OPEN ACCESS Freely available online

PLOS | COMPUTATIONAL BIOLOGY

Transformation of Stimulus Correlations by the Retina

Kristina D. Simmons[1☯], Jason S. Prentice[2,3☯], Gašper Tkačik[4], Jan Homann[2], Heather K. Yee[5], Stephanie E. Palmer[5], Philip C. Nelson[2], Vijay Balasubramanian[1,2,6,7]*

1 Department of Neuroscience, University of Pennsylvania, Philadelphia, Pennsylvania, United States of America, **2** Department of Physics, University of Pennsylvania, Philadelphia, Pennsylvania, United States of America, **3** Princeton Neuroscience Institute, Princeton University, Princeton, New Jersey, United States of America, **4** Institute of Science and Technology Austria, Klosterneuburg, Austria, **5** Department of Organismal Biology and Anatomy, University of Chicago, Chicago, Illinois, United States of America, **6** Laboratoire de Physique Théorique, cole Normale Supérieure, Paris, France, **7** Initiative for the Theoretical Sciences, CUNY Graduate Center, 365 Fifth Avenue, New York, New York, United States of America

Abstract

Redundancies and correlations in the responses of sensory neurons may seem to waste neural resources, but they can also carry cues about structured stimuli and may help the brain to correct for response errors. To investigate the effect of stimulus structure on redundancy in retina, we measured simultaneous responses from populations of retinal ganglion cells presented with natural and artificial stimuli that varied greatly in correlation structure; these stimuli and recordings are publicly available online. Responding to spatio-temporally structured stimuli such as natural movies, pairs of ganglion cells were modestly more correlated than in response to white noise checkerboards, but they were much less correlated than predicted by a non-adapting functional model of retinal response. Meanwhile, responding to stimuli with purely spatial correlations, pairs of ganglion cells showed increased correlations consistent with a static, non-adapting receptive field and nonlinearity. We found that in response to spatio-temporally correlated stimuli, ganglion cells had faster temporal kernels and tended to have stronger surrounds. These properties of individual cells, along with gain changes that opposed changes in effective contrast at the ganglion cell input, largely explained the pattern of pairwise correlations across stimuli where receptive field measurements were possible.

Citation: Simmons KD, Prentice JS, Tkačik G, Homann J, Yee HK, et al. (2013) Transformation of Stimulus Correlations by the Retina. PLoS Comput Biol 9(12): e1003344. doi:10.1371/journal.pcbi.1003344

Editor: Matthias Bethge, University of Tübingen and Max Planck Institute for Biologial Cybernetics, Germany

Received February 26, 2013; **Accepted** September 11, 2013; **Published** December 5, 2013

Copyright: © 2013 Simmons et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Funding: This work has been supported by National Institutes of Health (www.nih.gov) training grants 5-T90-DA022763-05 and 5-T32-EY007035-32; National Institutes of Health grant P30 EY001583; National Science Foundation (www.nsf.gov) grants PHY-1058202, EF-0928048, and DMR08-32802; the C.V. Starr Foundation (http://www.starrfoundation.org/); the Aspen Center for Physics (http://www.aspenphys.org/); and the Fondation Pierre Gilles de Gennes (www.fondation-pgg.org). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

* E-mail: vijay@physics.upenn.edu

☯ These authors contributed equally to this work.

Introduction

An influential theory of early sensory processing argues that sensory circuits should conserve scarce resources in their outputs by reducing correlations present in their inputs [1–3]. At the same time, recent work has clarified that some redundancy in the retinal output is useful for hedging against noise [4,5]. Moreover, sensory outputs with varying amounts of correlation can engage cortical circuits differently and thus result in a different sensory "code" [6]. Thus, some degree of redundancy appears to be useful to the brain when dealing with response variability and making decisions based on probabilistic input [7]. Indeed, correlations between neurons in visual cortex are largely unchanged between unstructured and naturalistic visual stimuli [8]. In the absence of neural mechanisms supporting adaptation to the structure of sensory inputs, increased stimulus correlation would induce a corresponding change in neural correlation. Alternatively, the retina may dynamically adjust its coding strategy to represent the new stimulus class efficiently. To explore this possibility, we characterized the impact of stimulus structure on retinal output correlation. Previous studies have examined pairwise correlations amongst retinal ganglion cell spike trains in specific stimulus conditions [9–14] but did not report the changes in correlation for the same pairs across stimuli.

Are there mechanisms that might allow the retina to adjust its functional properties when stimulus correlations change? Traditionally, retinal ganglion cells (RGCs) have been described by a fixed linear receptive field followed by a static nonlinearity [15], where surround inhibition acts linearly to suppress pairwise correlations in natural visual input [2,3]. In this view, the receptive field and nonlinearities might vary dynamically with stimulus correlations, possibly by changing the strength of lateral inhibition to maintain a fixed amount of output correlation. Indeed, correlation-induced changes in receptive fields have been observed in the LGN and visual cortex [16,17].

To test these ideas, we performed a series of experiments in which we presented the retina with several stimuli with varying degrees of spatial and temporal correlations. The retina never fully decorrelated its input; even for the least correlated white noise checkerboard stimuli, some correlations were present between pairs of retinal ganglion cell spike trains. Responding to natural movies, however, output correlations were only moderately increased compared to correlations in responses to white noise checkerboards, despite the dramatic difference in input-induced correlations. Specifically, the differences in output correlations were much less than those predicted by a non-adapting linear-nonlinear functional model responding to these stimuli. We found

Author Summary

An influential theory of early sensory processing argues that sensory circuits should conserve scarce resources in their outputs by reducing correlations present in their inputs. Measuring simultaneous responses from large numbers of retinal ganglion cells responding to widely different classes of visual stimuli, we find that output correlations increase when we present stimuli with spatial, but not temporal, correlations. On the other hand, we find evidence that retina adjusts to spatio-temporal structure so that retinal output correlations change less than input correlations would predict. Changes in the receptive field properties of individual cells, along with gain changes, largely explain this relative constancy of correlations over the population.

a similar result for spatio-temporal exponentially correlated stimuli, with an even smaller change in output correlations. In this way, pairwise output correlations change by a relatively small amount as compared to the expected input-induced change in response to stimuli that span a broad range of spatio-temporal correlations. Stimuli with only spatial correlations, on the other hand, produced increases in output correlations similar to the input-induced changes predicted by a static, non-adapting functional model. In the extreme case, for temporally uncorrelated full-field flicker, the output correlation increased substantially. These results suggest a key role for temporal processing in maintaining the level of output correlations. Indeed, we observed a robustly faster response timecourse and a modest skew towards stronger inhibitory surrounds in response to spatio-temporally correlated stimuli. These changes were sufficient to largely explain the observed relative suppression of pairwise correlations in the retinal output for those experimental conditions where receptive field measurements could be made.

Results

Simultaneous measurements of ganglion cell responses

We used a multi-electrode array to measure simultaneous responses from groups of ~ 40 retinal ganglion cells in guinea pig; data and stimuli are available at [18]. Each recording interleaved 10-minute blocks of white noise checkerboard stimuli with 10-minute blocks of correlated stimuli. Example frames from each stimulus are shown in Fig. 1B, together with their respective spatial and temporal correlation functions. We probed retinal responses to natural movies, which allowed us to determine properties of ganglion cell population activity during natural vision. However, natural movies contain strong correlations in time (trace under "natural" stimulus in Fig. 1B) and space (Fig. 1A, B). There are challenges with reliably estimating receptive fields from natural stimuli due to these strong correlations and the highly skewed natural intensity distribution (see Methods). We therefore also assessed the effect of spatio-temporal correlations in a more

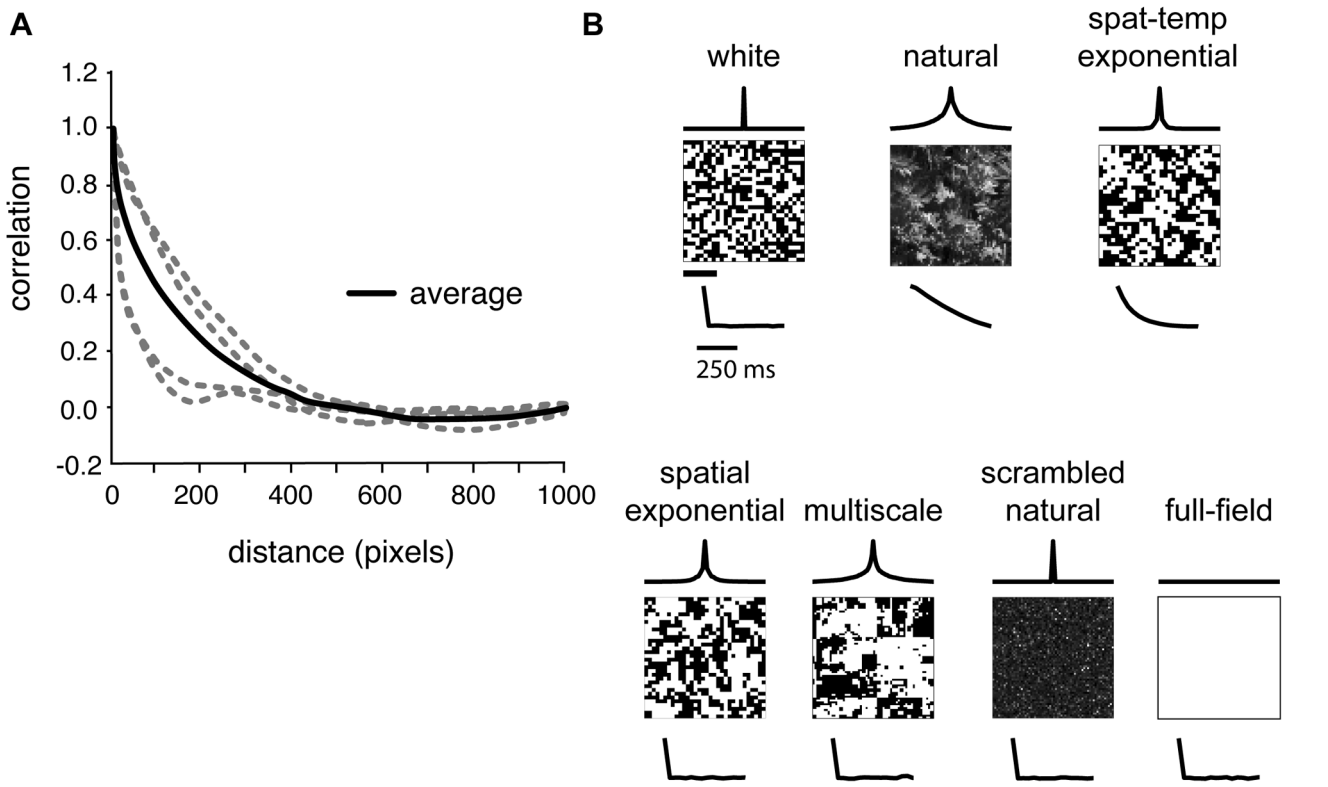

Figure 1. Natural and artificial stimuli vary in correlation structure. (**A**) Spatial correlation functions from four natural images (at higher resolution than the stimuli used in our experiments), in gray. Black line shows average correlation function over a large database of natural images. Although all images' correlation functions have the same general shape, there are clear differences between images. (**B**) Examples of the stimuli used in this work. Traces above frames show the spatial correlation function of each stimulus; traces below frames show the temporal correlation function. Stimuli were displayed at 30 Hz in alternating 10-minute blocks. Spatial scale bar (below white noise frame) for stimulus frames and spatial correlation functions is 400 μm; temporal scale bar for temporal correlation functions is 250 ms.
doi:10.1371/journal.pcbi.1003344.g001

