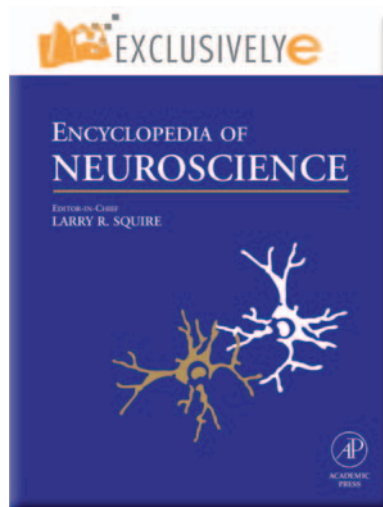


Provided for non-commercial research and educational use.
Not for reproduction, distribution or commercial use.

This article was originally published in the *Encyclopedia of Neuroscience* published by Elsevier, and the attached copy is provided by Elsevier for the author's benefit and for the benefit of the author's institution, for non-commercial research and educational use including without limitation use in instruction at your institution, sending it to specific colleagues who you know, and providing a copy to your institution's administrator.



All other uses, reproduction and distribution, including without limitation commercial reprints, selling or licensing copies or access, or posting on open internet sites, your personal or institution's website or repository, are prohibited. For exceptions, permission may be sought for such use through Elsevier's permissions site at:

<http://www.elsevier.com/locate/permissionusematerial>

Goldman M S, Compte A and Wang X -J (2009) Neural Integrator Models. In: Squire LR (ed.) *Encyclopedia of Neuroscience*, volume 6, pp. 165-178. Oxford: Academic Press.

Neural Integrator Models

M S Goldman, University of California at Davis, Davis, CA, USA

A Compte, Institut d'Investigacions Biomèdiques August Pi i Sunyer, Barcelona, Spain

X-J Wang, Yale University School of Medicine, New Haven, CT, USA

© 2009 Elsevier Ltd. All rights reserved.

Introduction

Since the 1906 publication of Charles Sherrington's *The Integrative Action of the Nervous System*, the concept of neurons as integrators of synaptic inputs has become a central tenet of neuroscience. Neural circuits integrate information in various ways: summing spatially, receiving converging inputs from different sensory modalities, bridging sensory and motor pathways, combining reward and cognitive signals, and so forth. Integration of information in the temporal domain, the topic of this article, has been increasingly recognized in recent years to be important for a variety of brain functions. Time integration is defined as the mathematical accumulation of an input signal so that the resulting output reflects the running total of the input. Networks in the nervous system can exhibit a great diversity in the 'integration time' during which they compute and store this running total. Typically, neurons are leaky devices with an integration time of tens of milliseconds. However, some neurons or neural networks can have a very short integration time constant (a few milliseconds or less); such 'coincidence detectors' can be activated only by synaptic events that are precisely synchronized. At the other extreme, neurons of certain brain areas behave as if their characteristic time constant was effectively infinite. Not only do they perform time integration of inputs without noticeable leak during stimulation but also they maintain the memory of the integrated signal after the stimulus is withdrawn, in the form of self-sustained persistent activity.

The first neural integrators were discovered in the oculomotor system in the 1960s when it was realized that motor command neurons signaled the velocity of eye movement, whereas downstream motor neurons encoded the eye position (Figure 1(a)). Therefore, an integration operation is required to convert velocity signals into position signals. Functionally, such neural integrators are well suited to implement oculomotor control, just like integrators commonly used in feedback control systems by engineers. Similarly, in the head direction system, which is part of the spatial

navigation system and signals the animal's directional heading, the inputs from the semicircular canals encode the head velocity, which is integrated by head direction cells into the positional signal that conveys the orientation of an animal's head in space (Figure 1(b)). When the head is still, head direction cells show elevated and directionally tuned persistent activity in the absence of external inputs.

These examples illustrate two encoding schemes of neural integrators – 'rate code' and 'location code.' In a rate code, exemplified by the oculomotor system (Figure 1(a)), the integrator accumulates the signal in the firing rate of the neurons and the persistent firing rate of each neuron varies monotonically with the encoded feature. By contrast, in a location code as used in the head direction system (Figure 1(b)), each neuron exhibits a bell-shaped tuning curve of an encoded feature with peak response at a particular 'preferred' feature. When a particular value of the encoded feature (e.g., a particular head direction) is being stored, neurons with preferred features similar to this value will be highly excited and neurons with more dissimilar preferred features will be less excited. This creates a localized activity profile across the population known as a 'bump' of neural activity, and the encoded feature (e.g., the current direction of the head) can be read out from the peak location of this bump. If velocity signals smoothly move this bump of activity around the neural population, then the location of the bump can represent the integral of the velocity input. In other words, the stimulus feature is specified by 'which neurons are active' in a location code and by 'how high the neural activities are' in a rate code. Mathematically, a perfect integrator is described by $dX/dt = I(t)$, where $I(t)$ is the input, and X is either the peak locus of network activity (in a location code) or the neural firing activity level (in a rate code).

Based on the mathematical properties of integrals in the sense of calculus, neural integrators are expected to be capable of certain computations of special importance to higher cognitive functions. For example, if a signal suddenly vanishes, its integral remains at exactly the same level as that before disappearing. This is suggestive of a memory system. Indeed, in a delayed discrimination task in which a monkey is required to actively remember the frequency of a vibrotactile stimulus during a memory period of a few seconds, neurons in the prefrontal cortex (a cortical area known to play a critical role in working memory) have been found to exhibit graded persistent activity that changes monotonically with the intensity of the transient somatosensory stimulus (Figure 1c). As another example, the integral

