

Heterogeneous Response Dynamics in Retinal Ganglion Cells: The Interplay of Predictive Coding and Adaptation

Sheila Nirenberg, Illya Bomash, Jonathan W. Pillow and Jonathan D. Victor
J Neurophysiol 103:3184-3194, 2010. First published 31 March 2010; doi:10.1152/jn.00878.2009

You might find this additional info useful...

Supplemental material for this article can be found at:

<http://jn.physiology.org/content/suppl/2010/05/13/jn.00878.2009.DC1.html>

This article cites 36 articles, 12 of which can be accessed free at:

<http://jn.physiology.org/content/103/6/3184.full.html#ref-list-1>

Updated information and services including high resolution figures, can be found at:

<http://jn.physiology.org/content/103/6/3184.full.html>

Additional material and information about *Journal of Neurophysiology* can be found at:

<http://www.the-aps.org/publications/jn>

This information is current as of September 6, 2011.

Heterogeneous Response Dynamics in Retinal Ganglion Cells: The Interplay of Predictive Coding and Adaptation

Sheila Nirenberg,^{1,2} Illya Bomash,⁴ Jonathan W. Pillow,⁵ and Jonathan D. Victor^{2,3}

¹Department of Physiology and Biophysics, ²Institute for Computational Biomedicine, ³Department of Neurology and Neuroscience, Division of Systems Neurology and Neuroscience, Weill Cornell Medical College; ⁴Tri-Institutional Training Program in Computational Biology and Medicine, Weill Cornell Medical College, New York, New York; and ⁵Center for Perceptual Systems, Institute for Neuroscience and Psychology, The University of Texas at Austin, Austin, Texas

Submitted 28 September 2009; accepted in final form 31 March 2010

Nirenberg S, Bomash I, Pillow JW, Victor JD. Heterogeneous response dynamics in retinal ganglion cells: the interplay of predictive coding and adaptation. *J Neurophysiol* 103: 3184–3194, 2010. First published March 31, 2010; doi:10.1152/jn.00878.2009. To make efficient use of their limited signaling capacity, sensory systems often use predictive coding. Predictive coding works by exploiting the statistical regularities of the environment—specifically, by filtering the sensory input to remove its predictable elements, thus enabling the neural signal to focus on what cannot be guessed. To do this, the neural filters must remove the environmental correlations. If predictive coding is to work well in multiple environments, sensory systems must adapt their filtering properties to fit each environment's statistics. Using the visual system as a model, we determine whether this happens. We compare retinal ganglion cell dynamics in two very different environments: white noise and natural. Because natural environments have more power than that of white noise at low temporal frequencies, predictive coding is expected to produce a suppression of low frequencies and an enhancement of high frequencies, compared with the behavior in a white-noise environment. We find that this holds, but only in part. First, predictive coding behavior is not uniform: most ON cells manifest it, whereas OFF cells, on average, do not. Overlaid on this nonuniformity *between* cell classes is further nonuniformity *within* both cell classes. These findings indicate that functional considerations beyond predictive coding play an important role in shaping the dynamics of sensory adaptation. Moreover, the differences in behavior between ON and OFF cell classes add to the growing evidence that these classes are not merely homogeneous mirror images of each other and suggest that their roles in visual processing are more complex than expected from the classic view.

INTRODUCTION

Sensory systems must confront the problem that their signaling capacity is limited. To make efficient use of this capacity, it has been hypothesized (Barlow 1961) that sensory systems exploit the fact that the environment has statistical regularities and is therefore in part predictable. The predictable aspects of the environment need not be explicitly represented in the neural stream since they are, by definition, not informative. By ignoring the predictable aspects of the environment, neurons can focus their signaling capacity on what is *not* predictable and thus convey sensory information in an efficient manner. However, this view immediately raises a question: as the organism moves from one environment to another, the statistical regularities of the sensory input

change. For the efficient coding hypothesis to hold, the sensory system must adapt appropriately to these changing statistics. Here we ask, does this happen?

Changes in the statistics of a sensory environment may consist of a change in the range of intensities, a change in its correlation structure, or both. Both kinds of changes are known to produce shifts in response properties that are at least qualitatively consistent with the efficient coding hypothesis. When the range of stimulus intensities increases, the neural response gain decreases (Brenner et al. 2000; Fairhall et al. 2001; Gaudry and Reinagel 2007a; Shapley and Victor 1978; Wark et al. 2009). When the correlation structure of the environment changes, neural filtering properties change, both in space (Hosoya et al. 2005; Lesica et al. 2007; Sharpee et al. 2006) and in time (Hosoya et al. 2005; Lesica et al. 2007). Our focus in this study is on the latter.

When one considers correlations, the efficient coding hypothesis translates into a design principle for sensory filters, known as “predictive coding” (Atick and Redlich 1990; Dan et al. 1996; Srinivasan et al. 1982): the design principle is that efficient coding is achieved by filters that remove the correlations in the sensory stream. The implication of this idea is that as an animal moves between environments with different correlation structures, the filters must change so that efficient coding is maintained.

With this in mind, and a focus on temporal correlations, we examined the dynamics of mouse retinal ganglion cells in environments with two very different kinds of temporal correlation structure: natural scenes and white noise. Qualitatively, and consistent with the results of others (Hosoya et al. 2005; Lesica et al. 2007), we find that filtering properties change in a manner that tends to maintain predictive coding across conditions. When switched between environments, neurons shift their dynamics, becoming more high-pass in the naturalistic environment than in white noise.

However, we also identify several aspects of the adaptation of retinal dynamics that are not anticipated from predictive coding. Most prominently, adaptation is not uniform across cell classes: in naturalistic conditions, the ON cell population reduces its gain by a factor of 10 at low frequencies, whereas the OFF cell population, on average, reduces its gain by a factor of only 2. At high frequencies, ON cells increase their gain by a factor of almost 2, whereas OFF cells show no increase. We find no statistical differences between bright and dark in the naturalistic stimulus to account for this difference between cell classes.

In addition to the differences in adaptive behavior between ON and OFF populations, adaptive behavior is also not uniform within the populations. As we will show, this heterogeneity is nontrivial:

Address for reprint requests and other correspondence: J. Victor, Department of Neurology and Neuroscience, Division of Systems Neurology and Neuroscience, Weill Cornell Medical College, New York, NY 10065 (E-mail: jdvicto@med.cornell.edu).

it does not arise merely because there are cells that are intermediate in behavior between ON and OFF.

In sum, we find that retinal dynamics shift in a manner that is consistent with the demands of predictive coding, but only in part. ON and OFF cells are not simply mirror images of each other, and the way that they adapt to changes in stimulus statistics accentuates this asymmetry and reveals heterogeneity within the cell classes.

METHODS

Preparation and recording

Recordings from central retinal ganglion cells of the isolated mouse retina were obtained via a flat array of 64 microelectrodes as described in Dedek et al. (2008). Briefly, spike waveforms were recorded using a Multichannel Neuronal Acquisition Processor (Plexon Instruments, Dallas, TX). Two different standard spike-sorting methods were used: a window discriminator or a principal component analysis (PCA)-based waveform sorting algorithm implemented in Chronux (Mitra and Bokil 2008) based on the method of Fee et al. (1996).

Stimuli

The white-noise stimulus consisted of a pseudorandom binary sequence presented in a 20 × 18 stimulus array of 80 × 80-micron (2.6 × 2.6°) square checks, updated every 67 ms. The naturalistic sequence consisted of a movie of a ground-level animal’s view of the landscape in New York City’s Central Park. The luminance within each 67 ms block was averaged, rescaled to an 8-bit range [0, 255], and presented in the same array as was used for the white-noise stimulus. The two levels used for the binary stimulus were 92 and 202, chosen so that the two stimuli were matched for mean (148) and SD (54). The mean luminance for both stimuli at the retina was 0.24 μW/cm². Each stimulus sequence was presented for 9,000 frames, lasting 600 s.

In addition to the above-cited stimulus samples used for constructing the model, we measured responses to repeated presentation of a validation (out-of-sample) sequence. For white noise, this consisted of 60 to 200 repeats of a 90-frame (6 s) sequence (seven retinae); for the naturalistic stimulus (all 82 cells), this consisted of 70 to 180 repeats of a 90-frame sequence (one retina) or a 150-frame (10 s) sequence (eight retinae).

