

Chapter 6

Firing-rate models and network dynamics

In order for us to understand the functioning of a brain composed of billions or tens of billions of neurons, we must first understand simpler systems. In the earlier chapters, we simplified by looking at the processing of just a single neuron or of two to three neurons connected together. Another method of simplification is to ignore the idiosyncrasies of individual neurons and to aggregate the behavior of many neurons with similar properties. Such methods enable us to gain insight so that we can generate hypotheses for the mechanisms by which a variety of cognitive processes arise in neural circuits—insight that would be impossible when surveying voltage traces or spike trains from hundreds or thousands of neurons in a simulation or experiment.

Therefore, in this chapter we explore models of network dynamics at a level of detail that does not require simulation of the membrane potential of every neuron. If the information processed by a network of neurons is represented by the number of spikes produced by different subsets of neurons in small time-windows, then a simplified description of the network is possible. Such a description is the firing-rate model.

6.1. Firing-rate models

The key variables of a firing rate model are the firing rates, which correspond to the average number of spikes per unit time of a subset of similarly responsive cells. This is in contrast to spiking models in which the key variables are the membrane potentials of individual cells. The subset of similarly responsive cells in a firing rate model can be called a unit, such that the unit's firing rate variable represents the mean firing rate of its constituent neurons.

Box 6.1. Unit: In a firing rate model, a group of neurons whose mean activity is simulated.

In firing rate models, the impacts of all the spikes from one unit on the mean firing rate of another unit combine to produce an effective connection strength. Self-connections (*i. e.*, from a unit to itself) are possible, indeed common, in firing rate models, because neurons within a unit, being similarly responsive and often spatially proximal to each other, are often connected with each other. Therefore, the spikes from neurons within a unit alter the firing rate of other neurons within the unit, so the unit's mean firing rate is impacted by a contribution that depends on its own firing rate. Even if no individual neuron were connected with itself, the units of a firing rate model would do so. Such recurrent feedback is an important feature of many models of neural circuits that we will consider in more detail in this chapter.

Box 6.2. Recurrent feedback: Input to a unit that depends on the unit's own activity.

Firing rate models are valid if the effects of spikes from different neurons in the unit can be combined together linearly and if the spike times from different neurons in the unit are distributed uniformly across the typical time-window. The time-window of importance corresponds to the timescale over which the network can change its firing rate, which can be on the order of the neural membrane time constant for increases in rate, or on the order of the synaptic time constant for decreases in rate.

Firing rate models have two main advantages over membrane-potential models. First, and most importantly, being simpler, they allow us to better understand the behavior of many model circuits. Second, having fewer variables and with slower dynamics, they allow for simulations that are orders of magnitude faster than the corresponding spiking network model. The second advantage enhances the first, as rapid simulations allow us to explore large regions of parameter space, in some cases exhaustively so that we can exclude certain hypotheses. In many cases a firing rate model can be treated by mathematical analysis without simulations, so that the direct effect of a parameter—such as connection strength on mean firing rate or network stability—can be seen in the resulting formula. Thus, firing rate models provide an important step in allowing us to *explain* network behavior, when in more complicated simulations we might just be able to *describe* the behavior.

Of course, firing rate models have corresponding disadvantages. In reality spike times do matter, and the response of a downstream neuron to a barrage of spikes does depend on the correlations between those spikes on a short time scale, not just the total number within time-windows on the order of 10ms or more, as assumed in firing-rate models.

Moreover, the combining of many neurons into a unit is only valid if either: (1) The neurons have identical firing rate curves; or (2) if neurons with different firing rate curves have a fixed ratio of connection strengths to such different neurons in other units. If both of these two requirements are broken, then the effective connection strength from one unit to another depends on which subsets of cells within a unit have been excited, so a single mean firing rate is a poor description of the unit. Therefore, distinct classes of neurons (which typically have different connectivity patterns), as well as neurons with different tuning curves, should be treated as different units.

Finally, behavior that does depend on the membrane potential—for example, the conductance of the NMDA receptor channel (Section 5.1.2)—is omitted from the firing rate model, though in principle it can be included if mean membrane potential can be directly extracted from mean firing rate (*cf.* Figure 2.9C).

Box 6.3. Mean-field theory: An analysis that ignores correlations between variables so that the effect of fluctuations can be “smoothed out” by averaging and the system’s behavior described according to the mean effect of all variables that comprise a unit.

Any firing rate model is a mean-field theory in the sense that only the mean firing rate is represented—where the mean is the rate averaged across all neurons in a unit. That is, the influence of one unit on another is assumed to depend only on the activity averaged across all neurons in the unit. The assumption can be valid if spike times are asynchronous in the limit of large numbers of neurons per unit.

One effect of the finite number of neurons per unit that is absent in a standard firing-rate model formalism, is the impact of individual spikes from the unit. As we have seen in Chapter 5, each spike from a presynaptic cell produces a rapid change in synaptic conductance of any postsynaptic cells, which results in a small uptick in the mean input conductance of any connected unit. These discrete events at random times can be incorporated into a firing-rate model with the addition of input noise to each unit. Since the mean firing rate of a neuron depends on the level of its input noise (see Tutorial 2.1), in more sophisticated firing rate models such noise should be incorporated in a consistent

manner. However, within this course, we will just add a noise term in the differential equations and treat its magnitude as a parameter that can be adjusted to the level needed to produce realistic network behavior.

6.2. Simulating a firing-rate model

Firing rate models are based on the input-output function of a neuron (Figure 6.1), which describes its firing rate as a function of either its total input current (*cf.* Figures 2.8 & 2.9B) or as a function of its excitatory and inhibitory incoming synaptic conductance. The second formulation is perhaps better since the spikes of one cell directly impact the synaptic conductance rather than synaptic current of another cell. In many cases a generic input-output function is used, such as a “sigmoid”, which increases from zero at very low or very negative input and saturates at a maximum for large positive input. Alternatively, a power-law can be used, or an empirical fit can be made to the response of either a real neuron or a spiking model neuron.

Box 6.4. Sigmoid: Any function of the form $1/[1 + \exp(-x)]$, often used in firing rate models.

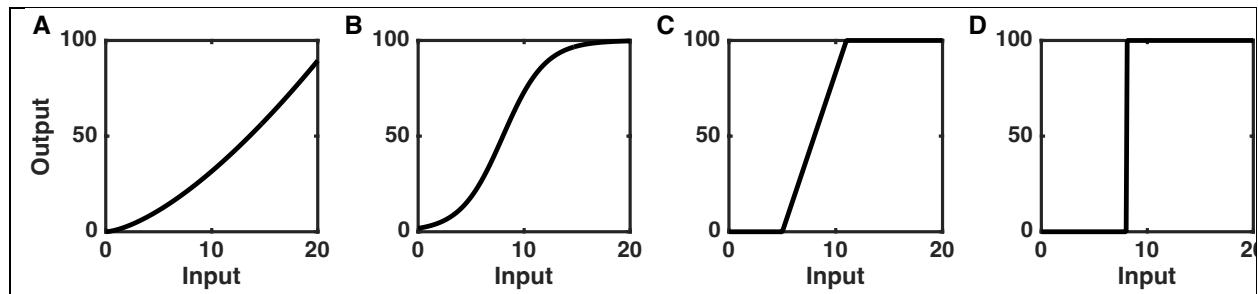


Figure 6.1. Examples of input-output functions for firing rate models. **A)** Power-law function, $y = A[I - I_0]_+^\alpha$ with amplitude $A = 1.2$, exponent $\alpha = 1.5$ and threshold $I_0 = 0$. The square brackets $[I - I_0]_+$ equate to 0 if $I < I_0$ and to $I - I_0$ otherwise. **B)** Sigmoid function, $y = \frac{r_{max}}{1+\exp[-(\frac{I-I_{0.5}}{\sigma})]}$ with maximum output, $r_{max} = 100$, input for half-max, $I_{0.5} = 8$, and inverse steepness, $\sigma = 2$. **C)** Threshold linear response function, with $y = 0$ for $I < I_0$, $y = \frac{r_{max}(I-I_0)}{\Delta I}$ for $I_0 \leq I \leq I_0 + \Delta I$, and $y = r_{max}$ for $I > I_0 + \Delta I$. Saturation is at $r_{max} = 100$, responsive range has $\Delta I = 6$, and the threshold is $I_0 = 5$. **D)** Binary response function, with $y = 0$ for $I < I_0$ and $y = r_{max}$ for $I > I_0$. Saturation is at $r_{max} = 100$ and the threshold is $I_0 = 8$.

As well as the unit’s input-output function, which is its firing rate response, it is also important to know how the input to other units depends on the firing rate of a particular unit. In firing rate models, such input is a product of a fixed connection strength and a synaptic gating function. The inter-unit connection strength depends on the numbers and mean-strengths of excitatory and inhibitory connections between the neurons that constitute each unit. The synaptic gating function depends on the mean firing rate of cells in the unit supplying input and on the synaptic dynamics (*e.g.*, Eq. 5.21).

Simulation of a firing rate model entails solving a set of coupled ordinary differential equations just as in prior simulations. The underlying procedure follows two steps. First,

the inputs to the units determine their firing rates. Second, the firing rates of units determine the new inputs of the units to which they are connected. Iteration by repeated cycles of this two-step process reveals the network's behavior in a simulation. If the behavior is a steady state of constant firing rates, or one of regular oscillation, the properties of these states may also be revealed by mathematical analysis of the underlying coupled equations.

The general form of the dynamics of a firing rate model is:

$$\tau_{r_i} \frac{dr_i}{dt} = -r_i + f_i(\{W_{ji}s_j\}) \quad \text{Eq. 6.1}$$

coupled with

$$\tau_{s_i} \frac{ds_i}{dt} = -s_i + F(r_i), \quad \text{Eq. 6.2}$$

where W_{ji} indicates the strength of connection from unit j to unit i , s_i denotes the fraction of downstream synaptic channels open as a function of incoming firing rate, and r_i is the mean firing rate of cells in each unit labeled i .

Writing the firing rate function as $f_i(\{W_{ji}s_j\})$, means the firing rate of unit i depends on each other unit, j , in a manner determined by the connection strength, W_{ji} , from unit j to unit i , multiplied by the synaptic gating variable, s_j , of unit j . The general form allows for different types of unit to have different types of effect on each other—for example input via some synaptic connections could add to the firing rate, whereas input via others could cause a multiplicative or divisive change in firing rate.

However, most often, and in this course, the inputs are just summed so that

$$f_i(\{W_{ji}s_j\}) = f_i\left(\sum_j W_{ji}s_j\right) = f_i(S_i). \quad \text{Eq. 6.3}$$

Eq. 6.3 means that the firing rate depends on the total synaptic input to a unit, $S_i = \sum_j W_{ji}s_j$, which is calculated by adding together the effects of all connected units. Inhibitory connections between units can then just be treated with negative values for the corresponding connection strength, W_{ji} .

The synaptic gating function, $F(r_i)$ in Eq. 6.2, is usually taken as a linearly increasing function of the presynaptic firing rate. This is an approximation, since synaptic gating, being the fraction of synaptic channels open, should saturate (*i.e.*, never surpass a maximum of 1) even if the firing rate were to increase to very high levels. However, in practice, if the firing rate itself saturates at a maximum value, the synaptic input implicitly has a corresponding maximum value too. In this case, we can write $F(r_i) = r_i/r_i^{(max)}$, where $0 \leq F(r_i) \leq 1$, which is a linear dependence with an inherent maximum value.

Finally, if one of the time constants is assumed to be much shorter than the other—for example if synaptic input changes on the timescale of 1ms following a change in presynaptic firing rate, but the rate takes 10ms to respond to changes in synaptic input—then one can approximate the shorter time constant as zero and require the right-hand side of the corresponding equation to evaluate to zero. For example, if $\tau_{s_i} \ll \tau_{r_i}$, then from Eq. 6.2: $-s_i + F(r_i) \approx 0$. This allows us to avoid any simulation of Eq. 6.2 simply by setting $s_i = F(r_i)$ and using this value in Eq. 6.1. If we proceed in this manner and incorporate the

linear approximation for $F(r_i)$, then $s_i = r_i/r_i^{(max)}$ and the set of differential equations to solve in Eq. 6.1 becomes (see Figure 6.2):

$$\tau_{r_i} \frac{dr_i}{dt} = -r_i + f_i \left(\sum_j W_{ji} \frac{r_j}{r_j^{(max)}} \right). \quad \text{Eq. 6.4}$$

In Eq. 6.4 the number of variables that must be simulated is equal to the number of units.

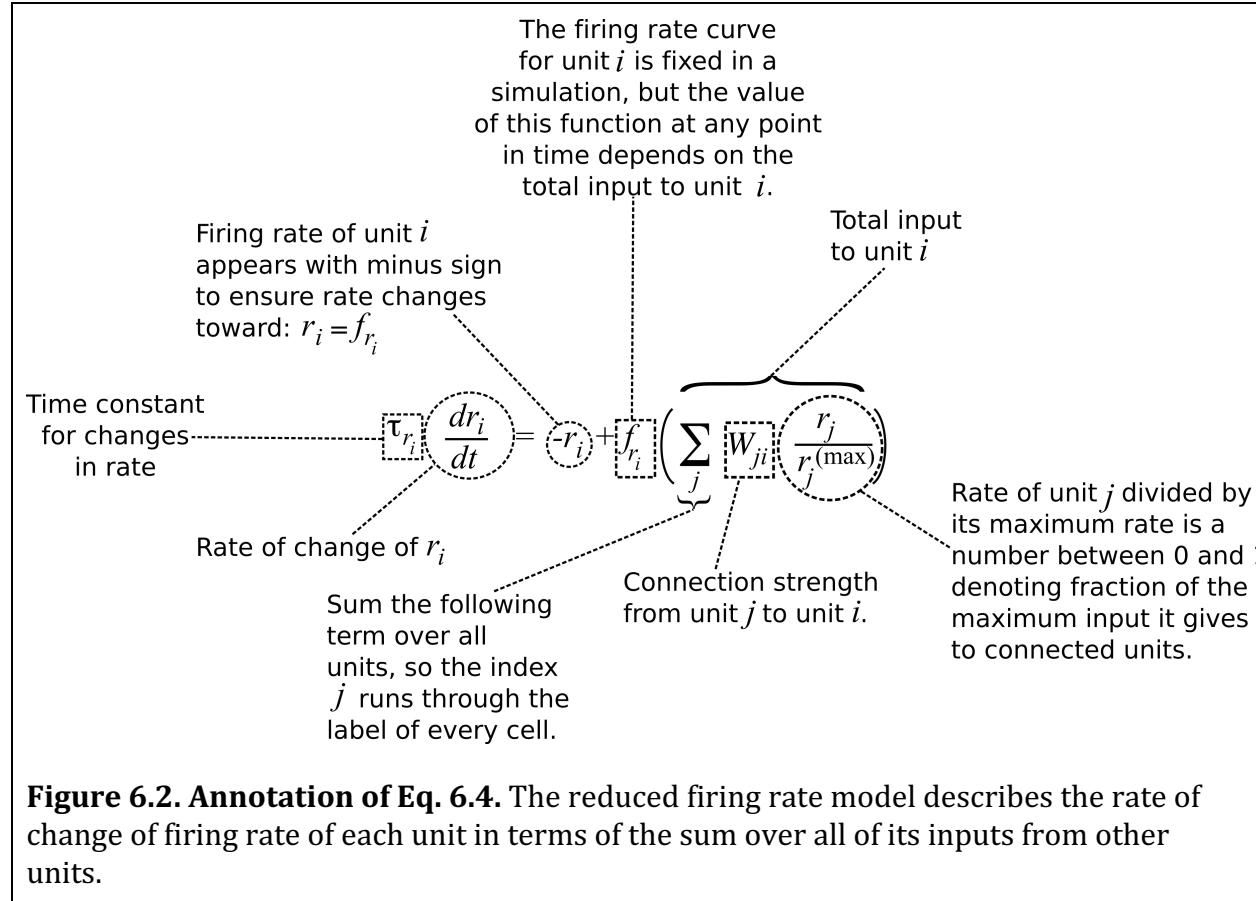


Figure 6.2. Annotation of Eq. 6.4. The reduced firing rate model describes the rate of change of firing rate of each unit in terms of the sum over all of its inputs from other units.

6.2.1. Meaning of a unit and Dale's Principle

Many individual neurons, especially in cortical circuits, have either a predominantly excitatory or predominantly inhibitory effect on all other neurons to which they provide input. Whether the effect is excitatory or inhibitory depends on the dominant type of neurotransmitter they release, which is conserved across a neuron's axonal terminals in nearly all cases according to Dale's Principle (Section 5.1.2). The principle can be extended to state that neurons are either excitatory or inhibitory, a classification which is valid for many circuits. When valid, such separation into two classes requires the connectivity matrix between neurons to contain rows of either all non-negative entries (for excitatory neurons) or all non-positive entries (for inhibitory neurons).

However, such a restriction on the connectivity matrix need never apply to firing-rate models, because each entry in the matrix can represent the net impact of a unit made up of both excitatory and inhibitory neurons on another such unit. For example, if unit A

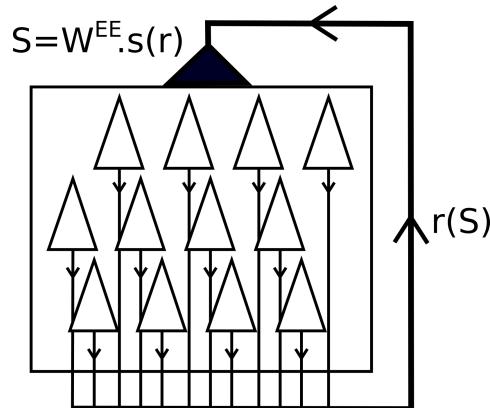


Figure 6.3. Firing-rate model unit with recurrent feedback. A unit of many neurons (triangles) produces spikes at a mean rate, r , which is a function of the total synaptic input, S . The total synaptic input, S , depends on the mean rate, r , via the term $s(r)$, which corresponds to the fraction of open synaptic channels when the presynaptic cells fire at the mean rate, r . The fraction of open channels, $s(r)$, which varies between 0 and 1, is multiplied by the maximum effective feedback strength, W^{EE} , to produce the total synaptic input, S . The superscript ' EE ' in W^{EE} denotes excitatory feedback to excitatory cells in this model.

connects to unit B through a preponderance of excitatory connections to excitatory cells then that inter-unit connection, A-to-B, would be excitatory. Yet unit A could also connect to unit C through a preponderance of inhibitory connections to excitatory cells, in which case the inter-unit connection, A-to-C, would be inhibitory (such an inhibitory inter-unit effect can also arise from excitatory connections to inhibitory cells—see Chapter 6). Thus, if units are comprised of a mixture of excitatory and inhibitory cells then there is no restriction on the signs of connections between units.