controlled stimulus with short-range exponential correlations in time and space and a binary intensity distribution (Fig. 1B, "spat-temp exponential"). Additional stimuli allowed us to vary the spatial correlation over a broad range, without temporal structure, in order to test the hypothesis that surround strength adapts to remove correlations in nearby parts of an image. Thus, we examined spatial correlations, in the absence of temporal structure, of increasing extent: spatially exponential, a "multi-scale" naturalistic stimulus featuring structure over many spatial scales, and full-field flicker (Fig. 1B, bottom row). The multiscale stimulus was designed to mimic the scale invariance of natural scenes in a controlled binary stimulus, featuring both small and large patches of correlated checks (such as the white area near the center). Its construction is detailed in Methods. In one experiment, we also compared responses to low-contrast white and multiscale stimuli to their high-contrast counterparts. Finally, to control for the effect of the skewed natural intensity distribution, we also conducted experiments presenting scrambled natural movies lacking spatial and temporal correlation while preserving the intensity distribution. The mean luminance and single-pixel variance were matched across all stimuli other than natural movies, scrambled natural movies, and low-contrast stimuli. Over 30 minutes of recording in each stimulus condition, the typical cell fired ~ 7000 spikes. This was sufficient to assess spike train correlations and to measure receptive fields for the white and exponentially correlated stimuli.

For preliminary analyses, we measured the spike-triggered average (STA) from each ganglion cell's response to white noise. The resulting receptive fields typically gave good coverage of the sampled visual field (Fig. 2A) and clustered into classes on the basis of their response polarity and temporal properties (Fig. 2B; details in Methods). The four basic classes that we consistently identified across experiments were fast-ON and fast-OFF, distinguished by the transient and biphasic nature of their temporal filter, and slow-ON and slow-OFF, which had longer integration times and often less prominent biphasic filter lobes. It is possible that each of the functionally identified cell classes comprises multiple types of cells. Separating cells by class did not qualitatively change many of the results reported below; in these cases, we combined all cells to improve statistical power.

To probe the effect of stimulus correlation on ganglion cell response properties in detail, we applied a standard functional model, the linear-nonlinear (LN) model. In this model, the visual stimulus is filtered with a linear kernel that represents the spatio-temporal receptive field (STRF) of the cell. The filter output is then passed through a nonlinear transfer function to generate a predicted firing rate. The nonlinearity encompasses thresholding and saturation, as well as any gain on the linear response. For white noise stimuli, the STA is a good estimator of the STRF [19]. However, this simple property does not hold for correlated stimuli, and so we fit the STRFs and other LN model parameters by maximum likelihood estimation (see Methods). For the weakly correlated spatio-temporal exponential stimulus, this technique reliably extracted receptive fields (Fig. 2C).

Variation of output correlations with stimulus condition

We computed the correlation coefficient between spike trains (binned at 33 ms) for all pairs of simultaneously recorded neurons. In response to natural movies, correlations between most pairs of cells increased in magnitude when compared with the correlations between the same pairs when viewing white noise (Fig. 3A). We quantified the size of this increase by finding the least-squares best fit line (Fig. 3B, gray lines) and defining the "excess correlation" of a population as the slope of this line minus one (see Methods). If all cell pairs had, on average, the same correlation in both stimulus conditions, the excess correlation would be zero. Excess correlation was not strongly dependent on bin size (Fig. S2B). In the case of natural movies, the excess correlation was 0.32 ± 0.20 (95% confidence interval computed using bootstrap resampling, as explained in Methods; see Table 1), modestly different from zero (and significantly nonzero at the 95% confidence level).

Because the retinal ganglion cell output is a highly transformed representation of its input, it is not trivial to formulate a naïve expectation for the magnitude of output correlation given an input correlation. In particular, simply evaluating the input correlation between stimuli at the receptive field centers of a pair of cells provides a misleading picture, since it neglects the spatial extent and possible overlap of receptive fields. We therefore chose to quantify the output correlation expected for a given input in a simple null model: the LN model fit to the white noise responses. This model captures correlation due to receptive field overlap and simple nonlinear processing, while neglecting correlations due to shared circuitry and more complex nonlinear behavior, such as adaptation. For cells which had sufficiently well-estimated white noise LN model parameters (as described in Methods) we were able to compare the excess correlation predicted by the model to that observed in the data. In order to separate effects that might arise due solely to changes in firing rate between conditions (see Fig. S3A) from changes specifically in pairwise correlations between cells, we adjusted the threshold of each model neuron separately under each stimulus to match predicted average firing rates to their empirical values. All other parameters, namely the spatio-temporal receptive field and the gain, were unchanged between stimuli. This "non-adapting" model predicted a significantly larger excess correlation in response to natural movies (gray bars in Fig. 3D and Fig. 4A), suggesting that the low observed excess correlation value under natural stimulation is a consequence of nontrivial processing in the retina.

In addition to strong correlations, however, natural stimuli are also characterized by a skewed distribution with many dark pixels and a few extremely bright pixels, whereas our white noise stimulus, included equal numbers of bright and dark pixels. To disentangle effects of correlations from effects due to intensity distribution, we presented the same retinae with a scrambled natural movie. In this stimulus, we started with natural movies and randomly shuffled the pixels in space and time to maintain the intensity distribution but remove correlations. The excess correlation in response to this stimulus was consistent with zero in both the measured and simulated responses (Fig. 4A, left bars), suggesting that the skewed natural intensity distribution does not significantly affect output correlations. Moreover, comparing the natural movie and scrambled natural movie directly, we found a small excess correlation consistent with that in the natural movie vs. white noise case. The non-adapting model again predicted that the relative similarity of output correlations was nontrivial (Fig. 4A, right bars). Thus, the retina greatly suppresses changes in correlations of natural visual stimuli.

We found a similar set of results for the more weakly correlated spatio-temporal exponential stimulus (Fig. 3B). In particular, the excess correlation was low (0.12 ± 0.05) compared to the increase predicted by the non-adapting model (excess correlation of 0.67; Fig. 3D). We also examined the results of experiments in which we presented stimuli with varying degrees of spatial correlation. As shown in Fig. 3C, many stimuli produced only a modest increase in output correlations. Some stimuli with strong spatial correlations, particularly the multiscale and full-field flicker stimuli, resulted in a clear increase in output correlations when compared to white noise. When we varied the contrast of a white noise

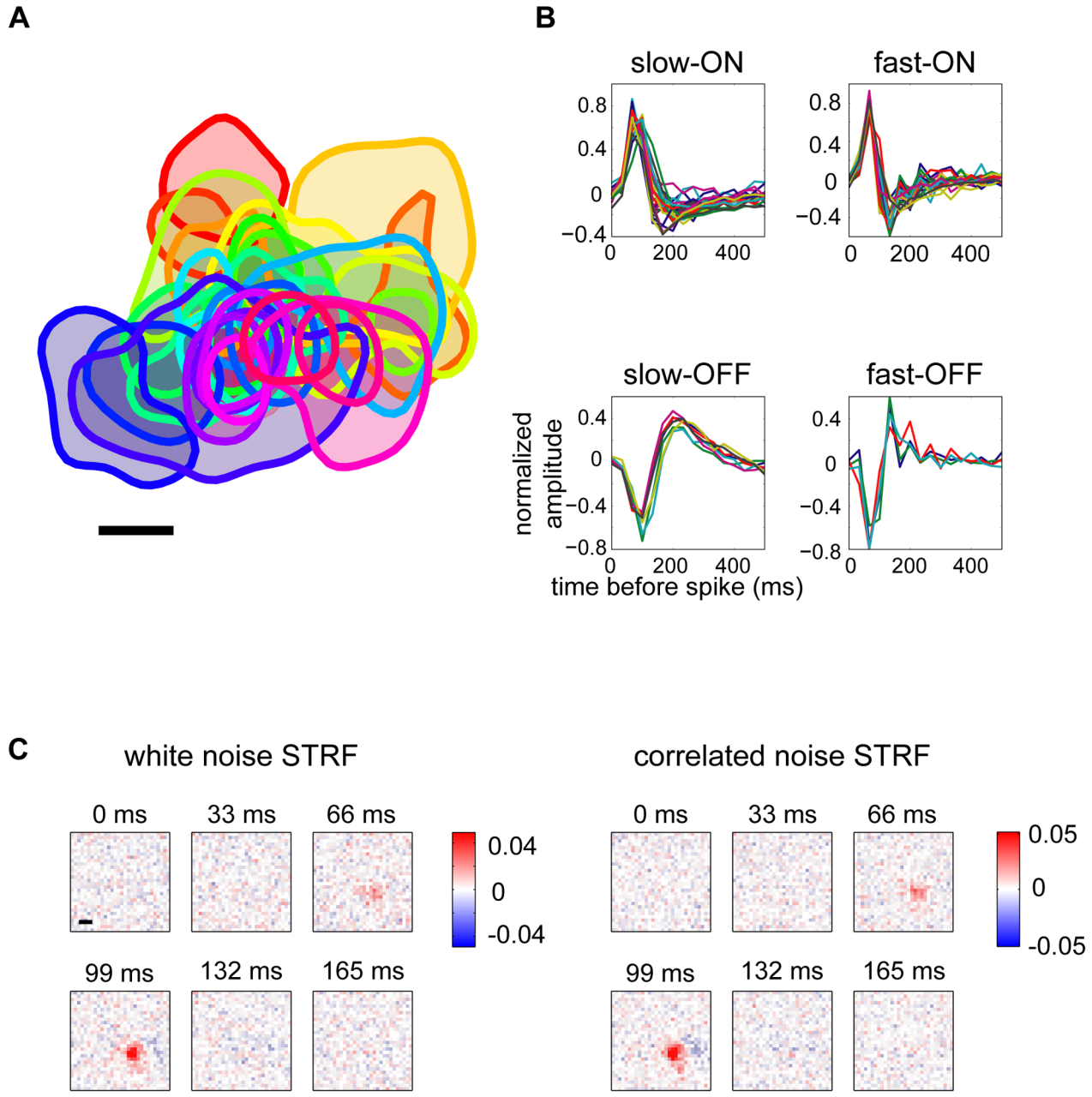

Figure 2. Retinal ganglion cell receptive fields measured using a multi-electrode array. (**A**) Receptive field locations of 31 cells recorded simultaneously from guinea pig retina. Each curve shows the 70% contour line of one receptive field. Scale bar is 200 μm. (**B**) Best-fitting temporal kernels for 75 cells, clustered into four classes. Classes were obtained by manually clustering temporal filters on the basis of the projection onto their first three principal components. (**C**) Maximum likelihood estimates of spatio-temporal receptive fields (STRFs) for an example cell. STRFs were computed separately using responses to white noise (left) or exponential spatio-temporally correlated stimuli (right). Scale bar is 200 μm.
doi:10.1371/journal.pcbi.1003344.g002

stimulus, output correlations decreased when the contrast was lowered while all other stimulus properties were kept fixed. Thus, the degree of correlation in the retinal output is not a reflection of stimulus correlations alone. On the other hand, decreasing the contrast of the multiscale stimulus did not significantly affect the output correlations, suggesting that stimulus correlation and contrast interact to shape output correlations.

