms. Without oscillatory drive (top), the memory fades in a few cycles. Periodic stimulation (bottom) preserves the memory. Rightmost plots show the phase-plane dynamics during the readout period $t \in [1000, 1500)$. The top system, without drive, has returned to rest. The driven system shows a stable limit cycle, with higher firing rates in the population that initially received the stimulus.
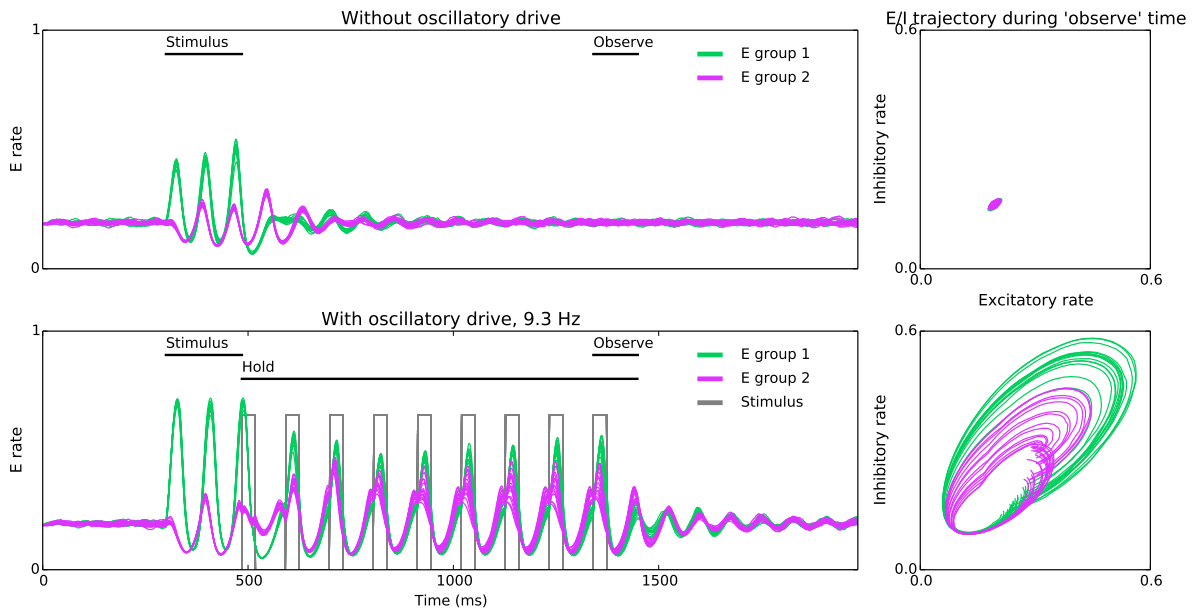
3

Figure 4: *Many-population generalization of the two-population simulated memory experiment of Figure 3.* Memory retention is similar to the two-population case. However, the dynamics of this larger population of neural oscillators is less stable: rates fluctuation for different oscillators within each group.
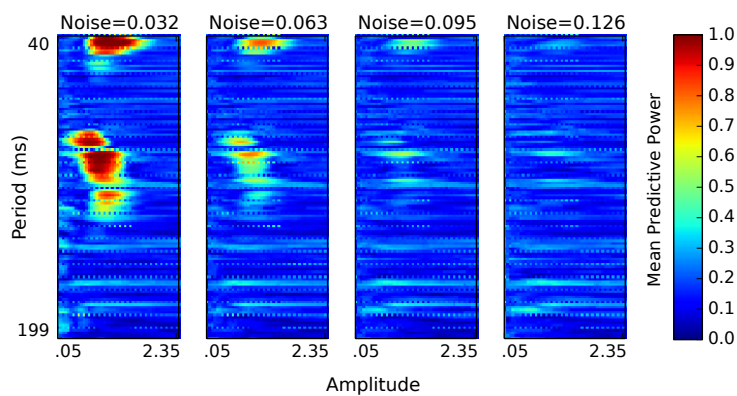

Figure 5: *A survey of memory robustness reveals two islands of stability.* We repeated the 30-population experiment of Figure 4, varying the noise level, as well as the amplitude and period of the synchronous oscillation used to stabilize the memory. Two islands of stability are apparent, at periods of 40 and 80 ms. These correspond to the symmetric and asymmetric regimes of Rule et al. (2011), respectively.

3 Switched linear system analogy

We would like to better understand the basic principles underlying oscillatory memory stabilization in the experiments. However, it is difficult to analyze higher-dimensional nonlinear stochastic systems. To provide at least some qualitative intuition, we now move to a model that replaces the nonlinear system with a switched linear system. This system is linear at almost all time-points, but we will model the driving oscillation as a switching of its parameters between one of two regimes.

