

26 Adaptation and Natural Stimulus Statistics

ADRIENNE FAIRHALL

ABSTRACT In trying to understand the structure and dynamics of sensory neural representations, it has been productive to examine the complex structure of natural images to discover the nature of the task that the nervous system performs. Two different forms of adaptation help sensory systems to handle the large variations and multiple spatial and temporal scales of natural inputs. Slow changes in excitability can be understood as an encoding of slowly varying stimulus properties that highlights change, while adaptive changes to neural coding strategies lead to more efficient representations of stimuli drawn from time-varying distributions.

Animals must parse a sea of sensory data to extract information that informs their behavioral interactions with the world. Early sensory systems must provide sufficiently resolved data to extract, produce, and manipulate representations of meaningful structure in the world. The composition of much of the natural world in terms of discrete events and objects at varying distances imposes a strong spatial and temporal structure on sensory stimuli from all modalities. Several properties are common: a multiplicity of spatial and temporal scales, a prevalence of sharp transitions or edges, and complex correlations. It has long been hypothesized that this special structure might be reflected in the organization of sensory neural systems.

Sensory neural coding describes the mapping of a sensory input to its representation in neuronal activity. While the classic notion of a tuning curve implies that this mapping is stationary, several different processes combine to create what can appear to be a dynamic relationship between input and output. These nonstationarities span a wide range of time scales, from neuronal adaptations that are almost instantaneous to the reconfiguration of brain circuitry that underlies long-term memory.

We focus here on the shorter time-scale processes referred to as adaptation. We will discuss how properties of adaptation may be relevant to the precise and efficient coding of inputs with the special structure of natural inputs. The term *adaptation* is used in two broadly defined ways. In the first sense, it describes the transience of many neural responses: the fact that a

neuron or neural system's response to a specific stimulus often reduces over time. While this is often thought of as a decay in the response, it can be interpreted as part of the dynamical coding properties of the neural system. *Adaptation* is also used to describe more general nonstationarities in the neural code: shifts in the form of neural representation over time. These shifts often appear to allow neural systems to represent inputs more efficiently.

Natural scenes and their challenges

Natural stimuli have a distinctive structure that results from the properties of the natural world (Ruderman, 1997; Simoncelli & Olshausen, 2001). These properties pose particular challenges for sensory systems, and also opportunities for selecting efficient strategies for encoding.

LARGE VARIATIONS Natural stimuli (e.g., figure 26.1A–D) exhibit modulations in amplitude over many orders of magnitude in intensity. Audible, naturally occurring sounds span a range from the threshold of hearing at around 10 dB to 120 dB, around 10 orders of magnitude, with little alteration in our ability to detect and identify sound sources (figure 26.1A). Olfactory stimuli are detectable by humans at fractions of a part per million (figure 26.1B), and odorants are also uniquely identified over a very wide range of concentrations. Light level and contrast can vary by 10 orders of magnitude, yet permit accurate viewing and stable object recognition (figure 26.1C). Given that the activity of primary sensory receptors typically increases with the intensity of the input, this tremendous dynamic range poses two questions. First, how can sensory systems represent such a wide range of inputs without saturation, and second, how do they achieve the ability to extract information about object identity that is invariant with intensity?

STRUCTURE AT MANY SCALES Because visual scenes contain structure in the form of objects and

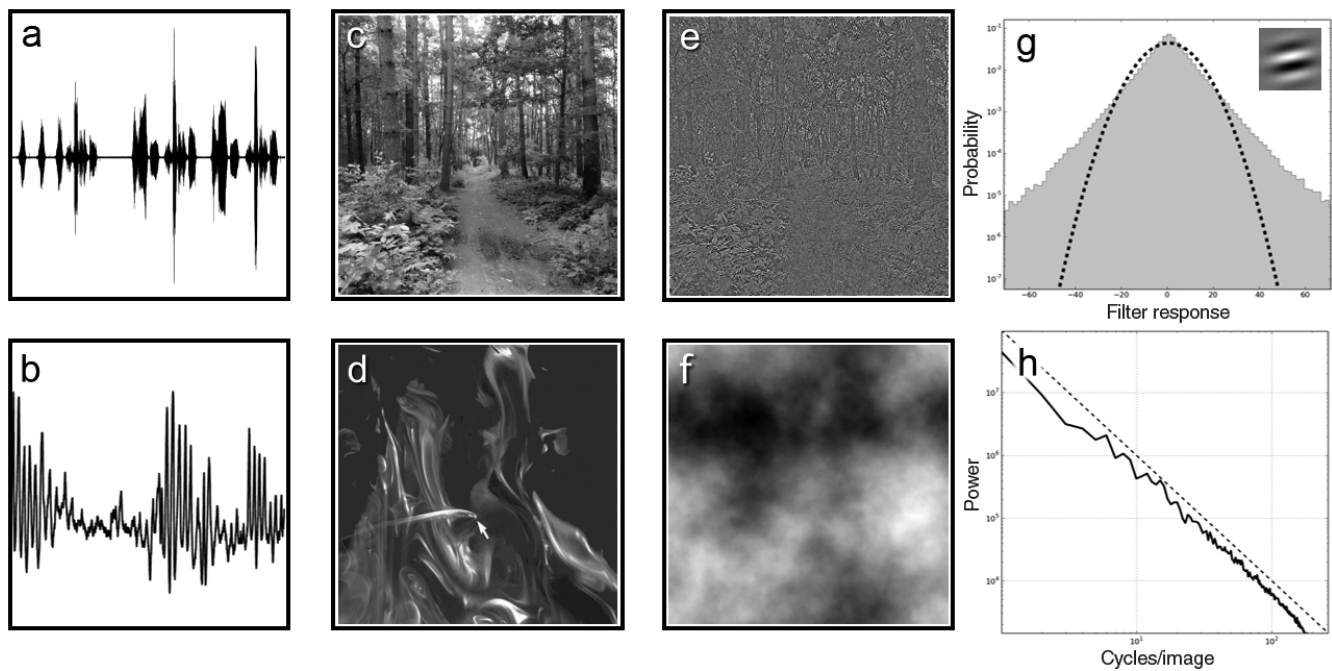


FIGURE 26.1 Examples and properties of natural stimuli. (A) A natural auditory stimulus, the sound pressure wave of a zebra finch's song (M. Fee). (B) The position of a rat's vibrissa during natural whisking (data courtesy of D. Kleinfeld). (C) A natural image (photograph by Ruben Holthuisen). (D) An underwater olfactory plume passing a lobster antenna (photograph by M. Koehl). (E) The image in (C),

whitened. (F) A Gaussian random image with spectral characteristics matching those of natural scenes. (G) The probability density, gray, of the output of a Gabor filter (inset) acting on the image in (C); differs strongly from Gaussian (dotted). (H) Natural images have a close to $1/f^2$ power spectrum. (See color plate 22.)

arrangements that can be viewed from a range of distances and that exist in a wide range of sizes (figure 26.1C), images exhibit self-similarity: as one zooms through different scales, the statistical composition of the scene is relatively invariant (Ruderman & Bialek, 1994; Simoncelli & Olshausen, 2001). For an image with intensity as a function of space given by $I(x, y)$ we can compute the image's autocorrelation function, $C(\delta x, \delta y) = I(x + \delta x, y + \delta y)I(x, y)$, where the average is over all locations x and y . As images are translation-invariant, one can equivalently compute this function in the Fourier domain to obtain the power spectrum. This two-dimensional function can be reduced to a function of spatial frequency f by averaging over phase. For natural images, this function behaves as a power law, $1/f^\alpha$, where the exponent α is very close to 2 (figure 26.1H; Ruderman & Bialek, 1994; Tolhurst, Tadmor, & Chao, 1992). The power-law form of the power spectrum is likely a result of the scale invariance of natural scenes; the particular exponent of 2 may be a result of the prevalence of edges, as edges themselves have a power spectrum of $1/f^2$. This property is not unique to vision: under ideal conditions, odorants that are advected by a turbulent airflow also become

self-similar and exhibit a power-law spectrum (Catrakis & Dimotakis, 1996).

