Charles University in Prague
First Faculty of Medicine

PhD Thesis

Principles of information processing in neuronal models

Lubomír Košťál

Thesis advisor: Doc. RNDr. Petr Lánský, CSc.
Institute of Physiology,
Academy of Sciences of the Czech Republic

I am obliged to Doc. RNDr. Petr Lánský, CSc. for sharing with me his insight into the field of neural computation, for his support and friendly atmosphere during countless discussions that made this work possible.

The thesis is based on 5 published papers (enclosed), that represent the main part of my research performed in the Institue of Physiology, Academy of Sciences of the Czech Republic, Prague, between September 2004 and March 2007. The research was a part of the PhD studies of Biomedicine at the First Faculty of Medicine, Charles University in Prague, with specialization in Medical Biophysics. I declare, that the presented results are original, unless clearly stated otherwise.

Prague, June 5, 2007, Lubomír Košťál

Contents

1 Introduction

Neurons communicate via chemical and electrical synapses, in a process known as synaptic transmission. The crucial event that triggers synaptic transmission is the action potential (or spike), a pulse of electrical discharge that travels along the axon excitable membrane. The shapes and durations of individual spikes generated by a given neuron are very similar, therefore it is generally assumed that the form of the action potential is not important in information transmission. The series of action potentials in time (spike trains) can be recorded by placing an electrode close to or inside the soma or axon of a neuron. Since individual spikes in a spike train are usually well separated, the whole spike train can be described as a series of all-or-none point events in time (Gerstner & Kistler, 2002). The lengths of interspike intervals (ISIs) between two successive spikes in a spike train often vary, apparently randomly, both within and across trials (Gerstner & Kistler, 2002; Shadlen & Newsome, 1998; Stein *et al.*, 2005). In order to describe and analyze neuronal firing, statistical methods and methods of probability theory and stochastic point processes have been widely applied (Cox & Lewis, 1966; Kass *et al.*, 2005; Moore *et al.*, 1966; Tuckwell, 1988).

One of the most fundamental questions in neuroscience has been the problem of neuronal coding, i.e., the way information about stimuli is represented in spike trains (Perkel & Bullock, 1968; Softky, 1995; Strong *et al.*, 1998). To answer this question, methods to compare different spike trains are needed first (Bhumbra *et al.*, 2004; Buracas & Albright, 1999; Nemenman *et al.*, 2004; Paninski, 2003; Rieke *et al.*, 1997; Victor & Purpura, 1997).

1.1 Rate coding scheme

In the rate coding scheme information sent along the axon is encoded in the number of spikes per observation time window (the firing rate) (Adrian, 1928). In most sensory systems, the firing rate increases, generally non-linearly, with increasing stimulus intensity (Kandel *et al.*, 1991). Any information possibly encoded in the temporal structure of the spike train is ignored. Consequently, rate coding is inefficient but highly robust with respect to the ISI 'noise' (Stein *et al.*, 2005). The question whether the temporal structure of ISIs is due to unavoidable fluctuations in spike generation or whether it represents an

informative part of the neuronal signal is not yet fully resolved (Gerstner & Kistler, 2002; Shadlen & Newsome, 1994; Stein *et al.*, 2005) and leads to the idea of temporal coding.

1.2 Temporal coding scheme

Temporal codes employ those features of the spiking activity, that cannot be described by the firing rate. For example, time to first spike after the stimulus onset, characteristics based on the second and higher statistical moments of the ISI probability distribution, or precisely timed groups of spikes (temporal patterns) are candidates for temporal codes (Buracas & Albright, 1999; Gerstner & Kistler, 2002; Rieke *et al.*, 1997). Possibility of information transmission by changes in ISIs serial correlation has been reported in crayfish interneurons (Sugano & Tsukada, 1978; Wiersma & Adams, 1950). For a classic overview of temporal coding see Perkel & Bullock (1968), for a more recent discussion see Abeles (1994); Rieke *et al.* (1997); Shadlen & Newsome (1994); Stein *et al.* (2005); Theunissen & Miller (1995).

1.3 Motivation and aims of the thesis

Fast information transfer in neuronal systems rests on series of action potentials, the spike trains, conducted along axons. Methods that compare spike trains are crucial for characterizing different neuronal coding schemes. While the description of neuronal activity from the rate coding point of view is relatively straightforward, the temporal coding allows infinite number of possibilities. Spike trains which are equivalent from the rate coding perspective may turn out to be different under various measures of their temporal structure. The purpose of this text is to describe a measure of randomness of the neuronal activity. We discuss properties of this measure with respect to rate and temporal coding schemes and its application to experimental data. We show, that spiking randomness is capable to capture characteristics that would otherwise be difficult to obtain with conventional methods. The notion of randomness is very different from that of variability, even though these terms are sometimes interchanged. Furthermore, since the definition of randomness is based on the concept of entropy (Shannon & Weaver, 1998), relation with other information-theoretic quantities can be established.

2 Theory and methods

2.1 Probabilistic description of neuronal activity

Spike train consists of times of spike occurences $\tau_0, \tau_1, \ldots, \tau_n$. For the purpose of further analysis it is advantageous to describe such spike train equivalently by a set of n ISIs $t_i = \tau_i - \tau_{i-1}$, $i = 1 \ldots n$. Arguably the most important characteristics calculated from t_i is the estimate $\bar{t}$ of the mean ISI,

$$\bar{t} = \frac{1}{n} \sum_{i=1}^{n} t_i. \tag{1}$$

Since $\sum_{i=1}^{n} t_i = \tau_n - \tau_0$, the average $\bar{t}$ is computed without recourse to particular interval lengths and thus presents the lowest level of ISI analysis (Moore *et al.*, 1966). Other common parameters, coefficient of variation and standard deviation of ISIs, require all measurements, t_i, and both rely on the estimate s^2 of the ISI variance,

$$s^2 = \frac{1}{n-1} \sum_{i=1}^{n} (t_i - \bar{t})^2. \tag{2}$$

However, $\bar{t}$ and s^2 are meaningful only if the spiking activity is stationary, i.e., if the major probability characteristics of the firing are invariant in time (Cox & Lewis, 1966; Landolt & Correia, 1978). Stationary neuronal firing is typically observed in the spontaneous activity, or under constant stimulus conditions (Gerstner & Kistler, 2002; Moore *et al.*, 1966; Tuckwell, 1988).

The probabilistic description of the spiking results from the fact, that the positions of spikes cannot be predicted deterministically, only the probability that a spike occurs can be given (Gerstner & Kistler, 2002). By far the most common probabilistic descriptor is the ISI probability density function $f(t)$, where $f(t)\,dt$ is the probability that spike occurs in an interval $[t, t+dt)$ (Moore *et al.*, 1966). Probability density function is usually estimated from the data by means of histograms.

There are several functions completely equivalent to $f(t)$, that characterize the spiking activity (Cox & Lewis, 1966; Landolt & Correia, 1978). The

cumulative distribution function $F(t)$,

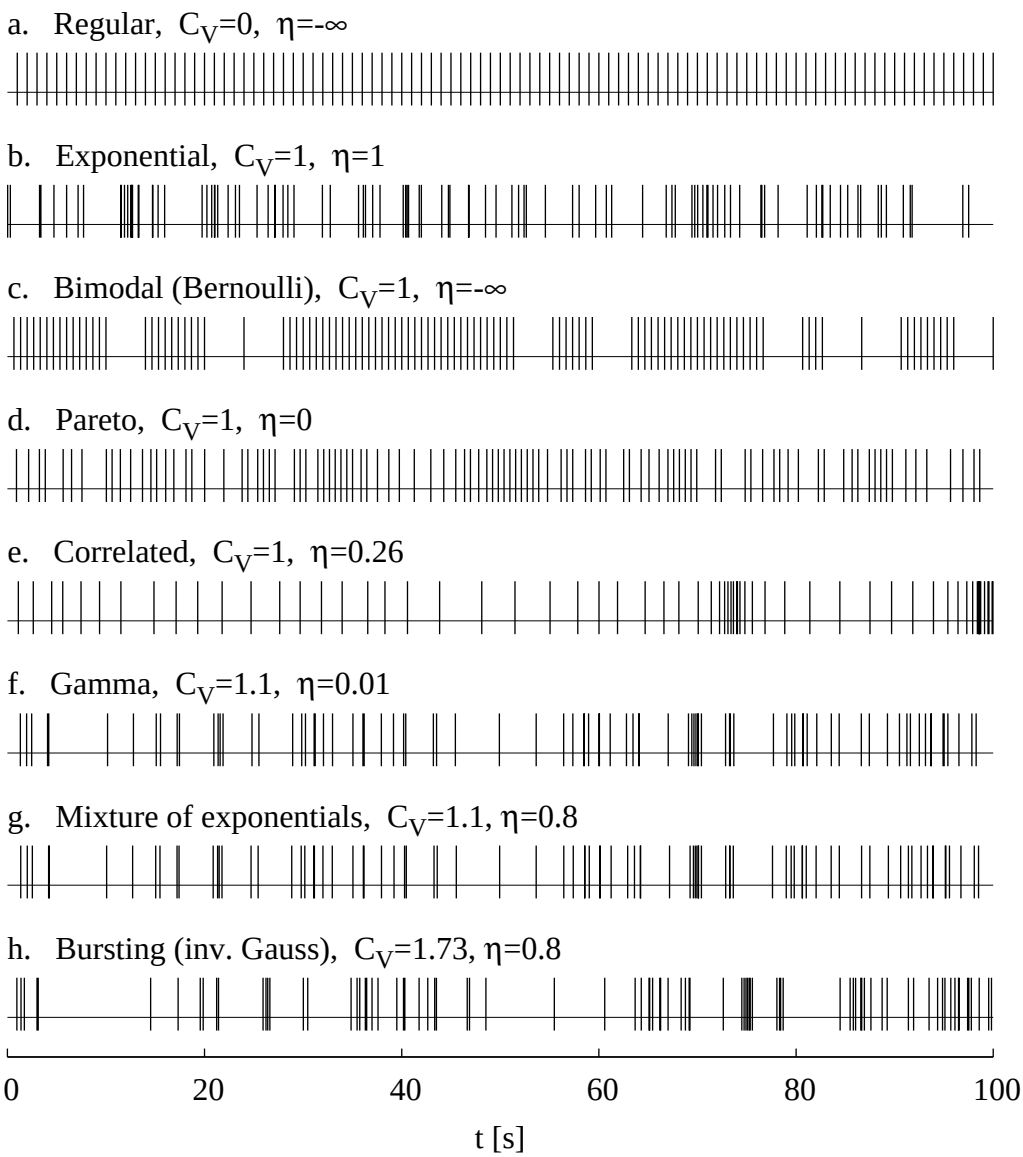
$$F(t) = \int_0^t f(z)\, dz, \tag{3}$$

gives the probability that the ISI will have a length not greater than t. $F(t)$ is easily estimated from data by means of empirical cumulative distribution function (Cox & Lewis, 1966; Duchamp-Viret *et al.*, 2005), which serves as a basis for some differential entropy estimators. The final probability descriptor we mention is the hazard rate $r(t)$,

$$r(t) = \frac{f(t)}{1 - F(t)}. \tag{4}$$

The hazard rate determines the probability $r(t)\, dt$ of spike occurring in a time interval $[t, t + dt)$ under the condition that there was no firing in $[0, t)$. The hazard rate characterizes the "imminency" of spiking (Tuckwell, 1988) and it has been traditionally employed in neuronal data analysis (Adrian *et al.*, 1964; Moore *et al.*, 1966; Poggio & Viernstein, 1964) to provide a different point of view from $f(t)$ and $F(t)$.

The mentioned decriptors, $f(t)$, $F(T)$ and $r(t)$, do not depend on the ordering of ISIs, i.e., they completely describe the firing when ISIs are mutually independent realizations of a positive random variable T, with mean ISI $E(T)$ and variance $Var(T)$ estimated by formulas (1) and (2). Such firing is called renewal process of ISIs (Cox & Lewis, 1966; Gerstner & Kistler, 2002). The plausibility of renewal models under steady-state stimulus conditions is supported by observation, that after a spike is emitted, the membrane potential of the cell returns to its (approximately) constant resting value (Gerstner & Kistler, 2002; Landolt & Correia, 1978; Stein, 1967; Tuckwell, 1988). Sometimes, however, there might be a dependency structure between the observed ISIs (Chacron *et al.*, 2001; Lansky & Rodriguez, 1999; Lindner, 2004; Longtin & Racicot, 1997; Ratnam & Nelson, 2000; Sakai *et al.*, 1999). The dependence may arise, for example, due to incomplete resetting of the membrane potential after the spike is emitted, which is experimentally observed especially in the distal parts of the neuron (Abeles, 1982). Such type of neuronal firing is not a renewal process, although the ISI probability distribution is invariant in time (due to the stationarity of spiking).



Caption on next page.

Consequently, the mean ISI is constant in time and therefore $E(T)$ carries all the information from the rate coding point of view, since $E(T)$ is inversely proportional to the (mean) firing rate (Gerstner & Kistler, 2002; Moore *et al.*, 1966). Basic observation reveals, however, that even if the firing rates are the same, the resulting spike trains can have very different appearances (Fig. 1). See Appendix for description of models employed in the figure.

Fig. 1: Examples of different simulated spike trains. Mean interspike interval is $E(T) = 1\,\mathrm{s}$ in all cases, thus the spike trains (a–h) are equivalent in the rate coding scheme. Temporal coding scheme must be used to classify the apparent differences. The interspike interval (ISI) variability (measured by coefficient of variation, C_V) is not sufficient for (b–e). The ISIs described by the exponential probability distribution (b) have many different lengths (i.e., they are 'variable'), but the same variability is achieved for a regular spiking disturbed by appropriately long pauses (c). Activity described by the Pareto distribution (d) has the same variability ($C_V = 1$), though it lacks short ISIs. Finally, (e) contains the same individual ISIs as (b) but ordered in a particular way (Markov chain, first-order serial correlation $\varrho = 0.9$). The measure that describes the differences is randomness, η, defined as the 'choice' of possible ISIs when reconstructing the spike train 'spike by spike'. Spike trains (f–h) were simulated using the same random seed to make the visual comparison easier. The differences in randomness between cases (f) and (g) with equal C_V are not as apparent as in (b–e). The same level of randomness, in (g) and (h), is reached with different variability and results in different spike trains. Bursting activity (h) is more variable than the exponential case (a), though its randomness is lower.

2.2 Spiking variability

One of the most frequently used characteristics of renewal neuronal firing is the ISI variability. The variability may be measured simply using the ISI variance, $Var(T)$, but variance depends on the mean ISI. Usually, it is required to characterize the spike train differences from the temporal coding point of view, in other words to describe properties which are distinct from the mean ISI. To achieve this the ISI lengths are rate-normalized, i.e., individual ISIs are divided by the mean ISI,

$$\Theta = \frac{T}{E(T)}, \tag{5}$$

so we obtain a new dimensionless random variable Θ with mean $E(\Theta) = 1$. Variance of Θ is equal to the coefficient of variation of the original random variable T, $Var(\Theta) = C_V$, where

$$C_V = \frac{\sqrt{Var(T)}}{E(T)}. \tag{6}$$

The main advantage of C_V as a measure of spiking variability (compared to variance) is that C_V is dimensionless and its value does not depend on the choice of units of ISIs (e.g., seconds or miliseconds) and thus ISI probability distributions with different means can be compared meaningfully (Softky & Koch, 1993). Furthermore, the C_V of ISIs is related to the variability coding hypothesis (Perkel & Bullock, 1968). The coding characterized by C_V has been proposed as a possible mechanism to transmit information about light intensity as well as adaptation state in the same spike train. A dark adapted cell has a larger C_V of ISIs than a light-adapted cell (for a given mean ISI) (Ratliff *et al.*, 1968). Changes in the level of bursting activity, characterized by values $C_V > 1$, are reported to be the proper code for edge detection in certain units of visual cortex (Burns & Pritchard, 1964) and also in hippocampal place cells (Fenton *et al.*, 2002). The variability of ISIs generated by the leaky integrate-and-fire model (Burkitt, 2006; Gerstner & Kistler, 2002) was recently a topic for a very extensive discussion initiated by Softky & Koch (1993).

Nevertheless, $E(T)$ and C_V are not sufficient to describe all possible differences between spike trains (Fig. 1b–e, where $E(T) = 1\,\mathrm{s}$ and $C_V = 1$). The spike trains described by the renewal processes of equal variability may have ISI probability distributions that differ in higher than second statistical moments. Additionaly, C_V does not account for statistical dependency between ISIs (by definition), and thus spike trains with the same marginal probability distributions of ISIs have the same variability.

Instead of employing characteristics based on higher statistical moments of the probability distributions involved and serial correlation coefficients of the ISIs we propose to measure the randomness of the spiking activity. Spiking randomness accounts automatically for differences in both marginal probability distributions and serial dependence of ISIs.

2.3 Spiking randomness

The randomness of spiking can be defined as the measure of 'choice' of different ISI lengths that appear in the spike train and the measure of 'freedom' in their serial ordering. Bigger choice of ISIs and more freedom in their ordering results, intuitively, in greater randomness of spiking. We first overview the concept of entropy (Shannon & Weaver, 1998), on which the measure of spiking randomness is based.

For a discrete random variable X with the set of possible states $\{x_1, x_2, \ldots, x_n\}$ and the corresponding probability mass function $p_i = \text{Prob}\{X = x_i\}$, the entropy $H(X)$ is defined as (Shannon & Weaver, 1998)

$$H(X) = -\sum_{i=1}^{n} p_i \ln p_i. \tag{7}$$

The entropy $H(X)$ is positive or equal to zero with equality if only one option is possible (no randomness). Maximum randomness (maximum $H(X)$) is reached when all p_i's are the same. If the logarithm base in formula (7) is 2, $H(X)$ can be interpreted as the average length in bits of the shortest description of X (Cover & Thomas, 1991). The entropy is a unique measure of randomness satisfying a set of intuitive conditions (Jaynes & Bretthorst, 2003; Shannon & Weaver, 1998), however, it is applicable to discrete systems only.

The extension of formula (7) for continuous probability distributions is impossible because the value of $H(X)$ diverges (Cover & Thomas, 1991). Therefore, the differential entropy $h(T)$ of the ISI probability density function $f(t)$ is defined as

$$h(T) = -\int_0^\infty f(t) \ln f(t)\, dt. \tag{8}$$

Differential entropy $h(T)$ does not have all the properties and intuitive interpretation of the entropy $H(X)$. The value of $h(t)$ changes with coordinate transforms, e.g., depends on the time units of ISIs. Probability density function $f(t)$ has a physical dimension (it is a derivative of probability with respect to time), therefore $h(T)$ has the dimension of its logarithm, e.g., logarithm of a milisecond. These facts show, that the differential entropy cannot be used to quantify the randomness of spiking activity. To overcome this problem, a discretization method has been adopted in literature (Rieke *et al.*, 1997; Strong

et al., 1998) which converts the task back to formula (7), however, the results depend on the discretization factor (Chacron *et al.*, 2001).

Here we proceed in a different way, avoiding the discretization. We want the randomness to characterize the spike train differences from the temporal coding point of view, in a similar way to C_V. Thus formula (5) is employed to rate-normalize the ISI lengths and the spiking randomness η is defined as the differential entropy of the random variable Θ. The following relation holds (Kostal *et al.*, 2007)

$$\eta = h(T) - \ln E(T). \tag{9}$$

Before discussing the properties of η we mention another approach to defining randomness. The exponential probability density function, $f_{\exp}(t)$, is given as

$$f_{\exp}(t) = \lambda \exp(-\lambda t), \tag{10}$$

where $\lambda > 0$ is the inverse of its mean, $\lambda = 1/E(T)$. An important property of $f_{\exp}(t)$ is, that it achieves maximum differential entropy among all ISI probability distributions with the same mean ISI (Cover & Thomas, 1991). The exponential model $f_{\exp}(t)$ represents the 'zero point' on the differential entropy scale for all ISI probability density functions with the same means. Kullback-Leibler (KL) distance $D(f, f_{exp})$ given by formula (Cover & Thomas, 1991)

$$D(f, f_{exp}) = \int_0^\infty f(t) \ln \frac{f(t)}{f_{\exp}(t)}\, dt \tag{11}$$

measures the deviation between probability density functions $f(t)$ and $f_{\exp}(t)$. Therefore, $D(f, f_{\exp})$ can be used to quantify the randomness of probability density function $f(t)$, if $f(t)$ has the same mean as $f_{\exp}(t)$ (Kostal & Lansky, 2006c). It can be shown that η is related to $D(f, f_{\exp})$ by a simple formula (Kostal *et al.*, 2007)

$$\eta = 1 - D(f, f_{\exp}) \tag{12}$$

and thus both proposed measures of randomness, η and $D(f, f_{\exp})$, are equivalent in their properties because their values differ only in a sign and a constant. Definition (9) can be naturally extended to account for non-renewal spiking activity (Kostal & Lansky, 2006a).

Finally, by employing the hazard rate from formula (4), we provide

an independent justification for maximum randomness of the exponential distribution. Intuitively, the most random firing makes the time to the first spike (since the last observed spike) most unpredictable. In other words, the probability of first spike occurring in $[t, t + dt)$ must be independent on the elapsed time t and consequently the hazard rate must be constant. (Any dependence of $r(t)$ on t would lead to increased predictability of firing due to more probable first spike occurrence after certain elapsed times.) The only ISI probability distribution with constant $r(t)$ is the exponential distribution defined in equation (10), then $r(t) = 1/\lambda$. Even though the randomness η can be determined from $r(t)$ (and not vice versa), we see two main advantages of η over $r(t)$ as a measure of randomness. First, $r(t)$ is a function not a number (contrary to η), and therefore comparison of randomness of different ISI distributions by means of $r(t)$ is difficult (Kostal & Lansky, 2006c). Second, while η in its general form accounts also for non-renewal spiking activity, $r(t)$ is used only in the renewal case (Moore *et al.*, 1966).

2.4 Randomness of non-renewal firing

In the stationary, but non-renewal spiking activity, the successive ISIs are realizations of identical, statistically dependent random variables $\{T_i\}$ and the activity is fully described by the joint probability density function $f(t_1, t_2, \dots)$ of ISIs. For the mean ISI, $E(T)$, thus holds $E(T) = E(T_i)$ (Cox & Lewis, 1966). The appropriate generalization of differential entropy $h(T)$ is given by the differential entropy rate $\bar{h}(f)$ (Cover & Thomas, 1991)

$$\bar{h}(f) = -\lim_{n\to\infty} \frac{1}{n} \int_0^\infty \cdots \int_0^\infty f(t_1, \dots, t_n) \ln f(t_1, \dots, t_n)\, dt_1 \dots dt_n. \tag{13}$$

Equation (13) provides the general form of differential entropy rate which can be significantly simplified for many cases of interest. For example, neuronal firing which is described by the first-order Markov chain (example in Fig. 1e) is fully characterized by the joint probability density function $f(t_1, t_2)$ of two adjacent ISIs (Cover & Thomas, 1991). Equation (13) then reads

$$\bar{h}(f) = -\int_0^\infty \int_0^\infty f(t_1, t_2) \ln f(t_2|t_1)\, dt_1\, dt_2, \tag{14}$$

where $f(t_2|t_1) = f(t_1, t_2)/f(t_1)$ is the conditional probability density function (Cox & Lewis, 1966).

The joint probability density function $f(t_1, t_2, \dots)$ describes also a general, non-stationary neuronal activity. By observing the first n spikes (from the stimulus onset) over many trials we may estimate the n-dimensional probability density function $f(t_1, \dots, t_n)$ that governs the immediate response of a neuron (or a population of neurons). The definition (13) holds without the limit (for a finite n) and we may formally put $\mu = \frac{1}{n}\sum_{i=1}^{n} E(T_i)$ instead of $E(T)$. Randomness of any activity may be calculated according to formula (9), however, due to non-stationarity the interpretation of η with respect to the rate coding scheme becomes unclear.

2.5 Properties of spiking randomness

Here we summarize basic properties of the spiking randomness η, and compare it with the properties of variability as measured by the coefficient of variation C_V.

- Due to rate-normalization of the ISI probability distribution, the randomness η is a dimensionless quantity and does not depend on coordinate transformations (Kostal *et al.*, 2007). Consequently η allows to compare different stationary spiking activities in the same way as C_V.

- Maximum spiking randomness is generated only by the renewal process with exponential probability distribution of ISIs (Poisson process, Fig. 1a). Substituting formula (10) into formula (9) gives $\eta = 1$. Any non-renewal spiking activity with exponential marginal probability distribution of ISIs must have $\eta < 1$, since less freedom in serial ordering of ISIs results in smaller randomness (Kostal & Lansky, 2006a).

- Coincidentally, both $\eta = 1$ and $C_V = 1$ for exponential distribution. Many non-exponential probability distributions can have $C_V = 1$, but their randomness is always $\eta < 1$. The equality $\eta = 1$ completely characterizes the exponential distribution of ISIs.

- Equally variable spike trains may differ in their randomness. However, the same spiking randomness may be achieved with different spiking variabilities (Fig. 1g and h). Thus, randomness provides an alternative

rather than superior characteristic of neuronal firing compared to variability (Kostal *et al.*, 2007).

- C_V is limited from below by $C_V = 0$ (regular spiking, Fig. 1a) but there is no maximum spiking variability. Values $C_V > 1$ are characteristic of bursting activity (Fig. 1h). On the other hand, there is no unique minimal randomness probability distribution, because $\eta = -\infty$ for any discrete random variable (Fig. 1a and c). However, discrete probability distributions are not valid models of spiking activity (ISI 'noise' is always present), and the fact that η may not be finite is of little practical consequence (Kostal & Lansky, 2006a).

- Spiking randomness is an information-theoretic measure, related to entropy and KL distance. The strength of information-theoretic measures lies in their ability to reveal non-linear dependencies (Cover & Thomas, 1991; Rieke *et al.*, 1997; Yamada *et al.*, 1993). Recently, KL distance has been used in the field of neuronal coding from the classification theory point of view Johnson *et al.* (2001) and as a predictor of purely rate coding models (Johnson & Glantz, 2004). Renormalized entropy (a special case of KL distance) has been shown to provide additional information over traditional tools in EEG record analysis (Kopitzki *et al.*, 1998; Quiroga *et al.*, 2000; Thakor & Tong, 2004).

2.6 Estimation of spiking randomness from data

The definition of randomness in formula (9) depends on the differential entropy. The problem of differential entropy from data estimation is well exploited in literature, see, e.g., Beirlant *et al.* (1997); Tsybakov & van der Meulen (1994) for an overview of available techniques. It is preferable to avoid estimations based on data binning (histograms), because discretization affects the results greatly. The support of ISI distributions is always positive, which makes the application of kernel estimators problematic.

If the neuronal firing is described by the renewal process our experience shows, that the simple and well researched Vasicek's estimator (Vasicek, 1976) gives good results on a wide range of data (Ebrahimi *et al.*, 1992; Esteban *et al.*, 2001; Miller & Fisher III, 2003). The Vasicek's estimator is based on the empirical cumulative distribution function. Given the n ranked ISIs

$\{t_{[1]} < t_{[2]} < \cdots < t_{[n]}\}$ the Vasicek's estimator $\hat{h}$ of differential entropy reads

$$\hat{h} = \frac{1}{n}\sum_{i=1}^{n} \ln\left[\frac{n}{2m}\left(t_{[i+m]} - t_{[i-m]}\right)\right] + \varphi_{\text{bias}}. \tag{15}$$

The positive integer parameter $m < n/2$ is set prior to computation and the two following conditions hold: $t_{[i-m]} = t_{[1]}$ for $(i-m) < 1$ and $x_{[i+m]} = x_{[n]}$ for $(i+m) > n$. The particular values of m corresponding to various values of n were determined by Ebrahimi *et al.* (1992). The bias-correcting factor is

$$\varphi_{\text{bias}} = \ln\frac{2m}{n} - \left(1 - \frac{2m}{n}\right)\Psi(2m) + \Psi(n+1) - \frac{2}{n}\sum_{i=1}^{m}\Psi(i+m-1), \tag{16}$$

where $\Psi(z) = \frac{d}{dz}\ln\Gamma(z)$ is the digamma function (Abramowitz & Stegun, 1965). Our experience with simulated data shows, that for sample sizes $n \approx 500$ the error of estimation is relatively small (Kostal & Lansky, 2006b), the positive bias with respect to true values is not important for small samples (Ebrahimi *et al.*, 1992; Esteban *et al.*, 2001) and the value of m may be approximated by an integer closest to $\sqrt{n}$. The disadvantage of Vasicek's estimator is, that it cannot be easily extended to non-renewal processes.

Non-renewal sustained neuronal activity is described by multi-dimensional joint probability distributions and so more elaborate techniques have to be employed in differential entropy estimation. One popular approach (Kraskov *et al.*, 2004; Victor, 2002) is realized by the Kozachenko-Leonenko binless estimator (Kozachenko & Leonenko, 1987), which is asymptotically unbiased and consistent, but the dimension of the problem must be known beforehand, and the underlying probability density function must be continuous. If the spiking is described by a d-dimensional probability density function, each vector $(t_j, t_{j+1}, \ldots, t_{j-1+d})$ of consequent ISIs represents a point in a d-dimensional space. If the observed spike train consists of N ISIs then total $n = N - d + 1$ of such points may be obtained (if the firing is stationary). The estimate $\hat{h}$ then reads

$$\hat{h} = \frac{d}{n}\sum_{i=1}^{n}\ln\lambda_i + \ln\left[\frac{(n-1)\sqrt{\pi^d}}{\Gamma(d/2+1)}\right] + \gamma, \tag{17}$$

where λ_i is the Euclidean distance of the i-th point to its nearest neighbour,

$\gamma = -\int_0^\infty e^{-z} \ln z \, dz \approx 0.5772$ is the Euler-Mascheroni constant and $\Gamma(z)$ is the gamma function (Abramowitz & Stegun, 1965). It must be stated, however, that 'reasonable' estimation of differential entropy of non-renewal spiking activity usually requires large amounts of data, often not available in experimental recordings.

