Behavioral/Systems/Cognitive

The Structure of Large-Scale Synchronized Firing in Primate Retina

Jonathon Shlens,[1,3] Greg D. Field,[1] Jeffrey L. Gauthier,[1] Martin Greschner,[1] Alexander Sher,[2] Alan M. Litke,[2] and E. J. Chichilnisky[1]

[1]The Salk Institute, La Jolla, California 92037, [2]Santa Cruz Institute for Particle Physics, University of California, Santa Cruz, California 95064, and [3]University of California, Berkeley, California 94720

Synchronized firing among neurons has been proposed to constitute an elementary aspect of the neural code in sensory and motor systems. However, it remains unclear how synchronized firing affects the large-scale patterns of activity and redundancy of visual signals in a complete population of neurons. We recorded simultaneously from hundreds of retinal ganglion cells in primate retina, and examined synchronized firing in completely sampled populations of ~50–100 ON-parasol cells, which form a major projection to the magnocellular layers of the lateral geniculate nucleus. Synchronized firing in pairs of cells was a subset of a much larger pattern of activity that exhibited local, isotropic spatial properties. However, a simple model based solely on interactions between adjacent cells reproduced 99% of the spatial structure and scale of synchronized firing. No more than 20% of the variability in firing of an individual cell was predictable from the activity of its neighbors. These results held both for spontaneous firing and in the presence of independent visual modulation of the firing of each cell. In sum, large-scale synchronized firing in the entire population of ON-parasol cells appears to reflect simple neighbor interactions, rather than a unique visual signal or a highly redundant coding scheme.

Introduction

Much of what is known about neural coding arises from studies of single neurons. However, in many neural circuits, the concerted activity of multiple cells cannot be predicted based on recordings from single cells, because the firing of different cells is not statistically independent. For example, retinal ganglion cells (RGCs), which transmit all visual information from the eye to the brain, exhibit strong synchronized firing—a tendency to fire nearly simultaneously more frequently than expected by chance (Arnett, 1978; Arnett and Spraker, 1981; Mastronarde, 1983a,b,c; Mastronarde, 1989; Meister et al., 1995; DeVries, 1999). Numerous studies based on recordings from pairs of nearby RGCs reveal that synchronized firing reflects a combination of common synaptic inputs and gap junction coupling (Mastronarde, 1983a,b,c; Mastronarde, 1989; Dacey and Brace, 1992; Jacoby et al., 1996; Stafford and Dacey, 1997; Brivanlou et al., 1998; Hu and Bloomfield, 2003; Hidaka et al., 2004; Amthor et al., 2005; Ishikane et al.,

2005; Ackert et al., 2006; Trong and Rieke, 2008). Theoretical and empirical studies suggest that synchronized firing in the retina could play an important role in processing and transmitting information (Meister, 1996; Nirenberg et al., 2001; Schneidman et al., 2006; Pillow et al., 2008), as it may in other sensory, memory, and motor systems (Usrey and Reid, 1999; Laurent, 2002; Harris, 2005; Narayanan et al., 2005).

However, to understand how synchronized firing affects the neural code of the retina, it is necessary to know how entire collections of RGCs—rather than just pairs—fire in concert (Schnitzer and Meister, 2003). Recent work suggests that synchronized firing in local clusters of 7–10 RGCs can be understood on the basis of pairwise interactions (Schneidman et al., 2006; Shlens et al., 2006). However, almost nothing is known about concerted activity in larger collections of RGCs, leaving open two central questions about how synchronized firing affects the neural code of the retina.

First, it is unclear whether large, distinctive patterns of synchronized firing could be produced by spatially extended sources of common input, potentially providing a unique visual signal to the brain (Schnitzer and Meister, 2003). In the mammalian retina, a diverse collection of wide-field amacrine cell types provide extensive lateral connections to RGCs (Dacey and Brace, 1992; Stafford and Dacey, 1997; MacNeil and Masland, 1998). These cells have been proposed to underlie visual functions such as the segregation of object and background motion (Masland, 2001; Baccus and Meister, 2002; Olveczky et al., 2003; Baccus, 2007), by simultaneously influencing activity in large collections of RGCs. The spatial extent of amacrine cell connectivity suggests that their impact may be difficult to observe in the small, local groups of RGCs previously examined.

Second, it is unclear whether the pairwise interactions in a complete population of cells may combine to dominate the activity of any individual cell (Schneidman et al., 2006; Bethge and Berens, 2007; Nirenberg and Victor, 2007; Tang et al., 2008). Extrapolations based on recordings from up to 10 cells have suggested that that the activity of each RGC could be strongly dictated by the surrounding population activity even if individual pairwise interactions are weak, implying high network redundancy (Schneidman et al., 2006). This prediction has never been tested in a population large enough or complete enough that the hypothesized redundancy could be observed.

Using large-scale electrophysiological recordings in the primate retina (Litke et al., 2004; Frechette et al., 2005), we probed the spatial scale and structure of synchronized firing in complete populations of ~50–100 ON-parasol RGCs, which uniformly sample visual space and form a major projection to the lateral geniculate nucleus. We report that patterns of synchronized firing in this population are significantly larger than previously appreciated. However, we also find that network activity can be summarized accurately on the basis of pairwise interactions between neighboring cells (Schneidman et al., 2006; Shlens et al., 2006), and therefore that no additional widespread interactions are implicated. Based on pairwise interactions, we show that no more than 20% of the activity each ON-parasol cell is accounted for by the activity of the network, indicating that although synchrony is powerful, the resulting signals are not massively redundant. These findings hold in the presence of visual stimulation, indicating that the observed structure of synchronized firing remains relevant during visual signaling.

Materials and Methods

Recordings. Preparation and recording methods have been described previously (Chichilnisky and Kalmar, 2002; Litke et al., 2004; Frechette et al., 2005). Briefly, eyes were obtained from 12 deeply and terminally anesthetized macaque monkeys (*Macaca mulatta*) used by other experimenters in accordance with institutional guidelines for the care and use of animals. Immediately after enucleation, the anterior portion of the eye and vitreous were removed in room light, and the eye cup was stored in darkness for at least 20 min before dissection, in a bicarbonate buffered Ames' solution (Sigma) bubbled with 95% O_2, 5% CO_2 at 32–34°C, pH 7.4. Under infrared illumination pieces of peripheral retina 3–5 mm in diameter, isolated from the retinal pigment epithelium, were placed flat against a planar array of 512 extracellular microelectrodes, covering an area of ~1800 × 900 μm. Data were obtained from 15–120 min periods of recording. The preparation was perfused with Ames' solution.

The voltage on each electrode was digitized at 20 kHz and stored for off-line analysis. Details of recording methods and spike sorting have been given previously (Litke et al., 2004). Briefly, spikes were identified using a threshold of four times the voltage SD. For each spike, the waveform of the spike and the simultaneous waveforms on six adjacent electrodes were extracted. Three to five waveform features were identified using principal components analysis. A mixture of Gaussians model was fit to the distribution of features using expectation maximization (Duda et al., 2001). The number of clusters and initial conditions for the model was determined automatically using an adapted watershed transformation (Castleman, 1996; Roerdink and Meijster, 2001). All clusters were visually inspected and when necessary, a mixture of Gaussian models was fitted using manually selected initial conditions. Clusters with a large number of refractory period violations (>10% estimated contamination based on refractory period violations), or spike rates <1 Hz, were excluded from further analysis.

Stimulation and receptive field analysis. An optically reduced stimulus from a gamma-corrected cathode ray tube computer display refreshing at 120 Hz was focused on the photoreceptor outer segments. The mean intensity was adjusted to a low photopic light level by including neutral density filters in the light path. The mean photon absorption rate for the

long (middle, short) wavelength sensitive cones was approximately equal to the rate that would have been caused by a spatially uniform monochromatic light of wavelength 561 (530, 430) nm and intensity 4300 (4200, 2400) photons/μm^2/s, incident on the photoreceptors. For the collection of ON-parasol cells shown in Figure 1, *A* and *B*, the mean firing rate during exposure to a steady, spatially uniform display at this light level was 11.8 ± 3.6 Hz and 22.4 ± 4.8 Hz for each preparation respectively. Across all preparations examined firing rates varied from ~5 to 20 Hz.

