Published paper

Song, Z and Juusola, M (2014) Refractory sampling links efficiency and costs of sensory encoding to stimulus statistics. The Journal of Neuroscience, 34 (21). 7216 - 7237. Doi: 10.1523/JNEUROSCI.4463-13.2014

Systems/Circuits

Refractory Sampling Links Efficiency and Costs of Sensory Encoding to Stimulus Statistics

Zhuoyi Song[2,3] **and Mikko Juusola**[1,2]

[1]National Key Laboratory of Cognitive Neuroscience and Learning, Beijing, Beijing Normal University, Beijing 100875, China, [2]Department of Biomedical Science, University of Sheffield, Sheffield S10 T2N, United Kingdom, and [3]Vision Research Laboratory, Cognitive, Perceptual and Brain Sciences University College London, London, WC1E 6BT, United Kingdom

Sensory neurons integrate information about the world, adapting their sampling to its changes. However, little is understood mechanistically how this primary encoding process, which ultimately limits perception, depends upon stimulus statistics. Here, we analyze this open question systematically by using intracellular recordings from fly (*Drosophila melanogaster* and *Coenosia attenuata*) photoreceptors and corresponding stochastic simulations from biophysically realistic photoreceptor models. Recordings show that photoreceptors can sample more information from naturalistic light intensity time series (NS) than from Gaussian white-noise (GWN), shuffled-NS or Gaussian-1/f stimuli; integrating larger responses with higher signal-to-noise ratio and encoding efficiency to large bursty contrast changes. Simulations reveal how a photoreceptor's information capture depends critically upon the stochastic refractoriness of its 30,000 sampling units (microvilli). In daylight, refractoriness sacrifices sensitivity to enhance intensity changes in neural image representations, with more and faster microvilli improving encoding. But for GWN and other stimuli, which lack longer dark contrasts of real-world intensity changes that reduce microvilli refractoriness, these performance gains are submaximal and energetically costly. These results provide mechanistic reasons why information sampling is more efficient for natural/naturalistic stimulation and novel insight into the operation, design, and evolution of signaling and code in sensory neurons.

Key words: *Drosophila*; information theory; phototransduction; sampling; stochasticity; vision

Introduction

Information about real-world similarities and differences is critical for successful behaviors (Barlow, 1961). Sampled and integrated by sensory neurons into graded macroscopic responses to drive synaptic transmission or action potential generation, this information limits an animal's perception and actions in the world. Physiological findings along sensory pathways suggest that stimulation, which mimics the structure of natural signals, may generate information-richer neural responses than Gaussian white noise (GWN) stimulation (Rieke et al., 1995; van Hateren, 1997; Lewen et al., 2001; Juusola and de Polavieja, 2003), which

lacks real-world correlations but maximizes information within its bandwidth and variance (Shannon, 1948). Thus, GWN stimulation may drive sensory systems submaximally or inefficiently, which calls into question its usefulness to study neural performance and energy consumption. Nonetheless, why, how, and where the statistical structure of stimulation changes the efficiency and costs of sensory-neural signals have not been investigated systematically.

Many photo-, olfactory-, and mechano-receptors adapt continuously, generating responses with complex amplitude and phase correlations to stimulus changes in natural environments (Rieke et al., 1995; van Hateren, 1997; Juusola and de Polavieja, 2003; Smear et al., 2011). These sensory neurons have transduction reactions compartmentalized in membrane elaborations that work as sampling units, such as cilia or microvilli (Fain et al., 2010), and may use stochastic adaptive sampling to encode discrete information (Song et al., 2012). Experiments and theory suggest that stochastic adaptive sampling happens in fly photoreceptors' light sensors, the rhabdomeres (Fig. 1A), wherein thousands of microvilli transduce single photon energies to elementary responses, quantum bumps. The size and latency of these samples vary stochastically and adapt to light changes, summing up the macroscopic response (Henderson et al., 2000). But after each bump, the light-activated microvilli are rendered briefly (50–300 ms) refractory (Song et al., 2012). Therefore, with brightening intensity, their sample rate begins to saturate because fewer are available to generate the next bumps (Howard et al.,
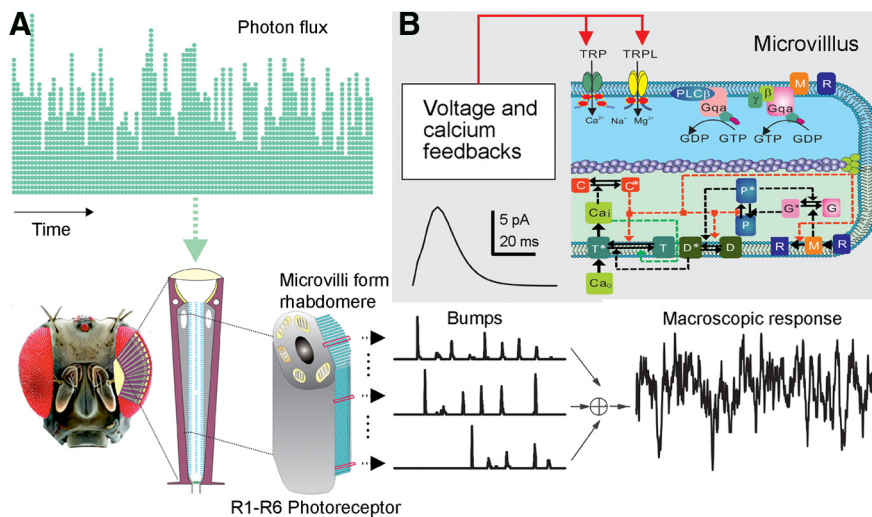
Figure 1. Stochastic adaptive sampling of light information by *Drosophila* photoreceptors. *A*, Each photoreceptor samples photon influx by ~30,000 microvilli, which together form the photo-sensitive light-guide, the rhabdomere. Single-photon-responses (bumps) from individual microvilli integrate a macroscopic response. *B*, Top, Each microvillus contains full phototransduction reactions, generating one bump (sample) to an absorbed photon at a time; voltage and Ca^{2+}-dependent feedbacks regulate sample size and speed. Bottom, Stochastic processes simulate bump generation. Molecular participants in microvillar phototransduction reactions. M*, Metarhodopsin; C*, Ca^{2+}-dependent negative feedback to multiple targets; D*, DAG; P*, G protein-PLC complex. Red and green dotted arrows indicate negative and positive feedbacks, respectively, as used in the stochastic model (Song et al., 2012). DAG (D*), Yet unresolved gating mechanism that includes production of DAG, InsP3, proton, and physical microvillar contraction (Hardie and Franze, 2012).

1987; Song et al., 2012). Conversely, the bump waveform (sample size) in each microvillus is set by its phototransduction cascade gain, as regulated by Ca^{2+}- and voltage-dependent memory (feedbacks) of the past bumps (Fig. 1B). These adaptations, which facilitate robust encoding of behaviorally relevant information in flies' habitats (Song et al., 2012), likely evolved through natural selection for sustainable visual lifestyles and energy costs (Niven et al., 2007; Gonzalez-Bellido et al., 2011).

Although information transmission capacity is clearly important in determining neural function, we lack both quantitative and mechanistic understanding on the relationship between stimulus statistics and sensory-neural signals. In this study, we systematically investigate how information transmission in photoreceptor recordings and in biophysically realistic photoreceptor models changes with stimulus bandwidth and intensity distributions. We show that stochastic sampling of light information by finite refractory microvilli populations determines why and how fly photoreceptors encode different stimulus statistics differently, with different efficiencies and costs. Specifically, our results suggest that longer dark contrasts, which characterize naturalistic stimuli, help to recover more refractory microvilli than equally bright stimuli without these features, improving neural information capture while lowering its metabolic costs.

Materials and Methods

Flies. Two- to 10-d-old wild-type red-eyed (Canton-S) and white-eyed (*w* Oregon R) female fruit flies (*Drosophila melanogaster*), and adult female killer flies (*Coenosia attanuata*) were used in the experiments. *Drosophila* were raised at 18°C in a 12 h/12 h dark/light cycle and fed on standard medium in our laboratory culture. *Coenosia* were captured from greenhouses in Almeria (Spain) and used within 3 d. During their captivity, *Coenosia* were kept at room temperature (~23°C) and fed with *Drosophila*.

Electrophysiology. Flies were immobilized inside a fly holder (see Fig. 4A) with their heads fixed with beeswax, as described previously (Juusola and Hardie, 2001a). For the recording microelectrode, a small hole, the

size of few ommatidia, was cut in the dorsal cornea and sealed with Vaseline to prevent the eye from drying. Intracellular voltage responses in *Drosophila* and *Coenosia* R1-R6 photoreceptors were recorded to various light stimuli (see below) using conventional sharp microelectrodes.

Sharp quartz or borosilicate microelectrodes (120–220 MΩ) were fabricated with a Sutter Instruments P2000 puller (Sutter Instruments) and filled with 3 M KCl solution. A blunt reference electrode, filled with fly ringer, was inserted into the fly head capsule close to the ocelli (see Fig. 4A). The head temperature of the flies was kept at 25 ± 1°C or 20 ± 1°C by a feedback-controlled Peltier device. The recordings were performed after 2–5 min of dark adaptation, using the discontinuous (switched) clamp method with a switching frequency of up to 40 kHz. The capacitance of the electrodes was compensated by monitoring the head-stage output voltage. To minimize effects of damage and external noise, such as instrumental noise or extrinsic neural/muscle activity, on the analysis, only stable recordings of low-noise and high sensitivity were chosen for this study. Such photoreceptors typically had resting potentials <−60 mV in darkness and >45 mV responses to saturating test light pulses (compare Fig. 4A).

Light stimulation. In most experiments, a high power "white"-light emitting diode (Seoul Z-Power LED P4 star, white, 100 Lumens) was used to stimulate photoreceptors. "White" LED was chosen because its red component wavelengths reduced prolonged depolarizing afterpotential effects; these are induced when stimulation lacks long wavelengths to convert green-activated visual pigments (metarhodopsin) back to their resting state (rhodopsin) (Minke, 2012). The LED was connected to a randomized quartz fiber optic bundle (spectral transmission range: 180–1200 nm), fitted with a lens and a pinhole (~1° as seen by the flies) (Gonzalez-Bellido et al., 2011), and attached onto a Cardan arm system, providing accurate positioning of the stimuli. The light source was driven by an OptoLED (Cairn Research), which utilizes a feedback circuitry with a light sensor to regulate the light output of the LED. Light pulses and stimuli with special statistics (see below) were played to a photoreceptor at the center of its receptive field (see Fig. 4A). Both the stimuli and the resulting voltage responses were filtered at 500 Hz (KEMO VBF/23 low pass elliptic filter) and sampled together at 1–10 kHz using a 12-bit A/D converter (National Instruments), controlled by a custom-written software system, Biosyst (Juusola and Hardie, 2001a; Juusola and de Polavieja, 2003) in MATLAB (MathWorks) environment. The stimulation regimen comprised the following light intensity time series (each seen by a fly as a flickering light point with specific brightness statistics):

GWN stimuli. To test whether and how the frequency bandwidth of light changes affects encoding, we used 2-s-long GWN stimuli, which had "flat" power spectrum up to 20, 50, 100, 200, or 500 Hz (Fig. 2A,B) low-pass filtered by MATLAB's filter toolbox. A bright (daylight) level (~10^6 photons/s) was switched on for 7–20 s, followed by repeated presentations (20–150 times) of a GWN stimulus superimposed upon it (see Fig. 4A). Because each of these stimuli had Gaussian amplitude distribution (Fig. 2A) between the contrast minimum of −1 (darkness) and the maximum of 1 (double the mean), their average contrast was effectively the same: $c = \Delta I/I$ = SD/mean ~ 0.32, where ΔI is the change and I mean intensity over time.

Naturalistic stimulation (NS). To test whether and how encoding of naturalistic light sequences differ from that of band-limited GWN, we selected 1-s-long pattern (NS; 10,000 points) from van Hateren natural stimulus collection (van Hateren, 1997) on the basis of its rich variability and dynamic range (Fig. 2E, blue; see Results). Its power spectrum be-
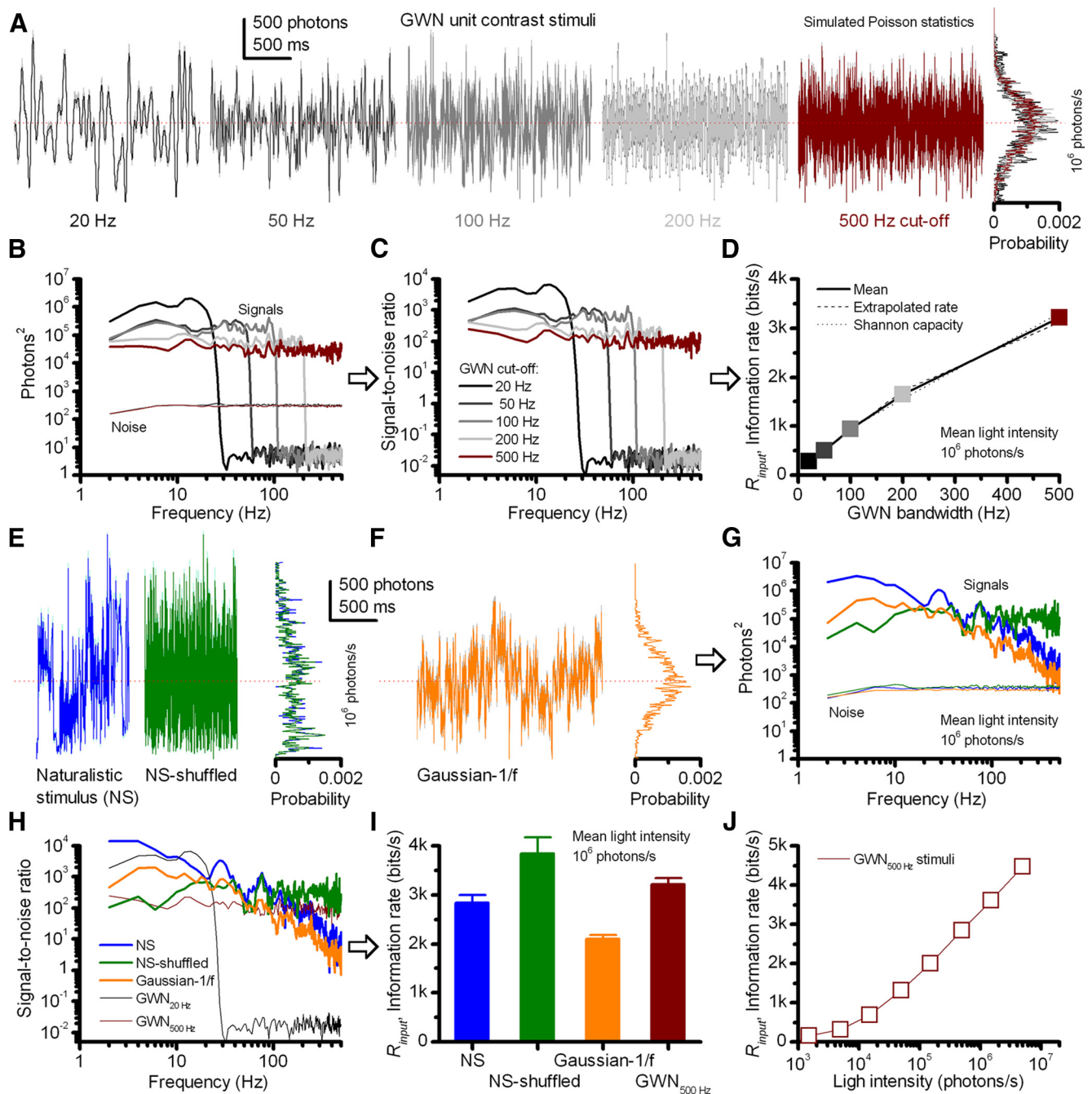
Figure 2. Light stimuli and their information rates, R_{input}, are bound by Poisson statistics. **A**, Mean and 10 simulated traces (light gray) of 20, 50, 100, 200, and 500 Hz GWN stimuli, highlighting inherent variations in light input; attributable to shot noise in photon emission from the source. Each stimulus has the same mean light output of 10^6 photons/s and contrast of 0.32. **B**, The mean bandwidth and noise power spectra (variance) of the stimuli. **C**, Their corresponding SNRs. **D**, Information transfer rate of Poisson GWN stimuli increases with bandwidth; shown for mean intensity of 10^6 photons/s. **E**, NS (blue) and its shuffled intensity values (green) have the same contrast (0.58) and probability density distribution (note variation in individual distributions). **F**, Gaussian stimulus with a $1/f$ power spectrum (contrast of 0.40; orange). **E, F**, Means and 10 simulated traces. These specific "broadband stimuli" were adjusted to have the same mean light intensity of 10^6 photons/s as GWN stimuli (**A**). **G**, Signal and noise power spectra of these stimuli. **H**, NS (blue) and Gaussian 1/f stimulus (orange) have broadly comparable SNRs, with their low stimulus frequencies having the highest values (or information densities). Their SNR maxima are similar to that of 20 Hz GWN (dotted). Shuffling (randomizing) intensity values of NS reduces its SNR to a lower mean (green), being still higher than that of 500 Hz GWN (wine). **I**, All the "broadband stimuli" had intrinsically high rate of information transfer ($\sim \geq 2000$ bits/s); as has 500 Hz GWN. **J**, Information transfer rate of light stimulation increases with its intensity (photon rate); shown for 500 Hz GWN.

haved approximately as $1/f$ (f: temporal frequency). NS was played back at 10 kHz and repeated 20 – 450 times. The stimulus intensity was adjusted to have the same mean as the GWN stimuli above ($\sim 10^6$ photons/ s). But because it contained some higher intensity values and prolonged temporal correlations, it had a higher mean contrast (~ 0.58) than the GWN stimuli (~ 0.32). In this study, we call such stimulation naturalistic (instead of natural) because flies in the wild rarely experience several

repeated presentations of the same stimulus. Furthermore, the stimulation lacked spatial and chromatic correlations of the natural environment, which likely adapt lateral information flow within the retina/ lamina network differently (Wardill et al., 2012).

Shuffled naturalistic stimulation (shuffled-NS). To test whether and how adaptation to temporal correlations in naturalistic light intensity time series advance encoding, we further performed control experiments

using a shuffled-NS sequence (above). The intensity values in its time bins were rearranged in a random sequence order (Fig. 2E, green). Although the shuffling "whitened" the stimulus sequence, minimizing time-dependent intensity correlations and maximizing information content (Fig. 2I), this did not alter its intensity distribution or contrast (0.58).

Gaussian noise 1/f-stimulation (Gaussian-1/f). To test to what degree encoding depends upon the global intensity distribution, we used a Gaussian noise light stimulus, in which power spectrum approximated that of NS (Fig. 2E, orange). In Fourier domain, we conducted uniform random phase-shifts to the NS 1/f-power spectrum, ensuring that it stays the same. The Gaussian-1/f stimulus was then generated by the inverse Fourier transformation of the random phase-shifted NS (Theiler et al., 1992).

GWN stimuli at different light intensity levels. To test how adaptation to different mean luminance affects encoding (see Fig. 12), we used a combined light source of four green and three red LEDs (Marl Optosource) (to reduce depolarizing afterpotential effects, see above), driven by a custom-built LED driver. The light intensity range of 5 log units was calibrated by counting the number of single photon responses, bumps (Lillywhite and Laughlin, 1979), during prolonged dim illumination (Juusola and Hardie, 2001a). The light output was attenuated by neutral density filters (Kodak Wratten) to provide seven light levels in half log unit steps. The lowest was estimated to be ~300 effective photons/s and the highest ~3×10^6 photons/s.

