Neuron, Vol. 26, 259–271, April, 2000, Copyright ©2000 by Cell Press

Stability of the Memory of Eye Position in a Recurrent Network of Conductance-Based Model Neurons

H. Sebastian Seung,*†‡ Daniel D. Lee,† Ben Y. Reis,† and David W. Tank†
*Department of Brain and Cognitive Sciences
Massachusetts Institute of Technology
Cambridge, Massachusetts 02139
†Bell Labs
Lucent Technologies
Murray Hill, New Jersey 07974

Summary

Studies of the neural correlates of short-term memory in a wide variety of brain areas have found that transient inputs can cause persistent changes in rates of action potential firing, through a mechanism that remains unknown. In a premotor area that is responsible for holding the eyes still during fixation, persistent neural firing encodes the angular position of the eyes in a characteristic manner: below a threshold position the neuron is silent, and above it the firing rate is linearly related to position. Both the threshold and linear slope vary from neuron to neuron. We have reproduced this behavior in a biophysically plausible network model. Persistence depends on precise tuning of the strength of synaptic feedback, and a relatively long synaptic time constant improves the robustness to mistuning.

Introduction

In many vertebrates, oculomotor behavior with the head stationary consists of static periods of fixation punctuated by saccadic eye movements (see Figure 1A). During fixation in mammals, premotor neurons in the prepositus hypoglossi and the medial vestibular nucleus exhibit maintained discharge, the frequency of which changes with every saccade to a new angular eye position (Lopez-Barneo et al., 1982; Escudero et al., 1992; McFarland and Fuchs, 1992; Scudder and Fuchs, 1992). Each change in frequency is stimulated by a transient pulse of input from command neurons and persists for the duration of the fixation. This paper is concerned with a fundamental question: what are the physiological mechanisms that allow a transient input to give rise to a persistent change in neural activity?

The two brainstem nuclei mentioned above are part of the "neural integrator" for horizontal eye position in the oculomotor system of mammals (Robinson, 1989; Moschovakis, 1997), where the term "integrator" is used in the sense of calculus. The saccadic transformation of a pulse into a step change in eye position is just one example of integration. More generally, other types of eye movements also involve the integration of velocity signals to produce changes in eye position.

Lesion (Cheron et al., 1986a, 1986b; Kaneko, 1997, 1999), inactivation (Cannon and Robinson, 1987; Cheron and Godaux, 1987; Straube et al., 1991; Godaux et al., 1993; Mettens et al., 1994b), and microstimulation (Cannon and Robinson, 1987; Godaux et al., 1989; Yokota et al., 1992) studies indicate that the prepositus hypoglossi and medial vestibular nucleus participate in the maintenance of persistent activity and do not simply read it out from some other area of the brain. Despite a large number of experimental and theoretical studies of the integrator, the mechanisms responsible for maintaining persistent neural activity are still unclear.

These mechanisms are not just of interest in the oculomotor system, because persistent neural activity has been observed in a wide variety of brain areas (Fuster, 1995), generally in association with short-term memory tasks. From this broader perspective, the integrator can be viewed as storing a short-term memory of the current eye position (Seung, 1996). Pharmacological studies have shown that fixation is impaired when the integrator is inactivated (Cannon and Robinson, 1987; Cheron and Godaux, 1987; Straube et al., 1991; Godaux et al., 1993; Mettens et al., 1994b), a disorder that can be interpreted as loss of the ability to store a short-term memory of eye position.

For physiological investigations, the goldfish is a convenient experimental preparation with fixation performance comparable to that of mammals (Mensh et al., 1997, Soc. Neurosci., abstract). A brainstem nucleus has previously been identified as part of the neural integrator for horizontal eye position in goldfish (Pastor et al., 1994). The present paper describes a mathematical model of the goldfish integrator with the biophysical realism necessary for fruitful interaction with experimental research. An early version of this work has appeared elsewhere (Lee et al., 1997).

In the model, saccades are stimulated by bursts of input from command neurons. Each burst of feedforward input changes the firing rates of neurons in the network, and this change is maintained by recurrent excitation after the feedforward input is over. Signals from integrator neurons lead to the oculomotor plant, so that persistent changes in these signals cause persistent changes in the angular position of the eyes. In short, the model network is able to hold the eyes still during fixation, because it maintains a short-term memory of the current eye position in the form of a persistent pattern of neural activity. The recurrent excitatory connections in the network are crucial for this persistence. Their strengths are precisely tuned in the model, so that activity neither decays nor increases during periods of fixation.

A number of properties of biological integrators are reproduced by the model. First, the encoding of eye position in neural activity is in accord with the results of single-unit recordings of integrator neurons (Lopez-Barneo et al., 1982; Escudero et al., 1992; McFarland and Fuchs, 1992; Scudder and Fuchs, 1992). Namely, each integrator neuron in the model exhibits a linear relationship between firing rate and eye position when it is active. However, there is also a threshold eye position

‡To whom correspondence should be addressed (e-mail: seung@mit.edu).

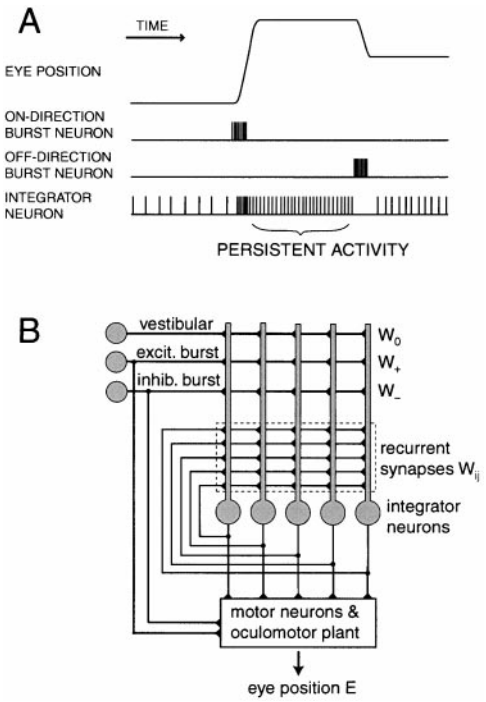

Figure 1. Signals in the Oculomotor System and the Architecture of a Recurrent Network Model

(A) Schematic of eye position and neural activity versus time in the oculomotor system.

(B) Network model of the integrator based on recurrent excitation. All integrator neurons receive feedforward input from vestibular, excitatory burst, and inhibitory burst neurons. To make the diagram compact, only 5 integrator neurons are shown, but 15 were actually used in the simulations. Signals from the burst and integrator neurons are relayed by the motor neurons to the oculomotor plant, to produce eye position.

below which it is silent. The linear slope and the threshold vary from neuron to neuron.

Second, persistence is not perfect, even in a well-tuned model. There is some drift of neural activity with time, which leads to drift in eye position during fixation. The drift velocity depends systematically on eye position, generally in a nonlinear manner. Similar systematic dependences have been experimentally observed in measurements of fixation behavior in the dark (Becker and Klein, 1973; Hess et al., 1985; Mensh et al., 1997, Soc. Neurosci., abstract).

Third, the persistence of neural activity degrades when synaptic strengths are mistuned, neurons are destroyed, or the strength of feedback is otherwise perturbed. This qualitatively reproduces the experimental results of pharmacological inactivation of the integrator, in which persistence degrades with the extent of inactivation (Cannon and Robinson, 1987; Pastor et al., 1994). It also suggests that the changes in fixation performance observed after visual–vestibular adaptation are due to changes in the strength of feedback in the integrator (Tiliket et al., 1994).

Both intrinsic and synaptic conductances of the neurons are modeled in a biophysically plausible way. The model neurons fire action potentials repetitively in response to applied current, at a rate that is an approximately threshold linear function of the current (Shriki et al., 1998, Soc. Neurosci., abstract). Real integrator neurons measured in vitro exhibit similar discharge properties (Serafin et al., 1991; du Lac and Lisberger, 1995; Wang et al., 1998, Soc. Neurosci., abstract). Since little is known about synaptic transmission in the goldfish integrator, a generic synaptic model is used. The model synapses are assumed to have a saturation nonlinearity, which is commonly observed in neurons in a wide variety of brain areas. The time constant of recurrent synapses is set at 100 ms. This value could correspond, for example, to the time constant of the NMDA receptor, if the recurrent synapses of the goldfish integrator turn out to be glutamatergic.

The strength of feedback was tuned through a careful design process, which involved approximation of the spiking, conductance-based model by a nonspiking, reduced model obtained using an averaging method similar to those introduced previously (Rinzel and Frankel, 1992; Ermentrout, 1994). The synaptic strengths of the reduced model were tuned so that it would maintain persistent neural activity, and these tuned values were placed in the conductance-based model. This tuning procedure was convenient but was not intended to be biologically plausible. Alternatively, tuning could conceivably be accomplished in biological integrators by adaptive mechanisms of synaptic plasticity (Arnold and Robinson, 1992, 1997; Seung, 1997, Soc. Neurosci., abstract).

Previously, there have been other recurrent network models of the integrator (Rosen, 1972; Kamath and Keller, 1976), most of which were formulated as networks of linear elements (Cannon et al., 1983; Galiana and Outerbridge, 1984; Cannon and Robinson, 1985; Arnold and Robinson, 1991, 1992, 1997). Like the present model, they were based on tuned positive feedback, although they were lacking in biophysical realism. The present model is different because it contains biologically relevant nonlinearities. Furthermore, it is based on a relatively long synaptic time constant of 100 ms and is therefore more robust than previous models, other things being equal. These points are explained more fully in the Discussion section.

Results

Persistent Activity in a Conductance-Based Model

As shown in Figure 1B, the model contains 15 integrator neurons, which interact with each other by recurrent excitatory synapses. They also receive feedforward input from three neurons. The vestibular input neuron is tonically active at a constant rate, simulating the background activity present in vestibular afferents when the head is stationary. The excitatory and inhibitory burst neurons are normally silent, except for occasional brief bursts of action potentials that cause saccadic eye movements.

Our choice of 15 integrator neurons in the model is comparable to estimates of 25–40 neurons in the goldfish oculomotor integrator (Pastor et al., 1994). Our methods could also be used to construct models of the integrator with larger numbers of neurons, to be consistent with the estimated sizes of mammalian integrators.