Response modeling

To compare ganglion cell dynamics in white-noise and naturalistic conditions in the context of the predictive coding hypothesis, we require a filter-based description of the stimulus–response relationship in each of the two environments. To provide these descriptions, we fitted separate linear-nonlinear-Poisson (LNP) models (Chichilnisky 2001) to data obtained under the two experimental conditions. The parameter-fitting strategy was based on the maximum-likelihood approach of Pillow et al. (2008) and was tailored to the needs of this study as described in the next section.

Model structure

Neuronal firing is modeled as an inhomogeneous Poisson process whose intensity (firing probability) $p(t)$ is the result of filtering the stimulus S by a linear filter L , and then applying a nonlinearity N

$$p(t) = N[(L * S)(t)] \tag{1}$$

where $*$ represents spatiotemporal convolution, formally defined by

$$(L * S)(t) = \int \int \int L(\tau, x, y) S(t - \tau, x, y) d\tau dx dy \tag{2}$$

Since the stimulus is constant on pixels of size $\Delta x \times \Delta y = 80 \times 80$ microns and frames of length $\Delta t = 67$ ms, we replace this integral by a discrete sum

$$(L * S)(t) = \sum_{i=1}^T \sum_{n_x} \sum_{n_y} S(t - i\Delta\tau, x_0 + n_x\Delta x, y_0 + n_y\Delta y) \times L(i\Delta\tau, n_x\Delta x, n_y\Delta y) (\Delta\tau\Delta x\Delta y) \tag{3}$$

where i covers $T = 18$ time steps (1.2 s) and the spatial indices n_x and n_y each cover 10 contiguous integers, corresponding to a 10×10 array of pixels covering and approximately centered on the receptive field. Here, the range of stimulus intensities S is mapped to $[-0.5, 0.5]$, with -0.5 corresponding to black and 0.5 corresponding to the maximum luminance, i.e., $0.42 \mu\text{W}/\text{cm}^2$.

To reduce the number of free parameters, we assume that the linear filter L is separable into a product of a temporal kernel $G_{temp}(\tau)$ characterized by 18 parameters, one for each lag, and a spatial kernel $G_{spat}(x, y)$ characterized by 100 parameters, one for each pixel [Pillow et al. (2008) used two such terms]. To remove the ambiguity of multiplicative constants shared by these factors, the kernels $G_{temp}(\tau)$ and $G_{spat}(x, y)$ are normalized by $\sum_{i=1}^T [G_{temp}(i\Delta\tau)]^2 = 1$ and $\sum_{n_x} \sum_{n_y} [G_{spat}(n_x\Delta x, n_y\Delta y)]^2 = 1$ and the overall size of L is brought into a single factor k_{mult}

$$L(\tau, x, y) = k_{mult} G_{temp}(\tau) G_{spat}(x, y) \tag{4}$$

To further reduce the number of free parameters and to provide for a principled extrapolation of the temporal kernel to long times without incurring an artifactual transient after the longest lag explicitly considered (1.2 s), the temporal kernel was constrained to be a sum of 10 basis functions (Pillow et al. 2008). The first five basis functions e_j ($j = 1, \dots, 5$) were set to 1 on the j th time bin (of length $\Delta\tau$) and 0 elsewhere. The last five basis functions e_j ($j = 6, \dots, 10$) were single lobes of raised cosine functions in logarithmic time

$$e_j(\tau) = \lambda_j \left\{ 1 + \cos \left[\frac{\pi}{\alpha_j} \left(\log \frac{\tau}{\Delta\tau} - \beta_j \right) \right] \right\} \tag{5}$$

for

$$\left| \log \frac{\tau}{\Delta\tau} - \beta_j \right| < \alpha_j$$

where α_j ranged from 0.22 to 0.55, β_j ranged from 1.8 to 2.6, and λ_j was chosen for unit normalization. This extrapolation had a negligible effect on the estimated transfer functions, since impulse responses had largely returned to zero at 1.2 s.

To allow for exploration of a wide range of kinds of neural responses, the nonlinearity N of Eq. 1 was parameterized as a cubic spline (i.e., a piecewise cubic polynomial with continuous second-order derivatives), which was not constrained to be monotonic. This allowed us to capture the behavior of neurons with ON–OFF characteristics. As we show, this model also provided for accurate estimation of response dynamics under both white-noise and natural scene conditions (Figs. 1B and 2B), as is required to study adaptation.

Model fitting

Instances of the preceding model were independently fit to responses to white-noise and naturalistic stimulation by maximizing the likelihood of the observed responses. To carry out this fit, we began with an exponential nonlinearity $N(x) = \exp(x + k_0)$, that is

$$p(t) = \exp[(L * S)(t) + k_0] \tag{6}$$

because for this nonlinearity, there are no local maxima (Paninski 2004), thus facilitating the parameter estimation. We then replaced the exponential nonlinearity by an approximating 6-knot cubic spline. The model parameters were fit by coordinate ascent (i.e., alternating stages of maximizing the log-likelihood with respect to (i) the spline coefficients and (ii) the filter parameters, until a maximum was reached). The spline parameters were not restricted to monotonic nonlinearities. Empirically, this procedure did not incur a drift in the values of k_{mult}

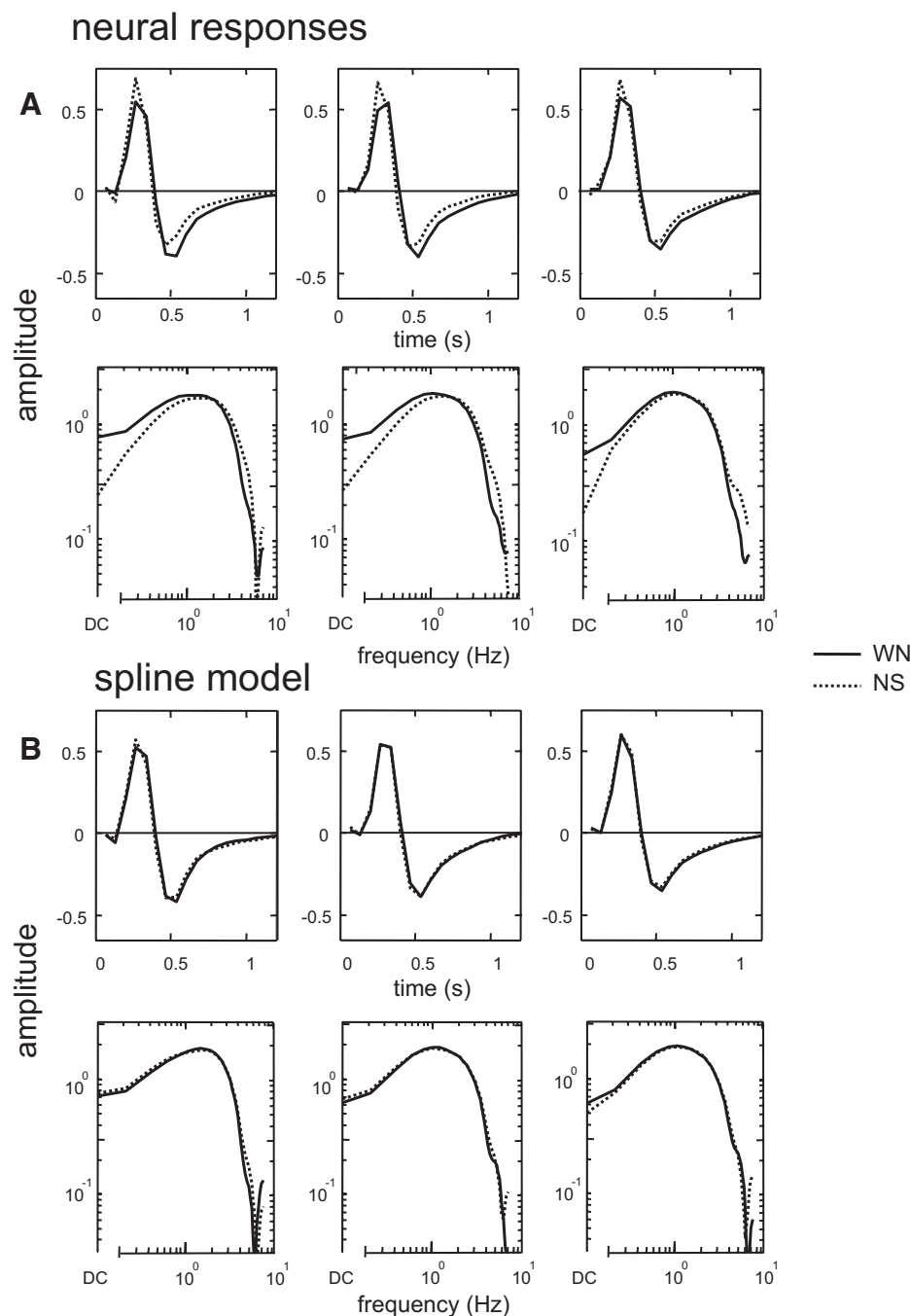


FIG. 1. *A*: temporal kernels (*top row*) and their Fourier transforms (*bottom row*) of 3 example ON retinal ganglion cells under white-noise (solid) and naturalistic (dotted) conditions. Under naturalistic conditions, the response at high frequencies is enhanced and the response at low frequencies is attenuated. Temporal kernels are normalized (root mean squared) to unity power; Fourier transforms maintain this normalization. Bias indices for the 3 cells (from *left to right*): 0.87, 0.89, 0.99. *B*: parallel analysis of model linear-nonlinear-Poisson (LNP) neurons whose parameters are chosen to match the white-noise behavior of the neurons in *A*. For model LNP neurons, the temporal characteristics are identical under white-noise and naturalistic conditions.

