Available online at www.sciencedirect.com
ScienceDirect
Journal of Mathematical Psychology 51 (2007) 183–197
Journal of Mathematical Psychology
www.elsevier.com/locate/jmp

Detection of visual stimuli in correlated noise

U. Mortensen*

FB Psychologie und Sportwissenschaften, Institut III, Westfälische Wilhelms Universität, Germany

Received 17 November 2004; received in revised form 31 January 2007
Available online 6 April 2007

Abstract

The hypothesis of detection by temporal probability summation (TPS) as characterised in Mortensen [(2007). An analysis of visual detection by temporal probability summation, submitted for publication] is critically evaluated, considering data from an experiment of Roufs and Blommaert [(1981). Temporal impulses and step responses of the human eye obtained psychophysically by means of a drift-correcting perturbation technique. *Vision Research*, *21*, 1203-1221], who derived the impulse and the step response for a sustained type of channel. The assumed approximate linearity of the channel is discussed with respect to recent findings from neurophysiological investigations. The data are shown not to be compatible with the TPS-postulate. Further, a model of a cell assembly, consisting of a homogeneous set of neurons, is presented that allows for a natural interpretation of random fluctuations in case of temporal peak detection (TPD). The model may be discussed with respect to TPS as well as to TPD; in any case, the model allows to integrate some results concerning the effects of attentional focussing on the detection process.

1. Introduction

In Mortensen (2007), detection by temporal probability summation (TPS) was discussed and contrasted to detection by temporal peak detection (TPD). If detection is by TPS, the stimulus is detected if, within a certain time interval $J = [0, T]$, the maximum of the activity of the detecting visual "channel" reaches or even exceeds a certain level. If detection is by TPD, the stimulus is detected if the maximum of the mean response to the stimulus reaches or exceeds a critical level. A psychometric function ψ was derived assuming that the noise is Gaussian, stationary, and can be characterised by the value of the second spectral moment λ_2, meaning that an almost arbitrary autocorrelation function of the noise is allowed.

In this paper, the use of this psychometric function will be illustrated with respect to data from Roufs and Blommaert (1981). Although somewhat of age, these data are most suited for a test of the hypothesis put forward by Watson (1982), namely that TPD is a special, though unlikely case of TPS. In his criticism of the TPD-assumption Watson referred only to an impulse response for transient channels, proposed by Roufs et al., but not to the step response for sustained channels. While the data for impulse responses can indeed be approximated by a TPS-model (implying, however, biased estimates of the fitted impulse response), there is no way to explain the data for the step input in terms of a TPS-model. The point here is not that detection is never by TPS; rather, it seems that subjects may be able to choose among the alternatives of detection by TPS or by TPD, depending on instructions and experimental conditions.

If detection is by TPD, it should be possible to define a corresponding psychometric function. This may be done introducing a random variable, say η, and postulate, for instance, that detection occurs if $g_{\max} + \eta > S$, $g_{\max}$ the maximum of the mean response to the stimulus, and S some internal threshold. The question now is what η represents. Since the detecting channel may be conceived as a cell assembly (Harris, 2005), a model of such a population of neurons will be presented that is not only more specific than the channel model underlying the TPS model employed to characterise ψ, but allows to identify η with a randomly varying activity component that is due to inputs from other assemblies. The numerical evaluation of the model suggests that $g_{\max}$ and η are inversely related.

*Fax: +49 251 83 39105.
E-mail address: mortens@psy.uni-muenster.de.

0022-2496/$ - see front matter © 2007 Elsevier Inc. All rights reserved.
doi:10.1016/j.jmp.2007.02.002

This relation allows to connect the results of this paper with results from experiments on the effect of attentional focussing, namely noise depression (Yeshurun & Carrasco, 1998, 1999).

The assumption of linearity of the detecting channel is not essential for a general discussion of TPS versus TPD. The sustained channel probed by Roufs et al. will, in general, be a nonlinear system, which may, however, be approximated by some suitably chosen linear system. In fact, the data provided by Roufs et al. can be described perfectly well with respect to such a linear approximation; the discussion of a nonlinear system with respect to these data would not make much sense since the data would not allow a sufficient specification of the parameters characterising the nonlinearity. The possibility of a linear approximation may not appear to be trivial in the light the neurophysiological results, which will be briefly reviewed in the following.

Zaghloul, Boahen, and Demb (2003) show that the classical view, according to which the retinal ON and OFF cells are driven by similar presynaptic circuits, does not hold; instead, the circuits are quite asymmetrical implying asymmetric contrast sensitivity as observed in the spiking behaviour. The spike generation is often modelled in terms of the linear–nonlinear Poisson (LNP)-model (Paninski, 2004), where, however, the refractory period is not taken into account. The model accounts well for spike rates. However, increasing the contrast reduces the sensitivity of the linear filter (Zaghloul, Boahen, & Demb, 2005); for each contrast value, a different linear filter has to be estimated. This is a problem when the input is a natural stimulus with continuously changing contrasts. On the next level, the neurons exhibit strong nonlinear mechanisms of adaptation (see, e.g. Baccus & Meister, 2002); the influence of cortical inputs was already characterised in context with the role of noise (Wolfart, Debay, Le Masson, Destexhe, & Bal, 2005).

V1 simple and complex cells are often taken as building blocks for psychophysical models. According to a characterisation by Hubel and Wiesel (1962) simple-cells are meant to be defined by (i) a division into excitatory and inhibitory regions, (ii) summation within these two parts, (iii) an antagonism between excitatory and inhibitory regions, and (iv) to allow a prediction of responses to stationary or moving spots of various shapes from a map of the two areas. If one of these attributes cannot be found in a neuron it is meant to be a complex cell. Spatiotemporal summation in simple-cells is assumed to be linear; this assumption was employed in Watson's (1987) psychophysical model. As Carandini et al. (2005, p. 10583) say: "The simple-cell definition offers so much that we are reluctant to ask whether it really works." While a number of studies showed that the spatiotemporal receptive field of simple-cells predicts the optimal orientation and spatial frequency of sinusoidal gratings of the neuron, nonlinearities showed up (a linear summing stage, followed by a nonlinear stage). However, the relative magnitude of responses to non-optimal stimuli is poorly predicted, i.e. there is an overestimation of the bandwidth of orientation and spatial frequency tuning. Further, there is a nonlinearity resulting from response saturation for high contrasts. Heeger (1992) proposed a model according to which a linear first part is postulated, followed by a nonlinear part defined by half-squaring the response of the linear part, and introducing a divisive inhibition from all other neurons whose receptive fields cover the same part of the visual field. This divisive operation gives rise to contrast normalisation and was employed for the purposes of psychophysics by Watson and Solomon (1997). Carandini et al. (2005) point out that the dichotomy of simple and complex cells may not hold; the authors argue that a given cell may be positioned on a continuum from simple to complex.

Complex cells are considered nonlinear units; Movshon, Thompson, and Tolhurst (1978a, 1978b) characterised them in terms of a static nonlinearity, a model which was replaced by the sandwich model of Carandini, Mechler, Leonhard, and Movshon (1996); according to this approach, a neuron can be a low pass linear filter, a rectification stage, followed by high pass linear filter. One method to investigate the properties of complex cells is the spike-triggered covariance (STC) analysis of the receptive field, exposed to either (spatial) white noise or natural images. According to this analysis, complex cells can be approximated by an oriented Gabor function; their function can, for the cat, be described by the energy model of Adelson and Bergen (1985). For the monkey, additional excitatory and suppressive effects have been found (Rust, Schwartz, Movshon, & Simoncelli, 2005). Further, there appear to exist nonlinear effects resulting from contextual modulation by stimuli outside the classical receptive field (Fitzpatrick, 2000; Freeman, Durand, Kiper, & Carandini, 2002), and natural images are more effective in driving these cells (Touryan, Felsen, & Dan, 2005). Vinje and Gallant (2000) concluded from their data that linear temporal filters cannot account for the temporal nonlinearities; David, Vinje, and Gallant (2004) arrived at similar conclusions.

So on the level of individual neurons linear models of neuronal "channels" appear to be quite inadequate. At a psychophysical level, however, linear models can often be fitted to detection data quite well. Surely these models have to be understood as representing linearisations of nonlinear mechanisms. If detection is assumed to be by a cell assembly, one may additionally assume that detection is that of a modulation or perturbation of the ongoing activity of the assembly. It is this modulation, not the detailed mechanisms controlling the activity of a single neuron, that can be approximated by a linear mechanism.

Overview: In Section 2, the TPS-model proposed in Mortensen (2007) and the notion of TPD is briefly presented; in Section 2.3 in particular the data of Roufs et al. will be discussed. In Section 3, some channel models are presented. In Section 3.1, models assuming additive and multiplicative noise are compared. In Section 3.2

a population model is presented which yields the interpretation for the random variable η required in TPD-models.

2. Temporal probability summation and temporal peak detection

2.1. Detection by TPS

This detection model is based on the following three assumptions:

(A1) Let $J=[0,T]$ be the time interval of a trial. The sample paths X_t of the stochastic process that represents the activity of the process in the detecting channel satisfy, for each $t\in J$, the condition $X(t)=g(t)+\xi(t)$, with $g(t)=E[X(t)]$ and $\xi(t)$ the value of a sample path of a stochastic process representing noise.

(A2) The noise can be represented by a stationary Gaussian process with autocorrelation function $R(\tau)$ and second spectral moment $\lambda_2=R''(0)=d^2R(\tau)/d\tau^2|_{\tau=0}$, with $0<\lambda_2<\infty$.

(A3) The stimulus is detected if $X^+=\max_{t\in J}X(t)>S$; the psychometric function is defined as $\psi(c)=1-P(X^+\leqslant S|c)$.

Comments: λ_2 and S are free parameters, as are parameters which could characterise the deterministic function g. Intuitively, λ_2 represents the speed of fluctuations of $\xi(t)\in\xi_t$: for small λ_2 the fluctuations are slow, for large λ_2 the fluctuations are fast (i.e. for $\lambda_2\to\infty$, $R(\tau)$ becomes a Dirac function).

