Nonlinear network models of the oculomotor integrator

D. D. Lee, B. Y. Reis, H. S. Seung, and D. W. Tank

{ddlee,reis,seung,dwt}@physics.bell-labs.com

Bell Laboratories
Lucent Technologies
Murray Hill, NJ 07974

Abstract

The neural integrator of the oculomotor system is modeled as a network of spiking, conductance-based model neurons. The static function of the integrator, holding the eyes still when the head is fixed in space, is the focus of the modeling. The synaptic weight matrix, which is of outer product form, is tuned by minimizing the mean squared drift velocity of the eyes over a range of eye positions, leading to an approximate line attractor dynamics. The conductance-based model is reduced to a rate-based one to simplify the tuning procedure.

Introduction

In the time intervals between saccades, the eyes are held still by a neural network integrator located in the hindbrain [1, 2]. With the head fixed in space, constant neural activity in the integrator is necessary for gaze stabilization. The pattern of activity is different for each eye position, even though the visual and vestibular inputs to the integrator remain the same. The origin of the integrator's ability to maintain multiple persistent patterns of activity is not obvious, since its constituent neurons do not appear to possess long intrinsic persistence times. A network mechanism is thus needed to explain the observed long time constants associated with gaze stabilization in the dark.

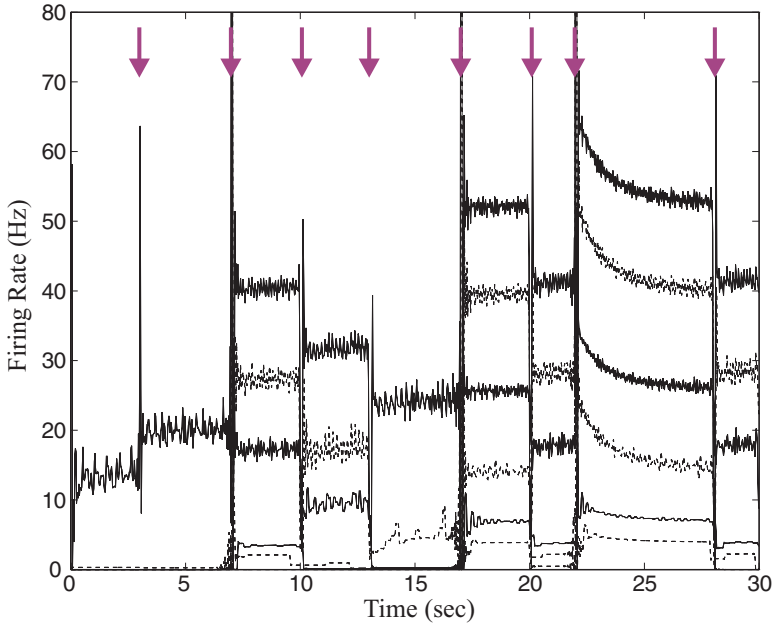

Figure 1: Spiking conductance-based model of the integrator with 40 neurons. The firing rates of six of the neurons are shown being driven by pulses of input from saccadic command neurons at the times indicated by the arrows.

Conductance-based model

Previous models of the integrator have relied on positive feedback in networks of completely linear neurons [3, 4, 5, 6]. In contrast, the present work uses conductance-based model neurons as described by Hansel and Sompolinsky [7]. Their nonlinear current-discharge relationships qualitatively match intracellular measurements of real integrator neurons [8, 9]. Synapses were modeled as current sources, with a time dependence described by the relation $e^{-t/\tau_1} - e^{-t/\tau_2}$, where $\tau_1 = 150$ ms and $\tau_2 = 20$ ms. The synaptic currents were assumed to sum linearly and the synaptic strengths were tuned by an optimization procedure that is described in the next section.

Figure 1 depicts the response of the model network to short pulses of input from saccadic command neurons at the times indicated by the arrows. Shown are the instantaneous firing rates (1/ISI) of the neurons. Eye position is not shown in the figure, but was modeled as a weighted linear sum of the rates of the neurons. Thus the firing rates and eye position all increase and decrease in steps, and are roughly constant in time between the quick saccadic movements that take place at the arrows. The reason for the "sag" in response to the second to last saccadic input will be explained in a later section.