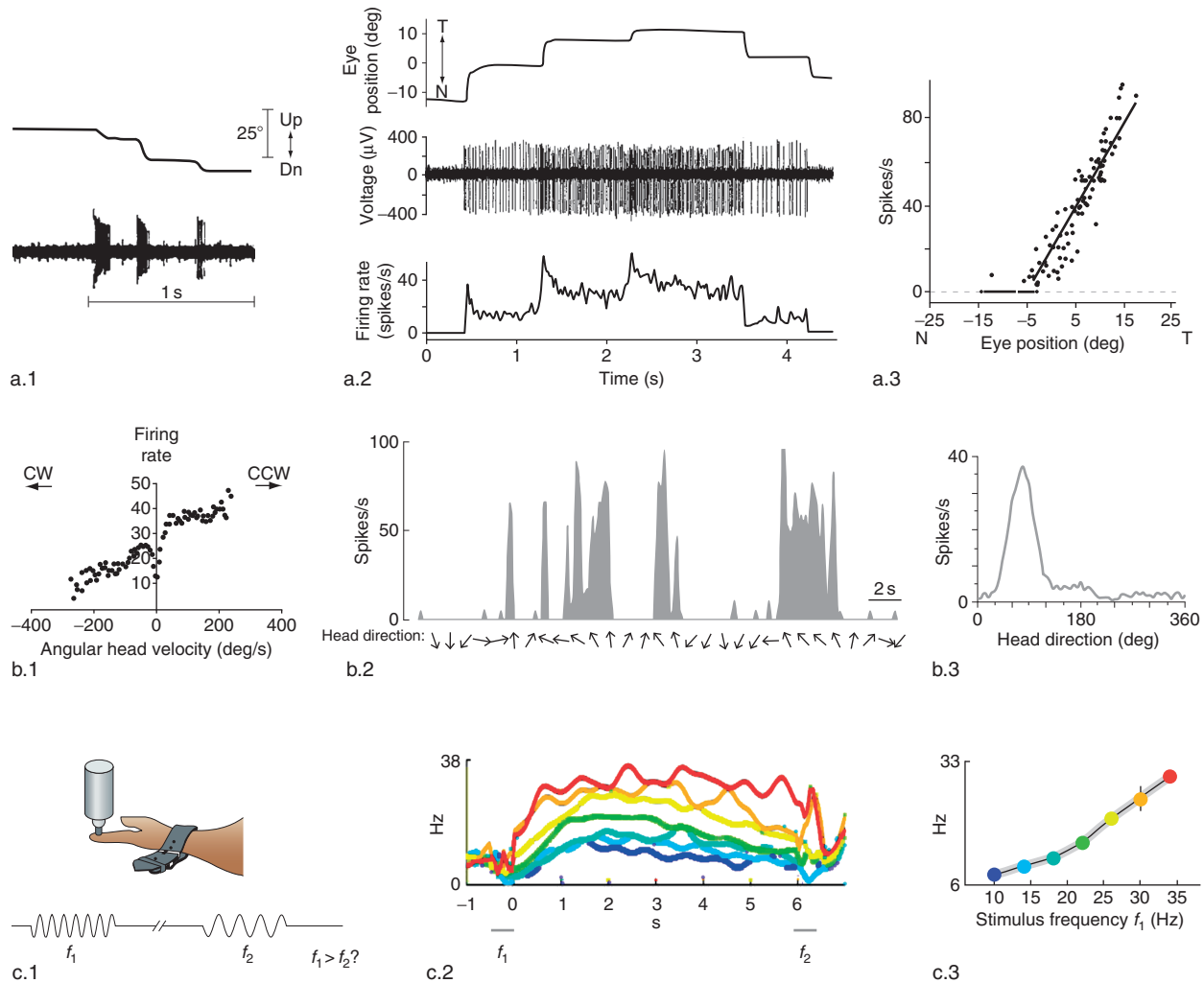


Figure 1 Experimental evidence for neural integrators in various systems reveals two principal coding schemes: 'rate code' and 'location code.' (a) In the oculomotor system, eye velocity-encoding motor commands from neurons in brain stem nuclei occur in the form of bursts at the time of each saccade but not during gaze fixation. However, neurons downstream fire tonically during fixations to maintain the eye muscles' tension and, thereby, the eye position (a.2). The neural integrator of the oculomotor system was hypothesized to explain how transient inputs as in a.1 (bottom) result in sustained responses as in a.2. Neurons in the oculomotor neural integrator show a rate code: The eye position is encoded in their average rate of sustained firing (a.3). (b) Input to head direction (HD) neurons comes from neurons in the brain stem that fire at the time of head rotations, with the firing rate encoding the angular velocity of head rotations. (b.1). HD neurons encode the integral of this signal and fire maximally only when the head points in a specific direction (b.2). The ensuing code, when plotting the average rate of the cell as a function of the HD, is a location code where neurons firing maximally encode the current HD (b.3). (c) In a task in which the vibration frequency f_1 of a tactile stimulus needs to be remembered over a delay period in order to compare it with another frequency f_2 (c.1), neurons in the prefrontal cortex present sustained firing during the delay period, which depends on f_1 (c.2; colors correspond to different values of f_1 as indicated in c.3). Using a rate code, these cells encode in their average delay firing rate the frequency f_1 (c.3). (a.1: top, eye position; bottom, extracellular voltage of motor command neuron) Reproduced from Kaneko CR, Evinger C, and Fuchs AF (1981) Role of cat pontine burst neurons in generation of saccadic eye movements. *Journal of Neurophysiology* 46: 387–408. (a.2, a.3) Reproduced from Aksay E, Baker R, Seung HS, and Tank DW (2000) Anatomy and discharge properties of premotor neurons in the goldfish medulla that have eye-position signals during fixations. *Journal of Neurophysiology* 84: 1035–1049. (b.1) Reproduced from Taube JS and Bassett JP (2003) Persistent neural activity in head direction cells. *Cerebral Cortex* 13: 1162–1172. (b.2, b.3) Adapted from Taube JS (1995) Head direction cells recorded in the anterior thalamic nuclei of freely moving rats. *Journal of Neuroscience* 15: 70–86. (c.1–c.3) Reproduced from Romo R, Brody CD, Hernández A, and Lemus L (1999) Neural correlates of parametric working memory in the prefrontal cortex. *Nature* 399: 470–473.

of a constant signal is a linearly increasing function. A neural integrator will thus be able to accumulate evidence in favor of alternative propositions in a decision process. When alternatives with a barely

discernible difference in their sensory inputs are presented over an extended period of time, an integrator can accumulate the small differences over time until it can eventually discriminate the alternatives through

its ever growing output. This accumulation dynamics is at the heart of the 'drift diffusion' model for reaction time decisions in the psychological literature. Neurophysiological studies using behaving monkeys have found single neuron activity in the posterior parietal cortex that is correlated with accumulation of sensory information during perceptual decisions. Also, drawing on the response of a neural integrator to a constant input, a clock can be built. If the neural integrator is reset each time it reaches a prespecified value, the resets will occur at uniform intervals as the ticks of a clock. Some authors have hypothesized that this scenario represents a candidate neural mechanism by which time intervals can be estimated.

This body of experimental evidence reveals that integration is a fundamental neural computation in the brain, and quantitative neural models have begun to yield clues as to what likely mechanisms underlie this operation in neural circuits and what theoretical concepts need to be explored experimentally. In this article, we present an overview of the computational models of neural integrators. We first define and discuss key concepts that underlie the construction of these models, and then we present some representative models that operate through several specific mechanisms. Finally, we discuss challenges and perspectives for future research on neural integrators.

From Concepts to Cellular Mechanisms

Time integration involves two basic operations. First, during stimulation, a neural integrator needs to additively increase or decrease its output in response to excitatory or inhibitory input signals, respectively. Second, a neural integrator should have a memory capability to store the running total of inputs. Neither integration over a long timescale nor memory of the running total is obvious from a neurophysiological standpoint because leakage of neuronal currents out of the membrane tends to make a neuron forgetful. As a result, the duration over which a neuron can temporally sum incoming synaptic potentials is limited by its membrane time constant. To achieve long integration times, most models therefore incorporate tuned positive feedback that serves to offset the intrinsic leakage of membrane currents. This positive feedback is most often in the form of synaptic feedback in a recurrently connected network, but it has also been implemented as feedback loops in the intracellular signal transduction pathways within a single neuron.

To illustrate how positive feedback can lead to integration, consider a simple model of a neuronal pool with a mean firing rate r and a synaptic connection to itself of weight w (Figure 2a). A pulse of

external input to the neuronal pool leads to an immediate change of firing rate. When the input terminates, the induced change in firing rate is sustained by reverberating input arriving through the recurrent synaptic connection. In this manner, the network both summates and stores external signals.

Quantitatively, integration by this simple network can be described by the firing rate equation:

$$\tau_{\text{neuron}} \frac{dr}{dt} = -r + wr + I(t) \quad [1]$$

where r represents the mean firing rate of the neuronal pool, w is the strength of the synaptic connection, $I(t)$ is the external input to be integrated, and τ_{neuron} is the intrinsic decay time constant of neurons in the pool. The first term on the right-hand side represents the intrinsic neuronal 'leak' processes which, in the absence of feedback ($w = 0$), lead to rate decay with time constant τ_{neuron} . The second term, wr , gives the total positive feedback, which is proportional to the strength of the feedback connection and increases with the neural activity level. The effective time constant of the network is given by

$$\tau_{\text{integrator}} = \frac{\tau_{\text{neuron}}}{|1 - w|} \quad [2]$$

For instance, when a constant external input is presented, activity in this model exponentially approaches a steady-state level (given by $I/(1-w)$) with time constant $\tau_{\text{integrator}}$. This means that the system cannot integrate an input I for a time much longer than $\tau_{\text{integrator}}$, beyond which the neural activity level becomes insensitive to the duration of the stimulus presentation. Moreover, after the input is turned off, the system decays back to zero, and hence the memory about the input is lost, with the same time constant. The system's integration time is fairly short in the absence of positive feedback (when $\tau_{\text{integrator}} = \tau_{\text{neuron}}$), but the integration time can be increased with stronger positive feedback (a larger w).

Perfect integration is achieved when the positive feedback balances the intrinsic neuronal leak processes (Figure 2a, $w = 1$) so that

$$\tau_{\text{neuron}} \frac{dr}{dt} = I(t) \Rightarrow r = \frac{1}{\tau_{\text{neuron}}} \int I(t) dt \quad [3]$$

This balancing of feedback and leak implies that such models require fine-tuning in order to achieve perfect integration. The tuning criterion can be viewed graphically by plotting the magnitude of the leak term (r) and the feedback term (wr) on the same set of axes (Figure 2, top row). For firing rates at which these terms balance, the network performs perfect integration. When feedback is less than leak, the stored total

