SENSORY INTERPRETATION OF NEURAL EVENTS

Gerard Hesselmans
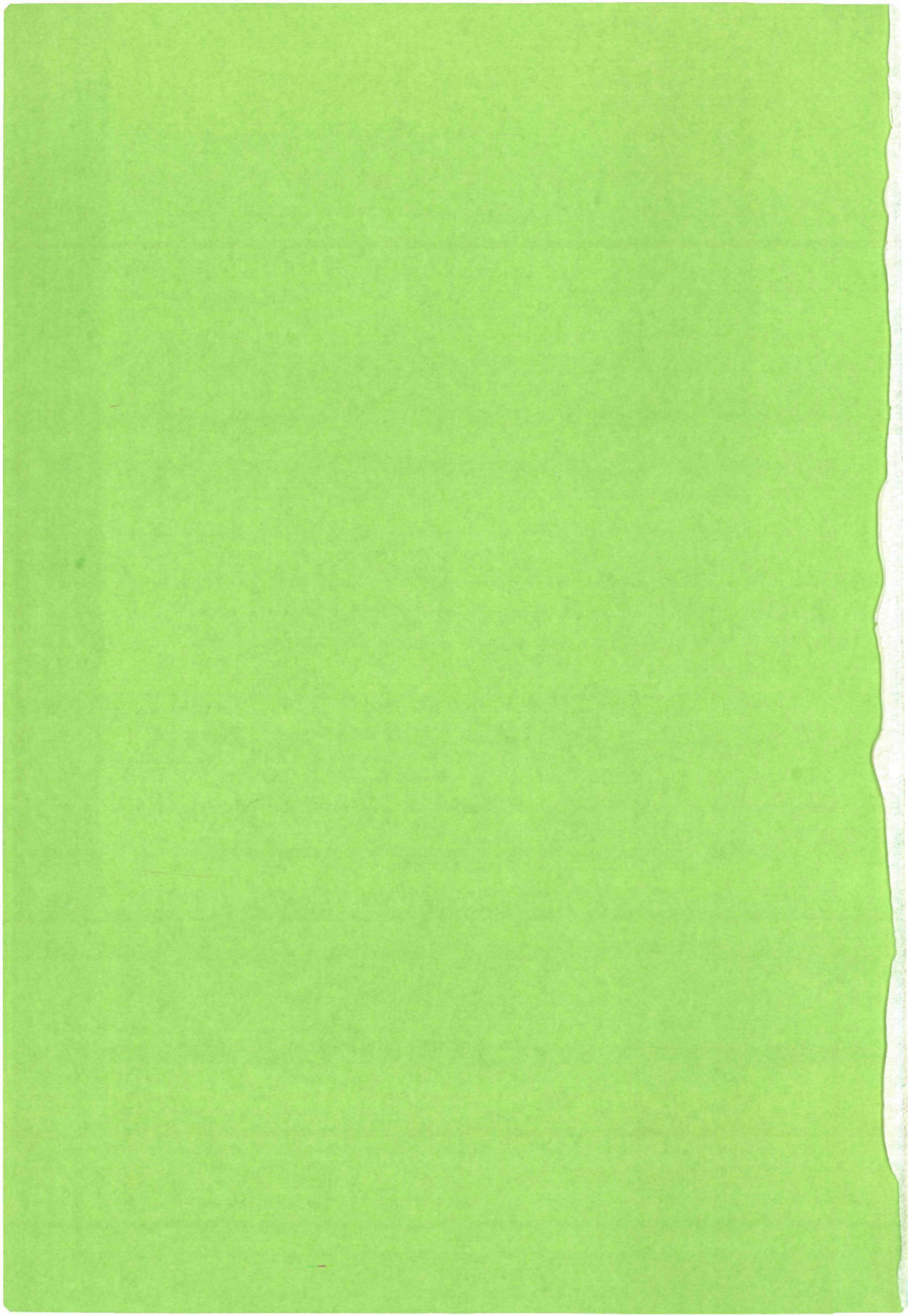
SENSORY INTERPRETATION OF NEURAL EVENTS

een wetenschappelijke proeve op het gebied van
de wiskunde en natuurwetenschappen

PROEFSCHRIFT

ter verkrijging van de graad van doctor
aan de Katholieke Universiteit te Nijmegen,
volgens besluit van het college van decanen
in het openbaar te verdedigen op donderdag 21 april 1988,
des namiddags te 1.30 uur precies

door

Gerardus Henricus Franciscus Maria Hesselmans

geboren 6 juni 1959 te Drunen

Contents

Introduction

The interaction of an animal with its environment is under the control of the nervous system. Information concerning the state of the external world is registered by the sensory system of the animal. The sensory information is passed to the nervous system for further analysis as to its relevance for the animal. After evaluation appropriate command signals may be sent to the motor system to act in response to the external stimuli.

The influence of a sensory stimulus depends on the properties of the sensory system. The sensitivity to a sensory modality is species-dependent. In this thesis, attention is focused on acoustic stimuli and their transformation by the auditory system of a lower vertebrate, the grassfrog (*Rana temporaria L.*). The grassfrog is the animal used to investigate the auditory system, and neural data will be presented from this animal.

The auditory system of the grassfrog is sensitive to frequencies between 100 and 5000 Hz. Sound enters the ear of the grassfrog at the ear-drum (tympanic membrane). Vibrations of the tympanic membrane are transported to the inner ear. In contrast to higher vertebrates, frogs have two inner ear organs sensitive to sound: the amphibian papilla processes sound with frequencies in the range 100-1000 Hz, and the basilar papilla sound with frequencies above 1000 Hz. Hair cells located at the amphibian and basilar papilla change the mechanical vibrations into electrical action potentials suited for further processing by the brain.

The brain consists of numerous neurons (Katz 1966) which form a network. Structurally a neuron can be divided into three parts: the dendrites (input), its soma, and an axon (output). The dendritic tree combines the messages of many other connecting neurons and transforms them into a generator potential.
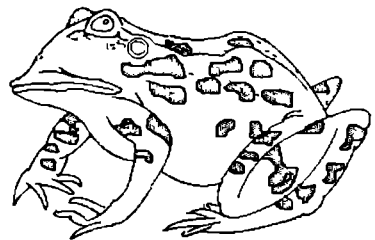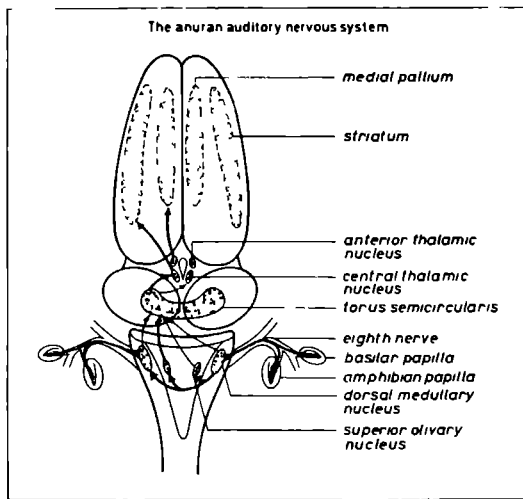
Figure 1: *Schematic overview of the auditory nervous system of the grassfrog (an anuran). For convenience only ascending neural pathways entering and leaving the left torus semicircularis are indicated (with kind permission of W. Epping (1985)).*

The generator potential controls the generation of action potentials at the soma. The output channel of a neuron is formed by an axon along which the action potential is transported to other parts of the nervous system. Interaction between neurons occurs at the synapse, the junction between different neurons. In this perspective the nervous system is an intricate network of cell bodies connected by axons and dendrites. Along this wiring neural communication takes place.

Neurons are to a certain degree grouped into functional units or nuclei. Some of them are more specialized in the processing of a certain sensory modality. In general nuclei in peripheral parts of the brain receive input from only one sensory modality e.g. vision or hearing. At higher levels, the various information sources are integrated. The largest auditory nucleus of the grassfrog is the torus semicircularis (TS), which is located in the midbrain and contains about 30000 neurons.

Neural activity is recorded by means of micro-electrodes (for experimental procedures the reader is refered to Epping and Eggermont 1985). Here attention will be focused upon the electrical pulses or action potentials related to sensory stimuli. An action potential or *neural event* lasts about 1 millisecond. During the refractory period of several milliseconds directly after the generation of an action potential, new action potentials cannot be generated; moreover neural activity is suppressed for another few milliseconds. The action potential is

2

the manifestation of ionic currents through the nerve membrane. The action potential occurs at the axon-hillock peripheral to the soma, and propagates along the axon away from the soma. The action potential is combined with output from other neurons in the dendritic tree. The result of the interaction of post-synaptic potentials is the generator potential and current. When they cross a critical level a new action potential is likely to be generated.

In peripheral parts of the sensory system, neural activity is often clearly related to a stimulus. For instance, many neurons in the TS of the grassfrog respond with characteristic activity patterns to a mating call (B-call) (see Chapter 4, figure 13). A common method to study the stimulus-(neural)event relation is by means of correlation procedures. An experiment-oriented approach is the reverse correlation method (de Boer and Kuyper 1968, Eggermont et al. 1983) which describes average properties of stimuli which have evoked a neural event. The reverse correlation method is put on a sound formal base through the concept of the pre-event stimulus ensemble (Aertsen et al. 1980, Johannesma and Aertsen 1982). Reverse correlation functions are obtained by suitable averages over the pre-event stimulus ensemble. From the pre-event stimulus ensemble the spectro-temporal receptive field (STRF) (Aertsen et al. 1980) can be determined which yields the frequency content of the stimulus as a function of time. However this characteristic in general depends on the particular stimulus ensemble used(Aertsen and Johannesma 1981).

The behavior of neurons in lower auditory stations can be related to simple characteristics (spectral, temporal, and spatial) of sound. In the midbrain of the grassfrog the various aspects are intricately coupled. This finding provides evidence for the presence of interneuronal connections. Functional connections can be determined by observing the activity of several neurons simultaneously. These functional connections might explain the capabilities of the grassfrog to identify and localize various (species-specific) sounds. Multi-unit activity was studied by Epping and Eggermont (1987). The correlation of the activity of different neurons appears to be stimulus-dependent.

The correlation between two time series of neural events requires a distance function. However, no objective distance function or measure for point processes is as yet available. In most experimental studies time is divided into "bins" and the choice of the binwidth is left to the experimenter's intuition of what is appropriate.

The work mentioned above is single-spike oriented: given the occurrence of an action potential what are the properties of the acoustic stimulus? Here this question will be generalized to actitivity patterns of several neurons: *given the neural activity, what might have been the stimulus or stimulus characteristics?* This question has been posed before in a more limited form. Given the neural activity, a small set of stimulus parameters, like intensity and frequency, has been estimated. See Chapter 2 for references. The validity of such an approach is of course limited to cases dealing with a low-dimensional stimulus ensemble.
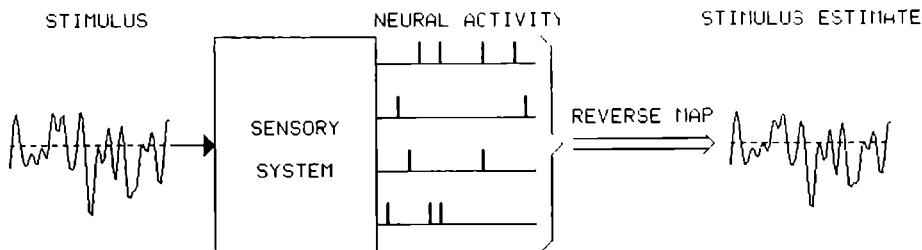
3

Figure 2: *The stimulus is transformed into neural activity by the sensory system and mapped back into the stimulus space.*

Here all stimulus aspects will be investigated, although the examples presented are low-dimensional, limited as they are by the two dimensions of this paper. The reverse mapping from the neural activity into the stimulus space is called the *sensory interpretation* of the neural activity (see figure 2). This implies that the evoked activity in the nervous system is interpreted as an internal image of the outer world which is decoded by the reverse map (Johannesma 1981, Johannesma et al. 1986).

Neural activity is interpreted within the context of a model of sensory transformation. A complete description of the activity of a single neuron already requires a model depending upon many variables. However, due to e.g. the Hodgkin-Huxley equations these variables are interdependent, and a reduction to a few descriptive variables might be feasible. The model used in this thesis is based upon previous work by Johannesma and van den Boogaard (1985), and consists of two parts: 1) a low-order Volterra system followed by 2) an exponential pulse generator. The Volterra system incorporates the deterministic and dynamic aspects of neural interaction: dendritic integration. The output of the Volterra system is the generator potential. The generator potential represents the state of the neuron. The exponential pulse generator reflects the stochastic nonlinear process of action potential generation at the axon-hillock. The description of the action potential is reduced to the moment of its occurrence. The waveform is not taken into account and is used only for identification of action potentials in multi-unit recordings (Epping and Eggermont 1987). Such an abstracted sequence of action potentials is called a point process. For an extensive treatise concerning point processes in relation to the present model the reader is refered to van den Boogaard 1985.

Once a model is set up, the neural activity pattern can be related to the time-space of sensory stimuli. A stochastic approach is taken to describe the set of stimuli. The model and the realization of the neural activity pattern determine the conditional probability distribution of the stimulus.
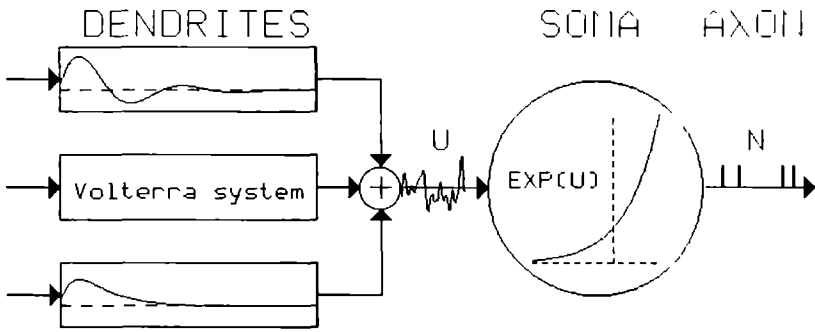
4

Figure 3: *Schematic representation of the neuron model.*

The reverse mapping from neural activity into sensory stimuli is probabilistic. The conditional distribution may be characterized by its maximum: the most plausible stimulus (Chapter 2). However this method fails when the maximum is not unique. Many neurons in the TS are insensitive to variations in the absolute phase of the stimulus signal. Therefore a more general approach is needed. This is done in Chapter 3, where the characteristic functional is introduced.

From the characteristic functional experimental quantities such as correlation functions can be obtained easily. Correlation functions can be ordered into a series of increasing complexity. In general all correlation functions are required to describe a set of signals completely. However if the ensemble is Gaussian distributed only two correlation functions are needed: average and covariance. In Chapter 4 the stimulus distributions are assumed to be approximately Gaussian. In this case, the average equals the most plausible stimulus. Since the most plausible stimulus is inappropriate to characterize the mapping by neurons in the TS, the second correlation function is used in Chapter 4. After a few approximations and a linear transformation of the second correlation function, a spectro-temporal representation of the stimulus is obtained. The validity of the approximations is supported by neural data. The approximations exclude neural interaction. This omission is corrected by the theoretical analysis of Chapter 5. As an addendum, in the final Chapter 6, the 'optimal' model given the experimental correlation functions is derived. Here 'optimal' is to be understood as maximally random given the restraints. The outcome turns out to be the model introduced by Johannesma and van den Boogaard (1985) and used in this thesis.

5

References

1. A. M. H. J. Aertsen, P. I. M. Johannesma, and D. J. Hermes, Spectro-temporal receptive fields of auditory neurons in the grassfrog. II. Analysis of the stimulus-event relation for tonal stimuli, *Biol. Cybernet.* 38:235-248 (1980).

2. A. M. H. J. Aertsen and P. I. M. Johannesma, A comparison of the spectro-temporal sensitivity of auditory neurons to tonal and natural stimuli, *Biol. Cybernet.* 42:145-156 (1981).

3. II. F. P. van den Boogaard, *Neural Interaction Equations: Transformation of Point Processes*, Thesis, University of Nijmegen, 1985.

4. J. J. Eggermont, P. I. M. Johannesma, and A. M. H. J. Aertsen, Reverse-correlation methods in auditory research, *Qt. Rev. Biophys.* 16:341-414 (1983).

5. W. J. M. Epping, *Auditory Information Processing in the Midbrain of the Grassfrog*, Thesis, University of Nijmegen, 1985.

6. W. J. M. Epping and J. J. Eggermont, Single unit characteristics in the auditory midbrain of the immobilized grassfrog, *Hearing Res.* 18:223-243 (1985).

7. W. J. M. Epping and J. J. Eggermont, Coherent neural activity in the auditory midbrain of the grassfrog, *J. Neurophysiol.* Vol. 57:1464-1483 (1987).

8. P. I. M. Johannesma, Neural representation of sensory stimuli and sensory interpretation of neural activity. Neural communication and control. *Adv. Physiol. Sc.* 30:103-125 (1981).

9. P. I. M. Johannesma and H. F. P. van den Boogaard, Stochastic formulation of neural interaction, *Acta Appl. Math.* 4:201-224 (1985).

10. P. I. M. Johannesma and A. M. II. J. Aertsen, Statistical and dimensional analysis of the neural representation of the acoustic biotope of the frog, *J. of Medical Systems*, Vol. 6-4:399-421 (1982).

11. P. I. M. Johannesma, A. M. H. J. Aertsen, H. F. P. van den Boogaard, J. J. Eggermont, and W. J. M. Epping, From synchrony to harmony: ideas on the function of neural assemblies and on the interpretation of neural synchrony, in: *Brain Theory,* (G. Palm and A. Aertsen, Eds.) Springer-Verlag, Berlin, 1986, pp. 25-47.

12. B. Katz, *Nerve, Muscle, and Synapse*, McGraw-Hill, New York, 1966.

Sensory Interpretation of Neural Activity Patterns

C. C. A. M. Gielen   G. H. F. M. Hesselmans
P. I. M. Johannesma

Abstract

Interaction of an animal with its environment depends on motor activity and sensory stimuli from the environment. Sensors (e.g. visual, auditory, vestibular) give a transformation of stimuli into neural activity patterns which can be used to construct or update the internal representation. This procedure requires an evaluation of the neural activity in terms of sensory processes.

This paper presents a theoretical approach for the sensory interpretation of neural activity based on a Bayesian estimation. The procedure implies a maximum likelihood estimation of the sensory stimulus that could have induced the observed neural activity pattern. The theoretical procedure is tested in a model study simulating the stochastic activity of a set of auditory nerve fibers. The results of the reconstructed stimuli are in good agreement with the stimuli that induced the neural activity in the auditory nerve fibers.

1   Introduction

Interaction of an animal with it on the sensory stimuli induced by the motor activity and/or generated by the environment. Sensors allow the animal to transform stimuli into a neural signal which is used to construct an internal representation of the environment. Effectors, such as muscles, allow the animal to interact with its environment. The role of the nervous system is to evaluate the signals from the sensors and to transform the information into a command signal for the effector system. As such the properties of the sensors determine the characteristics and limitations of the internal representation and, consequently, the behavioral repertoire.

These considerations have led to the question to what extent bounds can be defined for the internal representation, such as range and resolution, from knowledge of the peripheral transducer elements, which transform a stimulus into a neural code.

A large amount of information is available concerning the properties of sensory neurons of different modalities. For most sensory systems the main transduction properties of the sensors or receptors are known (see e.g. [21], [27], [31], [32]). Given the mapping of the stimulus into an activity pattern of sensory neurons the question arises, whether the inverse map, from a neural activity pattern into a sensory stimulus, exists and can be determined. This inverse mapping is a necessary requisite for an effective interaction of the organism with its environment. Apart from properties of the sensory transducer elements this inverse mapping is also determined by the topological structure of the neural elements. For each sensory modality part of the stimulus space is covered by a large number of elements, which together give a neural representation of the stimulus. Just as with the receptive fields in the retina ([22]) each point of the stimulus space is covered by the receptive field of more than one sensor, such that neural activity in different neurons is correlated. Since the distribution of receptive fields of the sensors is not homogeneous ([30, 21]) qualitative and quantitative aspects of the neural representation of the stimulus will not be homogeneously represented. This puts some constraints on the interpretation of multi-unit activity, i.e. the map from the neural activity in a large number of neurons to the stimulus space.

If the inverse map can be defined and determined in relation to experimental data, then what type of insight does it supply, We suggest that this inverse map can be considered as a sensory interpretation of the neural activity. It limits the perceptual resolution, which is constrained by the characteristics of the sensory neurons, and may indicate the relevant 'dimensions' in perception.

Proposals for a "sensory interpretation" of the activity of sensory neurons have been presented before. However in these publications the stimulus-space is reduced to a low-dimensional one, e.g. stimulus intensity and frequency in the auditory field ([6, 36]) and intensity, saturation and hue perception in vision research ([5, 8]). All these investigations were concerned with the estimation and discriminability of few parameters of simple stimuli, e.g. sustained tones or clicks. This paper presents a way to deal with time varying stimuli and to construct dynamic estimates of the complete stimulus.

In section 2 a set of formal definitions are presented to describe the stimulus space and the neural activity of a population of neurons. The characteristics of these sensory neurons will be assumed to be known. Next a Bayesian estimation procedure is developed to find the map ϕ, which represents the sensory interpretation of the neural activity, i.e. a map from the space of neural states into the stimulus space. Subsequently, the most plausible stimulus associated with this pattern of neural activity is defined and approximatively computed. This estimation procedure uses probability functions, which are conceptually closely related to the information processing by the nervous system, since the generation of action potentials by neurons is a stochastic process. In section 3 the theory to find the most plausible stimulus is further elaborated. Relevant questions in this context concern existence, uniqueness and characteristics of

the map ϕ. It will be shown that under some general conditions at least one, but in general more than one most plausible stimulus exists, thus leading to ambiguity in sensory interpretation.

In section 4 the theoretical procedure, outlined in sections 2 and 3, is applied to a model of a part of the nervous system, which is reasonably well understood: the auditory nerve, which carries the auditory information into the brain from the cochlea, where the mechanical vibrations, induced by a sound, are transformed into neural activity of the approximately 30,000 auditory nerve fibers.

2 Probabilistic description of sensory stimulus and neural activity

Neural activity and sensory stimuli will be described as variables in discrete time. The discretization of time is represented by intervals Δt. The restrictions on Δt will become clear in the text.

The neural activity at time t in population of K neurons is given by the vector $\mathbf{z} = \mathbf{z}(t)$ of dimension K with

$$z_k(t) \in \{0,1\}. \tag{1}$$

In this equation $z_k(t)$ is 1 or 0 corresponding to presence or absence of an action potential in neuron k in the time interval $[t, t + \Delta t)$. Since the description in this paper is based on probability functions, it implies that valuable results can be obtained only if the number K of neurons is sufficiently large. One reason for the time discretization is given by the discontinuous character of the neural activity. A single neuron can be considered as being in one of two states: it generates $(z_k = 1)$ or no $(z_k = 0)$ an action potential. The state of a neuron depends on the combined effect of all input signals and on time. *The probability for the neuron to generate an action potential increases monotonically with the sum of all input signals.* After an action potential has been generated the probability for another action potential to occur is lowered for a short while (approximately 1 msec). Because of the stochastic nature of this process the probability function, describing the probability of the neural states as a function of the stimulus will be considered. The definition (1) imposes the restriction on Δt to be so small, that a neuron can generate at most one action potential in the time interval Δt. Therefore Δt should be in the order of 1 ms, definitely not longer.

In the following the vector $\mathbf{x}$ denotes a stimulus. Depending on the sensory system under investigation $\mathbf{x}$ represents the pressure fluctuations of the air in the ear as a function of time for the auditory system, or a two-dimensional spatial intensity distribution as a function of time for the visual system. The time discretization will be introduced here too: the ith component of the vector $\mathbf{x}$ is given by

$$x_i(t) = \mathbf{x}(t - (i - 1)\Delta t). \tag{2}$$

9

Notice that the ith component $x_i(t)$ is not necessarily a scalar. For a visual stimulus $x_i(t)$ may represent the two-dimensional intensity distribution of luminous and chromatic contrast. Δt should be so small that the complete information remains present and that the phenomenon of aliasing does not occur. A finite choice of Δt is based on the assumption that an upper frequency can be found for each stimulus modality. Clearly the size of Δt is determined by the upper frequency W_1 in the stimulus ensemble and by the upper frequency W_2 in the frequency sensitive range of the sensory system under investigation. If W represents the upper frequency relevant for the analysis $(W = \max(W_1, W_2))$ then Δt is defined by

$$\Delta t = \min \left\{ 1 \text{ ms} , \frac{1}{2W} \right\}. \tag{3}$$

The dimension of x is given by the smallest upper bound of the range of the autocorrelation time T_1 of stimuli in the stimulus ensemble and the memory span of time constant T_2 of the sensory system under investigation. This leads to an upper limit for the stimulus duration $t = \max(T_1, T_2)$, which influences the generation of an action potential at time t. It defines the dimension of x as N where

$$N = \text{integer} \left(\frac{1}{\Delta t} \right) + 1. \tag{4}$$

Clearly x is a single stimulus from a collection of possible stimuli, which together form the stimulus space S, which forms a subspace of R^N for auditory stimuli. A real-valued function f on S will be defined, which describes a probability function over the stimulus space: $f(x)$ is the a $priori$ probability of occurrence of stimulus x.

In a similar way a real-valued function h, defined on the space of all neural states, is defined, which gives the probability distribution of all possible states z. We further define the neural response function $g(z|x)$. It is the probability function to find the neural state z provided a stimulus x was presented. Finally the probability distribution of a sensory stimulus given a neural activity pattern is represented by ϕ. It is the probability function to find the stimulus x given the fact that the neural state z was measured. The probability functions defined above are mutually related by the theorem of Bayes:

$$f(x)g(z|x) = h(z)\phi(x|z). \tag{5}$$

Now we can state the problem of this paper in a formal way. The best strategy to find the most plausible stimulus is to look for the stimulus x which gives a maximum for $\phi(x|z)$ given a neural state z. A necessary condition for the most plausible stimulus is, that it satisfies the equation

$$\begin{aligned} 0 &= \nabla_x \ln \phi(x|z) \\ &= \nabla_x [\ln f(x) + \ln g(z|x) - \ln h(z)], \end{aligned} \tag{6}$$

10

where $\nabla_{\mathbf{x}}$ is the gradient operator with respect to the stimulus $\mathbf{x}$. Since $\nabla_{\mathbf{x}}[\ln h(\mathbf{z})] = 0$ the term $\ln h(\mathbf{z})$ in Equation (6) may be dropped without loss of generality. In principle the remaining terms can be known by the experimenter.

The first term in Equation (6) is related to the distribution of the stimuli in the stimulus space S. The stimulus space S is a subspace of the space R^N, which contains all possible stimuli. S contains those stimuli, which when presented to man or animal can elicit a behavioral response. For example, the human auditory system has an upper frequency sensitivity of about 20 kHz. Therefore, the stimulus space for auditory stimuli is limited to frequencies below 20 kHz. The human eye has a temporal (flicker fusion frequency of about 60 Hz [19], [20]) as well as a spatial upper bound (small details up to 1 min. of arc can be discerned in the fovea [28]). These limits of sensory systems have been investigated extensively in psychophysics and electrophysiological experiments and form the boundaries of the stimulus space S. For animals these upper bounds are investigated with behavioral response tests and with electrophysiological means. So in principle we know the stimuli, which form the stimulus space S. Moreover, the stimuli from S are observable, which allows one to determine the probability $f(\mathbf{x})$ that a stimulus $\mathbf{x}$ is found.

Further calculations will be limited to a Gaussian distribution of stimuli. This amounts to the assumption that first and second-order stimulus characteristics supply a sufficient description. As a consequence

$$f(\mathbf{x}) = f_0 \exp\left[-\frac{1}{2}(\mathbf{x} - \mathbf{a})^T B^{-1}(\mathbf{x} - \mathbf{a})\right], \tag{7}$$

where

$f_0 =$ normalization factor, such that $\int d\mathbf{x}\, f(\mathbf{x}) = 1$,
$\mathbf{a}\ -$ expected or mean valued of stimulus,
$B\ =$ covariance matrix of stimulus ensemble.

Since B is the covariance matrix of the stimuli from S, B is positive definite and the number of eigenvalues of B equals the dimension of S. Therefore the mapping on S, which corresponds with the matrix B, is a homomorphism. Moreover there exists an inverse mapping B^{-1}, such that $B^{-1}B = I$, where I represents the identity matrix.

Although theoretically all stimuli from R^N may occur, in practice, sound pressure level of auditory stimuli will be limited to an appropriate upper value E. Therefore, only stimuli $\mathbf{x}$ with

$$(\mathbf{x}, \mathbf{x}) \leq E \tag{8}$$

will be considered where $(\mathbf{x}, \mathbf{x})$ represents the inner product of vector $\mathbf{x}$ with itself. This choice restricts the stimulus space to a compact (i.e. closed and bounded) and convex subset U of R^N.

For neurons of the nervus acousticus evidence has been presented that direct neural interactions between neurons are absent ([17]). Correlations in activity of these neurons are caused by common stimulus properties and overlapping receptive fields but not by neural interactions. This also appears to hold for the nervus opticus. Therefore $g(\mathbf{z}|\mathbf{x})$ can be deduced from the neural response functions $g_k(z_k|\mathbf{x})$ of the individual neurons. Formally this can be expressed as

$$g(\mathbf{z}|\mathbf{x}) = \prod_{k=1}^{K} g_k(z_k|\mathbf{x}) . \tag{9}$$

The neural response function $g_k(z_k|\mathbf{x})$ gives the conditional probability of finding neuron k in the state z_k ($z_k = 1$ if an action potential; $z_k = 0$ if no action potential is generated) after stimulus $\mathbf{x}$ is presented. By this definition the value of the function $g_k(z_k|\mathbf{x})$ will be between zero and one. In fact the probability of finding neuron k in one of the two possible states equals one ($z_k \in \{0,1\}$). Since the probability of finding an action potential is close to zero for weak stimuli ($|\mathbf{x}| \ll 1$), increasing gradually to one for very intense stimuli ($|\mathbf{x}| \gg 1$), the probability $g_k(z_k|\mathbf{x})$ can be written without loss of generality in the form

$$g_k(z_k|\mathbf{x}) = \frac{\exp G_k(z_k,\mathbf{x})}{\sum_{\xi_k} \exp G_k(\xi_k,\mathbf{x})} , \tag{10}$$

$$g_k(\mathbf{z}|\mathbf{x}) = \frac{\exp G_k(\mathbf{z},\mathbf{x})}{\sum_{\xi} \exp G_k(\xi,\mathbf{x})} . \tag{11}$$

The summation in the denominator of g_k and g over all neural activity patterns $\mathbf{z}$ gives the normalization to values between zero and one. Combination of Equations (9), (10), and (11) leads to the expression

$$G(\mathbf{z},\mathbf{x}) = \sum_{k=1}^{K} G_k(z_k,\mathbf{x}) . \tag{12}$$

Because of the binary nature of z_k and $\mathbf{z}$ the following concrete forms can be given to G_k and G

$$G_k(z_k,\mathbf{x}) = z_k v_k(\mathbf{x}) \tag{13}$$

$$G(\mathbf{z},\mathbf{x}) = \sum_{k=1}^{K} z_k v_k(\mathbf{x}) = (\mathbf{z}, \mathbf{v}(\mathbf{x})) , \tag{14}$$

where $(\mathbf{z}, \mathbf{v}(\mathbf{x}))$ represents the inner product between the vectors $\mathbf{z}$ and $\mathbf{v}(\mathbf{x})$. The function $\mathbf{v}(\mathbf{x})$ is related to the generator potential of the neuron. Comparison with Equation (10) shows that for increasing values of $v_k(\mathbf{x})$ the probability to find an action potential in neuron k ($z_k = 1$) increases and that the probability to find no action potential ($z_k = 0$) decreases.

12

The assumption to define the function $G_k(z_k, \mathbf{x})$ by the product of the neural state z_k and the stimulus $\mathbf{x}$ excludes nonlinearities in the spike generating mechanism such as the absolute refractory period. This restriction does not affect the generality of our approach since the refractory period influences the sequence of action potentials only for very effective, large amplitude stimuli, giving rise to a saturation of the response. A similar saturation is described by Equations (10) and (11) which prohibits very short spike intervals for physiologically relevant stimuli.

Related models were formulated by Siebert [36] in the auditory system and Buchsbaum [5] in vision. The scheme proposed by Buchsbaum consists of linear filters corresponding to the three colour mechanisms of the retina. The status of the output of these filters is comparable with the generator potential. The rate of the neural Poisson generators is however linearly controlled by the output of the filters. The model formulated by Siebert can easily be separated into two similar blocks. The first one being a distributed system of linear filters whose output, after a short-time mean square procedure is equivalent to the generator potential $v_k(\mathbf{x})$ presented in this paper. The second block consists of a no-memory spike generating mechanism, however with a different nonlinear saturating behavior. Due to the mean-square procedure, the validity of this model is limited to high frequency neurons without phase-lock. Colburn [6] indicates how to generalize this model to transient stimuli.

The fact that the function $G_k(z_k, \mathbf{x})$ does not depend explicitly on time disregards effects of prolonged stimulation such as depletion of transmitter. It is equivalent to the assumption that the system under study behaves stationary. For example, it assumes that time constants related to modification of the neural circuitry in ontogenetic development are long relative to the duration of the experiment.