Stimulus calibration in simulations. In Figures 6, 7, 8, 9, and 10, we simulated responses to GWN, NS, shuffled-NS, and Gaussian-1/f, stimuli, each having the mean intensity of 10^5 photon/s, whereupon the photon input to each microvillus was generated by the random-photon-absorption model (see below). This light intensity level, which is ~10 times dimmer than the daylight level used in the *in vivo* recordings (see Fig. 4 at 25°C; see Fig. 13 at 20°C), was chosen to bring closer the effective bump production rates of the simulations to those of the recordings. Markedly, the photoreceptor models (see below) lacked intracellular pupil, which we estimate could reduce photon influx to microvilli by ~10-fold (see Fig. 12). *Drosophila* has less intense screening pigmentation than *Calliphora*, in which intracellular pupil may reduce photon flux by 100-fold at very bright daylight intensities (Vogt et al., 1982; Howard et al., 1987; Roebroek and Stavenga, 1990; Stavenga, 2004), or *Coenosia* (Gonzalez-Bellido et al., 2011). However, in parallel with screening pigments, stochastic microvilli refractoriness reduces quantum efficiency, settling the information transfer of bright NS (~$>10^5$ photon/s) to similar rates, even without the intracellular pupil (Song et al., 2012). For example, transmission of NS in stochastic simulations in Figure 8D (10^5 photon/s; ~400 bits/s) is very close to that of 10^6 photon/s: ~420 bits/s for the same waveform (Song et al., 2012). Nonetheless, without the intracellular pupil, 30,000 microvilli begin to saturate to bright GWN stimuli, eventually reducing the signal-to-noise ratio (SNR) and information transfer of *Drosophila* photoreceptor output (see Fig. 12E).

Statistical properties of light stimuli. Photons are emitted by the light source, such as the LEDs above, at random, exhibiting detectable statistical fluctuations (shot noise) that can be modeled by Poisson statistics. Therefore, as each light stimulus trace differs from any other, with their mean equaling their variance (Fig. 2), we could estimate through simulations their average signals and noise (Fig. 2B, G), SNRs (Fig. 2C, H), and entropy and information rates (Fig. 2D, I, J; Table 1) by using Equations 2–5 as explained below.

Data analysis. All data analyses were performed with MATLAB (MathWorks).

Voltage responses of photoreceptors were recorded continuously to repeated stimulation. In many cases, because of their strong short-term adaptive trends, the first 5–20 traces were rejected from the analysis. Analytical and information theoretical methods for quantifying voltage responses of approximately steady-state-adapted fly photoreceptors to GWN and NS stimuli have been described in detail previously (Juusola et al., 1994; Juusola and Hardie, 2001a; Juusola and de Polavieja, 2003; Faivre and Juusola, 2008; Gonzalez-Bellido et al., 2011; Song et al., 2012). Below is a brief summary of the key approaches used in this study.

The linear frequency response, or transfer function $T(f)$, between the average response, or *signal* $s(t)$, and the GWN contrast stimuli, $c(t)$, was

Table 1. Entropy and information transfer rates of light stimuli used in this study, estimated from stochastic simulation^a

Stimuli	Cut-off frequency	Mean contrast (SD)	Entropy rate, R_s, mean ± SD (bits/s)	Information rate, R_{input}, mean ± SD (bits/s)	Information rate, C_{input}, mean ± SD (bits/s)
		Statistical properties of light stimuli (Poisson shot-noise) at mean intensity of 10^6 photon/s			
GWN	20 Hz	0.327	399 ± 9	273 ± 9	282
GWN	50 Hz	0.327	708 ± 20	489 ± 20	524
GWN	100 Hz	0.327	1346 ± 225	927 ± 225	953
GWN	200 Hz	0.327	2607 ± 172	1731 ± 172	1560
GWN	500 Hz	0.327	4045 ± 355	3114 ± 355	3302
NS	1/f	0.584	3314 ± 130	2744 ± 130	2949
NS-shuffled	"white"	0.584	4299 ± 99	3604 ± 99	4069
Gaussian-1/f	1/f	0.403	2948 ± 155	2049 ± 155	2157

^{a}Apart from 500 Hz GWN; NS had higher information transfer rate than the GWN stimuli. For the simulated stimuli, the table gives the mean ± SD (error) of the extrapolated rate (Eq. 5) and Shannon capacity (Eq. 2).

calculated using their 500-points-long spectral estimates, $S(f)$ and $C(f)$, respectively, as follows:

$$T(f) = \frac{\langle S(f) \cdot C^\star(f) \rangle}{\langle C(f) \cdot C^\star(f) \rangle} \quad (1)$$

Where $<>$ indicate the average over the different stretches, and $\star$ the complex conjugate; the spectral estimates were calculated using MATLAB's fast Fourier transform algorithm.

The gain part of $T(f)$ (e.g., see Fig. 12B) was used for estimating the 3 dB cutoff frequencies of the obtained voltage responses to different bandwidth GWN. Its phase part (see Fig. 12C) indicates lag between the input and output frequencies.

Signal-to-noise ratio (SNR). In each recording (e.g., see Fig. 4C), simulation (e.g., see Fig. 6C), or Poisson light stimulus series (Fig. 2; Table 1), the mean was the *signal*, whereas the *noise* was the difference between individual traces and the *signal*. Hence for an experiment using n trials (with $n = 20–450$), there is one *signal* trace and n *noise* traces. For the analysis, the *signal* and *noise* traces were divided into 50% overlapping stretches and windowed with a Blackman–Harris 4-term window (Harris, 1978), each giving three 500-points-long samples. As typically 20–50 consecutive traces were used (e.g., the most stable continuous segment in the recorded voltage responses), we obtained 60–150 spectral samples for the *noise* and 3 spectral samples for the *signal*. These were averaged, respectively, to improve the estimates.

$SNR(f)$ of the recording, simulation, or Poisson light stimulus series was calculated from their *signal* and *noise* power spectra, $<|S(f)|^2>$ and $<|N(f)|^2>$ (see Fig. 4D or Fig. 6D), respectively, as their ratio (see Fig. 4E or Fig. 6E), where $||$ denotes the norm and $<>$ the average over the different stretches (Juusola et al., 1994; Juusola and Hardie, 2001a). To eliminate data size bias in individual recording series from the same cell, which could contain responses to all GWN and NS (compare Fig. 4) or NS, NS-shuffled, and Gaussian-1/f stimuli (compare Fig. 5), the same number of traces (typically 20–50) was used for calculating its $SNR(f)$ estimates.

Information transfer rate estimation. To estimate information transfer rate of each recording (e.g., see Fig. 4C), simulation (e.g., see Fig. 6C), or Poisson light stimulus (Fig. 2) series, we used both (1) the classic Shannon formula (Shannon, 1948) and (2) the triple extrapolation method (Juusola and de Polavieja, 2003), which has been shown to obtain robust estimates from continuous responses (Juusola and de Polavieja, 2003; de Polavieja et al., 2005). Both of these methods require ergodic output; thus, we analyzed steady-state-adapted recordings and simulations, in which each response (or stimulus trace) is expected to be equally representative of the underlying encoding (or statistical) process. Yet, both methods have their known limitations, which can affect the accuracy of their estimates. Comparing the estimates side by side ensured that the information transfer rates were calculated consistently and accurately in this study.

Shannon formula. From $SNR(f)$, the information transfer rate is estimated as follows:

$$C = \int_{min}^{max} \log_2(SNR(f) + 1)df \qquad (2)$$

For both *Drosophila* and *Coenosia* data, we used minimum = 2 Hz and maximum = 500 Hz (resulting from 1 kHz sampling rate and 500 points window size).

The voltage responses of fly photoreceptors to NS are nonlinear and often non-Gaussian (van Hateren, 1997; Juusola and de Polavieja, 2003). Here, the Shannon formula, which assumes that the stimulus is Gaussian and that the signal and the noise are Gaussian and additive (Shannon, 1948), may either overestimate or underestimate the information content, depending on how well the data satisfies these different initial conditions. However, GWN evokes responses, in which amplitude distribution is essentially Gaussian (Juusola et al., 1994; Juusola and Hardie, 2001a), providing more accurate information estimates. The estimated information transfer rates are further influenced by the number and resolution of spectral signal and noise estimates and the finite size of the used data (van Hateren and Snippe, 2001; Juusola and de Polavieja, 2003).

Triple extrapolation method. Macroscopic responses of *Drosophila* (e.g., see Fig. 4C) or *Coenosia* photoreceptors (e.g., see Fig. 13A) or individual Poisson light stimulus traces (Fig. 2; Table 1) were first digitized (compare Fig. 3A,B) by dividing these into time intervals, T_w, that were subdivided into smaller intervals of $t_w = 1$ ms. This procedure selects "words" of length T_w with T_w/t_w "letters." The mutual information between the response S and the stimulus is then the difference between the total entropy, H_s:

$$H_S = - \sum_i P_S(s_i)\log_2 P_S(s_i) \qquad (3)$$

where $P_S(s_i)$ is the probability of finding the i-th word in the response, and the noise entropy H_N:

$$H_N = - \left\langle \sum_{i=1} P_i(\tau)\log_2 P_i(\tau) \right\rangle_\tau \qquad (4)$$

where $P_i(\tau)$ denotes the probability of finding the i-th word at a time t after the initiation of the trial. This probability $P_i(\tau)$ was calculated across trials of identical NS. The values of the digitized entropies depend on the length of the "words" T, the number of voltage levels v (v), and the size (as %) of the data file, $H^{T,v,size}$. The rate of information transfer was obtained taking the following three successive limits (Fig. 3C–E, respectively):

$$R = R_S - R_N = \lim_{T_w \to \infty} \frac{1}{T_w} \lim_{v \to \infty} \lim_{size \to \infty} (H_S^{T_w,v,size} - H_N^{T_w,v,size}) \qquad (5)$$

These limits were calculated by extrapolating the values of the experimentally obtained entropies. A typical response matrix for the analysis contained 1000 points × 100 trials. The total entropy and noise entropy of both recordings and simulations were then obtained from the response matrices using linear extrapolation within the following parameter ranges: size = 5/10, 6/10,...,10/10 of data; v = 4, 5,...,12 voltage levels; T^{-1} = 2, 3,..., 7 points. As adaptation in photoreceptors approaches steady state, their output varies progressively less (Juusola and de Polavieja, 2003). Similarly, the entropies of their responses, when digitized to ≤12 voltage levels, ceases to increase with increasing data size, enabling their limits to be extrapolated in control by linear fits (Fig. 3C–E). Consequently, as few as 20 response traces (each 1000 points long) typically provided similar information rate estimates to 100 traces of the same recording series. For photoreceptor outputs of narrow bandwidths (e.g., at 20°C; Tables 2 and 3) or low SNR (e.g., see Fig. 11), the data were down-sampled to 250 Hz before the extrapolations, giving t_w = 4 ms, which better represented their slow dynamics. However, for estimating the information transfer rates of Poisson light stimuli, which obviously are nonadaptive, the total entropy and noise entropy extrapo-

lation for size and v (Fig. 2; Table 1) were performed with second-order Taylor series, as such fits approximated these limits more accurately (Juusola and de Polavieja, 2003).

Although the triple extrapolation method (Juusola and de Polavieja, 2003) is not founded on statistical assumptions about the response and noise, errors can crop up as it extrapolates to the infinite limit of the three finite parameters (Fig. 3F). The estimation error, used in Figures 6J and 13B, E, is the SD of the extrapolated information transfer rates.

Despite their different principles and assumptions, Equations 2 and 5 produced consistent and, in many cases, similar estimates from the same finite data for relative comparisons (compare Fig. 3G; Tables 1, 2, and 3). However, because both methods require user input for data and parameter selection, we provide both estimates and/or their mean in the results (e.g., see Fig. 7D) to reduce bias in the analysis and its interpretation. Furthermore, to eliminate data size bias within individual recording series from the same cell, we used the same number of response traces (typically 20–100) to estimate their C (Eq. 2) and R (Eq. 5) to the different test stimuli (compare Fig. 4F,J). The number of response traces in each recording series, naturally, reflected an unavoidable experimental tradeoff, as we had a limited time to collect data (to up to eight test stimuli) while stable intracellular recording conditions persisted. Therefore, the average information rate estimates to each tested stimuli can have some (likely small) errors. But more importantly, the differences between these estimates should be realistic and largely unbiased.

Statistics. Test responses were compared with their controls by performing two-sided t tests.

Number of microvilli in rhabdomeres. EM images suggest that there are 30,000 microvilli in *Drosophila* R1-R6 outer photoreceptors and approximately the same amount in *Coenosia* R1-R6 photoreceptors (Gonzalez-Bellido et al., 2011).

Biophysical model of Drosophila *photoreceptor at 25°C.* We used a recently established biophysical *Drosophila* photoreceptor model to simulate voltage responses to time series of light intensities (Song et al., 2012). This model has four modules: (1) random photon absorption model, which regulates photon absorptions in each microvillus, following Poisson statistics; (2) stochastic bump model, in which stochastic biochemical reactions inside a microvillus captures and transduces the energy of photons to variable bumps or failures (compare Fig. 1B); (3) summation model, in which bumps from 30,000 microvilli integrate the macroscopic light-induced current (LIC) response; and (4) Hodgkin–Huxley (HH) model of the photoreceptor plasma-membrane, which transduces LIC into voltage response (see Fig. 6B).

Remarkably, this modeling approach does not require full knowledge of all molecular players and dynamics in the phototransduction to generate realistic responses. From a computational viewpoint, the exactness of the simulated molecular interactions is not critical (Song et al., 2012). As long as the photoreceptor model contains the right number of microvilli, each of which is a semiautonomous sampling unit, and their stochastic bump dynamics (average waveforms, latency distribution, and refractory period) approximate those in the real recordings, it will sample and process information much like a real photoreceptor (Song et al., 2012). The germaneness of the stochastic adaptive sampling framework to represent photoreceptors' neural information processing is further supported by the following observations:

Although the model parameters in the current study were obtained from R1-R6 photoreceptors used in the previous study (Song et al., 2012), the model output closely followed the response waveforms and information transfer rates of the new recordings. Hence, cross-validation of the model to the parameters of individual cells was unnecessary.

The photoreceptor plasma membrane was modeled deterministically by continuous functions (HH model, above), which act effectively as an adjustable scalar during light stimulation. Such filtering does not lose information (data processing theorem), unless there are round-up errors, data clipping, or erroneous singularities in the implementation of these functions (Lazar, 2007; Song et al., 2012; see also further tests below). In real cells, instead, information can be lost when LIC charges up the voltage response. For example, stochastically operating voltage-gated potassium channels, if few in number, may generate noise during this translation. However, similar information transfer rates of
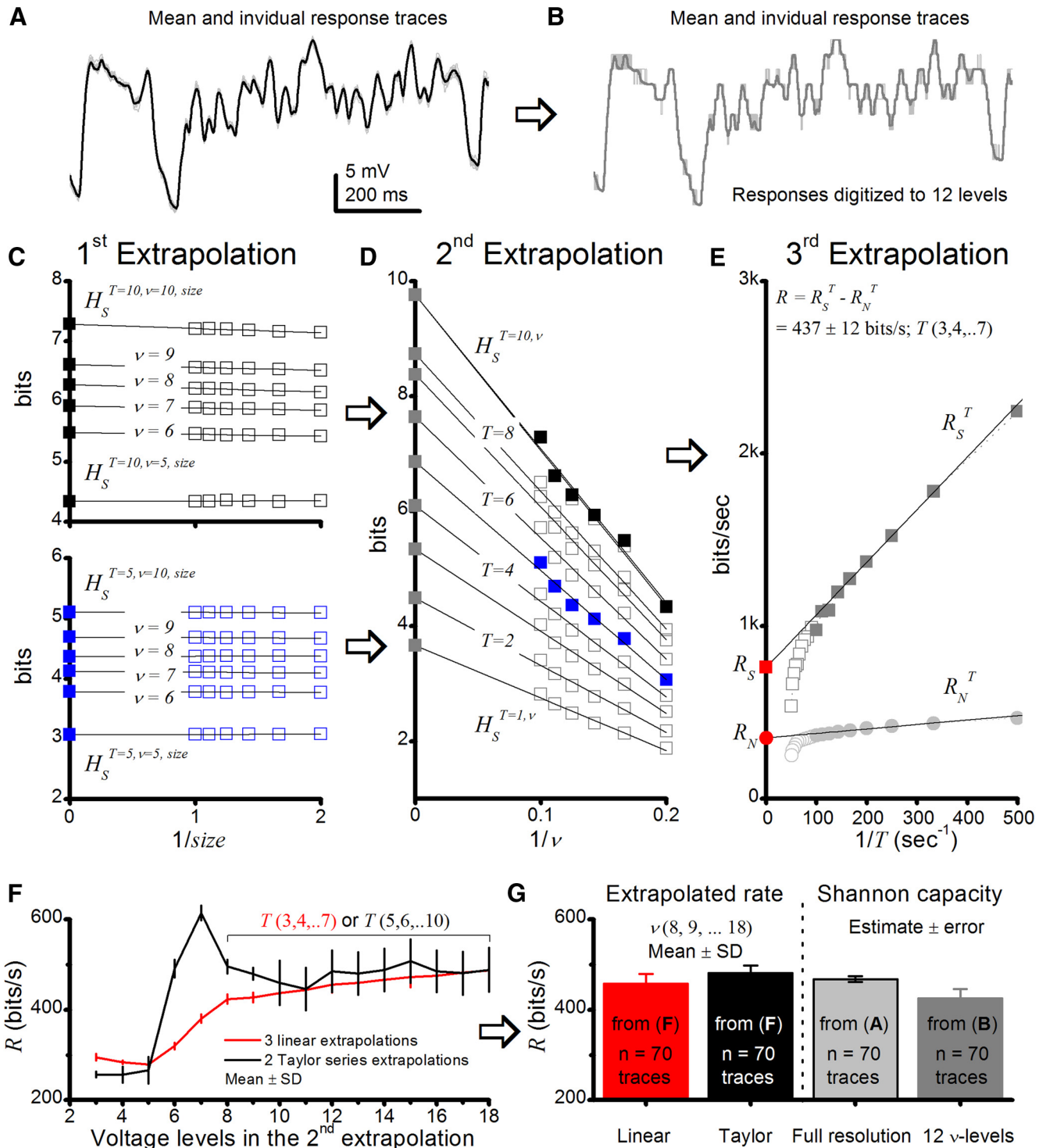
Figure 3. Using linear extrapolations to estimate entropy rate, R_S, noise entropy rate, R_N, and information transfer rate, R, of photoreceptor output. **A**, Mean (black) and 70 voltage responses (light gray) of a *Drosophila* R1–R6 photoreceptor to a 1-s-long naturalistic light intensity time series. **B**, The responses were digitized to 2–18 voltage levels, ν; shown for 12 levels. Entropy, H_S, and noise entropy H_N, are calculated for T-letters-long words, in which each 1-ms-long letter is a voltage level, ν, as explained previously (Juusola and de Polavieja, 2003). **C**, First extrapolation to infinite data size. Entropies of the 10 letter words (top) and 5 letter words (bottom) for 5–10 voltage levels fitted with linear trends. Thus, $H_S^{T=10, \nu}$ and $H_S^{T=5, \nu}$ (black and blue ■, respectively, for $\nu = 5$–10) are obtained from extrapolation of $H_S^{T=10, \nu, size}$ and $H_S^{T=5, \nu, siz}$ for $size \to \infty$ ($1/size \to 0$). Here, the probability of 5 letter words is similar for 50 –100% of data so size corrections in $H_S^{T=5, \nu}$ are negligible, but for 10 letter words size corrections impact $H_S^{T=10, \nu}$ slightly more. **D**, Second extrapolation to infinite voltage levels. $H_S^{T, \nu}$ is shown for words of 1–10 letters, each fitted with its linear trend. H_S^{T} (gray ■s for $T = 5$–10) is obtained from the extrapolation of $H_S^{T, \nu}$ when $\nu \to \infty$ ($1/\nu \to 0$); $H_S^{T=5}$ = blue ■; $H_S^{T=10}$ = black ■. **E**, Third extrapolation. Entropy rates obtained from extrapolations to infinitely long words. The total entropy rate, R_S (red ■), is obtained from a linear extrapolation when $T \to \infty$ ($1/T \to 0$). R_N (red ●) for the same data. Both R_S and R_N collapse to 0 when the data are insufficient to provide an adequate extrapolation of H_S^{T} and H_N^{T} for long words and high voltage resolutions. The graph, however, shows enough linearly aligned points for accurate estimations of R_S, R_N, and R. **F**, Effect of the number of voltage levels ν used in the second extrapolation on R. For $\nu \geq 8$, the first point for the second extrapolation is the fifth voltage level. Linear fits (red) and second-order Taylor series (black) give similar estimates ($<$10% difference) when $\nu = 8$–18 for these data. **G**, Average R estimates obtained from linear (red) or second-order Taylor series (black) fits by the triple extrapolation method (Eq. 5) and from Shannon equation (Eq. 2). These estimates for data in **A** are similar. For 12 voltage level data (**B**), Shannon capacity estimate is only ∼10% less than the estimates for the full response waveforms (**A**), implying consistency in these estimation methods.