3 Summary of main results

3.1 Model spiking activity

Probabilistic models of stationary spiking activity may be divided in two categories: statistical and biophysical. The statistical models are described by probability density functions which are simple enough to manipulate and adequately describe experimentally observed data but no other connection with neurophysiological reality is required. The biophysical models, on the other hand, result from attempts to describe the behavior of real neurons at different levels of abstraction (Gerstner & Kistler, 2002; Tuckwell, 1988). However, mathematical expressions for biophysical models are rarely given in a closed form and one has to rely on numerical approximations.

The analysis of several statistical and biophysical renewal process models was performed in Kostal & Lansky (2006b,c); Kostal *et al.* (2007). Here we present an overview of the main results. The statistical models are represented by gamma and log-normal ISI probability distributions, both are commonly used for experimental data description (Duchamp-Viret *et al.*, 2005; Levine, 1991; Mandl, 1992; McKeegan, 2002; Rospars *et al.*, 1994). The inverse Gaussian distribution (Chhikara & Folks, 1989) results from a simple point stochastic neuronal model (perfect integrator) where the approach of the membrane potential towards the threshold is described by the Wiener process with a positive drift (Berger *et al.*, 1990; Levine, 1991). The inclusion of leakage current into this model results in the more realistic diffusion leaky IF model (Burkitt, 2006; Tuckwell, 1988) where the membrane potential evolution is described by the Ornstein-Uhlenbeck process. The parameters of the leaky IF model determine two firing regimes, depending on the ratio $\xi = S/(\mu\tau)$ of the threshold membrane potential S to the neuronal input μ and membrane time constant τ (Burkitt, 2006; Kostal *et al.*, 2007). The sub-threshold regime is characterized by $\xi > 1$, supra-threshold by $\xi < 1$. The gamma, lognormal

and inverse Gaussian distributions are completely characterized by $E(T)$ and C_V and therefore it is possible to calculate unique value of η for each value of C_V (note that η is independent of $E(T)$). Similar calculation is possible for the leaky IF model once μ, τ and S are known (determining the supra- or sub-threshold regimes) because the amount of input 'noise' σ^2 controls the actual value of C_V (Kostal *et al.*, 2007). The following inference can be made based on Fig. 2, where the randomness of each model is plotted against the corresponding variability.

- The randomness-variability curves of the investigated models are often U-shaped with high randomness values distributed around $C_V = 1$. The notable exceptions are the sub-threshold regime of the leaky IF model and the statistical Pareto model (not included, Kostal & Lansky (2006a)).

- While small variability generally implies low randomness, high variability in the firing may result in both low as well as high randomness.

- It is well known that the lognormal, inverse Gaussian and leaky IF supra-threshold ISI distributions never become exponential, but in addition their maximum randomness (minimal KL distance from the exponential model) is not located at $C_V = 1$.

- The behavior of the leaky IF model in the supra-threshold regime is comparable to the perfect integrator (inverse Gaussian model). In the sub-threshold regime, the effect of coherence resonance (Lindner *et al.*, 2002) is demonstrated by local decrease of C_V for $\eta \approx 0.96$. However, there is no corresponding local decrease in spiking randomness, i.e., the coherence resonance for certain (high) values of ξ is observable in C_V but not in η. Though it is known that the degree of coherence resonance depends on the measure employed (Lindner *et al.*, 2004), the apparent disappearance of the effect on the η scale raises the question of what is the proper measure of ISI coherence (Kostal *et al.*, 2007).

Several statistical models of non-renewal spiking activity described by first-order Markov chains (Cox & Lewis, 1966) were examined in Kostal & Lansky (2006a). Markov structure in experimental data is reported in literature (Ratnam & Nelson, 2000), and even the first-order case makes the existence

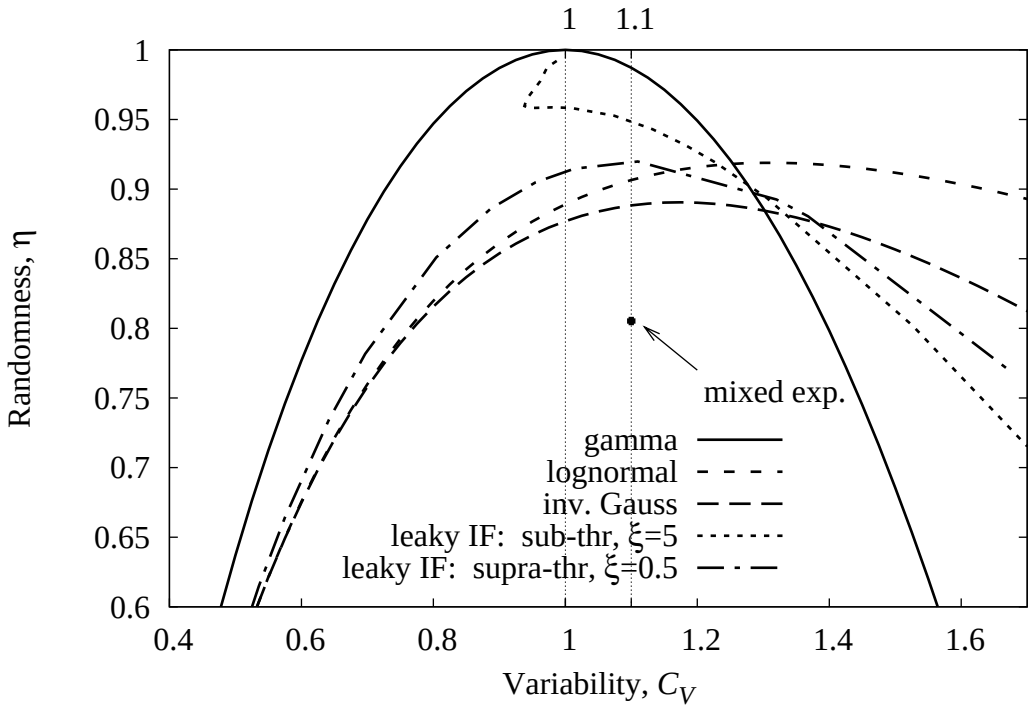


Fig. 2: Randomness vs. variability for some widely used renewal models of neuronal activity. For $C_V = 1$ the gamma distribution becomes exponential and thus $\eta = 1$. The inverse Gaussian and lognormal models never become exponential their maximum randomness (minimal KL distance from the exponential model with the same mean ISI) is not located at $C_V = 1$. The behavior of leaky IF model in the supra-threshold case is similar to the behavior of inverse Gaussian and lognormal models. The sub-threshold activity exhibits the effect of coherence resonance (local decrease of C_V for $\eta \approx 0.96$). The dot shows, that for $C_V = 1.1$ the randomness of the exponential mixture model is lower than that of any other model considered here.

of certain (short) sequences of ISIs more probable than others (basic temporal pattern formation, Kostal & Lansky (2006a)). Although the examined models were not used in data analysis previously, the results show that the serial correlation coefficient (Cox & Lewis, 1966) is a weak indicator of the true ISI serial dependence. For example, the Lawrance and Lewis model with exponential marginal probability distribution of ISIs (Lawrance & Lewis, 1977) can achieve either randomness $\eta = 0.97$ or $\eta = 0.82$ for the same value of first-

order serial correlation $\varrho = 0.17$. It follows, that randomness η or mutual information between ISIs (Cover & Thomas, 1991) should be employed when deciding on renewal or non-renewal character of experimental data. However, estimation of these information-theoretic quantities requires large amounts of data which are usually not available in experimental recordings. Finally we note, that the maximum order of non-zero serial correlation coefficient does not coincide with the dimension of the joint probability density function describing the activity. For example, first-order moving average process is non-renewal with all second- and higher-order serial correlations equal to zero. However, joint probability density function of two adjacent ISIs does not describe such process, since the Markov property, $\mathrm{Prob}\{T_n \leq t_n | T_{n-1} = t_{n-1}, \ldots, T_1 = t_1\} = \mathrm{Prob}\{T_n \leq t_n | T_{n-1} = t_{n-1}\}$ (Cox & Lewis, 1966), does not hold in this case.

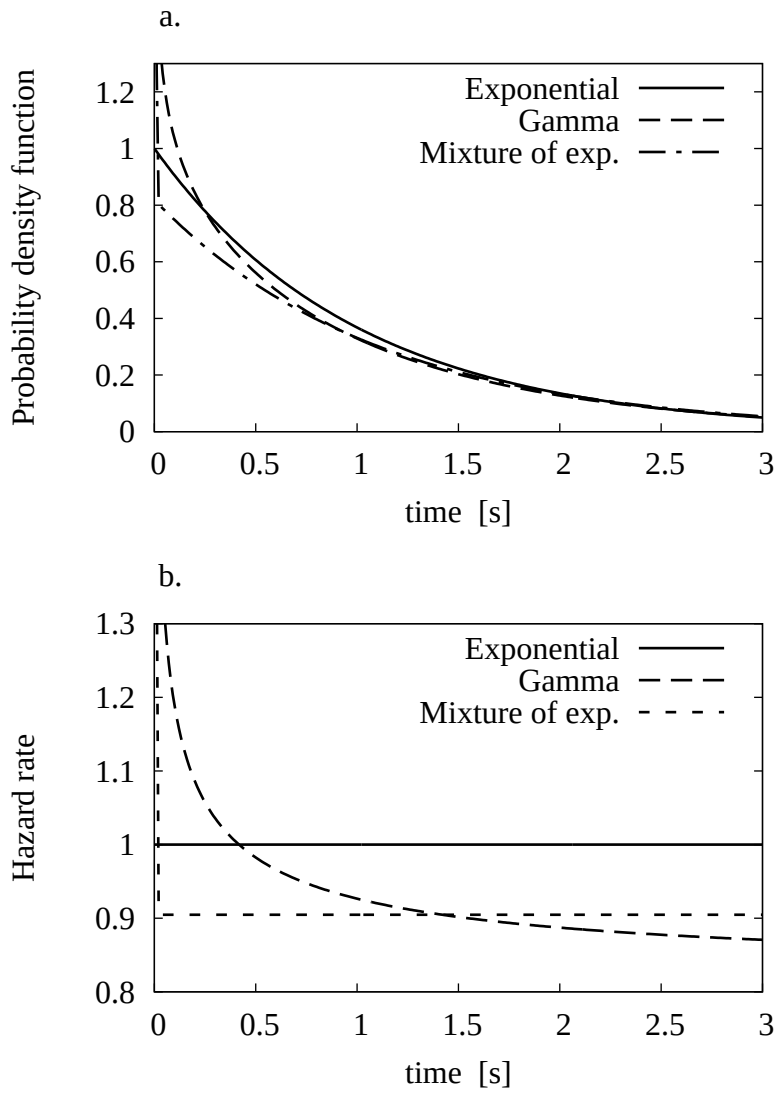
3.2 Simulated and experimental data

Here we provide an example with practical consequences, showing that estimates of randomness from two spike trains may differ significantly although their appearance is very similar (Fig. 1f and g) and their histograms are almost identical.

Bursting neuronal activity consists of runs of short ISIs (bursts) interspersed among comparatively longer ISIs. Bursting is usually characterized by $C_V > 1$ and it is often reported in experimental data analysis. (Bhumbra *et al.*, 2004; Rospars *et al.*, 1994; Duchamp-Viret *et al.*, 2005). The bursting activity of neuron is usually described by a mixture of two distributions, one for interburst ISIs and the other for intraburst ISIs. A common model of bursting activity is given by a probability density function of the mixture of two exponential (ME) distributions (Smith & Smith, 1965; Tuckwell, 1988)

$$f(t) = pae^{-ax} + (1-p)be^{-bx}, \tag{18}$$

where $p \in (0,1)$ and $a > 0$, $b > 0$, $a \neq b$. The parameters a, b and p are independent and consequently a whole range of different randomness values can be achieved for a fixed mean ISI and $C_V > 1$ (Kostal & Lansky, 2006b). We compare two simulated spike trains with $E(T) = 1\,\mathrm{s}$, $C_V = 1.1$: the first generated by the gamma model (Fig. 1f) and the second generated by the ME distribution (Fig. 1g). The theoretical value of η for the gamma model



Caption on next page.

in this case is $\eta = 0.99$. The parameters in formula (18) were set so, that $\eta = 0.80$ for the ME model. Fig. 2 shows, that the randomness of the ME distribution with variability $C_V = 1.1$ is the lowest of the considered models with the same C_V. The histograms of ISIs constructed from $n = 200$ spikes are, however, hardly distinguishable due to the striking similarity of spike trains in Fig. 1f and g. The estimated C_V values are (mean $\pm$ standard deviation): $\hat{C_V} = 1.1 \pm 0.06$ (gamma) and $\hat{C_V} = 1.104 \pm 0.05$ (ME). The estimates of randomness, $\hat{\eta}$ according to formula (15) (see Appendix B) with $\varphi_{\text{bias}} = 0$ and $m = 14$, averaged over several runs give $\hat{\eta} = 0.91 \pm 0.05$ (gamma) and $\hat{\eta} = 0.77 \pm 0.06$ (ME). The error of estimation is acceptable, because even

Fig. 3: Comparison of the mixture of exponentials (ME) and gamma models with parameters $E(T) = 1\,\mathrm{s}$ and $C_V = 1$. Exponential model with $E(T) = 1\,\mathrm{s}$ is also shown. Probability density functions (a) of ME and gamma models are almost identical for times greater than 0.5 s. However, the calculated spiking randomness η differs significantly (Fig. 2), which is supported independently by comparing the hazard rates of ME and gamma models. The gamma model approaches the constant hazard rate quickly and therefore its spiking activity is more random than that described by the ME model with monotonously decreasing hazard rate.

for 200 spikes the Vasicek's estimator clearly marks the difference in spiking randomness. Theoretical probability density functions of the exponential, gamma and ME models (Fig. 3a) differ for very short ISIs, however, histograms with wide-enough bins hide this difference. Visual comparison of hazard rates, however, provides an independent proof that the spiking randomness of the ME model is indeed different from both exponential and gamma model (Fig. 3b). The hazard rate of the gamma model rapidly approaches the constant value and thus confirms the small deviation in randomness from the exponential distribution, contrary to the ME distribution which has monotonously decreasing hazard rate. We conclude, that even though conventional analysis of two spike trains reveals no difference, the spike trains may still differ in their randomness and the difference is tractable even with limited amount of data. Values $\eta = 0.99$ and $\eta = 0.80$ may also characterize visually different spike trains. This is confirmed by comparing Fig. 1f and h with $\eta = 0.8$ but different levels of variability. The bursting activity described by the inverse Gaussian model contains more of both longer and shorter ISIs than the less variable gamma model.

Duchamp-Viret *et al.* (2005) estimated the randomness from the spontaneous activity recordings done on olfactory receptor neurons of freely breathing and tracheotomized rats. The recordings were obtained under steady-state conditions and it was shown that in the majority of cases the firing can be considered renewal. It was demonstrated, that the spontaneous activity is less variable but more random in the case of tracheotomized animals

than in those freely breathing. This effect is further enhanced if η is adjusted for different spiking rates by considering instead the ratio $\eta/E(T)$, i.e., the 'amount' of randomness per time unit (Kostal & Lansky, 2006a).

4 Concluding remarks

Comparison of neuronal spiking activity under different conditions plays a key role in resolving the question of neuronal coding. The spiking activity of a neuron is usually not deterministic, therefore ISI is described by means of probability distributions. We proposed an information-theoretic measure of spiking randomness, η, which can be related to the differential entropy or KL distance from the maximum entropy distribution. Conceptually, the spiking randomness can be best compared with the more often used ISI variability measured by the coefficient of variation, C_V. However, the properties of randomness and variability are different. Namely, small variability generally implies low randomness, but high variability of firing may not result in high level of randomness. Simultaneously, the same level of randomness can be reached by different values of variability, depending on the probabilistic model of the spiking.

Appendix A: Simulated spike trains

In this section we describe models of neuronal activity that were used to create Fig. 1 and that are not discussed in the previous text. We employed the standard transformation method (Devroye, 1986) for generating the ISIs from known probability density functions.

- ad c. The spike train contains ISIs of two possible lengths, θ_1 and θ_2, distributed according to the Bernoulli distribution

$$\mathrm{Prob}(T = \theta_1) = 1 - \mathrm{Prob}(T = \theta_2) = p, \tag{A-1}$$

  where $p \in [0, 1]$. It holds

$$E(T) = p\theta_1 + (1-p)\theta_2, \tag{A-2}$$

$$C_V = \frac{\sqrt{(1-p)p}|\theta_1 - \theta_2|}{(\theta_1 - \theta_2)p + \theta_2}. \tag{A-3}$$

  From $E(T) = 1\,\mathrm{s}$, $C_V = 1$ and by choosing $p = 1/10$ follows $\theta_1 = 4\,\mathrm{s}$ and $\theta_2 = 2/3\,\mathrm{s}$, which was used to generate the shown spike train.

- ad d. The probability density function of the Pareto distribution is

$$f(t) = \begin{cases} 0, & t \in (0, b) \\ ab^a t^{-a-1}, & t \in [b, \infty) \end{cases} \tag{A-4}$$

  with parameters $a > 2$ and $b > 0$. The following relations hold: $C_V = 1/\sqrt{(a^2 - 2a)}$ and $E(T) = ab/(a-1)$.

- ad e. The first-order Markov chain was generated by the Downton bivariate exponential model (Downton, 1970)

$$f(t_1, t_2) = \frac{a^2}{1-\varrho} \exp\left[\frac{a(t_1+t_2)}{\varrho - 1}\right] I_0\left(\frac{2a\sqrt{t_1 t_2 \varrho}}{1-\varrho}\right), \tag{A-5}$$

  where $\varrho \in (0, 1)$ is the first-order serial correlation, $a = 1/E(T)$ and $I_\nu(z)$ is the modified Bessel function of the first kind (Abramowitz & Stegun, 1965).

- ad f. Probability density function of the gamma distribution,

parameterized by $\mu = E(T)$ and C_V is

$$f(t) = \left(\frac{1}{C_V^2 \mu}\right)^{1/C_V^2} \Gamma(1/C_V^2)\, t^{1/C_V^2 - 1} \exp\left(-\frac{t}{C_V^2 \mu}\right), \tag{A-6}$$

where $\Gamma(z)$ is the gamma function.

- ad h. Probability density function of the inverse Gaussian distribution (Chhikara & Folks, 1989), parameterized by $\mu = E(T)$ and C_V is

$$f(t) = \sqrt{\frac{\mu}{2\pi C_V^2 t^3}} \exp\left[-\frac{1}{2C_V^2 \mu}\frac{(t-\mu)^2}{t}\right]. \tag{A-7}$$

References

Abeles, M. (1982) *Local cortical circuits: Studies of brain function*, vol. 6. Springer-Verlag, Berlin.

Abeles, M. (1994) Firing rates and well-timed events in the cerebral cortex. In Domany, E., Schulten, K. & van Hemmen, J. L. (eds.), *Models of Neural Networks II*. Springer, New York, pp. 121–138.

Abramowitz, M. & Stegun, I. (1965) *Handbook of Mathematical Functions, With Formulas, Graphs, and Mathematical Tables*. Dover Publications, New York.

Adrian, E. (1928) *The basis of sensation*. W. W. Norton & Co., New York.

Adrian, H., Goldberg, J. & Smith, F. (1964) Response of neurons of the superior olivary complex of the cat to acoustic stimuli of long duration. *J Neurophysiol*, **27**, 706–749.

Beirlant, J., Dudewicz, E., Gyorfi, L. & van der Meulen, E. (1997) Nonparametric entropy estimation: An overview. *Int J Math Stat Sci*, **6**, 17–39.

Berger, D., Pribram, K., Wild, H. & Bridges, C. (1990) An analysis of neural spike-train distributions: determinants of the response of visual cortex neurons to changes in orientation and spatial frequency. *Exp Brain Res*, **80**, 129–134.

Bhumbra, G., Inyushkin, A. & Dyball, R. (2004) Assessment of spike activity in the supraoptic nucleus. *J Neuroendocrinol*, **16**, 390–397.

Buracas, G. & Albright, T. (1999) Gauging sensory representations in the brain. *Trends Neurosci*, **22**, 303–309.

Burkitt, A. (2006) A review of the integrate-and-fire neuron model: I. homogeneous synaptic input. *Biol Cybern*, **95**, 1–19.

Burns, B. & Pritchard, R. (1964) Contrast discrimination by neurons in the cat's visual cerebral cortex. *J Physiol*, **175**, 445–463.

Chacron, M., Longtin, A. & Maler, L. (2001) Negative interspike interval correlations increase the neuronal capacity for encoding time-dependent stimuli. *J Neurosci*, **21**, 5328–5343.

Chhikara, R. & Folks, J. (1989) *The Inverse Gaussian Distribution: theory, methodology, and applications*. Marcel Dekker.

Cover, T. & Thomas, J. (1991) *Elements of information theory*. Wiley, New York.

Cox, D. & Lewis, P. (1966) *The statistical analysis of series of events*. Chapman and Hall, London.

Devroye, L. (1986) *Non-uniform random variate generation*. Springer-Verlag, New York.

Downton, F. (1970) Bivariate exponential distributions in reliability theory. *J Roy Stat Soc B*, **32**, 408–417.

Duchamp-Viret, P., Kostal, L., Chaput, M., Lansky, P. & Rospars, J. (2005) Patterns of spontaneous activity in single rat olfactory receptor neurons are different in normally breathing and tracheotomized animals. *J Neurobiol*, **65**, 97–114.

Ebrahimi, N., Habibullah, M. & Soofi, E. (1992) Testing exponentiality based on kullback-leibler information. *J Roy Stat Soc B*, **54**, 739–748.

Esteban, M., Castellanos, M., Morales, D. & Vajda, I. (2001) Monte carlo comparison of four normality tests using different entropy estimates. *Communications in Statistics: Simulation and Computation*, **30**, 761–785.

Fenton, A., Lansky, P. & Olypher, A. (2002) Properties of the extra-positional signal in hippocampal place cell discharge derived from the overdispersion in location-specific firing. *Neuroscience*, **111**, 553–566.

Gerstner, W. & Kistler, W. (2002) *Spiking Neuron Models: Single Neurons, Populations, Plasticity*. Cambridge University Press, Cambridge.

Jaynes, E. & Bretthorst, G. (2003) *Probability Theory: the logic of science*. Cambridge University Press, Cambridge.

Johnson, D. & Glantz, R. (2004) When does interval coding occur? *Neurocomputing*, **59**, 13–18.

Johnson, D., Gruner, C., Baggerly, K. & Seshagiri, C. (2001) Information-theoretic analysis of neural coding. *J Comput Neurosci*, **10**, 47–69.

Kandel, E., Schwartz, J. *et al.* (1991) *Principles of neural science*. Elsevier, New York.

Kass, R. E., Ventura, V. & Brown, E. N. (2005) Statistical issues in the analysis of neuronal data. *J Neurophysiol*, **94**, 8–25.

Kopitzki, K., Warnke, P. & Timmer, J. (1998) Quantitative analysis by renormalized entropy of invasive electroencephalograph recordings in focal epilepsy. *Phys Rev E*, **58**, 4859–4864.

Kostal, L. & Lansky, P. (2006a) Classification of stationary neuronal activity according to its information rate. *Network: Comp Neur Sys*, **17**, 193–210.

Kostal, L. & Lansky, P. (2006b) Similarity of interspike interval distributions and information gain in a stationary neuronal firing. *Biol Cybern*, **94**, 157–167.

Kostal, L. & Lansky, P. (2006c) Variability and randomness in stationary neuronal activity. *Biosystems*. doi:10.1016/j.biosystems.2006.05.010.

Kostal, L., Lansky, P. & Zucca, C. (2007) Randomness and variability of the neuronal activity described by the ornstein-uhlenbeck model. *Network: Comp Neur Sys*, in print.

Kozachenko, L. & Leonenko, N. (1987) Sample estimate of the entropy of a random vector. *Prob Inf Trans*, **23**, 95–101.

Kraskov, A., Stögbauer, H. & Grassberger, P. (2004) Estimating mutual information. *Phys Rev E*, **69**, 66138–16.

Landolt, J. & Correia, M. (1978) Neuromathematical concepts of point process theory. *IEEE Trans Biomed Eng*, **25**, 1–12.

Lansky, P. & Rodriguez, R. (1999) Two-compartment stochastic model of a neuron. *Physica D*, **132**, 267–286.

Lawrance, A. & Lewis, P. (1977) An exponential moving-average sequence and point process (ema1). *J App Prob*, **14**, 98–113.

Levine, M. (1991) The distribution of the intervals between neural impulses in the maintained discharges of retinal ganglion cells. *Biol Cybern*, **65**, 459–467.

Lindner, B. (2004) Interspike interval statistics of neurons driven by colored noise. *Phys Rev E*, **69**, 22901.

Lindner, B., Garcia-Ojalvo, J., Neiman, A. & Schimansky-Geier, L. (2004) Effect of noise in excitable systems. *Phys. Rep.*, **392**, 321–424.

Lindner, B., Schimansky-Geier, L. & Longtin, A. (2002) Maximizing spike train coherence or incoherence in the leaky integrate-and-fire model. *Phys Rev E*, **66**, 31916.

Longtin, A. & Racicot, D. (1997) Assessment of linear and nonlinear correlations between neural firing events. In Cutler, C. & Kaplan, D. (eds.), *Nonlinear Dynamics and Time Series: Building a Bridge between the Natural and Statistical Sciences*. Fields Institute Communications, pp. 223–239.

Mandl, G. (1992) Coding for stimulus velocity by temporal patterning of spike discharges in visual cells of cat superior colliculus. *Vision Res*, **33**, 1451–1475.

McKeegan, D. (2002) Spontaneous and odour evoked activity in single avian olfactory bulb neurones. *Brain Res*, **929**, 48–58.

Miller, E. & Fisher III, J. (2003) Ica using spacings estimates of entropy. *JMLR*, **4**, 1271–1295.

Moore, G., Perkel, D. & Segundo, J. (1966) Statistical analysis and functional interpretation of neuronal spike data. *Annu Rev Physiol*, **28**, 493–522.

Nemenman, I., Bialek, W. & de Ruyter van Steveninck, R. (2004) Entropy and information in neural spike trains: Progress on the sampling problem. *Phys Rev E*, **69**, 056111.

Paninski, L. (2003) Estimation of entropy and mutual information. *Neural Comput*, **15**, 1191–1253.

Perkel, D. & Bullock, T. (1968) Neural coding. *Neurosci Res Prog Sum*, **3**, 405–527.

Poggio, G. & Viernstein, L. (1964) Time series analysis of impulse sequences of thalamic somatic sensory neurons. *J Neurophysiol*, **27**, 517–545.

Quiroga, R., Arnhold, J., Lehnertz, K. & Grassberger, P. (2000) Kulback-leibler and renormalized entropies: Applications to electroencephalograms of epilepsy patients. *Phys Rev E*, **62**, 8380–8386.

Ratliff, F., Hartline, H. & Lange, D. (1968) Variability of interspike intervals in optic nerve fibers of limulus: Effect of light and dark adaptation. *Proceedings of the National Academy of Sciences*, **60**, 464–469.

Ratnam, R. & Nelson, M. (2000) Nonrenewal statistics of electrosensory afferent spike trains: Implications for the detection of weak sensory signals. *J Neurosci*, **20**, 6672–6683.

Rieke, F., Steveninck, R., Warland, D. & Bialek, W. (1997) *Spikes: Exploring the Neural Code*. MIT Press, Cambridge.

Rospars, J., Lansky, P., Vaillant, J., Duchamp-Viret, P. & Duchamp, A. (1994) Spontaneous activity of first-and second-order neurons in the frog olfactory system. *Brain Res*, **662**, 31–44.

Sakai, Y., Funahashi, S. & Shinomoto, S. (1999) Temporally correlated inputs to leaky integrate-and-fire models can reproduce spiking statistics of cortical neurons. *Neural Networks*, **12**, 1181–1190.

Shadlen, M. & Newsome, W. (1994) Noise, neural codes and cortical organization. *Curr Opin Neurobiol*, **4**, 569–579.

Shadlen, M. & Newsome, W. (1998) The variable discharge of cortical neurons: Implications for connectivity, computation, and information coding. *J Neurosci*, **18**, 3870–3896.

Shannon, C. & Weaver, W. (1998) *The Mathematical Theory of Communication*. University of Illinois Press, Illinois.

Smith, D. & Smith, G. (1965) A statistical analysis of the continual activity of single cortical neurones in the cat unanaesthetized isolated forebrain. *Biophys J*, **5**, 47–74.

Softky, W. (1995) Simple codes versus efficient codes. *Curr Opin Neurobiol*, **5**, 239–247.

Softky, W. & Koch, C. (1993) The highly irregular firing of cortical cells is inconsistent with temporal integration of random epsps. *J Neurosci*, **13**, 334–350.

