Modeling thalamus as a non-rectifying predictive comparator

William Softky

May 9, 2003
(based on work commencing at Math Research Branch, NIH 1995)

855 14th Ave
Menlo Park, CA 94025 USA
bill@softky.com

Abstract

A thalamic relay cell may act as a comparator, sending the brain the difference between sensory input and a dynamic, anticipatory prediction of it. Such predictive inhibition would implement several useful perceptual principles: nonlinear reduced-redundancy coding in space and time; compensation for feedback's inevitable processing and propagation delays; and refining the brain's representation by emphasizing its mistakes.

A comparator must transmit both positive and negative signals, while thalamic cells, like most neurons, are thought to fire action potentials only in response to depolarizing synaptic inputs (i.e. they rectify, transmitting positive but not negative signals). But a thalamic relay cell might yet transmit both sighs of prediction-input mismatch--an **un**rectified error signal--by using its two distinct spiking modes: single (tonic) spikes and post-inhibitory-rebound (PIR) bursts. Reverse-correlations of random excitatory and inhibitory input to a simulated relay cell showed, over a wide range of firing rates, that on average single spikes signal excess excitation and reduced inhibition, while bursts signal approximately the opposite, thereby implementing an approximate comparator function.

Introduction

Because thalamus lies on the path from most sensory input to cortex, its computational function could be crucial. That function, while still unknown, may include attentional gating of input [1], the control of processing nonlinearity [2], regulation of temporal smoothing [3], or the modulation of specific sensory signals by local context [4, 5]... all possibilities which employ the known top-down feedback from cortex to thalamus. A complementary role for thalamus may lie in forming an information-efficient, reduced-redundancy recoding of sensory input [6, 7, 8, 3, 9], so that each thalamic cell signals independently of its neighbors in space and time (e.g. by local decorrelation).

These disparate tasks of feedback-gating and redundancy-reduction could merge if the net feedback influence on each relay cell were predictive inhibition, with the interneuron/reticular/cortical circuit (treated as a prediction-generating "black box") activating hyperpolarizing inhibition on each relay cell just before sensory input arrives. For example, predictive inhibition of static images would suppress any LGN activity extrapolated in space along local contours [9, 5]; predictive inhibition of dynamic stimuli [10] would suppress LGN activity extrapolated in time from local motion. In either case, predicted sensory input would be blocked, and thalamus would relay to cortex the remaining difference between predicted and actual input (a sustained, perfect prediction would be unstable, because it would block the subsequent input necessary to refresh and continue the prediction).

Because a signal is only predictable when it contains redundant correlations over time [11], a mechanism which suppressed predictable signals would also *de facto* implement a very general and nonlinear form of redundancy-reduction, which is one of the few well-established first

principles for efficient coding of sensory information [12, 6, 8]. Such a prediction/subtraction strategy is analogous to the comparator portion [9] of the widely used Kalman filter [13, 14], which updates its model of the environment based on the difference between its predictions and its new input (as weighted by its confidence in the new input; see Fig. 1).

Generating a prediction of dynamic sensory input would be a complex and daunting task, which might evolve through simple order-sensitive Hebbian synapses [15] and/or by complex network algorithms, which can learn to predict a dynamic input vector [10, 11, 16, 17, 18] or reconstruct a static one [19, 20, 21, 14]. Once a prediction is (somehow) generated, subtracting it from the subsequently arriving input would highlight unexpected inputs or bad predictions (i.e. "mistakes"), signals which could be used to improve the generative model and its subsequent predictions. For static inputs, this process corresponds to successive approximation, iteratively improving the match between the "prediction" and the (past) sensory input [19, 20, 21, 14]. But dynamic, anticipatory predictions would probably serve mammalian perception better than static "predictions" (i.e. reconstructions), for three reasons: 1) natural sensory inputs change with time, often at a time-scale comparable to or faster than the natural processing times of cortical cells (e.g. tens of milliseconds for synaptic latencies [22] and dendritic delays [23]), so that a successful prediction must be generated in advance to coincide with the input; 2) the future is intrinsically harder to predict than the past, and hence would better test a generative model (as, by analogy, scientific theories should predict experimental results in the future as well as those in the past); and 3) using predictive inhibition to suppress action potential firing in model cells works best if the inhibition arrives at least a few milliseconds before the predicted sensory excitation.

In visual thalamus (LGN), each relay cell typically receives input from only a single sensory (retinal) cell [24], so it is well-situated to compare sensory input with local (interneuron) and higher-order (cortically-driven reticular) inhibitory signals at the most specific scale possible. But to act like a true comparator a relay cell must be capable of conveying two distinct signals of opposite sign: (1) sensory input without prediction, and (2) prediction without sensory input. This may be possible through those cells' two characteristic modes of firing, which are generated by distinct cellular mechanisms: single, isolated (tonic) spikes, vs. brief bursts of spikes (which ride on the low-threshold calcium action potential sometimes following membrane hyperpolarization). I assume (with [2]) that such bursts exist in awake animals, and result from PIR mechanisms as they do in anesthetized animals [25]; this paper tests whether the two spiking modes might signal different kinds of synaptic input.

Simulations

I simulated an established [26], off-the-shelf single-compartment model thalamic relay cell containing ordinary and PIR-bursting Hodgkin-Huxley kinetics (with the original parameters), while adding fluctuating "synaptic" conductances. This simple model is meant to capture only the most basic properties, without including real thalamic cells' detailed anatomy (e.g. dendrites, glomeruli), modality, or receptive-fields (e.g. X,Y,W, lagged, etc.). The pseudorandom excitatory (G_e) and inhibitory (G_i) conductances drove the cell model to fire action potentials, which were sorted into single spikes and bursts by established criteria [25] (Fig. 3). Reverse

correlation histograms (RCHs) of G_e and G_i, as triggered both by bursts and by single spikes, showed (in addition to the sorting method's small artifactual kinks near —40 ms and ±4 ms) several distinct features. First, both single spikes and bursts showed sharp peaks of excitation (RCH area 0.82-1.5 excitatory pulses (EPs) above baseline) in the few *ms* before t = 0, indicating that one or more EPs was typically necessary to initiate a response. But preceding those peaks, the two different response modes were associated with very different synaptic histories. Excitation was typically raised or flat to the left of the single-spike peak, and inhibition weaker than baseline (because single spikes fire most easily with less inhibition). Bursts, on the other hand, which require hyperpolarization to de-inactivate low-threshold calcium conductances [25, 26], tended to follow periods of prolonged (50-150 *ms*) weaker excitation and stronger inhibition.