Not all of the structure in a natural scene is captured by the power spectrum: one can easily construct a random image with a $1/f^2$ spectrum that has no object-like properties (figure 26.1F). The existence of higher-order correlations is apparent when one examines the distribution of two-point differences. If images were fully explained by the mean and power spectrum, these distributions would be Gaussian; however, they are typically strongly peaked at zero and have long tails (figure 26.1G), showing that higher-order moments, or equivalently, multipoint correlations, are needed to fully describe such images.

Despite the overall self-similarity, natural inputs often have a separation of meaning between large-scale and small-scale features (Attias & Schreiner, 1998; Joris, Schreiner, & Rees, 2004). Large-scale features might be termed *objects*, whereas small-scale features may be termed *textures*; features on these different scales convey different types of information. Here, "scale" can refer to either time or space. For example, the large-scale features that identify object boundaries are filled with finer-scale textural structure that conveys distinct

information both in visual images and in the temporal signals created by tactile exploration (Hill, Curtis, Moore, & Kleinfeld, 2011). Speech, the auditory input with which humans are most concerned, consists of a large-scale amplitude modulation or envelope that carries most of the information about word meaning, while fine textural details of frequency composition or pitch convey information about the speaker, nuance, and emotion. This general separation between large- and small-scale features suggests that different coding schemes might be used at different time or length scales.

Given these general properties, how might neural encoding be optimized for natural scenes? One hypothesis is that the coding properties of neurons should be organized to represent natural stimuli as efficiently as possible. *Efficient coding* means that a neuron's response properties are adjusted to the distribution of possible stimuli in such a way as to maximize the neuron's representational capacity (Attneave, 1954; Barlow, 1961). This idea has been generalized to neuronal populations through the notion of *sparse coding*, which imposes an additional constraint on the number of active neurons (Barlow, 2001; Bell & Sejnowski, 1997; Olshausen & Field, 1996a, 1996b, 1997, 2004). These approaches have led to a qualitative understanding of the form of receptive fields in a number of sensory systems. However, in the course of natural behavior and through transitions in environments and natural diurnal cycles, the statistics of stimuli continuously vary in space and time. An extension of the efficient coding hypothesis suggests that neural systems should dynamically adjust their coding strategies to best represent these time-varying stimulus statistics.

Adaptation as temporal decay

The term *adaptation* is often used to describe the tendency of neural responses to decay over time. When a stimulus is presented after no input, neuronal responses often initially rapidly increase, then gradually decrease to a lower steady state (Adrian, 1928). This ubiquitous effect occurs through a variety of mechanisms. In the periphery, sensory transduction often occurs through second messenger cascades that contain several opportunities for feedback that diminishes response to constant stimuli over time (Soo, Detwiler, & Rieke, 2008); further, receptors tend to inactivate (Sato et al., 2008). In spiking neurons, this effect is known as spike-frequency adaptation and is due to a range of ion channels, including a noninactivating K^+ current known as the M current (Brown & Adams, 1980); the medium afterhyperpolarization current, mAHP, due to

Ca^{2+} -dependent K^+ channels (Adelman, Maylie, & Sah, 2012); Na^+ -dependent K^+ channels (Schwindt, Spain, & Crill, 1989); and slow inactivation of sodium channels (Marom, 1998). Beyond single-neuron properties, short-term synaptic plasticity (Zucker & Regehr, 2002) can also reduce input strength in an activity-dependent way (Castro-Alamancos & Connors, 1996).

How one interprets the coding role of such adaptation depends on the method used to characterize the neural code. If one presents a sequence of inputs I that are fixed for some period of time, then change, one can construct an input-output curve derived from the responses immediately following each change in input, $f_0(I)$, and a different input-output curve from the steady state part of the response, $f_\infty(I)$ (Fernandez & White, 2010; Higgs, Slee, & Spain, 2006). This leads to a characterization of schemes of neural coding as either instantaneous or steady-state relative to an onset time. Benda and Herz (2003) developed a general description of adapting dynamics whereby the firing rate f relaxes between these two coding regimes according to the dynamics of an adaptation variable A :

$$f = f_0(I - A) \quad (1)$$

$$\tau \frac{dA}{dt} = f_\infty^{-1}(f) - f_0^{-1}(f) - A. \quad (2)$$

In this model, the adaptation variable A is driven by the neuronal activity f and relaxes with a fixed time constant of τ . When the two f - I curves are linear, with slopes f'_0 and f'_∞ , these dynamics correspond to a high-pass filter imposing gain $G(f_c)$ on the input component at f_c ,

$$G(f_c) = f'_\infty \sqrt{\frac{1 + (2\pi f_c \tau_{eff} f'_0 / f'_\infty)^2}{1 + (2\pi f_c \tau_{eff})^2}}. \quad (3)$$

The time constant τ_{eff} is given by $\tau f'_\infty / f'_0$ and corresponds to the measured time course of decay to a step of input. In this model, Eq. (1) captures the rapid onset of the response seen experimentally, and Eq. (2) the slow relaxation. This model predicts the behavior of electroreceptors in the electric fish, which respond with high gain to chirps embedded in a more slowly varying amplitude modulation signal (Benda, Longtin, & Maler, 2005).

