Mean Field Models

Richard Bertram

Department of Mathematics and

Programs in Neuroscience and Molecular Biophysics

Florida State University

Tallahassee, Florida 32306

The Wilson-Cowan Model

The biophysical models that we have primarily focused on thus far can tell us a lot about single neurons and small populations of coupled neurons. But brains contain an almost unimaginable number of neurons. Just the neocortex (the part involved in higher-order brain functions, such as sensory perception and cognition) of a rat brain contains roughly 10 million neurons, while that of a human is roughly 1 billion neurons. Each of these neurons may make hundreds of synaptic connections with other neurons, so in human neocortex we have a network with about 1 billion nodes and 100 billion or more edges. Throughout the human brain, estimates are that there are at least **100 trillion synaptic connections**. This is 1000 times the number of starts in our galaxy! You can't expect to describe this with the biophysical neurons we have used. Even the simpler integrate-and-fire neurons are of little use. To describe realistic neural structures, one must resort to phenomenological models that capture some aspects of the structure while disregarding most others. An example is a mean field model, also called a firing rate model. Such a model uses a single variable to describe the fraction of firing neurons in a neural population at each point in time. In a few cases, one can derive a mean field model from a network of single-neuron models, taking the limit as the number of neurons goes to infinity. But in practice this is rarely done, and

instead the mean field model is developed in an ad hoc fashion. The best example, which we'll explore, is the **Wilson-Cowan model** (1972) for interacting populations of excitatory and inhibitory neurons. There is a variable for the fraction of active excitatory neurons and a variable for the fraction of active inhibitory neurons.

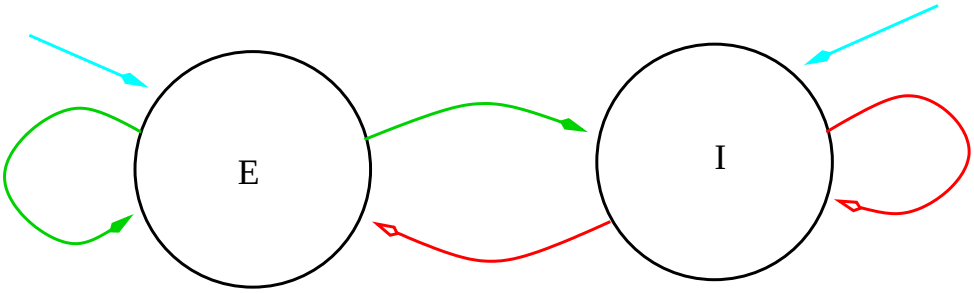

Figure 1: Two neural populations in the Wilson-Cowan model. One (E) is excitatory, the other (I) is inhibitory.

Such a model can be used profitably if there is no special spatial or temporal structure within a subpopulation. For example, it is useful if the E neurons are randomly connected, but would not be useful if the E neurons are clustered into interconnected layers.

Consider first only the excitatory population, with no I neurons. Then

$$\frac{dE}{dt} = [-E + f_E]/\tau_E \tag{1}$$

where τ_E is a time constant, $-E$ describes the first-order decay of E towards 0, and f_E describes the input into E. If one used a linear function for f_E, then the system could experience uncontrolled growth. Therefore,

a saturating function is used:

$$f_E = \frac{1}{1+e^{-x}} \tag{2}$$

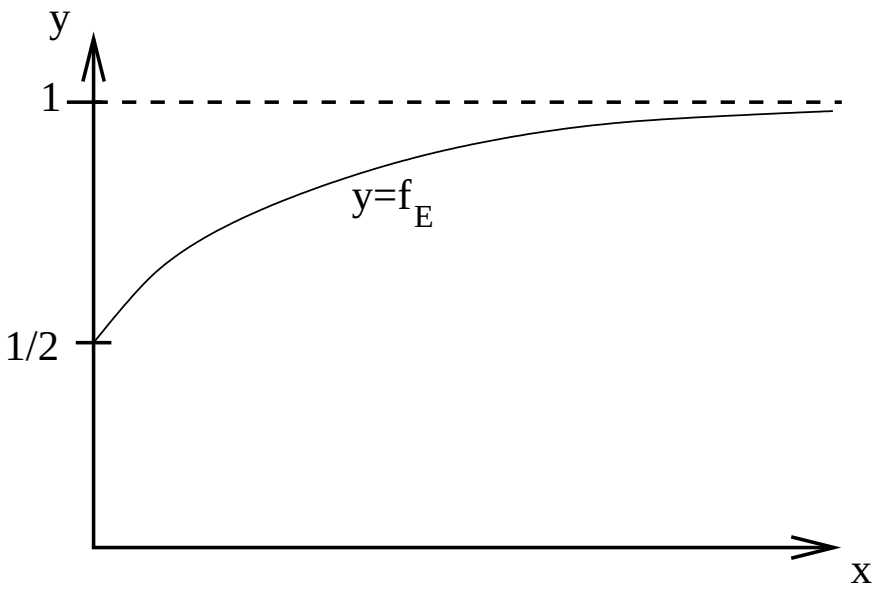

Figure 2: The input function in the Wilson-Cowan model.

The argument x of the input function includes both the **autofeedback** of E onto itself, and a constant term (p_E) representing extrinsic input:

$$x = aE + p_E \quad . \tag{3}$$

Together,

$$\frac{dE}{dt} = \left[-E + \frac{1}{1+e^{-(aE^*+p_E)}}\right] / \tau_E \quad . \tag{4}$$

This is a 1-dimensional system, with equilibria satisfying

$$E^* = f_E = \frac{1}{1+e^{-(aE^*+p_E)}} \quad . \tag{5}$$

A nice graphical method to solve this nonlinear algebraic equation is to plot $y = E$ and $y = f_E$ and look for intersections (Fig. 3).

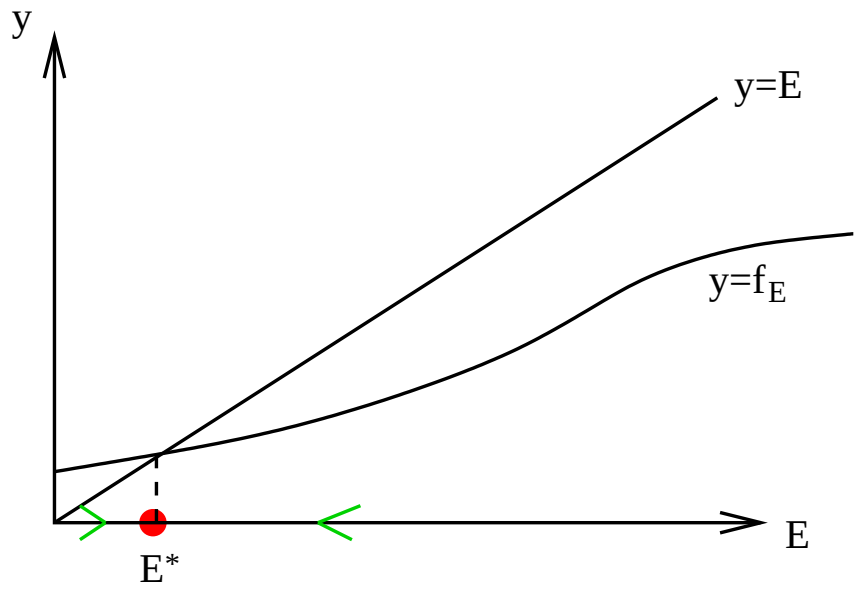

Figure 3: Single steady state when $a = 6$.

The single steady state is **stable** since $E > f_E$ when $E > E^*$, so $\frac{dE}{dt} < 0$ in this case. Also, $E < f_E$ when $E < E^*$, so $\frac{dE}{dt} > 0$. Together, these prove that the steady state is stable.

If the strength of the autofeedback, a, is increased, then the $y = f_E$ curve is deformed and two new steady states are born (Fig. 4). The system is now **bistable**; the outer steady states are stable, the inner one is unstable. The existence of the larger stable steady state E_3^* reflects the regenerative nature of this system due to the presence of positive feedback.

Now we add the inhibitory neurons into the system. These are just like the excitatory neurons, except that their output has the opposite polarity.