Spatiotemporal receptive fields were measured using a dynamic white noise stimulus in which the intensity of each display phosphor at each pixel location was selected randomly and independently over space and time from a binary distribution. RMS stimulus contrast was 96%, stimulus duration was 30 min. The pixel size (60 μm in 10 preparations, 96 μm and 120 μm in two others) was selected to accurately capture the spatial structure of parasol cell receptive fields. For each RGC, the spike-triggered average stimulus was computed; this summarizes how the cell integrates visual inputs over space and time (Marmarelis and Naka, 1972; Chichilnisky and Kalmar, 2001). An elliptical two-dimensional Gaussian function was fitted to the spatial profile; outlines in Figure 1 represent the boundary for these fits at the SD values specified in the figure caption. The average of the major and minor axes of the 1 SD contour of each parasol cell was ~200 μm.

Fraction of spikes in multi-neuron firing patterns. The following method was used to calculate the fraction of spikes in each cell that were synchronized with those of any of its immediate neighbors in the mosaic, over and above the chance expectation. For a given cell, consider the observed probability p with which the cell fired simultaneously (±5 ms) with any of its immediate neighbors. Also consider the probability q that this would occur if the neurons fired independently, obtained by multiplying the respective firing probabilities of the cells. The probability that the cell participates in unexpected synchronized firing is thus $p - q$. Dividing by the total firing probability of the cell r yields the fraction of spikes that were part of an unexpected firing pattern, $(p - q)/r$. Across 20 cells in each of four preparations, this value was 31 ± 5%, 22 ± 4%, 15 ± 2%, and 14 ± 4%, respectively (mean ± SD).

Measuring deviations from spatial isotropy. The invariant spatial properties of synchronized firing were quantified by measuring the degree of directional preference in the strength of synchrony. For each cell of interest in the mosaic interior, a polygon was constructed in which the angle and magnitude of each point measures the angle and strength of synchrony with neighboring neurons. The angle and magnitude of these points were uncorrelated ($R^2 = 0.02$, $n = 286$), and the major and minor axes of an ellipse fitted to all the polygons simultaneously differed by <7%, consistent with largely isotropic patterns of synchronized firing. Note that this analysis of cell pairs does not rule out violations of isotropy and translation invariance in higher-order interactions. The observation that firing patterns was isotropic was consistent across all preparations examined.

Identifying adjacent neurons in mosaic. Several statistics as well as the formulation of the pairwise-adjacent model required defining adjacent cells in the ON-parasol mosaic. A nonparametric technique was used to define adjacency based on receptive field center locations. Each location in space was associated with the nearest receptive field center. This definition delineates a collection of boundaries between territories associated with neighboring cells, known as a Voronoi tessellation (Voronoi, 1907; Okabe et al., 2000). Pairs of cells with shared boundaries are connected with a line in Figure 4*A*.

The Voronoi tessellation misidentifies adjacency if a hole exists in the mosaic due to an unrecorded cell (e.g., Fig. 1*A*, bottom left). Likewise, the tessellation misidentifies adjacency between cells along the edge of the mosaic. To deal with the latter problem, a secondary stipulation was added that adjacency requires a shared edge within the convex hull of the collection of receptive field centers.

Maximum entropy. Maximum entropy methods are used in statistical inference to identify an unknown distribution given several constraints which are insufficient to fully specify the answer. The method involves selecting the distribution with the greatest entropy consistent with the constraints,

$$\hat{P}(x) = \arg\max_{P(x)} \left[H[P] + \sum_{i=1}^{N} \lambda_i (E[F_i(x)] - k_i) \right],$$

where $H[P] = -\sum_x P(x)\log_2 P(x)$ is the entropy of a firing pattern distribution $P(x)$, N is the number of constraints, E is the expectation over $P(x)$, $\{F_i\}$ are linear functions of the firing patterns, $\{k_i\}$ are the measured constraints, and $\{\lambda_i\}$ are Lagrange multipliers enforcing the constraints (Jaynes, 1957a,b; Cover and Thomas, 1991; Amari, 2001). In our application the specified constraints are the firing rates expressed as probabilities and the joint firing patterns $P(x_i, x_j)$ for pairs of neurons x_i, x_j that are adjacent. The Lagrange multipliers enforcing each pairwise-adjacent constraint specify the sensitivity of the maximum entropy solution to slight changes to the strength of synchrony (see Fig. 4A, red lines). Note that a model with pairwise-adjacent constraints where a spike (or no spike) is labeled with a +1 (or −1) is identical to a nearest-neighbor Ising model (Schneidman et al., 2006; Tang et al., 2008). The interaction terms of this model (see Fig. 4A) do not directly reflect the strength of synchronized firing between a pair of neurons: for example, interaction terms can propagate correlation to explain observed synchronized firing between nonadjacent cells (Shlens et al., 2006).

A dual form of this problem is to recognize that computing the maximum entropy distribution is equivalent to computing the maximum likelihood estimate of the Lagrange multipliers $\{\lambda_i\}$ given the data (Csiszár, 1975; Berger et al., 1996). Calculating the maximum likelihood estimate of the parameters is a special case of estimating the parameters of a discrete Markov random field (Jordan, 1998) or equivalently, a Boltzmann machine (Hinton and Sejnowski, 1986). This problem is recognized to be log-concave everywhere, meaning that a unique global maximum exists and can be found with techniques from convex optimization (Darroch and Ratcliff, 1972; Press et al., 1988; Berger et al., 1996; Malouf, 2002), simulated annealing (Hinton and Sejnowski, 1986; Hertz et al., 1991), or approximations to the full likelihood (Perez, 1998; Hinton, 2002). In the case of large models, estimating model parameters is an area of active research (Lan et al., 2006).

A drawback of these methods is that most optimization techniques require an explicit calculation of the probability distribution, which in turn requires normalizing across all possible firing patterns. In the case of synchronized firing, the firing pattern distribution across n neurons consists of a set of binary random variables $x = (x_1, x_2, ..., x_n)$ with 2^n elements. Hence, determining the normalization constant across the firing pattern distribution for the entire neural population (e.g., $2^{104} \approx 10^{30}$ patterns) is not practical. The parameters of the pairwise-adjacent model can be estimated by making draws from the unnormalized probability distribution by exploiting Markov Chain Monte Carlo (MCMC) techniques (Kennel et al., 2005; Gamerman and Lopes, 2006). Traditional MCMC sampling can be slow, thus a specialized form of MCMC sampling was used (Swendsen and Wang, 1987). MCMC samples of the estimated distribution were drawn to calculate the likelihood gradient (Perez, 1998). To assure convergence, an adaptive learning procedure was used, which is equivalent to a form of simulated annealing (Jordan, 1998).

Because estimating the model parameters is complex, validation of the results was performed in two ways. First, the convergence of the procedure was examined by checking its ability to reproduce the desired constraints. In the case of the pairwise-adjacent model this amounts to matching the firing rates of individual neurons and the degree of synchronized firing between adjacent neurons. The trained model accurately matched the observed firing rate and strength of synchrony ($R^2 > 0.99$). The remaining error may reflect the convergence criterion selected and statistical variability between the training and testing data sets. Second, the overall ability of the learning procedure to correctly infer the parameters of a truly pairwise-adjacent distribution was tested. Thirty minutes worth of samples were generated from a pairwise-adjacent model, and a new pairwise-adjacent model was estimated on these simulated data samples. The newly trained model recovered an unbiased estimate of the parameters $\{\lambda_i\}$ from the simulated data ($R^2 > 0.99$), indicating that the learning procedure correctly inferred the appropriate model structure.

Likelihood and information theoretic analysis. Models were compared using likelihood analysis. Details of this analysis have been given previ-

ously (Shlens et al., 2006). Briefly, the likelihood of a data set under a given model is the probability of having observed that data set under the assumption that the model is correct. For each of the statistics examined (see Fig. 3), the number of observations was counted, producing a histogram $c = \{c_i\}$, where $m = \sum_i c_i$ is the total number of measurements. This produces an estimate of the probability of each value of the statistic, $P = \{p_i\}$, where $p_i \equiv c_i/m$. For a given model $Q = \{q_i\}$, the probability of having observed the histogram counts c if the model Q actually generated the observations is the multinomial likelihood (Duda et al., 2001)

$$L \propto \prod_i q_i^{c_i}.$$

The likelihood shrinks multiplicatively with increasing measurements (larger m). A quantity that is invariant to the number of measurements is the average likelihood $\bar{L} = L^{1/m}$. The Kullback–Leibler divergence from information theory (D_{KL}) bears a simple relationship to the average likelihood:

$$D_{KL}(P \| Q) = -\log_2 \bar{L}$$

in the limit of infinite measurements ($m \to \infty$). Estimating the KL divergence in finite data is nontrivial and subject to severe biases, especially in high-dimensional data. To address this issue, techniques for estimating entropy efficiently were used (Krichevsky and Trofimov, 1981; Nemenman et al., 2002; Paninski, 2003; Orlitsky et al., 2004).

