Model for a robust neural integrator

Alexei A. Koulakov[1,2], Sridhar Raghavachari[3], Adam Kepecs[3] and John E. Lisman[3]

[1] Salk Institute for Biological Studies, 10010 North Torrey Pines Road, La Jolla, California 92037, USA
[2] Department of Physics, University of Utah, 115 South 1400 East, Salt Lake City, Utah 84112, USA
[3] Volen Center for Complex Systems, Mailstop 013, Brandeis University, Waltham, Massachusetts 02454, USA
Correspondence should be addressed to A.A.K. (akula@physics.utah.edu)

Published online: 22 July 2002, doi:10.1038/nn893

Integrator circuits in the brain show persistent firing that reflects the sum of previous excitatory and inhibitory inputs from external sources. Integrator circuits have been implicated in parametric working memory, decision making and motor control. Previous work has shown that stable integrator function can be achieved by an excitatory recurrent neural circuit, provided synaptic strengths are tuned with extreme precision (better than 1% accuracy). Here we show that integrator circuits can function without fine tuning if the neuronal units have bistable properties. Two specific mechanisms of bistability are analyzed, one based on local recurrent excitation, and the other on the voltage-dependence of the NMDA (*N*-methyl-D-aspartate) channel. Neither circuit requires fine tuning to perform robust integration, and the latter actually exploits the variability of neuronal conductances.

Neurons in the temporal[1], parietal[2,3] and frontal[4] cortices show persistent spiking activity after the brief presentation of a stimulus. This firing is thought to play an important role in working memory[5]. In certain working memory circuits, neurons can maintain a range of firing rates that represent a graded stimulus parameter such as intensity or frequency[6]. This is a form of parametric working memory. Theoretical explanations for such processes have focused on understanding the mechanisms underlying graded persistent activity[7].

Insights into this problem have come from the study of integrator circuits, which are an example of parametric working memory. These circuits transform brief inputs into sustained firing that represents the integral of previous inputs. The best studied integrators are brainstem nuclei that are part of the oculomotor system. Their function is to integrate velocity signals from the inner ear and produce maintained, graded, compensatory eye movements[8–13] (**Fig. 1**).

One proposed mechanism for integrator function is based on tuned recurrent networks[11,13–15]. According to these models, firing rates are kept constant by reverberating neuronal activity through a system of recurrent excitatory synapses. As the neurons in such networks are mutually excitatory, there is a potential for instability due to runaway positive feedback. Theoretical work, however, has shown that under the right conditions, such networks can sustain a graded level of activity with reasonable stability. A major success of this class of models is that they can account for the observed 'recruitment' properties of the oculomotor integrator: different cells start firing at different angles of gaze and then increase their firing rate as gaze angle increases[13].

Despite the success of these models, it remains unclear whether the proposed mechanism is sufficiently robust to be physiologically feasible. The main difficulty is the requirement for fine tuning the strengths of recurrent synapses to keep feedback in exactly the right range. Using linear systems theory, it is possible to calculate the time constant of drift (τ) as a function of the mistuning (ε) of the synapses and the time constant of feedback (τ_f) via the recurrent synapses[11]:

$$\tau \approx \tau_f/\varepsilon \qquad (1)$$

It has been proposed[13] that the feedback time constant, τ_f, could be lengthened if the time constant of the synaptic conductance were long. However, even with the longest known synaptic conductances (~100 ms for NMDA conductance), the observed level of gaze drift implies that synaptic strengths have to be tuned to an accuracy better than 1%. It is unclear whether physiological mechanisms exist that could produce such precise tuning.

The question addressed in this study is whether fine tuning is necessary for integrator function. We describe a new class of integrator networks that do not require fine tuning. In previous models, each neuron of the network is assumed to be a simple unit whose firing frequency is uniquely determined by the input current. However, voltage-gated intrinsic conductances, synaptic conductances, or particular network configurations can cause neurons to show more complex firing behaviors. In particular, neurons can exhibit 'bistable' properties such that they are either silent or firing at the same input current, depending on the previous history of inputs into the neuron. Here we show that a recurrent network composed of bistable neuronal units can function as an integrator that does not require fine tuning. Thus, we describe a new design principle that may produce robustness in parametric working memory circuits as well as in biochemical and genetic networks that have similar multi-stable properties.

Results

Determinants of stability in finely tuned models

A framework for understanding the stability of recurrent integrator networks has previously been developed[11,14], which can be illustrated by an ensemble of identical neuronal units with recurrent connections (**Fig. 2a**). In an isolated unit—without recurrent connections—the neuronal firing rate (f) as a func-

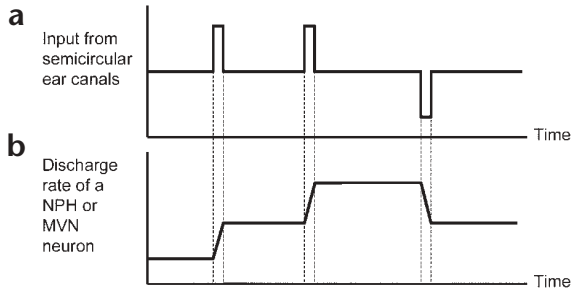

Fig. 1. The oculomotor system as an example of an integrator. (**a**) Head movements produce idealized signals reflecting head velocity. The sign depends on the direction of head movement. These signals represent the input to the integrator. (**b**) Discharge rate of a neuron in the integrator (medial vestibular nucleus (MVN) or nucleus prepositus hypoglossi (NPH) in the brainstem). The integrator shows a sustained increase in firing after each of two brief excitatory inputs and a decrease after a brief inhibitory input.

tion of the input current (I) is assumed to be approximately threshold linear (**Fig. 2b**, red curve); and the recurrent current depends on the total firing rate of the cells (**Fig. 2b**, blue line). The network is stable if the current provided by the feedback synapses exactly equals the current required to maintain the unit at its current firing level. This will happen if the red and blue curves overlap exactly, resulting in a continuum of stable points. The mistuning of the network is represented graphically as the non-overlap of the two curves, which is determined by the strength of the feedback synapses. Mistuning leads to drift, which continues until the network relaxes to the null position, losing the memory trace (**Fig. 2b**). If the feedback line is tuned to exactly match the neuronal response, the system is capable of storing information indefinitely. In practice, the tuning cannot be exact; the greater the mistuning, the greater the rate of drift (equation 1). In networks of this kind, synapses must be very finely tuned (to better than 1% accuracy) to account for the observed stability of the oculomotor integrator.