These trends were compared over wide ranges of excitation (30-fold) and inhibition (300-fold) by subtracting RCH(burst)-RCH(single) (excluding $|t|<5$ *ms*), and dividing the result by the maximum excursion (unsigned) of either one from its baseline. Values near ±1 were observed in the normalized histograms of both G_e and G_i, indicating that single spikes and bursts preferred very different "stimuli" (Fig. 5). Because both bursts and single spikes typically required a single EP for initiation, the average number n of EPs (above or below baseline) preceding a burst was also calculated. Where $n < 0$ (the case for most parameter combinations), the single EP triggering the burst was overbalanced by a larger deficit of EPs in the preceding tens of ms, so that a burst resulted from fewer-than-average EPs (Fig. 4). In summary, single spikes followed excess excitation and reduced inhibition, while bursts followed roughly the opposite pattern.

But bursts responded to synaptic input spread over a wider time window (5-150 *ms*) than did single spikes (1-30 *ms*). For LGN, these simulations would predict directly that a cell's linear spatiotemporal receptive field (RF) [27], as reconstructed from bursts alone, should be different from the RF reconstructed from single spikes alone. That burst RF should instead be similar to the convolution of the single-spike RF with a temporal dip-and-peak function like the burst RCHs of Fig. 5a (e.g. the burst RF should have a stimulus-to-output latency somewhat longer than the single-spike RF, and a richer temporal structure).

But beyond this straightforward analysis of RF properties lies a deeper hypothesis about sensory processing in general. The hypothesis of predictive inhibition arises from "first principles"-- the belief that efficient coding is central to sensory processing [6, 8, 28]—and is also broadly consistent with experiments. For example, the observed center-surround inhibition and transient temporal responses of LGN cells [29] and retinal cells [7] could result from inhibitory circuits which counteract natural images' simplest predictabilities, such as images' two-point brightness correlations in space and time [30. 3]; furthermore, the higher-order correlations of real images (like contours, discontinuities, and motion) may likewise contribute to the inhibitory "sharpening" effects of cortical feedback [5]. In addition to explaining receptive field properties, predictive inhibition could also account for cells' irregular firing, because cells driven with nearly equal and opposite excitation and inhibition are very sensitive to input fluctuations, and fire quite "noisily" [31, 28].

Possible experimental tests

The predictive comparator hypothesis for thalamic function does not follow from the single-cell simulations above; indeed, the simulations are just a preliminary test of the hypothesis. So an experimental search for predictive inhibition in thalamus would require extending visual stimuli from static displays to motion; from the traditional isolated and unnatural images (spots, bars, gratings) to whole-field, naturally predictable images; and from pairwise correlations to the higher-order correlations of shape-and motion-extrapolation [4, 32] that cortex might compute. In such a rich stimulus paradigm, a "predictive comparator" role for thalamus would predict the following experimental results: (1) that bursts should occur at least occasionally in natural, alert vision (because overprediction in a well-balanced system ought to occur along with underprediction); (2) that a RF-sized, briefly-moving contour which elicits excess single spikes from a cell ought to elicit fewer single spikes if it is preceded by and part of a full-field naturally moving image (i.e. if it is easily predictable); and (3) that the sudden disappearance of such a contour from a moving image should elicit more bursts from that cell (being now predicted but not present) than without the moving image.

Conclusion

In summary, this work presents three associated proposals: that predictive inhibition could implement reduced-redundancy coding; that the error signal from predicting the future is a plausible strategy for learning and improving primary sensory representations; and that thalamic relay cells might implement such predictive inhibition through their two distinct spiking modes. But othis work only presents biophysical simulations in support of the last of these ideas.

Even if thalamus does use bursts and spikes as a non-rectifying comparator signal, could cortex interpret them? Cortical circuitry is probably capable of discriminating single spikes from bursts through the combination of such fast synaptic mechanisms as facilitation and depression [33]. So it is tempting to speculate that cortex may have a specific function for bursts--both those from thalamic relay cells and from their cortical counterparts, the intrinsically-bursting layer-5 output cells--whatever that function proves to be.

Figures

Figure 1: *Left*: a schematic of a Kalman filter, which updates its internal model of the environment using the difference between that model and new data (as weighted by its confidence in the new data). *Right*: thalamus might implement an analogous comparator function. For input with high confidence (e.g. low noise), thalamus would relay an approximate difference between prediction and input, by using cellular mechanisms discussed below. This work does not not consider additional, possibly excitatory mechanisms involved with low-confidence input, nonlinearity-modulation, and attention, nor the "black box" (of cortical networks and learning rules) postulated to generate the dynamic, high-order predictions and to interpret the input-prediction mismatch.