Stein, R. (1967) Some models of neuronal variability. *Biophys. J*, **7**, 37–68.

Stein, R., Gossen, E. & Jones, K. (2005) Neuronal variability: noise or part of the signal? *Nat. Rev. Neurosci.*, **6**, 389–397.

Strong, S., Koberle, R., de Ruyter van Steveninck, R. & Bialek, W. (1998) Entropy and information in neural spike trains. *Phys. Rev. Lett.*, **80**, 197–200.

Sugano, N. & Tsukada, M. (1978) Effect of correlated adjacent interspike interval sequences of the excitatory motor axon on the opening movement of the crayfish claw opener muscles. *Biol Cybern*, **29**, 63–67.

Thakor, N. & Tong, S. (2004) Advances in quantitative electroencephalogram analysis methods. *Annu. Rev. Biomed. Eng.*, **6**, 453–495.

Theunissen, F. & Miller, J. (1995) Temporal encoding in nervous systems: A rigorous definition. *J Comput Neurosci*, **2**, 149–162.

Tsybakov, A. & van der Meulen, E. (1994) Root-n consistent estimators of entropy for densities with unbounded support. *Scand. J. Statist*, **23**, 75–83.

Tuckwell, H. (1988) *Introduction to Theoretical Neurobiology.* Cambridge University Press, New York.

Vasicek, O. (1976) A test for normality based on sample entropy. *J Roy Stat Soc B*, **38**, 54–59.

Victor, J. (2002) Binless strategies for estimation of information from neural data. *Phys Rev E*, **66**, 51903–51918.

Victor, J. & Purpura, K. (1997) Metric-space analysis of spike trains: theory, algorithms and application. *Network: Comp Neur Sys*, **8**, 127–164.

Wiersma, C. & Adams, R. (1950) The influence of nerve impulse sequence on the contractions of different crustacean muscles. *Physiol Comp*, **2**, 20–33.

Yamada, S., Nakashima, M., Matsumoto, K. & Shiono, S. (1993) Information theoretic analysis of action potential trains. *Biol Cybern*, **68**, 215–220.

List of publications

[I] Duchamp-Viret, P., Kostal, L., Chaput, M., Lansky, P. & Rospars, J. (2005) Patterns of spontaneous activity in single rat olfactory receptor neurons are different in normally breathing and tracheotomized animals. *J Neurobiol*, **65**, 97–114. (IF: 4.170)

[II] Kostal, L. & Lansky, P. (2006a) Classification of stationary neuronal activity according to its information rate. *Network: Comp Neur Sys*, **17**, 193–210. (IF: 2.055)

[III] Kostal, L. & Lansky, P. (2006b) Similarity of interspike interval distributions and information gain in a stationary neuronal firing. *Biol Cybern*, **94**, 157–167. (IF: 1.398)

[IV] Kostal, L. & Lansky, P. (2006c) Variability and randomness in stationary neuronal activity. *Biosystems*, in print (IF: 1.144)

[V] Kostal, L., Lansky, P. & Rospars, J.-P. (2006d) Encoding of pheromone intensity by dynamic activation of pheromone receptors. *Neurocomputing*, in print (IF: 0.790)

[VI] Kostal, L., Lansky, P. & Zucca, C. (2007) Randomness and variability of the neuronal activity described by the ornstein-uhlenbeck model. *Network: Comp Neur Sys*, in print. (IF: 2.055)

Active participation in meetings with abstracts

[i] Kostal L, Lansky P (2004) *The comparison of model and experimental ISI distributions using Kullback-Leibler distance*, (poster), Computational Systems Biology of the Neuronal Cell, SISSA-ICTP, Trieste, Italy

[ii] Kostal L, Lansky P (2005) *A method for analyzing stationary neuronal activity from the information-theoretic point of view*, (poster), ECMTB 2005 Dresden, Germany

[iii] Kostal L (2005) Lansky P, *Classification of stationary neuronal activity according to its information rate*, (contributed talk), Neural Coding 2005, Marburg, Germany

[iv] Kostal L, Lansky P (2005) *Variability and randomness in stationary neuronal activity*, (poster), Fifth Conference of the Czech Neuroscience Society, Prague, Czech Republic

[v] Kostal L, Lansky P (2006) *Comparison of variability and randomness in stationary neuronal firing*, (contributed talk), Information Theory, Neurobiology and Cognition, MPI MIS Leipzig, Germany

[vi] Kostal L, Lansky P, Rospars J-P (2006) *Pheromone plumes provide optimal signal for the olfactory sensory neuron*, (poster), CNS 2006, Edinburgh, UK

[vii] Kostal L, Lansky P (2006) *Characterization of stationary neuronal activity by its varability and randomness*, (poster), Prague Stochastics 2006, Prague, Czech Republic

[viii] Kostal L, Lansky P (2007) *Variability and randomness in stationary neuronal activity*, (contributed talk), Stochastic analysis and its applications III, MFF UK, Prague, Czech Republic

Reprints of published papers

Research papers [I–IV] and [VI] listed on p. 33 represent the main part of the thesis, to which the previous text provides a summary and, in some cases, extension. The copies of the papers are attached below.

Patterns of Spontaneous Activity in Single Rat Olfactory Receptor Neurons Are Different in Normally Breathing and Tracheotomized Animals

Patricia Duchamp-Viret,[1] Lubomir Kostal,[2] Michel Chaput,[1] Petr Lánsky,[2] Jean-Pierre Rospars[3]

[1] Neurosciences & Systèmes Sensoriels, Université Claude Bernard Lyon 1, CNRS UMR 5020, 50 avenue Tony Garnier, 69366 Lyon Cedex 07, France

[2] Institute of Physiology, Academy of Sciences, Videnska 1083, 14220 Prague 4, Czech Republic

[3] Physiologie de l'Insecte: Signalisation et Communication, UMR 1272 UPMC – INRA – INA-PG, INRA 78026 Versailles Cedex, France

Received 18 March 2005; accepted 18 March 2005

ABSTRACT: **Spontaneous firing of olfactory receptor neurons (ORNs) was recently shown to be required for the survival of ORNs and the maintenance of their appropriate synaptic connections with mitral cells in the olfactory bulb. ORN spontaneous activity has never been described or characterized quantitatively in mammals. To do so we have made extracellular single unit recordings from ORNs of freely breathing (FB) and tracheotomized (TT) rats. We show that the firing behavior of TT neurons was relatively simple: they tended to fire spikes at the same average frequency according to purely random (Poisson) or simple (Gamma or Weibull) statistical laws. A minority of them were bursting with relatively infrequent and short bursts. The activity of FB neurons was less simple: their firing rates were more diverse, some of them showed trends or were driven by breathing. Although more of them were regular, only a minority could be described by simple laws; the majority displayed random bursts with more spikes than the bursts of TT neurons. In both categories bursts and isolated spikes (outside bursts) occurred completely at random. The spontaneous activity of ORNs in rats resembles that of frogs, but is higher, which may be due to a difference in body temperature. These results suggest that, in addition to the intrinsic thermal noise, spontaneous activity is provoked in part by mechanical, thermal, or chemical (odorant molecules) effects of air movements due to respiration, this extrinsic part being naturally larger in FB neurons. It is suggested that spontaneous activity may be modulated by respiration. Because natural sampling of odors is synchronized with breathing, such modulation may prepare and keep olfactory bulb circuits tuned to process odor stimuli.** © 2005 Wiley Periodicals, Inc. J Neurobiol 65: 97–114, 2005

Keywords: **olfactory receptor unit; spontaneous firing; rat; single unit; free breathing; statistical analysis**

INTRODUCTION

Since the pioneering studies of Hubel and Wiesel in the 1970s (Hubel and Wiesel, 1970), the spontaneous activity present in several areas of the central nervous system has been shown to play an essential role in the refinement of neural connections, especially during development (Shatz and Stryker, 1988; Erzurumlu and Kind, 2001; McLaughin et al., 2003; Grubb et al.,

Correspondence to: J.-P. Rospars (rospars@versailles.inra.fr).

Contract grant sponsor: Academy of Sciences of the Czech Republic; contract grant number: 1ET400110401.

Contract grant sponsor: Barrande; contract grant number: 9146QL.

Published online 19 August 2005 in Wiley InterScience (www.interscience.wiley.com).
DOI 10.1002/neu.20177

2003; Hanson and Landsmesser, 2004). In the olfactory system, the organization of the connections between the primary olfactory receptor neurons (ORNs), located in the olfactory mucosa, and the second order neurons (i.e, mitral cells, MC), located in the olfactory bulb (OB), is very precise (Mombaerts et al., 1996; Ressler et al., 1993; Vassar et al., 1993, 1994) and is maintained despite the continual renewal of ORNs during the life span. ORNs that express the same odorant receptor are broadly dispersed across the olfactory epithelium, but their axons converge onto one or a few bulbar glomeruli (Mombaerts et al., 1996; Ressler et al., 1993; Vassar et al., 1993). ORN projection is governed by the receptor expression itself (Wang et al., 1998; Cutforth et al., 2003, Zou et al., 2004) and creates an OB topographic map of receptor activation that is one of the bases of the neural code allowing odor discrimination. ORN activity was not thought to be involved in the development and maintenance of the ORN connectivity map until recent reports indicated such a role of ORN spontaneous (Yu et al., 2004) and odor-evoked (Zou et al., 2004) activities.

Consistent with other observations (Lin et al., 2000; Zheng et al., 2000), Yu et al. (2004) showed that the development of the map was apparently normal if all ORNs were inactivated. However, a selective suppression of spontaneous activity or synaptic transmission in the ORNs that express the P2 odorant receptor resulted in the dissolution of the corresponding glomeruli and the gradual depletion of the inactive P2 ORNs. These results indicate a bidirectional influence of spontaneous activity in the ORNs, which would ensure both the ORN survival and the synaptic stability in glomeruli. Furthermore, Zou et al. (2004) showed that, once established, odor-evoked activity could refine the sensory map. They also proposed that, in a competitive environment, silent neurons, providing neither spike emission nor synaptic transmission, could be switched to transcribe a different receptor gene. Such a selective pressure exerted by surrounding active neurons raises the hypothesis that neural activity may regulate not only neurotransmitter expression (Borodinsky et al., 2004) but also the expression of olfactory receptors.

In frogs (Blank and Mozell, 1976; Duchamp et al., 1974; Duchamp-Viret et al., 2000; Gesteland, 1971; Gesteland et al., 1963; 1965; Juge et al., 1979; O'Connell and Mozell, 1969), salamanders (Bayling, 1979; Trotier and MacLeod, 1983), tortoises (Matthews, 1972), birds (McKeegan, 2002), and mice (Sicard, 1986) ORNs have been shown to fire action potentials more or less regularly at a frequency rate varying from a few spikes per minute to a few spikes per second in absence of odor stimulation. In most species only the average spontaneous firing rate or the histogram of firing frequencies is known. We showed in frogs (Rospars et al., 1994), that the spontaneous activity of ORNs can be described by a completely random (Poisson) process. The basic random event can be either the firing of an individual action potential or the occurrence of a burst of action potentials, that is series of spikes with interspike intervals (ISIs) much shorter than the interburst intervals. Spontaneous activity of frog ORNs was furthermore shown to influence the activity of second-order neurons because a partial deafferentation of the OB from the mucosa increased their spontaneous firing rate. This effect was not in accordance with the simple expectation that decreasing the excitatory inputs of mitral cells should decrease their firing rate. It raised the hypothesis that the partial deafferentation of the OB from the mucosa resulted in a relative decrease of the number of excitatory inputs on inhibitory GABAergic neurons in the glomerular layer. This hypothesis was supported by experiments (Duchamp-Viret et al., 1993) where the pharmacological antagonization of the inhibitory GABAergic network induced an increase in MC spontaneous rate, suggesting that ORN spontaneous activity influences the OB network stability.

The recent evidence that spontaneous spiking is involved in both the shaping of neuronal circuits in glomeruli and the ORN maintenance makes more salient the lack of description and characterization of this activity in mammalian ORNs. A statistical and modeling study of the spontaneous activity of rat single ORNs was thus undertaken from recordings gathered during studies of ORN reactivity to odors in anesthetized rats (Duchamp-Viret et al., 1999, 2000, 2003). ORNs were recorded in freely breathing and in tracheotomized animals, which offered an opportunity to examine the role of respiration rhythm in the genesis and the patterning of ORN spontaneous activity. In the present article the following questions are addressed: What are the qualitative and quantitative characteristics of the spontaneous activity in normally breathing rats? How is it modified by tracheotomy? How does it compare with that in other species (frogs, fishes, and birds)? What is the origin of the spontaneous activity in ORNs? Can it be ascribed to intrinsic (thermal noise) or extrinsic (odorant molecules) factors?

METHODS

Surgical Methods

All experiments were carried out in accordance with the European Communities Council Directive of November 24th, 1986 (86/609/EEC) for the care and use of laboratory

animals and all efforts were made to minimize animal suffering and to reduce the number of animals used. Adult Wistar rats (250–300 g) were used. Rats were anesthetized by an intraperitoneal injection of Equithesine (mixture of pentobarbital sodium and chloral hydrate) at the initial dose of 3 mL/kg. Anesthetic was then supplemented as necessary to maintain a deep level of anesthesia, as determined by the depth and rate of the respiratory rhythm of the rat and the lack of withdrawal of the leg in response to a moderately intense toe pinch. Rectal temperature was maintained at 37 ± 0.5°C by a homeothermic blanket (Harvard Apparatus, USA). The incision on the scalp was regularly infiltrated with 2% procaine. For recordings, anesthetized animals were left freely - breathing or they were tracheotomized. Recordings were performed in the Endoturbinate II. Access to the olfactory mucosa was gained by removing the nasal bones and then gently slipping aside the dorsal recess underlying these bones; no procaine was applied to this area.

Electrophysiological Recordings and Data Analysis

Single-unit action potentials were recorded using glass micropipettes filled with an alloy of Wood's metal (80%) and indium (20%) (Gesteland et al., 1959). Their impedances ranged from 2 to 4 MΩ. The signal was recorded on line through a conventional amplifier (band pass 300–3000 Hz) to a Data Tape Recorder (Biologic, France) and a CED-1401 data acquisition system (Cambridge Electronic Design Ltd., United Kingdom) connected to a computer. Spike activity was sampled at 15 kHz on the CED-1401 data acquisition system. The single unit nature of each recording was first examined during the experiment by triggering the recorded spikes near the background-noise on a storage oscilloscope. This allowed us to evaluate the characteristics of the polyphasic spike of each neuron studied in order to ensure that only one neuron was recorded throughout all the experimental procedure. Off line, recordings were further analyzed using the facilities offered by the Spike2 language associated to the CED-1401 system. Spikes were detected using a waveform template and then visually inspected. The single-unit nature of each recording was checked for the constancy of the total amplitude of the spike, the constancy of the ratio between the positive and negative phases, and the absence of action potentials within the absolute refractory period. All recordings that might have involved more than one neuron were discarded. [The number of records retained in each population were 19 (FB) and 15 (TT).] For each record t_i denotes the time of the ith spike occurrence and the ith 1SI is defined as $x_i = t_{i+1} - t_i$.

In parallel to single recordings, the mucosa field potential or electro-olfactogram (EOG) was recorded by using a large glass micropipette broken to about 50 μm OD and filled with saline solution (NaCl 9% + agar-agar or gelatin 5%). The signal was amplified and filtered between DC and 30 Hz. In absence of odor stimulation, this signal oscillated in phase with the respiration and was utilized to monitor the respiratory rhythm.

Tests of Stationarity

A spike train is nonstationary if the ISIs present a trend, that is, a progressive lengthening or shortening of ISIs within a record. Nonstationarity was systematically checked because it may affect the reliability of all quantitative characteristics derived from ISIs. For studying stationarity the linear regression of ISIs, x_i, on their serial numbers, i, was computed. The slope of the regression line was tested against zero ($p < 0.05$). Nonparametric tests based on median runs (Wald-Wolfowitz test) were also performed ($p < 0.05$) (Siegel, 1956), large negative values of the test statistics indicating the presence of a trend. Records for which at least one of the tests yielded a significant value were regarded as nonstationary.

Periodicity

To discover possible periodicities in the spike trains we employed a spectral density estimation. Spike trains, in which each spike is a point event, were transformed into waveforms in which each spike was approximated by a Gaussian curve instead of a delta function. So, from the times of occurrences $\{t_1 < t_2 < \ldots < t_n\}$ of n spikes we constructed a new "spike train" function $g(t)$:

$$g(t) = \sum_{i=1}^{n} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(t-t_i)^2}{2\sigma^2}\right) \qquad (1)$$

The width σ of the Gaussian curves was chosen as one-fourth the shortest ISI in the record in order to avoid overlapping curves. Fourier transform of a Gaussian curve is easy to compute, so the spectral power density $P_g(f)$ of $g(t)$ can be written as

$$P_g(f) = \frac{1}{n^2}\left[e^{-2\pi^2 f^2 \sigma^2}\right]^2 \left|\sum_{i=1}^{n} e^{-i2\pi f t_i}\right|^2 \qquad (2)$$

The choice of the norm coefficient $1/n^2$ is arbitrary because only the location of the peak of the periodogram, not its height, is of interest. This method presents the advantage of avoiding numerical computation of Fourier transform.

The presence of periodicities was also indicated by the Wald-Wolfovitz test used for testing the stationarity of spike trains. In contrast with trend detection, significantly large positive values of the test statistics suggest that long and short ISIs alternate with some regularity.

Tests of Independence between Successive Intervals

Serial dependence or independence of ISIs is an important criterion to characterize spontaneous activity. Dependence means that spike triggering is influenced by the firing of previous spikes (no wiping out of the "memory" of the neuron or of its input) and that the length of the current ISI is

related to the previous ones. Coefficients of serial correlation have been the most widely used measure of the serial dependency of ISIs (Cox and Lewis, 1966). For example, a value of the first-order serial coefficient r_1 close to 1 indicates that long ISIs are often followed by long ones and short ISIs by short ones. On the contrary, a value of r_1 close to –1 means that long ISIs are often followed by short ones and vice versa (negative correlation).

The significance of only the first-order coefficients was determined. To test whether the serial coefficient r_1 for a given record was significant, we used the asymptotical method based on the property that $r_1\sqrt{n-1}$, where n is the number of ISIs, follows the standard normal law if ISIs are not correlated.

Intensity and Variability of Activity

If a trend is detected, the mean ISI changes in time and estimating it by a constant is therefore inappropriate. For this reason all the records with a trend in mean ISI were shortened to find a section of the record that can be considered as stationary. The firing intensity of a stationary record was characterized by the sample mean ISI, $\bar{x}$, and the sample median, x_M, which are both inversely related to the firing frequency. The variability of the activity was expressed by the sample standard deviation s of ISIs, which is also in seconds, and by the sample coefficient of variation $CV = s/\bar{x}$, which is dimensionless. In addition to these parametric measures we used for the same purpose their nonparametric equivalents, the interquantile range *IQR*, which is the difference (in s) between the first quartile (the ISI value for which 25% of ISIs of a given neuron are shorter) and the third quartile (75% of ISIs are shorter) and the dimensionless quantity $CV_M = IQR/x_M$.

Comparison and Testing of Neuron-Wise Statistical Series

For visualizing a series of m statistics y_i (the statistics being $\bar{x}$, s, CV, etc., and m the number of neurons), each coming from a different neuron in a given category (FB or TT), the y_i values were ranked and a stairstep graph was drawn with identical steps equal to $1/m$ added at each y_i (in abscissa). Unlike the histogram, this empirical cumulative distribution function (or cumulated histogram) is unique, does not depend on the arbitrary choice of the bin width, and is much more efficient than histograms for detecting small differences between samples of small sizes. Comparison of these graphs from two samples can reveal slight differences in location (i.e., different central values) and in slope (i.e., different variabilities). Difference in location was tested using the Wilcoxon test, which is based on the sum of ranks. Difference in slope was tested with the Fisher test for comparing variances. In all cases the 5% significance level was selected. This procedure was used for all statistics based on a single value per neuron to compare the two categories FB and TT.

Probability Density Functions

The simplest distribution for describing ISIs is the exponential one, which is unimodal with a maximum at the origin. Its density is $f(x) = \alpha\exp(-\alpha x)$, where $x \geq 0$ is the duration of an ISI and $\alpha > 0$ is a parameter (firing rate). Here both the mean and standard deviation of ISIs are equal to $1/\alpha$, so that CV is always 1 and does not depend on α. Its cumulative distribution function is $F(x) = 1 - \exp(-\alpha x)$. It can be calculated with $F(x) = 0.25$, 0.5, and 0.75 for the first quartile, median, and third quartile, respectively, that $CV_M = 1.58$ is also independent of α. The exponential distribution characterizes a Poisson process, that is, a process in which events (action potentials) occur at any moment independently of the past firing pattern. In a Poissonian neuron, no mechanism is present that would give a structure to the successive occurrences of action potentials and thus it can be called purely random. When describing neuronal activity, such a process can be considered only as an approximation, because the mechanisms that normally trigger action potentials are expected to deviate from these assumptions, as evidenced by refractoriness and afterhyperpolarization, for example. This approximation can be expected to work only for slowly firing neurons when the instantaneous triggering of an action potential after a previous one can be neglected.

The Gamma distribution is another suitable description of ISIs. Its density is $f(x) = \lambda^k x^{k-1}\exp(-\lambda x)/\Gamma(k)$, where $\Gamma(k)$ is the gamma function (generalization of the factorial to real numbers) and $\lambda > 0$ related to mean $E(X)$ and $k > 0$ related to shape are parameters, with $E(X) = k/\lambda$. This distribution is characterized by a linear relationship between its mean and standard deviation, $\sigma = \sqrt{k}/\lambda$, and thus $\sigma = E(X)/\sqrt{k}$. For $k = 1$, the Gamma distribution becomes the exponential distribution and when k is large, it tends to a normal distribution with increasing mean and variance. When $k < 1$, it is more skewed than the exponential distribution, being infinity at $x = 0$. For integer $k > 1$, the Gamma distribution can result from various mechanisms. For example, a Poisson process of rate λ in which all events are suppressed except every kth event results in a Gamma distribution of interevent times.

The Weibull distribution of ISI was also utilized. It is not so commonly tested on experimental data as Gamma distribution, but McKeegan (2002) reported its agreement with data recorded from ORNs in hens. This distribution is characterized by the family of hazard functions $h(t) = t^{-\kappa}\lambda$, and associated cumulative distribution functions are

$$G(t) = 1 - \exp\left(-\frac{\lambda}{1-\kappa}t^{1-\kappa}\right), \quad \kappa < 1 \qquad (3)$$

Weibull and Gamma distributions can have the same shape as the exponential (with maximum at the origin) or have a single peak remote from the origin.

We fitted our data to these three distributions. Parameters α, λ, k and κ were estimated and then the goodness of fit was tested at the significance level $p = 0.05$ by the one-sample Kolmogorov-Smirnov test (Siegel, 1956), which is based on the empirical cumulative distribution function.

Entropy

The entropy concept was introduced into statistical information theory by Shannon (1948), and we employed it as an alternative measure for deviations of a given spike train from the Poisson process. In contrast to statistical testing, it can be used for ordering the trains with respect to their distance to the template (Poisson process). The differential entropy I of a continuous distribution $f(t)$ on $t \in [0, \infty)$ is given as

$$I(f(t)) = -\int_0^\infty f(t) \ln f(t) dt \tag{4}$$

Of all possible probability distributions on $[0, \infty)$ with fixed mean the exponential one has the maximum entropy

$$I_{\text{exp}} = 1 - \ln \alpha \tag{5}$$

where $1/\alpha = E(X)$ is the mean value of the ISIs. This is intuitively well understandable because the Poisson process is known as completely random and any deviation from such a process suggests some systematic influence on the spike generating system. Thus, after estimating the entropy I_{data} of the ISIs, the difference

$$D = 1 + \ln \bar{x} - I_{\text{data}} \tag{6}$$

is always non-negative ($D \geq 0$) and it measures the deviation of the ISI distribution with respect to the exponential one, that is the deviation of the spike train from the Poisson process of the same mean. D is called Kullback-Leibler distance. The quantity D, in contrast to CV, characterizes a Poisson process. The reason for this difference is that a necessary and sufficient condition for having a Poisson process is that $D = 0$, whereas the condition $CV = 1$ is not sufficient (all Poisson processes have $CV = 1$ but $CV = 1$ does not imply a Poisson process). In compensation, for calculating the CV it is enough to know the mean and the standard deviation but the complete distribution is needed to determine D.

The estimation of the entropy of a continuous distribution is a nontrivial problem. Given the n ranked ISIs $\{x_{[1]} < x_{[2]} < \ldots < x_{[n]}\}$ we used the entropy estimator proposed by Vasicek (1976)

$$I_{\text{data}} = \frac{1}{n}\sum_{i=1}^{n} \ln\left\{\frac{n}{2m}\left(x_{[i+m]} - x_{[i-m]}\right)\right\} \tag{7}$$

The positive integer parameter $m < n/2$ is set before computation and the two following conditions also hold: $x_{[i-m]} = x_{[1]}$ for $(i - m) < 1$ and $x_{[i+m]} = x_{[n]}$ for $(i + m) > n$. The particular values of m we used were determined by Ebrahimi et al. (1991).

Bursting

Neurons with an ISI distribution not fitting an exponential law were further investigated using the assumption that only the shortest ISIs were deviating from exponentiality. To find whether an ISI distribution had an exponential tail, we applied the following procedure: ISIs longer than a certain (small) value x_{cut} were tested against exponentiality. This procedure was tried sequentially with increasing x_{cut} by step of 1 ms starting from $x_{\text{cut}} = 0$, until the distribution of ISIs longer than x_{cut} was found to be an exponential. This procedure is meaningful only if the proportion of ISIs longer than x_{cut} is sufficiently large.

Two different cases can be found. In the first one, short ISIs are abundant with respect to what would be expected from an exponential distribution. Thus, such a neuron is bursting. In the second case the tail of the ISI distribution follows an exponential law but there is a lack of short ISIs with respect to the same law. These ISIs shorter than x_{cut} can also be declared as bursts but in contrast to the first case they appear rarely. In both cases the ISIs within the bursts of spikes are distinguished from ISIs between bursts using the assumption that bursts occur completely at random, that is, are distributed according to an exponential density. As a first hint for the presence of frequent (first case) or rare bursts (second case), CV larger or smaller than one can be used, respectively.

RESULTS

Slow and Fast Neurons

Spontaneous activity was recorded from 19 neurons in freely breathing animals (FB neurons) and from 15 neurons in tracheotomized animals (TT neurons). The mean and median of the ISIs of each neuron were determined. Histograms of the median ISIs are shown in Figure 1. Mean ISIs (not shown) ranged from 87 to 210 s and median ISIs ranged from 0.026 to 226 s. Six neurons (18%) had a mean ISI greater than 3.1 s and a median ISI greater than 2 s, as visible in the right part of the histograms in Figure 1. These "slow" neurons fired too few spontaneous action potentials (less than 50 per neuron) to permit quantitative analyses and were not studied.

Further analyses were thus performed on 14 TT and 14 FB neurons, whose firing characteristics (mean, variance, CV, and others) are summarized in Table 1. Spike trains of some of them are presented in Figure 2 showing the diversity of the firing patterns found in the samples. It can be seen in this figure that TT neurons had a global tendency to fire less spikes than FB neurons. This change in behavior, and others, induced by tracheotomy is quantitatively analyzed in the following subsections.

Trend, Periodicity, and Serial Dependency

Neurons showing a trend, periodicity, or serial dependency (e.g., spikes appearing in couples) had to

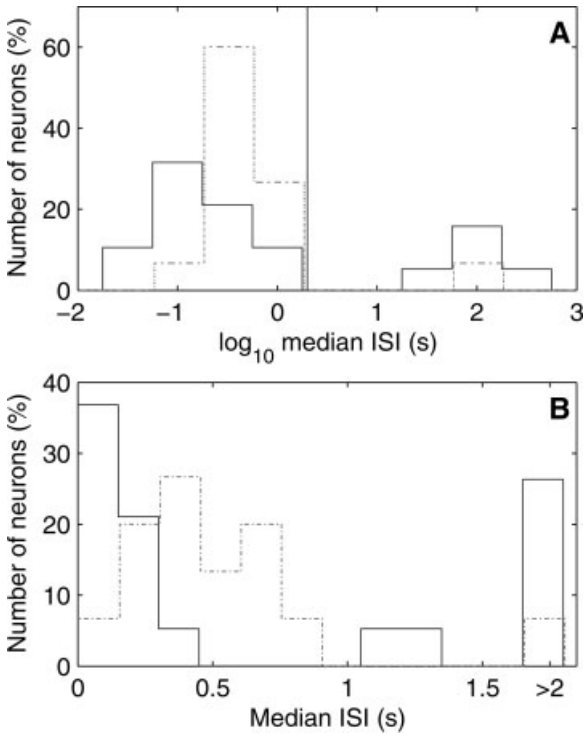


Figure 1 Histograms of median interspike intervals (ISI) of receptor neurons in freely breathing (FB) control animals (19 FB neurons, solid lines) and in tracheotomized (TT) animals (15 TT neurons, dotted lines). (A) Histograms in log-scale, all log median values shown. The vertical line at 2 s separates fast neurons on the left from slow ones on the right. (B) Histograms in linear scale, all median values are shown but neurons with median ISI greater than 2 s [on the right of the vertical line in (A)] are gathered in a single bar. In all other figures, only fast neurons, that is, with median ISI smaller than 2 s, are studied (14 neurons in each category).

be detected because such firing patterns would hide the more simple and basic firing behaviors that we wished to focus on. Only after removing these neurons, or the portion of the recordings displaying them, could the other analyses be performed reliably.