or k_0 , even though in principle they are redundant with spline parameters. This procedure was vetted by demonstrating that it properly recovered the parameters of model LNP systems (including systems with nonmonotonic nonlinearities) from their responses to both white-noise and naturalistic inputs (see Figs. 1*B* and 2*B*).

Calculations were carried out in Matlab, using code adapted from that of Pillow et al. (2008).

Analysis of model parameters

To calculate the frequency response, we Fourier-transformed the temporal kernel $G_{temp}(\tau)$ numerically. Since the impulse response was measured up to $T = 1.2$ s, frequency resolution is bounded by $1/T = 0.8$ Hz. That is, the DC value of the frequency response, which is the

average of the impulse response over the previous $T = 1.2$ s (18 time points), captures the average behavior over frequencies from 0 to 0.8 Hz.

In Figs. 1–4 we show in addition, for the interested reader, the frequency response at a step beyond the guaranteed resolution. To calculate this, we used the fact that the impulse response approaches zero at $T = 1.2$ s. Thus we tapered it to zero beyond $T = 1.2$ s by projecting it on the basis functions indicated in Eq. 5 and calculated frequency responses by Fourier transformation over the interval from 0 to $4T$. As a check, we also used zero-padding and found nearly identical results. Note that the conclusions of this study do not depend on knowledge of this fine structure.

To measure the goodness of fit of the model, we compared the ability of the model response to predict the neuron's response to the neuron's intrinsic reproducibility. To measure the model's predictive ability, we

used the variance explained by the model's prediction of the response to the out-of-sample validation trials (Chichilnisky 2001). To measure the neuron's intrinsic reproducibility, we compared the neural responses to the first and second halves of the validation trials. Our measure of goodness of fit (GOF) was the ratio of these quantities

$$GOF = \frac{V(M, R_1) + V(M, R_2)}{2V(R_1, R_2)} \quad (7)$$

where V is the fraction of variance explained, M is the model response, R_1 is the firing rate for the first half of validation-sequence trials, and R_2 is the firing rate for the second half of validation-sequence trials. Firing rates R_1 and R_2 were calculated by convolving the raw spike train sequence with a Gaussian whose standard deviation was half the bin width, 33 ms. We report results for the goodness of fit to the natural scene stimulus, which was measured in all cells; similar results were obtained for the goodness of fit measured with the white-noise stimulus (53/82 cells).

Bias index

The bias index (BI) is a measure of the overall balance of ON and OFF inputs, defined by

$$BI = (R_{ON} - R_{OFF}) / (R_{ON} + R_{OFF}) \quad (8)$$

where R_{ON} is the peak firing rate following a step increase in luminance and R_{OFF} is similarly defined for the offset transient. Note that for a cell that has only an ON response ($R_{ON} > 0$, $R_{OFF} = 0$), then $BI = 1$. Similarly, for a cell that only has an OFF response, $BI = -1$.

R_{ON} and R_{OFF} were measured by adapting the procedure of Carcier et al. (2003) for use with random binary stimulation. Specifically, we first identified the "optimal" check—i.e., the one that was closest to the receptive field center and covered most of it. Then, to measure R_{ON} , we calculated a poststimulus time histogram (PSTH) triggered by portions of the stimulus sequence in which optimal check was off for five frames (333 ms) and then on for two frames (133 ms). R_{ON} was taken as the maximum of the smoothed (Gaussian standard deviation 20 ms) PSTH in the window from 230 to 330 ms (see kernel functions in Figs. 1 and 2) following the onset transient. R_{OFF} was analogously measured from segments of the stimulus sequence in which the optimal check was on for five frames and then off for two frames. With this method, the population distribution of BI values was nearly identical to that reported by Carcier et al. (2003) who used an "optimal spot," rather than the optimal check used here.

RESULTS

Our goal is to determine whether ganglion cells adapt to changes in the statistics of their inputs, focusing on temporal changes. To do this, we used a model similar to that of Pillow et al. (2008) to characterize ganglion cell response dynamics in two environments: white-noise checkerboards and naturalistic stimuli with the same spatial discretization and the same mean and variance. This model describes the neural response as the result of three transformations: a linear spatiotemporal filter L , followed by a static nonlinearity N , followed by Poisson spike generation. Because the dynamics are described by the filter L , we focus on it.

We first consider individual example cells and then the behavior of the ON and OFF cell populations as a whole. Data are presented from 82 ganglion cells (38 ON, 44 OFF) in nine retinæ.

Example cells

Figure 1A shows that in typical ON cells, a change in the temporal statistics of the sensory input induces a change in

response dynamics. In the *top row*, we compare the temporal kernel of the linear filter measured under white-noise conditions (solid) and naturalistic stimulation (dotted). Visual inspection shows that the temporal kernel under naturalistic conditions has a larger initial transient and a smaller undershoot, suggesting that a change in dynamics has occurred. To understand the implications of this shape change for frequency tuning, we Fourier-transformed the temporal kernels (Fig. 1A, *second row*). The results showed that sensitivity is augmented at high frequencies (3 to 5 Hz), by a factor of 2 in these examples, corresponding to the larger initial transient. Sensitivity is strikingly attenuated at low frequencies, by a factor of up to 3 to 5 in these three examples and by a factor of 10 on average, as shown in the next section. This loss of low-frequency sensitivity corresponds to the change in the undershoot, resulting in an equalization of the area under the two lobes of the temporal kernel.

To be sure that this change represents an adaptation of ganglion cell dynamics, rather than a behavior built into the LNP model or a result of errors in estimating model parameters, we carried out a simple simulation: we used the LNP model obtained from white-noise conditions, simulated its responses to the white-noise and the naturalistic inputs, and then estimated model parameters from those responses according to the procedures described in METHODS. This shows what one would expect for a cell whose intrinsic characteristics do not change and controls for errors in estimating model parameters from the naturalistic stimulus ensemble (Sharpee et al. 2006) due to its correlation structure. Results of this analysis are shown in Fig. 1B. In contrast to the corresponding panels of Fig. 1A, the temporal kernels (*top row*) and their Fourier transforms (*bottom row*) obtained from white-noise and naturalistic stimulation superimpose. This shows that the changes observed in Fig. 1A cannot be explained by bias or imprecision in the fitting procedure: when the fitting procedure is applied to simulated spike trains generated by LNP neurons, the temporal kernels do not change. In other words, the changes seen in Fig. 1A reflect changes in the neurons' filtering characteristics, not artifacts of the fitting procedure.

OFF cells (Fig. 2A) show a behavior that contrasts dramatically with ON cells. First, the initial response transient changes very little between white-noise and naturalistic conditions. Correspondingly, as Fourier analysis shows, there is essentially no enhancement of the response in the high-frequency range (≥ 3 Hz) under naturalistic conditions. Second, although there is a change in the size of the undershoot in the two environments, the Fourier analysis shows that this results in a smaller change in the low-frequency sensitivity than that seen in ON cells of Fig. 1A. The reason for this is that although the undershoot amplitude changes, its area changes very little.

For completeness, we performed the same computational control for OFF cells as that for ON cells (Fig. 2B), confirming that the changes in the measured temporal kernels in the two stimulus conditions indeed reflect a change in filtering characteristics.