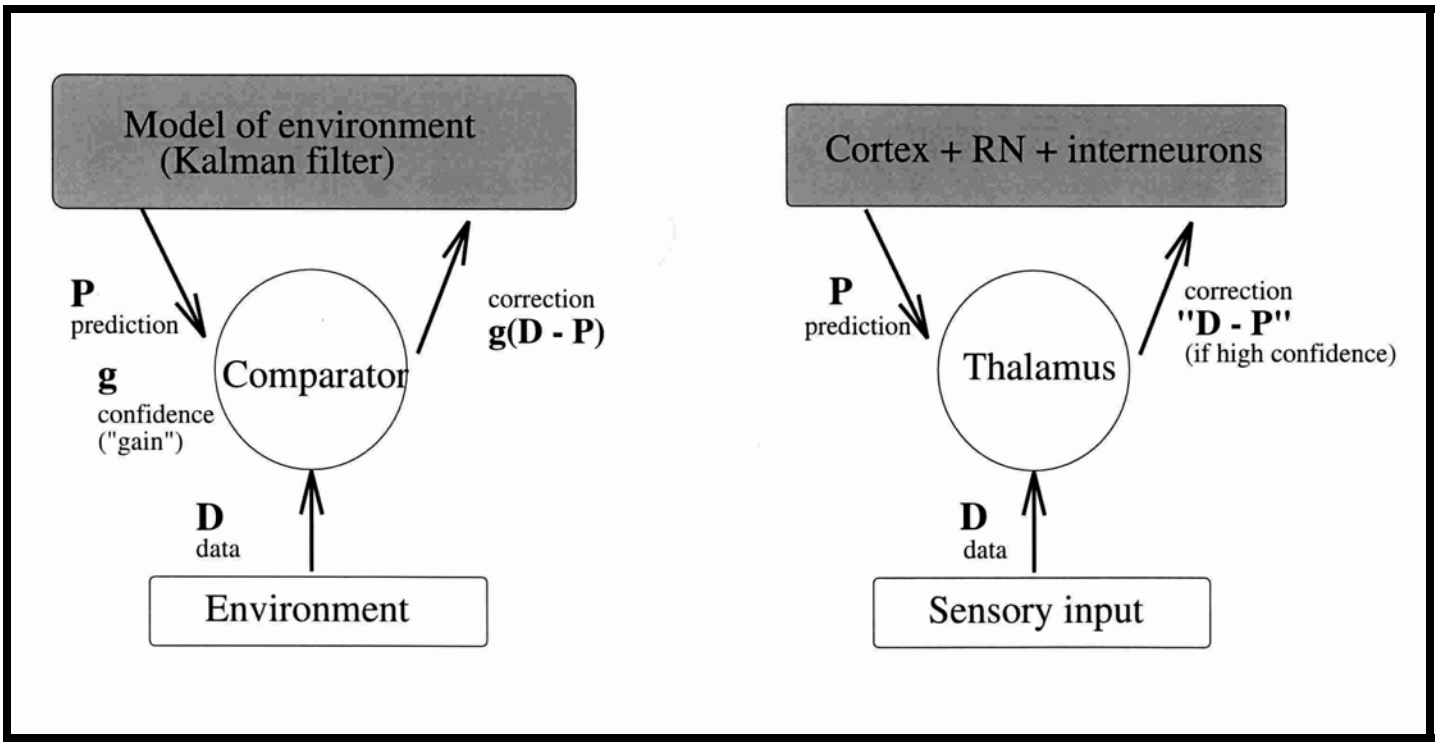


Figure 2: Thalamus might implement this comparator function by using predictive inhibition, which would block predicted sensory signals, allow the relay of unpredicted signals, and generate post-inhibitory-rebound (PIR) bursts when a prediction was unfulfilled by sensory input. Simulations showed that most PIR bursts were immediately preceded not only by prolonged hyperpolarization and missing excitation, but also by a single excitatory pulse (EP, dotted spike).