Integrator networks composed of bistable units

When bistable (hysteretic) units are used, the integrator can function even with loosely tuned feedback. To make this point clear, we first present a highly simplified model that can be solved exactly. The building block of this model is a neuronal unit whose firing rate is shown as a function of total input current (**Fig. 3a**). Firing begins abruptly at a given value of input current, $I^{\uparrow}$. If the current is then decreased, firing persists until the current falls below $I^{\downarrow}$. When the input current is within the interval ($I^{\downarrow}$, $I^{\uparrow}$), the unit can be either on (active) or off (silent), depending on the history of previous inputs. This implies that the unit's response is double-valued and exhibits hysteresis. The range within which the unit's response is bi-valued is determined by the properties that underlie bistability and is termed the hysteresis current, ΔI.

Robustness of the model can first be understood qualitatively by considering an ensemble of many units with the same ΔI, whose bistable regions overlap (**Fig. 3b**). Assume that the value of the input current is within the bistable range of one of the units (unit 2 in **Fig. 3b**). Further assume that unit 2 is active as are all the units below, such as unit 1. The units above, such as unit 3, are silent. If the input current changes within the range shown by the gray segment, none of the units changes state (units 1 and 2 remain on, while unit 3 stays off). Thus, if the state of the system is determined by the number of active units, this state is robust with respect to parameter variations, such as the input current. This type of robustness stems from bistability: the wider the bistable region, the larger the parameter variations that are tolerated. Because synaptic strength affects the input current, it follows that the system can also be stable despite substantial variations in synaptic strength.

Although this qualitative explanation is not rigorous, the model can be more exactly solved by considering an ensemble of many units that have the same ΔI but differ systematically in the threshold currents, $I^{\uparrow}$ (**Fig. 4a** and **b**). We first consider the response of this ensemble to the input current assuming no feedback connections. (Recall that this approach of separating the response of neuronal units and feedback was used for the finely-tuned model (**Fig. 2b**, red and blue lines respectively) and has been used in earlier studies[16].) In this model, the same value of the input current I is applied to all units in the ensemble. The response of this ensemble is described by the number of units, which are presently on. The number of such active units is denoted by n. If the input current increases, the number of active units can be found from the distribution of $I^{\uparrow}$ (**Fig. 4a**) and is equal to the number of units whose $I^{\uparrow}$ is smaller than the given value of input current I. Thus, as the input current is gradually increased, the number of active units increases gradually (line 1-2 in **Fig. 4c**). Importantly, when the current starts to decrease, as at point 2, there is a range (ΔI) in which no change in the number of active units occurs (line 2-3 in **Fig. 4c**). This is because such a decrease is necessary to switch the last excited unit off (**Fig. 3a**). The existence of this range makes this integrator design robust. If one further decreases the input current, beyond the value of ΔI, the dependence of the number of active units on the current follows the line 3-4, which is determined by the distribution of

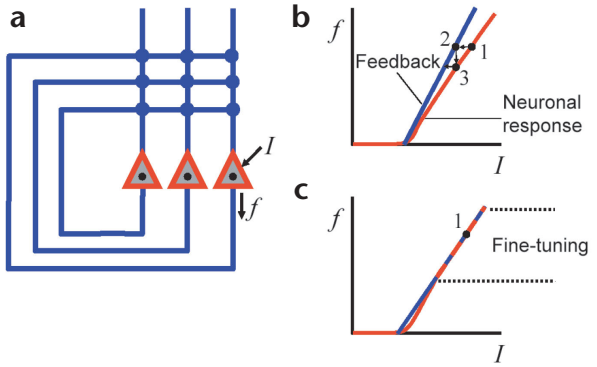

Fig. 2. The fine-tuning hypothesis. (**a**) The recurrent neuronal circuit: identical neuronal units (red triangles) with recurrent connections (blue lines). (**b**) Neuronal response (red) and recurrent feedback (blue) as a function of input. The red line gives the firing rate of a neuron as a function of input current, whereas the blue line shows the relationship between neuronal firing rate and the current, which is returned to the system by the feedback. The sequence of points 1-2-3 shows a trajectory of the system. (**c**) Precisely tuned feedback occurs when the blue and red lines approach each other, which is equivalent to ε approaching 0. If the red and blue line exactly coincide (neuronal response and feedback are perfectly tuned, $\varepsilon = 0$), as in the dashed region, the system never escapes from point 1, and is therefore capable of storing memory for a long time. As each point between the dashed lines has this property, the system is multistable.

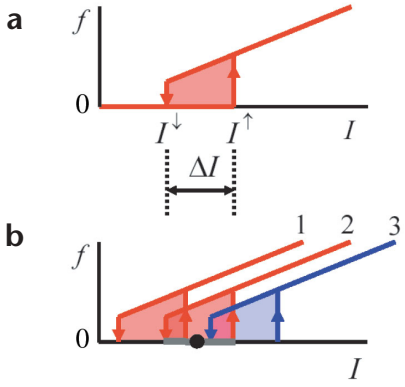

Fig. 3. Qualitative explanation of the new model containing bistable units. (**a**) The basic unit of the circuit exhibits hysteresis, the property that response depends on the history of inputs. At a given level of input, the unit can be either in a silent or firing state, depending on the history of inputs. Above $I^{\uparrow}$, units must turn on; below $I^{\downarrow}$, they must turn off. (**b**) Ensemble of three identical units, whose bistable regions overlap. The active units are shown by red lines; the response of an inactive unit is blue. The hysteretic loops of three units are shaded for visibility. If the input current (black dot) varies within the range shown by the gray segment, none of the units changes its state (switches on or off). This is the source of robustness of our model with respect to parameter variation.