For further analysis, we focused on the subset of stimuli shown in Fig. 3D, all of which were presented in experiments where we also obtained robust estimates of white noise receptive fields. Here we again simulated responses of an LN model using fixed receptive fields measured under white noise. For most stimuli, the model neurons showed changes in correlation at least as large as those observed in recordings. However, unlike the spatio-temporally correlated exponential and natural stimuli discussed above, the stimuli which had correlations in space only (spatial exponential and multiscale) or no correlations (scrambled natural movie) produced similar excess correlation values in recorded cells and in our non-adapting model. This suggests that a fixed linear filter, as in the non-adapting model, is largely sufficient to explain the effect of *spatial* correlations, whereas higher-order processing is necessary to suppress the impact of *temporal* stimulus correlations on output

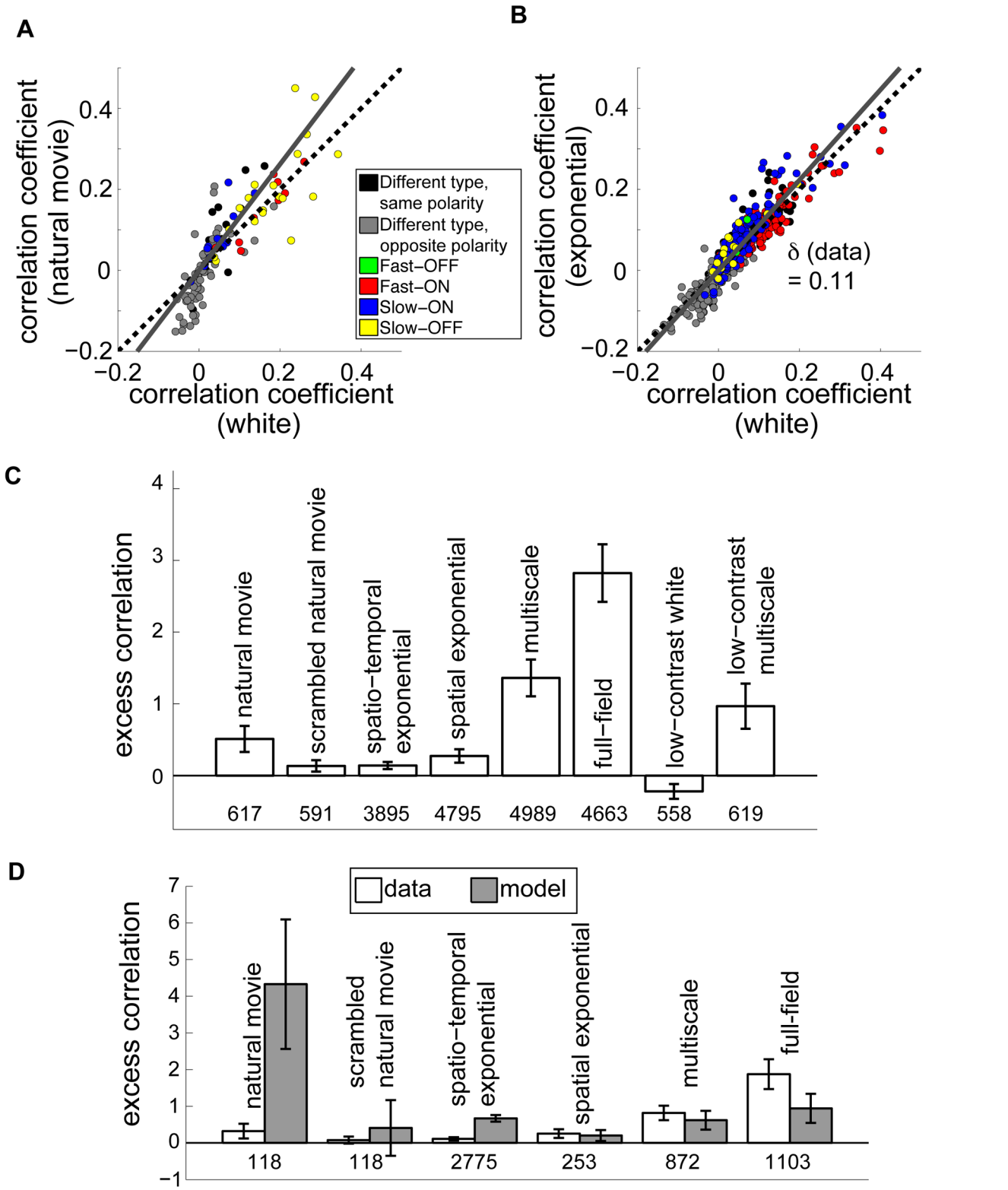

Figure 3. Retinal output correlations are largely constant between stimulus conditions. (**A**) Instantaneous spike train correlation coefficients between pairs of ganglion cells, comparing responses to natural movies and to white noise. Dashed black line is the diagonal. Cell pairs of the same class are indicated by colors in the legend. Different- class pairs are separated into ON-OFF (gray) and ON-ON or OFF-OFF pairs (black). The excess correlation, δ, is the deviation of the slope of the best fit line (gray) from the diagonal. (**B**) Same as (A) but for spatio-temporal exponentially correlated stimulus. (**C**) Excess correlation measured from ganglion cells responding to the indicated stimulus, compared to white noise. Numbers below bars indicate the number of cell pairs in each condition; all recorded cells are included. Error bars are 95% bootstrap confidence intervals computed over 50,000 random samples with replacement from the set of cell pairs. (**D**) Comparison of measured excess correlation (white) to non-adapting model predictions (gray) for the indicated stimuli. Model values were derived from LN neurons with parameters fit to white noise data. Only

cells whose receptive fields met a quality threshold are used here, in contrast to (C).
doi:10.1371/journal.pcbi.1003344.g003

correlation. For the spatially uniform full-field stimulus, output correlations appear to increase more than expected from the non-adapting model. Note, however, that the full-field data were collected as part of other experiments in which we presented white noise checkerboards for a shorter time (10 minutes, as opposed to 30 minutes). Thus the receptive fields are less well estimated and further studies are needed to verify with confidence the predictive performance of a non-adapting model.

As discussed above, we were able to identify the cell classes for many of our recorded neurons. In response to spatio-temporally exponentially correlated noise and natural movies, cell class had a modest effect on output correlations (Fig. 3A, B). Cells with opposite ON- or OFF- polarities (gray points) tended to have negative correlations, whereas cells of the same polarity (black and colored points) generally had positive correlations. Several opposite-polarity pairs did have positive correlation; these tended to have non-overlapping receptive field centers (Fig. S5). Moreover, pairs with opposite polarity showed a greater-than-average excess correlation, particularly in response to natural movies. Under natural movies, opposite-polarity pairs had an excess correlation of 1.5; under the spatio-temporal exponential stimulus their excess correlation was 0.38 (See Fig. S4A, B). Within same- class pairs, slow-ON and slow-OFF pairs (blue and yellow) tended to show a greater excess correlation than fast-ON and fast-OFF pairs (red and green). Pairs of slow cells had an excess correlation of 0.29 in the natural stimulus and 0.28 in the spatio-temporal exponential, while fast pairs were measured as 0.01 and -0.02 for the two stimuli, respectively. All of these class-dependent excess correlations were small compared to the overall non-adapting model predictions (excess correlations of 4.33 and 0.67 for natural and spatio-temporal exponential stimuli). We also assessed the relationship between receptive field separation and output correlation (Fig. 4C). Pairwise correlations tended to decay with distance, but the average change in output correlation between the correlated and white stimuli did not depend on distance.

Adaptation of temporal filters

We next sought to determine whether receptive fields adapt to stimulus correlations and whether this adaptation can explain the observed pattern of output correlations. As noted above, we were able to obtain STAs from responses to white noise. STAs computed in response to correlated stimuli, however, will be artificially blurred by the stimulus correlations. To obtain a better estimate of the spatio-temporal receptive field (STRF), we used maximum likelihood estimation to fit a LN model separately for the white and exponentially correlated stimuli [20]. Examples of STRFs obtained in this way for one cell are shown in Fig. 2C. The strongly correlated structure of the multiscale stimulus and the natural movies precluded robust, unbiased STRF estimation with limited data (see Methods). For this reason, we restricted any STRF computations to white noise and exponentially correlated noise. The latter stimulus is only weakly correlated and thus we would expect at most weak changes in the receptive fields between the conditions; indeed, receptive fields are hard to distinguish by eye for many cells. Measuring such weak changes requires high-quality receptive fields whose locations can be unambiguously determined (see Methods), as was the case for 75 neurons ($\sim$60% of the neurons recorded under spatio-temporal exponential correlated conditions). Cells that did not meet this standard were likely to include types that do not respond as well to checkerboard stimuli, e.g., direction selective ganglion cells and uniformity detectors. We included such cells in the analysis of Fig. 3C in order to maximize the generality of our results and to allow for the possibility that these neurons had qualitatively different output correlations. For the neurons that did pass the quality threshold, we found that the parameters of the LN model (for each neuron, a linear filter and a nonlinearity gain and threshold) changed with the stimulus.

Spike trains with sparse, transient firing events tend to be more decorrelated [14]. Motivated by this finding, together with our observation that temporally correlated stimuli yielded excess correlation in the non-adapting model that was higher than in the data, we analyzed adaptation in the temporal filtering properties of retinal ganglion cells. To isolate changes in temporal processing, we examined each neuron's STRFs (estimated separately under the white and exponentially correlated stimulus conditions) and extracted the temporal components (see Methods). These temporal profiles were faster for the correlated stimulus than for white noise (Fig. 5A). To quantify this difference, we computed the power spectrum of each neuron's temporal filter under each stimulus (Fig. 5B and 5D, top) and found a systematic increase in high frequencies under the correlated stimulus,

Table 1. Number of cells recorded in each condition.

stimulus	retinae	all cells	quality RFs[b]	fast ON[c]	fast OFF[c]	slow ON[c]	slow OFF[c]	excess corr.	conf. int.[d]
natural movie	3	84	34	12	0	9	13	0.51	±0.18
scrambled natural movie	3	82	34	12	0	9	13	0.13	±0.08
spatio-temporal exponential	5	212	75	29	4	31	8	0.14	±0.05
spatial exponential[a]	17	510	46	-	-	-	-	0.27	±0.09
multiscale[a]	16	513	62	-	-	-	-	1.36	±0.26
full-field[a]	14	483	276	-	-	-	-	2.82	±0.40
low-contrast white	1	47	-	-	-	-	-	−0.22	±0.10
low-contrast multiscale	1	49	-	-	-	-	-	0.97	±0.31

