Efficiency of Information Transmission by Retinal Ganglion Cells

Kristin Koch,[1] Judith McLean,[1] Michael Berry,[2]
Peter Sterling,[1] Vijay Balasubramanian,[3]
Michael A. Freed[1],*
[1]University of Pennsylvania
Department of Neuroscience
Philadelphia, Pennsylvania 19104
[2]Princeton University
Department of Molecular Biology
Princeton, New Jersey 08544
[3]University of Pennsylvania
Department of Physics
Philadelphia, Pennsylvania 19104

Summary

Background: Different types of retinal ganglion cells convey different messages to the brain. Messages are in the form of spike patterns, and the number of possible patterns per second sets the coding capacity. We asked if different ganglion cell types make equally efficient use of their coding capacity or whether efficiency depends on the message conveyed.
Results: We recorded spike trains from retinal ganglion cells in an in vitro preparation of the guinea pig retina. By calculating, for the observed spike rate, the number of possible spike patterns per second, we calculated coding capacity, and by counting the actual number of patterns, we estimated information rate. Cells with "brisk" responses, i.e., high firing rates, and a general message transmitted information at high rates (21 $\pm$ 9 bits s^{-1}). Cells with "sluggish" responses, i.e., lower firing rates, and specific messages (direction of motion, local-edge) transmitted information at lower rates (13 $\pm$ 7 bits s^{-1}). Yet, for every type of ganglion cell examined, the information rate was about one-third of coding capacity. For every ganglion cell, information rate was very close (within 4%) to that predicted from Poisson noise and the cell's actual time-modulated rate.
Conclusions: Different messages are transmitted with similar efficiency. Efficiency is limited by temporal correlations, but correlations may be essential to improve decoding in the presence of irreducible noise.

Introduction

The best-studied retinal ganglion cells are termed "brisk," because they fire at short latencies and high rates. To a repeated stimulus, they fire at nearly the same time and with a nearly constant number of spikes, suggesting that they might encode information in the precise timing of individual spikes [1]. Brisk cells have been thoroughly studied, because in addition to their liveliness, they are large and relatively easy to record [2]. Yet, brisk cells comprise only about half of all gan-

*Correspondence: michael@retina.anatomy.upenn.edu

glion cells [3–11], and it is essential to ask how the other half transmits information.

The remaining cells are termed "sluggish," because they tend to fire at longer latencies and lower rates [2]. Their long latencies and weak responses suggest that they might be considerably less precise, and thus, with respect to a timing code, "informationally disadvantaged." Sluggish cells have been less studied, because in addition to their weak responses, they are small and harder to record. Furthermore, they comprise numerous types, many selective for a specific stimulus feature such as direction of motion or a local-edge [2, 12, 13]. If "meaning" is established by this selectivity, then sluggish cells must transmit many different messages, all different from the messages of brisk cells.

We wondered if, given such diversity, ganglion cell coding might be governed by any general principles. In particular, we wondered how efficiently a ganglion cell exploits its capacity to encode information and whether brisk and sluggish cells have the same efficiency. Thus, we asked, what is the maximum information rate that an ideal ganglion cell could transmit, and how closely do real cells approach this limit?

Information rates were estimated by the direct method [14], which counts the number of different spike patterns, accounts for the deleterious effect of noise on information, and requires long, stable recordings. This was achieved, even for small cells, by viewing the intact guinea pig retina in vitro with infrared-DIC optics and by selecting cells under visual control with a loose-patch electrode. We found that every cell uses about one-third of its capacity and that the relation between information and firing rate is described by a single logarithmic function. When a cell's mean rate was used to modulate the rate of a Poisson process, the information rates of both brisk and sluggish cells approximated the information rate of the modulated Poisson process. Thus, brisk and sluggish cells have similar coding efficiency, suggesting that common principles govern their signaling.

Results

Brisk Cells Fire More Precisely and at Higher Rates

We recorded spikes from ganglion cells in an in vitro preparation of the mammalian (guinea pig) retina by forming "loose" seals (<1 GΩ) with low-resistance glass pipettes ($\sim$5 MΩ). Cells were classified as brisk (N = 19) or sluggish (N = 23) based on their responses to visual stimuli and their morphology (see Experimental Procedures). A major distinction is that brisk cells responded to a spatially uniform stimulus that included the receptive field surround, while some sluggish cells were suppressed (Figure 1). Thus, we used a 715 $\times$ 715 μm field of independently flickering checks (25 cells) or a single flickering check (17 cells, Figure 2). For both stimuli, check intensity was randomly chosen every 33 ms (30 Hz) from a Gaussian distribution with a mean
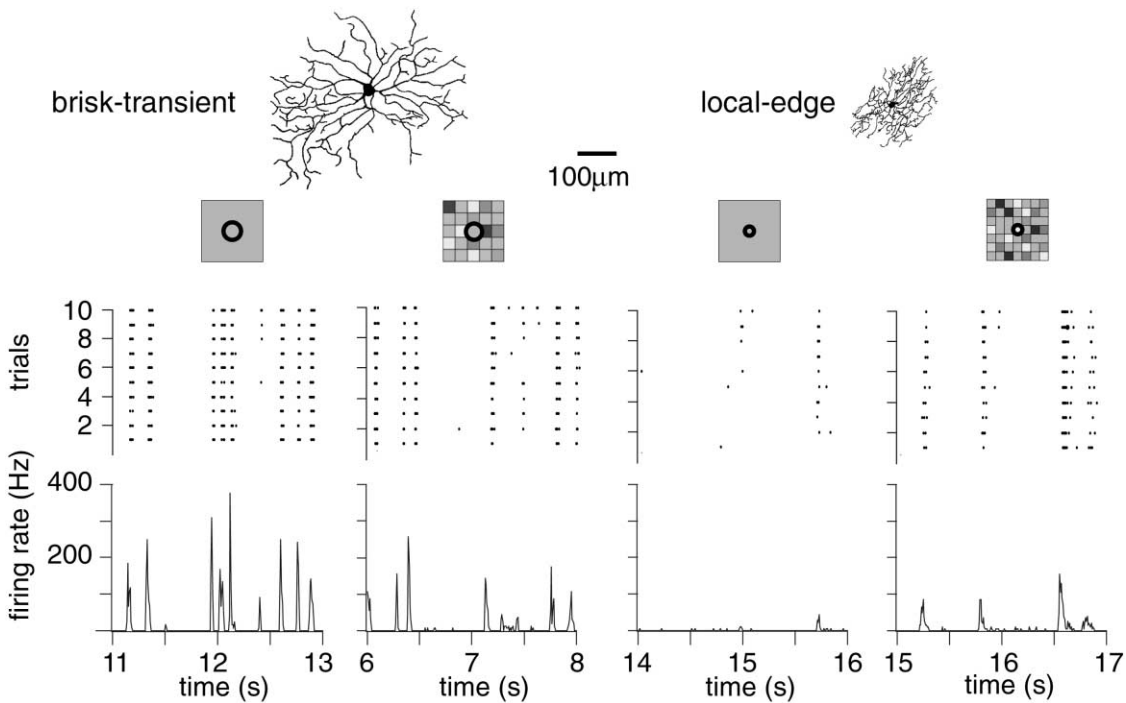
Figure 1. Standard Stimuli Drive Brisk and Sluggish Cells at Different Rates