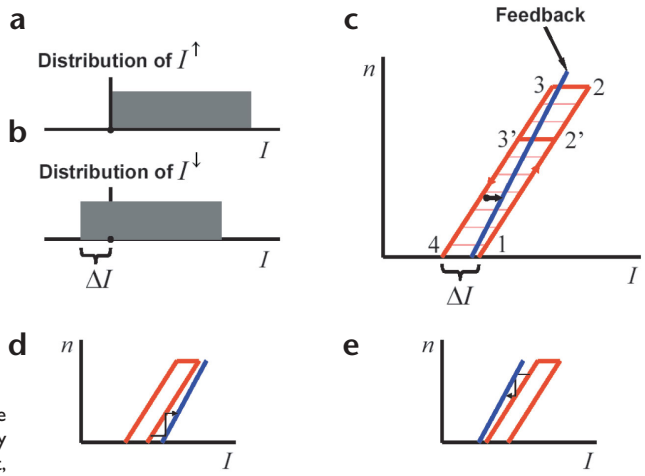

Fig. 4. Multistability in circuits with hysteretic units. (**a**, **b**) In the ensemble of many identical units, the thresholds for activation ($I^{\uparrow}$) and deactivation ($I^{\downarrow}$) are distributed uniformly within a certain range of input currents. (**c**) Number of active units as a function of external current. Assuming that the units are not connected, and therefore act independently, the number of active units for increasing current is determined by the distribution in (**a**) and is shown by the line 1-2. When the current decreases, the number of active units does not change unless the decrease exceeds the value of hysteresis, ΔI. After that, n is determined by the distribution in (**b**) and is shown by the line 3-4. Provided that the recurrent feedback (blue line) is between the the red lines (1-2 and 3-4), each point on the blue line is stable, and the system is multistable. Within the hysteretic loop, only horizontal movements are possible (pink lines). If a fluctuation in the input current displaces the system to a nearby point (black dot), the system must move horizontally to the original point (determined by value of feedback, black arrow) because the number of active units does not change between lines 1-2 and 3-4. (**d**) When external excitatory input arrives, it transiently offsets the feedback line beyond the hysteretic loop. Thus, new neurons get recruited. (**e**) For external inhibitory input, the feedback line is moved transiently to the left of the hysteretic loop, bringing some neurons below threshold.

thresholds $I^{\downarrow}$ (**Fig. 4b**). We conclude that the hysteretic response of individual units results in hysteresis in the response of the whole ensemble, given by the loop 1-2-3-4.

An important feature of this ensemble is that the region inside the hysteretic loop 1-2-3-4 can be reached with appropriate manipulation of the input current. This can be achieved if the input current is decreased at point 2' instead of at point 2 (**Fig. 4c**). The number of active units between point 2' and 3' is constant, similar to line 2-3. Thus we conclude that the number of active units cannot be changed as long as the system stays between lines 1-2 and 3-4, namely within the hysteretic loop (**Fig. 4c**). Thus, within the hysteretic loop, only horizontal movements are possible (pink lines in **Fig. 4c**). To change the number of active units, the borders of the hysteretic loop (lines 1-2 or 3-4) have to be reached.

Let us now take feedback connections into consideration. In our model, recurrent synapses provide the same amount of current into each unit. The value of feedback current is directly proportional the number of active units (n). This assumption simplifies the solution and at the same time mimics the saturation of recurrent synapses that use slow synaptic conductances[13]. Thus the contribution of a unit to the feedback current does not depend on its firing rate, and feedback current is simply proportional to the number of active units (**Fig. 4c**, blue line). If the feedback line is situated within the borders of the hysteretic loop, each point on the blue line is a stable point of the system. Indeed, assume that a fluctuation in the input current displaces the system to a nearby point (black dot in **Fig. 4c**). Because the number of active units does not change between lines 1-2 and 3-4, the system must move horizontally to the original point, determined by value of feedback: the original point on the blue line (black arrow in **Fig. 4c**). Thus, the system is multistable if the feedback line is positioned between lines 1-2 and 3-4. This can be achieved for a range of values of feedback strength (slope of the blue line). Therefore, the system does not have to be fine-tuned to have graded persistent activity.

For the system to function as an integrator, the external current must be able to change the number of active units. This can be achieved if an external current pulse moves the feedback line beyond the hysteretic loop. Thus, vestibular input is determined in our model by the offset of the feedback line or feedback-independent current, similar to previous models[13]. The trajectories of the system show that the longer the external current pulse, the further the system shifts up (**Fig. 4d**) or down (**Fig. 4e**). The increase in the number of active units is proportional to the duration of excitatory or inhibitory input current, and thus the system performs temporal integration.

Implementations with different forms of bistability

Bistability can result from a number of different mechanisms, two of which we have explored through simulated networks of conductance-based spiking neurons. In one of these, 'circuit-based bistability,' each unit is composed of a group of neurons that strongly excite one another[4,5,16,17]. This group can be either active or inactive as a result of recurrent excitation within the group. In the second mechanism, 'NMDAR-based bistability,' each unit is composed of only one neuron. Here the bistability arises from the voltage-dependent properties of the NMDA conductance at recurrent synapses.