6.3. Recurrent feedback and bistability

In this section, we will see how excitatory recurrent feedback can generate bistability. In this case, bistability means that neurons can maintain spiking activity at two distinct firing rates when receiving the same level of input (which is typically zero). The lower firing rate state is either quiescence, the absence of spikes, or activity in the range of 0 to 10Hz corresponding to the spontaneous activity of neurons in the absence of a stimulus. The higher firing rate state—typically in the tens of Hz in models, but lower in empirical data—can be initiated by an excitatory stimulus and persists once the stimulus is removed.

Such a bistable unit acts as a simple memory circuit, since its level of activity is history-dependent. Its firing-rate is high if its most recent input was excitatory (to switch on activity), and its firing-rate is low if its most recent input was inhibitory (to switch off its activity). Amazingly, neurons with similar stimulus-specific sustained responses have been recorded in monkeys during short-term memory tasks¹. Moreover, an antagonist of NMDA receptors (which disrupts their normal function) prevents such sustained responses². These data suggest that small circuits with excitatory feedback are the basis of short-term memory maintenance in the brain.

6.3.1. Bistability from positive feedback

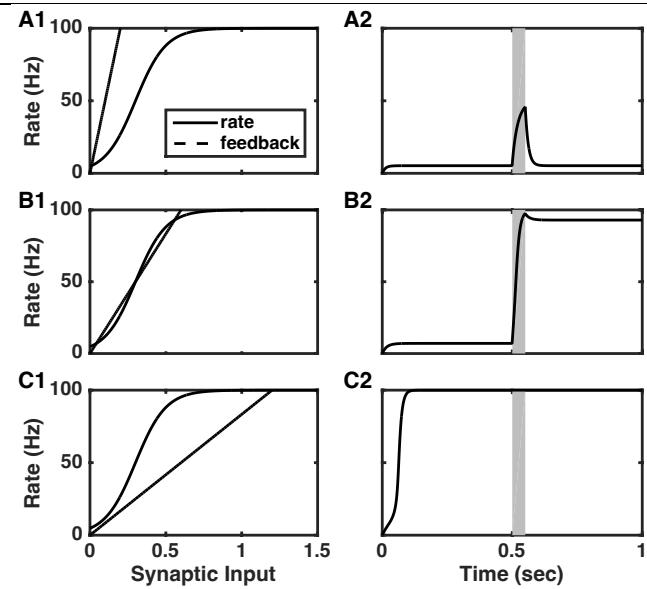


Figure 6.4. Recurrent feedback strength determines the stable states of a circuit. **A1, B1, C1)** Firing rate response of the unit is a sigmoidal function (solid line). For the feedback function (dashed line) the firing rate on the y-axis determines the level of synaptic input (x-axis). The stronger the connection strength, the greater the synaptic input produced by a given firing rate, so the shallower the feedback curve. **A2, B2, C2)** Dynamics of firing rate as a function of time for the models with rates initialized at zero and a pulse of input between 0.5s and 0.55s (gray bar). Note the switch in activity, representing memory of the stimulus in B2. **A)** With weak feedback, $W^{EE} = 0.2$, only a low-rate state is stable. **B)** With medium feedback, $W^{EE} = 0.6$, both a low-rate and a high-rate state are stable so a transient stimulus can switch the system between states. **C)** With strong feedback, $W^{EE} = 1.2$, only a high-rate state is stable. Curves are produced by the online code `bistable1.m`.

Bistability requires a process which, following a small shift of a system away from equilibrium, acts to accelerate the system even further from equilibrium until a tipping point is reached and a return to the prior equilibrium is not possible. For a new equilibrium to be reached—as it must for the system to be bistable rather than unstable—a second process must kick in to limit the overall change. The first, accelerating process is positive feedback, as it increases the rate of change in the same direction as the change. The second, limiting process is negative feedback, as further change decreases the rate of change.

In neural circuits with excitatory feedback (Figure 6.3), we can consider how much extra synaptic input is produced by an initial small change in firing rate. If that extra synaptic input is sufficient to produce a bigger change in firing rate than the initial small change, then even more feedback ensues. Greater firing rates followed by even greater feedback would then follow in a positive feedback loop. The initial firing rate at which such positive feedback dominates is a point of no return.

The amount of positive feedback arising within a circuit depends on the steepness of the firing-rate curve and the amount of extra synaptic input per spike. At low firing rates (such as zero), the firing-rate curve is almost flat, since a small amount of extra current does not generate more spikes when neurons are far below threshold. Therefore, a low firing rate state with minimal feedback can be a stable state in which the positive feedback is too small to cause runaway excitation. However, as the input and the firing rate

increases, firing-rate curves become supralinear (of increasing slope). It is possible for them to reach sufficient steepness that the positive feedback dominates—each additional spike produces enough input current to generate more than one additional spike—and a point of no return is reached. The neurons in the unit then begin to fire ever more rapidly. The firing rates are prevented from rising inexorably by one or more limiting processes, whose impact we will consider in some detail.

Box 6.5. Supralinear function: A function whose gradient increases along the x-axis, also called a convex function.

The simplest of limiting processes is firing rate saturation—neurons and all reasonable models of neurons have a maximum firing rate, on the order of several tens of Hz for excitatory pyramidal cells and on the order of 100-200Hz for many ‘fast-spiking’ inhibitory interneurons. Once the limit is reached, in simple models the firing-rate curve is flat, with spiking activity remaining at its maximum rate. The flaw in such simple models is that empirically spiking activity can cease with excessive input, *e.g.*, due to a loss of sodium deactivation (Section 4.2.3).

Figure 6.4 depicts the behavior of a single unit, i , with excitatory feedback of different strengths, $W_{ii} = W^{EE}$. In the left-hand panels the firing-rate as a function of input is plotted as a solid curve with a sigmoidal shape. In the absence of input the rate is low, the gradient of the curve increases with input at low rates, but decreases at high rates until it saturates at the maximum rate (100Hz here). The dashed lines in each of the left panels indicate the recurrent excitatory feedback. These lines are plotted using a flipped orientation of the axes, because the synaptic input plotted on the x-axis is due to recurrent feedback, so depends on the firing rate, which is plotted on the y-axis. When the firing rate is zero there is no synaptic input from recurrent feedback, so the dashed lines pass through the origin. If feedback is weak (panel A1) the synaptic input is low even when the firing rate is maximal. At higher feedback strength (panels B1 and C1) a given firing rate produces greater synaptic input, so the lines are less steep.

Box 6.6. Fixed point: A set of values for all of the variables of a system, at which their time-derivatives (their rates of change with time) are all zero, so the system does not change when at the fixed point.

The points of intersection of the two curves (dashed line and solid curve) in panels A1, B1, and C1, of Figure 6.4, indicate fixed points of the system, meaning the firing rate and the synaptic input do not change if they are set to those values (for more on fixed points see Chapter 7). This can be seen in the corresponding right-hand panels (A2, B2, and C2) where the firing rates are stationary when they reach their values at those intersections in the left-hand panels.

In panel A2, the temporary stimulus between 0.5s and 0.55s causes a temporary increase in firing rate. However, once the stimulus is removed, the firing-rate returns to its prior stable value. We can understand this from the curves in Figure 6.4, panel A1, where we see from the dashed line that at the higher firing rate the synaptic input is relatively low. We then see from the solid line of panel A1 that the low synaptic input generates a low firing rate, which leads to even lower synaptic input. Eventually, at the point of intersection, the small amount of synaptic input produced is exactly the amount needed to

maintain the low firing rate and the system does not change further. The single stable state, with low firing rate, is reached.

Box 6.7. Stable fixed point: A fixed point, toward which the system returns following any small changes in its variables.

Box 6.8. Unstable fixed point: A fixed point away from which the system changes following the slightest change in one or more of its variables. An unstable fixed point can be the point of no return when a system changes from one stable fixed point to another.

In the bistable system of Figure 6.4, panel B1, the two curves intersect at three points. Therefore, the system has three fixed points, that is, three distinct sets of firing rate and synaptic input at which their rates of change are zero. However, the system only has two stable states because the intermediate fixed point is an unstable one. An unstable fixed point is akin to a pencil standing vertically on its tip—while theoretically it could be aligned so that its center of mass is exactly above the tip, in which case it has no preferred direction in which to fall, in practice such perfect alignment is impossible and random air movements or vibrations would disturb any perfect alignment were it possible. Such unstable fixed points would not be observed in a simulation (as in panel B2) unless the simulation were started with the exact values of the fixed point and were noise-free.

That the intermediate fixed point is unstable can be understood from panel B1, by considering the system's response to small deviations above or below the fixed point. Following any deviation from a curve, the system will move back toward the two curves, with the firing-rate curve (solid) indicating the component of change along the y-axis, and the feedback curve (dashed) indicating the component of change along the x-axis. That is, a point on the figure representing the instantaneous firing rate and synaptic input would move vertically toward the solid curve and horizontally toward the dashed curve. Wherever the firing-rate curve is above (greater y-coordinate) and the feedback curve is to the right (greater x-coordinate) the net effect is an increase in both firing rate and feedback until the curves cross again. Wherever the firing rate curve is below (smaller y-coordinate) and the feedback curve is to the left (smaller x-coordinate) the net effect is a decrease in both firing rate and feedback until the curves cross again. These two effects combine to mean that the state of the system moves away from the intermediate fixed point following any small deviation.

The intermediate fixed point is, in fact, the point of no return mentioned earlier. If the system, having been in the low-firing rate state, is provided enough stimulus to pass the intermediate fixed point then it will head to the high-firing rate state, even after any stimulus is removed.

6.3.2. Limiting the maximum firing rate reached

In the previous section a leveling off of the firing-rate curve at its maximum value provided the only limit on run-away activity. Indeed, as can be seen from Figure 6.4, panel B1, if the linear feedback curve intersects the firing-rate curve at three points, the high-rate point is inevitably close to the unit's maximum rate (of 100Hz in this example). However, *in vivo*, firing rates are observed to switch by smaller amounts (10-30Hz), even though the individual neurons can reach rates on the order of 100Hz given sufficient input. One

possible cause of a reduction in the system's maximum persistent rate is a saturation of synaptic feedback.

One limit on synaptic feedback is the number of receptors available to be bound by neurotransmitter at the surface of the postsynaptic neuron. Once all of the receptors at feedback synapses are bound by ligand and their corresponding channels are opened, there is no further possible increase in synaptic input due to feedback. Therefore, once neurons are firing at sufficient rate to keep the postsynaptic receptors bound, no further feedback input is possible, so the positive feedback remains constant and the rate does not increase any further. If such feedback is predominantly mediated via receptors from which neurotransmitter is slow to dissociate, or whose associated ion channels are slow to close once opened, then the presynaptic rate at which feedback begins to saturate can be low. Thus, NMDA receptors, with a slow time-constant of 50-75ms (at *in vivo* temperatures) for unbinding of glutamate neurotransmitter, could play an important role in stabilizing persistent activity at low rates³ (Figure 6.5).

A second limit on synaptic feedback occurs at the axonal terminals of the presynaptic neuron. The limited number of release-ready vesicles that causes synaptic depression (Section 5.3.1) also limits the amount of sustained synaptic input to postsynaptic cells. Once a presynaptic cell is producing action potentials at a rate quicker than vesicles can "dock" in the membrane or be replenished, any further increase in presynaptic firing rate does not increase the rate of release of vesicles with neurotransmitter, so does not produce more synaptic input to postsynaptic cells (Eq. 5.27 and below). The firing rate at which synaptic depression begins to limit the synaptic feedback is proportional to the rate of replenishment of release-ready vesicles, to the total number of vesicles that can be release-ready at one time, and to the probability, for each release-ready vesicle, that an action potential causes its release.

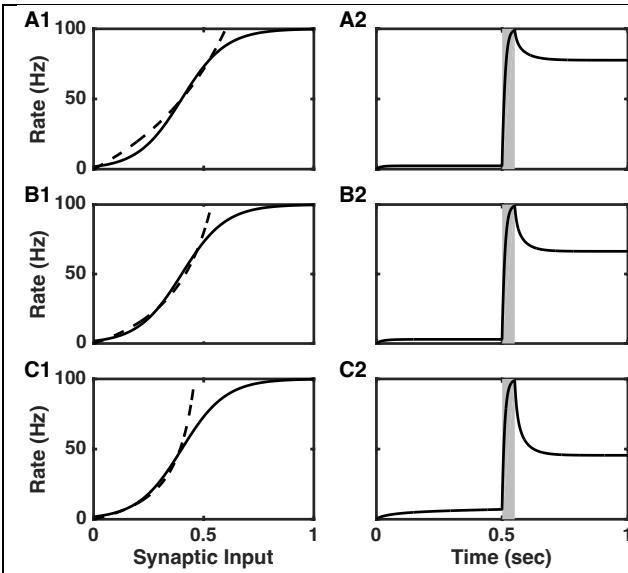


Figure 6.5. Persistent activity at lower rates via synaptic saturation. When synaptic feedback saturates (dashed curves become vertical at high rates) bistability is possible without the more active state's firing rate being so unrealistically high. Synaptic saturation occurs at lower rates (bottom) if the fraction of receptors bound per spike, α , is high rather than low. **A1-A2**) $\alpha = 0.2$, $W^{EE} = 1.2$. **B1-B2**) $\alpha = 0.5$, $W^{EE} = 0.75$. **C1-C2**) $\alpha = 1$, $W^{EE} = 0.55$. For all curves, $p_r = 1$ and $\tau_s = 50\text{ms}$. Curves are produced by the online code `bistable2.m` and are in the same format with the same stimuli as Figure 6.4.

In the following section, we will consider how such synaptic saturation can be incorporated into a firing-rate model, so that we can simulate the processes and assess their impact on the rate of sustained activity in Tutorial 6.1.

6.3.3. Dynamics of synaptic response

In the Appendix of Chapter 5 (Section 5.10.1), we showed that the synaptic time constant, τ_s , limits the mean synaptic response, $\langle s \rangle$, due to a presynaptic neuron firing spikes as a Poisson process at rate r via

$$\langle s \rangle = \frac{\alpha p_r r \tau_s}{1 + \alpha p_r r \tau_s}, \quad \text{Eq. 6.5}$$

where p_r is the release probability of each docked vesicle per spike, and α is the maximum fraction of postsynaptic receptors bound by neurotransmitter when all presynaptic docked vesicles are released. In a firing rate model that omits the presynaptic dynamics of depression and facilitation, the mean synaptic response at steady state, $\langle s \rangle$, can be used as $F(r)$ (Eq. 6.2) when the synaptic gating variable is set instantaneously as $s = F(r)$.

Alternatively, if the dynamics of s are simulated, the corresponding equation (whose steady state matches Eq. 6.5) is:

$$\frac{ds}{dt} = -\frac{s}{\tau_s} + \alpha p_r r (1 - s) \quad \text{Eq. 6.6}$$

where the first term on the right is the decay of synaptic conductance due to unbinding of neurotransmitter, and the second term is the amount of increase in s per spike multiplied by the mean rate of spike arrival.

The effective time constant for changes in s following changes in r becomes $\tau_{eff}(r) = \tau_s / (1 + \alpha p_r r \tau_s)$. In the absence of spikes ($r = 0$), the effective time constant approaches τ_s , the timescale for decay of synaptic input to zero. When the firing rate is high, the effective time constant is smaller, allowing s to increase more quickly in proportion to the rate of spike production.

In a circuit with excitatory feedback, we can re-plot the two steady state curves, $r(S)$ and $s(r)$, where $S = W^{EE} s$ and see that the rate of sustained activity is reduced if τ_s is relatively large (Figure 6.5).

6.3.4. Dynamics of synaptic depression and facilitation

Synaptic depression and facilitation can also be incorporated in firing-rate models. In the Appendix of Chapter 5 (Sections 5.10.2-3) we showed how the steady state values for the depression variable, D , and the facilitation variable, F , depend on presynaptic firing rate for a Poisson process. These variables modify respectively the amplitude of synaptic response, $\alpha = \alpha_0 D$, and the individual vesicle release probability, $p_r = p_0 F$. Their dynamics can be modeled according to the equations (cf. Eqs 5.9-5.10)

$$\frac{dD}{dt} = \frac{1 - D}{\tau_D} - p_r D r \quad \text{Eq. 6.7}$$

and

$$\frac{dF}{dt} = \frac{1 - F}{\tau_F} + f_F (F_{max} - F) r. \quad \text{Eq. 6.8}$$

These dynamical equations have stable steady states at fixed rate, r , that match the mean response to spikes arriving as a Poisson process of rate r (cf. Eq. 5.27):

$$D(r) = \frac{1}{1 + p_r r \tau_D} \quad \text{Eq. 6.9}$$

and (cf. Eq. 5.30)

$$F(r) = 1 + \frac{(F_{max} - 1)f_{fac}r\tau_F}{1 + f_{fac}r\tau_F}. \quad \text{Eq. 6.10}$$

Synaptic depression causes the mean steady state synaptic response to saturate at high presynaptic firing rates in a manner mathematically equivalent to synaptic receptor saturation. However, the rate at which such saturation becomes important is on the order of $1/p_r\tau_D$, which is typically lower than $1/\alpha p_r\tau_s$, the corresponding rate for saturation of the postsynaptic response. While such saturating feedback can limit the firing rate of persistent activity states to the observed lower levels, the slow time constant of synaptic depression tends to destabilize the persistent state and can generate oscillations (Figure 6.6A). We will investigate this process in Tutorial 6.1.