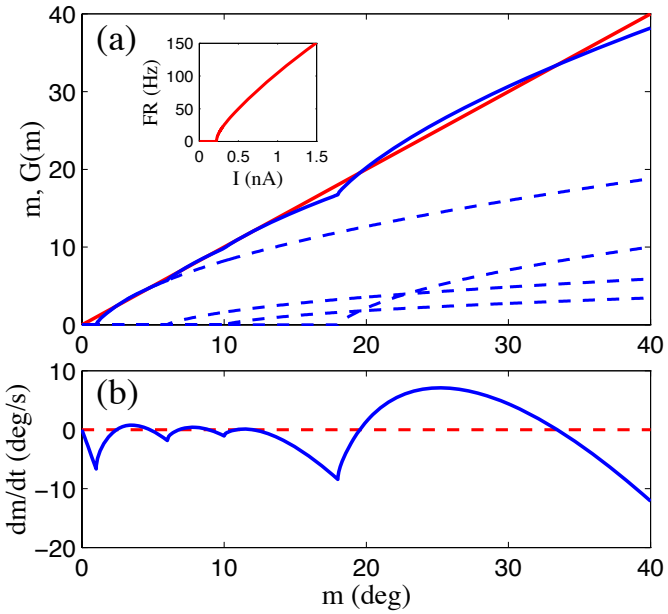

Figure 2: a) Approximation of a line attractor using 4 nonlinear neurons. The inset shows the rate-current relationship of an individual neuron. b) The resulting drift velocity of the eye as a function of eye position.

The model captures two essential aspects of integrator operation that are observed experimentally. The first is that the firing rate of each individual neuron is directly related to eye position, except when the neuron is below threshold. The second is that the eye position signal is intrinsically generated by the network itself, not read out from some other area in the brain. This is clear from the fact that the inputs to the model network only briefly pulse during saccades and otherwise remain constant over time, yet the network maintains persistent neural activity between saccades that is proportional to the different eye positions.

Rate-based model

The task of tuning the synaptic weights of the conductance-based model was simplified by reducing it to a rate-based model. To carry out this reduction, the steady-state current-discharge relationship of an individual conductance-based model neuron was measured from numerical simulations and is shown in the inset to Fig. 2(a). Thus the biophysical complexities of the conductance-based model were reduced to a simple functional relationship between the instan-

taneous firing rate v_i of a neuron and the sum total u_i of its synaptic input currents

$$v_i = g(u_i) \ . \tag{1}$$

The dynamics of the synaptic currents u_i, in turn, depend linearly on the firing rates v_j of the presynaptic neurons

$$\tau_s \frac{du_i}{dt} + u_i = \sum_{j=1}^{N} W_{ij} v_j + h_i \ . \tag{2}$$

The synaptic weight W_{ij} describes the influence of neuron j firing on another neuron i. In addition, the feedforward inputs coming from outside the integrator onto neuron i are lumped together in the h_i term. The time scale τ_s of all the synapses was taken to equal $\tau_1 = 150$ ms in the conductance-based model. This reduced rate model is an excellent approximation to the conductance-based model when the firing rates change slowly.

The synaptic weight matrix of the model was taken to be the outer product form $W_{ij} = \xi_i \, \eta_j$. Such a form emerges naturally when a Hebb-like learning rule is used to train the network, as will be discussed in a future publication. This type of weight matrix causes all dynamical trajectories to be attracted to a line in state space of the form

$$u_i = m\xi_i + h_i \ . \tag{3}$$

Relaxation onto this line occurs on the short time scale τ_s. Once the network is on the line, its dynamics is described by the scalar variable m which serves as the internal representation of eye position. The dynamical equation for m can be found by substituting Eq. (3) into Eqs. (1) and (2) to yield

$$\begin{aligned} \tau_s \frac{dm}{dt} &= \sum_{i=1}^{N} \eta_i \, g(m\xi_i + h_i) - m \\ &\equiv G(m) - m \ . \end{aligned} \tag{4}$$

The tuning procedure adjusts the synaptic weights such that the right hand side of Eq. (4) is very small, i.e. $G(m) \approx m$. Then the decay rate $dm/dt \approx 0$ and the variable m is approximately constant in time.