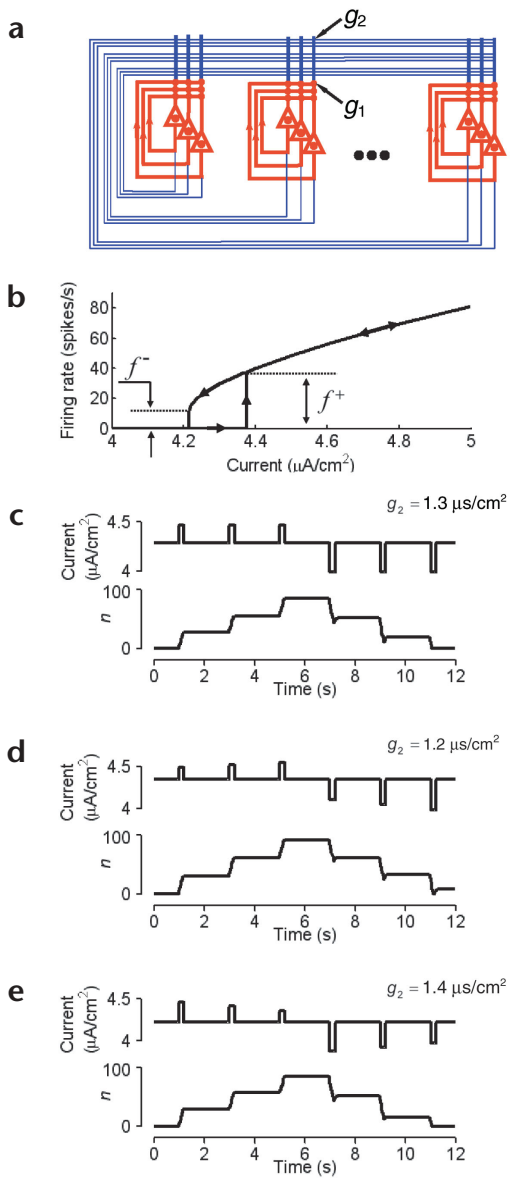

Fig. 5. Robust integrator network using circuit-based bistability. (**a**) The neural circuit combines two types of recurrent connections: local (red), which form units containing three neurons each, and global (blue), which allow integration. The local connections are formed by strong NMDA synaptic conductances g_l = 22 μs/cm^2. (**b**) The local connections make each unit bistable. This bistability is illustrated by the hysteretic dependence of the firing rate on the input current into each unit. The dependence is similar to the simple model for the unit in **Fig. 3a**. The hysteretic loop is described by different onset and offset firing rates (f^+ and f^-). (**c**) When NMDA conductances of the global feedback are set to g_2 = 1.3 μs/cm^2, the circuit operates as a neural integrator. That is, the number of the units in the active state (**c**, bottom trace) reflects the integral of the input current injected into the dendritic compartments (**c**, top trace). (**d**, **e**) The system can sustain considerable changes in the value of g_2 between 1.2 μs/cm^2 (**d**) and 1.4 μs/cm^2 (**e**) without the loss of multi-stability. Note that the heights of the input pulses are adjusted in the simulation to produce a uniform integration. This is equivalent to an adjustment of the gain of the integrator, when the feedback parameters are changed. Thus, in this model, the gain and robustness are disentangled: the network is robust, whereas the gain has to be adjusted by an external system.

Circuit-based bistability

In this model, we assume that each unit consists of a small group of neurons (three are used in the simulations) which form strong local recurrent connections with each other (**Fig. 5a**, red). These short-range recurrent connections make the response of this group of neurons hysteretic (bistable), similar to conventional short-term memory circuits[4,5,16]. Two stable states in the bistable region correspond to two self-sustained states of the network with positive feedback (**Fig. 5b**). In the silent state, the neurons do not provide feedback; in the active state, neuronal firing is sustained by strong recurrent feedback current.

In addition to the short-range connections, the bistable units are connected by long-range recurrent connections (**Fig. 5a**, blue). The value of $I^{\uparrow}$ varies systematically for each unit. This difference could arise in many different ways; we have implemented it generically as a difference in the feedforward current (Methods). Following previous work[13,18,19], all recurrent synapses are assumed to be slow (Methods).

A brief external excitatory input current to all the units produced a sustained increase in the number of active cells (**Fig. 5c–e**). The number of persistently firing cells increased with the number of excitatory input pulses and decreased with number of inhibitory pulses. When the synaptic strengths were changed by ±8%, the system displayed no drift in the number of active units between inputs (**Fig. 5d** and **e**). The system is therefore stable despite large parameter variation (16%). This implies that fine tuning is not necessary for stable operation of this integrator (**Fig. 5c–e**). These figures also demonstrate recruitment: the number of active units is determined by the value of the integral.

The network responded to prolonged pulses of input current as does the oculomotor integrator during smooth eye moments (**Fig. 6a**). We plotted the discharge rate of individual units as a function of the number of active units, which may be related to eye position for the case of the oculomotor integrator (**Fig. 6b**). When a neuron starts to fire, there is a sudden upward jump in firing rate (f^+); as the neuron stops firing, there is a sudden drop in firing frequency (f^-). However, because of hysteresis, the drop in frequency is smaller than the upward jump. This conclusion is independent of the solution of the read-out problem: how exactly the output of the circuit is related to the eye position. Thus, similar observation can be made by plotting the firing frequency of an individual neuron as a function of the average firing frequency of the entire network (**Fig. 6c**). Note that the onset (f^+) and offset (f^-) of firing are the same as in **Fig. 5b**. This property may help to identify this class of models in experiments (see Discussion).

NMDA receptor–based bistability

Because the NMDA conductance has a negative resistance region, the current–voltage curve of a neuron can be bistable if the NMDA conductance has an appropriate value relative to the leak conductance[20,21]. The origin of this bistability lies in the voltage-dependence of the NMDA current, produced by the Mg^{2+} block of the receptors[22,23]. In the down state, the leak conductance dominates and the neuron is at the resting potential with the NMDA conductance almost completely blocked. In the active state, the Mg^{2+} block of the NMDA conductance is relieved and this conductance is dominant. This state is stable because a hyperpolarizing deflection leads to an increase in inward current that restores the voltage to its original level. Conversely, a depolarizing deflection results in a decrease in driving force and thus less current flows until the original state

is restored. **Fig. 7a** shows current–voltage (*I–V*) relationships for neurons with a particular amount of dendritic leak and different magnitudes of the NMDA conductance. When the NMDA conductance is in a particular range, the *I–V* curves show two stable states (zero crossings with positive slope). If the NMDA conductance is too low, only the hyperpolarized state is stable; if it is too high, only the depolarized state is stable. Thus, the NMDA-dependent bistability of neurons requires recurrent synaptic input to bring the NMDA conductance into the appropriate range and is thus conditional upon the firing of other cells in the network (**Fig. 7b**). External inputs will switch the neuron into a firing state only if appropriate NMDA recurrent input from the network is present.