Synaptic facilitation increases the effective synaptic strength when firing rate increases, so it can stabilize a bistable system. The additional stability arises because the low firing rate state has effectively weaker synapses—making it harder for firing rate to increase—while the high firing rate state has effectively stronger synapses—making it harder for firing rate to decrease. Because of this, synaptic facilitation allows the spontaneous, low-firing rate state of a bistable system to have activity that is significantly above zero (Figure 6.6B) without the excitatory feedback causing a runaway increase in firing rate.

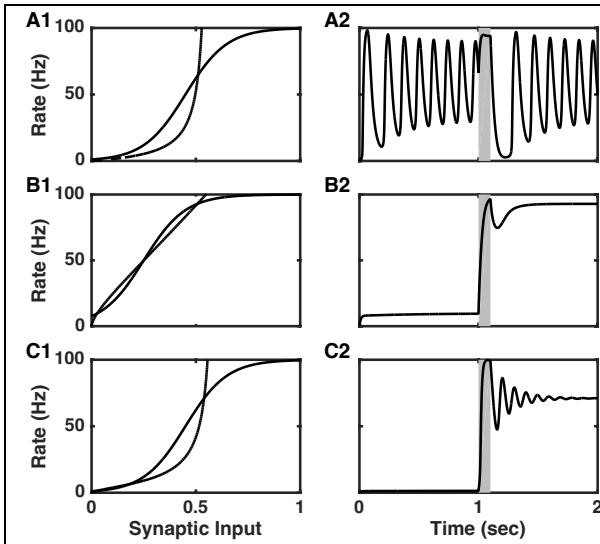


Figure 6.6. Activity states in a circuit with excitatory feedback and short-term synaptic dynamics. A) Synaptic depression alone can destabilize the persistent activity state and leads to slow oscillations. **B)** Synaptic facilitation enhances the stability of a bistable system, so the low-rate state can have activity above 10Hz. **C)** Facilitation and depression combine to generate a bistable system even when synaptic saturation is absent. Format of the figure is the same as that of Figure 6.3 and with the same stimulus. Parameters and codes that produced these figures are available online as [bistable_fac_dep.m](#).

6.3.5. Integration and parametric memory

In bistable systems, the feedback curve crossed the firing rate curve at three points, two of which indicated the firing rate and synaptic input of the two stable states (Figure 6.4B). In another scenario, rather than crossing, the two curves could lie on top of each other for a range of firing rates. In this case, any points in the range that lies on both curves correspond to values of the firing rate and synaptic input that the system would remain at if set to have those values. Moreover, if the curves overlap or cross nowhere else, then the system's variables move to a point where the curves overlap. In this situation, the system possesses a line attractor (also called a continuous attractor or a marginal state).

Box 6.9. Line attractor (also called a continuous attractor, or marginal state): a continuous range of fixed points that can be plotted as a curve or line. The system returns to the line, but not generally to the original point on the line following any small deviation.

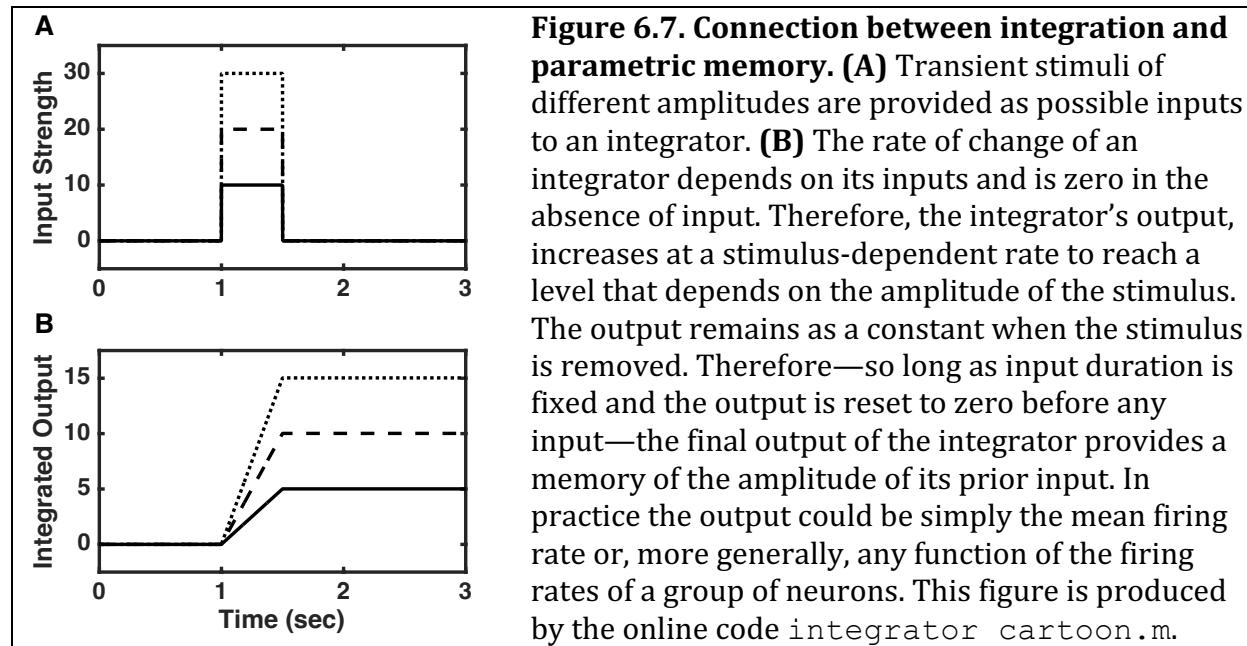
The line attractor is said to possess marginal stability, because following perturbations of the system the firing rate and synaptic input return to a point on the line, suggesting stability, but they may not return to the original point on the line. Movement along the line is easy.

Such easy movement can be beneficial, as the system can respond to—even sum up, or integrate—a series of small inputs, each of which kick the system’s state a little further along the line. Indeed, if a property of the circuit, such as its final firing rate, depends monotonically on the amount of external excitatory input the unit receives, then the circuit is acting as an integrator. That is, a line attractor can act as an integrator⁴.

Box 6.10. Integrator: A circuit which integrates its inputs in the mathematical sense, summing the values of discrete inputs.

Box 6.11. Parametric memory: Memory of a continuous quantity, or parameter of a stimulus.

An integrator provides memory of a continuous quantity—namely the integral over time of the stimulus—because when the stimulus is removed, the integral over time no longer changes. Therefore, the memory of the integrated value at stimulus offset remains (Figure 6.7). We shall investigate such integration in an excitatory feedback circuit in Tutorial 6.1.



Box 6.12. Self-excitation: Excitatory recurrent feedback to a group of cells or a unit.

Box 6.13. Cross-inhibition: An inhibitory interaction between two cell-groups or units that respond to different or opposite stimuli.

Integration has received considerable attention in computational neuroscience because integrators appear to be necessary in some small circuits (*e.g.* to maintain a fixed eye position in some species⁵), they play a large role in the field of decision-making^{6,7}, and they appear to underlie the ability of animals to retain short-term memory of continuous parameters^{8,9}. However, a circuit based on self-excitation of a single unit (or cross-inhibition between two units) requires fine-tuning in order for firing-rate and synaptic-feedback curves to align to an even greater extent than that shown in Figure 6.5.C1. Whether neural circuits possess the necessary mechanisms for the fine-tuning needed to produce a line attractor remains an open question.

6.4. Tutorial 6.1. Bistability and oscillations in a firing-rate model with feedback.

Neuroscience Goal: learn how synaptic dynamics impact the behavior of a bistable circuit with excitatory feedback; see how an approximate line attractor can behave as an integrator.
Computational Goal: practice at simulating firing rate models.

In this tutorial, we will explore the conditions for bistability and the effects of short-term synaptic plasticity in a firing-rate model with excitatory feedback. The three, coupled variables for the single unit will be firing rate, r , depression variable, D , and synaptic gating variable, s . The system of equations that you will simulate are an extension of Eq. 6.1:

$$\begin{aligned} \tau_r \frac{dr}{dt} &= -r + f(S) \\ \frac{dD}{dt} &= \frac{1-D}{\tau_D} - p_r D r \\ \frac{ds}{dt} &= -\frac{s}{\tau_s} + \alpha_0 D p_r r \cdot (1-s) \end{aligned} \quad \text{Eq. 6.11}$$

with the additional condition $r \geq 0$ and where the total synaptic input, S , is given by $S = W^{EE}s + s_{in}$. The firing rate curve, $f(S)$, in the top equation corresponds to the steady state firing rate as a function of total synaptic input. It is given as a power law with saturation:

$$f(S) = r_0 + r_{max} \frac{S^x}{S^x + \sigma^x} \text{ if } S > 0, \\ \text{and } f(S) = r_0 \text{ if } S \leq 0.$$

The exponent is $x = 1.2$ and the rate with no input is $r_0 = 0.1\text{Hz}$ in the first simulation (but r_0 will vary thereafter). The maximum possible firing rate is $r_0 + r_{max}$, with $r_{max} = 100\text{Hz}$. The input needed to reach the midpoint of the firing-rate range is given by $\sigma = 0.5$, and the time constant for changes in firing rate is $\tau_r = 10\text{ms}$. Other parameters in the model are α_0 , p_r , W^{EE} , and τ_D , and τ_s , which all vary from question to question.

For each question with a distinct parameter set, you will plot the steady state response functions in the absence of input: $f(W^{EE}s)$ as a function of s , and $s(r) = \frac{\alpha_0 D p_r r \tau_s}{1 + \alpha_0 D p_r r \tau_s}$ as a function of r . In the calculation of $s(r)$, you will need to use the steady state response of the depression variable, $D(r) = \frac{1}{1 + p_r r \tau_D}$. After plotting the steady state values as two distinct curves (whose intersections are fixed points of the system) you will then simulate the system and its response to transient inputs, s_{in} .

- 1) Fix the depression variable so that it remains unchanged from 1 and set the other parameters as follows: $\alpha_0 = 0.5$, $W^{EE} = 8$, $p_r = 1$, and $\tau_s = 2\text{ms}$.
- Plot $f(W^{EE}s)$ for a range of values of s from 0 to 1. On the same figure, for a range of values of r from 0 to r_{max} plot $s(r)$, with r on the y-axis and s on the x-axis.
 - Simulate the full set of three coupled ODEs (Eq. 6.11) for 20 seconds, with initial conditions $s = 0$ and $r = 0$. Add a temporary input of strength $s_{in} = 0.05$ for a duration of 50ms beginning at a time of 10 seconds. Plot the resulting firing rate versus time. Check that the stable rates correspond to the crossing points of the two curves in a).
- 2) Include synaptic depression with $\tau_D = 250\text{ms}$ and $p_r = 0.2$; keep $r_0 = 0.1\text{Hz}$ and $\alpha_0 = 0.5$; and increase W^{EE} to 60 (to compensate for the reduced release probability and for the weakening of synapses due to depression). Repeat parts 1a) and 1b), except set the temporary input to be of strength $s_{in} = 0.002$ and of duration 2 seconds. Comment on any behaviors you see, in particular any difference from 1).
- 3) Increase p_r to 0.5, reduce W^{EE} to 35, and set r_0 to -0.1Hz . Reset the temporary input to $s_{in} = 0.05$ for a duration of 50ms. Keeping all other parameters the same as in 2), repeat parts a) and b) and comment on your results.
- Set the initial value of the firing rate to 9Hz and the initial values of all other variables to their steady state at that rate (the steady state can be calculated rather than simulated). Repeat the simulation of part b) with the new initial conditions. Comment on your results.
- 4) Oscillations are enhanced with slow negative feedback. To see this, repeat 3a and 3b with the following parameters: $\tau_D = 125\text{ms}$ (half of its prior value); $\alpha_0 = 0.25$ (half of its prior value); and $p_r = 1$ (twice its prior value). You should find the steady state curves are identical to those in Question 3). Try to explain why the active state is now stable (oscillations decay) whereas in Question 3 it was unstable (oscillations grow).
- OPTIONAL
- 5) If the feedback curve overlaps with the firing-rate curve then the system can be almost stable at a lot of different firing rates. This allows for the firing rate to slowly accumulate when a stimulus is applied and for the final rate to be an approximately continuous readout of the time-integrated stimulus. You will omit synaptic depression, fixing $D = 1$ as in Question 1, and set the following parameters: $\tau_s = 50\text{ms}$, $W^{EE} = 1.6$, $\alpha_0 = 1$, $p_r = 0.1$, and $r_0 = 0.5\text{Hz}$. Repeat 1a) with these parameters.
- Repeat 1b) with these parameters using an input of strength $s_{in} = 0.001$ and a duration of 6 seconds.
 - Repeat 5b) except with a series of 30 inputs, each of amplitude $s_{in} = 0.005$ of duration 100ms, starting each input every 500ms. Comment on the behavior of the circuit.
 - Suppose that following an unknown number of such stimuli, you are given the firing rate of the unit 500ms after the onset of the last stimulus. Estimate how many distinct numbers of stimuli you could distinguish using the firing rate at this time.

6.5. Decision-making circuits

The importance of decision-making for us as humans is widely acknowledged because our decisions have consequences that shape our lives. However, it can be hard to define decision-making from a biological perspective, when we study a system from the outside and just consider how the functional response of a system depends on its inputs. Biologists can reduce the meaning of the word “decision-making” to include random events pertaining to inanimate matter. For example, it is common to suggest that a bacterium “decides” which protein to express, or to name the genetic event of a cell becoming one type or another a “cell fate decision”. Given such examples, one could equally ask whether a storm “decides” which path to take.

The difficulty lies in that decisions do appear to include a random component—when an animal is given identical choices, the time taken to make a choice and the actual choice made vary from trial to trial. The deterministic components—such as the level of an ongoing stimulus, any prior belief based on the animal’s experiences, the value of options based on perhaps genetic disposition as well as history—conspire together to shape the probability of different choices being made, while allowing (it appears) each individual choice to be undetermined.

In general usage, a decision suggests a point of no return—a commitment to a particular course of action—so has a lot in common with action selection. Often the selection is between discrete alternatives: Shall I eat one fruit or the other? Shall I accept a new job offer or stay in my current job? Shall I buy this house or keep looking? Other decisions are from a continuous range of possibilities: In which direction shall I point? How fast shall I drive? How long shall I wait?

Models of decision-making address three distinct issues. First, theorists produce models so as to account for the experimental data. In this case, the data include behavioral measures, such as the proportion of different responses and the times taken to make those responses¹⁰. They also include electrophysiological measures, in particular the dynamics of neural activity during decisions^{11,12}. A good model accounts for the distribution of these features over many trials and for how these distributions change as the stimulus changes.

Second, theorists analyze different models to ascertain which would produce more optimal behavior^{13,14}. Addressing the second issue is valuable, because it can provide an explanation for why animals would make decisions a certain way. It also allows us to test the degree of optimality achieved by any animal in any task.

Third, theorists produce models constrained by the properties of neurons and neural circuits engaged in decision-making¹¹. While psychologists may be content with a verbal or purely mathematical model, which accurately describes a behavior, neuroscientists are not satisfied until they understand the mechanisms by which the brain generates the behavior. Theorists can then proffer explanations for any suboptimal behavior observed in terms, perhaps, of the underlying biological constraints¹⁵.

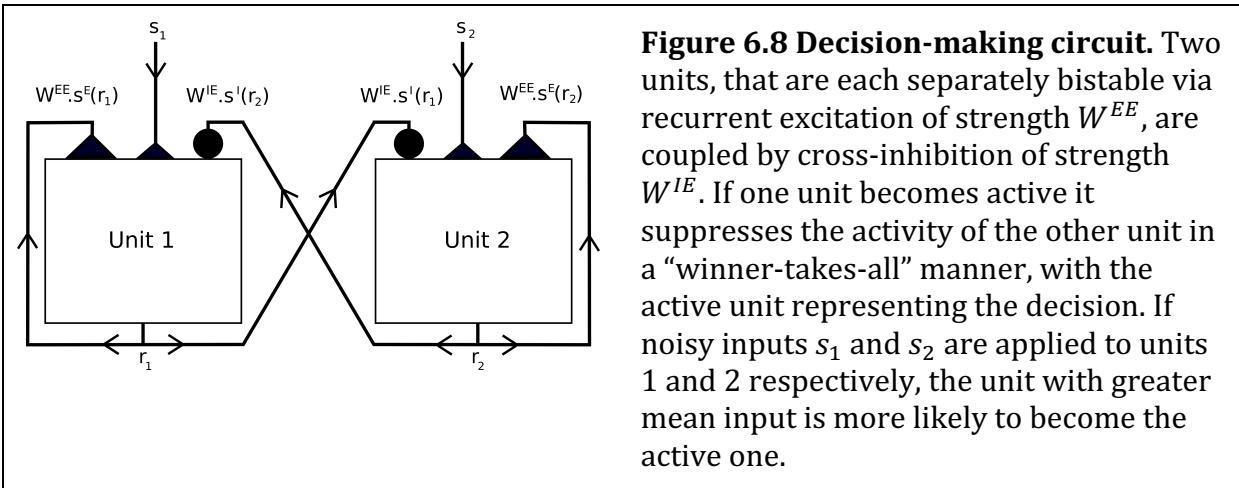


Figure 6.8 Decision-making circuit. Two units, that are each separately bistable via recurrent excitation of strength W^{EE} , are coupled by cross-inhibition of strength W^{IE} . If one unit becomes active it suppresses the activity of the other unit in a “winner-takes-all” manner, with the active unit representing the decision. If noisy inputs s_1 and s_2 are applied to units 1 and 2 respectively, the unit with greater mean input is more likely to become the active one.

Perceptual decisions are ones that require identification of an ambiguous stimulus—the action is then completely determined by a previous learned rule that matches stimulus to action. Historically, one of the most common stimuli used has been a display of randomly flickering and moving dots on a screen with a net bias or drift in their motion toward one direction over another¹⁶. The animal—in this task, typically monkey or human—indicates the direction of motion it perceives. Often only two choices are available, in which case the design is known as a two-alternative forced choice task.

