Inhibitory control by an integral feedback signal in prefrontal cortex: a model of discrimination between sequential stimuli.

Paul Miller and Xiao-Jing Wang

Supporting Information

Download. *Xppaut* files used to simulate the linear rate model, and the C++ code used in the full spiking network simulations are available for download from
www.wanglab.brandeis.edu/people/miller/discriminator.html .

On the website are instructions for downloading, compiling, running, and altering the C++ code.

Network simulations. This subsection is an expanded version of *Network Simulations* from the main text and refers to Fig. 1*c*.

We simulate the model using a network of interconnected leaky integrate-and-fire neurons. The network architecture is shown schematically in Fig. 1*c*. Subpopulations of 400 identical excitatory neurons (Fig. 1*c*, squares) or 200 identical inhibitory neurons (Fig. 1*c*, circle) all receiving independent background noise input but common network input form five components of the network.

Comparison. The comparison (C) unit was modeled as a population of 400 excitatory neurons. During the cue, each C neuron receives a separate Poisson spike train at a rate of 90 times the vibrational frequency (f1 or f2), representing afferent excitation projected from the secondary somatosensory cortex. Following the standard experimental scheme, the comparison stimulus frequency f2 is given by f2 = f1 + 8 Hz or f2 = f1 − 8 Hz.

ON. The stimuli also excite a group of "ON" cells (see Fig. 1*c*), which have strong, saturating recurrent excitation. The ON cells fire during the task but are not tuned to f1. Such untuned, task-dependent cells are observed and, in our network, provide extra excitation to the C neurons during the task. The excitation is necessary for delay firing rates of C neurons to be greater than their spontaneous rate (see Fig. 2).

Memory. The memory network is based on a published model (1) with the simplification of containing only positively monotonic excitatory neurons. The memory (M) neurons are connected to form a discrete integrator (2). Twelve bistable subpopulations (each of 400 cells) have a range of excitabilities and so become active consecutively in response to above-threshold input from C neurons. Once active, strong recurrent excitation maintains the activity within a subpop-

ulation. Cross-excitation to other subpopulations is the strongest to the populations with closest excitability, facilitating the activation of the next subpopulation (as in ref. (1)). The number of active subpopulations after the input represents a memory that is a discrete approximation to the temporal integral of the input (2). We adjusted strengths of excitatory cross-connections so that, with a given number n of active populations, the input required to activate the next $(n+1\text{th})$ population (the threshold in Fig. 3*a*) is approximately constant (for all n). This input threshold for M neurons sets an activity threshold, Θ, for C neurons which provide the input. M neurons only increase their activity while the firing rate of C neurons $> \Theta$ (dashed line in Fig. 2*b*).

Readout. The 12 bistable subpopulations excite a subpopulation of 400 excitatory readout cells (Fig. 1*c*), whose firing rates encode the memory.

Inhibition A single interneuron population (of 200 cells) receives input from the readout population and inhibits the C neurons with a strength approximately proportional to the activity of readout cells.

Our purpose here is to investigate how a discrete integrator affects the circuitry of integral feedback control and not to address the mechanisms needed to create such an integrator, which could arise from single cell properties (3,4) and from strong recurrent feedback between cells (2).

Single Neuron Properties for Spiking Network Code. We simulate the individual cells as leaky integrate-and-fire neurons (5). The membrane potential V_i of cell i obeys the current balance equation

$$C_{\mathrm{M}}\frac{dV_i}{dt} = -g_{\mathrm{L}}\left[V_i - V_{\mathrm{L}}\right] - \left[g_{\mathrm{AMPA}} + g_{\mathrm{NMDA}}(V_i)\right] S_{\mathrm{E},i}\left[V_i - V_{\mathrm{E}}\right] - g_{\mathrm{I}} S_{\mathrm{I},i}\left[V_i - V_{\mathrm{I}}\right] - g_{\mathrm{ext}} s_{\mathrm{ext},i}\left[V_i - V_{\mathrm{E}}\right] - g_{\mathrm{cue}} s_{\mathrm{cue}}\left[V_i - V_{\mathrm{E}}\right] \quad [1]$$

where g_{L} is the leak conductance, V_{L} the leak potential; g_{AMPA} and $g_{\mathrm{NMDA}}(t)$ are the conductances of AMPA and NMDA channels, respectively, with excitatory reversal potential V_{E}; and g_{I} and V_{I} are the conductance and reversal potential for inhibitory channels. g_{ext} and g_{cue} are the fixed conductances for background noisy input and applied, stimulus-dependent input, respectively, and s_{ext} and s_{cue} are the corresponding time-dependent gating variables (see below). When the membrane potential reaches a threshold, V_{thr}, the neuron spikes, and the membrane potential is reset at V_{reset} for an absolute refractory period, τ_{ref}, before continuing to follow Eq. **1**.