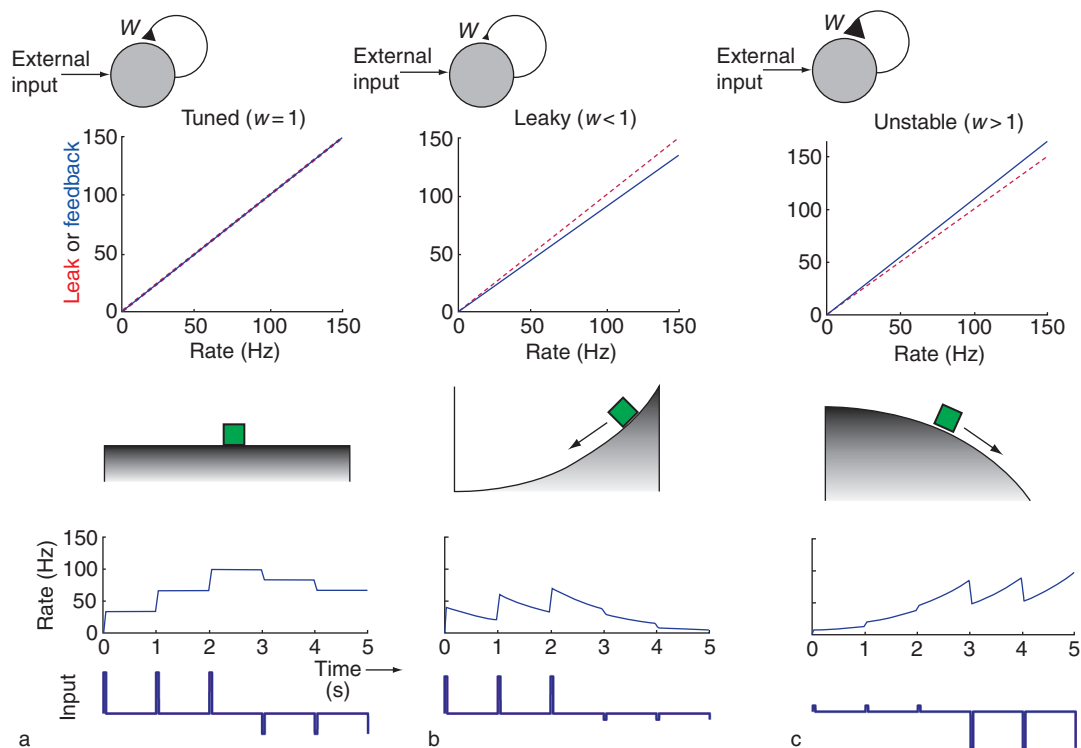


Figure 2 Positive feedback as a mechanism for neural integration. (a) Simple model of a neuronal population with positive feedback onto itself carried by a synaptic connection of weight w . (Top) When properly tuned, the feedback at any firing rate offsets the intrinsic leak processes that pull the firing rate back toward zero. Dashed red line, intrinsic leak; solid blue line, synaptic feedback. (Middle) The integrated total of external inputs can be conceptually represented by the location of a block along a surface. External inputs that are accumulated by the integrator move the block along the surface. When well tuned, the integrator can stably maintain any accumulated value, corresponding to the block not slipping along the flat surface. (Bottom) Accumulation of pulsed excitatory or inhibitory inputs into sustained firing rates. (b, Top) Model in which the synaptic weight is decreased so that the positive feedback is less than the intrinsic leak at all firing rates. Firing rate drift is proportional to the difference between the feedback and leak traces. (Middle) The model firing rate behavior can be represented by the location of a block on a curved surface, where the curvature is proportional to the firing rate drift in the absence of input. (Bottom) Decrease in the synaptic weight leads to exponential decay of firing rates in the absence of input. (c) A model in which feedback is too large leads to exponential growth of firing rates.

in the integrator leaks away at a rate proportional to the difference between the leak and feedback. In the absence of input, this leads to exponential decay on a timescale $\tau_{\text{integrator}}$ (leaky integrator; [Figure 2\(b\)](#)). Similarly, when feedback is greater than leak, activity exponentially increases with a time constant given by the same formula (unstable integrator; [Figure 2\(c\)](#)).

In a useful conceptual description, the network's output can be represented by the location of a block being pushed along a surface, with the movement of this block representing the dynamics of the network. External inputs slide the block with a velocity proportional to the size of the input. For a flat and level surface ([Figure 2\(a\)](#)), corresponding to a well-tuned network, the position of the block thus represents the integral of the external input. In the absence of input, the block stays in place, corresponding to successful maintenance of the accumulated input. For a mistuned network, the slope of the surface is

given by the degree of mistuning – that is, by the difference between the feedback and leak. In this case, even in the absence of input, the position of the block drifts with a velocity proportional to the mistuning. For the leaky network, the mistuning brings the activity back to the origin ([Figure 2\(b\)](#), bottom), whereas for the unstable network, the mistuning leads to unbounded increase in activity ([Figure 2\(c\)](#), bottom).

Although the previously discussed model is highly simplified, it captures the essence of the positive feedback mechanism used by more biologically detailed network models. Because of the strong coupling between neurons that mediates the positive feedback in such networks, only certain activity patterns can be stably maintained internally in the absence of external inputs. These activity patterns are typically assumed to form a continuous, one-dimensional line in the high-dimensional space of all possible firing

patterns; the distance along this line (analogous to the distance of the block along the surface in [Figure 2](#)) represents the accumulated total of inputs being stored by the integrator. Activity patterns not along the line decay rapidly to a nearby point along the line so that the line is said to be an ‘attractor,’ or ‘line attractor,’ of the networks dynamics.

Specific biological networks differ in their implementation of positive feedback, which may be through excitatory pathways or disinhibitory (hence, effectively excitatory) pathways. They also differ in the topographic organization of the pathways conveying this feedback and in the means by which tuning of this feedback is accomplished: in networks utilizing a rate code, the positive feedback provided to each neuron typically changes monotonically as input is accumulated, and the magnitude of feedback connections is tuned to ensure that activity does not leak or exhibit unstable growth. In networks that utilize a location code, the pathways that mediate the positive feedback are spatially structured to allow a localized bump of activity to be maintained, and tuning of the network is typically accomplished by making symmetric connections between neurons so that the bump of activity can be maintained equally well at any location. Finally, the biophysical substrate of the positive feedback can differ between models, from synaptic feedback in network models to feedback between intracellular variables such as intracellular calcium concentration, membrane potential, and conductance levels in single neuron integrator models. Next, we discuss each of these specific mechanisms and their applications to neural integrators.

Diversity of Model Scenarios

Network Integrator Models

Rate code It is often observed in neural integrator networks that two opposing populations of cells are responsible for accumulating the input. Inputs to these populations are typically arranged in a ‘push–pull’ manner such that when one population is excited (pushed) by the external input, the other population is inhibited (pulled). Both the oculomotor integrator, which is located in the brain stem, and neocortical integrator networks display such a push–pull organization. In the oculomotor neural integrator, external commands that encode the velocity of an intended eye movement are sent through excitatory connections to one cell population and through inhibitory connections to the opposing population. The network integrates these inputs into sustained differences in firing rate between the two populations, with the excited population exhibiting

above-background firing rates and the inhibited population exhibiting below-background firing rates.

Such networks contain two potential synaptic pathways to mediate positive feedback. Excitation between neurons within a population could provide a source of excitatory positive feedback. Inhibition between opposing populations could provide a source of mutual-inhibitory positive feedback whereby the excited population inhibits the opposing population and thereby disinhibits itself.

In the simplest models of this circuit, two neurons representing the opposing populations are assumed to be connected by mutual inhibition that provides a disinhibitory positive feedback loop ([Figure 3a](#), middle). Such a circuit requires tonic excitation to drive neuronal activity. Thus, external inputs (of vestibular origin) consist of two components: a tonic component responsible for the background firing rate of neurons and a modulated component that is excitatory to one population and inhibitory to the opposing population. Common, same-direction changes in the tonic background input to each population lead to a corresponding change in the background firing rates of the two populations but are not integrated over time ([Figure 3\(a\)](#), first step in inputs). Because the eye position in this system is determined by the difference in the outputs of the two populations, such inputs do not lead to a change in eye position. By contrast, oppositely directed changes in inputs to the two populations about this tonic level do lead to an integrated response. The population that receives increased external input exhibits an upward ramp in activity, whereas the opposing population exhibits a downward ramp ([Figure 3\(a\)](#), second step in inputs). The difference between these two outputs, which is a ramp of twice the slope of either individual population, represents the integral of the difference between the commands received by either population. An advantage of this network is that it properly integrates such differences while remaining robust against same-direction perturbations in input rates that might be caused by global changes in background rates. In linear network theory, the integration of input differences and nonintegration of common inputs are termed ‘differential mode’ integration and ‘common mode’ rejection, respectively.

Alternatively, integration in the oculomotor neural integrator could be instantiated by excitatory feedback within each population, as illustrated in [Figure 2](#). This idea has been quantitatively examined using realistic conductance-based spiking neurons and synaptic connections. [Figures 3\(b\) and 3\(c\)](#) show a model of a single population of the goldfish oculomotor neural integrator network, based on finely tuned recurrent excitatory positive feedback. Neurons were