Applying results from the theory of extreme values of dependent variables, one then finds that for "large" values of S (e.g. $S>3$) the psychometric function is given by the expression

$$\psi(c)=1-\exp\left[-\frac{\sqrt{\lambda_2}}{2\pi}\int_0^T\exp\left[-\frac{(S-g(t;c))^2}{2}\right]dt\right] \tag{1}$$

(Mortensen, 2007). A second expression, derived by Ditlivsen (1971), employing different principles and allowing for non-stationary noise, was also presented in Mortensen (2007). The predictions based on Ditlivsen's expression are identical to those by Eq. (1), so it is sufficient to concentrate on the latter.

Free parameters and the probability of a false alarm: The probability of a false alarm is given by

$$\psi(0)=1-\exp\left[-\frac{T\sqrt{\lambda_2}}{2\pi}\exp(-e^{-S^2/2})\right]. \tag{2}$$

For given values of $\psi(0)$ and T, S is a function of λ_2 or, vice versa, λ_2 is a function of S, so there is only one free parameter. For simplicity the possibility of false alarms due to guessing has been neglected here.

2.2. Detection by TPD

If detection is by TPS, ψ is determined by the complete course of g during the trial as represented by the interval $J=[0,T]$. If detection is by TPD, the effect of the stimulus can only be expressed the value of $g_{\max}(c)=\max_{t\in J}g(t,c)$. Let X^* be the decision variable; X^* is a function of $g_{\max}(c)$ and a random variable η, and detection occurs if $X^*>S$, S being some threshold value. There are (at least) two alternatives for a further specification of X^*:

$$X^*=\begin{cases} g_{\max}(c)+\eta & \text{(a)},\\ (g_{\max}(c)+x_0)/\eta, \quad \eta\neq 0 & \text{(b)}.\end{cases} \tag{3}$$

These definitions of X^* requires some *Comments*:

1. *The role of time*: Activity is always a process extended in time, but no time variable occurs in the definition of X^*. A straightforward interpretation e.g. of $X^*=g_{\max}(c)+\eta$ would be to say that detection may occur at the time $t_{\max}$, $g_{\max}=g(t_{\max})$, and that $\eta=\xi(t_{\max})$. For instance, Tyler and Chen (2000) refer to $X^*=g_{\max}+\eta$ as the *instantaneous internal response* and thus appear to have adopted this interpretation (they write r instead of X^*, though). This interpretation is, however, problematic, since it immediately leads to the question how the visual system singles out the activity precisely at $t_{\max}$. The assumption that detection depends—apart from the distribution of η—only on the value of $g_{\max}$ does not mean that detection can only occur at $t_{\max}$, the time at which g assumes its maximum $g_{\max}$. The event $X^*>S$ implies then either $g_{\max}(c)>S-\eta$ or $g_{\max}>S\eta-x_0$, depending on which alternative in (3) is chosen. This implies that g is larger than a certain critical level for some small interval of time around $t_{\max}$. It follows that the shape of g in the neighbourhood of $t_{\max}$ may indeed influence detection, so that the claim that the stimulus has an influence detection only via $g_{\max}$ may be too radical, and in this respect the assumption of TPD does represent an approximation. However, the following discussion of detection data will show that the effect of g being above a critical level for some finite amount of time may be negligible.
2. *The interpretation of* η: Corresponding to the foregoing interpretation of TPD it does not appear to make sense to relate η to the value of $\xi(t_{\max})$. It seems to be more plausible to relate η to the background activity of the detecting channel, which may, to a good degree of approximation, be about constant during a trial. This possibility will be explored in greater detail in Section 3. It should also be realised that there is no need to associate steep psychometric functions with detection by TPD. The steepness of ψ depends on the parameters of the distribution function of η, and the definition of detection by TPD does not imply any restrictions of these parameters; see Sections 3.3 and 4.

3. *Additive noise*: The alternative (3a) corresponds to the standard assumption in psychophysics, namely that the activity is additively composed of a deterministic part due to the stimulus, and another part representing noise. Since it is assumed that detection occurs if $X^* > S$, the noise is implicitly assumed to lift the activity above the threshold; for a given value of $g_{\max}$, detection is more likely to occur for a larger value of η. However, if η is related to the background activity of the detecting channel, it is conceivable that the response to the stimulus gets buried in this activity. This possibility is catered for in Eq. (3b). Here, detection is more likely to occur for a smaller value of η, and $x_0 > 0$ allows for false alarms. A motivation for considering this alternative will be given in Sections 3 and 4.

2.3. Comparison with data: brief pulses and step stimuli

Roufs and Blommaert (1981) introduced what they called the perturbation method, which allows to measure the temporal signal response to stimuli, in particular to brief (5 ms) pulses and to step inputs, provided the assumption of peak detection holds. A sketch of the method is provided in the Appendix. Here, we concentrate on stimuli defined as small, circular discs; this allows to neglect any possible spatial probability summation effects. Roufs et al. estimated the impulse and the corresponding step response. They fitted the function

$$h(t) = b(at)^p \exp(-at), \tag{4}$$

to the empirically determined impulse response, with estimates $\hat{a} = \frac{1}{12.66} \approx 0.079$, $\hat{p} = 3$. The normalisation with $\hat{b} = 0.742$ implies that $\max_t h(t) = 1$, see Fig. 1. The step response corresponding to (4) is given by

$$H(t) = \int_0^t h(u)\,du = \frac{b(\Gamma(1+p) - \Gamma(1+p, at))}{a}, \tag{5}$$

with $\Gamma(x) = \int_0^\infty t^{x-1}e^{-t}\,dt$, and $\Gamma(x,y) = \int_y^\infty t^{x-1}e^{-t}\,dt$.

Let $\hat{h}$ and $\hat{H}$ be estimates of h and H, found by employing the perturbation method. If detection is by TPD, then (34) in the Appendix implies that the step response can be predicted from the impulse response data by integration, that is,

$$\hat{H}(t) = \int_0^t \hat{h}(u)\,du, \tag{6}$$

neglecting any measurement errors. If this relation between $\hat{h}$ and $\hat{H}$ does not hold we can conclude that detection is not by TPD, provided the measurement error is sufficiently small. Then TPS is a possible alternative. With respect to Eq. (1), the function g is defined by Eq. (28) in the Appendix. If detection is by TPS, the threshold amplitudes c_0 for the stimulus alone and c_τ for the superposition of the stimulus and a brief pulse, presented with a delay τ, are determined as solutions of $\psi(c_0) = 0.5$ and $\psi(c_\tau) = 0.5$, where ψ is given by (1). These expressions for ψ imply that the estimates $\hat{h}$ and $\hat{H}$, computed on the basis of the values of c_0 and c_τ, will not be related by the linear operation (6). Instead, the data would suggest that the relation between the impulse and the step responses is nonlinear.

The results are presented in Fig. 2. In panel (a), the impulse response h, defined in (4), is shown together with the estimated impulse response given that detection is by TPS (black squares). If only the impulse response had been determined, it would be difficult to decide whether detection is by TPS or by TPD; the function (4) can also be fitted to the data (see Fig. 2a) and be interpreted in terms of TPD. TPS would then generate a distorted version of the impulse response. However, the results concerning the step response shown in Fig. 2b do not allow for such an interpretation. The solid squares represent the step response if detection is by TPS, and the open squares represent the step response predicted from the estimated impulse response (solid squares in Fig. 2a) making use of (6). To allow for an interpretation in terms of TPS, the step response estimated on the basis of the impulse response, as

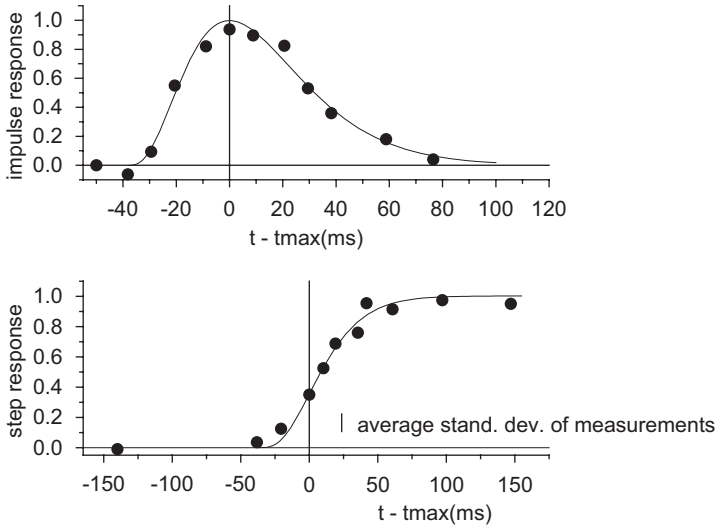

Fig. 1. The data of Roufs and Blommaert (1981); $t_{\max}$ is the time at which the impulse response (at top) assumes its maximum value.

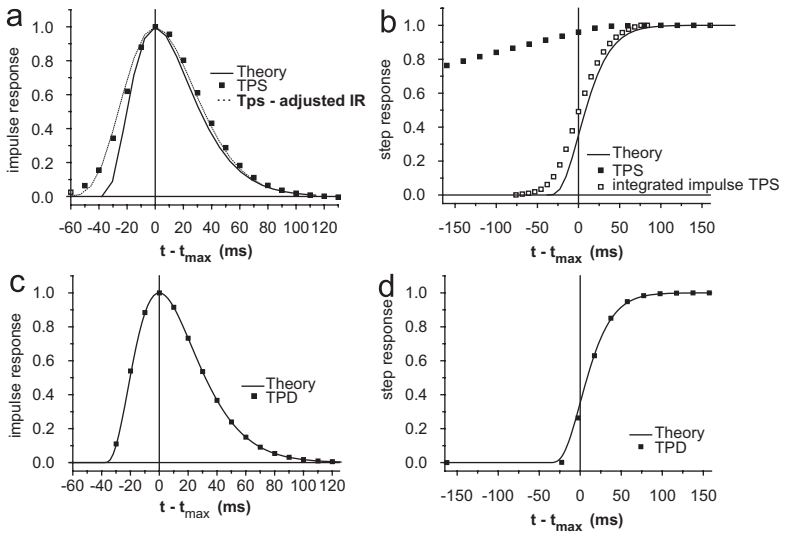

Fig. 2. Predictions of impulse and step responses in case of TPS and TPD. "Theory" refers to the analytic impulse (4) and step response (5) with parameters as determined by Roufs and Blommaert (1981): $\hat{a} = 0.07899$. $\hat{p} = 3$. Panel (a) shows the estimate of the impulse response in case of TPS (solid squares), and the impulse response (4) with parameters estimated to fit the TPS estimates: $\hat{a} = 0.115$ and $\hat{p} = 8.75$ (dotted line). Panel (b) shows the estimates in case of TPS (solid squares), and the step response resulting from integrating the TPS estimate of the impulse response (open squares). The predictions for TPS do not depend on the value of λ_2; see text for further discussion. Panels (c) and (d) show estimates and theory in case of TPD.

estimated when detection is by TPS, should not deviate from the curve indicated by the open squares (integrated impulse response TPS). The differences between predicted, estimated and "true" step response are sufficiently clear to reject the hypothesis of TPS. Interestingly, the value of λ_2 is irrelevant for the estimation of h and H: the estimates $\hat{h}$ and $\hat{H}$ are identical for different values of λ_2. Obviously, expression (34) in the Appendix implies that the effect of a particular value of λ_2 is cancelled. The predictions of the measurements under TPS conditions in Fig. 2 are determined by the analytic expression for the psychometric function (1), not by particular values of the parameters.

