

in *Computational Neuroscience: Trends in Research, 1998*
edited by J. M. Bower (Plenum, New York, 1998), pp. 491–496.

INFORMATION EXCHANGE BETWEEN PAIRS OF SPIKE TRAINS IN THE MAMMALIAN VISUAL SYSTEM

Steven B. Lowen,¹ Tsuyoshi Ozaki,² Ehud Kaplan,³ Malvin C. Teich⁴

¹ Department of Electrical & Computer Engineering
Boston University
8 Saint Mary's St., Boston, MA 02215
Email: lowen@bu.edu

² The Rockefeller University
1230 York Ave., New York, NY, 10021
Email: yoshi@camelot.mssm.edu

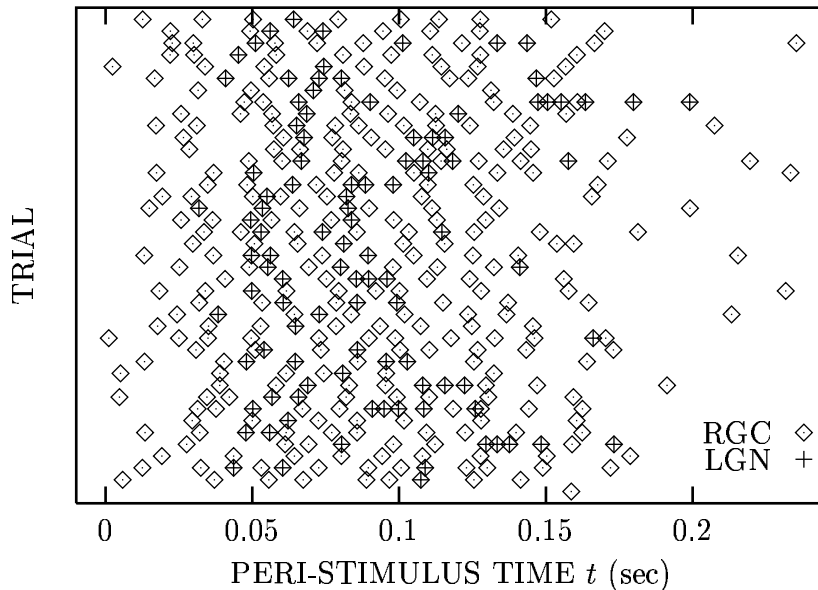
³ Department of Ophthalmology
Mt. Sinai School of Medicine
One Gustave Levy Pl., New York, NY, 10029
Email: kaplane@rockvax.rockefeller.edu

⁴ Departments of Electrical & Computer Engineering
and Biomedical Engineering
Boston University
8 Saint Mary's St., Boston, MA 02215
Email: teich@bu.edu

ABSTRACT

We have studied the neural firing patterns of retinal ganglion cells (RGCs) and their target lateral geniculate nucleus (LGN) cells. Reliable information transmission coexists with fractal fluctuations which appear in RGC and LGN firing patterns. Unexpectedly, these fluctuations appear not to be independent across LGN cells; information is also shared among pairs of LGN spike trains. Over short time scales, we find that clusters of spikes in the RGC neural firing pattern appear at the LGN output essentially unchanged, while isolated RGC firing events are more likely to be eliminated; thus the LGN action-potential sequence is not simply a *randomly* deleted version of the RGC spike train. Employing information-theoretic techniques, we estimate the information efficiency of the LGN neuronal output — the proportion of the variation in the LGN firing pattern that carries information about its associated RGC input — to be in the vicinity of 50% over counting windows below about 10 ms. We develop a new information-theoretic measure which helps determine at what time scale a neural spike train changes from a time code to a rate code.