Can we view this as a damped, driven linear system? We begin by returning to the two-population scenario. To build a linear system similar to the two-population model, start with a 4D linear system. Ensure that its largest eigenvalue pair reflects an oscillation with the target frequency, and is stable $\Re(\lambda) < 0$. Ensure, then, that the eivenvector for this state reflect an asynchronous mode.

Figure 6 shows a simulation of a damped linear resonator. In these simulations, we provide external drive in the form of a train of impulses, delivered every 100 ms to both populations. The transient response of the switched linear system qualitatively resembles that of the nonlinear model (Figure 3). However, its response to a shared driving oscillation is very different: instead of exciting instability, we see a continued decay of any initial differences in firing rate, superimposed on a synchronous oscillation.
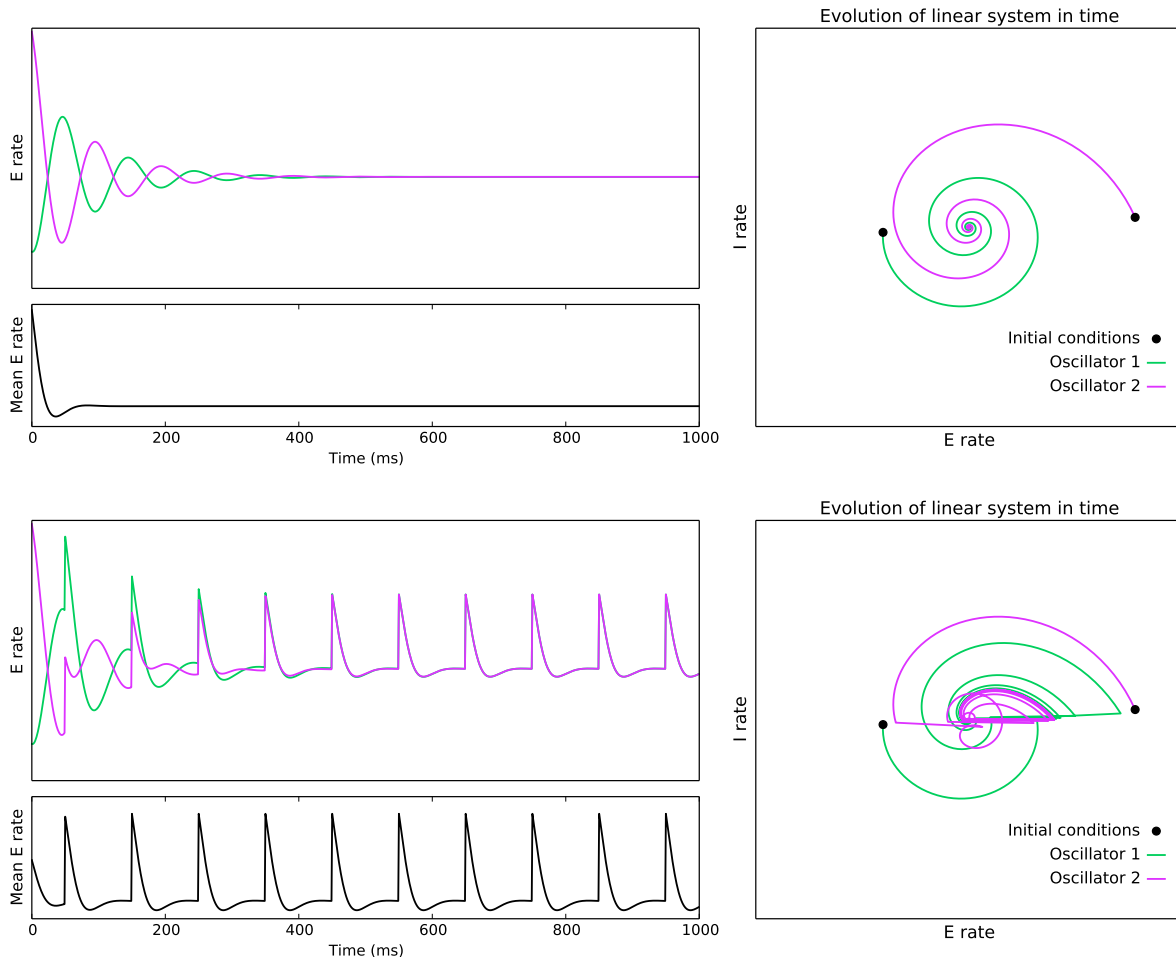

Figure 6: *Analogy of the two-population model as a damped, driven linear system.* We model the nonlinear neural oscillator as a damped linear oscillator. In the absence of external periodic drive (top), this system qualitatively captures the decay of network activity to a transient input. However, this system does not resemble the nonliear model in the presence of oscillatory drive: synchronous drive cannot excite an asymmetric mode, and both populations converge to identical activity.