Fig. 2, panels (c) and (d), shows the results when detection is by TPD. No attempt was made to incorporate measurement errors. The results for the case of TPD show that the perturbation method provides bias-free estimates of h and g. To summarise, one can say that when the step response can be predicted from the impulse response via (6), the data support the hypothesis of TPD, but not TPS. In this case the perturbation method yields bias-free estimates of h and g.

3. Channel models: illustrations

The assumption underlying the expression (1) is that detection is by a single channel without further specific assumptions concerning the channel. A channel may be conceived as some assembly of neurons, rather than a single neuron. For instance, a single neuron may not be able to provide sufficiently precise information about the stimulus. As an example, consider the identification of orientations. A single neuron responds to a large range of orientations, whereas psychophysical experiments indicate a much higher precision of orientation identification and discrimination (Westheimer, Shimamura, & McKee, 1976). Paradiso (1988) and Seung and Sompolinsky (1993) proposed models for the identification and discrimination of orientations according to which this precision is arrived at by appropriately combining the activity of neurons in certain populations or assemblies of neurons; for a generalisation of this model see Jazayeri and Movshon (2006). For the data provided by Roufs et al. no model for the readout, i.e. the interpretation of the data, as required e.g. when the orientation of a line element has to be identified, so simply the activation of a population of a set of neurons need to be considered.

One may generally assume that detection of a stimulus results from monitoring of the responses of a set of sensory neurons by some population of other neurons, e.g. a population of the prefrontal cortex (PFC) (Kim & Shadlen, 1999). The activation of PFC neurons thus depends on a sufficient activation of the sensory neurons, so it may suffice to model the activation of the latter. For the present purpose it is not necessary to precisely define which sensory neurons are meant: retinal ganglion cells, LGN cells, V1 or V2 cells, etc. In any case, the input to the neurons will be stochastic. The input may thus be described in terms of a mean-value function plus some random function.

In the models of Paradiso and Seung et al. mentioned above no explicit reference to spike trains or spike rates is made. In the following Section 3.1 this in a certain sense loose way of specifying the activity is kept, and only the notions of additive and multiplicative noise are illustrated. Further, a possible way of specifying the random variable η in Eq. (3a) is indicated. In Section 3.2 activity will be defined in terms of spike rates; this model leads to an interpretation of η corresponding to Eq. (3b).

3.1. Simple models: additive and multiplicative noise based on the integrate-and-fire model

According to assumption (A1) the activity of the detecting channel is given by $X(t) = g(t) + \xi(t)$, with $g(t) = E(X(t))$ the mean value function and ξ a trajectory of Gaussian noise. This representation of the activity does not yet mean that the noise is additive, it only means that $\xi(t) = X(t) - g(t)$ is the difference between the actual and the mean activity. In this section, a simple model for X is provided that allows to specify ξ as being either additive or multiplicative noise. To this end, X is assumed to be the solution of a stochastic differential equation (SDE) of the form

$$dX(t) = -(kX(t) - g(t) - \eta_0)\,dt + \sigma^2(t)\,dW(t). \tag{7}$$

Here, $k>0$ is a constant reflecting the time constant of the system, σ^2 defines the variance of the fluctuations, and $dW(t)$ is the formal derivative of a trajectory of the standard Wiener process[1] which is a formal representation of Brownian motion; although $dW(t)$ may be interpreted as Gaussian white noise (for details see e.g. Kloeden & Platen, 1992), the stochastic process having the X_t as sample paths is not a white noise process. If σ^2 is independent of X, the noise is additive, otherwise the noise is multiplicative (Honerkamp, 1990, p. 82).

η_0 is a constant within a trial, but may vary randomly between trials, and may be interpreted as representing an activity that results from input from other assemblies which is not generated by the stimulus; a more explicit reason for this interpretation will be given in Section 3.2. The effect of the stimulus is represented by g. Given (7), the mean value function of X is known to be given by

$$dm(t) = -km(t) + g(t) + \eta_0, \tag{8}$$

which is a deterministic equation well known to have the solution

$$\begin{aligned} m(t) &= m_0 e^{-kt} + \int_0^t e^{-k(t-\tau)}(g(\tau) + \eta_0)\,d\tau \\ &= \eta e^{-kt} + \int_0^t e^{-k(t-\tau)} g(\tau)\,d\tau + \frac{\eta_0}{k}(1 - e^{-kt}) \end{aligned} \tag{9}$$

(Arrowsmith & Place, 1982). Fig. 3 shows the responses of the channel as defined by (7), for (i) additive and (ii) multiplicative noise. The mean value function g was computed according to (9); g is the same for additive and multiplicative noise. Since additive noise is independent of g, the response appears as being less noisy the larger the amplitude of the mean value function. As can be seen from (9), $m(t)$ does not only depend on the mean response g to the stimulus, but also on $\eta_0(1 - e^{-kt})/k \to \eta_0/k$; thus one may say that for given value of the amplitude c and the threshold S the stimulus will be detected if the random variable η_0 assumes a value such that $\max_t m(t) > S$, S some threshold value. It turns out that except for small values of t and $\eta_0 = 0$, the functions $g(t)$ and $m(t)$ differ only by a proportionality constant $\alpha_1 = \max_t g(t)/\max_t m(t)$, so that $m(t) \to \alpha_1 g(t)$ for increasing values of t. For $\eta_0 > 0$, one has $m(t) \to \alpha_1 g(t) + \alpha_2$, with $\alpha_2 = \eta_0/k$. One is thus lead to a definition of the decision variable X^*:

$$X^* = \alpha_1 \max_{t \in J} g(t) + \eta, \quad \eta = \eta_0/k. \tag{10}$$

The convergence of $m(t)$ towards $\alpha_1 g(t) + \alpha_2$ is very fast, the error resulting from equating $m(t)$ and $\alpha_1 g(t) + \alpha_2$ is negligible. Since g can always be determined only up to a proportionality constant, this equation is actually equivalent to the definition of peak detection in Eq. (3a). The psychometric function can be defined in terms of the distribution function of the random variable $\eta = \eta_0/k$, reflecting the input from cell assemblies other than the one that signals the presence of the stimulus. The relation between $g(t)$ and $m(t)$ is illustrated in Fig. 4. Thus one has actually two kinds of "noise": one component is of diffusive type, which enters via the term dW_t in Eq. (7), the other one, represented by the random variable η, is due to input from other assemblies. It is this type of noise that defines the psychometric function.

3.2. The mean spike rate of a cell assembly

The following notions are taken from Gerstner (2000), see also Gerstner and Kistler (2002). Let $n_a(t, t+\Delta t)$ be the number of active (firing) neurons in a set of N neurons, in the interval $[t, t+\Delta t)$. The population activity is defined by

$$A(t) = \lim_{\Delta t \to 0} \frac{1}{\Delta t} \frac{n_a(t, t+\Delta t)}{N} = \frac{1}{N} \sum_{j=1}^{N} \sum_f \delta(t - t_j^{(f)}), \tag{11}$$

where δ is the Dirac function and $t_j^{(f)}$ the firing time of the jth neuron. The network is homogeneous if all neurons have the same input resistance R and the same membrane time constant τ_m. The input current I_i of the ith neuron is given by

$$I_i = \sum_{j=1}^{N} \sum_f w_{ij} \alpha(t - t_j^{(f)}) + I^{\text{ext}}(t), \tag{12}$$

w_{ij} being the synaptic coupling to the jth neuron, and $\alpha(t - t_j^{(f)})$ the time course of postsynaptic current generated by an input spike at $t_j^{(f)}$. Here $I^{\text{ext}}(t)$ is assumed to be defined by the mean response of the sensory neurons which are being monitored by the population of detecting neurons. If $w_{ij} = \kappa_0/N$, with κ_0 a free parameter, one has homogeneous all-to-all coupling; for $\kappa_0 > 0$ the coupling is excitatory, for $\kappa_0 < 0$ it is inhibitory, whereas $\kappa_0 = 0$ implies that all neurons are independent.

Let us consider integrate-and-fire neurons. The membrane potential of the ith neuron is then given by

$$\tau_m \frac{du_i}{dt} = -u_i + RI_i(t), \quad i = 1, \ldots, N, \tag{13}$$

[1] The standard Wiener process is a Gaussian process with independent increments, $E[W(t)] = 0$ for $t \geqslant 0$, $\text{Var}[W(t) - W(s)] = t - s$, for $0 \leqslant s \leqslant t$, or $\text{Kovar}[W(s), W(t)] = \min(s, t)$. The Wiener process represents Brownian motion for the special case that no frictional forces exist.