Trend. All TT neurons were found to be stationary with respect to the mean. For FB neurons a trend in the length of ISIs was detected in two records (FB29 and FB34). In FB29 the end of the record showed a slowing down of the activity. When we removed this part, the record became stationary. The neuron FB34 showed a more complicated behavior (see below). This trend was also removed by keeping approximately only the first 100 s of the record for further statistical evalution.

Periodicity. Spectral densities were computed for both categories of data. In the TT category no apparent peaks appeared, so no TT neuron had a periodic firing. Three neurons of the FB category were found to be significantly periodic close to 1 Hz (Fig. 3): FB35 and FB31 (clearly visible) and FB30 (less apparent). The first two were also detected by the Wald-Wolfowitz test (with $Z = 2.955$, $p = 0.003$ for FB35 and $Z = 2.430$, $p = 0.015$ for FB31). In all three cases the location of the peak of the periodogram corresponded exactly to the breathing frequency: 1.14 ± 0.06 (mean $\pm$ standard deviation, FB30), 1.10 ± 0.04 (FB31), and 1.48 ± 0.12 (FB35) Hz.

In general, for records with *CV* close to one the spectrum appears flat. For records with large *CV* the spectra have decreasing shape (similar to $1/f$ noise) and for records with *CV* lower than one the spectra are unimodal.

Serial Dependency. ISIs in six neurons were found to be serially correlated, three positively and three negatively. The neurons with positively correlated ISIs were FB34, TT52, TT73 (where FB34 was already problematic in trend testing). These neurons were also irregular with high values of *CV* (respectively 3.16, 1.27, and 1.10), which is indicative of a bursting tendency. On the contrary, the neurons with negatively correlated ISIs (FB25p, FB3l, and TT75) presented a tendency to fire relatively more regularly (their *CV*s were 0.77, 0.80, and 0.91, respectively) and with no apparent periodicity in two of them.

Central Value and Variability of ISIs

To characterize the central value and the variability of the ISIs in a single neuron, two types of measures were used: parametric (mean and standard deviation of ISIs) and nonparametric (median and interquartile range of ISIs). The parametric measures are sensitive to the absolute magnitude of the data (hence their sensitivity to outliers) whereas the nonparametric ones are not because they are based only on the ranks of data.

Central values (measured by means or medians) and variabilities (measured by standard deviations or interquartile ranges) of all "fast" neurons, either stationary (after removal of the nonstationary part of the record for FB29 and FB34) or periodic, were considered. Inclusion of periodic activity is justified because periodicity does not affect the average rate and increases variability in a deterministic way. Studying fit to distributions (see next subsection) was done following a more selective approach. FB and TT categories were compared based on cumulative histograms because interpretation of these graphs is relatively straightforward. Two cumulative histograms have the

Table 1 Global Characteristics of Fast Neurons*

	FB†	TT‡
Number of neurons	14	14
Number of ISIs	769 (71–2128)	310 (128–788)
Record duration	221 (184–280)	156 (63–212)
Central values $\bar{x}$	0.64 (0.087–3.0)	0.65 (0.19–1.7)
x_M	0.16 (0.026–l.3)	0.37 (0.l2–0.84)
Variabilities s	0.88 (0.0 69–4 .1)	0.73 (0.20–2.1)
IQR	0.26 (0.059–4.8)	0.71 (0.25–2.1)
CV	1.2 (0.57–2.9)	1.1 (0.84–1.3)
CV_M	1.9 (0. 78–3.8)	1.7 (1.2–2.5)

* Values given are for "fast" neurons (i.e., with median ISI smaller than 2 s) in the corresponding categories. The mean number of ISIs per neuron and the mean record duration are indicated as well as their extreme values. For each neuron i the mean ($\bar{x}_i$), standard deviation (s_i), and coefficients of variation (CV_i) of ISIs were determined, as well as their median (x_{Mi}), interquartile range (IQR_i), and the ratio IQR_i/x_{M_i} (CV_{M_i}) The table gives the means of x_i, s_i, and CV_i and the medians of x_{Mi}, IQR_i and CV_{Mi}, each based on 14 values (except CV_M of FB neurons, one outlier excluded). It gives also the minimum and maximum of each characteristic.

† Freely breathing animals.

‡ Tracheotomized animals.

same central value if they are superimposed or if they cross close to their 50% levels. They have the same variability if their global slopes, especially in the 25–75% levels, are the same. Note that the two characteristics, central value (location of the 50% level) and variability (slope), are a priori independent from one another.

Central Value of ISIs. The cumulative histograms of the means and medians of ISIs for FB and TT neurons are shown in Figure 4 (A,B). Both histograms had the same central values (location of the histograms) regarding means [Fig. 4(A), p = 0.37] but not regarding medians, the TT medians being significantly longer than the FB ones [Fig. 4(B), $p = 0.027$].

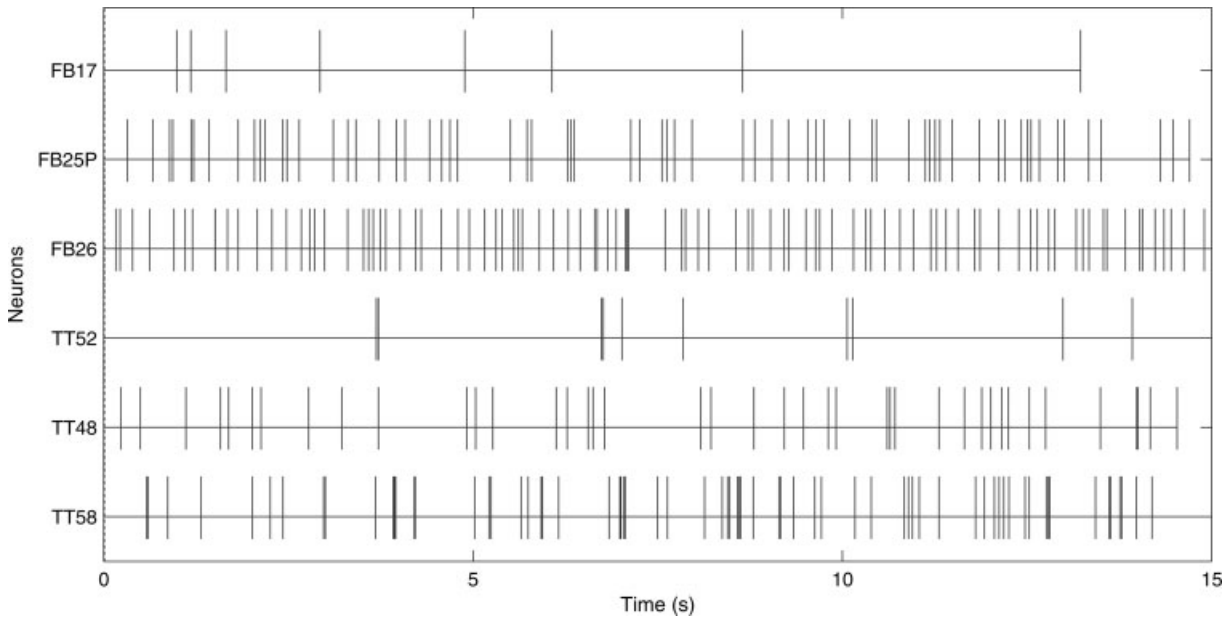


Figure 2 Examples of plots derived from recorded spike trains for the two categories (14 control neurons FB and 14 neurons TT from tracheotomized rats) of neurons with median ISI smaller than 2 s (see Fig. 1). Action potentials are represented by vertical lines during 15 s, from 60 to 75 s after beginning of recording. The time scale is the same for all graphs. The three neurons shown in each category illustrate long (FB17: 1.12 s, TT52: 0.84 s), median (FB25P: 0.18 s, TT48: 0.33 s), and short ISIs (FB26: 0.16 s, TT58: 0.12 s).

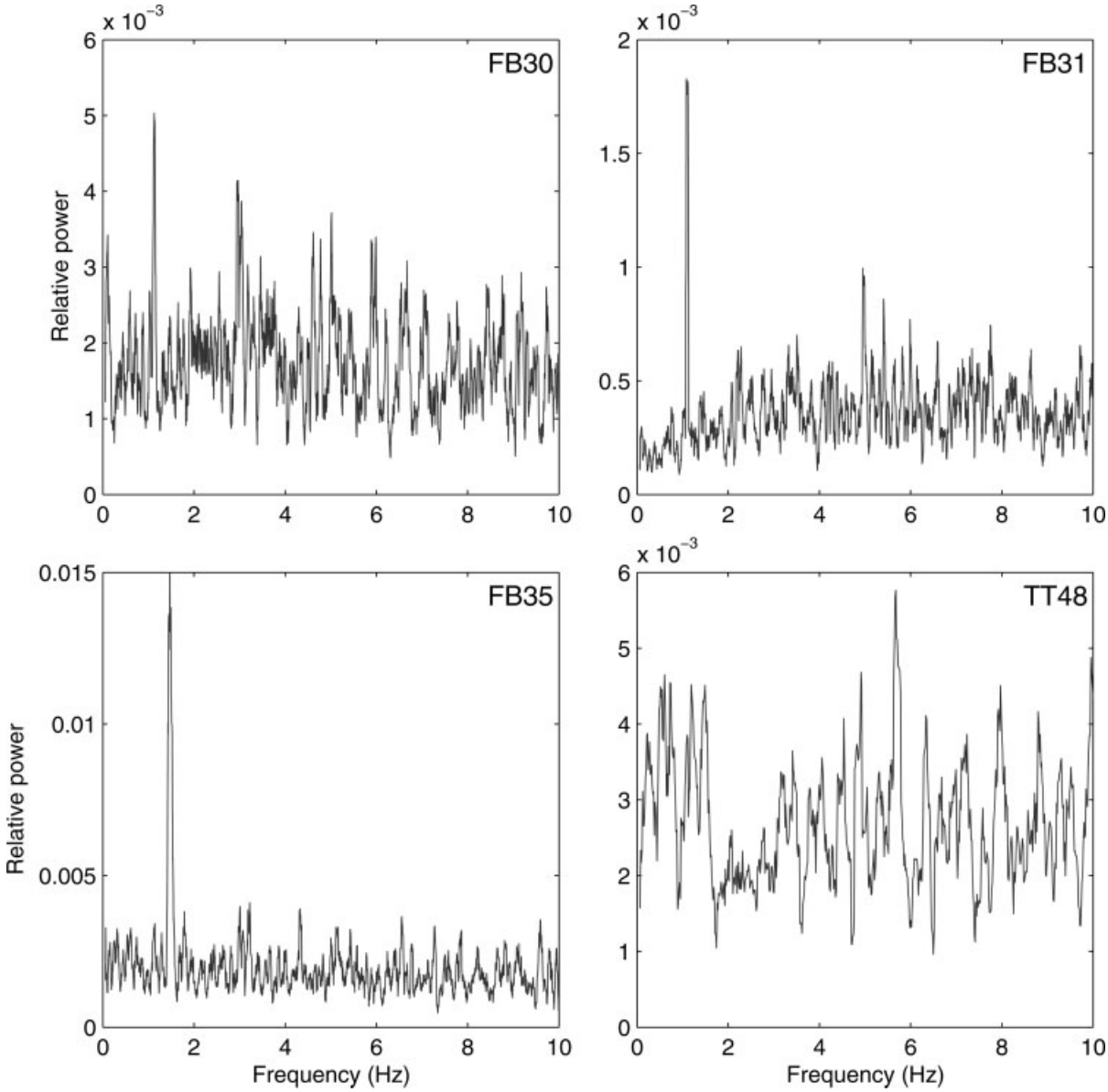


Figure 3 Spectral power densities of spike trains. The peaks indicate that the dominant firing frequencies are 1.13 Hz (FB30), 1.10 Hz (FB31), and 1.46 Hz (FB35). These values were found to be the respective breathing frequencies of the animals recorded from. The fourth periodogram (TT48), without dominant firing frequency, is representative of all other neurons.

Figure 4 Cumulative histograms of central values (A,B), variabilities (C,D), and relative dispersions (E,F) of ISIs for neurons in control (FB, solid line) and tracheotomized rats (TT, dashed line). (A) Means of ISIs $\bar{x}$ (in seconds). The cumulative histograms in both categories have same locations (significance probability $p = 0.37$) but different variabilities ($p = 0.0016$). (B) Medians of ISIs $\bar{x}$ (in seconds). The TT and FB histograms are significantly different in position ($p = 0.027$) and variability ($p = 0.036$). (C) Standard deviations, s, of ISIs (in seconds). The locations of FB and TT histograms are not significantly different ($p = 0.54$) but their variabilities are ($p < 10^{-3}$). (D) Interquartile ranges, *IQR*, of ISIs (in seconds). Two FB neurons with *IQR* greater than 2.5 s are not shown. The histograms have the same location (p value 0.61) but different variabilities ($p < 10^{-3}$). (E) Standard coefficient of variation of ISIs $CV = s/\bar{x}$ (dimensionless). In a Poisson process, CV is 1 (thin vertical line). Both histograms have the same locations ($p = 0.79$) but different variabilities ($p = 1.2 \times 10^{-3}$). (F) Relative dispersion of ISIs based on interquartile range and median, $CV_M = IQR/x_M$ (dimensionless); in a Poisson process, $CV_M = 1.58$ (thin vertical line). Both histograms have the same locations ($p = 0.61$) but different variabilities ($p < 10^{-3}$).

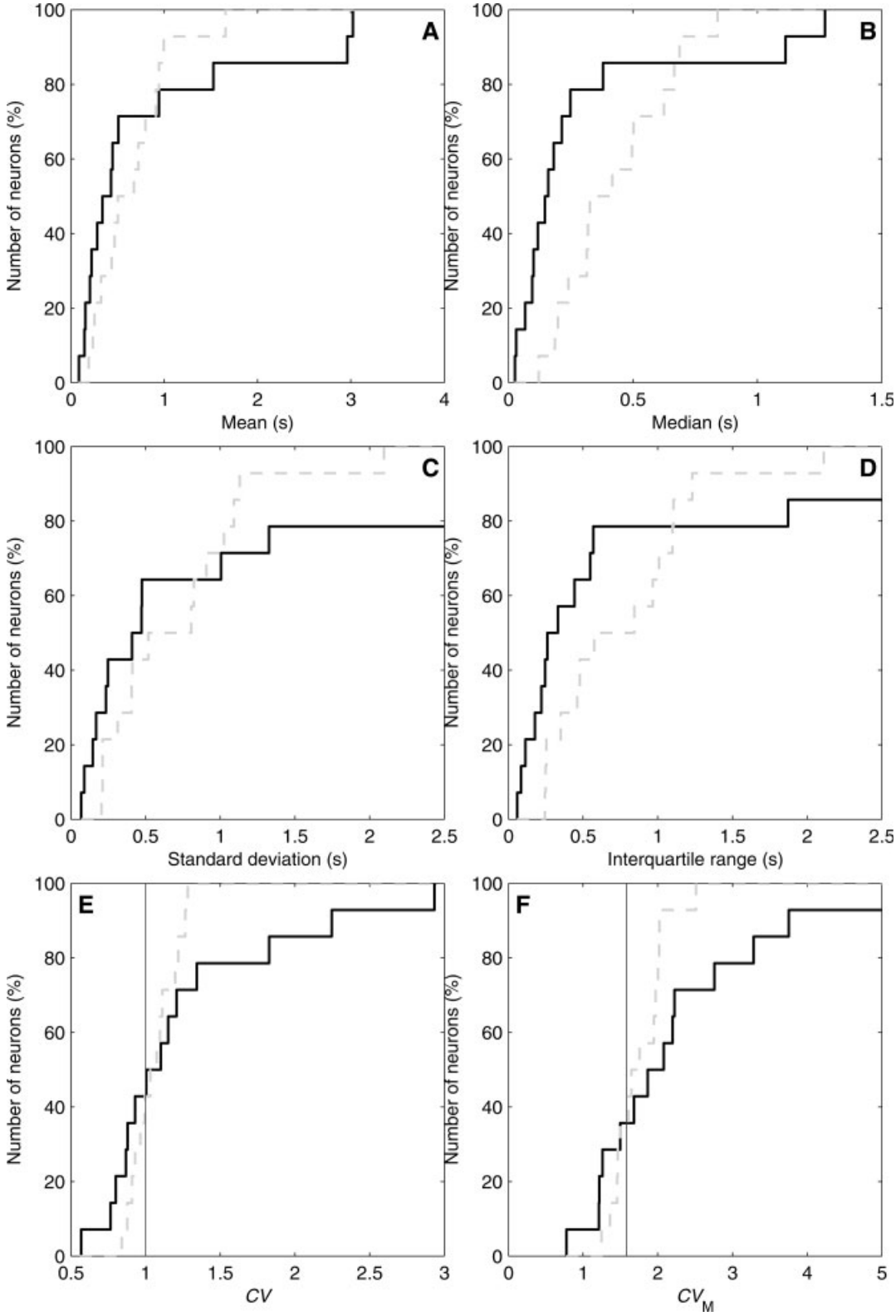

A
B
C
D
E
F
Number of neurons (%)
Mean (s)
Median (s)
Standard deviation (s)
Interquartile range (s)
CV
CV_M

Figure 4

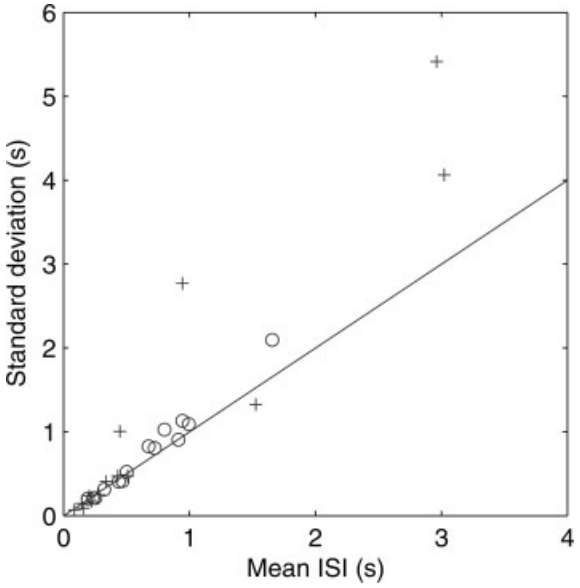


Figure 5 Scatter plot of standard deviation, s, against mean ISI $\bar{x}$ of the neurons' stationary activity of the two categories (crosses, control FB; circles, tracheotomized TT). Two outliers (for s) not shown. In a Poisson process, the points would lie on the line of slope 1.

Both means [Fig. 4(A), $p = 0.0016$] and medians [Fig. 4(B), $p = 0.036$] had different variabilities (slopes of the histograms) as judged by the Fisher test. This means that the variability of the central values in TT neurons was significantly lower than those in FB neurons.

Variability of lSIs. The cumulative histograms of the standard deviations s and interquartile ranges IQR of the ISIs in the two categories of neurons are compared in Figure 4(C,D), those of CV and CV_M are compared in Figure 4(E,F). The central values (locations) of FB and TT histograms were not significantly different as judged by either s or IQR but their variabilities (slopes), tested by Fisher test, were ($p = 1.1 \times 10^{-4}$ for s and $p = 1.2 \times 10^{-4}$ for IQR). The same conclusion holds for CV and CV_M with $p < 10^{-3}$. This means that neurons in the FB category were more different from one another than those in the TT category.

Relation between Average of ISIs and Their Variability. For TT as well as FB cells, the relation between the means and the standard deviations of ISIs in each neuron was well described by linear regressions with positive slopes (Fig. 5), which means that faster neurons tended to fire more regularly than slower neurons. The slopes of the regression lines for both FB and TT were close to one. This is compatible with a Poisson process although it is not a sufficient condition. So, the values of CV (see above) suggest that the deviation from exponential distribution went in both directions, overdispersion ($CV > 1$) and underdispersion ($CV < 1$, see Fig. 4).

Stochastic Processes Describing Spontaneous Activity

Up to now, the description of the neuronal activity was based on single parameters (mean, CV of each neuron, etc.). Now we want to propose a more detailed characterization of this activity based on functions that aim at describing the mechanisms responsible for generating spontaneous spikes. This analysis can be meaningfully done only on stationary neurons, which led us to remove the neurons FB30, FB31, and FB35, which presented a periodic activity.

Differential Entropy of ISIs. Entropy was estimated for both FB and TT neurons, and then compared to the theoretical entropy of the Poisson process with the appropriate mean. Figure 6 shows that the ISI distributions were closer to exponential in tracheotomized rats than in freely breathing ones.

Fitting of Spontaneous ISIs to Theoretical Distributions. Two approaches were followed for finding theoretical distributions liable to describe spontaneous ISIs.

In the first approach the observed ISI distribution in each record was compared to simple probability laws (exponential, Gamma, and Weibull). Table 2 confirms that the activity of FB neurons did not follow a Poisson process. The fit to the exponential distribution was rejected in 91% of FB neurons but in only 50% of TT neurons. Thus, there was a greater tendency of TT neurons to fire spikes purely at random than FB neurons. Using Gamma and Weibull distributions improved the fitting, although more than 73% of FB trains were still not accounted for. The situation was markedly different for tracheotomized animals because the three simple laws were rejected in only 29% of records. It can be noted that, for both FB and TT categories taken together, all seven exponential spike trains were of course accounted for by the Gamma law. Similarly, all 12 Gamma trains also fitted to the Weibull law, although a single train was fitted to the Weibull distribution only.

The second approach is based on the exponential law alone. Indeed, a closer examination of ISI histograms (Fig. 7) suggests that the relative failure of the exponential distribution was often due to abundance of the shortest ISIs, which can be interpreted as due to bursts of spikes. To test this hypothesis we exam-

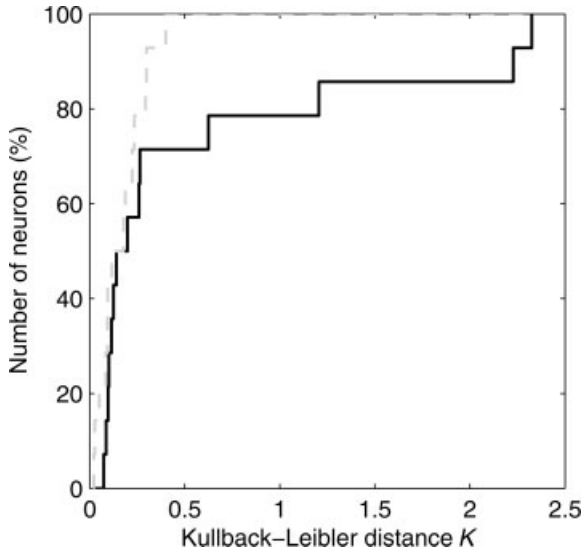


Figure 6 Cumulative histograms of deviations of ISI distributions from the exponential distribution. The comparison is based on Kullback-Leibler distance K, which measures the difference between the entropy of a spike train and the entropy of the Poisson process with the same mean ISI, therefore $K = 0$ implies Poissonian firing. The average distances of both categories FB (solid line) and TT (dashed line) are not significantly different ($p = 0.21$) but their variabilities are ($p < 10^{-7}$). Three neurons with periodic activity were not included.

ined whether an ISI length x_{cut} existed such that the ISIs longer than x_{cut}, that is, forming the tail of the distribution, follow an exponential shape. After excluding the shortest ISIs, the tails of the distributions could be fitted to exponentials (difference not significant at level 5%) in all stationary neurons, including 12 of them (Table 2) for which none of the three tested distributions (exponential, Gamma, and Weibull) was fitted. Of course, this procedure is meaningful only if x_{cut} is very small. Actually this is what was found: in all neurons, the values of x_{cut} never exceeded 150 ms (with a single exception) in FB neurons and 40 ms in TT neurons.

According to this second procedure neurons could be classified into three groups: those with random, regular, and bursting spike trains respectively. In the first group ($x_{cut} = 0$) the spikes can be considered as fired according to a purely random (Poisson) process (they are the same as the "exponential not rejected" of Table 2). The two other groups ($x_{cut} > 0$) had either less short ISIs (underdispersed spikes) or more short ISIs (overdispersed) than expected from a purely random process. The underdispersed trains can be called "regular" and the overdispersed ones "bursting". Table 3 shows that globally 56% of neurons were bursting and the others were not. However, the regular firing was absent in FB animals and appeared only in TT animals.

The second classification is not independent of the first one: all nonbursting (random and regular) neurons except one were fitted to at least one of the theoretical distributions tested (exponential, Gamma, or Weibull), whereas most bursting neurons (79%) could not be fitted to any of them (Table 4).

Crossing the two classifications, six firing patterns can be distinguished: exponential, Gamma, Weibull, regular, bursting, and a sixth pattern that is both Gamma, according to the first classification, and bursting, according to the second one (Table 5). If one excepts the intermediate latter group (three neurons, all in the TT category), the majority of FB neurons were bursting and the majority of TT neurons were not, that is, they were regular or followed a simple theoretical law. Two other facts can be noted. First, the Gamma and Weibull laws can accomodate a limited number of short intervals; if short ISIs are too many, they fail. Second, the regular neurons, with less short ISIs than expected from exponential, do not fit always to the same simple law; of the three such trains, one is described by the Gamma law, another by the Weibull law, and the third one by none.

Analysis of Bursts. In both groups with too many or too few bursts, the bursts of spikes, not the individual spikes, are exponentially distributed. The actual values of the maximum possible ISI in a burst, x_{cut}, depended on the neuron. Their cumulative histograms are shown separately for FB and TT neurons in Figure 8(A). Statistical testing indicated that these histograms are significantly different ($p \leq 10^{-3}$) both in position and variability, x_{cut} being shorter and less variable in TT than in FB neurons, although testing may be influenced by the low numbers of observations in both categories: 7 and 10. This confirms that FB neurons presented often a bursting spontaneous activity whereas

Table 2 Fittings of Observed ISIs to Theoretical Distributions*

Groups	FB	TT
Stationary neurons†	11 (100%)	14 (100%)
Exponential not rejected	1 (9%)	6 (43%)
Gamma not rejected	2 (18%)	10 (71%)
Weibull not rejected	3 (27%)	10 (71%)
All simple models rejected‡	8 (73%)	4 (29%)

* Kolmogorov-Smirnov test at significance level 5%.
† No trend and no periodicity of ISIs.
‡ Exponential, Gamma, and Weibull rejected.

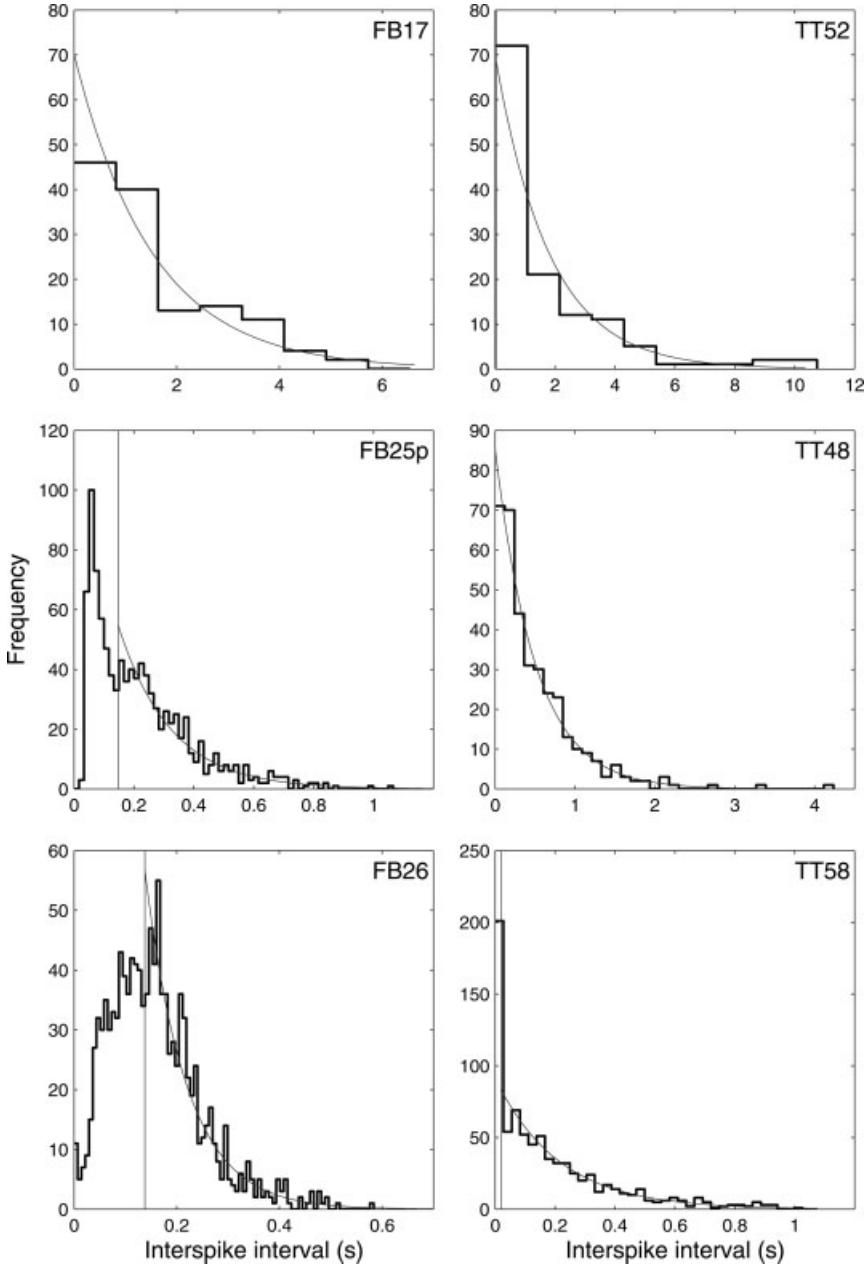


Figure 7 Examples of histograms of interspike intervals for the neurons shown in Figure 2 in each category FB (left column) and TT (right columns). The histograms of neurons FBI7, TT5, and TT48 fit directly to exponential distributions; they are not bursting and action potentials are fired at random. Those of neurons FB25p, FB26, and TT58 do not fit directly. After the smallest ISIs (corresponding to bursts, on the left of the vertical line at x_{cut} shown for FB25p, FB26, and TT58) are removed, the tails fit to exponential distributions. FB25p and FB26 are "regular" neurons (too few bursts), and TT58 is a bursting neurons (too many bursts), whose bursts are fired at random.

fewer TT neurons fired in bursts. When present, these bursts were furthermore less conspicuous.