Table 2. Rate of information transfer of *Drosophila* R1-R6 photoreceptors at 20°C to different stimuli, estimated from *in vivo* recordings[a]

$\sim 10^6$ photons/s: stimulus	Recordings: complete recording series at 20°C			
	Cut-off frequency	Mean $\pm$ SD (bits/s)	Encoding efficiency (%) R/R_{input}	n (cells)
GWN	20 Hz	215 $\pm$ 18	$\sim$77	8
GWN*	50 Hz	223 $\pm$ 28	$\sim$44	8
GWN	100 Hz	207 $\pm$ 27	$\sim$22	8
GWN	200 Hz	138 $\pm$ 26	$\sim$8	8
GWN	500 Hz	117 $\pm$ 23	$\sim$4	8
NS*	1/f	298 $\pm$ 17	$\sim$10	10 (8 + 2)
NS-shuffled	$\sim$500 Hz	194 $\pm$ 1	$\sim$5	2
Gaussian-1/f	1/f	214 $\pm$ 3	$\sim$7	2

[a]For the simulated responses to each stimulus, the table gives the mean and SD (error) of the extrapolated rate (Eq. 5) and Shannon capacity (Eq. 2). The mean light intensity level used in the model simulations is 10 times lower than that used in the recordings, as it is corrected for the missing screening pigments of the eye (see Materials and Methods).

*Responses to NS had significantly higher information transfer rate than the maximum for GWN stimulation: 50 Hz cut-off ($p = 4.6 \times 10^{-5}$).

Table 3. Rate of information transfer of *Drosophila* R1-R6 photoreceptors at 20°C to different stimuli, estimated from stochastic simulation[a]

10^5 photons/s: stimulus	Stochastic simulations at 20°C			
	Cut-off frequency	Mean $\pm$ SD estimation error (bits/s)	Encoding efficiency (%) R/R_{input}	n (estimates)
GWN	20 Hz	224 $\pm$ 11	$\sim$82	2
GWN	50 Hz	204 $\pm$ 8	$\sim$40	2
GWN	100 Hz	202 $\pm$ 0	$\sim$21	2
GWN	200 Hz	141 $\pm$ 1	$\sim$9	2
GWN	500 Hz	119 $\pm$ 9	$\sim$4	2
NS	1/f	295 $\pm$ 26	$\sim$10	2

[a]For the simulated responses to each stimulus, the table gives the mean and SD (error) of the extrapolated rate (Eq. 5) and Shannon capacity (Eq. 2). The mean light intensity level used in the model simulations is 10 times lower than that used in the recordings, as it is corrected for the missing screening pigments of the eye (see Materials and Methods).

intracellular voltage responses and corresponding simulations (Song et al., 2012) suggest that the photoreceptor plasma membrane adds little noise to encoding.

Full details, including the parameter values of the stochastic model, are given by Song et al. (2012).

Different model outputs. The stochastic photoreceptor model can provide output at three different processing levels: its "bump count" response (see Fig. 8) and macroscopic LIC (see Fig. 11) and voltage responses (see Fig. 6). Previous simulations to daylight stimulus intensities (Song et al., 2012) suggested that *Drosophila* photoreceptor's information rate predominantly reflects the "bump count" response ($\sim$90%), whereas the extra $\sim$10% information in LIC and voltage responses should be carried by bump waveform dynamics, as a memory of the past microvilli activations (i.e., the first bump of a microvillus is on average bigger than the following ones). Therefore, a "bump count" response should have $\sim$10% lower information rate than its corresponding LIC response. Furthermore, because LIC is filtered by continuous HH functions, which affect signal and noise equally (see above), the resulting voltage response should have the same information rate as its LIC counterpart ($\pm$ few bit(s) estimation error). We tested these predictions and found them matching closely with the estimated information rates of the simulated responses. For example, to bright NS-shuffled stimulus (10^5 photons/s) at 25°C, these were as follows: "bump count" response, 239 bits/s; LIC, 257 bits/s; and voltage response, 260 bits/s ($+/- \sim$10 bits/s estimation errors). Likewise, to Gaussian-1/f stimulus, these were as follows: bump count" response, 240 bits/s; LIC, 270 bits/s; and voltage response, 272 bits/s. In both cases, the slightly lower information rates of the "bump count" responses support the idea of bump waveforms carrying extra information, agreeing with our previous results obtained by an alternative analysis (Song et al., 2012).

Biophysical model of Drosophila *photoreceptor at 20°C.* We adjusted stochastic bump shape and latency distribution in the photoreceptor model output by rescaling the corresponding master parameters, n_s and l_a (Song et al., 2012), by the measured Q_{10} values of these processes (Juusola and Hardie, 2001b) (see Fig. 13C,D; Tables 2 and 3).

Stochastic photoreceptors models with different numbers of microvilli. The models were analogous to the biophysical model of *Drosophila* photoreceptor at 25°C, expect that in the simulations we used fewer sampling units (mirrovilli): 300, 900, 3000, or 9000 (see Fig. 11). By using the same bright light inputs (10^5 photons/s) as with the full model, these models helped us to quantify how the number of sampling units limits photoreceptor output (and its information rate) to stimuli with different statistical contents.

Killerfly R1-R6 photoreceptor model. We used the published killerfly (*C. attenuata*) photoreceptor model (Song et al., 2012) (see Fig. 13A,B). Similar to the *Drosophila* photoreceptor model, its macroscopic light current was integrated from current bumps from 30,000 stochastically operating microvilli, responding to either NS or GWN. The average bump shape, latency distribution, and refractoriness are much briefer than those of *Drosophila* (Gonzalez-Bellido et al., 2011; Song et al., 2012). These were obtained from recorded voltage signal and noise estimates and implemented similar to the *Drosophila* model (above). *Coenosia* cell body membrane was modeled using HH formalism (Song et al., 2012), based on *in vivo* current-injection recordings (Gonzalez-Bellido et al., 2011; Song et al., 2012). The simulated macroscopic LIC was injected to the cell body membrane model to obtain the corresponding voltage response (see Fig. 13A).

Stochastic adaptive sampling versus deterministic sampling. Photoreceptors transduce changes in light input into graded changes in their voltage output. Because light information is quantal, carried by stochastic photon arrivals, this constitutes counting (sampling) at the fundamental level (Song et al., 2012; Juusola et al., 2014). Experiments and simulations strongly suggest that a single microvillus transduces single photons into quantum bumps, which sum up macroscopic neural responses (Wong et al., 1982; Hochstrate and Hamdorf, 1990; Henderson et al., 2000; Juusola and Hardie, 2001a; Song et al., 2012). A microvillus can generate only one bump at a time, after which it is rendered briefly refractory (Song et al., 2012), limiting its maximum bump production rate. If bumps are individual samples and microvilli sampling units, then, following Shannon's information theory (Shannon, 1948), the maximum information transfer rate of a photoreceptor is determined by the number of its sampling units and the speed and reliability of their sampling. The close match between simulations and recordings suggests that these considerations hold true at least when photoreceptors have adapted to a relative steady state (Song et al., 2012). But changes in stochastic adaptive sampling may also explain generic changes in photoreceptor output in more dynamic stimulus conditions.

Accordingly, the shape of a photoreceptor's macroscopic response at any time instant can be considered to depend upon four sampling parameters: the number of microvilli (sampling units), the shape of bumps (sample size), the refractory period (determines sample rate), and the latency distribution (determines sample integration precision) (Song et al., 2012).

To obtain a better understanding of how the interdependency of the bump parameters influences the macroscopic response, we built a bump integration model. It uses heuristic rules to incorporate the predetermined sampling parameters. Instead of obtaining bump series through the phototransduction cascade model, these were generated separately so that: (1) each bump had a fixed shape (the average of all bumps from the stochastic phototransduction cascade model) and (2) predetermined latency (stochastic or of prefixed value); (3) no bumps were allowed to emerge in the middle of an ongoing bump response or during its refractory period; (4) the refractory periods had either fixed values or were generated from predefined distributions; and (5) the macroscopic response summed all bump series.

Because this model mimicked encoding of NS and GWN stimuli well, we used it explicitly to study how the sampling parameters affect photoreceptor output and encoding capabilities. Importantly, because of its inherit simplicity of having only four bump parameters, we could fix
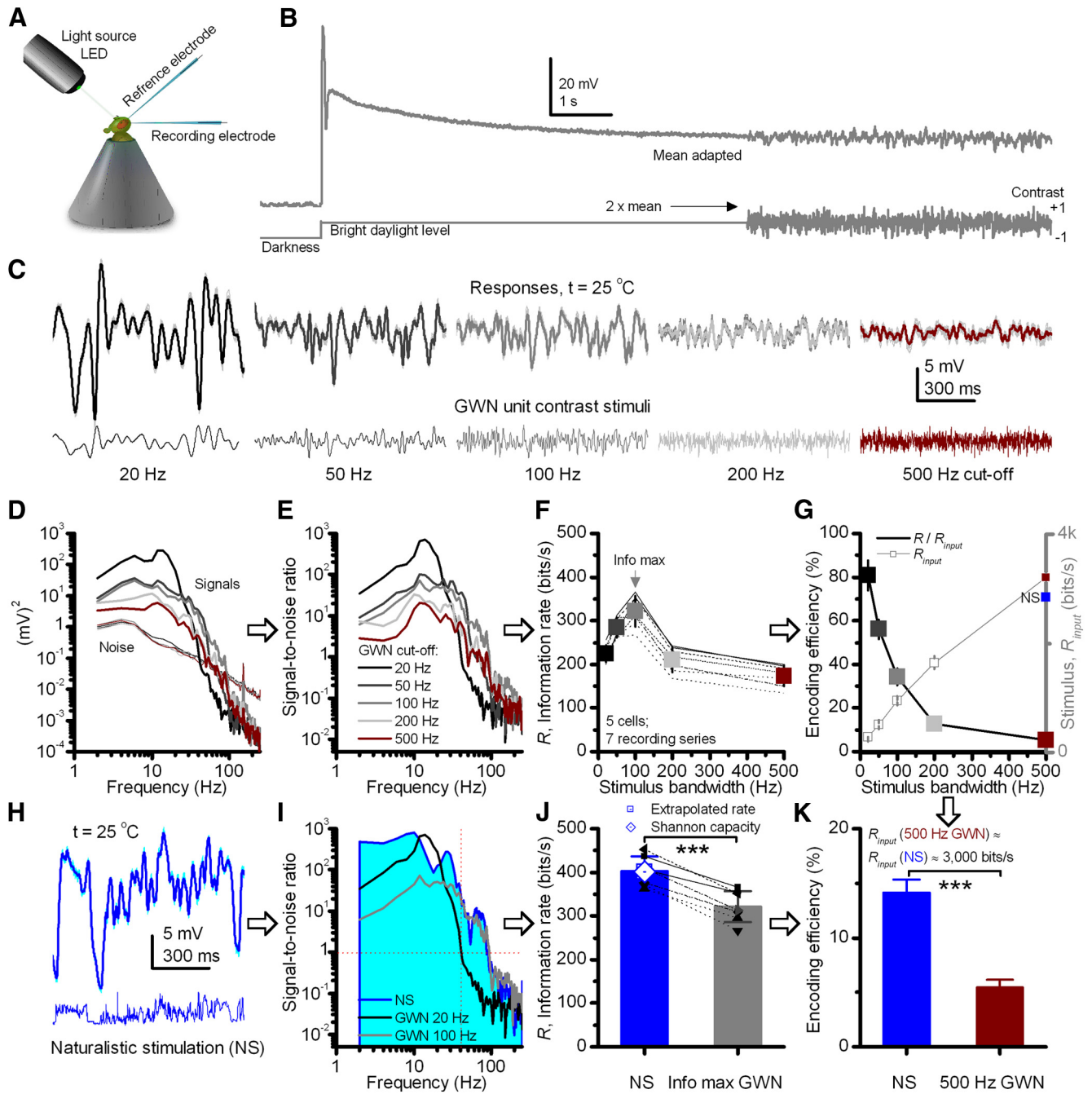
Figure 4. *Drosophila* R1–R6 photoreceptors can sample more information from NS than from GWN stimuli. *A*, Schematic of recording R1–R6 photoreceptors' voltage responses to light stimuli by conventional sharp microelectrodes *in vivo*. *B*, Photoreceptors were adapted to a daylight level (~10^6 photon/s) before GWN stimulation with peak-to-peak ±1 (unit) contrast modulation. *C*, Intracellular voltage responses to unit-contrast GWN stimuli with 20, 50, 100, 200, and 500 Hz cutoffs. Narrow-band (20 Hz) GWN evokes the largest responses; broadband (500 Hz) the smallest. Recordings from the same cell: means (colored), individual responses (thin gray). *D*, Voltage signals (thick) adapt but noise (thin) is unchanged by GWN modulation; power spectra calculated from the difference between individual responses and their mean (signal) for each stimulus. *E*, SNR of photoreceptor output is the highest during 20 Hz GWN. Broadening GWN bandwidth reduces the maximum but widens reliable response bandwidth (SNR > 1). *F*, Broadening GWN bandwidth reduces responses' information rate; decays from ~320 to ~190 bits/s. *G*, With stimulus information rate, R_{input}, increasing steadily with broadening GWN bandwidth, photoreceptors' encoding efficiency (R/R_{input}) decreases exponentially. *H*, Naturalistic light intensity time series (NS) evokes large responses; mean and 10 responses. *I*, Photoreceptors' SNR to NS (cyan) exceeds that to GWN for all the tested bandwidths. *J*, In every photoreceptor, information transfer to NS was >20% higher than to the maximal GWN stimulation (100 Hz cutoff); mean ± SD, $p = 4.2 \times 10^{-5}$. *K*, Photoreceptors encode more efficiently NS than 500 Hz GWN of similar input information rates (~3000 bits/s); mean (R/R_{input}) ± SD, $p = 9.4 \times 10^{-8}$. *F, J*, Different lines indicate individual recordings. In the same cells, information transfer rates between different recordings were very similar (*F, I*, compare two continuous and dotted lines), indicating stable adaptation and recording conditions. All recordings at 25°C. *F, G, J, K*, Data are mean ± SD.

three of them to investigate the role of the fourth one in shaping photo-receptor output to various stimuli.

Drosophila photoreceptor models with and without refractory periods. To study the role of refractory periods on sampling light information, we compared the continuous bump counts of the stochastic model with

those of a deterministic model, having no refractory periods, for both GWN and NS stimuli (see Figs. 8, 9, and 10*A*). In the deterministic simulations, all bump parameters and their corresponding distributions were predefined and obtained from the real stochastic simulations at the same light level (10^5 photon/s). These procedures and assumptions were

used: (1) Each bump had a fixed shape; the average of all bumps in the corresponding real stochastic simulation. (2) The bumps were generated after a predetermined stochastic latency, which followed the same latency distribution generated by the stochastic phototransduction cascade model. (3) The refractory period was set to 0 ms, although no bumps were allowed to emerge in the middle of an ongoing bump response. (4) The macroscopic response was generated by summing all 30,000 bump series, representing simultaneous outputs of all individual microvilli. For both models (with and without the refractory periods), to eliminate the role of bump shapes, the bump counts were obtained by counting the number of bumps at each time point across the 30,000 microvilli.

Mock stochastic model with bump refractory periods taken randomly from the real distribution. These procedures and assumptions were used (see Fig. 10B,C, Random): (1) Each bump had a fixed shape; the average of all bumps in the real stochastic simulation at the bright light level (10^5 photon/s). (2) The bumps were generated after a predetermined stochastic latency, which followed the latency distribution of the stochastic phototransduction cascade model. (3) No bumps were allowed to emerge in the middle of an ongoing bump response or during its refractory period. (4) The refractory periods were generated for the same predefined distribution, obtained from the real stochastic simulations. (5) The macroscopic response summed all bump series. By having latencies and refractory period drawn from a predefined distribution, these parameters are randomly shuffled along time, and hence eliminate any long-term light-adaption in these parameters (information carried by memory of past events in the real stochastic simulations).

Microvilli usage. To compare how different stimuli used microvilli over time, we calculated the microvilli usage from the mock bump integration model (see Fig. 10D). The light inputs were 2 s of 20, 50, 100, 200, and 500 Hz GWN and NS. In all cases, the photon input to each microvillus was generated by the random photon absorption model (Song et al., 2012).

In the mock stochastic model simulations, all bump parameters and their corresponding distributions were obtained from real stochastic simulations at 10^5 photon/s. The following procedures and assumptions were used: (1) Each bump had a fixed shape (the average of all bumps in the corresponding stochastic simulation). (2) The bumps were generated after a predetermined stochastic latency, drawn from the same latency distribution generated by the stochastic phototransduction cascade model. (3) No bumps were allowed to emerge in the middle of an ongoing bump response or during its refractory period. (4) The refractory periods were generated for the same predefined distribution, obtained from the corresponding stochastic simulations. (5) The macroscopic response was generated by summing all bump series. (6) The output state of each microvillus (e.g., generating a bump, refractory) was counted at each 1 ms time bin. This procedure was repeated across 30,000 microvilli to obtain their usage dynamics.

Estimating ATP consumption for information transmission in Drosophila *photoreceptors.* While the microvilli, which form the photosensitive part of a *Drosophila* photoreceptor (Fig. 1), generate the LIC, the photoinsensitive part of the plasma membrane uses many voltage-gated ion channels to adjust the LIC-driven voltage responses. In response to LIC, these open and close, regulating the ionic flow across the plasma membrane. But to maintain the desired ionic concentrations inside and outside, photoreceptors depend upon other proteins, such as ion cotransporters, ion exchangers, and ion pumps, to uptake or expel ions in and out. The work of the pumps in moving ions against their electrochemical gradients consumes energy (ATP). ATP consumption of a *Drosophila* photoreceptor thus much depends upon the ionic flow dynamics through its ion channels (Laughlin et al., 1998). To approximate these dynamics during light responses, we constructed a HH model of the photoreceptor body. This electrical circuit models the ion channels as conductances.