Average behavior of ON and OFF cells

Figure 3 summarizes the preceding analysis across the population of 82 retinal ganglion cells. When shifted from white-noise to naturalistic stimuli, average ON cell ($n = 38$)

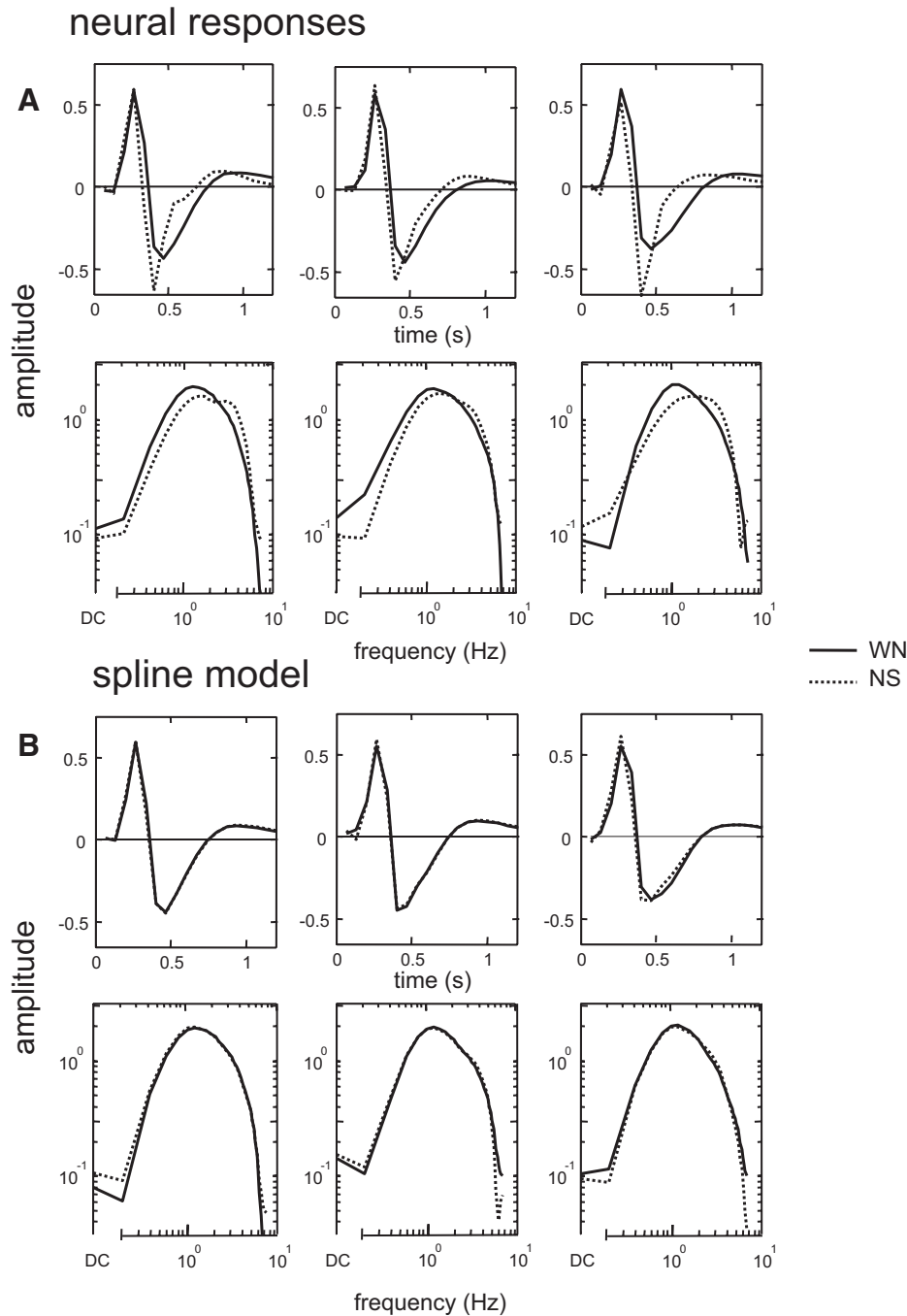


FIG. 2. *A*: temporal kernels (*top row*) and their Fourier transforms (*bottom row*) of 3 example OFF cells. Under naturalistic conditions, there is virtually no change in the response at the highest frequencies, and only very modest changes in the responses at low frequencies. Both behaviors contrast with that of the ON cells in Fig. 1. Bias indices for the 3 cells (from left to right): -0.27 , -0.26 , -0.81 . Other details as in Fig. 1. *B*: parallel analysis of model LNP neurons whose parameters are chosen to match the white-noise behavior of the neurons in *A*. For model LNP neurons, the temporal characteristics are identical under white-noise and naturalistic conditions.

sensitivity is reduced on average by a factor of 10 at the lowest frequencies and increases on average by a factor of 1.7 at 5 Hz.

In contrast, the OFF cell population ($n = 44$) shows, on average, much less adaptation: there is a reduction in sensitivity by a factor of <2 at low temporal frequencies and no appreciable increase at high temporal frequencies.

The shift seen in ON cells makes sense in terms of predictive coding: sensitivity to low frequencies is reduced in an environment in which low-frequency correlations are more prominent. However, this behavior is only qualitatively consistent with the behavior expected from predictive coding.

To work out the quantitative prediction, we determine how the temporal kernel must change so that it removes stimulus correlations as the environment changes. It is convenient to work in the frequency domain and to characterize stimulus correlations by the stimulus power spectrum $P(f)$. Assuming linearity and a neural transfer function $A(f)$, the power spectrum of the response is then $R(f) = P(f)|A(f)|^2$. Thus to maintain $R(f)$ constant as the power spectrum changes from $P_1(f)$ in environment 1 to $P_2(f)$ in environment 2, we must adjust $A(f)$ so that $P_1(f)|A_1(f)|^2 = P_2(f)|A_2(f)|^2$. This is equivalent to

$$\frac{|A_2(f)|}{|A_1(f)|} = \sqrt{\frac{P_1(f)}{P_2(f)}} \quad (9)$$

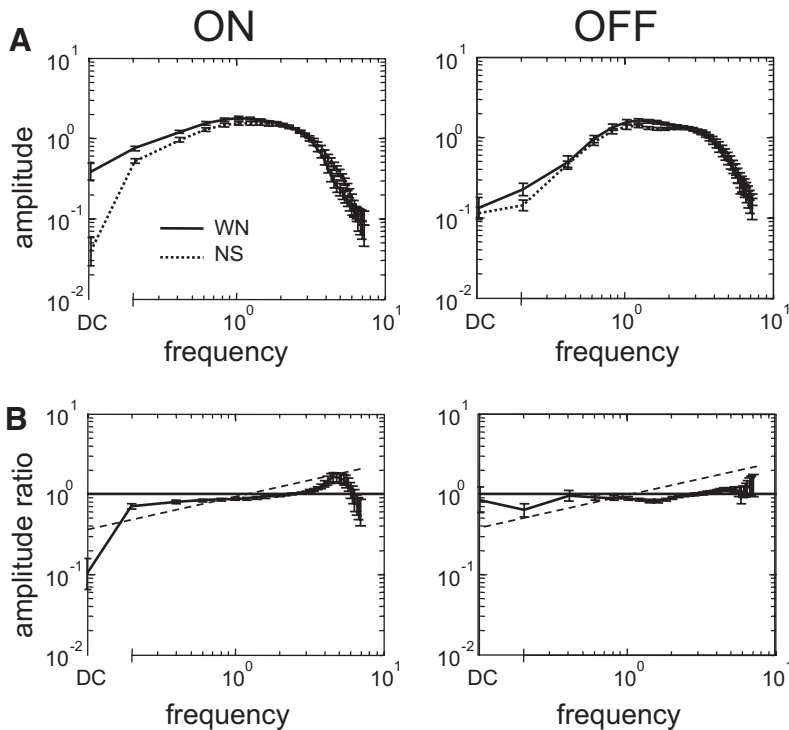


FIG. 3. Changes in temporal response characteristics across the population of ON (left) and OFF (right) ganglion cells. A: geometric mean of the Fourier transforms of the temporal kernels under white-noise (solid) and naturalistic (dotted) conditions. B: ratios. The dashed line indicates a slope of 0.41, the quantitative expectation of predictive coding. Error bars: 95% confidence limits (*t*-test).

That is, the change in the transfer function should be inversely proportional to the square root of the change in the power spectrum.