Stimulus was either a spatially homogeneous flicker or a flickering checkerboard optimized for the receptive field center. Stimulus sequence was typically repeated 80 times to obtain jitter, peak rates for firing events, and the overall mean rate for the entire recording.

OFF brisk-transient cell: Spatially homogeneous stimulus-evoked reproducible firing events with a firing precision (jitter) of 3 ms. Peak rate during an event and average rate for entire response were high (375 and 8.3 spikes s^{-1}). Checkerboard evoked a similar number of events with the same precision (3 ms) and similar peak and average rates (369 and 9 spikes s^{-1}).

Local-edge cell: Spatially homogeneous stimulus evoked few events, which were reproducible with a precision of 11 ms. Peak and average rates were low (44 and 0.4 spikes s^{-1}). Checkerboard evoked more firing events with a similar precision (10 ms). Peak and average rates were greater than for the spatially homogeneous stimulus (170 and 3.3 spikes s^{-1}).

intensity of about 10^4 photons s^{-1} μm^{-2} and a standard deviation equal to one-third of the mean. Check size was adjusted to produce the maximal response. In general, and unless otherwise stated, measures did not sig-

nificantly differ between stimuli. There were no interactions between cell class and stimulus (analysis of variance, Bonferroni/Dunn post hoc test, $p < 0.01$, Table 1).

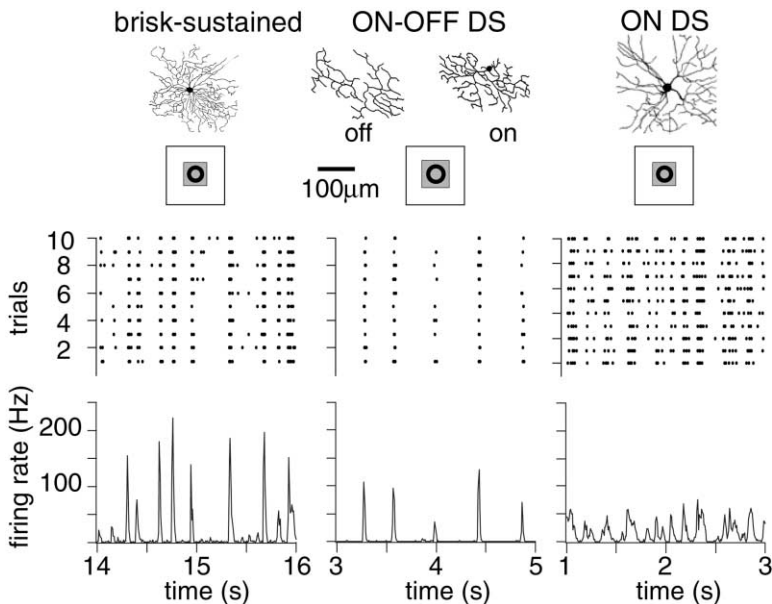The checkerboard and single-check stimuli were re-

Figure 2. Responses of Brisk-Sustained and Directionally Selective Cells to Flicker over the Receptive Field Center

Check size was optimized to evoke the highest firing rate from each cell.

OFF brisk-sustained cell fired precisely (4 ms) with high peak and average firing rates (241 and 32 spikes s^{-1}).

ON-OFF DS cell (dendrites shown separately for off and on strata) responded about as precisely as brisk cell (5 ms) but at lower rates (141 and 2.7 spikes s^{-1}).

ON DS cell responded less precisely than brisk cell (12 ms) and at lower rates (77 and 16 spikes s^{-1}). Generally, sluggish cells fired less precisely than brisk cells and at lower rates.