Our HH model included these ion transporters: $3Na^+/2K^+$-pump, $3Na^+/Ca^{2+}$-exchanger and $Na^+/K^+/2Cl^-$ mechanisms to balance the intracellular ionic fluxes. $Na^+/K^+/2Cl^-$ cotransporter balances with the voltage-dependent Cl^- and Cl^- leak conductances, maintaining intracellular Cl^- concentration. Ca^{2+} influx in the LIC (~41%) is then expelled by $3Na^+/Ca^{2+}$-exchanger in 1:3 ratio in exchange for Na^+ ions. Although there is K^+ influx in LIC (~24%), this is not enough to compensate K^+ leakage through voltage-gated K^+ conductances and K^+ leaks. Apart from a small amount of K^+ intake through $Na^+/K^+/2Cl^-$ cotransporter, $3Na^+/2K^+$-pump is the major K^+ uptake mechanism. It consumes 1 ATP molecule to uptake 2 K^+ ions and extrudes 3 Na^+ ions. Because it is the major energy consumer in the cell, we use only the pump current (I_p) to estimate the ATP consumption.

From the equilibrium of K^+ fluxes, I_p can be calculated as follows:

$$I_p = \frac{1}{2}(I_{shaker} + I_{shab} + I_{new} + I_{Kleak} - I_{LIC\_K}) - \frac{1}{4}(I_{cl} + I_{cleak})$$

(6)

where I_{shaker}, I_{shab}, I_{new}, and I_{Kleak} are the currents through shaker, shab, new, and K_{leak} channels, respectively, I_{LIC_K} is the K^+ influx in LIC and I_{Cl} and I_{Cleak} are the currents through the voltage-gated Cl^- and Cl^- leak channels, respectively. These currents can be calculated from the reverse potential of individual ions and their HH model produced conductances using Ohm's law:

$$I_{shaker} = (E_m - E_k)g_{shaker} \qquad (7)$$

$$I_{shab} = (E_m - E_k)g_{shab}$$

$$I_{new} = (E_m - E_k)g_{new}$$

$$I_{Kleak} = (E_m - E_k)g_{Kleak}$$

$$I_{cl} = (E_m - E_{cl})g_{cl}$$

$$I_{cleak} = (E_m - E_{cl})g_{cleak}$$

Using I_p, the number of ATP molecules hydrolyzed per second can be calculated:

$$ATP\ molecules\ s^{-1} = \frac{\int_0^T I_p dt}{T} \times \frac{N_A}{F} \qquad (8)$$

where N_A is Avogadro's constant and F is Faraday's constant. The number of ATP molecules per bit of information was calculated by dividing the estimated number of ATP molecules hydrolyzed in a second by the estimated information transfer rates (bits/s).

We did not model the dynamics of these pump mechanisms because, for the purpose of calculating ATP, only the time-integrated ionic fluxes count, not the time constants.

Previously, because of lack of a complete model for the photosensitive membrane, the LIC has only been estimated at the steady state (Laughlin et al., 1998; Niven et al., 2007), when the sum of all currents across the model membrane equals zero:

$$I_{shaker} + I_{shab} + I_{new} + I_{Kleak} + I_{cleak} + I_{cl} + I_p + I_{LIC} = 0 \quad (9)$$

Because we estimated LIC directly from the stochastic phototransduction model (above), we could calculate a photoreceptor's energy cost in response to any arbitrary light pattern, including naturalistic stimulation (Table 4). Thus, our phototransduction cascade model provides the functional equivalence to the light-dependent conductance used in the previously published models (Laughlin et al., 1998; Niven et al., 2007).

Results

How efficiently are different signals sampled and transmitted in the nervous system? How do neural functions reflect the stimulus conditions they have adapted to? We began analyzing these fundamental questions by measuring with intracellular microelectrode recordings (Fig. 4A) how R1-R6 photoreceptors of common slow-flying fruit fly (*D. melanogaster*) encode GWN stimuli of different bandwidths. These cells form the major sampling matrix of the fly eye and adapt to encode relative changes in light intensity (i.e., contrasts), initiating the achromatic visual

Table 4. Energy usage of *Drosophila* R1-R6 photoreceptors at 25 °C when transmitting information about different stimuli[a]

Mean light intensity of the stimulation at microvilli level (10^5 photon/s)	*Drosophila* R1-R6 photoreceptor model without refractory period			Stochastic *Drosophila* R1-R6 photoreceptor model with refractoriness		
	Information (bits) s^{-1}	ATP molecules s^{-1}	ATP molecules/bit	Information (bits) s^{-1}	ATP molecules s^{-1}	ATP molecules/bit
NS	455	7.538×10^9	1.657×10^7	403	4.633×10^9	1.150×10^7
GWN 20 Hz	271	5.261×10^9	1.941×10^7	267	2.999×10^9	1.123×10^7
GWN 50 Hz	356	5.269×10^9	1.480×10^7	319	3.013×10^9	0.945×10^7
GWN 100 Hz	364	5.265×10^9	1.446×10^7	325	3.014×10^9	0.927×10^7
GWN 200 Hz	282	5.275×10^9	1.871×10^7	231	3.017×10^{9b}	1.306×10^{7b}
GWN 500 Hz	242	5.268×10^9	2.177×10^7	191	3.013×10^{9b}	1.577×10^{7b}

[a]Mean information transfer rate estimates and corresponding energy expenditure for naturalistic light intensity time series, NS, and unit-contrast GWN, with different cut-off frequencies at the mean illumination of 10^5 photon/s. The middle columns indicate the information rates and energy consumption of the photoreceptor model without a refractory period; the right columns summarize the information and energy for the stochastic photoreceptor model. Stochastically operating microvilli reduce energy consumption on average by ~42% and ATP/bit by ~35%.

[b]Transmission of NS is cheaper than transmission of broad-band GWN (≥200 Hz) but more expensive than narrow-band GWN (≤100 Hz).

pathway to the fly brain (Joesch et al., 2010; Wardill et al., 2012). Thus, they constrain how well and fast a fly can see.

Encoding efficiency decreases with broadening stimulus bandwidth

In each experiment, a photoreceptor was adapted to the same daylight level before starting unit-contrast GWN modulation (Fig. 4B), in which frequency range was low-pass filtered to 20, 50, 100, 200, and 500 Hz, while its amplitude range was flanked by contrasts of 1 (twice the mean) and −1 (darkness). This gave each stimulus the average contrast of ~0.32. Theoretically, each of these light intensity time series is maximum entropy among stimuli with the same bandwidth and variance. Because light emission from the source can be modeled by Poisson statistics, we could calculate from photon fluctuation estimates their input information rates during repetitive stimulation (Fig. 2A–D; Table 1; see Materials and Methods). Then, the encoding efficiency of photoreceptors (their recorded outputs) could be assessed as the ratio between the corresponding output and input information rates (R/R_{input}) for each stimulus.

Intracellular recordings (Fig. 4C) showed that the narrow bandwidth stimulus (20 Hz cutoff) evoked the largest responses, whereas broadening the bandwidth reduced the responses. The effect was similar to that caused by increasing the playback velocity of GWN stimulation (Juusola and de Polavieja, 2003), implying that, when the stimulus bandwidth exceeded that of the responses (27.8 ± 1.0 Hz half-maximum cutoff, mean ± SD, 4 photoreceptors), more of its information allocated (i.e., was wasted on) frequencies too fast for the fly to see. This, however, did not affect voltage noise (Juusola et al., 1994) (Fig. 4D), which mostly reflects the average quantum bump (sample) size that integrates the mean responses (Juusola and Hardie, 2001a). Consequently, the mean response (signal; Fig. 4D) and SNR diminished with the broadening bandwidth (Fig. 4E), although, with too narrow a bandwidth, the stimulus lacked higher-frequency information that the fly eye can process and transmit (Zheng et al., 2006; Wardill et al., 2012). Therefore, these findings could much explain why photoreceptors' information transfer rate (Fig. 4F) first rose to ~320 bits/s at 100 Hz stimulus cutoff (where GWN encoding was maximal) before falling to ~190 bits/s at 500 Hz cutoff.

Information in GWN stimulation (R_{input}) increases steadily with broadening bandwidth (Fig. 4G, light gray). Interestingly, however, photoreceptors' encoding efficiency (R/R_{input}; black) neither reached a peak along their information capture of 50–100 Hz GWN (Fig. 4F, gray squares) nor plateaued thereabouts, as could happen if their rhabdomeres sampled light changes with a 100% photon-to-bump conversion rate (quantum efficiency).

Instead, photoreceptors' encoding efficiency decayed exponentially with broadening stimulus bandwidth: from 80% to 90% for 20 Hz stimulation to ~5% for 500 Hz GWN (Fig. 4G, black). These concurrent but opposing trends suggest that quantum efficiency of sampling adapted to the way light information was distributed within the stimulus bandwidth, further contributing to the photoreceptors' encoding efficiency.

Photoreceptors can sample more information from naturalistic stimuli than GWN

To determine whether the maximum information rate to 100 Hz GWN (Fig. 4F), having minimal temporal correlations, represented photoreceptors' capacity, or whether their information sampling could increase with input correlations, we examined encoding of naturalistic stimulation (NS).

Neighboring pixels in natural scenes most probably belong to the same object or background, reflecting similar light intensities, whereas object boundaries or edges reflect differently, separating the world into darker and lighter features (Field, 1987; Ratliff et al., 2010). A naturalistic light intensity time series, as a slice across these features, is not random but contain structured asymmetric contrast variations, which correlate strongly and drive photoreceptor output vigorously (van Hateren, 1997). But NS can also contain less dynamic sequences, such as intensities scanned over a smooth surface; these adapt responses to lower information rates (van Hateren and Snippe, 2001). Therefore, we chose a highly variable NS sequence (Fig. 2E, blue) that included large step-like transitions between longer dark and bright contrasts (van Hateren, 1997; (Zheng et al., 2009) for this test. To attribute potential improvements in encoding to the temporal structure (intensity correlations) of the light patterns (contrasts), NS stimulation was scaled to have the same mean daylight intensity as the GWN.

The given NS sequence (Fig. 4H) consistently evoked large voltage responses with higher SNRs (Fig. 4I) and information transfer rates (Fig. 4J) than the GWN stimuli. The recordings were performed sequentially and occasionally repeated in the same photoreceptor (e.g., Fig. 4F,J shows information rates of 7 complete recordings series from 5 cells). Because the responses were highly reproducible, it became evident that the highest information rate estimate (~320 bits/s) to GWN stimuli did not represent the capacity, whereas, equally, the higher estimate (~400 bits/s) to the NS was unlikely the maximum either. Presumably, there are other light intensity time series, which could evoke responses with even more information. Nonetheless, the findings indicated that *Drosophila* photoreceptors, overall, encoded the NS sequence more efficiently than broadband GWNs (200–500 Hz) of high R_{input} (Fig. 4K), but less efficiently than
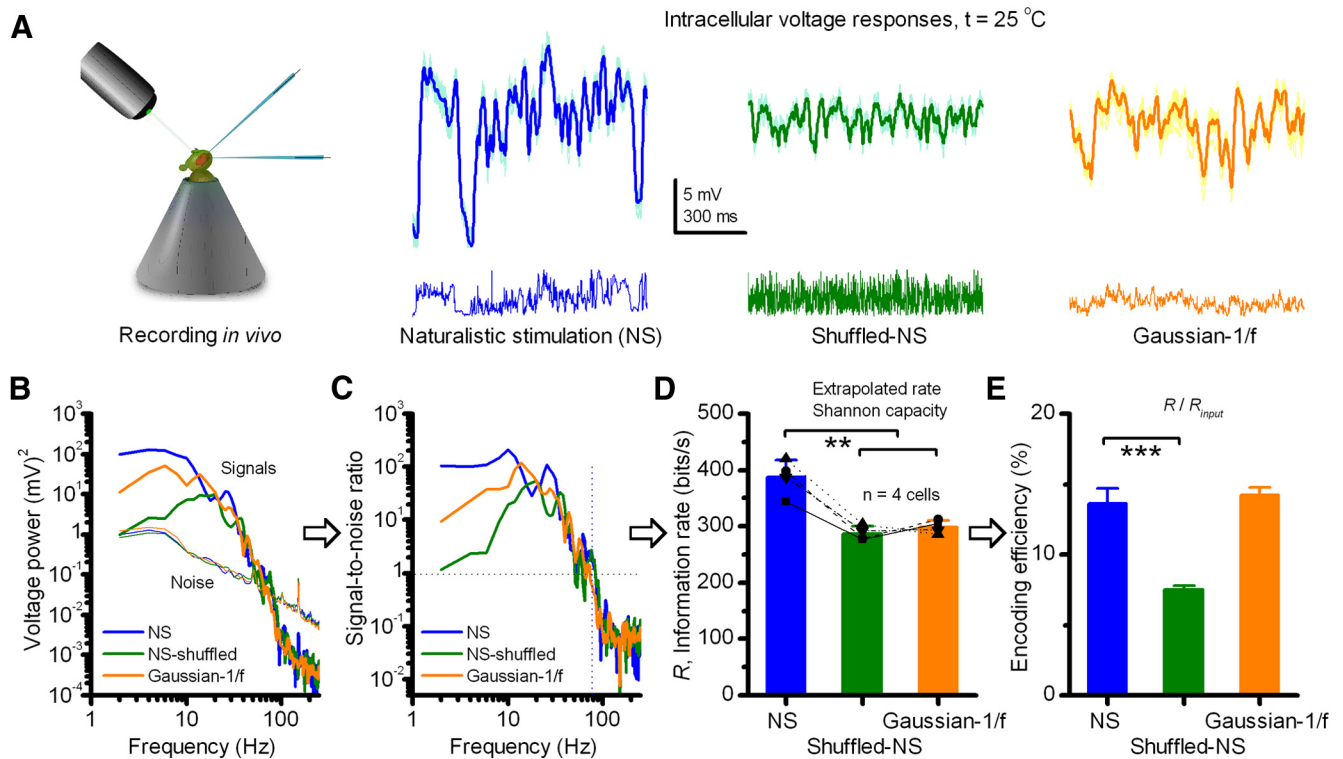
Figure 5. *Drosophila* R1–R6 photoreceptors' responses to naturalistic light intensity time series (NS, blue) are larger and carry more information than responses of the same cells to shuffled-NS (green) or Gaussian 1/f (orange) intensity series. ***A***, Means (thick) and SD (thin) of intracellularly recorded voltage responses to the repeated stimuli. Each stimulus had the same mean light intensity (10^6 photons/s). ***B***, Signal (average response) power to NS is larger than to shuffled-NS or Gaussian-1/f stimuli, but their noise powers are similar. ***C***, SNR of the responses is the highest to NS and the lowest to shuffled-NS. ***D***, Accordingly, the information rate of the responses is the greatest to NS (**p = 2.0 –3.8 × 10^{-2}). Different lines indicate individual recordings from the same cells. ***E***, Encoding efficiency is equally high for naturalistic and Gaussian-1/f stimuli but lower for shuffled-NS (***p = 8.9 × 10^{-4}). ***D***, ***E***, Data are mean ± SD.

narrow-band GWN stimuli (20–100 Hz) of lower R_{input} (Fig. 4G). Specifically, encoding of ≤50 Hz GWN was submaximal because even their input information rates (Table 1) were below or about what photoreceptors can (or are expected to) sample from information-rich NS (Fig. 4J), whereas encoding of ≥200 Hz GWN was inefficient (Fig. 4K) because much of this information was inaccessible to photoreceptors, too fast to be sampled reliably.

To confirm that these encoding characteristics were independent of the ambient recording conditions, we also performed the experiments in flies with ~5°C lower head temperatures (~20°C) (Tables 2 and 3). Predictably, because of more sluggish phototransduction reactions in cooler photoreceptors (Juusola and Hardie, 2001b), responses were slower and their information transfer rates and encoding efficiencies lower (Juusola and Hardie, 2001b) than at the flies' preferred temperature (Sayeed and Benzer, 1996) of 25°C (Fig. 4). However, the photoreceptors' information capture from naturalistic stimulation was again higher than that from GWN stimuli by the same margins (NS: ~300 bits/s; GWN: ~220 bits/s).

Encoding retains sensitivity to highly-structured local contrast changes
To what degree does photoreceptors' high information capture from NS reflect the immediate time order of light intensities (local contrast changes) rather than their global amplitude or frequency distributions? To explore this question, which is about efficiency of information sampling over different time scales, we added two different light intensity time series of equal means to the stimulation regimen. In addition to NS (Fig. 2E, blue), we

now recorded responses to a stimulus (green) that had a randomized time order of NS intensities (decorrelated; spectrally "white"; Fig. 2G) but the same distribution. We also recorded responses to a Gaussian stimulus (Fig. 2F, orange), in which "pink" (1/f) frequency distribution (Fig. 2G) approximated that of NS. These three distinctive stimuli were presented successively to each tested photoreceptor. Our aim, through comparative quantitative analysis, was then to link differences in encoding (these stimuli) to differences in (their) information allocation.

We discovered that the responses (Fig. 5A) to the original NS (blue) were larger than those to the shuffled (green) or "pink" stimuli (orange). Again, all the responses had similar noise power (Fig. 5B), indicating equivalent average bump (sample) size, as expected after adaptation to the same mean intensity (compare GWN experiments above) (Juusola and Hardie, 2001a). But because the signal (average response) power and thus SNRs (Fig. 5C) were lower in responses to shuffled-NS and Gaussian-1/f stimuli, photoreceptors sampled less information from them. The respective information transfer rates were ~74% (286 bits/s; Fig. 5D) and ~78% (301 bits/s) of that to NS (386 bits/s), and these relationships remained the same also at 20°C (Tables 2 and 3). Hence, once photoreceptors adapted to stimulus repetition, the fine time course of light intensity fluctuations (local contrast changes) largely determined their high information capture, with sampling being less sensitive to the global amplitude distribution. Unsurprisingly, the encoding efficiency to shuffled-NS, which had the highest information content but minimal correlations between successive intensity values, much like GWN, was only 7%, half of that to NS (Fig. 5E). Shuffling reduced especially the conspicuous longer intensity fluctuations (phasic or "edge-like"

contrast changes) that evoked the largest responses to NS. Instead, by now flipping (too) fast between intensities, much of this stimulus information became invisible to photoreceptors, akin to high-frequency GWN. More surprisingly, however, we found that the encoding efficiency to NS and Gaussian-1/f stimuli was approximately the same (~15%), despite their different mean contrasts (Fig. 2E,F) and information contents (Fig. 2I). This implies that, while adapting to 1/f frequency distribution of the stimuli, photoreceptors retain sensitivity to the most representative contrast changes, lasting ~≥30 ms (Fig. 5A–C; 2–30 Hz stimulus frequencies).

These findings, together with the results from the broadband GWN stimulation trials (Fig. 4J,K), imply that *Drosophila* photoreceptors can sample more visual information from the more-structured ("bursty" or "phasic") high-contrast changes of appropriate duration than from decorrelated or symmetric (Gaussian) contrast distributions. Thus, GWN, shuffled-NS, or Gaussian-1/f stimulation, regardless of their bandwidth and amplitude modulation and the retina temperature, underestimates the potential information capacity of photoreceptors, which can encode more information from natural-like light intensity fluctuations.

Stochastic photoreceptor model encodes realistically

What is the biophysical basis for these encoding differences? Why did Gaussian or decorrelated light inputs of high information content (Fig. 2I) cause clearly submaximal neural output? To gain mechanistic insight into these open questions, we needed to understand how the different stimuli were sampled and processed at the level of microvilli. Therefore, we first applied the recently established stochastic *Drosophila* photoreceptor model (Song et al., 2012) to simulate responses for the stimuli used in the recordings. The model integrated quantum bumps from 30,000 microvilli to macroscopic LIC (Fig. 6A). This was converted to voltage responses through a HH-type cell body membrane model (Fig. 6B) (Vähäsöyrinki et al., 2006; Song et al., 2012), which also regulated the electromotive force for the light-sensitive current across all microvilli, similar to real cells. The bump latency and waveform dynamics were adjusted by those of the mean adapted photoreceptors (Juusola and Hardie, 2001a; Song et al., 2012), with free parameters fixed (Song et al., 2012).