FIGURE 1: RGC AND LGN JOINT RASTER PLOT



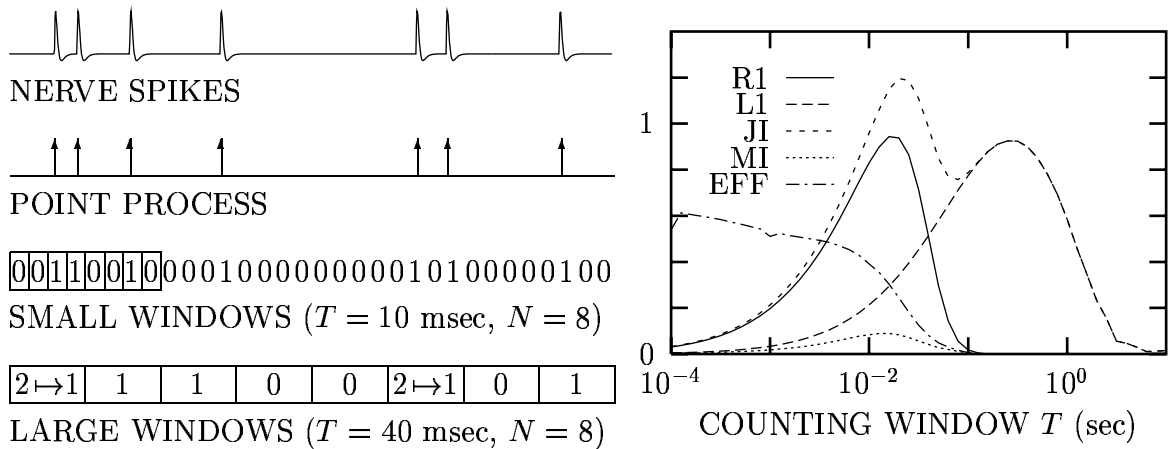
INTRODUCTION

The relationship between spike trains at the RGC and its target LGN cells have yielded to increased understanding over the course of many years' research (see the excellent review article by Funke and Wörgötter, 1997). One observation that repeatedly emerges is that conventional measures of noisiness, such as interspike-interval coefficient of variation and signal-to-noise ratio, generally increase at the LGN relative to the RGC (see, e.g., Tables 1 and 2 in Teich et al., 1997). This observation indicates that other important characteristics of the transformation should be sought and investigated; indeed many researchers, typically using a linear-systems approach, have carried out studies along these lines (Kaplan and Shapley, 1984; Levine and Troy, 1986; Funke and Wörgötter, 1997, to cite only a few). However, a thorough exploration of the inhibition and shunting mechanisms used by the LGN in its role as gatekeeper requires an analysis in which the joint statistical features of RGC and LGN spike trains are examined.

SELECTIVE PASSAGE OF RGC SPIKES BY THE LGN

The LGN spike train derives from the RGC spike train by a deletion mechanism. Aside from occasional clusters of calcium spikes, every LGN spike occurs simultaneously with an RGC spike at the input to the LGN. How are these spikes selected and why? The joint raster plot in Fig. 1 shows 40 stimulus-cycle trials for an ON/X cat RGC/LGN pair driven by a 4.2 Hz drifting grating at 100% contrast with a mean luminance level of 50 cd/m². Spikes from the RGC are denoted by diamonds, while crosses signify the associated LGN action potentials. Rather than forming a randomly deleted version of the RGC spikes, in which each RGC spike undergoes a Bernoulli trial (coin flip) to determine whether it survives as a spike at the output of the LGN, the LGN action potentials appear to occur either at the end of or during clusters of RGC action potentials. Figure 1 reveals several LGN spike triplets closely spaced in time. Other LGN firings tend to occur after a sequence of closely spaced RGC spikes. The result is that the LGN neuron tends to fire when the RGC is most active, near the peak of the

**FIGURE 2: STANDARD INFORMATION-THEORETIC MEASURES
FIXED NUMBER OF WINDOWS N**



RGC peri-stimulus time (PST) histogram, and when the signal is strongest. This can actually sharpen timing reliability in the LGN, by some measures, relative to that of the RGC (Funke and Wörgötter, 1997).

INFORMATION EXCHANGE BETWEEN RGC/LGN SPIKE-TRAIN PAIRS: STANDARD APPROACH

The signal-processing benefit that an LGN cell might afford an input spike train can be evaluated by examining several measures of the input and output spike trains. We consider the mutual information which provides an indication about the nature of the code used by the LGN and why clusters might be favored. We subsequently consider information exchange across pairs of LGN spike trains.