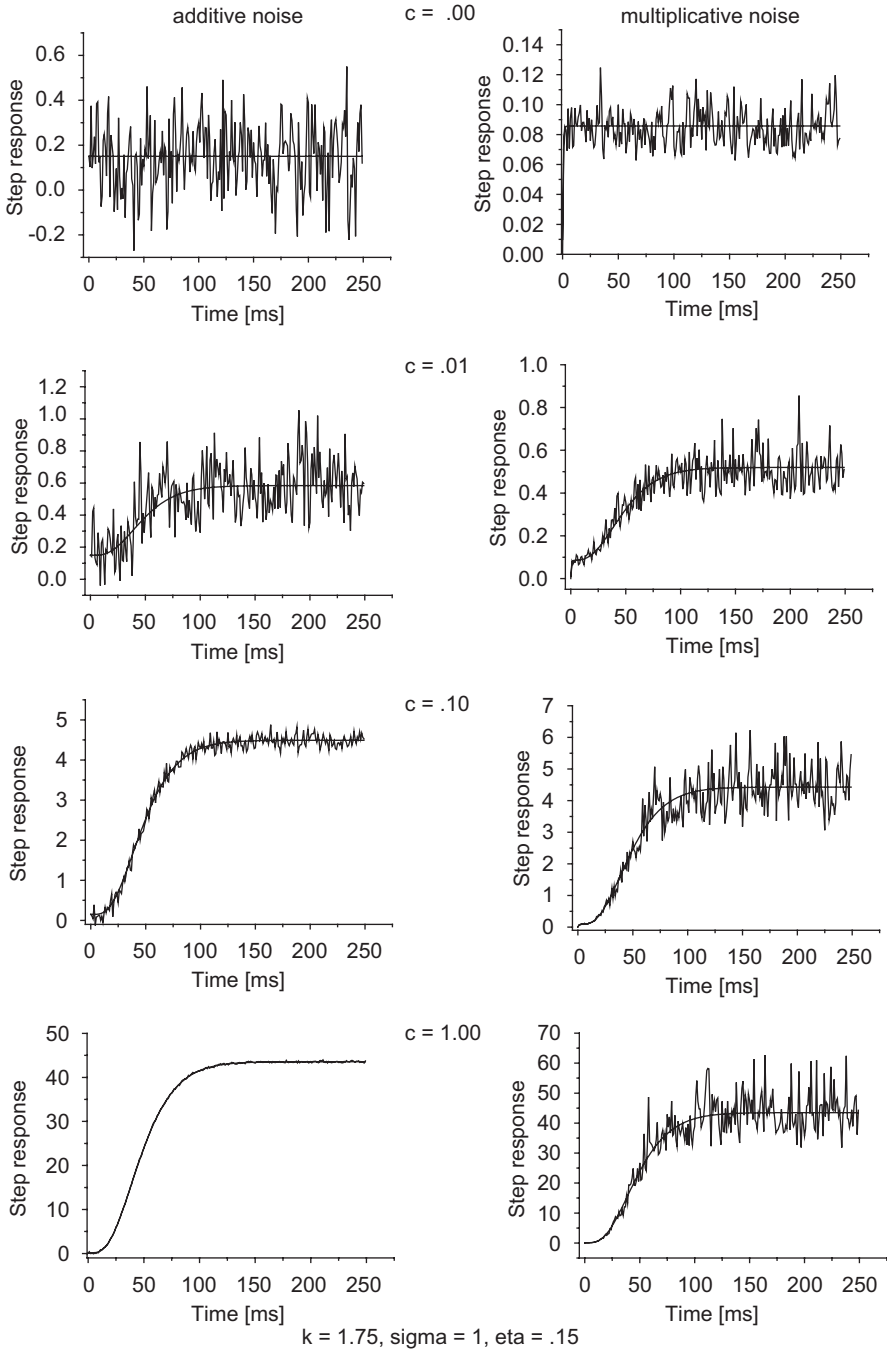

Fig. 3. Additive and multiplicative noise for different values of stimulus contrast. In all cases $k = 1.75$, $\sigma = 1$ and $\eta_0 = 0.15$.

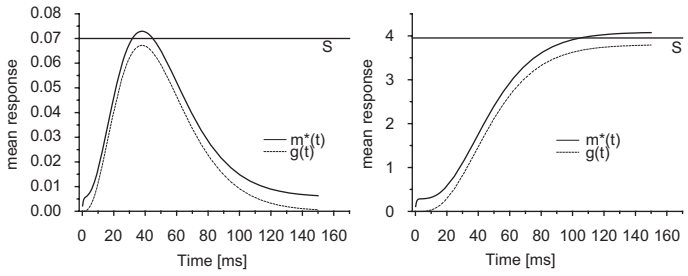

Fig. 4. Peak detection: $m^*(t) = \alpha_1 m(t) + \alpha_2$, where $m(t)$ is given by (9) and $\alpha_1 = \max_t g(t)/\max_t m(t)$, $\alpha_2 = \eta_0/k$; the functions $m^*(t)$ and $g(t)$ differ only by α_2. See text for further explanation.

where R is the input resistance and $\tau_m = RC$ is the membrane time constant. If the potential reaches a threshold value ϑ the membrane potential is reset to $u_r < \vartheta$. Let $n(u_0, u_0 + \Delta u)$ be the number of neurons with membrane potential $u_0 < u_i(t) \leqslant u_0 + \Delta u$ at time t. The proportion of neurons with membrane potential between u_0 and $u_0 + \Delta u$ for $N \to \infty$ is then

$$\lim_{N\to\infty}\left(\frac{n(u_0, u_0+\Delta u)}{N}\right) = \int_{u_0}^{u_0+\Delta u} p(u,t)\,du, \tag{14}$$

where $p(u,t)$ is the membrane potential density, defined as the density of membrane potentials in a large population of neurons. Furthermore, $p(u,t)$ satisfies the condition

$$\int_{-\infty}^{\vartheta} p(u,t)\,du = 1 \quad \text{for all } t \tag{15}$$

which implies that, at any moment, all neurons have a membrane potential below or equal to the threshold ϑ. The fraction of neurons with membrane potential reaching ϑ per unit of time equals $A(t)$. Since the membrane potential is reset to u_r when the threshold ϑ is reached, the membrane potential after the reset will increase at a rate proportional to the population activity $A(t)$, so one gets a term $A(t)\delta(u - u_r)$, which plays the role of a "source", see Eq. (16) below; δ is, as usual, the Dirac function.

Neurons are activated by incoming spikes via synapses. There exist different types of synapses, e.g. inhibitory and excitatory; for a more explicit characterisation see Gerstner and Kistler (2002, Section 2.4). Suppose an input spike at a synapse of type j causes a jump of the membrane potential of size v_j. The *effective spike arrival rate* for synapses of type j is v_j, where v_j may be interpreted as the mean value function of a Poisson process. For small jump amplitudes v_j one may then derive the diffusion approximation for $p(u,t)$:

$$\begin{aligned}\tau_m \frac{\partial}{\partial t}p(u,t) = &-\frac{\partial}{\partial u}\left\{\left[-u + RI^{\text{ext}}(t) + \tau_m \sum_j v_j(t) v_j\right] p(u,t)\right\} \\ &+ \frac{1}{2}\left[\tau_m \sum_j v_j(t) v_j^2\right]\frac{\partial^2}{\partial u^2}p(u,t) \\ &+ \tau_m A(t)\delta(u - u_r) + O(v_j^3);\end{aligned} \tag{16}$$

see Gerstner et al., p. 211. The firing threshold ϑ is interpreted as an absorbing boundary so that $p(\vartheta, t) = 0$, and one finds

$$A(t) = -\frac{\sigma^2(t)}{2\tau_m}\frac{\partial p(u,t)}{\partial u}\bigg|_{u=\vartheta} \tag{17}$$

with

$$\sigma^2(t) = \tau_m \sum_j v_j(t) v_j^2 \tag{18}$$

(Gerstner & Kistler, 2002, p. 211). Let

$$\alpha(t) = \tau_m \sum_j v_j(t) v_j. \tag{19}$$

In order to find $A(t)$ for given I^{ext} one has to find a solution for $p(u,t)$ and to compute $A(t)$ according to (17). This requires additional assumptions concerning $\alpha(t)$ and $\sigma^2(t)$, i.e. $v_j(t)$ and v_j. The following simplification will be adopted: we consider $p(u,t)$ for $u_r < u \leqslant \vartheta$; the term $\tau_m A(t)\delta(u - u_r)$ in (16) can then be neglected (see however, the definition of spike rate below). If one also neglects $O(w_k^3)$ and divides (16) by τ_m, after substituting (18) and (19), one arrives at the Fokker–Planck equation (FPE)

$$\begin{aligned}\frac{\partial}{\partial t}p(u,t) = &-\frac{\partial}{\partial u}\{[-u + RI^{\text{ext}}(t) + \alpha(t)]p(u,t)\} \\ &+ \sigma^2(t)\frac{1}{2}\frac{\partial^2}{\partial u^2}p(u,t).\end{aligned} \tag{20}$$

This FPE can be solved numerically provided that $\alpha(t)$ and $\sigma^2(t)$ are specified. However, for the purpose of this paper a different approach appears to be more direct. For a particular FPE, there exists a corresponding SDE, see e.g. Gardiner (1990), or Kloeden and Platen (1992). With respect to (20) one gets

$$\begin{aligned}du(t) = &[-a_0 u(t) + \alpha_0 g(t) + \alpha(t)]\,dt \\ &+ \sigma^2(t)\,dW(t), \quad u_r < u(t) \leqslant \vartheta\end{aligned} \tag{21}$$

with $a_0 = 1/\tau_m, \alpha_0 = R/\tau_m$, $g(t) = I^{\text{ext}}(t)$, and $W(t)$ being a standard Wiener process. Eq. (21) is linear and represents the case of additive noise. This equation has already the structure of Eq. (7); however, in (7) the trajectories X_t are not restricted to some interval, as are the trajectories u_t.

Given an input function $I^{\text{ext}}(t) = g(t)$ and a specification of $\alpha(t)$ and $\sigma^2(t)$, (21) can be evaluated numerically. A solution to (21) is a trajectory u_t of the stochastic process $\{u_t, t \in J\}$. The trajectories will differ from trial to trial, because of different noise sample paths. However, they share certain invariant features that are relevant for a discussion of the detection process, as will be illustrated in the following section.

3.3. Response to brief pulses

The following simplifying assumptions will be made:

(B1) $I^{\text{ext}}(t) = g(t) = cb(at)^p \exp(-at)$, if the stimulus is a brief pulse, or

$$g(t) = c\int_0^t b(a\tau)^p \exp(-a\tau)\,d\tau,$$

in case the stimulus is a step function, where a, b and p are the parameters estimated by Roufs and Blommaert (1981), see (4) and c is a factor to scale I^{ext} to the variation of the spike rate.

(B2) $\alpha(t)$ and $\sigma^2(t)$ are varying slowly if compared to $u(t)$. From the definition of $\alpha(t)$ and $\sigma^2(t)$, one may put $\sigma^2(t) = k_0\alpha(t)$ with k_0 some constant, and $\alpha(t) \approx \eta$ is a constant during $J = [0, T]$. In other words, η is a random variable varying only between trials.