The function $v_k(\mathbf{x})$ needs further clarification. A general way to characterize neuronal properties is given by the Wiener-Volterra expansion ([23, 25]). With this expansion the neural characteristics can be represented by a series of functionals of increasing complexity. Studies in the auditory system ([27, 3]), visual system ([9, 34, 26]) and vestibular system [29] have shown, that for a good approximation this series may be truncated after the second-order term. Therefore we write

$$v_k(\mathbf{x}) = b_k + (\mathbf{c}_k, \mathbf{x} - \mathbf{a}) + \frac{1}{2}(\mathbf{x} - \mathbf{a})^T D_k(\mathbf{x} - \mathbf{a}), \qquad (15)$$

where $\mathbf{c}_k$ and D_k are related to the first and second-order Wiener kernels of neuron k. Evidence for the fact, that neural responses can be predicted with this choice for $v_k(\mathbf{x})$ can be found in [10], [11], [18] and [26]. For linear systems the first two terms of $v_k(\mathbf{x})$ suffice for a full characterization of the system. The third term in Equation (15) is required for second degree nonlinear systems. At this moment the exact properties of $\mathbf{c}_k$ and D_k are irrelevant, however, it is important to notice that $v_k(\mathbf{x})$ is a continuously differentiable real-valued function,

defined on the stimulus space S and that $g(\mathbf{z}|\mathbf{x})$ can be known by calculation of the Wiener kernels for the individual neurons. To obtain an estimate of the sensory interpretation of the neural activity $\mathbf{z}$ several lines of thought can be followed. One possible approach is to determine the most plausible stimulus, which is the stimulus $\mathbf{x}$, which gives a maximum for $\phi(\mathbf{x}|\mathbf{z})$ for a given $\mathbf{z}$. Another estimate is obtained if one calculates the average plausible stimulus, given by $\sum_{\mathbf{x}} \mathbf{x}\,\phi(\mathbf{x}|\mathbf{z})$. In general, these estimates will not be the same if the function $\phi(\mathbf{x}|\mathbf{z})$ is asymmetric with respect to the maximum. In this study the approach will be adapted to determine the most plausible stimulus.

3 Determination of most plausible stimulus

A necessary condition on the most plausible stimulus is that it satisfies Equation (6). Since $\nabla_{\mathbf{x}} \ln h(\mathbf{z}) = 0$, (6) leads to

$$
\begin{aligned}
\mathbf{x} &= \mathbf{a} + B\nabla_{\mathbf{x}}[\mathbf{z}, \mathbf{v}(\mathbf{x})] - B\frac{\sum_{\xi}\{\nabla_{\mathbf{x}}(\xi, \mathbf{v}(\mathbf{x}))\}\exp(\xi, \mathbf{v}(\mathbf{x}))}{\sum_{\xi}\exp(\xi, \mathbf{v}(\mathbf{x}))} \\
&= \mathbf{a} + BV(\mathbf{z} - \mathbf{n}(\mathbf{x})),
\end{aligned}
\tag{16}
$$

where

$\mathbf{a}\quad$ = expected value of stimulus ensemble,
$B\quad$ = covariance matrix of stimulus ensemble,
$V\quad$ = $\nabla_{\mathbf{x}}\mathbf{v}(\mathbf{x})$ is $N \times K$ matrix of functions $\frac{\partial v_k(\mathbf{X})}{\partial \tau_n}$ relating
$\qquad$ the N stimulus components with the K neurons,
$\mathbf{n}(\mathbf{x})$ = expected neural activity pattern for sensory stimulus $\mathbf{x}$ defined by

$$
\mathbf{n}(\mathbf{x}) = \frac{\sum_{\xi} \xi \exp(\xi, \mathbf{v}(\mathbf{x}))}{\sum_{\xi}\exp(\xi, \mathbf{v}(\mathbf{x}))}\ .
\tag{17}
$$

Equation (16) shows that the most plausible stimulus is determined by *a priori* information about the stimulus ensemble ($\mathbf{a}$ and B), by properties of the neural transducer elements (V), by the actual neural activity pattern ($\mathbf{z}$) and by the expected neural activity pattern $\mathbf{n}(\mathbf{x})$ of the population neurons for the stimulus $\mathbf{x}$. A similar equation was obtained by Johannesma ([13]), who made the approximation $\mathbf{n}(\mathbf{x}) = 0$; this led to a simplification for the solution of the most plausible stimulus.

The general solution of Equation (16) is complicated by the fact that V and $\mathbf{n}(\mathbf{x})$ are functions of $\mathbf{x}$. Especially the relation between $\mathbf{n}(\mathbf{x})$ and $\mathbf{x}$ is complicated. In the following the existence and nature of the solution of Equation (16) will be discussed. By the choice of the stimulus ensemble, as given in Equation (7) and Equation (8), all stimuli are embedded in a compact and convex subset U of R^N. Then a mapping $\mathbf{M}$ is introduced by

$$
\mathbf{M} : \mathbf{x} \to \mathbf{a} + BV(\mathbf{z} - \mathbf{n}(\mathbf{x}))\,.
\tag{18}
$$

14

In a first approximation we assume

$$v_k(\mathbf{x}) = b_k + (\mathbf{c}_k, \mathbf{x} - \mathbf{a}) . \qquad (19)$$

This assumption implies that a first-order, linear approximation describes the 'potential' $v_k(\mathbf{x})$. Solutions for Equation (16), if $v_k(\mathbf{x})$ is nonlinearly related to $\mathbf{x}$ are discussed later.

Since $z_k \in \{0, 1\}$ and $0 < n_k(\mathbf{x}) < 1$, it follows that $0 < |z_k - n_k(\mathbf{x})| < 1$. Moreover $B\mathbf{c}_k$ and $\mathbf{a}$ are limited, resulting in an upper bound of the energy of the image of $\mathbf{x}$

$$E' = |\mathbf{a}| + \sum_{k=1}^{K} |B\mathbf{c}_k| . \qquad (20)$$

Choosing the upper energy E of the stimulus space U larger than E', $\mathbf{M}$ becomes a map from U to itself. From the definition of $n_k(\mathbf{x})$ and Equation (15), it follows that the function in Equation (18) is continuous and differentiable. Then the following theorem of Brouwer guarantees at least one fixed point for map $\mathbf{M}$, i.e. a vector $\mathbf{x}$ such, that $\mathbf{M}(\mathbf{x}) = \mathbf{x}$.

Theorem of Brouwer *Every continuous mapping from a compact and convex set to itself has at least one fixed point.*

A proof of this theorem can be found in any textbook on Topology (see e.g. [35]). This theorem guarantees the existence of an extreme (most or least plausible) stimulus.

With the approximation in Equation (19), the gradient of $\ln \phi(\mathbf{x}|\mathbf{z})$ is given by

$$B\nabla_{\mathbf{x}} \ln \phi(\mathbf{x}|\mathbf{z}) = \mathbf{x} + \mathbf{a} + \sum_{k=1}^{K} B\mathbf{c}_k \cdot \{z_k - n_k(\mathbf{x})\} . \qquad (21)$$

Since B is positive definite and n_k is increasing in v_k, for any $\mathbf{y} \in U$ each line $\mathbf{x} = \alpha\mathbf{y} + \beta$ in U ($\alpha \in R, \beta \in S$) has *at most* one stimulus such, that the Equation (21) equals zero. This can easily be verified by substituting $\mathbf{x} = \alpha\mathbf{y} + \beta$ in Equation (21) and multiplying by $\mathbf{y}B^{-1}$, then the right-hand side is a decreasing function in α. Therefore, each line contains at most one most plausible stimulus. Since this result applies to *each* line in U and since there is at least one most plausible stimulus, with the approximation given by Equation (19) there is exactly one most plausible stimulus.

Since $|\sum_{k=1}^{K} B\mathbf{c}_k| < E$ by the choice of E, Equation (21) becomes zero only for $\mathbf{x} \in U$ and not for stimuli $\mathbf{x}$ with $|\mathbf{x}| \geq E$. Consequently, the restriction to look for most plausible stimuli with a power less than E does not influence the results.

A way to find the most plausible stimulus is given by the following procedure. Let $\mathbf{x}_0$ be some arbitrary vector from U. Then we define a series of vectors in U by

$$\mathbf{x}_m = \mathbf{M}(\mathbf{x}_{m-1}) . \qquad (22)$$

Using the definition of M given in Equation (18) this becomes

$$\mathbf{x}_m = \mathbf{a} + BV(\mathbf{z} - \mathbf{n}(\mathbf{x}_{m-1})) = \mathbf{a} + B\sum_{k=1}^{K} \mathbf{c} \cdot \{z_k - n_k(\mathbf{x}_{m-1})\}. \qquad (23)$$

If this series does not converge, then one can resort to the series:

$$\mathbf{x} = (1 - \varepsilon)\mathbf{x}_{m-1} + \varepsilon \mathbf{M}(\mathbf{x}_{m-1}) \qquad (24)$$

which will always converge for ε small enough because the function n_k has a sigmoid shape, increasing monotonically as a function of $(\mathbf{c}_k, \mathbf{x})$. The rate of convergence depends on whether each function n_k increases fast or slow from 0 to 1 as a function of $\mathbf{x}$.

If instead of the linear approximation of $v_k(\mathbf{x})$ as given in Equation (19) the nonlinear expression of Equation (15) is used, the gradient of $\ln \phi(\mathbf{x}|\mathbf{z})$ is given by

$$
\begin{aligned}
B\nabla_{\mathbf{x}} \ln \phi(\mathbf{x}|\mathbf{z}) &= -(\mathbf{x} - \mathbf{a}) &+ B\sum_{k=1}^{K} \mathbf{c}_k \{z_k - n_k(\mathbf{x})\} \\
& &+ B\sum_{k=1}^{K} D_k(\mathbf{x} - \mathbf{a})\{z_k - n_k(\mathbf{x})\} \\
&= -(\mathbf{x} - \mathbf{a}) &+ B\sum_{k=1}^{K} \{\mathbf{c}_k + D_k(\mathbf{x} - \mathbf{a})\}\{z_k - n_k(\mathbf{x})\} \quad (25)
\end{aligned}
$$

Depending on the properties of D_k, there may be more than one most plausible stimulus.

For example, suppose that the stimulus ensemble is Gaussian white noise, then $\mathbf{a} = 0$ and $B = I$. Moreover, suppose that the nonlinearity is a pure two-sided rectification; then $D_k = I$ and $\mathbf{c}_k = 0$. In that case, Equation (25) reduces to

$$\nabla_{\mathbf{x}} \ln \phi(\mathbf{x}|\mathbf{z}) = -\mathbf{x} + \sum_{k=1}^{K} \mathbf{x}\{z_k - n_k(\mathbf{x})\} \qquad (26)$$

with

$$n_k(\mathbf{x}) = \frac{\exp(\mathbf{x}, \mathbf{x})}{1 + \exp(\mathbf{x}, \mathbf{x})} \quad \text{for } k \in \{1, 2, \ldots, K\}. \qquad (27)$$

This shows, that there is more than a single stimulus $\mathbf{x}$ for which expression (26) is zero. Two cases arise.

1. $\mathbf{x} = 0$ is a trivial solution. This is a minimum of $\phi(\mathbf{x}|\mathbf{z})$ and not a most plausible stimulus.
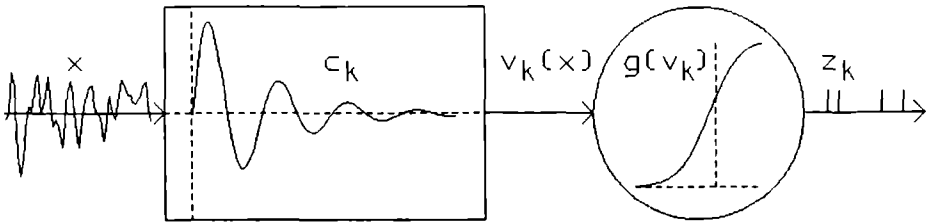
2. If $\mathbf{x} \neq 0$, the solution of

Figure 1: *Illustration of the model, which transforms the stimulus* x *into a sequence of events* z_k. *The event generating system consists of a linear filter followed by an instantaneous, nonnegative nonlinearity.*

$$\mathbf{x} = \mathbf{x} \sum_{k=1}^{K} \left(z_k - \frac{\exp(\mathbf{x}, \mathbf{x})}{1 + \exp(\mathbf{x}, \mathbf{x})} \right) \tag{28}$$

consists of a set of stimuli on a hypersphere. These solutions can be interpreted as 'most' plausible stimuli.

Grashuis ([11]) demonstrated that for some neurons in the auditory system the neural activity was determined by $(\mathbf{c}_k, \mathbf{x})^2 + (\tilde{\mathbf{c}}_k, \mathbf{x})^2$, where $\tilde{\mathbf{c}}_k$ is the Hilbert transform of the vector $\mathbf{c}_k$ and is orthogonal to $\mathbf{c}_k$. This implies that all stimuli for which $(\mathbf{x}, \mathbf{x})$ is constant in the hyperplane, spanned by the orthogonal vectors $\mathbf{c}_k$ and $\tilde{\mathbf{c}}_k$, give an equal probability for an action potential. This nonlinearity also results in the fact that not one but a set of stimuli in a closed curve in S form a set of most plausible stimuli.

Here is shown, that for each sensory system satisfying the assumptions described in Section 2, we can find at least one stimulus, most likely to have elicited the neural activity z. Clearly the precise properties of $f(\mathbf{x})$ and $g(z|\mathbf{x})$ determine whether there is only one such stimulus or whether more stimuli are equally plausible. In the next section the theory is used to find the most plausible stimulus for a set of simulated neurons in the auditory nerve, which transfers information from the cochlea, where a separation of frequencies takes place, to central brain structures.

4 Stimulus reconstruction: model simulation

We simulated a set of neurons in the auditory nerve, since the information processing by these neurons is well understood. A model can be made, which simulates these neurons with reasonable accuracy in several aspects. The model used in this study is given in Figure 1.

The process of sound transmission from the ear-drum via the middle ear and then travelling along the basilar membrane in the cochlea can be considered mainly as a mechanical filtering. Vibration of the basilar membrane

stimulates the hair cells which effect on the transduction process is a further filtering narrowing the band-pass characteristics of the mechanical filter [33]. The combined effect of these filtering stages is represented by the filter c_k. The release of transmitter generating post synaptic potentials on the dendrites of the auditory nerve fibers is related to the output of filter c_k. The input on the dendrites of each neuron is summated in the soma and related to this summated input action potentials are generated peripheral to the soma.

The probability of creating an action potential in neuron k, $g_k(z_k = 1|\mathbf{x})$ depends on the summated input: the generator potential $v_k(\mathbf{x})$.

$$v_k(\mathbf{x}) = b_k + (\mathbf{c}_k, \mathbf{x}), \tag{29}$$

$$g_k(z_k = 1|\mathbf{x}) = \frac{\exp(v_k(\mathbf{x}))}{1 + \exp(v_k(\mathbf{x}))}. \tag{30}$$

The approximation introduced by Equation (29) is supported by the result of Moller [27] and Grashuis [11] that neurons in the auditory nerve with a characteristic frequency below 1 kHz are characterized by their first-order Wiener kernels c_k. For a more detailed description see [4], [7], and [15].

Since there are no data available about the precise structure of the stimulus space S of all auditory stimuli, we assume that all stimuli with equal energy are equally likely to occur and that stimuli with higher intensity are less probable according to a Gaussian distribution. This reduces Equation (7) to

$$f(\mathbf{x}) = f_0 \exp\{-\frac{1}{2}\mathbf{x}I\mathbf{x}\} = f_0 \exp\{-\frac{1}{2}(\mathbf{x}, \mathbf{x})\}, \tag{31}$$

where I represents the identity matrix which is the covariance matrix of Gaussian white noise.

A set of 64 neurons with characteristic frequencies from 200 Hz to 2000 Hz equidistant on a linear scale was simulated. The experimental procedure was that to each neuron was presented the same stimulus and the neural response was recorded. The set of stimuli consisted of a frequency sweep of constant amplitude from 0 to 2000 Hz, a series of positive and negative clicks and a set of γ-tones [1]. Then in every time interval $[t, t + \Delta t)$ a stimulus $\mathbf{x}'$ was calculated according to the equation

$$\mathbf{x}' = \sum_{k=1}^{K} \mathbf{c}_k z_k . \tag{32}$$

These separate signals were simply added, following the additive procedure suggested by Johannesma [16]. This implies that Equation (32) is a first-order approximation of the most plausible stimulus and that for every action potential of neuron k the first-order kernel c_k is substituted.

Results for the frequency sweep are shown in Figure 2. The reconstructed stimulus appears to have a reasonable resemblance to the presented stimulus. Only during the onset of the frequency sweep there is a considerable difference.

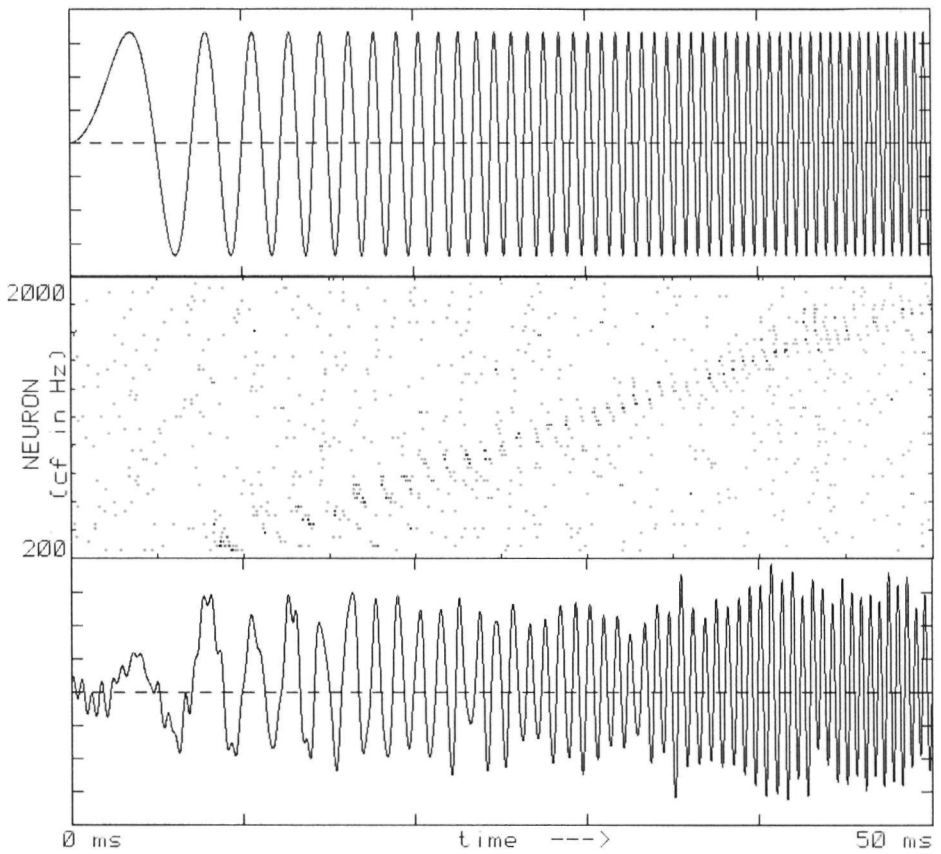Figure 2: *Frequency sweep. Top: the stimulus, a frequency sweep up to 2000 Hz as presented to the model. Middle: the neural activity pattern associated with the frequency sweep. The neural activity pattern of each neuron is plotted as a sequence of dots on a line. Activity of neurons with a different characteristic frequency (cf) is plotted at different lines. Bottom: the estimated most plausible stimulus.*
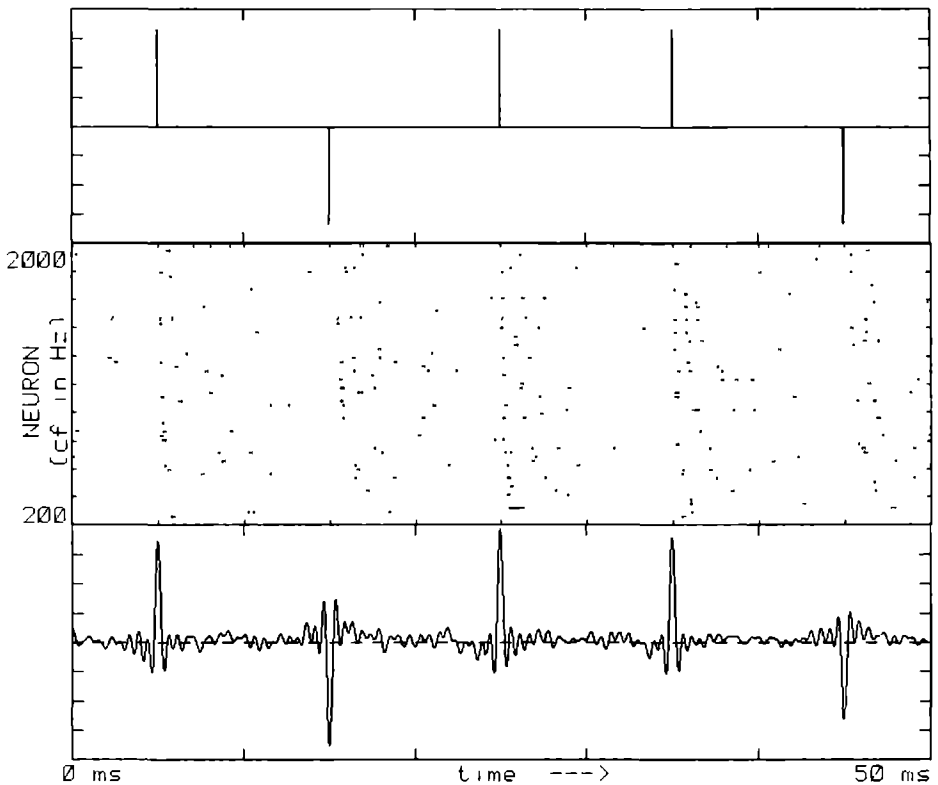
Figure 3: *Click sequence. Top: the stimulus consisting of 5 clicks (three positive, two negative). Middle: the neural activity pattern. The neural activity pattern of each neuron is plotted as a sequence of dots on a line. Activity of neurons with a different characteristic frequency (cf) is plotted at different lines. Bottom: the estimated most plausible stimulus.*

This can be explained by the lack of neurons in the model with a characteristic frequency below 200 Hz. The fluctuation in the overall amplitude of the reconstructed signal can be explained by two arguments. One reason is that there are frequency gaps of about 30 Hz between the consecutive neurons thus producing gaps in the frequency content of the reconstructed stimulus. Another reason is that 60% of the approximately 1000 simulated events can be attributed to the spontaneous incoherent activity of the neurons. Increasing the number of neurons in the neural population reduces this stochastic effect. The stimulus reconstruction for the clicks and γ-tones are shown in Figures 3, 4, and 5.

The most obvious failure of the reconstructions is that low and high frequencies are lost in the reconstructed signal. The explanation for this has been given above. However, we conclude that although quantitative differences exist, there is a general qualitative similarity.
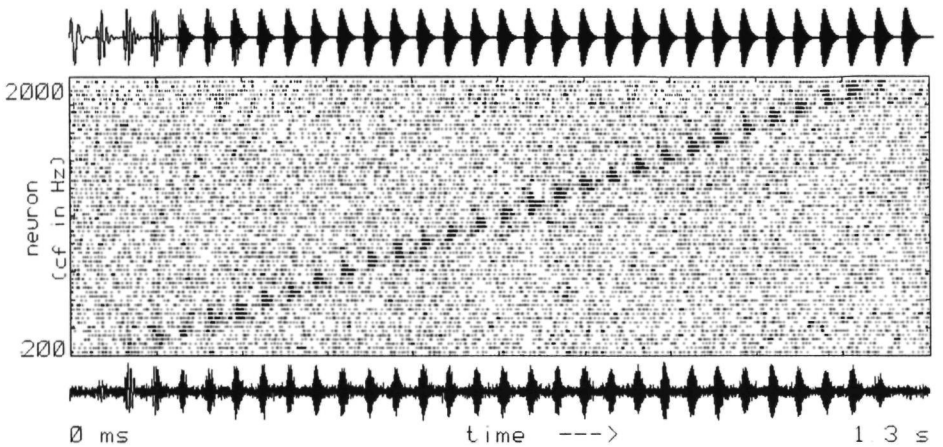
Figure 4: γ-tones. Top: the stimulus consisting of 32 γ-tones with characteristic frequencies from 96 Hz to 2080 Hz, with increments of 64 Hz. Middle: the neural activity pattern. The neural activity pattern of each neuron is plotted as a sequence of dots on a line. Activity of neurons with a different characteristic frequency (cf) is plotted at different lines. Bottom: the estimated most plausible stimulus.

Real neurons with a characteristic frequency above 1 kHz demonstrate an increasing loss of phase-lock. As a consequence these neurons cannot be completely characterized by a first-order kernel c_k but require a second-order kernel D_k for an adequate characterization. This loss of phase-lock is a result of the filtering properties of the inner hair cells and of the synaptic transmission, which acts as a rectifying mechanism followed by a low-pass filter with a cut-off frequency near 1 kHz. Physiologically the need of a second-order kernel is related to the fact, that the neural response is not anymore related to the phase of the stimulus, a phenomenon called loss of 'phase-lock' [21]. As a consequence no first-order Wiener-Volterra kernel exists and no stimulus reconstruction would be possible. This phenomenon is precisely predicted by the theory: if second-order kernels are involved, the most plausible stimulus consists of a set of stimuli, a hyperellipsoid in the stimulus space S, which contains all stimuli with the same frequency content but with variable phase relations between the frequency components. In this condition the first-order approximation (32) for the reconstructed signal is not satisfactory anymore. Therefore, for neurons with a characteristic frequency above 1 kHz it has been proposed to introduce the analytic signal ξ in (9), (10), (11), (13) and (14) instead of $\mathbf{x}$. The analytic signal is defined by:

$$\xi = \mathbf{x} + i\tilde{\mathbf{x}} \tag{33}$$

where $\tilde{\mathbf{x}}$ is the Hilbert transform of the signal $\mathbf{x}$ ([1]).

Although it does not add more information, it might suggest an alternative representation, which incorporates the phase of the stimulus.
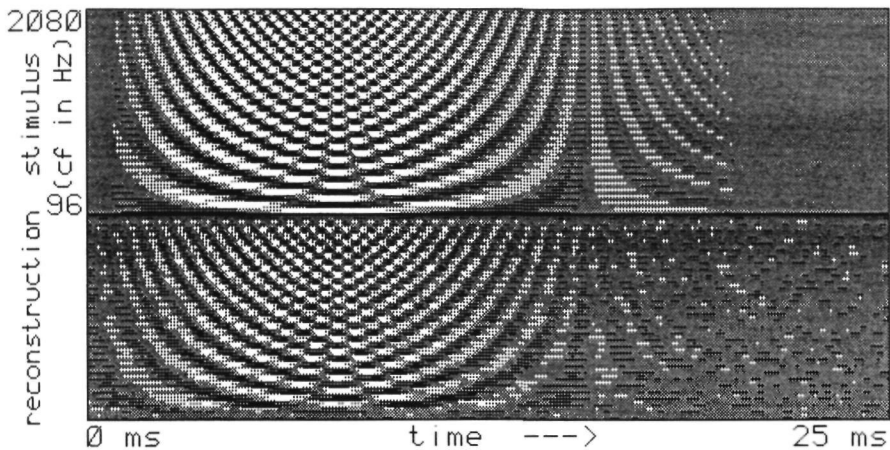
21

Figure 5: *γ-tones (grey scale). Enlargement of stimulus and stimulus reconstruction as shown in figure 4. Each of the 32 rows represents one tonepip (γ-tone). The instantaneous amplitude $x(t)$ of the signal is projected on a linear grey scale (white negative, dark positive). Top: the stimulus, 32 γ-tones. Bottom: the estimated most plausible stimulus.*

In fact the reconstructed stimulus obtained from Equation (32) is a first approximation to the asymptotic solution given by Equation (28). Evidently how well this first-order approximation describes the reconstructed stimulus will depend on the type of stimulus. However, the fact that for the three types of stimuli used in this simulation the result is in reasonable agreement with the presented stimulus indicates that the result does not critically depend on the type of stimulus.

5   Discussion

The aim of this paper is to present a tentative solution for the problem how to interprete a neural activity pattern. In traditional single-unit recordings a relation between stimulus and neural activity can be obtained by correlation techniques. However, it is generally accepted that study of neural information processing in central stages of the nervous system requires the evaluation of *multi-unit* activity. Procedures to interpret the physiological significance of this multi-unit activity are suggested in this paper. It does not pretend to present a solution used by the brain but rather to present a solution that is useful for the experimenter.

The number of assumptions which form the base of this approach is rather small. One important assumption is that neurons do not influence each other, neither by direct mutual neural connections, nor by some kind of feedback mech-

22

anism. This assumption does not exclude a correlation between neural activity in auditory nerve fibers. A correlation may be induced by common input to auditory nerve fibers, e.g. from identical hair cells. This is particularly true for auditory nerve fibers which characteristic frequency if very close. The present assumption excludes a correlation due to neural interactions by collaterals. Although this assumption has been confirmed for the auditory nerve under stationary stimulus conditions, such as Gaussian white noise [17], this assumption is in general not correct. In the auditory system there is an efferent feedback to the muscles of the inner ear, which acts as an automatic gain control and which protects the cochlea from damage by excessive sound volumes. In the visual system there is the pupil, which compensates to a moderate extent the rapid variations in light intensity. However, the assumption of independency of neural activity is convenient for mathematical reasons, but it is not essential to this theory. The main reason, which led us to this assumption, is the lack of knowledge about neural interactions. If neural interactions are incorporated, an additional term has to be added to Equations (13, 14):

$$G(\mathbf{z},\mathbf{x}) = (\mathbf{z},\mathbf{v}(\mathbf{x})) + (\mathbf{z},W(\mathbf{x})\mathbf{z}) \tag{34}$$

in which the neural interactions are incorporated in the second term. Especially if one tries to describe neural information processing in more central structures in the nervous system the second term will gain in importance. A special type of neural interaction which can be included in the second term of Equation (34), refers to the refractory period of a neuron. In fact, this is equivalent to a nonlinear neural interaction of a neuron with itself. However, as explained earlier this nonlinearity does not affect the responses appreciably for physiologically relevant stimuli, i.e. as long as stimulus intensities are not excessively high.

The fact that in general a nonlinear expression for the neural response properties in Equation (15) gives rise to more than one most plausible stimulus, may be a common phenomenon in nature. For example, auditory neurons, which participate in sound localization may be triggered equally well by stimuli in the same direction but at different distances, whatever the frequency of the sound. Very likely, as soon as general features are extracted from the general neural representation of the stimulus, some stimulus properties are lost and a reinterpretation will give rise to a class of stimuli with some common properties.

Another assumption was that the probability function, which describes the probability that a specific stimulus occurs, is characterized by the first two moments of the stimuli in the stimulus space S. This assumption clearly is a simplification since the function $f(\mathbf{x})$ may depend on previous stimuli with some weighting factor in time and probably also requires a term, which describes the interaction between stimuli from different sensory modalities, such as for example auditory, tactile, visual, etc. Other evidence for a more complicated function f comes from theories, supported by Mackay [24] and Sommerhoff [37], which state that a subject builds up an internal representation of the outer world. In this view the subject has some expectancy about the probability of

occurrences in this internal representation and incoming sensory information is used to check whether it corresponds, within certain error boundaries, with this internal representation. In this view the function f is a very complicated function of present and previous stimuli in a broad context. Up till now not enough is known about the psychological processes, involving perception and the creation of an internal representation, to specify the mapping in more detail. The aim of this paper is to present a possible interpretation of neural activity obtained in neurophysiological experiments. It does not pretend to simulate processes.

A practical disadvantage of this method might be, that it requires knowledge of the Wiener-Volterra kernels and the neural response to a particular stimulus for all neurons in a sensory system, something which is impossible to realize in normal experimental situations. One way to solve this problem is to use the fact that neurons in one nucleus of the brain do differ in their quantitative properties but usually have common qualitative properties. Even though neighbouring neurons differ in morphology and electrotonic properties, giving rise to very different response properties, usually a rather limited number of cell types can be distinguished within a particular nucleus. Therefore, a good insight in the response properties may be obtained by recording from a rather limited sample of cells in a nucleus. Moreover the quantitative properties of neurons within a particular cell type always change gradually, never discontinuously. If the variations in the quantitative properties are known, one may extrapolate neural properties obtained from a limited sample of neurons to the whole population. A similar problem is seen in statistical physics, where the state of a gas (macro state) is defined by a probability function, which depends on the micro-states of the individual elements in a statistical way.

A fundamental problem in the sensory interpretation of neural activity patterns is the relation between the *a priori* distribution of the stimulus ensemble as presented to the animal and the *a posteriori* distribution of the ensemble of stimuli associated with the occurrence of this pattern of neural activity. The relation of the original distribution and the one estimated from neural activity is in general complex. A less ambitious approach is to use just a few descriptive parameters of both distributions. Which type of parameters to choose, should be guided by *a priori* knowledge about the stimulus and sensory system. Possible descriptive quantities are moments and cumulants. If the distribution is Gaussian then the first moment or average signal equals the most plausible signal. The second moment, which is closely related to the concept of the receptive field, may be applied in case of neurons without phase-lock. Use of the estimation of the second moment leads to a spectro-temporal characterization of the stimulus [14]. An approach complementary to the detailed study of the estimation of a particular stimulus, is to look into the more global transformation properties of the sensory system. A comparison of the entropy of the *a priori* stimulus ensemble and the *a posteriori* stimulus ensemble should supply a measure for the information content of the neural activity.

References

1. A. M. J. H. Aertsen and P. I. M. Johannesma, Spectro-temporal receptive fields of auditory neurons in the grassfrog, *Biol. Cybernet.* 38:223-234 (1980).

2. G. von Békésy, *Experiments in Hearing*, McGraw-Hill, New York, 1960.

3. E. de Boer and H. R. de Jongh, On cochlear encoding: potentialities and limitations of the reverse-correlation technique, *J. Acoust. Soc. Am.* 63:115-135 (1978),

4. H. F. P. van den Boogaard, G. H. F. M. Hesselmans, and P. I. M. Johannesma, System identification based on point processes and correlation densities. I. The nonrefractory model, *Mathematical Biosciences* 80:143-171 (1986).

5. G. Buchsbaum and J. L. Goldstein, Optimum probabilistic processing in colour perception. Part I. Colour discrimination, Part II. colour vision as template matching, *Proc. Roy. Soc. London B.* 205:229-266 (1979).

6. H. S. Colburn, Theory of binaural interaction based on auditory-nerve data. I. General strategy and preliminary results on interaural discrimination. *J. Acoust. Soc. Am.* 54:1458-1470 (1973).

7. H. Davis, Biophysics and physiology of the inner ear, *Physiol. Rev.* 37:1-49 (1957).

8. R. Fitzhugh, A statistical analyzer for optic nerve messages, *J. Gen. Physiol.* 41(4):675-692 (1958).

9. C. C. A. M. Gielen, J. A. M. van Gisbergen, and A. J. H. Vendrik, Characterization of spatial and temporal properties of monkey LGN Y-cells, *Biol. Cybern.* 40:157-170 (1981).