Time scale of analysis. The selection of time scale (or bin size) for defining synchronized firing could influence the results. Ideally, the choice of bin size should match the width of the cross-correlation peak [~10 ms; Shlens et al. (2006), their Fig. 2B], and encompass a significant fraction of the refractory period, to accurately reflect the strength of correlation while avoiding statistical dependencies associated with refractoriness. Selecting too small a bin size could produce a poor estimate of correlation strength due to limited counting statistics. Selecting too large a bin size could mask synchrony by averaging over times outside the window of synchronization.

A twofold change in the bin size did not significantly affect the results. For the preparation of Figure 3A (top panels), using bin sizes of 5 and 20 ms, the pairwise-adjacent model accounted for, respectively, (99.2%, 99.4%, 99.5%) and (99.0%, 98.2%, 98.6%) of the departures from independence in the three statistics examined, similar to the results obtained with a 10 ms bin size (see Fig. 3A, see insets). The fraction of firing entropy of an individual RGC predictable from its neighbors was also little affected by bin size. For the retina in Figure 4A, using bin sizes of 5 and 20 ms, $7.3 \pm 0.9\%$ and $5.1 \pm 1.1\%$ of firing entropy, respectively, were predictable from the firing of neighboring cells. Similar results were obtained in the presence of visual stimulation.

Failures of the pairwise-adjacent model. The pairwise-adjacent model failed to explain ~1–2% of the deviations from statistical independence (see Fig. 3). Several potential sources of the deviations were examined.

First, when the waveforms of spikes in different neurons overlap in time, the recorded (summed) voltage waveform differs from the waveform produced by either cell alone, potentially causing spikes from both cells to be missed in the spike sorting procedure and leading to an underestimate of synchronized firing events. To test for this possibility, the following control procedure was performed. Thirty minutes of data was sampled from a pairwise-adjacent model fitted to recorded data. Then artifacts were introduced that mimicked the hypothetical spike sorting bias. In a first test, ~5% of pairwise synchronous events between adjacent neurons were removed. In a second test, 5% of triplet synchronous events between adjacent neurons were removed. A new pairwise-adjacent model was then fitted to and compared with the altered data sets. In both manipulations, no systematic biases were found in the three statistics of interest (see Fig. 3) and the model accounted for 99.3, 99.6, and 99.6% of the failures of statistical independence. Thus, artifacts associated with the spike sorting procedure are unlikely to entirely account for the observed failures of the model.

Second, to assess whether the failures were attributable to nonstationarity or finite counting statistics, the accuracy of the model was compared with that of a second firing pattern distribution (P_{emp}) that consists of the
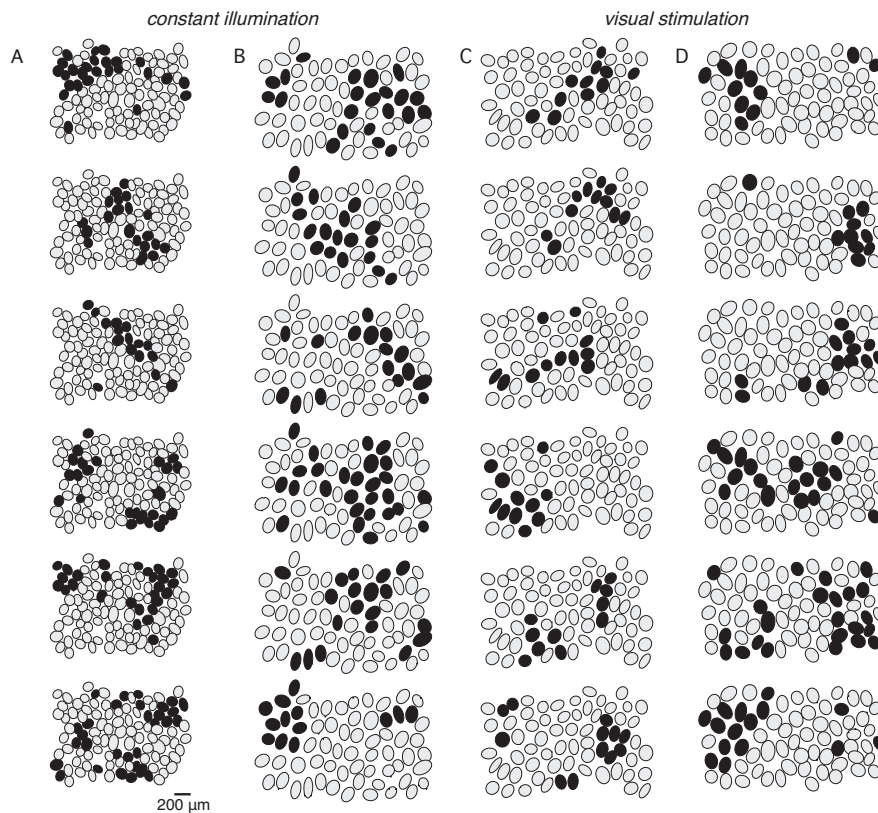
constant illumination   *visual stimulation*

200 μm

Figure 1. Islands of large-scale synchronized firing in populations of primate ON-parasol RGCs across four preparations (**A–D**). Each oval represents the 0.9, 1.0, 1.0, and 1.0 SD outline of the Gaussian fit to the receptive field of a single cell in each preparation, respectively. If a neuron spikes within a 10 ms interval, the receptive field is colored black. Shown are six selected frames to highlight the spatial scale of synchronized firing from each preparation under constant illumination (**A, B**; supplemental Movie 1, available at www.jneurosci.org as supplemental material) and visual stimulation (**C, D**; supplemental Movie 2, available at www.jneurosci.org as supplemental material).

firing probabilities estimated directly from an independent period of recording from the same neurons. The accuracy of P_{emp} as a model for P_{obs} reveals the reproducibility of the experiment. P_{emp} always accounted for at least 99.5% of the deviations from statistical independence. Thus, the failures of the model cannot be explained entirely by counting statistics or nonstationarity of the recordings.

Third, the possibility of systematic deviations between the model and data was examined in more detail (see Fig. 3, bottom rows). In 11 of the 12 preparations examined, the model did not systematically overpredict or underpredict any of the three statistics examined. In one preparation (Fig. 1A), the model overpredicted the frequency of large firing patterns but not small ones, and the overprediction grew systematically with size of the firing pattern. Thus, although the results were not entirely consistent across data sets, systematic deviations could not explain the failures of the model in all data sets.

Finally, a potential source of the discrepancy between data and model is an implicit assumption that was made about the temporal structure of the data: that the sequential time samples from the recording are independent observations. This assumption does not take into account interactions over time in the firing of cells. A notable temporal interaction is action potential refractoriness, which reduces the probability that a neuron will generate a spike for a period of several milliseconds after the occurrence of a spike. For temporal interactions to have a significant effect on the results, the distribution of firing statistics must depend substantially on the pattern of activity in preceding time samples. To test this, time samples were divided into two groups: those for which the preceding time sample contained more than the median number of neurons firing, and those for which the preceding time sample contained fewer than the median. Each of the three statistics was then examined separately for the two groups of firing patterns. The effect of the preced-

ing time sample was large. For example, in the preparation in Figure 1B, the Kullback–Leibler divergence between the full distribution of numbers of neurons firing and the group with high and low preceding firing was 1.74×10^{-4} and 1.30×10^{-4} bits/neuron, respectively. These values represent 2.4% and 3.2% of the magnitude of the departures from statistical independence in the data, respectively. Similar departures were observed for the other statistics examined. Thus, the effect of temporal interactions on the measured firing pattern statistics was potentially large enough to account for the observed failures of the model.

Results

Action potentials were recorded extracellularly from RGCs in isolated macaque monkey retina perfused with physiological saline solution, using an array of 512 recording electrodes (Litke et al., 2004; Frechette et al., 2005). The receptive field of each RGC was identified using reverse correlation with a white noise stimulus (Chichilnisky and Kalmar, 2001). Cells were segregated into distinct functional classes according to their receptive field characteristics. ON-parasol cells, which project to the magnocellular layers of the lateral geniculate nucleus, were identified by their distinctive receptive field size, density, and light response kinetics (Chichilnisky and Kalmar, 2002; Field et al., 2007). As expected from previous work, the receptive fields of ON-parasol cells formed a regular lattice, or mosaic, uniformly sampling the region of retina recorded (Fig. 1) (Devries and Baylor, 1997; Chichilnisky and Kalmar, 2002; Frechette et al., 2005). The completeness of the recorded mosaics indicated that nearly every ON-parasol cell in this 4 × 8 degree region of retina was recorded in each of 12 preparations ($n = 104, 48, 56, 56, 58, 66, 68, 75, 81, 66, 65,$ and 54 neurons). Subsets of one of these data sets were used in previous work (Shlens et al., 2006).