Table 1. Comparison of Brisk and Sluggish Ganglion Cell Classes

	Brisk	Sluggish
Firing rate (spikes s^{-1})	13 ± 7	8 ± 5[a]
Firing precision (ms)	6 ± 2	10 ± 3[a]
Information rate (bits s^{-1})	21 ± 9	13 ± 7[a]
Bits per spike	2.0 ± 0.7	2.1 ± 0.6
Information efficiency (% of capacity)	34 ± 7	32 ± 6
Total entropy (% of capacity)	78 ± 13	90 ± 9[a]
Noise entropy (% of capacity)	44 ± 13	59 ± 10[a]
$Z_{entropy}$	−0.33 ± 0.22	−0.12 ± 0.11[a]
Fano factor	0.9 ± 0.1	0.9 ± 0.1

[a] p < .01

peated 60 to 80 times; each stimulus repeat lasted 20 s. Brisk and sluggish cells produced "firing events," i.e., bursts of spikes that repeated for every stimulus [15] (Figures 1 and 2). The average firing rate, including firing events and intervening silent periods, was higher for brisk cells than for sluggish cells (13 ± 7 versus 8 ± 5 Hz).

Because coding capacity depends on the precision of spike timing, we estimated the standard deviation of spike timing across repeats ("jitter," see Experimental Procedures) [1]. Jitter was less for the checkerboard than it was for the single check (7 ± 3 versus 10 ± 4 ms). For both stimuli, brisk cells had less jitter than the sluggish cells did (6 ± 2 versus 10 ± 3 ms).

Brisk Cells Transmit at Higher Information Rates
We estimated information rates by using the direct method [14] (see Experimental Procedures). To accomplish this, we divided the spike train into time bins, counted the number of spikes in a 5 ms bin to form a "letter," and concatenated these letters into "words." We tracked the frequency of each word's occurrence and calculated the entropy of the frequency distribution.

The information rate was calculated as the total entropy of the spike train minus its noise entropy [14, 16]. The total entropy (unconditional entropy) was calculated for the distribution of words in the entire recording. The noise entropy (entropy conditional upon the stimulus) was calculated for words that occurred at the same time across repeats and was averaged across times. Both total and noise entropies were extrapolated to infinite data size and word length [16](see Experimental Procedures). In general, we found that cells with higher firing rates had higher information rates, total entropy, and noise entropy (Figures 3A–3C). Accordingly, brisk cells had higher information rates than sluggish cells (21 ± 9 versus 13 ± 7 bits s^{-1}).

Spikes from Brisk and Sluggish Cells Carry Equal Amounts of Information
To estimate the amount of information transmitted by a spike, we divided each cell's information rate by its average firing rate. Information per spike decreased with increasing firing rates from nearly 4 bits at 1 Hz to about

1 bit at 25 Hz (Figure 3D). Although brisk cells had higher firing rates, suggesting that they transmit less information per spike, their firing rates sufficiently overlapped those of sluggish cells so that both cell classes transmitted approximately the same amount of information per spike (2.0 ± 0.7 versus 2.1 ± 0.6 bits).

Brisk and Sluggish Cells Have Similar Transmission Efficiency
To estimate coding efficiency, we compared the information transmitted by a spike train to the maximum possible entropy at the observed spike rate, which sets an upper limit on information rate [17–19]. This limit, termed "coding capacity," is reached when the number of spikes in each time bin is independent of all other time bins (i.e., no temporal correlations), and the spike train is perfectly reproducible across trials (i.e., no noise) [19]. The probability of a spike occurring in any bin will be:

$$P_1 = R\Delta t \tag{1}$$

R is the mean firing rate and Δt is the bin width. Therefore, the probability of not observing a spike will be:

$$P_0 = 1 - R\Delta t \tag{2}$$

Therefore, the entropy per bin is:

$$H_{bin} = -R\Delta t \log_2(R\Delta t) - (1 - R\Delta t)\log_2(1 - R\Delta t) \text{ bits} \tag{3}$$

Dividing by Δt converts this into coding capacity in bits per second:

$$C(R, \Delta t) = \frac{-R\Delta t \log_2(R\Delta t) - (1 - R\Delta t)\log_2(1 - R\Delta t)}{\Delta t} \text{ bits s}^{-1} \tag{4}$$

We calculated coding capacity $C(R, \Delta t)$ for each cell by using its mean firing rate (R) and a bin width (Δt) of 5 ms. We defined information efficiency E as the percentage of coding capacity used to transmit information. For both brisk and sluggish cells combined, E was about 33 ± 7%.

To compare efficiency across firing rates, we graphed each cell's information rate against its firing rate R and fit the points with the equation $E \cdot C(R, \Delta t)$ (Figure 3A). The equation fit well because deviations of the information rate from this equation accounted for only 18% of its total variance (K^2 = coefficient of nondetermination). We also calculated total entropy for each cell as a percentage of its capacity. For brisk and sluggish cells combined, total entropy was 85 ± 12% of capacity, and the function $0.85 C(R, \Delta t)$ was a good fit (K^2 = 13%, Figure 3B). Noise entropy was 52 ± 14% of capacity, and the function $0.52 C(R, \Delta t)$ was an adequate fit (K^2 = 45%, Figure 3C).

Brisk and sluggish cells were equally efficient (34 ± 7 versus 32 ± 6% of capacity). Yet, the total entropy of brisk cells fell somewhat farther from their capacity than the total entropy of sluggish cells (78 ± 13 versus 90 ± 9% of capacity). Similarly, the noise entropy of brisk cells fell farther from capacity than the noise entropy of sluggish cells (44 ± 13 versus 59 ± 10% of capacity). Our interpretation is that brisk cells use less of their