Comments:

1. $I^{\text{ext}}(t) = g(t)$ is interpreted as a mean value function of the activity representing the input to the detecting population of neurons.
2. The assumption of slowly varying functions $\alpha(t)$ and $\sigma^2(t)$ is motivated by the work of Leopold, Murayama, and Logothetis (2003). These authors argued that, while the neural dynamics related to cognition and behaviour are usually fast, i.e. on a scale of milliseconds to at most a second, there also exist fluctuations on a slower scale. These reflect the influence of hemodynamic and metabolic processes of large-scale networks which contribute to the variability that is observed in particular cortical areas. If the stochastic fluctuations in the population code relevant for the detector are small relative to the mean activity generated by the stimulus, these large-scale fluctuations may correspond to the fluctuations of $\alpha(t)$ and $\sigma^2(t)$, meaning that their variation may be negligible within a trial and be slow even between trials. This may justify the approximation $\alpha(t) \approx \eta$ during a trial, and it would explain the observation that in a detection experiment one often has sequences of trials when the subject detects the stimulus, followed by sequences of trials when the subject does not detect the stimulus. This interpretation of η applies also to the random variable η introduced in Eq. (7).

If one adopts assumptions (B1) and (B2), the SDE (21) takes the form

$$du(t) = [-a_0 u(t) + \alpha_0 g(t) + \eta]\, dt + k_0 \eta\, dW(t), \quad u_r < u(t) \leqslant \vartheta. \tag{22}$$

For a given trial in a Monte-Carlo simulation, the value of $dW(t)$ is determined by a random-number generator, such that for a given value of t, $dW(t)$ is zero-mean Gaussian with variance σ^2. Above, η was introduced as a random variable. However, since only single trajectories will be considered it is not necessary to introduce random values of η. Instead, trajectories for "large" and "small" values of η will be considered in order to exemplify the influence of η on the trajectories.

Computation of spike rates:

1. *Numerical solution of the SDE*: The interval $J = [0, T]$ is sub-divided into n intervals of width $\Delta = T/n = 2^{-k}$, with $k \in \mathbb{N}$. The impulse response g (cf. (4)) derived by Roufs and Blommaert (1981) approaches 0 for $t \to 150$ ms, so $T = 150$ was chosen. A fairly realistic picture of the activity emerged for $k = 10$, implying $\Delta = 0.000976563$ and $n = T/2^k = 153\,600$ sub-intervals. This value of k is large and indicates that a very fine resolution of the time scale is necessary, in agreement with a finding of Hansel, Mato, Meunier, and Neltner (1998). The trajectories are discretised so that $du(t)/dt \approx (u((i+1)\Delta) - u(i\Delta))/\Delta$ for $t = (i+1)\Delta$. Thus $u((i+1)\Delta) = u(i\Delta) - (a_0 u(t) - \alpha_0 c g(i\Delta) - \eta)\Delta + k_0 \eta \Delta W_i$ with $\Delta W_i = W_{i+1} - W_i$, and W_{i+1} and W_i values of a Wiener process at times $(i+1)\Delta$ and $i\Delta$.
2. *Spike rate*: Let y_t be the number of times that u reaches the value ϑ within the tth interval; y_t will be referred to as the spike rate.

Recall that $u(t)$ represents the membrane potential at time t, averaged over the neurons of the population. The larger (i.e. closer to ϑ) the value of $u(t)$, the more neurons will fire at this time, and the larger the population activity $A(t)$ will be.

Mean value function: It is of interest to consider the mean value function of the spike rate. In case of the linear SDE (22) with unrestricted $u(t)$, for a given trial, the mean value function $m(t)$ is given by the deterministic differential equation

$$\frac{dm(t)}{dt} = -a_0 m(t) + \alpha_0 g(t) + \eta, \tag{23}$$

which corresponds to Eq. (8). Note that the mean value function depends on η, which is constant within trials but possibly changes between trials so $m(t)$ may have a different course in different trials. Unfortunately, the condition $u_r < u(t) \leqslant \theta$ does not allow to determine $m(t)$ from (23). However, given the chosen parameter values for a_0 and α_0 it is sufficient to consider an approximative solution taking into account the following considerations: apart from some transient effects, the function m can be approximated well by a linear transformation of g, so that $m(t) \approx \alpha_1 g(t) + \alpha_2$. It turns out that this approximation is also a first-order approximation of the mean value function m_y of the spike rate y_t, although the parameter values of α_1 and α_2 are different, so we can write

$$m_y(t) \approx \hat{m}_y(t) = \alpha_1 g(t) + \alpha_2. \tag{24}$$

The spike rate y_t thus provides an impression of $A(t)$. Fig. 5 shows some results for a pulse and a step input. The values of parameters were chosen such that the spike rate corresponds roughly to plausible values (cf. for instance Fig. 6.5 in Gerstner & Kistler, 2002, p. 218). In all simulations, the reset value u_r was put equal to zero, the threshold $\vartheta = 0.05$, the scale factor $\alpha_0 = 1.95$, $a_0 = 0.15$ and $k_0 = 1/2$. The trajectories for the pulse and corresponding step responses were generated with different values of η. Obviously, the smaller the value of η, the more pronounced the effect of the input I^{ext}. At the same time, the activity fluctuates within smaller intervals. Recall that η represents the input other than I^{ext} to the population of neurons. If this input is large, the response to $I^{\text{ext}}(t) = g(t)$ appears to get buried under the response to η.

The mean value function was fitted by eye, which is sufficient for the purpose of this paper.[2] Note that the value

[2] A more correct way of estimating the mean value function would have been to run, for a fixed value of η, the Monte-Carlo simulation of a trajectory N times, yielding N trajectories $u_t^1, \ldots, u_t^N$. An estimate of the

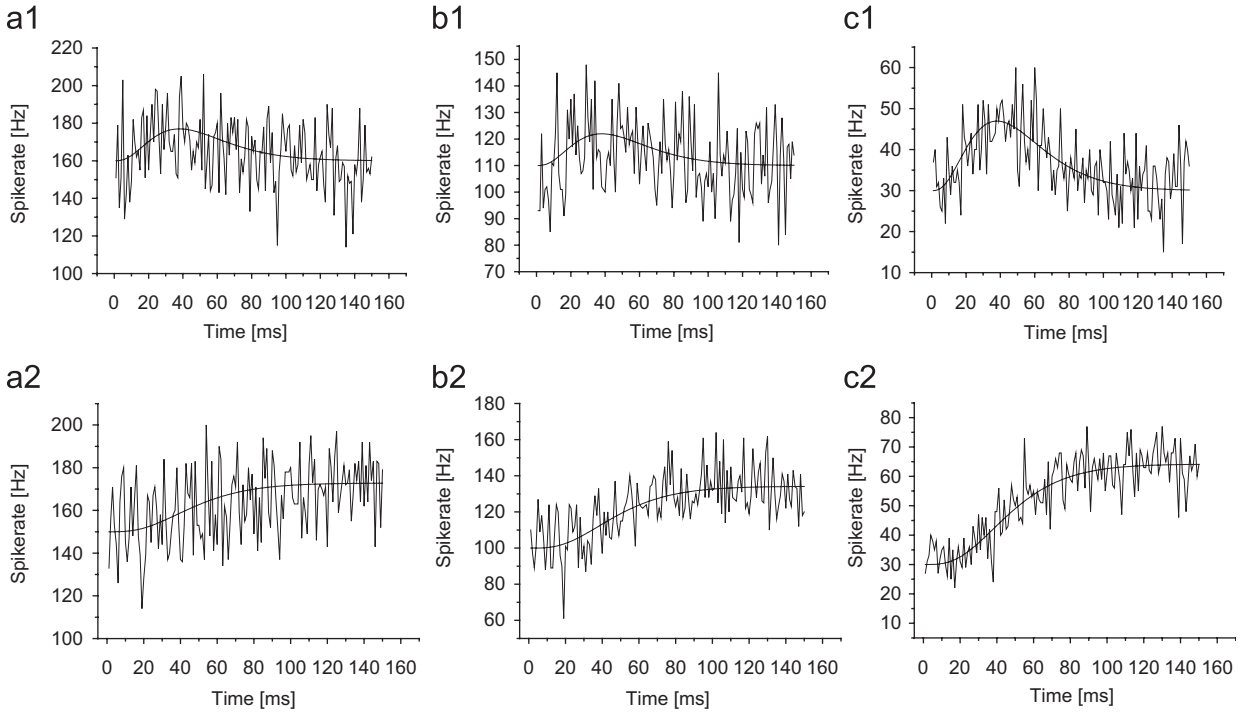

Fig. 5. Simulated spike rates in case for a brief pulse (a1), (b1), (c1), and step inputs (a2), (b2), and (c2). Threshold $\vartheta = 0.05$; (a1) and (a2): $\eta = 15.5$, $\alpha_2 = 160$ for the pulse and $\alpha_2 = 150$ for the step response; (b1) and (b2): $\eta = 7.5$, $\alpha_2 = 110$ for the pulse and $\alpha_2 = 150$ for the step response; (c1) and (c2): $\eta = 2.5$, $\alpha_2 = 30$ for the pulse as well as for the step response. For the pulse response, $\alpha_1 = 17$ in all cases, and for the step response, $\alpha_1 = 45$. For an explanation see text.

of α_1 is constant for the pulse response ($\alpha_1 = 17.0$) and for the step response ($\alpha_1 = 0.45$). For a given input type, the approximations $\hat{m}_y(t)$ for m_y thus differ only by an additive constant. Let the amplitude of the mean response be defined by

$$A_g = \max_{t\in J} m_y(t) - m_y(0) \approx \alpha_1 \max_{t\in J}[g(t)]. \tag{25}$$

The interesting finding from the calculations is that, for a given input type, A_g is independent of η to a fair degree of approximation, since $\alpha_1 =$ constant. So the mean responses $\hat{m}_y(t) = \alpha_1 g(t) + \alpha_2$ differ only by an additive constant. The variance of the spike rate is a monotone function of η, implying that $\hat{m}_y(t)$ appears to be more pronounced for smaller values of η. If one now considers detection by TPD, one is lead to the definition

$$X^* = \frac{\max_{t\in J} m_y(t)}{\eta}, \quad \eta > 0. \tag{26}$$

This definition of X^* takes into account the observation from above, namely that if the input represented by η is large, the signal $g(t)$ is likely to get buried under the response to η: a small value of η implies a large value of X^* and vice versa, and one is lead to the assumption that detection occurs if $X^* > S$, $S > 0$ some threshold value. If no stimulus is presented one has $\max_{t\in J} m_y(t) = m_y(0)$, and $P(m_y(0)/\eta > S)$ is the probability of a false alarm. Thus Eq. (26) corresponds to Eq. (3b).