We characterized further the firing of bursting neurons by determining the median duration of ISIs within bursts, that is, with a duration less than x_{cut} [Fig. 8(B)]. The proportion of ISIs appearing within bursts was between 23 and 77% in FB neurons and fell down to at most 19% in TT neurons [Fig. 8(C)]. The mean number of ISIs per burst ranged from 1.2 to 4.9 in FB neurons against 1 to 1.4 in TT neurons

Table 3 Fittings of Truncated ISI Distributions to Exponential Law*

Groups	FB	TT
Purely random trains†	1 (9%)	7 (50%)
Regular trains‡	3 (27%)	0 –
Bursting trains§	7 (64%)	7 (50%)

*Stationary neurons only.

† Fit to exponential law without truncating, that is, $x_{cut} = 0$ (see Fig. 7). Same as "Exponential not rejected" of Table 2, plus one TT neuron that fits for $x_{cut} = 0.004$ s smaller than any of its ISIs.

‡ Fit after truncating $x_{cut} > 0$ and with a number of ISIs $< x_{cut}$ smaller than the number F expected from the exponential law fitted to the tail (ISIs $> x_{cut}$).

§ Trains with $x_{cut} > 0$ and a number of ISIs $< x_{cut}$ greater than F.

[Fig. 8(D)]. (Note that the number of spikes is the number of ISIs plus one, so a burst of one ISI includes two spikes.) On these three criteria the FB and TT categories appeared significantly different ($p \leq 0.015$). However, the mean burst durations (0.12 to 1.92 s in FB neurons and 0.18 to 1.40 s in TT) were not significantly different [Fig. 8(E)].

DISCUSSION

Spontaneous Firing Rate and Its Variability, Trend, and Periodicity

Neurons from FB animals were more often slowly firing (26% with firing frequency below 0.5 Hz) than those of TT animals (7%). However, among the "fast" neurons (above 0.5 Hz) the FB neurons were globally more active than the TT neurons [as judged by medians of ISIs, Fig. 4(B)]. This apparent contradiction can be explained when it is realized that the variability of the firing frequency among "fast" FB neurons was significantly higher than that of TT neurons (see Fig. 4). In summary TT neurons displayed a simpler behavior, because their activity was remarkably similar to one another, all of them tending to fire at the same average frequency. Normal breathing changed this behavior: although the average firing seemed to remain approximately the same, some neurons tended to decrease their activity, resulting in the increase of the proportion of "slow" neurons, whereas an equal number tended to increase their activity resulting in the observed shift of the average frequency among "fast" FB neurons and the corresponding increase in their variability.

This diversity among FB neurons was further confirmed by the fact that systematic changes of the firing frequency during the recording session (trend) as well as periodic firing were observed only in this category. We showed that when a periodicity was detected it was precisely synchronized with breathing. Unsurprisingly, no periodic activity was found in TT neurons.

An unknown proportion of "silent" neurons, without any spontaneous activity and therefore with a high probability of not being recorded at all, must be added to the known proportion of excluded "slow" neurons. This proportion could be decreased in further studies, but not reduced to zero because there is likely no solution of continuity between slowly, rarely, and never firing neurons. The major part of this study, discussed below, is concerned with the "fast" neurons only, that is, those with a median ISI smaller than 2 s, which are presumably the most influential from a physiological point of view.

Patterns of Spontaneous Activity

Three types of basic firing patterns could be distinguished: purely random, less dispersed than purely random (called here "regular"), and more dispersed than purely random (called here "bursting"). Although in the logic of our analysis the regular and bursting neurons can be further subdivided, in three and two subtypes, respectively (see Table 5), this is an unessential refinement because as no two recordings are exactly alike, the number of subtypes is a matter of convenience and sample size. However, the diversity of firing patterns is worth noting because it is a clear indication that the spontaneous activity is influenced by several independent factors (see below), which likely include the diverse ages and physiological characteristics of ORNs.

More important than the number of subtypes is the fact that not all basic firing patterns are equally fre-

Table 4 Comparison* of the Classifications in Tables 2 and 3

		Simple Distributions†		
		EGW	Not EGW	Total
Bursting‡ (or not)	Purely random + regular	10	1	11
	Bursting	3	11	14
Total		13	12	25

* Without distinguishing categories FB and TT. This contingency table departs significantly from equiprobable distribution between cells ($\chi^2 = 9.81$, p-value $= 0.17\%$).

† Classification used in Table 2. EGW: fits to either exponential (E), Gamma (G), or Weibull (W) law.

‡ Classification used in Table 3.

Table 5 Global Classification of Spontaneous Firing Patterns*

Groups†	FB	TT
Purely random (exponential)	1 (9%)	6 (43%)
Gamma, not bursting	1 (9%)	1 (7%)‡
Weibull, not bursting	1 (9%)	0 –
No simple law, regular	1 (9%)	0 –
Gamma, bursting	0 –	3 (21%)
No simple law, bursting	7 (64%)	4 (29%)

* Stationary neurons only.

† Based on crossing the groups defined in Tables 2 and 3.

‡ This is the neuron mentioned in note † of Table 3 for which $x_{cut} = 0.004$ s.

quent, their relative proportions being correlated with the breathing conditions. FB neurons displayed the three basic patterns, purely random, regular, and bursting, in respective proportions 9, 2, and 64%, whereas no TT neurons presented the regular pattern, the two others being in equal proportions (see Tables 2 and 3). All observations were consistent with a greater simplicity of spontaneous activity in neurons from TT rats. Half of them fired spikes completely at random according to a Poisson process, that is, the corresponding ISIs followed exponential distributions with modes close to the origin (Table 2, Fig. 8). More than 70% of these TT neurons fired spikes according to a Gamma law (which includes the exponential one) or a Weibull law. The remaining TT neurons could be considered as bursting neurons that fired bursts of spikes or isolated spikes according to a Poisson process. However these bursts were relatively infrequent and usually composed of a single short ISI; it was very rare that they included more than one ISI.

In normally breathing animals (FB) the situation was different. Not only trend and periodicity synchronized with breathing appeared, but even in the records that could be considered as stable (the majority), the activity resembled the Poisson process in only one of them. Less than 30% of records followed a Gamma or Weibull distribution, which means that more than 70% of FB neurons could not be modeled by simple laws. So the proportions of modeled neurons were inverted in the two categories. Furthermore, the departure from randomness in FB neurons occurred in both possible directions, more regular and more irregular (bursting). In bursting FB neurons, the proportion of ISIs within bursts was greater than in the bursting TT [Fig. 8(C)] neurons and the bursts were more frequently composed of two ISIs or more [Fig. 8(D)].

In conclusion, the spontaneous activity in FB and TT neurons was qualitatively and quantitatively different. Some FB neurons had trend in their activity or were driven by breathing. Moreover, they rarely fitted to a simple homogeneous generator of spikes but often fired bursts of spikes in a completely random manner. On the other hand, TT neurons had activity that often closely resembled a Poisson process and if it deviated from this description, it was only slightly, either towards another simple process (Gamma or Weibull) or towards a bursting process involving small bursts.

Comparison with Other Species

Table 6 summarizes the mean firing rate and mean ISI observed in various species (frogs, hen, mouse, and rat). It must be noted that 1.25 spike/s in rat is derived from the mean ISI. The estimate of the firing rate based on the median ISI gives a rather different central value (4 spikes/s) because in a positively skewed distribution, that is, with the right tail longer than the left tail, as all distributions of ISIs are, the median is always smaller than the mean. Unfortunately, median spontaneous frequencies were not given in previous reports, except for a frog species.

Spontaneous activity is higher in the rat than in the frog. Part of this difference might be related to the temperature of the animal, which is higher in the rat (37–38°C) than in the frog (13°C). However, temperature by itself cannot account for the very high activity recorded in the olfactory nerve layer of the olfactory bulb in hens (40°C) by McKeegan (2002), although ascertaining whether only ORNs were recorded is difficult to do at this recording site. Spontaneous activity in the mouse (Sicard, 1986) seems higher than in the rat and less variable. It was recorded in conditions minimizing the influence of air movements due to respiration, which likely explains its lesser variability.

The proportion of (presumably regular) neurons following Gamma or Weibull law was 10% (hen), 28% (frog), and 27% (rat FB). After removal of bursts, the interburst intervals were found to follow a complete random (Poisson) process in all studied neurons in both frogs (Rospars et al., 1994) and rats (this study).

Biological Interpretation

We describe here for the first time spontaneous activity of ORNs *in vivo* in mammals. Our data obtained in TT and breathing rats show that the pattern of spontaneous firing is profoundly influenced by the

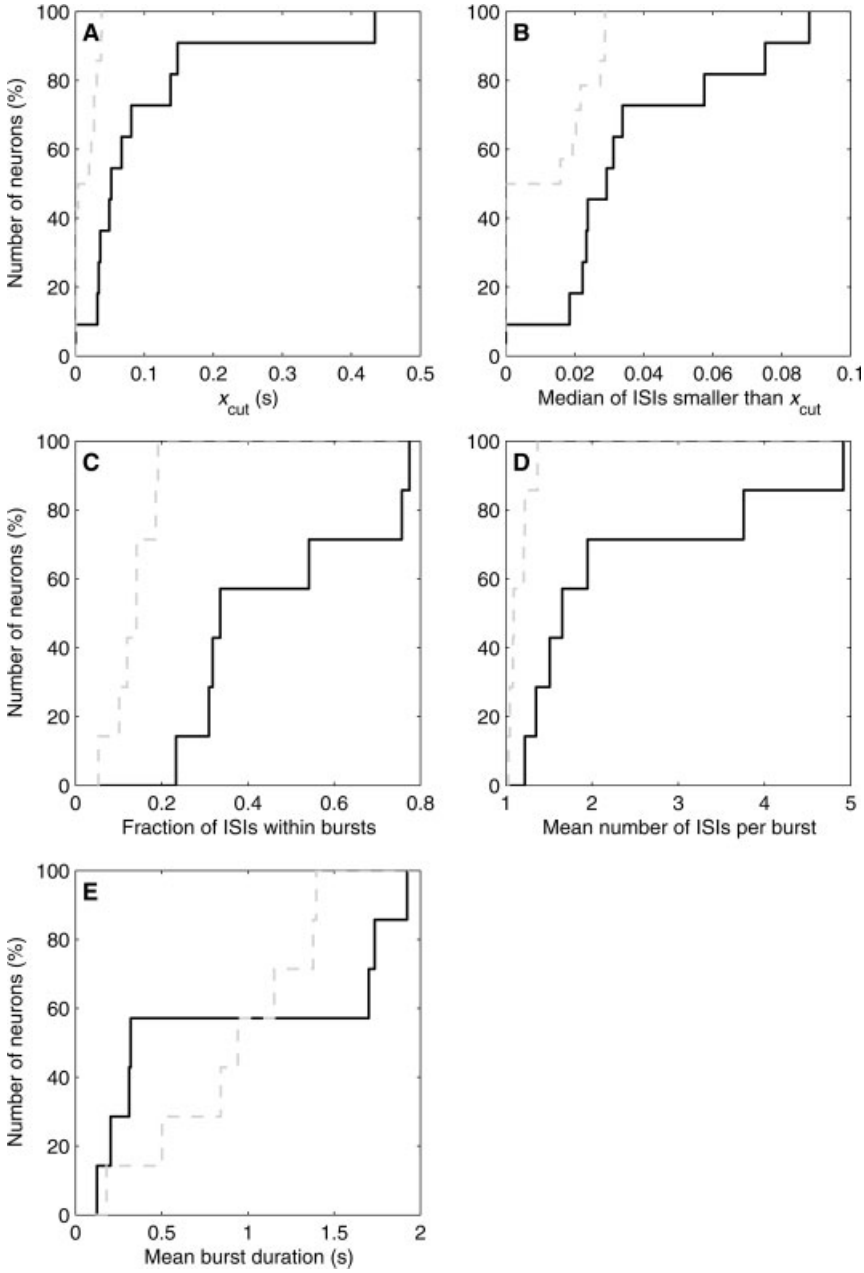


Figure 8 Cumulative histograms of x_{cut}, (A,B) and burst characteristics (C–E) for control animals (FB, solid line) and tracheotomized ones (TT, dashed line). (A) In all neurons of both categories, when ISIs shorter than x_{cut} are excluded from a record, the remaining ISIs (forming the tail of the distribution) follow an exponential law (see Table 2). Both histograms have different locations ($p = 10^{-3}$) and variabilities ($p < 10^{-8}$). (B) Comparison of the median duration of ISIs within bursts. Locations ($p = 0.005$) and variabilities ($p = 0.015$) are different. (C) Ratio of the number of ISIs within bursts to the total number of ISIs in a given neuron. The histograms differ both in location ($p < 10^{-3}$) and variability ($p = 0.0016$). (D) Mean number of ISIs per burst. The difference in location of histograms ($p = 0.003$) and their variability ($p < 10^{-5}$) are significantly different. (E) Mean burst duration. The histograms are not significantly different ($p > 0.15$).

respiratory conditions. The situation in TT rats can be considered as the simplest one because the air above the olfactory epithelium is not renewed. That condition diminishes or abolishes the occurrence of mechanical, thermal, or chemical stimulation. On the contrary, in normally breathing rats the surrounding air is renewed following the respiratory rhythm, which changes periodically the thermal and chemical

Table 6 Average Spontaneous Activity of ORNs in Different Species

Species	Firing Rate*	Mean ISI†	*n*	References
Frogs	0.4–0.5	2–2.5	16–112	Several authors (Cf. Rospars et al., 1994)
Frog‡	0.40	2.49	28	Rospars et al. (1994)
Hen	3.8	0.25	11	McKeegan (2002)
Mouse§	2.08	0.47	34	Sicard (1986)
Rat‖	1.25	0.73	28	This study

* In spikes per second.
† In seconds.
‡ "Fast" neurons only (mean ISI < 5.2 s).
§ Calculated from an histogram.
‖ "Fast" neurons in both freely breathing and tracheotomized rats (both have same mean, see Fig. 4).

environment of ORNs. The "spontaneous" activity can thus be interpreted as the summation of two processes: an intrinsic process, due to noise in the transduction machinery firing mostly single spikes, and an extrinsic process, due to the uncontrolled occurrence of either mechanical, thermal, or chemical (odorant molecules) effects from the environment, firing bursts of spikes. Among intrinsic processes, the random opening of the cyclic-nucleotide-gated channels might play a major role. It is related to the high basal level of cAMP in the neuron and contribute up to 40% to the resting conductance of the ORN (Pun and Kleene, 2003).

The observed differences between TT and breathing animals are consistent with this dual process interpretation. Intrinsic and extrinsic processes would be active in both conditions, but the extrinsic process is expected to be minimized in the TT condition. This hypothesis is in perfect agreement with our observations. However, normal breathing does not systematically induce an increase in firing rate and bursting; sometimes it induces a decrease (see Figs. 1 and 4) and apparition of more regular firing patterns (Table 3). These results suggest that, in both experimental conditions, extrinsic stimulation led with comparable probabilities to spike triggering and spike suppression (or delay). These effects were unexpected. Indeed, they cannot be easily interpreted as resulting from a direct, simple stimulation of neurons due to mechanical, thermal, or chemical changes accompanying air movements. The mechanism behind this smaller activity and greater regularity in the nonbursting fraction of FB neurons remains to be investigated.

If all bursts result from air movement as suggested above, this raises the problem of understanding why only 21% of FB neurons were clearly periodic and firing bursts synchronized with respiration. A possible explanation is that the difference between the periodic and nonperiodic neurons lies in the probability of firing a burst at each respiratory cycle, the probability being one in the periodic neurons and less than one in the others. So, a nonperiodic neuron would merely miss some cycles at random whereas a periodic one would miss none. However, at each cycle, the number of discharging neurons connecting a given mitral cell would be large enough to generate a periodic global spontaneous input.

If spontaneous activity is globally driven by respiration in this way, interesting new possibilities are opened for its physiological role. For example, in FB rats odor stimulation was shown to generate successive bursts of spikes and EOG waves that were both synchronized with respiration (Chaput et al., 2002). This indicates that in breathing animals, olfactory sensory inputs must be *de facto* analyzed in the bulb according to the temporal rhythm driven by respiration. So, in the absence of odor, the role of spontaneous activity may be to keep the bulbar neural network synchronized with respiration. It may also be to select and maintain a synaptic organization in the bulbar network capable of processing efficiently periodic sensory inputs, as suggested by experiments in the visual system (Stryker and Harris, 1986). Understanding the role of spontaneous activity in the survival of ORNs and in the development and maintenance of the glomerular circuitry should benefit from a better physiological and pharmacological characterization of the intrinsic and extrinsic processes that underlie this activity.

This work was supported by a grant from Academy of Sciences of the Czech Republic (Information Society, 1ET400110401), and by program Barrande (N° 9146 QL), between France and Czech Republic (to J.P.R. and P.L.). The authors acknowledge Professor Edwin Griff (Department of Biological Sciences, University of Cincinnati) for his careful reading of a previous version of this manuscript.

REFERENCES

Bayling F. 1979. Temporal patterns and selectivity in the unitary responses of olfactory receptors in the tiger salamander to odor stimulation. J Gen Physiol 74:17–36.

Blank DL, Mozell MM. 1976. Olfactory receptor response characteristics: a factor analysis. Brain Res Bull 1:185–92.

Borodinsky LN, Root CM, Cronin JA, Sann SB, Gu X, Spitzer NC. 2004. Activity-dependent homeostatic specification of transmitter expression in embryonic neurons. Nature 429:523–530.

Chaput MA, Duchamp-Viret P, Duchamp A. 2001. Recordings from olfactory receptor neurons in the rat. In: Simon SA, Nicolelis MAL, editors. Methods in chemosensory research. Boca Raton: CRC Press, p79–90.

Cox DR, Lewis PAW. 1966. The Statistical Analysis of Series of Events. London: Methuen.

Cutforth T, Moring L, Mendelsohn M, Nemes A, Shah NM, Kim MM, Frisen J, Axel R. 2003. Axonal ephrin-as and odorant receptors: Coordinate determination of the olfactory sensory map. Cell 114:311–322.

Duchamp A, Revial MF, Holley A, Mac Leod P. 1974. Odor discrimination by frog olfactory receptors. Chem Sens 1:213–233.

Duchamp-Viret P, Chaput MA, Duchamp A. 1999. Odor response properties of rat olfactory receptor neurons. Science 284:2171–2174.

Duchamp-Viret P, Delaleu JC, Duchamp A. 2000. GABA(B)-mediated action in the frog olfactory bulb makes odor responses more salient. Neurosci 97:771–777.

Duchamp-Viret P, Duchamp A, Chaput M. 1993. GABAergic control of odor-induced activity in the frog olfactory bulb. Electrophysical study with pricrotoxin and bicuculline. Neuroscience 53:111–120.

Duchamp-Viret P, Duchamp A, Chaput MA. 2003. Single olfactory sensory neurons simultaneously integrate the components of an odor mixture. Eur J Neurosci 18:2690–2696.

Ebrahimi N, Habibullah M, Soofi ES. 1991. Testing exponentiality based on Kullback-Leibler Information. JR Statist Soc B 54:739–748.

Erzurumlu RS, Kind PC. 2001. Neural activity: sculptor of 'barrels' in the neocortex. TINS 24:589–595.

Gesteland RC. 1971. Neural coding in olfactory receptor cells. In: Beidler LM, editor. Handbook of Sensory Physiology IV Chemical Senses. Berlin: Springer-Verlag.

Gesteland RC, Howland B, Lettvin JY, Pitts VH. 1959. Comments on microelectrodes. Proc Inst Radio Eng 47:1856–1862.

Gesteland RC, Lettvin JY, Pitts VH, Rojas H. 1963. Odor specificities of the frog's olfactory receptors. In: Zotterman Y, editor. Olfaction and Taste I. Oxford: Pergamon, p 19–34.

Grubb MS, Rossi FM, Changeux JP, Thompson ID. 2003. Abnormal functional organization in the dorsal lateral geniculate nucleus of mice lacking the beta 2 subunit of the nicotinic acetylcholine receptor. Neuron 40:1161–1172.

Hanson MG, Landmesser LT. 2004. Normal patterns of spontaneous activity are required for correct motor axon guidance and the expression of specific guidance molecules. Neuron 43:687–701.

Hubel DH, Wiesel TN. 1970. The period of susceptibility to the physiological effects of unilateral eye closure in kittens. J Physiol 206:419–436.

Juge A, Holley A, Delaleu JC. 1979. Olfactory receptor cell activity under electrical polarization of the nasal mucosa in the frog. I Spontaneous activity. J Physiol Paris 75:919–927.

Lin DM, Wang F, Lowe G, Gold GH, Axel R, Ngai J, Brunet L. 2000. Formation of precise connections in the olfactory bulb occurs in the absence of odorant-evoked neuronal activity. Neuron 26:69–80.

Mathews DF. 1972. Response patterns of single-neurons in the tortoise olfactory epithelium and olfactory bulb. J Gen Physiol 60:166–180.

McKeegan DEF. 2002. Spontaneous and odour evoked activity in single avian olfactory bulb neurones. Brain Res 929:48–58.

McLaughlin T, Torborg TL, Feller MB, O'Leary DDM. 2003. Retinotopic map refinement requires spontaneous retinal waves during a brief critical period of development. Neuron 40:1147–1160.

Mombaerts P, Wang F, Dulac C, Chao SK, Nemes A, Mendelsohn M, Edmonson J, Axel R. 1996. Visualizing an olfactory sensory map. Cell 87:675–686.

O'Connell RJ, Mozell MM. 1969. Quantitative stimulation of frog olfactory receptors. J Neurophysiol 32:51–63.

Pun RYK, Kleene SJ. 2003. Contribution of cyclic-nucleotide-gated channels to the resting conductance of olfactory receptor neurons. Biophysical J 84:3425–3435.

Ressler KJ, Sullivan SL, Buck LB. 1993. A zonal organization of odorant receptor gene expression in the olfactory epithelium. Cell 73:597–609.

Rospars JP, Lansky P, Vaillant J, Duchamp-Viret P, Duchamp A. 1994. Spontaneous activity of first- and second-order neurons in the frog olfactory system. Brain Res 662:31–44.

Shannon C. 1948. A Mathematical Theory of Communication. Bell Sys Tech 27:379–423.

Shatz CJ, Stryker MP. 1988. Prenatal tetrodotoxin infusion blocks segregation of retinogeniculate afferents. Science 242:87–89.

Sicard G. 1986. Electrophysiological recordings from olfactory receptor cells in adult mice. Brain Res 397:405–408.

Siegel S. 1956. Nonparametric Statistics for the Behavioral Sciences. Tokyo: McGraw-Hill Kogakusha.

Strycker MP, Harris WA. 1986. Binocular impulse blockade prevents the formation of ocular prevents the formation of ocular dominance columns in cat visual cortex. J Neurosci 6:2117–2133.

Trotier D, Mac Leod P. 1983. Intracellular recordings from salamander olfactory receptor cells. Brain Res 268:225–237.

Vasicek O. 1976. A Test for Normality Based on Sample Entropy. JR Statist Soc B 38:54–59.

Vassar R, Chao SK, Sitcheran R, Nunez JM, Vosshall LB, Axel R. 1994. Topographic organization of sensory projections to the olfactory bulb. Cell 79:981–991.

Vassar R, Ngai J, Axel R. 1993. Spatial segregation of odorant receptor expression in the mammalian olfactory epithelium. Cell 74:309–318.

Wang F, Nemes A, Mendelsohn M, Axel R. 1998. Odorant receptors govern the formation of a precise topographic map. Cell 93:47–60.

Yu CR, Power J, Barnea G, O'Donnel S, Brown HE, Osborne J, Axel R, Gogos JA. 2004. Spontaneous neural activity is required for the establishment and maintenance of the olfactory sensory map. Neuron 42:553–566.

Zheng C, Feinstein P, Bozza T, Rodriguez I, Mombaerts P. 2000. Peripheral olfactory projections are differentially affected in mice deficient in a cyclic nucleotide-gated channel subunit. Neuron 26:81–91.

Zou DJ, Feinstein P, Rivers AL, Mathews GA, Kim A, Greer CA, Mombaerts P, Firestein S. 2004. Postnatal refinement of peripheral olfactory projections. Science 304:1976–1979.

Biol Cybern (2006) 94: 157–167
DOI 10.1007/s00422-005-0036-6

ORIGINAL PAPER

Biological Cybernetics

Lubomir Kostal · Petr Lansky

Similarity of interspike interval distributions and information gain in a stationary neuronal firing

Received: 3 February 2005 / Accepted: 24 October 2005 / Published online: 29 November 2005

Abstract The Kullback–Leibler (KL) information distance is proposed for judging similarity between two different interspike interval (ISI) distributions. The method is applied by a comparison of four common ISI descriptors with an exponential model which is characterized by the highest entropy. Under the condition of equal mean ISI values, the KL distance corresponds to information gain coming from the state described by the exponential distribution to the state described by the chosen ISI model. It has been shown that information can be transmitted changing neither the spike rate nor coefficient of variation (CV). Furthermore the KL distance offers an indication of the exponentiality of the chosen ISI descriptor (or data): the distance is zero if, and only if, the ISIs are distributed exponentially. Finally an application on experimental data coming from the olfactory sensory neurons of rats is shown.

1 Introduction

The discharge activity of neurons is composed of the series of events called action potentials (or spikes). It is generally accepted that these action potentials form the dominant mode of communication in the central nervous system of living organisms. Since Shannon developed his general and rigorous theory of communication and information transmission in electro-engineering systems (Shannon 1948), many scientists of various background (biologists, engineers, mathematicians) have tried to apply it to the study of the properties of neural systems. Probably the most fundamental questions point to a problem of neuronal coding, (Rieke et al. 1997).

L. Kostal (✉) · P. Lansky
The Institute of Physiology,
Academy of Sciences of The Czech Republic,
Videnska 1083, 142 20 Prague 4,
The Czech Republic
Tel.: +420-2-41062276,
Fax: +420-2-41062488,
E-mail: E-mail: kostal@biomed.cas.cz

The classical results in early neuroscience (Adrian 1928) show that the number of spikes per a time period (firing rate) is related to the stimulus intensity, i.e., the firing rate increases with increasing stimulus intensity (generally non-linearly). The idea that most of the information is transferred by this rate coding is probably the oldest existing, however, many related questions arise; for an overview see, e.g., Gerstner and Kistler (2002). There are examples of situations in which time averaging (counting) is hardly possible. This all leads to a different view of neural coding, where exact timing of spikes or their temporal pattern plays a key role (Buracas and Albright 1999; Johnson and Glantz 2004; Stevens and Zador 1995). While rate codes and temporal spike codes are shown to be compatible under many circumstances (Gerstner and Kistler 2002), it is clear that infinitely more different spike records may have the same rate. Therefore, there is a need for the quantification of the differences among these firing patterns and the information they transfer (Bhumbra et al. 2004; Buracas and Albright 1999; Nemenman et al. 2004; Paninski 2003; Rieke et al. 1997; Stein 1967; Strong et al. 1998; Zador 1998).

Neuronal firing under constant conditions is often described as a renewal process of interspike intervals (ISI). Then the ISIs are realizations of a positive random variable T and are fully characterized by the probability density function, $f = f(t)$, where $f(t)\,\mathrm{d}t = \mathrm{Prob}(T \in [t, t + \mathrm{d}t))$ (Cox and Lewis 1966). The renewal character of the ISIs implies stationarity of the neuronal activity. Other characteristics, for example the autocorrelation function (renewal density), can be easily computed (see Perkel et al. 1967 for details) and indicate important features of the mechanism behind neuronal firing, e.g., it is constant for the exponential model of ISI.

The basic statistical description of the ISIs can give some information related to the above-mentioned possibilities of information transfer. The terms (firing) rate, mean rate, frequency or mean/instantaneous frequency are used differently by different authors, depending on the context (Lansky et al. 2004). We use the term "firing frequency", ν, for the reciprocal value of the mean interspike interval, $\nu = 1/E(T)$,

where the mean value $E(T)$ is estimated by the average length of ISI. The coefficient of variation CV (a ratio of standard deviation to the mean value) gives preliminary information about the temporal dispersion of neuronal discharge within a spike train; for a review see (Christodoulou and Bugmann 2001). An important characteristic of the CV is that it gives an indication of the exponential ISI distribution – the most prominent of all ISI distributions – when $CV = 1$.