A information rate

B total entropy
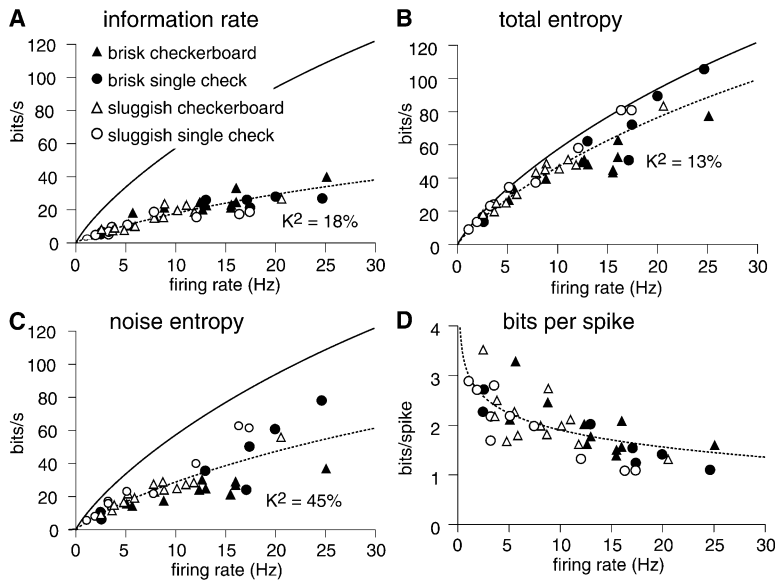
C noise entropy

D bits per spike

Figure 3. Brisk and Sluggish Cells Transmit Information with Similar Efficiency

(A–C) Solid line was calculated from the equation for coding capacity $C(R, \Delta t)$ versus average firing rate (R) with bin width $\Delta t = 5$ ms (see text). Dashed line is a constant percentage of $C(R, \Delta t)$, adjusted for best fit to information rate, total entropy, or noise entropy. For example, in A, dashed line is 0.33 * $C(R, \Delta t)$, denoting that information efficiency is 33%. K^2 is the coefficient of nondetermination, i.e., the percentage of the total variance comprised by the vertical deviations of data points from the dashed line.

(D) Information per spike declines with increasing rate. Dashed line is the equation for information efficiency divided by firing rate, 0.33 $C(R, \Delta t)/R$.

capacity for signaling (total entropy) than sluggish cells do, but they also lose less of this capacity to noise, leaving them with approximately equal capacity to transmit information.

Coding capacity increases as the bin width Δt decreases (Equation 4). In estimating both information and capacity, we chose a bin width of $\Delta t = 5$ ms, which was less than the average jitter of either brisk or sluggish classes, but it was between the absolute minimum and maximum jitter ($4 < \Delta t < 17$ ms). Bin widths significantly greater than the jitter (maintaining 4, 5, and 6 bins in a word) under-represented capacity and inflated efficiency; for example, a bin width of 18 ms gave an efficiency of 44%. Bin widths significantly less than the jitter (0.5, 1, and 3 ms), reduced efficiency only slightly (33 to 31%). But for any reasonable value of bin width, brisk and sluggish cells were equally efficient.

The standard stimulus refresh rate (30 Hz) effectively stimulated sluggish cells, but higher rates produced weak responses that precluded accurate information calculations. Brisk cells can respond to higher temporal frequencies, and thus, in a control experiment, we recorded from seven brisk cells and used a 120 Hz refresh rate. Efficiency measured in the standard way averaged $33 \pm 9\%$, which was not statistically different from the efficiency measured for brisk cells with the standard refresh rate ($34 \pm 7\%$).

Temporal Correlations Reduce Entropy

Temporal correlations between bins would prevent total entropy from reaching capacity. To measure these correlations, we computed the difference in total entropy for single bins and for infinitely long words, and then we divided this by the entropy for long words ($Z_{entropy} = [H_{long} - H_{short}]/H_{long}$) [20]. $Z_{entropy}$ was significantly below zero for both brisk cells and sluggish cells, indicating that they both had temporal correlations (a $Z_{entropy}$ of zero would indicate no correlations). Brisk cells had significantly more negative values than sluggish cells; the latter exhibited a more skewed distribution of values con-

centrated just below zero (median: -0.32 versus -0.08, skew: -0.58 versus -0.71), indicating that brisk cells had a greater degree of temporal correlation (Figure 4A). Because temporal correlations reduce the total entropy, more negative $Z_{entropy}$ for brisk cells is consistent with their total entropy departing further from coding capacity (78 ± 13 versus $90 \pm 9\%$ of capacity).

Ganglion Cell Coding Efficiency Approaches that of a Rate-Modulated Poisson Process

Our calculation of $Z_{entropy}$ shows that temporal correlations prevent ganglion cells from reaching their coding
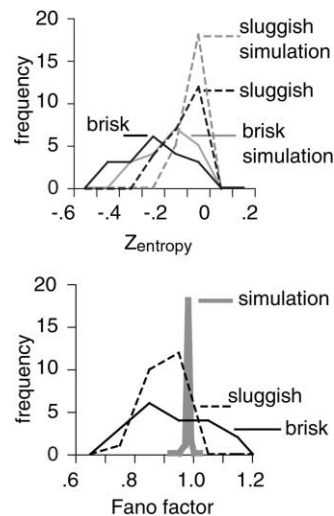

Figure 4. Quantifying Spike Train's Temporal Correlations and Noisiness

Upper graph: Negative $Z_{entropy}$ was greater for brisk than for sluggish cells. $Z_{entropy}$ for both classes resembled that of their respective rate-modulated Poisson simulation, reflecting similar temporal correlations.