[a]For our measurements of output correlation (Fig. 3C), we include additional data from experiments performed as part of other studies in which receptive field structure was not probed. For model correlations and other analyses, we only used the subset of retinae and cells for which we obtained robust receptive field estimates.
[b]We used a stringent requirement that receptive fields (RFs) be of high quality for any analyses in which we used receptive field estimates (e.g. in Fig. 3D).
[c]Cells were only divided into classes if they had high-quality receptive fields and were recorded in response to stimuli chosen for detailed analysis.
[d]95% confidence intervals were computed by boot-strapping. See Methods for details.
doi:10.1371/journal.pcbi.1003344.t001

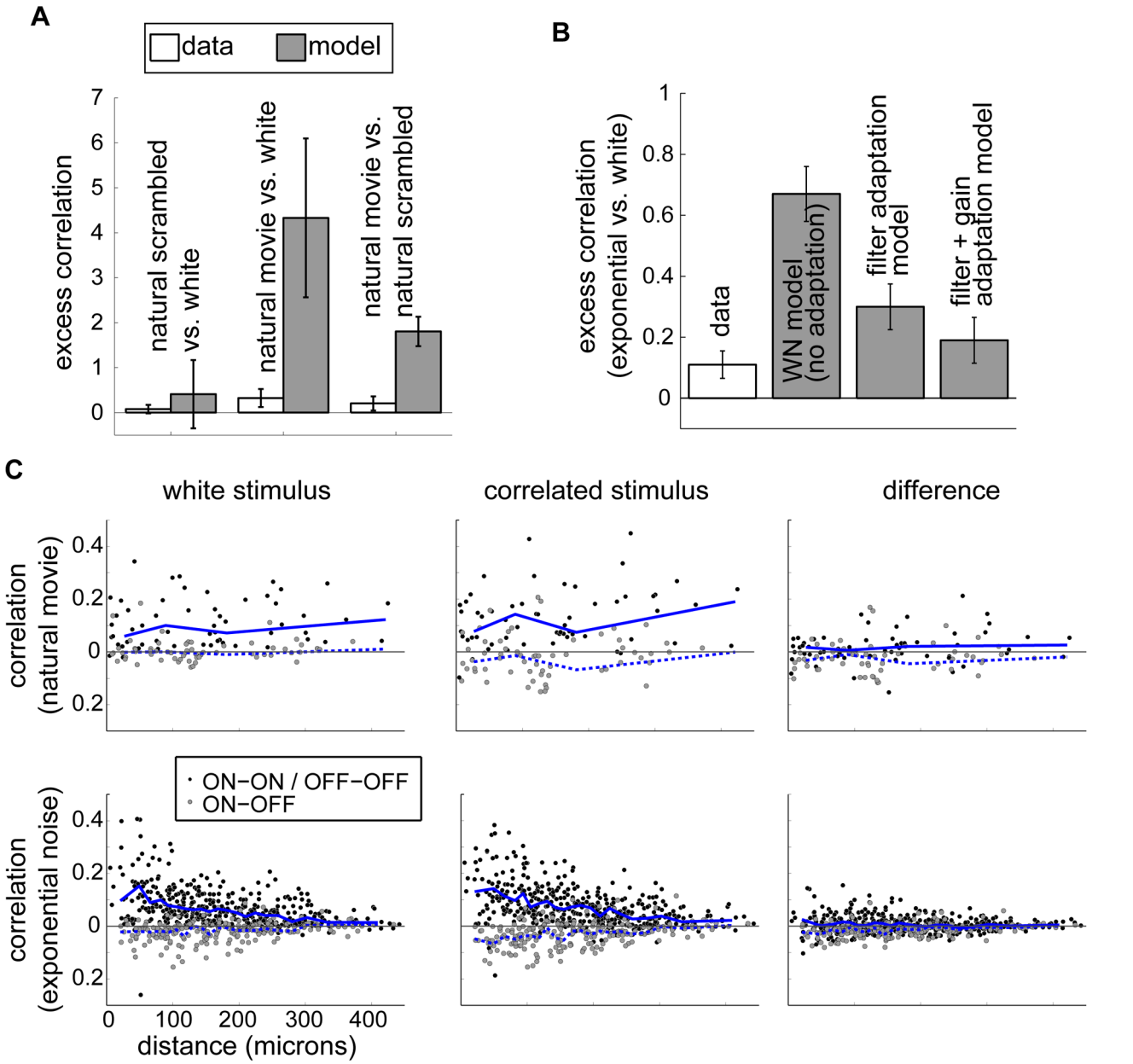

Figure 4. Analysis of pairwise correlations. (**A**) Excess correlations for natural stimuli. Left and middle bars show excess correlation when scrambled natural movies and intact natural movies, respectively, are compared to white noise in the data and in a population of non-adapting model neurons. Right bars show excess correlation when responses to natural movies are compared to scrambled natural movies directly. A non-adapting model predicts larger output correlations in response to the correlated natural input than seen in the data. (**B**) Output correlations under the spatio-temporal exponential stimulus compared with white noise as predicted by LN models with parameters fit to the data. The two leftmost bars (''data'' and ''WN model (no adaptation)'') reproduce the spatio-temporal ''data'' and ''model'' bars in Fig. 3D. (Note the difference in scale.) For the other bars, we simulated a population of neurons using linear filters measured from each stimulus but gains measured only from white noise (''filter adaptation model'') or using experimentally derived estimates of both linear filters and gains for each stimulus (''filter+gain adaptation model''). In the fully adapted model, excess correlations are consistent with the data. (**C**) Pairwise output correlation as a function of the distance between receptive field centers. Top row: Output correlations for white noise checkerboard (left) and natural movies (middle) and the difference in correlation between these conditions (right) for experiments where natural movies were presented. Bottom row: Output correlations for white noise checkerboard (left) and spatio-temporal exponential noise (middle) and the difference in correlation between these conditions (right) for experiments where spatio-temporal exponential noise was presented. Each point corresponds to one simultaneously recorded cell pair; within a row, the same pairs are represented in all three panels. Blue lines are the median correlation within bins chosen to contain 30 cell pairs each. Solid lines are median correlations for same-polarity cell pairs; dashed lines are for opposite-polarity pairs.
doi:10.1371/journal.pcbi.1003344.g004

indicating a shift toward high-pass filtering (Fig. 5E). As the correlated stimulus had relatively more power at low frequencies compared to the white stimulus, this form of adaptation compensates for differences in the power spectrum and, hence, tends to equalize output auto-correlations. In contrast, a non-adapting model with a filter estimated from white noise acting on

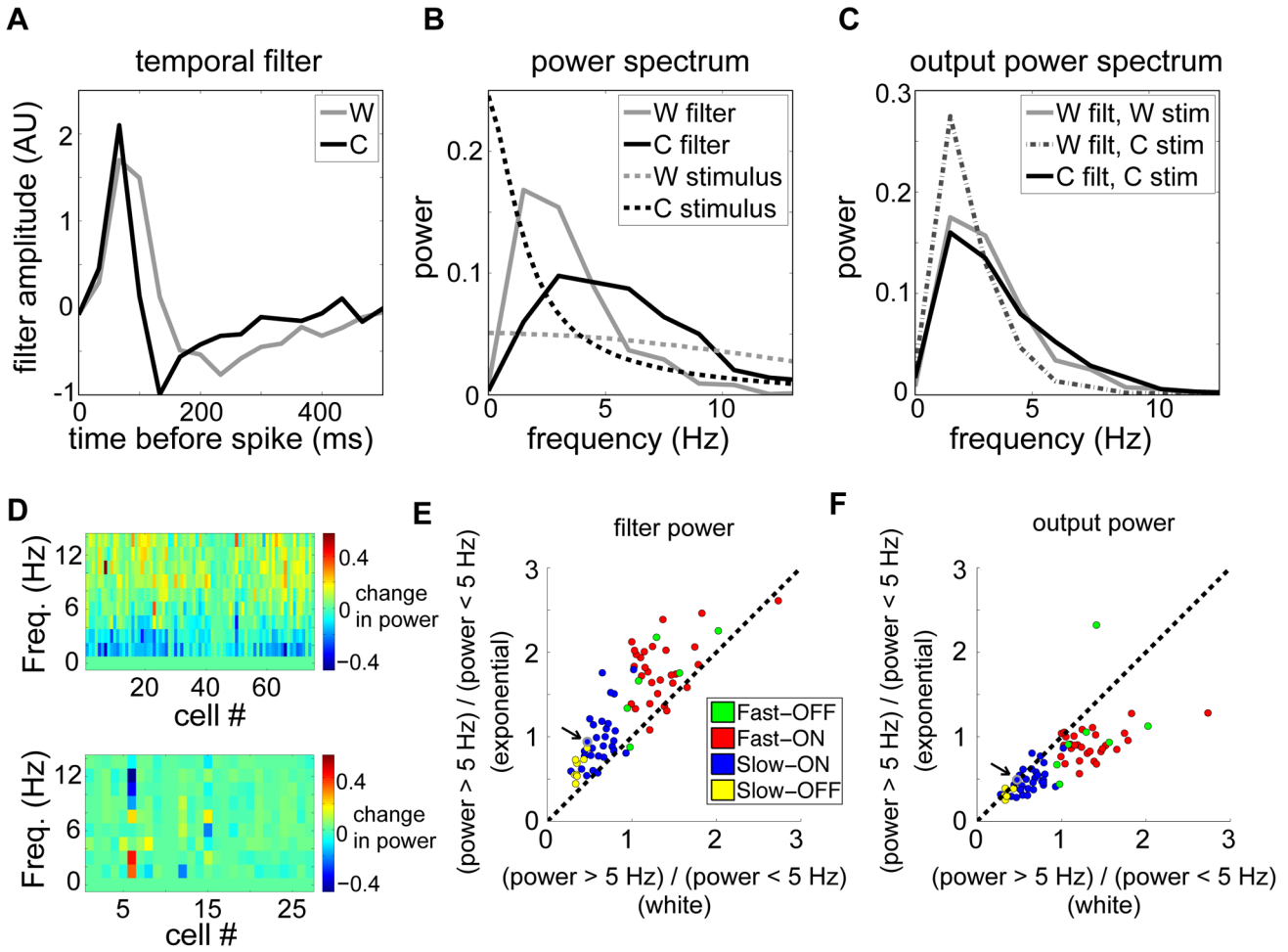

Figure 5. Adaptation of the linear temporal filter. (**A**) Temporal filters are faster under spatio-temporal exponentially correlated noise (C) than white noise (W). (**B**) Power spectrum of correlated noise input (C, black dashed line) has more low frequency power than white noise (W, gray dashed line). The power spectrum of the temporal filter for correlated noise (C, black solid line) has more high frequency power. (**C**) Power spectra of filter outputs: White-noise filter acting on white stimulus (solid gray); White-noise filter acting on correlated stimulus (dashed); Adapting correlated-noise filter acting on correlated stimulus (solid black). In adapted cases, output power spectra are similar between stimuli – i.e., temporal kernels compensate to maintain invariant output autocorrelation. (**D**) The difference in normalized filter power spectra between the correlated and white stimuli, for spatio-temporal (top) and spatial (bottom) exponential experiments. The power spectra of all filters in each stimulus were normalized by removing the DC component and dividing by the sum of squared amplitudes. The population change in temporal filters shows a consistent increase in high-frequency power relative to low-frequency power for the spatio-temporal, but not the spatial, stimulus. (**E**) Total power above 5 Hz divided by total power below 5 Hz for filters computed in response to correlated vs. white noise stimuli shows a shift towards high-pass signaling across the population. Arrow and gray circle indicate the pair shown in A–C. (**F**) Same analysis as in (E) applied to the filter output in (C). Points near the diagonal indicate near-complete compensation for stimulus changes; points below the diagonal indicate incomplete compensation.
doi:10.1371/journal.pcbi.1003344.g005

the correlated stimulus predicts large changes in the output power spectrum (Fig. 5C). Indeed, this compensation was nearly exact for many cells (Fig. 5C), though generally incomplete over the full population (Fig. 5F). These results, combined with those of [14], may indicate that the observed consistency of output correlations is produced by an increase in response transience when stimulus correlations increase.