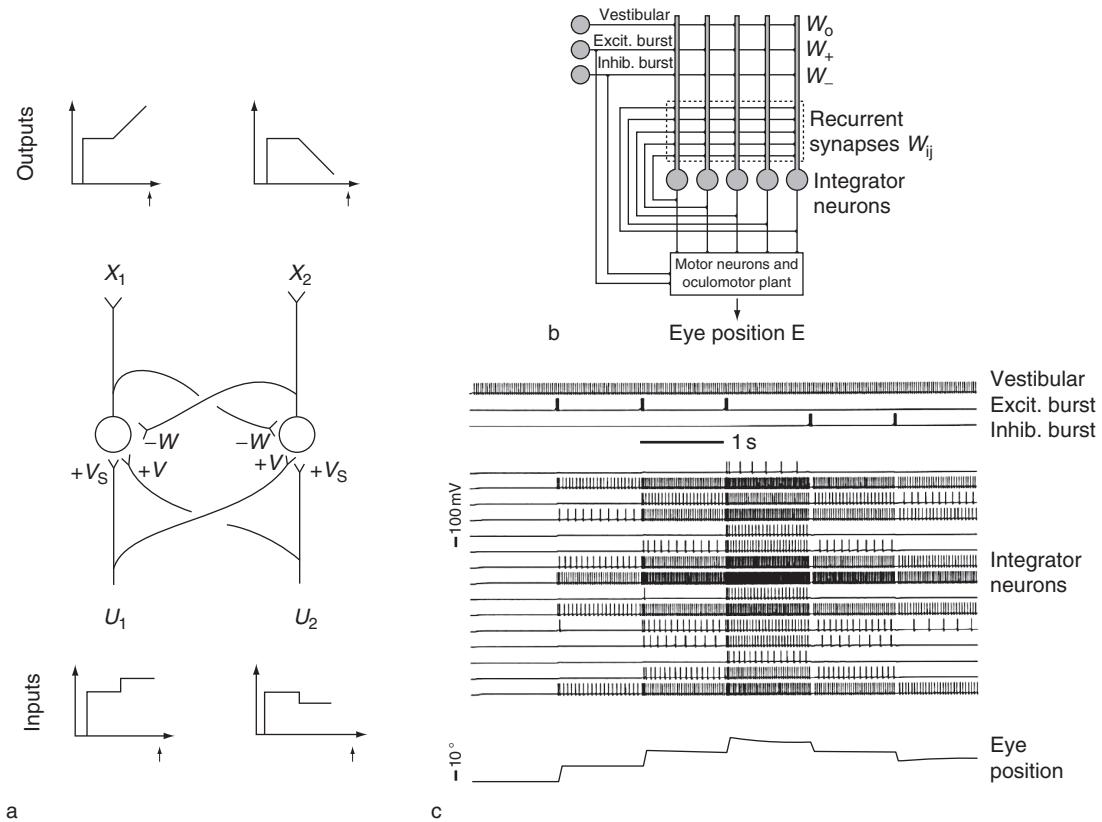


Figure 3 Rate code models of the oculomotor neural integrator. (a) Network model based on recurrent inhibition. Two neurons connected by mutual inhibition receive external inputs from ipsilateral and contralateral sources. Positive feedback is provided by each neuron inhibiting its neighbor and therefore disinhibiting itself. External inputs to the network consist of two components – a ‘common’ background input that is identical for each of the two neurons and a ‘differential’ command input that is excitatory to one neuron and inhibitory to the other. Changes in the common component (bottom, first step in inputs) lead to a change in background output levels but are not integrated (top traces). Differences in inputs (bottom, second step in inputs) are integrated by the network, in this case from a step into a ramp. In the oculomotor system on which this model was based, eye position is controlled by a difference between the outputs of two populations represented by these cells. Thus, the network output is robust to common changes in background inputs but is sensitive to differential inputs. (b, c) Integration by a network based on recurrent excitation. (b) Model network consisting of a population of neurons with recurrent excitatory synaptic connections. External commands from excitatory and inhibitory bursting neurons provide the velocity-encoding input for saccades, and vestibular signals provide tonic background input. Eye position is decoded from the network via a first-order model of the motor neurons and oculomotor plant. (c) Integration of bursts into steps of firing rate and eye position by the tuned network. Steps of firing rate in the absence of external command inputs are maintained by recurrent excitation. Different neurons begin firing at different eye positions due to differences in their background input levels. (a) Adapted from Cannon SC, Robinson DA, and Shamma S (1983) A proposed neural network for the integrator of the oculomotor system. *Biological Cybernetics* 49: 127–136. (b, c) Adapted from Seung HS, Lee DD, Reis BY, and Tank DW (2000) Stability of the memory of eye position in a recurrent network of conductance-based model neurons. *Neuron* 26: 259–271.

connected in an all-to-all manner, with each neuron receiving excitatory and inhibitory command inputs to be integrated as well as tonic vestibular background inputs that differed from neuron to neuron (Figures 3(b) and 3(c)). As in Figure 2, active neurons exchanged positive feedback through recurrent synaptic excitation, allowing them to persistently maintain changes in firing rate in response to transient command inputs. However, unlike the simple model of Figure 2, the size of the active population also changed in response to inputs because neurons were sequentially recruited into the network with

increasing network activity (with recruitment order determined by the amount of background input each neuron received). To offset the extra source of positive feedback that newly recruited neurons provided to their neighbors, the model assumed that synaptic transmission saturates as presynaptic firing rates increase. This idea of synaptic saturation balancing recruitment of neurons represented a significant step in the realistic modeling of rate-code integrators because it demonstrated how linear accumulation of inputs could occur even in systems with realistic nonlinearities. Experimentally, studies of the goldfish

oculomotor integrator have suggested that integration in this system is generated within each population of the integrator and not through recurrent inhibition between populations. This result is consistent with the idea of integration through excitatory positive feedback, but it is unknown whether single neurons in the system are endowed with intrinsic properties that could also contribute to the integration.

Location code Figure 4 shows schematically a typical neural integrator model in which the encoded feature is specified by the peak location of the network activity profile (location code). Neural networks that show this type of encoding has been used to model many neurobiological computations, from encoding of spatial location by place cells in the hippocampus to spatial working memory in dorso-lateral prefrontal cortex. The realization of a neural integrator that utilizes a location code requires two ingredients. First, a continuous family of localized persistent activity patterns (bumps; Figure 4(b)) must be generated, each representing the graded value of the accumulated signal. Such localized activity profiles can be sustained by excitatory reverberation among neighboring neurons as the positive feedback mechanism (Figure 4(a)) or, alternatively, using tonic excitation combined with structured synaptic inhibition that sculpts the network activity profile. Second, a shifting mechanism must enable transient inputs to smoothly move the bump in such a way that the moving speed is linearly proportional to the input amplitude so that the final location of the bump accurately represents the time integral of the input (Figure 4(c)). The head direction (HD) system captures especially well these two aspects of neural integrators using a location code. When an animal

turns its head, the angular velocity signal carried by vestibular inputs (Figure 1(b.1)) is integrated over time by HD cells into a positional signal, and the latter is sustained internally when the animal keeps the HD fixed (Figure 1(b.2)). HD cells are selective for angular HD according to a Gaussian (bell-shaped) tuning curve (Figure 1(b.3)).

Interestingly, evidence indicates that HD cells are generated in a neural circuit characterized by a paucity of local excitatory connections. Consistent with this observation, Figure 5 shows a model of HD neurons based on a mechanism without recurrent excitation in which the most active neurons push down their neighbors' activities through cross-directional inhibition rather than lifting up their own activities through isodirectional excitation. Baseline neuronal activities in this model are provided by constant-rate excitatory background inputs, and the direction-selective 'bump of persistent activity' reflects inhibition-driven decreases in activity below these externally driven levels. Moreover, the model surmises that this bump of activity is shifted around the network by 'rotation cells' consisting of two inhibitory cell populations (Figure 5(a)). When the HD is fixed, the inputs from the two inhibitory neural populations are balanced with each other, and the bump of activity in the excitatory neural population is maintained fixed (Figure 5(b), time epochs between input pulses). As the head turns, the angular velocity signal increases the firing of one inhibitory population while decreasing the firing of the other. The resulting asymmetric inhibitory inputs induce the activity pattern in the excitatory neural population to move at a constant speed proportional to the input amplitude (angular velocity) (Figure 5(b), time epochs during input pulses).

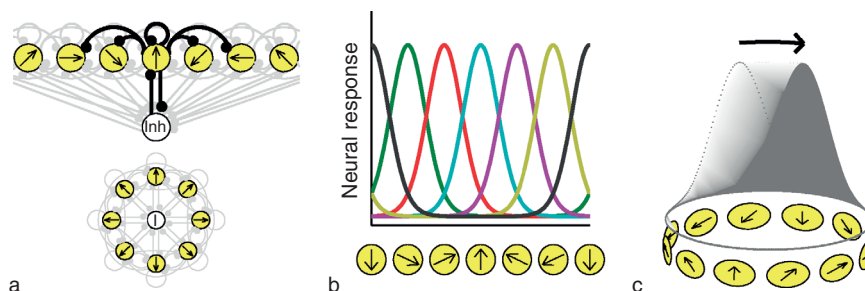


Figure 4 A bump attractor ring model as the substrate for a location code integrator. (a) Example of connectivity between neurons in a ring model attractor network. Recurrent excitation among near neighbors underlies positive feedback in the network, whereas growth of firing rates is controlled by nonselective feedback inhibition. When neurons at either end of the network (top) are made to coincide, a ring model network is obtained (bottom). (b) Given appropriate connectivity strengths, the network shows a whole family of stable neural activity states in the absence of external input ('bump' states, in different colors), each of which peaks at a different location on the network. The x-axis represents different neurons in the population, labeled by their preferred feature. For each curve, the y-axis values represent a possible distribution of firing rates across the population. (c) Bump states remain stably at their locations but can be shifted along the network smoothly by a 'push.' The final location of the shifting bump will be a read-out of the integral of this push force over the time during which it acts on the bump.