10. C. C. A. M. Gielen, J. A. M. van Gisbergen, and A. J. H. Vendrik, Reconstruction of cone-system contributions to responses of colour-opponent neurones in monkey lateral geniculate, *Biol. Cybern.* 44:211-221 (1982).

11. J. L. Grashuis, *The Pre-Event Stimulus Ensemble: An Analysis of the Stimulus-Response Relation for Complex Stimuli Applied to Auditory Neurons*, Thesis, University of Nijmegen, 1974.

12 D. D. Greenwood, Approximate calculations of the dimensions of the travelling-wave envelopes in four species, *J. Acoust. Soc. Am.* 34:1364-1369 (1962).

13. P. I. M. Johannesma, Neural representation of sensory stimuli and sensory interpretation of neural activity, Neural communication and control, in *Advances in Physiol. Sciences*, Vol. 30: *Neural Communication and Control.* pp. 103-126, Pergamon Press, 1981.

14. P. I. M. Johannesma and A. M. H J Aertsen, The phonochrome—a coherent spectro-temporal representation of sound, in *Localization and Orientation in Biology and Engineering*, pp. 42-45, (Varju and Schnitzler Eds.), Springer-Verlag, Berlin, 1984.

15. P. I. M. Johannesma and H. F. P. van den Boogaard, Stochastic formulation of neural interaction, *Acta Appl. Math.* 4:201-224 (1985).

16. P. I. M. Johannesma, A M. H. J. Aertsen, H. F. P. van den Boogaard, J. J. Eggermont, and W. J. M. Epping, From synchrony to harmony: Ideas on the function of neural assemblies and on the interpretation of neural synchrony, in *Brain Theory* (G Palm and A. Aertsen, Eds.), Springer, 1986.

17. D. H. Johnson and N. Y. S. Kiang, Analysis of discharges recorded simultaneously from pairs of auditory nerve fibers, *Biophys. J.* 16.719-734 (1976).

18. H. R. de Jongh, *Modelling the Peripheral Auditory System*, Thesis, University of Amsterdam, 1978.

19. D. H. Kelly, Spatio-temporal frequency characteristics of color-vision mechanisms, *J. Opt. Soc. Am.* 64:983-990 (1974).

20. D. H. Kelly and D. van Norren, Two-band model of heterochromatic flicker, *J. Opt. Soc. Am.* 67:1081-1091 (1977).

21. N. Y. S. Kiang, *Discharge Patterns of Single Fibers in the Cat's Auditory Nerve* M. I. T. Press, Cambridge, Mass., 1965.

22. J. J. Koenderink and A. J. van Doorn, Visual detection of spatial contrast; influence of location in the visual field, target extent and illuminance level, *Biol. Cybern.* 30:157-167 (1978).

23. Y. W. Lee and M. Schetzen, Measurement of the Wiener-kernels of a nonlinear system by crosscorrelation, *Int. J. Control*, 2:237-254 (1965).

24. D. M. Mackay, *Information, Mechanism and Meaning*, M. I. T. Press, Cambridge, Mass., 1970.

25. P. Z. Marmarelis and V. Z. Marmarelis, *Analysis of Physiological Systems*, Plenum Press, New York, 1978.

26. P. Z. Marmarelis and K. I. Naka, Nonlinear analysis and synthesis of receptive-field responses in the catfish-retina. I. Horizontal cell-ganglion cell chain, *J. Neurophysiol.* 36:605-618 (1973).

27. A. R. Møller, Dynamic properties of the responses of single neurons in the cochlear nucleus of the rat, *J. Physiol.* 259:63-82 (1976).

28. F. L. van Nes and M. A. Bouman, Spatial modulation transfer in the human eye, *J. Opt. Soc. Am.* 57:401-406 (1967).

29. D. P. O'Leary, R. F. Dunn, and V. Honrubia, Analysis of afferent responses from isolated semicircular canal of the guitar fish using rotational acceleration white noise inputs, Part I: Correlation of response dynamics with receptor innervation, Part II: Estimation of linear system parameters and gain and phase spectra, *J. Neurophysiol.* 39:631-659 (1976).

30. G. Österberg, Topography of the layer of rods and cones in the human retina. *Acta Ophthal.* (Kbh). Supp. 6 (1935).

31. R. W. Rodieck and J. Stone, Responses of cat retinal ganglion cells to moving visual patterns, *J. Neurophysiol.* 28:819-832 (1965a).

32. R. W. Rodieck and J. Stone, Analysis of receptive fields of cat retinal ganglion cells. *J. Neurophysiol.* 28:833-849 (1965b).

33. I. J. Russell and P. M. Sellick, Intracellular studies of haircells in the mammalian cochlea, *J. Physiol.* (Lond.) 284:261-290 (1978).

34. N. A. M. Schellart and H. Spekreijse, Dynamic characteristics of retinal ganglion cell responses in goldfish, *J. Gen. Physiol.* 59:1-21 (1972).

35. H. Schubert, *Topologie, eine Einführung*, B. G. Teubner Verlagsgesellschaft, Stuttgart, 1964.

36. W. M. Siebert, Stimulus transformation in the peripheral auditory system, in *Recognizing Patterns*, (P. A. Kolers and M. Eden, Eds.), M. I. T. Press, Cambridge, Massachusetts, 1968.

37. G. Sommerhoff, *Logic of the Living Brain*, Wiley & Sons, New York, 1974.

The Characteristic Functional of the Peri-event Stimulus Ensemble

Gerard H. F. M. Hesselmans   Henk F. P. van den Boogaard
Peter I. M. Johannesma

Abstract

A neuron is considered as a stimulus selective system. Stimuli from the stimulus ensemble (SE) presented to the neuron, which are associated with the generation of an action potential, form the peri-event stimulus ensemble (PESE). The relation between PESE and SE is analysed on the basis of characteristic functionals. Examples are given for different models of the neuron. For a multiplicative model a simple relation between the characteristic functional of stimulus ensemble and peri-event stimulus ensemble exists.

1 Introduction

In neurophysiological experiments the activity of one or more neurons is measured in response to sensory stimuli. Neural responses consist of sequences of electric pulses (spikes or action potentials), whose shape may be characteristic for the type of neuron. Although differences of shape can be used to separate the activity of simultaneously recorded neurons (Eggermont et al. [9]), the form of the action potential is assumed to be of secondary physiological importance. Since the duration is short compared with the time interval between successive spikes, the occurrence of an action potential may be considered as an event in time. As such, a sequence of action potentials can be seen as a realization of a stochastic point process $n(\cdot)$.

Analysis of the transformation of the stimulus to a point process can be made by reverse correlation methods (Boer and Kuyper [4]; Eggermont et al. [10]). In this analysis the peri-event stimulus ensemble (PESE) (Johannesma [13]; Aertsen [1, 2]; Hermes [12]) is compared with the stimulus ensemble (SE) presented to the animal. If the stimulus influences the neural activity, then the structure of the ensemble selected by the neuron (X_e) differs from the one presented to the neuron (X). Therefore the neuron may be considered as a system mapping the SE into the PESE by selection; this point of view is illustrated in Figure 1.
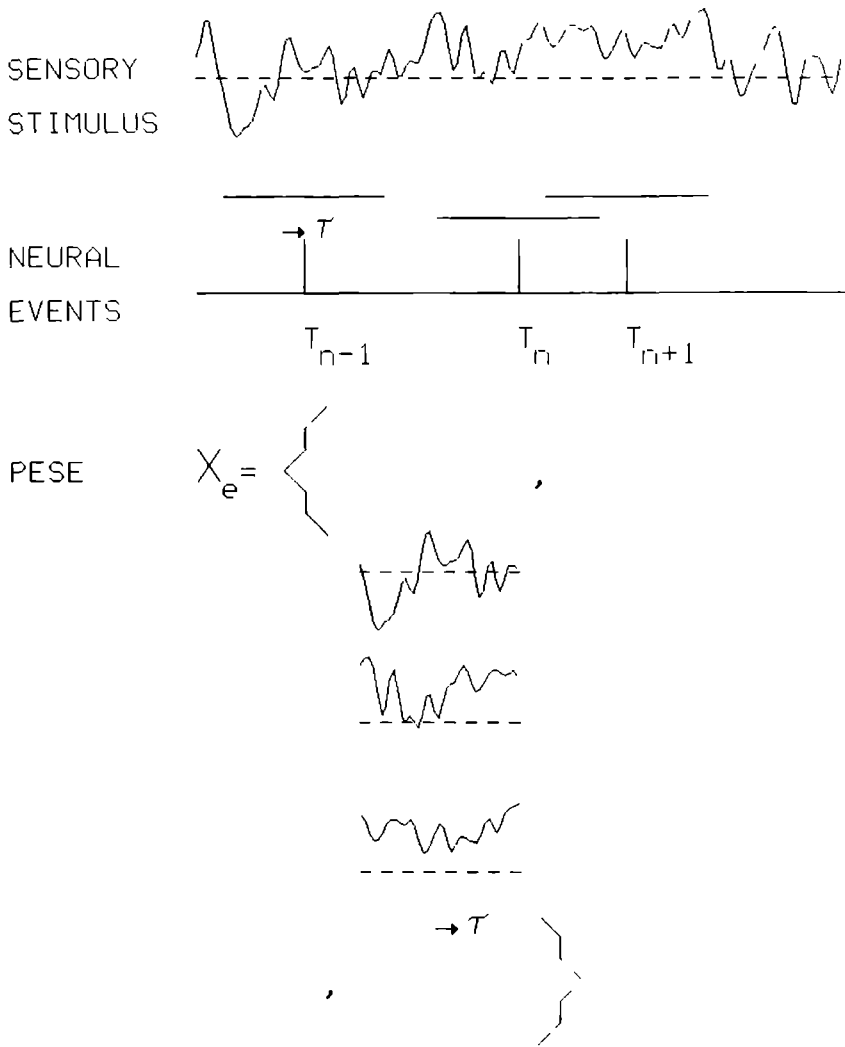
Figure 1: *Construction of the PESE from a realization of the stimulus process* **x**(·).

Since this map is determined by the dynamics of the neuron, comparison of SE and PESE may give insight into properties of the neuron. The sensitivity of the neuron to certain stimuli can be established, and sometimes even system identification is possible.

The goal of this investigation is to find an exact mathematical description of the relation of PESE and SE for certain classes of neurons. This relation then forms a complete characterization of the neuron and a formal base for the concept of receptive field, as used widely in neurophysiology. The discussion is based on product densities and the characteristic functional, which are introduced in Section 2. The general result given in Section 3 imposes almost no restrictions upon the model. Some examples for a model of the neuron are included in Section 4.

2 Representation of a stochastic process

the characteristic functional

The output process of the neuron, a sequence of action potentials, will be represented by a point process; however, the stimulus input can still be a continuous (Gaussian) process or a (Poisson) point process. Several descriptions of a stochastic process are available. One that completely determines an arbitrary stochastic process is the characteristic functional (Snyder [19]; van Kampen [15]; Srinivasan [20]), which is a generalization of the characteristic function (Lukacs [17]). Let the stimulus $x(\cdot)$ be a continuous stochastic process defined on the sample space X with probability distribution P, and $\xi(\cdot)$ an appropriate complex-valued test function. Then the characteristic functional $\Phi[\xi]$ is defined by the following expectation value:

$$\Phi[\xi] := E_X \left[\exp \int ds\, x(s)\xi(s) \right] , \tag{1}$$

where the expectation is taken with respect to $\chi = \{X, P\}$. Knowledge of the characteristic functional $\Phi[\xi]$ for an appropriate set of test functions $\xi(\cdot)$ supplies a complete description of the stochastic process $x(\cdot)$. An equivalent characterization of the process is given by the logarithm of the characteristic functional

$$\Psi[\xi] := \ln \Phi[\xi] = \ln E_X \left[\exp \int ds\, x(s)\xi(s) \right] . \tag{2}$$

On the other hand, the stimulus might be a point process, e.g. the sequence of action potentials generated by another neuron. In that case $x(\cdot)$ is a sequence of delta functions:

$$x(t) = \sum_{j \in Z} \delta(t - t_j) , \tag{3}$$

30

where the $\{t_j\}_{j \in Z}$ denote the (stochastic) times of occurrence of an input event. In the literature on point processes it is common to introduce the associated counting process defined by

$$\mathbf{N}(t) = \int_0^t \mathbf{x}(s)ds. \tag{4}$$

As a consequence $\mathbf{x}(s)ds$ in Equation (1) is replaced by $d\mathbf{N}(s)$, and in the format of point-process theory the characteristic functional becomes

$$\Phi[\xi] = E_\mathbf{x} \left[\exp \int d\mathbf{N}(s)\xi(s) \right]. \tag{5}$$

moments and product densities

In theory the space of test functions $\xi(\cdot)$ is infinite-dimensional. However, this space has a denumerably infinite base; therefore a reduction to a finite-dimensional subspace is possible. The dimensionality of the subspace is determined by the point process under consideration. Since the number of relevant dimensions is usually large, the characteristic functional Φ of a stochastic process $\mathbf{x}(\cdot)$ is a function of many variables. As a consequence it is difficult to estimate from experimental data. Therefore we look for characteristics of the process more directly related to experimental observables; these can be found in moments or cumulants. The nth moment of a stochastic process $\mathbf{x}(\cdot)$ is defined as

$$m_n(t_1, \ldots, t_n) := E_\mathbf{x} \left[\prod_{i=1}^n \mathbf{x}(t_i) \right]. \tag{6}$$

The moments m_n are the expansion 'coefficients' of the characteristic functional $\Phi[\xi]$ at $\xi = 0$. From $\Phi[\xi]$ the moments can be derived by functional differentiation (Stratonovich [21]). Therefore the characteristic functional is sometimes called moment generating functional (Billingsley [3]; van Kampen [15]):

$$m_n(t_1, \ldots, t_n) = \left(\frac{\delta}{\delta\xi} \right)^n \Phi[\xi] \Big|_{\xi=0}, \tag{7}$$

where the following abbreviation has been used:

$$\left(\frac{\delta}{\delta\xi} \right)^n = \frac{\delta}{\delta\xi(t_1)} \frac{\delta}{\delta\xi(t_2)} \cdots \frac{\delta}{\delta\xi(t_n)}. \tag{8}$$

The cumulants c_n are the combinations of moments such that information contained in low-order cumulants is, to a large extent, eliminated in high-order cumulants. The cumulants are the expansion 'coefficients' of $\Psi[\xi]$. They can be derived from Ψ by functional differentiation

$$c_n(t_1, \ldots, t_n) = \left(\frac{\delta}{\delta\xi} \right)^n \Psi[\xi] \Big|_{\xi=0} \tag{9}$$

31

If the stochastic process $x(\cdot)$ is stationary, then the moments and cumulants no longer depend on absolute moments of time, but only on differences of time:

$$m_n(t_1,\ldots,t_n) \quad \rightarrow \quad m_n(\tau_1,\ldots,\tau_{n-1}),$$
$$c_n(t_1,\ldots,t_n) \quad \rightarrow \quad c_n(\tau_1,\ldots,\tau_{n-1})$$

with $\tau_i = t_{i+1} - t_i$.

In the case of a point process $n(\cdot)$ moments are replaced by (factorial) product densities (Stratonovich [21]). At different times these are defined by

$$f_n(t_1,\ldots,t_n) := \lim_{\max \Delta t_i \downarrow 0} E \left[\prod_{j=1}^{n} \frac{\Delta N(t_j)}{\Delta t_j} \right], \tag{10}$$

where $\Delta N(t)$ is the increment of the counting process $N(\cdot)$ over the time increment Δt . Factorial product densities will be used instead of ordinary product densities because they show no delta functions on their diagonals. They are related to the characteristic functional by (see Kuznetsov and Stratonovich [16])

$$\Phi[\xi] = 1 + \sum_{n=1}^{\infty} \frac{1}{n!} \int ds_1 \cdots ds_n \, f_n(s_1,\ldots,s_n) \prod_{j=1}^{n} \left[e^{\xi(s_j)} - 1 \right]. \tag{11}$$

The factorial product densities $f_n(\cdot)$ can be found by functional differentiation of $\Phi[\xi]$ (Stratonovich [21]). In the same way the cumulants are translated by factorial cumulant densities. The factorial cumulant densities are the expansion 'coefficients' of $\Psi[\xi]$:

$$\ln \Phi[\xi] = \Psi[\xi] = \sum_{n=1}^{\infty} \frac{1}{n!} \int ds_1 \cdots ds_n \, h_n(s_1,\ldots,s_n) \prod_{j=1}^{n} \left[e^{\xi(s_j)} - 1 \right]. \tag{12}$$

Cumulant densities and product densities are related to one another in the same way as cumulants and moments, e.g.

$$h_1(t) = f_1(t), \tag{13}$$
$$h_2(t,s) = f_2(t,s) - f_1(t)f_1(s). \tag{14}$$

3 Representation of the peri-event stimulus ensemble

the characteristic functional

The aim of this paper is to find the characteristic functional $\Phi^r[\xi](t)$ of the PESE, the characteristic functional of the SE, defined in Equation (1) ,is known:

$$\Phi^r[\xi] := E_{\chi_e} \left[\exp \int ds \, x(s)\xi(s) \right], \tag{15}$$

where the expectation is taken with respect to $\chi_e = \{X, P_r\}$. Realizations of the PESE are elements of the same sample space X with different probability density P_r [see Equation (1)]. χ_e is the conditional measure, conditioned on the occurrence of an output event. The right-hand side of Equation (15) depends implicitly upon the time t at which the output event is supposed to occur via the PESE χ_e . The explicit time dependence becomes clear from

$$\Phi^r[\xi] := E_\chi \left[\exp \int ds\, \mathbf{x}(s)\xi(s) \,\middle|\, \Delta\mathbf{M}(t) = 1 \right] , \ \Delta t \text{ small enough}, \qquad (16)$$

where $\mathbf{M}(\cdot)$ is a realization of the output counting process, with an event about time t.

In order to arrive at explicit results assumptions have to be made about the nature of the SE and/or event generation. The first assumption concerns the SE: the stochastic process $\mathbf{x}(\cdot)$ forming the input to the neuron is assumed to be stationary. While this assumption does not appear to be necessary, it does simplify the relations. Since the time t indicating the time of occurrence of the output event can be set equal to zero without loss of generality, the label t will be omitted whenever possible. The second assumption is related to the neuron: only neurons whose output point process is orderly for arbitrary input are considered (Daley [8]; Snyder [19]). This is expressed in the condition

$$P[\Delta\mathbf{M}(t) = 1|x(\cdot)] \quad = \quad I[x](t)\,\Delta t + o(\Delta t)\,, \qquad (17)$$

$$P[\Delta\mathbf{M}(t) > 1|x(\cdot)] \quad = \quad o(\Delta t)\,, \qquad (18)$$

where

$x(\cdot)$ = realization of the stochastic input process $\mathbf{x}$,
$\Delta\mathbf{M}(t)$= increment of the counting process $\mathbf{M}(\cdot)$ over the time increment Δt ,
$I(\cdot)$ = nonnegative functional,
$I[x](\cdot)$ = intensity function [7].

Through Equations (17) and (18) the property of the output process being orderly has been made explicit; the probability that more than one action potential will occur in a small time increment Δt is negligible compared to the probability of there occurring just one event.

Under the assumption of orderliness and using conditional expectation, the characteristic functional of the PESE can be evaluated by taking expectation values over the SE (see Appendix):

$$\Phi^r[\xi] = \frac{E_\chi \left[I[\mathbf{x}] \exp \int ds\, \mathbf{x}(s)\xi(s) \right]}{E_\chi \left[I[\mathbf{x}] \right]} . \qquad (19)$$

Equation (19) forms the fundamental equation of this paper; it relates the SE, event generation, and the PESE. If the statistics of the input process $\mathbf{x}(\cdot)$
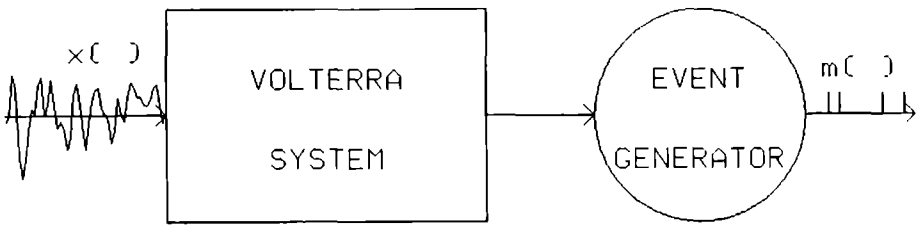
Figure 2: *Illustration of the model: the transformation of the stimulus process* $x(\cdot)$ *into a point process* $m(\cdot)$. *The event-generating system consists of a Volterra system followed by an instantaneous, nonnegative nonlinearity.*

and the event generating system $I[x](\cdot)$ are known, then the characteristic functional of the PESE can be calculated.

The denominator $E_x[I[x]]$ in Equation (19) is the intensity of the output point process. The characteristic functional $\Phi[\xi]$ of the PESE is the average of $\exp \int ds\, x(s)\xi(s)$ weighted by the intensity function $I[x](\cdot)$. If $\exp \int ds\, x(s)\xi(s)$ and $I[x]$ are uncorrelated, then input and output process are independent and $\Phi''[\xi] = \Phi[\xi]$.

moments and product densities

Using functional differentiation and Equation (19), the moments of the PESE can be derived from the characteristic functional; they are given by

$$m_n^c(t_1,\ldots,t_n) = \frac{E_x\left[I[x]\prod_{j=1}^{n} x(t_j)\right]}{E_x\left[I[x]\right]} . \tag{20}$$

The factorial product densities of the PESE in the case of an input point process $n(\cdot)$ are given by

$$f_n^c(t_1,\ldots,t_n) = \lim_{\max \Delta t_i \downarrow 0} \frac{E_x\left[I[n]\prod_{j=1}^{n} \frac{\Delta N(t_j)}{\Delta t_j}\right]}{E_x\left[I[n]\right]} . \tag{21}$$

A possible neural model in accordance with Equations (17) and (18) is a Volterra system followed by an event generator (see Section 4). In that case Equations (17) and (18) specify that the sequence of action potentials can be represented by a Poisson process whose intensity is governed by the stochastic stimulus process $x(\cdot)$, and the output process $m(\cdot)$ is a doubly stochastic Poisson process (Cox and Isham [7]). In Figure 2 an illustration of such a model is given.

It should be stressed here that while the stimulus ensemble is stationary and the neuron is time-invariant, the peri-event stimulus ensemble is nonstationary: its statistical characteristics depend on the time with respect to the occurrence of the action potential.

4 Examples

input process

In order to arrive at explicit results, assumptions have to be made about the nature of the SE and the event generation. We will restrict the input to Gaussian processes in the case of continuous processes, and to (Gauss-)Poisson processes in the case of point processes, since their characteristic functionals are relatively simple. In case of the Gauss-Poisson process (Milne and Westcott [18]) the first two cumulant densities are sufficient for a complete description of the process; higher-order cumulants are zero. The characteristic functional takes the form

$$\Psi[\xi] = \int ds\, h_1(s) \left[e^{\xi(s)} - 1 \right] + \frac{1}{2} \int dr \int ds\, h_2(r,s) \left[e^{\xi(r)} - 1 \right] \left[e^{\xi(s)} - 1 \right] . \quad (22)$$

For the simpler Poisson processes, $h_2(r,s) = 0$ and Equation (22) reduces to

$$\Psi[\xi] = \int ds\, h_1(s) \left[e^{\xi(s)} - 1 \right] . \quad (23)$$

Comparison of $f_2(r,s)$ and $f_1(r) \cdot f_1(s)$ gives an impression of the deviation of the process from the Poisson process [see Equation (14)]. If, as already assumed, the stochastic input process is stationary, then $f_1(s) \to f_1$ and $f_2(r,s) \to f_2(r-s) = f_2(s-r)$.

The formula (22) is comparable to the one for the Gaussian process, a continuous stochastic process (van Kampen [15]), since both processes are described by their first two cumulants:

$$\Psi[\xi] = \int ds\, c_1(s)\xi(s) + \frac{1}{2} \int dr \int ds\, c_2(r,s)\xi(r)\xi(s) . \quad (24)$$

the model

The model for the transformation of the input process into an output point process is chosen according to Johannesma and van den Boogaard [14] as elaborated by van den Boogaard et al. [5,6]. Briefly we recall that in this model the neuron performs a linear or nonlinear integration of the neural input $x(t)$. The result of this integration is a stochastic process $u(t)$ which represents the state of the neuron (generator potential). This analogue variable governs the generation of an output event. The mathematical formulation of the preceding reads
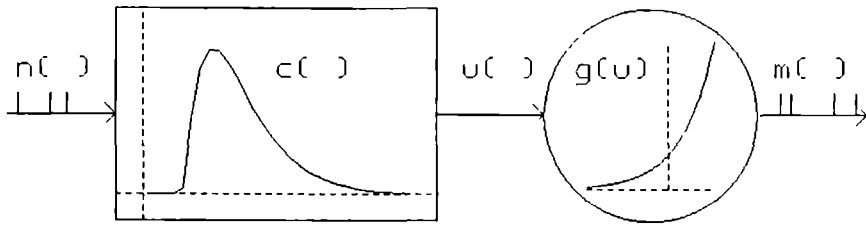
Figure 3: *Illustration of the model: the transformation of a point process* $\mathbf{n}(\cdot)$ *into another point process* $\mathbf{m}(\cdot)$. *The event-generating system consists of a linear filter followed by an instantaneous, exponential nonlinearity.*

$$U[\mathbf{x}] \quad = \quad \mathbf{u}(t) \tag{25}$$

$$= \quad u_0 + \int ds\, c(t-s)\mathbf{x}(s) + \frac{1}{2}\int dr \int ds\, D(t-r,t-s)\mathbf{x}(r)\mathbf{x}(s) + \cdots,$$

$$I[\mathbf{x}](t) \quad = \quad g(U[\mathbf{x}]) = g(\mathbf{u}(t)), \tag{26}$$

where

$$
\begin{array}{ll}
u_0 & = \text{working point of the neuron,} \\
c(t-s) & = \text{linear filter,} \\
D(t-r,t-s) & = \text{nonlinear filter of second degree,} \\
g(\cdot) & = \text{instantaneous, nonnegative nonlinearity.}
\end{array}
$$

In this paper the nonsaturating exponential pulse generation is chosen for the event-generating nonlinearity g. This is combined with a linear or a second-degree filter.

Example A (Linear filter and exponential event generation). For the event generation an exponential form has been chosen :

$$g(\mathbf{u}) = e^{\mathbf{u}}, \tag{27}$$

and as a consequence

$$I[\mathbf{x}] = e^{U[\mathbf{x}]}. \tag{28}$$

Substitution of Equation (28) into Equation (19) gives for the characteristic functional of the PESE

$$\Phi^r[\xi] = \frac{E_\mathbf{x}\left[\exp\left(U[\mathbf{x}] + \int ds\, \mathbf{x}(s)\xi(s)\right)\right]}{E_\mathbf{x}\left[\exp U[\mathbf{x}]\right]}. \tag{29}$$

A general relation for an arbitrary SE can be formulated if the filter $U[\mathbf{x}]$ is linear:

$$\mathbf{u}(t) = u_0 + \int ds\, \mathbf{x}(s)c(t-s). \tag{30}$$

36

If the moment of occurrence of the event is defined as $t = 0$, then Equation (30) can be written as

$$U[\mathbf{x}] = u_0 + \int ds\, \mathbf{x}(s)\gamma(s)\,, \tag{31}$$

where γ is the time-inverted form of $c(s)$:

$$\gamma(s) = c(-s)\,. \tag{32}$$

Substitution of Equation (31) into Equation (32) leads to the elimination of u_0, and the result is

$$\Phi^e[\xi] = \frac{E_x\left[\exp \int ds\, \mathbf{x}(s)\left\{\gamma(s) + \xi(s)\right\}\right]}{E_x\left[\exp U[\mathbf{x}]\right]}\,. \tag{33}$$

Recall of Equation (2) gives the relation

$$\Phi^e[\xi] = \frac{\Phi[\gamma + \xi]}{\Phi[\gamma]}\,, \tag{34}$$

or

$$\Psi^e[\xi] = \Psi[\gamma + \xi] - \Psi[\gamma]\,. \tag{35}$$

Equation (34) shows that for exponential event generation preceded by a linear filter the characteristic functional of the PESE can be expressed as the characteristic functional of the SE taken at a different position in ξ-space; this property holds for an arbitrary SE. Since moments or cumulants of the PESE can be expressed as functional derivatives of Φ^e or Ψ^e at $\xi = 0$, it follows from Equation (34) that these moments or cumulants are derivatives of the characteristic functional Φ or its logarithm Ψ evaluated at $\xi = \gamma$:

$$\left(\frac{\delta}{\delta\xi}\right)^n \Phi^e[\xi]\Bigg|_{\xi=0} = \left(\frac{\delta}{\delta\xi}\right)^n \frac{\Phi[\gamma + \xi]}{\Phi[\gamma]}\Bigg|_{\xi=0} = \left(\frac{\delta}{\delta\xi}\right)^n \frac{\Phi[\xi]}{\Phi[\gamma]}\Bigg|_{\xi=\gamma}\,. \tag{36}$$

In order to make this more explicit we consider a stationary Poisson process as SE. The logarithm of its characteristic functional is given by Equation (23); for the stationary Poisson process this becomes

$$\Psi[\xi] = f_1 \int ds \left[e^{\xi(s)} - 1\right]\,. \tag{37}$$

The associated PESE is then characterized by

$$\Psi^e[\xi] = f_1 \int ds\, e^{\gamma(s)} \left[e^{\xi(s)} - 1\right]\,. \tag{38}$$

Comparison of Equation (37) and Equation (38) indicates that the PESE is again a Poisson process, but now modulated with the intensity function

$$f_1^e(\tau) = f_1 e^{\gamma(\tau)}\,, \tag{39}$$

37

where τ is the time relative to the neural event.

To illustrate the feasibility of the above model, theoretical calculations and a simulation are compared with experimental results. The neural data were obtained from a single unit (unit 270-0.1-1) in the auditory midbrain (torus semicircularis) of the grassfrog (*Rana temporaria L.*). For details about the experimental set-up see Epping and Eggermont [11]. Equation (39) relates the intensity functions of SE and PESE to the filter. The filter was adapted in order to yield the least possible rms difference. The SE was a homogeneous Poisson point process with intensity 16 pulses per second. The linear filter was

$$c(s) = \begin{cases} 19\exp(-35.7s)[1 - \exp(-6.67s)] & \text{if } s > 0 \\ 0 & \text{if } s \le 0 \end{cases} \cdot \qquad (40)$$

Furthermore a delay of 37 msec was introduced. The results are shown in Figure 4. The first-order product densities in the first row of Figure 4 are quite similar, as should be expected from the above procedure. However, the agreement between the second-order product densities as obtained from the experimental data on the one hand and those obtained from theory and simulation on the other hand is in general not to be expected, unless there is some correspondence between the model and the behavior of the real neuron.

Secondly we consider a stationary Gauss-Poisson process as SE. Substitution of Equation (22) results in the logarithm of the characteristic functional of the PESE. From its general form it can be shown that the PESE is again a Gauss-Poisson process but, as should be expected, nonstationary. The (factorial) cumulant densities of the PESE can be computed in the normal way from $\Psi^c[\xi]$; they turn out to be

$$h_1^c(t) = h_1 e^{\gamma(t)} \left(1 + \int ds \, \frac{h_2(t,s)}{h_1} \left[e^{\gamma(s)} - 1\right]\right) \qquad (41)$$

$$h_2^c(t,s) = h_2(t,s)e^{\gamma(t)+\gamma(s)}. \qquad (42)$$

Simulations have been made to illustrate the theoretical results. For the SE a Gauss-Poisson process has been chosen with intensity f_1 of 8 events/second and second-order cumulant density

$$h_2(t,s) = \begin{cases} 40 & \text{if } |t - s| < 0.1 \text{ sec} \\ 0 & \text{if } |t - s| \ge 0.1 \text{ sec} \end{cases} \qquad (43)$$

The linear filter was

$$c(s) = \begin{cases} e^{-10s} & \text{if } s > 0 \\ 0 & \text{if } s \le 0 \end{cases}, \qquad (44)$$

and the event generator

$$g(u) = e^{u}. \qquad (45)$$

38
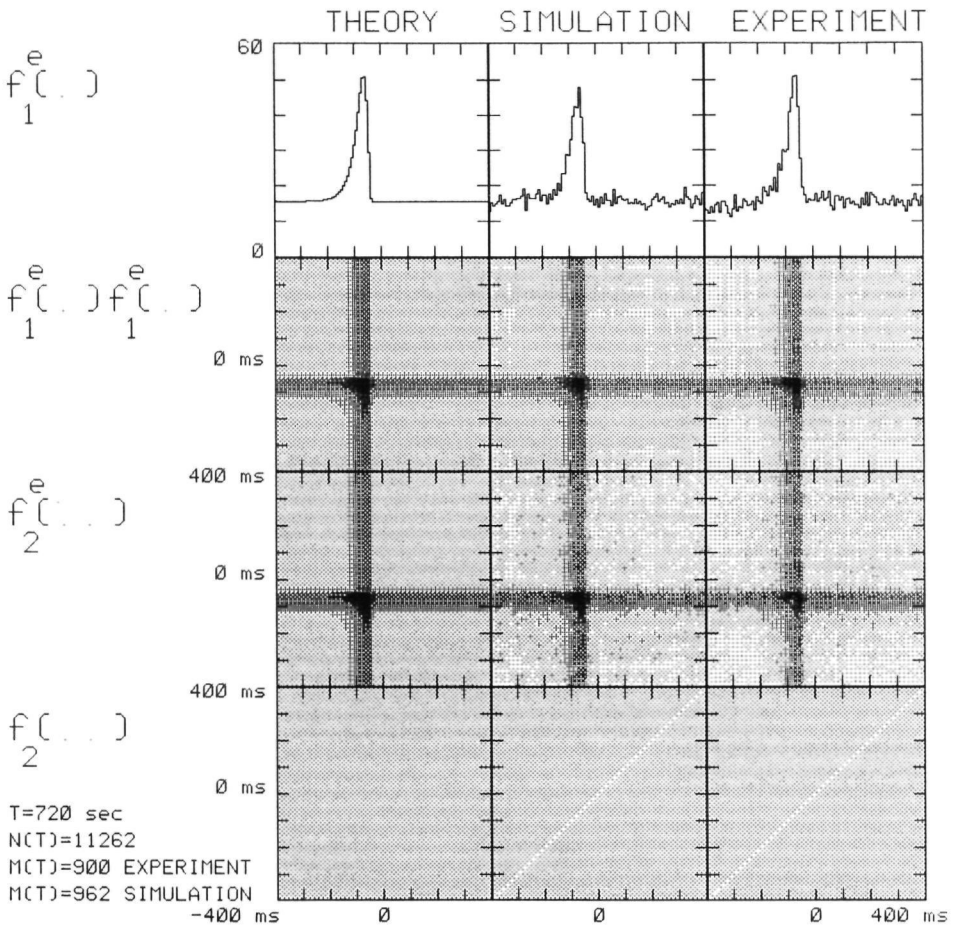
Figure 4: *Comparison of the product densities of the PESE as derived from theory, simulation, and experiment. The model used in theory and simulation consisted of a linear filter followed by an exponential event generator. The same stimulus, a homogeneous Poisson process, was used during the simulation and the experiment.*

39

Figure 5: *The transformation of a Gauss-Poisson process* $\mathbf{n}(\cdot)$ *(SE) into a Gauss-Poisson process* $\mathbf{m}(\cdot)$ *(PESE) by a multiplicative neuron. The results of a computer simulation are compared with calculations based on equations (41, 42).*

The output intensity was approximately 7.3 events/second.