The first- and second-order statistical features are easy to estimate from experimental data. From the statistical point of view, the higher moments provide information about the shape of the distribution (ISI), that cannot be obtained using only the first- and second-order moments. Thus, the higher moments help us to quantify and measure some of the features, that are visible in histograms but not in the mean and CV. The higher moments represent the step between estimate of ISI density and the mean and CV only. On the other hand, the estimates of higher moments cannot be reliably determined from samples of relatively small size as is normal with the neuronal data. Thus, the attempts to use higher moments are not so frequent as using mean and CV (Han et al. 1998; Lewis et al. 2001; Ruskin et al. 2002; Shinomoto et al. 2002). There is a strong demand for measuring the variablity of neuronal discharge and CV is not the only existing method. Based on the results of Holt et al. (1996), Shinomoto et al. (2003) introduce a local measure of spike train variability, L_V, which is not based on the statistical moments. Some properties of L_V are similar to those of CV, e.g., they are both zero for a regular ISI sequence and for a Poisson process $L_V = 1$. Analogously to CV, L_V is not unique in the sense that $L_V = 1$ does not imply Poisson nature of spikes.

There exist many neuronal models of different types and for their mutual comparison, or their fitting to data, various methods have been applied. Most of these methods are closely related to the already mentioned coding strategies. The aim of this paper is to propose a measure of deviation between two models of ISI, respectively, between a model and data and investigate its properties. We will use tools provided by information theory, employing the Kullback–Leibler (KL) distance for the purpose and show explicitly how the most common ISI descriptors differ from the exponential one using this measure. Also an example on experimental data will be presented to provide a practical illustration.

2 Theory and methods

2.1 Entropy and the Kullback–Leibler distance

The concept of entropy was introduced into statistical information theory by Shannon (1948). For a discrete probability distribution with n possible states, each with a probability $p_i, i \in \{1, \ldots, n\}$, the entropy H is defined as

$$H = -\sum_{i=1}^{n} p_i \ln p_i. \tag{1}$$

The entropy H measures the 'randomness' of the distribution, see Shannon (1948) for details. It is maximized when all p_i values are equal. In information theory the logarithm base is chosen to be 2, as it is related to 'bits' of information. For the simplicity of calculation, we use a natural logarithm. The units are then commonly called 'nats'.

The differential entropy h (Cover and Thomas 1991) of a continuous probability distribution f defined on $[0, \infty)$ is

$$h(f) = -\int_0^\infty f(t) \ln f(t)\, \mathrm{d}t. \tag{2}$$

It is possible to rewrite the Eq. (2) into a form based on the cumulative distribution function, $F(t) = \int_0^t f(x)\, dx$, by using the method of Ling and He (1993), which gives

$$h(f) = -\int_0^1 \ln \frac{\mathrm{d}F(t)}{\mathrm{d}t}\, \mathrm{d}F(t). \tag{3}$$

Both definitions, (2) and (3), play their role from the point of view of numerical estimation in dependency whether the density or the cumulative distribution function is employed.

Though formula (2) is analogous to Eq. (1), differential entropy does not have the same properties and intuitive interpretation as entropy H of a discrete random variable. Namely, h given by formula (2) cannot be used directly to measure the information content of a random variable as it may become negative and depends on the scale of the random variable (Cover and Thomas 1991; Shannon 1948). To overcome these difficulties we measure the information content of a random variable with a density f 'against' some reference state given by another random variable with density g (both defined on $[0, \infty)$) as a KL distance of these two distributions,

$$\mathrm{KL}(f, g) = \int_0^\infty f(t) \ln \frac{f(t)}{g(t)}\, \mathrm{d}t. \tag{4}$$

This approach was proposed by Tarantola (1994); Tarantola and Valette (1982). The reference state described by the probability density function g is termed the 'state of null information' or the 'state of total ignorance' (Jaynes 1968; Tarantola 1994). The quantity $\mathrm{KL}(f, g)$ provides a measure for the information content of f (or information 'gained' from f when knowing g). Though it depends on the template distribution g, it is invariant with respect to a transformation of variable and always non-negative. This quantity can also be interpreted as a 'coding inefficiency' when using distribution g to 'encode' distribution f (more details in Cover and Thomas 1991).

Generally, the KL distance is a measure of the 'closeness' between two probability density functions (Cover and Thomas 1991), though it does not form a measure in the metric sense. It is not symmetrical, however, one can use its symmetric extension (resistor-averaged KL distance), see e.g., Rozell et al. (2004). This proves to be useful when comparing more than two probability densities simultaneously and also when it is not clear what is the template (ideal) distribution to which the others are related.

2.2 The KL distance between the exponential and general model

The exponential distribution plays a key role in neuronal modeling. Its probability density function f_e is

$$f_e(t) = a\exp(-at), \tag{5}$$

where for the parameter $a > 0$ holds $E(T_e) = 1/a$. We see that the firing frequency ν completely characterizes the distribution, $\nu = 1/E(T_e) = a$. For the exponential distribution holds $CV = 1$, independently of the parameter a, but the reverse implication that $CV = 1$ guarantees exponentiality is not true.

One of the most important features of exponential distribution, in the context of information theory, is that among all probability distributions on the real positive half-line with fixed $E(T)$, the exponential distribution maximizes the entropy h. The entropy $h(f_e)$ of the exponential distribution is

$$h(f_e) = 1 - \ln a. \tag{6}$$

Thus, among all possible stationary neuronal discharge activities the one described by the exponential distribution of ISI is the most random. It is natural to choose it as a reference state, the state of null information.

We will measure the information 'gained' when changing from the response state described by the exponential distribution f_e given by Eq. (5) to a response state described by a probability density function f with mean value $E(T)$. For this purpose we employ formula (4), in which by substituting $g = f_e$ we obtain

$$\mathrm{KL}(f, f_e) = aE(T) - \ln a - h(f). \tag{7}$$

If the mean values of f and f_e differ, then the spike trains generally carry different information in the concept of rate coding. However, as we want to analyze the possibilities for a mechanism different from rate coding, let us suppose that the means of f and f_e are equal, i.e., $E(T) = E(T_e) = 1/a$. Then for the KL distance of these two distributions we obtain the following difference in entropies (Soofi et al. 1995)

$$\mathrm{KL}(f, f_e) = h(f_e) - h(f) = 1 + \ln E(T) - h(f). \tag{8}$$

The formula (8) for the information gained when changing from the reference state is also intuitively consistent with the notion of information as a reduction in entropy (Shannon 1948; Borst and Theunissen 1999).

We further precise the interpretation of the KL distance as a measure of information using the notion of mutual information since it is commonly used in the neuronal context. The mutual information $I(S; R)$ (Cover and Thomas 1991) determines the dependence between stimuli S and responses R (Borst and Theunissen 1999). Let the set of stimuli $S = \{s_i\}_{i=1}^{n}$ be discrete and the set of responses continuous $R = \{T, T > 0\}$, where T is an ISI. Mutual information can be formally expressed as $I(S; R) = \sum_i p(s_i)\, i(R|s_i)$, where $i(R|s_i)$ is called the specific information due to the stimulus s_i. Analogously to DeWeese and Meister (1999) we express $i(R|s_i)$ as

$$i(R|s_i) = h(R) - h(R|s_i). \tag{9}$$

It follows from formula (9) that the specific information is large for those stimuli that have only a few different responses associated with them, because $h(R|s_i)$ is the uncertainty in response given a particular stimulus s_i. In the limits, if there is only a single response related to the stimulus s_i, then $h(R|s_i) = -\infty$. We have limited ourselves to the case in which the ISIs are described by a renewal process with probability density function f and the stimuli conditions are constant in time. Under these two assumptions we can assign to each ISI distribution with density f a chosen stimulus s_i. The uncertainty in response given stimulus s_i then becomes $h(R|s_i) = h(f)$, i.e., the smaller the value of $h(R|s_i)$, the greater the information gain due to s_i. The remaining term in formula (9), the marginal entropy $h(R)$, depends on the distribution of stimuli. This distribution affects the absolute value of $i(R|s_i)$; however, the relative encoding efficiency for different s_i's remains unchanged. It is useful to view $h(R)$ as the entropy of the spontaneous neuronal activity, see Chacron et al. (2001) for details. If this activity is described by the Poisson process, then formula (9) corresponds to the expression for the KL distance (8). Furthermore, using formula (8) on two renewal processes described by distributions f_A and f_B yields

$$\mathrm{KL}(f_B, f_e) - \mathrm{KL}(f_A, f_e) = h(f_A) - h(f_B). \tag{10}$$

The KL distance thus also provides a way to determine the specific information in cases when the spontaneous activity is not a Poisson process.

It follows from the definition (4) that the KL distance may sometimes tend to infinity. This is due to the fact that a continuous random variable generally carries an infinite amount of information (van der Lubbe 1997, p. 171). Nevertheless, this fact can be considered as merely formal and without consequences, realizing that in practice we are always working with finite precision on a finite time scale. It also follows from Eq. (8) that the KL distance between any density and the exponential one with the same mean is positive and equal to zero if, and only if, $f = f_e$. This statement of equivalence, in contrast to equality $CV = 1$, allows us to judge exponentiality. In statistical literature the KL distance also appears as one of the approaches to exponentiality testing (Ebrahimi et al. 1992; Choi et al. 2004).

2.3 Evaluation of the KL distance from data

It follows from Eq. (8) that the evaluation of the KL distance from experimental data is reduced to the problem of estimating the entropy of the involved distributions. Estimation of the entropy of exponential distribution is equivalent to the estimation of the mean value $E(T_e)$. Therefore, the only open problem is the estimation of the entropy $h(f)$. Two possible approaches to $h(f)$ estimation exist: the parametric one, where from a preselected model its parameters are estimated and then the entropy is calculated, and the nonparametric one, where $h(f)$ is estimated directly from data without specifying the model.

The first approach has been recently exploited, e.g., by Reeke and Coop (2004), for the case of shifted gamma distribution (with three independent parameters). To illustrate this approach we will estimate the parameters of several common ISI descriptors. The goodness of fit of the data to the models will be checked by the standard one-sample Kolmogorov–Smirnov (KS) test (Gibbons 1971) using the estimated parameters. In the cases when the null hypothesis cannot be rejected, the KL distance will be evaluated using this parametric approach. In general, however, the application of the KL distance method is not conditioned by the KS test.

As regards the second approach, many possibilities to estimate $h(f)$ non-parametrically exist and have been discussed in literature, see e.g. Beirlant et al. (1997) and Tsybakov and Meulen (1996) for an overview. We may divide non-parametric entropy estimators into two groups. The so-called "plug-in" estimators using formula (2) directly, thus some estimate of density has to be constructed, employing, e.g., histograms, kernel-smoothed probability densities and the estimators where the probability density function does not appear explicitly, usually derived using formula (3) in some way. To this group belong e.g., the estimator of Vasicek (Vasicek 1976), or estimators based on nearest sample spacings, see Beirlant et al. (1997). The choice of concrete estimator is strongly situation and purpose dependent. Histogram estimates or smoothed kernel densities are well known to be greatly affected by bin-size resp. bandwidth. The fact that the estimated probability density function has limited support to the real positive half-line may in some cases be the next reason why not to employ kernel-smoothed density estimates. In our case explicit evaluation of probability density function f is not needed so we avoided "plug-in" entropy estimators. To illustrate the non-parametric KL distance evaluation on the experimental data, we use the simple binless entropy estimator of Vasicek (1976), which is known to converge and behave well for various types of data (Ebrahimi et al. 1992; Miller and Fisher III 2003). The examples of estimated entropy values relevant to our calculations are included in Fig. 1. Given n ISIs $\{t_1, t_2, \ldots, t_n\}$, we sort them with respect to their length $\{t_{[1]}, t_{[2]}, \ldots, t_{[n]}\}$. Then the estimated entropy, $h(\mathrm{data})$, is

$$h(\mathrm{data}) = \frac{1}{n}\sum_{i=1}^{n} \ln\left[\frac{n}{2m}(t_{[i+m]} - t_{[i-m]})\right], \tag{11}$$

with a free parameter $m < n/2$ and $t_{[j]} = t_{[1]}$ for $j < 1$ and $t_{[j]} = t_{[n]}$ for $j > n$. Note that estimator (11) is based on discretization of formula (3) using empirical cumulative distribution function. The additional parameter m allows the avoidance of possible numerical problems resulting from such discretization. The relation between n and m was determined by Ebrahimi et al. (1992). For sample sizes $n \geq 200$ holds $m = 13$, which was the value of m we used in the entropy estimation.

3 Results and discussion

This section is divided into two parts. First, we investigate the KL distances of several commonly used ISI distributions from the exponential one under the condition of equal mean values. All the tested distributions are described by two independent parameters related to their mean and variance. It follows from the formula (8), valid under equality of the mean values, that the KL distance should also depend on two parameters. Employing the scaling property of differential entropy (Cover and Thomas 1991, p. 233) together with formula (8) yields that in our case the KL distance does not depend on $E(T)$, as will be shown explicitly in the studied examples. To provide a unified view of the results, we chose coefficient of variation CV as an independent variable. This particular choice also seems to be the most natural for the purpose of insight into possible mechanisms of neural coding mentioned earlier. In the light of our setting, the dependence $\mathrm{KL}(CV)$ allows to judge possible amount of information being transmitted between two particular states of neuronal firing as a function of spike train variability (while the firing frequency is not changing). Secondly, an application of the KL distance on the experimental data is shown. The obtained results are related to those from the first part.

3.1 Gamma distribution

Gamma distribution is one of the most frequent statistical descriptors of ISIs (Hentall 2000; Levine 1991; McKeegan 2002; Mandl 1992). Its probability density function is

$$f(t) = \frac{b^a t^{a-1} \mathrm{e}^{-bt}}{\Gamma(a)}, \tag{12}$$

where $\Gamma(z) = \int_0^\infty t^{z-1}\exp(-t)\,\mathrm{d}t$ is the gamma function and $a > 0$, $b > 0$ are the parameters. From formula (12) these relations follow

$$E(T) = \frac{a}{b}, \quad CV = \frac{1}{\sqrt{a}}. \tag{13}$$

Using formula (2), the entropy of gamma distribution is

$$h(f) = a + (1-a)\Psi(a) - \ln b + \ln\Gamma(a), \tag{14}$$

where $\Psi(a) = \frac{\mathrm{d}}{\mathrm{d}a}\ln\Gamma(a)$ is the digamma function. By substituting Eqs. (13) and (14) in formula (8), we find the KL distance of the gamma distribution from the exponential one in the explicit form as a function of CV

$$\mathrm{KL}(CV) = 1 - \ln CV^2 - \ln\Gamma\left(1/CV^2\right) + \frac{\Psi(1/CV^2) - 1}{CV^2} - \Psi(1/CV^2). \tag{15}$$

This dependence is shown in Fig. 1 together with the estimated values and standard deviations from simulated data. We see that the error of estimation is relatively small and a possible positive bias with respect to true values is negligible for tested sample sizes. The density f given by formula (12) is

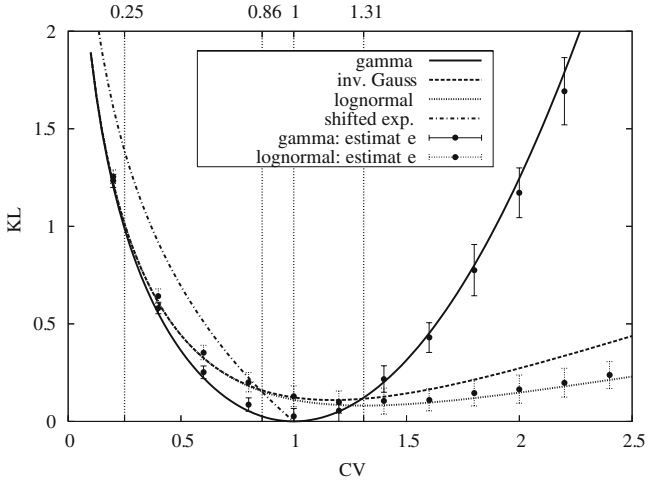


Fig. 1 The Kullback–Leibler (KL) distance as a function of coefficient of variation (CV) for the tested models. Four values of CV (marked on the minor horizontal axis) are chosen to produce Fig. 2. The KL distance of gamma and shifted exponential (available only for $CV \leq 1$) distributions is zero for $CV = 1$, implying that at this point they become exponential. For $CV \to 0$ and for $CV \to \infty$ the KL distance tends to infinity for the tested distributions. With CV increasing from zero the KL distances of lognormal, inverse Gaussian and gamma distributions are initially the same. Then gamma branches off at $CV \approx 0.25$ and the lognormal and inverse Gaussian depart at $CV \approx 1$. For lognormal and inverse Gaussian distributions the distance never reaches zero and even the minima are not located at $CV = 1$ implying that for no combination of parameters these distributions become exponential. A common feature of the tested models is that near $CV = 1$ the values of KL distances are generally low, however, minimum for the lognormal, resp. inverse Gaussian distribution is located at $CV \approx 1.31$, resp. $CV \approx 1.17$. Furthermore, for these two distributions, especially for the lognormal, the KL distance grows very slowly with increasing CV, compared to the gamma distribution. Estimated values from simulated data with sample sizes $n = 500$ and their standard deviations over 100 trials are shown for comparison with theoretical results

exponential for $a = 1$ and therefore $\mathrm{KL}(CV = 1) = 0$. The KL distance tends to infinity for $CV \to 0$ and $CV \to \infty$. We can see from the figure that $\mathrm{KL}(CV)$ increases rapidly for $CV > 1$ especially if compared to the other models presented here. For $CV < 0.25$ (approximately) the KL distances of gamma, lognormal and inverse Gaussian distributions become the same. Probability densities of investigated models for the values of CV selected with respect to results in Fig. 1 and with mean ISI equal to one are presented in Fig. 2. We can see that the gamma distribution ranges from the shape similar to the Gaussian distribution at $CV = 0.25$ through the "typical" shape of gamma distribution at $CV = 0.86$ to the exponential distribution at $CV = 1$. Finally, for $CV > 1$ the gamma distribution is characterized by a majority of very short intervals and long tail of the distribution. This feature, at least at the first approximation, looks like bursting type of neuronal activity.

The KL distance provides a different approach to the comparison of two distributions if compared with the KS test: it is 'global', i.e., it takes into account the whole shapes of the compared density functions, while the key parameter for the KS method is the extreme local 'distance' of the curves. Thus, one can construct cases for which the results given by KL and KS are contradictory. The most striking example is when the template distribution is zero on some interval where the other function is not. Then the KL distance is infinite, while the KS statistics may be even an infinitely small one. Thus, KS represents an alternative to KL, however, with completely different properties.

3.2 Inverse Gaussian distribution

The inverse Gaussian distribution (Chhikhara and Folks 1989) is often used to describe neural activity (Iynegar and Liaom 1997) and fitted to experimentally observed ISIs (Gerstein and Mandelbrot 1964; Levine 1991). This distribution results from the Wiener process with positive drift (the depolarization has a linear trend to the threshold) as a stochastic neuronal model (Berger et al. 1990; Berger and Pribram 1992; Levine 1991). The probability density of inverse Gaussian distribution can be expressed as a function of two parameters $a > 0$ and $b > 0$

$$f(t) = \sqrt{\frac{a}{2\pi b t^3}} \exp\left[-\frac{1}{2b}\frac{(t-a)^2}{at}\right], \tag{16}$$

with

$$E(T) = a, \quad CV = \sqrt{b}. \tag{17}$$

The fact that the mean and coefficient of variation can be easily related to the parameters a, b, which act as a scale and shape parameter, respectively, of the curve, is of practical advantage.

The KL distance of the inverse Gaussian distribution from the exponential one as a function of CV is

$$\mathrm{KL}(CV) = \frac{1}{2}\ln\frac{\mathrm{e}}{2\pi} - \ln CV + \frac{3}{\sqrt{2\pi}}\frac{\mathrm{e}^{1/CV^2}}{CV} K^{(1,0)}_{\frac{1}{2}}(1/CV^2), \tag{18}$$

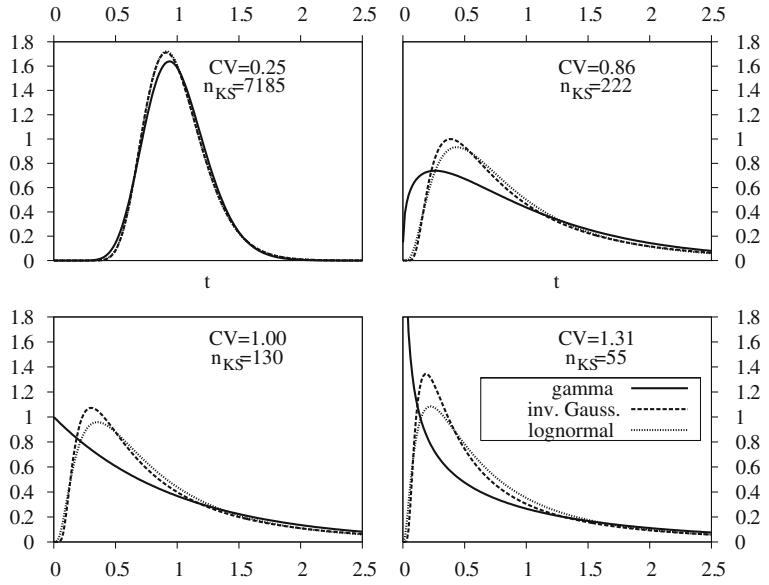


Fig. 2 Probability density functions $f(t)$ of the gamma, inverse Gaussian and lognormal distributions with $E(T) = 1$ s for four different values of CV indicated in Fig. 1. The lowest number n_{KS} of interspike intervals (ISIs), required for the Kolmogorov–Smirnov (KS) test to "distinguish" between gamma and inverse Gaussian distributions at 5% significance level is given. We can see that for $CV = 0.25$ the three probability density functions are hardly distinguishable. The number n_{KS} is too large compared to usual sample sizes obtained in the experiments. The probability density functions start to differ more for $CV = 0.86$ (this is the value, for which the KL distances of lognormal, inverse Gaussian and shifted exponential distribution approximately equal), n_{KS} falls into the average length of experimental records. At $CV = 1$ the gamma distribution becomes exponential (KL $= 0$), while the lognormal and inverse Gaussian do not, and they are even not "close" to exponential as much as possible. The minimal KL distance of lognormal distribution (corresponding to its maximum "closeness" to exponential distribution) is at $CV = 1.31$. At this point the KL distance of inverse Gaussian and gamma distributions is roughly the same, though their probability densities differ strikingly (n_{KS} is merely 55)

where $K_\nu^{(1,0)}(z)$ is the derivative of the modified Bessel function of the second kind (Abramowitz and Stegun 1972)

$$K_\nu^{(1,0)}(z) = \frac{\partial}{\partial \nu} K_\nu(z). \tag{19}$$

The dependence is shown in Fig. 1. Due to the fact that neither for any combination of parameters nor asymptotically the inverse Gaussian distribution is exponential, the KL distance is not zero for $CV = 1$. This fact can be interpreted in this way, following formula (8): even if there is no change in $E(T)$ and $CV = 1$ there still may be some gain of information coming from the reference state (described by exponential distribution). Following Fig. 1 we can judge the information gain under the condition of fixed CV. The minimum of KL(CV) for the inverse Gaussian distribution is located at $CV \approx 1.173$. We can see a difference here, compared to the previous case of the gamma distribution. It has been already noted that the condition $CV = 1$ does not imply exponentiality, but in this case even the minimal distance of the inverse Gaussian to exponential distribution is not located at $CV = 1$. In Fig. 2 there are ISI probability densities for four selected values of CV. The important fact is that for $CV = 0.25$ the models are practically indistinguishable. Even at $CV = 0.86$ and $CV = 1$, lognormal and inverse Gaussian distributions are of very similar shape. Finally, compared to gamma, inverse Gaussian lacks very short ISIs which can be considered as positive feature of this distribution, if refractoriness should be taken into account.

3.3 Lognormal distribution

The lognormal distribution of ISI, with some exceptions (Bershadskii et al. 2001), is rarely presented as a result of a neuronal model. However, it represents quite a common descriptor in ISI data analysis (Levine 1991), e.g., a mixture of two lognormal distributions has been used recently (Bhumbra et al. 2004).

The lognormal distribution is given by the probability density function

$$f(t) = \frac{1}{t\sigma\sqrt{2\pi}} \exp\left[-\frac{(\ln t - m)^2}{2\sigma^2}\right], \tag{20}$$

where m and $\sigma > 0$ are the parameters. As the variable $\ln T$ is normally distributed, it follows

$$E(T) = \exp(m + \sigma^2/2), \quad CV = \sqrt{\exp\sigma^2 - 1}. \tag{21}$$

We use formula (8) to compute the KL distance of lognormal distribution from the exponential one. Expressing the KL distance as a function of CV, we come to a formula

$$\mathrm{KL}(CV) = \frac{1}{2}\left[\ln\frac{CV^2+1}{\ln(CV^2+1)} + \ln\frac{\mathrm{e}}{2\pi}\right]. \tag{22}$$

From there it follows that

$$\frac{\mathrm{d(KL)}}{\mathrm{d}(CV)} = \frac{CV[\ln(CV^2+1) - 1]}{(CV^2+1)\ln(CV^2+1)} \tag{23}$$

and the minimum is at $CV = \sqrt{\mathrm{e}-1} \approx 1.311$. The dependence is shown in Fig. 1 together with the estimated values and standard deviations from simulated data, which shows good correspondence between theoretical and numerical approaches. The estimate is not systematically biased and the relative error is very small. Again, as in the case of inverse Gaussian distribution, we see that even if $CV = 1$ the distribution is not exponential. Yet again – the minimal possible deviation of lognormal distribution from exponential one is not at $CV = 1$. It is interesting that for $CV < 1$ (approximately) there is no difference in lognormal and inverse Gaussian distributions from the perspective of the KL distance. The equality in the KL distance, of course, does not imply that these distributions are identical, see Fig. 2. Nevertheless, their similarity is very high.

3.4 Distributions involving a refractory period

The refractory phase is such a state of a neuron, coming immediately after a spike was generated, when it is impossible for another spike to be emitted. In more detail, one can distinguish the absolute refractory phase, when the generation of the next spike is absolutely impossible and the relative refractory phase, when it is merely not probable. The typical duration of the absolute refractory phase is around 2–4 ms, while the relative one may last around 10–20 ms, depending on the definition of "not probable", Gerstner and Kistler (2002). The discussion on the topic of refractory phases and their importance is still ongoing (Berry II and Meister 1998).

Recently, a shifted gamma distribution was used as a generally suitable ISI probability distribution for parametric entropy estimation by Reeke and Coop (2004). The absolute refractory phase is, in this model, described by a shift in time for $\tau > 0$, while parameters a, b are kept the same as in Eq. (12). Correspondingly, $E(T)$ and CV change in a simple way but the entropy of such distribution is independent of τ as follows from formula (2). In the case of shifted gamma distribution, we have three independent parameters and it is no longer possible to describe the KL distance from the exponential model just as a function of one parameter (CV). On the other hand, comparing the shifted gamma model to the exponential distribution shifted by the same value τ gives naturally the same results for $\mathrm{KL}(CV)$ as given by Eq. (15). The same is true for any shifted distribution. However, it might be interesting to compare two exponentials; one with and the other one without refractory phase, as follows.

The probability density function of the shifted exponential distribution with parameter $a > 0$ and a shift $\tau \geq 0$ is

$$f(t) = \begin{cases} 0, & t \leq \tau \\ a\mathrm{e}^{-a(t-\tau)}, & t > \tau. \end{cases} \tag{24}$$

Then it follows

$$E(T) = \frac{1+a\tau}{a}, \quad CV = \frac{1}{1+a\tau}. \tag{25}$$

It is obvious that in this case it is always $CV < 1$ for $\tau > 0$. The evaluation of the KL distance of the shifted exponential distribution from the exponential one under the condition of the same means using Eqs. (8) and (25) gives

$$\mathrm{KL}(CV) = -\ln(CV). \tag{26}$$

For $\tau = 0$ is $CV = 1$ and f given by formula (24) is exponential which is confirmed by $\mathrm{KL} = 0$. The function $\mathrm{KL}(CV)$ is shown in Fig. 1.

The shape of $\mathrm{KL}(CV)$ given by Eq. (26) differs from the KL distances of gamma, inverse Gaussian and lognormal models with the corresponding CV. We see that this curve, despite taking the value of zero at $CV = 1$, is the steepest among all of them. Thus we can ask, what is the critical value of CV such that for smaller values the KL distance of the shifted exponential distribution is greater than that of any other tested model? Following Fig. 1 we estimate the critical value of CV as the one for which the KL distances of the lognormal and shifted exponential distributions equal. This results in (a very nice) equation

$$\frac{(CV^2+1)CV^2}{\ln(CV^2+1)} = \frac{2\pi}{\mathrm{e}}, \tag{27}$$

which yields $CV \approx 0.86$. Realizing that

$$CV = 1 - \frac{\tau}{E(T)} \tag{28}$$

and using the critical value of CV, we receive a critical ratio of the refractory phase to the mean value as $\tau/E(T) \approx 14.4\,\%$. For ratios greater than the critical the KL distance of the shifted exponential distribution is greater than that of other models tested here. For example, if the average absolute refractory phase is 3 ms, the corresponding critical mean value is approximately 21 ms.