The firing rate nonlinearity is important   We have confirmed that a linear system cannot capture the effects that we see: there is no way for synchronous drive to excite an asynchronous mode. The key to this phenomenon must therefore lie in the firing-rate nonlinarity (Figure 7).
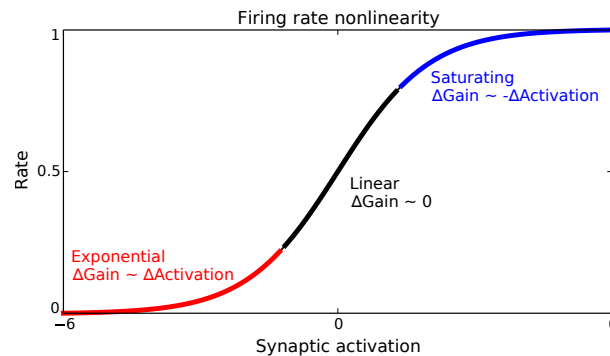

Figure 7: Sigmoidal firing rate nonlinearity. This transfer function is linear in the vicinity of 0, but resembles the exponential for large negative inputs, and saturates for large positive inputs. The nonlinear system occupies the exponential portion of the firing rate nonlinearity, in which an increase in drive leads to an increase in gain.

Importantly, the logistic nonlinearity has positive curvature for low amounts of activation. For the nonlinarity chosen here, in fact, it resembles the exponential function. Any increase in amplitude, therefore, causes a corresponding increase in gain. This periodically-driven gain increase is responsible for destabilizing the synchronous state, and thereby preserving the asynchronous mode corresponding to the initial memory.

Linear analogy to two coupled oscillator model   Figure 8 illustrates a working linear analogy to the nonlinear model. Here, rather than provide a periodic impulse, we instead periodically increase the gain. This makes the asynchronous modes unstable. Left unchecked, this would lead to diverging behavior. We therefore capture the saturating aspects of the nonlinearity by limiting the maximum firing rates. This provides a piecewise linear model of the nonlinear system that exhibits similar dynamics and stability.
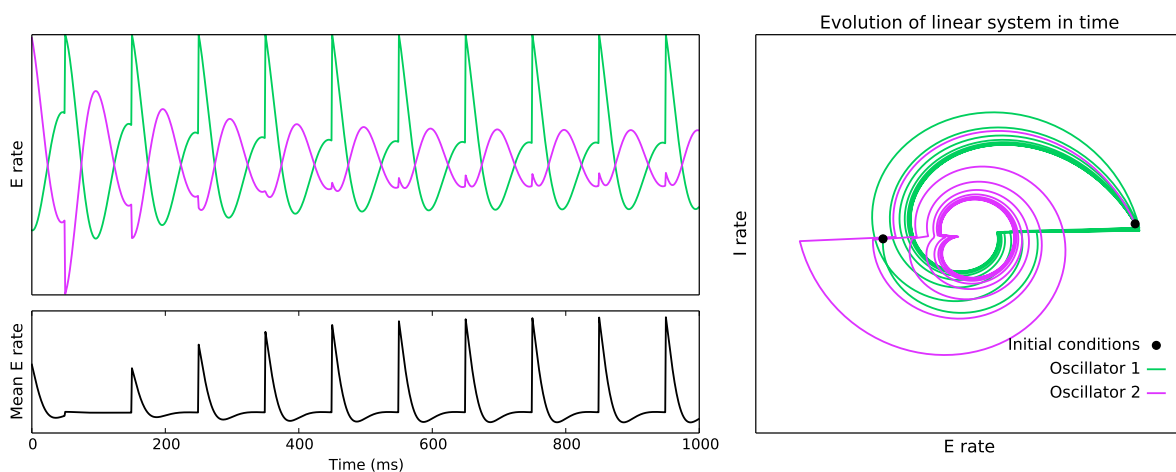

Figure 8: Periodic modulation of E-E coupling *and* periodic forcing, with a limit on the maximum rate, qualitatively resembles nonlinear system

4 Summary:

Selective, short-term maintenance of attractor states is important for working memory in neural systems. Neural networks, however, are typically stabilized by feedback inhibition. The delays in this inhibition inherently lead to unavoidable oscillations. Here we have explored a scenario where the attractor dynamics associated with a working memory are not fixed points, but instead fixed limit cycles embedded within a population oscillation.