The simulated voltage responses (Fig. 6C) to the band-limited GWN stimuli closely matched the real recordings (Fig. 4C), showing similar dynamics with the estimated noise power, SNR, and information transfer rate (Fig. 6D–F), suggesting equivalent encoding. Differences were minor and predictable, mostly attributable to the missing recording noise, microsaccadic eye movements, long-term adaptation, and intracellular pupil mechanism that dynamically limits photon influx to microvilli (Song et al., 2012), making the simulations proportionally less noisy at lower frequencies (Figs. 4D and 6D). The simulations further lacked information fed via feedback synapses and gap junctions from the neighboring cells (Zheng et al., 2006; Wardill et al., 2012). These extra inputs likely broadened the bandwidth in real recordings, particularly to 20 Hz GWN (Figs. 4H and 6H, thin red lines).

Nonetheless, most importantly for this analysis, the model output to naturalistic stimulation (Figs. 6H and 7A, blue) differed as predicted from those to GWN, shuffled-NS (Fig. 7A, green) and Gaussian-1/f stimuli (Fig. 7A, orange), closely following the behavior of the real recordings *in vivo* (Figs. 4 and 5). Aptly, the simulated responses to NS had a higher signal power, SNR and information transfer rate than to the other stimuli (Fig. 6I,J: GWN; Fig. 7B–D: shuffled-NS and Gaussian-1/f), whereas the noise powers of the simulations showed virtually identical fre-

quency distributions, as expected for the same mean intensity stimuli (Juusola and Hardie, 2001a) (Figs. 6D and 7B). Furthermore, for the tested high-input information stimuli (Fig. 2I), the simulations had the highest encoding efficiency with NS (Fig. 7E), which evoked the largest response fluctuations. Tables 2 and 3 show these correspondences at 20°C.

Information increases with sample rate modulation

Because the stochastic photoreceptor model sampled and processed light information much like its real counterparts, we could assess with simulations the contribution of microvilli refractoriness to the observed differences in encoding. We did this systematically by comparing, for GWN stimuli and NS, the bump (sample) counts of the stochastic model (Fig. 8B) to the bump counts of a deterministic model (Fig. 8A). In the stochastic model, refractory microvilli cannot produce bumps; whereas in the deterministic model, microvilli had no refractoriness, converting practically every absorbed photon to a bump. In the simulations, the inputs (photon counts) to the two models were identical and the bumps (outputs) generated by their 30,000 microvilli had the same average shape and latency distribution (see Materials and Methods). Therefore, the synchronized ratios between the model outputs (their bump counts) should provide us unique but representative dynamic quantum efficiency estimates for each stimulus (Fig. 8C).

At bright illumination, many stochastically operating microvilli failed to respond to photons because photons arrived/were absorbed during the refractory period. This fall in quantum efficiency (Fig. 8C; bump-to-photon ratio) reduced dramatically the overall sample count (Fig. 8B), here by ~65%. Importantly, the simulations revealed that the diminishing information transfer rates with broadening GWN bandwidth (Figs. 4F and 6F) resulted from diminishing bump fluctuations, henceforth called sample rate modulation. This was caused by the diminishing contrast visible to *Drosophila* and further modulated by the fall and fluctuations in quantum efficiency (Fig. 8C). Nonetheless, the average sample rate to the different bandwidth GWN stimuli remained the same (Fig. 8B, red dotted line). Thus, with the bumps to different GWN being of the same average size (Juusola et al., 1994; Juusola and Hardie, 2001a) (Fig. 6D), the smallest output modulation to 500 Hz GWN simply comprised the fewest bumps (Fig. 8B, wine). As the photoreceptor membrane affects the signal and noise equally during bump integration (Juusola and de Polavieja, 2003; Song et al., 2012) and data processing theorem (Shannon, 1948), the smallest sample rate modulation must have directly contributed to the lowest information transfer rate (Figs. 4F and 6F). At the opposite end of this scale, naturalistic stimulation (blue), which incorporated larger contrast changes in bursty sequences, used microvilli more efficiently. It evoked the largest sample rate modulation (Fig. 8B), causing the highest SNR (Figs. 4H and 6H) and rate of information transfer (Figs. 4I and 6I) in the photoreceptor output.

These observations were quantified by basic statistical metrics. The relationship between the size of sample rate modulation, as measured by SD, and GWN bandwidth (Fig. 8D) traced photoreceptors' rate of information transfer (Figs. 4F and 6F) to the same inputs. The weakest correlation was with 20 Hz GWN (Fig. 8E). Whilst this low-frequency stimulus evoked large sample rate modulation, with its range approaching that of NS, its high SNR allocated only a fraction of photoreceptors' full frequency range. This penalized it against the broader bandwidth responses, with integration over its narrower bandwidth (Eq. 2) giving a proportionally lower information rate.
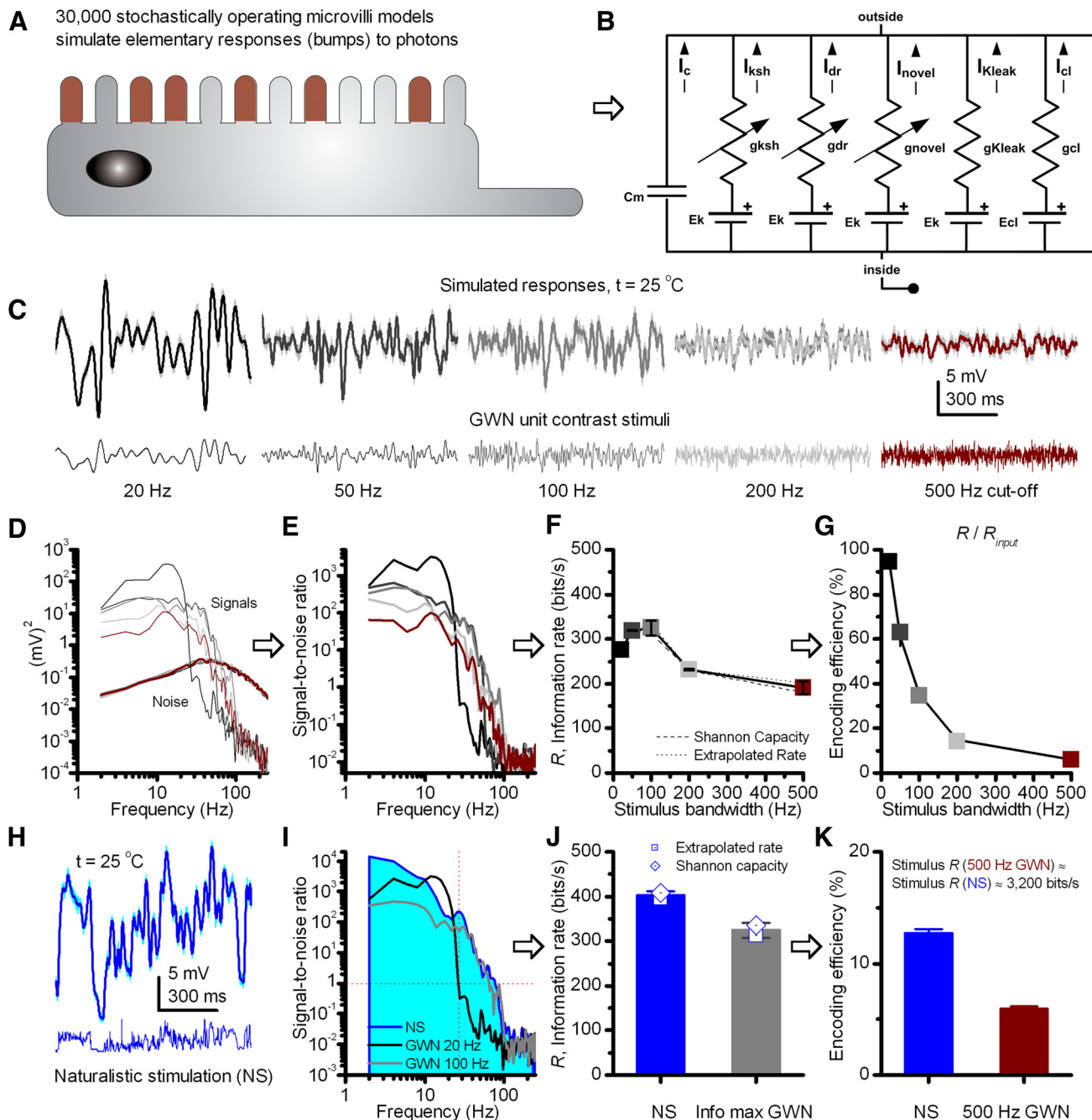
Figure 6. Stochastic *Drosophila* photoreceptor model encodes light information much like a real photoreceptor. *A*, The model generates LIC bumps, mimicking 30,000 stochastically operating microvilli. *B*, Bumps sum up a macroscopic LIC that charges up voltage responses, V$_m$, on an HH-type photoreceptor membrane model, having capacitance, C$_m$, voltage-gated K$^+$ conductances (g$_{ksh}$, g$_{dr}$, g$_{novel}$), and K$^+$ and Cl$^-$ leaks (g$_{Cl}$ and g$_{Kleak}$) (Vähäsöyrinki et al., 2006; Song et al., 2012). *C*, 20 Hz GWN evokes the largest simulated responses; 500 Hz the smallest. Plots show the means (colored) and individual responses (thin gray). The simulations used the same band-limited unit-contrast GWN stimuli as the recordings in Figure 4*C*, but at the light level of 10^5 photon/s. This is because the model lacks the intracellular screening pigments (a pupil mechanism), which reduce light input by ~10-fold (see Fig. 12). *D*, Signals (thin traces) adapt but noise (thick) is unchanged by GWN modulation; power spectra calculated from the time series difference between individual responses and their mean (signal) for each stimulus. Because simulations lack recording noise, muscle activity, and intracellular pupil, there is less low-frequency noise power than in recordings (compare Fig. 4*D*). *E*, SNR of the model output is the highest for 20 Hz GWN. Broadening the GWN bandwidth reduces the maximum but widens the bandwidth of reliable responses (SNR > 1). *F*, The simulations' information transfer rates vary with the GWN bandwidth; from 190 to 320 bits/s. *G*, Encoding efficiency (*R/R$_{input}$*) of simulated photoreceptor output decreases with broadening GWN bandwidth, much like in the real recordings (compare Fig. 4*G*). *H*, NS evokes large simulated responses; mean (blue) and 10 responses (cyan). *I*, SNR of the model output to NS (cyan area) exceeds that to GWN for all tested bandwidths. *J*, Information transfer to NS is >20% higher than to the maximum GWN (100 Hz cutoff), having the same relative ratios as the recordings (compare Fig. 4*J*). *K*, Photoreceptors encode more efficiently NS than 500 Hz GWN, similar to the recordings (compare Fig. 4*K*).

These results support earlier suggestions that photoreceptors' adaptation to different light intensity levels largely reflects a divisive steady-state nonlinearity (van Hateren and Snippe, 2001) and propose refractory information sampling as an important mechanism for it. A photoreceptor's sensitivity is set by its quantum efficiency (mean bump/photon conversion rate). At dim conditions, sensitivity is high (~100% quantum efficiency) because there are many more available microvilli than incoming photons are being sampled (to bumps). Whereas, at very bright conditions, sensitivity is low (≪50% quantum efficiency) be-
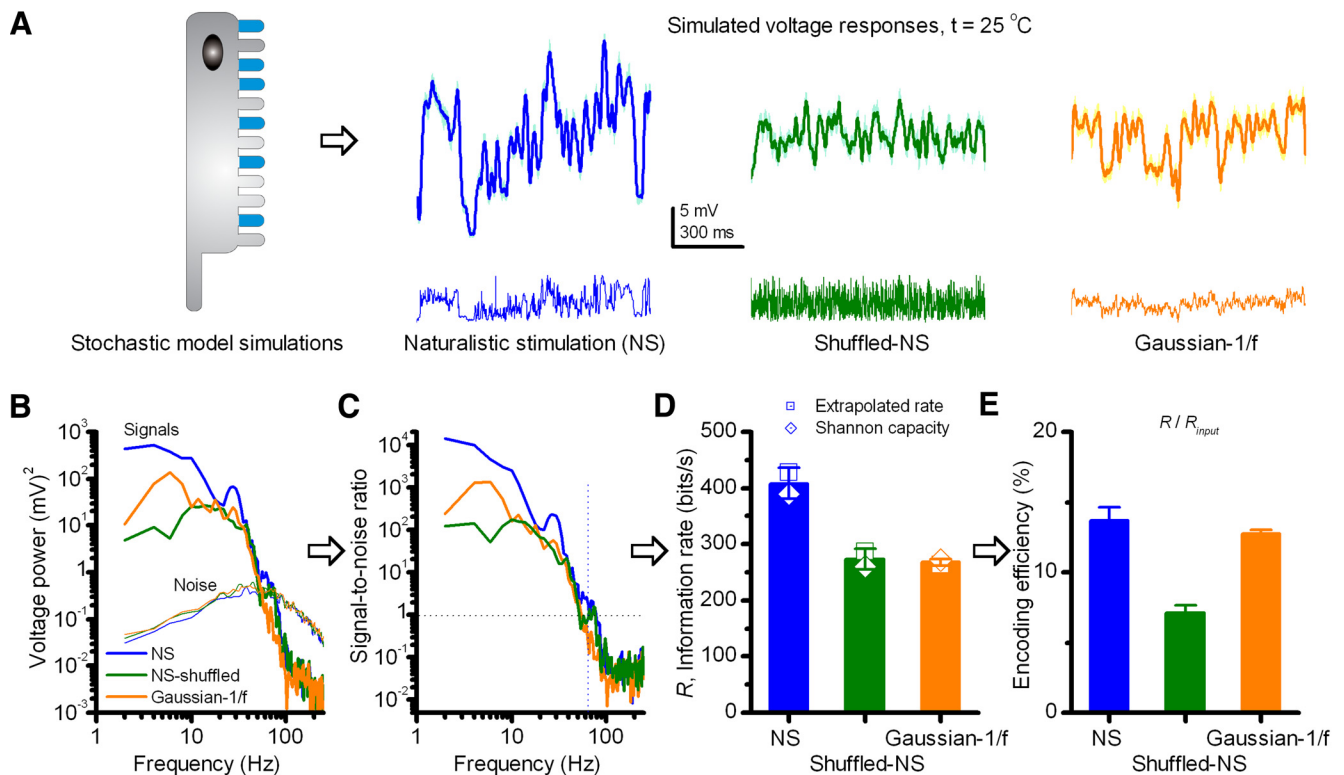
Figure 7. Simulated voltage responses of *Drosophila* R1–R6 photoreceptors to naturalistic light intensity time series (NS) are larger and carry more information than those to shuffled-NS or Gaussian 1/f intensity series of the same mean (10^5 photons/s). ***A***, Means (thick) and SD (thin) of the simulated responses to the repeated stimuli. ***B***, Signal (average response) power to NS is larger than to shuffled-NS or Gaussian-1/f stimuli, but their noise powers are similar. ***C***, SNR of the responses is highest for NS and lowest for shuffled-NS. ***D***, Accordingly, the information rate of the responses is the greatest to NS. ***E***, Encoding efficiency is the highest to NS and lowest to shuffled-NS. The simulated responses behaved very similar to the real recordings (compare Fig. 5). A different Gaussian-1/f stimulus sequence (of the same statistics) was used in real recordings (compare Fig. 5A), but this made little difference to signaling performance (as predicted by the stochastic adaptive photoreceptor model).

cause many microvilli are refractory and most incoming photons are not sampled (to bumps). Markedly, these effects are further augmented by subsequent adaptation in average sample size; the bumps are large in dim and small in bright illumination (Wong et al., 1982; Juusola and Hardie, 2001a). Dynamic changes in their waveforms also contribute, but to a lesser extent, to photoreceptors' information transfer (Song et al., 2012).

Refractoriness accentuate intensity changes in image pixels
We further discovered how stochastic refractoriness of light-activated microvilli, through exerting a memory of past events in bump integration (Song et al., 2012), accentuated certain stimulus features relative to others. To elucidate this phasic nonlinearity and how it contributed to information sampling, we first highlight simulated and recorded examples where fluctuations in refractoriness, as measured by quantum efficiency, shaped responses (Fig. 9A) to 20 Hz GWN (left), 100 Hz GWN (middle), and NS (right). In these plots, sensitization is indicated in green and desensitization in orange.

The simulations in left and middle panels of Figure 9A show how, after longer darker stimulus periods (bottom, dark contrasts, black), transient increases in quantum efficiency (middle; QE_1 and QE_4) sensitized photoreceptor output (top; local voltage responses: R_1 and R_4) to light increments (bottom; light contrasts: C_1 and C_4). Equally in the same traces, brighter periods of stimulation reduced quantum efficiency (QE_2 and QE_3) to same-sized (C_2) or larger (C_3) light increments, desensitizing photoreceptor output (R_2 and R_3). In other words, after darker periods, fewer microvilli were refractory and a photoreceptor sampled

more bumps, integrating larger macroscopic responses to comparable light increments than after brighter periods. As expected, fluctuations in stochastic microvilli refractoriness were prominent to NS (right), in which dark contrasts are more representative (Ratliff et al., 2010). Thus, these features (e.g., C_{5-6}), by strongly modulating the number of microvilli that participates in the response, increased phasic nonlinearities in quantum efficiency (QE_{5-6}) and photoreceptor output (R_{5-6}), comparable with what we see in the real recordings (top) (compare Juusola, 1993; Song et al., 2012). The simulations also revealed how photoreceptors' information transfer rate increases with quantum efficiency fluctuations (Fig. 9B,C). These basic correlative analyses further imply that quantum efficiency fluctuations modulate bump integration (sample rate modulation: Fig. 8D,E), supported by strong relationships between the magnitude of these fluctuations, simulation bandwidth, and corresponding information transfer rates.

These results imply that stimulus-dependent changes in microvilli refractoriness are likely the main reason why *Drosophila* photoreceptors sample visual information from different stimulus statistics differently. However, the weakness of our observations and interpretations is that these much depend upon the accuracy of the dynamic quantum efficiency ratio, estimated between the two models: with and without the refractory microvilli. Therefore, to reduce potential bias in our conclusions, we next compared the normalized outputs (bump counts) of the models (Fig. 10A,B) directly. The normalization of each model output was done for each stimulus; by dividing the respective bump count (at each 1 ms time bin) by its maximum (across the time bins).

The normalized sample counts of the stochastic simulations (colored traces) appeared clearly more differentiated than the deterministic ones (dark yellow histograms) (Fig. 10A), which essentially depict normalized counts of the photons reaching the microvilli. Most obviously, as expected, the normalized bump counts of the stochastic model to NS (right) accentuated contrast fluctuations more than those of the deterministic model ($SD_{stochastic} > SD_{deterministic}$; $p = 2.3 \times 10^{-3}$; $n = 20$ nonoverlapping 50-ms-long samples), but these effects were noticeable even in the bump counts to 500 Hz GWN stimulus (left). Refractoriness thus caused sample rates to react more rapidly and strongly (arrows) to sudden light increments and decrements, exerting a short period of amplification, relative to the foregoing rate. Specifically, the early on-responses to light increments comprised more samples, and thus had better SNRs than subsequent responses, for which fewer microvilli were activated. Likewise, the early off-responses to light decrements were larger as more microvilli remained quiescent than in subsequent responses because initially many were still refractory. The polarity of the rate change is not critical; the bigger and faster the increment or decrement, the more information it can potentially carry. Consequently, a photoreceptor's information transfer is higher at large dim-to-bright or bright-to-dim stimulus transitions and then reduces in correlation with the adaptation to the stimulus, as quantified by previous experiments (compare Juusola and de Polavieja, 2003, their Fig. 4).