The continuous variation of natural stimuli suggests an alternative perspective: that the dynamics of adaptation is an essential part of the encoding of the stimulus. The transient responses are a strong response to a *change* in stimulus: that is, the firing rate can be considered as a combined function of the new stimulus value and the derivative of the stimulus. As the response $R(t)$ to a stimulus stream s at any time t depends on the

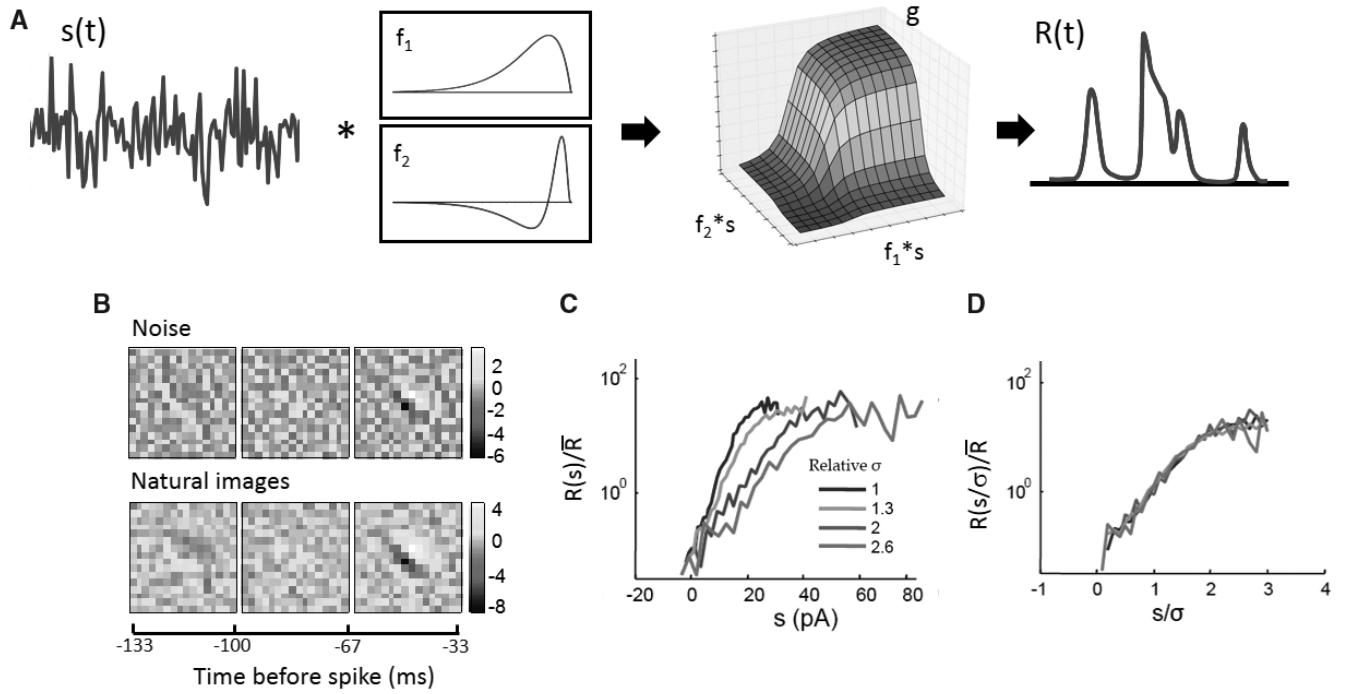


FIGURE 26.2 Linear-nonlinear models and adaptation. (A) Multidimensional linear-nonlinear cascade model: the stimulus $s(t)$ is filtered through one or more filters f_i and passed through a nonlinear function $g(f_1 * s, f_2 * s, \dots)$. Here, f_1 is a leaky integrating filter while f_2 is closer to a derivative. The filters f_i and g can both change with adaptation to stimulus context. (B) Example of filter adaptation: spatiotemporal filters from V1 computed using the maximally informative dimension method (Sharpee et al., 2004) when driven either

by a flickering noise input or a natural image ensemble. The filters show changes in low-frequency content (adapted from Sharpee et al., 2006). (C) The nonlinearity g can also adapt. Input-output curves from a cortical neuron driven by a filtered white noise current input (R. Mease). The stimulus s is the current filtered through the spike-triggered average. (D) The same input-output curves plotted against the stimulus normalized by the standard deviation of the current. (See color plate 23.)

recent past of the stimulus (for example, the fact that it just underwent a step, or that it has been constant for some time), one can approximate it as some function f of the recent history of the stimulus:

$$R(t) = f(s(t' \leq t)).$$

In the simplest case, this response might be a linear function:

$$R(t) = f * s = \int_0^\infty d\tau f(\tau) s(t - \tau), \quad (4)$$

where here, $f(\tau)$ is a set of coefficients that weights the stimulus at times in the recent past, denoted by τ . Equivalently, $f(\tau)$ is a *filter* acting on the stimulus. One can think about $f(\tau)$ as defining a feature of the input to which the system is sensitive, and the operation performed in Eq. (4) is a projection of the stimulus onto that feature. The nature of $f(\tau)$ reveals the dependence of the output on variations in $s(t)$: for example, the response might simply be an average of the recent stimulus, in which case $f(\tau)$ will have a form like f_1 in figure

26.2A. If, as is often observed, the output activity responds transiently to changes in the stimulus, $f(\tau)$ is likely to have a differentiating nature, as in f_2 in figure 26.2A. Generally, one might expect the response to have a sensitivity to both the recent mean stimulus and its derivative, so that $f(\tau)$ might be a composition of an averaging filter and a differentiating filter. If the response is truly approximately linear, these components will simply add. It is likely, however, that the response is a nonlinear function of this weighted sum:

$$R(t) = g\left(\int_0^\infty d\tau f(\tau) s(t - \tau)\right), \quad (5)$$

where $g(\cdot)$ is a nonlinear function that relates the firing rate output to the stimulus projected onto the feature $f(\tau)$. In the case that two or more features contribute to the response nonlinearly, the response could then be approximated as

$$R(t) = g\left(\int_0^\infty d\tau f_1(\tau) s(t - \tau), \int_0^\infty d\tau f_2(\tau) s(t - \tau), \dots\right), \quad (6)$$

where now $s_1 = \int_0^\infty d\tau f_1(\tau) s(t - \tau)$ and $s_2 = \int_0^\infty d\tau f_2(\tau) s(t - \tau)$ are the components of the entire temporal history of s , projected onto features f_1 and f_2 and $g(\cdot, \cdot)$ is a two-dimensional function of these two variables (figure 26.2A). The features to which a given neuron is sensitive, and the set of features to which a population of neurons is sensitive, define a basis set that captures the components of an input to which the neural population responds (Schwartz, Pillow, Rust, & Simoncelli, 2006).

In this formulation, the initial rapid increase followed by an exponential decay of the firing rate of the neuron in response to a constant stimulus corresponds to convolving the input with a high-pass filter like that defined in Eq. (3). It is important to realize that the decaying dynamics observed in $R(t)$ can be usefully interpreted not as a fatigue of the response, but as part of a well-defined coding property of the system in which the neuron's response is not an instantaneous input-output function but incorporates history dependence and sensitivity to derivatives (Lundstrom, Higgs, Spain, & Fairhall, 2008; Nagel & Wilson, 2011).