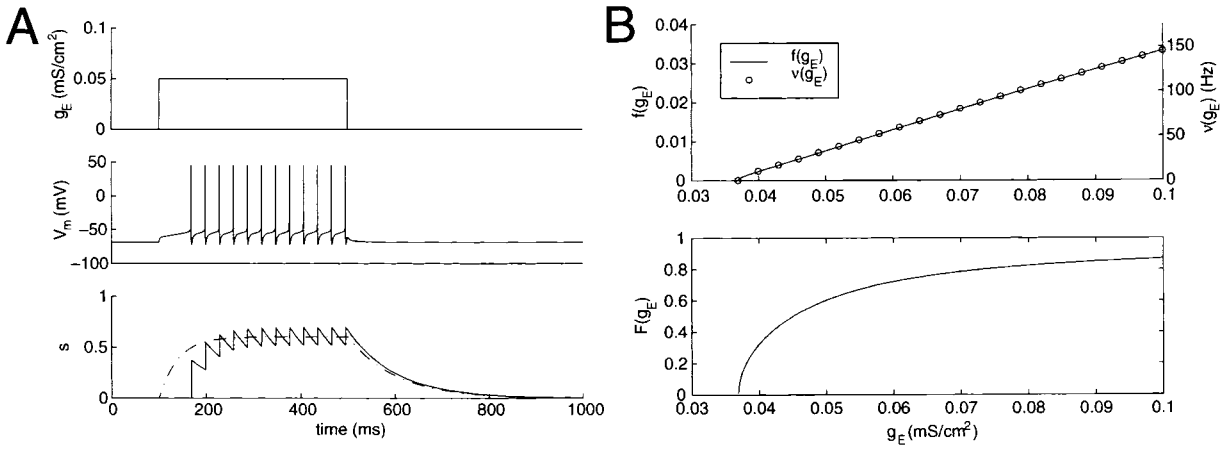

Figure 2. Repetitive Firing of the Conductance-Based Model Neuron

(A) Dynamics of membrane voltage *V* and synaptic activation *s* during repetitive firing.

(Top) A step in excitatory synaptic conductance g_E starting at 100 ms and ending at 500 ms.

(Middle) The first action potential occurs with a latency of about 68 ms. Thereafter, convergence to periodic behavior is rapid, with little spike frequency adaptation.

(Bottom) Synaptic summation for $\tau_{syn} = 100$ ms. Each action potential causes a jump in the synaptic activation *s*, with rise time equal to the width of the action potential. Between action potentials, *s* decays exponentially, as seen clearly in the uninterrupted decay after the last action potential. The jump sizes decrease as *s* increases, due to saturation. The broken line shows the behavior of the reduced dynamics in Equation 12.

(B) Response functions characterizing output of a neuron during repetitive firing as a function of excitatory synaptic input g_E with inhibitory synaptic input $g_I = 0$.

(Top) The neural response function *f* (defined in Equation 10) is almost exactly proportional to firing rate ν. Except for some rounding of the curves near threshold, both functions are roughly linear in g_E over the range in the graph.

(Bottom) The neural response function *F* also has a threshold but includes in addition the effects of saturation nonlinearity (see Equation 13).

We utilized a conductance-based model neuron, devised by Shriki et al. (1998, Soc. Neurosci., abstract), with a single compartment containing leak, voltage-dependent transient sodium, delayed rectifier potassium, and A-type potassium conductances. The equations governing these intrinsic conductances are described in the Experimental Procedures and Appendix. The model neuron has a threshold linear relationship between firing rate and applied current, resembling real integrator neurons measured in vitro (Serafin et al., 1991; du Lac and Lisberger, 1995; Wang et al., 1998, Soc. Neurosci., abstract).

Synaptic transmission is modeled with the simplifying assumption that a presynaptic spike train produces the same fraction of open receptors at every postsynaptic target. This fraction, called the synaptic activation, is incremented by each presynaptic action potential and decays exponentially with some time constant. There are two classes of synapses in our simulations, slow (100 ms) and fast (5 ms). The recurrent and vestibular synapses are slow and show temporal summation as in Figure 2A. The excitatory and inhibitory burst synapses are fast. Saturation nonlinearity was also included in the synaptic model (see Figure 2B). Whether such saturation exists in biological integrator synapses is not yet known, but similar saturation effects have been observed in other types of synapses throughout the nervous system.

The membrane potentials of the three input neurons are shown in Figure 3A. The vestibular neuron fires at a constant rate of roughly 40 Hz; the excitatory burst neuron fires at 1, 2, and 3 s after the beginning of the simulation; and the inhibitory burst neuron fires at 4 and 5 s.

The membrane potentials of the 15 integrator neurons are shown in Figure 3B. Each excitatory burst evokes a burst in the integrator neurons, and each inhibitory burst causes them to pause briefly from firing. These transient effects of transient stimulation are not surprising. More remarkably, the transient stimulation has a persistent effect: the firing rates immediately after a burst are different than they were before, and these new rates persist until the next burst a second later.

This persistence is not caused by feedforward synaptic input. Since the burst synapses are fast, burst input quickly decays back to zero after the burst. Furthermore, the vestibular input never changes: it is the same both before and after the burst. Therefore, the persistence is the result of recurrent synaptic excitation of the integrator neurons.

During each interburst interval, there is a persistent pattern of activity distributed across the population of integrator neurons, as shown in Figure 3B. In most intervals, the activity pattern consists of both active and inactive neurons. But in the first interval all neurons are inactive, and in the fourth interval all neurons are active. Within the subset of active neurons, the firing rates are heterogeneous, rather than uniform.

In the oculomotor system, the integrator and burst neurons drive extraocular motor neurons, which in turn drive the oculomotor plant (Belknap and McCrea, 1988). Here, this signal transformation is modeled as a linear combination of integrator and burst synaptic activations, which is then low-pass filtered with the viscoelastic time constant of the oculomotor plant. The resulting eye position is shown as a function of time in Figure 3C. Each excitatory burst causes a positive step in eye

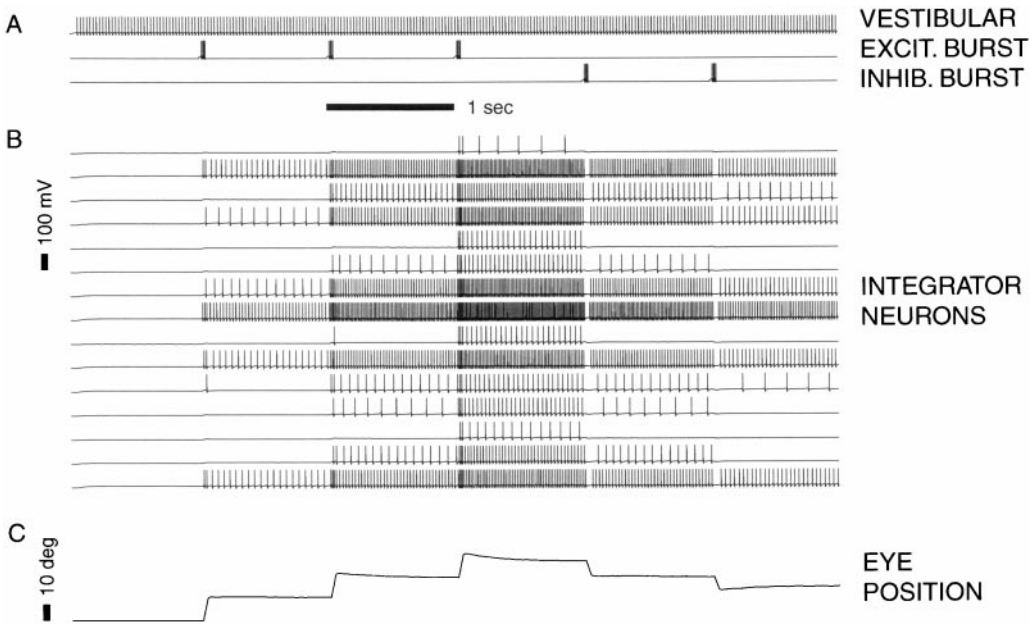

Figure 3. Persistent Changes of Neural Activity and Eye Position Induced by Transient Burst Inputs, as Simulated in the Conductance-Based Network Model

The weights from the burst neurons to the integrator neurons are $W_+ = 0.03$ and $W_- = 0.15$ for all i (see Equation 6).
(A) Vestibular, excitatory burst, and inhibitory burst neurons. The membrane potentials are plotted versus time.
(B) Integrator neurons. Each transient burst of input causes a persistent change in neural activity.
(C) Eye position. The persistence during interburst intervals is due to the persistence of neural activity in the integrator.

position, and each inhibitory burst causes a negative step. The persistence of eye position is due to the persistence of neural activity in the integrator.

Tuning of Synaptic Strengths

The persistent neural activity evident in Figure 3 is not a generic property of the recurrent network architecture used here. Rather, it arose from a careful design process in which the strength of feedback was precisely tuned. First, the spiking, conductance-based model was approximated by a *reduced model* that dispenses with the dynamics of action potential generation (see Equation 14 in the Experimental Procedures). The reduced model is formally similar to the "neural network equations" popular in brain modeling and computer science (Wilson and Cowan, 1972; Hopfield and Tank, 1986; Grossberg, 1988) but possesses a precise biophysical interpretation by virtue of its derivation from the conductance-based model using an averaging method similar to those introduced previously (Ermentrout, 1998b). Intuitively, the reduction is possible because slow recurrent synapses filter the neural spike trains, making possible a smooth, nonspiking description of network dynamics (Rinzel and Frankel, 1992; Ermentrout, 1994).

In many traditional neural network models, the nonlinearity of neural response is summarized by a single sigmoidal function, the precise form of which is a matter of convention. In contrast, the reduced model used here relies on two response functions, both of which are calculated from the dynamics of action potential generation in the conductance-based model. One response function is essentially the firing rate as a function of synaptic input (Figure 2B, top). It has a threshold nonlinearity but is approximately linear above threshold. The other response function includes the effects of synaptic saturation (Figure 2B, bottom). Both response functions are important for the model.