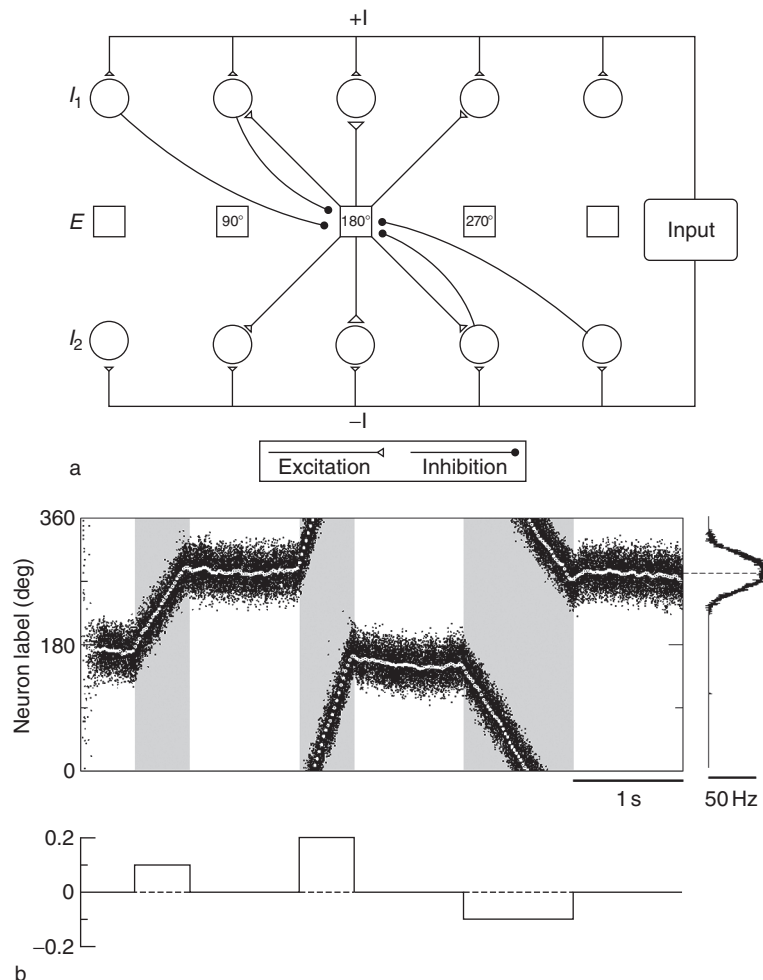


Figure 5 Time integration by a bump attractor network. (a) Model scheme. An excitatory neural network, encoding a directional angle (0–360°), receives inputs from two inhibitory neural populations that are balanced with each other when there is no input. A velocity signal increases input to one of the inhibitory populations (+I) and decreases input to the other (–I), leading to a bias for the excitatory network. (b) Network firing pattern (top) in response to a series of input signal steps (bottom). Neurons in the network are aligned along the y-axis, labeled by their preferred directional angles. The x-axis represents time. Each dot is a spike. The network activity pattern has the form of a bell-shaped profile (top right), with peak location encoding the directional angle (white line). In the absence of an external input, the directional information is maintained by a persistent firing pattern. An input induces the bump of activity to move at a speed proportional to the stimulus amplitude; hence, the network performs a time integral of the input. Simulation by P Song using the model published in Song P and Wang X-J (2005) Angular path integration by moving ‘hill of activity’: A spiking neuron model without recurrent excitation of the head-direction system. *Journal of Neuroscience* 25: 1002–1014.

Computation performed by the model is quite similar to an integral operation in calculus. When the input intensity is doubled (second vs. first input pulse), the hill of activity moves twice as fast. Moreover, the network can integrate both positive (first and second pulses) and negative (third pulse) inputs. In fact, with the third input of an amplitude half that of the second input, but lasting twice as long, the bump of activity moves back to the position before the second pulse, as expected by perfect integration.

Single Neuron Integrator Models

Neural integrators can also reside in the internal biochemical machinery of a single neuron. Experimental

evidence that single cells can exhibit a range of graded persistent activity has been based on recordings in an *in vitro* preparation of rat entorhinal cortical slices (Figure 6(a)). A number of computational models have been built to explore the mechanistic basis of this activity (Figures 6(b)–6(d)). All of these alternative models implicate two principal actors: the muscarinically activated calcium-dependent nonselective cation current (I_{CaN}) and the concentration of calcium in the cytoplasm. The interaction between these two substrates can subserve integration through a positive feedback loop. During neural activity intracellular calcium increases because of the influx of calcium ions through high-threshold calcium channels, increased

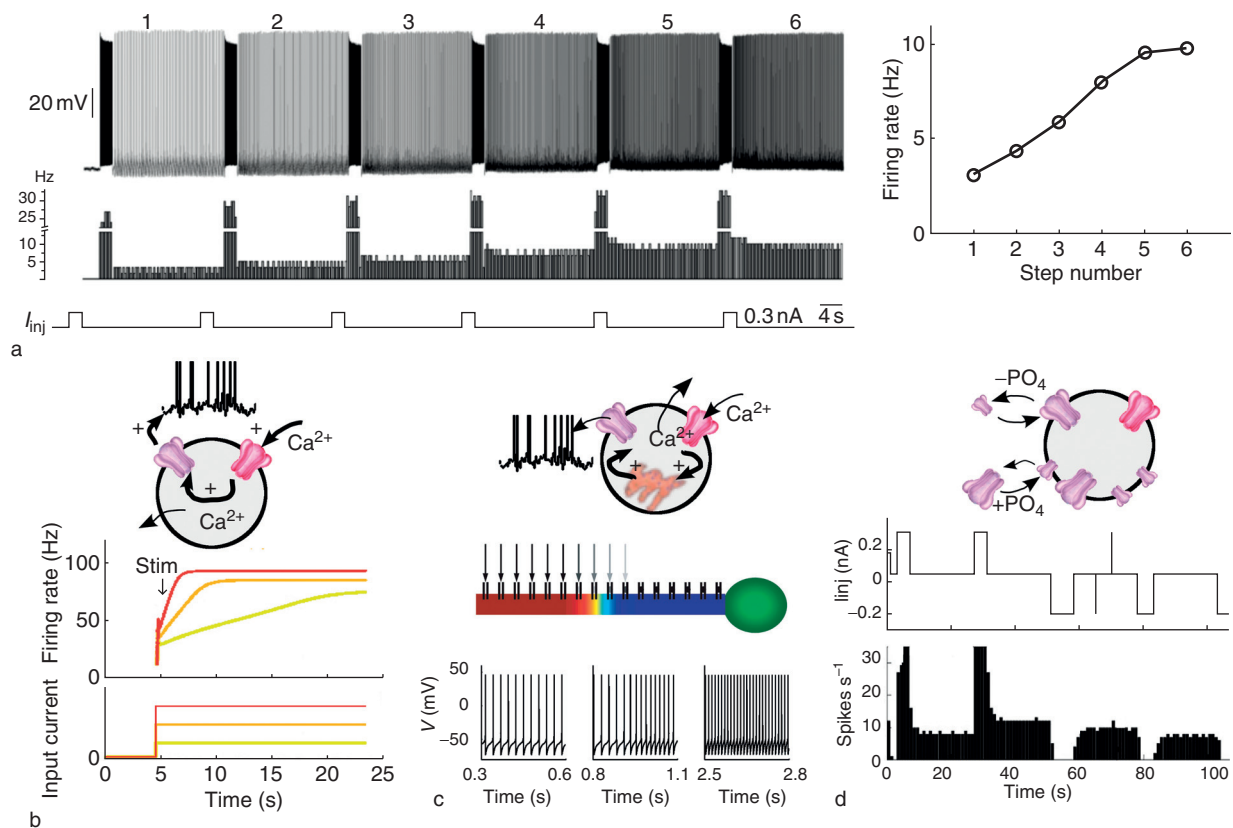


Figure 6 Mechanisms of integration in single neurons. (a) When activated cholinergically, synaptically isolated neurons of the entorhinal cortex integrate successive, brief pulses of current injection (I_{inj} ; bottom left) into stable graded firing rate levels (top left). This also corresponds to a rate code, where the average rate of the neuron is linearly related to the number of stimuli applied (right). (b) Positive feedback (thick arrows indicated with a + sign) between intracellular calcium entering through high-threshold calcium channels (pink) and calcium-dependent nonspecific cation channels (purple). (Lower panels) Linearly increasing neural response (up to integrator saturation limit) for three distinct step pulse currents. (c) Positive feedback (thick arrows indicated with a + sign) between intracellular calcium and calcium release from internal stores (orange). (Middle) With uniform distribution of these feedback units along a dendrite, stable localization of the calcium front (color code in dendrite) determines the total amount of current I_{CaN} (channels open on the dendrite) summated at the soma. (Bottom) Stable neural firing rate after three consecutive pulses of current. (d) During external stimulation, phosphorylation/dephosphorylation of I_{CaN} channels modifies the balance of intrinsic currents in the neuron and its spontaneous firing rate. No positive feedback is involved in this mechanism. (Lower panels) A computational model based on this mechanism reproduces faithfully the experimental results in (a) (top graph, injected current; bottom graph, neural output). (a) Adapted from Egorov AV, Hamam BN, Fransen E, Hasselmo ME, and Alonso AA (2002) Graded persistent activity in entorhinal cortex neurons. *Nature* 420: 173–178. (b) Adapted from Durstewitz D (2003) Self-organizing neural integrator predicts interval times through climbing activity. *Journal of Neuroscience* 23: 5342–5353. (c) From Loewenstein Y and Sompolinsky H (2003) Temporal integration by calcium dynamics in a model neuron. *Nature Neuroscience* 6: 961–967. (d) Adapted from Fransén E, Tahvildari B, Egorov AV, Hasselmo ME, and Alonso AA (2006) Mechanism of graded persistent cellular activity of entorhinal cortex layer V neurons. *Neuron* 49: 735–746.