These differences were equally distinct compared with sample rate modulation of a closely related deterministic model, in which bump refractory periods were random (thereby abolishing any memory)

Figure 8. Stochastic microvilli refractoriness reduces sensitivity (sample numbers) through a large fall in quantum efficiency (bump-to-photon conversion ratio). **A**, Sample (bump) count of a deterministic photoreceptor model to unit-contrast GWN (20, 50, 100, 200, and 500 Hz cutoffs) and NS of the same mean brightness (10^5 photon/s). Every absorbed photon causes a bump, except those hitting the same microvillus before the previous bump has ended. Here, only ~7% failed. **B**, In the steady-state-adapted stochastic model, microvilli refractoriness reduces sample (bump) counts to GWN and NS by ~65% (dotted red lines); compare the lower sample counts with the same stimuli (**A**). Photon absorbed in the same microvillus before the previous bump ends or during refractory period are lost. **C**, Quantum efficiency was estimated as the ratio between the sample counts of the stochastic (**B**) and deterministic (**A**) photoreceptor models at each moment of time; it shows coordinated fluctuations for each stimulus, which represent a dynamic account of successful microvilli activations. **D**, Sample (bump) rate modulation of stochastic photoreceptor model (**B**) was quantified as SD of the sample counts to each stimulus. Mean ± SD to each stimulus was calculated from three data sections (50% overlapping 500-ms-long windows) that matched those used in the information rate estimation (see Materials and Methods). **E**, Sample rate modulation (as SD) to the given stimuli correlated clearly with their corresponding information rate estimates (from Fig. 6 F, J).

but from the same distribution (Fig. 10B), indicating that refractory periods contain information of their past bumps. Again, the phasic accentuation of bump counts was visually most distinctive to NS (right: $SD_{stochastic} > SD_{random}$; $p = 1.2 \times 10^{-2}$; $n = 20$ nonoverlapping 50-ms-long samples), especially to long and large contrast fluctuations, but it was also evident to larger transient changes in GWN stimulation (left: $SD_{stochastic} > SD_{random}$; $p = 1.6 \times 10^{-3}$, $n = 20$). Notably, the enhancement was independent of stochastically changing bump shapes, which also carry memory (Song et al., 2012), as only the bump times were counted here. Because of this accrued information and the sample rate modulation being still large (Fig. 8B), the stochastic microvilli refractoriness reduced information in photoreceptor output only by ~12% (Fig. 10C). If, instead, the bump refractoriness was assigned randomly, the information transfer fell by ~17%. Thus, in bright stimulation, microvilli refractoriness stores some information (or loses less) when it exchanges sensitivity (reduction in sample
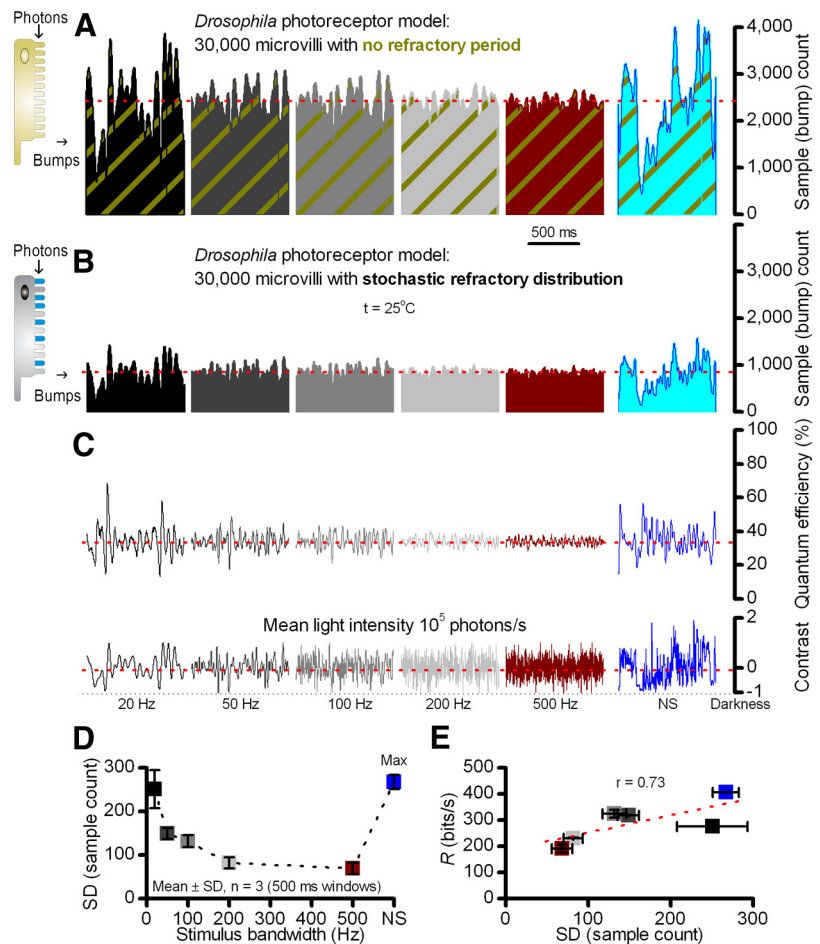
numbers) for contrast resolution (relative increase in sample rate modulation), sculpting more phasic responses. This nonlinear neural enhancement of intensity changes in image pixels (Burton and Laughlin, 2003; Brinkworth et al., 2008; Zheng et al., 2009) likely improves motion (Joesch et al., 2010) and features (Marr and Hildreth, 1980) detection downstream.

Encoding efficiency depends upon microvilli recovery

The mean microvilli usage over time, including both refractory and activated microvilli, was ~40% for both GWN and naturalistic stimuli (Fig. 10D), here mostly reflecting the same mean light intensity. With the fall in quantum efficiency, the photoreceptor output adapted close to the midpoint, which theoretically enables near-maximal responsiveness to light increments and decrements (Song et al., 2012) while being much protected from saturation. Conspicuously, naturalistic stimulation, with its larger and bursty contrast changes, was driving the microvillar
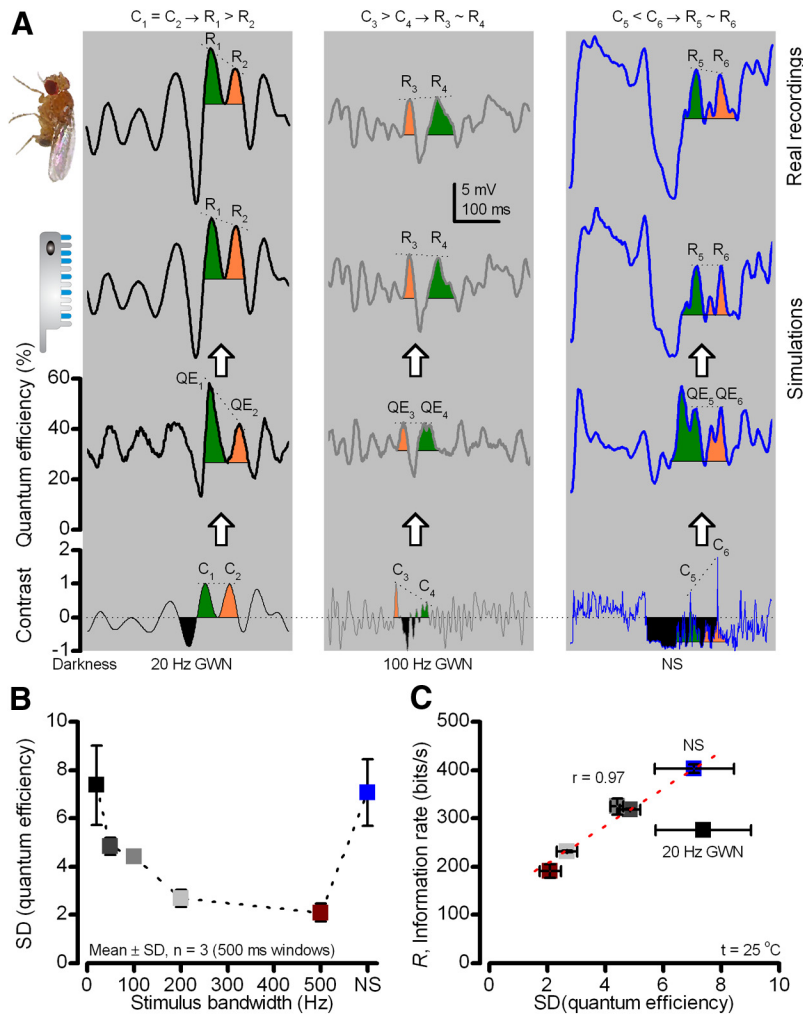
Figure 9. Fluctuations in quantum efficiency accentuate certain stimulus features relative to others. More microvilli recover from refractoriness (and are available to generate bumps) during dark than light contrasts. The resulting increase in quantum efficiency (QE) enhances bump production rate to a subsequent light increment (positive contrast: C), sensitizing photoreceptor output (R). ***A***, Gray panels (left, middle, right) represent how quantum efficiency shapes voltage responses to 20 Hz and 100 Hz GWN and NS, respectively. In each case, the recorded voltage responses (top first row) are compared with the corresponding simulations (second row) of the stochastic photoreceptor model. The quantum efficiencies were estimated as in Figure 8. Signals sensitized by increase in quantum efficiency are shown in green; desensitized ones in orange. Left, Following a larger dark contrast (black area), increase in quantum efficiency QE_1 enhances the response R_1 to the first contrast C_1, whereas subsequent quantum efficiency QE_2 reduces, desensitizing the second response R_2 to an equal-sized contrast C_2. Middle, Contrast C_3 is ∼3 times C_4, but with more dark contrasts (black areas) preceding C_4, $QE_3 = QE_4$, responses R_3 and R_4 reach the same size. Right, Dark contrasts are more representative during NS, causing more dynamic quantum efficiency fluctuations. Contrast C_6 is ∼2 times of C_5. Dark contrasts (black areas) before C_5 are larger than those before C_6, fluctuating quantum efficiency (QE_{5-6}) so that responses R_5 and R_6 become approximately equal. ***B***, SD of quantum efficiency is larger during NS than 50 –500 Hz GWN stimuli, but not for 20 Hz GWN. ***C***, SD of quantum efficiency during different stimuli correlates with the corresponding information transfer rate of photoreceptor output. Average SD to each stimulus was calculated from three data sections (50% overlapping 500-ms-long windows) that matched those used in the information rate estimation (compare Fig. 8 *D, E*).

population far more dynamically than GWN stimuli, which lack these amplitude and phase correlations.

To summarize, our results from the *Drosophila* photoreceptor models (Figs. 7, 8, 9, and 10; for daylight conditions) suggest that the longer dark contrasts of naturalistic light intensity time series facilitate recovery of refractory microvilli. Because NS further contains large light contrasts, photoreceptors can sample up larger and information-richer responses than from GWN, shuffled-NS, or Gaussian-1/f stimuli, all of which have less these features. Dark contrasts thus increase encoding efficiency by improving quantum efficiency and microvilli usage over time.

Importantly, however, stochastic refractoriness shapes nonlinearly photoreceptor output to all stimulus statistics. It exerts sensitivity control on information sampling in two parallel ways. On one hand, it exerts divisive steady-state nonlinearity by adapting the mean bump/photon conversion ratio (quantum efficiency) to current light conditions. On the other hand, it exerts phasic nonlinearity on bump integration by enhancing sudden stimulus fluctuations in the expense on background, the mean bump production rate. With the overall bump count reduced (by the fall in quantum efficiency), refractory microvilli inevitably sacrifice some information. However, this reduction in information transfer is relatively small, as bump fluctuations in each time moment still contain several hundreds of samples, in which integration generates a smooth response of very high SNR.

We next examined how three related physical factors determine photoreceptors' signaling performance through bump (sample) rate modulation: (1) the number of sampling units (fewer vs more microvilli), (2) light intensity (fewer vs more photons), and (3) speed of sampling (slower vs faster microvilli).

Performance improves with increasing microvilli numbers

What happens to signaling performance if photoreceptors had fewer microvilli? Rhabdomeric photoreceptors of insects show structural and functional adaptations, including different microvilli numbers, associated with special habitats and lifestyles (Gonzalez-Bellido et al., 2011). Moreover, insects go through a metamorphosis whereupon their larvae and adults have quite differently structured photoreceptors (Frolov et al., 2012). These adaptations seemingly support unique behaviors in prevailing light conditions of often quite different visual environments. Many questions about visual behaviors or capabilities thus remain but with limited quantitative evidence how these relate to photoreceptors' ability to sample light information. For this purpose, we analyzed by simulations how microvilli numbers may contribute to insect vision.

A photoreceptor's signaling performance depends upon sample rate modulation, originating from light statistics (Juusola and Hardie, 2001a; Juusola and de Polavieja, 2003) and microvilli availability (Howard et al., 1987; Song et al., 2012). The more light entering the photoreceptor and the fewer microvilli available to encode it, the more saturation could compromise the performance. To start untangling the contributions of these factors in encoding, we changed systematically the bandwidth of unit-contrast GWN stimulus and microvilli numbers in model

simulations. Crucially, we further tested how well these models encoded naturalistic stimulus of the same daylight level. The stochastic phototransduction in each microvillus was modeled by that of *Drosophila*, generating current bumps that integrated macroscopic LIC responses.

The results showed that responses increase with microvilli numbers for all stimulus bandwidths (Fig. 11*A*), improving the SNR of the photoreceptor output (Fig. 11*B*). Overall, microvilli resisted complete saturation amazingly well; as always, some were stochastically returning to the pool of available ones. In particular, if the stimulus contrast was predominantly within the visible low-frequency range (i.e., occurred slower than a photoreceptor's integration time), as few as 300 stochastically operating microvilli could generate responses to bright narrow band GWN (20 Hz cutoff) and NS with >10 times more signal than noise. However, the more microvilli the models contained, the finer stimulus changes these could resolve as different, broadening the output range, shown for 100 Hz GWN stimulus (Fig. 11*C*).

Remarkably, the model with only 300 microvilli could encode nearly 100 bits/s from NS and 20 Hz GWN (Fig. 11*D*), yet it performed poorly with high-frequency stimuli (~20 bit/s), generating minuscule noisy responses (Fig. 11*A*). Nonetheless, the increasing microvilli numbers and information capture from the different GWN stimuli formed characteristic relationships of diminishing returns, suggesting transition points between the affordable slow and the more expensive fast vision for diurnal insects. For instance, the models with 300–3000 microvilli encode approximately the same amount of information from NS and 20 Hz GWN, and only with >3000 microvilli, encoding of NS become more profitable. Moreover, having 10 times more microvilli (from 3000 to 30,000) only improves low-frequency (20 Hz cutoff) contrast detection by ~30% (from 188 to 267 bits/s) but should increase maintenance costs greatly. Consequently, slow diurnal insects or larvae need not invest in a myriad of microvilli to see slow contrast changes well (compare Frolov et al., 2012). However, to encode more fast information, a fly photoreceptor may invest in more microvilli; here, to have at least 9000 of the modeled kind.

Photoreceptors' encoding efficiency increased with microvilli numbers for all the tested light stimuli (Fig. 11*E*). This implied that, with more microvilli available, a smaller fraction of them were refractory, and the photoreceptors could generate larger sample rate modulations to given light intensity fluctuations. Specifically, we found that efficiency was consistently the greatest for 20 Hz GWN of the lowest input information content, approaching 100%. Thus, increasing microvilli numbers beyond 30,000 could not improve neutral representations of low-frequency contrasts any
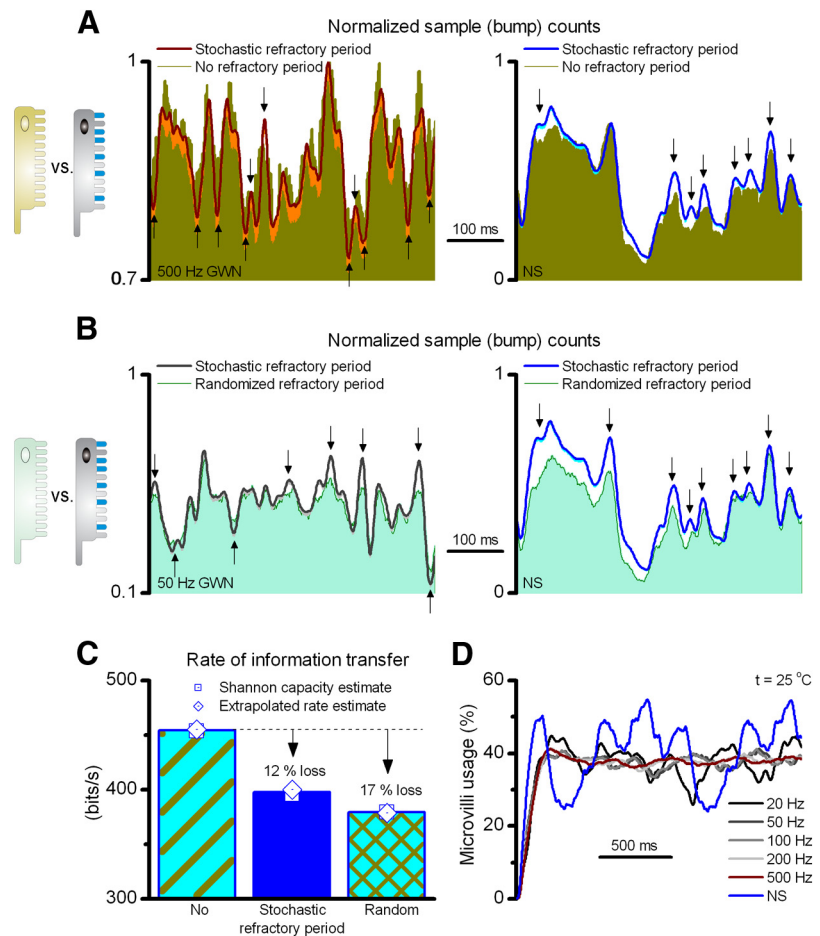

Figure 10. *A*, Stochastic refractoriness enhances relative sample rate modulation to contrast changes. Compare the normalized bump counts of the stochastic (means: wine and blue) and deterministic no refractory period (dark yellow area) models to unit-contrast 500 Hz GWN (above) and NS (below). Arrows indicate enhanced phasic components to sudden contrast changes; SDs (orange and cyan). The model outputs are time-aligned. *B*, Stochastic refractory period in *Drosophila* photoreceptor model enhances neural representations of sudden light changes (i.e., intensity changes in image pixels). This is clearly seen when comparing normalized bump counts of the stochastic model (colored lines; mean ± SD) with that of a mock model (green histogram), in which refractory periods have taken randomly from the same distribution. *C*, Stochastic refractoriness reduces information in photoreceptor output to NS only by ~12% (403 bits/s) with respect to the model without the refractoriness (*A*, 455 bits/s). The loss of information for the other tested stimuli was similar: 20 –100 Hz GWN (7 ± 5%); 200 –500 Hz GWN (17 ± 1%), NS-shuffled (13%), and Gaussian-1/f (13%). This limited information loss is attributable to enhanced responses to contrasts, encoded by relatively large changes in sample numbers. In a deterministic model, where the bumps were randomly assigned refractory periods from the same distribution, more information was lost (17%; 379 bits/s). Again, responses to the other stimuli showed comparable losses: 20 –100 Hz GWN (11 ± 10%); 200 –500 Hz GWN (23 ± 3%), NS-shuffled (18%), and Gaussian-1/f (17%). *D*, Mean microvilli usage is ~40%, but NS drove stochastically operating microvilli population more dynamically than GWN stimuli, which lack longer dark or large bursty contrasts.

further but instead be metabolically wasteful. It is noteworthy that this stimulus, because of its low-frequency content, contained longer dark contrasts than the other GWN. These further helped to recover refractory microvilli, and by that increase quantum efficiency of sampling.