The reduced model was further simplified by constraining the recurrent synaptic weight matrix to be of the outer product form. Such structure emerges spontaneously from Hebbian synaptic plasticity in associative memory models (Hopfield, 1982). Related learning rules have also been proposed for the integrator (Arnold and Robinson, 1991, 1992, 1997; Seung, 1997, Soc. Neurosci., abstract; Xie and Seung, 2000) and could provide a way for approximate outer product structure to arise spontaneously in biological integrators.

With the outer product constraint, tuning the strength of feedback becomes equivalent to a problem in function approximation. This is described mathematically in the Experimental Procedures (Equation 18) and depicted graphically in Figure 4. Each curve in Figure 4A is a copy of the saturating response function at the bottom of Figure 2B and represents the contribution of a neuron to the feedback in the network. The threshold and scale of each curve are controlled by the strengths of recurrent and vestibular synapses connected to the corresponding neuron. These parameters have been adjusted so that the sum of all these curves approximates a straight line with unity slope, which is the condition for persistence through tuned feedback. The quality of the approximation is depicted in Figure 4B. The nth curve from the bottom is the cumulative sum of the curves in Figure 4A with the n leftmost thresholds. The

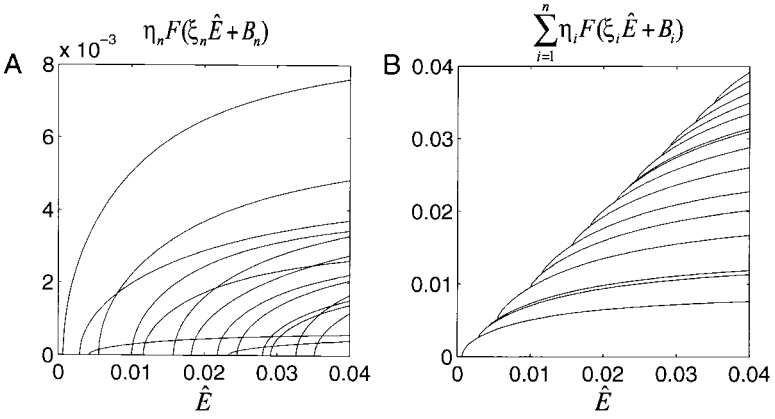

Figure 4. Tuning Positive Feedback by Balancing Saturation and Recruitment

(A) Graphical representation of the amount of feedback contributed by each integrator neuron to the network, which is given for neuron *i* by the term $\eta_i F(\xi_i \hat{E} + B_i)$ in Equation 18. Each curve is the same function *F* as in Figure 2B, but horizontally scaled and translated by ξ_i and B_i and vertically scaled by η_i.

(B) Cumulative sums of the same curves, in order of increasing threshold. The envelope of these curves is a good approximation to a straight line with unity slope, which is the condition for tuning positive feedback. The variable $\hat{E}$ is the internal representation of eye position (see Experimental Procedures for definition). As it increases, linearity is achieved by recruiting more neurons to compensate for saturation.

envelope of these curves is the full sum, which is a good approximation to the requisite straight line.

The approximation shown in Figure 4B yielded the synaptic strengths listed in Table 1, which were then used to perform the numerical simulations of the spiking, conductance-based model shown in Figure 3. The success of this procedure in yielding a network with persistent neural activity demonstrates that the reduced model is a good approximation to the spiking, conductance-based model.

Drift

Figure 3 demonstrated persistent changes in firing rate and eye position. To characterize persistence more quantitatively, we performed a longer simulation of the conductance-based model, driven by a pseudorandom sequence of saccades. Figure 5A shows the resulting 100 fixations, interrupted by one saccade per second. Six seconds of this trace are shown expanded in Figure 5B. The solid line is the result of a linear fit of eye position versus time during one fixation, excluding the time period immediately after the saccade. The slope of this line was used to calculate the drift velocity of the eye in each fixation. This drift velocity is graphed in the points of Figure 5C, and is generally less than a few degrees per second for eye positions less than 35°. The reduced model also predicts a systematic relationship between drift velocity and eye position (see Equation 23 in the Experimental Procedures), which is graphed as the solid line, and agrees well with the behavior of the conductance-based model.

Neural Coding

The encoding of eye position in firing rate was also quantified in the long simulation. During each intersaccadic interval, the mean firing rate and eye position were computed. The data points of Figure 5D show the rate–position relationship for three of the neurons in the conductance-based model. The solid lines show the rate–position relationships of all the neurons in the reduced model (see Equation 21 in the Experimental Procedures). Again, the conductance-based model and the reduced model agree very well.

For each neuron, the rate is zero for positions less than some threshold value but is approximately linear above this threshold. This is consistent with single-unit recordings of integrator neurons, which usually have a threshold in the normal oculomotor range. The slope of the rate–position relationship, or the position sensitivity, varies from neuron to neuron. Similar heterogeneity of position sensitivities is observed in single-unit recordings of integrator neurons.

Robustness

Our model required careful tuning to allow maintenance of persistent neural activity. Figures 6 and 7 show that deviations from this optimal design cause marked deteriorations in persistence.

In Figure 6A, the strengths of all integrator synapses are reduced to 90% of their tuned values. Fixation ability is impaired: after saccades in both directions, eye position is not persistent but converges rapidly to a null position, as shown at the top. The membrane potential of one of the integrator neurons is shown in the middle panel. After saccades, the firing rate of the neuron always converges to the same value, instead of persisting in time. Unstable behavior is produced if the integrator synapses are increased to 110% of their tuned strengths, as shown in Figure 6B. Both eye position and firing rate increase exponentially at low values, though this explosive instability is limited by saturation.

The simulations of Figure 7 were like those of Figure 5, except that a single neuron was removed. A graph of eye position versus time shows rapid decay at large eye positions (Figure 7A). This causes the range of eye positions in Figure 7A to be smaller than in Figure 5A. From the functional dependence of drift velocity on eye position (Figure 7C), it is evident that there is high drift at eye positions larger than the threshold of the missing neuron. At eye positions below threshold, the drift is unaffected, as can be seen by comparison with Figure 5C.

Discussion

The oculomotor integrator has been modeled as a network of conductance-based neurons interacting by recurrent excitation. The design of the spiking, conductance-based model relied on a nonspiking, reduced model derived using an averaging method. The strength

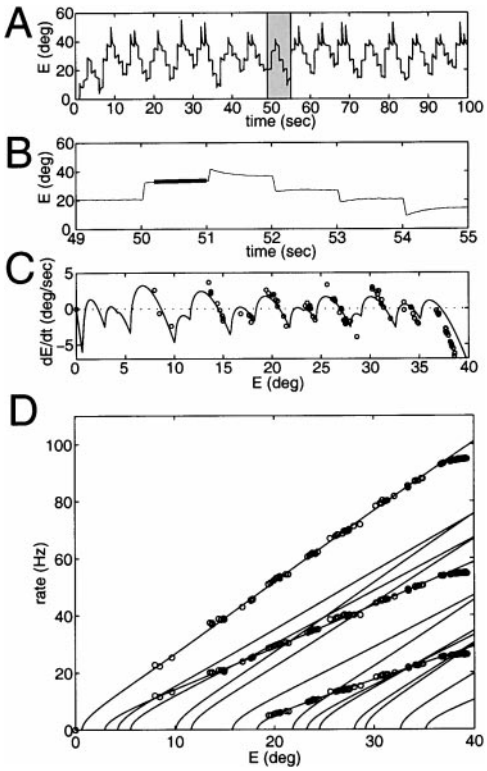

Figure 5. Numerical Simulation of Conductance-Based Model Driven by a Sequence of 100 Saccades at 1 s Intervals

The applied currents of the burst neurons are Gaussian random variables with mean 5 $\mu A/cm^2$ and standard deviation 1 $\mu A/cm^2$. The weights from the burst neurons to the integrator neurons are $W_+ = 0.02$ and $W_- = 0.18$ for all i.

(A) Eye position E versus time. Fixation performance is good over a range of eye positions, up to about 35°.

(B) An expanded view of 6 s of the same time series, along with a sample linear fit to $E(t)$ during an interburst interval, excluding the first 200 ms after the burst. The slope of this line was used to calculate dE/dt in every interburst interval.

(C) This resulted in a graph of the functional relationship between the drift dE/dt and E, shown in the data points. The points are in good agreement with the solid line from Equation 23 of the reduced model. In the range up to about 35°, the drift does not exceed a few degrees per second.

(D) Threshold linear encoding of eye position in the firing rates of integrator neurons. The mean firing rate and mean eye position were computed during each interburst interval for three integrator neurons and are shown in the data points. The solid lines are the rate–position relationships for all 15 integrator neurons, according to the reduced model (see Equation 21). The agreement between the conductance-based model and the reduced model is very good.

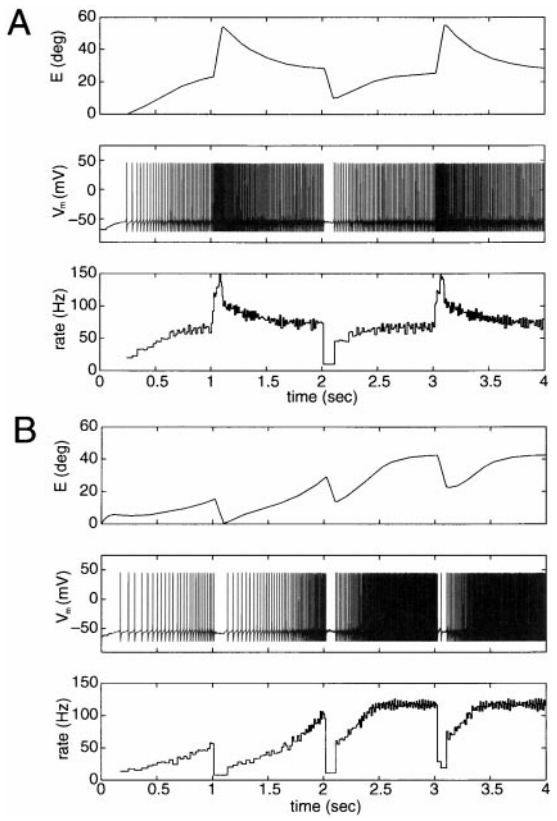

Figure 6. Loss of Persistence Caused by Mistuning of Synaptic Feedback

(A) Loss of persistence caused by reducing recurrent synaptic weights by 10% from their tuned values. The vestibular synapses have also been increased by 10%, to keep the range of eye positions roughly the same as before. The weights from the burst neurons to the integrator neurons are $W_+ = 0.03$ and $W_- = 0.15$ for all i. Each burst is stimulated by a current pulse lasting 100 ms.