Power-law adaptation

While modeling spike-frequency adaptation with a single exponential has been successful in several applications (Benda & Herz, 2003), the decay observed in a variety of different neurons exhibits multiple temporal components (French & Torkkeli, 2008; La Camera et al., 2006; Schwindt, Spain, Foehring, Chubb, & Crill, 1988; Thorson & Biederman-Thorson, 1974). The response following a stimulus step can be approximately power-law:

$$R(t) = R_0 t^{-\alpha}$$

A response function that can account for both the derivative-like rapid onset to a step and also the slow power-law decay can be expressed in the frequency domain as a transfer function of the form $R(\omega) = (i\omega)^\alpha$. If the exponent α is 1, this operation is differentiation: the system responds to the step with a pulse that decays immediately. When α is less than 1, this is known as fractional differentiation (Oldham & Spanier, 2006), and can be conceived of as a mixture of integration and differentiation. Such a system also responds with a sharp increase to a step, but following the step, the response exhibits a slow decay. Because of the power-law nature of the response, if such a system is driven by a switching stimulus that changes amplitude with some period T (Baccus & Meister, 2002; Kvale & Schreiner, 2004; Maravall et al., 2007; Smirnakis, Berry, Warland, Bialek,

& Meister, 1997), the time scale of the response appears to scale with T (Fairhall, Lewen, Bialek, & de Ruyter van Steveninck, 2001b; Lundstrom et al., 2008; Wark, Fairhall, & Rieke, 2009; Wark, Lundstrom, & Fairhall, 2007). This response function scales each input frequency component's amplitude by a power-law function of frequency ω^α and rotates each component in the complex plane by a constant amount, $i\alpha = e^{i\pi\alpha/2}$. This rotation corresponds to a frequency-independent phase shift of $\phi = \pi\alpha/2$. These dynamics are approximately realized by several sensory systems (Anastasio, 1994; Fairhall et al., 2001b; Lundstrom et al., 2008; Lundstrom, Fairhall, & Maravall, 2010); however, the constant phase property is not always observed along with power-law-like decays (Drew & Abbott, 2006; Pozzorini, Naud, Mensi, & Gerstner, 2013). Note that this response is a linear function of the time-varying input amplitude. For a linear function, the response simply scales with the size of the input; more biophysical (Nagel & Wilson, 2011; Ozuysal & Baccus, 2012) or normative (Wark et al., 2009) models have been evoked to account for nonlinear effects such as a change in the time constant of the response with a change in input amplitude.

What might power-law adaptation accomplish in terms of signal processing? Since many natural inputs have a power-law spectrum, it is possible that this adaptive transformation has the effect of whitening such inputs: following this transformation, different frequencies in the input are represented more uniformly, and thus, more efficiently, by the firing rate output. Some direct evidence for this was found by Pozzorini et al. (2013), who showed that the spectrum of the input to layer V cortical pyramidal neurons *in vivo* is indeed matched to the power-law filtering induced by multiple adaptation time scales in spike-frequency adaptation in these neurons, such that natural current inputs are whitened.

There are several possible biophysical mechanisms that might underlie power-law-like responses at the single-neuron and network level, including multiple time scales in Na^+ channel inactivation (Gilboa, Chen, & Brenner, 2005; Toib, Lyakhov, & Marom, 1998), spike-frequency adaptation currents (Lundstrom et al., 2008), and synaptic dynamics (Drew & Abbott, 2006; Puccini, Sanchez-Vives, & Compte, 2007). In principle, one or more of these mechanisms could work together to increase the power-law exponent (Lundstrom et al., 2010).

Adaptive coding with changing stimulus statistics

So far, we have considered how adaptive dynamics themselves constitute a channel for encoding stimulus

variations in the time domain. We will now use the descriptions of responses given in Eqs. (4–6) to consider a general approach to capturing the effects of different forms of adaptation on encoding. Eqs. (5) and (6) express the concept that a neuron's response can be approximated as a nonlinear function acting on one or more linear filters applied to the input. This is known as a linear-nonlinear cascade model (Schwartz et al., 2006). The filter can be generalized to act on arbitrary components of the stimulus: spatial, temporal, spectral, or other representations (figure 26.2). The two components of such a coding model, the linear filtering of the system and its nonlinear input-output mapping, may be particularly adapted to the statistics of the inputs. Furthermore, these components may be able to change in time to track nonstationary inputs.

OPTIMAL FILTERS First, the linear filter stage of the response can cancel out the correlations in the input so that responses represent only “surprising” components of the stimulus. This is known as *predictive coding*. Based on the spatial autocorrelation function of natural scenes, Srinivasan, Laughlin, and Dubs (1982) computed the amount of inhibitory surround required to cancel out these correlations and found that this prediction matched the filtering properties of first-order cells in the fly eye. Atick and Redlich (1990; Atick, 1992) used information theory along with a model of photoreceptor noise to derive filters that both optimally reduce the noise and whiten the power spectrum in natural scenes. The predicted filters qualitatively match the response characteristics of retinal ganglion cells and behavioral contrast sensitivity. Predictive coding can remove not only spatial correlations, but also temporal ones (Dong & Atick, 1995; van Hateren, 1992a, 1992b). For example, predictive coding has been invoked to explain adaptation to stimulus input in the encoding of electrical signals in cerebellum in electric fish and in dorsal root ganglia (Roberts & Portfors, 2008).

While visual processing in the retina and lateral geniculate nucleus (Dan, Atick, & Reid, 1996; Dong & Atick, 1995) may serve to whiten images and image streams, figure 26.1 shows that significant correlational structure still remains. Olshausen and Field (1996a) used natural image patches to learn a set of spatial patterns that best represent the images, but such that only a small number is needed to reconstruct any image. Constraining the number required for representing natural images makes this a “sparse” representation: if the spatial-filtering properties of neurons in cortical area V1 corresponded to these patterns, any given stimulus would activate as few neurons in the population as

possible. The filters that result from this procedure strongly resemble oriented Gabor-like receptive fields with a range of spatial scales, as are characteristic of V1 simple cells (Bell & Sejnowski, 1997). Similarly, Smith and Lewicki (2006) were able to learn filters with the same form and population distribution of filters of cochlear auditory neurons from natural sound ensembles using this sparseness constraint. Together, these studies suggest that neurons encode stimuli in such a way as to minimize correlations across neurons, and across time, with the additional condition that the population response is sparse (Olshausen & Field, 2004).

OPTIMAL INPUT-OUTPUT FUNCTIONS The second part of the standard coding model of figure 26.2 concerns the nonlinear mapping from filtered input to output, the function $g(\cdot)$ in Eq. (5). Information-theoretic concepts can also be applied to the form of this function. If the range of the output is limited, one should use all available output symbols (e.g., firing rates) as often as possible; under certain assumptions, this will maximize the entropy of the output and thus maximize information transmission through the coding channel (Barlow, 1961; Laughlin, 1981). In order for this to happen, the input-output relation of the system must depend on the probability distribution of the inputs; with a fixed maximum output, the input-output curve should be the cumulative integral of the input distribution. Laughlin (1981) measured the distribution of contrasts in natural scenes and showed that the input-output relations of fly large monopolar cells indeed matched this predicted form.

These results indicate that the structure of some neural codes can be understood in terms of optimizing response functions relative to the statistics of natural stimuli. However, as we have emphasized, a typical property of such inputs is their local variability in space and time. One might therefore expect that the relationship between input and output changes in time as stimulus statistics change. Indeed, many sensory neural codes display the ability to adapt to the current context of the stimulus.