Lower graph: Fano factor (a measure of noise) was smaller for real cells than for the Poisson simulation.
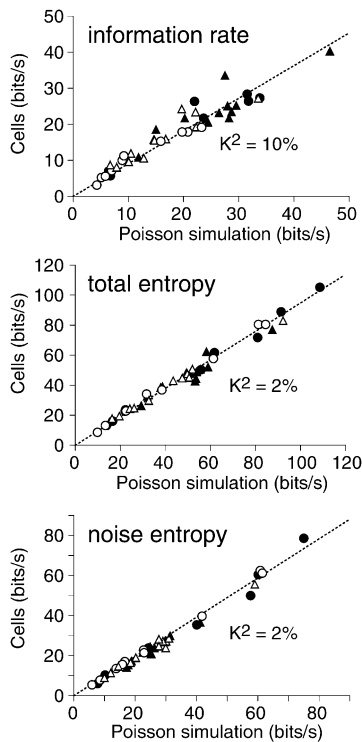
Figure 5. Spike Trains and Rate-Modulated Poisson Process Have Similar Information Rates, Total Entropy, and Noise Entropy

The information rate, total entropy, and noise entropy of brisk and sluggish cells are graphed against the same measures from the Poisson simulation (symbols same as Figure 3). Linear regression fits are constrained to pass through the origin. Deviation of information rate from a regression fit was only 10% of its total variance (K^2 = coefficient of nondetermination). Similarly, total entropy and noise entropy deviated by only 2%. By this standard, the efficiency of information transmission is similar for all cells.

capacity. Noise due to stochastic spike generation would also prevent ganglion cells from reaching capacity. Thus, we constructed a simple model of the spiking process that included both stimulus-evoked correlations and noise and used it to compare brisk and sluggish cells.

The model was a rate-modulated Poisson process. The time-varying firing rate was measured for each recorded cell in 5 ms bins across repeats (i.e., the poststimulus time histogram) and was used to set the instantaneous rate of a Poisson noise generator that determined the number of spikes in 0.5 ms bins. Then, simulated recordings were generated from this model by matching the number of repeats in simulated and actual recordings. The simulated recordings were evaluated by the direct method just as they were for the actual recordings to measure their total entropy, noise entropy, and information rate. In these measures, the actual recordings closely approximated the simulated recordings, reaching 96 ± 15% of the information rate, 97 ± 7% of the total entropy, and 98 ± 7% of the noise entropy. On all these measures, brisk and sluggish cells were not statistically different (total entropy: 95 ± 7 versus 99 ± 6%, noise entropy: 95 ± 7 versus 100 ± 6%, information rate: 95 ± 15 versus 97 ± 15%, Figure 5).

Ganglion Cell Spike Train Shows Temporal Correlations Similar to a Rate-Modulated Poisson Process

We calculated $Z_{entropy}$ for the simulated recordings just as we did for the actual recordings. Brisk cells had more negative values than sluggish cells did; sluggish cells exhibited a more skewed distribution of values concentrated just below zero (median: −0.15 versus −0.07, skew: −0.29 versus −0.54; Figure 4A). In this regard, the simulations duplicated the actual recordings. Indeed, for both cell classes the distributions of $Z_{entropy}$ closely matched that of their simulations. This suggests that the time-modulated firing rate largely accounted for differences in temporal correlations between brisk and sluggish cell classes.

Ganglion Cells Fire More Regularly Than a Rate-Modulated Poisson Process

To compare the noisiness of a recording and that of its simulation, we counted spikes in a sliding window of 5 ms incremented by 1 ms. We then calculated the ratio of the variance of this count to the mean count across repeats (Fano factor). For the recordings, the Fano factor ranged between 0.7 and 1.2. Most recordings had values less than one, leading to a value of 0.9 ± 0.1 over the entire population (Figure 4B). Brisk and sluggish cells had statistically indistinguishable Fano factors. For the simulations, as expected, the Fano factor was 1 ± 0.01. Thus, recordings had slightly less noise entropy than a rate-modulated Poisson process, partly because brisk and sluggish cells fired more consistently across repeats.

Discussion

To our initial question, "how efficiently does a ganglion cell exploit its capacity to encode information?" we now have an answer. First, we compared a cell's information rate to that of a spike train with the same mean rate but lacking temporal correlations and noise. This standard represents the greatest information rate that can be conveyed at a given spike rate, i.e., the coding capacity [17, 19]. Against this standard, a cell's estimated information rate was consistently about one-third of capacity (Figure 3A). Deviation from this fixed percentage was moderate (K^2 = 18%). The estimated total entropy was 85% of capacity and deviated less (K^2 = 13%), but the noise entropy was 52% of capacity and deviated more (K^2 = 45%). Brisk and sluggish cells filled their coding capacity to the same extent. The main distinction between the two classes is that brisk cells had greater temporal correlation (more negative $Z_{entropy}$) and correspondingly less total entropy (78 ± 13 versus 90 ± 9% of capacity). Similar results have been seen in the salamander and rabbit retinas (M.B., unpublished data).

Second, we compared a cell's information rate to the rate of a modulated Poisson process. This standard accounted for temporal correlations (by including the actual time-varying spike rate from each cell) and for noise (by including a Poisson noise generator). We omitted a relative refractory period because brisk and sluggish cells differ (~1 ms versus 5 ms), so including it

would have produced different standards. A cell's information rate approached the information rate of a modulated Poisson process, accounting for 96% of the simulation's information rate with a small deviation ($K^2 = 10\%$). A cell's estimated total entropy and noise entropy also closely matched the modulated Poisson process ($K^2 = 2\%$). Brisk and sluggish cells were indistinguishable by any of these measures.

Spike Rate Encodes Most Information

The match between actual information rates and the modulated Poisson process suggests that most of the information in spike trains is carried by a rate code with fine temporal modulation. At first glance, this might appear inconsistent with reports that spike timing conveys more information than the number of spikes [1]. However, we specified time-varying firing rate with a bin width very close to the spike train's timing precision and refractory period (5 ms). Thus, the high temporal precision with which we modulate our "rate code" is essentially equivalent to a "timing code," because most bins would contain only one well-timed spike.