Eye position can be read out linearly from the network,

$$E = \sum_i \eta_i v_i \ . \tag{5}$$

If $G(m) \approx m$ in Eq. (4), then $E \approx m$ when the network is on the line of Eq. (3). Then the actual eye position E is equal to the internal representation m in the integrator. Furthermore,

$$v_i \approx g(E\xi_i + h_i) \tag{6}$$

when the network is on the line, so that the relationship between v_i and E is approximately linear above the firing threshold with a slope proportional to ξ_i. In other words, the position sensitivity of neuron i is directly proportional to the parameter ξ_i.

Optimal approximation

The tuning procedure is to minimize a cost function equivalent to the mean square drift velocity averaged over a range of m,

$$\left\langle \left| \sum_{i=1}^{N} \eta_i \, g(m\xi_i + h_i) - m \right|^2 \right\rangle_m . \tag{7}$$

The ξ_i and h_i parameters, and the range of m are fixed so that the ranges of firing rates, currents, and eye positions seen in the model correspond roughly to what is seen experimentally. The minimization is done with respect to the η_i parameters only.

If $g(u)$ were linear, then η_i could be tuned so that $G(m) = m$ in Eqs. (4) and (7) exactly. Such linearity has been the basis of previous network models of the integrator [3, 4, 5, 6]. This implies that the drift velocity dm/dt is zero at all positions m and that every point on the line is a fixed point of the dynamics. Thus Eq. (3) describes an attractive line of fixed points, otherwise known as a *line attractor* [2].

Fig. 2 illustrates what happens to the line attractor when $g(u)$ is nonlinear. The upper part of the figure shows the optimal approximation $G(m) \approx m$ achievable with four nonlinear neurons. It is clear that superpositions of scaled and translated versions of the current-discharge relationship $g(\xi_i m + h_i)$ produce a nonlinear curve that cannot exactly match the linear function m over the whole continuous range of m. Instead, the optimal η_i generates a nonzero drift velocity dm/dt from Eq. 4 that is shown in Fig. 2(b). In this case, the integrator network contains a finite number of fixed points corresponding to the zero crossings in Fig. 2(b) and therefore is only an approximate line attractor.

The form of the optimal approximation can vary, depending on biological assumptions which enter through the choice of the parameters ξ_i and h_i and the constraints on the η_i optimization. Currently we are investigating a wide variety of model networks resulting from different biological assumptions. Fig. 2 shows just one example, which was chosen to illustrate a point about the interplay between recruitment and sublinearity. The ξ_i were chosen randomly within a positive range, giving all neurons the same on-direction. The inputs h_i were randomly chosen such that the thresholds of the neurons were distributed over only the lower half range of m for which the cost function was minimized.

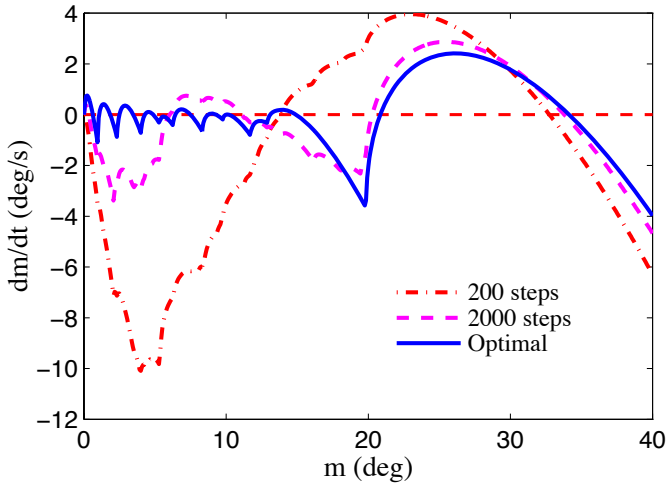

Figure 3: Training a network of 40 neurons using gradient descent. The drift velocity as a function of eye position is shown after 200 and 2000 iterations along with the optimal solution.

Furthermore, the η_i were constrained to be nonnegative resulting in a nonnegative least squares optimization that can be done by standard methods. This constraint along with the positivity of the ξ_i gives rise to a network with only excitatory connections.

