Adaptation and Information Transmission in Fly Motion Detection

Moshe N. Safran,[1,2] Virginia L. Flanagin,[3] Alexander Borst,[3,4] and Haim Sompolinsky[2,5,6]

[1]*Institute of Life Sciences,* [2]*Interdisciplinary Center for Neural Computation,* [5]*Racah Institute of Physics, Hebrew University, Jerusalem, Israel;* [3]*Bernstein Center for Computational Neuroscience, Munich, Germany;* [4]*Department of Systems and Computational Neuroscience, Max Planck Institute of Neurobiology, Martinsried, Germany; and* [6]*Center for Brain Science, Harvard University, Cambridge, Massachusetts*

Safran MN, Flanagin VL, Borst A, Sompolinsky H. Adaptation and information transmission in fly motion detection. *J Neurophysiol* 98: 3309–3320, 2007. First published October 10, 2007; doi:10.1152/jn.00440.2007. In this work, we studied the adaptation of H1, a motion-sensitive neuron in the fly visual system, to the variance of randomly fluctuating velocity stimuli. We ask two questions. *1*) Which components of the motion detection system undergo genuine adaptational changes in response to the variance of the fluctuating velocity signal? *2*) What are the consequences of this adaptation for the information processing capabilities of the neuron? To address these questions, we characterized the adaptation of H1 by estimating the changes in the parameters of an associated Reichardt motion detection model under various stimulus conditions. The strongest stimulus dependence was exhibited by the temporal kernel of the motion detector and was parametrized by changes in the model's high-pass time constant (τ_H). This time constant shortened considerably with increasing velocity fluctuations. We showed that this adaptive process contributes significantly to the shortening of the velocity response time-course but not to velocity gain control. To assess the contribution of time-constant adaptation to information transmission, we compared the information rates generated by our adaptive model motion detector with model simulations in which τ_H was held fixed at its unadapted value for all stimulus conditions. We found that for intermediate stimulus conditions, fixing τ_H at its unadapted value led to higher information rates, suggesting that time-constant adaptation does not optimize total information rates about velocity trajectories. We also found that, over the wide range of stimulus conditions tested here, H1 information rates are dependent on the amplitude of velocity fluctuations.

INTRODUCTION

Adaptation is usually defined as a change in the response properties of a system, making it better suited to cope with the present environment. One interesting form of adaptive behavior in sensory systems is the adaptive temporal integration of motion stimuli exhibited by neurons in both the fly (Borst et al. 2005) and the primate (Bair and Movshon 2004) visual systems. The time course of these neurons' responses to randomly fluctuating velocity stimuli, as measured in the velocity spike-triggered average (STA), shortens when the amplitude of velocity fluctuations is increased. Previous work (Borst et al. 2005) has shown that a simple model of motion detection, the Reichardt model, predicts an automatic shortening of the time course of the velocity STA with increasing velocity fluctuations, in qualitative agreement with experimental results, even when all model parameters are assumed to be fixed. Similarly,

the well-documented velocity gain control observed in H1 (Borst 2003; Brenner et al. 2000a; Fairhall et al. 2001) was shown to be an automatic consequence of the inherent nonlinearity of the Reichardt detector (Borst et al. 2005). Here we ask whether in addition to these effects, parameters of the fly motion detection system undergo genuine adaptational changes in response to the variance of the fluctuating velocity signal. To address this question, we modeled H1 as an array of Reichardt motion detectors, consisting of a high-pass filter (HPF), a low-pass filter (LPF), a multiplier, and a subtraction stage, followed by a static nonlinearity. Model parameters were fit to spike trains recorded from H1 under a wide range of stimulus conditions. We found that the HPF time constant (τ_H) is strongly dependent on the stimulus statistics, shortening considerably with increasing velocity fluctuations. This parameter change was interpreted as the signature of an adaptive process in the fly motion detection system and was related to the known adaptive properties of motion-sensitive neurons measured in other stimulus paradigms (Borst and Egelhaaf 1987; Borst et al. 2003; de Ruyter van Steveninck et al. 1986; Harris et al. 1999; Reisenman et al. 2003). Comparison of the observed velocity STAs to model simulations with τ_H fixed at its minimal and maximal observed values revealed that this adaptation contributes significantly to the shortening of the time course of the velocity response. The second model component that changed with the stimulus statistics was the static nonlinearity, which underwent a relatively weak reduction in slope, quantifying the contribution of additional adaptive processes beyond the automatic gain control (Borst et al. 2005) predicted by the Reichardt model.

Information-theoretic measures are often used to assess the functional role of sensory adaptation. For example, velocity gain control in H1 has been interpreted as an adaptive rescaling of the response that serves to maximize information transmission by enabling the system to use its full dynamic range under changing stimulus conditions (Brenner et al. 2000a). However, our results suggest that gain control is primarily an automatic consequence of the inherent nonlinearity of the motion detector, whereas the primary adaptive process is related to the system's temporal kernel. This adaptation is parametrized in our model by the observed reduction in τ_H with increasing velocity fluctuations. We therefore studied the effect of τ_H adaptation on information transmission in H1. To address this issue, we compared the information rates generated by our adaptive model motion detector with model simulations in which τ_H was held fixed at its minimal and maximal observed

Address for reprint requests and other correspondence: M. N. Safran, Dept. of Neurobiology, Inst. of Life Sciences, Silberman Bldg., Rm. 3-307, Hebrew Univ., Givat Ram, Jerusalem 91904, Israel (E-mail: moshe.safran@mail.huji.ac.il).

values for all stimulus conditions. We found that for intermediate stimulus conditions, longer τ_H led to stronger responses, which in turn led to higher information rates. These results suggest that the observed reduction in τ_H actually has a detrimental effect on overall information transmission. Similarly, we found that, over the wide range of stimulus conditions tested here, H1 information rates (both bits/s and bits/spike) are dependent on the amplitude of velocity fluctuations, indicating that the system does not optimally use its full dynamic range (bits/s) or operate at optimal efficiency (bits/spike) under all stimulus conditions.

METHODS

Experiments

Blowflies (*Calliphora vicina*, $n = 20$) were stimulated by a moving sinusoidal wave grating (22° spatial wavelength, 63% contrast, 14-cd/m² mean luminance), with a low-pass filtered white noise velocity profile. The correlation time of the velocity profile was either $\tau_0 = 20$ ms or $\tau_0 = 100$ ms ($n = 10$ flies for each value of τ_0; we shall refer to the two groups as "$\tau_0 = 20$ ms flies" and "$\tau_0 = 100$ ms flies," respectively). Each fly was stimulated with five different velocity signals, with SD of $\sigma = 0.1, 0.5, 1, 5,$ and 10 periods per second (Hz), respectively. To verify consistency of experimental conditions in the two groups of flies, all $\tau_0 = 20$ ms flies were presented with one additional stimulus with parameters $\tau_0 = 100$ ms and $\sigma = 5$ Hz, and we verified that results were similar to those of the $\tau_0 = 100$ ms flies for the same stimulus condition. For each stimulus condition, 75–110 sweeps of identical stimuli, each with a duration of 9 s, were presented, with a 1-s pause between stimulus presentations. The grating was presented on a cathode ray tube (Tektronix 608) by means of a Picasso image synthesizer (Innisfree, Cambridge, MA), at a frame rate of 200 Hz. The screen had a horizontal and vertical extent of 65 and 80°, respectively, as seen by the fly, and was positioned at a distance of 7.5 cm from the fly, 75% left-frontal and 25% right-frontal, to optimally stimulate the left H1 neuron. Spikes were recorded extracellularly from the left H1 neuron with a tungsten electrode inserted in the lobula plate, fed through a threshold device, and transferred at 1-kHz temporal resolution to a computer (Pentium II–based PC with a DAS16 I/O board, MetraByte, Tauton, MA). Data analysis was performed using custom-written MATLAB (The Math-Works, Natick, MA) programs on data rebinned to 2-ms temporal resolution, unless otherwise stated.

Model of motion detection

STIMULUS.   Similar to the experiment, the stimulus in the model consists of a moving sine grating, whose velocity $v(t)$ is expressed as a temporal frequency (in Hz) that reflects the number of spatial periods passing a given image location per second. The luminance level at a given angular location θ at time t is given by

$$L(\theta, t) = L_0\{1 + \rho \sin [2\pi\theta/\lambda - x(t)]\} \qquad (1)$$

where L_0 is the mean luminance level, ρ is the contrast of the grating, λ is its spatial wavelength, and $x(t) = 2\pi \int_0^t v(u)du$ is the total displacement of the grating from time 0 to time t. The units of $x(t)$ are such that $x = 2\pi$ corresponds to displacement by one spatial period of the grating. The velocity profile $v(t)$ is generated by low-pass filtering of gaussian white noise with zero mean. We write the time-lagged velocity autocorrelation as $\sigma^2 c(t)$, where $c(t) = \exp(-|t|/\tau_0)$ is the normalized autocorrelation at time lag t, τ_0 is its time constant, and σ is the SD of the velocity.