A network of these neurons (all-to-all connected without self-connections) can maintain multiple memory states if each cell is activated at a different level of synaptic feedback. This requires that each neuron have a different threshold for activation. This can be achieved by uniformly spacing the dendritic leak conductance of each neuron, leading to variations in the thresholds for bistability. Such a network can function as an integrator (**Fig. 7c**). A burst of inputs (see Methods) will turn on only those neurons that already receive sufficient recurrent excitation to be in the bistable range (middle curve in **Fig. 7a**). Once active, these neurons will continue to fire in the absence of the trigger input. Successive inputs recruit additional neurons, while raising the firing rate of previously active neurons (**Fig. 7d**). Neurons that were activated early now receive large NMDA input, which leaves only the depolarized membrane state stable (bottom curve in **Fig. 7a**). Thus, these neurons cannot be reset until they are brought back into the bistable range by a reduction in NMDA input. This model is robust to changes in feedback gain, tolerating ±20% changes in the NMDA conductance values without compromising integrator function.

Systematic variations in leak conductance, however, are not the only way to achieve variation in the threshold. Notably, we found that even choosing all the variables that control bistability (the leak conductance, the slow voltage-gated conductances and the strengths of the NMDA conductances) from random distributions (see **Supplementary Methods** online) resulted in a functional integrator network (**Fig. 8a**). Although the quantitative details of individual randomized network instantiations differ somewhat, our model functioned as an integrator for each set of parameters (**Fig. 8b**). Note that the overlap of bistable ranges required for integrator function (**Fig. 2b**) can occur even when conductances are chosen at random (**Fig. 8c**). Thus, rather than requiring precise tuning, this form of integrator network exploits the variability of synaptic and intrinsic properties to perform robust integration.

Discussion

Here we present a new class of integrator circuits that uses bistable units instead of simple threshold-linear neurons. We demonstrated the function of this class of integrators both through the analysis of simplified bistable units and through simulations of networks of compartmental neurons with more realistic properties. Although the particular mechanisms that underlie bistability are not as important as the existence of a bistable regime and differing regions of bistability among the units, we have shown that robust integrator function can result from two quite different mechanisms for generating bistability. In one, bistability is achieved by strong recurrent excitation between a small group of neurons. In the other, bistability arises from the voltage-dependence of NMDA receptors (NMDARs).

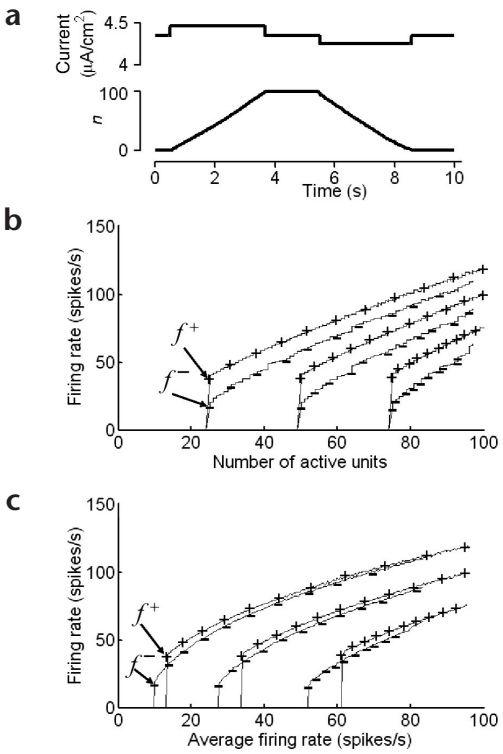

Fig. 6. Response to weak inputs. (**a**) Same model as in **Fig. 5c** subject to two long pulses of the input current of opposite sign (top). The number of active units (bottom) reflects the integral of the input current, similarly to **Fig. 5c**. The number of active units therefore may be treated as the eye position. (**b**) Firing rates for three neurons in the integrator, as a function of the number of active units. Plus and minus markers indicate the trace for monotonously increasing and decreasing number of active units, which correspond to eye movements in the on and off directions. These three neurons belong to unit numbers 25, 50 and 75 out of the total 100. The traces exhibit different onset and offset firing rates f^+ and f^-. These values are the same as in **Fig. 5b**. The dependence of the firing rate on the number of active units is close to linear above threshold. (**c**) The onset and offset firing rates are independent of the solution of the readout problem. Thus, if the firing rates for the same three neurons as in (**b**) are plotted versus the average firing rate of the whole population, the values of f^+ and f^- are the same as in (**b**).

An attractive feature of integrator networks with bistable units is that the strength of recurrent synapses need not be precisely tuned. Indeed, in the case of the NMDA-dependent bistability, we found that the synaptic strengths can be randomly selected from within a certain range. By eliminating the requirement for fine tuning, our work strengthens the case that recurrent excitatory networks can perform integrator function. Our models also retain the ability of previous models to explain the recruitment phenomenon seen in the oculomotor integrator of the goldfish[12]. These new integrator models are therefore both robust and physiologically plausible.