3.5 Experimental data

The data comes from extracellular recordings made from olfactory receptor neurons of freely breathing and tracheotomized rats using glass insulated tungsten microelectrodes. Spontaneous, single-unit action potentials were recorded using metal-filled glass micropipettes filled with an alloy of Wood's metal (80%) and indium (20%). The single unit nature of the recorded spikes was controlled during the experiment by triggering the recorded neuron near the background-noise on a storage oscilloscope. More details of the data acquisition are described in Duchamp-Viret et al. (2003). Here, we do not distinguish between the data that come from freely breathing or tracheotomized rats. A comparison of these two conditions are published elsewhere (Duchamp-Viret et al. 2005). The sample sizes range from (circa) $n = 100$ to $n = 2000$ and all records have been tested for non-stationarity (the Wald–Wolfowitz test, serial correlation, periodogram). We use these recordings to illustrate the estimation of the KL distance. Both parametric and non-parametric methods were applied. The results are summarized in Fig. 3, where estimated KL distances are plotted in dependency on the CV along with the theoretical curves of KL distances re-plotted from Fig. 1. Two different categories of data are

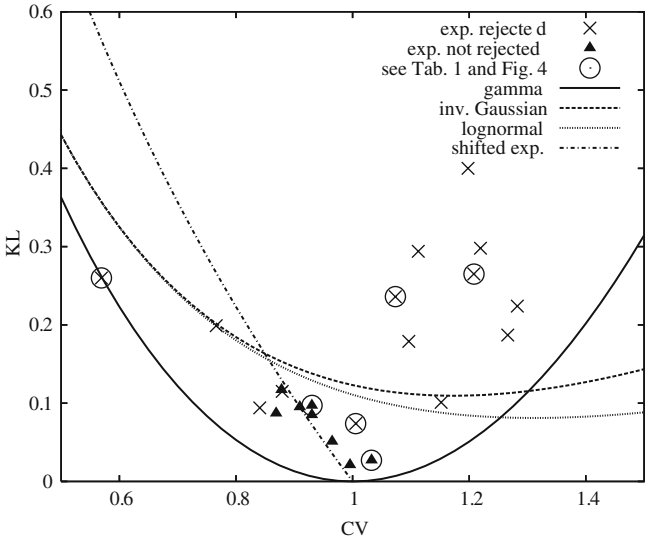


Fig. 3 The KL distance as a function of CV for the experimental data. Theoretical curves of the tested models are re-plotted from Fig. 1. Two different categories of data are distinguished based on the KS test of exponentiality at 5% level of significance. Spike trains and ISI histograms of the *encircled data* are shown in Fig. 4 and those with CV close to unity are given a more detailed treatment in Table 1. We see that even for $CV \approx 1$ there exist data that are not exponential. The points around the value $CV = 1$, where the hypothesis of exponentiality was not rejected, are asymmetric, closely following the theoretical curve for the shifted exponential distribution. Generally, we see that the data obey the rule indicated in Fig. 1 for the common ISI descriptors, i.e., the smallest values of the KL distance are around $CV = 1$. For $CV > 1$ the general "course" of the data is even steeper than that of gamma distribution

distinguished, exponentiality rejected or not rejected, based on the Kolmogorov–Smirnov test at 5% significance level. We can see that even if $CV \approx 1$ the exponentiality may be rejected. We note that the data obey the general feature of the tested models: the small values of the KL distance are distributed around $CV = 1$. Of particular interest is the asymmetrical distribution of possibly exponential data around $CV = 1$. Except for one spike train all records for which exponentiality is not rejected have $CV < 1$. Furthermore, these points (filled triangles) closely follow the theoretical curve for the shifted exponential distribution. Of course from this fact we cannot conclude that these ISIs follow this distribution.

For $CV > 1$ the calculated KL distances are far above all considered curves. Though the KL distance does not tell us which particular distribution to use, the difference between theoretical and estimated KL values suggests that the gamma distribution cannot describe the data well. We can expect that analogously to olfactory neurons in frogs (Rospars et al. 1994), there is a bursting character in this activity. The bursting activity of the neuron is usually described by a mixture of two distributions, one for interburst ISIs and the other for intraburst ISIs (Mandl 1992). For example Bhumbra et al. (2004) combines two lognormal distributions given by formula (20), which results in a very flexible model with five unknown parameters. An alternative, and more common, description could be a probability density function f of the mixture of two exponential distributions

$$f(t) = pa\mathrm{e}^{-ax} + (1-p)b\mathrm{e}^{-bx} \tag{29}$$

with $p \in (0,1)$ and $a > 0$, $b > 0$, $a \neq b$. We can ask, whether model (29) can be as far from the single exponential model as are the data in Fig. 3. To check this idea we first compute $E(T)$ and CV of the distribution (29)

$$E(T) = \frac{pb + (1-p)a}{ab},$$
$$CV = \sqrt{\frac{2pb^2 + 2(1-p)a^2}{(pb + (1-p)a)^2} - 1}. \tag{30}$$

Expressing p and b from Eqs. (30), we re-parameterize the original formula (29) using parameters $E(T)$, CV and a.

Table 1 Comparison of two pairs of spike trains (encircled in Fig. 3 with $CV \approx 1$)

Filename	Exp.	n	$\nu\,[\mathrm{s}^{-1}]$	CV	$\mathrm{KL}_{\mathrm{data}}$	KL_{γ}	$\mathrm{KL}_{\mathrm{shift}}$
RN20_01	−	1906	6.862	1.005	0.074	–	–
RT48_01	+	355	1.988	1.032	0.027	0.001	–
RT58_01	−	788	5.246	1.073	0.236	–	–
RN30_01	+	549	1.963	0.930	0.097	0.006	0.073

The Kullback-Leibler (KL) distances were estimated both from data (non-parametrically) and from theoretical models (parametrically). The "−" sign in the "Exp." column indicates that the hypothesis of exponentiality was rejected by the Kolmogorov–Smirnov (KS) test, while "+" states that it was not rejected (within a 5% significance level). n is the number of interspike intervals (ISIs) in the record, ν the firing frequency, CV is the computed coefficient of variation, $\mathrm{KL}_{\mathrm{data}}$ is estimated directly from data (non-parametrically), KL_{γ} is the corresponding distance of the gamma distribution and $\mathrm{KL}_{\mathrm{shift}}$ of the shifted exponential. We see that the KL distances computed parametrically and non-parametrically are rather different. This can be attributed both to wide confidence intervals of estimated parameters and to the behavior of Vasicek's estimator

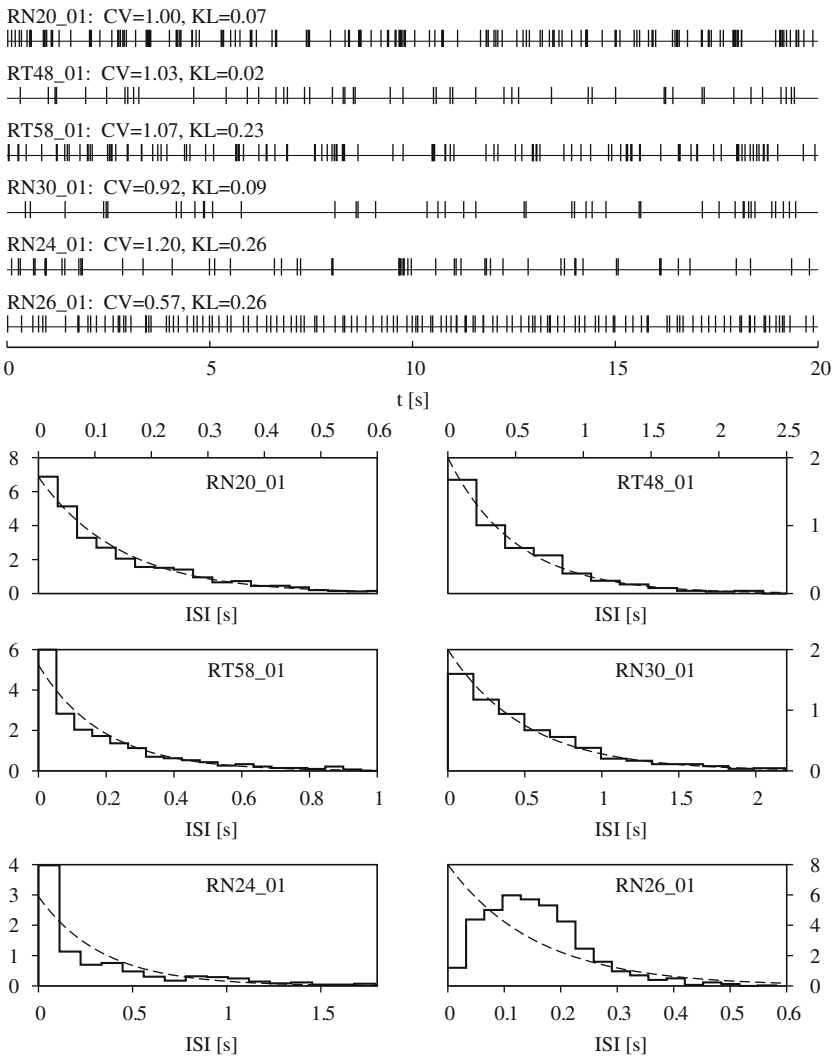


Fig. 4 First 20 s of the *spike trains* and ISI *histograms* (of the whole record) for the selected data encircled in Fig. 3 compared with exponential probability density with the corresponding mean ISI. The values of CV and KL distance estimated from data are also indicated. The *first four records* with CV closest to unity are explored in Table 1. The record RN24_01 was chosen to represent the group of data with $CV > 1$ and relatively large KL distance. The sharp increase in the frequency of very small ISI values together with relatively flat tail of the histogram and large KL distance value favors the double exponential model as the more probable ISI descriptor compared to the gamma model. On the other hand, the record RN26_01 was chosen for comparison only due to its small value of CV. Except RT48_01 all records have $n > 500$

Though the expressions grow in size quickly and get difficult to handle analytically, we can evaluate the KL distance $\mathrm{KL}(E(T), CV, a)$ from formula (8) numerically. The results for $CV = 1.2$ show that for any $E(T)$ we can find a value of parameter a such that the KL distance of the double exponential is greater than the KL distance of gamma distribution with the same mean and CV. Though it is not possible to fit the theoretical distributions to the data based solely on the KL number, we deduce that the double exponential distribution has a chance to describe the bursting behavior better than gamma distribution.

We chose two pairs of data sets with CV close to unity to show the situation in more detail. The results are summarized in Table 1. The parameters of the gamma, inverse Gaussian, lognormal and shifted exponential were estimated from the data. Then the Kolmogorov–Smirnov test was performed to test the goodness of fit at a 5% significance level. In the case of not rejecting the null hypothesis, the KL distance was estimated parametrically using the previous theoretical results. (Note that for all of these four data sets the inverse Gaussian and lognormal models were rejected.) Though all four data sets have $CV \approx 1$ there are differences in the KL distance from the exponential distribution. One may note that $\mathrm{KL}_{RN20\_01} > \mathrm{KL}_{RT48\_01}$ even though $|1 - CV_{RN20\_01}| < |1 - CV_{RT48\_01}|$. The spike trains and corresponding ISI histograms of the four above-mentioned data sets (together with two other records also encircled in Fig. 3) are shown for comparison in Fig. 4, each plotted together with the exponential distribution with the corresponding mean ISI.

4 Conclusions

The (KL) distance was proposed as a measure of similarity between two (ISI) distributions. Choosing the exponential one as the template we analyzed the KL distance from both data and models. We selected four common two-parametric distributions: gamma, lognormal, inverse Gaussian and shifted exponential.

Fixing the mean values of exponential and model distribution (or data) to be equal, the KL distance corresponds to information gain coming from a state described by the exponential distribution of ISI to another state described by the model distribution. Thus, it reveals a different mechanism from rate coding of information transmission. Furthermore, the KL distance is interpreted in terms of specific information and its usefulness when determining the efficiency of the stimulus encoding is shown. It is natural to express the information gain as a function of a spike train variability, commonly reflected by coefficient of variation (CV). For exponential distribution $CV = 1$, however, the reverse implication does not hold. The KL distance offers an alternative tool to CV to judge exponentiality of the model distribution (or data), because the exponentiality is guaranteed if, and only if, the distance is zero. Furthermore, while there are tools to measure the variability of spike trains, the KL distance measures a different characteristics – the randomness of the underlying process.

The following inference can be made on the basis of the KL distance of ISIs distributions:

1. Even if neither spike frequency nor coefficient of variation changes, the KL distances to the exponential distribution can be different for different models (data) and thus there is still a gain of information coming from one state to another.
2. It is well known that the lognormal and inverse Gaussian distributions never become exponential, but surprisingly their minimal KL distances to this distribution are not located at $CV = 1$.
3. For CV increasing from zero (regular spiking) the KL distances of lognormal, inverse Gaussian and gamma distributions are initially the same. Then gamma branches off at $CV \approx 0.25$ and the lognormal and inverse Gaussian depart at $CV \approx 1$. For low values of CV the differences among the distributions are hardly distinguishable for usual sample sizes available in neural spiking data studies.
4. For lognormal and inverse Gaussian distributions, the KL distance grows very slowly for $CV > 1$, compared to the gamma distribution, and their distances to the exponential distribution are practically the same for $CV = 1$ as for $CV < 2$.
5. The KL distance of shifted exponential ($CV < 1$) is the steepest from all of the investigated alternatives.
6. As shown in the experimental data, even if $CV \approx 1$, the ISI distribution may not be exponential, and this is confirmed by the Kolmogorov–Smirnov test. Although the data follow the general features indicated by theoretical results, the "course" of their KL distance for $CV > 1$ is steeper even than that of gamma distribution. This suggests/confirms the bursting character of this data.
7. The occurrence of data, where exponentiality cannot be rejected, is asymmetric around $CV = 1$ and closely follows the theoretical curve for the shifted exponential distribution.

Acknowledgements Helpful conversations with Priscilla E. Greenwood and Jean-Pierre Rospars are acknowledged. The authors wish to thank Patricia Duchamp-Viret and Michel Chaput for making their experimental data available. The authors are grateful to Nick Dorrel for linguistic correction and to an anonymous referee for comments which improved the quality of the manuscript. This work was supported by the Research project AV0Z 5011922, Center for Neuroscience LC554 and by Academy of Sciences of the Czech Republic Grant (Information Society, 1ET400110401).

References

Abramowitz M, Stegun IA (1972) Handbook of mathematical functions. Dover, New York
Adrian ED (1928) The basis of sensation. W. W. Norton, New York
Beirlant J, Dudewicz EJ, Gyorfi L, van der Meulen EC (1997) Nonparametric entropy estimation: an overview. Int J Math Stat Sci 6:17–39
Berger D, Pribram K, Wild H, Bridges C (1990) An analysis of neural spike train distributions: determinants of the response of visual cortex neurons to changes in orientation and spatial frequency. Exp Brain Res 80:129–134
Berger D, Pribram K (1992) The relationship between the gabor elementary function and a stochastic model of interspike interval distribution in the responses of visual cortex neurons. Biol Cybern 67:191–194
Bershadskii A, Dremencov E, Fukayama D, Yadid G (2001) Probabilistic properties of neuron spiking time series obtained in vivo. Eur Phys J B 24:409–413
Berry II MJ, Meister M (1998) Refractoriness and neural precision. J Neurosci 6:2200–2211
Bhumbra GS, Inyushkin AN, Dyball REJ (2004) Assessment of spike activity in the supraoptic nucleus. J Neuroendocrin 16:390–397
Borst A, Theunissen FE (1999) Information theory and neural coding. Nature Neurosci 2:947–957
Buracas GT, Albright TD (1999) Gauging sensory representations in the brain. Trends Neurosci 22:303–309
Chacron MJ, Longtin A, Maler L (2001) Negative interspike interval correlations increase the neuronal capacity for encoding time-dependent stimuli. J Neurosci 21:5328–5343
Chhikhara RS, Folks JL (1989) The inverse Gaussian distribution: theory, methodology and applications. Marcel Dekker, New York
Choi B, Kim K, Song SH (2004) Goodness-of-fit test for exponentiality based on Kullback-Leibler information. Commun Stat-Simul C 33(2):525–536
Christodoulou C, Bugmann G (2001) Coefficient of variation vs. mean interspike interval curves: what do they tell us about the brain? Neurocomputing 38–40:1141–1149
Cover TM, Thomas JA (1991) Elements of information theory. Wiley, New York
Cox DR, Lewis PAW (1966) The statistical analysis of series of events. Wiley, New York
DeWeese MR, Meister M (1999) How to measure the information gained from one symbol. Network: Comput Neural Syst 10:325–340
Duchamp-Viret P, Duchamp A, Chaput MA (2003) Single olfactory sensory neurons simultaneously integrate the components of an odor mixture. Eur J Neurosci 10:2690–2696
Duchamp-Viret P, Chaput MA, Kostal L, Lansky P, Rospars J-P (2005) Patterns of spontaneous activity in single rat olfactory receptor

neurons are different in normally breathing and tracheotomized animals. J Neurobiol 65:97–114
Ebrahimi N, Habibullah M, Soofi ES (1992) Testing exponentiality based on Kullback-Leibler Information. J Stat Soc B 3:739–748
Gerstein G, Mandelbrot B (1964) Random walk models for the spike activity of a single neuron. Biophys J 4:41–68
Gerstner W, Kistler W (2002) Spiking neuron models. Cambridge University Press, Cambridge
Gibbons JD (1971) Nonparametric Statistical Inference. McGraw-Hill, New York
Han YMY, Chan YS, Lo KS, Wong TM (1998) Spontaneous activity and barosensitivity of the barosensitive neurons in the rostral ventrolateral medulla of hypertensive rats induced by transection of aortic depressor nerves. Brain Res 813:262–267
Hentall ID (2000) Interactions between brainstem and trigeminal neurons detected by cross-spectral analysis. Neuroscience 96:601–610
Holt GR, Softky WR, Koch C, Douglas RJ, (1996) Comparison of discharge variability in vitro and in vivo in cat visual cortex neurons. J Neurophysiol 75:1806–1814
Iynegar S, Liaom Q (1997) Modeling neural activity using the generalized inverse Gaussian distribution. Biol Cybern 77:289–295
Jaynes ET (1968) Prior probabilities. IEEE Trans Syst Sci Cybern 3:227–241
Johnson DH, Glantz RM (2004) When does interval coding occur? Neurocomputing 58–60:13–18
Lansky P, Rodriguez R, Sacerdote L (2004) Mean instantaneous firing frequency is always higher than the firing rate. Neural Comput 16:477-489
Levine MW (1991) The distribution of the intervals between neural impulses in the maintained discharges of retinal ganglion cells. Biol Cybern 65:459–467
Lewis CD, Gebber GL, Larsen PD, Barman SM (2001) Long-term correlations in the spike trains of medullary sympathetic neurons. J Neurophysiol 85:1614–1622
Ling YB, He B (1993) Entropic analysis of biological growth models. IEEE Trans Bio-Med Eng 40:48–54
van der Lubbe JCA (1997) Information theory. Cambridge University Press, Cambridge
Mandl G (1992) Coding for stimulus velocity by temporal patterning of spike discharges in visual cells of cat superior colliculus. Vision Res 33:1451–1475
McKeegan DEF (2002) Spontaneous and odor evoked activity in single avian olfactory bulb neurons. Brain Res 929:48–58
Miller EG, Fisher III JW (2003) ICA using space estimates of entropy. J Mach Learn Res 4:1271–1295
Nemenman I, Bialek W, de Ruyter van Steveninck RR (2004) Entropy and information in spike trains: progress on the sampling problem. Phys Rev E 69:056111-1–056111-6
Paninski L (2003) Estimation of entropy and mutual information. Neural Comput 15:1191–1253
Perkel DH, Gerstein GL, Moore GP (1967) Neuronal spike trains and stochastic point processes. Biophys J 4:391–418
Rieke F, Warland D, de Ruyter van Steveninck RR, Bialek W (1997) Spikes: exploring the neural code. MIT Press, Cambridge
Reeke NR, Coop AD (2004) Estimating the temporal entropy of neuronal discharge. Neural Comput 16:941–970
Rospars J-P, Lansky P, Vaillant J, Duchamp-Viret P, Duchamp A (1994) Spontaneous activity of first- and second-order neurons in the frog olfactory system. Brain Res 662:31–44
Rozell CJ, Johnson DH, Glantz RM (2004) Measuring information transfer in the spike generator of crayfish sustaining fibers. Biol Cybern 90:89–97
Ruskin DN, Bergstrom D A, Walters JR (2002) Nigrostriatal lesion and dopamine agonists affect firing patterns of rodent entopeduncular nucleus neurons. J Neurophysiol 88:487–496
Shannon C (1948) A mathematical theory of communication. AT&T Tech J 27:379–423
Shinomoto S, Yutaka S, Hiroshi O (2002) Recording site dependence of the neuronal spiking statistics. Biosystems 67:259–263
Shinomoto S, Shima K, Tanji J (2003) Differences in spiking patterns among cortical neurons. Neural Comput 15:2823–2842
Soofi ES, Ebrahimi N, Habibullah M (1995) Information distinguishability with application to analysis of failure data. J Am Stat Assoc 90:57–668
Stein RB (1967) The information capacity of nerve cells using a frequency code. Biophys J 7:797–826
Stevens CF, Zador A (1995) Information through a spiking neuron. Adv Neural Inf Proc Syst 8:75–81
Strong SP, Koberle R, de Ruyter van Steveninck RR, Bialek W (1998) Entropy and information in spike trains. Phys Rev Lett 80:197–200
Tarantola A (1994) The inverse problem theory. Elsevier, Amsterdam
Tarantola A, Valette B (1982) Inverse problems = quest for information. J Geophys 50:159–170
Tsybakov AB, van der Meulen EC (1996) Root-*n* consistent estimators of entropy for densities with unbounded support. Scand J Stat 23:75–83
Vasicek O (1976) A test for normality based on sample entropy. J Stat Soc B 38:54–59
Zador A (1998) Impact of synaptic unreliability on the information transmitted by spiking neurons. J Neurophysiol 79:1219–1229

Network: Computation in Neural Systems
June 2006; 17: 193–210

Taylor & Francis
Taylor & Francis Group

Classification of stationary neuronal activity according to its information rate

LUBOMIR KOSTAL & PETR LÁNSKÝ

Institute of Physiology, Academy of Sciences of the Czech Republic, Videnska 1083, 142 20 Prague 4, The Czech Republic

(Received 6 July 2005; revised 21 November 2005; accepted 23 January 2006)

Abstract
We propose a measure of the information rate of a single stationary neuronal activity with respect to the state of null information. The measure is based on the Kullback–Leibler distance between two interspike interval distributions. The selected activity is compared with the Poisson model with the same mean firing frequency. We show that the approach is related to the notion of specific information and that the method allows us to judge the relative encoding efficiency. Two classes of neuronal activity models are classified according to their information rate: the renewal process models and the first-order Markov chain models. It has been proven that information can be transmitted changing neither the spike rate nor the coefficient of variation and that the increase in serial correlation does not necessarily increase the information gain. We employ the simple, but powerful, Vasicek's estimator of differential entropy to illustrate an application on the experimental data coming from olfactory sensory neurons of rats.

Keywords: *Spike train, stationary activity, information rate, Markov chain*

Introduction

It is generally accepted that the information in neuronal systems is transferred using the series of action potentials–the spike trains. There are two main hypotheses that attempt to classify possible ways in which the spike train may carry information: the frequency (rate) codes and the temporal spike codes (Theunissen & Miller 1995; Gerstner & Kistler 2002). Both hypotheses rely on the important assumption that single spikes are mutually indistinguishable. Therefore the spike trains can be considered as a time series of point events. This is a widely accepted simplification that allows further analysis, especially from the information-theoretic point of view.

The idea of frequency coding was proposed by Adrian (1928). He showed that the number of spikes per some time window (spike frequency) non-linearly increases with the increase of the stimulus intensity. The idea of temporal spike coding (Perkel & Bullock 1968; Theunissen & Miller 1995), on the other hand, employs the timing of the spikes or their particular temporal pattern. This is a consequence of the fact that there are situations where time averaging is not possible. Also the high time variability in neuronal discharge, particularly within the cortex, see, e.g., Buracas & Albright (1999), may play some role and cannot

Correspondence: Lubomir Kostal, Institute of Physiology, Academy of Sciences of the Czech Republic, Videnska 1083, 142 20 Prague 4, The Czech Republic. Tel: +420 2 4106 2276; Fax: +420 2 4106 2488; E-mail: kostal@biomed.cas.cz

ISSN: 0954-898X print / ISSN 1361-6536 online
DOI: 10.1080/09548980600594165

be simply "averaged out". The extreme case, which gives an upper bound on the possible amount of information encoded (Zador 1998), follows from the hypothesis that the precise value of each successive interspike interval (ISI) carries the information (in other words the information gain is maximized if there is no noise in the transmission). Though frequency codes and temporal spike codes are shown to be compatible in many cases (Gerstner & Kistler 2002), it is clear that infinitely many different spike records may have the same frequency.

If the the neuronal firing is stationary, the mean spike frequency (the inverse of the mean ISI, Lánský et al. (2004) carries the information from the frequency coding hypothesis point of view. However, as has been shown by Bialek et al. (1991) (see overview in Theunissen & Miller (1995)), the frequency codes in single neurons carry information about dynamic stimuli, which results in a non-stationary neuronal signal. The temporal coding, on the other hand, has been shown to occur almost exclusively under steady-state stimulus conditions (Fuller & Looft 1984; Middlebrooks et al. 1994).

At the first approximation, neuronal firing under steady-state conditions is often described as a renewal process. This is justified by the observation that the emitted spike resets the membrane potential of the neuron's body independently of the preceding synaptic processes (Abeles 1982). In this case, ISIs are described as independent realizations of a positive random variable T. The corresponding probability density function $f(t)$ is thus the complete descriptor of such neuronal activity with the expected value of T (the mean ISI) denoted by $E(T)$.

Most of the time, however, though the neuronal firing is stationary there is a dependency structure between the observed ISIs (Longtin & Racicot 1996; Chacron et al. 2001). The dependence may arise due to incomplete resetting of the membrane potential after the spike is emitted, which is experimentally observed especially in the distal parts of the neuron (Abeles 1982). Such a type of neuronal firing is not a renewal process. The successive ISIs $\{T_i\}$ are then statistically dependent with an expected value $E(T) = E(T_i)$. The activity is fully described by the joint probability density function $f(t_1, t_2, \ldots)$, see, e.g., Cox & Lewis (1966) for details. The importance of the dependence in the ISI structure is also reflected in recent efforts to include the effect of serial correlation into neuronal models (Sakai et al. 1999; Lánský & Rodriguez 1999; Lindner 2004). Despite this effort, we are not aware of any models where analytical results are available, except those vaguely mentioned in Lawrance (1972, p. 215) for the model developed by Lampard (1968).

In this article, we will try to quantify the information encoded by the temporal coding scheme, i.e., under the assumption of stationary stimuli conditions. The information theory, introduced by Shannon (1948) provides the mathematical basis for the task.

Theory and methods

Without loss of generality, information may be defined as a decrease in uncertainty (Shannon 1948). For discrete random variables, the entropy, H, measures the uncertainty and is thus closely related to the notion of information. The quantity that measures the uncertainty of a continuous random variable T with a probability density function $f(t)$ is called differential entropy $h(f)$, also denoted $h(T)$ (Cover & Thomas 1991):

$$h(f) = -\int_{-\infty}^{\infty} f(t) \ln f(t)\, dt. \tag{1}$$

Differential entropy $h(f)$ does not share the same properties and intuitive interpretation as the entropy H (Shannon 1948). Namely, it can be negative and its value changes with coordinate transforms. Thus, it cannot be used as an absolute measure of the information content. Nevertheless, the most "random" distribution is still the one that maximizes h under the given constraints. No discrete random variable appears in this article and for simplicity the term "entropy" is used for $h(f)$.

The measure of deviation of two probability density functions $f(t)$ and $g(t)$ that is related to the concept of entropy is the Kullback–Leibler (KL) distance (also relative entropy, (Cover & Thomas 1991))

$$\mathrm{KL}(f, g) = \int_{-\infty}^{\infty} f(t) \ln \frac{f(t)}{g(t)} dt. \tag{2}$$

The KL distance defined by formula 2 is not symmetric and does not satisfy the triangle inequality. Further, if there exists an interval such that $g(t) = 0$ while $f(t) \neq 0$ then $\mathrm{KL}(f, g) = \infty$. On the other hand, the KL distance is independent of coordinate transforms and $\mathrm{KL}(f, g) \geq 0$ with equality if, and only if, $f(t) = g(t)$. These properties make the KL distance suitable for measuring the information content described by the probability density $f(t)$ relative to the reference probability density $g(t)$, which maximizes uncertainty. The interpretation of KL distance as a measure of the information content was pioneered by Tarantolav & Valette (1982) in the general theory of inverse problems. The reference state $g(t)$ is often described as the state of null information or the "state of total ignorance" (Tarantola 1994). If the interpretation in terms of information is not needed, the KL distance can be still used as a tool to compare ISI distributions.