ELABORATED REICHARDT MOTION DETECTOR MODEL.   We model the fly's motion detection system (Fig. 1) as an array of local
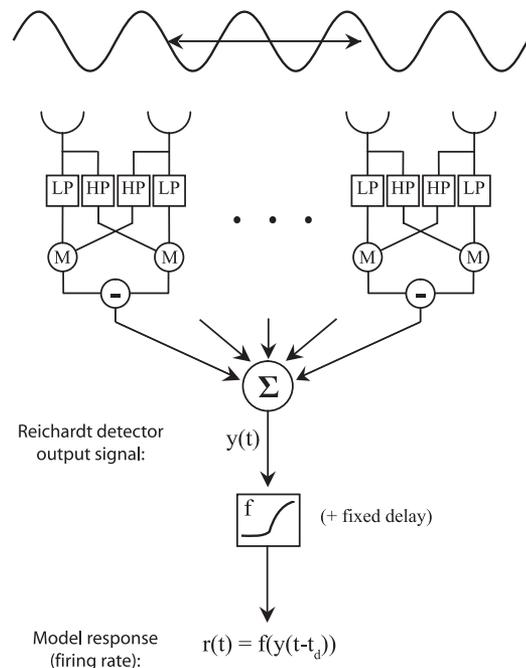

FIG. 1.   Elaborated Reichardt detector model. Luminance signals from neighboring locations are fed through a high-pass filter (HPF) and low-pass filter (LPF), respectively, and multiplied. This operation is done twice in mirror-symmetrical subunits, whose outputs are subtracted to produce a directionally sensitive signal of local motion. Outputs of an array of such local motion detectors are summed, delayed, and passed through a static nonlinearity, determining the instantaneous neuronal firing rate.

correlation-based motion detectors known as Reichardt detectors (Borst and Haag 2002; Egelhaaf 1989; Egelhaaf and Borst 1989; Egelhaaf and Reichardt 1987; Haag et al. 2004; Poggio and Reichardt 1973; Reichardt 1961, 1987; Single and Borst 1998). Reichardt detectors extract the direction of motion by multiplying[1] the brightness signals from neighboring image locations after asymmetric temporal filtering. This operation is done twice in mirror-symmetrical subunits, whose outputs are subtracted. Following Borst et al. (2003) and Bialek and de Ruyter van Stevenick (2005), we used a Reichardt detector with an LPF in one input line to the multiplier and an HPF in the cross arm. The output of a local Reichardt detector at angular location θ in response to a luminance field $L(\theta,t)$ is given by

$$y_{\text{local}}(\theta, t) = \int_0^\infty d\tau \int_0^\infty d\tau' K(\tau, \tau')L(\theta, t - \tau)L(\theta - \varepsilon, t - \tau') \qquad (2)$$

where $K(\tau, \tau') = K_H(\tau)K_L(\tau') - K_L(\tau)K_H(\tau')$ is the time response of the Reichardt detector, $K_H(\tau) = \delta(\tau) - \frac{1}{\tau_H} \exp(-\tau/\tau_H)$ and $K_L(\tau) = \frac{1}{\tau_L} \exp(-\tau/\tau_L)$ are the impulse responses of the HPF and LPF, respectively, τ_H and τ_L are their time constants, and ε is the angular spacing between the two arms of the motion detector.

The outputs of these local motion detectors then undergo spatial summation. For simplicity, we assume uniform summation over a visual field that is spanned by an integer number n of spatial periods of the sinusoidal grating stimulus (*Eq. 1*). Substituting *Eq. 1* into *Eq. 2* and integrating over θ, we find that the output signal of the Reichardt detector array in response to a moving sinusoidal grating is given by

---

[1] For possible mechanisms of neuronal multiplication, see Gabbiani et al. (2002) and Torre and Poggio (1978).

$$y(t) = y_0 \int_0^\infty d\tau \int_0^\infty d\tau' K_H(\tau) K_L(\tau') \sin[\Delta x(t - \tau, t - \tau')] \quad (3)$$

where $y_0 = n\lambda(L_0\rho)^2 \sin(2\pi\varepsilon/\lambda)$, and $\Delta x(t, t') = 2\pi\int_{t'}^t v(u)du$ is the total displacement of the grating from time t' to time t, in the same units as those used for $x(t)$ in *Eq. 1*. Importantly, the dependence of the Reichardt detector array's output signal on the stimulus history is proportional to the sine of Δx, rendering motion detection an inherently nonlinear process.[2]

Under our stimulus conditions, the mean of the Reichardt detector array's output signal is zero. Squaring *Eq. 3* and averaging over the Gaussian ensemble of velocity stimuli, we find that the variance of the output signal of the Reichardt detector array is given by

$$Var(y) = y_0^2 \int_0^\infty d\tau_1 \int_0^\infty d\tau_1' \int_0^\infty d\tau_2 \int_0^\infty d\tau_2' K_H(\tau_1) K_L(\tau_1') K_H(\tau_2) K_L(\tau_2')$$

$$\cdot \exp\{-2\pi^2\sigma^2[\Delta(\tau_1' - \tau_1) + \Delta(\tau_2' - \tau_2)]\}$$

$$\cdot \sinh[-4\pi^2\sigma^2\Gamma(\tau_1, \tau_1'; \tau_2, \tau_2')] \quad (4)$$

where $\Gamma(\tau_1, \tau_1'; \tau_2, \tau_2') \equiv \int_{\tau_1'}^{\tau_1} du_1 \int_{\tau_2'}^{\tau_2} du_2 c(u_1 - u_2)$ is the covariance of $\Delta x(\tau_1, \tau_1')$ and $\Delta x(\tau_2, \tau_2')$, normalized by $4\pi^2\sigma^2$. The normalized variance of $\Delta x(\tau, \tau')$ is given by $\Delta(\tau' - \tau) \equiv \Gamma(\tau, \tau'; \tau, \tau')$.

The output signal of the Reichardt detector array can assume both positive and negative values, producing a positive signal in response to preferred-direction motion stimuli and a negative signal in response to motion in the opposite direction. This signal can be interpreted as proportional to the total input received by H1 (Borst et al. 1995; Haag et al. 2004; Single and Borst 1998; Single et al. 1997). To compare our model to the spike trains of H1, we introduce an additional static nonlinearity, $f(\cdot)$, which transforms the Reichardt detector array's output signal, $y(t)$, into a positive firing rate, $r(t)$. We make no a priori assumptions about the shape of this static nonlinearity; instead, we calculate it directly from the data, as part of the parameter estimation process (see *Parameter estimation*). We also allow for an additional fixed delay in the system, t_d. The firing rate of our H1 model is therefore given by

$$r(t) = f[y(t - t_d)] \quad (5)$$

In this work, we shall refer to $y(t)$ as the Reichardt detector output signal and to $r(t)$ as the response or the firing rate of our H1 model.

Velocity response time course and gain

The time course of the velocity response is described by the stimulus-response cross-correlation function, defined as

$$c_{rv}(t) \equiv \langle r(t' - t) \cdot v(t') \rangle_{t'} \quad (6)$$

Under our stimulus conditions, this function is equivalent to the spike-triggered average of the velocity stimulus.

We define the velocity response function, $R_t(v)$, as the average firing rate at time $t' + t$, subject to the condition that the velocity at time t' is equal to v

$$R_t(v) \equiv \langle r(t' + t)|v(t') = v \rangle_{t'} \quad (7)$$

where $\langle \ldots \rangle_{t'}$ denotes average over time t'. Under stationary stimulus conditions, this is equivalent to averaging over the gaussian velocity stimulus ensemble. We define the velocity response gain, G_t, as the maximal slope of the velocity response function

$$G_t \equiv \max_v \frac{dR_t(v)}{dv} \quad (8)$$

In this work, we will use the velocity response function and gain at a time lag t_{peak} equal to the time of the peak of $c_{rv}(t)$.

Data analysis

PARAMETER ESTIMATION. To estimate the parameters of the fly's motion detection system, we minimize the mean-square error (MSE) between the peristimulus time histogram (PSTH) of H1, $r_{data}(t)$, and the firing rate predicted by the model, $r(t)$

$$MSE[f(\cdot), \tau_L, \tau_H, t_d] = \langle [r_{data}(t) - r(t)]^2 \rangle_t$$

$$= \langle \{r_{data}(t) - f[y_{\tau_L, \tau_H}(t - t_d)]\}^2 \rangle_t \quad (9)$$

where $y_{\tau_L, \tau_H}(t - t_d)$ is the (delayed) output signal generated by a model Reichardt detector array with time constants τ_L and τ_H, in response to the stimulus used in the experiment. As is explained below (*Performance and model selection*), the average $\langle \ldots \rangle_t$ is taken either over all stimulus conditions or only over times belonging to one particular value of σ, depending on the particular model variant. Similarly, τ_L and τ_H denote either a single time constant or a vector of five time constants, one for each σ, depending on the model variant.

To calculate the static nonlinearity $f_{\tau_L, \tau_H, t_d}^{est}(y')$ that minimizes *Eq. 9* for a particular choice of τ_L, τ_H, and t_d, we rewrite *Eq. 9* as

$$MSE[f(\cdot), \tau_L, \tau_H, t_d] = \langle [r_{data}(t)]^2 \rangle_t$$

$$+ \int dy' P_{\tau_L, \tau_H}(y')\{[f(y')]^2$$

$$- 2f(y')\langle r_{data}(t)|y_{\tau_L, \tau_H}(t - t_d) = y' \rangle\} \quad (10)$$

where $P_{\tau_L, \tau_H}(y')$ is the probability that $y_{\tau_L, \tau_H}(t - t_d) = y'$, and $\langle \ldots |y_{\tau_L, \tau_H}(t - t_d) = y' \rangle$, denotes an average over all times for which $y_{\tau_L, \tau_H}(t - t_d) = y'$. Minimizing *Eq. 10* with respect to the function $f(y')$, we find that the estimated static nonlinearity for a given choice of τ_L, τ_H, and t_d is equal to the average firing rate of the neuron, conditioned on y'

$$f_{\tau_L, \tau_H, t_d}^{est}(y') = \langle r_{data}(t)|y_{\tau_L, \tau_H}(t - t_d) = y' \rangle \quad (11)$$

In practice, $f_{\tau_L, \tau_H, t_d}^{est}(y')$ is estimated by binning the values of y' (binwidth = 3/7 Std(y_{τ_L, τ_H})) and calculating the average firing rate for all times t in which $y_{\tau_L, \tau_H}(t - t_d)$ falls into a given bin.