The total synaptic drive for excitation or inhibition (S_{E} or S_{I}) is the sum of synaptic inputs from all presynaptic neurons j,

$$S_i = \sum_j W_{j\to i} s_j(t), \quad [2]$$

where $W_{j\to i}$ is the relative synaptic weight from cell j to cell i, and s_j is the synaptic current gating variable activated by the presynaptic neuron j firing spikes at times $t_{\mathrm{spike},j}$. Specifically,

for excitatory synapses, we have

$$\frac{ds_j}{dt} = \alpha_s \cdot \overline{P_{\mathrm{R}}}(t)[1 - s_j]\delta(t - t_{\mathrm{spike},j}) - \frac{s_j}{\tau_s} \quad [3]$$

and for inhibitory synapses

$$\frac{ds_j}{dt} = \delta(t - t_{\mathrm{spike},j}) - \frac{s_j}{\tau_s} \quad [4]$$

with synaptic time constants τ_s. The probability of vesicular release, $\overline{P_{\mathrm{R}}}(t)$, is described in *Short-Term Plasticity of Excitatory Synapses.*.

Background noisy input to all neurons is simulated using uncorrelated Poisson spike trains at a rate, r_{ext}, through nonsaturating synapses, of conductance g_{ext}, which are gated in accordance to

$$\frac{ds_{\mathrm{ext}}}{dt} = \delta(t - t_{\mathrm{spike,ext}}) - \frac{s_{\mathrm{ext}}}{\tau_{\mathrm{ext}}} \quad [5]$$

with synaptic time constant τ_{ext} following spikes at times $t_{\mathrm{spike,ext}}$.

Similarly, during the stimulus, Poisson spike trains of rate λ generate additional excitation through AMPAR-mediated synapses of conductance g_{cue} multiplied by a gating variable s_{cue}, which follows

$$\frac{ds_{\mathrm{cue}}}{dt} = \delta(t - t_{\mathrm{spike,cue}}) - \frac{s_{\mathrm{cue}}}{\tau_{\mathrm{ext}}}. \quad [6]$$

In the network models presented here, background, feed-forward and stimulus inputs are mediated by AMPA receptors, with $g_{\mathrm{AMPA}} = 36\mathrm{nS}$, $\tau_{\mathrm{ext}} = 2\mathrm{ms}$, and $V_{\mathrm{E}} = 0\mathrm{mV}$.

Recurrent excitation within a population (between the numbered memory subpopulations and within the group of ON cells) is mediated by a combination of NMDA receptors (6,7) with maximum conductance $g^{\max}_{\mathrm{NMDA}} = 36$ nS, $\tau_s = 100$ ms and $V_{\mathrm{E}} = 0$ mV, as well as AMPA receptors with $g_{\mathrm{AMPA}} = 18$ nS, $\tau_{\mathrm{ext}} = 2$ ms, and $V_{\mathrm{E}} = 0$ mV.

Feedback inhibition is mediated through GABA$_A$ receptors with $g_{\mathrm{I}} = 12$ nS, $\tau_s = 10$ ms and $V_{\mathrm{I}} = -70$ mV.

Cellular parameters are for excitatory cells: $C_{\mathrm{M}} = 0.5$ nF, $g_{\mathrm{L}} = 38.4$ nS, $V_{\mathrm{L}} = -70$ mV, $V_{\mathrm{reset}} = -60$ mV, $V_{\mathrm{thr}} = -45$ mV, $\tau_{\mathrm{ref}} = 2$ ms, $g_{\mathrm{ext}} = 6$ nS, $r_{\mathrm{ext}} = 1.2$ kHz; and for inhibitory cells: $C_{\mathrm{M}} = 0.2$ nF, $g_{\mathrm{L}} = 17.6$ nS, $V_{\mathrm{L}} = -70$ mV, $V_{\mathrm{reset}} = -60$ mV, $V_{\mathrm{thr}} = -50$ mV, $\tau_{\mathrm{ref}} = 1$ ms, $g_{\mathrm{ext}} = 1.6$ nS, $r_{\mathrm{ext}} = 1.8$ kHz.