A dependence of the code on statistical context appears in the calculations of Atick and Redlich (1990; Atick, 1992), which predict that the receptive fields of ganglion cells will be low-pass at low light levels, or low signal-to-noise ratio, and more band-pass as the signal-to-noise ratio increases. This tendency of neural filters to shift toward higher frequencies with higher signal-to-noise as the amplitude of the stimulus increases is typical in many systems, even in single neurons (Hong, Aguera y Arcas, & Fairhall, 2007). Filter properties can also adapt to more complex statistical properties of the

input. In the retina, when stimulated with random inputs that contain spatiotemporal correlations, ganglion cell filters adjust over minutes to cancel out these correlations (Hosoya, Baccus, & Meister, 2005). These changes were proposed to be due to anti-Hebbian synaptic plasticity reweighting interneuron contributions to the receptive field. Similarly, slow adaptation of the filter characteristics of cortical V1 neurons occurs when stimuli are switched from Gaussian white noise to a natural image ensemble (Sharpee et al., 2006; figure 26.2B).

Along with changing filter properties, the nonlinear input-output function may also vary dynamically with changes in stimulus statistics. A method of maintaining high information transmission through a coding channel is to continuously match response range to a changing range of inputs (Brenner, Bialek, & de Ruyter van Steveninck, 2000). Suppose that the response function appropriately spans the dynamic range of a stimulus ensemble. If this input ensemble has a mean or variance that changes in time, the input-output curve should shift its center along the stimulus axis as the mean changes, and alter its slope to accommodate changes in the variance (Enroth-Cugell & Shapley, 1973; Shapley & Victor, 1979). In the fly visual system, the input-output curve adapts to stimuli with changing variance such that the stimulus appears to be encoded in units of its time-varying standard deviation (Brenner et al., 2000; Fairhall et al., 2001a). This rescaling of the input with standard deviation in neural response functions has been shown to occur, with different levels of precision, in visual (Baccus & Meister, 2002; Kim & Rieke, 2001), auditory (Kvale & Schreiner, 2004; Nagel & Doupe, 2006), and somatosensory neurons (Diaz-Quesada & Maravall, 2008; Maravall, Petersen, Fairhall, Arabzadeh, & Diamond, 2007), and even in developing cortical neurons responding to an injected input current (Mease, Famulare, Gjorgjieva, Moody, & Fairhall, 2013; figure 26.2C and D).

Multiple time scales of information representation

We have described two apparently distinct notions of adaptation: the decay processes that tend to encode large-scale stimulus variation with transient dynamics, and the ability of a coding scheme to adjust to the local statistics of a stimulus. While some systems appear to explicitly separate representations at different time scales into different neuronal populations, for example, in electrosensation in response to amplitude modulation versus the modulation envelope (Stamper, Fortune, & Chacron, 2013), in many cases, these different components are temporally multiplexed in the same

neurons (Baccus & Meister, 2002; Fairhall et al., 2001a; Fairhall et al., 2001b; Nagel & Wilson, 2011; Panzeri, Brunel, Logothetis, & Kayser, 2010; Wen, Wang, Dean, & Delgutte, 2012).

Recall the earlier separation of natural stimulus scales into “textures” and “objects.” For a time-varying stimulus amplitude, we can approximately decompose the stimulus as $s(t) = \sigma(t)\eta(t)$, where $\sigma(t)$ is the envelope and $\eta(t)$ is a fast-varying fluctuation. The slower-adapting dynamics encode the stimulus envelope, while an adapting coding scheme permits responses to stimulus details that maintain high fidelity in the face of changing stimulus statistics such as overall amplitude or correlation structure. When the input-output curve adapts to the scale σ of the stimulus, the response might be approximated as

$$R[s(t' \leq t)] = F[\sigma](t)g((f * \eta)(t)).$$

Here the rate factor $F[\sigma]$, a functional of the time-varying envelope, multiplies a linear-nonlinear model, Eq. (5), so that the fast fluctuations modulate the firing rate on short time scales and the stimulus is treated in units normalized to the stimulus standard deviation, while the overall gain depends on the slower-varying stimulus statistics. As the mean firing rate is a function of the amplitude—although potentially a nontrivial one—the inherent ambiguity of a normalized code is removed: information about $\sigma(t)$ is available, although precise spike timing encodes stimulus fluctuations in a variance-independent way (Fairhall et al., 2001a).

Thus, adaptive processes can play two different roles in information processing. The temporal envelope dynamics contribute to the scaling of the output, called “output gain” (Carandini & Heeger, 2012), while the normalization of the stimulus is an example of “input gain.” The coexistence of fast input adaptation with slow relaxation of the firing rate has been observed in the fly visual system (Fairhall et al., 2001a), the retina (Baccus & Meister, 2002), and the somatosensory system (Lundstrom et al., 2010; Maravall et al., 2007); in the auditory system, these time scales may be similar (Wen et al., 2012). At the single-neuron level, the channel dynamics that drive spiking define the feature that triggers spikes (Hong et al., 2007) and govern encoding of stimulus components at fast time scales through precise spike timing while the slower, activity-dependent time scales of adaptation encode slower stimulus variations in the time-averaged rate or membrane dynamics (Lundstrom et al., 2008; Nagel & Wilson, 2011). These processes are not necessarily independent: a biophysical model that incorporates history-dependent changes

in kinetics accounts well for both aspects of multiple time scale adaptation in the retina (Ozuysal & Baccus, 2012).

Predicting time scales of adaptation

Many of the experiments previously discussed imply that neural systems incorporate in their response properties an implicit model of time-varying stimulus statistics. Some of these coding changes occur almost instantaneously (Baccus & Meister, 2002; Fairhall et al., 2001a), suggesting that no explicit learning process is occurring and that such adaptive behavior can arise as a result of intrinsic neuronal nonlinearity, both with respect to the stimulus filter (Famulare & Fairhall, 2009) and the input-output function (Paninski, Lau, & Reyes, 2003; Rudd & Brown, 1997). This is consistent with experimental and modeling observations that the nonlinearities of ion-channel dynamics in single neurons contribute to this behavior (Kim & Rieke, 2003; Mease et al., 2013; Yu & Lee, 2003). Other changes, such as alterations in filtering properties in response to changes in spatiotemporal correlations in the stimulus, occur slowly, and without a fixed time scale that is easily mapped on to an intrinsic time constant (Hosoya et al., 2005; Smirnakis et al., 1997; Sharpee et al., 2006).

What sets the time scale for adaptation? While adaptation has clear benefits in reshaping coding strategies toward higher-fidelity representations as stimulus statistics change, there are also costs: adapting too rapidly can amplify input noise (Dunn & Rieke, 2006). Furthermore, there are statistical limits to how rapidly a system can track time-varying stimulus statistics. DeWeese and Zador (1998) considered a model of adaptation as a Bayesian update process whereby incoming samples are used to update a probability distribution over possible stimulus parameters (mean or variance). This approach predicts two experimental observations. First, it predicts that systems should adapt more rapidly to mean than to variance. Second, there should be an asymmetry between adapting to a higher variance than in adapting to a lower one: a sample from the lower-variance ensemble has a high likelihood of being drawn from the higher-variance ensemble, while the higher variance sample is likely to be an outlier in the lower-variance ensemble. This is indeed observed in the fly visual system (Snippe, Poot, & van Hateren, 2004). In this particular formulation, the time scales of adaptation were dominated by the internal model for the typical rate of change of the tracked parameter. By assuming no inherent time scale for this process, Wark et al. (2009) constructed a Bayesian framework that could

account for the observation that the time scale of adaptation in retinal ganglion cell inputs depends on stimulus length: a long period of observation of samples from a particular variance ensemble strengthens one's belief about the current stimulus parameter and slows the rate of adaptation to a new state. The idea of inference has also been applied to the problem of estimating a neuron's membrane potential through observation of spikes (Pfister, Dayan, & Lengyel, 2010); short-term synaptic plasticity emerges as a consequence of optimal estimation. These examples show that optimal statistical inference can serve as a normative model to qualitatively explain the temporal dynamics of adaptation (Lochmann & Deneve, 2011; Lochmann, Ernst, & Deneve, 2012), although these approaches do not yet provide a prediction of absolute time scale on theoretical grounds.