Substituting the left-hand side of *Eq. 11* into *Eq. 10*, we now find that minimizing the MSE is equivalent to maximizing the following objective function

$$V[\tau_L, \tau_H, t_d] \equiv \int dy' P_{\tau_L, \tau_H}(y')[f_{\tau_L, \tau_H, t_d}^{est}(y')]^2$$

$$= \langle \{f_{\tau_L, \tau_H, t_d}^{est}[y_{\tau_L, \tau_H}(t - t_d)]\}^2 \rangle_t \quad (12)$$

with respect to τ_L, τ_H, and t_d. As in *Eq. 9*, the averages in *Eqs. 11* and *12* are either over all stimulus conditions or over only one value of σ, and τ_L and τ_H are either scalars or vectors, depending on the model variant.

Because the mean of $f_{\tau_L, \tau_H, t_d}^{est}[y_{\tau_L, \tau_H}(t - t_d)]$ is always, by definition, equal to the mean firing rate $\langle r_{data}(t) \rangle$, maximizing V is equivalent to maximizing the variance of $f_{\tau_L, \tau_H, t_d}^{est}[y_{\tau_L, \tau_H}(t - t_d)]$. Our objective function V (*Eq. 12*) is equivalent to the one introduced by (Paninski 2003) in the context of estimation of the parameters of a linear-nonlinear (LN) model, consisting of a linear filter (K) followed by a static nonlinearity (f), under non-gaussian stimulus conditions. Here we show that *Eq. 12*, which was derived in Paninski (2003) using a ϕ-divergence technique, can also be derived from the principle of

---

[2] For nonsinusoidal periodic gratings, we find that $y(t)$ is a weighted sum over the spatial frequencies present in the grating, where each element of the sum is proportional to $\sin(k\Delta x)$, $k = 1, 2, \ldots$ (M.N.S., unpublished notes).

minimization of the MSE. We apply the method to our scenario, in which the filter is assumed to have the form of a Reichardt detector, $K = K_H(\tau)K_L(\tau')$, parametrized only by the two time constants τ_L and τ_H. The $\sin[\Delta x(t - \tau, t - \tau')]$ term in *Eq. 3* corresponds to the non-gaussian stimulus discussed in Paninski (2003). Both the stimulus and the filter are, in our scenario, vectors in the space of functions of two times, τ and τ'.

PERFORMANCE AND MODEL SELECTION. We consider three possible loci of adaptation: the two motion detector time constants τ_L and τ_H and the static nonlinearity $f(\cdot)$. To determine which, if any of these parameters adapt to the velocity variance (σ^2), we compare the performance of all $2^3 = 8$ possible variants of our model. In each model variant, some of the parameters are estimated separately for each value of σ, whereas the other(s) are assumed to be fixed for all values of σ (see the table in Fig. 3). For simplicity, the delay t_d is assumed to be independent of σ in all model variants.

The simplest model variant is *model A* (see table in Fig. 3), in which all parameters are fixed for all values of σ. For this model variant, the averages in *Eqs. 11* and *12* are taken over all stimulus conditions, and τ_L and τ_H are scalars. To estimate the parameters of this model, we compute $y_{\tau_L,\tau_H}(t - t_d)$, $f^{est}_{\tau_L,\tau_H,t_d}(\cdot)$, and $V[\tau_L, \tau_H, t_d]$ for all possible values of τ_L, τ_H, and t_d (upper and lower limits 2 ms $\le \tau_L$, $\tau_H \le 600$ ms, $0 \le t_d \le 60$ ms, sampled at 2-ms intervals), using *Eqs. 3*, *11*, and *12*. We maximize $V[\tau_L, \tau_H, t_d]$ by exhaustively screening this entire parameter space.

In *models B, C,* and *D*, $f(\cdot)$ is kept fixed for all values of σ, whereas τ_L and/or τ_H are allowed to adapt to the stimulus conditions. For these model variants, the averages in *Eqs. 11* and *12* are still over all stimulus conditions, but τ_L and/or τ_H become vector(s) of five time constants, one for each σ. Screening this high-dimensional parameter space exhaustively is prohibitively time consuming. We therefore maximize V for these model variants using a direct search algorithm, sampling progressively smaller (50 to 2 ms) intervals around the current estimated maximum. If no new maximum is found and the interval reaches 2 ms, the search is terminated; if a new maximum is found, the interval is increased and the search continues around the new maximum. The parameters that were estimated for *models F, G,* and *H*, respectively, are used as initial conditions for this algorithm for *models B, C,* and *D*, respectively.

For *models E–H*, in which $f(\cdot)$ is allowed to adapt, the averages in *Eqs. 11* and *12* are calculated separately for each stimulus condition, resulting in five separate objective functions $V_\sigma[\tau_L, \tau_H, t_d]$, $\sigma = 0.1$, 0.5, 1, 5, or 10 Hz, one for each stimulus condition. The arguments τ_L and τ_H of each V_σ are scalars. Each objective function is calculated exhaustively for all possible τ_L, τ_H, and t_d, as was done for *model A*. Each V_σ is maximized with respect to whatever time constant(s) are being allowed to adapt, obtaining the estimated adaptive time constant(s) for each possible choice of t_d and fixed time constant (if any). We maximize the sum of the resulting five $V_\sigma^{max}[\tau_{fixed}, t_d]$, to obtain the estimated fixed time constants(s) and t_d.

To compare the performance of the different model variants, we use a fivefold cross-validation procedure. Parameters for each model variant are estimated for a given fly using data from four fifths of the stimulus duration and tested on the remaining one fifth of the data from the same fly. The firing rate of the model in response to the test stimuli is calculated by applying *Eq. 3* to the test stimulus using the estimated time constants, yielding a Reichardt detector signal $y_{test}(t)$ and generating a firing rate $r_{test}(t)$ by linear interpolation of the estimated $f(y)$. We calculate the MSE between $r_{test}(t)$ and the response (PSTH) of H1 to the test stimulus. We repeat this procedure for five different choices of training and test datasets (folds), with each fold containing data from all five σ conditions. The resulting generalization MSEs are averaged over the five folds, yielding $n = 10$ generalization scores for each model, one for each fly. We perform a three-factor [τ_L, τ_H, and $f(\cdot)$] two-level (fix/ad) repeated-measures ANOVA (Keppel and Wickens 2004) to determine which, if any, of

the three model components significantly improve the generalization score when they are allowed to adapt. ANOVAs are performed using SPSS software (SPSS, Chicago, IL); effects with $P < 0.01$ are considered statistically significant. We also calculate the correlation coefficients between the predicted and actual firing rates in response to the various test stimuli.

CALCULATION OF VELOCITY RESPONSE TIME COURSE AND GAIN. Stimulus-response cross-correlations are calculated using *Eq. 6*. After subtracting the baseline value, measured over the range of $5 \times \tau_0 \le t \le 1$ s during the prestimulus period, we normalize the correlation functions by their peak values. We calculate the full-width at half-max and the peak latency of the resulting normalized correlation functions.

Velocity response functions (*Eq. 7*) are calculated at a time lag equal to the peak of the stimulus-response cross-correlation $c_{rv}(t)$ (*Eq. 6*). Velocity values are binned with a binwidth of $0.25 \times \sigma$. The velocity response gain is estimated by computing the maximal slope of the appropriate velocity response function. Slopes are computed using a velocity range of width $0.75 \times \sigma$.

CALCULATION OF INFORMATION RATES. We calculate the mutual information between the spike count $k(t)$ in a given window at time t, and the stimulus history $\{v(t + t')\}_{t'=-\infty}^{\infty}$, for each stimulus condition. We use window sizes of 2, 4, 10, and 20 ms for the $\tau_0 = 20$ ms flies and 2, 4, 10, 20, 40, and 100 ms for the $\tau_0 = 100$ ms flies. This analysis makes no assumptions about which aspects of the stimulus are being encoded by the spike counts. Following Strong et al. (1998), and assuming sufficient sampling of the stimulus space, we replace the average over all possible stimulus histories with an average over time, yielding

$$I[k(t); \{v(t + t')\}_{t'=-\infty}^{\infty}] = H(k) - H(k|\{v\})$$
$$= -\sum_k P(k)\log_2 P(k) + \sum_{k,\{v\}} P(k, \{v\})\log_2 P(k|\{v\})$$
$$= -\sum_k P(k)\log_2 P(k) + \frac{1}{T}\sum_{k,t} P(k|t)\log_2 P(k|t) \quad (13)$$

where $H(k) = -\sum_k P(k)\log_2 P(k)$ is the total entropy of the spike counts, and $H(k|\{v\}) = H(k|t) = -\frac{1}{T}\sum_{k,t} P(k|t)\log_2 P(k|t)$ is the noise entropy. $P(k|t)$ is the probability of observing k spikes at time t, and $P(k) = \frac{1}{T}\sum_t P(k|t)$ is the marginal distribution of k. The results of this analysis are presented in units of bits per second, and in units of bits per spike, calculated by dividing the result by the mean firing rate in each stimulus condition.