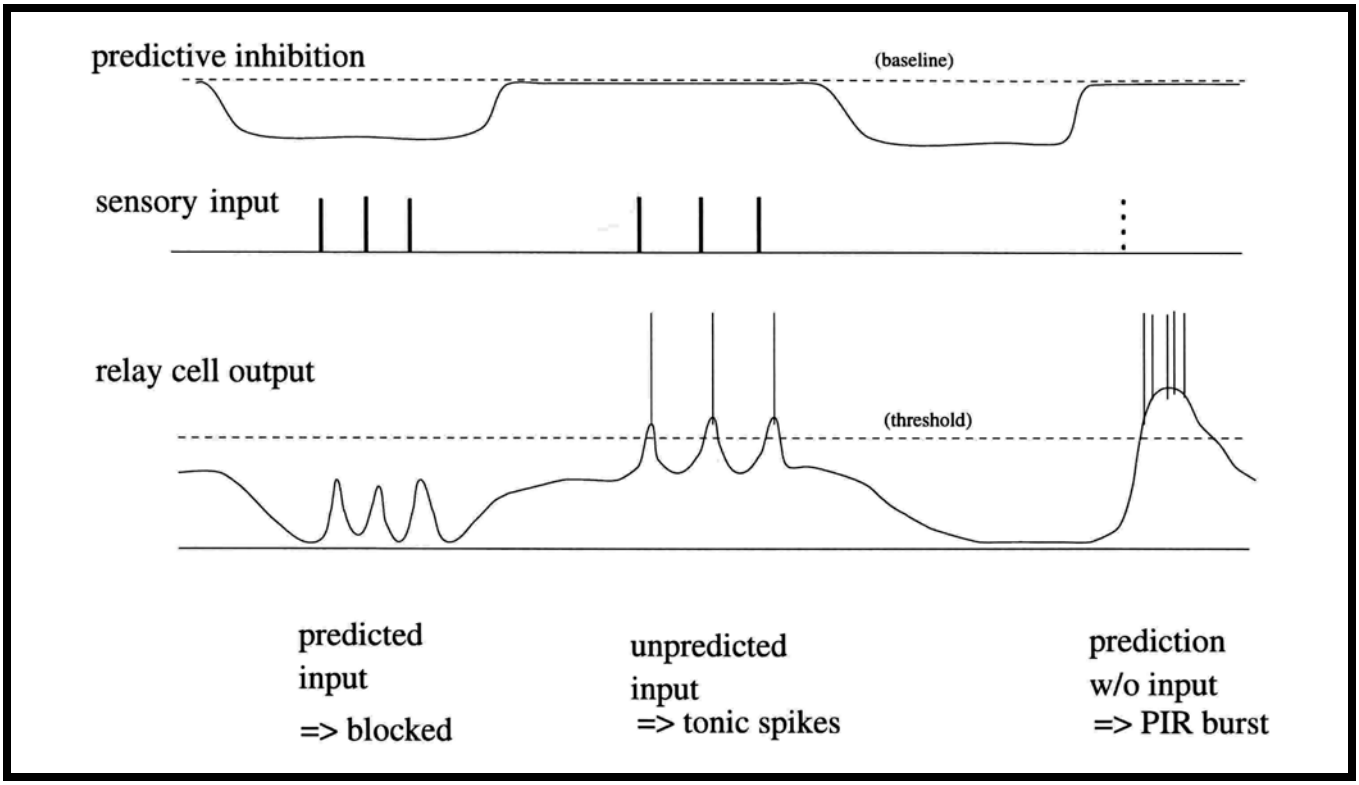
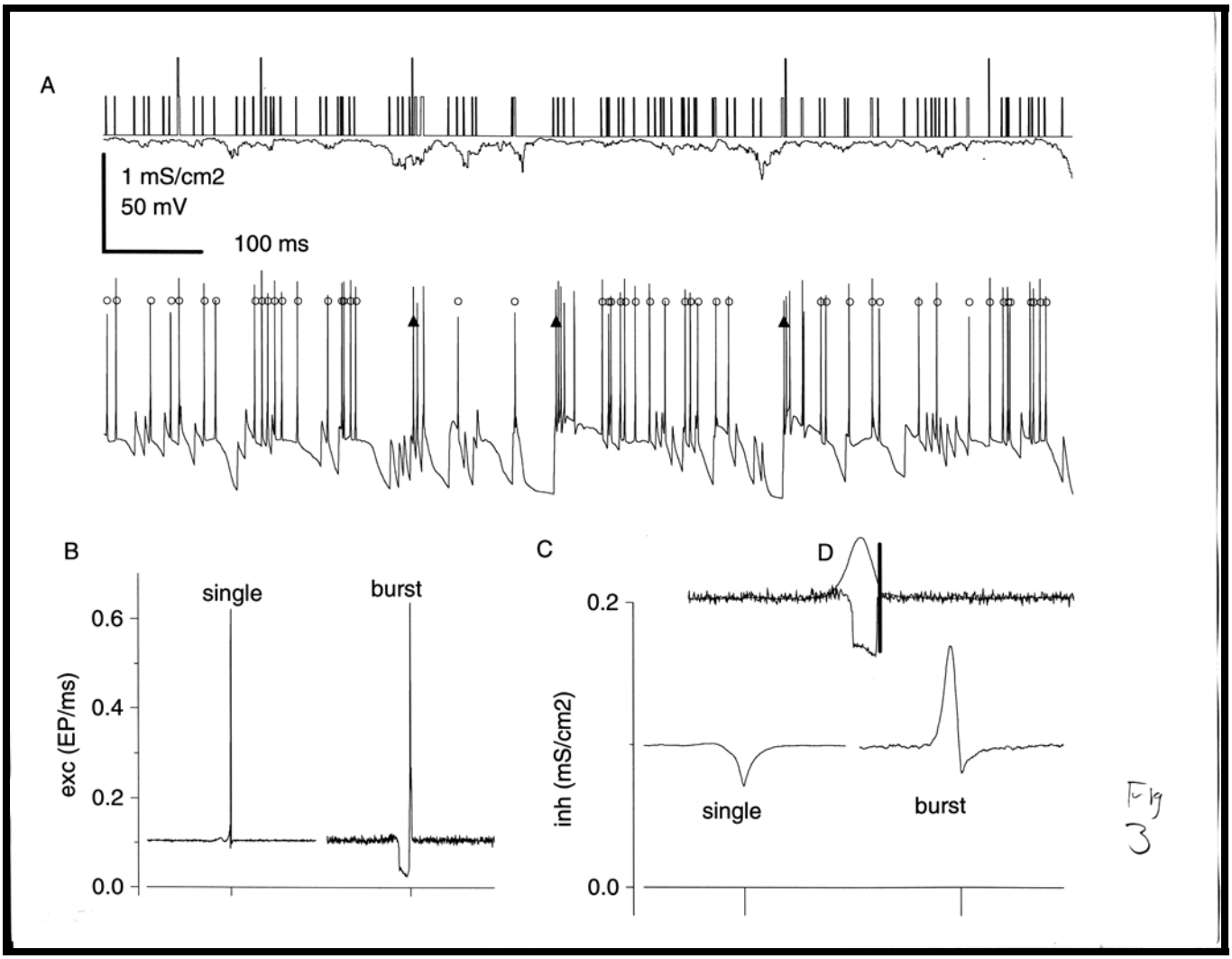


Figure 3: A simple, simulated relay cell with PIR kinetics was driven with random excitatory and inhibitory synaptic input; reverse-correlations of that input, as triggered by single (tonic) spikes or bursts, indicated that single spikes typically followed excess excitation and reduced inhibition, while bursts followed reduced excitation and excess inhibition.

METHODS: The simulated excitatory conductance $G_e(t)$ was composed of Poisson-distributed (rate *r*) square-wave pulses (EPs) of 0.4 (*mS/cm2*) amplitude, 1 *ms* duration, and reversal potential E_{exc} =0 *mV*; the inhibitory $G_i(t)$ was hyperpolarizing (E_{inh}=*-85 mV*), evolving with the stochastic differential equation $dG_i/(1\ ms)=0.18\sigma(t) -(G_i - \mathbf{G_i})/\tau$, in which $\sigma(t)$ was normalized gaussian noise, τ=30 *ms* gave the approximate autocorrelation timescale and $\mathbf{G_i}$ the mean amplitude (Gi(t) had approximately lognormal amplitude distribution). Simulation times ranged from 780 to 10,000 sec (10^8-10^9 10 μ*sec* timesteps). A "burst" was defined (following [25]) as any interspike interval (IS I) of 4 *ms* or less which followed an 1ST of 40 *ms* or more; only the timing of the first spike (of the 2-7 typical) in 30 *ms* was subsequently analyzed. All spikes not included in those 30 *ms* were considered single spikes (qualitatively similar results also held when bursts were defined as any ISI< 4 *ms* irrespective of preceding spikes).