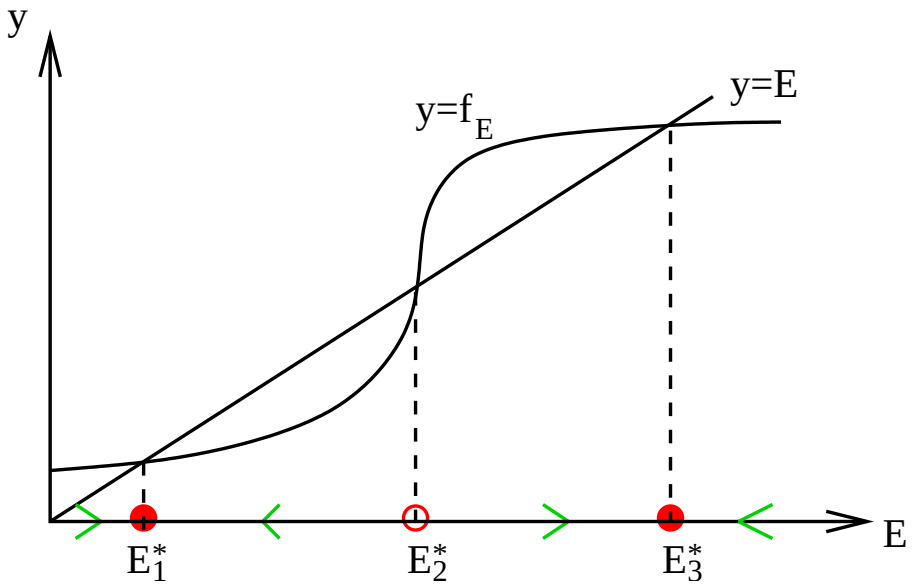

Figure 4: The system is bistable when $a = 10$.

The I neurons provide input to both the E and I neurons, so the equations for the two types are:

$$\frac{dE}{dt} = [-E + f_E(aE - bI + p_E)]/\tau_E \quad (6)$$

$$\frac{dI}{dt} = [-I + f_I(cE - dI + p_I)]/\tau_I \quad (7)$$

where f_I has the same form as f_E. This is the Wilson-Cowan model.

The E**-nullcline** is $E = f_E(aE - bI + p_E)$. The I**-nullcline** is $I = f_I(cE - dI + p_I)$. For parameter value $p_E = -5$ the phase portrait is shown in Fig. 5.

There is a single stable steady state, but if the phase point is perturbed past the middle branch of the E-nullcline, then there is a regenerative response, with E first increasing further before it begins to decrease back towards the steady state. This is very similar to an **impulse** produced

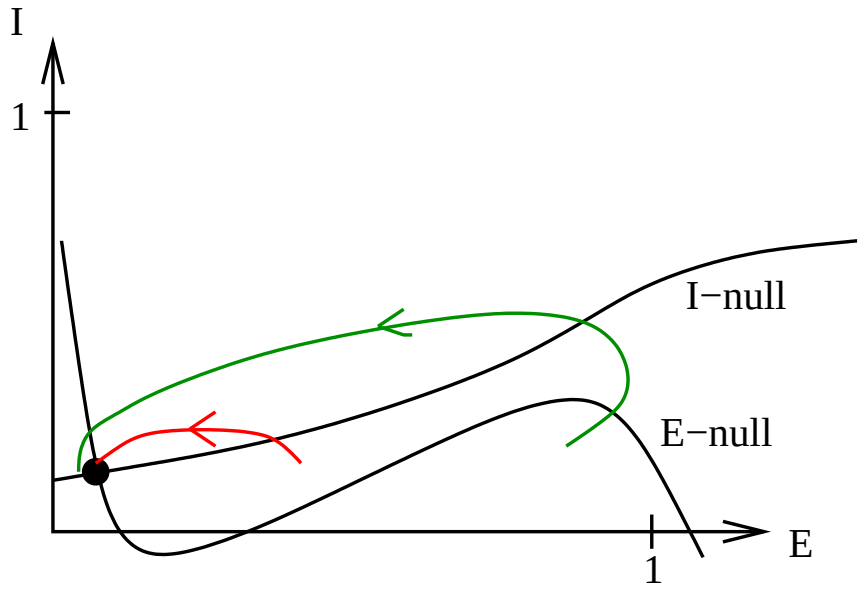

Figure 5: Wilson-Cowan phase portrait, excitable regime ($p_E = -5$).

by the Morris-Lecar model. In the context of Wilson-Cowan, this would correspond to a population spike. That is, a spike of activity in the population of E neurons followed by an I spike.