intracellular calcium activates the I_{CaN} current, which in turn admits cations into the neuron, depolarizing it and generating further neural activity (Figure 6(b)). Hence, this is a positive feedback mechanism similar to that of Figure 2, but it is realized at the single cell level rather than by a synaptically interconnected circuit. It has been proposed that a learning algorithm can be used to dynamically adjust the parameters of such a mechanism to generate linearly increasing responses to step current pulses of varying intensity (Figure 6(b)). Alternatively, the fundamental positive feedback

loop could be established between the intracellular calcium and intracellular stores of calcium (e.g., the endoplasmic reticulum). The positive feedback occurs because intracellular stores of calcium have been found to release calcium to the cytoplasm upon increases in intracellular calcium (calcium-induced calcium release). In this scenario, I_{CaN} is not part of the positive feedback loop but, rather, serves to translate the calcium level into neuronal activity by providing calcium-dependent depolarizing currents (Figure 6(c)). Models of neural integrators built around this

scenario require a spatial distribution of the positive feedback units in order to generate graded levels of activity. One possibility would be to distribute them continuously along the neural dendrites so that a front between high- and low-calcium regions stabilizes at arbitrary locations on the dendrite (which again requires fine-tuning). Strong bursts of action potentials would progressively move this front toward the soma, enhancing the total amount of I_{CaN} and the stable firing rate of the cell (Figure 6(c)).

The two mechanisms described previously (Figures 6(b) and 6(c)) rely on a biochemical feedback loop that is highly sensitive to small modifications of the neuronal properties (fine-tuning). However, experiments on isolated, cholinergically activated entorhinal cortex neurons (Figure 6(a)) have challenged this modeling framework. In these neurons, both the careful matching of membrane currents and the release of calcium from internal stores do not seem to be necessary for neural integration. This has prompted another mechanistic explanation that does not rely on a positive feedback loop. Instead, a computational model was built based on the hypothesis that phosphorylation or dephosphorylation of I_{CaN} channels is triggered by excessively high or low concentrations of intracellular calcium, respectively. During normal, autonomous neural function it was assumed that intracellular calcium levels do not reach the extreme values required to trigger phosphorylation or dephosphorylation in the model. However, during the long pulses of external current injection tested experimentally, the phosphorylation states of the I_{CaN} channels could be changed. Because the two forms of the I_{CaN} channel have significantly different maximal conductances, these changes permanently modify the balance of currents present in the neuron and thus the basal firing rate (Figure 6d). Furthermore, because the associated basal calcium concentrations are far from the thresholds for phosphorylation or dephosphorylation of the I_{CaN} channel, the normal positive feedback loop between voltage activity, calcium concentration, and I_{CaN} conductance state is broken and the firing rates remain perfectly robust against small fluctuations. However, it is unclear whether it is biologically plausible to assume that at high calcium levels phosphorylation is not counteracted by dephosphorylation so that effectively there is no 'biochemical leak.' Moreover, the robustness does occur at a cost relative to positive feedback models because it also implies that the network is incapable of integrating small inputs. This trade-off between robustness against small fluctuations and insensitivity to small inputs is a feature of several models of neural integrators that attempt to solve the fine-tuning problem.

Hybrid Models Incorporating Synaptic Feedback and Intrinsic Bistability

Integrator models based solely on network mechanisms require fine-tuning of synaptic connection weights in order to perform accurate integration over long timescales. Some models have demonstrated that this need for fine-tuning can be reduced if neurons contain one or more digital elements that only turn on when their inputs are sufficiently large and only turn off again when their inputs are sufficiently small. In the simplest realization of this scheme, an integrator network could be made from N 'bistable' neurons that can exist in either of two stable states for a given level of input – an 'on' (active) state and an 'off' (inactive) state. If external inputs sequentially turn neurons on or off at a rate proportional to the size of the input, then the number of active units at any given time reflect an integral of the external input.

Unfortunately for this simple scheme, experimentally observed integrator neurons have not been found to exhibit such discrete on or off neuronal firing. However, this observation does not preclude the possibility that some fraction of a neuron's inputs are transmitted through bistable elements such as dendritic plateau potentials. For example, Figure 7 shows a network in which a conventional recurrent excitatory positive feedback model was modified to include multicompartmental neurons that each contain multiple bistable dendrites (Figure 7(a)). Dendritic compartments were assumed to generate plateau potentials that turn on when synaptic input exceeds a value s_{on} but do not turn off until the input falls below a lower value s_{off} (Figure 7(b)). The firing rate of each neuron reflects the number of its dendrites that are active, in addition to background and external command inputs. Because the dendrites operate in a digital manner, stored activity in the integrator network can be maintained stably even if synaptic inputs are altered by moderate amounts. As shown in Figure 7(d), the network integrates pulses of inputs into steps of firing rate even when synaptic weights (and hence recurrent synaptic inputs) are altered by an amount that would cause gross decay in firing rates in fine-tuned networks (compare to Figure 2(b)). Conceptually, the effect of the bistability in this model is to add ridges to the bottom of the energy surfaces described in Figure 2 (Figure 7(c)). External inputs to the network, which can be represented by (large) tiltings of the surface, slide the block to a location representing the new accumulated value. Imprecision in tuning of the network corresponds to (smaller) warping of the surface. Unlike the fine-tuned network, this warping does not cause the network's