Shadow Voltage for NMDA Current. The conductances of NMDA channels are voltage-dependent and follow (8)

$$g_{\mathrm{NMDA}}(t) = \frac{g^{\max}_{\mathrm{NMDA}}}{1.0 + \exp(-62\, V_{\mathrm{shadow}})/3.57} \quad [7]$$

where V_{shadow} is the shadow voltage (in mV). The shadow voltage is given by Eq. **1**, like the membrane potential, but is not reset after spikes. The shadow voltage is an approximation of

the more continuous dendritic membrane potential and (unlike V_i) has the realistic property that its mean value increases with firing rate.

Short-Term Plasticity of Excitatory Synapses. All excitatory synapses exhibit short-term presynaptic facilitation and depression (9,10). We implement the scheme described by Matveev and Wang (11), which assumes a docked pool of vesicles containing neurotransmitter, where each released vesicle is replaced with a time constant τ_d. The finite pool of vesicles leads to synaptic saturation, as when the presynaptic neuron fires more rapidly than vesicles are replaced, and no extra excitatory transmission is possible. Such synaptic depression contributes to stabilizing persistent activity at relatively low rates.

We make the simplification that there are many synapses between each pair of connected neurons, such that the average release probability per synapse, $\overline{P_{\rm R}(t)}$, simply scales the amplitude of synaptic transmission, as shown in Eq. **3**. We assume that more than one vesicle can be released per spike, hence the release amplitude at any individual synapse, $P_{\rm R}(t)$, is

$$P_{\rm R}(t) = p_v(t) \cdot < n(t) > /n_0, \qquad [8]$$

where $p_v(t)$ is the release probability for an individual vesicle and $n(t)$ is the number of docked vesicles (smaller than a maximum n_0). Similarly, we do not keep track of a discrete $n(t)$ for every individual synapse but assume that, over several synapses between two neurons, we can treat the average $< n(T) >$ as a continuous variable obeying

$$\frac{d < n >}{dt} = \frac{n_0 - < n >}{\tau_d} - \overline{P_{\rm R}(t)}\delta(t - t_{\rm spike}) \qquad [9]$$

decreasing by $\overline{P_{\rm R}}$ after a spike at time $t_{\rm spike}$. The value calculated for $P_{\rm R}(t)$ is used in Eq. **3** as the amplitude of excitatory synaptic transmission.

The vesicular-release probability is given by the product of three gating variables, $p_v(t) = O_1(t)O_2(t)O_3(t)$. A gating variable $O_i(t)$ $(i = 1, 2, 3)$ increases because of calcium influx triggered by an action potential, followed by a decay with time constant τ_f^i between spikes. Specifically, the following simple update rule is used: A gating variable $O_i(t)$ $(i = 1, 2, 3)$ follows

$$O_i(n+1) = 1 - \left\{1 - O_i(n) \exp\left[-(t_{n+1} - t_n)/\tau_f^i\right]\right\} C_f^i. \qquad [10]$$

Our simulations use the following values for the parameters in the memory network: $n_0 = 10$, $\tau_d = 250$ ms $C_f^1 = 0.1$, $\tau_f^1 = 50$ ms, $C_f^2 = 0.2$, $\tau_f^2 = 200$ ms, $C_f^3 = 0.4$, and $\tau_f^3 = 2$ sec.

Connectivity Details in the Spiking Network. Five modules containing different types of neurons are included in the full spiking network model. These are (see Fig. 1*c*) comparison (C) neurons, ON neurons, memory (M) neurons, readout (R) neurons, and inhibitory (I) neurons. Connections between modules are all to all. Neurons are identical within all modules, except for the memory module, which contains 12 subpopulations (groups), each containing a set of

identical neurons, but with neurons in different subpopulations having different thresholds. The connections between groups, labeled $i = 1$ to 12 in the memory network, are described below.