The results in the left column of Figure 5 are computed from Equations (41, 42), while the right column of Figure 5 shows the results of the sim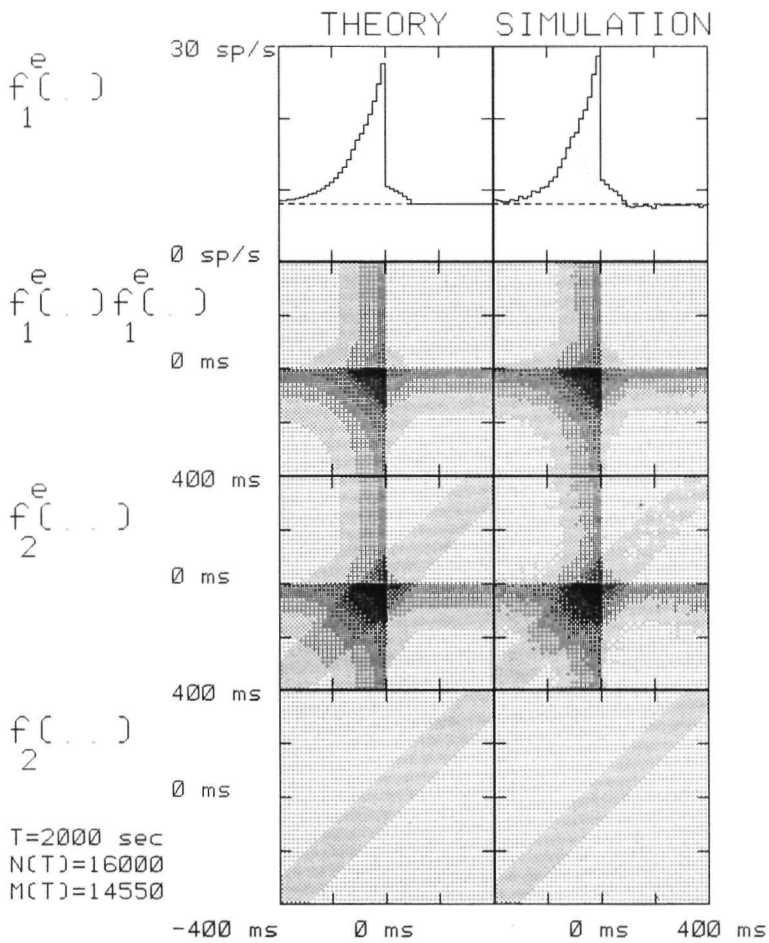ulation. Because both SE and model obey the theoretical assumptions, the agreement between left and right columns should be expected. The intensity of the PESE $f_1^r(\tau)$ reflects influence of the filter ($\tau < 0$) and of the structure of the SE ($\tau > 0$). The structure of the second-order product density $f_2^r(\sigma, \tau)$ appears to be a superposition of $f_1^r(\sigma) \cdot f_1^r(\tau)$ and the structure of the stimulus-ensemble $f_2(\sigma, \tau)$.

Example B (Second-degree filter and exponential event generation). In the previous section the filter was assumed to be linear as expressed in Equation (30). In this section we take a second-degree form

$$\mathbf{u}(t) = u_0 + \int ds\, c(t - s)\mathbf{x}(s) + \frac{1}{2} \int dr \int ds\, D(t - r, t - s)\mathbf{x}(r)\mathbf{x}(s)\,. \tag{46}$$

Again defining the moment of occurrence of the event to be $t = 0$, Equation (46) can be written

$$U[\mathbf{x}] = u_0 + \int ds\, \gamma(s)\mathbf{x}(s) + \frac{1}{2} \int dr \int ds\, \Delta(r, s)\mathbf{x}(r)\mathbf{x}(s)\,, \tag{47}$$

where

$$\gamma(s) = c(-s)\,, \quad \Delta(r, s) = D(-r, -s)\,. \tag{48}$$

In a compact notation we rewrite Equation (47) in the form

$$U[x] = u_0 + \gamma \circ \mathbf{x} + \frac{1}{2}\mathbf{x} \circ \Delta \circ \mathbf{x}\,. \tag{49}$$

Making use of Equation (29), we find for the characteristic functional of the PESE

$$\Phi^r[\xi] = \frac{E_\mathbf{x}\left[\exp\left\{\mathbf{x} \circ (\xi + \gamma) + \frac{1}{2}\mathbf{x} \circ \Delta \circ \mathbf{x}\right\}\right]}{E_\mathbf{x}\left[\exp\left\{\mathbf{x} \circ \gamma + \frac{1}{2}\mathbf{x} \circ \Delta \circ \mathbf{x}\right\}\right]}\,. \tag{50}$$

In this section the exact solution of Equation (50) will be given if $\mathbf{x}(t)$ is a Gaussian process with zero mean and covariance $B(t, s)$. Johannesma [13] solved the problem by discretizing time. In this way the stimulus is reduced to a n-dimensional subspace. Using the theory of multivariate normal distributions, the n-dimensional problem can be solved. However, the transition from discrete time back to continuous time is not clear, since the behavior of the process $\mathbf{x}(\cdot)$ between sample points is not taken into account.

Another way of solving the problem posed by Equation (50) is by expanding the stimulus $\mathbf{x}(\cdot)$ in the eigenfunctions $e_j(\cdot)$ of its covariance:

$$\mathbf{x}(t) = \sum_{j=1}^{\infty} \mathbf{x}_j e_j(t) \tag{51}$$

41

with

$$\mathbf{x}_j = \int dt\, \mathbf{x}(t) e_j(t) . \tag{52}$$

The expansion coefficients $\mathbf{x}_j$ are independent and normally distributed. The eigenfunctions $e_j(t)$ are assumed to be sorted by decreasing eigenvalue λ_j :

$$\lambda_j = E\left[\mathbf{x}_j^2\right] \geq \lambda_{j+1} = E\left[\mathbf{x}_{j+1}^2\right] , \text{ all } j \in N , \tag{53}$$

and the power of the stimulus is finite:

$$E\left[\int ds\, \mathbf{x}(s)^2\right] = \sum_{j=1}^{\infty} \lambda_j < A , \; A \in R^+ . \tag{54}$$

Under these assumptions it is possible to choose an n-dimensional subspace such that the power contained in its orthogonal complement is arbitrarily small:

$$\sum_{j=n+1}^{\infty} \lambda_j < \varepsilon . \tag{55}$$

The expectation values in Equation (50) can be evaluated in two independent steps. The contribution of the first n expansion coefficients can be calculated by using the same arithmetic operations as in Johannesma [13], while the influence of the other components is limited to order ε. Taking the limit $\varepsilon \downarrow 0, n \to \infty$ leads to

$$\Psi^e[\xi] = \frac{1}{2}(\xi + \gamma) \circ B^e \circ (\xi + \gamma) - \frac{1}{2} c \circ B^e \circ \gamma , \tag{56}$$

where

$$B^e = \left(B^{-1} - \Delta\right)^{-1} \tag{57}$$

is the covariance of the PESE. One can readily check the formula in case $D = 0$ [see Equation (24)], or with some difficulty if B and D share the same eigenfunctions. Note that a stationary stimulus $\mathbf{x}(\cdot)$ does not have a finite power; this can be circumvented by considering first a finite interval $[-T, T)$ and afterwards taking the limit T to infinity.

Equation (56) shows that $\Psi^e[\xi]$ is a second-degree functional in ξ and therefore the PESE is also a Gaussian process, with mean and covariance

$$m_1^e(t) = \int ds\, B^e(t,s) c(s-t) , \tag{58}$$

$$B^e(t,s) = \left[B(t,s)^{-1} - \Delta(t,s)\right]^{-1} . \tag{59}$$

The PESE, as should be expected, is nonstationary. To illustrate the theoretical results, simulations have been made. For the stimulus, zero-mean pseudorandom Gaussian white noise has been chosen. The linear filter was

$$c(s) = \begin{cases} 0.2 e^{-10s} & \text{if } s > 0 \\ 0 & \text{if } s \leq 0 \end{cases} , \tag{60}$$

the second-degree filter

$$D(t, s) = -c(s + 0.1)c(t + 0.1) \,, \tag{61}$$

and the event generator

$$g(u) = e^u \,. \tag{62}$$

The output intensity was approximately 9.7 events/second. The results of the simulation are compared with the theory in Figure 6.

The structure of the PESE can be split into two parts. From 100 ms before until the generation of an action potential, the average stimulus is completely described by the filter $c(s)$. Before that time the average stimulus is almost unchanged; however, the PESE shows a negative covariance. Since the stimulus is iid, the PESE is not different from the SE after the occurrence of the action potential.

5 Conclusions and discussion

In this paper we have shown that under the general assumption of orderliness the characteristic functional of the PESE can be calculated in principle. If a multiplicative neural model is chosen, a simple relation between the characteristic functionals of SE and PESE can always be established. Given the structure of the SE and the filter of the exponential system, a complete description of the PESE can be given. For example, in Example A a Gauss-Poisson process was transformed into another Gauss-Poisson process. In both examples both SE and PESE belong to the same class of processes; this property seems to be limited to multiplicative models. This class invariance is important, since it allows a comparison of the receptive field of the neuron (PESE) and the stimulus ensemble. Note that the theorem can be used, and in Example A has been used, for system identification, as can be understood from e.g. Equations (39) and (41, 42).

Appendix

In this section we will show that Equation (19) holds if the output process $\mathbf{M}(\cdot)$ is orderly for arbitrary input. We recall that we have to prove

$$\Phi^r[\xi] \quad := \quad E_{\mathbf{x}} \left[\exp \int ds\, \mathbf{x}(s) \xi(s) | d\mathbf{M}(t) = 1 \right] \tag{63}$$

$$= \quad \frac{E_{\mathbf{x}} \left[I[\mathbf{x}](t) \exp \int ds\, \mathbf{x}(s) \xi(s) \right]}{E_{\mathbf{x}} \left[I[\mathbf{x}](t) \right]} \,. \tag{64}$$
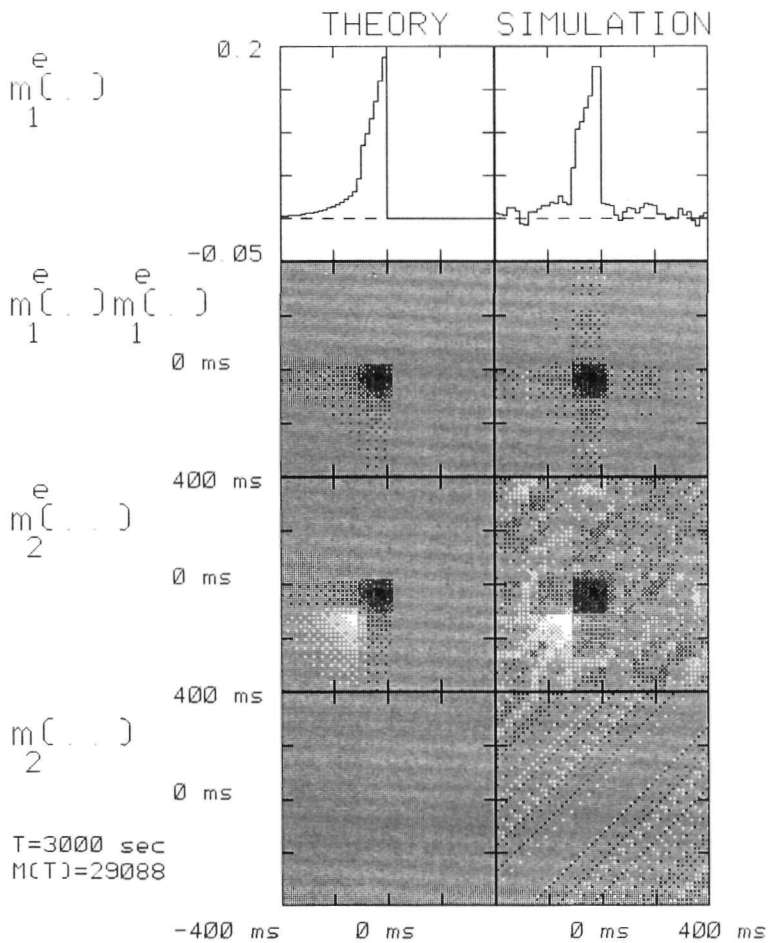
43

Figure 6: *The Transformation of a Gaussian process (SE) into another Gaussian process (PESE) by a multiplicative neuron. The results of a computer simulation are compared with calculations based on Equations (58) and (59).*

44

The validity of this equation can be shown starting from

$$E_\chi \left[\Delta \mathbf{M}(t) \exp \int ds\, \mathbf{x}(s)\xi(s) \right] . \tag{65}$$

Using conditional expectation [3], (65) equals

$$E_\chi \left[E[\Delta \mathbf{M}(t)|\mathbf{x}(\cdot)] \exp \int ds\, \mathbf{x}(s)\xi(s) \right] . \tag{66}$$

The orderliness of $\mathbf{M}$ implies (Daley [8])

$$E[\Delta \mathbf{M}(t)|\mathbf{x}(\cdot)] = P[\Delta \mathbf{M}(t) = 1|\mathbf{x}(\cdot)] + o(\Delta t) . \tag{67}$$

and substituting the model (17), (18), one gets

$$= E_\chi \left[\{ I[\mathbf{x}](t)\Delta t + o(\Delta t) \} \exp \int ds \mathbf{x}(s)q(s) \right] \tag{68}$$

$$= \Delta t E_\chi \left[I[\mathbf{x}](t) \exp \int ds\, \mathbf{x}(s)\xi(s) \right] + o(\Delta t) . \tag{69}$$

On the other hand, using conditional expectation again, (65) may be written as

$$= E_\chi \left[\exp \int ds \mathbf{x}(s)\xi(s) \big| \Delta \mathbf{M}(t) - 1 \right] P\left[\Delta \mathbf{M}(t) = 1 \right]$$
$$+ E_\chi \left[\Delta \mathbf{M}(t) \exp \int ds\, \mathbf{x}(s)\xi(s) \big| \Delta \mathbf{M}(t) > 1 \right] P\left[\Delta \mathbf{M}(t) > 1 \right] . \tag{70}$$

If the real part of $\xi(\cdot)$ is zero, then since $\mathbf{M}(\cdot)$ has already been assumed to be orderly, the second term of (70) can be approximated by

$$\left| E_\chi \left[\Delta \mathbf{M}(t) \exp \int ds\, \mathbf{x}(s)\xi(s) | \Delta \mathbf{M}(t) > 1 \right] P\left[\Delta \mathbf{M}(t) > 1 \right] \right|$$
$$< E\left[\Delta \mathbf{M}(t) | \Delta \mathbf{M}(t) > 1 \right] P\left[\Delta \mathbf{M}(t) > 1 \right] = o(\Delta t) . \tag{71}$$

Substitution of (71) in (70), followed by setting (69) and (70) equal, leads to

$$E_\chi \left[\exp \int ds\, \mathbf{x}(s)\xi(s) | \Delta \mathbf{M}(t) = 1 \right] E_\chi \left[I[\mathbf{x}](t) \right] \Delta t + o(\Delta t)$$
$$= E_\chi \left[I[\mathbf{x}](t) \exp \int ds\, \mathbf{x}(s)\xi(s) | \Delta \mathbf{M}(t) = 1 \right] \Delta t . \tag{72}$$

Taking the limit $\Delta t \downarrow 0$ completes the proof of the theorem of Section 3.

References

1. A. M. H. J. Aertsen, P. I. M. Johannesma, and D. J. Hermes, Spectro-temporal receptive fields of auditory neurons in the grassfrog. II. Analysis of the stimulus-event relation for tonal stimuli, *Biol. Cybernet.* 38:235-248 (1980).

2. A. M. H. J. Aertsen and P. I. M. Johannesma, The spectro-temporal receptive field. A functional characterization of auditory neurons, *Biol. Cybernet.* 42:133-143 (1981).

3. P. Billingsley, *Probability and Measure*, Wiley, New York, 1979.

4. E. de Boer and P. Kuyper, Triggered correlation, *IEEE Trans. Bio-Med. Eng.*, July 1968, pp. 169-179.

5. H. F. P. van den Boogaard and P. I. M. Johannesma, The master equation for neural interaction, *IMA J Math. Appl. Med. Biol.* 1(4):365-389 (1984).

6. H. F. P. van den Boogaard, G. H. F. M. Hesselmans, and P. I. M. Johannesma, System identification based on point processes and correlation densities. I. The nonrefractory model, *Math. Biosc.* 80:143-171 (1986).

7. D. R. Cox and V. Isham, *Point Processes*, Monographs on Applied Probability and Statistics, Chapman and Hall, London, 1980.

8. D. J. Daley, Various concepts of orderliness for point processes, in *Stochastic Geometry*, (E. F. Harding and D. G. Kendall, Eds.), Wiley, London, 1974, pp. 148-161.

9. J. J. Eggermont, W. J. M. Epping, and A. M. H. J. Aertsen, Stimulus dependent neural correlations in the auditory midbrain of the grassfrog *(Rana temporaria L.), Biol. Cybernet.* 47:103-117 (1983).

10. J. J. Eggermont, P. I. M. Johannesma, and A. M. H. J. Aertsen, Reverse correlation methods in auditory research, *Quart. Rev. Biophys.* 16(3):341-414 (1983).

11. W. J. M. Epping and J. J. Eggermont, Sensitivity of neurons in the auditory midbrain of the grassfrog to temporal characteristics of sound. I. Stimulation with acoustic clicks, *Hearing Res.* 24:37-54 (1986).

12. D. J. Hermes, J. J. Eggermont, A. M. H. J. Aertsen, and P. I. M. Johannesma Spectro-temporal characteristics of single units in the auditory midbrain of the lightly anesthetized grassfrog *(Rana temporaria L.)* investigated with noise stimuli, *Hearing Res.* 5:145-179 (1981).

13. P. I. M. Johannesma, Functional identification of auditory neurons based on stimulus-event correlation, in *Psychophysical, Physiological and Behavioural Studies in Hearing* (G. Brink and F. A. Bilsen, Eds.), Delft U.P., Delft, 1980, pp. 77-84.

14. P. I. M. Johannesma and H. F. P. van den Boogaard, Stochastic formulation of neural interaction, *Acta Appl. Math.* 4:201-224 (1985).

15. N. G. van Kampen, *Stochastic Processes in Physics and Chemistry*, North-Holland, Amsterdam, 1981.

16. P. I. Kuznetsov and R. L. Stratonovich, A note on the mathematical theory of correlated random points, in *Nonlinear Transformations of Stochastic Processes*, pp. 101-115. (P. I. Kuznetsov, R. L. Stratonovich, and V. I. Tikhonov, Eds.), Pergamon, Oxford, 1965.

17. E. Lukacs, *Characteristic Functions*, Griffin, London, 1960.

18. R. K. Milne and M. Westcott, Further results for Gauss-Poisson processes, *Adv. Appl. Prob.*, 4:151-176 (1972).

19. D. L. Snyder, *Random Point Processes*, Wiley, New York, 1975.

20. S. K. Srinivasan, *Stochastic Point Processes and Their Applications*, Griffin, London, 1974.

21. R. L. Stratonovich, *Topics in the Theory of Random Noise*, Vol. 1, Gordon and Breach, New York, 1963.

Spectro-Temporal Interpretation of Activity Patterns of Auditory Neurons

Gerard H. F. M. Hesselmans   Peter I. M. Johannesma

Abstract

Sensory receptors transform an external sensory stimulus into an internal neural activity pattern. This mapping is studied through its inverse. In a preceding paper we showed that within the context of a neuron model composed of a linear filter followed by an exponential pulse generator and a Gaussian stimulus ensemble a unique 'most plausible' first-order stimulus estimate can be constructed. This method, applicable only to neurons showing phase-lock, is extended to neurons without phase-lock. In this situation second-order spectro-temporal stimulus estimates are produced; examples are given from simulation. The method is applied to activity of neurons in the auditory system of the frog.

1 Introduction

Animals obtain information about the environment through sensory receptors. Stimuli from the external world are transformed into a neural signal leading to an internal representation of the environment. Given the mapping of the stimulus into an activity pattern of sensory neurons the question arises, whether the inverse map, from a neural activity pattern into a sensory stimulus, exists and can be determined.

In Gielen et al. 1987 these questions are investigated within the context of a neural model. In this model the neurons, conditional the sensory input, are assumed to behave independently. For primary neurons (nervus acousticus) this appears to be correct. It has been shown experimentally that there are no interactions between primary neurons; correlations in activity of these neurons are caused by common stimulus properties and overlapping receptive fields. Therefore the transformation of a stimulus into a neural activity pattern is completely determined by the neural response functions of the individual neurons. Gielen et al. considered the auditory neuron as a Volterra system followed by an exponential event generator.

In search of a most plausible first-order stimulus estimate it was concluded that one should distinguish between linear and nonlinear Volterra systems. If
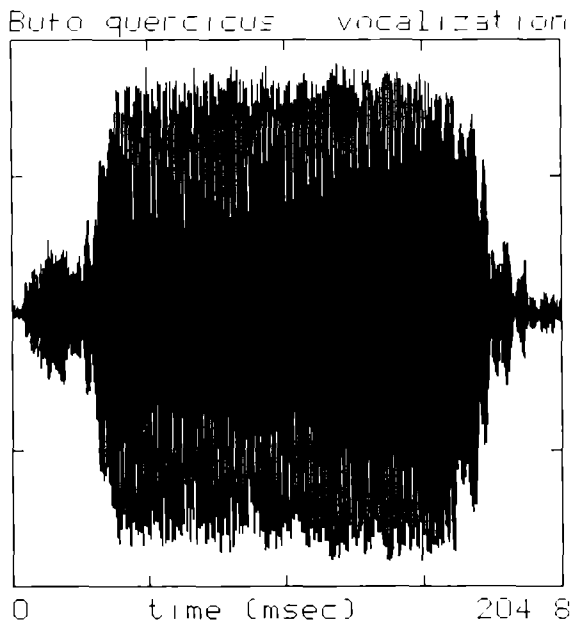
Figure 1: *Vocalization.* *Call of the Bufo quercicus*

49

the system is linear and the stimulus is Gaussian, then a unique most plausible first-order stimulus estimate exists, i.e. the mapping is invertible. If the Volterra system is second-order nonlinear, then the solution is a hyperplane in stimulus space. These different types of behavior are related to the phenomenon of phase-lock (Epping et al. 1986; Gielen et al. 1987).

Measurement of the correlation functions indicates that low-pass neurons can be approximated by a linear first-order Volterra system (Grashuis 1974). If the stimulus is Gaussian and all neurons are low-pass, then for every neural activity pattern, there exists at most one 'most plausible' stimulus. This most plausible stimulus can be approximated by replacing every action potential by its first-order correlation function. The effectiveness of this inverse map has been verified by computer simulations.

Above a critical frequency, auditory neurons show loss of phase-lock. As a consequence the second-order stimulus-event correlation is needed for the characterization of the neural response function. The second-order kernel can be considered as the spectro-temporal sensitivity of the neuron (Aertsen and Johannesma 1981). Due to loss of phase-lock the neural activity is primarily a result of the spectro-temporal content of the stimulus signal. The inverse map in this situation can no longer be based upon first-order properties. In this paper an inverse map based on the spectro-temporal sensitivity is constructed which results in an estimation of the spectro-temporal characteristics of the stimulus presented to the animal. It is an extension of the first-order method which is adapted to higher-order stimulus properties. Results of simulations and experimental data from the auditory midbrain of the grassfrog will be shown.

2   Description of sensory stimuli

The sensory environment of an animal is composed of several modalities. In this paper the auditory system will be investigated. In the following $x(\cdot)$ denotes a stimulus signal, it represents the pressure fluctuations of the air as a function of time for the auditory system. In Figure 1 the call of a toad is given.

In a natural environment many different acoustic sources are present. The composition of all the sounds in a given ecosystem is called the acoustic environment (Johannesma et al. 1986). In order to compare signals within the acoustic environment a representation like the one shown in Figure 1 seems to be far from optimal since it does not show the structure within the signal. Structure within a signal becomes visible in higher-order representations, e.g. product functions. The n-fold product functions $p_{ni}(\cdot)$ of a stimulus signal $x_i(\cdot)$ are defined by:

$$p_{ni}(t_1, \ldots, t_n) = \prod_{j=1}^{n} x_i(t_j).$$ (1)

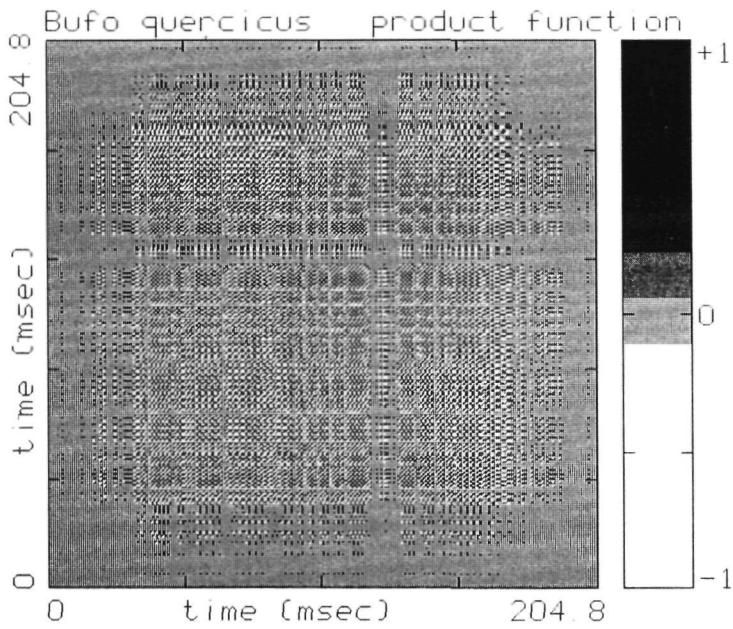The first product function of a signal is the signal itself (see Figure 1). The

Figure 2: *Second-order product function of the call of the Bufo quercicus, as shown in Figure 1. Amplitude is coded in grey.*

51
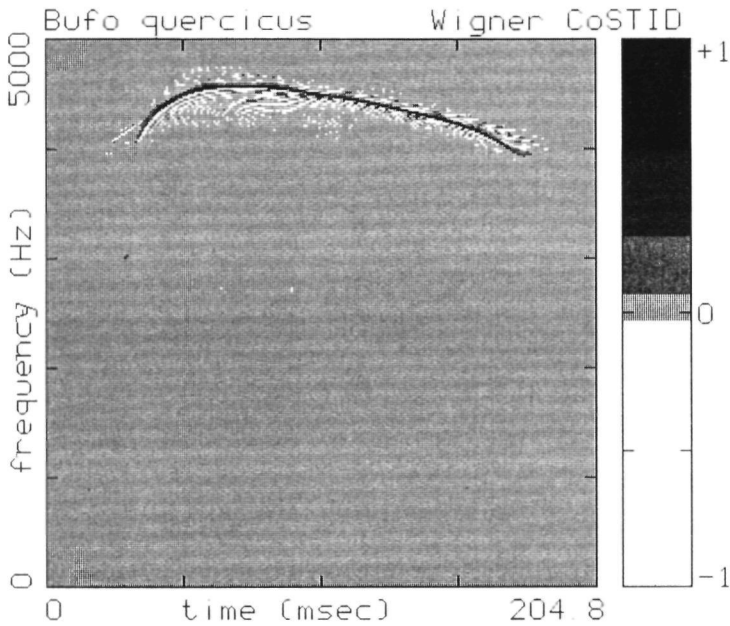
Figure 3: *Wigner CoSTID of the call of the Bufo quercicus. Amplitude is coded in grey.*

higher-order product functions show the structure within the signal. Figure 2 shows the second product function of (again) a call of a toad.

Although Figure 2 shows a large amount of structure in this rather complicated signal, it is difficult to interpret. An associated representation which is easier to display and perceptually more acceptable for the human investigator is the Wigner CoSTID (Johannesma et al. 1981; Hermes 1985; Yen 1987). It is a linear transformation of the second product function which leads to a real-valued coherent spectro-temporal representation of the signal (Aertsen et al. 1984). The calculation of the (Wigner) CoSTID starts from the analytic signal ξ.

$$\xi(t) = x(t) + i\tilde{x}(t) \tag{2}$$

with $\tilde{x}(t)$ the Hilbert transform of $x(t)$. The (lagged) product function (Johannesma et al. 1984; Aertsen et al. 198 bitemporal representation (Hermes 1985) of the analytic signal is defined by:

$$\Pi(t,\tau) = \xi^*\left(t - \frac{\tau}{t}\right)\xi\left(t + \frac{\tau}{t}\right) \tag{3}$$

52

where '*' denotes complex conjugation. Note that also the product function $p_2(\cdot,\cdot)$ of the real-valued stimulus signal $x(\cdot)$ and the product function $\Pi(\cdot,\cdot)$ of the complex-valued analytic signal $\xi(\cdot)$ are linearly related. By Fourier transformation of $\Pi(\cdot,\cdot)$ with respect to the time difference τ the Wigner CoSTID $\Xi(\omega,t)$ is obtained.

$$\Xi(\omega,t) = \int d\tau \, e^{-i\omega\tau} \Pi(t,\tau) \tag{4}$$

The Wigner CoSTID gives both frequency content and phase of the signal as a function of time. The power of the signal is given by the amplitude. Considering the signal as a superposition of tone-pips; a Gaussian amplitude modulated sinusoid (Hermes 1985), with either different carrier frequencies and/or different onset-times then the fluctuations or 'ghost-images' can be explained by their interference products. The frequency of the spectro-temporal fluctuations is proportional to the difference in carrier frequencies and onset-times of the tone-pips. Since different onset-times and phase differences are related, the phase differences are coded by the fluctuations. Note that only relative phase relations are known. The Wigner CoSTID determines the signal apart from one overall phase factor. Projections of the CoSTID upon time-axis and frequency-axis result in respectively the temporal-intensity and spectral-intensity. More about the interpretation of the Wigner CoSTID can be found in Hermes 1985 and Yen 1987. An example of a Wigner CoSTID is given in Figure 3.

The call of the toad consist mainly of frequency components around 4.4 kHz, with phase relations shifting in time. Compare Figures 2 and 3. Product functions of order higher than 2, share the interpretation problems known from the second order product function. Unfortunately no satisfactory transformations like the Wigner CoSTID are available.

Product functions can not only be used to distinguish between separate stimuli, but also to give a statistical description of the stimulus ensemble (SE) as a whole. The moments of the SE are found by averaging the product functions of the N individual signal elements of the SE:

$$m_n(t_1,\ldots,t_n) = -\frac{1}{N} \sum_{x_i \in SE} p_{ni}(t_1,\ldots,t_n). \tag{5}$$

In general, many moments are necessary for a complete statistical description. However, if the ensemble is Gaussian, the first two moments are sufficient (van Kampen 1981). The stimuli experienced by an animal under natural conditions are not Gaussian distributed, however the acoustical environment might be approximated by a Gaussian ensemble and during an electro-neurophysiological experiment the experimenter may choose such a SE.
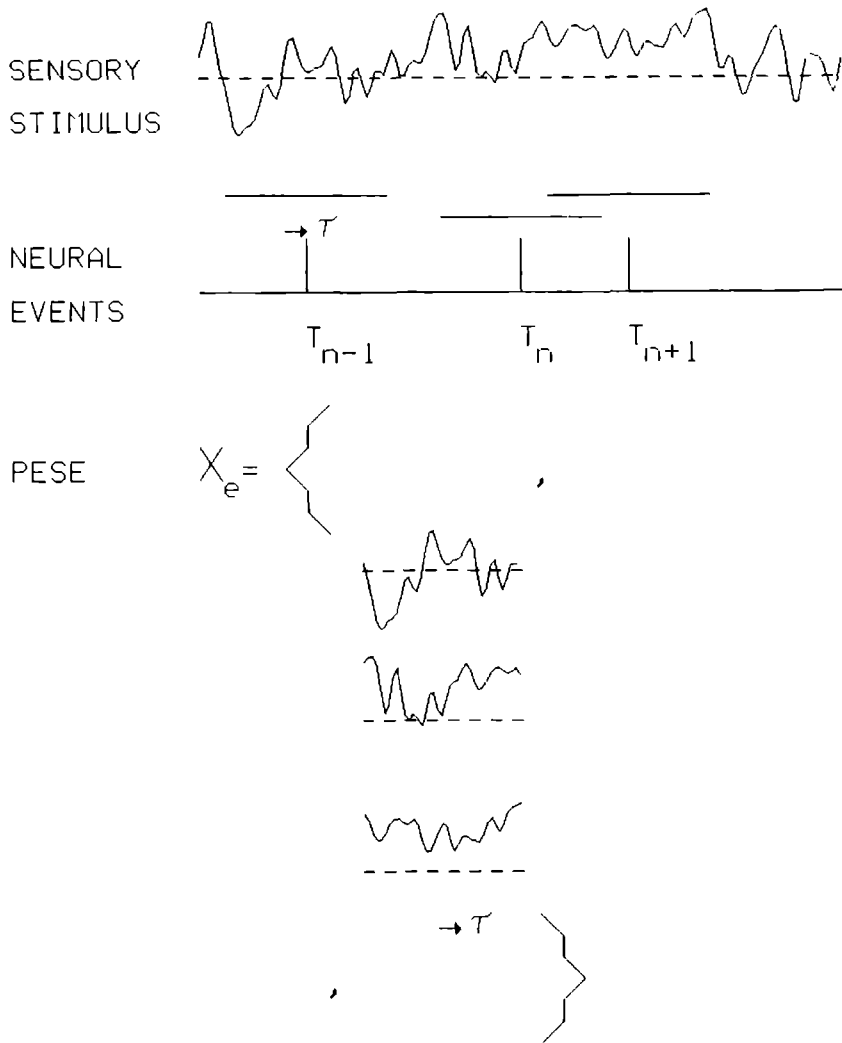
Figure 4: *Construction of the peri-event stimulus ensemble from the stimulus and recorded neural events.*

3 Description of response-related stimuli

The auditory system transforms the stimulus (e.g. a mating call) into a series of action potentials. To analyse this transformation reverse correlation methods (de Boer and Kuyper 1968; Eggermont et al. 1983) will be used. This analysis method compares the ensemble of all stimuli presented to the animal during the experiment with those associated with the occurrence of an action potential (neural event). How to construct such a peri-event stimulus ensemble (PESE) which seems to contain stimuli of special interest for the neuron, is shown in Figure 4.