Autocorrelation function: There is no explicit expression for the autocorrelation function of the spike rate. However, Eq. (21) defines, for $I^{\text{ext}}(t) \equiv 0$ and $\sigma^2(t) = \sigma^2$ being a constant, an Ornstein–Uhlenbeck process, for which the autocorrelation function is given by

$$R(\tau) = \exp(-a_0\tau)\sigma^2/2a_0$$

with $\tau \geqslant 0$. We can expect that the autocorrelation function of the spike rate will also depend on the values of a_0 and $\sigma^2 = \eta$, which means that the autocorrelation function will vary from trial to trial. However, simulations with different combinations of values of a_0 and η (not presented here) show that the effect of the value of η appears to be small and may thus be neglected for the purposes of this paper.

4. Summary and discussion

The main results may be summarised as follows:

1. It was investigated whether the TPS-model, proposed in Mortensen (2007), can cater with data that most likely reflect detection by TPD, namely those of Roufs and Blommaert (1981). To this end, the perturbation method employed by these authors was used to compute estimates $\hat{g}_\text{i}$ for the impulse response g_i and $\hat{g}_\text{s}$ for the

(*footnote continued*)
mean value function would result from averaging the u_t^j, $j = 1,\ldots,N$ at values t_k, $k = 1,\ldots,n$. However, this procedure is extremely time consuming, without giving more information about the mean value function than by fitting $\alpha_1 g(t) + \alpha_2$ as an approximation.

step response g_s. Roufs and Blommaert (1981) were able to show that the data satisfy the condition $g_s(t) = \int_0^t g_i(u)\,du$; this relation should necessarily hold if their hypotheses of (i) (approximate) linearity and (ii) detection by TPD are compatible with the data. Since the data agree with these hypotheses, they cannot be rejected. The question is whether, for a given probability of false alarm, a value of the second spectral moment λ_2 exist such that the threshold contrasts c can be computed from the condition $\psi(c,\tau) = \psi_0$ a constant (here: $\psi_0 = .5$) for different temporal distances τ of the perturbing pulse and the onset of the step stimulus, where $\psi(c,\tau)$ is given by (1). As Fig. 2 shows, there is no value for λ_2 that would allow to estimate the correct c-values. The question remains how one can determine whether detection is indeed by TPD. The simplest experimental way appears to be to test whether the threshold amplitudes are independent of the value of T, as long as T is larger than t_{max}, the time at which the deterministic response assumes its maximum. Unfortunately, Roufs et al. did not explicitly test for this.
Meinhardt and Mortensen (1998) also successfully predicted the data from one experiment from those of another on the basis of linear models and the assumption of TPD, which may be taken as further support for the linearity approximation and TPD as a mode of detection. However, the temporal form of the stimulus presentation was always identical.

2. The channel model underlying (1) is somewhat unspecified: the activity of the "channel" is assumed to be a Gaussian process X_t. For any $t \in J$, the activity is given by $X(t) = g(t) + \xi(t)$, where $\xi(t) = X(t) - g(t)$ may, in principle, be either additive or multiplicative. Eq. (1) was derived assuming additive noise. Fig. 3 shows how the responses to a step stimulus would look if the noise is either additive or multiplicative. Since in case of additive noise the amplitudes (or the variance) of $\xi(t)$ remain invariant as the amplitude of g increases, the effect of the noise should become smaller the larger the amplitude of g. In case of multiplicative noise, defined by the condition $\sigma^2(t, X(t)) \propto X(t)$ in (7), the amplitudes of $\xi(t)$ are about proportional to g. This suggests that the assumption of multiplicative noise would certainly not allow to cater for the data of Roufs and Blommaert (1981) if detection were indeed by TPS.
The definition of multiplicative noise underlying the plots in Fig. 3 corresponds to that given in Dayan and Abbott (2001). A more general definition defines noise to be multiplicative if, in Eq. (7), $\sigma^2(t)$ depends on $X(t)$. Therefore one could consider the relationship $\sigma^2(t) \propto 1/X(t)$, $X(t) \geqslant 0$ for all $t \in J$. In this case the noise amplitudes become minimal in the neighbourhood of t_{max}, $g_{max} = g(t_{max})$, which may allow for an approximation of TPD-detection in case of detection by TPS. Apart from the fact that such a specification of multiplicative noise is completely ad hoc and does not correspond to physiological results (see Mortensen, 2007, for a review of results) is clear that this type of noise does not help to rescue the hypothesis of detection by TPS with respect to the data of Roufs and Blommaert (1981): TPS effects would result in particular for values of τ within the range $(0, t_{max})$. In any case, the general considerations concerning the approximation of TPD in case of detection by TPS presented in Mortensen (2007) do not depend on the type of noise.
3. In order to overcome the lack of specificity of the channel model a cell-assembly model was proposed in Section 3.2. The model allows for inputs from other assemblies, where these inputs are represented by mean value functions of spike trains. So there are basically two types of "noise": dW_t, corresponding to diffusive noise (e.g. synaptic noise), and $\alpha(t)$ in (21). If $\alpha(t)$ assumes large values within a trial, the stimulus will get buried in the activity of the assembly, i.e. the stimulus is likely not to be detected. The reverse holds when $\alpha(t)$ assumes small values within a trial. So if one can assume $\alpha(t) \approx \eta$ a constant during a trial, one may define the psychometric function for the case of detection by TPD in terms of the distribution function of η, which is assumed, as a simplification, to vary only between trials.
4. It should be noted that the fast convergence of the mean response $m(t)$ towards $\beta g(t)$, β some proportionality constant, means that the mean response of the assembly is proportional to the mean input g. Thus according to the model the form of g appears to be determined by early processing stages; this would correspond to findings from fMRI studies, see e.g. Boynton, Demb, Glover, and Heeger (1999), where further references may be found.

To the extend the model represented by Eq. (21) is realistic, the inspection of the trajectories defined by Eq. (21) suggests the implausibility of a TPS model. The trajectories presented in Fig. 5 indicate that the assumption of detection by TPS implies that detection would depend on a sequence of peaks in the spike rate, where these peaks are of a very short duration if compared to the duration of the channel's impulse response. Phenomenologically, during trials in which the stimulus is detected, the visible response to the stimulus appears and decays smoothly; there is no sequence of extremely short pulses. So TPS would require some higher-level neuronal population that is activated by the population considered here, which should smooth the sequence of peaks of the spike rates over a time window. Actually, it would be this second neuronal population that mediates the detection response, i.e. that turns it into a smoothly varying conscious event. However, this brings us back to a TPD model, because detection will depend on the peak of the smoothed sequence of peaks in the spike rate.

It has been argued[3] that reaction time studies, e.g. the study by Tolhurst (1975), indicate the effect of TPS in

[3] By an anonymous reviewer.

detection tasks, meaning that a rejection of the TPS hypothesis is incompatible with these findings. While it is true that the expression (1) may be employed in one way or another in modelling reaction times, the results from studies on reaction times do not necessarily contradict the TPD-interpretation of the data from Roufs et al. In their experiment, no reaction times were measured. A pure detection task and a detection task combined with a reaction time measurement may yield different results since in reaction times the temporal structures (i) of the detection process and (ii) of the response process are combined. The considerations in this paper refer to detection processes only, and the point here is that TPS cannot be considered a general model of detection. Rather, subjects appear to be able to choose among a TPS or a TPD strategy, depending on experimental conditions.

The role of attention: It appears to be well established that attention has an effect on the detection process. Yeshurun and Carrasco (1998) (and further references there) argue that focussing attention on a position or a feature will reduce noise and enhance the response (here: g). This hypothesis is supported by more recent work (e.g. Martinez-Trujillo & Treue, 2004; Treue, 2003, 2004; Yeshurun & Carrasco, 1999). Reynolds and Desimone (2003) found that attention increases the contrast gain in V4 neurons. When two stimuli are projected onto the same receptive field, one weaker than the other, the response is determined by the one with the greater contrast. If attention is focussed on the weaker, the response is greater than that to the stimulus with greater contrast.

To summarise, the task is to model the reduction of the noise, and possibly a proportional increase of $g(t)$. The simulations based on Eq. (22) suggest that the probability of detection is increased by a decrease of η, and therefore of σ^2, whereas the amplitude A_g, defined in (25), remains invariant with respect to changes in η. The finding that attentional focussing may reduce the noise can be modelled as a reduction of the amplitude of the random function $\sigma(t)$ in (21), i.e. with respect to the simplified model considered here as a reduction of the expected value of η. This interpretation appears to be plausible since one may argue that the attentional focus on one aspect will reduce the activity due to other, not attended aspects of a stimulus. The model proposed by Hamker (2004) may provide a conceptual framework for a model of detection that encompasses the attentional modulation of $\alpha(t)$ in (19) as well as the proportional increase of g.

Luck, Chelazzi, Hillyard, and Desimone (1997) reported an *increase* of spontaneous firing rates in V2 and V4 neurons, of 30–40%, when attention was directed inside a receptive field of these neurons. Ress, Backus, and Heeger (2000) also reported an *increased* activity of V1, V2 and V3 neurons after attentional focussing. These findings certainly complicate matters and indicate that the model requires sophisticated modifications in order to be able to account for data from various experiments.