To calculate the total information rates of the H1 spike trains (*Eq. 13*), we estimate the distribution $P(k|t)$ empirically by constructing a histogram of spike counts measured at a given time during different trials. Following Strong et al. (1998) and Borst (2003), we control for sampling bias by calculating $H(k|t)$ based on data fractions of 1/10, 1/7, 1/5, 1/4, 1/3, 1/2, and 1 of the trials and performing a quadratic extrapolation to infinite sample size. To control for sampling bias of $H(k)$ caused by finite stimulus duration, we calculate $H(k)$ based on subsegments of 1/10, 1/7, 1/5, 1/4, 1/3, 1/2, and 1 of the stimulus duration, average over all subsegments for each fraction, and perform a quadratic extrapolation to infinite stimulus duration. As expected (Strong et al. 1998) for datasets of similar size (Brenner et al. 2000b), this correction is small.

We also calculate the information rates predicted by the Reichardt model of motion detection. For the smallest window size of 2 ms, we can safely assume that only one spike per time window can be produced, because of the refractory period of the H1 neuron. We therefore interpret the firing rate of the Reichardt model, $r(t)$, expressed in units of spikes per bin, as an instantaneous probability of

spiking, allowing us to evaluate *Eq. 13* using $P(k = 1|t) = r(t)$ and $P(k = 0|t) = 1 - r(t)$. For larger window sizes, it is necessary to incorporate a description of the trial-to-trial variability of the spike counts into our model. The responses of H1 to time-varying velocity signals such as ours are known to be highly reliable across trials, rendering the often-used inhomogeneous Poisson model a poor description of the spiking statistics (de Ruyter van Steveninck et al. 1997; Haag and Borst 1997). To incorporate a more accurate description of the spiking statistics into our model, we generalize the static nonlinearity $f(y)$ (*Eq. 11*), by calculating the full distribution of spike counts for a given value of the Reichardt detector output signal, $P(k|y)$. This distribution is calculated directly from the spike trains of H1 for each fly and for each stimulus condition, using the output signal $y_{\tau_L, \tau_H}(t - t_d)$ of a Reichardt detector with the τ_L, τ_H calculated in the parameter estimation procedure and delayed by the estimated t_d. Values of y are discretized with a binwidth of $\frac{3}{7} \cdot \text{Std}(y_{\tau_L, \tau_H})$. The mean of $P(k|y)$ is equal, by definition, to the value of the static nonlinearity $f^{est}_{\tau_L, \tau_H, t_d}(y)$ (*Eq. 11*), whereas its overall shape describes the trial-to-trial variability of the spike counts for a given value of the Reichardt detector output signal. This technique retains the assump-

tion that the spiking statistics are determined exclusively by the instantaneous value of the Reichardt detector output signal, y while "borrowing" the detailed behavior of the trial-to-trial variability from the actual data. For a description of a similar technique, see Brenner et al. (2000a).

To calculate the information rates predicted by the Reichardt model, we use *Eq. 3* to calculate the Reichardt detector output signal $y_{\tau_L, \tau_H}(t - t_d)$ generated in response to the stimuli used in the experiment, using the parameters τ_L, τ_H, and t_d previously estimated for each fly, averaged over training sets. For each stimulus condition (σ), we use the appropriate adaptive value of τ_H, as determined by the fitting procedure for model variant G. We calculate the generalized model response $P_{model}(k|t)$ by linear interpolation of the appropriate $P(k|y)$ at $y = y_{\tau_L, \tau_H}(t - t_d)$ [negative $P_{model}(k|t)$ values resulting from extrapolation beyond the edges of $P(k|y)$ were set to zero and the distribution was renormalized]. The resulting model response is then substituted into *Eq. 13*, yielding the model's information rate. To quantify the contribution of adaptation to information transmission, we repeat this analysis with τ_H fixed at its minimal and maximal observed values for all stimulus conditions and compare the results.
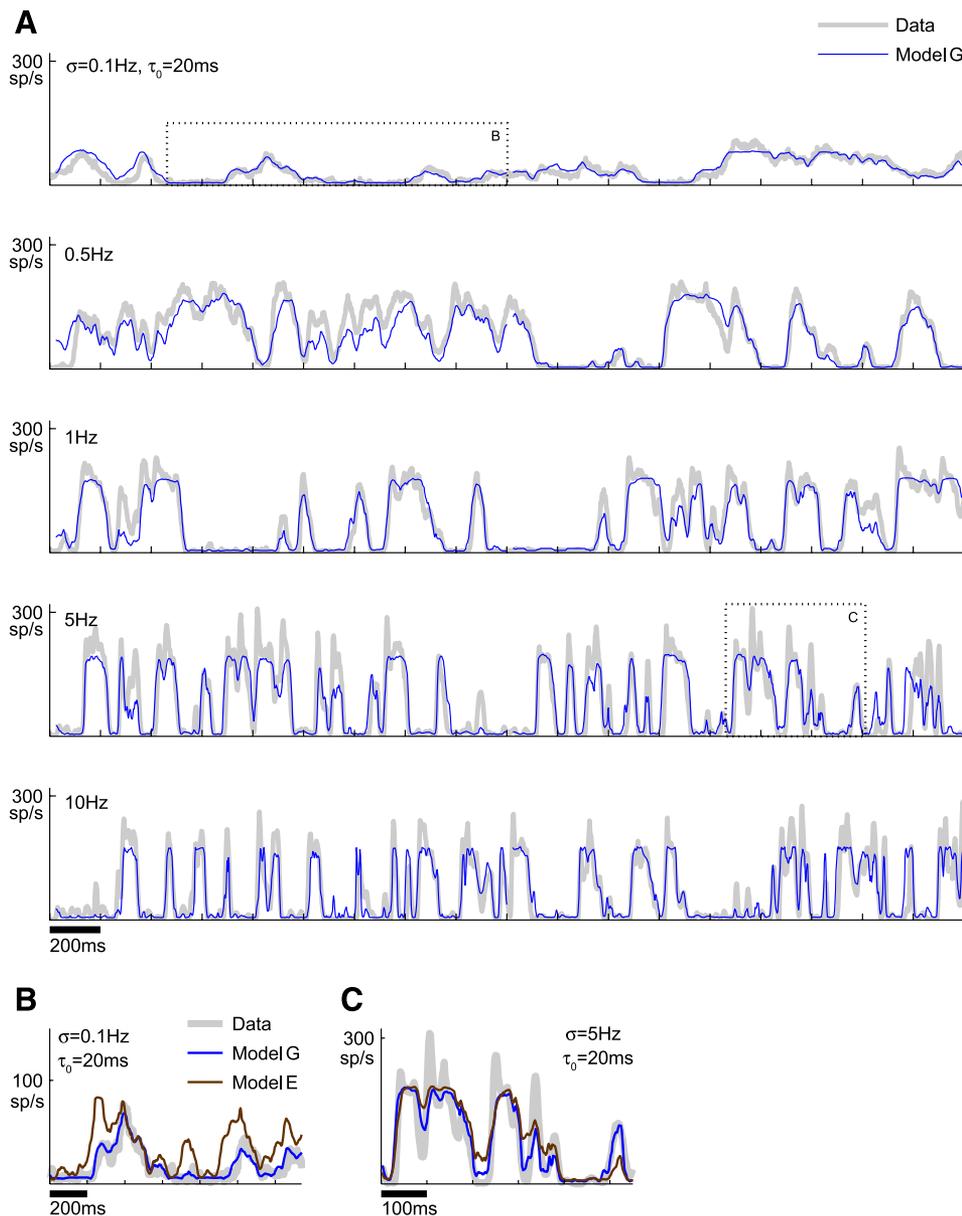

FIG. 2. H1 neuron and Reichardt model responses. *A*: peristimulus time histogram (PSTH) segments, $\tau_0 = 20$ ms, $\sigma = 0.1$, 0.5, 1, 5, and 10 Hz. Actual responses of an H1 neuron (gray; smoothed with a 10-ms moving average for display purposes only) and adaptive Reichardt model response (variant G, blue). *B* and *C*: PSTH segments (magnification of boxes in *A*); data (gray; smoothed with a 10-ms moving average for display purposes only), model variant G (blue), and model variant E (brown); $\sigma = 0.1$ (*B*) and 5 Hz (*C*); $\tau_0 = 20$ ms. See table in Fig. 3 for definition of model variants.

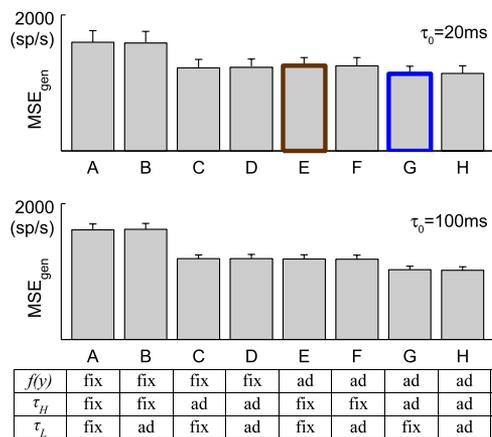|	A	B	C	D	E	F	G	H
$f(y)$	fix	fix	fix	fix	ad	ad	ad	ad
τ_H	fix	fix	ad	ad	fix	fix	ad	ad
τ_L	fix	ad	fix	ad	fix	ad	fix	ad

FIG. 3. Mean-square generalization error for all model variants (mean and SE between flies). Table shows which components were kept fixed (fix) and which were estimated separately (adapted) for different values of σ (ad) for each model variant.