Here, the power spectrum of the white-noise stimulus is flat [$P_1(f) = K_1$], whereas the power spectrum of the naturalistic stimulus is approximately $P_2(f) = K_2 f^{-a}$, where $a = 0.83$, as

determined empirically from our stimulus (see Supplemental Fig. S1).¹ Thus, the preceding analysis predicts that adaptation to the naturalistic stimulus would result in a change in the Fourier transform of the temporal kernel proportional to $f^{a/2}$

¹ The online version of this article contains supplemental data.

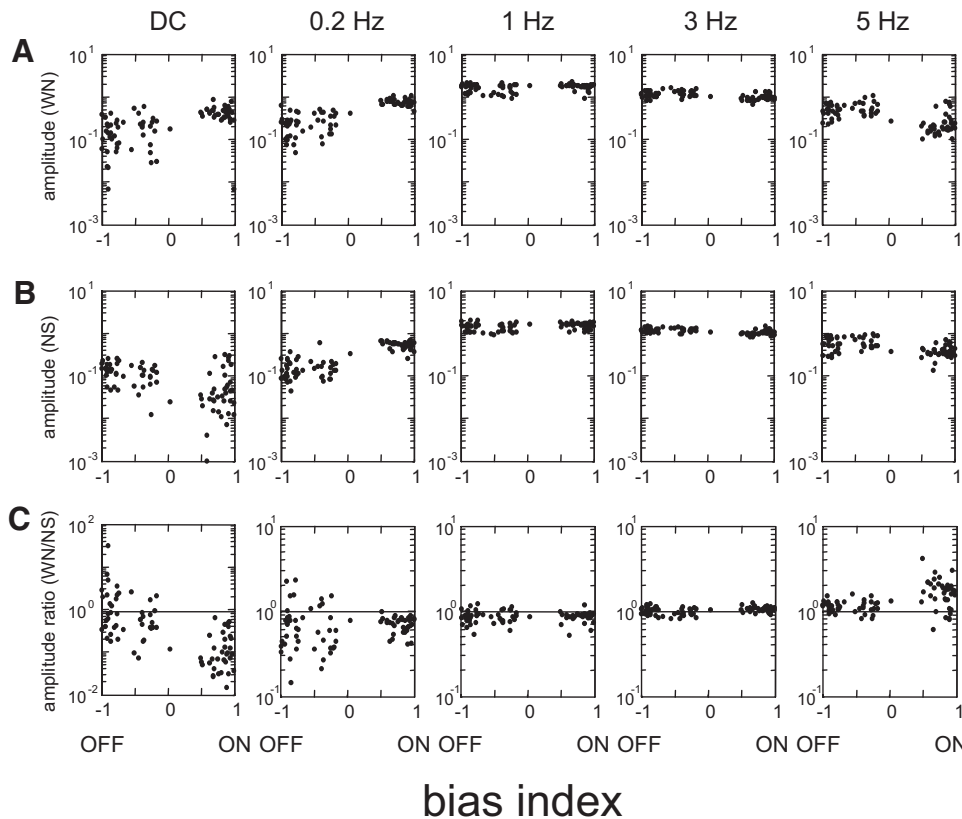


FIG. 4. Temporal tuning and adaptation along the ON–OFF continuum. Abscissa is bias index (see METHODS), which spans the gamut from -1 (pure OFF) to $+1$ (pure ON). Ordinate is the Fourier transform of the temporal kernel measured under white-noise conditions (A), naturalistic stimulation (B), and the ratio (C). Note the different ordinate scale on the leftmost panel in the bottom row (C).

and that the behavior of ON and OFF cells would be similar to each other. As a comparison with the line of slope $a/2$ shows (Fig. 3), ON cell behavior is qualitatively consistent with this expectation, but departs quantitatively. OFF cells, on the other hand, show almost no shift, inconsistent with the expectations of predictive coding.

Heterogeneity of adaptive behavior within ganglion cell populations

At this point, we have seen that adaptive behavior is distributed in a strikingly nonuniform manner across ganglion cells: ON cells adapt in a manner that is at least qualitatively consistent with predictive coding, but OFF cells, on average, show little adaptation. However, the ON versus OFF distinction, although fundamental to any discussion of ganglion cell behavior (Dowling 1970; Rodieck 1973), is a simplified picture of the diversity of ganglion cell classes. In particular, it represents a dichotomy along a continuum of a range of weightings of ON and OFF inputs (Carcieri et al. 2003).

Thus given that ON and OFF cells, on average, have distinct adaptive behaviors (Fig. 3), one can anticipate that individual ganglion cells will differ in their adaptive behavior, reflecting the balance of their ON and OFF inputs. To test this hypothesis, we analyzed how adaptive behavior of individual cells depended on the relative contribution of their inputs. The results of this analysis are shown in Fig. 4. As the figure shows, there is heterogeneity in adaptive behavior, although it substantially exceeds what would be expected by a mixing of ON and OFF signals.

To show this, we used the “bias index” (BI, Eq. 8) to quantify the relative contribution of ON and OFF inputs to each ganglion cell (Carcieri et al. 2003). A BI of -1 indicates that a cell receives exclusively OFF input; a BI of $+1$ indicates that a cell receives exclusively ON input. As seen in Fig. 4C, the BI accounts for only a small fraction of the cell-to-cell variation in adaptive behavior. Moreover, the cells with BI near $+1$ or -1 (those that are likely to have exclusively ON or OFF input) showed a range of behavior just as large as that of the cells whose BI was far from the extremes.

We next formalize these observations: (i) that there is a substantial difference in adaptive behavior between ON and OFF cell classes and (ii) that there is heterogeneity within cell classes that is not explained by the mixture of ON and OFF inputs.

With regard to the first observation, we compared the log of the adaptation ratio (the ratio of the Fourier transforms of the temporal kernels measured under the two conditions) between cell classes. Using a bipartite subdivision into OFF and ON subgroups as in Fig. 3, there was a highly significant difference at DC and 5 Hz ($P < 0.001$). The difference at 3 Hz was less significant ($P < 0.01$) and there was no significant difference at the intermediate frequencies (0.2 and 1 Hz) shown in Fig. 4. As noted earlier, the ON versus OFF subdivision is likely an oversimplification: as described by Carcieri et al. (2003), the BI distribution is trimodal, with a cutpoint between OFF and ON-OFF subsets at a BI of approximately -0.6 . However, this refinement does not change our conclusions: with the tripartite subdivision, we found that ON cells differed substantially from both OFF and ON-OFF cells at DC and 5 Hz ($P < 0.001$), whereas OFF and ON-OFF cells differed only minimally from each other ($P < 0.05$ at DC, NS at 5 Hz).

With regard to the second observation, heterogeneity within cell classes, we used the BI as a measure of the mixture of ON and OFF inputs and regressed the log of the adaptation ratio against the BI, separately within each cell class. Within each cell class (using either the bipartite or tripartite division), no more than 10% of the variance could be accounted for by the BI. We conclude that the heterogeneity within cell classes was not simply a result of a mixture of ON and OFF inputs.

It is interesting to note that the heterogeneity within categories does not merely represent random variation of responses: there is structure to the distribution of adaptive behavior. Specifically, ganglion cells that show adaptation at low temporal frequencies also tend to show adaptation at high temporal frequencies. Figure 5 shows this, by comparing the adaptation ratio at DC and at 5 Hz. Across all cells, these ratios are strongly negatively correlated ($r = -0.48$, $P < 0.001$). This covariation is concentrated within the ON subpopulation ($r = -0.40$, $P < 0.01$).

Finally, we rule out the possibility that modeling error accounts for the observed heterogeneity. To do this, we examined the adaptation ratio as a function of goodness of fit of the response model (see METHODS). As seen in Fig. 6, across the range of cells that were fit well by the model and cells that were fit poorly, a much smaller effect cells decreased their responses at DC and increased their responses at 5 Hz, whereas OFF cells showed a much smaller effect. Correspondingly, there was no correlation between adaptation ratio and goodness of fit (DC: $r = -0.02$; 5 Hz: $r = 0.12$, both $P > 0.05$).

In sum, we found that retinal ganglion cells adjust their dynamics so that their low-frequency sensitivity is reduced and their high-frequency sensitivity is enhanced, when moved from a white-noise environment to a naturalistic one. However, this difference is only qualitatively consistent with the expectations of predictive coding and there is substantial nonuniformity, both between and within cell classes.