The mean response and interactions between neurons: The model proposed in Section 3.2 is meant to capture very elementary aspects of the temporal course of neuronal activation that is generated by the presentation of spatially very simple stimuli. Spatially more complex stimuli will require a more explicit discussion of further facets of neural activity, like e.g. local, nonlinear summation of responses, which simulates dendritic fields of grouping cells (du Buf, 1992, 2005). Here, the coupling between neurons was assumed to be given by $w_{ij} = \kappa_0/N$ for all (i,j); no specific assumption was made as to the values of κ_0 and N. This postulate looks like a very drastic simplification. However, in the light of empirical work and theoretical computations, the assumption may serve as a first approximation. Averbeck and Lee (2004) pointed out that, according to some experiments employing multi-electrode recordings, correlations between activities of neurons are not relevant to the information extracted for the detection response, whereas other investigations suggest that coherent oscillations play an important role. Averbeck et al. argue that the long time scale of most correlations appears to be inconsistent with the fast propagation of information. The role of coherent oscillations may differ from that of noise correlations: while some work suggests that such oscillations may not even play a role in the read-out process when the task is to segregate textures (Lamme & Spekreijse, 1998), Woelbern, Eckhorn, Frien, and Bauer (2002) reported that such oscillations are relevant. Golledge et al. (2003) found in the context of feature binding that between 89% and 96% of the information was carried by firing rates, and only 4–11% of the information was carried by correlations in the primary visual cortex. Averbeck and Lee (2004) suggested that coherent oscillations may be related to the gating of information flow in the cortex, and thus to attentional focussing. Romo, Hernádez, Zainos, and Salinas (2003) argued that, while it may be true that positively correlated noise decreases the coding precision of similarly tuned neurons, positively correlated fluctuations of differently tuned neurons actually increases coding efficiency. Meinhardt, Schmidt, Persike, and Röers (2004) and Persike and Meinhardt (2006) found strong synergy effects in the detection of Gabor stimuli. These effects may be explained in terms of lateral interactions among neurons responding to individual elements of the Gabor stimuli. However, Johnson (2004) demonstrated that synergy measures, i.e. measures of entropy and mutual information, do not necessarily assess cooperation. Non-cooperative populations, which only have stimulus-induced dependencies, can always represent stimulus features with arbitrary fidelity, regardless of the neural code, if only the population size is sufficiently large. Such findings may justify the assumption that the values of w_{ij} are small and that particular patterns of w_{ij} need not be considered.

Related work: Models assuming that the neural activity can be represented by some diffusion process and that detection occurs when the activity exceeds some threshold

value appear to be related to the TPS model. For instance, Ratcliff and Rouder (1998), Ratcliff, van Zandt, and McCoon (1999) and Ratcliff (2001) assumed that a detection response is triggered when the amount of accumulated evidence exceeds a critical level. However, this model employs the notion of "evidence" in a very general way and modifications must be made in order to be able to determine responses to impulse and step functions. Another model, called LATER (Linear Approach to Threshold with Ergodic Rate) was proposed by Reddi and Carpenter (2000). There it is assumed that a decision signal X develops linearly from an initial value S_0 at some rate r, and r is subject to Gaussian perturbation. A response is triggered when a threshold value S_T is reached. A similar model was proposed by Reeves, Santhi, and Decaro (2005). All these models are based on rather elegant simplifications of mathematically complex diffusion models; whether they can serve to "explain" data like those of Roufs and Blommaert (1981) remains to be investigated.

There exist a few other aspects of the detection process that may have to be considered in more detail. In models of population coding (Georgopoulos, Kettner, & Schwartz, 1988; Pouget, Dayan, & Zemel, 2003) the question of optimal detection is discussed. In this context, Deneve, Latham, and Pouget (1999) proposed recurrent network architectures with broad tuning curves that allow to approximate an ideal observer. Their networks estimate the orientation ϕ of bars in a way that approximates maximum-likelihood estimations of the parameter ϕ. The authors employed deterministic, nonlinear equations which represent the mean output of neural units, plus additive noise that does not depend on the deterministic stimulus response. The deterministic response reflects the mean spike rate. The subject is assumed to respond with respect to the maximally activated unit, but temporal aspects were not explicitly discussed; one may say that implicitly detection by TPD was assumed. The model proposed in Section 3 differs from the population coding models in so far as the spike rate is considered as a stochastic process developing in time, and is similar to these models in so far as detection is assumed to depend on the maximum of the mean spike rate. It may therefore be interesting to investigate the population model of Section 3, combined with the TPD assumption, with respect to the ideal observer.

There is one aspect that different detection models have in common: the assumption that detection occurs if some decision variable exceeds a threshold value. In a TPS model it is the raw activity, in a TPD model it is the mean response, and in other models it may be the likelihood of a trajectory given a stimulus was shown, or some measure of information or "evidence". So implicitly some interpretation of the activity of neuronal populations is postulated, with at least two alternatives: (a) either yet another neuronal population is required to interpret the interpreting population, which in the end means that one runs into an infinite regress, or (b) the interpreting population itself generates the conscious percept. As Shadlen and Newsome (2001, p. 1916) remarked, we have some knowledge of the sensory and motor processing stages but we know virtually nothing about the key cognitive stage of decision formation. So there is some future work to look forward to.

Acknowledgment

I am greatly indebted to Prof. Hans du Buf, University of the Algarve, for his critical comments and constructive discussions of the manuscript.

Appendix A. The perturbation method

Let $s_\delta = c\Delta t\delta(t)$ be a brief pulse of duration Δt, $\delta(t)$ the Dirac-Delta-function, and let $s_\sigma(t)$ denote a stimulus with arbitrary temporal course. Let $H(t) = 1$ for $t > 0$ and $H(t) = 0$ for $t \leqslant 0$. Consider a stimulus with wave form

$$s(t) = \begin{cases} c[\Delta t\delta(t) + qH(t-\tau)s_\sigma(t-\tau)], & \tau > 0 \text{ (a)}, \\ c[qs_\sigma(t) + \Delta tH(t+\tau)\delta(t+\tau)], & \tau \leqslant 0 \text{ (b)}, \end{cases}$$
$$t \geqslant 0,\ q \ll 1. \tag{27}$$

In (a), the pulse s_δ precedes the presentation of s_σ by the interval τ, in (b) s_σ precedes the pulse *ssd* by the interval $-\tau$, $\tau < 0$. Negative τ-values are introduced only to allow for the graphical representation chosen by Roufs et al. Suppose the detecting channel can be approximated by a linear system for sufficiently small stimulus amplitudes c. Let g be the deterministic response to s, and let h denote a unit response, i.e. a response when the contrast equals 1, with $h(t) = 0$ for $t < 0$ because of causality. It follows that

$$g(t) = \begin{cases} c[h_\delta(t) + qH(t-\tau)h_\sigma(t-\tau)], & \tau > 0 \\ c[qh_\sigma(t) + H(t+\tau)h_\delta(t+\tau)], & \tau \leqslant 0 \end{cases} \tag{28}$$

Assumptions:

1. The stimulus is detected if $g_{\max} = k_0$, with k_0 a certain constant (this is the peak-detection assumption).
2. Let $g_\delta(t) = c_0h_\delta(t)$, and let t_{ex} be the time at which $g_\delta(t)$ assumes its maximum. Let t_m be the time at which g assumes its maximum, i.e. let

$$g(t_m) = c(h_\delta(t_m) + qh_\sigma(t_m - \tau)), \tag{29}$$
$$g_\delta(t_{ex}) = c_0h_\delta(t_{ex}). \tag{30}$$

If the value of q is sufficiently small, then

$$t_m \approx t_{ex}. \tag{31}$$

The assumption $t_m \approx t_0$ may be motivated as follows:

$$\left.\frac{dg(t)}{dt}\right|_{t=t_m} = c\left(\frac{dh_\delta(t)}{dt} + q\,\frac{dh_\sigma(t-\tau)}{dt}\right)\Bigg|_{t=t_m} = 0, \tag{32}$$

which implies

$$\frac{dh_\delta(t_m)}{dt} = -q\,\frac{dh_\sigma(t_m - \tau)}{dt}. \tag{33}$$

For q sufficiently small one has $dh_\delta(t_m)/dt \approx 0$, meaning that for $q \ll 1$ one may assume $t_m \approx t_{ex}$. The approximation is the better, the more peaked h_δ is around t_{ex}. One has then the approximation

$$\frac{k_0}{q}\left(\frac{1}{c_s(\tau)}-\frac{1}{c_0}\right)=\hat{h}_\sigma(t_{ex}-\tau)\approx h_\sigma(t_{ex}-\tau). \quad (34)$$

Proof. Detection occurs with $\psi=\psi_0$ if $g(t_m)=g(t_{ex})=k_0$, so that $h_\delta(t_{ex})=k_0/c_0$, k_0 a constant whose value depends on ψ_0. Substituting t_{ex} for t_m and k_0/c_0 for $h_\delta(t_{ex})$ in (29) one is immediately led to (34); this holds for positive and negative values of τ. □

Comment: Let $t'=t_{ex}-\tau$, Roufs et al. plot the estimates $\hat{h}_\sigma(t')$ versus the corresponding values of $-\tau$, allowing for negative τ-values. If $\tau=t_{ex}$, one gets $t'=0$, so the value of $\hat{h}_\sigma(0)$ plotted versus $-\tau=-t_{ex}$. In order to get the value for $h_\sigma(150)$ one gets $\tau=t_{ex}-150$, which will be negative if $t_{ex}<150$, etc.

References

Adelson, E. H., & Bergen, J. R. (1985). Spatiotemporal energy models for the perception of motions. *Journal of the Optical Society of America, A*, *2*, 284–299.

Arrowsmith, D. K., & Place, C. M. (1982). *Ordinary differential equations*. London: Chapman & Hall.

Averbeck, B. B., & Lee, D. (2004). Coding and transmission of information by neural ensembles. *Trends in Neurosciences*, *27*(4), 225–230.

Baccus, S. A., & Meister, M. (2002). Fast and slow contrast adaptation in retinal circuitry. *Neuron*, *36*, 909–919.

Boynton, G. M., Demb, J. B., Glover, G. H., & Heeger, D. J. (1999). *Vision Research*, *39*, 257–269.

Carandini, M., Demb, J. B., Mante, V., Tolhurst, D. J., Dan, Y., Ohlshausen, B. A., et al. (2005). Do we know what the early visual system does? *The Journal of Neuroscience*, *25*, 10577–10597.

Carandini, M., Mechler, F., Leonhard, C. S., & Movshon, J. A. (1996). Spike train encoding by regular spiking cells in the visual cortex. *Journal of Neurophysiology*, *76*, 3425–3441.

David, S. V., Vinje, W. E., & Gallant, J. L. (2004). Natural stimulus statistics alter the receptive field structure of V1 neurons. *The Journal of Neuroscience*, *24*, 6991–7006.

Dayan, P., & Abbott, L. F. (2001). *Theoretical neuroscience. Computational and mathematical modeling of neural systems*. Cambridge: The MIT Press.