The relative simplicity of the task, whose difficulty is easily titrated by adjusting the coherence of the dot motion, has produced a wealth of behavioral and electrophysiological data. The distribution of response times for correct choices and errors, as well as their relative probability, was first described by mathematical models such as the drift diffusion model, based on integration of noisy evidence to a threshold¹⁷. In these models the evidence refers to the difference between two stimuli, one representing rightward motion, the other representing leftward motion. The threshold corresponds to a point of no return at which the decision is made. Such a model is optimal in many conditions.

Box 6.14. Integration of evidence: The optimal method of accumulating successive samples of a stimulus such that information from each sample is equally weighted.

In this section we shall consider a circuit (Figure 6.8), based on the Wang model¹¹, which accounts for much of the behavior and neural activity observed in perceptual decisions between two alternatives, and is constrained by the properties of neural circuits. The circuit can be extended trivially to model more than two alternatives when necessary. The Wang model showed how two groups of neurons corresponding to the two alternatives could both reproduce the behavior and the ramping activity of the “winning” group of neurons using a “winner-takes-all” circuit (see Figure 6.9).

Box 6.15: Winner-takes-all circuit: A circuit with multiple distinct cell-groups or units, within which only one unit can sustain high activity, because in so doing it suppresses activity of other units through cross-inhibition.

The winner-takes-all circuit (Figure 6.8) is based on a combination of self-excitation and cross-inhibition. The self-excitation is sufficient that each unit alone would be bistable (as in Figures 6.4B and 6.5), able to remain in a state of low spontaneous activity or, once

excited, to persist in a state of high activity. The cross-inhibition (*i.e.*, reciprocal inhibition) between the two units is strong enough to ensure that if one of the units is in the active state the other cannot be highly active. Thus cross-inhibition ensures the processing of the circuit is competitive and can only produce a single winner, generating a binary outcome when two units are involved.

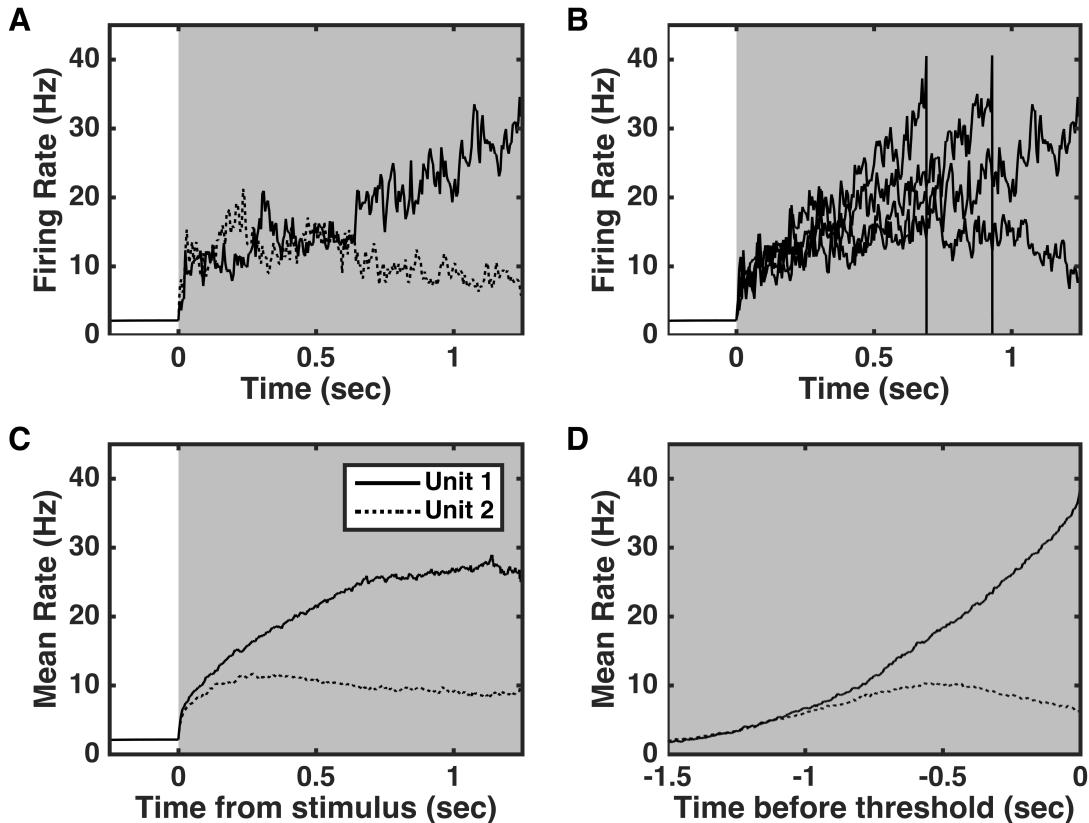


Figure 6.9. Dynamics of a decision-making circuit. **A.** Upon stimulus delivery, the firing rate of Unit 1 (solid) rises gradually over the course of almost a second until the threshold rate of 40Hz is reached. The firing rate of Unit 2 (dotted) rises initially, but is then suppressed by inhibition from Unit 1. The gray shaded region indicates the time of input, during which Unit 1 receives slightly greater excitatory input than Unit 2. **B.** The firing rate of Unit 1 on four successive trials demonstrates considerable variability in the time to reach threshold. **C.** The mean rate as a function of time from onset of the stimulus. **D.** The mean rate as a function of time aligned to the threshold-crossing point—the curve of Unit 1 is steeper, and the suppression of the rate of Unit 2 is clearer, than in C (dotted). These figures are produced using the circuitry of Figure 6.8, by the online code “decision making dynamics figure.m”.

6.5.1. Decisions by integration of evidence

During perceptual decisions, neural activity in parietal cortex appears to ramp up gradually over a time period that can be on the order of a second¹². Achieving such a slow time-course of variation in a circuit comprised of neurons and synapses with time constants on the order of a few milliseconds requires fine-tuning (as with the integrator in Tutorial 6.1, Q.5). The degree of fine-tuning is reduced if feedback within the circuit is predominantly

through NMDA receptors with a long time constant (50ms is a reasonable value *in vivo*). The necessary degree of fine-tuning can be understood from a linear firing-rate model, such as the leaky-competing accumulator model (LCA¹⁸). The LCA is based on a circuit like that of Figure 6.8, but with the following simplified dynamics for the two competing groups:

$$\tau \frac{dr_1}{dt} = -r_1 + W_s r_1 + W_x r_2 + i_1 \quad \text{Eq. 6.12a}$$

$$\tau \frac{dr_2}{dt} = -r_2 + W_s r_2 + W_x r_1 + i_2, \quad \text{Eq. 6.12b}$$

where W_s is a positive parameter representing self-excitation and W_x is a negative parameter representing cross-inhibition. If we define the difference between inputs as $i_d = i_1 - i_2$ and difference between rates as $r_d = r_1 - r_2$ then subtracting Eq. 6.12b from Eq. 6.12a leads to:

$$\tau \frac{dr_d}{dt} = -r_d + (W_s - W_x)r_d + i_d. \quad \text{Eq. 6.13}$$

This can be rearranged to give

$$\frac{dr_d}{dt} = -\frac{r_d}{\tau_{eff}} + \frac{i_d}{\tau} \quad \text{Eq. 6.14}$$

where the effective time constant, $\tau_{eff} = \frac{\tau}{1-(W_s-W_x)}$, determines the timescale that the firing rate decays to zero in the absence of input. The effective time constant is a circuit property, longer than the neural time constant, τ , because $W_s > W_x$ in these circuits.

Notice that when $W_s - W_x = 1$ the effective time constant is infinite and the system is a perfect integrator—*i. e.*, r_d is proportional to the integral over time of i_d . However, in the absence of such perfect tuning, the system's time constant, τ_{eff} , is extended by a factor proportional to the closeness of $W_s - W_x$ to 1. For example, when $W_s - W_x = 0.95$ then the circuit's time constant, τ_{eff} , is 20 times greater than τ , the longest cellular time constant¹⁹. So, with fine-tuning at this 5% level (0.95 being 5% away from a perfect 1) the system's time constant would be 1 second if feedback were via NMDA receptors with a time constant of 50ms. This would set 1 second as the timescale over which activity leaks away and over which any prior information about a stimulus encoded in the circuit's activity is forgotten.

6.5.2. Decision-making performance

Decision-making performance can be judged by different criteria, depending on the details of the task. If the stimulus duration is fixed (forced response) then the fraction of correct responses is the appropriate metric. If the subject chooses when to respond (free response) then performance can be judged via the mean time taken per decision when a required fraction of correct responses is produced. Alternatively, when a correct response yields more reward for an animal than an incorrect one, then the rate of reward is a suitable metric.

Box 6.16. Forced response paradigm: A study in which the time of a response is fixed.

Box 6.17. Free response paradigm: A study in which the subject is free to make a response whenever ready.

The relative cost of an incorrect response versus a slower correct response determines whether faster decisions with more errors or slower decisions with fewer errors are optimal. When the goal is to maximize reward rate, an experimenter can adjust the inter-trial intervals to manipulate a subject's speed or accuracy of responses. Slower, more accurate responses are favored if the delay between trials is lengthened.

When simulating a model neural circuit, a correct decision is indicated when the activity of the group of neurons receiving greater mean stimulus is above the activity of the other group of neurons at the decision time. In the free response paradigm, the decision time is determined by the activity of either group of neurons reaching a response threshold.

For any given model circuit, the response threshold can be varied to determine the circuit's optimal performance. If the threshold is raised then decisions are slower and more accurate. They are slower because it takes activity more time to reach threshold. They are more accurate, because noise accumulates as the square-root of time, while the signal accumulates linearly with time in an integrator—so given sufficient time the signal can dominate over the noise.

Box 6.18. Speed-accuracy tradeoff: The common situation in which faster responses are less accurate, while slower responses are more accurate.

In a neural circuit a change in response threshold could be achieved by modulation of synaptic strengths at the outputs of the decision-making circuit such that a higher level of activity in the circuit is required to produce a response. As an alternative to altering the threshold, the inputs could be scaled to achieve the same speed-accuracy tradeoff. Such an example can be seen in Figure 6.10 by comparing the solid curves with the dashed curves. For an integrator, a scaling down of its inputs acts exactly like a scaling up of its threshold, assuming both the stimulus and noise terms to be integrated scale identically.

Regardless of whether scaling is at the inputs or outputs of a circuit, the ability to modulate the strength of synaptic conductance allows an animal to adapt to circumstances and produce rapid inaccurate decisions or slower more accurate ones as desired. The neuromodulator norepinephrine has been proposed to serve this purpose^{20,21}.

6.5.3. Decisions as state-transitions

An alternative to the gradual ramping of neural activity arising from integration is the possibility that neural activity changes by abrupt transitions between different stable states. The abruptness might correspond to the suddenness with which percepts can change and ideas can “pop” into our heads. If the times of such transitions vary from trial to trial, such “jumping” behavior could reproduce the observed gradual ramping once neural activity is averaged over many trials.

Box 6.19. State-transition: A rapid change in neural activity from one state, such as one stable fixed point, to another.

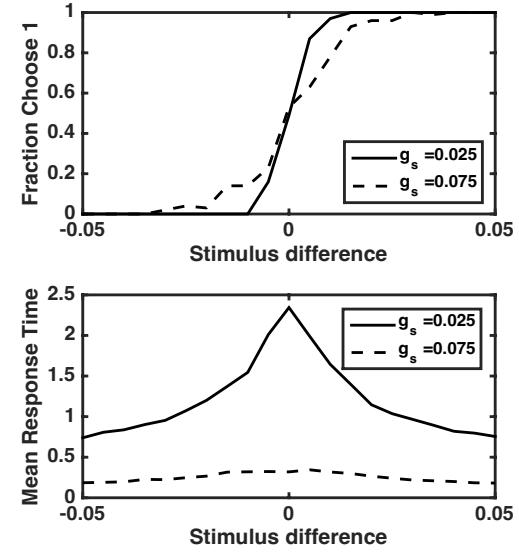


Figure 6.10. The speed-accuracy tradeoff. **A.** The fraction of trials in which Unit 1 reaches threshold first increases in an approximately sigmoidal manner from 0 to 1, as the stimulus difference, $s_1 - s_2$, increases. With reduced input conductance (solid line versus dashed line) a smaller stimulus difference is required for accurate responses. **B.** The cost of improved accuracy is a longer time to reach threshold when input conductance is weaker (solid line versus dashed line). Responses are slowest when the stimuli are most equal. These figures are produced using the circuitry of Figure 6.8, by the online code “decision_making_vary_g_stim.m”.

Transitions between stable states depend on noise fluctuations, so as the level of noise increases the time to make a transition decreases, corresponding to faster decisions. Random noise can enhance the performance of such a decision-making circuit, because the speeding of response-time can more than compensate for the decrease in response-accuracy when low levels of random noise are added. Such behavior of the circuit produces “stochastic resonance”—an optimal level of noise for producing responses in a given time—which is in contrast to the usual expectation that random noise detracts from information processing.

It is worth noting that perfect integration of evidence (Eqs. 6.13 and 6.14 with $W_s - W_x = 1$ and $\tau_{eff} \rightarrow \infty$) is mathematically the optimal method for discerning the greater of two noisy inputs. In perfect integration, there is neither forgetting (or leak) of early information, nor an over-reliance on the initial input, either of which can occur in other models. Rather, in perfect integration, input at all time-points is weighted equally. However, when biological constraints are included so that the decision-making circuit contains its own intrinsic noise (so is never a perfect integrator) and the threshold cannot be raised indefinitely (because neural firing rates are bounded) it turns out that decisions by state transitions can be optimal¹⁵. In Tutorial 6.2 we compare these two different types of decision-making in a single circuit.

6.5.4. Biasing Decisions

If, ahead of any stimulus, recent history or other prior information indicates that one choice is more likely to be the correct one then we should bias our choice according to such prior information (see Bayes’ Theorem, Eq. 1.32). Similarly, if one of our choices is more rewarding than the other, we should bias our choice toward the more rewarding outcome (*cf.* Eq. 3.6). As we saw in Section 3.5.2, the bias impacts the probability of a choice being correct multiplicatively, so if either the reward for one response, or its prior probability, is double that of the other stimulus, we should be twice as likely to make the corresponding response, all other things being equal.

Box 6.20. Bias: In decision making, a tendency to prefer one alternative over others before the stimulus is presented. Such bias can improve performance if appropriate for the task.

One can achieve such bias qualitatively in a model by applying additional constant input to the favored alternative. Such bias simply produces a horizontal shift in the choice probability curve (shown without bias in Figure 6.10A) such that if the stimulus difference were zero the probability of choosing the favored alternative would be greater than 0.5.

Alternatively, the circuit could be initialized in a state that is closer to one threshold than the other. For a perfect integrator, the ratio of the distances to threshold for the two competing alternatives is the ratio of choice probabilities produced when the stimuli are identical. For example, if the circuit were initialized to require half the change in rate to produce response A compared to that needed to produce response B then response A would be produced on 2/3 of trials and response B on 1/3 of trials given a stimulus with equal support for A and B.

6.6. Tutorial 6.2. Dynamics of a decision-making circuit in two modes of operation.

Neuroscience goals: see how noise can lead to significant trial-to-trial variability in behavior; compare the behavior of an integrator circuit with a state-transition circuit.

Computational goals: simulate coupled firing-rate units with different types of noise; record event times and realign vectors with different event times and pad them for further analysis.

In this tutorial, we will produce a decision-making circuit of the type shown in Figure 6.8 and study its response to inputs. The circuit contains two units, labeled $i=1,2$. When one of the two units reaches its decision-threshold we assume one of the two alternatives has been chosen.

We will be able to switch the mode of decision-making by altering the connection strengths and thresholds of the units. In order to produce a perfect integrator, the firing-rate curve of the units will be a linear function of their inputs,

$$\tau \frac{dr_i}{dt} = -r_i + I_i - \Theta,$$

with the constraint $0 \leq r_i \leq r_{max}$. In the above equation, Θ is the firing-threshold (the minimum input needed to cause the unit to fire, not to be confused with the decision-threshold) and I_i is the input to unit i given by

$$I_i(t) = W_s \cdot r_i(t) + W_x \cdot r_j(t) + G \cdot s_i(t) + \sigma_{int} \cdot \eta_i(t),$$

where $r_j(t)$ is the rate of the other cell (so if $i = 1$ then $j = 2$ and if $i = 2$ then $j = 1$). Each unit will receive independent internal noise of strength σ_{int} as well as a scalable independent noise within the stimulus $s_i(t)$, of strength σ_s , as described below.

The speed-accuracy tradeoff is achieved by altering the parameter G , which scales the input synapses to the circuit (like an input conductance).

The stimuli, when present, are given by $s_1(t) = \bar{s} + \Delta s/2 + \sigma_s \cdot \eta'_1(t)$ and $s_2(t) = \bar{s} - \Delta s/2 + \sigma_s \cdot \eta'_2(t)$, with mean $\bar{s} = 1$, and variable difference Δs .

The terms $\eta_1(t)$, $\eta_2(t)$, $\eta'_1(t)$, and $\eta'_2(t)$ each represent unit variance noise to be generated independently on each trial, as described in question 2 below.

For the integration mode, set $W_s = 0.975$, $W_x = -0.025$, and $\Theta = -0.5$ for each unit.

For the jumping mode, set $W_s = 1.05$, $W_x = -0.05$, and $\Theta = 4$ for each unit.

For both circuits set a decision threshold of 50Hz and the maximum rate, $r_{max} = 60\text{Hz}$. In all simulations simulate until either the threshold is reached by one of the units or a maximum time of 10s is reached. Use a time constant of $\tau = 10\text{ms}$.

Answer all questions with the circuit in integration mode, then repeat them all with the circuit in jumping mode. Comment on any similarities or differences between the two sets of results.

1) In integration mode set $G = 1$ and in jumping mode set $G = 2.5$. Simulate a single trial without noise ($\sigma_{int} = 0$ and $\sigma_s = 0$) as follows:

- a) Initialize the firing rates of the two units to their identical steady state value in the absence of a stimulus or noise by solving for r in

$$\tau dr/dt = -r + W_s \cdot r + W_x \cdot r - \Theta = 0$$

and enforcing the bounds on the range of allowed values of r .