RESULTS

Performance and model selection

We find that the elaborated Reichardt model accounts well for the complex responses of H1 to our randomly fluctuating velocity stimuli. Figure 2A shows 3.6-s segments of the experimentally measured firing rate (PSTH) of an H1 neuron under the five velocity variance conditions compared with the responses predicted by the Reichardt model (variant G, see table in Fig. 3). The model response follows the actual PSTH quite closely, despite the fact that model parameters for each 1.8-s segment were estimated using the remaining portion of the data (see *Methods*). The average correlation coefficients (CCs) between actual responses and model (variant G) predictions were between 0.76 and 0.96, depending on the stimulus condition.

To determine which parameters of the motion detection system adapt to stimulus statistics, we compare the performance of eight model variants. In each model variant, some

model parameters are estimated separately for each velocity variance condition, thereby allowing for adaptation to stimulus conditions, whereas the other parameter(s) were assumed to be fixed for all values of σ. Figure 3 shows the MSEs for each of the eight model variants, for the $\tau_0 = 20$ ms and $\tau_0 = 100$ ms flies. The table indicates which components are kept fixed (fix) and which are allowed to adapt to stimulus statistics (ad) in each model variant. Comparing *models A and B* with *models C–F*, we observe that allowing adaptation of either τ_H or $f(\cdot)$ leads to a significant (3-way repeated-measures ANOVA, $n = 10$, $P \leq 0.001$, main effects) reduction in the generalization error of the model. In contrast, estimating τ_L separately for each stimulus condition does not reduce the generalization error ($P = 0.829$ for $\tau_0 = 20$ ms, $P = 0.189$ for $\tau_0 = 100$ ms), as can be seen by comparing *model A* to *B*, *C* to *D*, etc. Allowing both τ_H and $f(\cdot)$ to adapt to the stimulus conditions (*model G*) results in a further reduction of the generalization error, as can be seen by comparing *model G* to *models C–F*. These findings are consistent for both of the stimulus correlation times (τ_0) used in our experiments [for $\tau_0 = 20$ ms, there was also a significant ($P < 0.001$) negative interaction between τ_H and $f(\cdot)$, reflecting the fact that the contributions of these 2 parameters to the MSE sum sublinearly]. We conclude that model variant G, with adaptive τ_H and $f(\cdot)$, is the best fit to H1 responses under our experimental conditions. This model variant also outperformed model variant E [adaptive $f(\cdot)$, fixed τ_H] under each individual stimulus condition (Wilcoxon sign-rank tests for generalization MSEs and CCs, $P < 0.01$), with the exception of $\tau_0 = 100$ ms, $\sigma = 10$ Hz, where the differences were not statistically significant.

Figure 2, *B* and *C,* compares segments of the experimentally measured firing rate (PSTH) with the responses of model variants E [adaptive $f(\cdot)$, fixed τ_H, brown] and G [adaptive $f(\cdot)$ and τ_H, blue] for a low-variance ($\sigma = 0.1$ Hz; Fig. 2B) and a high-variance ($\sigma = 5$ Hz; Fig. 2C) stimulus condition. In both cases, *model G* is a better match to the actual response, because of adaptation of the high-pass time constant τ_H, which is long for the low-variance condition and short for high-velocity
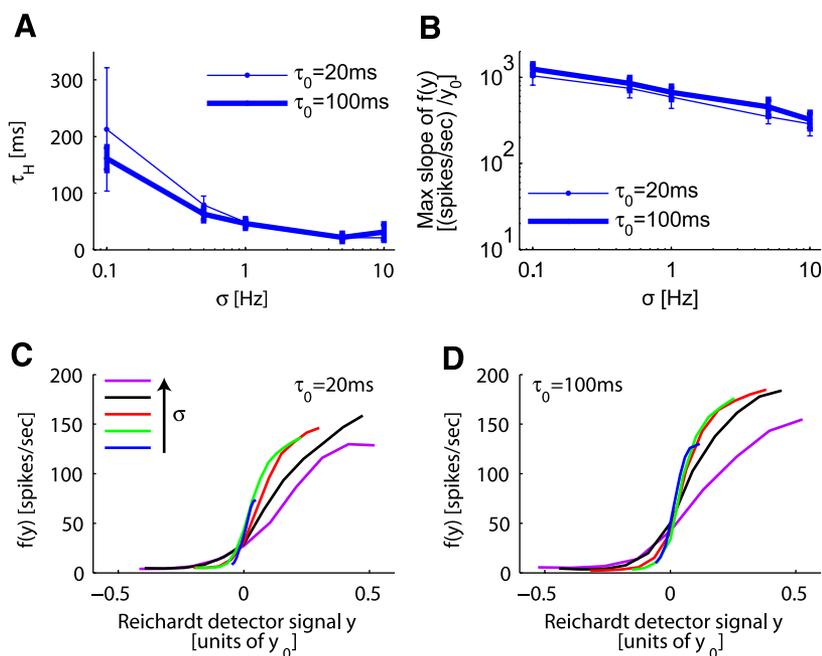

FIG. 4. Estimated model parameters (variant G). *A*: estimated high-pass time constants, $\tau_0 = 20$ (thin line) and 100 ms (thick line), mean and SD between flies. Two lines overlap at $\sigma = 1, 5$ Hz. *B*: maximal slope of the estimated static nonlinearities (mean and SD between flies). Slopes were calculated using 4 data points. Scale of y-axis is same as in Fig. 6B to facilitate comparison. *C* and *D*: estimated static nonlinearities averaged over all flies and training sets: $\sigma = 0.1$ (blue), 0.5 (green), 1 (red), 5 (black), and 10 Hz (purple); $\tau_0 = 20$ (*C*) and 100 ms (*D*).

variance (see *Motion detector parameters* and Fig. 4*A*). The fixed τ_H estimated under model variant E had an intermediate value, leading to an overly brisk model response to the low-variance stimulus (Fig. 2*B*) and an excessively sluggish response under the high-variance condition (Fig. 2*C*). For the remainder of this paper, we will therefore present results for model variant G, which we will refer to as the adaptive Reichardt model.

Motion detector parameters

The strongest dependence on stimulus statistics is exhibited by the HPF time constant τ_H (Fig. 4*A*), which shrinks from ~200 ms to as little as 20 ms with increasing σ. The behavior of τ_H is consistent for the two stimulus correlation times used in our experiments. The low-pass filter time constant (τ_L) amounted to ~30 ms, and the fixed delay (t_d) was ~20 ms. The second model component that exhibits statistically significant adaptation is the static nonlinearity (Fig. 4, *C* and *D*). The maximal slope of this function decreases by a factor of 3–4 as σ is increased (Fig. 4*B*). Model parameters estimated from $\tau_0 = 20$ ms flies' responses to a $\sigma = 5$ Hz, $\tau_0 = 100$ ms control stimulus were consistent with the behavior of the $\tau_0 = 100$ ms flies at $\sigma = 5$ Hz, although their mean firing rates were slightly lower than those of the $\tau_0 = 100$ ms flies.

In conclusion, we observe significant and systematic adaptation of motion detector parameters to the variance of our random velocity stimuli, as parametrized by the changes in the HPF time constant τ_H and the static nonlinearity $f(\cdot)$ in model variant G. In the following sections we will analyze the contribution of these parameter changes to the system's velocity response properties and information transmission.

Contribution of time-constant adaptation to velocity response time course and gain

We quantify the effective time course of the H1 velocity response by calculating the normalized stimulus-response cross-correlation function (*Eq. 6*). Figure 5*A* shows the shape of this function for the various stimulus conditions, calculated from the H1 spike trains and from the responses of model variant G. As previously reported in both the fly (Borst et al. 2005) and primate (Bair and Movshon 2004) visual systems, the width and peak latency of this function decrease with increasing velocity fluctuations (Fig. 5, *B* and *C*; black traces). The adaptive Reichardt model (Fig. 5, *B* and *C*; blue traces) reproduces this effect faithfully.

To quantify the contribution of τ_H adaptation to this effect, we simulate the model's response to the various stimulus conditions, but with τ_H held fixed at the minimal (20 ms; Fig. 5, *B* and *C*, red traces) or maximal (200 ms; Fig. 5, *B* and *C*,
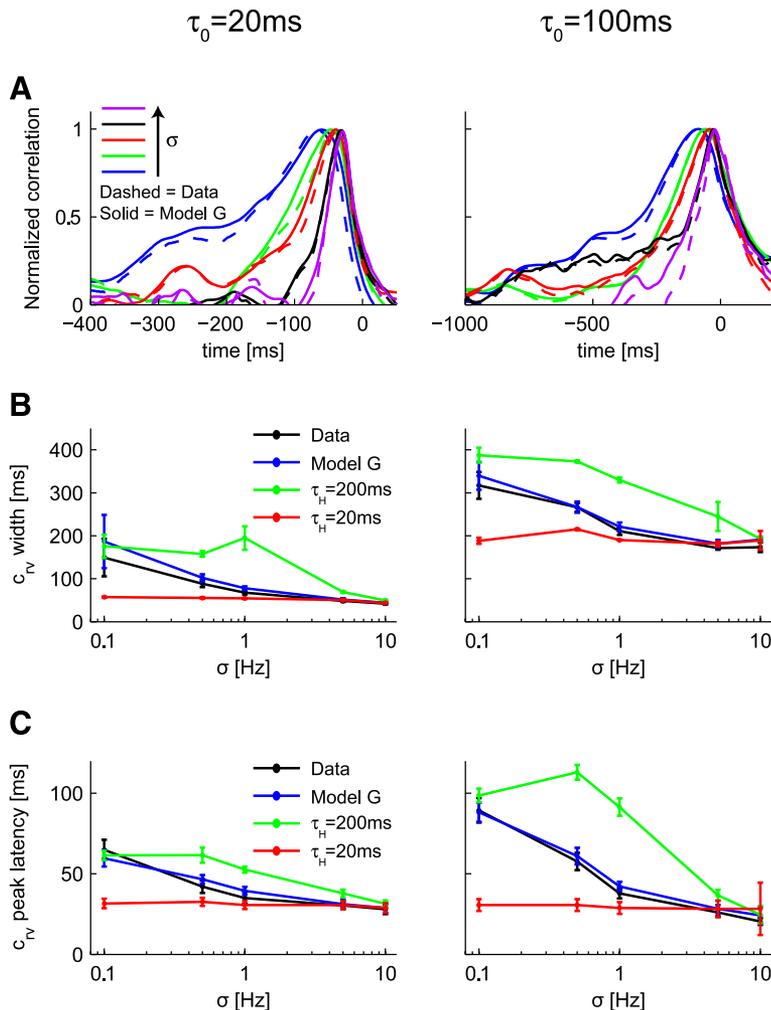