Synchronized firing in large populations

To examine multi-neuron synchronized firing, the activity of the entire population of ON-parasol cells was recorded in the presence of spatially uniform, steady photopic illumination. The prevalence of synchronized firing was determined by calculating the fraction of spikes in each cell that were synchronized with those of any of its immediate neighbors in the mosaic, over and above chance expectation (see Materials and Methods). Across four preparations this value ranged from 14% to 31%. Thus, a substantial fraction of the spikes produced by any given ON-parasol cell reflect synchronized firing.

To examine the spatial structure of multi-neuron synchronized firing, movies were generated in which each frame represented the activity of all cells in a 10 ms time bin. The receptive field of each cell was filled if the cell fired a spike (supplemental Movie 1, available at www.jneurosci.org as supplemental material). Synchronized firing was examined at the 10 ms time scale because this time scale captured much of the autocorrelation within a cell and cross-correlation between cells (42% and 83% of
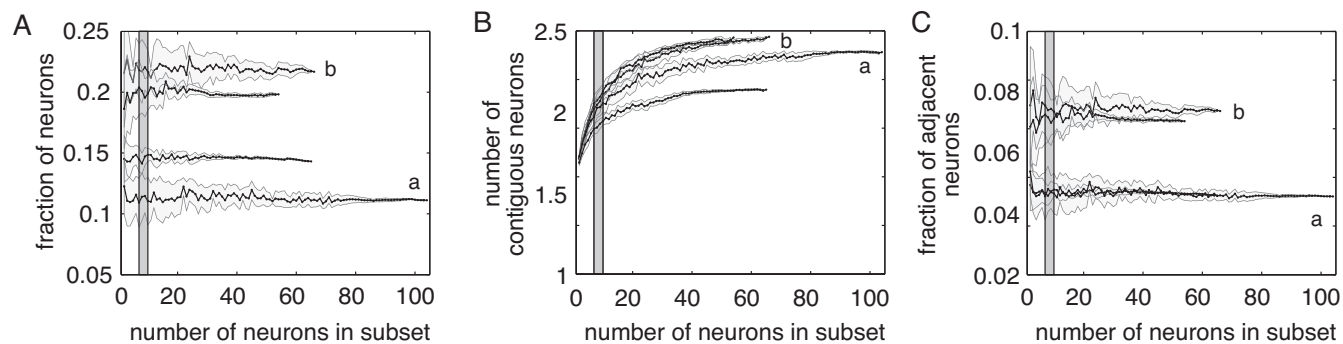
Figure 2. Subsampling analysis reveals that 512 electrode recordings capture scale of synchronized firing. Complete mosaic of ON-parasol cells in four preparations (including Fig. 1 *A, B*; see letter labels) is systematically enlarged from center. Black (gray) lines represent mean (± 1 SD) across all observed firing patterns. *A*, Fraction of neurons firing in subsampled population. *B*, Number of neurons constituting an island of contiguous synchronized firing in subsampled population. *C*, Fraction of pairs of adjacent cells firing in subsampled population (see Results). Gray bar represents size of multi-neuron firing patterns examined previously (Schneidman et al., 2006; Shlens et al., 2006).

variance, respectively; see Materials and Methods for expanded discussion of time scale) (Shlens et al., 2006). Several snapshots of activity are shown for two recordings in Figure 1, *A* and *B*; these data were selected to emphasize synchronized firing. Synchronized firing occurred at spatial scales substantially larger than the distance between nearest neighbors in the mosaic. Specifically, synchronized firing events including up to 44% of the population, and up to 26 spatially contiguous neurons, were observed in four preparations examined. In one preparation (Fig. 1 *A*), firing patterns of this magnitude were observed every 30 min. If the cells fired independently, such patterns would occur only once every ~6000 years. Therefore, the observed patterns of activity clearly deviate from the predictions of independence, and significantly exceed the scale of multi-neuron synchronized firing previously reported in the retina (Schnitzer and Meister, 2003; Schneidman et al., 2006; Shlens et al., 2006).

Some regularities in the patterns of synchronized firing are apparent from visual inspection (Fig. 1 *A, B*; supplemental Movie 1, available at www.jneurosci.org as supplemental material). First, as expected from previous work (Shlens et al., 2006), synchronized firing was spatially localized: active cells tended to be clustered in a local region, the size of which covaried with the average receptive field size of the ON-parasol cells recorded. Second, synchronized firing was spatially isotropic: the strength of synchrony was approximately invariant to the orientation of the line joining a cell pair, exhibiting deviations from circular symmetry no larger than 7% (see Materials and Methods; data not shown). Third, synchrony was approximately translation invariant: its spatial extent was approximately independent of the location of the recorded cells: no distinct boundaries or stereotyped motifs appeared in recorded firing patterns, and each cell was equally likely to exhibit synchronized firing with each of its neighbors. These features of synchronized firing were observed consistently in all preparations examined.

Measuring the complete pattern of activity

The above findings suggest that synchronized firing in ON-parasol cells may be summarized parsimoniously, provided that the measurements encompass a sufficiently large region of retina. The sufficiency of the recording was tested by examining three statistics of synchronized firing. The first statistic was the fraction of neurons firing in a given time sample. For example, in the bottom panel of Figure 1 *B*, a total of 13 neurons fired (20%). The second statistic was the number of firing neurons in spatially contiguous groups consisting of more than one neuron (see Ma-

terials and Methods). For example, in the bottom panel of Figure 1 *B*, two contiguous firing groups were observed, sizes 10 and 3. The third statistic was the fraction of pairs of adjacent firing cells. For example, in the bottom panel of Figure 1 *B*, there were 18 pairs of adjacent firing cells (11%). [Note that a more complete approach to assess synchronized firing would be to estimate the frequency of all 2^n possible firing patterns in n neurons (Shlens et al., 2006), which is impractical when n is in the range of 50–100 as in the present data.]

If the present recordings were sufficient to capture the full extent of synchronized firing, then the values of these statistics should reach asymptotes with a collection of cells no larger than the total number of cells recorded. To test this prediction, these statistics were calculated using spatially contiguous subsets of recorded cells in four preparations (Fig. 2). For small groups of cells similar to those examined in previous work (Schneidman et al., 2006; Shlens et al., 2006) (e.g., 7–10 neurons), the three statistics either exhibited large variance or assumed mean values significantly different from the values observed with the entire population (gray bar). However, these statistics approached asymptotic values, with gradually shrinking variance, with ~40 cells.

Large synchronized firing patterns appeared not to represent singular events, but instead samples from a continuum, as shown by the distributions of the three statistics in Figure 3 *A* (top row, gray bars). These distributions also reveal the spatial extent of synchronized firing. For example, in the recording of Figure 1 *A*, for >99.9% of the recorded patterns, no more than 25 neurons (40%), and no more than 15 spatially contiguous neurons (23%), fired synchronously. Finally, the three statistics may be used to summarize quantitatively the departures from statistical independence in the population (Fig. 3 *A*, top row, dashed curves; supplemental Movie 4, available at www.jneurosci.org as supplemental material). The large observed deviations reflect the prominence of synchronized firing in the population.