- b) Set $\Delta s = 0.1$ for a stimulus that commences at 0.5s and remains on. Simulate the differential equations with a time-step of 0.5ms then plot the firing rate of both units as a function of time, and separately plot the difference in their firing rates.

2) Add independent noise to the entire duration of each stimulus, by repeatedly choosing two random numbers (one for each stimulus) from the unit Gaussian distribution (*e.g.*, by using numpy's `random.randn`) and keeping these two noise values fixed for each 2ms of the stimulus duration. Multiply each noise value by $\sigma_s = 0.2$ and divide by the square-root of 0.002, before adding to the stimulus. Add intrinsic noise in a similar manner, again keeping it fixed every 2ms, and using the magnitude $\sigma_{int} = 0.25$. Ensure the intrinsic noise is present across all of the time simulated (not just when the stimulus is on). Simulate 200 trials, for each trial recording the firing rate as a function of time.

- a) Plot the mean firing rate aligned to stimulus onset for both units, being careful to include (and normalize) at each time-point only those trials yet to produce threshold-crossing.
- b) Plot the mean firing rate during the decision-making period, aligning across trials to the time of decision-threshold crossing. Again, be careful to either include at each time point only those trials being simulated at that time point, or to set the rates to the initial value at all time points before the simulation started. This is necessary, since the firing rate arrays for each trial will be offset from each other by large amounts according to the range of threshold-crossing times.
- c) Record which unit first crosses the decision threshold and the time of threshold-crossing. Plot a histogram of the times to threshold on trials for which the unit 1 (which received greater input) was the first to reach threshold.

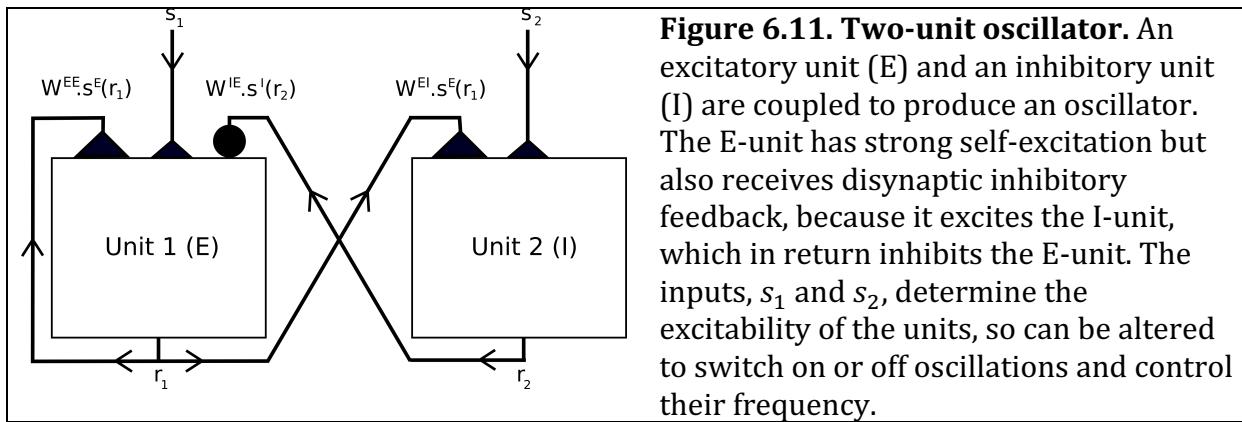
3) Now repeat 2c), looping through values of $\Delta s = 0, 0.25, 0.5, 0.75, 1$. Record the mean time for unit 1 to reach threshold (on trials when it does do so) as well as the fraction of trials in which unit 1 is the winner for each value of Δs . Plot these two quantities separately against stimulus-difference, Δs .

4) Now repeat 1)-3) with $G = 0.5$ using the parameters for integration mode then $G = 2$ using the parameters for jumping mode.

Note 1: Notice that in integration mode, when the input is reduced, integration is only possible over a small range of rates before the weaker unit is silenced. This limits a circuit's ability to perform perfect integration over a longer period simply by weakening input synapses. Such deviation from perfect integration explains why it is possible to find the jumping mode to be both faster and more accurate in question 4, something that should not be possible when compared to perfect integration (except perhaps by "luck" from the randomness inherent in these simulations).

Note 2: The methods used here simulate each trial individually to build up statistics of the times to threshold and probability for each unit to reach threshold first. It is possible, using the methods of partial differential equations to simulate the probability of each unit having a given rate at a given time, and thus generate the probability that a unit crosses threshold at a given time. Such methods generate exact probabilities, akin to calculating the probability of two coin tosses being heads as $\frac{1}{2} \times \frac{1}{2} = \frac{1}{4}$, rather than simulating 1000 coin tosses to obtain approximately 250 pairs of double-heads. These mathematical tools can be acquired in courses in advanced calculus or stochastic processes²².

6.7. Oscillations from excitatory and inhibitory feedback

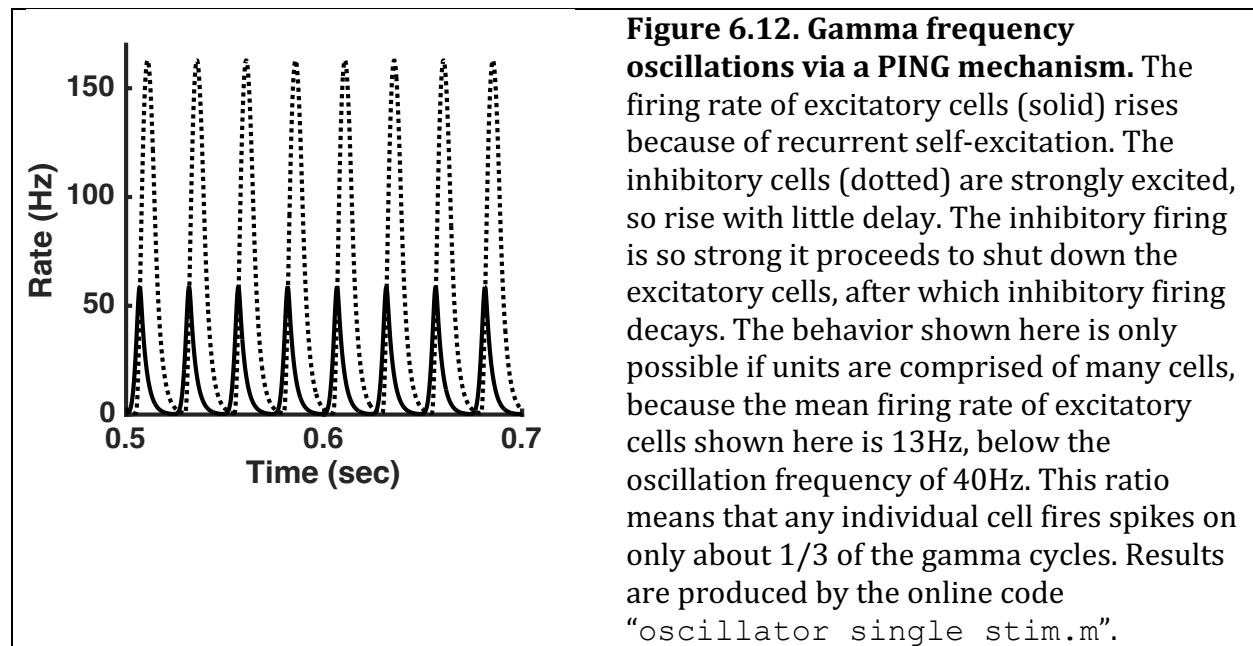


If sensitive electrodes are placed on our heads, they are likely to reveal oscillatory electric fields. Such external fields arise when oscillating neural activity is synchronized among many neurons aligned in a similar direction. We have seen in Chapter 4 how the spiking of individual neurons can be an oscillatory process, or how specific intrinsic channels can lead to periodic bursts of spikes. We have also seen in Chapter 5 how the interplay between two neurons can produce oscillations that would not arise in either cell alone. All of these oscillations depended on specific types of intrinsic channels within the neurons, so even in examples where the coupling between cells was essential, the oscillations could not be attributed entirely to the network effect. In this section, we will consider oscillations that are entirely due to network properties, so can be simulated in firing rate models, in which each unit has no internal dynamics.

The frequency, power, and coherence of oscillations vary with our mental state and are task dependent. A relatively high-frequency oscillation is the gamma rhythm, in the range 30-80Hz, which is associated with increased attention and memory processing²³. For example, when a subject attends to a visual stimulus, the power of gamma oscillation in the

responsive area of visual cortex increases²⁴, whereas when the subject attends to an auditory signal, the gamma power in auditory cortex increases²⁵.

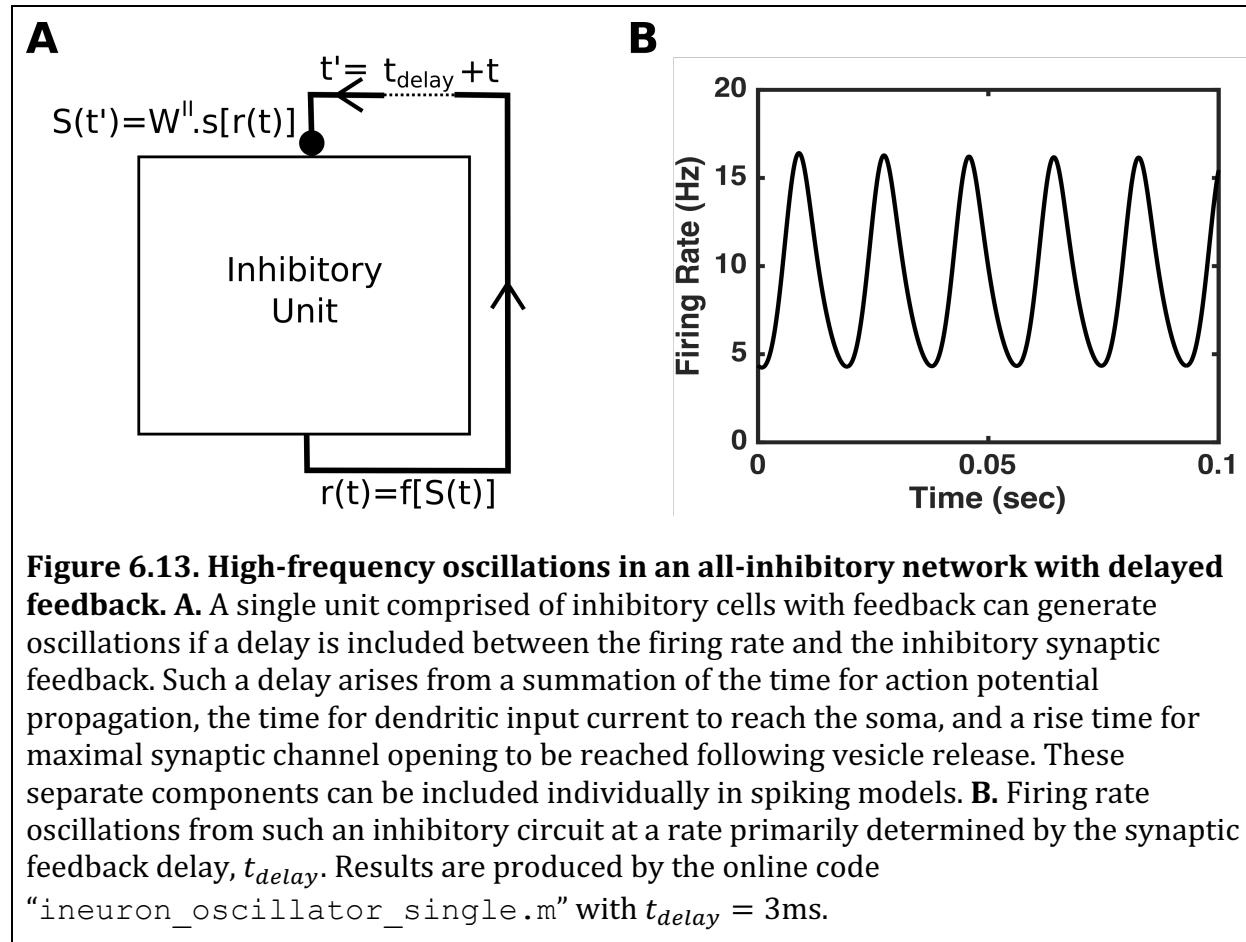
Box 6.21. Gamma rhythm: oscillatory neural activity in the range of 30-80Hz, which can be localized to specific brain areas, indicative of attention and cognitive processing.



One subset of models of the generation of gamma oscillations relies on the coupling between excitatory neurons (pyramidal cells) and inhibitory neurons (fast-spiking interneurons)²⁶. The operation of these PING models (standing for Pyramidal-InterNeuron Gamma) can be understood using a firing-rate model (Figure 6.11), because the oscillation depends entirely on network feedback, not on any intrinsic oscillation properties of the cells. In a PING model, self-excitation among excitatory pyramidal neurons destabilizes any low-level of spontaneous activity, generating a burst of excitatory activity. The burst of excitation causes the interneurons to fire even more actively. The high interneuron activity produces delayed feedback inhibition to suppress the burst of excitation among the pyramidal cells. Once the excitation ends in this manner, the interneurons are no longer excited and their activity also drops, allowing a new burst to arise once the inhibitory input to pyramidal cells decays (Figure 6.12).

In alternative models, oscillations of a similar frequency can arise among inhibitory neurons alone^{27,28}. Such network oscillations require inhibitory neurons to have high spontaneous activity in the absence of input from other neurons in the oscillating circuit, because any periodic inhibitory input from other neurons can only produce periodic dips in a neuron's firing rate. The high spontaneous activity can be due to intrinsic properties of the cells themselves (high conductance of sodium and/or calcium channels relative to the conductance of potassium channels) or due to tonic (meaning constant, non-oscillating) input from excitatory cells from outside their circuit. In such a purely inhibitory circuit, the neurons can synchronize to produce pulses of inhibition that hyperpolarize all cells transiently. The time to recover from inhibition is determined predominantly by the

synaptic time constant. If this is similar enough for all cells then many of them will, upon recovery, produce action potentials at a similar time. These synchronized action potentials generate a new pulse of inhibition that prevents further spiking for a short amount of time (Figure 6.13).



It is important in the firing-rate model of interneuron gamma to incorporate a synaptic delay—corresponding to a delay in time from when the soma of a neuron produces a spike to the time that the inhibitory current impacts the soma of a downstream cell. Without such a delay, as the firing rate of an inhibitory unit rises, it suppresses its own firing rate, preventing further rise. With a delay, the rate has time to rise before the inhibition kicks in. In a spiking model, the time for the membrane potential to recover from an inhibitory pulse would add to any delay in a synaptic input. In a firing rate model, the total delay can be mimicked if the synaptic feedback is set to depend on the firing rate of the unit a short time, t_{delay} , earlier (Figure 6.13A).

Box 6.22. Power spectrum: A plot of oscillatory power versus frequency, to indicate the dominant frequencies in any time-dependent signal with some oscillating components.

Box 6.23. Fourier transform: A process by which a time-dependent function can be separated into the contributing amplitudes of oscillating components of different frequencies, which together combine to produce the original function.

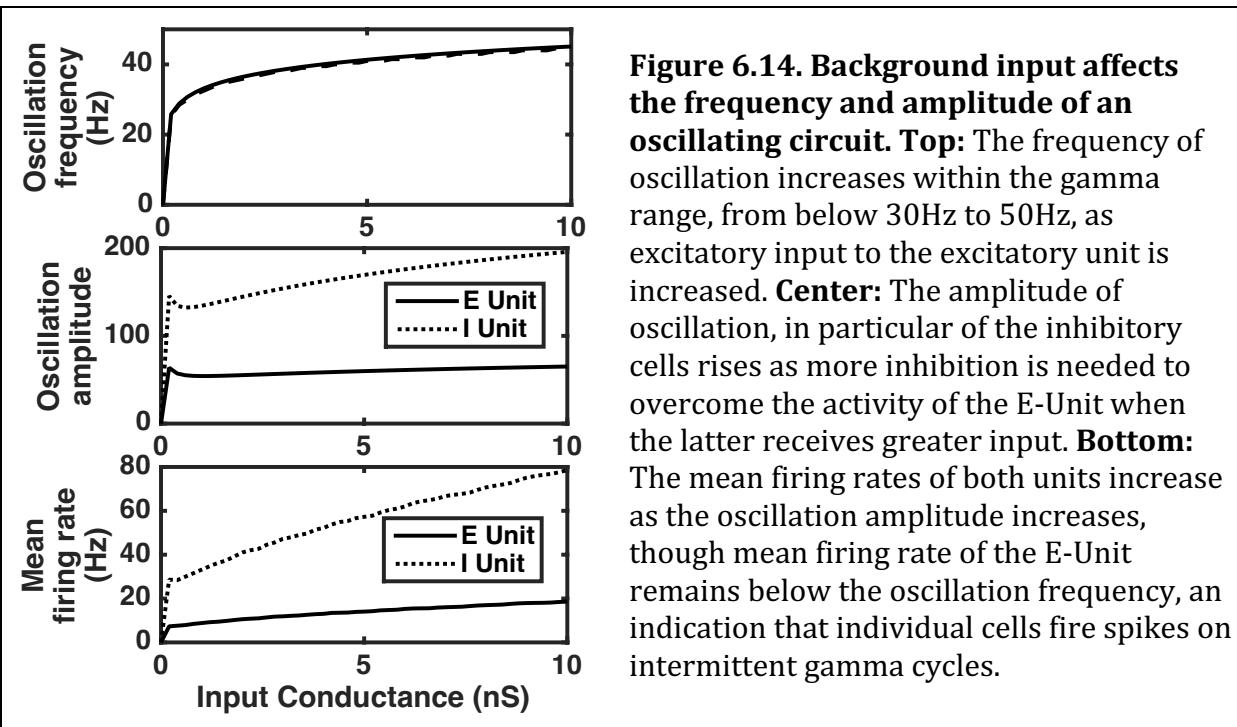


Figure 6.14. Background input affects the frequency and amplitude of an oscillating circuit. **Top:** The frequency of oscillation increases within the gamma range, from below 30Hz to 50Hz, as excitatory input to the excitatory unit is increased. **Center:** The amplitude of oscillation, in particular of the inhibitory cells rises as more inhibition is needed to overcome the activity of the E-Unit when the latter receives greater input. **Bottom:** The mean firing rates of both units increase as the oscillation amplitude increases, though mean firing rate of the E-Unit remains below the oscillation frequency, an indication that individual cells fire spikes on intermittent gamma cycles.