To characterize the stimuli in the PESE again a statistical description by means of moments will be used. The moments of the PESE are found by averaging the product functions of the N_r selected individual signal elements:

$$\mu_n(t_1,\ldots,t_n) = -\frac{1}{N_e} \sum_{x_i \in PESE} p_{ni}(t_1,\ldots,t_n). \tag{6}$$

The moment calculated using only the stimuli associated with the generation of an action potential by a particular neuron. The moments of the PESE are found by averaging over a subset of the SE. In Equation (6) properties of the PESE are obtained by means of an ensemble average. For practical reasons this is often replaced by a time average which is allowed in case of a stationary stimulus and a time-invariant system.

For auditory neurons, the stimuli in the PESE are in someway different from those in the SE. Both ensembles can be compared by looking at their respective moments. If SE and PESE are Gaussian, then differences, if any, should be visible in their first two moments, or equivalently their average signal and Wigner CoSTID.

4 Estimation of stimulus characteristics from neural activity

In the previous sections attention has been focused upon stimulus description. However inside the brain not the signals but sequences of action potentials are present which code the relevant stimuli. In this section a method will be proposed for interpreting the neural activity patterns by means of stimulus characteristics.

An optimal stimulus reconstruction scheme to the animal is likely to enhance relevant stimulus aspects. Since beforehand the set of relevant stimuli is not known, the best one can do, is try to reconstruct the actual stimulus or its characteristics from the neural activity pattern. Due to adaptation by the animal, it is to be expected, that the set of relevant stimuli to a large extent covers the set of stimuli, that can be represented by the nervous system.
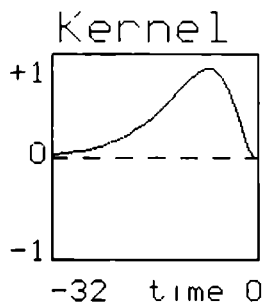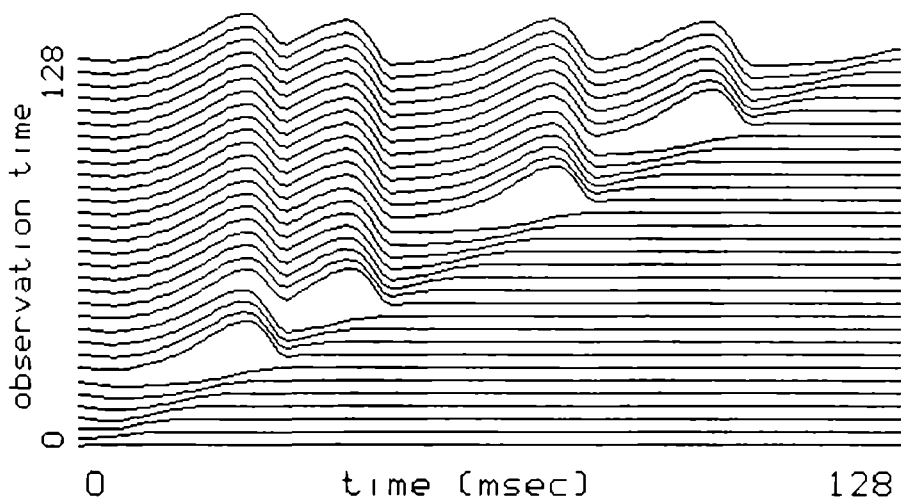
Figure 5: *Stylized example of the stimulus estimation procedure Top : First-order estimation kernel. Middle: Neural activity pattern. Bottom: Dynamic estimation of the stimulus. The first row shows the a priori first-order estimate. The subsequent rows show the change in this estimate as more information about the neural activity pattern is added.*

For an optimal estimation by an external observer of the sensory stimulus associated with a neural activity pattern, knowledge both of the structure of the nervous system and of the characteristics of the SE are necessary (Johannesma et al. 1986). Although knowledge of the SE is readily accessible for the observer, at least during a neurophysiological experiment, this is not the case for the nervous system. Without irreversibly damaging the brain it is impossible to obtain detailed information about neural connections. Therefore, it is assumed that

1. sufficient knowledge about the SE is available, and

2. correlation of the neurons within the observational set is determined by their stimulus dependency.

As a consequence each sequence of action potentials is modelled by a Poisson point process with a stimulus driven intensity (Cox and Isham 1980; Snyder 1975) and structure within the neural activity process is solely due to stimulus correlations. This approximation is quite realistic in peripheral parts of the nervous system if firing intensities are low.

Since neural events are assumed to be generated independently they are given equal weight in the stimulus estimation procedure. Each time an action potential occurs the estimate is changed by an amount suitable for the specific neuron. In order to balance this cumulative procedure the estimate should drift in the opposite direction during the absence of events. For a stimulus characterization by means of moments we propose:

$$\frac{d}{dt}\overline{m}_n(t_1,\ldots,t_n \mid \mathcal{H}(t)) = \sum_j \left[\frac{d}{dt}N_j(t) - \lambda_j\right] \Delta\overline{m}_n^j(t_1 - t,\ldots,t_n - t) \qquad (7)$$

where

$m_n(\cdots \mid \mathcal{H}(t))$ = estimate of the nth moment of the stimulus based
upon the observed history of neural activity $\mathcal{H}(t)$,
$\mathcal{H}(t)$ = symbolic denotation of the neural activity in the observation
interval,
$N_j(\cdot)$ = counting point process which models the neural activity of
neuron j
$dN_j(t)$ = the number of events in $[t, t + dt)$,
λ_j = average intensity of neuron j,
$\Delta\overline{m}_n^j(\cdots)$ = change in the estimate of the nth moment of the stimulus
caused by the occurrence of an event of neuron j.

Each neuron contributes in the same way to the updating. The change in the estimate is determined by the difference between actual and expected neural activity. The change is abrupt if an event occurs and gradual in the absence of events. On the average no change occurs, because $E[dN_j(t) - \lambda_j dt] = 0$. A
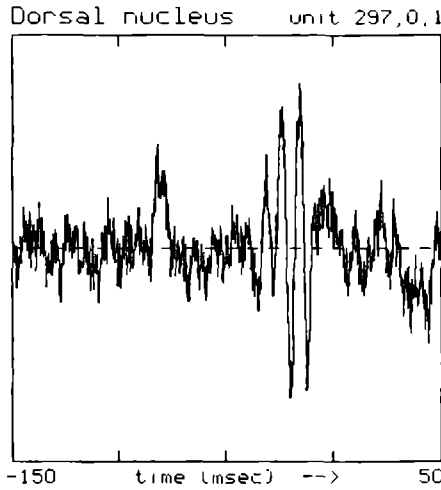
Figure 6: *First-order reverse correlation function of noise stimulus and neural activity in the dorsal nucleus of the grassfrog (Rana temporaria L.).*

logical starting-point of the estimate updating procedure is the *a priori* stimulus distribution. This results in the following integral formulation.

$$\overline{m}_n(t_1,\ldots,t_n|\mathcal{H}(t)) = m_n(t_1,\ldots,t_n) +$$
$$\sum_j \left[\int_A dN_j(t)\Delta\overline{m}_n^j(t_1-t,\ldots,t_n-t) - \int_A dt\lambda_j\Delta\overline{m}_n^j(t_1-t,\ldots,t_n-t)\right] \quad (8)$$

where

$m_n(\cdots)$ = *a priori* estimate of the nth moment of the stimulus ensemble,
A = observation interval.

The stimulus estimate consists of 1) an *a priori* estimate, 2) contributions of the individual neurons induced by the occurrence of an event and 3) a global opposite effect. During the recording of the neural activity the observation interval A increases in time. Figure 5 shows a dynamic application of formula (8) for an imaginary electro-neurophysiological recording, to obtain an estimate of the first-order moment.

Before the first event occurs, the stimulus estimate becomes more and more negative, indicating the absence of an excitatory stimulus or the presence of a suppressive stimulus. The estimation 'kernel' $\Delta\overline{m}_n^j(\cdots)$ relates the *a priori* properties of the stimulus with those leading to an event. Combining this notion

with Section 3 leads to:

$$\Delta\overline{m}_n^j(t_1,\dots,t_n) = \mu_n^j(t_1,\dots,t_n) - m_n(t_1,\dots,t_n). \tag{9}$$

The estimate of the stimulus is based upon knowledge of the SE and the PESE. An animal, *if* it should make use of a similar estimation procedure, may obtain this information while being in constant interaction with its environment and by inheritance. The investigator however can obtain *a priori* knowledge by invoking a learning phase.

If the stimulus has zero mean, then $m_1(\cdot) = 0$, and the first-order kernel in Equation (9) reduces to the reverse correlation function $\mu_1^j(\tau)$ (de Boer and Kuiper 1968). In case of primary neurons (nervus acousticus) or neurons from the dorsal nucleus this reverse correlation function often shows a band-pass behavior (see Figure 6). Therefore its average is approximately zero:

$$\int_A dt\,\Delta\overline{m}_1^j(t_1 - t) \simeq 0, \tag{10}$$

and the cumulative effect can be neglected. As a consequence Equation (8) reduces to:

$$\overline{m}_1^j(t_1|\mathcal{H}(t)) \simeq \sum_j \int_A dN_j(t)\,\Delta\overline{m}_1(t_1 - t). \tag{11}$$

Using a somewhat different approach the same result was obtained by Gielen et al. 1987.

Such a 'simplification' is not possible in the second-order case, since second-order functions in general are positive definite. The second-order formulation of Equation (8) reads:

$$\overline{m}_2(t_1,t_2|\mathcal{H}(t)) = m_2(t_1,t_2) + \tag{12}$$

$$\sum_j \left[\int_A dN_j(t)\,\Delta\overline{m}_2^j(t_1 - t, t_2 - t) - \int_A dt\,\lambda_j \Delta\overline{m}_2^j(t_1 - t, t_2 - t) \right].$$

Using the linear relations between the real-valued product function, the complex-valued product function of the analytic signal, and the Wigner CoSTID a transition from Equation (12) to a Wigner CoSTID format is possible.

$$\overline{\Xi}(\omega,t|\mathcal{H}(t)) = \Xi(\omega,t) + \sum_j \left[\int_A dN_j(s)\,\Delta\overline{\Xi}^j(\omega,t - s) - \int_A ds\,\lambda_j \Delta\overline{\Xi}^j(\omega,t - s) \right]$$

$$\tag{13}$$

A dynamic application of Equation (13) is given in Figure 7. The method is used to estimate the spectro-temporal stimulus properties.
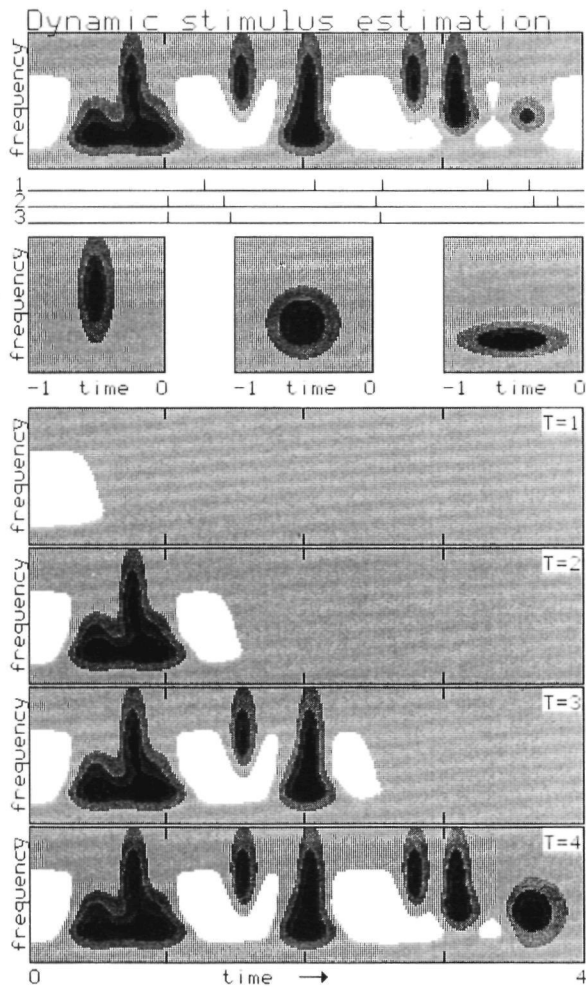
Figure 7: *Stylized example of the spectro-temporal stimulus estimation procedure From top to bottom: A) Wigner CoSTID of the stimulus. B) Neural activity pattern. C) estimation kernels. D) Dynamic estimation of the stimulus. Result of the estimation procedure based upon the observation of neural activity during 1 up to 4 time units.*

60
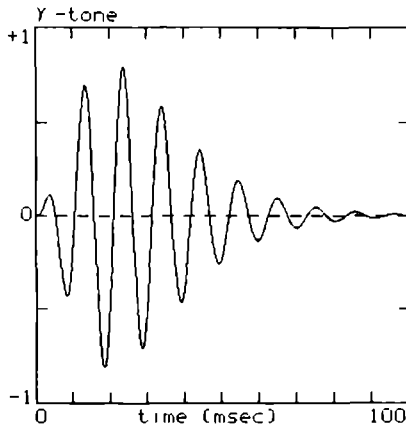
Figure 8: γ-Tone. $\gamma = 3$, $\tau = 10$ msec, $A = 1.5$ and $f = 96$ Hz.

5 Stimulus reconstruction: model simulation

The estimation procedure sketched in Section 4 has been applied in a computer simulation. The stimulus consisted of 8 γ-tones (Aertsen and Johannesma 1980). A γ-tone is given by:

$$m(t) = A \left(\frac{t}{\tau}\right)^{\gamma-1} \exp\left(-\frac{t}{\tau}\right) \sin(2\pi f t) . \tag{14}$$

All tone-pips had the same amplitude ($A = 1$), time constant ($\tau = 5$ msec) and form ($\gamma = 3$). The carrier frequencies were distributed between 96 and 544 Hz on a linear scale.

The model has been taken from van den Boogaard et al. 1985 and consists of 16 independent neurons. Each neuron is composed of a second-order band-pass filter followed by an exponential event generator (see Figure 9).

The transformation of the stimulus into a generator potential by the filter is given by:

$$u(t) = \int_{-\infty}^{t} ds\, x(s) w(t - s) \tag{15}$$

with

$x(\cdot) = $ stimulus,
$w(\cdot) = $ impulse response of the filter,
$u(\cdot) = $ generator potential.

The generator potential controls the probability for an event to be generated.

$$P[\Delta \mathbf{N}(t) = 1 | \mathbf{u}(t) = u] = \Delta t\, \nu\, e^u + o(\Delta) \tag{16}$$

$$P[\Delta \mathbf{N}(t) > 1 | \mathbf{u}(t) = u] = o(\Delta t), \quad \Delta t \text{ small enough} \tag{17}$$
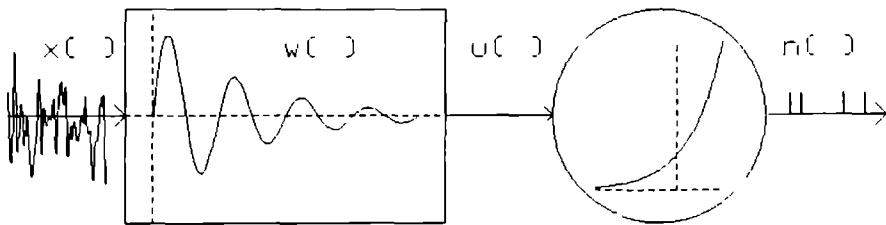
Figure 9: *The model. The transformation of the stimulus $x(\cdot)$ into a sequence of events. The event generating system consists of a linear filter followed by an instantaneous, exponential nonlinearity.*

with

$\Delta N(t)$ number of events in the interval $[t, t + \Delta t)$,
ν spontaneous firing rate in the 'absence' of a stimulus.

Given the stimulus the neural activity is described by an inhomogeneous Poisson point process (Cox and Isham 1980; Snyder 1975). The model parameters were choosen as follows:

    filters
        time constant τ : 5 msec,
        central frequencies   cf : 16 neurons distributed between 200 Hz
                                        and 500 Hz on an equidistant scale,
        amplitude A : 1800,

    event generator
        spontaneous activity ν : 200/sec.

    A similar set of parameters, although larger, has been used previously by Gielen et al. 1987 to estimate first-order stimulus properties. Here second-order properties are of interest.

    Figure 10 gives an overview of the results. The top row (part A) shows the sequence of γ-tones. Carrier frequencies increase from left to right. The next row (part B) shows this frequency increase more clearly in the Wigner CoSTID. Between the CoSTID's of the individual γ-tones some interference products are visible. However the calculation of the Wigner CoSTID of the series of 8 γ-tones should be done coherently and more interference products should be present, e.g. the interference product of the first and third γ-tone is located on top of the Wigner CoSTID of the second γ-tone. Since such a long range coherence is undesirable at this stage, the temporal integration mentioned in Equation (4) is
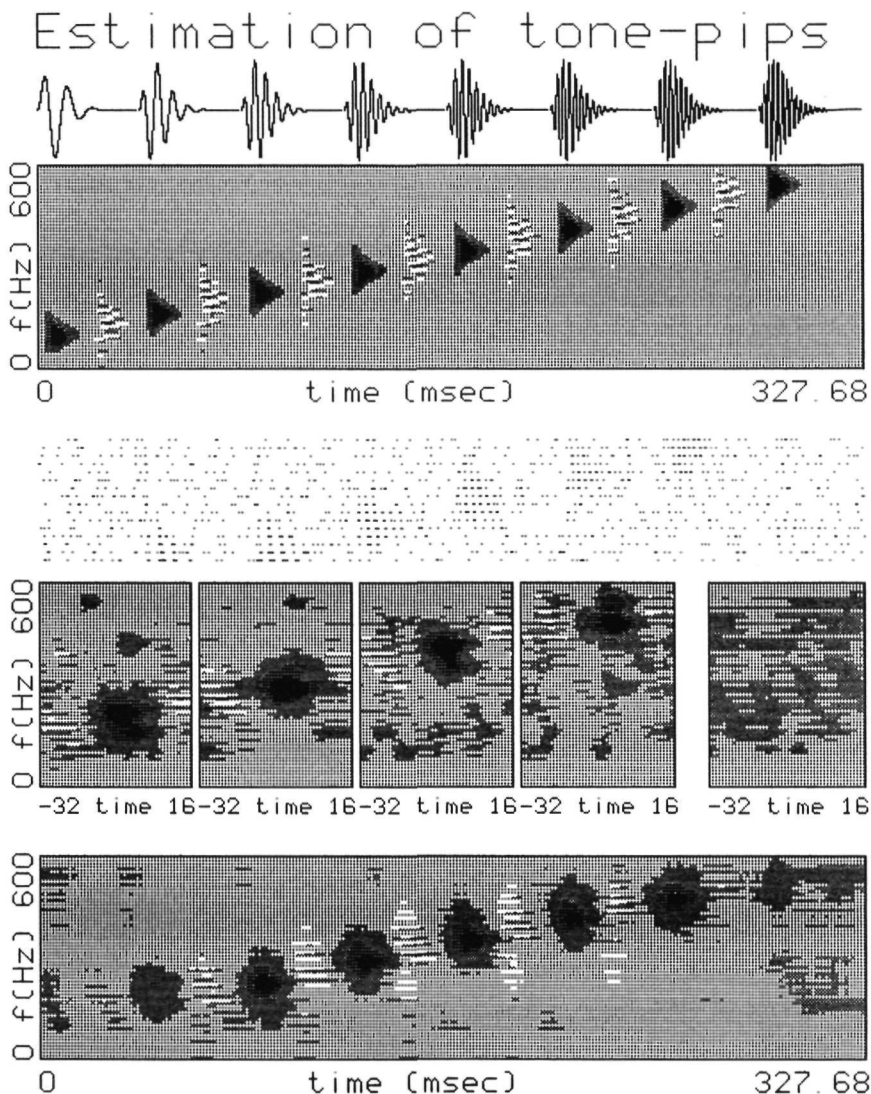
Figure 10: *Stimulus reconstruction based on computed simulated data. A) Sequence of γ-tones. B) Wigner CoSTID of the γ-tone sequence. C) Neural activity pattern. Events of neurons with a different characteristic frequency (cf) are plotted at different lines. Activity patterns are sorted by the cf. of the neurons; high-frequency neurons at the top and low-frequency neurons at the bottom. D) Wigner CoSTID of 4 out of 16 stimulus driven model neurons and one independently generated Poisson process. E) Stimulus estimate after 1 stimulus presentation.*

performed using a sliding window of width equal to the length of just one γ-tone. This procedure mimics short term coherence. Part C gives the neural events as a sequence of dots on a line. Activity of neurons with a different characteristic frequency (cf) is plotted at different lines. Neurons increase their activity if γ-tones with a similar carrier frequency are presented. Using the PESE of each neuron, all Wigner CoSTID's were calculated. Results of some representative neurons are shown in the last row but one (part D). Since the neurons only differ in their characteristic frequency, the only difference between the respective CoSTID's is their position along the vertical frequency-axis. Other differences are due to statistical fluctuations. The last column in this row shows the result based on a (stimulus independent) homogeneous Poisson process. This last CoSTID is an estimate of the spectro-temporal stimulus properties to expect *a priori*. The fifth and last row (part E) gives the stimulus estimate using the previously calculated Wigner CoSTID's and the *same* neural activity patterns, i.e. learning and 'estimation phase' use the same neural data. The first (low-frequency) and last (high-frequency) γ-tones are not or only weakly recovered, because they are outside the range of spectral-sensitivity of the model neurons; γ-tones within this range are well estimated. Note that the interference products are also present in the estimated Wigner CoSTID. This is due to the coherent presentation of the γ-tones in the 'learning-phase', and the presence of the interference products in the estimation kernels.

6 Stimulus reconstruction: applied to neural data

The reconstruction method has also been applied to data from the auditory midbrain (torus semicircularis) of the grassfrog (*Rana temporaria L.*). The response of 6 neurons in 6 different frogs was recorded during auditory stimulation with Gaussian noise and mating-call vocalizations. Since the recordings have been obtained from 6 different frogs, independence of the neural activity is guaranteed. For details about the experimental set-up see Epping et al. 1985.

During the first part of the experiment Gaussian noise (90 dB SPL) was presented. The estimation kernels were calculated. The resulting Wigner CoSTID's are shown in Figure 11. The neurons show a broad range of spectro-temporal sensitivity.

During the second part a mating call was presented each 2.8 seconds. In order to maintain the same adaptation level, the mating call was presented together with the Gaussian noise used in part one (see Figure 12). This mating call or B-call can be approximated by a series of γ-tones with varying amplitude.

The recorded neural events together with the Wigner CoSTID's were used to estimate the spectro-temporal stimulus characteristics. The results are shown in Figure 13.
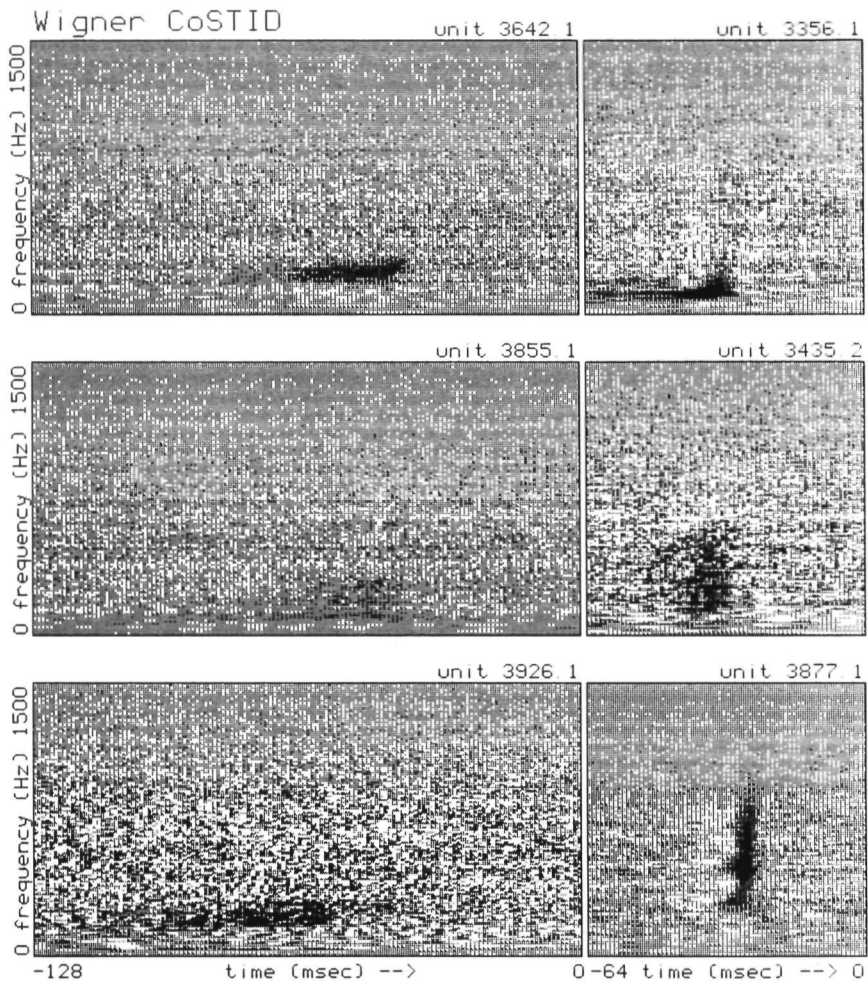
Figure 11: *Wigner CoSTID of 6 neurons from the midbrain of the grassfrog based on a Gaussian stimulus ensemble.*
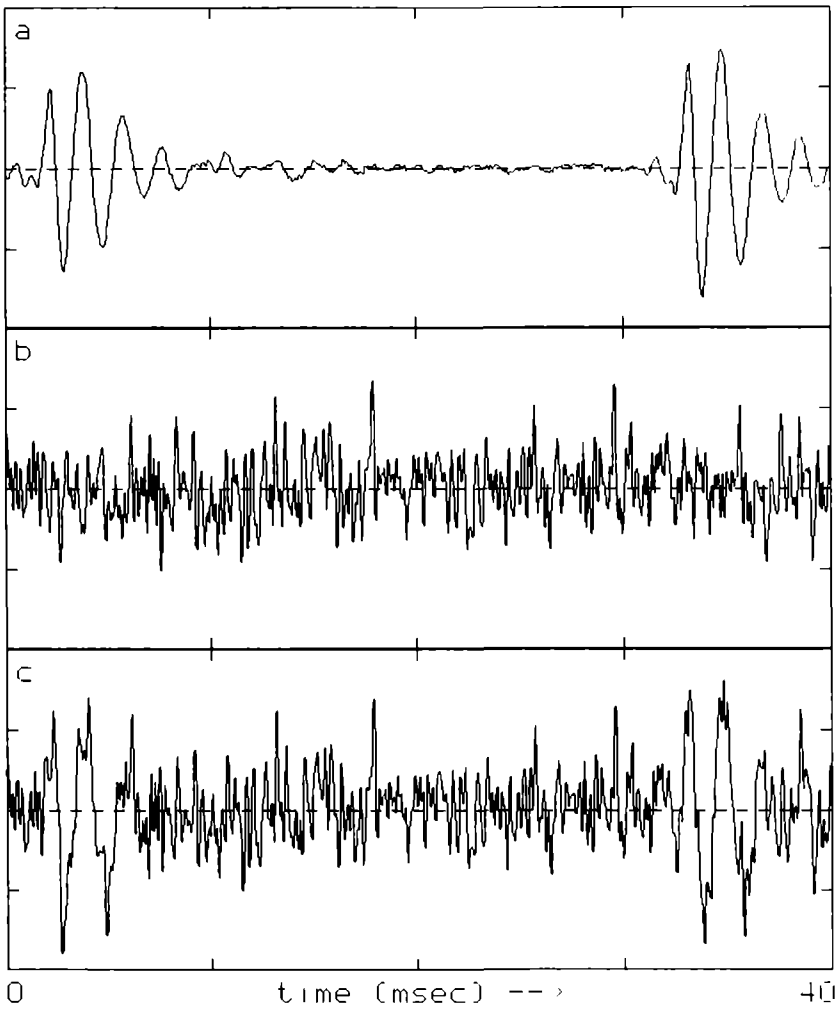
Figure 12: *Auditory stimulus A) Part of a mating call of the grassfrog. B) Part of the Gaussian noise. C) Superposition of mating call and Gaussian noise.*
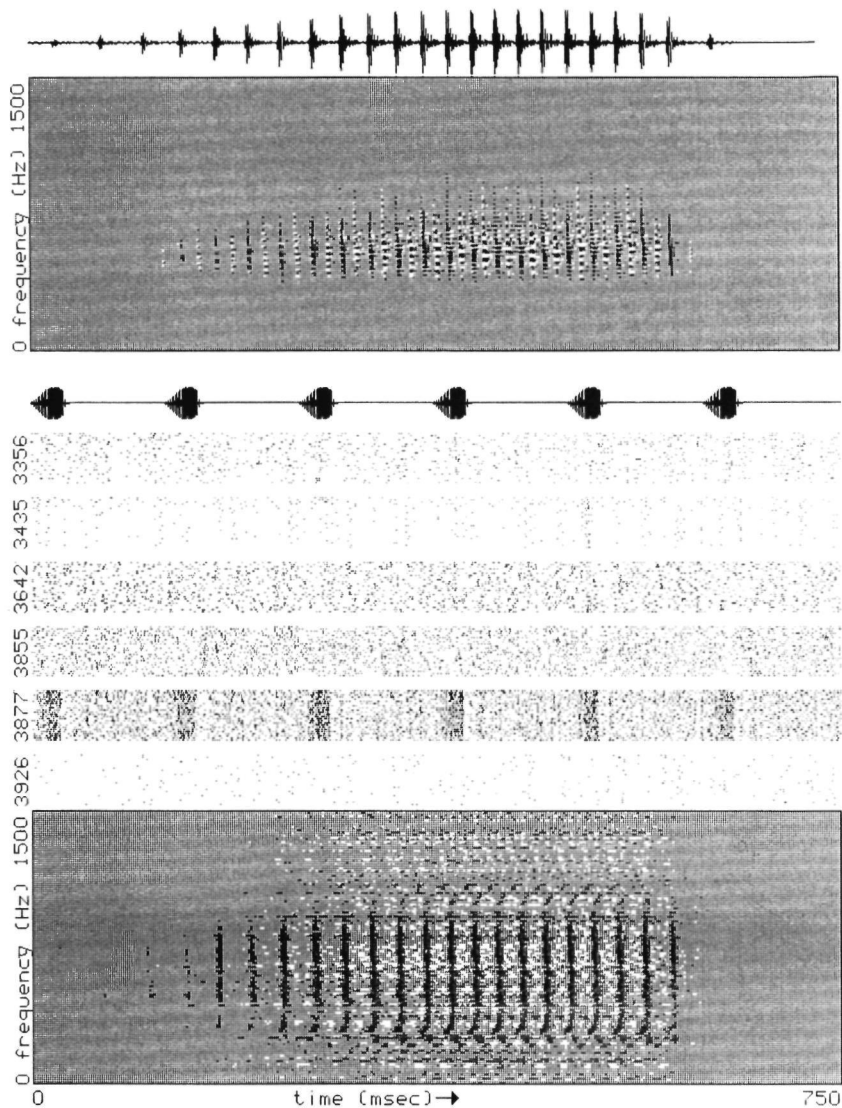
Figure 13: *Stimulus reconstruction based on experimental data. A) Mating call of the grassfrog B) Wigner CoSTID of the mating call C) Neural activity pattern. After the unit number of the neuron the neural activity is plotted at different lines. Each line shows the response to 6 presentations of mating call and noise. The time span of this part is 16864.4 msec. D) Average estimated stimulus Wigner CoSTID after 196 stimulus presentations.*

67

Part A gives the mating call and part B of this figure shows its Wigner CoSTID. Note the interference products. Part C gives the neural response. An estimate based upon just one stimulus presentation is rather poor, since it is based upon the activity patterns of just 6 neurons and it is also corrupted by the influence of the simultaneously presented noise sequence.

Since the Gaussian noise was presented incoherent with the mating calls, its influence can be suppressed by averaging the results of the 196 presentations of the frog-vocalization. The result is shown in part D of Figure 13. Since the Gaussian noise used in the learning-phase is an incoherent signal compared with the mating call, interference products are neither present in the kernels nor in the estimated Wigner CoSTID of the stimulus.

7   Conclusion and discussion

In this paper a procedure is proposed for the estimation of the sensory stimulus presented to an animal. It is based upon the observed activity of a set of neurons, *a priori* statistical knowledge of the stimulus, and statistical knowledge of the response of the neuron to the stimulus. The procedure is taken form the viewpoint of the experimenter who has statistical knowledge of both SE and PESE.