The dynamics of Wilson-Cowan are in fact very similar to the dynamics of Morris-Lecar:

$$E \Longleftrightarrow V$$

$$I \Longleftrightarrow w.$$

These are both prototypical models of an **excitable system**.

Like Morris-Lecar, Wilson-Cowan can produce limit cycle behavior. Increase the p_E parameter, translating the E-nullcline upward so that the intersection is on the middle branch.

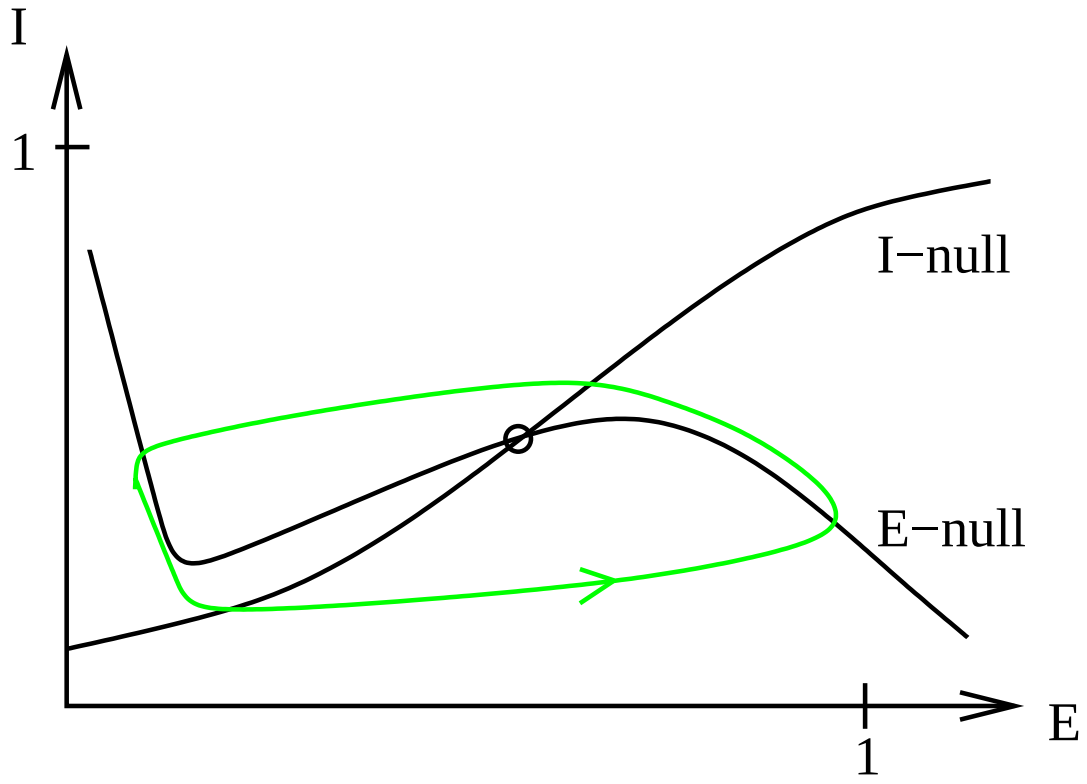

Figure 6: Limit cycle behavior, when $p_E = -1$.

This produces a periodic train of population spikes.