The only information not represented by a finely modulated rate would be meaningful patterns of spikes whose relative timing to one another is more precise than their timing relative to the stimulus. Such patterns are known: ganglion cells exhibit spike bursts that are not correlated with the stimulus. Bursts last about 50 ms in sluggish cells in which they apparently result from feedback between amacrine and bipolar cells [21]. Bursts last about 10 ms in brisk cells in which they are apparently caused by single transmitter quanta. Similarly, there are more correlated spikes between different ganglion cells than with the stimulus, providing spatial spike patterns [22]. At present, there is controversy as to whether such spike patterns convey extra information that is not present in the spike rate [22–24]. Although spike patterns do not appear to appreciably change the encoded information, it is still possible that the meaning of those spike patterns differs from the meaning encoded by spike rate [25].

Coding Capacity of a Single Spike

The information encoded by a single spike averages about 2 bits, similar to previous estimates (1.6, 1.7 bits) [26, 27] and declines with increasing spike rate. This decline can be explained by a simplified equation for coding capacity (Equation 4). For a brisk cell firing at about 13 spikes per second, the probability of a spike occurring in a given 5 ms bin is less than 0.07. Sluggish spike rates are lower, and thus the probability of a spike is even lower. Given a very low probability, Equation 4 simplifies to:

$$\text{capacity per spike} \approx -\log_2(R\Delta t) \text{ bits} \qquad (5)$$

Thus, a spike's capacity to convey information is proportional to the negative logarithm of the firing rate. Assuming that coding efficiency is relatively constant across spike rates, the actual information conveyed by a spike will be some constant proportion of this capacity (K) and thus equal to:

$$\text{information per spike} \approx -K\log_2(R\Delta t) \text{ bits} \qquad (6)$$

This explains why information per spike declines with increasing spike rate (Figure 4D) [28].

Functional Implications

A neuron's efficiency compared to a noiseless, uncorrelated standard has been estimated in several sensory systems at 20–60% (frog saccular hair cell, cricket mechanoreceptor, and salamander ganglion cell) [17]. Also, coding efficiency for retinal ganglion cells under naturalistic stimulation was found to be nearly 50% in salamander and 30% in rabbit (M.B., unpublished data). The present estimate of $33 \pm 7\%$ falls squarely within the previously estimated range; thus, efficiency is broadly consistent across diverse sensory modalities. The present contribution, beyond its focus on a mammalian system, is to show that efficiency is highly consistent across different cell types within a broad class (ganglion cells). Furthermore, when efficiency is estimated in a way that takes into account a cell's average firing rate, efficiency is preserved across different spatial patterns and different refresh rates. The white-noise stimuli used here lack spatio-temporal correlations that might favor brisk or sluggish cells. Rather, they contain a wide variety of patterns that evoke a broad ensemble of spike patterns in both cell classes. It is possible that natural scenes, which have well-known spatio-temporal correlations, might result in responses with a different efficiency for one or both cell classes. However, if the primary distinction between responses to our stimuli and to natural scenes is a change in the firing rate, we expect that one-third efficiency will continue to hold.

Why would efficiency be consistent across cells? Efficiency may result from an optimization that transmits the most information for the fewest spikes. There would be strong selective pressure for this, because spiking is metabolically expensive [29–34]. Thus, efficiency may be consistent because the same optimizing principles apply to all cells.

Why can't a cell completely fill its capacity? The main losses of efficiency are attributable to noise and temporal correlations. Spikes, because they depend on ion channels, are inherently noisy. Temporal correlations conveying redundant information could be removed [35–37], but they are retained because they allow a noisy signal to be partially decoded. Redundancy improves decoding by allowing an estimate of the average rate during the interval over which the correlations occur. Thus, temporal correlations may be retained because they improve decoding in the presence of noise.

Experimental Procedures

Electrophysiology

All procedures were performed in accordance with University of Pennsylvania and National Institutes of Health guidelines. The retina, attached to pigment epithelium, choroid, and sclera, was flattened by cutting radial slits and applied sclera side down to filter paper. The retina was placed in a chamber on the stage of an upright microscope and superfused with oxygenated Ames medium (37°C). Extracellular recordings were sampled at 5 kHz and high-pass filtered at 100 Hz (Axoclamp-2A, Digidata 1322A, Clampex 8.1 software, Axon Instruments, Foster City, CA). Spikes were detected when the recording's first derivative exceeded a threshold.

Stimuli

Visual stimuli were generated in Matlab (MathWorks, Natick, MA), using the Psychophysics Toolbox to control a video card [38, 39]. The stimulus was displayed on a 1-inch CRT monitor with a P43 phosphor (545 nm, 640 × 480 pixels, 60 Hz frame rate, Lucivid MR1-103A; MicroBrightfield, Inc., Colchester, VT), projected through the top port of the microscope, and focused onto the retina at a magnification of 7.6×. The relationship between voltage and monitor intensity was linearized in software with a lookup table.

Cell Morphology

At the end of recording, the loose patch pipette was replaced with a sharp pipette (30MΩ) backfilled with 1% DiI in 100% EtOH (1,1′-dioctadecyl-3,3,3′,3′-tetramethylindocarbocyanine perchlorate, Molecular Probes, Eugene, OR). Cells were injected with DiI with a positive current pulse for 5–10 min. The tissue was subsequently fixed in 4% formaldehyde for 15 min and mounted from 0.1M phosphate buffer (pH 7.4). Cell bodies were counterstained to locate the borders of the inner-plexiform layer by immersing the retina for 3 min in a buffered 0.1% solution of the nucleic acid dye, SYTO13 (Molecular Probes). Cell morphology and stratification were recorded by confocal microscopy.

Cell Classification

Brisk and sluggish cell classes were identified by spike responses and morphology [2, 40–44]. When stimuli extended beyond the receptive field center, brisk cells were weakly antagonized while sluggish cells were strongly suppressed. To a single flickering check, brisk cells had higher peak rates (>100 spikes s^{-1}). For brisk cells, the autocorrelogram of spike times showed a narrower peak (<5 ms) [21, 45]. By these criteria, "brisk" included brisk-sustained, brisk-transient, and delayed-off types [45]. Sluggish cells included local-edge, on DS, and on-off DS types.