The stimulus is assumed to be stationary; it may be Gaussian noise, a set of randomly distributed artificial tones or an ensemble of natural sounds derived from the acoustic biotope of the animal. Although the neurons may be connected, they are assumed to have no functional interaction; the correlation of the neurons within the observational set is determined only by their stimulus dependency. Due to this stimulus dependency or stimulus selection property of the neurons, the PESE is nonstationary. In the first part of the procedure statistical knowledge has to be acquired about the PESE: distribution, moments or cumulants. In line with experimental practice the first two moments or cumulants are determined. The first cumulant forms the characteristic stimulus of the neuron, the second one is equivalent to its spectro-temporal receptive field. If SE and PESE are Gaussian distributed, then their first two moments or cumulants are sufficient for a complete statistical characterization of both ensembles. For a system theoretical analysis a Gaussian ensemble is attractive, for a neuro-ethological approach a natural ensemble of stimuli based on the acoustic biotope opens the possibility to investigate the base of natural forms of auditory behavior. In the latter case an analysis based upon just the first two cumulants is bound to be incomplete.

Since the estimation procedure is based upon statistical characteristics of the SE, its applicability is limited to a single stimulus ensemble. A transfer from one SE to another is in principle only possible through a full characterization of the relevant part of the nervous system (e.g. by its Volterra kernels). The updating process is essentially linear. No scaling of the estimate is necessary, if one remains within the same stimulus ensemble. Using a least square error

criterion and a linear estimation procedure, one can easily show that the proposed estimate is optimal, if correlation is absent. Poisson processes have this property. Experimental data show that if intensities are low, neural activity often can be described by a Poisson process.

Two final remarks are related to the appreciation of the method. Firstly no neural interaction has been taken into account. As a consequence the function of the nervous system considered here is filtering and representation of the stimulus. Active aspects of perception such as extrapolation were not regarded. Secondly this method of sensory interpretation of neural activity is not intended to represent an internal image present in the animal but to give a systematic description for an external observer of the observational bounds of the animal and of the limits to perception based on the patterns of neural activity. Its predictions however can be tested in behavioral experiments.

References

1. A. M. II. J. Aertsen and P. I. M. Johannesma, Spectro-temporal receptive fields of auditory neurons in the grassfrog. I. Characterization of tonal and natural stimuli, *Biol. Cybernet.* 38:223-234 (1980).

2. A. M. H. J. Aertsen and P. I. M. Johannesma, The spectro-temporal receptive field; a functional characteristic of auditory neurons, *Biol. Cybernet.* 42:133-143 (1981).

3. A. M. H. J. Aertsen, P. I. M. Johannesma, J. Bruijns, and P. W. M. Koopman, Second-order representation of signals, in *Localization and Orientation in Biology and Engineering*, (Varju and Schnitzler Eds.), Springer-Verlag, Berlin Heidelberg, 1984, pp. 38-41.

4. E. de Boer and P. Kuyper, Triggered correlation, *IEEE Trans. on Bio-Med. Eng.*, July 1968, pp. 169-179.

5. H. F. P. van den Boogaard, G. H. F. M. Hesselmans, and P. I. M. Johannesma, System identification based on point processes and correlation densities. I. The nonrefractory model, *Math. Biosci.* 80:143-171 (1986).

6. D. R. Cox and V. Isham, *Point Processes*, Monographs on Applied Probability and Statistics, Chapman and Hall, London, 1980.

7. J. J. Eggermont, P. I. M. Johannesma, and A. M. H. J. Aertsen, Reverse correlation methods in auditory research, *Quart. Rev. Biophys.* 16(3):341-414 (1983).

8. W. J. M. Epping and J. J. Eggermont, Single unit characteristics in the auditory midbrain of the immobilized grassfrog, *Hearing Res.* 18:223-243 (1985).

9. W. J. M. Epping and J J. Eggermont, Sensitivity of neurons in the auditory midbrain of the grassfrog to temporal characteristics of sound. II Stimulation with amplitude modulated sound, *Hearing Res.* 24:55-72 (1986).

10 C. C. A. M. Gielen, G. H. F. M. Hesselmans, and P. I. M. Johannesma, Sensory Interpretation of Neural Activity Patterns *Math. Biosci.* 87: (1987).

11. J. L. Grashuis, *The pre-event stimulus ensemble: an analysis of the stimulus-response relation of complex stimuli applied to auditory neurons*, Ph.D. Thesis, Nijmegen 1974.

12. D. J. Hermes, Separation of time and frequency, *Biol. Cybernet.* 52:109-115 (1985).

13. G. H. F. M. Hesselmans, H. F. P. van den Boogaard, and P. I. M. Johannesma, The characteristic functional of the peri-event stimulus ensemble, *Math. Biosci.* 85:211-230 (1987).

14. P. I. M. Johannesma, A. M. H. J. Aertsen, L. Cranen, and L. J. Th. O. van Erning, The phonochrome: A coherent spectro-temporal representation of sound, *Hearing Res.* 5:123-145 (1981).

15. P. I. M. Johannesma and A. M. H. J. Aertsen, The phonochrome: A coherent spectro-temporal representation of sound, in *Biology and Engineering*, (Varju and Schnitzler Eds.), Springer-Verlag, Berlin Heidelberg, 1984, pp. 42-45.

16. P. I. M. Johannesma and H. F. P. van den Boogaard, Stochastic formulation of neural interaction, *Acta Appl. Math.* 4:201-224 (1985).

17. P. I. M. Johannesma, A. M. II. J. Aertsen, H. F. P. van den Boogaard, J. J. Eggermont, and W. J. M. Epping, From synchrony to harmony: ideas on the function of neural assemblies and on the interpretation of neural synchrony, in *Brain theory* (G. Palm and A. Aertsen, Eds.), Springer-Verlag, Berlin, 1986, pp. 25-47.

18. N. G. van Kampen, *Stochastic Processes in Physics and Chemistry*, North-Holland, Amsterdam, 1981.

19. D. L. Snyder, *Random Point Processes*, Wiley, New York, 1975.

20. N. Yen, Time and frequency representation of acoustic signals by means of the Wigner distribution function: Implementation and interpretation, *J. Acoust. Soc. Am.* 81(6):1841-1850 (1987).

Distribution of Sensory Stimulus Estimated from Neural Activity

Gerard H. F. M. Hesselmans

Abstract

The relation between the *a priori* distribution of the stimulus ensemble as presented to the animal and the *a posteriori* distribution of the ensemble of stimuli associated with the occurrence of the induced neural activity pattern is studied within the context of a model. In three successive steps the model is further specified; orderly, multiplicative, and a Volterra expansion. A dynamic method is presented, yielding an estimate of the complete *a posteriori* stimulus distribution.

1 Introduction

By means of its sensory system an animal obtains information about its environment. Sensors transform stimuli into neural activity patterns. The responses of primary neurons to stimuli of different modality provide the only representation of the environment to the central nervous system. The neural signals can be used to construct an internal representation of the environment. After an evaluation by the nervous system command signals for the effector systems are generated. Effectors, such as muscles, allow the animal to interact with its environment.

The neural activity patterns are comprised by trains of action potentials propagating in the nerve-fibers. The short wave forms (1 ms) of the action potentials are to a good approximation identical in case of one neuron. Therefore a neural activity pattern is almost completely described by the times of occurrence of identical neural events. A mathematical tool to describe such an abstracted phenomenon is provided by point-process theory (Snyder 1975).

The analysis of a series of events by means of point-process theory leads to a description of the train of events, e.g. it can be modelled by a renewal process or by a Gauss-Poisson process. For the animal it is however more relevant to know the relation between the neural activity pattern and the possible stimulus, i.e. a sensory interpretation of neural activity. Proposals for a stimulus based explanation of neural activity have been presented before (for references

see e.g. Gielen et al. 1987). However in these publications the stimulus-space is projected onto a low-dimensional subspace, e.g. stimulus intensity and frequency in the auditory field and intensity, saturation and hue perception in vision research.

Johannesma (1981) gave an impulse to a more general sensory interpretation of neural activity and formulated an estimation method based upon the characteristic stimuli of a set of independent low frequency auditory neurons. Along the same line Gielen et al. (1987) estimated the most plausible stimulus. This paper presents a way to estimate the complete stimulus.

The structure of this paper is as follows. In Section 2 the neural activity is idealized by an orderly point process. Under this rather mild restriction with respect to the neuron model, general results describing the evolution of the stimulus estimate can be formulated. In fact orderliness is a property of neural activity which is a direct consequence of the refractory period: neurons do not fire more than once within about 1 msec. The conditional distribution changes as more and more neural activity is observed in time. In Sections 3 and 4 the neural model is limited to more concrete forms. First, the sensory input and the neural feedback mechanism are separated. As a result the neural feedback does not appear explicitly in the equations. Second, an explicit model of stimulus dependency is given. The model that will be used is adopted from Johannesma and van den Boogaard (1985). The description of the output point process is intensity based (Cox 1972; Snyder 1975). The intensity $\lambda(t)$ depends upon a record of the data available at time t, including the history of the output point process. In the analysis we prefer the more intuitive and more tractable intensity approach above the alternative martingale description (Segall and Kailath 1975; Bremaud 1981). Consequently the intensity approach is used in this paper.

2 Time dependent description of sensory stimuli

The aim of this paper is to characterize the stimulus given the neural activity in a certain time interval $[0,t)$. The stimulus $x(\cdot)$ is stochastic and not directly observable, but its *a priori* distribution $f[x]$ is known. The dimensionality of the stimulus is as yet of no importance, but depends upon the sensory system under investigation and the number of sensors, e.g. one or two eyes. As a function of time it may be a scalar for the monaural auditory system or a two-dimensional spatial intensity distribution for the visual system. It is in the final model that we will restrict ourselves two the one dimension of sound: pressure fluctuations of the air as a function of time at one ear.

The neuron responds to the stimulus by generating action potentials. It is through the neural activity pattern that an *a posteriori* estimate of the stimulus can be obtained. Some stimuli result in characteristic neural activity patterns. The inverse also holds: a neural activity pattern might point at a specific stimulus or a specific set of stimuli. The presence of the pattern changes the *a*

priori statistics $f[x]$ of the stimulus ensemble. The stochastics of the 'new' conditional ensemble given the activity pattern $N(\cdot)$ on $[0, t)$ can be described by a conditional probability distribution $f[x|N(\tau), \tau \in [0, t)]$. If the condition is to be written in full in each equation, expressions will be awkward. In order to improve the readability of the coming expressions, the history of the neural activity from time zero till the end of the observation interval at time t will be abbreviated by $\mathcal{H}(t)$. Using this new convention the conditional stimulus distribution is denoted by $f[x|\mathcal{H}(t)]$. We will first consider a situation, where we have one stimulus input and one sensory neuron.

We assume that the probability of an event to occur is influenced both by the stimulus process $x(\cdot)$ and by the previous events. In general different stimuli mean different neural activity properties. E.g. the autocorrelation density of the sequence of action potentials is stimulus dependent (Eggermont et al. 1983). In order to arrive at an explicit result the class of neural models is limited.

Assumption 1 *The neural activity can be modelled by an orderly output point process $N(\cdot)$, depending on the sensory stimulus $x(\cdot)$ and history $\mathcal{H}(t)$ of the point process.*

This is expressed by:

$$P[\Delta N(t) = 0|x, \mathcal{H}(t)] \quad = \quad 1 - \lambda[x, \mathcal{H}(t)](t)\Delta t + o(\Delta t) \tag{1}$$

$$P[\Delta N(t) = 1|x, \mathcal{H}(t)] \quad = \quad \lambda[x, \mathcal{H}(t)](t)\Delta t + o(\Delta t) \quad \text{as } \Delta t \downarrow 0 \tag{2}$$

$$P[\Delta N(t) > 1|x, \mathcal{H}(t)] \quad = \quad o(\Delta t) \tag{3}$$

where

$\Delta N(t)$ = number of neural events in the interval $[t, t + \Delta t)$.
$\lambda[\cdot](\cdot)$ = intensity function (Cox and Isham 1980).

Equations (1,2,3) show the characteristic property of orderly point processes (Daley 1974). The probability of 2 or more events to occur in a small time increment can be neglected (Snyder 1975; Cox and Isham 1980). Multiple events are ruled out. From neurophysiological data it is known that neurons exhibit inhibition: directly after the occurrence of an event, the generation of another event is completely suppressed during about 1 ms and partially the next few milliseconds. As a result the right-hand side of Equation (3) equals zero if Δt is less than the absolute refractory period; $P[\Delta N(t) > 1|x, \mathcal{H}(t)] = 0$.

Equations (1,2,3) give the probability of events to occur given both stimulus input $x(\cdot)$ and neural history $\mathcal{H}(t)$. The conditional probabilities given only the neural history are given by:

$$P[\Delta N(t) = 0|\mathcal{H}(t)] \quad = \quad 1 - \lambda[\mathcal{H}(t)](t)\Delta t + o(\Delta t) \tag{4}$$

$$P[\Delta N(t) = 1|\mathcal{H}(t)] \quad = \quad \lambda[\mathcal{H}(t)](t)\Delta t + o(\Delta t) \quad \text{as } \Delta t \downarrow 0 \tag{5}$$

$$P[\Delta N(t) > 1|\mathcal{H}(t)] \quad = \quad o(\Delta t) \tag{6}$$

with $\lambda[\mathcal{H}(t)](t) = E[\lambda[\mathbf{x}, \mathcal{H}(t)](t)|\mathcal{H}(t)]$ the average intensity function. The 'unconditional' point process is also orderly. Given the past of the process, the probability of two or more events to occur in a small time increment can still be neglected.

Neural activity is not limited to a certain time span, it evolves in time. Given the estimated stimulus distribution $f[x|\mathcal{H}(t)]$ at time t and the number of events $\Delta N(t)$ in the next time increment $[t, t + \Delta t)$, we would like to estimate the stimulus distribution $f[x|\mathcal{H}(t + \Delta t)]$ at time $t + \Delta t$. Note that the neural history up to time $t + \Delta t$ can be described by $\mathcal{H}(t + \Delta t)$ or by $\mathcal{H}(t)$ combined with $\Delta N(t)$. By applying a few times Bayes' theorem concerning conditional probabilities:

$$P[A|B] \cdot P[B] = P[AB] = P[B|A] \cdot P[A], \tag{7}$$

the stimulus distribution at time $t + \Delta t$ can be related to the one at time t.

$$f[x|\mathcal{H}(t + \Delta t)] = \frac{P[\Delta N(t)|x, \mathcal{H}(t)]}{P[\Delta N(t)|\mathcal{H}(t)]} f[x|\mathcal{H}(t)] \tag{8}$$

The influence of the addition of neural information is determined by knowledge from the past. Note that Equation (8) holds for all $\Delta N(t)$. Substitution of the orderly model (see Equations (1, 2, 3) and (4, 5, 6) yields difference equations for updating the estimated stimulus distribution:

$$\Delta f[x|\mathcal{H}(t), \Delta N(t)] := f[x|\mathcal{H}(t), \Delta N(t)] - f[x|\mathcal{H}(t)] = \tag{9}$$

$$\begin{cases} -\Delta\lambda[\mathcal{H}(t)](t)\left[\frac{\lambda[x,\mathcal{H}(t)](t)}{\lambda[\mathcal{H}(t)](t)} - 1\right] f[x|\mathcal{H}(t)] + o(\Delta t) & \text{if } \Delta N(t) = 0 \\ \left[\frac{\lambda[x,\mathcal{H}(t)](t)}{\lambda[\mathcal{H}(t)](t)} - 1\right] f[x|\mathcal{H}(t)] + o(1) & \text{if } \Delta N(t) = 1 \end{cases}$$

The values of the conditional density are irrelevant in case $\Delta N(t) \geq 2$, as they always will be multiplied by factors of order $P[\Delta N(t) \geq 2] = o(\Delta t)$. Equation (9) shows how the conditional distribution is changed by the addition of neural information present in the activity pattern. The change in the probability distribution can be described by a jump if an action potential occurs and by a small step proportional to Δt if no neural activity is present. During the absence of neural activity the stimulus properties drift at a rate of $\lambda[\mathcal{H}(t)](t)$ in the direction opposite to the one imposed by the possible neural events.

A similar result in case $\Delta N(t) = 1$ was obtained by Johannesma (1980) and Hesselmans et al. (1987) for the Fourier/Laplace transform of the distribution. This transform is known as the characteristic functional (see Section 5). Strictly speaking their results were only formulated in the context of a stationary stimulus. However following the rationale given by Hesselmans et al. (1987) no reason could be found why this result should not be applicable in the nonstationary case as well.

Note that during an absolute refractory period no events occur. As a consequence the evolution of the conditional distribution is completely described

by the case $\Delta N(t) = 0$ of Equation (9), in which moreover the intensity $\lambda[x, \mathcal{H}(t)](t) = 0$. So we may conclude that the conditional distribution $f[x|\mathcal{H}(t)]$ does not change during the absolute refractory period.

Without loss of precision Equation (9) can be integrated to a single version:

$$\Delta f[x|\mathcal{H}(t), \Delta N(t)] = \tag{10}$$
$$\{\Delta N(t) - \Delta t \, \lambda[\mathcal{H}(t)](t)\} \left[\frac{\lambda[x,\mathcal{H}(t)](t)}{f[x|\mathcal{H}(t)]} - 1 \right] + o(\Delta t) \quad \text{if } \Delta N(t) \leq 1 \,.$$

The change of distribution is determined by what actually happens: $\Delta N(t)$, and what is expected to happen: $\lambda[\mathcal{H}(t)](t)$. Unless $\lambda[\mathcal{H}(t)](t) = 0$ these two quantities cannot be equal.

Since $E[\Delta N(t)|\mathcal{H}(t)] = \Delta t \lambda[\mathcal{H}(t)](t)$, the conditional distribution is on the average not changed by additional neural information.

In addition the size of the change depends on the ratio of $\lambda[x, \mathcal{H}(t)](t)$ and $\lambda[\mathcal{H}(t)](t)$. If the neural activity does not depend upon the stimulus then $\lambda[x, H(t)](t) = \lambda[H(t)](t)$ and the distribution is not changed.

Equation (10) is the fundamental equation of this paper and allows, at least in principle, a solution of the distribution problem. More explicit results concerning the quantities $\lambda[x, \mathcal{H}(t)](t)$ and $\lambda[\mathcal{H}(t)](t)$ will be given in the next sections.

3 A multiplicative model

The results of Section 2 are quite general. In order to arrive at more specific results the generality of Section 2 will be restricted by additional assumptions about the model. The model is inspired by Johannesma and van den Boogaard (1985). In the model two types of variables occur: action and state variables. The action variable is the action potential. The neural activity and the sensory input are integrated deterministically to a state variable: the generator potential, which is stochastic as a function of a stochastic phenomenon. The generator potential governs the generation of an output event. In order to separate the neural input from the stimulus input we propose (see also Cox 1972; Brillinger 1975; Johnson and Swami 1983):

Assumption 2 *The contribution of neural and sensory input to the intensity function is multiplicative.*

The mathematical formulation of the preceding reads:

$$\lambda[x, \mathcal{H}(t)](t) = \exp U[x, \mathcal{H}(t)](t) \tag{11}$$
$$U[x, \mathcal{H}(t)](t) = V[x](t) + W[\mathcal{H}(t)](t) \tag{12}$$
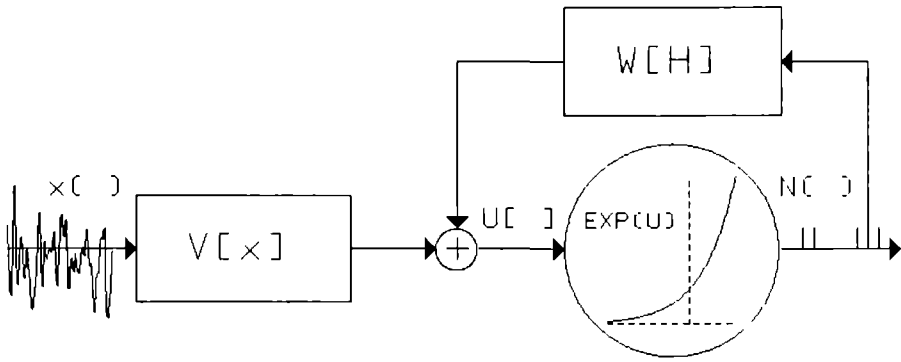
where

75

Figure 1: *Multiplicative model. The generator potential $U[x, \mathcal{H}]$ is the sum of the sensory and neural contributions, $V[x]$ respectively $W[\mathcal{H}]$. The output intensity depends exponentially upon the generator potential.*

$U[x, \mathcal{H}(t)](t)$ = generator potential
$V[x](t)$ = stimulus contribution to the generator potential
$W[\mathcal{H}(t)](t)$ = neural contribution to the generator potential.

A flow diagram of the model is given in figure 1. Substitution of Equations (11, 12) in the two quantities of interest $\lambda[\mathcal{H}(t)](t)$ and $\lambda[x, \mathcal{H}(t)](t)$ yields respectively:

$$
\begin{aligned}
\lambda[\mathcal{H}(t)](t) &= E[\lambda[\mathbf{x}, \mathcal{H}(t)](t) | \mathcal{H}(t)] \\
&= E[\exp\{V[\mathbf{x}](t) + W[\mathcal{H}(t)](t)\} | \mathcal{H}(t)] \\
&= E[\exp V[\mathbf{x}](t) | \mathcal{H}(t)] \exp W[\mathcal{H}(t)](t), \quad (13)
\end{aligned}
$$

and

$$
\begin{aligned}
\frac{\lambda[x, \mathcal{H}(t)](t)}{\lambda[\mathcal{H}(t)](t)} &= \frac{\exp V[x](t) + W[\mathcal{H}(t)](t)}{E[\exp V[\mathbf{x}](t) | \mathcal{H}(t)] \exp W[\mathcal{H}(t)](t)} \\
&= \frac{\exp V[x](t)}{E[\exp V[\mathbf{x}](t) | \mathcal{H}(t)]} \quad (14)
\end{aligned}
$$

Substitution of Equation (14) into Equation (10) leads to:

$$
\Delta f[x|\mathcal{H}(t), \Delta N(t)] = \quad (15)
$$
$$
\{\Delta N(t) - \Delta t \lambda[\mathcal{H}(t)](t)\} \left[\frac{\exp V[x](t)}{\int dx \, f[x|\mathcal{H}(t)] \exp V[x](t)} - 1 \right] f[x|\mathcal{H}(t)] + o(\Delta t)
$$
$$
\text{if } \Delta N(t) \leq 1.
$$

The jump in the conditional stimulus distribution induced by the occurrence of an event does not explicitly depend upon the neural feedback mechanism $W[\mathcal{H}(t)](t)$. Only indirectly via the intensity $\lambda[\mathcal{H}(t)](t)$ and through the change of the distribution in the absence of neural events the feedback is of importance.

4   A Volterra model

In this section a more explicit form is given of the model by Johannesma and van den Boogaard (1985). The stimulus component $V[x](t)$ is chosen to be a Volterra system, whose order is limited to kernels of maximal degree 2. The addition of higher degree kernels would lead to mathematically unmanageable formulas, while an analysis by Grashuis (1974) shows that the activity of primary auditory neurons can be described by just 2 kernels. Note that Grashuis (1974) did not include refractory mechanisms. Results are likely to improve, when this neural feedback is taken into account.

Assumption 3 *The stimulus is transformed by a second degree Volterra system.*

The mathematical formulation of the preceding results in:

$$V[x](t) = v_0 + \int d\sigma\, c(\sigma)x(t-\sigma) + \frac{1}{2} \int d\sigma \int d\tau\, D(\sigma,\tau)x(t-\sigma)x(t-\tau) \quad (16)$$

where

v_0 = working point of the neuron
$c(\cdot)$ = linear filter
$D(\cdot,\cdot)$ = nonlinear filter of second degree

The contribution of the sensory input is processed nonlinearly. Given the conditional stimulus distribution at time t, the influence of the occurrence of an event in the next time increment on the conditional distribution can be solved by means of Equation (15).

The relation between the original distribution and the one estimated from neural activity is in general complicated. However if *Gaussian distributions* are assumed, it suffices to relate the *a priori* and *a posteriori* mean and covariance or the first two moments, which completely determine both distributions. First and second-order properties have simple meanings of which the first needs no comment. The second has two interpretations. Firstly, the covariance, gives the uncertainty with respect to the average or expected stimulus value and secondly the second moment is directly related to the concept of a spectro-temporal receptive field (Aertsen and Johannesma 1981).

If the stimulus distribution given the neural activity up to time t is Gaussian then it is fully described by its average $m(\cdot|\mathcal{H}(t))$ and covariance $B(\cdot,\cdot|\mathcal{H}(t))$. Substitution of Equation (16) into Equation (15) yields the *a posteriori* average and covariance (see Johannesma (1980); Hesselmans et al. (1987)):

$$m_t^e = B_t^e[B_t^{-1}m_t + c_t] \quad (17)$$
$$B_t^e = [B_t^{-1} - D_t]^{-1} \quad (18)$$

where

$$m_t(\tau) = m(\tau|\mathcal{H}(t))$$
$$B_t(\tau,\sigma) = B(\tau,\sigma|\mathcal{H}(t))$$
$$m_t'(\tau) = m(\tau|\mathcal{H}(t), \Delta N(t) = 1)$$
$$B_t'(\tau,\sigma) = B(\tau,\sigma|\mathcal{H}(t), \Delta N(t) = 1)$$
$$c_t(\tau) = c(t - \tau)$$
$$D_t(\tau,\sigma) = D(t - \tau, t \quad \sigma).$$

The stimulus remains Gaussian with mean and covariance given by Equations (17, 18). The addition of an extra neural event increment $[t, t + \Delta t)$ changes the stimulus parameters with a jump but leaves the stimulus class unaltered. Unfortunately this class-invariance property is lost during a period of no neural activity. This can be shown through a study of the case $\Delta N(t) = 0$ of Equation (9).

Although the assumptions of a second-order Volterra system and a Gaussian stimulus distribution have some advantages, the loss of class-invariance limits the applicability of the Gaussian approximation to neurons with weak connectivities. This problem cannot be overcome by limiting the Volterra system to just one kernel. Other explicit formulations of stimulus distributions that have some realism are not available. Therefore, in the next section, the distribution format will be put aside in favour of its Fourier/Laplace transform, the characteristic functional. In this domain it is possible to drop the Gaussian assumption.

5 The characteristic functional

A stimulus description equivalent with the distribution formulation is the characteristic functional (Srinivasan 1974; Snyder 1975; van Kampen 1981). The characteristic functional $\Phi[\xi]$ of the stochastic continuous stimulus process $\mathbf{x}(\cdot)$ is defined by the following expectation value:

$$\Phi[\xi] := E[\exp \int d\sigma\, \mathbf{x}(\sigma)\xi(\sigma)]. \tag{19}$$

Knowledge of the characteristic functional $\Phi[\xi]$ for an appropriate set of test functions $\xi(\cdot)$ supplies a complete description of the stochastic process $\mathbf{x}(\cdot)$. The conditional characteristic functional given the neural activity is given by:

$$\Phi[\xi|\mathcal{H}(t)] := E[\exp \int d\sigma\, \mathbf{x}(\sigma)\xi(\sigma)|\mathcal{H}(t)]. \tag{20}$$

Although descriptions by means of the distribution and the characteristic functional are equally powerful, expressions in either format do not have to be equally simple. A straightforward transformation of Equation (15) yields:

$$\Delta\Phi[\xi|\mathcal{H}(t), \Delta N(t)] = \tag{21}$$
$$\{\Delta N(t) - \Delta t\lambda[\mathcal{H}(t)](t)\}\left[\frac{E[\exp\{V[\mathbf{X}](t) + \int ds\, \mathbf{X}(s)\xi(s)\}|\mathcal{H}(t)]}{E[\exp V[\mathbf{X}](t)|\mathcal{H}(t)]} - \Phi[\xi|\mathcal{H}(t)]\right]$$

if $\Delta N(t) \leq 1$, with (recall Equation (13))

$$\lambda[\mathcal{X}(t)](t) = E[\exp V[\mathbf{x}](t)|\mathcal{X}(t)] \exp W[\mathcal{X}(t)](t) . \qquad (22)$$

Equations (21, 22) give the time evolution of the characteristic functional under the influence of neural activity. The *a priori* stimulus distribution is still arbitrary. It is of course possible to assume as in the previous section a second degree Volterra system, but this would make Equations (21, 22) unsolvable or reduce the stimulus class to approximate Gaussian distributions. Observe the linear form in the exponent of the characteristic functional. This linear form fits well with a first degree Volterra system.

$$V[x](t) = v_0 + \int d\sigma \, c(\sigma)x(t-\sigma) = v_0 + \int d\sigma \, c_t(\sigma)x(\sigma) \qquad (23)$$

Substitution of Equation (23) into Equations (21 ,22) yields:

$$\Delta\Phi[\xi|\mathcal{X}(t), \Delta N(t)] = \{\Delta N(t) - \Delta t \lambda[\mathcal{X}(t)](t)\} \left[\frac{\Phi[\xi + c_t|\mathcal{X}(t)]}{\Phi[c_t|\mathcal{X}(t)]} - \Phi[\xi|\mathcal{X}(t)] \right] \qquad (24)$$

if $\Delta N(t) \leq 1$, with

$$\lambda[\mathcal{X}(t)](t) = \Phi[c_t|\mathcal{X}(t)] \exp\{v_0 + W[\mathcal{X}(t)](t)\} . \qquad (25)$$

Equations (24, 25) form a closed set, which describes the evolution of a stimulus estimate with arbitrary *a priori* distribution.

6   Stimulus description based on multi-unit activity

The results obtained in this paper in case of one neuron, can easily be generalized to the multi-unit case. Due to the orderliness of the set of point processes, the probability of 2 events of different neurons to occur in the same small time increment can be neglected. The n-unit equivalent of Equation (8) is:

$$f[x|\mathcal{X}(t), \Delta N_j(t)] = \frac{P[\Delta N_j(t)|x, \mathcal{X}(t)]}{P[\Delta N_j(t)|\mathcal{X}(t)]} f[x|\mathcal{X}(t)] \qquad (26)$$

where $\mathcal{X}(t)$ is a description of *all* point processes under consideration. If more units are considered, then Equation (26) is the new starting point of the analysis. As a result the basic Equation (9) is modified as follows. The influence of the addition of an event is described by a set of formulas, one for each neuron. The behavior in the absence of neural events is determined by an equation where

the right-hand side is replaced by a sum of individual contributions.

$$\Delta f[x|\mathcal{H}(t), \Delta N_j(t) \text{ all } j] = \qquad (27)$$

$$\begin{cases} -\Delta t \sum_{j=1}^{n} \lambda_j[\mathcal{H}(t)](t) \left[\frac{\lambda_j[x.\mathcal{H}(t)](t)}{\lambda_j[\mathcal{H}(t)](t)} - 1 \right] f[x|\mathcal{H}(t)] + o(\Delta t) \\ \qquad \qquad \text{if } \sum_j \Delta N_j(t) = 0 \\ \left[\frac{\lambda_j[x.\mathcal{H}(t)](t)}{\lambda_j[\mathcal{H}(t)](t)} - 1 \right] f[x|\mathcal{H}(t)] + o(1) \\ \qquad \qquad \text{if } \Delta N_j(t) = 1 \text{ and } \sum_i \Delta N_i(t) = 1 \end{cases}$$

with $\lambda_j[\cdot]$ the intensity function of process j. The first case of Equation (27) applies if no event occurs in the time interval $[t, t + \Delta t)$, the second case if exactly one event occurs, whatever its type. Recall that the occurrence of more events has probability of smaller order than Δt. A more compact formulation of Equation (27) is:

$$\Delta f[x|\mathcal{H}(t), \Delta N_j(t) \text{ all } j] = \qquad (28)$$

$$\sum_j \left\{ \Delta N_j(t) - \Delta t \lambda_j[\mathcal{H}(t)](t) \right\} \left[\frac{\lambda_j[x \ \mathcal{H}(t)](t)}{\lambda_j[\mathcal{H}(t)](t)} - 1 \right] f[x|\mathcal{H}(t)] + o(\Delta t)$$
$$\text{if } \sum_j \Delta N_j(t) \le 1.$$

The change in distribution equals the sum of the contributions of the individual neurons. In general Equation (28) cannot be separated explicitly, since neural activity is correlated through direct neural interaction and/or indirectly via the stimulus.

7   Discussion

In this paper a dynamic estimation procedure based on neural activity is presented. This estimate changes stepwise if an action potential is detected, while the behavior is continuously in the interval between two neural events.

Starting with a Gaussian distributed stimulus and a neuron model with connectivities which are not too large, the *a posteriori* stimulus distribution associated with a neural activity pattern is also approximately Gaussian. In this case the time dependence of the stimulus parameters: mean and covariance, can be determined.

The results obtained in this paper in case of one neuron, can easily be generalized to the multi-unit case. If more units are considered, then the basic Equation (10) should be modified. The history $\mathcal{H}(t)$ should be extended to a description of *all* point processes under consideration. In general, an estimate of the stimulus based upon several neurons cannot be separated into contributions of estimates based upon the activity patterns of the individual neurons. If neural interaction is present then $\lambda_j[x, \mathcal{H}(t)](t) \ne \lambda_j[x, \mathcal{H}_j(t)](t)$, with $\mathcal{H}_j$ the history of the neural activity of neuron j, and $\mathcal{H}$ the history of the activity pattern of the set of neurons. Even if neural connections are absent then still

$\lambda_j[\mathcal{H}(t)](t) \neq \lambda_j[\mathcal{H}_j(t)](t)$ due to stimulus correlation. Sufficient conditions that allow complete separation in Equation (28) appear to be:

1. neural interaction is absent,

2. neurons respond to different stimulus features; receptive fields do not overlap,

3. stimulus features for which neurons are sensitive are independent.

If more neurons are considered it is increasingly unlikely that these conditions are met. Unfortunately these conditions appear to be necessary for a separation of the multi-unit estimate into a combination of single-unit estimates.

In this paper the rationale is provided for a more detailed sensory interpretation of neural activity. Since the evolution of the distribution can be traced, the behavior of stimulus characteristics like moments, cumulants or most plausible stimulus can be investigated as well. Gielen et al. (1987) proposed a first-order estimate for a most plausible stimulus. Under the assumption of Gaussian distributions this is equivalent to the estimation of the first moment of the *a posteriori* stimulus distribution. This approach can easily be extended to the second-order case. Within what range of parameters the method is valid may be investigated by simulations. Afterwards it can be applied to real neural data.

References

1. A. M. H. J. Aertsen and P. I. M. Johannesma, The spectro-temporal receptive field; a functional characteristic of auditory neurons, *Biol. Cybernet.* 42:133-143 (1981).