DISCUSSION

Predictive coding in a changing environment

Sensory systems have limited signaling capacity and must operate in many kinds of environments. To confront the prob-

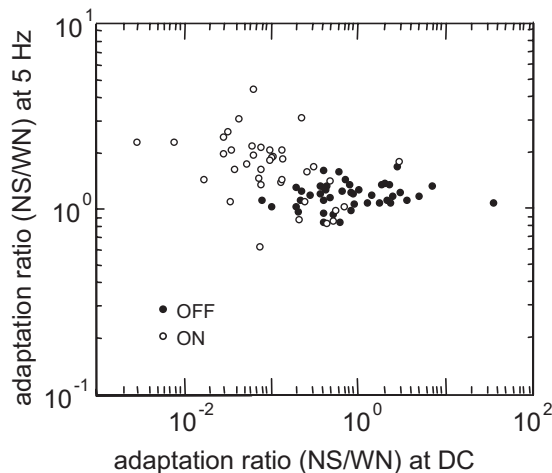


FIG. 5. Cell-to-cell variation in the adaptive change at high frequencies (ordinate) is correlated with the adaptive change at low frequencies (abscissa) for the entire population (correlation coefficient = -0.48 , $P < 0.001$) and for ON cells (correlation coefficient = -0.40 , $P < 0.01$).

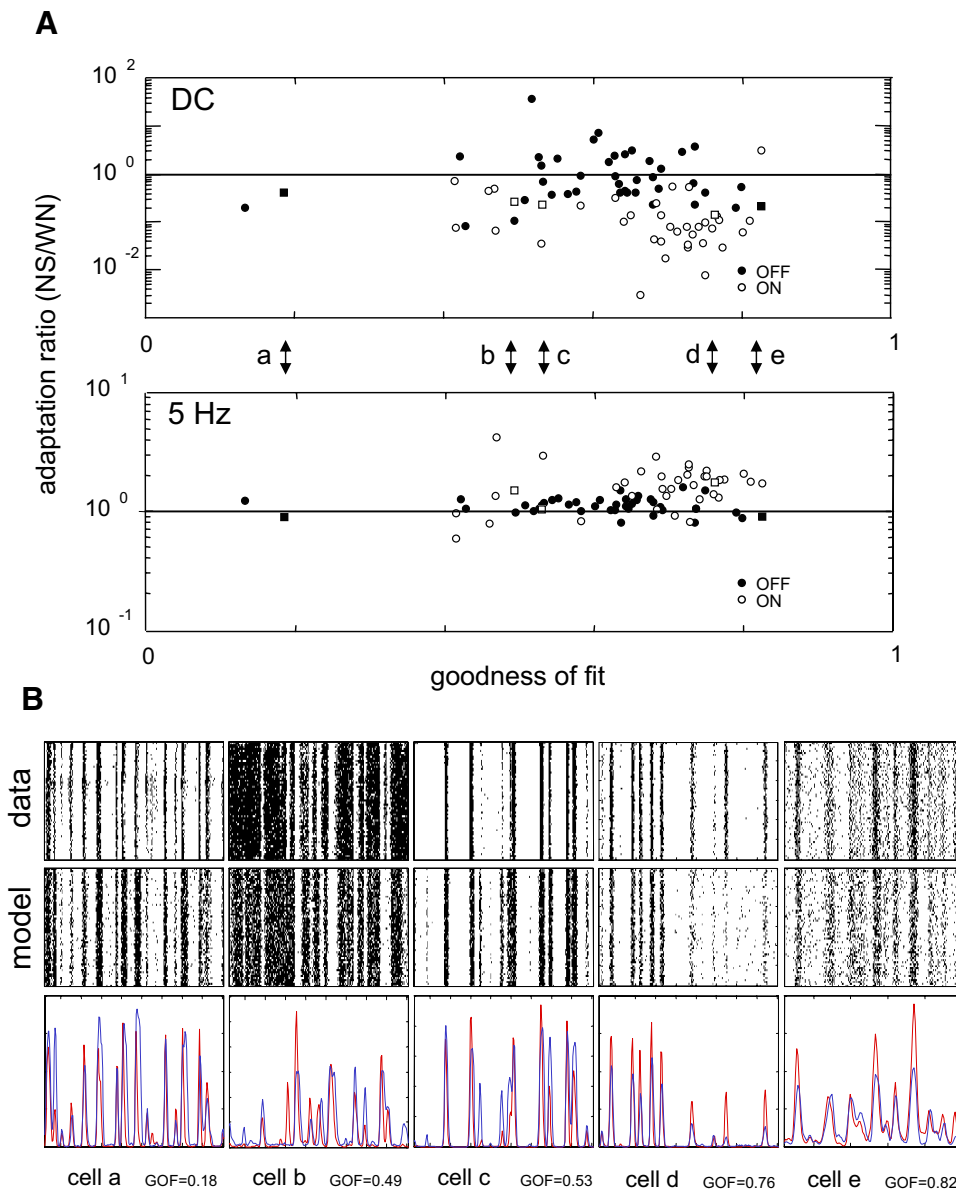


FIG. 6. *A*: heterogeneity of adaptation ratio (ordinate) is nearly independent of goodness of fit (GOF) of model (abscissa). *Top panels*: adaptation ratio at DC. *Bottom panels*: adaptation ratio at 5 Hz. See METHODS (Eq. 7) for details for calculation of goodness of fit. ON cells are plotted with open symbols; OFF cells with filled symbols. *B*: example cells with a range of values of GOF, as indicated below each set of responses. a: OFF cell; b: ON cell; c: ON cell; d: ON cell; e: OFF cell. The rasters show 60 trials (first trial at the top); the PSTHs show smoothed firing rates (33 ms half-width) for neural responses (red) and model (blue). Full-scale values and time bases are: 80 impulses/s \times 10 s (a), 300 impulses/s \times 10 s (b), 100 impulses/s \times 10 s (c), 80 impulses/s \times 10 s (d), and 80 impulses/s \times 6 s (e). The example cells in *B* are plotted with square symbols in *A*.

lem of limited signaling capacity, sensory systems make use of predictive coding (Barlow 1961; Laughlin 1981; Srinivasan et al. 1982). The idea of predictive coding is that a sensory system need not signal what is predictable about its environment and thus can concentrate its capacity on the aspects of the incoming sensory signals that cannot be guessed. To implement this, the filtering properties of a sensory system need to decorrelate the input signal—i.e., remove what is predictable about its environment. This, indeed, is what is found experimentally, both in the spatial domain (Atick and Redlich 1990, 1992; Hosoya et al. 2005; Sharpee et al. 2006; Srinivasan et al. 1982) and in the temporal domain (Dan et al. 1996; Hosoya et al. 2005; Lesica et al. 2007; Srinivasan et al. 1982).

However, sensory systems must also operate in many different environments and these environments may have different statistical characteristics. If predictive coding is to be effective as the environment changes, the filtering properties of the sensory system must adapt to its changing correlation properties.

Our main finding is that this expectation holds (Hosoya et al. 2005; Lesica et al. 2007), but in a heterogeneous fashion across and within cell classes. Specifically, when shifted from a white-noise environment to a naturalistic one, ON cells, on average, change their gain in a manner that approximately compensates for the statistics of natural scenes, but OFF cells, on average, do not (Fig. 3). Moreover, we also find that adaptive behavior is heterogeneously distributed *within* the cell classes (Figs. 4 and 5). We now consider the functional implications of these two levels of heterogeneity.

Heterogeneity between cell classes

The simplest functional hypothesis that might account for the differences between ON and OFF cells is that it merely reflects statistical asymmetries in the environment. For example, if the bright regions of natural scenes were temporally correlated but the dark regions were not, then predictive coding would in fact account for this asymmetry. However, although there are notable asymmetries in the spatial correlation struc-

ture of bright and dark regions (Balasubramanian and Sterling 2009), no such differences in the temporal correlation structure have been identified (Dong and Atick 1995a) and our analysis confirms this (Supplemental Fig. S1). Thus the observed difference in adaptive behavior between ON and OFF cells does not appear to be driven merely by the needs of predictive coding.

An alternative hypothesis is suggested by the recognition that adaptation has potential disadvantages (Wark et al. 2009). If a sensory system completely adapts to the statistics of its sensory environment, then its ability to communicate those statistics has been lost. Similarly, if a sensory system adapts based on an incorrect inference that the environment has changed, it becomes less efficient than if it did not adapt. These losses are mitigated if there is a population of ganglion cells that does not adapt (or that adapts differentially), which is what OFF cells appear to do.

ON and OFF classes are not mirror images

The differences in adaptive behavior reported here are part of the larger picture of growing evidence that ON and OFF pathways are not simply mirror images, but differ in a more fundamental manner. These differences encompass not only the way that ON and OFF cells adapt, but also the basic aspects of their behavior in a single environment.