Next, we found separate temporal profiles for the center and surround and computed the latency, measured as time to peak, of each. Surround latencies did not differ between white noise and spatio-temporally exponentially correlated noise (t-test, $p=.7$, $n=75$). However, center latencies l were shorter for correlated noise. We quantified the shift in terms of an adaptation index $(l_{corr}-l_{white})/(l_{corr}+l_{white})$. The histogram of the adaptation index (Fig. 6A; mean $=-0.03$, std $=0.03$; t-test $p<10^{-12}$, $n=75$; Wilcoxon signed rank text $p<10^{-10}$) showed a robust tail toward shorter center latency for correlated stimuli (skewness $=-0.53$). Moreover, almost every cell from which we obtained receptive fields had a longer latency for white noise than for correlated noise (Fig. 6B; mean change $=6.1$ ms). This was true across cell classes.

To determine whether these changes in temporal filtering were due to the presence of temporal correlations in this particular stimulus (unlike many of the other stimuli we examined), we also measured receptive fields from a separate population of ganglion cells responding to white noise and to a stimulus that was exponentially correlated in space but not in time. In this case, filters did not show a systematic change in power spectra (Fig. 5D, bottom), but the center latencies were shorter for the correlated stimulus (Fig. 6C; mean change $=7.2$ ms). Again, computing adaptation indices indicated that this effect was significant (mean $=-0.04$, std $=0.03$; t-test $p<10^{-10}$, $n=37$; Wilcoxon signed rank test $p<10^{-7}$). Thus, spatial correlations in the

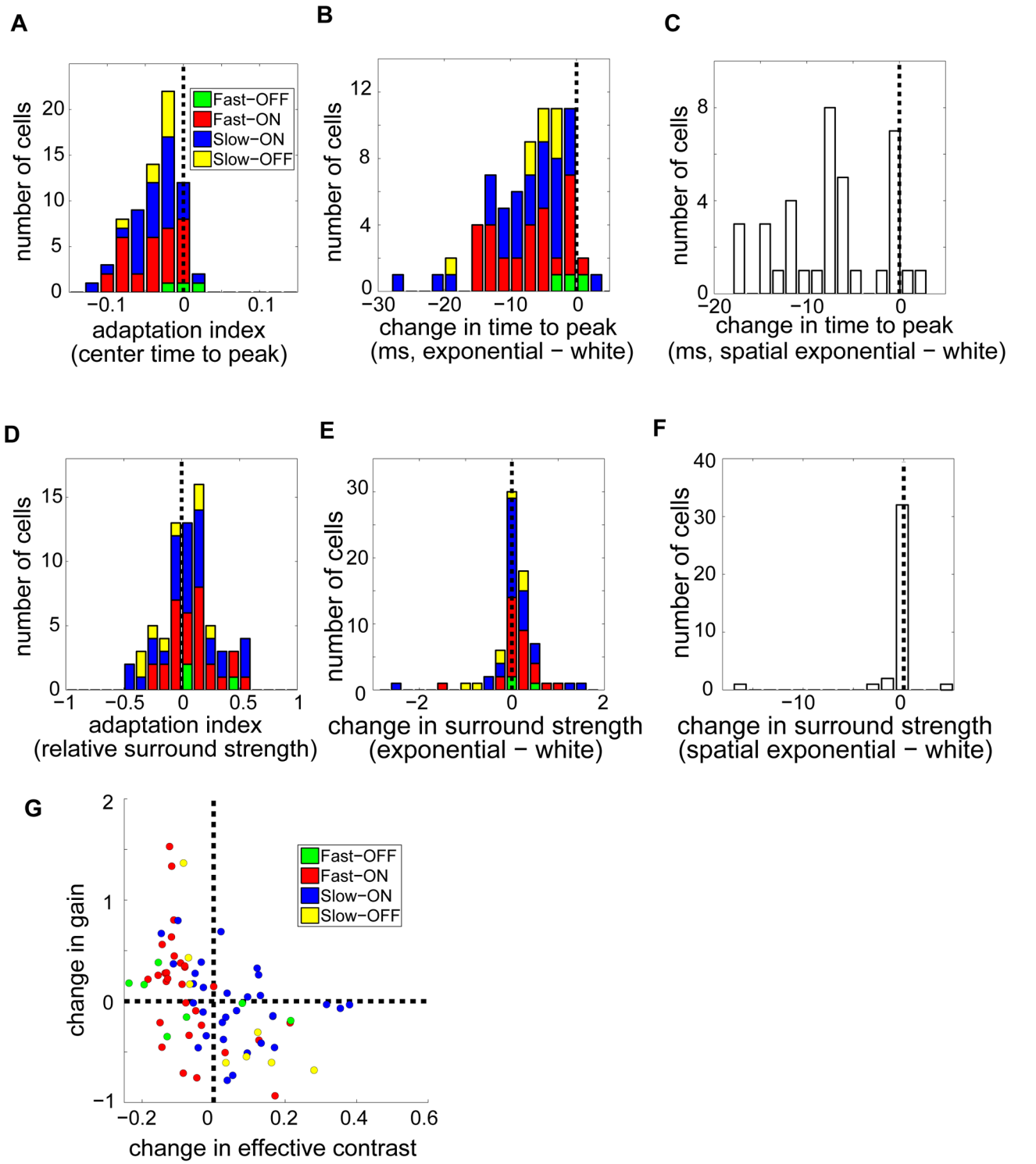

Figure 6. Adaptation of the spatio-temporal receptive field and gain. (**A**) Center latency (time to peak of the temporal kernel) is shorter for spatio-temporal exponentially correlated noise. Histogram shows adaptation indices (corr−white)/(corr+white) for center latency ($n=75$). (**B,C**) Changes in center latency (corr−white) for spatio-temporally correlated (B) and temporally correlated (C) stimuli, in milliseconds. Almost all cells have a decreased time to peak when responding to a correlated stimulus. (**D**) Adaptation indices, computed as in (A), for relative surround strength (surround/center ratio) show a slight skew toward a stronger surround for spatio-temporally correlated noise ($n=75$). (**E, F**) Difference in surround strength for the spatio-temporal (E) and spatial (F) exponential stimuli. (**G**) Gain adaptation. Gains were defined as the slope of the LN model nonlinearity, and obtained separately for the response to white noise and to the spatio-temporally correlated exponential stimulus. Effective contrast, the standard deviation of the linear filter output, was similarly measured in both stimuli. The difference in gain, correlated value minus white noise value, is plotted against the difference in effective contrast. Increases in effective contrast tend to invoke compensating decreases in gain ($n=75$).
doi:10.1371/journal.pcbi.1003344.g006

stimulus appear to produce a shift in the timing of the response without changing the shape of the filter (as measured by the power spectrum). The lack of an effect in the power spectrum may explain why output correlations for the spatially exponential stimulus are not reduced relative to the change in correlation for a non-adapting model (Fig. 3D).

Adaptation of spatial receptive fields and nonlinearity gain

The conventional view of retinal circuitry suggests that adaptive decorrelation arises from stronger or wider surround inhibition during viewing of correlated stimuli. We thus computed the amplitudes of the surround and center components of each neuron's STRFs in both white noise and spatio-temporally exponentially correlated noise. Defining the relative surround strength, k, as the ratio of surround amplitude to center amplitude (details in Methods), we computed an adaptation index for each cell as $(k_{corr}-k_{white})/(k_{corr}+k_{white})$. This adaptation index has a modestly positive mean (Fig. 6D; $\text{mean}=0.075$, $\text{std}=0.24$; two-tailed t-test, $p=0.008$, $n=75$; Wilcoxon signed rank test $p=.008$), as do the changes in surround strength themselves (Fig. 6E). In addition, the cells with the greatest degree of surround adaptation had a robust tendency to increase in surround strength ($\text{skewness}=0.15$). There was no discernible dependence on cell class. Interestingly, the surround strength showed only a marginally significant change when spatial correlations (but not temporal correlations) were added to white noise (Fig. 6F; $\text{mean adaptation index}=-0.087$, $\text{std}=0.26$; two-tailed t-test, $p=0.05$, $n=37$; Wilcoxon signed rank test $p=.02$). Thus, while we do find some evidence for an increase in surround strength with stimulus correlation, the effect is subtle. This outcome is surprising given the common view since the work of Barlow [1,2] that surround inhibition is primarily responsible for decorrelation of visual stimuli. However, it is possible that the exponential correlations that permitted us to estimate receptive fields are too weak to evoke strong lateral inhibition.

Finally, we examined the gain g of each neuron, defined as the maximum slope of the logistic nonlinearity fit to each neurons' response (see Methods). Since the gain enters the nonlinearity after the stimulus passes through the linear filter, we normalized the filter to unit euclidean norm in order to obtain an unambiguous definition of g. We found that the gains of individual neurons changed when the stimulus was more correlated, but there was not a systematic change between stimuli. Recall that the gain of many sensory neurons, including retinal ganglion cells, is known to change with the contrast of the stimulus [21,22]. To test for a possibly related mechanism at work in our data we first defined "effective contrast," σ_{white} and σ_{corr}, as the standard deviation of the normalized linear filter output in each stimulus, respectively. This notion of effective contrast roughly captures the variability of the ganglion cells' input, taking presynaptic processing into account. Any nonlinear gain control mechanism in the ganglion cell layer should therefore be sensitive to this quantity. For some cells σ_{corr} exceeded σ_{white}, while for others the reverse was true. Measuring the gains in both stimulus conditions (g_{white} and g_{corr}), however, we found systematic adaptation opposing the change in effective contrast: gain tended to increase when effective contrast decreased and vice-versa. Specifically, the quantities $\Delta g=g_{corr}-g_{white}$ and $\Delta\sigma=\sigma_{corr}-\sigma_{white}$ were significantly anticorrelated (Fig. 6G; Spearman's $\rho=-0.54$, $p=10^{-6}$, $n=75$).