Pairwise-adjacent interactions largely explain synchronized firing

Can the observed patterns of synchronized firing be explained with a simple model? One possibility is that they can be understood entirely on the basis of measured interactions between pairs of cells (Schneidman et al., 2006; Shlens et al., 2006; Tang et al., 2008), or (most simply) between pairs of adjacent cells in the mosaic (Shlens et al., 2006). This hypothesis was tested using the observed firing pattern statistics (Fig. 3 *A*, top row).
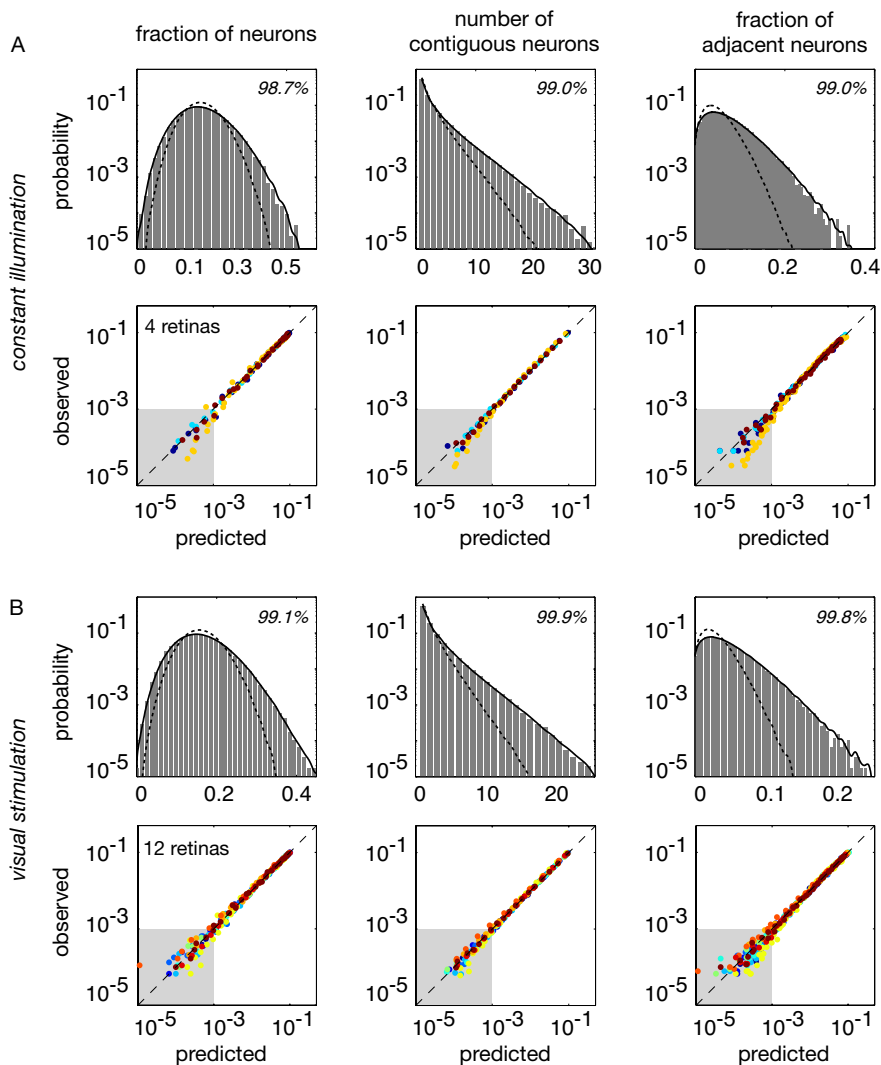
Figure 3. Large-scale statistical features of neural populations under constant illumination (***A***) and visual stimulation (***B***). ***A***, Top row, Distribution of fraction of neurons firing (left), number of contiguous neurons firing (middle), and fraction of adjacent neurons firing in cross-validated data sets (bars) and predicted from statistical independence (dashed line) and pairwise-adjacent model (solid line). Data set is the same as in Figure 1*B*. Bottom row, Summary comparison between observed probability and prediction from pairwise-adjacent model across four retinas. Each point represents a firing pattern and each color corresponds to one of four retinas. The gray boxes highlight infrequent, large firing patterns encompassing approximately >40% of neurons firing, >15 contiguous neurons firing, or >15% of adjacent cells firing, respectively. ***B***, Same as ***A***, but in the presence of visual stimulation. Top row, The data set from Figure 1*D*. Bottom row, Summary across 12 retinas. Note logarithmic axes.

To generate predictions of population responses based purely on pairwise and adjacent interactions, the maximum entropy methodology developed in recent work was used (Schneidman et al., 2003b, 2006; Shlens et al., 2006). To introduce the approach, let $x_i \in \{1, 0\}$ indicate whether neuron i fires (or not) during a single time bin, and let $P_{obs}(x_i)$ indicate the probability that it fires (or not). Similarly, let $P_{obs}(x_i, x_j)$ represent the joint probability of two cells each firing (or not), and let $P_{obs}(x_1, \ldots, x_n)$ indicate the joint probability of any given pattern of firing in the entire population of cells. The goal is to produce and test a model of $P_{obs}(x_1, \ldots, x_n)$, solely from the knowledge of $P_{obs}(x_i)$ and $P_{obs}(x_i, x_j)$, assuming no further structure in the population activity. A natural solution, borrowed from statistical mechanics (Jaynes, 1957a,b), is to select a pairwise model $P_{model}(x_1, \ldots, x_n)$ which has maximum entropy, or randomness, subject to the constraints $P_{model}(x_i) = P_{obs}(x_i)$ and $P_{model}(x_i, x_j) = P_{obs}(x_i, x_j)$ (see Materials and Methods). A simpler model is one in which the constraints

$P_{obs}(x_i, x_j)$ are only obtained from immediate neighbors in the mosaic; this will be referred to as the pairwise-adjacent model (Shlens et al., 2006). A visualization of the pairwise-adjacent model fitted to data from the preparation in Figure 1*B* is given in Figure 4*A* (see Materials and Methods).

To test the model, simulated firing patterns were generated and compared with the data. Qualitatively, the accuracy of the model predictions may be assessed by comparing movies generated from the model and data (supplemental Movies 1, 3, available at www.jneurosci.org as supplemental material). Quantitatively, the firing patterns produced by the model closely matched all three measured statistics of the data (Fig. 3*A*, top row, solid curves). Strikingly, the model correctly predicted the frequency of spatially contiguous islands of up to at least 30 firing neurons, even though model parameters were constrained only by the firing probabilities of pairs of adjacent cells. Across four preparations, the observed frequency of firing patterns matched that predicted from the fitted pairwise-adjacent model (Fig. 3*A*, bottom row). Large events encompassing >40% of neurons firing or >15 contiguous neurons firing (Fig. 3*A*, bottom row, gray box) occurred infrequently and were subject to large counting variability. Despite this variability, the pairwise-adjacent model produced an unbiased prediction of each firing pattern across preparations.

To summarize these findings, the similarity of the distributions of the three statistics for the model (P_{model}) and the data (P_{obs}) was measured using likelihood analysis. The likelihood associated with a particular model P_{model} is the probability of observing the data, given the model. The negative logarithm of the average likelihood asymptotically approaches the Kullback–Leibler divergence $D_{model} \equiv D_{KL}(P_{obs}\|P_{model})$, and expresses the similarity of the data and model distributions (see Materials and Methods). The fraction of deviations from statistical independence accounted for by a particular model, $F = 1 - D_{model}/D_{ind}$, measures the success of the model in explaining synchronized firing. The value of F was computed for the data across four preparations (including Fig. 1*A*,*B*), and for each of the three statistics. For the data in Figure 3*A*, the values of F are given inset. In all cases the pairwise-adjacent model accounted for 98–99% of the deviations from statistical independence observed (see Materials and Methods for an analysis of the time scales and deviations not accounted for by the model).

Measuring the influence of the network on each cell
The accurate fit of the pairwise-adjacent model made possible an analysis of the influence of shared signals on the activity of each cell. A natural measure of shared signals is the information that the activity of the surrounding population contains about the

activity of the reference cell. Specifically, the information about the activity of a reference cell, y, contained in the activity of its neighbors, $x_1, \ldots, x_n$, is measured by the reduction in uncertainty about the activity of the reference cell obtained from knowledge of the neighbor activity, $I(y; x_1, \ldots, x_n) = H(y) - H(y \mid x_1, \ldots, x_n)$, where $H(\cdot)$ represents the entropy, or uncertainty, of a random variable. The pairwise-adjacent model makes possible an unbiased estimate, because it permits the generation of requisitely large simulated data sets exhibiting approximately the same statistics as the data.

First, the information that the activity of a single neighboring cell contained about the activity of a reference cell was estimated, i.e., $I(y; x_1)$. For five example cells, this quantity was small, reflecting weak pairwise coupling (Fig. 4B, leftmost points). Next, the analysis was repeated for increasing numbers of neighbors, accumulated in order of distance, up to $n = 15$ cells. For cells in Figure 4B, the amount of information in the surrounding activity increased with the number of neighbors considered, reaching an asymptotic maximum of 0.06 ± 0.01 bits with 15 cells. Across all reference cells in this data set (excluding cells with fewer than six recorded neighbors, to avoid edge effects), the total information about the activity of the reference cell contained in its 15 nearest neighbors was 0.056 ± 0.0087 bits. In three other preparations this value was 0.094 ± 0.025 bits, 0.08 ± 0.011 bits, and 0.039 ± 0.01 bits (Fig. 4C). Just six immediate neighbors provided the large majority (90% ± 7.4%) of the information about the firing of the reference cell (Fig. 4C).