2. P. Bremaud, *Point Processes and Queues; Martingale Dynamics*, Springer, New York, 1981.

3. D. R. Brillinger, The identification of point process systems, *Ann. Probab.* 3:909-924 (1975).

4. D. R. Cox, The statistical analysis of dependencies in point processes, in *Stochastic Point Processes*, (P. A. Lewis Ed.), Wiley, New York, 1972, pp. 55-66.

5. D. R. Cox and V. Isham, *Point Processes*, Monographs on applied probability and statistics, Chapman and Hall, London, 1980.

6. D. J. Daley, Various concepts of orderliness for point-processes, in *Stochastic Geometry*, (E. F. Harding and D. G. Kendall Eds.), Wiley, London, 1974, pp. 148-161.

7. J. J. Eggermont, W. J. M. Epping, and A. M. H. J. Aertsen, Stimulus dependent neural correlations in the auditory midbrain of the grassfrog (Rana temporaria L.), *Biol. Cybernet.*, 47:103-117 (1983).

8. C. C. A. M. Gielen, G. H. F. M. Hesselmans, and P. I. M. Johannesma, Sensory interpretation of neural activity patterns, *Math. Biosci.* (1987) (in press)

9. J. L. Grashuis, *The Pre-Event Stimulus Ensemble: An Analysis of the Stimulus-Response Relation of Complex Stimuli Applied to Auditory Neurons*, Thesis, University of Nijmegen, 1974.

10. G. H. F. M. Hesselmans, H. F. P. van den Boogaard, and P. I. M. Johannesma, The characteristic functional of the peri-event stimulus ensemble, *Math. Biosci.* 85:211-230 (1987).

11. P. I. M. Johannesma, Functional identification of auditory neurons based on stimulus event correlation, in *Psychophysical, Physiological and Behaviour Studies in Hearing* (G. Brink and F. A. Bilsen, Eds.), Delft U. P., Delft, 1980, pp. 77-84.

12. P. I. M. Johannesma, Neural representation of sensory stimuli and sensory interpretation of neural activity, in *Advances in Physiol. Sciences*, Vol 30: *Neural communication and control.* pp.103-126, Pergamon Press, 1981.

13. P. I. M. Johannesma and H. F. P. van den Boogaard, Stochastic formulation of neural interaction, *Acta Appl. Math.* 4:201-224 (1985).

14. D. H. Johnson and A. Swami, The transmission of signals by auditory-nerve fiber discharge patterns, *J. Acoust. Soc. Am.* 74(2):493-501 (1983).

15. N. G. van Kampen, *Stochastic Processes in Physics and Chemistry*, North-Holland, Amsterdam, 1981.

16. A. Segal and T. Kailath, The modeling of randomly modulated jump processes, *IEEE Transactions on information theory*, IT-21:135-143 (1975).

17. D. L. Snyder, *Random Point Processes*, Wiley, New York, 1975.

18. S. K. Srinivasan, *Stochastic Point Processes and Their Applications*, Griffin, London, 1974.

Maximum Entropy, Correlation Functions, and Exponential Models

Gerard H. F. M. Hesselmans

Abstract

Input-output relations of a neuron can be studied through correlation functions. Input and output processes are assumed to be stationary. Based on a maximum-entropy argument we conclude that the interpretation of correlation functions can be done in an 'optimal' form by assuming a model consisting of a Volterra system followed by an exponential event generator.

1 Introduction

To investigate complex neural networks correlation techniques (de Boer and Kuiper 1968, Eggermont et al. 1983) have been used. Conclusions about functional connections are made, based upon these correlations, but no explicit model is formulated. Conclusions drawn from the correlations functions, however, often assume implicitly a linear or low-order Volterra system. The aim of this paper is to form an explicit description of the neural transducer, based on the correlation functions obtained from experiments. The neuron is described completely by *all* joint correlation functions of its input and output. However, only the first can be estimated reliably from experimental data. A prudent method is required to fill in the remaining uncertainty: making minimal constraints on the neural activity given some correlation functions or equivalently making the process as random as possible.

Neural activity can be modelled by a set of point processes. Point processes are fully described by product densities. The product densities are the theoretical counterparts of the experimentally observed correlation functions. Probability distributions of a set of stochastic point processes are determined by their joint product densities.

The above problem is reformulated mathematically as follows. Determine an optimal set of probability distributions given just a few product densities. Optimal is as indicated before to be interpreted as: maximizing the increase in entropy of the point processes. The entropy increase is maximized at every

time instant. Of course results obtained at different times need to be consistent. The definition of entropy as given by McFadden (1965) will be used. McFadden determined the process yielding a maximum increase of entropy given only the first product density or intensity of the process, and found it to be the (in)homogeneous Poisson point process. In this paper the problem is generalized to constraining product densities of higher order.

The use of a maximum-entropy increase estimate formalizes an assumption of ignorance about aspects of the distribution other than the functions explicitly used as constraints. The maximum-entropy approach yields a set of extremization conditions. These conditions result directly in a generator of the point process, or less formally, a model of the neuron. The maximum-entropy distribution itself, however is not derived, but is determined indirectly by the model. Also the error introduced by using estimates of the product densities instead of their exact values is not discussed. We will argue that as a consequence of the decision to use correlation functions a low-order Volterra system is not an optimal choice, but should be followed by an exponential event generator. For clarity this statement is first made in case of one point process. Then it is extended to two point processes followed by a general conclusion. A treatise concerning continuous stochastic processes in which the total entropy is maximized, is given by Victor and Johannesma (1986), and yields distributions which are described by an exponentiation of a low-order Volterra series.

In Section 2 the product densities and sample function densities are introduced. They are related to each other and their time dependence is discussed. In Section 3 the time dependence of the point process is given. In Section 4 the concept of entropy as a measure of disorder is introduced. A definition based upon sample function densities is given and in Section 5 the increase of entropy is maximized under the restriction of known correlations. In Section 6 the results of the previous sections are generalized to two and more processes.

2   Densities

Consider a point process $N(\cdot)$. This process is observed during a time interval $[0, t)$. The events which constitute realizations of the point process occur at random instants T_i, $0 \leq T_1 \leq T_2 \leq \ldots$. The number of events in intervals $[0, t)$ is given by the random variable $N(t)$, $N(t) \geq 0$. For ease of notation vector variables like $(t_1, \ldots, t_n)$ will be abbreviated by $\mathbf{t}_n$. The point process on $[0, t)$ can be completely determined in two ways. The product densities $f_n(\mathbf{t}_n)$ are defined by (Snyder 1975):

$$f_n(\mathbf{t}_n) := \lim_{\max \Delta t_i \downarrow 0} E \left[\prod_{i=1}^{n} \frac{\Delta N(t_i)}{\Delta t_i} \right]. \tag{1}$$

The sample function densities $p_n(t_n; t)$ are defined on $[0, t)^n$ by (Snyder 1975):

$$p_n(t_n) := \lim_{\max \Delta t_i \downarrow 0} E \left[1_{[N(t)=n]} \prod_{i=1}^{n} \frac{\Delta N(t_i)}{\Delta t_i} \right],$$ (2)

where $1_{[N(t)=n]}$ is the indicator function of the outcome $[N(t) = n]$.

The product densities yield the probability to find events at different times t_i irrespectively of what happens between events. This in contrast with the sample function densities for which we explicitly require that no intermediate events occur. The sample function density gives the joint probability for a realization of $N(t) = n$ events which occur at different times $T_i = t_i$. We assume that all the n-dimensional functions of both densities exist. Observe that the random instants T_i are ordered, while the arguments of $f_n(t_n)$ and $p_n(t_n; t)$ are not. For convenience of notation, we have extended them to functions on $[0, t)^n$ which are invariant for permutation of the components of t_n. Without this extension formulae (3) and (4) below would become rather awkward. As a consequence of this extension, the weights assigned to $[0, t)^n$ by $f_n(t_n)$ and $p_n(t_n; t)$ have to be reduced by a factor $n!$.

Note that the product densities $f_n(t_n)$ do not depend upon a particular time interval $[0, t)$, this makes them easier to estimate from experimental data and more suitable for a first analysis of the experimental data. The product densities are the theoretical counterparts of the correlation functions. However the sample function densities are more suited for a theoretical analysis when a model is known. The sample function densities yield a more transparent definition of the entropy (see next section).

The product densities $f_n(t_n)$ are related to the sample function densities $p_n(t_n; t)$ by (Stratonovich 1963):

$$f_n(t_n) = \sum_{k=0}^{\infty} \frac{1}{k!} \int_{[0,t)^k} du_k \, p_{n+k}(t_n, u_k; t).$$ (3)

The inverse is given by:

$$p_n(t_n; t) = \sum_{k=0}^{\infty} \frac{(-1)^k}{k!} \int_{[0,t)^k} du_k \, f_{n+k}(t_n, u_k; t).$$ (4)

Henceforth $\int du_k$ for $k = 0$ is to be interpreted as evaluation of the integrand at t, i.e., $p_n(t_n; t)$ in (3).

3 Time evolution of the process

The evolution of the point process is determined by the underlying model. In view of present understanding of neural behavior the point process is represented

as a generalized birth process or a conditional Poisson process. The probability of an event to occur in the time increment $[t, t + \Delta t)$ given the history of $N(t) = n$ events having occurred at times $\mathbf{t}_n$ is given by:

$$P[\Delta N(t) = 0 | N(t) = n; \mathbf{T}_n = \mathbf{t}_n] \quad = \quad 1 - \lambda_n(t|\mathbf{t}_n) + o(\Delta t) \tag{5}$$

$$P[\Delta N(t) = 1 | N(t) = n; \mathbf{T}_n = \mathbf{t}_n] \quad = \quad \lambda_n(t|\mathbf{t}_n) + o(\Delta t) \qquad \text{as } \Delta t \downarrow 0, \tag{6}$$

$$P[\Delta N(t) > 1 | N(t) = n; \mathbf{T}_n = \mathbf{t}_n] \quad = \quad o(\Delta t) \tag{7}$$

with $\lambda_n(t|\mathbf{t}_n)$ the conditional intensity of the point process. The occurrence of more than one event in a small time increment is of smaller order than Δt and can be neglected. This is called orderliness (Daley 1974). The property of the process being conditionally Poisson can be used to obtain two equations of continuity. For all $n \geq 0$, we have the following (McFadden 1965):

$$\frac{d}{dt} p_n(\mathbf{t}_n; t) \quad = \quad -\lambda_n(t|\mathbf{t}_n) p_n(\mathbf{t}_n; t) \tag{8}$$

$$p_{n+1}(\mathbf{t}_n, t; t) \quad = \quad \lambda_n(t|\mathbf{t}_n) p_n(\mathbf{t}_n; t). \tag{9}$$

The process changes from a state of n events into a state of $n + 1$ events at a rate proportional to $\lambda_n(t|\mathbf{t}_n)$ and the probability densities of these two states are related via $\lambda_n(t|\mathbf{t}_n)$. The $\lambda_n(t|\mathbf{t}_n)$ are the 'generators' of the probability distribution of the point process. Together with the initial conditions and Equations (8,9), they determine the process completely.

Note that the model cannot be chosen arbitrarily. It is subject to the constraints imposed by the experimental observations: the correlation functions or estimated product densities. The product densities $f_n(\mathbf{t}_n)$ and the conditional intensity $\lambda_n(t|\mathbf{t}_n)$ are related as follows:

$$f_{n+1}(t, \mathbf{t}_n) \quad = \quad \sum_{k=0}^{\infty} \frac{1}{k!} \int_{[0,t)^k} d\mathbf{u}_k \, p_{n+k+1}(t, \mathbf{t}_n, \mathbf{u}_k; t) \tag{10}$$

$$= \quad \sum_{k=0}^{\infty} \frac{1}{k!} \int_{[0,t)^k} d\mathbf{u}_k \, \lambda_{n+k}(t|\mathbf{t}_n, \mathbf{u}_k) p_{n+k}(\mathbf{t}_n, \mathbf{u}_k; t). \tag{11}$$

In the first step (10) the relation between sample function densities and product densities as expressed by Equation (3) is used. In the second step (11) the equation of continuity (9) has been inserted.

The remaining task is to provide a satisfactory form of the 'model' $\lambda_n(t|\mathbf{t}_n)$ constrained by (11). If m product densities $f_n(\mathbf{t}_n)$ are estimated from experimental data then the conditional intensities $\lambda_n(t|\mathbf{t}_n)$ are subject to the m constraints given by Equation (11).

4   Entropy

The entropy of a process is a measure of its disorder, or equivalently, the uncertainty about which realization is to be expected. It indicates the amount of

information needed to fix or predict a particular realization of the process. The entropy $H(t)$ of a point process in the interval $[0, t)$ is defined by (McFadden 1965):

$$H(t) = -\sum_{n=0}^{\infty} \frac{1}{n!} \int_{[0.t)^n} dt_n \, p_n(\mathbf{t}_n; t) \log p_n(\mathbf{t}_n; t). \tag{12}$$

Note that a different notation was used by McFadden (in our version of (12) the integration domain has been extended to all of $[0, t)^n$). In order to investigate the increase of the entropy we now consider the time derivative:

$$\frac{d}{dt} H(t) = \quad - \sum_{n=0}^{\infty} \frac{1}{n!} \int_{[0.t)^n} dt_n \, p_{n+1}(\mathbf{t}_n, t; t) \log p_{n+1}(\mathbf{t}_n, t; t)$$

$$\quad - \sum_{n=0}^{\infty} \frac{1}{n!} \int_{[0.t)^n} dt_n \, \frac{d}{dt} p_n(\mathbf{t}_n \ ; t) \left[1 + \log p_n(\mathbf{t}_n \ ; t)\right]. \tag{13}$$

The first term results from the differentiation of the multiple integral with respect to the outer limit t and shifting the summation variable n, and the second term results from the differentiation of the integrand. Substitution of the equations of continuity (8,9) into (13) leads to:

$$\frac{d}{dt} H(t) = \sum_{n=0}^{\infty} \frac{1}{n!} \int_{[0.t)^n} dt_n \, p_n(\mathbf{t}_n; t) \lambda_n(t|\mathbf{t}_n) \left[1 - \log \lambda_n(t|\mathbf{t}_n)\right]$$

$$= E\left[\lambda_N(t|\mathbf{T}_N)\left(1 - \log \lambda_N(t|\mathbf{T}_N)\right)\right]. \tag{14}$$

The expectation in Equation (14) is taken over all random variables $\mathbf{T}_N$, including N. The increase of the entropy of the process at time t depends upon the distribution of the process up to time t, and upon the growth of the process via the conditional intensity $\lambda_n(t|\mathbf{t}_n)$. We may extend the usual convention $0 \log 0 := 0$ for entropies to expressions with $\lambda_n(t|\mathbf{t}_n)$ at the place of 0.

5 Maximum-entropy increase

Up to this point we have followed closely the reasoning by McFadden. We will generalize McFaddens approach to the case where:

Assumption 1 *The first m product densities are given (also after time t).*

McFadden assumed only the intensity to be known ($m = 1$). The first product density depends upon just one time argument. The stochastic properties of the process at different times are independent. In this paper, higher-order product densities are considered as well. Therefore the stochastics of the process at time t depends on the realization of the process before time t. As a consequence the time derivative of the entropy at time t depends also on the distribution of

the point process before time t. In order to get *consistent* results with respect to the distribution of the process we impose the following relation:

$$p_n(\mathbf{t}_n; t) = \sum_{k=0}^{\infty} \frac{(-1)^k}{k!} \int_{[t,t+\tau)^k} d\mathbf{u}_k \, p_{n+k}(\mathbf{t}_n, \mathbf{u}_k; t + \tau). \qquad (15)$$

for $\tau \geq 0$. The distribution of the process on $[0, t)$ is not changed on the average by what happens on $[t, t+\tau)$. The distribution on $[0, t + \Delta t)$ can, with the help of the continuity equations (8,9), be related to the distribution on $[0, t)$.

$$p_n(\mathbf{t}_n; t + \Delta t) = [1 - \lambda_n(t|\mathbf{t}_n)] \, p_n(\mathbf{t}_n; t + \Delta t) + o(\Delta t) \qquad (16)$$

$$p_{n+1}(\mathbf{t}_{n+1}; t + \Delta t) = \lambda_n(\mathbf{t}_{n+1}|\mathbf{t}_n) p_n(\mathbf{t}_n; t) + o(\Delta t), \qquad (17)$$

as $\mathbf{t}_n \in [0, t)^n$ and $t_{n+1} \in [t, t + \Delta t)$. Note that Equation (16) describes the case of zero events in $[t, t+\Delta t)$ and (17) the case of one event in $[t, t+\Delta t)$. Recall that the occurrence of more events has probability of smaller order than Δt, and can be neglected. With the aid of Equations (16,17) the sample function densities $p_n(\mathbf{t}_n; t)$ can be obtained by recursion. The initial values are: $p_0(t = 0) = 1$ and $p_n(\mathbf{t}_n; t = 0) = 0$ for $n \geq 1$ and $\mathbf{t}_n = \mathbf{0}$. Due to the consistency relation (15) and Equations (16,17) the maximization of the entropy increase at time t should be done only with respect to $\lambda_n(t|\mathbf{t}_n)$ once the distribution on $[0, t)$ is determined.

Consider as fixed the quantities: t, $\mathbf{t}_n$, and $p_n(\mathbf{t}_n; t)$ for $n \in N$, $\mathbf{t}_n \in [0, t)^n$. We will maximize $\frac{dH}{dt}$ in (14) for varying $\lambda_n(t|\mathbf{t}_n)$ $n \in N$, under the restraints in (11) for $n = 1, 2, \ldots, m$ and $\mathbf{t}_n \in [0, t)^n$. We may and will restrict our attention to vectors $\mathbf{t}_n$ with all components different. Consider the $\mu_n(\mathbf{t}_n; t)$ as "Lagrange multipliers" in the maximization procedure of:

$$\frac{d}{dt} H(t) \;-\; \sum_{n=0}^{m-1} \int_{[0,t)^n} d\mathbf{v}_n \, \mu_n(\mathbf{v}_n; t) \times$$
$$\left[f_{n+1}(t, \mathbf{v}_n) - \sum_{k=0}^{\infty} \int_{[0,t)^k} d\mathbf{u}_k \, \lambda_{n+k}(t|\mathbf{v}_n, \mathbf{u}_k) p_{n+k}(\mathbf{v}_n, \mathbf{u}_k; t) \right] (18)$$

Let, for fixed n and $\mathbf{t}_n$, $\delta_n(\cdot - \mathbf{t}_n)$ be the sum of the $n!$ Dirac δ-functions with peaks at $\mathbf{t}_n$ and all its permutations. Replace in (18) the $\lambda_n(t|\mathbf{u}_n)$ by $\lambda_n(t|\mathbf{u}_n) + \varepsilon \delta_n(\mathbf{u}_n - \mathbf{t}_n)$, differentiate to ε, and obtain the result for $\varepsilon = 0$. Solve $\mu_n(\mathbf{t}_n; t)$ from all these results for varying n and $\mathbf{t}_n$ to find:

$$\left[\log \lambda(t|\mathbf{t}_n) - \sum_{j=0}^{m-1} \mu_j(\mathbf{t}_n; t)^* \right] p_n(\mathbf{t}_n|t) = 0 \qquad (19)$$

Here "*" expresses that a sum over all subsets of size j out of the n time arguments $\mathbf{t}_n$ should be taken, e.g. $\mu_3(\mathbf{t}_3; t)^* = \mu_2(t_1, t_2; t) + \mu_2(t_1, t_3; t) +$

$\mu_2(t_2, t_3; t)$. The reader who verifies Equation (19) might come up with an additional factor $j!$. This depends on how the symmetry of the time arguments t_j is taken care of. However this additional factor can easily be absorbed in the $\mu_j(t_j; t)$. The $\mu_j(t_j; t)$ are determined by (19) and (11) for $n = 0, 1, \ldots, m - 1$. If $p_n(\mathbf{t}_n; t) = 0$, then the value of $\lambda_n(t|\mathbf{t}_n)$ is irrelevant for the distribution of the point process, so we may simplify (19) in all cases to:

$$\lambda_n(t|\mathbf{t}_n) = \exp \sum_{j=0}^{m-1} \frac{1}{j!} \mu_j(\mathbf{t}_n; t)^+ \tag{20}$$

$$= \exp \sum_{j=0}^{m-1} \int_{[0\,t)^j} dN(\mathbf{s})_j \, \mu_j(\mathbf{s}_j; t) \tag{21}$$

where $dN(\mathbf{s})_j$ is an abbreviation of $dN(s_1) \cdots dN(s_j)$. Equation (20,21) determines a class of generators $\lambda_n(t|\mathbf{t}_n)$ of the point process. Less formally it gives an 'optimal' model which might underly the correlation functions estimated from experiments. Of course the parameters of the model still have to be determined. The Lagrange multipliers $\mu_j(\mathbf{t}_j; t)$ determine via Equations (8,9) the distribution $p_n(\mathbf{t}_n|t)$ of the process, now also an infinitesimal time beyond t. Substitution of these model-dependent sample function densities in Equations (11) yield a set of constraining equations according to which the parameter functions $\mu_j(\mathbf{t}_j; t)$ should be chosen.

The case $m = 1$, corresponds to a result of McFadden (1965). The entropy increase is a maximum if the process is the (inhomogeneous) Poisson process. Note that the increase of entropy is not maximized by varying the distribution of the point process in the past, but by a proper choice of the conditional intensities $\lambda_n(t|\mathbf{t}_n)$, which regulate the distribution in $[0, t + dt)$ at its right end.

At this point we introduce a second assumption:

Assumption 2 *The point process has a finite memory h.*

In case of a strictly finite memory the integral in Equation (21) can be restricted to $[0, h)$ for $t > h$:

$$\lambda_n(t|\mathbf{t}_n) = \exp \sum_{j=0}^{m-1} \frac{1}{j!} \int_{[0\,h)^j} dN(t - \mathbf{u})_j \, \mu_j(t - \mathbf{u}_j; t) \tag{22}$$

where $t - \mathbf{u}_j$ is an abbreviation of $t - u_1, \ldots, t - u_j$. Note that some elements in $\mathbf{t}_n$ have become obsolete. The intensity of the point process at time t depends only on those elements of $\mathbf{t}_n$ for which $t_i > t - h$. In formula, if $\tilde{\lambda}_k(t|\mathbf{u}_k)$ is the conditional intensity at t given $N(t) \quad N(t - h) = k$ and occurrences of the point process at $t - \mathbf{u}_k$, then

$$\lambda_n(t|\mathbf{t}_n) = \tilde{\lambda}_k(t|\mathbf{u}_k) \text{ if } N(t) - N(t - h) = k \text{ and } \mathbf{u}_k := t - \mathbf{t}_j; \mathbf{t}_j > t - h \tag{23}$$
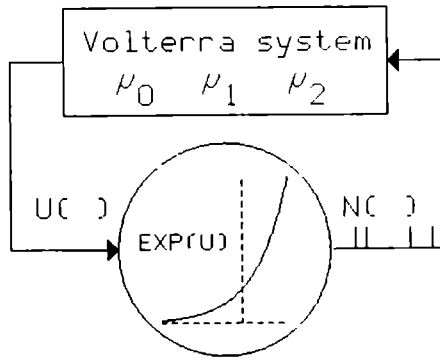
If we assume that:

Figure 1: *The optimal single unit model. It consists of a Volterra system and an exponential event generator*

Assumption 3 *The point process is stationary,*

then $\tilde{\lambda}_k(t|\mathbf{u}_k)$ in (23) does not depend on t. Consequently, $\mu_j(\mathbf{t}_j;t)$ in (22) is a function of $\mathbf{u}_j = t - \mathbf{t}_j$, which, by abuse of notation, we will write as $\mu_j(\mathbf{u}_j)$. We have found that in a time-invariant neuron model with finite memory the dependence on the past is a matter of relative times. Note that the results hold for an arbitrary time-horizon h. Therefore the results can be extended to processes with an infinite memory which can be approximated by a finite-memory process. Such processes are said to have a *fading memory*, i.e. two input signals which are close in the *recent past*, but not necessarily close in the *remote past* yield present outputs which are close (Boyd and Chua 1985).

In many analyses of experimental data only the intensity of the process $f_1(t) = f_1$, the autocorrelation $f_2(t_1, t_2) = f_2(t_2 - t_1)$, and sometimes the joint occurrence density or second-order Poisson kernel based on $f_3(t_1, t_2, t_3) = f_3(t_3 - t_1, t_3 - t_2)$ are established. Based on such experimental observations and the criterion of maximal entropy increase the following model is 'optimal':

$$
\begin{aligned}
\lambda_n(t|\mathbf{t}_n) &= \exp\left[\mu_0 + \sum_{i=1}^{n}\mu_1(t - t_i) + \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n}\mu_2(t - t_i, t - t_j)\right] \quad (24)\\
&= \exp\left[\mu_0 + \int_0^t dN(\sigma)\,\mu_1(t - \sigma) + \frac{1}{2}\int_0^t dN(\tau)\int_0^t dN(\sigma)\,\mu_2(t - \sigma, t - \tau)\right]
\end{aligned}
$$

The model consists of a Volterra system and an exponential event generator (see figure 1). The neuron model is time-invariant, and the dependence upon the past is a matter of relative times. The output is a conditional Poisson process. This model which is generalized in the next section, has been introduced before by Johannesma and van den Boogaard (1985) and elaborated by van den

90

Boogaard (1986). In general we get:

$$\lambda_n(t|\mathbf{t}_n) = \exp \sum_{k=0}^{\infty} \frac{1}{k!} \int_{[0.t)^k} dN(\mathbf{v}_k)\, \mu_k(t - \mathbf{v}_k) \tag{25}$$

with $\mu_k(\mathbf{u}_k) = 0$ if $f_{k+1}(\mathbf{t}_{k+1})$ is unknown (not estimated),
and $\mu_k(\mathbf{u}_k)$ properly chosen if $f_{k+1}(\mathbf{t}_{k+1})$ is known.

The $f_{k+1}(\mathbf{t}_{k+1})$ need not be consecutive. The relations between the different quantities are summarized in the following scheme:

$$f_n \overset{(3,4)}{\Leftrightarrow} p_n \overset{(9)}{\Leftrightarrow} \lambda_n \overset{(25)}{\Leftrightarrow} \mu_n \tag{26}$$

The product densities f_n on the left represent the experimental data. They determine via the sample function densities p_n and the conditional intensities λ_n the "Lagrange multipliers" μ_n on the right. In the transformation from product densities into sample function densities *all* densities are required. However the set of product densities f_n is incomplete, therefore the maximum-entropy increase argument is introduced.

6 Multi-unit model

Results of the previous sections were related to only one neuron. In this section we first extend the results to two neurons and then results in the general case are given. The joint entropy of two point processes $N_1(\cdot)$ and $N_2(\cdot)$ is defined by (see Equation (12)):

$$H_{12}(t) = -\sum_{n=0}^{\infty} \frac{1}{n!} \int_{[0.t)^n} dt_n \sum_{m=0}^{\infty} \frac{1}{m!} \int_{[0.t)^m} ds_m p_{nm}(\mathbf{t}_n, \mathbf{s}_m; t) \log p_{nm}(\mathbf{t}_n, \mathbf{s}_m; t)$$
$$\tag{27}$$

where $p_{nm}(\mathbf{t}_n, \mathbf{s}_m; t)$ is the joint probability density to find events of point process $N_1(\cdot)$ at times $\mathbf{t}_n$ and events of point process $N_2(\cdot)$ at times $\mathbf{s}_m$. The joint product densities $f_{nm}(\mathbf{t}_n; \mathbf{s}_m)$ and sample function densities are related in a similar way as mentioned in Section 2. The sums should be replaced by double sums.

If the two point processes are modelled as generalized birth processes, then 3 equations of continuity can be formulated. For all $n \geq 0$, $m \geq 0$ we have the following (see Equations (8,9)):

$$\frac{d}{dt} p_{nm}(\mathbf{t}_n, \mathbf{s}_m; t) = -\left[\lambda^1_{nm}(t|\mathbf{t}_n, \mathbf{s}_m) + \lambda^2_{nm}(t|\mathbf{t}_n, \mathbf{s}_m)\right] p_{nm}(\mathbf{t}_n, \mathbf{s}_m; t) \tag{28}$$

$$p_{n+1.m}(t, \mathbf{t}_n, \mathbf{s}_m; t) = \lambda^1_{nm}(t|\mathbf{t}_n, \mathbf{s}_m) p_{nm}(\mathbf{t}_n, \mathbf{s}_m; t) \tag{29}$$

$$p_{n.m+1}(\mathbf{t}_n, t, \mathbf{s}_m; t) = \lambda^2_{nm}(t|\mathbf{t}_n, \mathbf{s}_m) p_{nm}(\mathbf{t}_n, \mathbf{s}_m; t) \tag{30}$$

with $\lambda^1_{nm}(t|t_n,s_m)$ and $\lambda^2_{nm}(t|t_n,s_m)$ the intensity functions of the respective point processes. Note that none of the Equations (28,29,30) contains a cross-term. This is a consequence of the fact that the bivariate process is also a conditional Poisson process. The occurrence of an event of one process and another event of the other process in a small time increment can be neglected. With the help of the new relations (27) and (28,29,30) the time derivative of the entropy of the combined process can be found (see also Equation (14)).

$$\frac{d}{dt}H_{12}(t) = \sum_{n=0}^{\infty}\frac{1}{n!}\int_{[0,t)^n}dt_n \sum_{m=0}^{\infty}\frac{1}{m!}\int_{[0,t)^m}ds_m\, p_{nm}(t_n,s_m;t) \times \qquad (31)$$

$$\left\{\lambda^1_{nm}(t|t_n,s_m)\left[1-\log\lambda^1_{nm}(t|t_n,s_m)\right]+\lambda^2_{nm}(t|t_n,s_m)\left[1-\log\lambda^2_{nm}(t|t_n,s_m)\right]\right\}$$

The extreme points of this derivative with respect to $\lambda^1_{nm}(t|t_n,s_m)$ and $\lambda^2_{nm}(t|t_n,s_m)$ have to be found subject to the following constraining relations (see Equation (11)).

$$f_{n+1,m}(t,t_n,s_m) = \qquad (32)$$

$$\sum_{k=0}^{\infty}\frac{1}{k!}\int_{[0,t)^k}du_k \sum_{j=0}^{\infty}\frac{1}{j!}\int_{[0,t)^j}dv_j\, \lambda^1_{n+k,m+j}(t|t_n,u_k,s_m,v_j)p_{n+k,m+j}(t_n,u_k,s_m,v_j;t)$$

$$f_{n,m+1}(t_n,t,s_m) = \qquad (33)$$

$$\sum_{k=0}^{\infty}\frac{1}{k!}\int_{[0,t)^k}du_k \sum_{j=0}^{\infty}\frac{1}{j!}\int_{[0,t)^j}dv_j\, \lambda^2_{n+k,m+j}(t|t_n,u_k,s_m,v_j)p_{n+k,m+j}(t_n,u_k,s_m,v_j;t)$$

If we assume again that the product densities $f_{nm}(t_n,s_m)$ have been estimated from experimental data up to order 3; $m+n \leq 3$, then by the calculus of variations we will find another similar set of conditions on extreme points of the constrained maximum entropy (see Equation (19)). As a consequence of the previous reasoning we find for the model of the first neuron (see Equation (24)):

$$\lambda^1_{nm}(t|t_n,s_m) = \exp\left[\mu^{12}_{00}\; + \int_0^t dN_1(\sigma)\mu^{12}_{10}(t-\sigma)+\int_0^t dN_2(\sigma)\mu^{12}_{01}(t-\sigma)\right.$$

$$+ \frac{1}{2}\int_0^t dN_1(\tau)\int_0^t dN_1(\sigma)\,\mu^{12}_{20}(t-\sigma,t-\tau)$$

$$+ \frac{1}{2}\int_0^t dN_2(\tau)\int_0^t dN_2(\sigma)\,\mu^{12}_{02}(t-\sigma,t-\tau)$$

$$\left. + \int_0^t dN_1(\tau)\int_0^t dN_2(\sigma)\,\mu^{12}_{11}(t-\sigma,t-\tau)\right] \qquad (34)$$

The superscript i in μ^i_j refers to the point process under consideration and the subscript j to the order of the constraining product density or correlation function. The model can be represented by a Volterra system followed by an exponential event generator (see figure 2).
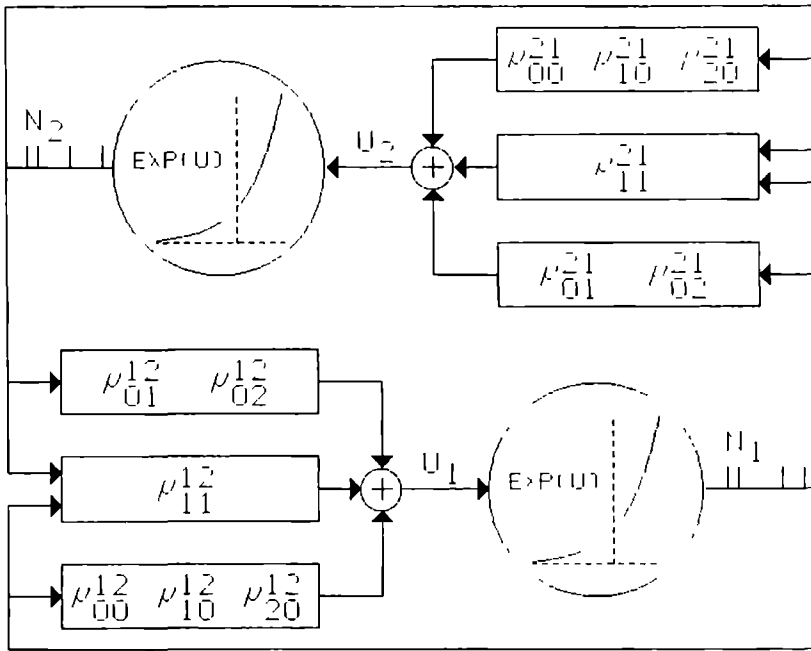
Figure 2: *The optimal double unit model. Neural input is processed by Volterra systems and event generation depends exponentially upon the output of the Volterra system.*

The order of the Volterra system corresponds to the order of the constraining functionals. This conclusion can be extended to relations between an arbitrary number of neurons and arbitrary order of the correlations.