Examples of each cell type were too few to identify new correspondences between spiking and morphology [46], but we confirmed the known ones: brisk-transient cells had a radiating dendritic tree (α morphology, Figure 2), and brisk-sustained cells had bushy dendrites (β morphology, Figure 3) [47]. The local-edge detector had narrow monostratified dendrites, ramifying in the middle of the inner plexiform layer, and the on-off directionally selective cell had a small bistratified morphology [43, 48, 49]. The on directionally selective cell had a dendritic arbor almost as large as the on brisk-transient cell, but it was more convoluted.

Calculation of Timing Precision (Jitter)

Firing precision was calculated first by constructing the shuffled autocorrelation function, which represents correlations between spikes in different stimulus repeats [50, 51]. This removed primary correlations between spikes in the same repeat. Yet, such primary correlations would produce spurious secondary correlations between spikes across repeats. Thus, we corrected for these spurious correlations by deconvolving the shuffled autocorrelation function with the autocorrelation function. We fit the corrected correlation function with a Gaussian function, and we used the standard deviation of this function (in milliseconds) as a measure of jitter.

Extrapolation and Data Adequacy

We extrapolated our estimates of total and noise entropy to infinite data size according to the method of Strong et al. [16]. Stimulus repeats were divided into shorter segments, and the estimated entropy was graphed against the number of segments, s. We then fit this graph with a third-order polynomial ($H_0 + H_1 s + H_2 s^2$) and took H_0 as the extrapolated entropy. Extrapolated entropy was plotted against inverse-word length (1 to 6 bins) and extrapolated to zero inverse-word length or, in other words, infinite-word length. To assure we had enough data to accurately estimate information, we compared the estimated total and noise entropy to the entropies measured for our longest word length (6 bins). If the estimated total or noise entropy was less than 15% of the entropies calculated for the longest word, we then excluded that recording (16 cells).

References

1. Berry, M.J., Warland, D.K., and Meister, M. (1997). The structure and precision of retinal spike trains. Proc. Natl. Acad. Sci. USA *94*, 5411–5416.

2. Cleland, B.G., and Levick, W.R. (1974). Brisk and sluggish concentrically organized ganglion cells in the cat's retina. J. Physiol. *240*, 421–456.

3. Wässle, H., Levick, W.R., and Cleland, B.G. (1975). The distribution of the alpha type of ganglion cells in the cat's retina. J. Comp. Neurol. *159*, 419–438.

4. Hughes, A., and Wassle, H. (1976). The cat optic nerve: fibre total count and diameter spectrum. J. Comp. Neurol. *169*, 171–184.

5. Stone, J. (1978). The number and distribution of ganglion cells in the cat's retina. J. Comp. Neurol. *180*, 753–771.

6. Hughes, A. (1981). Population magnitudes and distribution of the major modal classes of cat retinal ganglion cell as estimated from HRP filling and a systematic survey of the soma diameter spectra for classical neurones. J. Comp. Neurol. *197*, 303–339.

7. Leventhal, A.G., Rodieck, R.W., and Dreher, B. (1985). Central projections of cat retinal ganglion cells. J. Comp. Neurol. *237*, 216–226.

8. Wässle, H., Boycott, B.B., and Illing, R.B. Morphology and mosaic of on- and off-beta cells in the cat retina and some functional considerations. Proc R Soc Lond [Biol] 1981, 212:177–195.

9. Wässle H, Peichl L, Boycott BB: Morphology and topography of on- and off-alpha cells in the cat retina. Proc. R. Soc. Lond. [Biol.] 1981, 212:157–175.

10. Fukuda, Y., Hsiao, C.F., and Watanabe, M. Morphological correlates of Y, X and W type ganglion cells in the cat's retina. Vision Res. 1985, 25:319–327.

11. Stein, J.J., Johnson, S.A., and Berson, D.M. (1996). Distribution and coverage of beta-cells in the cat retina. J. Comp. Neurol. *0372*, 597–617.

12. Stone, J., and Fukuda, Y. (1974). Properties of cat retinal ganglion cells: a comparison of W-cells with X-cells and Y-cells. J. Neurophysiol. *37*, 722–748.

13. Rowe, M.H., and Palmer, L.A. (1995). Spatio-temporal receptive-field structure of phasic W cells in the cat retina. Vis. Neurosci. *12*, 117–139.

14. van Steveninck, R.R.D., Lewen, G.D., Strong, S.P., and Koberle, R. (1997). W. B: Reproducibility and variability in neural spike trains. Science *275*, 1805–1808.

15. Berry, M.J., and Meister, M. (1998). Refractoriness and neural precision. J. Neurosci. *18*, 2200–2211.

16. Strong, S.P. (1998). de Ruyter van Steveninck RR, Bialek W, Koberle R: On the application of information theory to neural spike trains. Pac. Symp. Biocomput. *80*, 621–632.

17. Rieke, F., and Warland, D. Stevenick RdRv, Bialek W: *Spikes Exploring the Neural Code*. Cambridge, MA: MIT Press; 1997.

18. Reinagel, P., Godwin, D., Sherman, S.M., and Koch, C. (1999). Encoding of visual information by LGN bursts. J. Neurophysiol. *81*, 2558–2569.

19. MacKay, D., and McCulloch, W.S. (1952). The limiting information capacity of a neuronal link. Bull. Math. Biophys. *14*, 127–135.

20. Reinagel, P., and Reid, R.C. (2000). Temporal coding of visual information in the thalamus. J. Neurosci. *20*, 5392–5400.