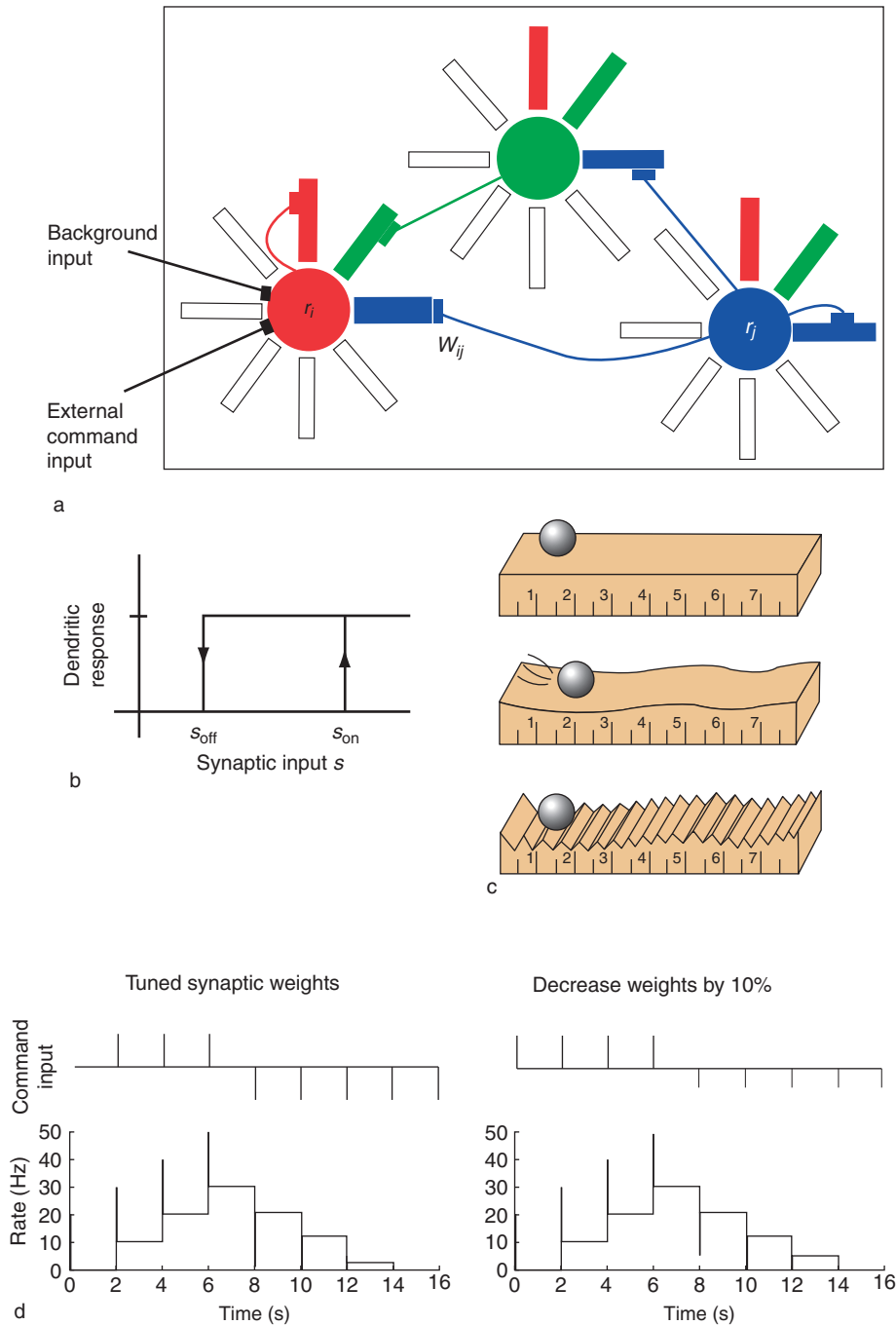


Figure 7 A hybrid intrinsic network mechanism of integration. (a) Network of multicompartment model neurons. Each neuron projects recurrent excitation of weight W to a single bistable dendrite of each of its neighboring neurons. Neurons also receive external background and command inputs at the soma. (b) Dendrites have a bistable input–output relationship characteristic of a plateau potential. Dendrites are activated when synaptic input exceeds a value s_{on} and deactivated when input falls back below a value s_{off} . (c) The bistable nature of the dendrites can be conceptualized as creating ridges in the surface along which the ball representing the state of the system moves (see [Figure 2](#)). These ridges make the network output resistant to drift that otherwise would be caused by small perturbations in synaptic connection strength that correspond to warping and tilting of the surface. (d) Robust maintenance of activity in the network. Pulses of input are accumulated into different levels of stable output (rate code). Firing rates can be maintained in the absence of command inputs even when weights are changed by an amount that would cause severe drift in a network without bistable elements (right; compare to [Figure 2b](#)). (a, b, d) Adapted from Goldman MS, Levine JH, Major G, Tank DW, and Seung HS (2003) Robust persistent neural activity in a model integrator with multiple hysteretic dendrites per neuron. *Cerebral Cortex* 13: 1185–1195. (c) Adapted from Pouget A and Latham P (2002) Digitized neural networks: Long-term stability from forgetful neurons. *Nature Neuroscience* 5: 709–710 (comment on Koulov AA, Raghavachari S, Kepecs A, and Lisman JE (2002) Model for a robust neural integrator. *Nature Neuroscience* 5: 775–782).

representation of its integrated input to slip as long as the warping is insufficiently large to knock the block out of the divots. However, this robustness against imprecise tuning also implies that the network is unresponsive to small inputs that do not tilt the surface enough to cause the block to slide. A challenge for future models incorporating bistability is to overcome this inherent trade-off between robustness to imperfect tuning and inability to integrate small external inputs.

Accumulation of Information versus Graded Memory

So far, our discussion has focused on perfect time integration in the sense of calculus, which is exemplified by oculomotor neural integrators and HD cells. However, it is important to note that different components of the nervous system are likely to exhibit different types of temporal integration, some less perfect than others, to satisfy a variety of computational demands. In particular, it is useful to distinguish two types of time integration: approximately linear summation of inputs during stimulation (e.g., slow ramping activity in response to a prolonged constant input) and parametric working memory (graded conversion of pulse-like input to plateau-like persistent activity). [Figure 8\(a\)](#) shows a line attractor network model of parametric working memory, which was developed to simulate the monkey delayed somatosensory comparison experiment ([Figure 1\(c\)](#)). Like the prefrontal neuron shown in [Figure 1\(c\)](#), this model produces graded persistent activity, which is approximately proportional to the strength of a transient input pulse and is self-sustained after the stimulus offset. The mechanism underlying graded persistent activity is precisely the tuned positive feedback described in [Figure 2](#) but implemented with a realistic circuit model of spiking neurons. Positive feedback in the model was mediated by NMDA receptors, which have a relatively slow decay time constant of 50–100 ms, reducing the requirement for fine-tuning of synaptic weights. Such a line attractor network can sum inputs in the sense of calculus. However, working together with other neural circuits, or using additional gating input signals, such a network has been shown to be capable of encoding the intensity rather than the duration of a stimulus; that is, the graded persistent activity is proportional to the input amplitude, but if the stimulus is prolonged for increasingly longer durations, the level of neural firing would not grow linearly with time, contrary to a neural integrator. In the somatosensory delayed discrimination task the subject has to encode and remember across a delay the frequency of a somatosensory vibration, not the

number of vibrations applied on a fingertip. Thus, the behavioral task requires parametric working memory but not strict time integration of input during stimulation.

Conversely, a neural circuit capable of nearly perfect integration of inputs during stimulation does not necessarily exhibit graded self-sustained persistent activity when the stimulus is withdrawn. This is illustrated in a neural circuit model of decision making ([Figure 8\(b\)](#)). In simulations of a two-alternative forced choice task, firing activity of neurons increases quasi-linearly over time (up to a point) in response to a constant and long-lasting stimulus. Such ramping activity gradually accumulates information about the evidence in favor of one of the two alternative choices, and the rate of ramping activity increases with the input amplitude (corresponding to the quality of sensory evidence) ([Figure 8\(b\)](#)). Later in time, however, competition between neural populations representing the two alternatives ultimately leads to a categorical choice that is manifested by the divergence in the neural activity of the two neural populations. One of them increases while the other decays. Thus, if a decision-based motor response must be delayed, persistent activity during working memory maintenance would be binary. The mnemonic neural firing pattern would not depend on the strength of input in a graded manner but reflect only the categorical choice made during stimulation. This was exactly what was found in lateral intraparietal area neurons in the two-alternative forced-choice version of a visual motion discrimination task and in a discrete attractor network model of decision making and working memory. Graded time integration during stimulation and binary persistent activity during a memory period are not incompatible with each other because the former is a property of neural responses to external inputs, whereas the latter reflects stable states in the absence of inputs. Thus, time integration in decision making does not necessarily require line attractors. Of course, some decision tasks require graded integration of information across temporal gaps, which engage both time integration and parametric working memory. Novel task designs are required to identify the neural activity underlying such decision processes.

Challenges and Future Directions

How neural integrators can be realized robustly by plausible biophysical mechanisms remains a topic of active current research. Current models are based on three ideas. First, positive feedback can create long time constants from much faster decaying biophysical processes. This scenario requires a need for

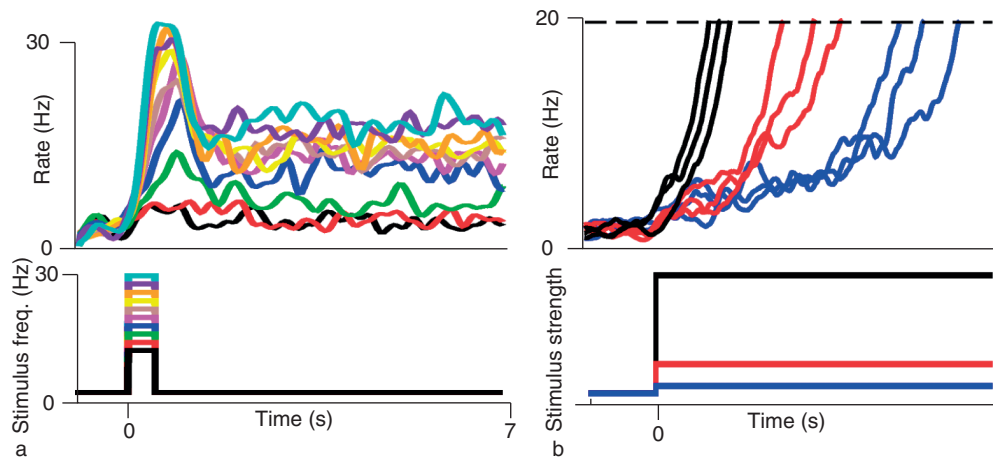


Figure 8 Graded memory and accumulation of information. (a) Graded memory in a network that converts a pulse-like input into persistent activity. The transient input is presented with a range of amplitudes, and the induced persistent activity after the stimulus offset increases approximately linearly with the input strength. The simulations are from a line attractor model of spiking neurons for the somatosensory delayed discrimination experiment (Figure 1(c)). (b) Linear ramping activity during stimulus presentation. Neural activity increases quasi-linearly in time, and it varies stochastically from trial to trial (three trials are shown for each input amplitude). The ramping rate increases with the input amplitude. Such 'ramping-to-threshold' probabilistic neural activity can serve to accumulate information for decision making, or can be used to estimate a time interval. The simulation results are from a discrete attractor network model of perceptual decision making and working memory. (a) Adapted from Miller P, Brody C, Romo R, and Wang X-J (2003) A recurrent network model of somatosensory parametric working memory in the prefrontal cortex. *Cerebral Cortex* 13: 1208–1218. (b) Adapted from Lo CC and Wang X-J (2006) Cortico-basal ganglia circuit mechanism for a decision threshold in reaction time tasks. *Nature Neuroscience* 9: 956–963.