We first use the standard information-theoretic approach for determining the information content of a neural spike train as a measure of the relative reliability of these signals (Meister, 1996; Rieke *et al.*, 1997; Saleh and Teich, 1987). This method treats segments of spike trains as neural codes, and indicates how information is transmitted in a quantifiable manner without making any assumptions about the nature of these codes. A spike train, for example RGC activity transmitted to an LGN cell by S potentials (Kaplan and Shapley, 1984), is first divided into a sequence of counting windows of the same duration T . As shown in Fig. 2, the code is generated by writing a 1 if at least one point-process event (idealized nerve spike) occurs in a window, and a 0 if not, and this process is repeated for all counting windows. The binary stream thus produced is grouped into overlapping segments of $N = 8$ windows each. Each such segment forms one particular codeword out of a total of $M = 2^N$ possibilities. A histogram is then formed from the codewords, as an estimate of their probabilities of occurrence. The entropies are then calculated (Rieke *et al.*, 1997) for this RGC spike train, as well as for the corresponding LGN sequence of action potentials, and also for the two spike trains jointly (JI), from which the mutual information (MI) may be obtained.

An efficiency (EFF) can be defined (Rieke *et al.*, 1997) as the ratio of the MI to the LGN spike-train entropy; this serves to quantify the proportion of information available in the LGN spike train that is shared with the RGC input. A higher efficiency does not necessarily mean improved performance, however, since an ideal axon that transmits

action potentials perfectly, but without any processing, has a theoretical efficiency of unity. The LGN apparently selects part of the signal transmitted from the RGC, discarding some entropy which presumably does not correspond to useful information for the task at hand. In the process, the relative spike-time reliability can increase, since the LGN preferentially transmits clusters of spikes. The procedure described above is carried out for a variety of counting times T . To examine the implementation of this approach, Fig. 2 presents curves of the three entropies and the mutual information as functions of the counting window, for an RGC/LGN spike train pair recorded under maintained-discharge conditions (ON/X cat RGC/LGN pair driven by a blank screen of luminance 50 cd/m^2) and with a data segment length of $N = 8$ windows per codeword. This choice of N yields the most details about RGC/LGN spike train behavior given the data size available.

For larger values of N the number of possible joint RGC/LGN codewords 2^{2N} becomes comparable in magnitude with the number of such codes available in the data, and spuriously large values of the efficiency result. The values are normalized by dividing by N , to give entropies and mutual information in bits per window. As seen in Fig. 2, both the RGC and LGN curves approach the theoretical maximum of 1 bit per window, but at different times; the times correspond to approximately half the average interevent intervals of each spike train as expected. For large window durations, all windows contain a spike, and thus the information is identically zero; this limitation is fundamental to the counting method used. The joint entropy lies below the theoretical maximum of two bits per window (one for each spike train) and above the entropies of both the RGC and LGN spike trains, as it must. The mutual information remains below 0.2 bits per window, but the efficiency exceeds one-half for counting windows less than 10 msec, indicating that most of the variation in the LGN spike train represents information transmitted from the RGC at these window sizes.

MUTUAL INFORMATION (BITS/WINDOW)