21. Freed, M.A., Smith, R.G., and Sterling, P. (2003). Timing of quantal release from the retinal bipolar terminal is regulated by a feedback circuit. Neuron *38*, 89–101.

22. Meister, M., Lagnado, L., and Baylor, D.A. (1995). Concerted signaling by retinal ganglion cells. Science *270*, 1207–1210.

23. Nirenberg, S., Carcieri, S.M., Jacobs, A.L., and Latham, P.E. (2001). Retinal ganglion cells act largely as independent encoders. Nature *411*, 698–701.

24. Schneidman, E., Bialek, W., and Berry, M.J., 2nd. (2003). Synergy, redundancy, and independence in population codes. J. Neurosci. *23*, 11539–11553.

25. Fairhall, A.L., Lewen, G.D., Bialek, W., and de Ruyter Van Steveninck, R.R. (2001). Efficiency and ambiguity in an adaptive neural code. Nature *412*, 787–792.

26. Warland, D.K., Reinagel, P., and Meister, M. (1997). Decoding visual information from a population of retinal ganglion cells. J. Neurophysiol. *78*, 2336–2350.

27. Passaglia, C.L., and Troy, J.B. (2004). Information transmission rates of cat retinal ganglion cells. J. Neurophysiol. *91*, 1217–1229.

28. Rieke, F., Bodnar, D.A., and Bialek, W. (1995). Naturalistic stimuli increase the rate and efficiency of information transmission by primary auditory afferents. Proc R Soc Lond B Biol Sci. *262*, 259–265.

29. Attwell, D., and Laughlin, S.B. (2001). An energy budget for signaling in the grey matter of the brain. J. Cereb. Blood Flow Metab. *21*, 1133–1145.

30. Balasubramanian, V., and Berry, M.J., 2nd. (2002). A test of metabolically efficient coding in the retina. Network *13*, 531–552.

31. Levy, W.B., and Baxter, R.A. (1996). Energy efficient neural codes. Neural Comput. *8*, 531–543.

32. Laughlin, S.B., van Steveninck, R.R.D., and Anderson, J.C. (1998). The metabolic cost of neural information. Nat. Neurosci. *1*, 36–41.

33. Balasubramanian, V., Kimber, D., and Berry, M.J., 2nd. (2001). Metabolically efficient information processing. Neural Comput. *13*, 799–815.

34. de Polavieja, G.G. (2002). Errors drive the evolution of biological signalling to costly codes. J. Theor. Biol. *214*, 657–664.

35. Barlow, H.B. Possible principles underlying the transformation of sensory messages. In: *Sensory Communication*. pp. 217–234. Cambridge, MA: MIT Press; 1961: 217–234.

36. Srinivasan, M.V., Laughlin, S.B., and Dubs, A. (1982). Predictive coding: a fresh view of inhibition in the retina. Proc. R. Soc. Lond. B. Biol. Sci. *216*, 427–459.

37. Atick, J.J., Griffin, P.A., and Redlich, A.N. (1996). Statistical approach to shape from shading: reconstruction of three-dimensional face surfaces from single two-dimensional images. Neural Comput. *8*, 1321–1340.

38. Brainard, D.H. (1997). The Psychophysics Toolbox. Spat. Vis. *10*, 443–446.

39. Pelli, D.G., and Zhang, L. (1991). Accurate control of contrast on microcomputer displays. Vision Res. *31*, 1337–1350.

40. Caldwell, J.H., and Daw, N.W. (1978). New properties of rabbit retinal ganglion cells. J. Physiol. *276*, 257–276.

41. Cleland, B.G., and Levick, W.R. (1974). Properties of rarely encountered types of ganglion cells in the cat's retina and an overall classification. J. Physiol. *240*, 457–492.

42. Rockhill, R.L., Daly, F.J., MacNeil, M.A., Brown, S.P., and Masland, R.H. (2002). The diversity of ganglion cells in a mammalian retina. J. Neurosci. *22*, 3831–3843.

43. Berson, D.M., Pu, M., and Famiglietti, E.V. (1998). The zeta cell: A new ganglion cell type in cat retina. J. Comp. Neurol. *399*, 269–288.

44. Amthor, F.R., Takahashi, E.S., and Oyster, C.W. (1989). Morphologies of rabbit retinal ganglion cells with concentric receptive fields. J. Comp. Neurol. *280*, 72–96.

45. Devries, S.H., and Baylor, D.A. (1997). Mosaic arrangement of ganglion cell receptive fields in rabbit retina. J. Neurophysiol. *78*, 2048–2060.

46. Troy, J.B., and Shou, T. (2002). The receptive fields of cat retinal ganglion cells in physiological and pathological states: where we are after half a century of research. Prog. Retin. Eye Res. *21*, 263–302.

47. Boycott, B.B., and Wässle, H. (1974). The morphological types of ganglion cells of the domestic cat's retina. J. Physiol. *240*, 397–419.

48. Amthor, F.R., Takahashi, E.S., and Oyster, C.W. (1989). Morphologies of rabbit retinal ganglion cells with complex receptive fields. J. Comp. Neurol. *280*, 97–121.

49. Amthor, F.R., Oyster, C.W., and Takahashi, E.S. (1984). Morphology of on-off direction-selective ganglion cells in the rabbit retina. Brain Res. *298*, 187–190.

50. Perkel, D.H., Gerstein, G.L., and Moore, G.P. (1967). Neuronal spike trains and stochastic point processes II. Simultaneous spike trains. J. Neurophysiol. *7*, 419–440.

51. Ghose, G.M., Ohzawa, I., and Freedman, R.D. (1994). Receptive-field maps of correlated discharge between pairs of neurons in the cat's visual cortex. J. Neurophysiol. *71*, 330–346.