FIG. 5. Velocity response time course. *A*: normalized stimulus-response cross-correlation (*Eq. 6*), calculated from data (dashed lines) and model variant G (solid lines); $\sigma = 0.1$ (blue), 0.5 (green), 1 (red), 5 (black), and 10 Hz (purple); $\tau_0 = 20$ (*left*) and 100 ms (*right*). Curves become narrower and peak earlier as σ is increased. *B*: full-width at half-max of normalized c_{rv} as a function of σ; data (black), model variant G (blue), model simulations with $\tau_H = 200$ (green) or 20 ms (red) for all stimulus conditions; $\tau_0 = 20$ (*left*) and 100 ms (*right*). Mean and SD between flies. *C*: latency of peak of normalized c_{rv} as a function of σ; data (black), model variant G (blue), model simulations with $\tau_H = 200$ (green) and 20 ms (red); $\tau_0 = 20$ (*left*) and 100 ms (*right*). Mean and SD between flies.

green traces) values observed in our experiments. We measure the width and latency of the resulting c_{rv}. As previously reported (Borst et al. 2005), the width and latency decrease with σ even when τ_H is held fixed at 200 ms (Fig. 5, *B* and *C*, green traces; the slight nonmonotonicity of some of the $\tau_H =$ 200 ms simulation results shown in Fig. 5, *B* and *C*, is an artifact of the particular stimulus realizations used in our experiments and did not occur in model simulations when longer stimuli were used). However, the actual observed time-course control (Fig. 5, *B* and *C*, black traces) is clearly steeper, indicating that τ_H adaptation contributes significantly to this effect under most stimulus conditions. For large stimulus fluctuations (i.e., $\sigma = 10$ Hz), the inherent nonlinearity dominates, automatically suppressing the contribution of past velocity history regardless of the value of τ_H. Conversely, when τ_H is fixed at 20 ms (red traces), there is no noticeable automatic time-course control under our stimulus conditions, because of the weak history dependence of the model response.

The H1 neuron is also known to rapidly adapt its velocity response gain (*Eq. 8*) to the amplitude of stimulus velocity fluctuations (Borst et al. 2005; Brenner et al. 2000a; Fairhall et al. 2001). Figure 6A shows the velocity response functions (*Eq. 7*) for the various stimulus conditions, as calculated from the H1 spike trains and from the adaptive Reichardt model responses. The black and blue traces in Fig. 6B show the gain of this function (*Eq. 8*) for data and the adaptive Reichardt model, respectively. Over our range of σ, the velocity gain is reduced by a factor of 22 ($\tau_0 = 20$ ms) or 35 ($\tau_0 = 100$ ms) in H1 responses and by a similar factor in the adaptive Reichardt model. To determine whether time-constant adaptation contributes to the observed gain control, we calculated the velocity response gain from model simulations in which τ_H was held fixed at 20 or 200 ms for all stimulus conditions (Fig. 6B, red and green traces). The behavior of the gain is not altered when

τ_H is kept fixed at its unadapted value of 200 ms (green), indicating that time-constant adaptation does not contribute to velocity gain control.

The contribution of $f(\cdot)$ adaptation to velocity gain control is difficult to quantify without an exact description of the saturation properties of $f(\cdot)$ and their dependence on stimulus conditions. However, we note that the slope of $f(\cdot)$ decreases by a factor of 3–4 (Fig. 4B), whereas the velocity gain is reduced by a factor of 22–35 (Fig. 6B), suggesting that most of the observed gain control is explained by the inherent nonlinearity of the motion detector as described by the Reichardt model, independent of parameter change (Borst et al. 2005). The decrease in the slope of $f(\cdot)$ can be interpreted as the signature of additional processes of gain adaptation that are not accounted for by the Reichardt model.

Information transmission

The information rate of the H1 neuron has been found to be rather insensitive to the statistics of the velocity stimulus, whereas decreasing the size or the contrast of the grating results in a strong deterioration of information transmission (Borst 2003; Fairhall et al. 2001). However, these studies examined only a limited range of stimulus conditions. Here we calculate H1 information rates for a wide range of velocity variance conditions, spanning two orders of magnitude (Fig. 7, *A* and *B*, black traces). We find that over this wide range of σ, H1 information rates (bits/s, Fig. 7A; bits/spike, Fig. 7B, black traces) are not invariant to stimulus statistics. The adaptive Reichardt model (blue traces) exhibits similar behavior, albeit with lower overall information rates, reflecting the portion of response fluctuations that are not accounted for by the model.

We find that H1 information rates (bits/s) are correlated, across flies and stimulus conditions, with the mean ($r = 0.76$)
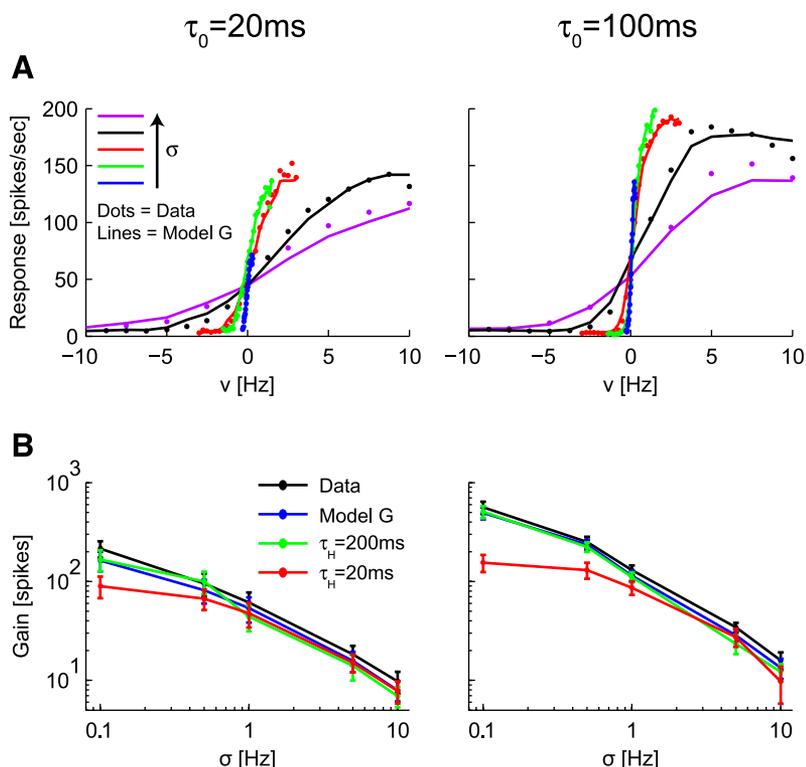