Output correlations in an adapting model

Finally, we assessed whether the receptive field changes reported above could account for the observed modest increase in output correlations between white noise and the spatio-temporal exponential stimulus. For experiments using spatio-temporally exponential noise, as discussed above, we measured the adaptation in LN model parameters fit to each stimulus. We then separately examined the effect of adaptation in different parameters on the excess correlations predicted by the LN models. Including adaptation of the linear filters, but not the gain, produced a significantly improved match between the model and the data (Fig. 4B, "filter adaptation model"). Additionally allowing the gain to adapt produced output correlations consistent with the data (Fig. 4B, "filter+gain adaptation model"). The contribution of gain adaptation to decorrelation is interesting in light of our observation that output correlations are lower for stimuli with lower contrast (Fig. 3C). Low contrast stimuli generally evoke lower firing rates, which could result in decreased pairwise correlations absent any change in linear filtering properties. (See Text S1 for a derivation of this result and Fig. S3B for evidence that excess correlation tends to increase with, but is not fully determined by, average firing rate.) At the same time, changes in contrast lead to gain control, wherein gain is higher for lower stimulus contrast. This gain adaptation could also affect output correlations, as in Fig. 4B. It would be interesting to know how gain control interacts with changes in other properties, such as the nonlinearity threshold and the shape of the linear filter, to set the correlations in the retinal response. Note that the LN model is fit to each neuron independently, without taking correlations between neurons into account. Its successful prediction of the change in pairwise correlations, without explicit introduction of inter-neural interactions, is therefore noteworthy. We conclude that observed adaptation in receptive fields and gains is adequate to explain the output correlations in responses to a spatio-temporally correlated stimulus.

Discussion

Our principal finding is that the mammalian retina maintains a moderate level of output correlation across a wide range of spatio-temporally correlated stimuli ranging from white noise checkerboards (with limited correlations) to natural movies (with wide spatial and temporal correlations). While the amount of output correlation varies between stimuli, the changes are much less than predicted by a non-adapting linear-nonlinear functional model. Our data also suggest a differential effect of spatial versus temporal correlations on the functional properties of the retinal output. We focused here on spatial variations in our stimuli, but it would be interesting to design future studies to explore the space-time differences more systematically. In response to spatio-temporal exponential noise, where the receptive fields could be estimated, we showed that the relative invariance of output correlations is largely accounted for by the observed changes in the linear receptive field (faster temporal kernels and slightly stronger surround inhibition for more correlated stimuli) and by changes in the nonlinear gain (anti-correlated to changes in effective contrast). The latter findings give an interpretation of the results in terms of a conventional functional model (here a linear-nonlinear cascade), but the measurement of output correlations is model-independent.

Classifying cells into classes revealed a slight dependence of excess correlation on cell class: most robustly, opposite polarity ON-OFF pairs showed the greatest increase in correlation magnitude when stimulus correlation increased. Indeed, if the retinal output is split across parallel functional channels, redundancy is likely to be highest within a channel due to shared circuit inputs. It may thus be advantageous, from an information

encoding perspective, for decorrelation to act *within* a channel, with residual correlations across classes signaling to downstream areas relevant relationships between the information in different channels.

Pitkow and Meister [14] showed that salamander retina partly decorrelates naturalistic inputs but that the response to white noise is more correlated than the input, in part due to receptive field overlap between ganglion cells. Here we demonstrated a similar phenomenon in mammalian retina: consistent with their results, we found that changes in output correlations were often smaller than changes in input correlations. We also extended their findings by showing explicitly that this partial decorrelation occurs in individual pairs of neurons. In [14], it was also suggested that the linear receptive field measured from white noise was insufficient to explain the amount of decorrelation seen for naturalistic stimuli, and the bulk of the decorrelation was attributed to changes in the threshold of a functional model of ganglion cells. However, the authors did not directly measure the (possibly different) receptive fields of ganglion cells responding to correlated stimuli, nor did they follow particular cell pairs across different stimuli. Our measurements suggest that the nonlinear processing proposed in [14] can be described in terms of adaptation of the linear receptive field and nonlinear gain with the net effect that output correlations are reduced relative to the expected input-induced correlations, as was observed in visual cortex by [8]. Our results also recall those of [16], [23], [17], and [24], who showed that receptive fields in LGN and primary visual cortex differ in structure when probed with natural movies versus random stimuli.

We also found that the gain of retinal ganglion cells responding to correlated stimuli changes with "effective contrast" σ_{white} and σ_{corr}, i.e. with the standard deviation of the input to the nonlinearity in a linear-nonlinear model of ganglion cells. In classical contrast gain control, firing rates and response kinetics adapt to temporal contrast and to the spatial scale of stimuli [21,22]. As increased stimulus correlation may produce a qualitatively similar input to the inner plexiform layer as increased contrast, some of the cellular mechanisms underlying contrast adaptation might also contribute to the phenomena we have uncovered. This provides an avenue for future study of the functional mechanisms underlying adaptation to correlations.

We have focused in the present work on the failure of a non-adapting linear-nonlinear model to capture the relatively small scale of observed excess correlations and have seen that adaptation in the linear filter might remedy this discrepancy. Alternatively, shared circuitry in the population of neurons may be engaged by correlated inputs and require explicit inclusion in any functional model of retinal responses to different classes of correlated stimuli [9,25]. Such shared circuitry leads to noise in one neuron being passed to multiple nearby neurons and is thus measured by "noise correlations." While addition of fixed, stimulus-independent noise correlation would not greatly change our results, a *change* in noise correlation with stimulus would provide a different candidate mechanism to account for our data [26]. This is another avenue for future work.

We have focused here on the effects of spatial correlations in an experimental design where we could compare receptive fields computed from responses to two different stimuli. Thus, we used relatively weak exponential correlations to ensure that we were not measuring artifacts of the stimulus correlations themselves. Recovering receptive fields from strongly correlated stimuli can require long recording times. Because our experimental design involved comparisons between several different stimuli, we were only able to recover receptive fields for moderately correlated stimuli. Future work could simply present each stimulus for a longer duration to assess receptive field changes at a population level rather than analyzing multiple stimuli in one experiment. In such experiments with more data from each cell, alternative methods of receptive field estimation such as Maximally Informative Dimensions [17] could potentially be applied. Further work could also include parallel studies with stimuli including temporal correlations only to complement our findings on responses to spatial correlations.

Finally, it would be interesting to determine the timecourse of the adaptations observed here. Knowing whether a change in stimulus correlations induces changes in receptive fields and output correlations within seconds, tens of seconds, or longer would help to clarify the relationship between processing of correlations and adaptation to other stimulus features such as contrast. Again, the design of our experiments precluded making these measurements – we focused on long segments to measure steady-state processing of correlations, whereas assessing the timecourse of changes requires finer and more systematic sampling of transitions between stimuli.

Why would the retina need to adapt, in the behaving animal, to variations in spatial correlations? While natural scenes are scale-invariant on average, the specific correlations do vary depending on the scene and the viewing distance (see Fig. 1A). Barlow originally suggested that sensory systems should decorrelate their inputs to make efficient use of limited neural bandwidth [1]. Consistent with this idea, we found that retina removes redundancies in spatio-temporally correlated stimuli but also that the retinal output is not completely decorrelated. Rather, the output correlations are reduced to a lower level, roughly similar to correlations in responses to white noise checkerboards when considered relative to the much larger input-induced correlations predicted by a non-adapting functional model of neural response. What drives this tradeoff? Recall that redundancy can be useful to protect against noise, to facilitate downstream computations, or to enable separate modulation of information being routed to distinct cortical targets. Thus, it may be that a certain degree of output correlation between retinal ganglion cells represents a good balance between the benefits of decorrelation and the benefits of redundancy [5]. Sensory outputs with varying amounts of correlation may also be decoded differently by cortex [6], in which case maintaining a fixed visual code might require that retinal output correlations are within the range expected by downstream visual areas. In these interpretations, it makes sense that the retina adapts to maintain correlation within a relatively narrow range across stimulus conditions, as we have found.

Methods

Ethics statement

This study was performed in strict accordance with the recommendations in the Guide for the Care and Use of Laboratory Animals of the National Institutes of Health, as well as the guidelines of the American Veterinary Medical Association. The protocol was approved by the Institutional Animal Care and Use Committee of the University of Pennsylvania (Permit Number 803091). All surgery was performed under anesthesia, and every effort was made to minimize suffering.

Neural recording

We recorded retinal ganglion cells from Hartley guinea pig using a 30-electrode array (30 μm spacing, Multi Channel Systems MCS GmbH, Reutlingen, Germany). After anesthesia with ketamine/xylazine (100/20 mg/kg) and pentobarbital

(100 mg/kg), the eye was enucleated and the animal was euthanized by pentobarbital overdose. The eye was hemisected and dark adapted. The retina was separated from the pigment epithelium, mounted on filter paper, and placed onto the electrode array, ganglion cells closest to the electrodes. Extracellular signals were recorded at 10 kHz. The retina was maintained in well-oxygenated bath of Ames' medium at a temperature of 37°C. The health of the preparation was monitored by tracking the average firing rates of active cells. Recording times were 2–4 hours, a typical duration over which the guinea pig retina preparation remains robustly responsive. We tested the consistency of responses offline by comparing activity levels near the beginning and end of the recording. We also measured the responses to a flash of light immediately before and after presentation of our main experimental stimuli. If any of these measures changed greatly, we took this as a sign of poor health and discarded the corresponding dataset. Spike times were extracted with the spike-sorting algorithm described in [27]; briefly, a subset of data was manually clustered to generate spike templates that were then fit to the remaining data using a Bayesian goodness-of-fit criterion.

Data are available from the Dryad Digital Repository: http://dx.doi.org/10.5061/dryad.246qg.

Stimulus generation

We displayed checkerboard stimuli (see Fig. 1B) at 30 Hz on a Lucivid monitor (MicroBrightField inc., Colchester, VT) and projected the image onto the retina. The mean luminance on the retina was $9000\ \text{photons/s}\cdot\mu\text{m}^2$ (low photopic); each check occupied an area between 50 μm×67 μm and 100 μm×133 μm. To make white noise and exponentially correlated stimuli, we first produced random checkerboards with intensities drawn from a Gaussian distribution. Spatio-temporally correlated stimuli were produced by filtering the Gaussian stimulus with an exponential filter with a time constant of three stimulus frames (99 ms) and a space constant of one check to match the scale of typical receptive fields. Stimuli with only spatial exponential correlations were constructed similarly, but with a time constant set to zero. To create the multiscale stimulus, we first generated gaussian white noise checkerboards at several power-of-two scales. The largest scale consisted of a single check filling the entire stimulus field, the next largest was a 2×2 checkerboard (with check size equal to half the stimulus field), the third largest was a 4×4 checkerboard (check size one quarter of the stimulus field), and so on. The checkerboards at all scales were then summed and thresholded to obtain a binary stimulus qualitatively mimicking the scale-invariant structure of spatial correlations in natural scenes (Fig. 1B). This stimulus did not contain temporal correlations. Natural movies of leaves and grasses blowing in the wind were collected with a Prosilica GE 1050 high-speed digital camera with a 1/2″ sensor (Allied Vision Technologies GmbH, Stadtroda, Germany) connected to a laptop running StreamPix software (NorPix Inc., Montreal, Canada) to grab frames at 60 fps. The camera resolution was set to 512×512 pixels, and movies were filmed from a fixed tripod about 5 feet from the trees and grass. Natural light was used to illuminate our outdoor scenes, and exposure time was set ($300-3000\ \mu\text{s}$) to capture variation in shadows and avoid saturation of our 8-bit luminance depth. Videos were collected for up to 30 minutes; 10 second to 5 minute segments with continuous motion were selected. Videos were downsampled to match the resolution and frame rate of our stimulus monitor. When we analyzed responses to movies taken from different settings, we did not see a sizable change in output correlation (Fig. S4C); thus, we combined all natural videos in our analysis. To produce a scrambled control for natural movies, pixels were randomly shuffled in space and time to remove all correlations. All stimuli other than natural movies (intact and scrambled) were thresholded at the median to fix the mean luminance and single-pixel variance and to maximize contrast. This binarization did not affect the power spectra significantly. For low-contrast stimuli, all deviations from the mean luminance were halved to give an overall contrast of 50%. Typically, we alternated 10-minute blocks of white noise with 10-minute blocks of a correlated stimulus.