fine-tuning that seems unreasonably strict. For neural integrators using a rate code, fine-tuning amounts to adjustment of synaptic weights in the networks, whereas for neural integrators using a location code, fine-tuning implies symmetry of connections and homogeneity of neuronal properties in networks. Moreover, such systems typically display random drifts of neural activity in the absence of external inputs, which would be functionally detrimental. Second, very slow biophysical and biochemical processes naturally provide long time constants. However, systems that are built on processes with very slow kinetics, such as the single neuron integrator cells of entorhinal cortex, suffer from a sluggish response to inputs that does not match the rapid accumulation of inputs exhibited in other systems. Third, nonlinearities such as plateau potentials or thresholds have been used to provide robust behavior in some models but also place constraints on the types of functions that can be integrated. Further progress depends on new experiments that place more constraints on, and establish a firmer mechanistic basis for, computational modeling in this area. What types of inputs a given system does or does not integrate may provide a guide to future models because different mechanisms may be tailored to the system's specific needs. Nevertheless, certain overarching themes, such as positive feedback and analog memory storage of accumulated input levels, are likely to provide a common framework in which to view many systems.

Although most of the current research on modeling of integrators has focused on neural dynamics and robustness to perturbations on fast timescales, an open question concerns what learning rules can assemble and maintain the tuning of such circuits. Experimental evidence for plasticity rules in neural integrators is lacking, and modeling offers the opportunity to suggest appropriate experiments to address this question. Work in the oculomotor system has shown that with altered visual feedback, the oculomotor integrator can be tuned to become leaky or unstable. This suggests that reinforcement learning signals, such as retinal slip for the oculomotor system, can be utilized to tune integrator dynamics. In addition, theoretical work on this system has suggested how spike timing-dependent plasticity could be used in an unsupervised manner to maintain stable firing in the absence of external inputs. In networks utilizing a location code, where proper performance depends on homogeneity across the network so that a bump of activity can be stably maintained at any location, theoretical work has shown how a homeostatic synaptic plasticity rule (synaptic scaling) can effectively create this homogeneity in a network of neurons with diverse intrinsic parameters. Moreover, a modeling study has proposed that line attractors using a rate code can be realized using a symmetry principle for the construction of the network connectivity. If so, homeostatic regulation mechanisms may also offer a possible solution to the fine-tuning

problem for neural integrators using a rate code. Other theoretical work has noted that the higher order statistics of neuronal voltage fluctuations can exhibit special properties when networks are tuned to a line attractor, and has suggested how a learning mechanism that monitors such fluctuations could be used to tune the integrator dynamics. Close collaboration between theory and experiments will be needed to test and further develop such hypotheses at both the single neuron and the systems levels.

See also: Attractor Network Models; Multisensory Convergence and Integration; Neural Coding of Spatial Representations; NMDA Receptors and Development; NMDA Receptor Function and Physiological Modulation; Oculomotor Control: Anatomical Pathways; Oculomotor System: Models; Olfactory Coding; Sensorimotor Integration: Models; Short Term and Working Memory; Working Memory: Capacity Limitations.

Further Reading

- Aksay E, Baker R, Seung HS, and Tank DW (2000) Anatomy and discharge properties of pre-motor neurons in the goldfish medulla that have eye-position signals during fixations. *Journal of Neurophysiology* 84: 1035–1049.
- Brody CD, Romo R, and Kepecs A (2003) Basic mechanisms for graded persistent activity: Discrete attractors, continuous attractors, and dynamic representations. *Current Opinion in Neurobiology* 13: 204–211.
- Cannon SC, Robinson DA, and Shamma S (1983) A proposed neural network for the integrator of the oculomotor system. *Biological Cybernetics* 49: 127–136.
- Durstewitz D (2003) Self-organizing neural integrator predicts interval times through climbing activity. *Journal of Neuroscience* 23: 5342–5353.
- Egorov AV, Hamam BN, Fransen E, Hasselmo ME, and Alonso AA (2002) Graded persistent activity in entorhinal cortex neurons. *Nature* 420: 173–178.
- Fransen E, Tahvildari B, Egorov AV, Hasselmo ME, and Alonso AA (2006) Mechanism of graded persistent cellular activity of entorhinal cortex layer V neurons. *Neuron* 49: 735–746.
- Goldman MS, Levine JH, Major G, Tank DW, and Seung HS (2003) Robust persistent neural activity in a model integrator with multiple hysteretic dendrites per neuron. *Cerebral Cortex* 13: 1185–1195.
- Kaneko CR, Evinger C, and Fuchs AF (1981) Role of cat pontine burst neurons in generation of saccadic eye movements. *Journal of Neurophysiology* 46: 387–408.
- Koulakov AA, Raghavachari S, Kepecs A, and Lisman JE (2002) Model for a robust neural integrator. *Nature Neuroscience* 5: 775–782.
- Lo CC and Wang X-J (2006) Cortico-basal ganglia circuit mechanism for a decision threshold in reaction time tasks. *Nature Neuroscience* 9: 956–963.
- Loewenstein Y and Sompolinsky H (2003) Temporal integration by calcium dynamics in a model neuron. *Nature Neuroscience* 6: 961–967.
- Machens CK and Brody C (2008) Design of continuous attractor networks with monotonic tuning using a symmetry principle. *Neural Computation* 20: 452–485.
- Machens CK, Romo R, and Brody C (2005) Flexible control of mutual inhibition: A neural model of two-interval discrimination. *Science* 307: 1121–1124.
- Major G and Tank D (2004) Persistent neural activity: Prevalence and mechanisms. *Current Opinion in Neurobiology* 14: 675–684.
- Miller P, Brody C, Romo R, and Wang X-J (2003) A recurrent network model of somatosensory parametric working memory in the prefrontal cortex. *Cerebral Cortex* 13: 1208–1218.
- Miller P and Wang X-J (2006) Inhibitory control by an integral feedback signal in prefrontal cortex: A model of discrimination between sequential stimuli. *Proceedings of the National Academy of Sciences of the United States of America* 103: 201–206.
- Pouget A and Latham P (2002) Digitized neural networks: Long-term stability from forgetful neurons. *Nature Neuroscience* 5: 709–710.
- Renart A, Song P, and Wang X-J (2003) Robust spatial working memory through homeostatic synaptic scaling in heterogeneous cortical networks. *Neuron* 38: 473–485.
- Robinson DA (1989) Integrating with neurons. *Annual Review of Neuroscience* 12: 33–45.
- Roitman JD and Shadlen MN (2002) Response of neurons in the lateral intraparietal area during a combined visual discrimination reaction time task. *Journal of Neuroscience* 22: 9475–9489.
- Romo R, Brody CD, Hernández A, and Lemus L (1999) Neural correlates of parametric working memory in the prefrontal cortex. *Nature* 399: 470–473.
- Seung HS (1996) How the brain keeps the eyes still. *Proceedings of the National Academy of Sciences of the United States of America* 93: 13339–13344.
- Seung HS (2003) Amplification, attenuation, and integration. In: Arbib MA (ed.) *The Handbook of Brain Theory and Neural Networks*, 2nd edn., pp. 94–97. Cambridge, MA: MIT Press.
- Seung HS, Lee DD, Reis BY, and Tank DW (2000) Stability of the memory of eye position in a recurrent network of conductance-based model neurons. *Neuron* 26: 259–271.
- Song P and Wang X-J (2005) Angular path integration by moving ‘hill of activity’: A spiking neuron model without recurrent excitation of the head-direction system. *Journal of Neuroscience* 25: 1002–1014.
- Taube JS (1995) Head direction cells recorded in the anterior thalamic nuclei of freely moving rats. *Journal of Neuroscience* 15: 70–86.
- Taube JS and Bassett JP (2003) Persistent neural activity in head direction cells. *Cerebral Cortex* 13: 1162–1172.
- Usher M and McClelland J (2001) On the time course of perceptual choice: The leaky competing accumulator model. *Psychological Review* 108: 550–592.
- Wong K-F, Huk AC, Shadlen MN, and Wang X-J (2007) Neural circuit dynamics underlying accumulation of time-varying evidence during perceptual decision-making. *Neuroscience Frontier of Cognitive* 1: 6, doi:10.3389/neuro.10/006.2007.
- Wang X-J (2002) Probabilistic decision making by slow reverberation in cortical circuits. *Neuron* 36: 955–968.
- Zhang K (1996) Representation of spatial orientation by the intrinsic dynamics of the head-direction cell ensemble: A theory. *Journal of Neuroscience* 16: 2112–2126.