(Top) Loss of fixation performance. Eye position is not persistent but decays toward a single null position after every saccade.

(Middle) Membrane potential of a single neuron in the network.

(Bottom) The instantaneous firing rate of the neuron is not persistent but decays to a single stable level after every burst input.

(B) Unstable behavior caused by increasing recurrent synaptic weights by 10% from their tuned values. The weights from the inhibitory burst neuron to the integrator neurons are $W_- = 0.20$ for all i. Each inhibitory burst is stimulated by a current pulse lasting 100 ms.

(Top) Eye position drifts toward positive values after every saccade but saturates at high values.

(Middle) Membrane potential of a single neuron in the network.

(Bottom) The instantaneous firing rate of the neuron increases after every inhibitory burst input. The increase looks exponential at low rates but saturates at high rates.

of synaptic feedback in the reduced model was tuned so that it would maintain persistent activity. When the resulting synaptic strengths (Table 1) were transferred to the conductance-based model, numerical simulations exhibited persistent activity, confirming the validity of the tuning procedure. The encoding of eye position in neural activity was threshold linear, in accord with single-unit recordings of integrator neurons.

Saturation and Recruitment

There are two types of nonlinearity in the reduced model, the threshold for action potential discharge and synaptic saturation. Both types are evident in the neural response function at the bottom of Figure 2B and in Figure 4A. Because of saturation, the amount of feedback contributed by a neuron becomes insufficient as its activity increases. In Figure 4B, this shortfall is compensated by recruitment of neurons that rise above threshold as eye position increases. Therefore, in a properly tuned network, a balance between recruitment and saturation keeps the amount of synaptic feedback sufficient to maintain persistent neural activity.

But the function approximation of Figure 4B is still

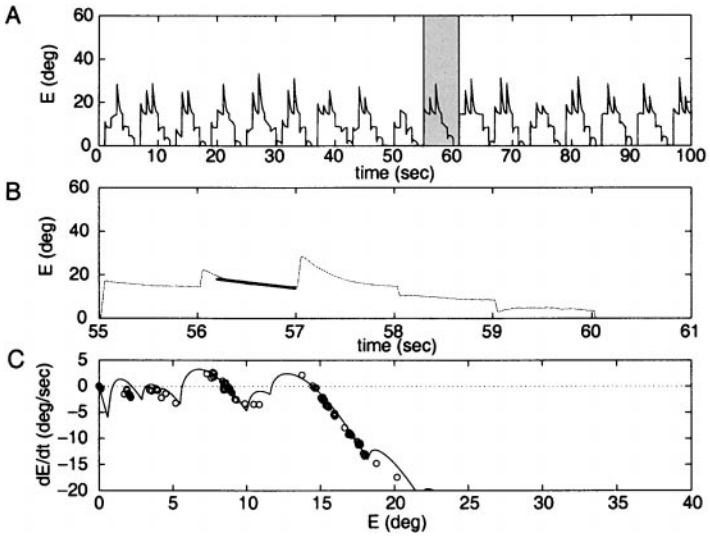

Figure 7. Loss of Persistence after Lesion of a Single Neuron

The network is driven by a pseudorandom sequence of saccades at 1 s intervals with stimulation parameters the same as in Figure 5.

(A) Eye position E versus time. The range of E has shrunk relative to Figure 5, due to increased drift at larger values of E.

(B) An expanded view of 6 s of the same time series, along with a sample linear fit to $E(t)$ during an interburst interval, excluding the first 200 ms after the burst. The slope of this line was used to calculate dE/dt in every interburst interval.

(C) This resulted in a graph of the functional relationship between the drift dE/dt and E, shown in the data points. The solid line is from Equation 23 of the reduced model. Compared to Figure 5C, the drift is left unchanged for E less than the threshold of the lesioned neuron, but is substantially increased for E greater than the threshold.

not perfect, even though the synaptic strengths have been well tuned. Consequently, there is residual drift in neural activity and eye position during fixation (Figure 5C). The drift velocity is fairly low in the range of eye positions between 0° and 35° but rises sharply for eye positions greater than 35°. This large drift velocity can be seen in Figures 5A and 5C. No neurons have thresholds for these large values of eye position, so it is not possible to counterbalance saturation by recruitment. Therefore, the present model predicts that the fixation performance of an integrator should be poor in parts of the oculomotor range where there are no thresholds of integrator neurons, assuming that the recurrent synapses are excitatory and strongly saturating.

Neural Coding

For most integrator neurons, the relationship between firing rate and eye position has a threshold nonlinearity, according to single-unit recordings (Lopez-Barneo et al., 1982; McFarland and Fuchs, 1992). This threshold is reproduced by the present model (Figure 5D) but was lacking in previous linear models. Above threshold, the firing rates of the model neurons encode eye position linearly, again consistent with single-unit recordings. The thresholds and slopes vary over the population of neurons.

The present model demonstrates that a nonsaturating rate–position relationship can be compatible with strongly saturating synapses. Previously, the empirically observed lack of saturation in the rate–position relationship was used to justify linear network models (Cannon et al., 1983) and to reject previous models based on elements with saturation nonlinearity (Rosen, 1972; Kamath and Keller, 1976), but this argument is not valid if synapses saturate.

Excitation versus Inhibition

The anatomical and physiological facts about the synaptic organization of biological integrators are still limited. There is some evidence that ipsilateral recurrent excitation and contralateral inhibition are basic features of this organization (McCrea and Baker, 1985; Escudero and Delgado-Garcia, 1988; Spencer et al., 1989; Escudero et al., 1992; Gamkrelidze et al., 1999, Soc. Neurosci., abstract). Previous integrator models have emphasized the role of contralateral inhibition, noting that mutual inhibition could create the positive feedback required for persistence (Cannon et al., 1983; Galiana and Outerbridge, 1984; Arnold and Robinson, 1997). However, the results of midline lesions have yielded conflicting evidence concerning this issue (Cheron et al., 1986b; Anastasio and Robinson, 1991; Pastor et al., 1994).

As the role of contralateral inhibition is still controversial, we have made our first generation model unilateral for the sake of simplicity. Since the model includes ipsilateral recurrent excitation only, its firing rates always increase with increasing eye position, so that all lines in Figure 5D have positive slope. In other words, pure recurrent excitation leads to position sensitivities all of the same sign, as was previously noted by Cannon et al. (1983). Likewise, single-unit recordings have shown that integrator neurons on the same side of the goldfish brain have position sensitivities of the same sign (Aksay et al., 1997, Soc. Neurosci., abstract).

Our model is a counterexample to the claim by Cannon et al. (1983) that a purely excitatory network cannot maintain persistent activity in the presence of tonic vestibular input. They assumed that the threshold for neural firing is at zero current. In our model neuron, there is a positive threshold for firing, so their argument does not apply.

The Intrinsic Cellular Time Constant

In the present model, as in previous models, the persistence of activity in a single neuron can be characterized by a single cellular time constant τ_{cell}. The role of synaptic feedback is to boost the network time constant τ_{net} to be longer than τ_{cell}. In their linear network models, Robinson and collaborators have assumed that τ_{cell} is a membrane time constant with a value of 5 ms (Cannon et al., 1983; Cannon and Robinson, 1985; Arnold and Robinson, 1991, 1992, 1997). In our model, τ_{cell} corresponds to the

time constant τ_{syn} of recurrent synapses, which we chose to be 100 ms. At present, there is little experimental data concerning synaptic transmission in biological integrators, so our choice is tentative. The value of 100 ms could correspond, for example, to the time constant of the NMDA receptor. Interestingly, it has been reported that injections of APV, an NMDA receptor antagonist, into the neural integrator cause deterioration of fixation performance in cats (Mettens et al., 1994a). However, it is not clear from this experiment whether the affected NMDA receptors belong to recurrent synapses of the integrator.

Whether τ_{cell} is 5 or 100 ms has significant consequences for the robustness of the network to mistuning of parameters. These consequences can be precisely quantified for a linear network as follows. To attain a network time constant τ_{net}, the global strength of synaptic feedback (as quantified by the largest eigenvalue of the synaptic weight matrix) must be tuned to unity with a precision of τ_{cell}/τ_{net} (Seung, 1996). According to measurements of goldfish oculomotor fixation in the dark, the network time constant is typically greater than $\tau_{net} =$ 10 s (Mensh et al., 1997, Soc. Neurosci., abstract). If $\tau_{cell} = 5$ ms, then the required precision is better than 1 part in 2000. If $\tau_{cell} = 100$ ms, only 1 part in 100 is required. Although this argument is strictly valid only for linear networks, it suggests that networks are generally more robust to mistuning when τ_{cell} is longer, other things being equal. As a result, the present model is more robust than the linear network models of Robinson and collaborators, for the most part (some of the subtleties involved in this kind of comparison are discussed by Seung, 1996). Nevertheless, the present model is still sensitive to perturbations, as shown in Figures 6 and 7, and so may not be as robust as biological integrators. Therefore, it seems important to look for mechanisms of synaptic plasticity that could improve robustness by maintaining tuning.

Continuous Attractors

The present model has a number of similarities to the ring model, a recurrent network that is able to store a memory of a directional variable by maintaining persistent activity patterns (Zhang, 1996; Camperi and Wang, 1998). The preferred directions of neurons in the model are evenly spaced on a ring, and the strength of synaptic interaction between two neurons depends only on the difference in their preferred directions (Ben-Yishai et al., 1995). The ring model can integrate angular velocity signals to produce changes in the stored directional variable (Zhang, 1996), much as the present model is able to integrate the pulses of input from command neurons to produce step changes in eye position.