Cell typing

We used reverse correlation to compute the spike-triggered average (STA) for each cell responding to white noise. We performed principal component analysis (PCA) on the best-fitting temporal kernels and split cells into two clusters based on the sign of the first component; the clusters were identified as ON and OFF classes based on the sign of their temporal kernels. (Our undersampling of OFF cells [4,28] may be due to electrode bias, as individual OFF cells are smaller and therefore less likely to be detected by our electrode array.) PCA was repeated for the ON and OFF groups separately. We manually identified clusters based on the projections onto the first three principal components; in this way we identified four functional classes, including slow-OFF, fast-OFF, fast-ON, and slow-ON (see Fig. 2B).

Maximum likelihood estimation of linear-nonlinear models

To obtain spatio-temporal receptive fields (STRFs) for both white and exponentially correlated stimuli, we used publicly available code (strflab.berkeley.edu; [20]) to carry out maximum likelihood estimation. We parameterized the model by a linear filter acting on the stimulus and a logistic nonlinearity, so that firing probability is $p(s)=1/(1+\exp(-g*(s-b)))$, where s represents the linear filter output, and g and b are gain and offset parameters. Gradient ascent with early stopping was used to compute a maximum likelihood estimate of the linear filter that best fit the data. We initialized the algorithm for each neuron using the spike-triggered average recorded in response to white noise. Many cells do not yield clear receptive fields when probed with white noise, either because this stimulus does not evoke a sufficiently strong response or because the response is not well modeled as a single linear filter. To avoid potential artifacts that could arise from including such cells in our receptive field and model analyses, we selected cells whose receptive fields had clearly visible centers. This classification of receptive fields as high-quality was done before any other data analysis in order to avoid biasing the selection. In datasets where we obtained receptive fields for both white noise and a correlated stimulus the designations of high-quality agreed between the two stimuli for 98% of cells. The subset of cells identified in this way also had center locations that were clearly delineated by our automated receptive field analysis, giving confirmation of our visual threshold.

The STRF baseline was poorly constrained by the maximum likelihood procedure, since an additive change in the STRF has a similar effect to a proportional shift in the offset parameter of the nonlinearity. We therefore normalized the STRFs by subtracting an estimate of the baseline: we allowed the fit to include components extending 100 ms after the spike — where the true filter must be zero by causality — and subtracted the mean of these frames. Inclusion of these post-spike frames also allowed us to verify that the temporal autocorrelations in our stimuli did not produce any acausal artifacts in the recovered STRFs. We normalized the estimated linear filters to have unit Euclidean norm (square root of the sum of squares of filter values) and then

used gradient ascent to separately fit the gain and offset of a logistic nonlinearity. Since the likelihood function in this case is convex, there was no possibility of local maxima. While we were able to compute unbiased estimates of STRFs from responses to stimuli with exponential correlations, our multiscale and natural movie stimuli were too correlated to estimate unbiased receptive fields with the number of spikes we were able to obtain in a single recording. Maximally Informative Dimensions, an important alternative receptive field estimation method [29], would similarly be constrained by the number of spikes obtainable when exploring multiple stimulus conditions in a single recording session, as we have done.

Correlation analysis

Correlations were measured as the correlation coefficient between pairs of simultaneously recorded neurons. Spike trains were divided into 33 ms bins; we assigned a bin a 1 if it had one or more spikes and a zero otherwise. The results reported above did not change if we used spike counts in each bin rather than binarizing. Indeed, 98% of timebins had one or fewer spikes and less than 0.05% had more than three spikes. We summarized the results by finding the best fit line of the form $\rho_{corr} = (1+\delta)\rho_{white}$, where ρ_{white} and ρ_{corr} are the pairwise correlations under the white and correlated stimuli, respectively. We estimated the excess correlation, δ, by the total least squares regression method and computed 95% bootstrap confidence intervals from 50,000 bootstrap resamples of the set of ganglion cell pairs.

Such instantaneous correlations are thought to combine slow stimulus-driven effects with fast intrinsic effects due to shared noise [13]. To verify that this did not affect our results, we isolated the stimulus-driven component, by noting that our cross-correlation functions can feature a short-timescale peak riding on a slow component and extracting the latter. Specifically, we binned the spike trains into 1 ms bins and computed cross-covariance functions between pairs. To isolate the stimulus-induced component, we smoothed the cross-covariance functions by fitting a cubic B-spline curve with knots spaced at 20 ms to suppress the fast noise component. We then found the shift, within a 200 ms window, which maximized the absolute value of the smoothed cross-covariance and estimated the correlation coefficient as the cross-covariance at this shift normalized by the product of the standard deviations. This gave excess correlation values consistent with those reported above (not shown).

We also computed the power spectra of the stimuli, the best-fitting temporal kernels, and the filter outputs (i.e. stimulus power spectra multiplied by filter power spectra). We summarized each power spectrum by computing the total power above 5 Hz divided by the total power below 5 Hz.

Measures of receptive field characteristics

Given a STRF estimated for one cell under one of the stimulus conditions, we first performed principal component analysis on the timecourses of the individual pixels. From the resulting set of "principal timecourses" we selected the one most similar to the timecourse of the pixel that achieves the peak value in the full STRF. The complete linear filter was collapsed into a single frame by finding the projection of each pixel onto this principal timecourse. This procedure is equivalent to finding the best (least squares) spatio-temporally separable approximation to the STRF: $K(x,t) = k(x)w(t)$, where $k(x)$ and $w(t)$ are the spatial and temporal components of the approximation. From the single frame $k(x)$, we extracted the center and surround regions. To find the center, we began with the peak pixel and then recursively expanded the region in a contiguous patch to include any pixels whose timecourses had at least a 50% correlation with already included pixels. We ended the recursive process after the first pass in which no pixels were added to the center. At this point, all pixels not included in the center were considered part of the surround for the purpose of assessing the surround strength.

Taking the center defined in this way as a mask for the full STRFs, we summed all pixel values within the center at each time point to generate a temporal profile for the central receptive field. To obtain temporal kernels with greater precision than the 30 Hz time scale of our STRFs, we used cubic spline interpolation with knots spaced every 33 ms. From the interpolated timecourses, we measured the time to peak under each stimulus for the center. In addition, the peak value of this temporal profile was taken to be the center weight of the receptive field. Similar computations yielded the surround time to peak and surround weight. We then quantified the relative surround strength as the ratio of surround weight to center weight.

In addition, we measured the gain g of each neuron under each stimulus condition. To test for contrast gain control, we defined "effective contrast," σ, as the standard deviation of the linear filter output. To avoid ambiguity between linear filter amplitude and gain, we normalized each STRF to have unit Euclidean norm before computing the gain and the effective contrast.

We used the analysis method described here because it gave the most robustly unbiased results in our simulations (see below), but we also wanted to verify that our results did not change dramatically with slightly different methods (see details in Text S2 and Table S1). Briefly, we made a series of modifications to our receptive field extraction method and repeated the analyses described in the main text for each modification.

Model validation of receptive field analysis

To validate our STRF analysis methods, we generated synthetic data using a linear-nonlinear (LN) model. We then applied STRF extraction and analysis methods identical to those applied to real data to check that the known LN parameters were extracted in an unbiased fashion. The linear filter was chosen to be spatio-temporally separable, with the temporal component taken from measured ganglion cell responses. The spatial filter was modeled as a difference-of-Gaussians, where the size and strength of the surround Gaussian relative to the center Gaussian were chosen to mimic receptive fields of real neurons. In each simulation, parameters for 100 model neurons were chosen independently. The results are summarized in Table 2.

In our first simulation, the surround radius (relative to center radius) was chosen from a Gaussian distribution with mean 2 and standard deviation 0.3, the relative surround strength from a Gaussian distribution with mean 1 and standard deviation 0.1, and the offset coordinates from Gaussian distributions with mean 0 and standard deviation 2 ("Standard model" in Table 2). For each model neuron, the same filter was applied to the spatio-temporal exponentially correlated and uncorrelated stimuli in order to simulate cases without adaptation. Across the population, our model neurons showed only a slight bias in center latency between the two stimuli (Fig. S1A). While this effect reaches significance (for $\alpha = .05$), the effect size is orders of magnitude smaller than that seen in the data and thus could not explain our experimental results. We also observed a tendency toward a slightly stronger relative surround strength under white noise than under correlated noise (Fig. S1B). Note that this is opposite the effect observed in our experimental results (Fig. 6D–F). Thus, if anything our results may be stronger than reported in the main text.

To further validate our analysis we ran simulations with an even wider range of model parameters. We first constructed model

Table 2. Model validation of receptive field analysis.

	AI (relative surround strength)				AI (center time to peak)			
	mean	**std**	***p***	**skew**	**mean**	**std**	***p***	**skew**
Standard model	−0.02	0.06	.004	−1.10	− 0.0003	0.002	.04	− 4.70
Small surround weight	−0.06	0.13	< .0001	−0.12	− 0.0003	0.002	.04	− 4.69
Large surround weight	−0.05	0.26	.08	−2.44	− 0.0014	0.004	.0003	− 2.66
Small surround radius	− 0.01	0.08	.19	−4.56	−0.0009	0.003	.0007	−2.49
Large surround radius	−0.02	0.05	<.0001	−0.24	−0.0005	0.002	.01	−3.71

Adaptation index (AI) in surround strength and center latency for different non-adapting control models. Columns labeled "mean," "std," and "skew" show the mean, standard deviation, and skewness of the adaptation indices for the indicated analysis; columns labeled "*p*" show the $p-$values from $t-$tests of each distribution against the null hypothesis of zero mean. **Standard model**: Surround radii (relative to center radii) had mean 2 and standard deviation 0.3; surround weights (relative to center weights) had mean 1 and standard deviation 0.1. **Small surround weight**: Surround weights had mean 0.5; all other parameters were the same as in the standard model. **Large surround weight**: Surround weights had mean 2. **Small surround radius**: Surround radii had mean 1. **Large surround radius**: Surround radii had mean 3.

doi:10.1371/journal.pcbi.1003344.t002

neurons with surround radii measured from Gaussian distributions with means of 1 ("Small surround radius" in Table 2) or 3 ("Large surround radius"), both with standard deviation 0.3, and all other parameters the same as in our original simulation. In separate simulations, we kept the mean surround radius at 2 but chose the relative surround strength from a Gaussian distribution with mean 0.5 ("Small surround weight") or 2 ("Large surround weight"), both with standard deviation 0.1. As with our original set of parameters, the models recovered from STRF analysis had at most slight biases toward weaker surrounds and shorter center times to peak under correlated noise (see Table 2).