Bistability and the role of NMDA channels

Single-neuron current injection experiments have tested for intrinsic bistability and found that it is not present in the goldfish medulla[24]. This finding is consistent with both implementations of our model. In the circuit-based mechanism, each unit consists of a

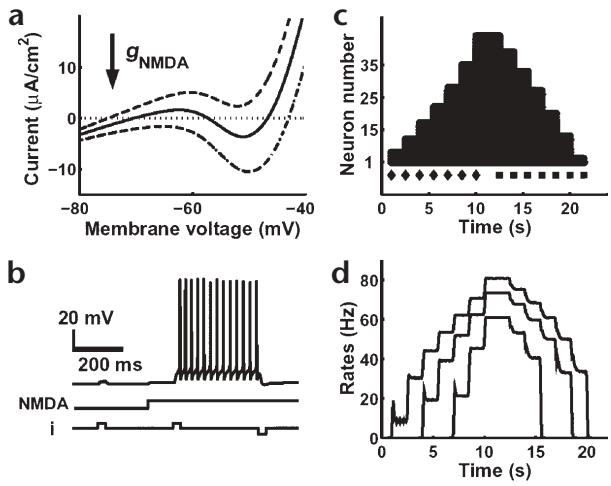

Fig. 7. Integrator network based on the voltage-dependence of the NMDA conductance. (**a**) Current-voltage curve demonstrating how bistability arises when the NMDA conductance has the appropriate value relative to leak. The currents plotted are total steady-state currents (intrinsic and synaptic currents, but excluding spiking currents) and with perfect voltage-clamp. (**b**) Persistent firing is conditional. An input pulse fails to initiate persistent firing in this cell. However, if recurrent input (NMDA) is active, the cell is now bistable and the same input pulse (i) can initiate persistent firing. A pulse of the opposite sign can return the cell to silent state. (**c**) The number of active neurons is shown as a function of time. Positive and negative inputs are shown below with diamonds and squares, respectively. This plot also shows how neurons participate in a recruitment process. (**d**) Shows the time evolution of firing rates for three example neurons.

small group of neurons and it is the unit that is bistable, not a single neuron. This implies that current injection into a single neuron would not reveal bistability. In the case of the NMDAR-based mechanism, current injection could potentially reveal bistability, but two conditions must be met for each recorded neuron. First, the injections have to be at the site where bistability is generated. This may be dendritic, since the NMDAR-based bistability is generated in dendrites. Second, at only the gaze angle at which firing begins is there enough NMDA conductance to produce bistability. This gaze angle could be identified operationally as the angle at which firing terminates during smooth pursuit. It is not clear whether these conditions were met in previous experiments[24].

The role of NMDA receptors in integrator function has previously been tested by global application of NMDA antagonists[25–27] that disrupt integrator function, as indicated by a greater drift rate of gaze. Although it is likely that the NMDA channel has an important role in integrator function, these studies do not address whether it is the long time constant of the NMDA channel, or its voltage-dependent activation, that is important. Circuit-based models require only the first property; NMDAR-based models depend primarily on the second property.

Experimental predictions

First, as the integrator models described here do not require precisely tuned synaptic weights, it should be possible to apply receptor antagonists in low doses and show that the integrator still functions despite a decrease in synaptic strength. A second prediction relates to the presence of hysteresis in cellular response during slow integration (**Fig. 6**). During slow eye movements in the 'on' direction, there is sudden onset of firing (f^+); during slow movement in the 'off' direction, there is sudden offset of firing (f^-) that is smaller than f^+. Thus, our model may be identified by the difference between f^+ and f^-. The absolute size of these jumps will depend on the conductances that control bistability; these properties have not yet been determined in any known integrator. More generally, some other form of hysteresis may be sufficent to make the integrator robust. Our suggestion relies on the presence of microscopic hysteresis in cellular response, rather than on its physiological basis. This is evident from the simplified model, which does not specify the exact nature of hysteresis but nevertheless displays robustness.

Third, as a minimal-threshold current is required to turn a bistable unit on or off, it should be possible to apply such inputs over and over again without any stable change in the network output. This is an essential difference between multistable systems of the kind proposed here and finely tuned networks that integrate all inputs including noise, a sensitivity that makes them vulnerable to drift.

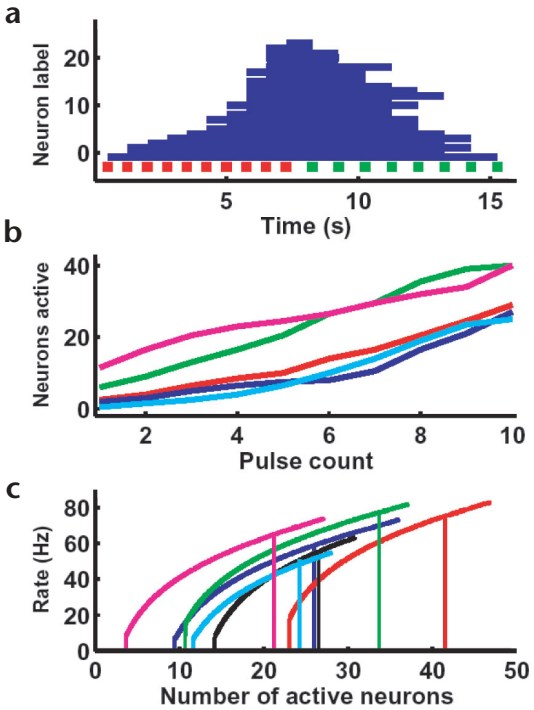

Fig. 8. Robust integrator function when conductances are randomly chosen. Leak and intrinsic conductances are chosen randomly as described (**Supplementary Methods**). (**a**) The number of active neurons is shown as function of time. Positive and negative inputs are shown below with red and green squares, respectively. The neurons are ordered by their dendritic leak conductance. Note that the order in which neurons turn on/off is determined by the combination of the leak conductance as well as the values of intrinsic and synaptic conductances. (**b**) The mean number of neurons active in the integrator network at a given pulse count (mathematical integral of on and off pulses). Different curves are the result of different simulations where the values of several conductances were randomly chosen (**Supplementary Methods**). In all simulations, we found a monotonic relationship between the number of active neurons and pulse count. (**c**) Heterogeneous response characteristics of neurons taken from one of the simulations shown in (**b**). Note the large range of bistability that underlies the robustness of this model.

A fourth aspect of our model that may be testable concerns the relationship of the conductances of single cells to their recruitment pattern. The model predicts that the magnitude of some intrinsic or synaptic conductance correlates with whether the cell fires early or late during the integration process. For example, cells firing late may either have weaker inputs from other units or enhanced leakage conductances that make them require strong input to switch to the up state.