Building neural models from natural inputs

Due to the nonlinearities of neural systems, standard methods for deriving models of neural coding from data—reverse correlation (Rieke, Warland, de Ruyter van Steveninck, & Bialek, 1997; Simoncelli, Paninski, Pillow, & Schwartz, 2004) and generalized linear models (Truccolo, Eden, Fellows, Donoghue, & Brown, 2005)—lead to an inherent dependence of the resulting model on the input statistics. Furthermore, those methods are guaranteed to be correct only for Gaussian input distributions (Paninski, 2003). However, while Gaussian white noise is convenient in that it samples responses to multiple frequencies in an unbiased way (Rieke et al., 1997; Simoncelli et al., 2004), in many cases it does not drive the brain in a way that resembles sensory function during normal behavior (Quiroga, Reddy, Kreiman, Koch, & Fried, 2005; Theunissen, Sen, & Doupe, 2000). Thus, it is necessary to develop methods to obtain reliable and unbiased estimates of coding properties that generalize across different stimulus ensembles. The method of maximally informative dimensions (Sharpee, Rust, & Bialek, 2004) searches for filter directions in stimulus space that maximize mutual information between the filtered stimulus and spikes; it does not depend on the use of a Gaussian input. In general, models that have been sampled with a given white or colored noise ensemble notoriously do not do well in predicting responses to natural scenes, presumably due to rapid adaptation to changing stimulus statistics (Rieke & Rudd, 2009) as well as the effects of nonlinearities prior to linear filtering (Schwartz & Rieke, 2011). While dynamical models can reproduce some of the characteristics of adapting systems (Ozuysal & Baccus, 2012), future coding models should aim to

capture a neural system's information representation properties while incorporating these adaptive properties.

ACKNOWLEDGMENTS Thanks to Rachel Wilson, Julijana Gjorgjieva, and Blaise Agüera y Arcas for very helpful comments and assistance in preparation of the manuscript. This work was supported by NSF grant 0928251.

REFERENCES

- ADELMAN, J. P., MAYLIE, J., & SAH, P. (2012). Small-conductance Ca²⁺-activated K⁺ channels: Form and function. *Annu Rev Physiol*, 74, 245–269.
- ADRIAN, E. D. (1928). *The basis of sensation: The action of the sense organs*. New York, NY: Norton.
- ANASTASIO, T. J. (1994). The fractional-order dynamics of brainstem vestibulo-oculomotor neurons. *Biol Cybern*, 72(1), 69–79.
- ATICK, J. J. (1992). Could information theory provide an ecological theory of sensory processing? *Network*, 3, 213–251.
- ATICK, J. J., & REDLICH, A. N. (1990). Towards a theory of early visual processing. *Neural Comput*, 2, 308–320.
- ATTIAS, H., & SCHREINER, C. E. (1998). Blind source separation and deconvolution: The dynamic component analysis algorithm. *Neural Comput*, 10(6), 1373–1424.
- ATTNEAVE, F. (1954). Some informational aspects of visual perception. *Psychol Rev*, 61(3), 183–193.
- BACCUS, S. A., & MEISTER, M. (2002). Fast and slow contrast adaptation in retinal circuitry. *Neuron*, 36(5), 909–919.
- BARLOW, H. (2001). Redundancy reduction revisited. *Network*, 12(3), 241–253.
- BARLOW, H. B. (1961). Possible principles underlying the transformation of sensory messages. In W. Rosenblith (Ed.), *Sensory communication* (pp. 217–234). Cambridge, MA: MIT Press.
- BELL, A. J., & SEJNOWSKI, T. J. (1997). The independent components of natural scenes are edge filters. *Vis Res*, 37(23), 3327–3338.
- BENDA, J., & HERZ, A. V. (2003). A universal model for spike-frequency adaptation. *Neural Comput*, 15(11), 2523–2564.
- BENDA, J., LONGTIN, A., & MALER, L. (2005). Spike-frequency adaptation separates transient communication signals from background oscillations. *J Neurosci*, 25(9), 2312–2321.
- BRENNER, N., BIALEK, W., & DE RUYTER VAN STEVENINCK, R. (2000). Adaptive rescaling maximizes information transmission. *Neuron*, 26(3), 695–702.
- BROWN, D. A., & ADAMS, P. R. (1980). Muscarinic suppression of a novel voltage-sensitive K⁺ current in a vertebrate neurone. *Nature*, 283(5748), 673–676.
- CARANDINI, M., & HEEGER, D. J. (2012). Normalization as a canonical neural computation. *Nat Rev Neurosci*, 13(1), 51–62.
- CASTRO-ALAMANCOS, M. A., & CONNORS, B. W. (1996). Short-term plasticity of a thalamocortical pathway dynamically modulated by behavioral state. *Science*, 272, 274–277.
- CATRAKIS, H. J., & DIMOTAKIS, P. E. (1996). Scale distributions and fractal dimensions in turbulence. *Phys Rev Lett*, 77(18), 3795–3798.
- DAN, Y., ATICK, J. J., & REID, R. C. (1996). Efficient coding of natural scenes in the lateral geniculate nucleus: Experimental test of a computational theory. *J Neurosci*, 16(10), 3351–3362.
- DEWEES, M., & ZADOR, A. (1998). Asymmetric dynamics in optimal variance adaptation. *Neural Comput*, 10, 1179–1202.
- DIAZ-QUESADA, M., & MARAVALL, M. (2008). Intrinsic mechanisms for adaptive gain rescaling in barrel cortex. *J Neurosci*, 28(3), 696–710.
- DONG, D. W., & ATICK, J. J. (1995). Temporal decorrelation: A theory of lagged and nonlagged responses in the lateral geniculate nucleus. *Network: Comput Neural Syst*, 6, 159–178.
- DREW, P. J., & ABBOTT, L. F. (2006). Models and properties of power-law adaptation in neural systems. *J Neurophysiol*, 96(2), 826–833.
- DUNN, F. A., & RIEKE, F. (2006). The impact of photoreceptor noise on retinal gain controls. *Curr Opin Neurobiol*, 16(4), 363–370.
- ENROTH-CUGELL, C., & SHAPLEY, R. M. (1973). Adaptation and dynamics of cat retinal ganglion cells. *J Physiol*, 233(2), 271–309.
- FAIRHALL, A. L., LEWEN, G. D., BIALEK, W., & DE RUYTER VAN STEVENINCK, R. R. (2001a). Efficiency and ambiguity in an adaptive neural code. *Nature*, 412(6849), 787–792.
- FAIRHALL, A. L., LEWEN, G. D., BIALEK, W., & DE RUYTER VAN STEVENINCK, R. (2001b). Multiple time scales of adaptation in a neural code. *Advances in Neural Information Processing Systems, Vol. 13* (pp. 124–130).
- FAMULARE, M., & FAIRHALL, A. L. (2009). Adaptation in simple neurons: Dependence of feature selectivity on stimulus statistics. *Neural Comput*, 22(3), 581–598.
- FERNANDEZ, F. R., & WHITE, J. A. (2010). Gain control in CA1 pyramidal cells using changes in somatic conductance. *J Neurosci*, 30(1), 230–241.
- FRENCH, A. S., & TORKKELI, P. H. (2008). The power law of sensory adaptation: Simulation by a model of excitability in spider mechanoreceptor neurons. *Ann Biomed Eng*, 36(1), 153–161.
- GILBOA, G., CHEN, R., & BRENNER, N. (2005). History-dependent multiple-time-scale dynamics in a single-neuron model. *J Neurosci*, 25, 6479–6489.
- HIGGS, M. H., SLEE, S. J., & SPAIN, W. J. (2006). Diversity of gain modulation by noise in neocortical neurons: Regulation by the slow afterhyperpolarization conductance. *J Neurosci*, 26, 8787–8799.
- HILL, D. N., CURTIS, J. C., MOORE, J. D., & KLEINFELD, D. (2011). Primary motor cortex reports efferent control of vibrissa motion on multiple time scales. *Neuron*, 72(2), 344–356.
- HONG, S., AGÜERA Y ARCAS, B., & FAIRHALL, A. L. (2007). Single neuron computation: From dynamical system to feature detector. *Neural Comput*, 19(12), 3133–3172.
- HOSOYA, T., BACCUS, S. A., & MEISTER, M. (2005). Dynamic predictive coding by the retina. *Nature*, 436, 71–77.
- JORIS, P. X., SCHREINER, C. E., & REES, A. (2004). Neural processing of amplitude-modulated sounds. *Physiol Rev*, 84(2), 541–577.
- KIM, K. J., & RIEKE, F. (2001). Temporal contrast adaptation in the input and output signals of salamander retinal ganglion cells. *J Neurosci*, 21, 287–299.
- KIM, K. J., & RIEKE, F. (2003). Slow Na⁺ inactivation and variance adaptation in salamander retinal ganglion cells. *J Neurosci*, 23(4), 1506–1516.