We have shown that the presence of an external oscillatory drive, perhaps e.g. an attention signal, can switch such a network from a "read-in" mode, in which the system responds to external input, and a "hold" model, in which the system retains a memory of its past states. These concepts may underlie flexible working memory solutions that can coexist with population oscillations.

To better understand the mechanisms underlying this, we explored a switched piecewise linear model that captured the qualitative dynamics of the original system. We found that firing-rate nonlinearities with positive curvature are important for allowing a synchronous (non-selective) external signal to excite the asymmetric network states associated with a memory trace.

5 Appendix 1: Analysis of linear system

Linear system Can we think of this as a damped, driven linear system?

$$X = (E_1, E_2, I_1, I_2)$$

$$\dot{X} = JX$$

- Synchronous oscillatory drive cannot create asynchronous winner-take-all dynamics in a linear system

Linear system: reduced parameters In the nonlinear system, the effect doesn't require E-E coupling or I-I coupling between different oscillators. Only a mutually inhibitory term d is needed.

$$J = \begin{pmatrix} a & 0 & c & d \\ 0 & a & d & c \\ e & 0 & g & 0 \\ 0 & e & 0 & g \end{pmatrix}$$

This system has the following eigenvalues:

$$J = \begin{pmatrix} a & 0 & c & d \\ 0 & a & d & c \\ e & 0 & g & 0 \\ 0 & e & 0 & g \end{pmatrix}$$

$$w_1 = \sqrt{(a-g)^2 + 4e(c-d)} \qquad \begin{aligned} \lambda_1 &= (a+g-w_1)/2 \\ \lambda_2 &= (a+g+w_1)/2 \end{aligned}$$

$$w_2 = \sqrt{(a-g)^2 + 4e(c+d)} \qquad \begin{aligned} \lambda_3 &= (a+g-w_2)/2 \\ \lambda_4 &= (a+g+w_2)/2 \end{aligned}$$

Re-parameterize the linear system to maping these coefficients onto the synaptic coupling coefficients A:

$$J = \begin{pmatrix} a & 0 & c & d \\ 0 & a & d & c \\ e & 0 & g & 0 \\ 0 & e & 0 & g \end{pmatrix} = \begin{pmatrix} A_{ee} & 0 & -\sigma A_{ei} & -(1-\sigma)A_{ei} \\ 0 & A_{ee} & -(1-\sigma)A_{ei} & -\sigma A_{ei} \\ A_{ie} & 0 & -A_{ii} & 0 \\ 0 & A_{ie} & 0 & -A_{ii} \end{pmatrix}$$

$$w_1 = \sqrt{(A_{ee} + A_{ii})^2 + 4A_{ie}A_{ei}(1-2\sigma)} \qquad \begin{aligned} \lambda_1 &= (A_{ee} - A_{ii} - w_1)/2 \\ \lambda_2 &= (A_{ee} - A_{ii} + w_1)/2 \end{aligned}$$

$$w_2 = \sqrt{(A_{ee} + A_{ii})^2 - 4A_{ie}A_{ei}} \qquad \begin{aligned} \lambda_3 &= (A_{ee} - A_{ii} - w_2)/2 \\ \lambda_4 &= (A_{ee} - A_{ii} + w_2)/2 \end{aligned}$$

Add in a weak coupling parameter α:

$$J = \begin{pmatrix} a & 0 & c & d \\ 0 & a & d & c \\ e & 0 & g & 0 \\ 0 & e & 0 & g \end{pmatrix} = \begin{pmatrix} A_{ee} - 1 & 0 & -\sigma A_{ei} & -(1-\sigma)A_{ei} \\ 0 & A_{ee} - 1 & -(1-\sigma)A_{ei} & -\sigma A_{ei} \\ A_{ie} & 0 & -A_{ii} - \alpha_i & 0 \\ 0 & A_{ie} & 0 & -A_{ii} - \alpha_i \end{pmatrix}$$

$$w_1 = \sqrt{(A_{ee} + A_{ii} + \alpha_i - 1)^2 + 4A_{ie}A_{ei}(1-2\sigma)} \qquad \begin{aligned} \lambda_1 &= (A_{ee} - A_{ii} - w_1 - \alpha_i - 1)/2 \\ \lambda_2 &= (A_{ee} - A_{ii} + w_1 - \alpha_i - 1)/2 \end{aligned}$$

$$w_2 = \sqrt{(A_{ee} + A_{ii} + \alpha_i - 1)^2 - 4A_{ie}A_{ei}} \qquad \begin{aligned} \lambda_3 &= (A_{ee} - A_{ii} - w_2 - \alpha_i - 1)/2 \\ \lambda_4 &= (A_{ee} - A_{ii} + w_2 - \alpha_i - 1)/2 \end{aligned}$$