FIG. 6. Velocity response function and gain. *A*: velocity response function (*Eq. 7*), calculated from data (dots) and model variant G (lines); $\sigma = 0.1$ (blue), 0.5 (green), 1 (red), 5 (black), and 10 Hz (purple); $\tau_0 = 20$ (*left*) and 100 ms (*right*). *B*: velocity response gain (*Eq. 8*) as a function of σ; data (black), model variant G (blue), model simulations with $\tau_H = 200$ (green) or 20 ms (red) for all stimulus conditions; $\tau_0 = 20$ (*left*) and 100 ms (*right*). Mean and SD between flies.
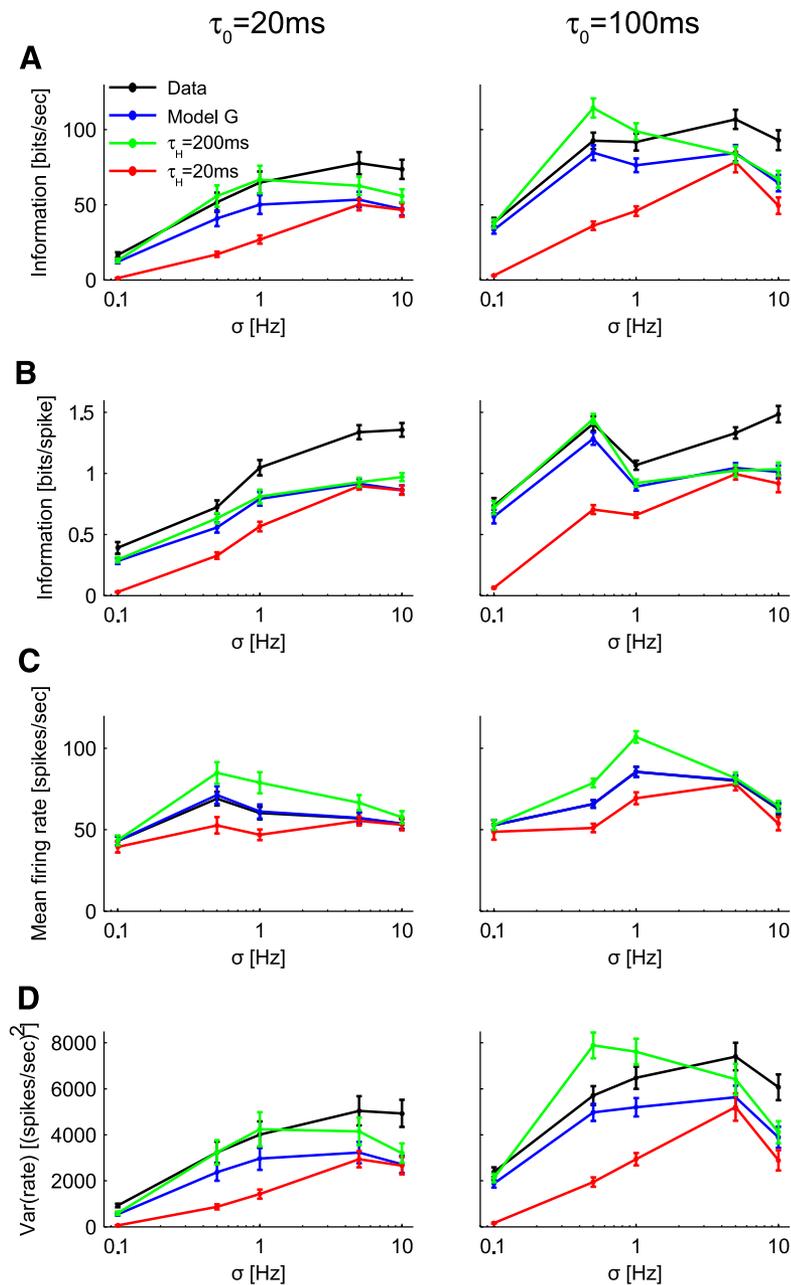
FIG. 7. Total information about stimulus history. *A*: mutual information between spike counts and entire stimulus history (*Eq. 13*), as a function of σ; data (black), model variant G (blue), model simulations with $\tau_H = 200$ (green) or 20 ms (red) for all stimulus conditions; $\tau_0 = 20$ (*left*) and 100 ms (*right*). Mean and SE between flies. Results shown are for spike count windows of 4 ms. *B*: same as *A*, but normalized by mean firing rate for each stimulus condition. *C*: mean firing rate as a function of σ. *D*: variance of firing rate as a function of σ.

and the variance ($r = 0.98$) of the firing rate. The mean firing rate (Fig. 7*C*, black) exhibits a bell-shaped dependence on σ, and the variance of the firing rate (Fig. 7*D*, black), as well as the information rate (bits/s; Fig. 7*A*), shows a similar trend for our range of stimulus conditions. Similar findings regarding the mean and variance of H1 firing rates have been previously reported in Flanagin (2006). The mean firing rates of our adaptive Reichardt model (Fig. 7*C*, blue) are by definition equal to those observed in the data, because of the way the static nonlinearity is defined (*Eq. 11*). The response variance of the adaptive Reichardt model (Fig. 7*D*, blue) exhibits a bell-shaped trend similar to that observed in the data.

Following the steps described in *Elaborated Reichardt motion detector model*, we derive an analytical expression for the variance of the Reichardt detector output signal, Var(*y*) (*Eq. 4*). This calculation reveals that Var(*y*) is a bell-shaped function of σ (Fig. 8). Thus the basic Reichardt model, without the static

nonlinearity, already predicts bell-shaped behavior of the mean firing rate, firing-rate variance, and information rate (bits/s).

The velocity gain control observed in H1 has been interpreted as an adaptive rescaling of the response, which serves to maximize information transmission by enabling the system to use its full dynamic range under changing stimulus conditions (Brenner et al. 2000a). However, our results suggest that gain control is primarily an automatic consequence of the inherent nonlinearity of the motion detector, whereas the primary adaptive process is related to the system's temporal kernel. This adaptation is parametrized in our model by changes in τ_H. We therefore ask whether τ_H adaptation maximizes information transmission. To address this question, we calculate the information rates (*Eq. 13*) predicted by the adaptive Reichardt model (Fig. 7*A*, blue traces). We compare these results to model simulations in which τ_H is held fixed at its minimal (20 ms; Fig. 7*A*, red traces) or maximal (200 ms; Fig. 7*A*, green
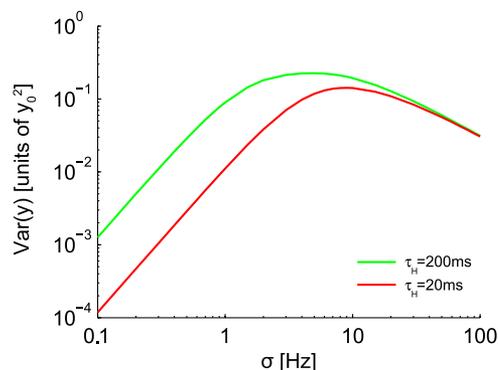
FIG. 8. Variance of the Reichardt detector output signal, as calculated in *Eq. 4*. $\tau_H = 200$ (green) and 20 ms (red); $\tau_0 = 20$ ms.

traces) observed value. We find that for intermediate values of σ, time-constant adaptation actually has a detrimental effect on information transmission, because information rates (bits/s) would have been higher had τ_H remained at its unadapted value of 200 ms. The information rate in bits per spike (Fig. 7B) is essentially the same with τ_H adaptation (blue trace) and with τ_H fixed at 200 ms (green trace), indicating that τ_H adaptation cannot be understood as an optimization of this measure, either. These results did not depend on the size of the time window used to count spikes (we show results for spike count window sizes of 4 ms, because these yielded the highest information rates at most σ for both values of τ_0), or on the extrapolation method used for $f(y)$.

Our analytical results (*Eq. 4*; Fig. 8) show that the variance of the Reichardt detector output signal, Var(y), is an increasing function of τ_H for our range of parameters and stimulus conditions. This increase leads to the higher firing rates, response variance, and information rates (bits/s) observed in the full model simulations (green traces compared with blue in Fig. 7, *A*, *C*, and *D*), where y is fed through the static nonlinearity, $f(y)$, to generate the model's firing rate. Thus our results suggest that saturation of the system's dynamic range, parametrized in our model by the shape of the static nonlinearity, is not an important limiting factor for information transmission under our stimulus conditions.

DISCUSSION

Adaptation of motion detector parameters

In this work, we studied the adaptation of the H1 neuron to stimulus velocity variance by fitting the neural responses to a simple model of motion detection. The model consisted of an array of Reichardt motion detectors, containing an HPF, an LPF, a multiplier, and a subtraction stage, followed by a static nonlinearity (Fig. 1). We found that the HPF time constant and static nonlinearity of the motion detector adapt to stimulus statistics (Figs. 3 and 4). The HPF time constant shortens considerably when stimulus fluctuations are increased, whereas the static nonlinearity shows a relatively small reduction in its slope. In contrast, we did not detect any adaptation of the LPF time constant (Fig. 3, *column A* vs. *B*, *C* vs. *D*, etc.). This behavior was highly consistent for the two stimulus correlation times (τ_0) used.

Our findings regarding the time constants of H1 are consistent with earlier results obtained using simple motion stimuli in

adapt-and-probe experimental paradigms. In its unadapted state, the H1 cell has been found to respond to a brief motion pulse with a sudden rise in activity, followed by an exponential decay with a time constant of ~ 300 ms. After prolonged exposure to constant motion stimuli, the decay time constant of the impulse response was found to shorten to values as low as 30 ms. The extent of the shortening was found to depend systematically on the velocity and the contrast of the adapting stimulus (de Ruyter van Steveninck et al. 1986). This adaptation was found to occur after exposure to motion in either the preferred or the null direction and even after exposure to flicker stimuli (Borst and Egelhaaf 1987). Adaptation was observed only when the test stimulus was presented in the same area of the visual field as the adapting stimulus, indicating that this is a spatially local process (de Ruyter van Steveninck et al. 1986; Reisenman et al. 2003). Analytical treatment of the Reichardt model with an HPF in one arm and an LPF in the other, as used in this work, shows that the impulse response time constant is equal to τ_H (Borst et al. 2003), indicating that this parameter undergoes motion adaptation. In this work, we found that τ_H shortens after exposure to white-noise velocity stimuli, with the amount of shortening depending systematically on the variance of the velocity fluctuations, extending the results of Borst and Egelhaaf (1987).

We did not detect any significant adaptation of the LPF time constant of the motion detection system for our stimulus conditions. This result is consistent with the behavior of the steady-state responses of H1 and other fly motion-sensitive neurons to constant motion stimuli. This response has been found to peak at stimulus velocities of ~ 2–10 Hz, independent of the stimulus contrast (Harris et al. 1999; Reisenman et al. 2003). Similarly, the steady-state responses of HS motion-sensitive neurons in the drone fly (*Eristalis tenax*), have been found to peak at ~ 7 Hz, independent of prior motion adaptation (Harris et al. 1999). These findings indicate that the relevant time constant for the steady-state velocity tuning is fixed at a few tens of milliseconds [i.e., $1/(2\pi \times 4 \text{ Hz}) \approx 40$ ms] and does not undergo significant motion adaptation. Analytical treatment of the Reichardt model shows that the location of the peak of the steady-state response depends primarily on τ_L, explaining the different behavior of the impulse and steady-state responses (Borst et al. 2003).