- KVALE, M. N., & SCHREINER, C. E. (2004). Short-term adaptation of auditory receptive fields to dynamic stimuli. *J Neurophysiol*, 91(2), 604–612.
- LA CAMERA, G., RAUCH, A., THURBON, D., LUSCHER, H. R., SENN, W., & FUSI, S. (2006). Multiple time scales of temporal response in pyramidal and fast spiking cortical neurons. *J Neurophysiol*, 96, 3448–3464.
- LAUGHLIN, S. B. (1981). A simple coding procedure enhances a neuron's information capacity. *Z Naturforsch C*, 36, 910–912.
- LOCHMANN, T., & DENEVE, S. (2011). Neural processing as causal inference. *Curr Opin Neurobiol*, 21(5), 774–781.
- LOCHMANN, T., ERNST, U. A., & DENEVE, S. (2012). Perceptual inference predicts contextual modulations of sensory responses. *J Neurosci*, 32(12), 4179–4195.
- LUNDSTROM, B. N., FAIRHALL, A. L., & MARAVALL, M. (2010). Multiple time scale encoding of slowly varying whisker stimulus envelope in cortical and thalamic neurons in vivo. *J Neurosci*, 30(14), 5071–5077.
- LUNDSTROM, B. N., HIGGS, M. H., SPAIN, W. J., & FAIRHALL, A. (2008). Fractional differentiation by neocortical pyramidal neurons. *Nat Neurosci*, 11, 1335–1342.
- MARAVALL, M., PETERSEN, R. S., FAIRHALL, A. L., ARABZADEH, E., & DIAMOND, M. E. (2007). Shifts in coding properties and maintenance of information transmission during adaptation in barrel cortex. *PLoS Biol*, 5(2), e19.
- MAROM, S. (1998). Slow changes in the availability of voltage-gated ion channels: Effects on the dynamics of excitable membranes. *J Membr Biol*, 161(2), 105–113.
- MEASE, R. A., FAMULARE, M., GJORGJEVA, J., MOODY, W. J., & FAIRHALL, A. L. (2013). Emergence of adaptive computation by single neurons in the developing cortex. *J Neurosci*, 33(30), 12154–12170.
- NAGEL, K. I., & DOUPE, A. J. (2006). Temporal processing and adaptation in the songbird auditory forebrain. *Neuron*, 51(6), 845–859.
- NAGEL, K. I., & WILSON, R. I. (2011). Biophysical mechanisms underlying olfactory receptor neuron dynamics. *Nat Neurosci*, 14(2), 208–216.
- OLDHAM, K. B., & SPANIER, J. (2006). *The fractional calculus: Theory and applications of differentiation and integration to arbitrary order*. Mineola, NY: Dover.
- OLSHAUSEN, B. A., & FIELD, D. J. (1996a). Emergence of simple-cell receptive field properties by learning a sparse code for natural images. *Nature*, 381(6583), 607–609.
- OLSHAUSEN, B. A., & FIELD, D. J. (1996b). Natural image statistics and efficient coding. *Network*, 7(2), 333–339.
- OLSHAUSEN, B. A., & FIELD, D. J. (1997). Sparse coding with an overcomplete basis set: A strategy employed by V1? *Vis Res*, 37(23), 3311–3325.
- OLSHAUSEN, B. A., & FIELD, D. J. (2004). Sparse coding of sensory inputs. *Curr Opin Neurobiol*, 14, 481–487.
- OZUYSAL, Y., & BACCUS, S. A. (2012). Linking the computational structure of variance adaptation to biophysical mechanisms. *Neuron*, 73(5), 1002–1015.
- PANINSKI, L. (2003). Convergence properties of three spike-triggered analysis techniques. *Network*, 14(3), 437–464.
- PANINSKI, L., LAU, B., & REYES, A. (2003). Noise-driven adaptation: In vitro and mathematical analysis. *Neurocomputing*, 52–54, 877–883.
- PANZERI, S., BRUNEL, N., LOGOTHETIS, N. K., & KAYSER, C. (2010). Sensory neural codes using multiplexed temporal scales. *Trends Neurosci*, 33(3), 111–120.
- PFISTER, J. P., DAYAN, P., & LENGUEL, M. (2010). Synapses with short-term plasticity are optimal estimators of presynaptic membrane potentials. *Nat Neurosci*, 13(10), 1271–1275.
- POZZORINI, C., NAUD, R., MENSI, S., & GERSTNER, W. (2013). Temporal whitening by power-law adaptation in neocortical neurons. *Nat Neurosci*, 16(7), 942–948.
- PUCCINI, G. D., SANCHEZ-VIVES, M. V., & COMPTE, A. (2007). Integrated mechanisms of anticipation and rate-of-change computations in cortical circuits. *PLoS Comput Biol*, 3(5), e82.
- QUIROGA, R. Q., REDDY, L., KREIMAN, G., KOCH, C., & FRIED, I. (2005). Invariant visual representation by single neurons in the human brain. *Nature*, 435(7045), 1102–1107.
- RIEKE, F., & RUDD, M. E. (2009). The challenges natural images pose for visual adaptation. *Neuron*, 64(5), 605–616.
- RIEKE, F., WARLAND, D., de RUYTER VAN STEVENINCK, R., & BIALEK, W. (1997). *Spikes: Exploring the neural code*. Cambridge, MA: MIT Press.
- ROBERTS, P. D., & PORTFORS, C. V. (2008). Design principles of sensory processing in cerebellum-like structures. Early stage processing of electrosensory and auditory objects. *Biol Cybern*, 98(6), 491–507.
- RUDD, M. E., & BROWN, L. G. (1997). Noise adaptation in integrate-and-fire neurons. *Neural Comput*, 9(5), 1047–1069.
- RUDERMAN, D. L. (1997). Origins of scaling in natural images. *Vis Res*, 37(23), 3385–3398.
- RUDERMAN, D. L., & BIALEK, W. (1994). Statistics of natural images: Scaling in the woods. *Phys Rev Lett*, 73(6), 814–817.
- SATO, K., PELLEGRINO, M., NAKAGAWA, T., NAKAGAWA, T., VOSSHALL, L. B., & TOUHARA, K. (2008). Insect olfactory receptors are heteromeric ligand-gated ion channels. *Nature*, 452(7190), 1002–1006.
- SCHWARTZ, G., & RIEKE, F. (2011). Perspectives on information and coding in mammalian sensory physiology: Nonlinear spatial encoding by retinal ganglion cells: When $1 + 1 \neq 2$. *J Gen Physiol*, 138(3), 283–290.
- SCHWARTZ, O., PILLOW, J. W., RUST, N. C., & SIMONCELLI, E. P. (2006). Spike-triggered neural characterization. *J Vis*, 6, 484–507.
- SCHWINDT, P. C., SPAIN, W. J., & CRILL, W. E. (1989). Long-lasting reduction of excitability by a sodium-dependent potassium current in cat neocortical neurons. *J Neurophysiol*, 61, 233–244.
- SCHWINDT, P. C., SPAIN, W. J., FOEHRING, R. C., CHUBB, M. C., & CRILL, W. E. (1988). Slow conductances in neurons from cat sensorimotor cortex in vitro and their role in slow excitability changes. *J Neurophysiol*, 59(2), 450–467.
- SHAPLEY, R., & VICTOR, J. D. (1979). The contrast gain control of the cat retina. *Vis Res*, 19(4), 431–434.
- SHARPEE, T. O., RUST, N. C., & BIALEK, W. S. (2004). Analyzing neural responses to natural signals: Maximally informative dimensions. *Neural Comput*, 16, 223–250.
- SHARPEE, T. O., SUGIHARA, H., KURGANSKY, A. V., REBRIK, S. P., STRYKER, M. P., & MILLER, K. D. (2006). Adaptive filtering enhances information transmission in visual cortex. *Nature*, 439, 936–942.
- SIMONCELLI, E. P., & OLSHAUSEN, B. A. (2001). Natural image statistics and neural representation. *Annu Rev Neurosci*, 24, 1193–1216.
- SIMONCELLI, E. P., PANINSKI, L., PILLOW, J., & SCHWARTZ, O. (2004). Characterization of neural responses with