To assess the relative strength of network interactions, the information $I(y; x_1, \ldots, x_n)$ was compared with the entropy of the firing of the reference cell, $H(y)$. For the preparation in Figure 4A, this comparison is plotted with red points in Figure 4D. Note that the variation in entropy across reference cells solely reflects the variation in firing rates of the individual cells. A best fit line (data not shown) indicates that the information was 7.0 ± 1.1% of the entropy of the firing of the reference cell (see Materials and Methods for analysis across bin sizes). Across three other preparations, the network activity accounted for 18.9 ± 3.2%, 11.2 ± 1.7%, and 6.5 ± 1.5% of the entropy. In summary, no more than 20% of the uncertainty in the firing of an individual cell was accounted for by the activity of the surrounding population (Fig. 4D, dashed line).

Synchronized firing in the presence of visual signals

The above results refer only to synchronized firing in the spontaneous activity of RGCs. However, the relevance of synchronized firing for retinal function ultimately depends on its role in visual signaling to the brain, and it is possible that the influence of the network on an individual cell changes when RGC firing is modulated by a visual stimulus (Fig. 1C,D; supplemental Movie 2, available at www.jneurosci.org as supplemental material). To test this possibility, the pairwise-adjacent model was fitted to data obtained in the presence of a white noise stimulus, in which the
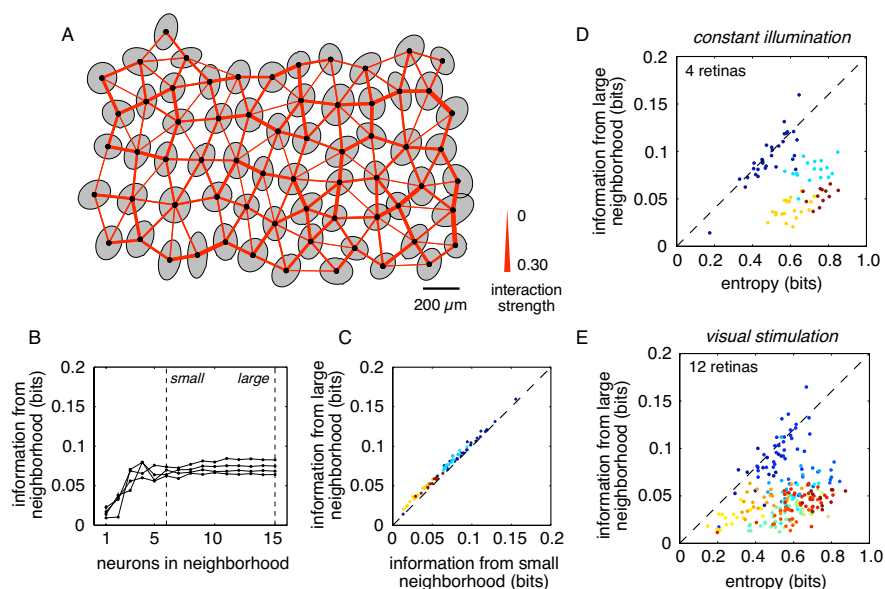

Figure 4. Strength of network interactions. ***A***, Visualization of inferred network architecture in pairwise-adjacent model for Figure 1B. Thickness of red lines indicates the interaction strength (see Materials and Methods). ***B***, Information about firing of four example cells from ***A*** available in the surrounding neighborhood as a function of the neighborhood size. ***C***, Comparison of the information about the firing of a single cell contained in small (6 neurons) and large (15 neurons) neighborhoods. Each point represents a single cell and each color corresponds to one of four preparations. ***D***, Comparison of the amount of information about the firing of a single cell in a large neighborhood to the entropy of the individual cell, under constant illumination. Dashed line indicates a slope of 0.2. Colors are the same as in ***C***. ***E***, Same as ***D***, but under white noise visual stimulation. Each color corresponds to 1 of 12 different retinas.

intensity and color of each pixel of the display varied randomly and independently over time. The stimulus was strong enough to drive firing in the cells: it was used to extract the receptive fields shown in Figure 1. However, the pixels were small enough that they generated no more than 8.4% additional synchronized firing in adjacent cells [as determined by shuffle correction (Perkel et al., 1967; Palm et al., 1988)]. Thus, this stimulus provided a test of whether visually modulated activity that is approximately independent in each cell alters the relative influence of the network.

Figures 3B and 4E show the results from 12 retinas. The pairwise-adjacent model accurately predicted the frequency of firing patterns in the retina from Figure 1D (Fig. 3B, top row) as well as 11 other retinas (Fig. 3B, bottom row). Importantly, the pairwise-adjacent model predicted the frequency of large-scale firing events (>15 contiguous neurons) (Fig. 3B, bottom row, gray box) and accounted for 98.0–99.8% of the deviations from statistical independence. For the same four preparations in Figure 4D, the activity in the neighbors accounted for 6.3 ± 1.0%, 18.6 ± 3.3%, 9.5 ± 1.5%, and 5.1 ± 1.4% of the activity of the reference cell, respectively (Fig. 4E). Thus, as the visual function of RGCs was engaged by independent stimulus modulations, the pairwise-adjacent model remained accurate and the relative influence of shared and unique signals in each cell remained roughly fixed.

Discussion

The present findings reveal the large-scale structure of multi-neuron synchronized firing in a population of RGCs that transmits a complete visual representation to the brain. ON-parasol cells fired synchronized spikes in large, spatially contiguous groups, and 15–30% of spikes in each cell were associated with multi-neuron synchronized firing not expected by chance. However, a simple model, based on pairwise interactions between

immediate neighbors in the population, accounted for 99% of the spatial scale and structure of synchronized firing. These neighbor interactions accurately predicted the frequencies of the largest patterns of electrical activity measured. Despite strong synchronized firing, the activity of any given cell was only weakly redundant: no more than 20% of the firing of each cell was predictable from the activity of its neighbors. These findings were virtually unchanged by independent visual stimulation of the cells. In summary, synchronized firing in a complete population of ON-parasol cells appears to reflect simple neighbor interactions, rather than a unique collective visual signal or a highly redundant coding scheme.

Synchronized firing in the retina

A previous finding in larval tiger salamander retina indicated that RGCs often fire in large groups, raising the possibility that synchronized firing patterns represent a unique, elementary symbol in the neural code of the retina (Schnitzer and Meister, 2003). However, because that study lacked information about the spatial structure and cell type organization of synchronized firing, it left open the question of what the unique symbol could be. Other studies have also examined multi-neuron synchronized firing, but at a scale significantly smaller than the patterns of activity in the neural population, providing an incomplete picture (Schneidman et al., 2006; Shlens et al., 2006).

The present work characterizes the spatial "footprint" of synchronized firing in one cell type at a sufficiently large scale to capture all the apparent structure. The data did not reveal singular and unique synchronized firing events in the ON-parasol population, nor any particular tendency toward large-scale synchronization. Instead, the data suggest that large-scale synchronized firing events form part of a continuum of patterns of collective electrical activity produced by neighbor interactions.

Redundancy in a complete neural population

Although synchronized firing was strong and structured, the activity of each ON-parasol cell was only weakly redundant with the activity of surrounding ON-parasol cells. This finding relates to two predictions from previous studies.

One study of RGCs in larval salamander retina predicted, based on an extrapolation from small groups of cells, that interactions among all RGCs in a small region of the retina could produce highly ordered or "freezing" behavior (Schneidman et al., 2006), a phenomenon not observed in the present work. However, two important differences in the studies bear mention. First, the previous study was performed using stimuli with substantial spatial redundancy, which would be expected to elevate synchronized firing [for discussion, see Bethge and Berens (2007), Nirenberg and Victor (2007), and Roudi et al. (2009)]. Second, the extrapolation from the data to the complete RGC population was based on the approximation that correlated activity between pairs of cells is independent of distance within the recorded region. This approximation implies that the total interaction strength would grow as the square of the number of cells. In the present work, examining an essentially complete collection of cells uniformly tiling a large region of retina, correlated activity declines systematically with distance (Shlens et al., 2006) and is captured by local interactions between neighboring cells. Thus the total interaction strength scales linearly with the number of cells, and the prediction of what could occur in a fully connected local population (Schneidman et al., 2006) remains untested.