(figure next page)



(Figure 3) **A)** One second of simulated random synaptic conductances (top: excitatory impulses of 0.4 *mS/cm2*, *r=0.1 EP/ms*; middle: inverted inhibitory G_i*= 0.1 mS/cm2*) caused irregular single spikes (open circles) and bursts (filled triangles). **B)** Reverse correlations (±300 *ms*) of excitatory impulses as triggered by single spikes (left) and bursts (right); tick marks show *t=0*. **C)** Reverse correlations of inhibition G_i. Deviations from baseline of these traces for *t>0* reflect the autocorrelation of the conductance, not acausality. **D (*inset*)** Reverse-correlation differences (burst - single, excluding $|t| < 5$ *ms* and normalizing by the maximum deviation of either from baseline; dark bar shows ±1 at *t=0*). Bursts are preceded by more inhibition (smooth trace) than are single spikes, and by fewer EPs (rough trace), so that-- if relay cells receive predictive inhibition of their sensory input-- a burst might signal "overprediction," and a single spike "underprediction."

Figure 4: Firing rates of the model thalamic relay cell for widely varying strengths of random excitation and inhibition. *(left)* Overall firing rate (length of bar), plotted by mean excitatory impulse rate *r* and inhibitory strength $\mathbf{G_i}$; scale bar is 20 *Hz*, and asterisk shows the simulation of Fig. 3. *(center)* Firing rates of single spikes (left bar) and bursts (right bar). These data span the regime of physiologically relevant single-spike and burst firing rates. Single-spike firing rates increase with excitation and decrease with inhibition, while burst rates do roughly the opposite. *(right)* The average number *n* of excess EPs preceding a burst--typically, a time-averaged deficit of EPs added to the triggering EP--shown as bars above or below each cross. Where single firing strongly exceeded burst firing (the most plausible regimes, to the right) bursts are preceded by a net deficit of excitatory input; in contrast, single spikes are typically preceded by 1-1.5 extra EPs.