A Firing Rate Model for Developing Neural Networks

Developing neural networks typically lack inhibitory connectivity. Even the synpses that release GABA, the standard inhibitory neurotransmitter, are excitatory early in development. This is because the reversal potential for chloride ions, which are one of the main type of ions flowing through GABA receptors, is around 0 mV early in development, and only later becomes very negative. Since all the coupling is excitatory, one would expect that the firing neurons would recruit other neurons to fire, so the whole population would fire pretty much continuously. However, this does not happen. Instead, there are population spikes of activity, much like what one gets from the Wilson-Cowan model when it exhibits limit cycle behavior. However, Wilson-Cowan has inhibitory neurons, which can explain the periods of silence between active episodes. How does it work in immature networks? One possibility is that episodes of activity are terminated by synaptic depression. We will examine the firing rate model that was developed to explore this possibility (published by Tabak *et al.*, *J. Neurophysiol.*, 103:2208, 2010.)

Assume that there is no particular structure to the neural connectivity, a safe bet in immature networks, and let a denote the activity level of the neural population. When a neuron in this population fires it affects the other neurons with a **synaptic strength** of ws, where w is the synaptic

weight and s is the synaptic efficacy. The weight is a parameter that reflects the number and size of the synapses, while the efficacy reflects the degree of presynaptic depression and is a variable that ranges from 0 to 1. When $s = 0$ the synapse is completely depressed, perhaps due to depletion of the releasable vesicle pool, and when $s = 1$ there is no depletion at all. The more the neuron fires the smaller s becomes. Now instead of thinking of a single neuron, apply this to the entire population, so that s becomes the synaptic efficacy of the population.

The differential equations are

$$\frac{da}{dt} = -a + a_\infty(wsa - \theta_a) + n(t) \tag{8}$$

$$\frac{ds}{dt} = \frac{s_\infty(a) - s}{\tau_s} \ . \tag{9}$$

The a_∞ function is an increasing sigmoid of a since the greater activity of the network means greater excitatory input across the population. The s_∞ function is a decreasing sigmoidal function of a since the greater the activity the greater will be the presynaptic depression. The functions are:

$$a_\infty(x) = \frac{1}{1 + \exp(\frac{-x}{k_a})} \tag{10}$$

$$s_\infty(x) = \frac{1}{1 + \exp(\frac{x-\theta_s}{k_s})} \tag{11}$$

The factor wsa in Eq. 8 represents the synaptic input (synaptic strength times the activity level), while θ_a is a parameter that adjust the effect

that the input will have on the activity of the population. The term $n(t)$ is random input or "noise" that reflects the random or stochastic nature of synaptic transmission. We have seen many instances of the first order kinetics used in Eq. 9.

This model, with appropriate choices for parameter values, produces almost-periodic episodes of activity, capturing what is observed in the data. This is shown in Fig. 7. During each episode the efficacy s declines, due to synaptic depression. As it declines the regenerative coupling in the "network" becomes weaker, leading eventually to the termination of an episode. At this point the efficacy slowly rises, simulating the buildup of the releasable vesicle pool in the neurons. The activity-dependent rise and fall of synaptic coupling s is what drives the activity episodes. The small oscillations during each episode and silent phase are due to the noise term in the a differential equation.

Because this model only has two variables it can be analyzed in the phase plane. Figure 8 shows the two nullclines; the a-nullcline has the familiar cubic shape of the Wilson-Cowan mean field model as well as the Morris-Lecar model for a single neuron. The activity variable changes rapidly relative to the efficacy variable, so the trajectory follows the a nullcline (the "fast nullcline") except for jumps between the lower and upper branches. Because the nullclines intersect only on the middle branch

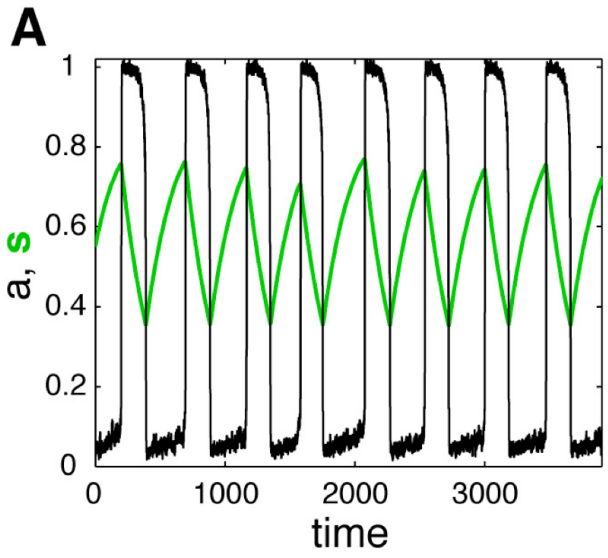

Figure 7: Almost-periodic episodes of activity in the firing rate model for an immature neural network.