Conclusion: *The 'optimal' model based on the measurement of correlation functions consists of Volterra systems followed by exponential event generators. The order of the kernels of the Volterra systems corresponds with the order of the experimentally obtained correlation functions.*

7 Conclusions and discussion

In this paper a model is proposed for a neural transducer. This model conforms to the experimental data, and the ignorance about the unknown aspects is formalized by the assumption of maximum-entropy increase. An approach along the line of a maximum-entropy increase instead of an overall maximum is taken since the latter leads to a maximum-entropy distribution and not to a 'model'. A transition from the distribution to a time-invariant model seems to be a difficult problem, since the distribution is influenced by the edges. The maximum-entropy increase approach suffers only from an onset phenomenon, the overall approach suffers also from offset problems. Furthermore in a stationary situation, we expect the entropy in the limit to be a linear function of time. In this situation both approaches may yield the same solutions.

The maximum-entropy distribution can be described by an exponentiation of a low-order Volterra series. This distribution is akin to the one given by Victor and Johannesma (1986). The distribution function here is a functional on the realizations of a point process, while the paper by Victor and Johannesma was concerned with continuous processes. In the second-order case the Volterra kernels of the latter are related to the cumulants of the continuous stochastic process. This relation does not exist for point process.

It might be possible to choose constraints in such a way that no maximizing model exists. Two results by van den Boogaard (1985) point at such a possibility. First, if we restrict ourselves to the single-unit case, the autocorrelation equals in first order the connectivity. Second, a 'positive' filter results in an explosion of activity and no point process to describe the model exists. Therefore a 'positive' autocorrelation will not match with a linear Volterra system followed by an exponential event generator. A way out of such a pathologic situation is to include more correlation functions or choose another model based upon other considerations.

The proposed model is only an approximation of the neuron under study. By incorporating the effect of higher-order correlations the model should approach the actual neuron. Since the presented model is basically an element of the subclass of Volterra systems which yield a nonnegative output, it will approach the 'real' model if this model is analytic.

In the literature the most common models belong to the class of Wiener-Volterra systems. The reason is the simplicity attributed to them. If we look at self-exciting systems, the only model solved explicitly is the linear model (Hawkes 1971). All elements of this type of population of mutually connected point processes are connected by linear filters and fire according to a linear pulse generator. Although the linear model allows the derivation of explicit forms of the filters, it is restricted to excitatory connectivities, which should be sufficiently bounded to avoid instabilities. These excitatory connections are in conflict with the inhibitory auto-connectivities of 'real' neurons which show refractory effects. If connectivity loops are included then both Wiener-Volterra systems and the model proposed in this paper can be solved only numerically. If connectivity loops are excluded then Wiener-Volterra systems seem to be preferable. However refractory mechanisms require feedback for a correct description of neural activity. In contrast with the presented model, nonnegative firing intensities, are not guaranteed by the Wiener-Volterra model. Therefore the argument that low-order Wiener-Volterra systems are mathematically easier to treat is invalid.

The author wishes to thank Peter Johannesma for his help in the interpretation of mathematical concepts, and Wim Vervaat for corrections on imprecise mathematical formulations.

References

1. E. de Boer and P. Kuyper, Triggered correlation, *IEEE Trans. Bio-Med. Eng.*, July 1968, pp. 169-179.

2. H. F. P. van den Boogaard, *Transformation of point processes, correlation functions and system identification. II. Self exciting point processes*, Ph.D. Thesis, Nijmegen, 1985

3. H. F. P. van den Boogaard, G. H. F. M. Hesselmans, and P. I. M. Johannesma, System identification based on point processes and correlation densities. I. The nonrefractory model, *Math. Biosci.* 80:143-171 (1986).

4. S. Boyd and L. O. Chua, Fading memory and the problem of approximating nonlinear operators with volterra series, *IEEE Transactions on circuits and systems*, VOL. CAS-32, No. 11, November 1985.

5. D. J. Daley, Various concepts of orderliness for point processes, in *Stochastic Geometry*, (E. F. Harding and D. G. Kendall, Eds.), Wiley, London, 1974, pp. 148-161.

6. J. J. Eggermont, P. I. M. Johannesma, and A. M. H. J. Aertsen, Reverse correlation methods in auditory research, *Quart. Rev. Biophys.* 16(3):341-414 (1983).

7. A. G. Hawkes, Spectra of some self-exciting and mutually-exciting point processes, *Biometrika* 58:83-90 (1971).

8. A. G. Hawkes, Point spectra of some mutually exciting point processes, *J. R. Statist. Soc.* 833:438-443 (1971).

9. P. I. M. Johannesma and H. F. P. van den Boogaard, Stochastic formulation of neural interaction, *Acta Appl. Math.* 4:201-224 (1985).

10. P. I. Kuznetsov and R. L. Stratonovich, A note on the mathematical theory of correlated random points, in *Nonlinear Transformations of Stochastic Processes*, pp. 101-115. (P. I. Kuznetsov, R. L. Stratonovich, and V. I. Tikhonov, Eds.), Pergamon, Oxford, 1965.

11. J. A. McFadden, The entropy of a point process, *J. Soc. Indust. Appl. Math.*, 13(4):988-994 (1965).

12. D. L. Snyder, *Random Point Processes*, Wiley, New York, 1975.

13. J. D. Victor and P. I. M. Johannesma, Maximum-entropy approximations of stochastic nonlinear transductions: an extension of the Wiener theory, *Biol. Cybern.* 54:289-300 (1986).

Summary and discussion

The nervous system can be represented by a network, a system of cell bodies (somata) connected by axons and dendrites. The axons and dendrites are the wiring along which the cell bodies communicate by means of short electrical pulses or action potentials. The triple of soma, axon and dendrites forms an anatomical unit called a neuron. Because of their brevity compared to other relevant phenomena, and the similarity of different action potential waveforms generated by a single neuron, a sequence of action potentials is sufficiently characterized by their times of appearance. Such an abstraction is called a point process. Point-process theory has been applied before to quite different kinds of random processes as queues and earthquakes.

A neural network can be studied through its input-output relations. The usual procedure is to present a sensory stimulus, e.g. a sound, and determine the response of the network, a realization of a point process. By presenting a representative sample of signals and recording the evoked response, the neural network can be characterized. However this method provides no answer to questions concerning the function of neural networks, such as: "What is the functional difference between two neural networks?". To answer this question it is more relevant to record the response of both networks upon the presentation of the same stimulus, and form an estimation of the plausibility of stimuli to have been the cause of this response. If both probability distributions are similar, then the function of the two neural networks may be said to be similar.

Note that a direct comparison of the neural output of the two neural networks is not feasible since a direct measure on point processes is not available at the moment. Furthermore it makes no sense to look at details of the neural activity. A comparison of letters in two books written in English and German tells very little about the subjects of both books.

The reverse mapping from neural activity patterns back to the stimulus-space is the central issue of this thesis. The conditional stimulus distribution given a realization of the neural activity is investigated. This transformation can be studied once a model is set up. The selected model of auditory neurons is based on experimental work by Grashuis (1974) and a general theoretical study by Johannesma and van den Boogaard (1985). Grashuis used the same special case of the general model to predict the response of auditory neurons in the cochlear nucleus of the cat (forward approach). In this thesis the auditory system of the grassfrog (*Rana temporaria L.*) is modeled and the model is used
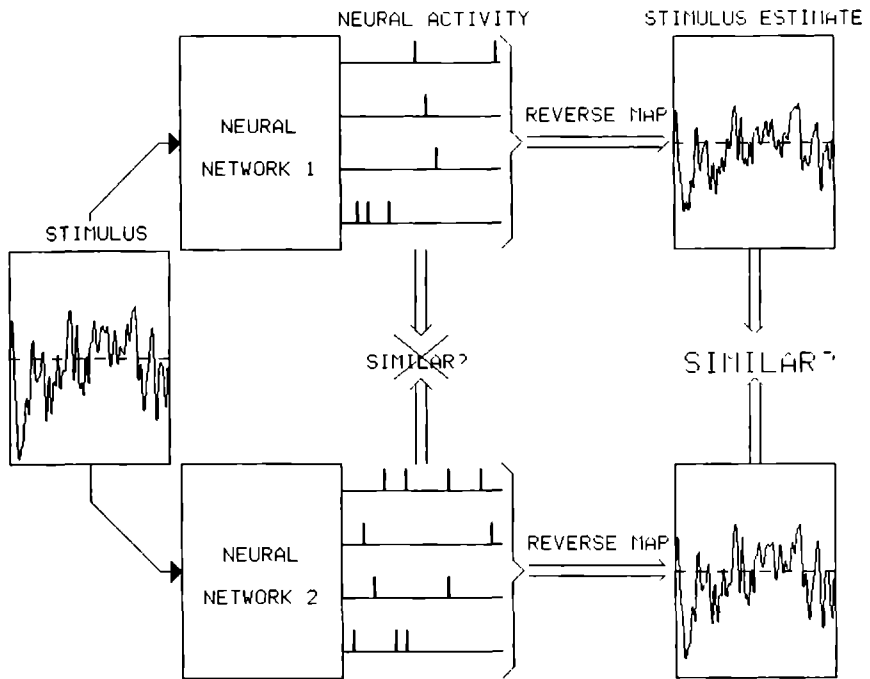
97

Figure 1: *The functional comparison of two neural networks done through reverse maps of the neural activity.*

in the reverse approach. The model consists of a Volterra system followed by an exponential event generator.

The implications by this model on the structure of neural processes are discussed in Chapter 6. Some information about the structure in a sequence of action potentials is provided by the experimentally determined correlation functions. It appears that the above model imposes the least possible additional structure on the neural activity.

Neural data from the grassfrog have been taken, because this animal has been studied extensively. Furthermore, since the grassfrog is a lower vertebrate, the mapping induced by its auditory system is expected to be not too complicated.

The reverse mapping given the recorded neural activity results in a probability distribution for possible sensory (auditory) stimuli. This probability distribution cannot be visualized in general, because of its high dimensionality. In Chapter 2 the maximum or most plausible stimulus is used as a one-dimensional characteristic of the distribution. The reverse mapping is determined for a set of sixty-four parallel model neurons, which are composed of a Volterra system followed by an exponential pulse generator.

Two cases are considered. In case of a linear system (first degree Volterra) a unique most plausible stimulus is shown to exist. In case of a second degree Volterra system, most plausible stimuli are shown to form a hyperplane. The linear case is further elaborated and a first approximation of the reverse map is given. In this approximation each neural event is substituted by the impulse response of the Volterra system of the associated neuron. The summation of impulse responses gives an estimate of the most plausible stimulus. Computer simulations using the responses of the set of parallel neurons after presentation of clicks, gamma-tones and a frequency sweep yield a satisfactory similarity of original and estimate.

Experimental data indicate that the linear model can be applied to primary neurons in the nervus acusticus or to neurons in the dorsal medullary nucleus, since these neurons show phase-lock. The generation of action potentials is synchronized to the phase of excitatory tonal stimuli. However phase-lock is largely lost in the subsequent stages in the processing of auditory stimuli, as neurons respond equally to stimuli with the same frequency but with different phase. Therefore a more general model, e.g. a second degree Volterra system is needed to model neurons in the torus semicircularis (TS). Furthermore, a unique most plausible stimulus can no longer be found when phase-lock is absent.

In Chapter 3 a general description of the stimulus distribution is provided by the characteristic functional. It can be considered as an infinite-dimensional Fourier/Laplace transform of the probability distribution. By functional differentiation moments and cumulants can be derived from the characteristic functional. The moments and cumulants are the theoretical counterparts of the experimental correlation functions. If the distribution is Gaussian, it is fully determined by its first two moments or cumulants. The characteristic func-

tional can not only be calculated for the stimulus ensemble (SE), but also for the set of stimuli associated with the generation of an action potential: the peri-event stimulus ensemble (PESE) which is an extension of the pre-event stimulus ensemble (Aertsen et al. 1980).

The characteristic functionals of SE and PESE are related by the model. This relation provides a means for system identification. In case of the linear Volterra system the impulse response of the filter can be found as follows. Present Gaussian white noise to the auditory neuron, record its response and determine the set of stimuli occurring prior to the action potentials. The average of this PESE gives the time-reverse of the linear filter.

Explicit relations between SE, PESE, and model are obtained for the 'second degree' model in combination with a Gaussian or Gauss-Poisson SE. The theoretical results are supported by computer simulations and applied to experimental data. In particular, the model is fitted to data from a single neuron in the TS of the grassfrog. Although the filter (linear Volterra system) is only adjusted to fit to the average conditional stimulus, second order properties (covariance) can be explained by the model as well. This adds to the credibility of the model.

In Chapter 4 the ideas presented in Chapters 2 and 3 are combined to characterize the mapping by neurons from the torus semicircularis. From Chapter 2 the idea is taken to substitute each neural event by a characteristic quantity. The impulse responses used previously had zero mean. This cannot be expected in general. Therefore the reverse mapping procedure is slightly modified by subtracting the characteristic quantity when no action potential occurs. From Chapter 3 we use the idea of taking for the characteristic quantity the difference between stimulus characteristics of PESE and SE, and the moment description. The basic elements are given by the difference of what is to be expected a priori and what is expected given the occurrence of a single event.

To estimate the first characteristic moment quantity, the average stimuli of SE and PESE should be subtracted from each other, and the difference should be used in the way mentioned above. An analogue procedure should be applied to estimate the higher order moments. For neurons in the TS without phase-lock, the averages of a Gaussian SE and PESE are equal and as a result the estimate of the first moment of the conditional distribution equals zero. In the Gaussian case the first moment equals the most plausible stimulus which is already studied in Chapter 2. Therefore the second moment is used to study the mapping by neurons in the TS (Chapter 5).

Actually, not the second moment is used, but rather the Wigner CoSTID. The Wigner CoSTID, a coherent spectro-temporal intensity density, yields the frequency content of the stimulus as a function of time, and can be obtained by a linear transformation of the second moment. The Wigner CoSTID is chosen, because it allows easier interpretation than the second moment. Elementary Wigner CoSTID's of 6 neurons in the TS are obtained during stimulation with

100

Gaussian noise. These are used to interpret the responses to a mating-call masked with noise. The conditional Wigner CoSTID resembles the a priori one, although some of the coherence in the original mating-call is lost (see Hermes 1985).

The estimation procedure can be divided into two parts. First, the characteristic quantities are determined. Second, these quantities are substituted for the neural events in the sensory interpretation of a neural activity pattern. To get an 'optimal' result, both steps require a SE with the same properties.

In Chapter 2, a Gaussian SE is used implicitly in the first step, while in the next step the response to three quite different stimuli is interpreted. The results indicate that the method is rather insensitive for a transfer from one SE to another.

In Chapter 4, again a Gaussian SE is used to characterize six neurons in the TS of the grassfrog. An ensemble of natural stimuli would be more appropriate, however these stimuli from the environment or biotope of the animal are very diverse (Aertsen et al. 1979). The acoustic biotope of the grassfrog includes wind, mating calls, the clatter of a stork, clicks etc. The obtained elementary Wigner CoSTID's seem to be good enough for interpreting the response to a mating-call in a noisy background.

With the interpretation of the response to the masked mating-call, most of the noise is lost. The estimated Wigner CoSTID resembles more the CoSTID of the pure mating-call than the corrupted one. The six neurons of the TS perform a kind of selection. This part of the nervous system of the grassfrog seems to be more sensitive to the mating-call than to noise.

In Chapter 5 the general solution to the interpretation problem is derived. This result does not change the qualitative properties of the sensory interpretation of neural activity. The occurrence of an action potential changes the estimate with a jump, while the estimate drifts in the opposite direction in the absence of events. However, the quantitative details are different in the exact procedure. The changes which are kept fixed in Chapters 2 and 4, depend upon the previous neural activity and the current estimate of the stimulus distribution. Furthermore the stimulus characteristics (moments) cannot be estimated independently, but should be considered simultaneously. The estimation procedures applied in Chapters 2 and 4 can be considered as first-order approximations.

Chapter 5 provides an exact solution to the interpretation of multi-unit activity and gives the conditions under which the previously sketched approximations might be applicable. They are:

1. neural interaction is absent,

2. neurons respond to different stimulus features: receptive fields do not overlap,

3. stimulus features for which neurons are sensitive are independent.

Examples in this thesis are chosen to satisfy these conditions as far as possible. In particular condition 1) is fulfilled all the time. Restrictions 2) and 3) do not require more explicit knowledge about the neuron model. Therefore an estimation procedure including overlapping stimulus sensitivities might be developed in the near future. Inclusion of neural interaction is harder to tackle. The introduction of interconnections in system identification leads to recursive relations. Partial solution of these equations is in general impossible. They should be solved simultaneously. Unfortunately full knowledge of the (functional) connections in the model is required for obtaining the exact conditional stimulus distribution given a realization of neural activity.

By comparing the presented stimulus and the conditional stimulus distribution, information about the selectivity and sensitivity of the nervous system is obtained. It is in the stimulus-space that functional measures on neural activity patterns can be found. Therefore the sensory interpretation of neural events gives a better understanding of the image the animal might have of its environment.

References

1. A. M. H. J. Aertsen, J. W. T. Smolders, and P. I. M. Johannesma, Neural representation of the acoustic biotope: On the existence of stimulus-event relations for sensory neurons, *Biol. Cybernet.* 32:175-185 (1979).

2. A. M. II. J. Aertsen, P. I. M. Johannesma, and D. J. Hermes, Spectro-temporal receptive fields of auditory neurons in the grassfrog. II. Analysis of the stimulus-event relation for tonal stimuli, *Biol. Cybernet.* 38:235-248 (1980).

3. J. L. Grashuis, *The pre-event stimulus ensemble: an analysis of the stimulus-response relation of complex stimuli applied to auditory neurons*, Ph.D. Thesis, Nijmegen 1974.

4. D. J. Hermes, Separation of time and frequency, *Biol. Cybernet.* 52:109-115 (1985).

5. P. I. M. Johannesma and H. F. P. van den Boogaard, Stochastic formulation of neural interaction, *Acta Appl. Math.* 4:201-224 (1985).

Samenvatting

De interactie tussen een dier en zijn omgeving wordt geregeld door het zenuw-stelsel. Informatie over de buitenwereld wordt door de zintuigen geregistreerd en doorgegeven aan de hersenen. Aldaar vindt een nadere analyse plaats, en worden indien nodig spieren geactiveerd.

In dit proefschrift wordt de relatie tussen de prikkel uit de buitenwereld en de gemeten activiteit in de hersenen bestudeerd. De verzameling prikkels die gebruikt is bestaat uit geluiden die voor het proefdier, de bruine kikker, en voor de (theoretisch-)onderzoeker van belang kunnen zijn: ruis, klikken, toonstootjes, paringsroepen enz. Gemeten is in de grootste auditive kern (torus semicircularis) in de middenhersenen van de bruine kikker (*Rana temporaria L.*).

De electrische activiteit in de hersenen wordt veroorzaakt door zenuwcellen. Een zenuwcel bestaat uit drie delen: de dendrietenboom, het cellichaam en het axon. De dendrietenboom verwerkt de signalen van andere zenuwcellen tot de generatorpotentiaal, die de toestand van de zenuwcel beschrijft. Afhankelijk van de waarde van de generatorpotentiaal wordt er in het cellichaam een actiepoten-tiaal gegenereerd. De actiepotentiaal wordt via het axon verder getransporteerd naar andere zenuwcellen of naar spieren.

De gedurende een experiment gemeten signalen in de hersenen van de bruine kikker bestaan uit actiepotentialen. Dit zijn spanningspulsjes met een duur van 1 milliseconde. De vorm van deze pulsjes varieert nauwelijks voor iedere zenuwcel afzonderlijk en de duur ervan is kort vergeleken met de tijdschaal waarop verschijnselen in de buitenwereld zich afspelen. Daarom kan bij de beschrijving van een rij actiepotentialen volstaan worden met de tijdstippen waarop ze optreden. De wiskundige theorie die hierop van toepassing is heet puntprocessentheorie. Deze theorie is al eerder toegepast op o.a. wachtrij-problemen en aardbevingen.

Het patroon van actiepotentialen in centrale delen van de hersenen ziet er in het algemeen chaotisch uit. In minder centrale delen, dat wil zeggen dichter bij een zintuig (oor), kan de neurale activiteit meer direct in verband gebracht wor-den met de prikkel (geluid). Ook hier geldt echter dat het door een enkel geluid opgewekte neurale activiteitspatroon sterk varieert. De vraag, hoe groot deze variatie is kan echter niet onmiddellijk worden beantwoord, omdat er nog geen goede afstandsmaat tussen reeksen actiepotentialen gevonden is. Afstanden tussen geluiden kunnen echter wel gegeven worden. Hierdoor wordt het indi-

rect toch mogelijk om over het verschil tussen twee reeksen actiepotentialen te spreken. Daarvoor is het nodig om het geluid dat geassocieerd kan worden met de neurale activiteit te kennen.

Hetzelfde probleem treedt op als men de functie van twee verschillende zenuwstelsels of delen daarvan wil vergelijken. Het direct vergelijken van de door beide zenuwstelsels voortgebrachtte reeksen actiepotentialen is een zinloze bezigheid. Zoals bij het vergelijken van Duitse en Franse woorden, waarbij men niet naar uitspraak of spelling moet kijken maar naar betekenis, is het van meer belang om datgene waarnaar de reeksen actiepotentialen verwijzen te onderzoeken.

Het *centrale probleem* van dit proefschrift luidt: welke prikkel of verzameling prikkels kan de oorzaak zijn van de gemeten activiteit in de hersenen.

Zonder een model kan bovenstaande vraag niet beantwoord worden. Daarom is er een model geformuleerd dat voldoende aansluit bij experimentele gegevens en dat wiskundig nog hanteerbaar is. Gegeven dit model is geprobeerd onder andere de meest waarschijnlijke prikkel, de gemiddelde prikkel en de frequentieinhoud van de geluidsprikkel te schatten. De geschatte prikkel gegeven de gemeten reeksen actiepotentialen noemen we de *sensorische interpretatie* van deze hersenactiviteit.

In eerste benadering blijken deze grootheden geschat te kunnen worden met een som van elementaire stukjes signaal. Schatting van bijvoorbeeld het meest waarschijnlijke geluid (Hoofdstuk 2) in de buitenwereld gebeurt door op momenten dat een actiepotentiaal optreedt een vast stukje signaal bij de bestaande schatting op te tellen en gedurende de afwezigheid van actiepotentialen de schatting in de omgekeerde richting bij te stellen.

De elementaire stukjes signaal worden gevonden uit het verschil van de vooraf verwachte prikkeleigenschap en de gemiddelde prikkeleigenschap waarop een actiepotentiaal volgt van de betreffende zenuwcel. In hoofdstuk 3 zijn deze grootheden door middel van het model met elkaar verbonden. Het model bleek een goede beschrijving te geven van metingen aan een zenuwcel in de bruine kikker.

In hoofdstukken 2 en 4 is de sensorische interpretatie van enkele activiteitspatronen bepaald. In hoofdstuk 2 is dit gedaan voor de opgewekte activiteit van enkele met behulp van een computer gesimuleerde zenuwcellen op drie verschillende geluiden: een in frequentie oplopende toon, een serie kliks en toonstootjes. Deze modelmatige zenuwcellen vormen een redelijke beschrijving voor zenuwcellen in de nervus acusticus. Hoewel de karakteristieke stukjes signaal bepaald waren met behulp van ruis, bleken de geschatte meest waarschijnlijke geluiden goed te lijken op de feitelijk aangeboden geluiden. Dit duidt op een zekere robuustheid van de procedure.

In hoofdstuk 4 zijn zes zenuwcellen in de torus semicircularis van de bruine kikker bestudeerd. Voor deze zenuwcellen geldt dat schattingen van het gemiddelde of het meest waarschijnlijke geluid dat het opgewekte activiteitspatroon

tot gevolg heeft, geen zinvol resultaat geven. Het is beter om naar de frequentieinhoud van het signaal als functie van de tijd te kijken. Dit kan met behulp van de Wigner CoSTID: een geavanceerde notenbalk, waarbij de fase-relaties tussen de verschillende tonale componenten behouden blijven. Op basis van ruis is de frequentiegevoeligheid van iedere zenuwcel bepaald. Deze frequentiegevoeligheid is daarna gebruikt om uit de gemeten reeksen actiepotentialen de bij de aangeboden paarroep behorende 'notenbalk' terug te schatten. Dit lukt heel aardig, echter een deel van de samenhang tussen de verschillende frequentiecomponenten in de paarroep gaat verloren.

Tenslotte wordt in hoofdstuk 5 een algemene oplossing gegeven. In principe kan, gegeven het model en algemene kennis over de omgeving(sgeluiden), bij ieder patroon van actiepotentialen de voorwaardelijke kansverdeling van de (geluids)prikkel gevonden worden. Uit de oplossing blijkt ook wanneer de bovengenoemde benadering geldig is. Tenminste dient aan de volgende voorwaarden voldaan te zijn:

1. er zijn geen verbindingen tussen zenuwcellen,

2. verschillende zenuwcellen dienen te reageren op onderling verschillende prikkeleigenschappen: verschillende frequenties,

3. de prikkeleigenschappen dienen onafhankelijk te zijn.

De uitgewerkte voorbeelden blijken achteraf gezien slechts aan voorwaarde 1) volledig te voldoen. Aan de beide andere voorwaarden is slechts gedeeltelijk voldaan. Omdat voorwaarden 2) en 3) niet meer kennis van het model vereisen, ligt een uitbreiding van de procedure in die richting binnen de mogelijkheden. De toevoeging van verbindingen tussen zenuwcellen stuit op ernstige practische en theoretische problemen. Zonder gebruik te maken van gegevens uit andere disciplines, als b.v. anatomie en neurofysiologie, lijkt schatten van de verbindingen niet mogelijk. Zonder extra beperkingen aan mogelijke zenuwmodellen zijn binnen de huidige opzet onrealistisch langdurige experimenten nodig. Ook de wiskunde biedt weinig soelaas. De introductie van koppelingen in een systeemidentificatieprocedure leidt in het algemeen tot recurrente relaties. Slechts in een enkel geval zijn die expliciet op te lossen. Dit betekent dit men vaak zijn toevlucht zal moeten nemen tot numerieke benaderingen met behulp van een computer. Kennis van de zenuwverbindingen is echter noodzakelijk voor het bepalen van de stimulus behorende bij een bepaald neuraal activiteitspatroon.

Door het vergelijken van de gepresenteerde prikkel met de geschatte prikkel is het mogelijk om de gevoeligheid en de selectiviteit van het zenuwstelsel te bepalen. De functie van het zenuwstelsel kan nu kwantitatief bestudeerd worden. Hierdoor draagt de sensorische interpretatie van neurale activeit bij aan het verkrijgen van een beter inzicht in het beeld dat het dier zich kan vormen omtrent zijn buitenwereld.

105

Curriculum vitae

Gerard Hesselmans werd op 6 juni 1959 geboren te Elshout (gemeente Drunen).

Zijn VWO opleiding begon hij in 1971 aan het Dr. Mollercollege in Waalwijk waar hij in 1978 het atheneum B diploma behaalde.

In 1978 begon hij aan zijn natuurkundestudie aan de Katholieke Universiteit te Nijmegen waar in 1981 het kandidaatsexamen met bijvak informatica werd afgelegd. Het doctoraalexamen in de experimentele natuurkunde, hoofdrichting biofysica, behaalde hij op 8 september 1983 cum laude. Het afstudeeronderzoek werd uitgevoerd op het Laboratorium voor Medische Fysica en Biofysica onder begeleiding van Dr. Peter Johannesma en Dr. Henk van den Boogaard, en betrof modelvorming van neurale wisselwerking binnen het wiskundige kader van puntprocessentheorie (titel doctoraalscriptie: "Correlation densities of conditional Poisson firing neurons in simple neural nets without connectivity loops").

Gedurende het laatste jaar van zijn natuurkunde studie volgde Gerard een cursus didactiek bij Dr. Hans Beltman. Na het behalen van het doctoraal werd deze cursus afgesloten met een stage op het Canisiuscollege Mater Dei in Nijmegen. In augustus 1984 behaalde hij de 1ste-graadslesbevoegdheid in natuurkunde.

Vanaf december 1983 tot december 1987 is hij als wetenschappelijk medewerker verbonden geweest aan de Faculteit der Wiskunde en Natuurwetenschappen van de Katholieke Universiteit te Nijmegen. In het kader van een promotieonderzoek bestudeerde hij op het Laboratorium voor Medische Fysica en Biofysica neurale activiteit met behulp van puntprocessen. In het bijzonder werd een sensorische interpretatie aan reeksen actiepotentialen gegeven.

Daarnaast was Gerard vanaf februari 1984 gedurende twee jaar in deeltijd werkzaam bij het Instituut voor Pathologische Anatomie (Prof. dr. G. P. Vooys) van de Faculteit der Geneeskunde en Tandheelkunde van de Katholieke Universiteit te Nijmegen. Daar bestudeerde hij de invloed van long- en baarmoederhalskanker op structuur- en vormparameters van cellen.

List of publications

1. H. F. P. van den Boogaard, G. H. F. M. Hesselmans and P. I. M. Johannesma, Point processes, correlation functions and system identification, (*abstract*) International Symposium on Mathematical Biology, November 10-15, 1985, Kyoto Japan.

2. P. S. Oud, J. A. M. Reubsaet-Veldhuizen, H. L. M. Beck, M. M. M. Pahlplatz, G. H. F. M. Hesselmans, H. G. Hermkens, J. Tas, J. James and G. P. Vooys, DNA and nuclear protein measurement in columnar epithelial cells of human endometrium, *Cytometry* 7, 325-330 (1986).

3. M. M. M. Pahlplatz, M. W. Katzko, G. H. F. M. Hesselmans, P. S. Oud and G. P. Vooys, Two methods for analyzing pleural smears for the presence of abnormalities, *Pattern Recognition Letters* 4, 405-411 (1986).

4. H. F. P. van den Boogaard, G. H. F. M. Hesselmans and P. I. M. Johannesma, System identification based on point processes and correlation densities. I. The nonrefractory neuron model, *Mathematical Biosciences* 80, 143-177 (1986).

5. G. H. F. M. Hesselmans, H. F. P. van den Boogaard and P. I. M. Johannesma, The characteristic functional of the peri-event stimulus ensemble, In: *Biofysica van Hersenen en Gedrag*, Nijmegen 1980-1985.
   K.U. Nijmegen 1986.

6. G. H. F. M. Hesselmans, H. F. P. van den Boogaard and P. I. M. Johannesma, The characteristic functional of the peri-event stimulus ensemble, *Mathematical Biosciences* 85:211-230 (1987).

7. C. C. A. M. Gielen, G. H. F. M. Hesselmans and P. I. M. Johannesma, Sensory interpretation of neural activity patterns. *Mathematical Biosciences* 87: (1987).

8. G. H. F. M. Hesselmans and P. I. M. Johannesma, Spectro-temporal interpretation of activity patterns of auditory neurons, *Mathematical Biosciences* (1987) (submitted).

9. G. H. F. M. Hesselmans, Distribution of sensory stimulus estimated from neural activity, *Mathematical Biosciences* (1987) (submitted).

10. G. H. F. M. Hesselmans, Maximum Entropy, Correlation Functions, and Exponential Models, *Mathematical Biosciences* (1988) (submitted).

Stellingen

behorende bij het proefschrift

SENSORY INTERPRETATION OF NEURAL EVENTS

1. In het model voor een zenuwcel bestaande uit een tweede-orde Volterra-systeem gevolgd door een exponentiële pulsgenerator leidt een Gaussisch stimulusensemble tot een eveneens Gaussisch peri-event-stimulusensemble (PESE). Deze klasse-invariantie gaat echter verloren zodra aanwezigheid van neurale activiteit als voorwaardelijke gebeurtenis in het PESE vervangen wordt door afwezigheid. De verdeling van de *a posteriori* stimulus geschat op basis van een rij actiepotentialen is derhalve niet-Gaussisch.
   Hoofdstuk 5

2. De beschrijving van een niet-Gaussisch symmetrisch stimulusensemble vereist minimaal nog een vierde- en een zesde-orde cumulant. Een op basis van tweede-orde eigenschappen berekende spectro-temporele beschrijving (CoSTID) is dan dus onvolledig.
   Hoofdstuk 3

3. De beschrijving van een puntproces door een rij van 'nullen' en 'enen' leidt tot conceptuele misverstanden. Binnen het computerjargon zijn 'nul' en 'een' namelijk uitwisselbaar en voegen ze evenveel informatie toe. Binnen de puntprocessen-theorie voegt in een infinitesimaal kort tijdsinterval slechts het optreden van een gebeurtenis informatie toe. De afwezigheid van een gebeurtenis bevat echter door zijn overweldigende waarschijnlijkheid geen informatie.
   P. Bremaud, *Point processes and queues*;
   martingale dynamics, Springer, New York, 1981.

4. Het toevoegen van extra (neurale) informatie leidt niet altijd tot een nauwkeuriger *a posteriori* schatting van de externe sensorische stimulus. Het is zelfs zo dat de onzekerheid, gegeven door de entropie van de schatting toe kan nemen.

5. Gezien het abstracte karakter van bepaalde passages uit dit proefschrift is een sensorische interpretatie van het bij de lezer dezes opgewekte neurale activiteitspatroon niet zinvol.

6. In deze tijd van automatisering moet het mogelijk zijn dat het controleren op spellingsfouten van drukproeven door promovendi achterwege kan blijven.

7. De bouw van eenpersoons-wooneenheden in afgelegen buitenwijken plaatst alleenstaanden in een isolement en past derhalve niet in een sociaal beleid.

8. Gezien het feit dat pathologische verschijnselen zich statistisch gezien slecht door normale verdelingen laten beschrijven, verdient het aanbeveling ruime aandacht te schenken aan niet-normale verdelingen in het statistiek-programma voor geneeskunde-studenten.

9. De wetenschappelijke inhoud van een promovendus wordt door zijn collegae afgemeten aan het aantal publicaties, de inhoud van zijn proefschrift door familie en vrienden aan de vormgeving.

10. Ter verkrijging van inzicht in neurale informatieverwerking is het *zelf* uitvoeren van (computer)simulaties van neurale netwerken door de wetenschappelijk onderzoeker een vereiste.

11. Bewondering begint te vaak daar waar begrip eindigt.

Nijmegen, april 1988   Gerard Hesselmans