At the most basic level, the synaptic mechanisms that provide input to ON and OFF ganglion cells differ in their linearity, their contrast–response characteristics (Zaghloul et al. 2003), and their temporal characteristics (Murphy and Rieke 2006). Their filtering properties differ as well: OFF cells are more numerous and have smaller receptive fields than those of ON cells (Balasubramanian and Sterling 2009) and this difference in spatial resolution persists into visual cortex (Zemon et al. 1988). Our data add a difference in the temporal filtering properties: ON cells are more broadband than OFF cells (Fig. 3), which corresponds to the observation that ON cell temporal kernels tend to be biphasic, whereas OFF cell temporal kernels tend to be triphasic (Figs. 1 and 2).

Further differences between ON and OFF cells are revealed by examining how they adapt to a changing environment, including (i) the striking differences reported here, (ii) the differences in adaptation rates following steps changes of white-noise stimuli reported by Wark et al. (2009), and (iii) evidence that ON cells adapt their dynamics much more than OFF cells during dark adaptation (Pandarinath et al. 2009).

Although it is tempting to think of the ON and OFF pathways as a means to subdivide positive and negative visual signals into components that spiking neurons can transmit, the fact that ON and OFF pathways differ in so many respects suggests that this view is an oversimplification. As mentioned earlier, the existence of a subset of ganglion cells that do not adapt, or that adapt differentially, can mitigate some of the deleterious effects of adaptation. However, the fact that ON and OFF cells differ even in their basic filtering properties suggests that they may play different roles in specific visual tasks or environments, even when adaptive behavior is not engaged.

Qualitative, not quantitative, agreement

Even at the level of the average behavior within the ON cell population, agreement with the expectations of predictive cod-

ing is only approximate (Fig. 3). In this section, we outline some reasons for this. We emphasize, though, that none of these considerations can account for the heterogeneity we observe, either between or within cell classes.

The main reason is that to determine the consequences of the efficient coding hypothesis for response dynamics, it is necessary to make simplifying assumptions: that the environment is Gaussian, and that neurons can be regarded as approximately linear. In principle, the efficient coding hypothesis also applies to scenarios in which the environmental signals have complex, non-Gaussian statistics (Geisler 2008) and neural channels that are nonlinear; however, to work out the consequences of the efficient coding hypothesis in this general setting is simply not possible. Therefore as have others (e.g., Atick and Redlich 1990, 1992; Dan et al. 1996; Dong and Atick 1995a,b; Hosoya et al. 2005; Sharpee et al. 2006; Srinivasan et al. 1982), we focused on the implications of the second-order correlation properties of the stimulus. This leads to the simple, straightforward prediction of predictive coding: that neural filtering will decorrelate (prewhiten) the stimulus.

Neurons have limited power and they also have noise that is frequency dependent, although within the realm of the above-cited approximation, this does not change the expectations of predictive coding. To see this, we first consider such a neuron confronted with an environment consisting of uncorrelated Gaussian noise. Intuitively, information transmission is maximized if the neuron devotes most of its power to frequencies at which its internal noise is low. Made rigorous, this intuition becomes the classic “water-filling” theorem (Shannon 1948): a recipe for the linear filter that optimizes the information transmission by a noisy channel in an uncorrelated environment. When the environment has correlations, information transmission can be further improved, by first filtering the environmental input to remove what is predictable. This yields an uncorrelated (i.e., white) signal, to which the filter specified by the water-filling theorem can then be applied. The optimal neural transformation is a product of these two components: one that depends only on the environment (the decorrelating filter) and one that depends only on the noise characteristics of the neuron (the water-filling filter). Because the first filter does not depend on the characteristics of the neuron, neuronal noise does not influence how we expect the decorrelating filter to adapt. For a more detailed analysis, see Diamantaras et al. (1999).

Heterogeneity within cell classes

Our second main finding is that even within ganglion cell classes, adaptive behavior is heterogeneous, with the heterogeneity not simply accounted for by a mixing of ON and OFF signals (Figs. 4 and 5). This conclusion holds independent of whether we consider the gamut from OFF to ON to be a continuous one quantified by the bias index, whether we dichotomize this range or whether we adopt the trichotomous classification of Carcieri et al. (2003). These authors also identified a subset of long-latency cells with ON and OFF responses and a subset of ON cells that were sustained. These distinctions support the notion of dynamic heterogeneity within the OFF and ON subsets. However, we are unable to make a precise connection between these distinctions and the present measurements of adaptation, since the long-latency and sustained subsets were defined by Carcieri et al. (2003) on the

basis of responses to sustained, optimal spots, a stimulus that was not used in the present experiments.

We now consider the functional implications of this heterogeneity. We begin by considering adaptation in its biological context. Both from the point of view of analytical convenience and the traditional approach to the analysis of neural transductions, it is natural to consider the “white-noise” condition to be the baseline and the naturalistic stimulus to be the perturbation to which ganglion cells adapt (as we have done earlier). However, from a biological perspective, the opposite view is more appropriate: the retina has evolved and developed in the animal’s environment and the white-noise stimulus is the one that represents a perturbation. From this perspective, it is not at all surprising that the adaptation ratio observed is only approximately what is expected from predictive coding—a white-noise environment is only rarely encountered and, consequently, one for which there is little selective pressure for efficiency.

Conversely, the kinds of environments encountered by an animal are likely to be more varied than just alternations between a stereotypical “natural scene” versus white noise, for example, dense brush, open space, and underground, in various admixtures, viewed in a range of lighting and atmospheric conditions. Each of these environments is characterized by distinctive statistics. Thus, to maintain predictive coding, the retina would be required to have a wide variety of adaptive behaviors, to compensate for a multidimensional range of changes in the environmental spectrum. The heterogeneity we observe may serve this purpose.

Interestingly, one can imagine how a single mechanism for adaptation, such as the well-described contrast gain control, could be varied among different cells (e.g., varied among different ON cell) to produce this heterogeneity. The action of the contrast gain control (Shapley and Victor 1978, 1981) is as follows: when contrast is high, low-frequency responses are attenuated and high-frequency responses are enhanced. Importantly, the setting of the contrast gain control is determined by a neural measure of contrast, not its physical measure, i.e., it is determined by the cells’ sensitivity to contrast, not by absolute contrast. This allows the contrast gain control to produce adaptation to temporal correlations: even though the naturalistic stimulus and the white-noise stimulus are equated for physical contrast, the sensitivity is higher for the naturalistic one because it contains more power at the frequencies that the mouse retina is most sensitive to: <3 Hz. Thus the contrast gain control is engaged more strongly for naturalistic stimuli than for white noise of the same physical contrast (Lesica et al. 2007) and could suffice to account for the basic phenomenon that low-frequency responses are attenuated for natural scenes. If this idea is correct, then the sensitivity to contrast should determine the extent of adaptation. This should hold not just between white-noise and naturalistic stimuli, but between any two environments: the extent of the change in dynamics should be governed, quantitatively, by the extent to which the frequency content of the environment activates the contrast gain control.

In this view, variability in individual neurons’ expression of the contrast gain control would allow different cells within a population to adapt to different kinds of environments. This heterogeneity in adaptive behavior could be accomplished simply: by a linear admixture of retinal signals that are, and are

not, influenced by the contrast gain control. This simplicity does have a limitation, though: it cannot adapt in an arbitrary fashion, since its repertoire is limited to this one-parameter family of mixtures. This kind of limitation may account for the imperfect adaptive behavior that we observe.

An implication of this idea is that the mechanism for contrast gain control might differ between the ON and OFF pathways. This follows from the observation that both ON and OFF cells show effects of the contrast gain control (Benardete and Kaplan 1999; Lesica et al. 2007; Ohzawa et al. 1985; Shapley and Victor 1978, 1981), but they differ in the way in which they adapt to the statistics of their inputs. As mentioned earlier, differences between ON and OFF pathways at a mechanistic level are now well documented (Murphy and Rieke 2006; Zaghloul et al. 2003) and a consequent difference in the mechanism of the gain control would not be unexpected.

Whatever its mechanistic basis, the observation that retinal ganglion cells differ in the extent to which they adapt raises an interesting possibility: that this heterogeneity provides a way to bootstrap adaptation along a single dimension at the level of individual ganglion cells (e.g., a contrast gain control) into multivariate adaptation at the population level. That is, we speculate that local populations containing a mix of neurons that differ in the extent to which they adapt may provide for efficient coding in a much wider range of environments and task demands than could be achieved by a population of neurons that all adapted similarly.