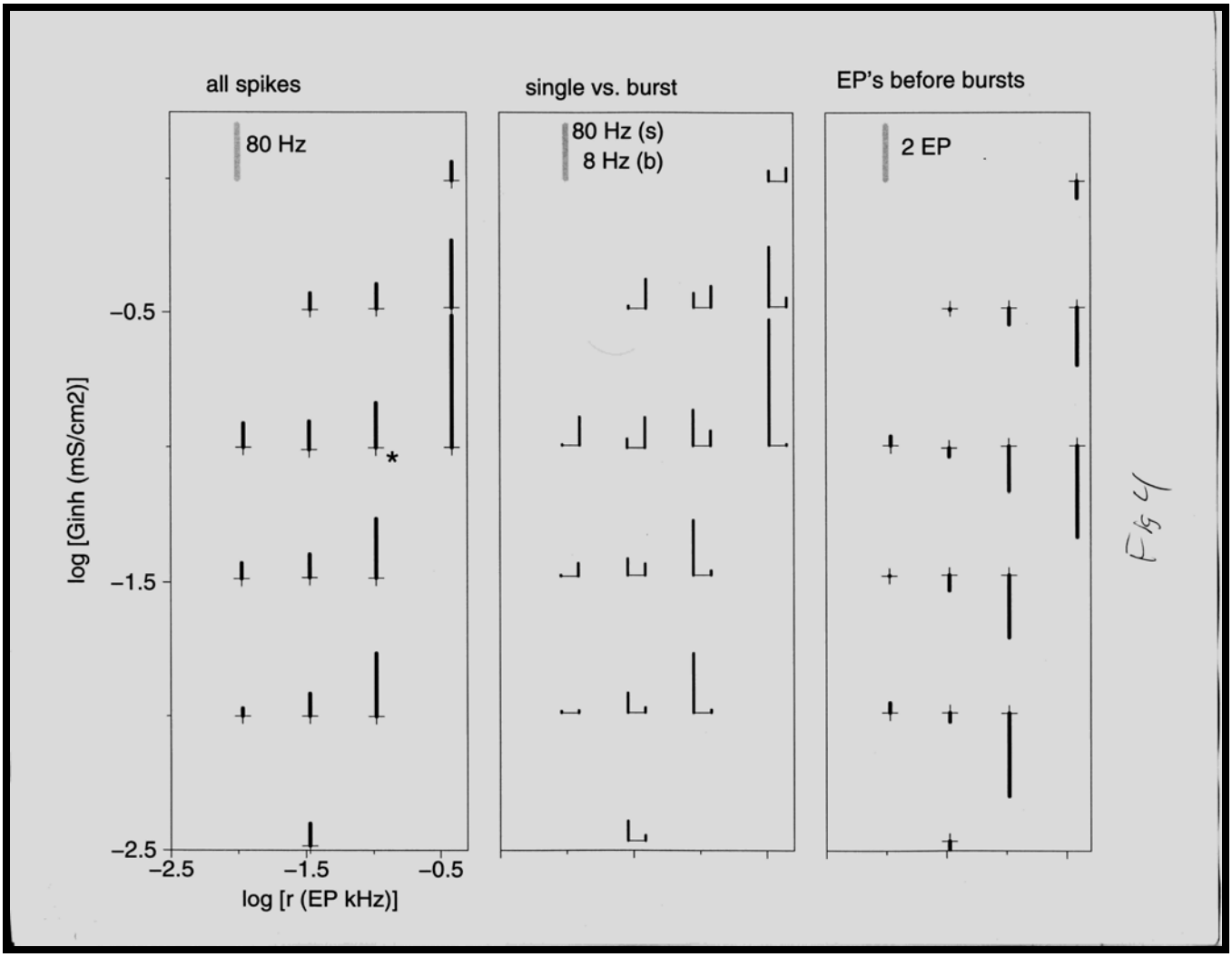
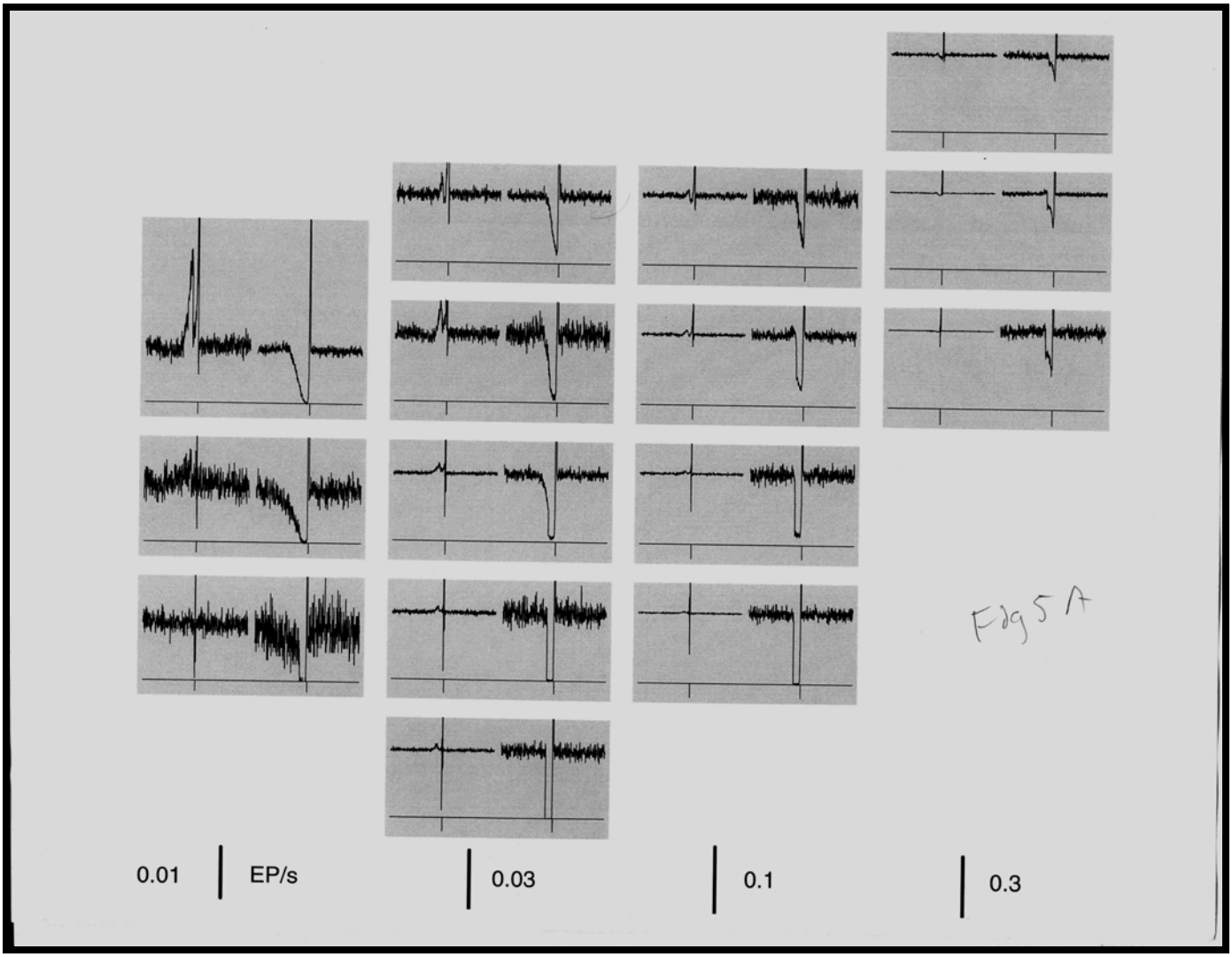
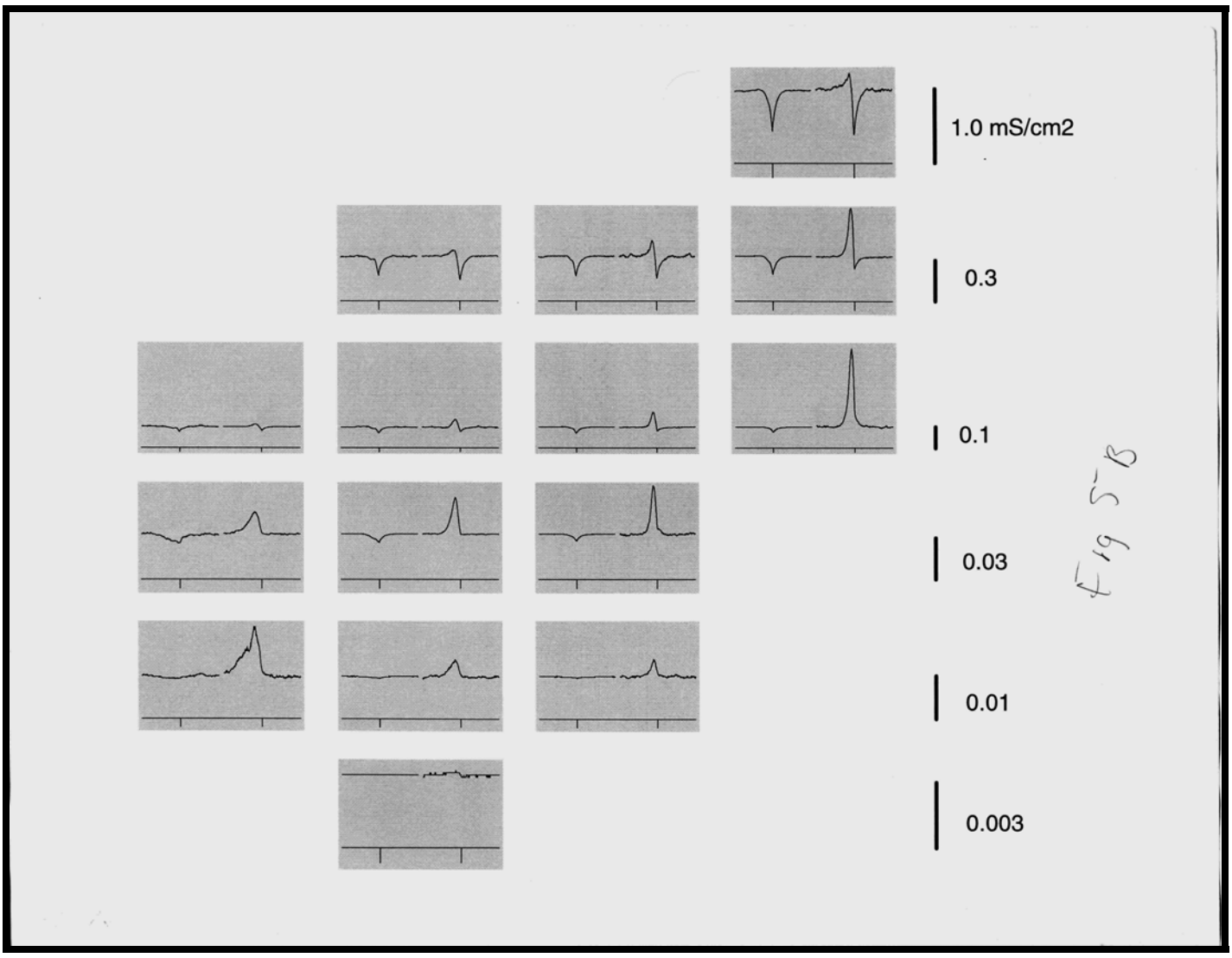