The ring model reproduces the fact that head direction (Zhang, 1996) and direction of a saccadic target (Camperi and Wang, 1998) are encoded in neural firing rates according to a bell-shaped tuning curve. This is analogous to the way in which the present model reproduces the experimental observation that eye position is encoded in firing rates in a threshold linear manner (Lopez-Barneo et al., 1982; Escudero et al., 1992; McFarland and Fuchs, 1992; Scudder and Fuchs, 1992).

In the nonspiking versions of both models, the persistent activity patterns form a continuous family parametrized by the remembered variable (see Equation 16). This continuous family can be idealized as a *continuous dynamical attractor*, a manifold of stable fixed points (Seung, 1996). Strictly speaking, this is just an idealization, as both models contain only a finite number of fixed points at which the drift truly vanishes. The rest of the persistent activity patterns are not truly fixed points but rather points at which drift is very slow (see Figure 5C). Based on considerations of structural stability, it can be argued that any recurrent network model with a continuous attractor must depend on precise tuning of synaptic feedback (Seung, 1996).

Accordingly, corruption of synaptic strengths in either the ring model or the present model results in drift of both neural activity and the remembered variable (compare Figure 6 of this paper and Figure 5 of Zhang, 1996). This shows that both models depend on precise tuning of synaptic feedback. At first glance, the ring model may not appear to require tuning, but in fact its ability to store a memory of a directional variable depends on rotational symmetry of synaptic interactions. With such symmetry, the model is able to maintain a persistent activity pattern centered around any direction on the ring, in the absence of any external signal containing directional information. It is difficult to see how rotational symmetry could be realized in a biological neural network without some mechanism for tuning synaptic strengths.

Dealing with the problem of tuning is the greatest challenge facing network models based on continuous attractors. The tuning procedure followed here is convenient but artificial. It could potentially be replaced by adaptive mechanisms of synaptic plasticity that enable the network to tune itself based on visual or other types of error signals (Arnold and Robinson, 1992, 1997; Seung, 1997, Soc. Neurosci., abstract; Xie and Seung, 2000).

Procedures

Synaptic feedback has been hypothesized to be the mechanism of persistent neural activity in a number of integrator models (Rosen, 1972; Kamath and Keller, 1976). Most of these models have been formulated as networks of linear elements (Cannon et al., 1983; Galiana and Outerbridge, 1984; Cannon and Robinson, 1985; Arnold and Robinson, 1991, 1992, 1997):

$$\tau_{cell}\frac{dv_i}{dt} + v_i = \sum_{j=1}^{N} W_{ij}v_j + B_i. \quad (1)$$

In these equations, the firing rates $v_1,\ldots,v_N$ of a network of N neurons are assumed to interact linearly through synaptic connections of strength W_{ij}, and τ_{cell} is generally interpreted as a membrane time constant. If synaptic strengths are properly tuned, such linear networks can maintain a short-term memory of eye position that is linearly encoded in persistent neural activity. The conditions for tuning can be expressed in terms of the eigenvalues and eigenvectors of the synaptic weight matrix (Seung, 1996). In some models, the synaptic connections are designed (Cannon et al., 1983; Cannon and Robinson, 1985), while in others they emerge spontaneously from a synaptic learning rule (Arnold and Robinson, 1991, 1992, 1997; Seung, 1997, Soc. Neurosci., abstract).

The linear models make two major simplifications that are questionable. They ignore action potentials, describing the activity of each neuron with a dynamical variable like firing rate. Although this

simplification is traditional in neural network models, the range of its validity has not been clear. Furthermore, the model neurons are assumed to be perfectly linear elements. This lends mathematical tractability, at the cost of ignoring the threshold and saturation nonlinearities that are typical of most neurons. Here, we describe the methods used to design a conductance-based model of the integrator, which is more biophysically realistic than previous linear models.

A gentler introduction to the methods described here is available in another paper about the dynamics of a single neuron with an excitatory autapse (Seung et al., 2000). The autapse model is simpler than the present network model, because feedback is localized to a single loop instead of distributed over a web of synaptic connections. The autapse model more simply illustrates the concept of persistence through tuned synaptic feedback, although it cannot reproduce the distributed neural codes that are observed in biological integrators. The autapse model has the further limitation that it relies on a nonsaturating synapse, whereas network models can work with either saturating or nonsaturating synapses (strongly saturating synapses are used here).

The Conductance-Based Model

Dynamics of Action Potentials

According to studies of integrator neurons in vitro, the relationship between firing rate and applied current is approximately linear above threshold over a large range (Serafin et al., 1991; du Lac and Lisberger, 1995; Wang et al., 1998, Soc. Neurosci., abstract). To reproduce this threshold linear relationship, we utilized a model neuron introduced by Shriki et al. (1998, Soc. Neurosci., abstract). This choice was not critical; other model neurons in the literature with similar linear behavior (Ermentrout, 1998a; Wang, 1998) could have been substituted. At present, threshold linearity is the only empirical constraint imposed on the model neuron. A true biophysical model of integrator neurons should become possible in the future, when more information about their intrinsic conductances becomes available.

The voltage V_i of the single compartment of neuron i obeys the current balance equation:

$$C_m \frac{dV_i}{dt} = -I_{L,i} - I_{Na,i} - I_{K,i} - I_{A,i} - I_{syn,i} + I_{app,i}\ , \tag{2}$$

where C_m is the membrane capacitance. The right-hand side contains a leak current I_L, a voltage-dependent transient sodium current I_{Na}, a delayed rectifier potassium current I_K, and an A-type potassium current I_A. The properties of these currents are described completely in the Appendix.

If the synaptic current I_{syn} is zero, and the applied current I_{app} is held constant above a threshold value of 2.046 $\mu A/cm^2$, then the model neuron converges to repetitive firing at a frequency that is approximately linear in applied current (Shriki et al., 1998, Soc. Neurosci., abstract).

The synaptic current $I_{syn,i}$ in neuron i is the sum of excitatory and inhibitory contributions:

$$I_i^{syn} = g_{E,i}\,(V_i - V_E) + g_{I,i}\,(V_i - V_I). \tag{3}$$

Here, $g_{E,i}$ denotes the sum total of excitatory synaptic conductances in neuron i, and $g_{I,i}$ is the sum total of inhibitory synaptic conductances. The voltages $V_E = 0$ and $V_I = -70$ mV are the reversal potentials of excitatory and inhibitory synapses, respectively.

Synaptic Transmission

As described above, the dynamics of intrinsic conductances generate action potentials in the model neuron. These action potentials lead to synaptic transmission and the opening of synaptic receptors in postsynaptic target neurons. The fraction of open channels at each of the synapses made by neuron j on its targets is described by a single dimensionless variable s_j called the *synaptic activation*, which takes values in the range from zero to one. We modeled its dynamics as (Wang and Rinzel, 1992):

$$\tau_{syn,j} \frac{ds_j}{dt} + s_j = \alpha_s \sigma\,(V_j)\,(1 - s_j). \tag{4}$$

Strictly speaking, there should be a synaptic activation s_{ij} and a synaptic time constant $\tau_{syn,ij}$ for each synapse from neuron j to i. However, we assumed that every synapse from neuron j has the same time constant, $\tau_{syn,ij} = \tau_{syn,j}$ for all i. It follows that $s_{ij} = s_j$ for all i, since they are all driven by the same presynaptic voltage V_j.

The synapse is gated by the presynaptic voltage V_j through the sigmoid function:

$$\sigma\,(V) = \frac{1}{1 + \exp[V - \theta_s)/\sigma_s]}. \tag{5}$$

Normally, this gating function is vanishingly small, but it approaches unity briefly whenever a presynaptic action potential drives the voltage over the threshold $\theta_s = -20$ mV. The parameter $\sigma_s = 2$ mV controls the sharpness of the voltage dependence.

Figure 2A illustrates the dynamics of synaptic summation in response to stimulation by a train of action potentials. Each presynaptic action potential causes a sharp rise in the synaptic activation s, roughly as fast as the width of the action potential. This is followed by exponential decay with time constant $\tau_{syn} = 100$ ms. Saturation nonlinearity of the synapse is evident from the fact that the first action potential causes a larger jump in s than later ones. This nonlinearity is due to the $1 - s$ factor on the right-hand side of Equation 4. The size of the increment is controlled by the parameter α_s, which was set at $\alpha_s = 200$ to produce strong saturation.

The time constant of the synaptic activation has been chosen as either slow ($\tau_{syn} = 100$ ms) or fast ($\tau_{syn} = 5$ ms) for each neuron. Recurrent and tonic feedforward excitation is modeled with slow synapses, while feedforward burst excitation is modeled with fast synapses. Likewise, burst inhibition is modeled with fast synapses. This means that all burst input is transient, decaying rapidly after the burst is over. It is convenient to make the burst input fast, in order to clearly illustrate the persistence of the recurrent network, but this is not an essential feature of the model. A more biophysically detailed synaptic model has not been used, as not much is known about synaptic transmission in the integrator. The most important property of the model is its relatively long time constant of 100 ms, which is important for improving the robustness of tuned positive feedback (Seung, 1996).

Synaptic Connectivity

The network simulations contained 15 integrator neurons and 3 input neurons: a vestibular neuron, an excitatory burst neuron, and an inhibitory burst neuron (see Figure 1B). The vestibular neuron was tonically active at a constant rate of roughly 40 Hz, simulating the background activity present in primary vestibular afferents when the head is stationary. This tonic activity was produced by an applied current of $I_{app,0} = 3$ $\mu A/cm^2$. The burst neurons were silent most of the time, except for occasional bursts of activity that stimulated saccadic eye movements. These bursts were produced by 50 ms pulses of applied current, with magnitudes specified in the figure captions.

The synaptic activations of the integrator neurons are denoted by $s_1,...,s_N$; those of the vestibular, excitatory burst, and inhibitory burst neurons are s_0, s_+, and s_-. The excitatory synaptic conductance of the ith integrator neuron is:

$$g_{E,i} = \sum_{j=1}^{N} W_{ij} s_j + W_{i0} s_0 + W_+ s_+, \tag{6}$$

which is composed of recurrent excitation from other integrator neurons, feedforward tonic excitation from the vestibular neuron, and feedforward burst excitation. The inhibitory synaptic conductance:

$$g_{I,i} = W_- s_- \tag{7}$$

comes solely from the inhibitory burst neuron.