Supporting Information

Figure S1 Receptive field results are validated with model neurons and are robust to analysis method changes. (**A**) Center time to peak for a population of non-adapting model neurons, plotted as in Figure 6A. Receptive fields were constructed as a difference of Gaussians. Surround radii (relative to center radii) had a mean of 2 and a standard deviation of 0.3. Surround weights (relative to center weights) had a mean of 1 and a standard deviation of 0.1. The model neurons do not show a large difference between stimuli in center time to peak. (**B**) Model neurons described in (A) show a slight bias toward smaller recovered relative surround strength under correlated noise compared to white noise. (**C**) Center time to peak is longer for white noise than for correlated noise when the surround only includes non-center pixels whose time courses are positively correlated with the time course of the peak negative pixel. (**D**) Relative surround strengths is marginally higher for correlated noise than for white noise under the same analysis as in (C). (**E**) Center time to peak is longer for white noise than for correlated noise when the center is not required to form a single contiguous component. The figure omits a few outliers that lie outside the range of the horizontal axis and have longer time courses under correlated noise. (**F**) Relative surround strengths is marginally higher for correlated noise than for white noise under the same analysis as in (E).
(TIF)

Figure S2 Correlation measurement controls. (**A**) As a control on the quality of the non-adapting LN model, we examined the difference between its predicted pairwise correlations under the white noise stimulus (the stimulus to which the model was fit) and the observed correlations. Since the model is a single-neuron model that does not attempt to capture pairwise correlations, we do not expect it to reproduce these correlations perfectly. Nevertheless, many cell pairs are well-predicted, indicating that their correlation is largely due to receptive field overlap. There is, however, a slight tendency for the model to underestimate correlations: this is likely due to its neglect of noise correlations. We note that such a bias will not effect the model's predicted *excess correlation*, unless it is very different in the two stimulus conditions. But such an effect would represent a form of non-trivial processing in its own right. (**B**) Our correlation measurements were based on binned spike trains. We measured excess correlation, in the spatiotemporal exponential dataset, for a variety of bin sizes. Its value is largely independent of bin size. Error bars represent 95% bootstrap confidence intervals. All correlations reported in the main text were estimated from spike trains binned at 33 ms.
(TIF)

Figure S3 Firing rates are within a physiologically normal range. (**A**) Firing rates are similar across stimuli. Each cell's average white noise firing rate was subtracted to aid in comparison across experiments. Red line indicates the median for each stimulus. Blue boxes show the middle 50% of values; black boxes, the middle 90%. Gray crosses represent the outlying 10% of cells. Numbers below each box show the number of cells included. (**B**) Comparison of the excess correlation for each correlated stimulus to the average response (relative to the white noise response) shows that excess correlation is not simply a reflection of firing rate.
(TIF)

Figure S4 Dependence of excess correlation on response polarity and type of scene. (**A, B**) Excess correlation values for natural movie (A) and spatio-temporal exponential (B) stimuli separated by polarity of the recorded pair. Correlated stimuli evoked a greater increase in output correlations for ON-OFF pairs than for ON-ON and OFF-OFF pairs (white bars). This increase was consistent with correlation changes predicted by a non-adapting model for opposite-polarity pairs responding to

natural scenes (A, rightmost bar) but was smaller than the expected correlation change in all other cases (gray bars). For the spatio-temporal exponential stimulus, where we could compute white noise and correlated noise receptive fields, the observed changes filter and gain were able to reproduce the smaller change in output correlations relative to input correlations (B, dark gray bars). (**C**) Excess correlation computed by comparing responses to different natural movies directly. Our natural stimulus ensemble was composed of movies taken in three outdoor settings. During an experiment, movies from these scenes were interspersed. Here we see that output correlation are similar for all three; thus, in the main text we group all natural movie data together.
(TIF)

Figure S5 Dependence of correlation on linear filter overlap for opposite-polarity cell pairs. Correlation coefficient between opposite-polarity cell pairs recorded under a white noise stimulus. The dot product of the two cells' spatiotemporal linear filters is used as a measure of the overlap between areas with the same polarity; a pair of cells whose receptive fields are identical in shape but have opposite polarity will have a strong negative filter dot product. The fact that positive correlation between opposite-polarity pairs' spike trains is strongly associated with positive filter dot products may be due to one cell's receptive field center lying in the other's surround or to differences in the temporal filtering of the two cells such that the RF centers do not overlap temporally.
(TIF)

Table S1 Variants on receptive field analysis method. Adaptation index (AI) in surround strength and center latency measured by variations in our analysis method. Columns labeled "mean," "std," and "skew" show the mean, standard deviation, and skewness of the adaptation indices for the indicated analysis; columns labeled "p" show the $p-$values from $t-$tests of each distribution against the null hypothesis of zero mean. **Standard analysis**: Receptive fields were analyzed as presented in the main text. **Disconnected center**: The center was not required to form a contiguous region. **Surround threshold**: A threshold criterion was used to find the surround so that only pixels positively correlated with the peak surround pixel were included. **No mean subtraction**: The mean of the frames after each spike was not subtracted from the STRFs. **First principal component**: The full STRF was collapsed onto a single frame by projecting onto the first principal timecourse rather than the principal timecourse most similar to the peak pixel. **Masks from WN**: Center and surround regions measured from white noise were used to analyze STRFs from both stimuli. **Masks from CN**: Center and surround regions measured from spatio-temporal exponentially correlated noise were used to analyze STRFs from both stimuli.
(PDF)

Text S1 Dependence of output correlation on gain and firing rate. In the main text we found that gain adaptation contributes to decorrelation in a population of LN neurons. In Text S1, we derive a mathematical relationship between the gain of model LN neurons and the pairwise correlations in the model population.
(PDF)

Text S2 Tests of robustness. To verify that our receptive field results were not an artifact of our specific analysis method, we repeated the analysis with a number of modifications. In Text S2, we report the results of these analyses.
(PDF)

Acknowledgments

We are grateful to Olivier Marre, Gautam Awatramani, David Schwab, Tatyana Sharpee, Michael Berry and E.J. Chichilnisky for useful comments at various stages of this work.

Author Contributions

Conceived and designed the experiments: KDS JSP GT PCN VB. Performed the experiments: KDS JSP. Analyzed the data: KDS JSP GT JH. Wrote the paper: KDS JSP VB. Produced the natural movie stimuli: HKY SEP.

References

1. Barlow HB (1961) Possible principles underlying the transformations of sensory messages. In: Rosenblith W, editor, Sensory Communication, Cambridge, MA: MIT Press, chapter 13. pp. 217–234.
2. Srinivasan MV, Laughlin SB, Dubs A (1982) Predictive coding: a fresh view of inhibition in the retina. Proc R Soc Lond B Biol Sci 216: 427–59.
3. Atick JJ (1992) Could information theory provide an ecological theory of sensory processing? Network 3: 213–251.
4. Borghuis BG, Ratliff CP, Smith RG, Sterling P, Balasubramanian V (2008) Design of a neuronal array. J Neurosci 28: 3178–89.
5. Tkačik G, Prentice JS, Balasubramanian V, Schneidman E (2010) Optimal population coding by noisy spiking neurons. Proc Natl Acad Sci U S A 107: 14419–14424.
6. Estebanez L, El Boustani S, Destexhe A, Shulz DE (2012) Correlated input reveals coexisting coding schemes in a sensory cortex. Nat Neurosci 15: 1691–1699.
7. Barlow H (2001) Redundancy reduction revisited. Network 12: 241–253.
8. Fiser J, Chiu C, Weliky M (2004) Small modulation of ongoing cortical dynamics by sensory input during natural vision. Nature 431: 573–578.
9. Ganmor E, Segev R, Schneidman E (2011) The architecture of functional interaction networks in the retina. J Neurosci 31: 3044–3054.
10. Puchalla JL, Schneidman E, Harris RA, Berry II MJ (2005) Redundancy in the population code of the retina. Neuron 46: 493–504.
11. Schneidman E, Bialek W, Berry II MJ (2003) Synergy, redundancy, and independence in population codes. J Neurosci 23: 11539–11553.
12. Trong PK, Rieke F (2008) Origin of correlated activity between parasol retinal ganglion cells. Nat Neurosci 11: 1343–1351.
13. Greschner M, Shlens J, Bakolitsa C, Field GD, Gauthier JL, et al. (2011) Correlated firing among major ganglion cell types in primate retina. J Physiol 589: 75–86.
14. Pitkow X, Meister M (2012) Decorrelation and efficient coding by retinal ganglion cells. Nat Neurosci 15: 628–635.
15. Rodieck RW (1965) Quantitative analysis of cat retinal ganglion cell response to visual stimuli. Vision Res 5: 583–601.
16. Lesica NA, Jin J, Weng C, Yeh CI, Butts DA, et al. (2007) Adaptation to stimulus contrast and correlations during natural visual stimulation. Neuron 55: 479–91.
17. Sharpee TO, Sugihara H, Kurgansky AV, Rebrik SP, Stryker MP, et al. (2006) Adaptive filtering enhances information transmission in visual cortex. Nature 439: 936–942.
18. Simmons K, Prentice J, Tkacik G, Homann J, Yee H, et al. (2013) Data from: Transformation of stimulus correlations by the retina. doi:10.5061/dryad.246qg.
19. Chichilnisky EJ (2001) A simple white noise analysis of neuronal light responses. Network 12: 199–213.
20. Theunissen FE, David SV, Singh NC, Hsu A, Vinje WE, et al. (2001) Estimating spatio-temporal receptive fields of auditory and visual neurons from their responses to natural stimuli. Network: Comput Neural Syst 12: 289–316.
21. Smirnakis SM, Berry II MJ, Warland DK, Bialek W, Meister M (1997) Adaptation of retinal processing to image contrast and spatial scale. Nature 386: 69–73.
22. Baccus SA, Meister M (2002) Fast and slow contrast adaptation in retinal circuitry. Neuron 36: 909–919.
23. David SV, Vinje WE, Gallant JL (2004) Natural stimulus statistics alter the receptive field structure of V1 neurons. J Neurosci 24: 6991–7006.
24. Touryan J, Felsen G, Dan Y (2005) Spatial structure of complex cell receptive fields measured with natural images. Neuron 45: 781–791.
25. Pillow JW, Shlens J, Paninski L, Sher A, Litke AM, et al. (2008) Spatio-temporal correlations and visual signalling in a complete neuronal population. Nature 454: 995–9.
26. Granot-Atedgi E, Tkačik G, Segev R, Schneidman E (2013) Stimulus-dependent maximum entropy models of neural population codes. PLoS Comput Biol 9: e1002922.

27. Prentice JS, Homann J, Simmons KD, Tkačik G, Balasubramanian V, et al. (2011) Fast, scalable, bayesian spike identification for multi-electrode arrays. PLoS One 6: e19884.
28. Ratliff CP, Borghuis BG, Kao YH, Sterling P, Balasubramanian V (2010) Retina is structured to process an excess of darkness in natural scenes. Proc Natl Acad Sci U S A 107: 17368–17373.
29. Sharpee T, Rust NC, Bialek W (2004) Analyzing neural responses to natural signals: maximally informative dimensions. Neural Comput 16: 223–250.