Connection strengths between neurons in the memory network depend only on their group numbers, and are all-to-all. All weights are normalized (*i.e.* divided) by the number of neurons in the presynaptic group, so that average network properties should be independent of the system size. The set of excitatory weights, $W_{\rm EE}$, follows:

$$W_{i\to j}^{\rm EE} = W_0^{\rm EE} \exp\left(\frac{-|i-j|}{N_{\rm grps}\sigma_i}\right) \qquad [11]$$

for $i \neq j$. The recurrent excitation within the same group is significantly stronger than between groups, so we define a separate set of parameters, $W_{i\to i}^{\rm EE} = W_i$.

The file connections_in.dat contains the set of values for connection strengths given below, and the file structure.cpp generates the weight matrix within the code. These files are available for download at
www.wanglab.brandeis.edu/people/miller/discriminator.html .

The full set of parameters are as follows for the discrete network: $W_0^{\rm EE} = 0.75$, $W_1 = 2.6$, $W_2 = 2.55$, $W_3 = 2.55$, $W_4 = 2.55$, $W_5 = 2.6$, $W_6 = 2.65$, $W_7 = 2.7$, $W_8 = 2.75$, $W_9 = 2.8$, $W_{10} = 2.85$, $W_{11} = 2.9$, $W_{12} = 3.0$; $\sigma_1 = 0.45$, $\sigma_2 = 0.45$, $\sigma_3 = 0.45$, $\sigma_4 = 0.55$, $\sigma_5 = 0.65$, $\sigma_6 = 0.75$, $\sigma_7 = 0.8$, $\sigma_8 = 0.8$, $\sigma_9 = 0.8$, $\sigma_{10} = 0.8$, $\sigma_{11} = 0.8$, $\sigma_{12} = 0.8$. Connections to the readout memory cell: $W_{1R} = 0.5$, $W_{2R} = 0.6$, $W_{3R} = 0.8$, $W_{4R} = 1.0$, $W_{5R} = 1.1$, $W_{6R} = 1.15$, $W_{7R} = 1.2$, $W_{8R} = 1.25$, $W_{9R} = 1.3$, $W_{10R} = 1.35$, $W_{11R} = 1.4$, $W_{12R} = 1.45$; and from the readout cell to inhibitory interneurons: $W^{\rm EI} = 2$. Feedforward excitation has $W_{\rm CM}^{\rm EE} = 0.33$ and feedback inhibition has $W_{\rm MC}^{\rm IE} = 1.7$.

The memory network includes an extra group of excitatory neurons which show cue-independent, but task-related activity. These ON neurons have self-excitation, $W_{\rm ON}^{\rm EE} = 1.4$ and excite the comparison cells, $W_{\rm ON,C}^{\rm EE} = 2.5$.

REFERENCES

1. Miller, P., Brody, C. D., Romo, R. & Wang, X.-J. (2003) *Cereb. Cortex* **13**, 1208–1218.

2. Koulakov, A. A., Raghavachari, S., Kepecs, A. & Lisman, J. E. (2002) *Nat. Neurosci.* **5**, 775–782.

3. Egorov, A. V., Hamam, B. N., Fransen, E., Hasselmo, M. E. & Alonso, A. A. (2002) *Nature* **420**, 173–178.

4. Goldman, M. S., Levine, J. H., Tank, G. M. D. W. & Seung, H. S. (2003) *Cereb. Cortex* **13**, 1185–1195.

5. Tuckwell, H. C. (1988) *Introduction to Theoretical Neurobiology.* (Cambridge University Press, Cambridge, U.K.).

6. Wang, X.-J. (1999) *J. Neurosci.* **19**, 9587–9603.

7. Wang, X.-J. (2001) *Trends in Neurosci.* **24**, 455–463.

8. Jahr, C. E. & Stevens, C. F. (1990) *J. Neurosci.* **10**, 3178–3182.

9. Varela, J. A., Sen, K., Gibson, J., Fost, J., Abbott, L. F. & Nelson, S. B. (1997) *J. Neurosci.* **17**, 7926–7940.

10. Hempel, C. M., Hartman, K. H., Wang, X. J., Turrigiano, G. G. & Nelson, S. B. (2000) *J. Neurophysiol.* **83**, 3031–3041.

11. Matveev, V. & Wang, X. J. (2000) *J. Neurosci.* **20**, 1575–1588.