Note: all A are non-negative

Note: In the linear system, dynamics depend on product $A_{ie}A_{ei}$, not these weights individually. One could write merge these into a single lumped positive feedback parameter $A_{fb} = A_{ie}A_{ei}$.

$$w_1 = \sqrt{(A_{ee} + A_{ii})^2 + 4A_{fb}(1 - 2\sigma)} \qquad \begin{aligned} \lambda_1 &= (A_{ee} - A_{ii} - w_1)/2 \\ \lambda_2 &= (A_{ee} - A_{ii} + w_1)/2 \\ \lambda_3 &= (A_{ee} - A_{ii} - w_2)/2 \\ \lambda_4 &= (A_{ee} - A_{ii} + w_2)/2 \end{aligned}$$
$$w_2 = \sqrt{(A_{ee} + A_{ii})^2 - 4A_{fb}}$$

We impose the following constraints on the linear system:

- The system should be stable without input ($\mathfrak{R}(\lambda) \le 0$)

- Weights A_{ee} and A_{ie} are not negative

- Weights A_{ii} and A_{ei} are not positive

- Coupling $1 - \sigma < \frac{1}{2}$

This implies:

- Term w_2 becomes imaginary when $(A_{ee} + A_{ii})^2 < 4A_{fb}$

  - In this case the conjugate pair λ_3, λ_4 is stable when $\frac{1}{2}(A_{ee} - A_{ii}) < 1$. Larger $A_{ee} + A_{ii}$ supresses relative to the feedback term supresses oscillations. Larger A_{ii} relative to A_{ee} is stabilizing.

- Term w_1 becomes imaginary when $(A_{ee} + A_{ii})^2 < 4A_{fb}(1 - 2\sigma)$. When $\sigma > \frac{1}{2}$, the RHS is negative and w_1 is never imaginary.

- When w_1 is real, $\lambda_1 < \lambda_2$ can give rise to exponential modes. Both will be transient if $\lambda_2 < 0$. This requires that $A_{ee} - A_{ii} + \sqrt{(A_{ee} + A_{ii})^2 - 4A_{fb}} < 0$

6 Appendix 2: Weakly coupled oscillators: effective coupling is stronger than individual coupling

For the $N > 2$ case, we have many oscillators to consider. Empirically we find that the dynamics are attractive: forces within each group pull trajectories together, and forces between each group maintain the distinct memory. Nothing about the connectivity assumptions ensures this: the coupling between each pair of neural oscillators is identical, and the groups are arbitrary. The memory stability therefore hinges on emergent dynamics.
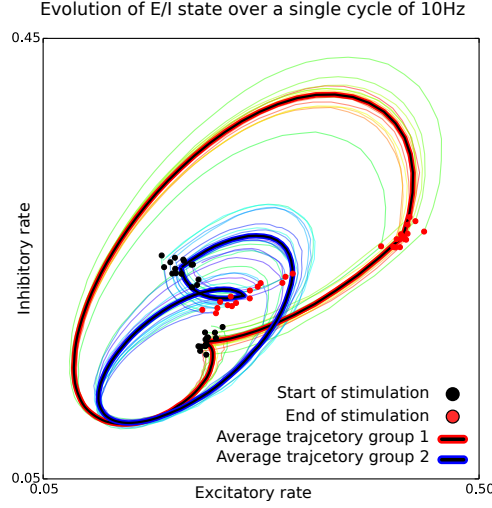

Figure 9: Empirically, we find that arbitrary stimulus-driven groupings within the population of neural oscilltors are stable. Starting from an initial condition that contains a small perturbation, the trajectories converge to a stable, fixed limit cycle (black traces).

Decompose effective excito-excitatory interation into the contribution from within and outside each neural oscillator:

$$S_e = \sigma_e \langle E_{other} \rangle + (1 - \sigma_e) E_{self}$$

In limit of large N, for the bistable mode, some fraction γ of E_i will be part of the same ensemble as E_j, call E_{self}, and the rest will occupy an E_{other}

$$S_e = \sigma_e (\gamma E_{self} + (1 - \gamma) E_{other}) + (1 - \sigma_e) E_{self}$$

$$S_e = \sigma_e (1 - \gamma) E_{other} + (1 - \sigma_e (1 - \gamma)) E_{self}$$

There is a new effective coupling constant $\sigma_e' = \sigma_e (1 - \gamma)$.

(unfinished)