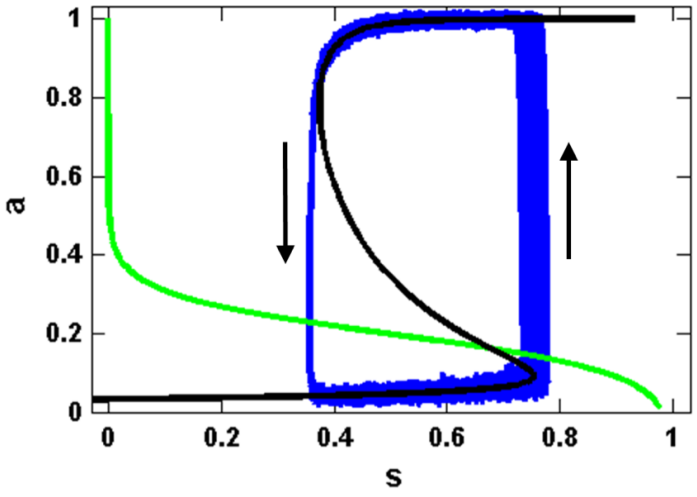

Figure 8: Noisy relaxation oscillations result from the time scale difference between the a and s dynamics, and the cubic-like a-nullcline.

of the a nullcline, the equilibrium is unstable. This results in relaxation oscillations. However, because of the noise term the trajectory is noisy, so we have noisy relaxation oscillations.

From Fig. 8 it appears that the noise is much greater at the end of a silent phase than at the end of an active phase. Why is this? To answer this, first note that the noise term is in the a ODE, and in the a nullcline equation,

$$a = a_\infty + n(t) \quad . \tag{12}$$

So one can look at the effects of positive and negative noise values on the a nullcline. With a positive noise value, $n = \Delta i$, the low knee moves significantly leftward, while it moves significantly rightward with negative noise, $n = -\Delta i$. In contrast, the high knee (HK) is barely effected by the noise. Since the trajectory follows the lower and upper branches of the a nullcline, and since the LK is the termination point of the silent phase, there will be a large variation in the duration of a silent phase. In contrast, since the HK is associated with the termination of an episode, there will be little variation in the point of termination of an episode and thus any variation in episode duration is due to variation in the starting point of the episode.

If there is a short silent phase, what effect will this have on the duration of the next episode? It will be short, since the starting point is close to

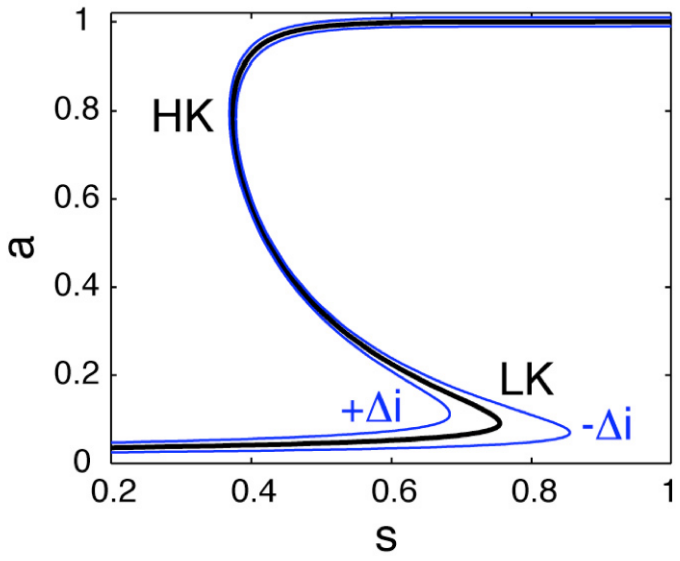

Figure 9: Positive and negative noise has a much greater effect on the low knee (LK) than the high knee (HK).

the termination point. So if one makes a plot of silent phase durations vs. subsequent episode durations there will be a positive correlation. If there is a short episode, what effect will this have on the duration of the next silent phase? None, since the silent duration is determined largely by the location of the termination point, and this is subject to random noise. So if one makes a plot of episode durations vs. subsequent episode durations there will be no correlation.