Finally, given the difference between the two forms of bistability that we have simulated (Methods), procedures that remove voltage-dependence (for example, lowering Mg^{2+}) but do not affect kinetics should be useful in distinguishing between the NMDA-based and the circuit-based mechanisms for bistability.

Exploiting heterogeneity to produce a robust integrator

Our integrator network models require that different units have different activation thresholds, and there are several ways in which this may arise. In the model based on the voltage-dependence of the NMDA receptor, we achieved the required spread by choosing relevant biophysical parameters of the model from random distributions. This result suggests that natural neuron-to-neuron variability could be sufficient to produce the differential thresholds called for in this class of integrator models. Thus, this design exploits the heterogeneity among neurons. It should be emphasized that the parameters of the distributions are set, and randomness refers to different draws from these distributions. This limitation is reasonable because physiological parameters cannot vary over an indefinitely wide range; they are constrained by the background genotype.

Implications for other systems

Integration has been most extensively studied in the oculomotor system[11] and in circuits that represent head direction[28,29], but there is increasing evidence that integrators may have even broader use. Neural firing during a decision task gradually increases before the time of decision, and that increase may reflect the operation of an integrator that accumulates both negative and positive evidence until a net evidence criterion is met for making a decision[30].

The neurons in our integrator models show persistent activity at graded levels. Such graded persistence has been observed in neurons in the prefrontal cortex and is termed parametric working memory. When monkeys are required to hold in memory a graded representation of a variable frequency stimulus, neurons fire persistently during the working memory period at a rate proportional to the stimulus variable remembered[6]. With an appropriate stimulus-to-input conversion in our network, the firing of a neuron can be brought to any arbitrary level and maintained there during a delay period. Thus this class of models is a good candidate for mediating parametric working memory.

Conceptually similar multistability may also occur in the case of the graded, persistent modification of synaptic strengths, which also implements temporal integration. Consider a connection with modifiable elements that are bistable and have different thresholds for synaptic strengthening (LTP)[31,32]. Initially, the plasticity induction protocol might be sufficient to trigger LTP only in the element with the lowest threshold for plasticity. However, the resulting LTP would make the same induction protocol generate more overall depolarization and hence trigger plasticity at elements with a higher plasticity threshold, owing to positive feedback.

The need for multistabilty may be widespread in biological systems, as it has been found in a range of biochemical systems[33] and may underlie the maintenance of phenotypic differences during development[34,35]. The need for hybrid models, involving both Boolean and continuous elements, has been suggested in genetic regulatory circuits[35]. Multistable networks have also been found in immune systems[36]. The question remains whether there are common rules by which bistable elements can interact to produce multistability in disparate systems.

Methods

For details of simulations, see **Supplementary Methods** online. Briefly, the circuit-based bistability model contains 300 two-compartmental excitatory neurons[20,37] connected all-to-all by NMDA synapses. This group of neurons is divided into 100 units, containing three neurons each. The synaptic conductances between each neuron and other neurons inside and outside the same unit are $g_1 = 22$ μs/cm^2 and $g_2 = 1.3$, 1.2, 1.4 and 1.3 for **Fig. 5c**, **d**, **e** and **Fig. 6**, respectively. An input current was applied to dendritic compartments of all neurons. The input current is the same for the neurons belonging to the same unit. Across units, it is uniformly spaced (distributed) in the interval 0–1.5 μA/cm^2 to reproduce distributions in **Fig. 4a** and **b**. The total input current into the neuron receiving the maximum current in the whole population is shown in **Figs. 5** and **6**. The network (40 neurons) using NMDAR-based bistability has all-to-all connections (without autapses). Each neuron was modeled using a somatic compartment with spike currents and a dendritic compartment containing persistent sodium current, slow and transient potassium currents and a leak[38]. Recurrent synapses had a large NMDA/AMPA ratio. The leak conductance in the dendrite is spaced uniformly in **Fig. 7** and chosen randomly from a uniform distribution in **Fig 8**. The network receives external inputs via fast excitatory and inhibitory synapses from burst neurons. A small (6-neuron) version of the simulation program can be found at http://www.bio.brandeis.edu/lismanlab/integrator.

Note: Supplementary information is available on the Nature Neuroscience website.

Acknowledgments

The authors are grateful to C. Kaneko and H. Sompolinsky for encouragement and to S. Seung and his group for numerous helpful suggestions. A.A.K. and A. K. were supported by fellowships from the Swartz and Alfred P. Sloan Foundations. We also acknowledge grants NIH 461434 and NSF 450578.

Competing interests statement

The authors declare that they have no competing financial interests.

RECEIVED 28 FEBRUARY; ACCEPTED 18 JUNE 2002

1. Miyashita, Y. & Chang, H. S. Neuronal correlate of pictorial short-term memory in the primate temporal cortex. *Nature* **331**, 68–70 (1988).
2. Gnadt, J. W. & Andersen, R. A. Memory related motor planning activity in posterior parietal cortex of macaque. *Exp. Brain Res.* **70**, 216–220 (1988).
3. DeSouza, J. F. *et al.* Eye position signal modulates a human parietal pointing region during memory-guided movements. *J. Neurosci.* **20**, 5835–5840 (2000).
4. Goldman-Rakic, P. S. Cellular basis of working memory. *Neuron* **14**, 477–485 (1995).
5. Fuster, J. M. *Memory in the Cerebral Cortex: An Empirical Approach to Neural Networks in the Human and Nonhuman Primate* (MIT Press, Cambridge, Massachusetts, 1995).
6. Romo, R., Brody, C. D., Hernandez, A. & Lemus, L. Neuronal correlates of parametric working memory in the prefrontal cortex. *Nature* **399**, 470–473 (1999).
7. Wang, X. J. Synaptic reverberation underlying mnemonic persistent activity. *Trends Neurosci.* **24**, 455–463 (2001).
8. Cannon, S. C., Robinson, D. A. & Shamma, S. A proposed neural network for the integrator of the oculomotor system. *Biol. Cybern.* **49**, 127–136 (1983).
9. Fukushima, K., Kaneko, C. R. & Fuchs, A. F. The neuronal substrate of integration in the oculomotor system. *Prog. Neurobiol.* **39**, 609–639 (1992).
10. Arnold, D. B. & Robinson, D. A. A learning network model of the neural integrator of the oculomotor. *Biol. Cybern.* **64**, 447–454 (1991).
11. Robinson, D. A. Integrating with neurons. *Annu. Rev. Neurosci.* **12**, 33–45 (1989).
12. Aksay, E., Baker, R., Seung, H. S. & Tank, D. W. Anatomy and discharge properties of pre-motor neurons in the goldfish medulla that have eye-position signals during fixations. *J. Neurophysiol.* **84**, 1035–1049 (2000).