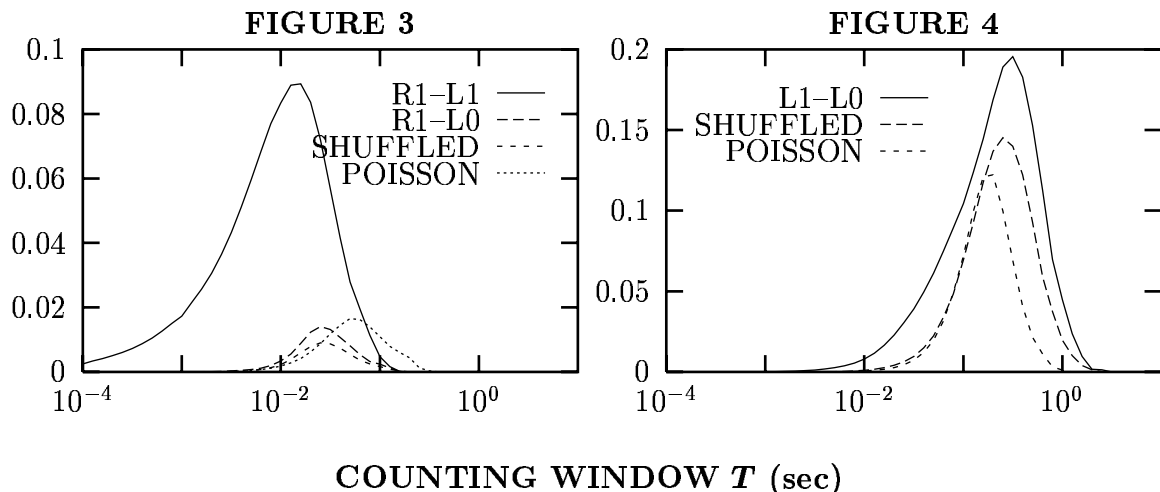


Figure 3 shows the mutual information as a function of the counting time, with a data segment length of $N = 8$, between 1) an RGC/LGN spike train pair (R1-L1) recorded under maintained discharge conditions (the same MI curve as shown in Fig. 2), 2) the same RGC spike train and an unrelated LGN spike train recorded at the same time (R1-L0), 3) a shuffled version of the spike trains used in 1) (same interspike intervals, but in a random order), and 4) homogeneous Poisson processes with the same mean rate as those in 1). The mutual information for the associated RGC/LGN spike-train pair significantly exceeds that of the two surrogates and of the same RGC

with an alternate LGN cell for counting times less than about 40 msec, so that the RGC/LGN spike trains likely exhibit reliable timing precision over these time scales. Indeed, the similar shapes of the shuffled surrogate and the unrelated RGC/LGN spike trains shows that this surrogate successfully captures the behavior of these spike trains when studied with this measure, while removing the effects of precision timing. For counting times greater than 80 msec, however, the Poisson surrogate exhibits a larger apparent mutual information than the original data, indicating that timing precision and reliability reported by this estimation procedure at this time scale likely arises from limitations in the procedure itself.

INFORMATION EXCHANGE ACROSS PAIRS OF LGN SPIKE TRAINS

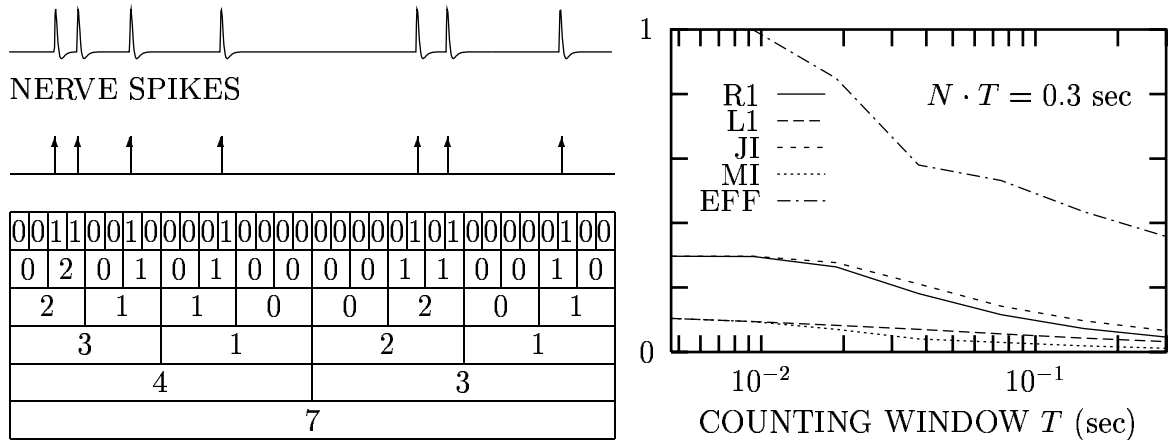
As shown in Fig. 3, significant mutual information exists between RGC/LGN pairs as expected, but not in unrelated RGC/LGN cells nor between two RGCs (not shown), for which the mutual information is small, roughly equal to that of the surrogates. However, evidence exists for mutual information between two simultaneously recorded LGN cells, although their inputs, the associated RGCs, do not appear to be related. Figure 4 shows the mutual information as a function of the counting time, with a data segment length of $N = 8$, between two LGN cells recorded under maintained discharge conditions, and for two surrogate data sets. The two LGN cells show a small but significant increase in the mutual information over that of the surrogates, especially for counting windows less than 30 msec. Evidently some process is acting jointly on the two LGN cells which introduces dependencies and thus joint reliability between the two spike trains, since no such dependency exists in the RGC spike trains to any significant degree. For driven cells, in contrast, the mutual information calculated between any combination of spike trains resembles that of the solid curve in Fig. 3. In that case, results for RGC/LGN pairs resemble those for RGC/RGC, LGN/LGN, and unrelated RGC/LGN pairs; all exhibit mutual information values which exceed those of the surrogates by significant amounts (not shown). We see that such relationships do not exist under maintained discharge conditions, so that it will be useful to investigate whether the close relation among all four of the spike trains under driven conditions arises from common fluctuations imposed by the stimulus or its presence, and how different stimulus conditions will affect these relationships.