6.8. Tutorial 6.3. Frequency of excitatory-inhibitory coupled unit oscillator and PING.

Neuroscience Goals: discover how the frequency and amplitude of oscillations in a circuit can depend on external inputs.

Computational Goals: addressing limiting cases of "0/0" analytically; introduction to the Fourier transform.

In this tutorial, we will simulate the circuit of an excitatory firing-rate unit (1) coupled to an inhibitory firing-rate unit (2), as in Figure 6.11, with parameters that lead to oscillations. We will measure the amplitude and frequency of the oscillations and in so doing we will see how a Fourier transform operates.

1) The firing rate function for each unit will be based on the empirical fit by Chance and Abbott to the response of leaky integrate-and-fire neurons²⁹:

$$f(V_{ss}) = \frac{V_{ss} - V_{th}}{\tau(V_{th} - V_{reset})\{1 - \exp[-(V_{ss} - V_{th})/\sigma_V]\}}.$$

In this model, V_{ss} is the steady state membrane potential for the LIF model in the absence of spiking (Eq. 2.9):

$$V_{ss} = \frac{G_L E_L + G_I E_I + G_E E_E}{G_L + G_I + G_E},$$

where G_L , G_I , and G_E are respectively the leak, inhibitory, and excitatory conductances and E_L , E_I , and E_E are their corresponding reversal potentials. V_{th} is the threshold of the membrane potential, V_{reset} is the reset potential, τ is the membrane time constant, and the parameter σ_V (the noise in the membrane potential) determines the steepness and sharpness of curvature of the firing-rate curve.

As written, the function $f(V_{ss})$ is undefined if $V_{ss} = V_{th}$ (both numerator and denominator are zero). However, one can show that in the limit of very small δ (that is if $\delta \ll \sigma_V/a$), if $V_{ss} = V_{th} + \delta$ or $V_{ss} = V_{th} - \delta$, then $f(V_{ss}) \sim \sigma_V/[\tau(V_{ss} - V_{reset})a]$. Therefore, the function $f(V_{ss})$ should be defined to have this limiting value when $\delta = 0$ (and $V_{ss} = V_{th}$). The computer cannot figure out such a limit though, so when evaluating the function, we must include a conditional statement in the code that produces the correct limiting value instead of dividing by zero whenever $V_{ss} = V_{th}$.

Use the following parameters: $V_{th} = -50\text{mV}$, $V_{reset} = -80\text{mV}$, $\sigma_V = 1\text{mV}$, $\tau = 3\text{ms}$, $E_L = -70\text{mV}$, $E_I = -65\text{mV}$, $E_E = 0\text{mV}$ and $G_L = 50\text{pS}$. The excitatory and inhibitory conductances, G_E and G_I , will be time-dependent and differ between the two cells:

$$\begin{aligned} G_E^{(1)}(t) &= W^{EE} s_E^{(1)}(t) + G_{in}^{(1)} \\ G_I^{(1)}(t) &= W^{IE} s_I^{(2)}(t) \\ G_E^{(2)}(t) &= W^{EI} s_E^{(1)}(t) + G_{in}^{(2)} \\ G_I^{(2)}(t) &= 0 \end{aligned}$$

The connection strengths are set as $W^{EE} = 25\text{nS}$, $W^{EI} = 4\text{nS}$, and $W^{IE} = 800\text{nS}$. The excitatory input conductances, $G_{in}^{(1)}$ and $G_{in}^{(2)}$ will vary across trials, but for questions 1)-3) set $G_{in}^{(1)} = 1\text{nS}$ and $G_{in}^{(2)} = 0\text{nS}$.

1) Set up the simulation of two units with parameters given above, such that for unit i , the firing rate varies according to:

$$\tau \frac{dr_i}{dt} = -r_i + f(V_{ss,i})$$

where $V_{ss,i}$ is the instantaneous steady state that depends on the values of conductances for unit i as described above. The synaptic variables depend on presynaptic rate according to:

$$\frac{ds_E^{(1)}}{dt} = -\frac{s_E^{(1)}}{\tau_E} + \alpha r_1 (1 - s_E^{(1)})$$

and

$$\frac{ds_I^{(2)}}{dt} = -\frac{s_I^{(2)}}{\tau_I} + \alpha r_2 (1 - s_I^{(2)}),$$

where the synaptic time constants are $\tau_E = 2\text{ms}$ and $\tau_I = 5\text{ms}$. For all synapses set $\alpha = 0.2$. Simulate the equations for 2.5 seconds using a time-step of $\Delta t = 0.1\text{ms}$, assuming initial values of zero for all variables, and plot the firing rates of each unit on the same graph. If the firing rates oscillate, what is the oscillation frequency?

2) In this section, you will attempt the first of two methods for automatically calculating the frequency of oscillation.

- a) Copy the firing rate of the first unit into a new vector that omits an initial transient period of 0.5 seconds (or longer than 0.5 seconds if the oscillations take longer to stabilize).
- b) Determine the maximum and minimum values of the firing rate during the oscillation and use those values to select two thresholds, one slightly below the maximum rate to indicate the rate is approaching a peak and the other slightly above the minimum rate to indicate the rate is approaching a trough.

- c) Define a variable, a threshold-crossing indicator, which you initialize with a value of 0, unless the first value of firing rate is above the higher threshold, in which case you initialize the indicator with a value of 1 (because the system is already above-threshold).
- d) Loop through the firing rate vector. Within the loop set a conditional such that if the threshold-crossing indicator is zero and the rate is greater than the higher threshold then the time-point is recorded and the indicator is set to one. If the rate is lower than the lower threshold, reset the threshold-crossing indicator to zero.
- e) Calculate the oscillation period using the times of the first and last crossings of the higher threshold and the number of oscillations in between these crossings.

3) In this section, you will use a computationally simpler method that is equivalent to taking the Fourier transform.

- a) Truncate the firing rate vector (as in 2a) to obtain the periodic portion of the trial.
- b) Produce a vector of frequency values, from 0Hz to 100Hz with steps of 0.5Hz, and set up a loop through all values of the vector.
- c) Within the loop, for each value of frequency, f , create two vectors, each of the same length as the truncated firing rate vector, one equal to $\sin(2\pi ft)$ and the other equal to $\cos(2\pi ft)$ where the time points, t , span the time interval of the truncated firing rate vector in steps of Δt .
- d) Find the overlaps between the sine vector and the truncated firing rate vector by multiplying each corresponding element of the two vectors and taking the mean of the result. Record this overlap in a vector of coefficients,

$$A(f) = \overline{\sin(2\pi ft) \cdot r_1(t)},$$

one value for each frequency (the overbar in the above equation indicates that the mean value is taken, in this case over the portion of the trial where $r_1(t)$ is periodic). Repeat to find the overlap of the firing rate vector with the cosine vector, to produce a set of coefficients,

$$B(f) = \overline{\cos(2\pi ft) \cdot r_1(t)}.$$

- e) You now have two vectors; one with coefficients that each represent how similar the oscillating firing rate is to a sine wave of a given frequency, the other with coefficients that each represent the similarity to a cosine wave. It turns out that the sum of the squares of these two coefficients provide a measure of the oscillating power that is independent of the phase offset—the oscillation can be a sine or a cosine or anywhere in between and the sum of the squares is not altered. Therefore, produce a power spectrum as a vector such that each element, $P(f)$, is the sum of the squares of the corresponding sine coefficient and cosine coefficient:

$$P(f) = A^2(f) + B^2(f).$$

- f) Plot $P(f)$ as a function of the frequency vector and extract the value of f other than zero for which $P(f)$ is greatest. Compare this frequency to the oscillation frequency you calculated in Question 2. Why does $f = 0$ produce the greatest value for $P(f)$ (so must be discarded) even when the rate oscillates? Do you see multiple peaks? If so, why?

- 4) Let the applied stimulus to excitatory cells range from $G_{in}^{(1)} = 0\text{nS}$ to $G_{in}^{(1)} = 10\text{nS}$ (while keeping the input to inhibitory cells at $G_{in}^{(2)} = 0\text{nS}$). Plot relevant figures (such as Figure 6.14) to determine how the oscillation frequency, oscillation amplitude for both excitatory and inhibitory cells, and the mean firing rates of both excitatory and inhibitory cells depend on $G_{in}^{(1)}$. Explain any trends or relationships you notice.
- 5) Fix $G_{in}^{(1)} = 2\text{nS}$ and let the applied stimulus to inhibitory cells range from $G_{in}^{(2)} = 0\text{pS}$ to $G_{in}^{(2)} = 25\text{pS}$ (note the smaller values). Plot the same figures as in Question 4 and explain any differences in the results of Questions 4) and 5).

6.9. Orientation selectivity and contrast invariance

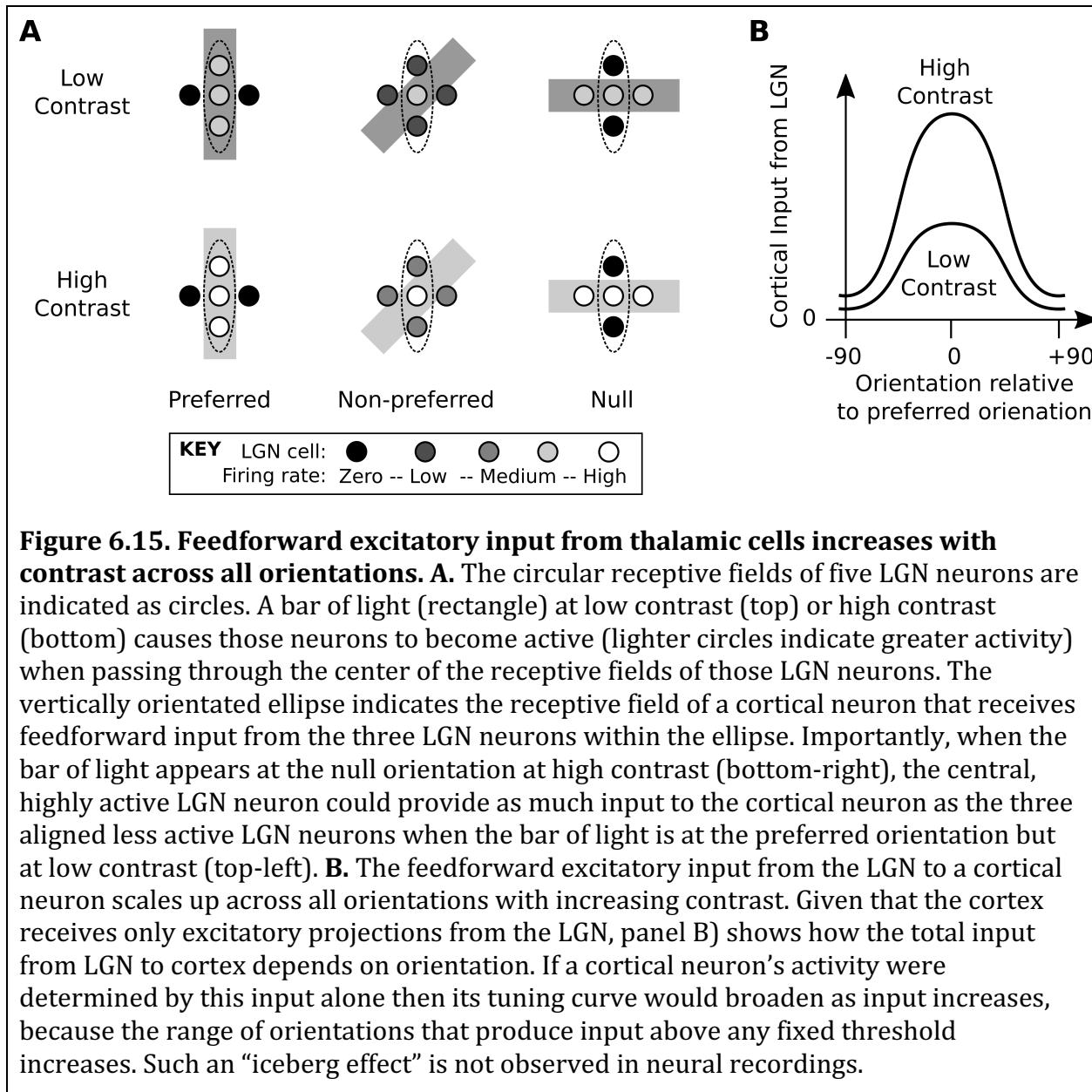
A notable feature of many neurons in primary visual cortex (V1) is their strong activity in response to an edge or a bar of a particular orientation within their visual receptive field (Figure 1.4B). As the orientation of a bar appears at an angle rotated from the preferred orientation (the one producing the greatest response in a neuron) a smooth decline in the neuron's firing rate is observed as a function of stimulus angle. When a bar is presented with the null orientation, which is perpendicular to the preferred orientation, the neuron's response can be lower than the spontaneous activity it produces in the absence of any stimulus.

Box 6.24. Orientation selectivity: The preferred firing of many neurons in V1 to visual inputs which are edges or gratings oriented at a particular angle.

Box 6.25. Primary visual cortex, V1: Also called occipital cortex, or striate cortex, the region of cortex at the back of the brain, which receives most direct visual input from the retina via the thalamus, with perhaps the most studied neural circuitry of mammals.

Box 6.26: Lateral geniculate nucleus, LGN: A region of the thalamus receiving inputs from the optic nerve, which sends excitatory outputs to V1.

Such orientation-tuned responses are observed in Layer 4 neurons, which receive excitatory, but not inhibitory, input from the lateral geniculate nucleus (LGN) of the thalamus. The firing rates of neurons in the LGN increase with stimulus contrast—early stages of visual processing from the pupil to the retina compensate for changes in overall brightness or luminance of a visual stimulus such that light-dark boundaries generate downstream responses.



One puzzling question has been how the responses of orientation-selective neurons in V1 can be explained in terms of their inputs from the LGN. In particular, can V1 responses be explained in terms of the summation of different inputs from the LGN, or are intracortical connections within V1 essential for shaping the V1 responses^{30,31}?

6.9.1. Ring models

In answering this question, one feature that has required explanation is the observation of contrast invariance. The tuning curve of a neuron in V1 does not change its shape—it just increases its amplitude multiplicatively—as the contrasts of stimuli increase. Since projections from the thalamus to the cortex are only excitatory, the observation of contrast invariance rules out any model in which the V1 neurons only receive input from thalamic neurons. In such models, as contrast increases the excitatory thalamic input to the V1

neurons increases, even when the stimulus is at non-preferred orientations. Therefore, while a tuning curve with a peak at a preferred orientation could be produced for a neuron, somewhat like an iceberg rising out of the water, the tuning curve would broaden with increased contrast (Figure 6.15).

Box 6.27. Contrast invariance: Tuning curves, which do not change their shape—such that the firing rates can be multiplicatively scaled—when the contrast of a set of visual stimuli is changed.

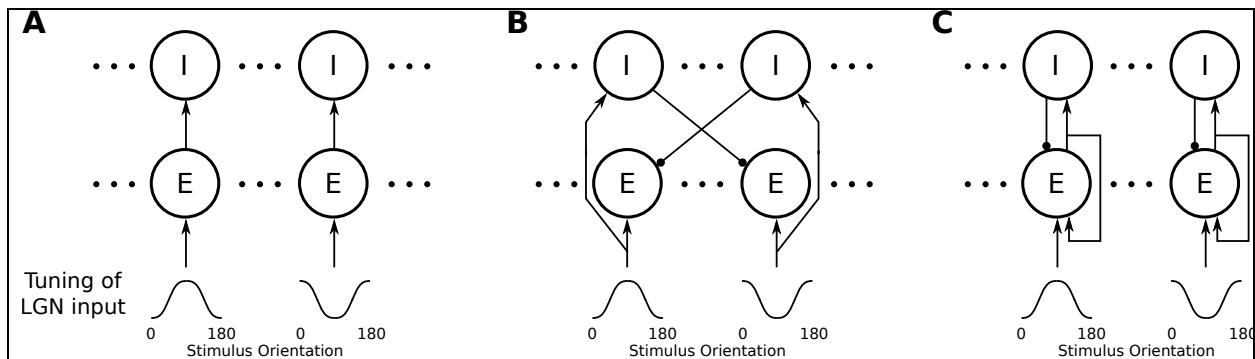
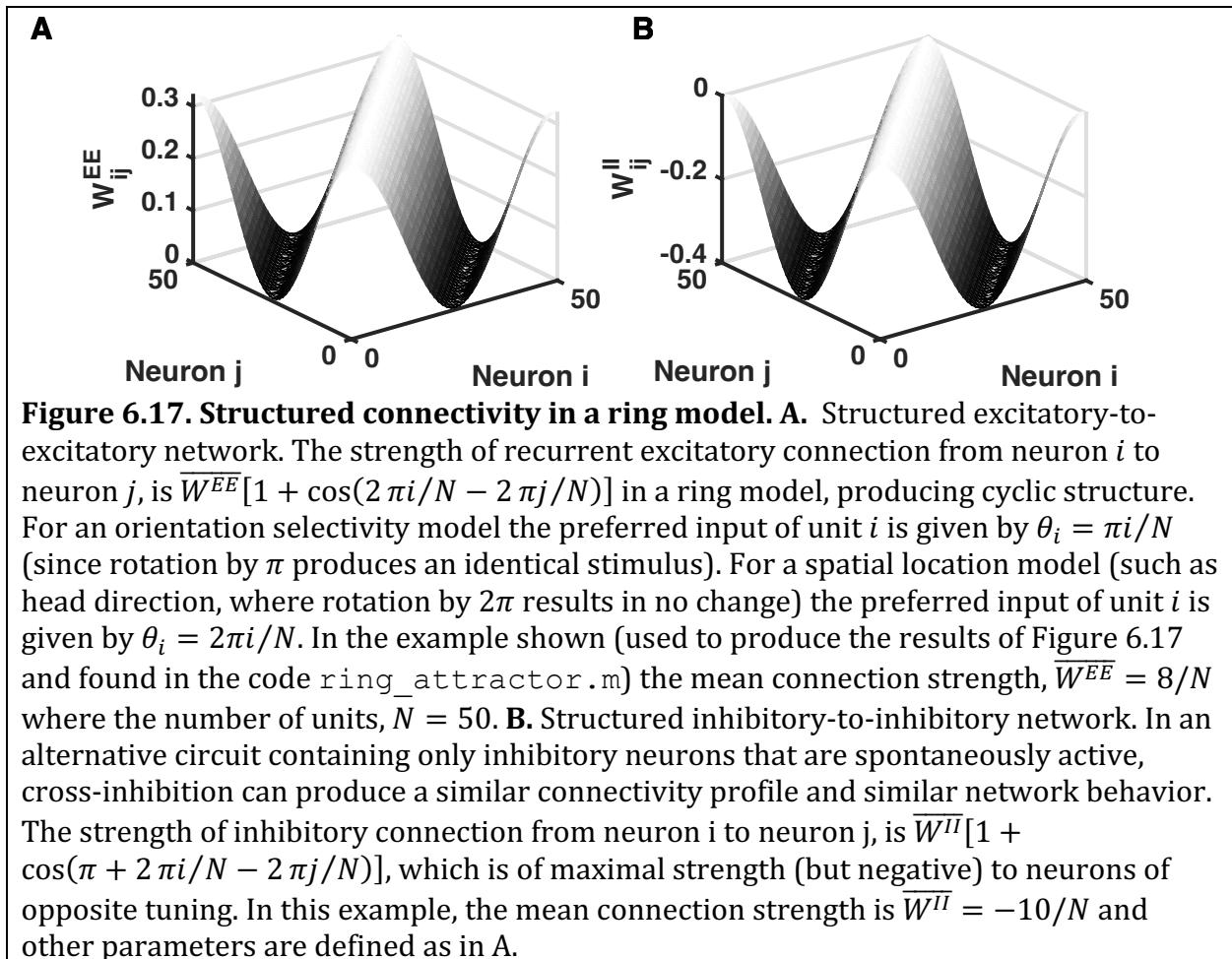