Performance improves with increasing light modulation

The easiest way to increase sample rate modulation, and thus photoreceptors' information transfer rate, is to increase stimulus intensity. However, mechanistic interpretation of such dynamics and its limiting factors is nontrivial. In dim illumination, Poisson statistics of photon emission make stimulation noisy (photon shot-noise), whereas the quantum efficiency of fly photoreceptors is virtually 100%, with only few microvilli refractory, by chance. This should cause a photoreceptor's encoding efficiency to be high but information transfer rate (R) low, as it follows the low information content
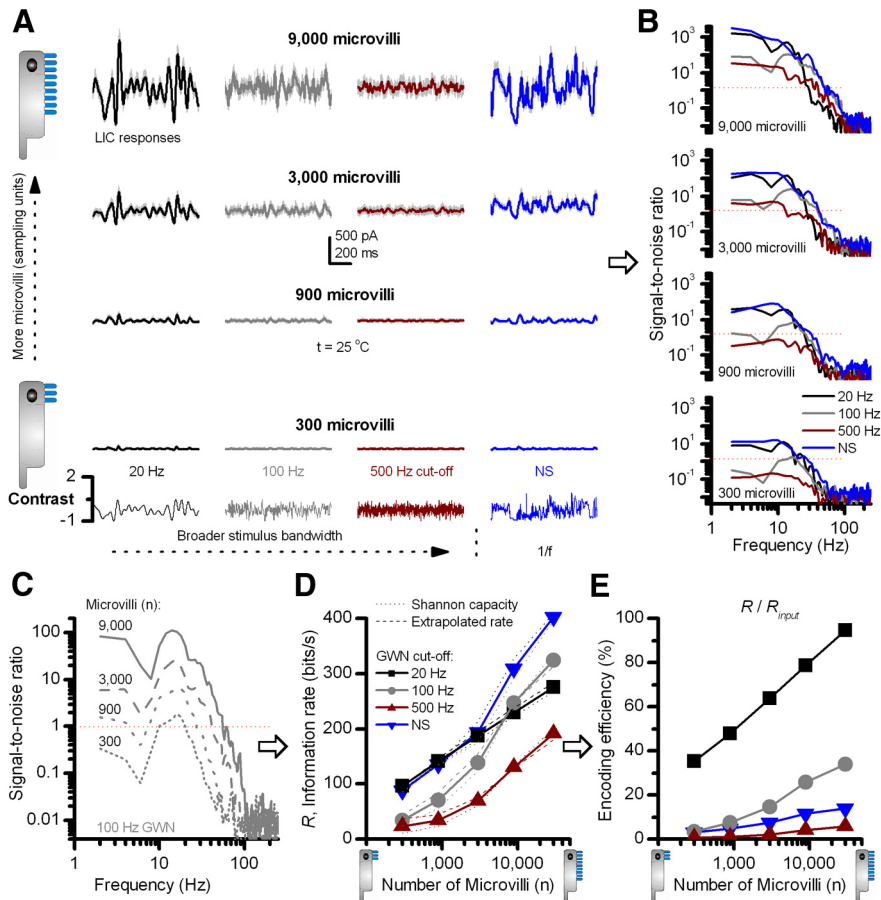
Figure 11. Stochastic simulations show how adding more microvilli (sampling units) enlarges photoreceptor output, its bandwidth, and information to GWN and NS at daylight level (10^5 photon/s). Thus, photoreceptors' bump (sample) rate modulation to light fluctuations increases with stochastically operating microvilli population. **A**, Having more microvilli increases responses to stimulus despite their bandwidth; shown for, 20, 100, and 500 Hz GWN and NS that has "pink" (1/f) power spectra. Macroscopic LIC: mean and 50 responses (light gray). If contrast modulation is too fast, photoreceptors appear saturated. **B**, The SNR of LIC depends on the number of microvilli and is higher for responses to stimuli with large low-frequency modulation (contrast). **C**, SNR to the same GWN stimulus increases with the microvilli population, which thus can generate larger sample rate modulation to contrast changes. **D**, A photoreceptor's information transfer increases by increasing the number of microvilli or the low-frequency contrast modulation (visible to the fly: NS and 20 Hz GWN). Different microvilli numbers lead to different information transfers for different GWN stimuli, suggesting evolutionary transition points where the cost of maintaining microvilli may outweigh any information gains of having faster vision. NS evokes responses with high information rate. **E**, Photoreceptors encode 20 Hz GWN, which has low information content (R_{input}), most efficiently; regardless of the microvilli numbers.

tion at lower light levels than wild-type ones. In comparable simulations, the average bump shapes and their latency distributions were incorporated into the stochastic model from measured values at each light level (Juusola and Hardie, 2001a), as described recently (Song et al., 2012). Notably, the average bumps are larger and slower at dim intensities and smaller and faster at brighter illumination, whereas their latency distributions differ little within the tested light intensity range (Juusola and Hardie, 2001a). By simulating bump production over all 30,000 microvilli, the model was used to predict macroscopic voltage responses during repeated GWN stimulation at each light intensity level. We then compared the signaling performance and dynamics of the real recording series with those of simulated responses.

We found that both recordings and simulations showed similar encoding dynamics (Fig. 12), explainable by theoretical concepts of stochastic adaptive sampling. Increase in light intensity level (and thus modulation) increased sample rate modulation to GWN stimulation, predicting photoreceptors' macroscopic response amplitude (Juusola et al., 1994; Juusola and Hardie, 2001a) and its eventual saturation (Howard et al., 1987). The simulations could also replicate the corresponding relative changes in contrast gain (Fig. 12*A*,*B*), SNR (Fig. 12*D*), and information transfer rate (Fig. 12*E*), although in the simulations, as in real recordings (Juusola and Hardie, 2001a; Song et al., 2012), adapting bump shapes further adjusted its broadening frequency response (Fig. 12*A*,*B*). Moreover, by recording both from wild-type and white-eye photoreceptors, we could further estimate

how far the eye's pigmentation protected microvilli from saturation, extending the efficient encoding range for bright GWN. The difference between the stimulus intensities evoking comparable maximum information transfer rates in these photoreceptors (Fig. 12*E*) suggested that the screening pigments reduced the effective photon input to microvilli by ~10-fold and thus extended the encoding range the same amount. Because the model photoreceptors lack this mechanism, daylight inputs in the simulations were systematically reduced by 10-fold to minimize potential bias to the corresponding recordings throughout this study (see Materials and Methods).

The simulations further characterized how encoding efficiency reduces with increasing mean light intensity, as shown for 500 Hz GWN stimulation (Fig. 12*F*), validating the prediction above. Markedly, as the stochastic photoreceptor model's information transfer rate approached saturation, encoding efficiency fell proportionally the most: from ~10% at 1.5×10^5 photons/s to ~5% at 5.1×10^5 photons/s (arrows).

in light input (R_{input}). But when the light intensity is increased (and so R_{input}), encoding efficiency should reduce with quantum efficiency (Song et al., 2012) as more microvilli become refractory. Ultimately, a photoreceptor's information rate to the brightening stimulation should increase with its increasing sample rate modulation until potential saturation, when its most microvilli are most of the time refractory. Saturation, however, much depends upon stimulus statistics. In particular, it should affect high-frequency GWN, which lack longer dark contrasts. Responses to NS, which has these features, are more immune to saturation effects (Fig. 11), even at very bright mean intensities, as we have shown previously (Song et al., 2012).

We tested these theoretical concepts experimentally by recording voltage responses of photoreceptors in wild-type and white-eye *Drosophila* to 200 GWN stimuli at different light levels, covering nearly 4 log intensities. The white-eye eyes lack screening pigments and intracellular pupil that give the wild-type eye its red appearance and reduce light influx into rhabdomeres. Therefore, white-eye photoreceptors should show signs of satura-

Performance improves with faster sampling

Alternatively, sample rate modulation could be increased by quickening the processing speed (Gonzalez-Bellido et al., 2011) and microvilli refractoriness (Song et al., 2012). To quantify this effect on GWN encoding, we recorded voltage responses of killer fly (*C. attenuata*) R1-R6 photoreceptors to a unit-contrast stimulus with 300 Hz cutoff (Fig. 13*A*; wine). We also simulated their outputs with a stochastic model, in which parameters were obtained from the best recordings. *Coenosia* is a small fast-flying aerial predator of the same *muscoid* family as *Drosophila* but separated by 120 million years of evolution (Gonzalez-Bellido et al., 2011). Its fast photoreceptors have presumably been adapted to prevent blurred vision during acrobatic hunts. However, similar to the slower fruit fly photoreceptors, each has ~30,000 microvilli. In *Coenosia*, though, their processing speed and recovery are ~3 times faster (Song et al., 2012). This enabled *Coenosia* photoreceptors to encode >5 times more information (Fig. 13*B*) from the same GWN stimulus than *Drosophila* photoreceptors (Fig. 13*C,D*).

In each photoreceptor, we also measured its encoding performance to the same natural light intensity time series. Predictably, both *Coenosia* and *Drosophila* photoreceptors extracted more information from the natural stimulus than from GWN, but very different amounts in relative terms (~1.3 and 2.2 times more, respectively). Thus, encoding efficiency (Fig. 13*C,F*) for the given stimuli (300 Hz GWN and NS) varies greatly in the photoreceptors of different fly species.

Last, using biophysical analyses, we examined how refractory microvilli affect the metabolic cost of visual information to different stimulus statistics.

Refractory microvilli reduces energetic encoding costs

Information transfer by neurons is inherently linked to their energy budget. For graded potential neurons, such as photoreceptors, maintaining ion gradients needed for generating continuous voltage responses consumes ATP (Laughlin et al., 1998). Because we used a full stochastic photoreceptor model, in which macroscopic response (LIC) drove the generation of signals on the voltage-sensitive plasma membrane (of HH type), we could estimate the relationship between the rate of transmitted information and its price in ATP generation directly for different stimulus statistics.
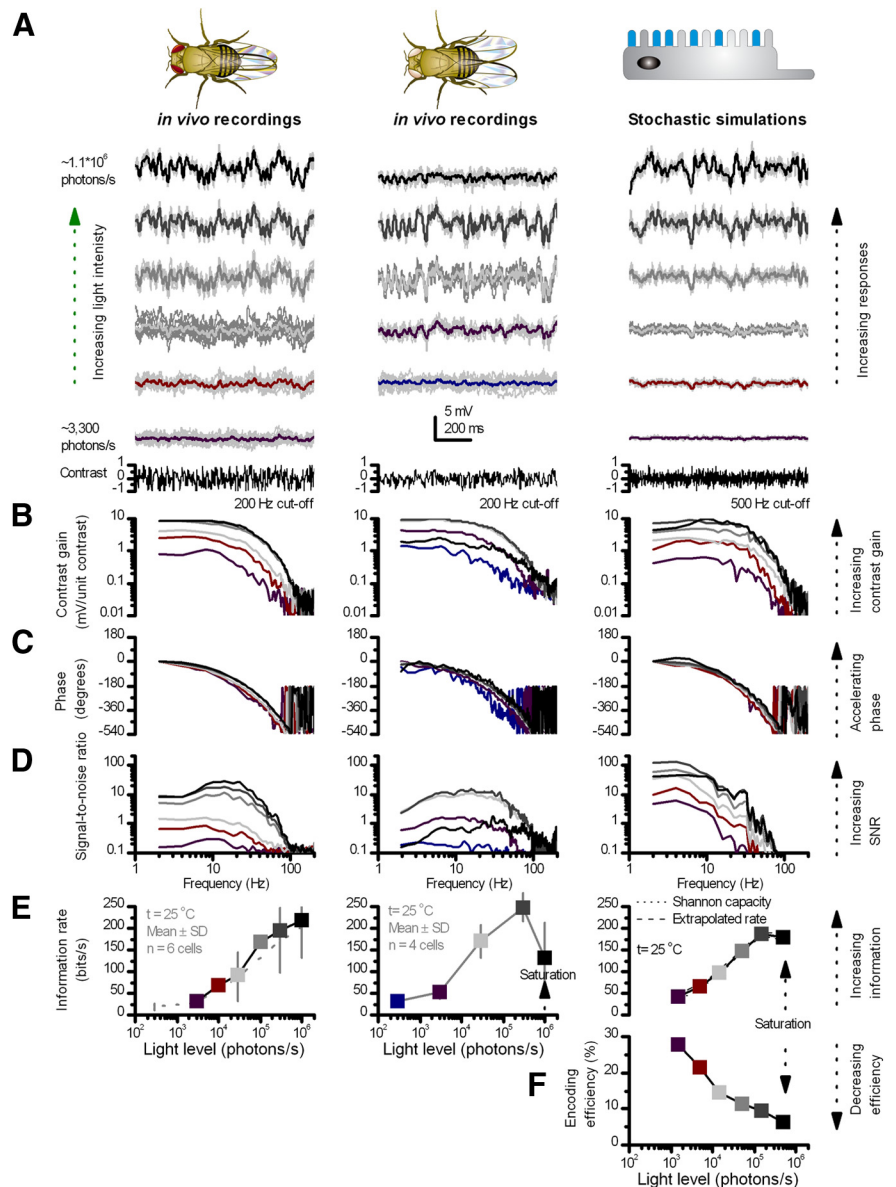

Figure 12. Photoreceptor output, its bandwidth, SNR, and information rate increases with the brightness of unit-contrast GWN stimulation. Bump (sample) rate modulation increases with increasing light modulation until saturation, and thus with it information in photoreceptor output. The stochastic model (right) predicts the experimental results (left, middle). ***A***, Responses of *Drosophila* R1–R6 photoreceptors increase with brightening light modulation, regardless of the stimulus bandwidth (here, GWN with 200 and 500 Hz cutoffs). Left column, R1–R6 WT red-eyed *Drosophila*. Middle, WT white-eyed *Drosophila*. Right, Stochastic model simulations. Mean and 10 responses (light gray). Responses in white-eyed *Drosophila* saturate at the brightest stimulation because its photoreceptors lack screening pigments and the intracellular pupil, which reduce light input in WT microvilli. Consequently, at bright light levels, the majority of microvilli are in a refractory state, and the bright fast GWN modulation drives them inefficiently. ***B***, The contrast gain of photoreceptor output increases with light intensity until bump (sample) rate modulation from microvilli saturate (at two brightest intensities). ***C***, The phase of a photoreceptor's linear frequency response accelerates with increasing light intensity. ***D***, SNR in photoreceptor output to the same GWN stimulus increases with light intensity until microvilli saturate (at two brightest intensities). In white-eyed *Drosophila*, this is most dramatic likely because their photoreceptors experience higher photon influx, including any scattered light. ***E***, Photoreceptors' information transfer rate increases with light intensity and narrowing GWN bandwidth (the contrast modulation in the stimulus visible to the fly), until bump (sample) rate changes from microvilli saturate (at two brightest intensities). Different stimulus bandwidth (compare Fig. 4*F*) causes different rates. But the overall light-induced increase in a photoreceptor's information transfer occurs regardless of stimulus bandwidth. ***F***, Photoreceptors' encoding efficiency to GWN stimuli reduced with increasing light intensity, caused by reduction in quantum efficiency (bump/photon conversion ratio).

We discovered that refractoriness lowers the metabolic cost of information. Every bump costs, and maintaining a higher membrane potential consumes more ATP (Laughlin et al., 1998). Without microvilli refractoriness, thousands of bumps at any one moment would integrate (Fig. 8*A*), yielding a mean potential

Song and Juusola • Refractory Information Sampling

J. Neurosci., May 21, 2014 • 34(21):7216–7237 • **7235**
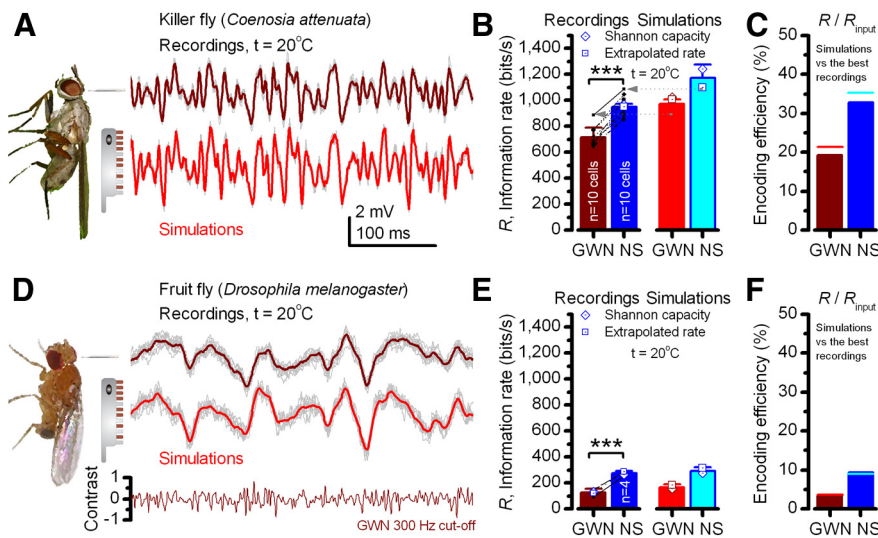
Figure 13. Photoreceptors' information transfer rate and encoding efficiency to GWN and NS are different in flies with fast and slow vision. Both fast-flying *Coenosia* and slow-flying *Drosophila* R1–R6 photoreceptors have 30,000 microvilli, but the former sample light changes and recover from them faster, resulting in higher information capture. ***A***, Voltage responses of a *Coenosia* photoreceptor (brown) and respective stochastic model simulations (red) to unit-contrast GWN stimulation with 300 Hz cutoff; light level: ~10^6 photons/s. ***B***, Information transfer of recorded and simulated *Coenosia* photoreceptors to GWN stimulus (***A***) and NS; these cells can capture from the GWN up to ~80% of the information they capture from NS. Recordings: mean ± SD, $p = 9.5 \times 10^{-7}$; simulations: mean ± estimation error. ***C***, *Coenosia* photoreceptors encoded ~30% of information in NS; encoding is ~1.7 times more efficiently than 300 Hz GWN; GWN R_{input} ~ 4600 bits/s and NS R_{input} ~ 3300 bits/s. ***D***, Voltage responses of a *Drosophila* photoreceptor (brown) and respective stochastic model simulations (red) to the same unit-contrast GWN stimulation as in ***E***. ***E***, Information transfer rates of recorded and simulated *Drosophila* photoreceptors to GWN (***D***) and NS; these cells capture less information from the GWN than NS. Recordings: mean ± SD, $p = 1.2 \times 10^{-4}$; simulations: mean ± estimation error. Furthermore, *Drosophila* photoreceptors encode less efficiently both the stimuli than *Coenosia* photoreceptors. ***F***, *Drosophila* photoreceptors encoded NS ~2.5 times more efficiently than 300 Hz GWN. ***B***, ***E***, Lines indicate recordings from same photoreceptors. In every cell, NS evoked higher information transfer. Here simulated *Coenosia* photoreceptor output carries proportionally more information than the average recordings because it is based on the best GWN and NS recordings (gray arrows). Simulations lack recording noise, muscle activity, and intracellular pupil, which reduce information in recordings. Data at 20°C as in Tables 2 and 3, not 25°C as in Figs. 3-12.

approximately twice as high (30.4 mV vs 16.1 mV), costing ~5.6×10^9 ATP molecules/s (Table 4). Remarkably, with refractoriness the price falls by >40% to ~3.3×10^9 ATP molecules/s, and the price for transmitting one bit of information by ~35%. For naturalistic stimulation, we estimate that a *Drosophila* R1-R6 photoreceptor can encode at least 400 bits/s (likely more, if the stimulus is sped up) (Juusola and de Polavieja, 2003), and consequently the estimated metabolic cost of information is ~10 million ATP molecules/bit. These calculations further suggest that information encoding of naturalistic stimuli is less expensive than encoding of broadband (>200 Hz cutoff) GWN stimulation (compare Laughlin et al., 1998; Niven et al., 2007). Therefore, GWN stimulation underestimates neural performance and can overestimate energy consumption.