Deneve, S., Latham, P. E., & Pouget, A. (1999). Reading population codes: A neural implementation of ideal observers. *Nature Neuroscience*, *2*, 740–745.

Ditlivsen, O. (1971). Extremes and first passage times with applications in civil engineering. *Some approximative results in the theory of stochastic processes*. Thesis, Technical University of Denmark, Copenhagen.

du Buf, J. M. H. (1992). Modeling spatial vision at the threshold level. *Spatial Vision*, *6*, 25–60.

du Buf, J. M. H. (2005). Modelfest and contrast-interrelation-function data predicted by a retinal model. *In 28th European conference on visual perception (ECVP). Perception*, *34*(Suppl.), 240.

Fitzpatrick, D. (2000). Seeing beyond the receptive field in primary visual cortex. *Current Opinion in Neurobiology*, *10*, 438–443.

Freeman, C. B., Durand, S., Kiper, D. C., & Carandini, M. (2002). Suppression without inhibition in visual cortex. *Neuron*, *35*, 759–771.

Gardiner, C. W. (1990). *Handbook of stochastic processes for physics chemistry and the natural sciences*. Berlin: Springer.

Georgopoulos, A. P., Kettner, R. E., & Schwartz, A. B. (1988). Primate motor cortex and free arm movements to visual targets in three-dimensional space, II. Coding of the direction of movement by a neuronal population. *Journal of Neuroscience*, *8*, 2928–2937.

Gerstner, W. (2000). Population dynamics of spiking neurons: Fast transients, asynchronous states and locking. *Neural Computing*, *12*, 43–89.

Gerstner, W., & Kistler, W. (2002). *Spiking neuron models—single neurons, populations, plasticity*. Cambridge: Cambridge University Press.

Golledge, D. R., Panzeri, S., Heng, F., Pola, G., Scannell, J. W., Giannicopoulos, D. V., et al. (2003). Correlations, feature-binding, and population coding in primary visual cortex. *NeuroReport*, *7*, 1045–1050.

Hamker, F. H. (2004). A dynamic model of how feature cues guide spatial attention. *Vision Research*, *44*, 501–521.

Hansel, D., Mato, G., Meunier, C., & Neltner, L. (1998). On numerical simulations of integrate-and-fire neural networks. *Neural Computation*, *10*, 467–483.

Harris, K. D. (2005). Neural signatures of cell assembly organization. *Nature Reviews*, *6*, 399–407.

Heeger, D. J. (1992). Normalization of cell responses in cat striate cortex. *Visual Neuroscience*, *9*, 181–197.

Honerkamp, J. (1990). *Stochastic dynamical systems*. New York: VCH Publishers, Inc.

Hubel, D. H., & Wiesel, T. N. (1962). Receptive fields, binocular interaction, and functional architecture of the cat's visual cortex. *Journal of Neurophysiology*, *68*, 1373–1383.

Jazayeri, M., & Movshon, J. A. (2006). Optimal representation of sensory information by neural populations. *Nature Neuroscience*, *9*, 690–696.

Johnson, D. H. (2004). Neural population structures and consequences for neural coding. *Journal of Computational Neuroscience*, *16*, 69–80.

Kim, J. N., & Shadlen, M. N. (1999). Neural correlates of a decision in the dorsolateral prefrontal cortex of the macaque. *Nature Neuroscience*, *2*, 176–185.

Kloeden, P. E., & Platen, E. (1992). *Numerical solution of stochastic differential equations*. Berlin: Springer.

Lamme, V. A., & Spekreijse, H. (1998). Neuronal synchrony does not represent texture segregation. *Nature*, *396*, 362–366.

Leopold, D. A., Murayama, Y., & Logothetis, N. K. (2003). Very slow activity fluctuations in monkey visual cortex: Implications for functional brain imaging. *Cerebral Cortex*, *13*, 1047–3211.

Luck, S. J., Chelazzi, L., Hillyard, S. A., & Desimone, R. (1997). Neural mechanisms of spatial selective attention in areas V1, V2 and V4 of Macaque visual cortex. *Journal of Neurophysiology*, *77*, 24–42.

Martinez-Trujillo, J. C., & Treue, S. (2004). Feature-based attention increases the selectivity of population response in primate visual cortex. *Current Biology*, *14*, 744–751.

Meinhardt, G., & Mortensen, U. (1998). Detection of aperiodic test patterns by pattern specific detectors revealed by subthreshold summation. *Biological Cybernetics*, *79*, 413–425.

Meinhardt, G., Schmidt, M., Persike, M., & Röers, B. (2004). Feature synergy depends on feature contrast and objecthood. *Vision Research*, *44*, 1843–1850.

Mortensen, U. (2007). An analysis of visual detection by temporal probability summation. Journal of Mathematical Psychology, in press, doi:10.1016/j.jmp2007.02003.

Movshon, J. A., Thompson, I. D., & Tolhurst, D. J. (1978a). Receptive field organization of complex cells in the cat's striate cortex. *Journal of Physiology*, *283*, 79–99.

Movshon, J. A., Thompson, I. D., & Tolhurst, D. J. (1978b). Spatial summation in the receptive fields of simple cells in the cat's striate cortex. *Journal of Physiology*, *283*, 53–77.

Paradiso, M. A. (1988). A theory for the use of visual orientation formation which exploits the columnar structure of striate cortex. *Biological Cybernetics*, *58*, 35–49.

Persike, M., & Meinhardt, G. (2006). Synergy of features enables detection of texture defined figures. *Spatial Vision*, *19*, 77–102.

Pouget, A., Dayan, P., & Zemel, R. (2003). Information processing with population codes. *Annual Review of Neuroscience*, *26*, 381–410.

Ratcliff, R. (2001). Putting noise into neurophysiological models of simple decision making. *Nature Neuroscience*, *4*, 336–337.

Ratcliff, R., & Rouder, J. N. (1998). Modelling response times for two-choice decisions. *Psychological Science*, *9*, 347–356.

Ratcliff, R., van Zandt, T., & McCoon, G. (1999). Connectionist and diffusion models of reaction times. *Psychological Review*, *106*, 261–300.

Reddi, B. A., & Carpenter, R. H. S. (2000). The influence of urgency on decision time. *Nature Neuroscience*, *3*, 827–830.

Reeves, A., Santhi, N., & Decaro, S. (2005). A random-ray model for speed and accuracy in perceptual experiments. *Spatial Vision*, *18*, 73–83.

Ress, D., Backus, B. T., & Heeger, D. J. (2000). Activity in primary visual cortex predicts performance in a visual detection task. *Nature Neuroscience*, *3*, 940–945.

Reynolds, J. H., & Desimone, R. (2003). Interacting roles of attention and visual salience in V4. *Neuron*, *37*, 853–863.

Romo, R., Hernádez, A., Zainos, A., & Salinas, E. (2003). Correlated neuronal discharges that increase coding efficiency during perceptual discrimination. *Neuron*, *38*, 649–657.

Roufs, J. A. J., & Blommaert, F. J. J. (1981). Temporal impulses and step responses of the human eye obtained psychophysically by means of a drift-correcting perturbation technique. *Vision Research*, *21*, 1203–1221.

Rust, N. C., Schwartz, O., Movshon, J. A., & Simoncelli, E. (2005). Spatio-temporal elements of macaque V1 receptive fields. *Neuron*, *46*, 945–956.

Seung, H. S., & Sompolinsky, H. (1993). Simple models for reading population codes. *Proceedings of the National Academy of Sciences, USA*, *90*, 10479–10753.

Shadlen, M. N., & Newsome, W. T. (2001). Neural basis of aperceptual decision in the parietal cortex (are LIP) of the rhesus monkey. *Journal of Neurophysiology*, *86*, 1916–1938.

Tolhurst, D. J. (1975). Reaction times in the detection of gratings by human observers—a probabilistic mechanism. *Vision Research*, *23*, 1143–1149.

Touryan, J., Felsen, G., & Dan, Y. (2005). Spatial structure of complex cell receptive fields measured with natural images. *Neuron*, *45*, 781–791.

Treue, S. (2003). Visual attention: The where, what, how and why of saliency. *Current Opinion in Neurobiology*, *13*, 428–432.

Treue, S. (2004). Perceptual enhancement of contrast by attention. *Trends in Cognitive Sciences*, *8*, 435–437.

Tyler, C. W., & Chen, C. C. (2000). Signal detection theory in the 2AFC paradigm: Attention, channel uncertainty and probability summation. *Vision Research*, *40*, 3121–3144.

Vinje, W. E., & Gallant, J. L. (2000). Sparse coding and decorrelation in primary visual cortex during natural vision. *Science*, *287*, 1273–1276.

Watson, A. B. (1982). Derivation of the impulse response: Comments on the method of Roufs and Blommaert. *Vision Research*, *22*, 1335–1337.

Watson, A. B. (1987). Efficiency of the human image code. *Journal of the Optical Society of America, A*, *4*, 2401–2417.

Watson, A. B., & Solomon, J. A. (1997). Model of visual gain control and pattern masking. *Journal of the Optical Society of America, A*, *14*, 2379–2391.

Westheimer, G., Shimamura, K., & McKee, S. P. (1976). Interference with line-orientation sensitivity. *Journal of the Optical Society of America*, *66*, 332–338.

Woelbern, T., Eckhorn, R., Frien, A., & Bauer, R. (2002). Perceptual grouping correlates with short synchronization in monkey prestriate cortex. *NeuroReport*, *13*, 1881–1886.

Wolfart, J., Debay, D., Le Masson, G., Destexhe, A., & Bal, T. (2005). Synaptic background activity controls spike transfer from thalamus to cortex. *Nature Neuroscience*, *8*, 1760–1767.

Yeshurun, Y., & Carrasco, M. (1998). Attention improves or impairs visual performance by enhancing spatial resolution. *Nature*, *396*, 72–75.

Yeshurun, Y., & Carrasco, M. (1999). Spatial attention improves performance in a spatial resolution task. *Vision Research*, *39*, 293–306.

Zaghloul, K. A., Boahen, K., & Demb, J. B. (2003). Different circuits for ON and Off retinal ganglion cells cause different contrast sensitivities. *The Journal of Neuroscience*, *23*, 2645–2654.

Zaghloul, K. A., Boahen, K., & Demb, J. B. (2005). Contrast adaptation in subthreshold and spiking responses of mammalian Y-type retinal ganglion cells. *The Journal of Neuroscience*, *26*, 860–868.