Use of LNP models

Recently, model structures other than LN cascades have been used to study issues related to adaptation (Borst et al. 2005; Famulare and Fairhall 2010; Gaudry and Reinagel 2007b) and this raises the question of why we used an LNP structure. The reason was that our goal was to examine ganglion cell response properties in the context of predictive coding. Since predictive coding makes predictions about filters, we needed a model that describes the response properties in those terms. The LNP served this purpose. If, for example, we had fit one of these alternative models to our data and found that it accounted for the differences in the responses in the two environments, it would not have provided information about whether predictive coding was occurring. To determine that, we would have to characterize the filtering properties of the alternative model under the two operating conditions. Thus we took the LNP approach because it is a direct way to get to the filters. Note that our intent is not to exclude these models; we just use the LN approach for a particular purpose.

Conclusion

We have identified adaptive behavior in retinal ganglion cells that is qualitatively consistent with the demands of predictive coding in a changing environment, but is characterized by heterogeneity at two levels: between ganglion cell classes and within ganglion cell classes. We speculate that this heterogeneity has a functional role: to provide for efficient coding in multiple kinds of environments and for multiple tasks, rather than to optimize efficiency for a single stereotypical “natural scene.”

ACKNOWLEDGMENTS

We thank M. Meytlis for providing the naturalistic stimulus and for assistance with the retinal recordings.

GRANTS

This work was supported by National Eye Institute Grants EY-12978 to S. Nirenberg and EY-9314 and EY-7977 to J. Victor.

DISCLOSURES

No conflicts of interest are declared by the authors.

REFERENCES

- Atick JJ, Redlich AN.** Towards a theory of early visual processing. *Neural Comput* 2: 308–320, 1990.
- Atick JJ, Redlich AN.** What does the retina know about natural scenes. *Neural Comput* 4: 196–210, 1992.
- Balasubramanian V, Sterling P.** Receptive fields and functional architecture in the retina. *J Physiol* 587: 2753–2767, 2009.
- Barlow HB.** Possible principles underlying the transformation of sensory messages. In: *Sensory Communication*, edited by Rosenblith WA. Cambridge, MA: MIT Press, 1961, p. 217–234.
- Benardete EA, Kaplan E.** The dynamics of primate M retinal ganglion cells. *Vis Neurosci* 16: 355–368, 1999.
- Borst A, Flanagan VL, Sompolinsky H.** Adaptation without parameter change: dynamic gain control in motion detection. *Proc Natl Acad Sci USA* 102: 6172–6176, 2005.
- Brenner N, Bialek W, de Ruyter van Steveninck R.** Adaptive rescaling maximizes information transmission. *Neuron* 26: 695–702, 2000.
- Carcieri SM, Jacobs AL, Nirenberg S.** Classification of retinal ganglion cells: a statistical approach. *J Neurophysiol* 90: 1704–1713, 2003.
- Chichilnisky EJ.** A simple white noise analysis of neuronal light responses. *Network* 12: 199–213, 2001.
- Dan Y, Atick JJ, Reid RC.** Efficient coding of natural scenes in the lateral geniculate nucleus: experimental test of a computational theory. *J Neurosci* 16: 3351–3362, 1996.
- Dedek K, Pandarinath C, Alam NM, Wellershaus K, Schubert T, Willecke K, Prusky GT, Weiler R, Nirenberg S.** Ganglion cell adaptability: does the coupling of horizontal cells play a role? *PLoS One* 3: , 2008.
- Diamantaras KI, Hornik K, Strintzis MG.** Optimal linear compression under unreliable representation and robust PCA neural models. *IEEE Trans Neural Netw* 10: 1186–1195, 1999.
- Dong DW, Atick JJ.** Statistics of natural time-varying images. *Network Comput Neural Syst* 6: 345–358, 1995a.
- Dong DW, Atick JJ.** Temporal decorrelation: a theory of lagged and non-lagged responses in the lateral geniculate-nucleus. *Network Comput Neural Syst* 6: 159–178, 1995b.
- Dowling JE.** Organization of vertebrate retinas: The Jonas M. Friedenwald Memorial Lecture. *Invest Ophthalmol Vis Sci* 9: 655–680, 1970.
- Fairhall AL, Lewen GD, Bialek W, de Ruyter Van Steveninck RR.** Efficiency and ambiguity in an adaptive neural code. *Nature* 412: 787–792, 2001.
- Famulare M, Fairhall A.** Feature selection in simple neurons: how coding depends on spiking dynamics. *Neural Comput* 22: 581–598, 2010.
- Fee MS, Mitra PP, Kleinfeld D.** Automatic sorting of multiple unit neuronal signals in the presence of anisotropic and non-Gaussian variability. *J Neurosci Methods* 69: 175–188, 1996.
- Gaudry KS, Reinagel P.** Benefits of contrast normalization demonstrated in neurons and model cells. *J Neurosci* 27: 8071–8079, 2007a.
- Gaudry KS, Reinagel P.** Contrast adaptation in a nonadapting LGN model. *J Neurophysiol* 98: 1287–1296, 2007b.
- Geisler WS.** Visual perception and the statistical properties of natural scenes. *Annu Rev Psychol* 59: 167–192, 2008.
- Hosoya T, Baccus SA, Meister M.** Dynamic predictive coding by the retina. *Nature* 436: 71–77, 2005.
- Laughlin S.** A simple coding procedure enhances a neuron's information capacity. *Z Naturforsch C* 36: 910–912, 1981.
- Lesica NA, Jin J, Weng C, Yeh CI, Butts DA, Stanley GB, Alonso JM.** Adaptation to stimulus contrast and correlations during natural visual stimulation. *Neuron* 55: 479–491, 2007.
- Mante V, Frazor RA, Bonin V, Geisler WS, Carandini M.** Independence of luminance and contrast in natural scenes and in the early visual system. *Nat Neurosci* 8: 1690–1697, 2005.
- Mitra P, Bokil H.** *Observed Brain Dynamics*. Oxford, UK: Oxford Univ. Press, 2008.
- Murphy GJ, Rieke F.** Network variability limits stimulus-evoked spike timing precision in retinal ganglion cells. *Neuron* 52: 511–524, 2006.
- Ohzawa I, Sclar G, Freeman RD.** Contrast gain control in the cat's visual system. *J Neurophysiol* 54: 651–667, 1985.
- Pandarinath C, Victor JD, Nirenberg S.** Symmetry breakdown in the ON and OFF pathways of the retina at night: functional implications. Program No. 848.6, 2009 Neuroscience Meeting Planner. Washington, DC: Society for Neuroscience, 2009. Online.
- Paninski L.** Maximum likelihood estimation of cascade point-process neural encoding models. *Network* 15: 243–262, 2004.
- Pillow JW, Shlens J, Paninski L, Sher A, Litke AM, Chichilnisky EJ, Simoncelli EP.** Spatio-temporal correlations and visual signalling in a complete neuronal population. *Nature* 454: 995–999, 2008.
- Rodieck R.** *The Vertebrate Retina: Principles of Structure and Function*. San Francisco, CA: Freeman, 1973.
- Shannon C.** A mathematical theory of communication. *Bell Syst Tech J* 27: 379–423, 1948.
- Shapley RM, Victor JD.** The effect of contrast on the transfer properties of cat retinal ganglion cells. *J Physiol* 285: 275–298, 1978.
- Shapley RM, Victor JD.** How the contrast gain control modifies the frequency responses of cat retinal ganglion cells. *J Physiol* 318: 161–179, 1981.
- Sharpee TO, Sugihara H, Kurgansky AV, Rebrik SP, Stryker MP, Miller KD.** Adaptive filtering enhances information transmission in visual cortex. *Nature* 439: 936–942, 2006.
- Srinivasan MV, Laughlin SB, Dubs A.** Predictive coding: a fresh view of inhibition in the retina. *Proc R Soc Lond B Biol Sci* 216: 427–459, 1982.
- Wark B, Fairhall A, Rieke F.** Timescales of inference in visual adaptation. *Neuron* 61: 750–761, 2009.
- Zaghloul KA, Boahen K, Demb JB.** Different circuits for ON and OFF retinal ganglion cells cause different contrast sensitivities. *J Neurosci* 23: 2645–2654, 2003.
- Zemon V, Gordon J, Welch J.** Asymmetries in ON and OFF visual pathways of humans revealed using contrast-evoked cortical potentials. *Vis Neurosci* 1: 145–150, 1988.