Figure 5: Reverse-correlation histograms (single/burst pairs as in Fig. 3 b-d) for various strengths of random excitation and inhibition (arrayed as in Fig. 4)..

A) Excitation histograms (truncated; scale bars below) show that single spikes are preceded primarily by narrow peaks in excitation, while bursts are preceded by peaks (0.82-1.5 EPs) and even stronger dips





(Figure 5 B) Inhibition histograms (scale bars at right) show that single spikes are usually preceded by a dip in inhibition, while bursts are preceded by excess inhibition .

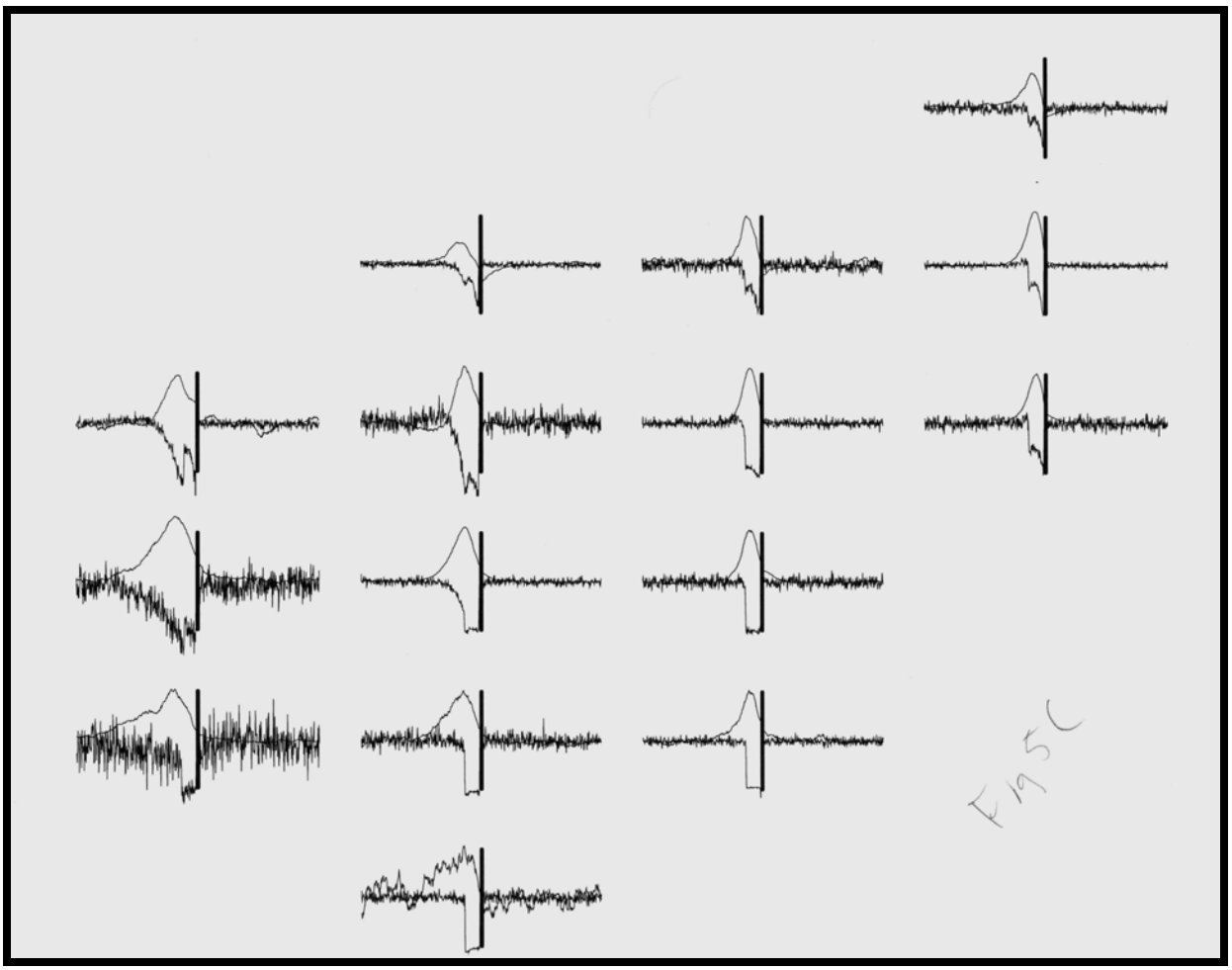


Figure 5C) The normalized difference (burst-single) histograms show the consistent form and relative amplitude of these patterns over a wide parameter range, indicating that single spikes and bursts result from approximately opposite patterns of synaptic activity, and thus might implement the two signs of a comparator function.