Figure 6.16. Circuit motifs for investigation of contrast-invariant orientation

selectivity in V1. **A.** Feedforward excitation is shown to two excitatory cells (labeled E) with opposite orientation preferences. Tuning curves of excitatory cells are unaffected by cortical connections. **B.** Feedforward excitation and inhibition to excitatory cells. Inhibitory input to excitatory units is from interneurons (labeled I) receiving oppositely tuned LGN input. **C.** Feedforward and feedback excitation and inhibition. Feedback is via recurrent excitatory and inhibitory connections from similarly tuned cells within the cortex. **A-C.** In all figures, just two pairs of cells are shown, whereas the complete set would be arranged according to orientation preference in a ring, with all orientations present. Only the strongest connections are shown, yet in all models the local connections also spread to neurons with similar, not just identical, orientation preferences, as shown in Figure 6.16.

Such a model (Figure 6.16.A) is a bit of a straw man (*i. e.*, expected to fail) because inhibitory neurons also receive thalamic input. Therefore, in an alternative feedforward model, the tuned excitatory V1 neurons could receive direct excitation from the thalamus, but also indirect inhibition from the thalamus. If the indirect inhibition were to target neurons with the opposite stimulus preference and were to scale with contrast in the same manner as the direct excitation (Figure 6.16.B), then the increase in inhibition with contrast could counteract the increase in excitation with contrast at non-preferred orientations. We will investigate this possibility in Tutorial 6.4.

The final type of model (Figure 6.16.C), which is now widely accepted, is one that includes recurrent feedback in V1. Again, this should not be a surprise when considering the known anatomy, because a typical V1 cell, even in layer 4, receives the vast majority of its inputs (approximately 90%) from other cortical neurons rather than from the thalamus³².



A key feature of models based on recurrent feedback is that excitatory connections between excitatory cells are more effective—being more prevalent, or stronger, or both—between neurons that have similar tuning curves (Figure 6.17). Such local excitatory feedback increases the gain of responses to input, enhancing any supralinearity of the response, such that the effect of increased contrast is much greater at the preferred orientation (where the response is high) than at non-preferred orientations.

Given that the connectivity between neurons in these models depends on the difference in their preferred orientations, neurons are labeled and arranged according to their preferred orientation. Such an arrangement produces a ring model, because as the bar-stimulus is rotated by π radians (*i.e.*, by 180°), the stimulus returns to the original one. Therefore, neurons with preferred orientations of slightly greater than 0 are strongly connected to neurons with preferred orientations of slightly less than π .

Here it is worth noting that a connectivity structure can be ring-like, without the neurons' being physically located in a ring-like structure. However, it turns out that for orientation-selective neurons in V1 a ring-like arrangement of the neurons' physical positions does arise, producing what are known as pinwheels (Figure 6.18)^{33,34}. Such correspondence between the response properties of neurons and their physical location is called a topographic map, a common feature of primary sensory areas (see also Figure 1.4B).

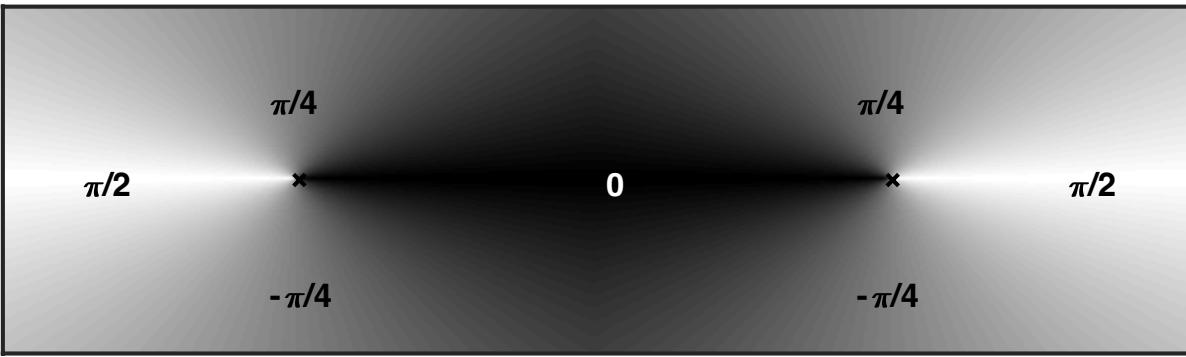
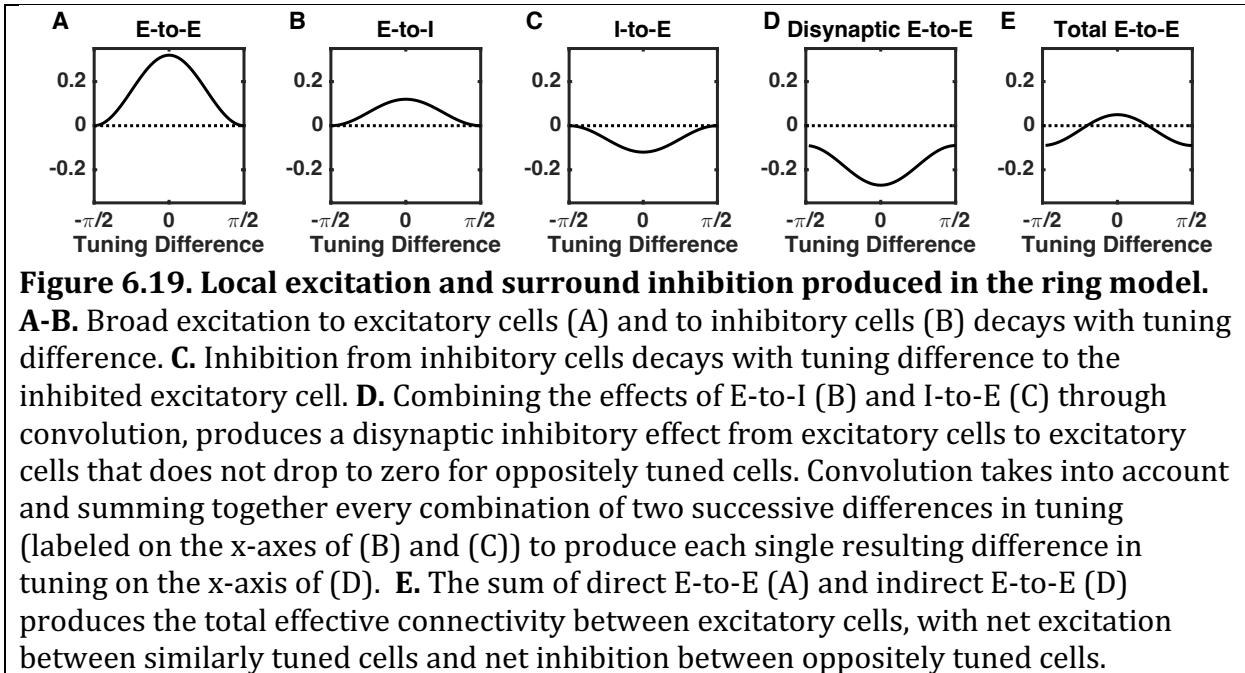


Figure 6.18. Pinwheel arrangement of neurons in V1. The figure represents an area of cortex containing numerous neurons shaded by their preferred orientations, with two pinwheel centers (marked by crosses) contained within the region. Neurons are arranged by orientation preference (indicated by labels on the figure) counter-clockwise around the left pinwheel center and clockwise around the right one. Brightness of shading indicates the absolute value of a neuron's preferred orientation, which ranges from $-\pi/2$ to $\pi/2$. The pinwheel arrangement observed in many mammals^{31,32} is efficient in ensuring the neurons most likely to be connected with each other in a ring model—those with similar orientation preferences—are nearby each other in the physical space of V1.

The second key component of contrast-invariant models of orientation selectivity is inhibitory feedback within the cortical network that spreads beyond the excitatory feedback. The inhibitory feedback must be sufficiently strong to prevent inactive units from becoming active whenever contrast increases and the firing rates of the most responsive units increase. In models of such “surround suppression” the extent of all connections can be identical (they are simply included as a cosine in our model). The inhibitory feedback spreads further than excitatory feedback when all connections have the same spatial spread, because feedback inhibition to excitatory cells is disynaptic—via an excitatory synapse to an inhibitory cell then via an inhibitory synapse to an excitatory cell. For example, if neurons connect to other cells whose preferred orientations differ by up to $\pm \frac{\pi}{4}$, then an excitatory cell can change the rate of an inhibitory cell with a $+\frac{\pi}{4}$ difference and that inhibitory cell inhibits a new excitatory cell with a $+\frac{\pi}{4}$ difference in preferred orientation from itself, so a $+\frac{\pi}{2}$ difference in preferred orientation from the original excitatory cell. In such a situation, activity in the original excitatory cell might increase the rate of all cells within $\pm \frac{\pi}{4}$ of its preferred orientation, but via the inhibitory cells, cause a decrease in rate of cells whose preferred orientation is between $+\frac{\pi}{4}$ and $+\frac{\pi}{2}$ (or $-\frac{\pi}{4}$ and $-\frac{\pi}{2}$).

With appropriate connections strengths, the result of the total network feedback is that cells with similar orientation preferences excite each other, while they inhibit cells with very different orientation preferences (see Figure 6.19).



6.10. Ring attractors for spatial memory and head direction

If the recurrent excitatory feedback is increased in a model for contrast-invariant orientation selectivity, then the bump of activity produced when the stimulus is on can remain after stimulus offset (Figure 6.20). The activity in such a network can indicate both whether a stimulus was recently presented (a binary memory) and the angle corresponding to the stimulus (memory of a continuous quantity, called parametric memory).

Box 6.28. Ring attractor: A network in which neurons can be labelled by a variable, such as orientation, with circular symmetry, and thus arranges in a ring, in which the network possesses stable states with peak activity anywhere on the ring.

In Figure 6.4 we saw the classes of stable states of the network progress from only an inactive state (Figure 6.4A), to both an active and an inactive state (Figure 6.4B), to only an active state (Figure 6.4C) when the degree of excitatory feedback to a single unit was monotonically increased. The same progression through classes occurs in a ring network with spatially distributed excitatory feedback. For the ring network, all three of the classes of activity can provide useful functional roles.

First, we saw in the previous section that even if the inactive state is the only stable state following stimulus offset, the circuit can produce contrast-invariant responses during stimulus presentation. Second, a circuit with both active and inactive states can provide memory for the location (on a circle) of a stimulus while remaining inactive in the absence of a stimulus (Figure 6.20A). Third, a circuit with no inactive state can provide important information to an animal about a parameter that should always exist, such as its head direction (Figure 6.20B). Indeed, we shall see later in this section that a ring attractor circuit can provide both a memory of direction and update the direction by near-perfect integration of angular velocity if the latter is provided as an input.

Box 6.29. Head-direction cell: A neuron, which is most active when an animal's head is facing in a particular direction, most likely produced by integration of angular velocity signals, since its tuning is present in the dark.

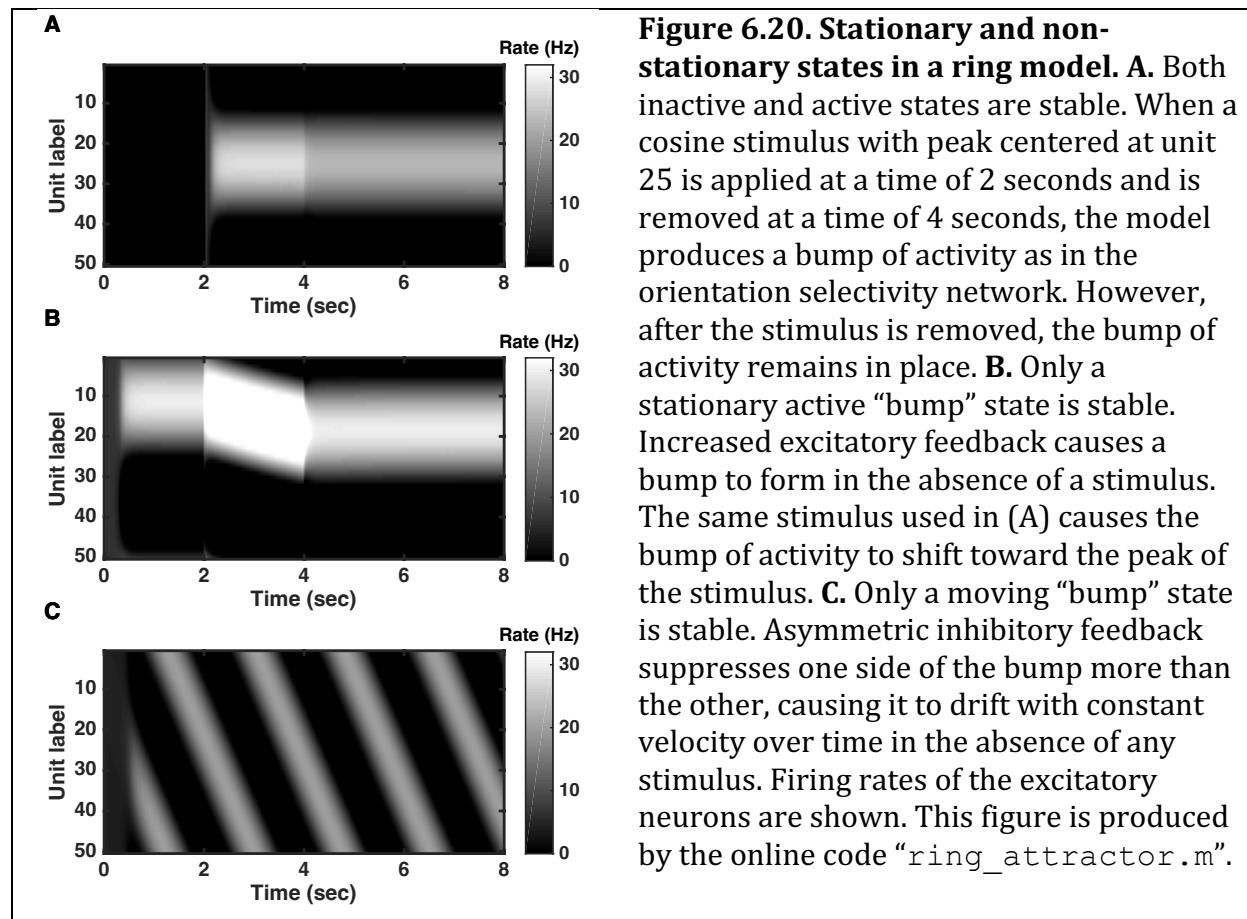


Figure 6.20. Stationary and non-stationary states in a ring model. **A.** Both inactive and active states are stable. When a cosine stimulus with peak centered at unit 25 is applied at a time of 2 seconds and is removed at a time of 4 seconds, the model produces a bump of activity as in the orientation selectivity network. However, after the stimulus is removed, the bump of activity remains in place. **B.** Only a stationary active “bump” state is stable. Increased excitatory feedback causes a bump to form in the absence of a stimulus. The same stimulus used in (A) causes the bump of activity to shift toward the peak of the stimulus. **C.** Only a moving “bump” state is stable. Asymmetric inhibitory feedback suppresses one side of the bump more than the other, causing it to drift with constant velocity over time in the absence of any stimulus. Firing rates of the excitatory neurons are shown. This figure is produced by the online code “ring_attractor.m”.