This is a very strong prediction of the model that comes directly from the phase plane analysis. Does it match the data from immature neural networks? Indeed it does. Similar correlation patterns have been observed in developing retina, developing cortical networks, and developing spinal cord (as shown in Fig. 11). This modeling prediction, validated by experi-

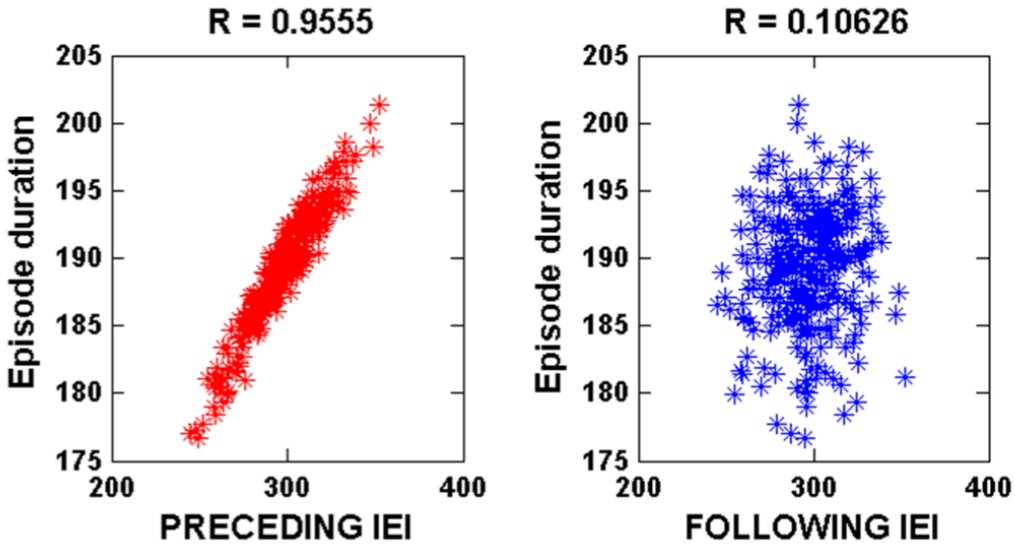

Figure 10: In the model, there is a positive correlation between preceding silent phase durations (IEI=inter-epsisode interval) and subsequent episode durations, but not between episode durations and subsequent silent phase durations.

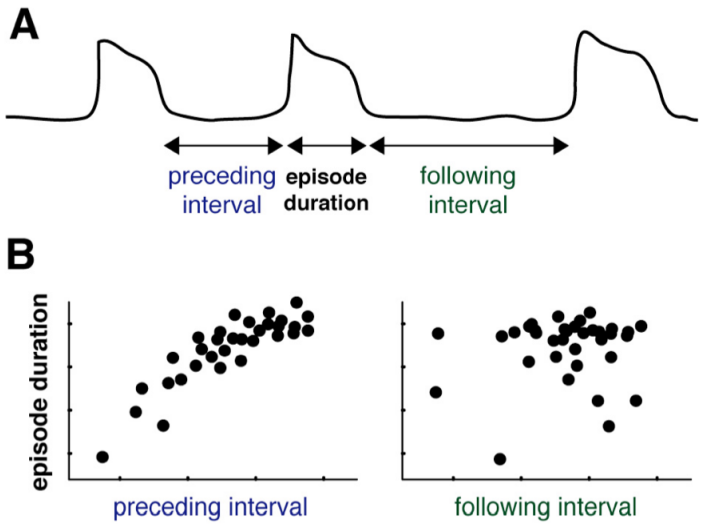

Figure 11: The model prediction regarding correlation patterns of episode and inter-episode durations is supported by data from the developing spinal cord. From Tabak *et al.*, *J. Neurophysiol.*, 103:2208, 2010.

mental studies, provides support for synaptic depression as the mechanism for episodic activity in developing neural networks.