Since synaptic activations are dimensionless variables, the synaptic weights have the dimensions of conductance. The weight W_{ij} is the maximal conductance of the synapse from neuron j to i, achieved only if all receptors are open ($s_j = 1$). The recurrent synaptic strengths are of the outer product form:

$$W_{ij} = \xi_i \eta_j, \tag{8}$$

where the values of the parameters $\xi_1,...,\xi_N$ and $\eta_1,...,\eta_N$ are listed in

Table 1. Tuned Values of Parameters in the Network Model

ξ_i	$100\eta_i$	$10B_i$
1.0700	0.9255	0.3623
0.6387	0.5109	0.3506
0.8641	0.0739	0.3328
0.7916	0.6413	0.3258
1.0348	0.4433	0.2651
0.9573	0.3464	0.2580
0.7739	0.4826	0.2472
0.5111	0.4848	0.2760
0.9928	0.3294	0.1526
0.7668	0.0668	0.1909
0.8693	0.3370	0.1563
0.9752	0.2616	0.0961
1.0531	0.2417	0.0627
0.9429	0.3549	0.0617
0.6058	0.3707	0.1563

The vectors ξ_i, η_i, and B_i for the integrator neurons in the network model. The neurons are listed in order of increasing threshold. The units of ξ_i and B_i are conductance (mS/cm^2), when η_i is dimensionless.

Table 1. These parameter values were found by an optimization procedure to be described later.

Every integrator neuron receives feedforward excitation from the vestibular neuron. The strengths of the vestibular synapses are specified in Table 1 in the form $B_i = W_{i0}<s_0>$, which is the time average of the vestibular contribution to the excitatory conductance in Equation 6. Like the recurrent synaptic connections, these were chosen using an optimization procedure to be described later. The average of the synaptic activation of the vestibular neuron is $<s_0> = 0.6465$.

Each integrator neuron receives feedforward excitation and inhibition from the burst neurons. All excitatory burst synapses have the same strength W_+, and all inhibitory burst synapses have the same strength W_-. The numerical values are given in the figure legends. This uniformity is convenient but is not important for the model.

Oculomotor Plant

The oculomotor integrator is a premotor area that projects to motor neurons that drive the extraocular muscles. We modeled the transformation performed by the motor neurons and the oculomotor plant with a first-order linear differential equation:

$$\tau_E \frac{dE}{dt} + E = c\left(\sum_{j=1}^{N} \eta_j s_j + \rho_+ s_+ + \rho_- s_-\right). \quad (9)$$

This amounts to a low-pass filtering of a linear combination of the synaptic activations of the integrator and burst neuron. This model should not be interpreted literally as saying that the integrator and burst neurons project directly to the extraocular muscles. Rather, it is a mathematical approximation of the effect of the neural circuitry and musculature intervening between these neurons and the eye.

To produce a step change in eye position, the oculomotor plant must be driven by "pulse step" input (Robinson, 1964). The pulse moves the eye to a new position, overcoming viscous drag forces. The step holds the new eye position. Here, the viscous drag is characterized by the time constant $\tau_E = 150$ ms of the oculomotor plant. The step is provided by the integrator neurons (s_1,...,s_N), and the pulse is provided by the burst neurons (s_+, s_-). The parameter values $\rho_+ = 0.12$ and $\rho_- = -0.07$ were set by hand so that burst inputs produced nearly perfect step changes in eye position. Note that the η_j in Equation 9 are the same parameters that appear in the outer product form (Equation 8) of the recurrent connections. The prefactor of $c = 1000°$ was set to give an oculomotor range of 40°.

The Reduced Model

An averaging method was applied to construct a nonspiking, reduced model from the conductance-based model. The method is similar to others that were proposed previously (Rinzel and Frankel, 1992; Ermentrout, 1994, 1998b) but performs better for strongly saturating synapses, which is the relevant case here.

We replaced $\sigma(V_j)$ in Equation 4 by its time average for fixed $g_{E,j}$. This replacement becomes exact in the limit of infinitely slow synapses ($\tau_{syn} \to \infty$) and is often a good approximation for large but finite τ_{syn}. The averaging method was applied to the conductance-based model during the time intervals between bursts, when the fast excitatory and inhibitory burst synapses are inactive. Therefore, only the recurrent synapses, which are slow and excitatory, are considered in the following discussion.

For g_E above a threshold value, the model neuron converges rapidly to repetitive firing at a constant frequency, as shown in Figure 2A. Both the membrane potential $V(t;g_E)$ and synaptic activation $s(t;g_E)$ are periodic in time. Let $<u(t)> = T(g_E)^{-1}\int_0^{T(g_E)} dt\, u(t)$ denote the time average of the periodic function $u(t)$ over a single interspike interval of length $T(g_E)$, and define:

$$f(g_E) = \frac{<(1 - s(t; g_E))\sigma(V(t; g_E))>}{<1 - s(t; g_E)>} = \frac{1}{\alpha_s}\frac{<s(t; g_E)>}{<1 - s(t; g_E)>} \quad (10)$$

In words, $f(g_E)$ is the time average of $\sigma(V(t;g_E))$, weighted by $1 - s(t;g_E)$. This definition differs from the unweighted average $<\sigma(V(t; g_E))>$ used in previous applications of the averaging method (Ermentrout, 1994, 1998b). A theoretical justification for our use of the weighted average (Equation 10) is outside the scope of this paper. The empirical justification is simply that the weighted average gave superior results. When the synaptic strengths of the conductance-based model were tuned using a reduced model obtained from the standard method of averaging, the persistence time of neural activity was poor. But tuning based on a reduced model obtained from our nonstandard method was successful, as illustrated by the results reported in this paper. Note that both versions of the method of averaging give the same result in the limit as $\tau_{syn} \to \infty$, since $1 - s(t;g_E)$ becomes constant in time.

The second definition in Equation 10 can be proven equivalent to the first by averaging Equation 4 over a period and is more convenient in some circumstances. It follows from the second definition that our nonstandard method of averaging yields neural network equations with the same fixed points as the mean field approximation (Shriki et al., 1998, Soc. Neurosci., abstract), another method of deriving neural network equations from conductance-based models. The standard method of averaging and the mean field approximation give the same fixed point equations for nonsaturating synapses, but not for saturating synapses.

The shape of f is shown at the top of Figure 2B, and was found by numerically simulating repetitive firing for various values of g_E and computing the time average in Equation 10. It turns out that the function $f(g_E)$ is almost exactly proportional to firing rate ν:

$$f(g_E) \approx 0.229\ \text{kHz}^{-1}\ \nu(g_E). \quad (11)$$

Both functions are approximately linear in g_E.

Substituting $f(g_{E,j})$ for $\sigma(V_j)$ in Equation 4 yields the averaged equation of motion:

$$\tau_{syn,j}\frac{ds_j}{dt} + s_j = \alpha_s\,(1 - s_j) f(g_{E,j}). \quad (12)$$

This equation can be regarded as a simplified description of the neural dynamics portrayed in Figure 2A. One way of interpreting this figure is to regard the neuron as two devices cascaded in series. The first device consists of the intrinsic conductances. It transforms the excitatory synaptic conductance g_E (Figure 2A, top) into a train of action potentials V (Figure 2A, middle). The second device is the synapse, which transforms the presynaptic action potentials into the synaptic activation s (Figure 2A, bottom).

The averaged equation of motion describes the transformation from g_E to s directly, omitting the intermediate stage of the membrane potential V. Because the complex dynamics of intrinsic conductances have been neatly packaged into the function f, the dynamics of the reduced model (Equation 12) is quite simple. If g_E is held constant in time, the synaptic activation s exponentially approaches:

$$F(g_E) = \frac{\alpha_s f(g_E)}{1 + \alpha_s f(g_E)} \tag{13}$$

Judging from the bottom of Figure 2A, the reduced model (Equation 12) accurately captures the overall exponential behavior of s but smooths over the sudden jumps due to action potentials. This smoothing is in the same spirit as the classical neural network equations (Wilson and Cowan, 1972; Hopfield and Tank, 1986), which replaced trains of discrete action potentials with continuous variables like instantaneous firing rate or temporally smoothed membrane potential. The effects of synaptic saturation are evident in F (Figure 2B), unlike in the function f.

Substituting Equation 6 for the synaptic conductance $g_{E,i}$ into Equation 12 yields:

$$\tau_{syn,i} \frac{ds_i}{dt} + s_i = \alpha_s (1 - s_i) f\left(\sum_{j=1}^{N} W_{ij} s_j + B_i \right). \tag{14}$$

The tonic input from the vestibular neurons is in the bias term $B_i = W_{i0} <s_0>$. Excitatory and inhibitory burst input have been neglected, as the equations are only meant to approximate the conductance-based model during the time intervals between bursts.

Tuning the Reduced Model

The fixed point equations of the reduced model are:

$$s_i = F\left(\sum_{j=1}^{N} W_{ij} s_j + B_i \right), \tag{15}$$

which is obtained by setting $ds_i/dt = 0$ in Equation 14 and using the definition in Equation 13. Since we have assumed the outer product form $W_{ij} = \xi_i \eta_j$ for the synaptic weight matrix (see Equation 8), the fixed point equations (Equation 15) are approximately satisfied if:

$$s_i \approx F(\xi_i \hat{E} + B_i), \tag{16}$$

$$\hat{E} = \sum_i \eta_i s_i. \tag{17}$$

These conditions can in turn be satisfied if:

$$\hat{E} \approx \sum_{i=1}^{N} \eta_i F(\xi_i \hat{E} + B_i). \tag{18}$$

Hence, we tuned parameters to minimize the squared difference between the two sides of Equation 18 over a range of values of $\hat{E}$. Each term in the sum is represented graphically in Figure 4A, and the cumulative sums are shown in Figure 4B.