6.10.1. Dynamics of the ring attractor

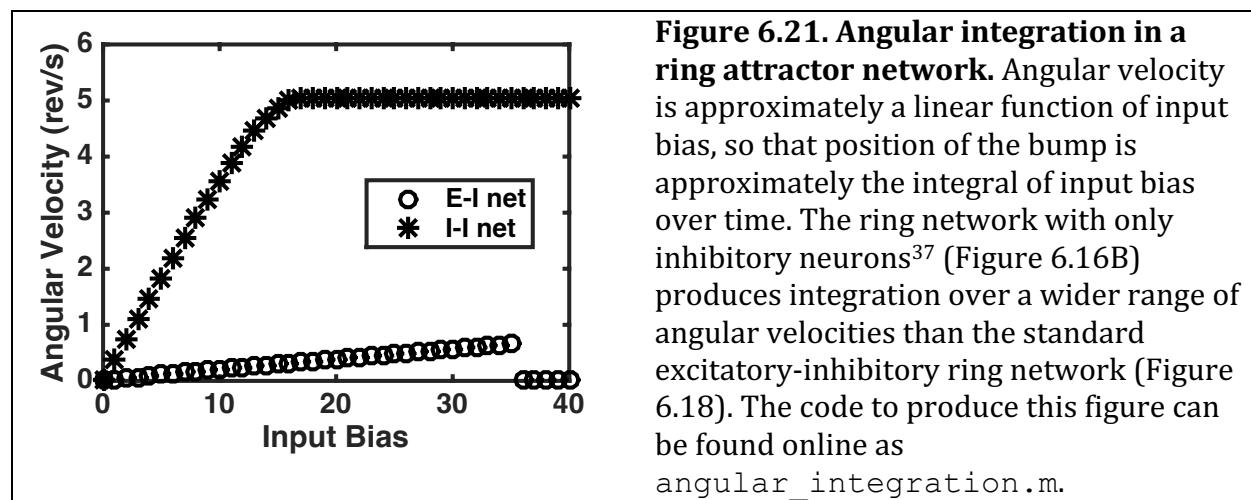
The ring attractor is so-called because of the ring-like symmetry to the system and because the active ‘bump’ of activity that remains following stimulus offset has a stable shape that reappears—the network’s activity is attracted back to it—after any temporary deviation. The position of the bump along the ring is marginally stable, as a low level of noise or a small asymmetric stimulus can cause it to drift gradually away from its initial position³⁵⁻³⁷. If the bump is initiated at a particular location (Figure 6.20A) thereafter the random drift causes the variance of its position to increase linearly in time. Such random walk-like behavior is typical of any line attractor (of which the ring-attractor is a special case).

The ring attractor was used to account for the visuo-spatial short-term memory³⁸ required in a task in which a monkey must recall the location of a stimulus—presented at a point on a circle—after a delay following stimulus offset. The monkey had to shift its eyes in a saccade to the remembered location following a cue at the end of the delay. Electrophysiology during the task revealed neurons in prefrontal cortex with firing rates in the delay period tuned to the prior stimulus location, just like neurons in a ring attractor.

As with the line attractor mentioned earlier in the chapter, the marginal stability of the ring attractor allows it to act as an integrator (Figure 6.7). Any asymmetry in the connectivity of the circuit causes the bump of activity to move toward whichever side is receiving more excitation or less inhibition. Such behavior can lead to instability of a symmetric, so otherwise stationary bump, if the neurons have strong adaptation or synapses are strongly depressing. In such cases, once noise causes the bump to move in a given direction, then the trailing edge of the bump—where neurons were recently more active—has effectively fewer excitatory neurons or weaker synapses than the leading edge of the bump, where neurons were recently inactive. Thus, the bump continues to move in one direction. Such behavior can be the cause of spontaneous waves in cortex.

More usefully, if the differences in excitability of the two sides of a bump can be controlled by an input signal, then the direction and velocity of the bump can be controlled. In this case, the bump can perform angular integration, moving more quickly in response to a large asymmetric input and more slowly in response to a weak asymmetric input. To control the level of asymmetry, the feedback should be via two populations one with a connection bias in one direction, the other with a connection bias in the opposing direction. Preferential input to one of these biased feedback populations causes the bump of activity to move in the corresponding direction³⁹⁻⁴¹.

Simulations suggest that such angular integration is more robust in a ring model with inhibitory feedback connections (Figure 6.17B) than with excitatory feedback connections (Figure 6.17A). The range of velocities over which near-perfect integration arises is greatly enhanced in the former case (Figure 6.21). Interestingly, the cells responsive to head direction (the integral of the head's angular velocity⁴¹) might arise in just such a ring attractor formed by inhibitory feedback³⁹.



6.11. Tutorial 6.4. Orientation selectivity in a ring model

Neuroscience goal: see how contrast-invariant gain can arise from different network configurations.

Computational goal: simulation of a large number of coupled ODEs, using arrays to store variables and their coupling constants.

In this tutorial, you will simulate a network of 100 firing-rate model units (50 excitatory and 50 inhibitory) to investigate how connections within the network shape neural responses to external input. We will compare a feed-forward model with a recurrent circuit model (see Figure 6.16). The tutorial follows the work of Ben-Yishai and Sompolinsky³⁵.

Set up two arrays, with N_t rows and 50 columns, where N_t is the number of time points to be simulated. One array is for 50 excitatory cells and the other for 50 inhibitory cells. A simulation of 300ms with a time step of 0.1ms is sufficient.

Define three 50×50 connectivity matrices, each of which can be initialized with entries of zero. These are respectively W^{EE} , W^{EI} , and W^{IE} , for the connection strengths E-to-E, E-to-I, and I-to-E respectively (we ignore I-to-I connections in this tutorial). For each part of the tutorial the firing rates will be treated as a linear function of inputs, so their dynamics follow:

$$\begin{aligned}\tau_E \frac{dr_i^E}{dt} &= -r_i^E + \frac{1}{N} \sum_j W_{ij}^{EE} r_j^E + \frac{1}{N} \sum_j W_{ij}^{IE} r_j^I + I_0^E + S_i^E \\ \tau_I \frac{dr_i^I}{dt} &= -r_i^I + \frac{1}{N} \sum_j W_{ij}^{EI} r_j^E + I_0^I + S_i^I\end{aligned}$$

with the additional conditions $r_i^E > 0$ and $r_i^I > 0$ for all units and all times (rates cannot be negative).

In the above equations, I_0^E and I_0^I represent baseline input to excitatory and inhibitory units respectively. A negative value for these quantities is equivalent to a positive threshold in the f-I curve. The stimulus-dependent inputs, S_i^E and S_i^I , depend on the orientation preference of the unit labeled by i . We define the orientation preference of each excitatory unit as $\theta_i = \pi i / N$, where $N = 50$ is the total number of units of a given type, and use $\theta_{cue} = \pi/2$ to denote the orientation of the stimulus. Given these definitions, we use:

$$\begin{aligned}S_i^E &= A_E c [1 + \varepsilon \cdot \cos(2\theta_{cue} - 2\theta_i)], \\ S_i^I &= A_I c [1 + \varepsilon \cdot \cos(2\theta_{cue} - 2\theta_i)],\end{aligned}$$

with c being the contrast, which will vary from 0 to 1. $\varepsilon = 0.5$ represents the degree of modulation of input with orientation, in this case producing a three-fold increase in input from the null direction (where $\theta_{cue} = \theta_i + \pi/2$, so $\cos(2\theta_{cue} - 2\theta_i) = -1$) to the preferred direction (where $\theta_{cue} = \theta_i$, so $\cos(2\theta_{cue} - 2\theta_i) = +1$).

Other parameters depend on the network being simulated as follows:

Network A. $\tau_E = \tau_I = 10$ ms, $I_0^E = I_0^I = -10$, $A_E = 40$, $A_i^I = 40$, $W_{ij}^{EE} = W_{ij}^{EI} = W_{ij}^{IE} = 0$.

Network B. $\tau_E = \tau_I = 10$ ms, $I_0^E = I_0^I = -5$, $A_E = 40$, $A_i^I = 40$, $W_{ij}^{EE} = W_{ij}^{EI} = 0$, $W_{ij}^{IE} = -[1 + \cos(\pi + 2\theta_i - 2\theta_j)]/N$.

Network C. $\tau_E = 50$ ms, $\tau_I = 5$ ms, $I_0^E = 2$, $I_0^I = 0.5$, $A_i^E = 100$, $A_i^I = 0$, $W_{ij}^{EE} = 5[1 + \cos(2\theta_i - 2\theta_j)]/N$, $W_{ij}^{EI} = 3[1 + \cos(2\theta_i - 2\theta_j)]/N$, $W_{ij}^{IE} = -4[1 + \cos(2\theta_i - 2\theta_j)]/N$.

Complete questions 1-4 using each of the networks A, B, and C.

- 1) With all firing rates initially at zero, simulate the network as a function of time, with the stimulus present the entire time, using contrasts of $c = 0, 0.25, 0.5, 0.75$, and 1 .

Note: Firing rates can be stored as arrays, with each row representing a different time point and each column a different cell. This allows the sum over all excitatory units providing input to all excitatory cells within the network to be calculated using matrix multiplication as: $rE(i-1, :) @ WEE$ (the $@$ multiplies the preceding vector by the following matrix if both are numpy arrays).

Each entry in the row vector produced by the above segment of code represents the total within-network excitatory input to a unit. The “ $i-1$ ” in the segment of code represents the previous time point, because firing rates at the previous time point are used when calculating input at the current time point, “ i ”.

- 2) Plot, using separate figures for the excitatory and inhibitory units, firing rate as a function of time for those units with preferred orientations of $\theta_i = \theta_{cue} = \pi/2$ and $\theta_i = \pi$. Ensure the same figures are used for all five contrasts.

- 3) At the last time point simulated plot the firing rates of the population of excitatory and inhibitory units as a function of their index on two separate figures (one for excitatory units, the other for inhibitory units). Ensure the same figures are used for all contrasts.

- 4) Produce rescaled versions of the plots produced in 3) by dividing the firing rates of each unit by the mean firing rate averaged over all units for that stimulus contrast. Comment on any contrast-invariant scaling observed, or lack thereof—*i. e.*, whether, when the response curves produced with different contrasts are rescaled by their mean response, they become identical to each other.

- 5) In addition to contrast invariance, further investigations showed that excitatory neurons received their greatest inhibitory input when the stimulus was at their preferred orientation, which is when they also received their greatest excitatory input. Which model(s) are compatible with this result?

- 6) Experiments typically involve the recording of a few neurons and acquire each neuron's tuning curve using a range of stimuli—in this case with stimuli of multiple orientations. How can we discuss the tuning curve (response of a single neuron to multiple orientations of stimuli) from our simulations when we only ever simulated stimuli with a single orientation?

Questions for Chapter 6.

- 1) Connections from the thalamus to V1 are purely excitatory. Why would it be reasonable in a firing-rate model to allow a connection from a unit in the thalamus to a unit in V1 to have a negative strength?
- 2) What happens to (i) the typical time to threshold and (ii) the accuracy of responses in an integrator-type decision-making circuit if:
 - a) Noise in the inputs is increased?
 - b) Input synapses are strengthened?
 - c) Thresholds of all units are increased?
- 3) A firing-rate model produces oscillations with a frequency of 20Hz with neurons never firing at a rate above 15Hz. Can such a model apply to real spiking neurons, and if it can, what would you expect to see in the neural spike trains?
- 4) The stationary bump state of a ring attractor can become unstable if synapses are depressing. Why does this happen, and what are the two possible stable bump states that are generated in this process?

References for Chapter 6

1. Goldman-Rakic PS. Cellular basis of working memory. *Neuron*. 1995;14:477-485.
2. Wang M, Yang Y, Wang CJ, et al. NMDA receptors subserve persistent neuronal firing during working memory in dorsolateral prefrontal cortex. *Neuron*. 2013;77(4):736-749.
3. Wang XJ. Synaptic basis of cortical persistent activity: the importance of NMDA receptors to working memory. *J Neurosci*. 1999;19(21):9587-9603.
4. Seung HS. How the brain keeps the eyes still. *Proc Natl Acad Sci U S A*. 1996;93(23):13339-13344.
5. Aksay E, Gamkrelidze G, Seung HS, Baker R, Tank DW. In vivo intracellular recording and perturbation of persistent activity in a neural integrator. *Nat Neurosci*. 2001;4(2):184-193.
6. Mazurek ME, Roitman JD, Ditterich J, Shadlen MN. A role for neural integrators in perceptual decision making. *Cerebral Cortex*. 2003;13(11):1257-1269.
7. Wong KF, Wang XJ. A recurrent network mechanism of time integration in perceptual decisions. *The Journal of neuroscience : the official journal of the Society for Neuroscience*. 2006;26(4):1314-1328.
8. Miller P, Brody CD, Romo R, Wang XJ. A recurrent network model of somatosensory parametric working memory in the prefrontal cortex. *Cerebral Cortex*. 2003;13(11):1208-1218.
9. Romo R, Brody CD, Hernandez A, Lemus L. Neuronal correlates of parametric working memory in the prefrontal cortex. *Nature*. 1999;399(6735):470-473.
10. Ratcliff R, McKoon G. The diffusion decision model: theory and data for two-choice decision tasks. *Neural Comput*. 2008;20(4):873-922.
11. Wang XJ. Probabilistic decision making by slow reverberation in cortical circuits. *Neuron*. 2002;36:955-968.
12. Shadlen MN, Newsome WT. Neural basis of a perceptual decision in the parietal cortex (area LIP) of the rhesus monkey. *Journal of Neurophysiology*. 2001;86(4):1916-1936.
13. Bogacz R. Optimal decision-making theories: linking neurobiology with behaviour. *Trends Cogn Sci*. 2007;11(3):118-125.
14. Drugowitsch J, Pouget A. Probabilistic vs. non-probabilistic approaches to the neurobiology of perceptual decision-making. *Curr Opin Neurobiol*. 2012;22(6):963-969.
15. Miller P, Katz DB. Accuracy and response-time distributions for decision-making: linear perfect integrators versus nonlinear attractor-based neural circuits. *Journal of Computational Neuroscience*. 2013;35(3):261-294.
16. Shadlen MN, Newsome WT. Motion perception: seeing and deciding. *Proceedings of the National Academy of Sciences of the United States of America*. 1996;93(2):628-633.
17. Ratcliff R, Smith PL, Brown SD, McKoon G. Diffusion Decision Model: Current Issues and History. *Trends Cogn Sci*. 2016;20(4):260-281.
18. Usher M, McClelland JL. The time course of perceptual choice: the leaky, competing accumulator model. *Psychol Rev*. 2001;108(3):550-592.
19. Seung HS, Lee DD, Reis BY, Tank DW. The autapse: a simple illustration of short-term analog memory storage by tuned synaptic feedback. *J Comput Neurosci*. 2000;9(2):171-185.

20. Eckhoff P, Wong-Lin KF, Holmes P. Optimality and robustness of a biophysical decision-making model under norepinephrine modulation. *The Journal of neuroscience : the official journal of the Society for Neuroscience*. 2009;29(13):4301-4311.
21. Shea-Brown E, Gilzenrat MS, Cohen JD. Optimization of decision making in multilayer networks: the role of locus coeruleus. *Neural Comput*. 2008;20(12):2863-2894.
22. Gillespie DT. *Markov Processes*. Academic Press; 1992.
23. Jensen O, Kaiser J, Lachaux JP. Human gamma-frequency oscillations associated with attention and memory. *Trends Neurosci*. 2007;30(7):317-324.
24. Fries P, Reynolds JH, Rorie AE, Desimone R. Modulation of oscillatory neuronal synchronization by selective visual attention. *Science*. 2001;291(5508):1560-1563.
25. Sokolov A, Pavlova M, Lutzenberger W, Birbaumer N. Reciprocal modulation of neuromagnetic induced gamma activity by attention in the human visual and auditory cortex. *Neuroimage*. 2004;22(2):521-529.
26. Whittington MA, Stanford IM, Colling SB, Jefferys JG, Traub RD. Spatiotemporal patterns of gamma frequency oscillations tetanically induced in the rat hippocampal slice. *J Physiol*. 1997;502 (Pt 3):591-607.
27. Whittington MA, Traub RD, Kopell N, Ermentrout B, Buhl EH. Inhibition-based rhythms: experimental and mathematical observations on network dynamics. *Int J Psychophysiol*. 2000;38(3):315-336.
28. Geisler C, Brunel N, Wang XJ. Contributions of intrinsic membrane dynamics to fast network oscillations with irregular neuronal discharges. *J Neurophysiol*. 2005;94(6):4344-4361.
29. Abbott LF, Chance FS. Drivers and modulators from push-pull and balanced synaptic input. *Prog Brain Res*. 2005;149:147-155.
30. Priebe NJ, Ferster D. Inhibition, spike threshold, and stimulus selectivity in primary visual cortex. *Neuron*. 2008;57(4):482-497.
31. Priebe NJ. Mechanisms of Orientation Selectivity in the Primary Visual Cortex. *Annu Rev Vis Sci*. 2016;2:85-107.
32. Peters A, Payne BR. Numerical relationships between geniculocortical afferents and pyramidal cell modules in cat primary visual cortex. *Cereb Cortex*. 1993;3(1):69-78.
33. Koch E, Jin J, Alonso JM, Zaidi Q. Functional implications of orientation maps in primary visual cortex. *Nat Commun*. 2016;7:13529.
34. Wilson DE, Whitney DE, Scholl B, Fitzpatrick D. Orientation selectivity and the functional clustering of synaptic inputs in primary visual cortex. *Nat Neurosci*. 2016;19(8):1003-1009.
35. Ben-Yishai R, Bar-Or RL, Sompolinsky H. Theory of orientation tuning in visual cortex. *Proc Natl Acad Sci USA*. 1995;92:3844-3848.
36. Camperi M, Wang XJ. A model of visuospatial short-term memory in prefrontal cortex: recurrent network and cellular bistability. *J Comput Neurosci*. 1998;5:383-405.
37. Compte A, Brunel N, Goldman-Rakic PS, Wang XJ. Synaptic mechanisms and network dynamics underlying spatial working memory in a cortical network model. *Cereb Cortex*. 2000;10:910-923.
38. Funahashi S, Bruce CJ, Goldman-Rakic PS. Mnemonic coding of visual space in the monkey's dorsolateral prefrontal cortex. *J Neurophysiol*. 1989;61(2):331-349.

39. Song P, Wang XJ. Angular Path Integration by Moving "Hill of Activity": A Spiking Neuron Model without Recurrent Excitation of the Head-Direction System. *J Neurosci*. 2005;25:1002-1014.
40. Zhang K. Representation of spatial orientation by the intrinsic dynamics of the head-direction cell ensembles: A theory. *J Neurosci*. 1996;16:2112-2126.
41. Turner-Evans D, Wegener S, Rouault H, et al. Angular velocity integration in a fly heading circuit. *Elife*. 2017;6.