13. Seung, H. S., Lee, D. D., Reis, B. Y. & Tank, D. W. Stability of the memory of eye position in a recurrent network of conductance-based model neurons. *Neuron* **26**, 259–271 (2000).
14. Seung, H. S. How the brain keeps the eyes still. *Proc. Natl. Acad. Sci. USA* **93**, 13339–13344 (1996).
15. Seung, H. S., Lee, D. D., Reis, B. Y. & Tank, D. W. The autapse: a simple illustration of short-term analog memory storage by tuned synaptic feedback. *J. Comput. Neurosci.* **9**, 171–185 (2000).
16. Wilson, H. R. & Cowan, J. D. Excitatory and inhibitory interactions in localized populations of model neurons. *Biophysical J.* **12**, 1–24 (1972).
17. Hopfield, J. J. Neural networks and physical systems with emergent collective computational abilities. *Proc. Natl. Acad. Sci. USA* **79**, 2554–2558 (1982).
18. Wang, X. J. Synaptic basis of cortical persistent activity: the importance of NMDA receptors to working memory. *J. Neurosci.* **19**, 9587–9603 (1999).
19. Koulakov, A. A. Properties of synaptic transmission and the global stability of delayed activity states. *Network* **12**, 47–74 (2001).
20. Lisman, J. E., Fellous, J. M. & Wang, X. J. A role for NMDA-receptor channels in working memory. *Nat. Neurosci.* **1**, 273–275 (1998).
21. Schiller, J. & Schiller, Y. NMDA receptor–mediated dendritic spikes and coincident signal amplification. *Curr. Opin. Neurobiol.* **11**, 343–348 (2001).
22. Nowak, L., Bregestovski, P., Ascher, P., Herbet, A. & Prochiantz, A. Magnesium gates glutamate-activated channels in mouse central neurons. *Nature* **307**, 462–465 (1984).
23. Mayer, M. L., Westbrook, G. L. & Guthrie, P. B. Voltage-dependent block by Mg^{2+} of NMDA responses in spinal cord neurons. *Nature* **309**, 261–263 (1984).
24. Aksay, E., Gamkrelidze, G., Seung, H. S., Baker, R. & Tank, D. W. *In vivo* intracellular recording and perturbation of persistent activity in a neural integrator. *Nat. Neurosci.* **4**, 184–193 (2001).
25. Godaux, E., Cheron, G. & Mettens, P. Ketamine induces failure of the oculomotor neural integrator in the cat. *Neurosci. Lett.* **116**, 162–167 (1990).
26. Mettens, P., Cheron, G. & Godaux, E. NMDA receptors are involved in temporal integration in the oculomotor system of the cat. *Neuroreport* **5**, 1333–1336 (1994).
27. Priesol, A. J., Jones, G. E., Tomlinson, R. D. & Broussard, D. M. Frequency-dependent effects of glutamate antagonists on the vestibulo-ocular reflex of the cat. *Brain Res.* **857**, 252–264 (2000).
28. Taube, J. S., Muller, R. U. & Ranck, J. B. Jr. Head-direction cells recorded from the postsubiculum in freely moving rats. I. Description and quantitative analysis. *J. Neurosci.* **10**, 420–435 (1990).
29. Blair, H. T. & Sharp, P. E. Anticipatory head direction signals in anterior thalamus: evidence for a thalamocortical circuit that integrates angular head motion to compute head direction. *J. Neurosci.* **15**, 6260–6270 (1995).
30. Shadlen, M. N. & Newsome, W. T. Neural basis of a perceptual decision in the parietal cortex (area lip) of the rhesus monkey. *J. Neurophysiol.* **86**, 1916–1936 (2001).
31. Lisman, J. E. A mechanism for memory storage insensitive to molecular turnover: a bistable autophosphorylating kinase. *Proc. Natl. Acad. Sci. USA* **82**, 3055–3057 (1985).
32. Petersen, C. C., Malenka, R. C., Nicoll, R. A. & Hopfield, J. J. All-or-none potentiation at CA3-CA1 synapses. *Proc. Natl. Acad. Sci. USA* **95**, 4732–4737 (1998).
33. Cimino, A. & Hervagault, J. F. Irreversible tansitions in a model substrate cycle. An experimental illustration. *FEBS Lett.* **263**, 199–205 (1990).
34. Novick, A. & Weiner, M. Enzyme induction as an all-or-none phenomenon. *Proc. Natl. Acad. Sci. USA* **43**, 553–566 (1957).
35. McAdams, H. H. & Arkin, A. Simulation of prokaryotic genetic circuits. *Annu. Rev. Biophys. Biomol. Struct.* **27**, 199–224 (1998).
36. Segel, L. A., Jager, E., Elias, D. & Cohen, I. R. A quantitative model of autoimmune disease and T-cell vaccination: does more mean less? *Immunol. Today* **16**, 80–84 (1995).
37. Pinsky, P. & Rinzel, J. Intrinsic and network rhythmogenesis in a reduced Traub model for CA3 neurons. *J. Comput. Neurosci.* **1**, 39–60 (1994).
38. Kepecs, A. & Raghavachari, S. in *Advances in Neural Information Processing Systems* Vol. 14 (eds. Dietterich, T. G., Becker, S. & Ghahramani, Z.) (MIT Press, Cambridge, Massachusetts, 2002).