- stochastic stimuli. In M. Gazzaniga (Ed.), *The new cognitive neurosciences* (pp. 327–338). Cambridge, MA: MIT Press.
- SMIRNAKIS, S. M., BERRY, M. J., WARLAND, D. K., BIALEK, W., & MEISTER, M. (1997). Adaptation of retinal processing to image contrast and spatial scale. *Nature*, *386*(6620), 69–73.
- SMITH, E. C., & LEWICKI, M. S. (2006). Efficient auditory coding. *Nature*, *439*(7079), 978–982.
- SNIPPE, H. P., POOT, L., & VAN HATEREN, J. H. (2004). Asymmetric dynamics of adaptation after onset and offset of flicker. *J Vis*, *4*(1), 1–12.
- SOO, F. S., DETWILER, P. B., & RIEKE, F. (2008). Light adaptation in salamander L-cone photoreceptors. *J Neurosci*, *28*(6), 1331–1342.
- SRINIVASAN, M. V., LAUGHLIN, S. B., & DUBS, A. (1982). Predictive coding: A fresh view of inhibition in the retina. *Proc R Soc Lond B Biol Sci*, *216*(1205), 427–459.
- STAMPER, S. A., FORTUNE, E. S., & CHACRON, M. J. (2013). Perception and coding of envelopes in weakly electric fishes. *J Exp Biol*, *216*, 2393–2402.
- THEUNISSEN, F. E., SEN, K., & DOUPE, A. J. (2000). Spectral-temporal receptive fields of nonlinear auditory neurons obtained using natural sounds. *J Neurosci*, *20*(6), 2315–2331.
- THORSON, J., & BIEDERMAN-THORSON, M. (1974). Distributed relaxation processes in sensory adaptation. *Science*, *183*(121), 161–172.
- TOIB, A., LYAKHOV, V., & MAROM, S. (1998). Interaction between duration of activity and time course of recovery from slow inactivation in mammalian brain Na channels. *J Neurosci*, *18*, 1893–1903.
- TOLHURST, D. J., TADMOR, Y., & CHAO, T. (1992). Amplitude spectra of natural images. *Ophthalmic Physiol Opt*, *12*(2), 229–232.
- TRUCCOLO, W., EDEN, U. T., FELLOWS, M. R., DONOGHUE, J. P., & BROWN, E. N. (2005). A point process framework for relating neural spiking activity to spiking history, neural ensemble, and extrinsic covariate effects. *J Neurophysiol*, *93*(2), 1074–1089.
- VAN HATEREN, J. H. (1992a). A theory of maximizing sensory information. *Biol Cybern*, *68*, 23–29.
- VAN HATEREN, J. H. (1992b). Theoretical predictions of spatiotemporal receptive fields of fly LMCs, and experimental validation. *J Comp Physiol A*, *171*, 157–170.
- WARK, B., FAIRHALL, A. L., & RIEKE, F. (2009). Time scales of inference in visual adaptation. *Neuron*, *61*(5), 750–761.
- WARK, B., LUNDSTROM, B. N., & FAIRHALL, A. (2007). Sensory adaptation. *Curr Opin Neurobiol*, *17*(4), 423–429.
- WEN, B., WANG, G. I., DEAN, I., & DELGUTTE, B. (2012). Time course of dynamic range adaptation in the auditory nerve. *J Neurophysiol*, *108*(1), 69–82.
- YU, Y., & LEE, T. S. (2003). Dynamical mechanisms underlying contrast gain control in single neurons. *Phys Rev E*, *68*, 011901.
- ZUCKER, R. S., & REGEHR, W. G. (2002). Short-term synaptic plasticity. *Annu Rev Physiol*, *64*, 355–405.