Figure 2 illustrates the importance of nonlinear recruitment in compensating for the sublinear behavior of neurons above threshold. This recruitment can also be seen in Fig. 1 as firing rates increase. In the region where there are no thresholds, the network is unable to compensate by recruiting additional neurons and the drift velocities become significantly worse. This results in the large drift at high firing rates which is seen in the simulation of the conductance-based model in Fig. 1.

Learned approximation

The least squares optimization of Eq. (7) was calculated using a nonbiological algorithm. As will be shown elsewhere, a biologically plausible synaptic learning rule can perform the same optimization in an online fashion. However, such learning rules generally lead to solutions that are only approximately optimal. A solution qualitatively similar to these suboptimal solutions is obtained by performing gradient descent on the cost function in Eq. (7) and stopping short of the true minimum. Figure 3 shows the result of 200 and 2000 projected gradient descent steps on a network of 40 neurons. Notice that the drift velocity tends

to be greater and the number of fixed points is less than in the optimal solution. Although the quality of the solution improves as more gradient descent steps are taken, the rate of improvement slows over time because the cost function in Eq. (7) is ill-conditioned.

Discussion

Our model suggests that a biological approximation to line attractor dynamics has two sources of error. One source is the finite error of the best approximation by nonlinear neurons. The other is the suboptimality of learned approximations. Both types of error lead to a drift velocity that depends on eye position. In our current research, we are trying to understand these two sources of error and their relative contribution in real networks.

Measurements of eye position in the dark generally reveal a position-dependent drift velocity [10], which can be compared with model results like those shown in Fig. 2(b). Our model is able to relate the physiological properties of neurons and synapses in the integrator with a behavioral quantity, the drift velocity of the eyes. Making experimentally testable predictions, however, is complicated because the drift velocity can vary depending on the biological assumptions that are made.

We are also exploring the effects of dilute connectivity on our model, along with the role of inhibition and excitation. The present model is able to qualitatively integrate its inputs in the sense that velocity-coded inputs cause changes in position. But achieving quantitatively correct linear integration of input signals from nonlinear neurons requires further tuning of the system [11, 12]. Furthermore, modeling integrator function in conditions when rates vary quickly, as during saccades and high velocity VOR, will require model neurons that exhibit spike-frequency adaptation [13].

Acknowledgements

We are grateful to B. Shraiman for his assistance in the early stages of this investigation. We would also like to thank E. Aksay, R. Baker, B. Mensh, and H. Sompolinsky for helpful discussions.

References

[1] D. A. Robinson. Integrating with neurons. *Annu. Rev. Neurosci.* 12, 33 (1989).

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

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

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

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

[6] D. B. Arnold and D. A. Robinson. A neural network model of the vestibulo-ocular reflex using a local synaptic learning rule. *Phil. Trans. R. Soc. Lond.* B337, 327 (1992).

[7] D. Hansel and H. Sompolinsky. Chaos and synchrony in a model of a hypercolumn in visual cortex. *J. Comput. Neurosci.* 3, 7 (1996).

[8] M. Serafin, C. de Waele, A. Khateb, P. P. Vidal, and M. Mühlethaler. Medial vestibular nucleus in the guinea-pig. I. Intrinsic membrane properties in brainstem slices. *Exp. Brain Res.* 84, 417 (1991).

[9] S. du Lac and S. G. Lisberger. Membrane and firing properties of avian medial vestibular nucleus neurons in vitro. *J. Comp. Physiol.* A176, 641 (1995).

[10] K. Hess, H. Reisine, and M. Dursteler. Normal eye drift and saccadic drift correction in darkness. *Neuro-ophthalmol.* 5, 247 (1985).

[11] F. Miles and S. Lisberger. Plasticity in the vestibulo-ocular reflex: a new hypothesis. *Annu. Rev. Neurosci.* 4, 273 (1981).

[12] S. du Lac, J. Raymond, T. Sejnowski, and S. Lisberger. Learning and memory in the vestibulo-ocular reflex. *Annu. Rev. Neurosci.* 18, 409 (1995).

[13] R. Quadroni and T. Knöpfel. Compartmental models of type A and type B guinea pig medial vestibular neurons. *J. Neurophysiol.* 72, 1911 (1994).