Acknowledgements

This work has benefltted from numerous discussions with J. Rinzel, along with useful comments by B. Pearlmutter, D. Ruderman, J. Atick, A. Pece, and S. M. Sherman.

References

[1] F. Crick. Function of the thalamic reticular complex: the searchlight hypothesis. Proc. Natl. Acad. Sci. USA, 81:4586-4590, 1984.

[2] S. M. Sherman. Dual response modes in lateral geniculate neurons: mechanisms and functions. Vis. Neurosci., in press, 1996.

[3] D. Doug and J. Atick. Temporal decorrelation - a theory of lagged and nonlagged responses in the lateral geniculate-nucleus. Network-coin, 6(2): 159—178, 1995.

[4] K. L. Grieve and A. M. Sillito. Differential properties of cells in the feline primary visual cortex providing the cortifugal feedback to the lateral geniculate nucleus and visual claustrum. J. Neurosci., 15:4868—4874, 1995.

[5] J. Cudeiro and A. M. Sillito. Spatial fequency tuning of orientation-discontinuity-sensitive corticofugal feedback to the cat lateral geniculate nucleus. J. Physiol., 490.2:481—492, 1996.

[6] H. B. Barlow. Possible principles underlying the transformations of sensory messages. In W. Rosenblith, editor, Sensory Communication, chapter 13, pages 217—234. MIT Press, Cambridge, 1961.

[7] M. V. Srinivasan, S. B. Laughlin, and A. Dubs. Predictive coding: a fresh view of inhibition in the retina. Proc. R. Soc. Lond. B, 216:427—459, 1982.

[8] J. Atick and N. Redlich. Towards a theory of early visual processing. Neural Computation, 2:308—320, 1990.

[9] A. Pece. Reduced redundancy of a gabor representation: a possible computational role for feedback from primary visual cortex to lateral geniculate nucleus. In I. Aleksander and J. G. Taylor, editors, Artificial Neural Nets 2, Proceedings of ICAAN-2, pages 865—868. Elsevier Science Publishers, 1992.

[10] W. Softky. Unsupervised pixel-prediction. In M. Hasselmo, editor, Advances in Neural and Information Processing Systems 8, pages 809—815. MIT Press, 1996.

[11] Learning complex, extended sequences using the principle of history compression. Jürgen Schmidthuber. Neural Computation, 4:234-242,1992.

[12] D. M. MacKay. Towards an information-flow model of human behavior. Brit. J. Psychol., 47:30-43, 1956.

[13] P. S. Maybeck. Stochastic models, estimation, and control, volume 1. Academic Press, Inc., New York, 1979.

[14] R. Rao and D. Ballard. Dynamic model of visual memory predicts neural response properties in the visual cortex. Technical Report (C.S./U. Rochester), 95.4:1-9, 1995.

[15] P. Read Montague, Peter Dayan, and T. Sejnowski. A framework for mesencephalic dopamine systems based on predictive hebbian learning. J. Neuroscz., 16:1936—1947, 1996.

[16] L. Rabiner. A tutorial on hidden markov models and selected applications in speech recognition. Proc. IEEE, 77(2) :257—286,1989.

[17] K. Fielding and D. Ruck. Recognition of moving light displays using hidden markov models. Pattern Recognition, 28:1415—1421, 1995.

[18] In A. Weigand and N. Gerschenfeld, editors, Time Series Prediction. Addison-Wesley, 1994.

[19] D. Mumford. Neuronal architectures for pattern-theoretic problems. In C. Koch and J. Davis, editors, Large-scale theories of the cortex, pages 125—152. MIT Press, 1994.

[20] S. Ullman. Sequence-seeking and counterstreams: a model for bidirectional information flow in cortex. In C. Koch and J. Davis, editors, Large-scale theories of the cortex, pages 257-270. MIT Press, 1994.

[21] G. Hinton, P. Dayan, B. Frey, and R. Neal. The wake-sleep algorithm for unsupervised neural networks. Science, 268:1158-1161, 1995.

[22] J. McClurkin, L. Optican, B. Richmond, and T. Gawne. Concurrent processing and complexity of temporally encoded neuronal messages in visual perception. Science, 253:675-677, 1991.

[23] H. Agmon-Snir and I. Segev. Signal delay and input synchronization in passive dendritic structures. J. Neurophysiol., 70:2066-2085, 1993.

[24] S. M. Sherman and C. Koch. Thalamus. In Gordon Shepherd, editor, The Synaptic organization of the brain, chapter 8, pages 246-278. Oxford University Press, 1990.

[25] S-M Lu, W. Guido, and S. M. Sherman. Effects of membrane voltage on receptive field properties of lateral geniculate neurons in the cat: contributions of the low-threshold CA2~ conductance. J. Neurophysiol., 68:2185-2198, 1992.

[26] M. Rush and J. Rinzel. Analysis of bursting in a thalamic neuron model. Biol. Cybern., 71:281-291, 1994.

[27] D. Golomb, D. Kleinfeld, R. Reid, R. Shapley, and B. Shraiman. On temporal codes and the spatiotemporal response of neurons in the lateral geniculate nucleus.
J. Neurophysiol., 72:2990-3003, 1994.

[28] W. Softky. Simple codes vs. efficient codes. Current Opinion in Neurbiology, 5:239-247, 1995.

[29] C. Reid and R. Shapley. Saptial structure of cone inputs to receptive fields in primate lateral geniculate nucleus. Nature, 356:715-717, 1992.

[30] D. Doug and J. Atick. Statistics of natural time-varying images. Network: Computation in Neural Systems, 6(3):345-358, 1995.

[31] M. Shadlen and W. Newsome. Noise, neural codes and cortical organization. Current Opinion in Neurobiology, 4:569-579, 1994.

[32] R. Nilhawan. Motion extrapolation in catching. Nature, 370:256-257, 1994.

[33] J. Deuchars and A. Thomson. Temporal and spatial properties of local circuits in cortex. TINS, 17:119-126, 1994.