In a similar vein, a second study of larval salamander RGCs revealed substantial (up to ~40%) redundancy in the visual sig-

nals carried by pairs of RGCs, particularly in cell pairs with highly overlapping receptive fields (Puchalla et al., 2005). In principle, this high pairwise redundancy could produce even higher redundancy in a full population of such cells. In contrast, the present results show that the information carried about the activity of an ON-parasol cell by all the surrounding ON-parasol cells is no more than ~20%. This value provides a bound on the visual signal redundancy in the entire ON-parasol population [Schneidman et al. (2003a), their Eq. 11]. However, it does not provide a measure of redundancy across all RGC populations in a local region of retina.

A simple hypothesis may reconcile the present and previous findings: firing and visual signals are largely nonredundant among RGCs of a single type, which exhibit limited receptive field overlap, but highly redundant across all RGCs of different types in a small region of retina, which exhibit substantial receptive field overlap. This would be predicted if the dominant source of noise in the retina were photoreceptor noise (see below), because RGCs with overlapping receptive fields presumably sample from the same photoreceptors.

It is worth noting that different RGC types project to distinct targets in the brain and probably subserve distinct visual functions (Berson, 2008). Therefore, high redundancy between different cell types in a local region may reveal little about the functional impact of synchronized firing, whereas low redundancy within a cell type over an extended region may be relevant for understanding downstream computations.

Circuits mediating synchronized firing

Previous studies have revealed that synchronized firing in RGCs of several species arises from a combination of gap junction coupling and common synaptic input (Mastronarde, 1989; Dacey and Brace, 1992; Jacoby et al., 1996; Stafford and Dacey, 1997; Brivanlou et al., 1998; Hu and Bloomfield, 2003; Hidaka et al., 2004). In principle, either of these mechanisms could mediate synchronized firing observed here. Recent work using paired intracellular recordings indicates that gap junction coupling [potentially through amacrine cells (Dacey and Brace, 1992)] and correlated synaptic input contribute to synchronized firing among ON-parasol cells, although the latter dominates (Trong and Rieke, 2008). Although the present work provides no additional insight into mechanism, it does suggest that the mechanisms underlying synchronized firing in pairs of neighboring ON-parasol cells may suffice to explain the patterns of activity in the entire network. One possibility is that a single noise source, namely photoreceptor noise, dominates both firing variability in individual RGCs and shared variability in neighboring RGCs (Trong and Rieke, 2008; F. Rieke, personal communication).

It remains possible that novel stimuli could activate distinct retinal circuits that would create more complex patterns of RGC activity than would be produced by neighbor interactions. For example, recent work has shown that looming stimuli (Ishikane et al., 2005) and reversing moving stimuli (Schwartz et al., 2007) generate unique patterns of synchronized firing in frog and salamander RGCs, respectively. Whether such effects are present in the ON-parasol population remains to be seen.

Future directions and relevance for other neural circuits

The present results highlight the value of statistical approaches to probe the effect of synchronized firing on visual signals. Although the maximum entropy approach has the benefit of relying on few assumptions, its utility may be limited because of the large data sets required to probe firing dynamics and stimulus dependencies. An

alternative approach is a parametric model of light response that includes interactions in the spiking activity of adjacent cells and also provides an explicit, optimal decoding algorithm for the neural signals (Pillow et al., 2008). This model may be useful for probing the temporal structure of synchronized firing and its role in visual signaling.

The regularity of the observed firing patterns made it possible to summarize their structure and scale with a few statistics and a simple model. Previous findings suggest that such a summary of network interactions may prove more difficult in cortical circuits (Callaway, 1998; Yoshimura and Callaway, 2005; Yoshimura et al., 2005), where the anatomical and functional organization may not be as simple or as precisely registered as in the retina. However, it remains possible that the functional organization of at least some central circuits is simple, even though their anatomical arrangement is complicated by factors such as multiple overlaid maps of sensory variables [e.g., primary visual cortex (Ohki et al., 2005, 2006)]. Furthermore, increasing knowledge about cell type organization in the brain (Callaway, 2005; Göbel et al., 2007; Mitchell et al., 2007) might reveal a more regular structure in patterned firing within each defined neural population.

References

Ackert JM, Wu SH, Lee JC, Abrams J, Hu EH, Perlman I, Bloomfield SA (2006) Light-induced changes in spike synchronization between coupled ON direction selective ganglion cells in the mammalian retina. J Neurosci 26:4206–4215.

Amari S (2001) Information geometry on hierarchy of probability distributions. IEEE Trans Inf Theory 47:1701–1711.

Amthor FR, Tootle JS, Grzywacz NM (2005) Stimulus-dependent correlated firing in directionally selective retinal ganglion cells. Vis Neurosci 22:769–787.

Arnett D, Spraker TE (1981) Cross-correlation analysis of the maintained discharge of rabbit retinal ganglion cells. J Physiol 317:29–47.

Arnett DW (1978) Statistical dependence between neighboring retinal ganglion cells in goldfish. Exp Brain Res 32:49–53.

Baccus SA (2007) Timing and computation in inner retinal circuitry. Annu Rev Physiol 69:271–290.

Baccus SA, Meister M (2002) Fast and slow contrast adaptation in retinal circuitry. Neuron 36:909–919.

Berger A, Della Pietra S, Della Pietra C (1996) A maximum entropy approach to natural language processing. Comput Linguist 22:1–36.

Berson D (2008) Retinal ganglion cell types and their central projections. In: The senses: a comprehensive reference (Masland R, ed). Oxford, UK: Elsevier.

Bethge M, Berens P (2007) Near-maximum entropy models for binary neural representations of natural images. Adv Neural Inf Process Syst 20:1–8.

Brivanlou IH, Warland DK, Meister M (1998) Mechanisms of concerted firing among retinal ganglion cells. Neuron 20:527–539.

Callaway EM (1998) Local circuits in primary visual cortex of the macaque monkey. Annu Rev Neurosci 21:47–74.

Callaway EM (2005) A molecular and genetic arsenal for systems neuroscience. Trends Neurosci 28:196–201.

Castleman KR (1996) Digital image processing. Englewood Cliffs, NJ: Prentice Hall.

Chichilnisky EJ (2001) A simple white noise analysis of neuronal light responses. Network 12:199–213.

Chichilnisky EJ, Kalmar RS (2002) Functional asymmetries in ON and OFF ganglion cells of primate retina. J Neurosci 22:2737–2747.

Cover T, Thomas J (1991) Elements of information theory (Wiley Series in Telecommunications). New York: Wiley.

Csiszár I (1975) I-divergence geometry of probability distributions and minimization problems. Ann Probab 3:146–158.

Dacey DM, Brace S (1992) A coupled network for parasol but not midget ganglion cells in the primate retina. Vis Neurosci 9:279–290.

Darroch J, Ratcliff D (1972) Generalized iterative scaling for log-linear models. Ann Math Stat 43:1470–1480.

DeVries SH (1999) Correlated firing in rabbit retinal ganglion cells. J Neurophysiol 81:908–920.

Devries SH, Baylor DA (1997) Mosaic arrangement of ganglion cell receptive fields in rabbit retina. J Neurophysiol 78:2048–2060.

Duda R, Hart P, Stork D (2001) Pattern classification. New York: Wiley.

Field GD, Sher A, Gauthier JL, Greschner M, Shlens J, Litke AM, Chichilnisky EJ (2007) Spatial properties and functional organization of small bistratified ganglion cells in primate retina. J Neurosci 27:13261–13272.

Frechette ES, Sher A, Grivich MI, Petrusca D, Litke AM, Chichilnisky EJ (2005) Fidelity of the ensemble code for visual motion in primate retina. J Neurophysiol 94:119–135.

Gamerman D, Lopes HF (2006) Markov chain Monte Carlo: stochastic simulation for Bayesian inference, Ed 2 (Texts in statistical science series). Boca Raton, FL: Chapman and Hall/CRC.

Göbel W, Kampa BM, Helmchen F (2007) Imaging cellular network dynamics in three dimensions using fast 3D laser scanning. Nat Methods 4:73–79.

Harris KD (2005) Neural signatures of cell assembly organization. Nat Rev Neurosci 6:399–407.

Hertz J, Krogh A, Palmer RG (1991) Introduction to the theory of neural computation. Redwood City, CA: Addison-Wesley.

Hidaka S, Akahori Y, Kurosawa Y (2004) Dendrodendritic electrical synapses between mammalian retinal ganglion cells. J Neurosci 24:10553–10567.

Hinton G, Sejnowski T (1986) Learning and relearning in Boltzmann machines. In: Parallel distributed processing: explorations in the microstructure of cognition, Vol 1, Foundations (Rumelhart D, McClelland J, eds), pp 45–76. Cambridge, MA: MIT.