Lindemann et al. (2005) recorded the graded membrane potential responses of the blowfly motion-sensitive neuron HSE to naturalistic motion stimuli and fitted them to a similar model of motion detection. Their analysis yielded seemingly nonsystematic changes in the estimated motion detector time constants, which only weakly improved the goodness of fit. In contrast, our experimenter-controlled velocity stimuli enabled us to observe a systematic reduction in τ_H with increasing velocity fluctuations, and to examine the relative contribution of this adaptation to H1's response time course and information transmission properties. In addition, our cross-validation procedure indicated that of the two time constants, only τ_H exhibits significant adaptation under our stimulus conditions.

In addition to the shortening of τ_H, we also found a decrease in the slope of the static nonlinearity of the motion detector (Fig. 4B). This would predict a reduction in the amplitude of the impulse and steady-state responses after motion adaptation. Interestingly, Harris et al. (1999) found that the amplitude of both the impulse and the steady-state responses of the HS

neuron was reduced after exposure to constant motion stimuli because of a hyperpolarizing afterpotential.

Contribution to velocity response time-course and gain control

By comparing our results to model simulations in which τ_H was fixed at its maximal and minimal observed values, we showed that time-constant adaptation contributes significantly to the shortening of the velocity response time course (c_{rv}) but not to velocity gain control (Figs. 5 and 6). Over our range of σ, the velocity gain was reduced by a factor of 22 ($\tau_0 = 20$ ms) to 35 ($\tau_0 = 100$ ms) (Fig. 6*B*). A large portion of this effect can be accounted for by the inherent adaptive properties of the motion detection system, as described in Borst et al. (2005). The contribution of additional processes of gain adaptation is parametrized by the static nonlinearity $f(\cdot)$, which decreases its slope by a factor of three to four as σ is increased (Fig. 4*B*).

Adaptation and information transmission

Our model simulations indicate that time-constant adaptation in H1 does not optimize information transmission, because information rates would have been higher (bits/s) or unchanged (bits/spike) if τ_H had remained at its maximal value of 200 ms for all stimulus conditions (Fig. 7, *A* and *B*, cf. green and blue traces). Our results suggest that, as stimulus fluctuations are increased, time-constant adaptation facilitates increased encoding of recent velocities, at the expense of the system's overall information rate. Identifying the specific stimulus features encoded by H1 under different stimulus conditions and quantifying the reasons for their selection from a functional point of view are interesting challenges for future experimental and theoretical studies.

Optimal information transmission also predicts that information rates should remain constant over changing stimulus conditions. This prediction was corroborated in Borst (2003), where information rates were found to be largely independent of the stimulus entropy (firing rates were fairly constant for the stimuli used in that work). Similarly, Fairhall et al. (2001) found that H1 information rates scale with the firing rate, implying that the system is operating at a constant, presumably optimal, level of efficiency (bits/spike). However, these studies examined only a limited range of stimulus conditions. In this work, we calculated H1 information rates for a wider range of velocity fluctuation amplitudes, spanning two orders of magnitude. We found that, over this range of stimulus conditions, H1 information rates (both bits/s and bits/spike) are dependent on the amplitude of velocity fluctuations.

Multiple adaptive processes

Adaptation is usually defined as a change in the response properties of a given system, making it better suited to cope with the present environment. The fly motion detection system is a classic example of multiple adaptive behaviors (Borst and Egelhaaf 1987; Borst et al. 1995, 2005; Brenner et al. 2000a; de Ruyter van Steveninck et al. 1986; Fairhall et al. 2001; Harris et al. 1999; Reisenman et al. 2003; Single et al. 1997). In this work, we focused on the adaptive temporal integration of motion signals in H1. We related the well-known adaptation of the H1 velocity impulse response observed in simple adapt-and-probe paradigms (Borst and Egelhaaf 1987; Borst et al. 2003; de Ruyter van Steveninck et al. 1986; Reisenman et al. 2003) to the narrowing of the velocity STA observed in experiments using random motion stimuli (Borst et al. 2005), through adaptation of the motion detector time constant τ_H. Our model simulations indicated that adaptation of τ_H does not optimize the system's overall information rate. In addition, we found that H1 information rates are not constant over our range of stimulus conditions.

REFERENCES

Bair W, Movshon JA. Adaptive temporal integration of motion in direction-selective neurons in macaque visual cortex. *J Neurosci* 24: 7305–7323, 2004.

Bialek W, de Ruyter Van Steveninck RR. Features and dimensions: motion estimation in fly vision http://arxiv.org/abs/q-bio/0505003.

Borst A. Noise, not stimulus entropy, determines neural information rate. *J Comput Neurosci* 14: 23–31, 2003.

Borst A, Egelhaaf M. Temporal modulation of luminance adapts time constant of fly movement detectors. *Biol Cybern* 56: 209–215, 1987.

Borst A, Egelhaaf M, Haag J. Mechanisms of dendritic integration underlying gain control in fly motion-sensitive interneurons. *J Comput Neurosci* 2: 5–18, 1995.

Borst A, Flanagin VL, Sompolinsky H. Adaptation without parameter change: dynamic gain control in motion detection. *Proc Natl Acad Sci USA* 102: 6172–6176, 2005.

Borst A, Haag J. Neural networks in the cockpit of the fly. *J Comp Physiol A Neuroethol Sens Neural Behav Physiol* 188: 419–437, 2002.

Borst A, Reisenman C, Haag J. Adaptation of response transients in fly motion vision. II. Model studies. *Vision Res* 43: 1309–1322, 2003.

Brenner N, Bialek W, de Ruyter van Steveninck RR. Adaptive rescaling maximizes information transmission. *Neuron* 26: 695–702, 2000a.

Brenner N, Strong SP, Koberle R, Bialek W, de Ruyter van Steveninck RR. Synergy in a neural code. *Neural Comput* 12: 1531–1552, 2000b.

de Ruyter van Steveninck RR, Zaagman WH, Mastebroek HAK. Adaptation of transient responses of a movement-sensitive neuron in the visual system of the blowfly Calliphora-Erythrocephala. *Biol Cybern* 54: 223–236, 1986.

de Ruyter van Steveninck RR, Lewen GD, Strong SP, Koberle R, Bialek W. Reproducibility and variability in neural spike trains. *Science* 275: 1805–1808, 1997.

Egelhaaf M. Motion detection and visual orientation: a combined behavioral, neuronal and theoretical analysis of the fly's visual system. In: *Neural Mechanisms of Behavior*, edited by Erber J, Menzel R, Pflüger HJ, and Todt D. Stuttgart, Germany: Georg Thieme Verlag, 1989.

Egelhaaf M, Borst A. Transient and steady-state response properties of movement detectors. *J Opt Soc Am A* 6: 116–127, 1989.

Egelhaaf M, Reichardt W. Dynamic response properties of movement detectors: theoretical analysis and electrophysiological investigation in the visual system of the fly. *Biol Cybern* 56: 69–87, 1987.

Fairhall AL, Lewen GD, Bialek W, de Ruyter Van Steveninck RR. Efficiency and ambiguity in an adaptive neural code. *Nature* 412: 787–792, 2001.

Flanagin VL. *Dynamic Adaptation in Fly Motion Vision* (PhD thesis). München, Germany: Department of Systems and Computational Neurobiology, Max-Planck Institute of Neurobiology, 2006.

Gabbiani F, Krapp HG, Koch C, Laurent G. Multiplicative computation in a visual neuron sensitive to looming. *Nature* 420: 320–324, 2002.

Haag J, Borst A. Encoding of visual motion information and reliability in spiking and graded potential neurons. *J Neurosci* 17: 4809–4819, 1997.

Haag J, Denk W, Borst A. Fly motion vision is based on Reichardt detectors regardless of the signal-to-noise ratio. *Proc Natl Acad Sci USA* 101: 16333–16338, 2004.

Harris RA, O'Carroll DC, Laughlin SB. Adaptation and the temporal delay filter of fly motion detectors. *Vision Res* 39: 2603–2613, 1999.

Keppel G, Wickens TD. *Design and Analysis : A Researcher's Handbook.* Upper Saddle River, NJ: Pearson Prentice Hall, 2004.

Lindemann JP, Kern R, van Hateren JH, Ritter H, Egelhaaf M. On the computations analyzing natural optic flow: quantitative model analysis of the blowfly motion vision pathway. *J Neurosci* 25: 6435–6448, 2005.

Paninski L. Convergence properties of three spike-triggered analysis techniques. *Network: Comput Neural Syst* 14: 437–464, 2003.

Poggio T, Reichardt W. Considerations on models of movement detection. *Kybernetik* 13: 223–227, 1973.

Reichardt W. Autocorrelation, a principle for the evaluation of sensory information by the central nervous system. In: *Sensory Communication*, edited by Rosenblith WA. New York: John Wiley, 1961, p. 303–317.

Reichardt W. Evaluation of optical motion information by movement detectors. *J Comp Physiol A* 161: 533–547, 1987.

Reisenman C, Haag J, Borst A. Adaptation of response transients in fly motion vision. I. Experiments. *Vision Res* 43: 1291–1307, 2003.

Single S, Borst A. Dendritic integration and its role in computing image velocity. *Science* 281: 1848–1850, 1998.

Single S, Haag J, Borst A. Dendritic computation of direction selectivity and gain control in visual interneurons. *J Neurosci* 17: 6023–6030, 1997.

Strong SP, Koberle R, de Ruyter van Steveninck RR, Bialek W. Entropy and information in neural spike trains. *Phys Rev Lett* 80: 197–200, 1998.

Torre V, Poggio T. A synaptic mechanism possibly underlying directional selectivity to motion. *Proc R Soc Lond B* 202: 409–416, 1978.