The components of the vector ξ_i were chosen randomly from a uniform distribution in the interval (0.5,1.1) and rounded to four decimal places. Thresholds θ_i were chosen from 0 to 0.037 in steps of 0.0025, and then a Gaussian noise with standard deviation 0.001 was added. The biases were computed via $B_i = 0.0368 - \theta_i/\xi_i$ and rounded to five decimal places. The η_i were determined by minimizing the squared difference between the two sides of the equation at values of $\hat{E}$ from 0 to 0.038 in steps of 0.0001. The η_i were rounded to the values shown in Table 1. It was verified that all parameters ξ_i, η_i, and B_i were nonnegative, in keeping with the assumption of excitatory connections. From ξ_i and η_i, the synaptic weight matrix was constructed via $W_{ij} = \xi_i \eta_j$. The strengths of the vestibular synapses $W_{i0} = B_i/<s_0>$ were calculated from the bias B_i.

The tuned values given in Table 1 are special, but not unique. There are other settings of parameters that give the network comparable persistence. Persistence does not require local tuning of each and every synaptic strength to a unique value, but only that the strength of feedback be globally tuned.

Neural Coding of Eye Position

The quantity $\hat{E}$ introduced in Equation 17 plays an important role in the reduced model: it is the *internal representation* of eye position. This can be seen by substituting Equation 17 in Equation 9 to obtain:

$$\tau_E \frac{dE}{dt} + E = c\hat{E}. \tag{19}$$

This is valid during intersaccadic time intervals, when the terms in Equation 9 containing s_+ and s_- can be neglected. Provided that $\hat{E}$ changes slowly, the approximation $E \approx c\hat{E}$ holds, so that the oculomotor plant "reads out" $\hat{E}$ to produce eye position E.

During intersaccadic intervals, the synaptic conductances of the integrator neurons are functions of $\hat{E}$:

$$g_{E,i} = \xi_i \hat{E} + B_i. \tag{20}$$

This expression is derived from Equation 6 by substituting the outer product form (Equation 8), using the definition (Equation 17), and neglecting the burst input. Combining this with Equation 11 yields $\nu_i \propto f(\xi_i \hat{E} + B_i)$. Upon substituting $\hat{E} \approx E/c$, we find the relationship between firing rate and eye position:

$$\nu_i \propto f(\xi_i E/c + B_i) \tag{21}$$

which is graphed in the solid lines of Figure 5D for all 15 integrator neurons. The lines are straight because f is approximately linear (Figure 2B, top).

This equation is remarkable, because it specifies how the parameters of the model relate to the properties that are measured in single-unit recordings of integrator neurons. The slope of each line is the *position sensitivity* of the neuron and is determined by the recurrent synaptic connections (through ξ_i). The threshold eye position depends not only on the recurrent synapses (through ξ_i) but also the feedforward vestibular synapses (through $B_i = W_{i0} <s_0>$).

Drift in Neural Activity and Eye Position

The difference between the two sides of Equation 18 has an important dynamical interpretation. To derive it, we note that the variable $\hat{E}$ obeys the equation:

$$\tau_{syn} \frac{d\hat{E}}{dt} + \hat{E} = \sum_i \eta_i (1 - s_i) f(\xi_i \hat{E} + B_i). \tag{22}$$

This is obtained by multiplying the reduced model (Equation 14) by η_i, summing over i, substituting the outer product form $W_{ij} = \xi_i \eta_j$ for the synaptic weight matrix, and using the definition (Equation 17) of $\hat{E}$. For a tuned network, the approximation (Equation 16) for s_i can be substituted, yielding:

$$\tau_{syn} \frac{d\hat{E}}{dt} + \hat{E} \approx \sum_i \eta_i F(\xi_i \hat{E} + B_i) \tag{23}$$

So the difference between the two sides of Equation 18 is proportional to the drift $d\hat{E}/dt$ in the variable $\hat{E}$.

According to Equation 23, $d\hat{E}/dt$ vanishes at the zero crossings of the solid line in Figure 5C. These correspond to fixed points of the reduced model. The zero crossings with negative slope are stable fixed points, as can be shown by a linear stability analysis of Equation 23. There is some clustering of the values of $\hat{E}$ around these stable fixed points in Figures 5C and 5D. The presence of these fixed points should not be overemphasized. They are not in themselves important to the function of the integrator. Only low drift is important for fixation, and fixed points are just a side effect of the tuning required to make $d\hat{E}/dt$ small.

The existence of multiple stable fixed points is possible because of the nonlinearity of F in Equation 18. For a linear network, the relationship between dE/dt and E would be linear, so that only zero or one fixed point would be possible. In an optimally tuned network, the number of stable fixed points presumably grows with the size of the network.

Rates versus Spikes

The success of the tuning procedure in producing persistent activity shows that the reduced model is a very good approximation to the conductance-based model, but it is by no means perfect. For example, we have found (but do not report fully here) oscillatory behaviors in the firing rates of the conductance-based model that are not present in the reduced model. This discrepancy is a sign of a breakdown in the method of averaging. It is well known that the method encounters problems when applied to systems in which natural frequencies are in whole-number ratios with each other, a phenomenon known as *resonance* (Sanders and Verhulst, 1985). If the spiking vestibular input is removed and replaced by a constant bias term in the synaptic conductance of the integrator neurons,

the oscillations become much weaker. This indicates that the main resonance in our simulations is between the integrator neurons and the vestibular neuron. However, the oscillations do not vanish completely when the synchronous drive of the vestibular neuron is removed, so apparently there are also resonances between integrator neurons. It would be interesting to look for such effects in experimental measurements of integrator neural activity.

The method of averaging was applicable because of the slow synapses in the model. If the synapses are made faster while the size of the network is fixed, the reduced model becomes an inaccurate approximation, so that the present tuning procedure breaks down, though other tuning procedures might be possible. The problem is mitigated if the number of neurons is increased, so that it becomes possible to construct an accurate reduced model by averaging over neurons instead of time (Shriki et al., 1998, Soc. Neurosci., abstract).

In the present work, the spiking nature of neural activity does not play a major role; the nonspiking, reduced model is a good approximation to the spiking, conductance-based model. It should be noted that there is a different conception of persistent activity based on synchronous volleys of spikes circulating in a network, though it is not clear how to make such a "synfire theory" (Abeles, 1991) compatible with the experimental facts about neural coding in the integrator.

Appendix: Intrinsic Conductances

Our simulations utilize a model neuron introduced by Shriki, Hansel, and Sompolinsky (Shriki et al., 1998, Soc. Neurosci., abstract). The only modification we have made is to increase the threshold by using a higher leak conductance. Unless otherwise noted, the measurement units are voltage (mV), conductance (mS/cm^2), current (μA/cm^2), and capacitance (μF/cm^2).

Current Balance Equation

$$C_m \frac{dV_i}{dt} = -I_L(V_i) - I_{Na}(V_i, h_i) - I_K(V_i, n_i) - I_A(V_i, b_i) - I_{syn}(V_i, \{s_j\}) + I_{app,i} \quad (24)$$

$C_m = 1$ μF/cm^2.

Intrinsic Conductances

Leak Current I_L.

$$I_L(V) = g_L(V - V_L) \quad (25)$$

$g_L = 0.2$, $V_L = -65$.

Sodium Current I_{Na}.

$$I_{Na}(V,h) = g_{Na} m_\infty^3(V) h (V - V_{Na}) \quad (26)$$

$$m_\infty(V) = \frac{\alpha_m(V)}{\alpha_m(V) + \beta_m(V)} \quad (27)$$

$$\alpha_m(V) = \frac{(V + 30)/10}{1 - \exp[-(V + 30)/10]} \quad (28)$$

$$\beta_m(V) = 4 \exp[-(V + 55)/18] \quad (29)$$

$$\phi_h^{-1} \frac{dh}{dt} = \alpha_h(V)(1 - h) - \beta_h(V)h \quad (30)$$

$$\alpha_h(V) = 0.07 \exp[-(V + 44)/20] \quad (31)$$

$$\beta_h(V) = \frac{1}{\exp[-(V + 14)/10] + 1} \quad (32)$$

$g_{Na} = 100$, $V_{Na} = 55$, $\phi_h = 10$.

Delayed Rectifier Potassium Current I_K.

$$I_K(V) = g_K n^4 (V - V_K) \quad (33)$$

$$\phi_n^{-1} \frac{dn}{dt} = \alpha_n(V)(1 - n) - \beta_n(V)n \quad (34)$$

$$\alpha_n(V) = \frac{(V + 34)/100}{1 - \exp[-(V + 34)/10]} \quad (35)$$

$$\beta_n(V) = \frac{1}{8} \exp[-(V + 44)/80] \quad (36)$$

$g_K = 40$, $V_K = -80$, $\phi_n = 10$.

A-Type Potassium Current I_A. The activation variable a is instantaneous, while the inactivation variable b has a relaxation time τ_b that is independent of voltage.

$$I_A(V) = g_A a_\infty^3 b (V - V_K) \quad (37)$$

$$a_\infty(V) = \frac{1}{\exp[-(V + 50)/20] + 1} \quad (38)$$

$$b_\infty(V) = \frac{1}{\exp[(V + 80)/6] + 1} \quad (39)$$

$$\frac{db}{dt} = \frac{b_\infty(V) - b}{\tau_b} \quad (40)$$

$g_A = 20$, $\tau_b = 20$.

We used the fourth-order Runge-Kutta method with step size 0.01 ms to integrate these equations, except in Figure 2B, where a step size of 0.002 ms was used. With no synaptic or applied current, the dynamical variables converge to a fixed point at $V = -68.3737$, $h = 0.9820$, $n = 0.0631$, and $b = 0.1259$.

Instantaneous rate functions were calculated from spike times defined as the downward zero crossings of the membrane potential. The rate between successive spikes at times t_a and t_{a+1} was defined as $1/(t_{a+1} - t_a)$. In other words, the rate function was constant during each interspike interval.

Acknowledgments

We thank O. Shriki, H. Sompolinsky, and D. Hansel for providing us with their model neuron, and X.-J. Wang for discussions on synaptic models. E. Aksay and X. Xie provided helpful comments on the manuscript.

Received June 18, 1999; revised February 8, 2000.

References

Abeles, M. (1991). Corticonics: Neural Circuits of the Cerebral Cortex (Cambridge: Cambridge University).

Anastasio, T.J., and Robinson, D.A. (1991). Failure of the oculomotor neural integrator from a discrete midline lesion between the abducens nuclei in the monkey. Neurosci. Lett. *127*, 82–86.

Arnold, D.B., and Robinson, D.A. (1991). A learning network model of the neural integrator of the oculomotor system. Biol. Cybern. *64*, 447–454.