Hinton GE (2002) Training products of experts by minimizing contrastive divergence. Neural Comput 14:1771–1800.

Hu EH, Bloomfield SA (2003) Gap junctional coupling underlies the short-latency spike synchrony of retinal alpha ganglion cells. J Neurosci 23:6768–6777.

Ishikane H, Gangi M, Honda S, Tachibana M (2005) Synchronized retinal oscillations encode essential information for escape behavior in frogs. Nat Neurosci 8:1087–1095.

Jacoby R, Stafford D, Kouyama N, Marshak D (1996) Synaptic inputs to ON parasol ganglion cells in the primate retina. J Neurosci 16:8041–8056.

Jaynes E (1957a) Information theory and statistical mechanics. Phys Rev 106:620–630.

Jaynes E (1957b) Information theory and statistical mechanics II. Phys Rev 108:171–190.

Jordan MI (1998) Learning in graphical models (adaptive computation and machine learning). Cambridge, MA: MIT.

Kennel MB, Shlens J, Abarbanel HD, Chichilnisky EJ (2005) Estimating entropy rates with Bayesian confidence intervals. Neural Comput 17:1531–1576.

Krichevsky R, Trofimov V (1981) The performance of universal coding. IEEE Trans Inf Theory 28:199–207.

Lan X, Roth S, Huttenlocher DP, Black M (2006) Efficient belief propagation with learned higher-order Markov random fields. In: European Conference on Computer Vision, Vol II, pp 269–282. London: Springer.

Laurent G (2002) Olfactory network dynamics and the coding of multidimensional signals. Nat Rev Neurosci 3:884–895.

Litke A, Bezayiff N, Chichilnisky E, Cunningham W, Dabrowski W, Grillo A, Grivich M, Grybos P, Hottowy P, Kachiguine S, Kalmar R, Mathieson K, Petrusca D, Rahman M, Sher A (2004) What does the eye tell the brain? Development of a system for the large scale recording of retinal output activity. IEEE Trans Nucl Sci 51:1434–1440.

MacNeil MA, Masland RH (1998) Extreme diversity among amacrine cells: implications for function. Neuron 20:971–982.

Malouf R (2002) A comparison of algorithms for maximum entropy parameter estimation. In: Proceedings of the Sixth Conference on Natural Language Learning (CoNLL-2002), pp 49–55. Stockholm, Sweden: Morgan Kauffman.

Marmarelis PZ, Naka K (1972) White-noise analysis of a neuron chain: an application of the wiener theory. Science 175:1276–1278.

Masland RH (2001) The fundamental plan of the retina. Nat Neurosci 4:877–886.

Mastronarde DN (1983a) Correlated firing of cat retinal ganglion cells. I. spontaneously active inputs to x- and y-cells. J Neurophysiol 49:303–324.

Mastronarde DN (1983b) Correlated firing of cat retinal ganglion cells. II. responses of x- and y-cells to single quantal events. J Neurophysiol 49:325–349.

Mastronarde DN (1983c) Interactions between ganglion cells in cat retina. J Neurophysiol 49:350–365.

Mastronarde DN (1989) Correlated firing of retinal ganglion cells. Trends Neurosci 12:75–80.

Meister M (1996) Multineuronal codes in retinal signaling. Proc Natl Acad Sci U S A 93:609–614.

Meister M, Lagnado L, Baylor DA (1995) Concerted signaling by retinal ganglion cells. Science 270:1207–1210.

Mitchell JF, Sundberg KA, Reynolds JH (2007) Differential attention-dependent response modulation across cell classes in macaque visual area v4. Neuron 55:131–141.

Narayanan NS, Kimchi EY, Laubach M (2005) Redundancy and synergy of neuronal ensembles in motor cortex. J Neurosci 25:4207–4216.

Nemenman I, Shafee F, Bialek W (2002) Entropy and inference, revisited. Adv Neural Inf Process Syst 14:95–100.

Nirenberg SH, Victor JD (2007) Analyzing the activity of large populations of neurons: how tractable is the problem? Curr Opin Neurobiol 17:397–400.

Nirenberg S, Carcieri SM, Jacobs AL, Latham PE (2001) Retinal ganglion cells act largely as independent encoders. Nature 411:698–701.

Ohki K, Chung S, Ch'ng YH, Kara P, Reid RC (2005) Functional imaging with cellular resolution reveals precise micro-architecture in visual cortex. Nature 433:597–603.

Ohki K, Chung S, Kara P, Hübener M, Bonhoeffer T, Reid RC (2006) Highly ordered arrangement of single neurons in orientation pinwheels. Nature 442:925–928.

Okabe A, Boots B, Sugihara K, Chiu SN (2000) Spatial tessellations: concepts and applications of Voronoi diagrams. Wiley series in probability and statistics. New York: Wiley.

Olveczky BP, Baccus SA, Meister M (2003) Segregation of object and background motion in the retina. Nature 423:401–408.

Orlitsky A, Santhanam N, Zhang J (2004) Universal compression of memoryless sources over unknown alphabets. IEEE Trans Inf Theory 50:1469–1481.

Palm G, Aertsen AM, Gerstein GL (1988) On the significance of correlations among neuronal spike trains. Biol Cybern 59:1–11.

Paninski L (2003) Convergence properties of three spike-triggered analysis techniques. Network 14:437–464.

Perez P (1998) Markov random fields and images. CWI Q 11:413–437.

Perkel DH, Gerstein GL, Moore GP (1967) Neuronal spike trains and stochastic point processes. II. Simultaneous spike trains. Biophys J 7:419–440.

Pillow JW, Shlens J, Paninski L, Sher A, Litke AM, Chichilnisky EJ, Simoncelli EP (2008) Spatio-temporal correlations and visual signalling in a complete neuronal population. Nature 454:995–999.

Press W, Flannery B, Teukolsky S, Vetterling W (1988) Numerical recipes in C. Cambridge, UK: Cambridge UP.

Puchalla JL, Schneidman E, Harris RA, Berry MJ (2005) Redundancy in the population code of the retina. Neuron 46:493–504.

Roerdink J, Meijster A (2001) The watershed transform: definitions, algorithms and parallelization strategies. Fundam Inform 41:187–228.

Roudi Y, Nirenberg S, Latham P (2009) Pairwise maximum entropy models for studying large biological systems: when they can work and when they can't. PLos Comput Biol, in press.

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

Schneidman E, Still S, Berry MJ 2nd, Bialek W (2003b) Network information and connected correlations. Phys Rev Lett 91:238701.

Schneidman E, Berry MJ 2nd, Segev R, Bialek W (2006) Weak pairwise correlations imply strongly correlated network states in a neural population. Nature 440:1007–1012.

Schnitzer MJ, Meister M (2003) Multineuronal firing patterns in the signal from eye to brain. Neuron 37:499–511.

Schwartz G, Taylor S, Fisher C, Harris R, Berry MJ 2nd (2007) Synchronized firing among retinal ganglion cells signals motion reversal. Neuron 55:958–969.

Shlens J, Field GD, Gauthier JL, Grivich MI, Petrusca D, Sher A, Litke AM, Chichilnisky EJ (2006) The structure of multi-neuron firing patterns in primate retina. J Neurosci 26:8254–8266.

Stafford DK, Dacey DM (1997) Physiology of the a1 amacrine: a spiking, axon-bearing interneuron of the macaque monkey retina. Vis Neurosci 14:507–522.

Swendsen RH, Wang JS (1987) Nonuniversal critical dynamics in Monte Carlo simulations. Phys Rev Lett 58:86–88.

Tang A, Jackson D, Hobbs J, Chen W, Smith JL, Patel H, Prieto A, Petrusca D, Grivich MI, Sher A, Hottowy P, Dabrowski W, Litke AM, Beggs JM (2008) A maximum entropy model applied to spatial and temporal correlations from cortical networks in vitro. J Neurosci 28:505–518.

Trong PK, Rieke F (2008) Origin of correlated activity between parasol retinal ganglion cells. Nat Neurosci 11:1343–1351.

Usrey WM, Reid RC (1999) Synchronous activity in the visual system. Annu Rev Physiol 61:435–456.

Voronoi G (1907) Nouvelles applications des paramètres continus à la théorie des formes quadratiques. J Reine Angew Math 133:97–178.

Yoshimura Y, Callaway EM (2005) Fine-scale specificity of cortical networks depends on inhibitory cell type and connectivity. Nat Neurosci 8:1552–1559.

Yoshimura Y, Dantzker JL, Callaway EM (2005) Excitatory cortical neurons form fine-scale functional networks. Nature 433:868–873.