Discussion

To survive, living systems must sample information from their environment continuously. They use accumulated information to adapt their structures and functions to work better while facing physical constraints and environmental changes (Barlow, 1961). So, under selection pressures, adaptation to prevailing conditions (Darwin, 1859) characteristically improves efficiency and performance of living systems (see also de Polavieja, 2004; Pérez-Escudero et al., 2009). Recent studies indicate that natural light environment is asymmetrical, containing more dark than light patches (Ratliff et al., 2010) and that this affects how retinae and brains of animals have adapted to represent the world neurally (Ratliff et al., 2010; Kremkow et al., 2014). In this work, we pre-

sented evidence that information sampling in fly photoreceptors may already use this asymmetry to accentuate responses to naturalistic light stimuli, improving vision and likely the flies' fitness.

Our results suggest that light stimuli with intermittent longer dark contrasts help to recover more refractory microvilli than stimuli without them. Therefore, naturalistic signals, which are more representative of such features, can generate larger changes in microvilli activation than equally bright stimuli of different statistical structure, as represented by macroscopic responses with the higher information transfer rate. Importantly, the biophysical analysis of the concurrent energy expenditure further suggests that refractoriness lowers the metabolic cost of neural information. Refractory stochastic information sampling by microvilli thus likely represents the earliest sensory-neural adaptation to the physical characteristics of the world, linking efficiency and costs of sensory encoding to stimulus statistics.

Refractoriness reduces redundancies in bright illumination

Because dark contrasts are more frequent than light ones in the world (Ratliff et al., 2010), they are partly predictable and thus redundant. Therefore, reducing some of their redundancy at the earliest stage, by sampling, may help the visual systems to optimize information transfer for the later processing stages (Barlow, 1961; Atick, 1992; van Hateren, 1992). In this viewpoint, microvilli refractoriness seems an adaptation to reduce related temporal redundancies in fly photoreceptor output. But to remove predictable information, sampling must be carefully adjusted to input information, which varies within scenes and vastly from night to day. Otherwise, photoreceptors integrate information suboptimally within their limited output range, generating noisy, inefficient, or metabolically costly responses. Appropriately, refractoriness (quantum efficiency) adapts to ambient illumination, helping photoreceptors to produce reliable neural estimates of light changes.

In the dim environment of low SNR, photoreceptors' sensitivity and quantum efficiency are high (~100%) with absorbed photons causing large bumps, in which sizes and timings vary considerably. But with photons few and far apart, only a small number of bumps at any time moment integrate a macroscopic response. Encoding is thus redundant and actively utilizes variations in sample latencies and sizes to increase the reliability of the estimated light changes (see also Heimonen et al., 2006; Padmanabhan and Urban, 2010; Angelo et al., 2012).

In the bright environment of high SNR, conversely, more microvilli become refractory. Dark and light contrasts modulate quantum efficiency, accentuating variations in sample (bump) numbers, which integrate the macroscopic response. Refractoriness thus improves contrast resolution, by reducing redundancies, but this comes at the expense of sensitivity. For naturalistic stimulation, this is a low price to pay, as information is not in the

mean voltage, which the refractoriness decreases, but largely in the relative modulation, which the refractoriness increases. Thus, a 65% drop in sensitivity reduced information in the responses by only 12%. Against these costs, photoreceptors can better encode changes in natural scenes, particularly phasic or bursty contrast information; and so likely enable brain circuits to better represent real-world similarities and differences.

High encoding efficiency to low-frequency stimulation

Surprisingly, bright 20 Hz GWN stimulation ($\sim10^6$ photons/s) generated responses with near 100% encoding efficiency in *Drosophila* photoreceptors (Figs. 4G and 6G); that is, low-frequency pattern information was converted into neural information with little loss. How is this possible when the corresponding quantum efficiency was only ~35%, with two-thirds of photons failing to generate a bump?

In photoreceptors, light intensity changes are processed logarithmically (van Hateren and Snippe, 2006) and, in this case, fully within their frequency range. So despite the large photon loss in sampling, the successful ones still transduced ~350 new bumps per millisecond on average. With this many samples (with their sizes and latencies varying stochastically), their population response became smooth and likely free of temporal aliasing (Juusola et al., 2014). The tradeoff of stochastic sampling is broadband noise, but this is reduced naturally as the bumps integrate the macroscopic response. Furthermore, the sample rate modulation was augmented by the relatively long-lasting dark contrasts of this stimulus, which dynamically recovered refractory microvilli (Fig. 9A), reducing saturation effects (Fig. 11A). The macroscopic response thus had a high SNR over the whole (low-frequency) stimulus bandwidth (recordings: Fig. 4D, $SNR_{max} = \sim720$; simulations; Fig. 6D, $SNR_{max} = \sim3220$), approaching the physical limit (Poisson stimulus: Fig. 2C: $SNR_{max} = \sim6600$). Because of $\log_2$ bit-conversion (Eqs. 2–4), the information rate estimates for the maximum SNR outputs and input are close (recordings: 255 bits/s; simulation: 267 bits/s and stimulus: 282 bits/s, respectively, giving 90%–95% encoding efficiency).

Although one may underestimate the light intensity level in the experiments, and so to overestimate encoding efficiency by some ratio, such bias seems unlikely here. Information rate in *Drosophila* photoreceptor output to naturalistic stimulation improves only marginally with increasing mean intensity after $\sim10^6$ photons/s, attributable to a steep fall in quantum efficiency (Song et al., 2012). This forces photoreceptors' sample rate modulation to approach similar dynamics at bright light levels ($>10^5$ photons/s), likely making the world consistent to a fly. Thus, encoding efficiency inevitably falls at brighter light levels, as information rate in light input increases with intensity (Fig. 2J), whereas that of photoreceptor output approaches a limit (compare Fig. 12F). But for stimuli with longer dark contrasts, including NS and 20 Hz GWN, this limit is reached at brighter intensities than for stimuli that lack these features. Therefore, we reason that encoding efficiency for 20 Hz GWN should have its maximum (upper limit) at $\sim10^6$ photons/s, and this may be as high as 90%–95%, as suggested by our results.

Decorrelated stimulation biases performance comparisons

When GWN bandwidth matches, or is just above, the characteristic microvillar sampling rate, a photoreceptor's information transfer rate approaches its maximum for decorrelated stimuli. However, because photoreceptors of different fly species are specialized for distinct visual lifestyles and habitats, having different microvilli numbers and phototransduction speeds (Gonzalez-

Bellido et al., 2011; Song et al., 2012), their sampling rates are different and adapt differently. Therefore, using the same GWN stimulus to measure photoreceptor performance in slow- and fast-flying flies leads to an unwarranted comparison, where their encoding is deficient but for different reasons. A slow-flying *Drosophila* sees 20 Hz GWN efficiently, whereas a fast-flying *Coenosia* rapidly adapts to its slow modulation more strongly and without using its high-frequency encoding capacity. Contrariwise, in 300 Hz GWN stimulation, *Coenosia* resolves high frequencies well, whereas *Drosophila* responds only to its residual low-frequency modulation. Information capture in every measurement, though, is submaximal and by different amounts.

The photoreceptors would encode more information if their microvilli were utilized better. Decorrelated and most Gaussian stimulation retains a large fraction of microvilli refractory, underutilizing their sampling capacity (Fig. 10D). Natural scenes embed information-rich amplitude and phase correlations of asymmetric (bursty) contrast changes (van Hateren, 1997), representing light objects and prolonged dark contrasts from shadows, dark objects, and edges. These drive microvilli and, thus, crucially, any interneural communication henceforth (van Hateren, 1997; Lewen et al., 2001; Juusola and de Polavieja, 2003; Zheng et al., 2006; Wardill et al., 2012), more efficiently. Thereby, decorrelated stimulation fails to test the information capacity of any neuron that receives its inputs from photoreceptors.

Generality of stochastic adaptive sampling

Refractoriness of sampling may have evolved also in other receptor cells to enhance encoding of real-world stimulus sequences, such as phasic concentration changes in natural odor plumes (Smear et al., 2011). In these scenarios, adapting the sampling dynamics of quantal events (e.g., scent molecules) to changing information is more important than any filtering process afterward. This is because neural integration of samples increases information and reduces noise, whereas filtering responses downstream cannot increase information (data processing theorem).

In conclusion, sensory neurons sample environmental information using specialized membrane elaborations, such as cilia or microvilli, capable of responding to energy quanta or single molecules. A central issue in sensory neuroscience has been to understand this first stage of encoding as it limits how animals sense the world. In this study, we showed, by using experimental and theoretical approaches, how thousands of sampling units (microvilli) in fly photoreceptors gather information from different stimulus statistics. We discovered that a striking adaptation in sampling, stochastic refractoriness of microvilli, improves contrast resolution in daylight conditions. But this came at the expense of sensitivity, affecting neural representations of different stimuli differently. Remarkably, refractoriness also lowered the metabolic cost of information. These findings provided mechanistic reasons why and how the earliest neural code depends upon the statistical context of the stimulation.

References

Angelo K, Rancz EA, Pimentel D, Hundahl C, Hannibal J, Fleischmann A, Pichler B, Margrie TW (2012) A biophysical signature of network affiliation and sensory processing in mitral cells. Nature 488:375–378. CrossRef Medline

Atick JJ (1992) Could information theory provide an ecological theory of sensory processing. Netw Comp Neural 3:213–251. CrossRef Medline

Barlow HB (1961) Possible principles underlying the transformation of sensory messages. In: Sensory communication (Rosenblith WA, ed). Cambridge: Massachusetts Institute of Technology.

Brinkworth RSA, Mah EL, Gray JP, O'Carroll DC (2008) Photoreceptor processing improves salience facilitating small target detection in cluttered scenes. J Vis 8.

Burton BG, Laughlin SB (2003) Neural images of pursuit targets in the photoreceptor arrays of male and female houseflies Musca domestica. J Exp Biol 206:3963–3977. CrossRef Medline

Darwin C (1859) The origin of species by means of natural selection: or, the preservation of favored races in the struggle for life: John Murray.

de Polavieja GG (2004) Reliable biological communication with realistic constraints. Phys Rev E70.

de Polavieja GG, Harsch A, Kleppe I, Robinson HP, Juusola M (2005) Stimulus history reliably shapes action potential waveforms of cortical neurons. J Neurosci 25:5657–5665. CrossRef Medline

Fain GL, Hardie R, Laughlin SB (2010) Phototransduction and the evolution of photoreceptors. Curr Biol 20:R114–R124. CrossRef Medline

Faivre O, Juusola M (2008) Visual coding in locust photoreceptors. PLoS One 3:e2173. CrossRef Medline

Field DJ (1987) Relations between the statistics of natural images and the response properties of cortical cells. J Opt Soc Am A 4:2379–2394. CrossRef Medline

Frolov R, Immonen EV, Vähäsöyrinki M, Weckström M (2012) Postembryonic developmental changes in photoreceptors of the stick insect Carausius morosus enhance the shift to an adult nocturnal life-style. J Neurosci 32:16821–16831. CrossRef Medline

Gonzalez-Bellido PT, Wardill TJ, Juusola M (2011) Compound eyes and retinal information processing in miniature dipteran species match their specific ecological demands. Proc Natl Acad Sci U S A 108:4224–4229. CrossRef Medline

Hardie RC, Franze K (2012) Photomechanical responses in Drosophila photoreceptors. Science 338:260–263. CrossRef Medline

Harris FJ (1978) On the use of the windows for harmonic analysis with the discrete Fourier transform. Proc IEEE 66:51–84. CrossRef

Heimonen K, Salmela I, Kontiokari P, Weckström M (2006) Large functional variability in cockroach photoreceptors: optimization to low light levels. J Neurosci 26:13454–13462. CrossRef Medline

Henderson SR, Reuss H, Hardie RC (2000) Single photon responses in Drosophila photoreceptors and their regulation by Ca^{2+}. J Physiol Lond 524:179–194. CrossRef Medline

Hochstrate P, Hamdorf K (1990) Microvillar components of light adaptation in blowflies. J Gen Physiol 95:891–910. CrossRef Medline

Howard J, Blakeslee B, Laughlin SB (1987) The intracellular pupil mechanism and photoreceptor signal-noise ratios in the fly Lucilia cuprina. Proc R Soc Lond B Biol Sci 231:415–435. CrossRef Medline

Joesch M, Schnell B, Raghu SV, Reiff DF, Borst A (2010) ON and OFF pathways in Drosophila motion vision. Nature 468:300–304. CrossRef Medline

Juusola M (1993) Linear and nonlinear contrast coding in light-adapted blowfly photoreceptors. J Comp Physiol A Neuroethol Sens Neural Behav Physiol 172:511–521. CrossRef

Juusola M, de Polavieja GG (2003) The rate of information transfer of naturalistic stimulation by graded potentials. J Gen Physiol 122:191–206. CrossRef Medline

Juusola M, Hardie RC (2001a) Light adaptation in Drosophila photoreceptors: I. Response dynamics and signaling efficiency at 25 degrees C. J Gen Physiol 117:3–25. CrossRef Medline

Juusola M, Hardie RC (2001b) Light adaptation in Drosphila photoreceptors: II. Rising temperature increases the bandwidth of reliable signaling. J Gen Physiol 117:27–42. CrossRef Medline

Juusola M, Kouvalainen E, Järvilehto M, Weckström M (1994) Contrast gain, signal-to-noise ratio, and linearity in light-adapted blowfly photoreceptors. J Gen Physiol 104:593–621. CrossRef Medline

Juusola M, Song Z, Hardie RC (2014) Biophysics of insect phototransduction. In: Encyclopedia of computational neuroscience. New York: Springer.

Kremkow J, Jin J, Komban SJ, Wang Y, Lashgari R, Li X, Jansen M, Zaidi Q, Alonso JM (2014) Neuronal nonlinearity explains greater visual spatial resolution for darks than lights. Proc Natl Acad Sci U S A 111:3170–3175. CrossRef Medline

Laughlin SB, de Ruyter van Steveninck RR, Anderson JC (1998) The metabolic cost of neural information. Nat Neurosci 1:36–41. CrossRef Medline

Lazar AA (2007) Information representation with an ensemble of Hodgkin-Huxley neurons. Neurocomputing 70:1764–1771. CrossRef

Lewen GD, Bialek W, de Ruyter van Steveninck RR (2001) Neural coding of naturalistic motion stimuli. Network 12:317–329. CrossRef Medline

Lillywhite PG, Laughlin SB (1979) Transducer noise in a photoreceptor. Nature 277:569–572. CrossRef Medline

Marr D, Hildreth E (1980) Theory of edge detection. Proc R Soc Ser B Biol 207:187–217. CrossRef Medline

Minke B (2012) The history of the prolonged depolarizing afterpotential (PDA) and its role in genetic dissection of Drosophila phototransduction. J Neurogenet 26:106–117. CrossRef Medline

Niven JE, Anderson JC, Laughlin SB (2007) Fly photoreceptors demonstrate energy-information trade-offs in neural coding. PLoS Biol 5:828–840. CrossRef Medline

Padmanabhan K, Urban NN (2010) Intrinsic biophysical diversity decorrelates neuronal firing while increasing information content. Nat Neurosci 13:1276–1282. CrossRef Medline

Pérez-Escudero A, Rivera-Alba M, de Polavieja GG (2009) Structure of deviations from optimality in biological systems. Proc Natl Acad Sci U S A 106:20544–20549. CrossRef Medline

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. CrossRef Medline

Rieke F, Bodnar DA, Bialek W (1995) Naturalistic stimuli increase the rate and efficiency of information transmission by primary auditory afferents. Proc Biol Sci 262:259–265. CrossRef Medline

Roebroek JGH, Stavenga DG (1990) Insect pupil mechanisms: 4. Spectral characteristics and light intensity dependence in the blowfly, Calliphora erythrocephala. J Comp Physiol A Neuroethol Sens Neural Behav Physiol 166:537–543.

Sayeed O, Benzer S (1996) Behavioral genetics of thermosensation and hygrosensation in Drosophila. Proc Natl Acad Sci U S A 93:6079–6084. CrossRef Medline

Shannon CE (1948) A mathematical theory of communication. AT&T Tech J 27:379–423.

Smear M, Shusterman R, O'Connor R, Bozza T, Rinberg D (2011) Perception of sniff phase in mouse olfaction. Nature 479:397–400. CrossRef Medline

Song Z, Postma M, Billings SA, Coca D, Hardie RC, Juusola M (2012) Stochastic, adaptive sampling of information by microvilli in fly photoreceptors. Curr Biol 22:1371–1380. CrossRef Medline

Stavenga DG (2004) Angular and spectral sensitivity of fly photoreceptors: III. Dependence on the pupil mechanism in the blowfly Calliphora. J Comp Physiol A Neuroethol Sens Neural Behav Physiol 190:115–129. CrossRef Medline

Theiler J, Eubank S, Longtin A, Galdrikian B, Doyne Farmer J (1992) Testing for nonlinearity in time series: the method of surrogate data. Physica 58:77–94.

Vähäsöyrinki M, Niven JE, Hardie RC, Weckström M, Juusola M (2006) Robustness of neural coding in Drosophila photoreceptors in the absence of slow delayed rectifier K^+ channels. J Neurosci 26:2652–2660. CrossRef Medline

van Hateren JH (1992) A theory of maximizing sensory information. Biol Cybern 68:23–29. CrossRef Medline

van Hateren JH (1997) Processing of natural time series of intensities in the early visual system of the blowfly. Perception 26:6–7. CrossRef Medline

van Hateren JH, Snippe HP (2001) Information theoretical evaluation of parametric models of gain control in blowfly photoreceptor cells. Vision Res 41:1851–1865. CrossRef Medline

van Hateren JH, Snippe HP (2006) Phototransduction in primate cones and blowfly photoreceptors: different mechanisms, different algorithms, similar response. J Comp Physiol A Neuroethol Sens Neural Behav Physiol 192:187–197. CrossRef Medline

Vogt K, Kirschfeld K, Stavenga DG (1982) Spectral effects of the pupil in fly photoreceptors. J Comp Physiol 146:145–152. CrossRef

Wardill TJ, List O, Li X, Dongre S, McCulloch M, Ting CY, O'Kane CJ, Tang S, Lee CH, Hardie RC, Juusola M (2012) Multiple spectral inputs improve motion discrimination in the Drosophila visual system. Science 336:925–931. CrossRef Medline

Wong F, Knight BW, Dodge FA (1982) Adapting bump model for ventral photoreceptors of limulus. J Gen Physiol 79:1089–1113. CrossRef Medline

Zheng L, de Polavieja GG, Wolfram V, Asyali MH, Hardie RC, Juusola M (2006) Feedback network controls photoreceptor output at the layer of first visual synapses in Drosophila. J Gen Physiol 127:495–510. CrossRef Medline

Zheng L, Nikolaev A, Wardill TJ, O'Kane CJ, de Polavieja GG, Juusola M (2009) Network adaptation improves temporal representation of naturalistic stimuli in Drosophila eye: I. Dynamics. PLoS One 4:e4307. CrossRef Medline