INFORMATION EXCHANGE BETWEEN RGC/LGN SPIKE-TRAIN PAIRS: NEW APPROACH

We have developed a new information-theoretic approach, shown in Fig. 5, in which the overall length of the data segment remains fixed ($N \cdot T = 0.3$ sec in this example), while the number of windows within the segment changes. Figure 5 presents the same quantities shown in Fig. 2. In the limit of a small window the relative precision attains a high value, and the spike train is analyzed as a precise *time code*. At the other limit the entire segment is a single window, and the spike train is analyzed as a *rate code*. This method has the advantage of keeping the data segment length fixed, rendering the effects of including varying amounts of data less important. The knee in the efficiency curve at 37.5 msec suggests a crossover time near which the time code can be said to change to a rate code. The neural system itself can, of course, function simultaneously at all scales, thereby transmitting time-code and rate-code information simultaneously.

Larger windows have higher counting resolution (rather than registering any positive number of spikes with a 1 symbol), extending the range of usefulness to longer

**FIGURE 5: NEW INFORMATION-THEORETIC MEASURES
FIXED DATA-SEGMENT DURATION $N \cdot T$**



window sizes. In particular, the RGC counts need not all assume a value of unity, and thus yield no useful information, when the LGN has appreciable information, as in Fig. 2 near 0.2 seconds.

Further, large numbers of windows are associated with small window sizes, alleviating the finite-data problems mentioned above. Indeed, we were able to extend the number of windows N to 64 before encountering the limitations of our data size. For this large value of N and the data sets we have used, each RGC code becomes unique; however, the LGN, having a much lower rate, does not generate unique codes. Since the LGN does not add any information to the RGC/LGN code pairs already present in the unique RGC codes, the joint information is equal to that in the RGC, and the MI assumes a value equal to that of the LGN. Thus in this case, the efficiency becomes 100%. This illustrates another limitation of the use of efficiency calculations, which require care in their interpretation; the spuriously high efficiency values shown here result from data limitations, and not from neural processing at the LGN.

REFERENCES

- Funke, K., and Wörgötter, F., 1997, On the significance of temporally structured activity in the dorsal lateral geniculate nucleus (LGN), *Prog. in Neurobiol.* 53:67.
- Kaplan, E., and Shapley, R. M., 1984, The origin of the S (slow) potential in the mammalian lateral geniculate nucleus, *Exp. Brain Res.* 55:111.
- Levine, M. M., and Troy, J. B., 1986, The variability of the maintained discharge of cat dorsal lateral geniculate cells, *J. Physiol. (London)* 330:339.
- Meister, M., 1996, Multineuronal codes in retinal signaling, *Proc. Natl. Acad. Sci. USA* 93:609.
- Rieke, F., Warland, D., de Ruyter van Steveninck, R., and Bialek, W., 1997, *Spikes: Exploring the Neural Code*, MIT Press, Cambridge, MA.
- Saleh, B. E. A., and Teich, M. C., 1987, Can the channel capacity of a lightwave communication system be increased by the use of photon-number-squeezed light?, *Phys. Rev. Lett.* 58:2656.
- Teich, M. C., Heneghan, C., Lowen, S. B., Ozaki, T., and Kaplan, E., 1997, Fractal character of the neural spike train in the visual system of the cat, *J. Opt. Soc. Am. A* 14:529.