Arnold, D.B., and Robinson, D.A. (1992). A neural network model of the vestibulo-ocular reflex using a local synaptic learning rule. Philos. Trans. R. Soc. Lond. B Biol. Sci. *337*, 327–330.

Arnold, D.B., and Robinson, D.A. (1997). The oculomotor integrator: testing of a neural network model. Exp. Brain Res. *113*, 57–74.

Becker, W., and Klein, H.M. (1973). Accuracy of saccadic eye movements and maintenance of eccentric eye positions in the dark. Vision Res. *13*, 1021–1034.

Belknap, D.B., and McCrea, R.A. (1988). Anatomical connections of the prepositus and abducens nuclei in the squirrel monkey. J. Comp. Neurol. *268*, 13–28.

Ben-Yishai, R., Bar-Or, R.L., and Sompolinsky, H. (1995). Theory of orientation tuning in visual cortex. Proc. Nat. Acad. Sci. USA *92*, 3844–3848.

Camperi, M., and Wang, X.J. (1998). A model of visuospatial working memory in prefrontal cortex: recurrent network and cellular bistability. J. Comput. Neurosci. *5*, 383–405.

Cannon, S.C., and Robinson, D.A. (1985). An improved neural-network model for the neural integrator of the oculomotor system: more realistic neuron behavior. Biol. Cybern. *53*, 93–108.

Cannon, S.C., and Robinson, D.A. (1987). Loss of the neural integrator of the oculomotor system from brain stem lesions in monkey. J. Neurophysiol. *57*, 1383–1409.

Cannon, S.C., Robinson, D.A., and Shamma, S. (1983). A proposed neural network for the integrator of the oculomotor system. Biol. Cybern. *49*, 127–136.

Cheron, G., and Godaux, E. (1987). Disabling of the oculomotor neural integrator by kainic acid injections in the prepositus-vestibular complex of the cat. J. Physiol. *394*, 267–290.

Cheron, G., Gillis, P., and Godaux, E. (1986a). Lesions in the cat

prepositus complex: effects on the optokinetic system. J. Physiol. *372*, 95–111.

Cheron, G., Godaux, E., Laune, J.M., and Vanderkelen, B. (1986b). Lesions in the cat prepositus complex: effects on the vestibulo-ocular reflex and saccades. J. Physiol. *372*, 75–94.

du Lac, S., and Lisberger, S.G. (1995). Membrane and firing properties of avian medial vestibular nucleus neurons in vitro. J. Comp. Physiol. [A] *176*, 641–651.

Ermentrout, B. (1994). Reduction of conductance-based models with slow synapses to neural nets. Neural Comput. *6*, 679–695.

Ermentrout, B. (1998a). Linearization of f-i curves by adaptation. Neural Comput. *10*, 1721–1729.

Ermentrout, B. (1998b). Neural networks as spatio-temporal pattern-forming systems. Rep. Prog. Phys. *61*, 353–430.

Escudero, M., and Delgado-Garcia, J.M. (1988). Behavior of reticular, vestibular and prepositus neurons terminating in the abducens nucleus of the alert cat. Exp. Brain Res. *71*, 218–222.

Escudero, M., de la Cruz, R.R., and Delgado-Garcia, J.M. (1992). A physiological study of vestibular and prepositus hypoglossi neurones projecting to the abducens nucleus in the alert cat. J. Physiol. *458*, 539–560.

Fuster, J.M. (1995). Memory in the Cerebral Cortex (Cambridge, MA: MIT Press).

Galiana, H.L., and Outerbridge, J.S. (1984). A bilateral model for central neural pathways in vestibuloocular reflex. J. Neurophysiol. *51*, 210–241.

Godaux, E., Cheron, G., and Gravis, F. (1989). Eye movements evoked by microstimulations in the brainstem of the alert cat. Exp. Brain Res. *77*, 94–102.

Godaux, E., Mettens, P., and Cheron, G. (1993). Differential effect of injections of kainic acid into the prepositus and vestibular nuclei of the cat. J. Physiol. *472*, 459–482.

Grossberg, S. (1988). Nonlinear neural networks: principles, mechanisms, and architectures. Neural Networks *1*, 17–61.

Hess, K., Reisine, H., and Dursteler, M. (1985). Normal eye drift and saccadic drift correction in darkness. Neuroophthalmology *5*, 247–252.

Hopfield, J.J. (1982). Neural networks and physical systems with emergent collective computational abilities. Proc. Nat. Acad. Sci. USA *79*, 2554–2558.

Hopfield, J.J., and Tank, D.W. (1986). Computing with neural circuits: a model. Science *233*, 625–633.

Kamath, B.Y., and Keller, E.L. (1976). A neurological integrator for the oculomotor control system. Math. Biosci. *30*, 341–352.

Kaneko, C.R.S. (1997). Eye movement deficits after ibotenic acid lesions of the nucleus prepositus hypoglossi in monkeys. I. Saccades and fixation. J. Neurophysiol. *78*, 1753–1768.

Kaneko, C.R.S. (1999). Eye movement deficits following ibotenic acid lesions of the nucleus prepositus hypoglossi in monkeys. II. Pursuit, vestibular, and optokinetic responses. J. Neurophysiol. *81*, 668–681.

Lee, D.D., Reis, B.Y., Seung, H.S., and Tank, D.W. (1997). Nonlinear network models of the oculomotor integrator. In Computational Neuroscience: Trends in Research, J.M. Bower, ed. (New York: Plenum Press).

Lopez-Barneo, J., Darlot, C., Berthoz, A., and Baker, R. (1982). Neuronal activity in prepositus nucleus correlated with eye movement in the alert cat. J. Neurophysiol. *47*, 329–352.

McCrea, R.A., and Baker, R. (1985). Anatomical connections of the nucleus prepositus of the cat. J. Comp. Neurol. *237*, 377–407.

McFarland, J.L., and Fuchs, A.F. (1992). Discharge patterns in nucleus prepositus hypoglossi and adjacent medial vestibular nucleus during horizontal eye movement in behaving macaques. J. Neurophysiol. *68*, 319–332.

Mettens, P., Cheron, G., and Godaux, E. (1994a). NMDA receptors are involved in temporal integration in the oculomotor system of the cat. Neuroreport *5*, 1333–1336.

Mettens, P., Godaux, E., Cheron, G., and Galiana, H.L. (1994b). Effect of muscimol microinjections into the prepositus hypoglossi and the medial vestibular nuclei on cat eye movements. J. Neurophysiol. *72*, 785–802.

Moschovakis, A.K. (1997). The neural integrators of the mammalian saccadic system. Front. Biosci. *2*, d552–d577.

Pastor, A.M., de La Cruz, R.R., and Baker, R. (1994). Eye position and eye velocity integrators reside in separate brainstem nuclei. Proc. Natl. Acad. Sci. USA *91*, 807–811.

Rinzel, J., and Frankel, P. (1992). Activity patterns of a slow synapse network predicted by explicitly averaging spike dynamics. Neural Comput. *4*, 534–545.

Robinson, D.A. (1964). The mechanics of human saccadic eye movement. J. Physiol. *174*, 245–264.

Robinson, D.A. (1989). Integrating with neurons. Annu. Rev. Neurosci. *12*, 33–45.

Rosen, M.J. (1972). A theoretical neural integrator. IEEE Trans. Biomed. Engin. *19*, 362–367.

Sanders, J.A., and Verhulst, F. (1985). Averaging Methods in Nonlinear Dynamical Systems: Applied Mathematical Sciences (New York: Springer-Verlag).

Scudder, C.A., and Fuchs, A.F. (1992). Physiological and behavioral identification of vestibular nucleus neurons mediating the horizontal vestibuloocular reflex in trained rhesus monkeys. J. Neurophysiol. *68*, 244–264.

Serafin, M., Waele, C.D., Khateb, A., Vidal, P.P., and Muhlethaler, M. (1991). Medial vestibular nucleus in the guinea-pig. I. Intrinsic membrane properties in brainstem slices. Exp. Brain Res. *84*, 417–425.

Seung, H.S. (1996). How the brain keeps the eyes still. Proc. Natl. Acad. Sci. USA *93*, 13339–13344.

Seung, H.S., Lee, D.D., Reis, B.Y., and Tank, D.W. (2000). The autapse: a simple illustration of short-term analog memory storage by tuned synaptic feedback. J. Comput. Neurosci., in press.

Spencer, R.F., Wenthold, R.J., and Baker, R. (1989). Evidence for glycine as an inhibitory neurotransmitter of vestibular, reticular, and prepositus hypoglossi neurons that project to the cat abducens nucleus. J. Neurosci. *9*, 2718–2736.

Straube, A., Kurzan, R., and Büttner, U. (1991). Differential effects of bicuculline and muscimol microinjections into the vestibular nuclei on simian eye movements. Exp. Brain Res. *86*, 347–358.

Tiliket, C., Shelhamer, M., Roberts, D., and Zee, D.S. (1994). Short-term vestibulo-ocular reflex adaptation in humans. I. Effect on the ocular velocity-to-position neural integrator. Exp. Brain Res. *100*, 316–327.

Wang, X.J. (1998). Calcium coding and adaptive temporal computation in cortical pyramidal neurons. J. Neurophysiol. *79*, 1549–1566.

Wang, X.J., and Rinzel, J. (1992). Alternating and synchronous rhythms in reciprocally coupled inhibitory model neurons. Neural Comput. *4*, 534–545.

Wilson, H.R., and Cowan, J.D. (1972). Excitatory and inhibitory interactions in localized populations of model neurons. Biophys J. *12*, 1–24.

Xie, X., and Seung, H.S. (2000). Spike-based learning rules and stabilization of persistent neural activity. Adv. Neural Info. Process. Syst. *12*, in press.

Yokota, J.I., Reisine, H., and Cohen, B. (1992). Nystagmus induced by electrical stimulation of the vestibular and prepositus hypoglossi nuclei in the monkey. Exp. Brain Res. *92*, 123–138.

Zhang, K. (1996). Representation of spatial orientation by the intrinsic dynamics of the head-direction cell ensemble: a theory. J. Neurosci. *16*, 2112–2126.