The above mentioned definitions can be extended into more dimensions, see Cover & Thomas (1991) for details. Let $\{T_1, \ldots, T_n\}$ be a set of n random variables described by the joint probability density function $f(t_1, \ldots, t_n)$. Then the (joint) entropy $h(T_1, \ldots, T_n)$, sometimes also denoted as $h(f)$, is defined by

$$h(T_1, \ldots, T_n) = -\int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} f(t_1, \ldots, t_n) \ln f(t_1, \ldots, t_n)\, dt_1 \ldots dt_n. \tag{3}$$

The joint entropy can be expressed as a sum of conditional entropies:

$$h(T_1, \ldots, T_n) = \sum_{i=1}^{n} h(T_i \mid T_{i-1}, \ldots, T_1), \tag{4}$$

where $h(T_i \mid T_{i-1}, \ldots, T_1) = -\int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} f(t_1, \ldots, t_i) \ln f(t_i \mid t_{i-1}, \ldots, t_1)\, dt_1 \ldots dt_i$. The independence bound for the joint entropy states that (Cover & Thomas 1991)

$$h(T_1, \ldots, T_n) \leq \sum_{i=1}^{n} h(T_i) \tag{5}$$

with equality if, and only if, the variables T_i are independent, i.e., when the conditional densities are equal to the marginal ones $f(t_i \mid t_{i-1}, \ldots, t_1) = f(t_i)$ for each i. Furthermore, it is possible (Cover & Thomas 1991, p. 273) to define the "entropy per variable," or, the entropy rate $\bar{h}$ of a stochastic process $\{T_i\}_{i=1}^{\infty}$ as

$$\bar{h}(f) = \lim_{n \to \infty} \frac{1}{n} h(T_1, \ldots, T_n). \tag{6}$$

The existence of the limit is guaranteed if the sequence $\{T_i\}_{i=1}^{\infty}$ is stationary.

Similarly to Formula "3", the KL distance of two joint probability density functions $f(t_1, \ldots, t_n)$ and $g(t_1, \ldots, t_n)$ is

$$\mathrm{KL}(f, g) = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} f(t_1, \ldots, t_n) \ln \frac{f(t_1, \ldots, t_n)}{g(t_1, \ldots, t_n)} dt_1 \ldots dt_n. \tag{7}$$

In agreement with Equation 6 we formally define the rate R of the KL distance as:

$$R = \lim_{n\to\infty} \frac{1}{n} \mathrm{KL}(f(t_1, \ldots, t_n), g(t_1, \ldots, t_n)). \tag{8}$$

Expanding R employing formulas 3 and 6 gives

$$R = -\lim_{n\to\infty} \left[\frac{1}{n} \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} f(t_1, \ldots, t_n) \ln g(t_1, \ldots, t_n)\, dt_1 \ldots dt_n \right] - \bar{h}(f). \tag{9}$$

In light of the KL distance as a measure of information, we interpret R as the "information rate" of a stochastic process described by the joint probability density function $f(t_1, t_2, \ldots)$ relative to the state of null information described by $g(t_1, t_2, \ldots)$. This concept of information differs from the standard notion of mutual information introduced by Shannon (1948), though under special conditions a relation between the two can be found, as we demonstrate later. We will employ this definition to classify the available information content in several particular models of neuronal activity.

Results and discussion

The spike train as a renewal process

We apply formula 9 to find the information rate R of a renewal process describing the ISI generation. First, however, we have to determine uniquely the appropriate state of null information. The least informative state is in other words the most random one, i.e., it bounds the entropy rate $\bar{h}$ of all other possible processes from above. Formula 5 implies, that such sequence of random variables must consist of independent and identically distributed variables maximizing the entropy h. The maximum entropy distribution on $(0, \infty)$ with fixed expected value is the exponential one described by the probability density function

$$g(t) = a \exp(-at), \tag{10}$$

where $a > 0$ and the expected value equals to $1/a$. The entropy of the exponential distribution is $h(g) = 1 - \ln a$. The state of null information is thus described by the Poisson process which holds a prominent position in the neuronal modeling (Gerstner & Kistler 2002). The Poisson-like firing has been experimentally observed in many situations, particularly in the cortical neurons, see Abeles (1982) for details. From Formulas 4 and 6 it follows, that the entropy rate of a renewal process is equal to the entropy of the probability density. Thus, the Equation 9 reduces to

$$R = aE(T) - \ln a - h(f). \tag{11}$$

As we quantify the information gained beyond the hypothesis of frequency coding, we let distributions $f(t)$ and $g(t)$ have the same mean values, $E(T) = 1/a$. Formula 11 finally results in

$$R = 1 + \ln E(T) - h(f). \tag{12}$$

As mentioned, the entropy $h(T)$ of a random variable T changes with the transformation $T \to sT$ (with a 'scaling' constant $s > 0$) as $h(sT) = h(T) + \ln s$ (Cover & Thomas 1991,

p. 233). Due to this "scaling property" the rate R in the Formula 12 does not depend on the actual $E(T)$, which also follows from the fact that the rate is defined "per ISI". We see that the information rate of the renewal process gained beyond frequency coding relative to the state of null information is reduced to the calculation (estimation) of $h(f)$.

In the following, we compare the information rates R of several ISI distribution models described by probability density functions $f(t)$ with two parameters. The rate R is parameterized by the coefficient of variation CV–the ratio of standard deviation to $E(T)$. While the advantage of CV is that for exponential distribution holds $CV = 1$ (independently of $E(T)$), the condition $CV = 1$ does not in general imply exponentiality of the distribution. This feature was explored in greater detail with respect to experimental data in Kostal & Lansky (2006).

Gamma model. The gamma distribution is one of the most frequent statistical descriptor of ISI (Levine 1991; Mandl 1992; Rieke et al. 1997). Its probability density function $f(t)$ is defined by

$$f(t) = \frac{b^a t^{a-1} \exp(-bt)}{\Gamma(a)}, \tag{13}$$

with parameters $a > 0, b > 0$ and $\Gamma(z) = \int_0^\infty t^{z-1} e^{-t}\, dt$ the gamma function. It holds for Equation 13: $E(T) = a/b$ and $CV = 1/\sqrt{a}$. Using Formula 12 we obtain the information rate $R(CV)$

$$R(CV) = 1 - \ln CV^2 - \ln \Gamma(1/CV^2) + \frac{\Psi(1/CV^2) - 1}{CV^2} - \Psi(1/CV^2), \tag{14}$$

where $\Psi(z) = \frac{d}{dz} \ln \Gamma(z)$ is the digamma function.

The rate $R(CV)$ from Equation 14 is plotted in Figure 1A. It is useful to view the dependence $R(CV)$ for every possible value of CV and R. To do this, we employ a conformal mapping $(CV, R) \to (\hat{CV}, \hat{R})$ described by

$$\begin{aligned} \hat{CV} &= \arctan CV, \\ \hat{R} &= \arctan R. \end{aligned} \tag{15}$$

Due to this transformation, the whole quadrant $(0, \infty) \times (0, \infty)$ is mapped onto section $(0, \pi/2) \times (0, \pi/2)$ and the points $\hat{CV} = \pi/2$, resp. $\hat{R} = \pi/2$ are identified with $CV = \infty$, resp. $R = \infty$. For convenience, labels on Figure 1A. correspond to the original scale (CV, R).

The gamma density $f(t)$ in Equation 13 is exponential for $a = 1$, implying $R(CV = 1) = 0$. The information rate R tends to infinity for $CV \to 0$. This is a general property which can be seen directly from Formula (12): for $CV = 0$ the variable T is described by a δ-function and the entropy is thus $h(f) = -\infty$, see Cover & Thomas, p. 229 (1991) for details. On the other hand, the limit $R(CV \to \infty) = \infty$ is true for the gamma distribution, but does not hold in general.

Inverse Gaussian model. The inverse Gaussian distribution of ISI can be obtained from the integrate-and-fire class of neuronal models and is also often fitted to experimental data (Gerstein & Mandelbrot 1964; Levine 1991). Its probability density is

$$f(t) = \sqrt{\frac{a}{2\pi b t^3}} \exp\left[-\frac{1}{2b}\frac{(t-a)^2}{at}\right], \tag{16}$$

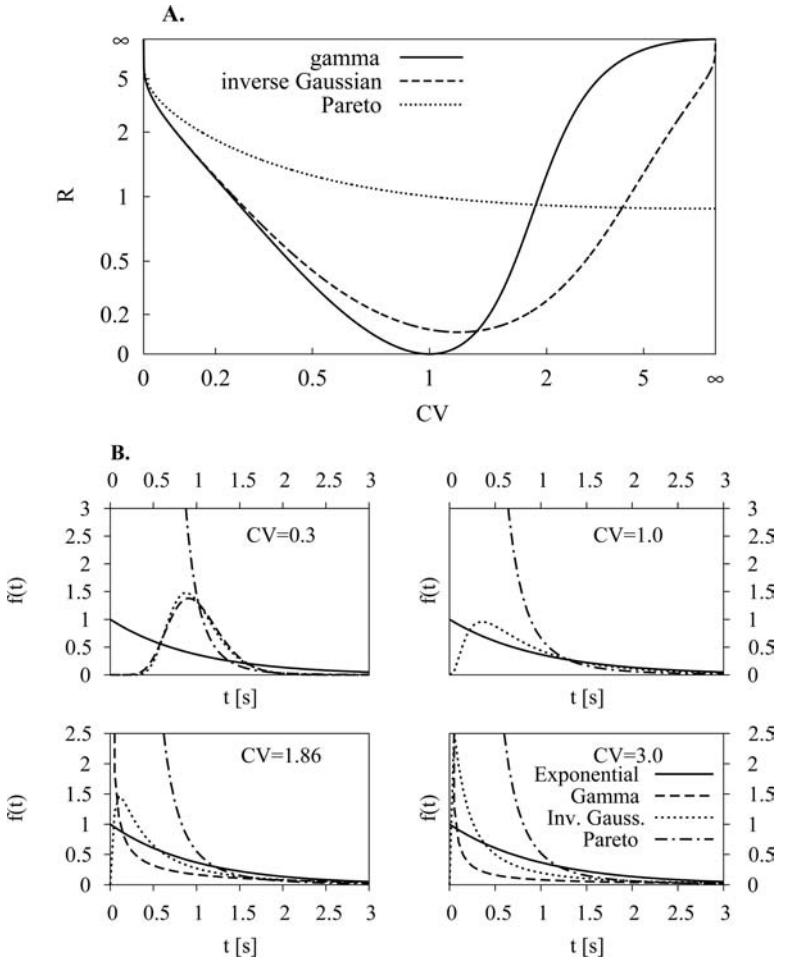


Figure 1. Information rates R of renewal processes in dependence on the CV (A). For $CV = 1$ the rate of gamma model is zero, implying that at this point its distribution becomes exponential. The inverse Gaussian model is never exponential ($R > 0$) and its minimal distance is at $CV \approx 1.17$. The Pareto model is also never exponential, but contrary to the gamma and inverse Gaussian cases its rate decreases with increasing CV, reaching the limit $R(CV \to \infty) \approx 0.89$. The information rate of inverse Gaussian grows slowly compared to that of the gamma distribution. The rates of gamma and inverse Gaussian are equivalent for $CV \approx 1.31$. For $CV \ll 1$ the gamma and inverse Gaussian rates are hardly distinguishable, though this fact cannot be used to judge the degree of their mutual similarity, see (B), where the three corresponding probability density functions with $E(T) = 1$ s are plotted for selected values of CV.

with parameters $a > 0$ and $b > 0$. In this case, $CV = \sqrt{b}$, $E(T) = a$ and using the same technique as in the case of the gamma model the information rate $R(CV)$ is obtained in the form

$$R(CV) = \frac{1}{2}\ln\frac{e}{2\pi} - \ln CV + \frac{3}{\sqrt{2\pi}}\frac{\exp(1/CV^2)}{CV}K_{1/2}^{(1,0)}(1/CV^2), \tag{17}$$

where $K_\nu^{(1,0)}(z)$ is the derivative of the modified Bessel function of the second kind $K_\nu^{(1,0)}(z) = \frac{\partial}{\partial\nu}K_\nu(z)$ (Abramowitz & Stegun 1972).

The resulting dependence $R(CV)$ from Equation (17) is plotted in Figure 1A using the transformation rules (Equation 15). The inverse Gaussian distribution is not exponential for $CV = 1$ thus always $R(CV) > 0$. Furthermore, the minimal rate (and thus the maximum similarity with the exponential model) does not occur at $CV = 1$ but at $CV \approx 1.17$. The rates of gamma and inverse Gaussian models become equivalent at $CV \approx 1.31$. In comparison with gamma model, we see that though $R(CV \to \infty) = \infty$ for both cases the gamma model approaches infinite information rate $R(CV)$ much faster than the inverse Gaussian. On the other hand, for very small values of CV both rates are undistinguishable.

Pareto model. The Pareto distribution is not resulting from any theoretic neuronal model and we are not aware of any attempt to fit it to experimentally observed ISIs. We present it here to show a different kind of behavior than that of the above mentioned models. The probability density function $f(t)$ of the Pareto distribution is

$$f(t) = \begin{cases} 0, & t \in (0, b) \\ ab^a t^{-a-1}, & t \in [b, \infty) \end{cases} \tag{18}$$

with parameters $a > 2$ and $b > 0$. In this case $CV = 1/\sqrt{(a^2 - 2a)}$ and $E(T) = ab/(a-1)$. Using Formula 12, we arrive at the equation for the information rate $R(CV)$ of the Pareto model in the form

$$R(CV) = CV^2 - CV\sqrt{1 + CV^2} + \ln\left(2 + \frac{1 + 2CV^2}{CV\sqrt{1 + CV^2}}\right), \tag{19}$$

for illustration see Figure 1A. The Pareto model is never exponential and we can see that for increasing values of CV the rate $R(CV)$ is slowly decreasing with the limit (which is also the minimal value of $R(CV)$): $R(CV \to \infty) = \ln 4 - 1/2 \approx 0.886$. The information rates of Pareto and gamma models are equal for $CV \approx 1.86$, while for gamma and inverse Gaussian models it is equal for $CV \approx 1.26$. Nevertheless this fact cannot cannot be used to judge the degree of similarity between the gamma and either the Pareto or inverse Gaussian models. The comparison of probability density functions $f(t)$ of the inverse Gaussian, gamma and Pareto models can be seen in Figure 1B for $E(T) = 1$ s and $CV \approx 1.86$.

R and specific information

Mutual information $I(\mathcal{S}; \mathcal{R})$ (Cover & Thomas 1991) determines the dependence between stimuli $\mathcal{S}$ and responses $\mathcal{R}$ (Borst & Theunissen 1999). The information gained from a particular stimulus is known once the variability of responses across the whole set of stimuli is determined. $I(\mathcal{S}; \mathcal{R})$ has no informative value if only one stimulus is presented. The coding efficiency of chosen stimulus can be judged according to the deviation of the response from the spontaneous activity (Chacron et al. 2001), i.e., the most informative stimuli cause the largest difference. The information rate R provides a natural measure for this difference. Furthermore, in the following we show that under certain conditions a link between R and $I(\mathcal{S}; \mathcal{R})$ may be established.

The set of stimuli $\mathcal{S} = \{s_i\}_{i=1}^n$ is discrete and the set of responses is realized by ISIs which can take any positive value. Mutual information can be formally expressed as $I(\mathcal{S}; \mathcal{R}) = \sum_i p(s_i)\, i(\mathcal{R} \mid s_i)$, where $i(\mathcal{R} \mid s_i)$ is called the specific information due to the stimulus s_i. Analogously to DeWeese & Meister (1999), we express $i(\mathcal{R} \mid s_i)$ as

$$i(\mathcal{R} \mid s_i) = h(\mathcal{R}) - h(\mathcal{R} \mid s_i). \tag{20}$$

From Formula 20 it follows that the specific information is large for those stimuli that have only a few different responses associated with them because $h(\mathcal{R} \mid s_i)$ is the uncertainty in response given stimulus s_i. If the stimulus s_i evokes only single possible response then it holds $h(\mathcal{R} \mid s_i) = -\infty$, because the probability density function of responses is realized by δ-function.

We have restricted ourselves to the case in which the ISIs are described by a renewal process with probability density function f and the stimuli conditions are stationary in time. Under these two assumptions, we can assign ISI distribution with density f to the stimulus s_i and the uncertainty in response becomes $h(f) = h(\mathcal{R} \mid s_i)$. The remaining term $h(\mathcal{R})$ in Formula 20 depends on the distribution of stimuli. It is possible to view $h(\mathcal{R})$ as the entropy of the spontaneous neuronal activity $h(\mathrm{Spon})$. The difference $i(\mathrm{Spon} \mid s_i) = h(\mathrm{Spon}) - h(\mathcal{R} \mid s_i)$ does not add up to mutual information in the Shannon's sense. Nevertheless, the information rate computed employing the spontaneous activity differs only by a constant from the true mutual information rate and therefore the two measures behave similarly, see (Chacron et al. 2001) and (Chacron et al. 2003) for details.

If the spontaneous activity is described by the Poisson process, then Formula 20 corresponds to the expression for the information rate (Equation 12) and R coincides with specific information. If the spontaneous activity differs from the Poisson firing, then $i(\mathrm{Spon}|s_i)$ can be obtained from formula

$$i(\mathrm{Spon} \mid s_i) = R(f) - R(\mathrm{Spon}) = h(\mathrm{Spon}) - h(f), \tag{21}$$

where $R(\mathrm{Spon})$ resp. $R(f)$ are the information rates of the spontaneous activity and the activity in question. Note that we cannot directly employ a general (non-Poisson) spontaneous activity as the state of null information (which is required to maximize the entropy) because we would lose the interpretation in terms of information.

The spike train as a Markov chain

First we employ Formula 9 to find the information rate R of a general stationary ISI model. Using the same reasoning as in the case of renewal process we find the state of null information to be the Poisson process. Simplifying Equation 9 in this situation yields an expression similar to Formula 11

$$R = aE(T) - \ln a - \bar{h}(f) \tag{22}$$

and putting the mean ISI of the stochastic process described by $f(t_1, t_2, \ldots)$ and that of the state of null information equal, it yields

$$R = 1 + \ln E(T) - \bar{h}(f). \tag{23}$$

The available information rate of the general stationary stochastic process (gained beyond the idea of frequency coding) relative to the state of null information is again reduced to the calculation (estimation) of the entropy rate. (Obviously the relationship between R and the specific information holds also in the general stationary case.)

The computation of the entropy rate is however hardly possible in the general case (Cover & Thomas 1991). Even though the limit is theoretically guaranteed to exist the convergence may be arbitrarily slow with increasing dimension of the joint probability density function. In the following, we restrict ourselves to the class of stationary stochastic processes satisfying the first-order Markov property (Cox & Lewis 1966)

$$\mathrm{Prob}\{T_n \le t_n \mid T_{n-1} = t_{n-1}, \ldots, T_1 = t_1\} = \mathrm{Prob}\{T_n \le t_n \mid T_{n-1} = t_{n-1}\}. \tag{24}$$

In other words, each ISI depends only on the immediately preceding one and is conditionally independent of all other preceeding ISIs. The Markov chain is therefore fully described by the joint probability density function $f(t_1, t_2)$ of the two adjacent ISIs. Condition 24 simplifies the expression for the entropy rate and analogously to Cover & Thomas (1991, p. 66), we write

$$\bar{h}(f) = h(T_2 \mid T_1) = -\int_0^\infty \int_0^\infty f(t_1, t_2) \ln f(t_2 \mid t_1)\, dt_1\, dt_2. \tag{25}$$

(To avoid indexing we denote $T_1 \equiv X$ and $T_2 \equiv Y$.)

Formula 23 for the information rate R of the first-order Markov chain becomes

$$R = 1 + \ln E(Y) - h(Y \mid X). \tag{26}$$

Equation 26 can be rewritten using the mutual information $I(X; Y)$, the symmetric quantity that measures dependence between two random variables X, Y (Cover & Thomas 1991)

$$I(X; Y) = h(Y) - h(Y \mid X). \tag{27}$$

Combining Equations 26 and 27 separates the total information rate R into two parts

$$R = R_1 + I(X; Y), \tag{28}$$

where $R_1 = h(g) - h(Y)$ is the information rate of the renewal process described by the marginal probability density function corresponding to the given Markov chain. The important property of mutual information is that $I(X; Y) = 0$ if, and only if, the variables X and Y are independent, i.e., the Markov chain is reduced to the renewal process. Note that the above derived results can be extended to the kth-order class of Markov chains (conditional dependence on the k preceeding states).

In the following, we compare the information rates R of several Markov chain ISI models described by the probability density functions $f(x, y)$. We parameterize the rate R in dependence on the serial correlation $\varrho = [E(XY) - E(X)E(Y)]/[\sqrt{Var(X)}\sqrt{Var(Y)}]$, which is frequently used to measure the dependence between variables X and Y. The resulting rate R is again independent of the expected value of the ISI.

Lawrance and Lewis model. The model introduced by Lawrance and Lewis (L–L) is described by the joint probability density function (Lawrance & Lewis 1977)

$$f(x, y) = \frac{a^2 b}{1 - b + b^2}\left\{U(bx - y)\frac{1 - b}{b^2}\exp\left[-\frac{ab(x + y) - ay}{b^2}\right]\right.$$
$$\left. + \exp\left[-\frac{a(x + by)}{b}\right] + U(y - bx)\frac{(b - 1)^2}{b}\exp[-a(x - bx + y)]\right\} \tag{29}$$

with parameters $a > 0$, $b \in (0, 1)$ and $U(x)$ the Heaviside unit step-function: $U(x < 0) = 0$ and $U(x \geq 0) = 1$. The (first-order) serial correlation ϱ is

$$\varrho = b(1 - b), \tag{30}$$

from which the limitation of the L–L model follows: serial correlation is confined in interval $(0, 1/4)$. The marginal distribution of the L–L model is exponential with parameter a, see Formula 10.

We investigate the information rate of the L–L model using formula 28. The term R_1 is zero and thus the total rate R is equivalent to the mutual information, $R = I(X; Y)$. Two

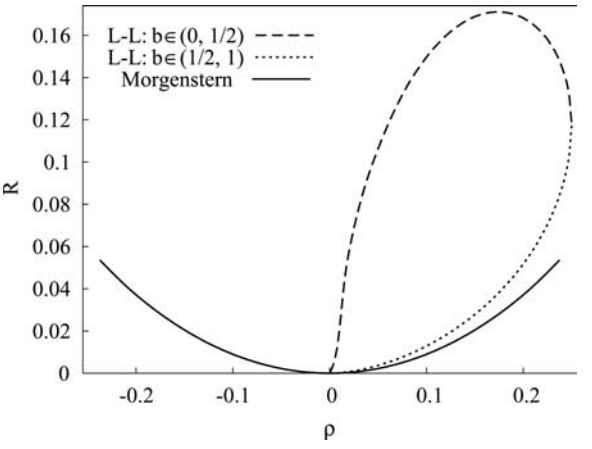


Figure 2. The information rates R in dependence on the serial correlation ϱ of the Lawrance and Lewis (L–L) and Morgenstern ISI models. For $b \in (1/2, 2)$ (the L–L model) the rate increases monotonically to its maximum $R(\varrho = 1/4) \approx 0.12$. For $b \in (0, 1/2)$ the rate reaches its maximum at $R(\varrho \approx 0.17) \approx 0.17$ and then decreases. The Morgenstern model exhibits simpler behavior: its rate is smaller compared to that of the L–L model at the corresponding value of ϱ.

values of parameter b lead to the same value of the serial correlation in Equation 30, thus we have to consider two solutions of R depending on which value of b was used

$$b \in (0, 1/2) : \frac{1}{2} - \frac{1}{2}\sqrt{1 - 4\varrho},$$

$$b \in (1/2, 1) : \frac{1}{2} + \frac{1}{2}\sqrt{1 - 4\varrho}.$$

The resulting $R(\varrho)$ was carried out numerically and is plotted in Figure 2. For $b \in (1/2, 1)$ the rate increases monotonically from zero ($R(\varrho = 0) = 0$, the ISIs are independent) to its maximum value $R(\varrho = 1/4) \approx 0.12$. A more interesting result comes from examining the behavior of R for $b \in (0, 1/2)$. The maximum value $R \approx 0.17$ does not correspond to the maximum value of serial correlation, but is located at $\varrho \approx 0.17$. From this value with increasing serial correlation we observe a decreasing rate, while the marginal distribution is still exponential. This seemingly paradoxical result comes from the fact that the value of serial correlation as a measure of dependency cannot be used to judge the degree of difference from the state of null information.

We can judge the qualitative behavior of the L–L model from the joint probability density plots shown in Figure 3A and B. Though serial correlation is often found in experimental data, two-dimensional histograms corresponding to Figure 3 are not presented. The length of the "immediately preceeding" ISI is given on the x-axis, the length of the current ISI is on the y-axis and the probability of their joint occurrence is indicated by the shade (a darker tone corresponds to a higher value, absolute numbers are not important). The positive serial correlation of the ISIs can be seen immediately for $b = 0.23$ (A): short ISIs tend to be followed by short ones (the dark region for $x < 0.5$ s and $y < 0.5$ s), longer ISIs by comparatively long ones. Note especially the narrow dark band of highly probable (x, y) pairs in the lower part of the plot. This feature makes the existence of certain (short) sequences of ISIs more probable than others, i.e., it can be regarded as a simple mechanism of temporal pattern formation. Unfortunately, the serial correlation in this case is not large enough to make the

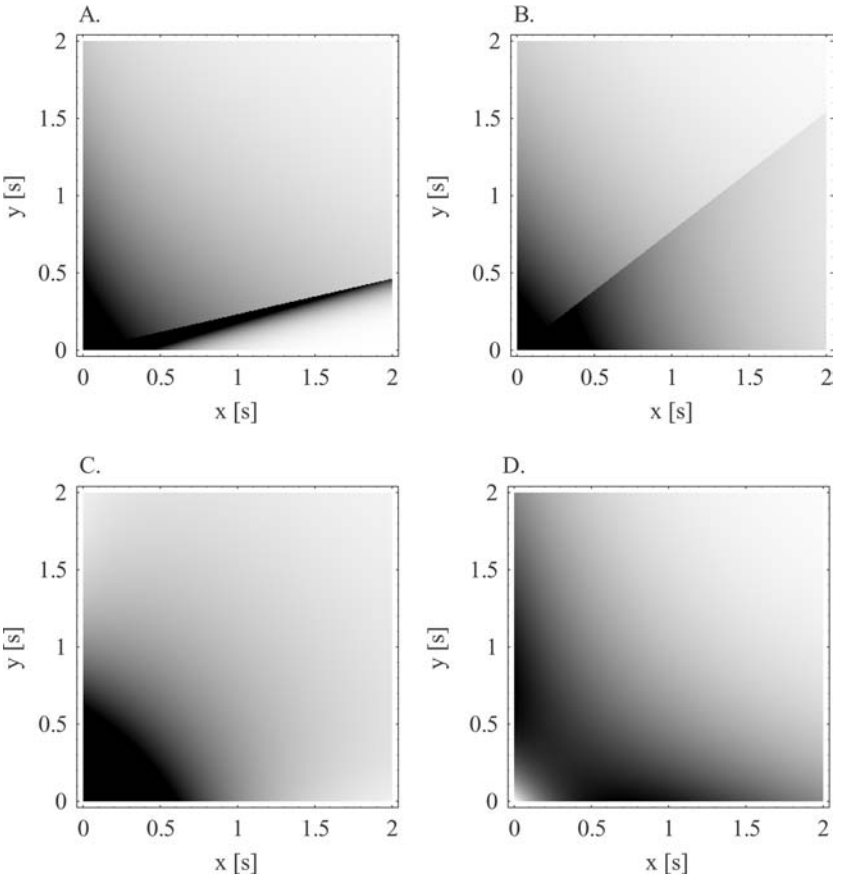


Figure 3. Density plots of the joint probability density functions $f(x, y)$ of the Markov chain ISI models with $E(Y) = 1$ s, darker tones correspond to higher values (absolute numbers are not important). The Lawrance and Lewis (L–L) model (A + B) is plotted for two different values of parameter b. Though the serial correlation ϱ is the same in both situations the shapes of $f(x, y)$ are different. This difference is captured by different values of the rate R. The Morgenstern model (C + D) is plotted for both extremal values of the serial correlation $\varrho = \pm 1/4$ with equal values of $R \approx 0.6$.

effect visually pronounced (see also Figure 4). Below the band of high probability one can see a triangle of nearly zero probability, e.g., $x > 1.5$ s can hardly be followed by $y < 0.25$ s. The behavior of the L–L model for $b = 0.77$ (B) differs in many aspects, though the value of serial correlation is the same as for $b = 0.23$. The sharp band of high probability is missing and the joint probability density is more diffused. Very short ISIs are often followed by even shorter ones than in the previous case (compare with Figure 4D), but the preference for long ISIs following the longer ones ($x > 1$ s) is not pronounced (note the darker triangle below the diagonal). The difference between the two cases is captured in different values of the information rates R.

The neuronal activity described by the L–L model is simulated in Figure 4D for $E(Y) = 1$ s and serial correlation $\varrho = 0.17$ ($b = 0.23$). Another realization corresponding to Figure 3B is shown in Figure 4E. Note that the simulations were always done with the same initial value of random seed. The relatively small differences from the Poisson case (Figure 4A) are captured in comparatively small values of information rate $R \approx 0.17$.

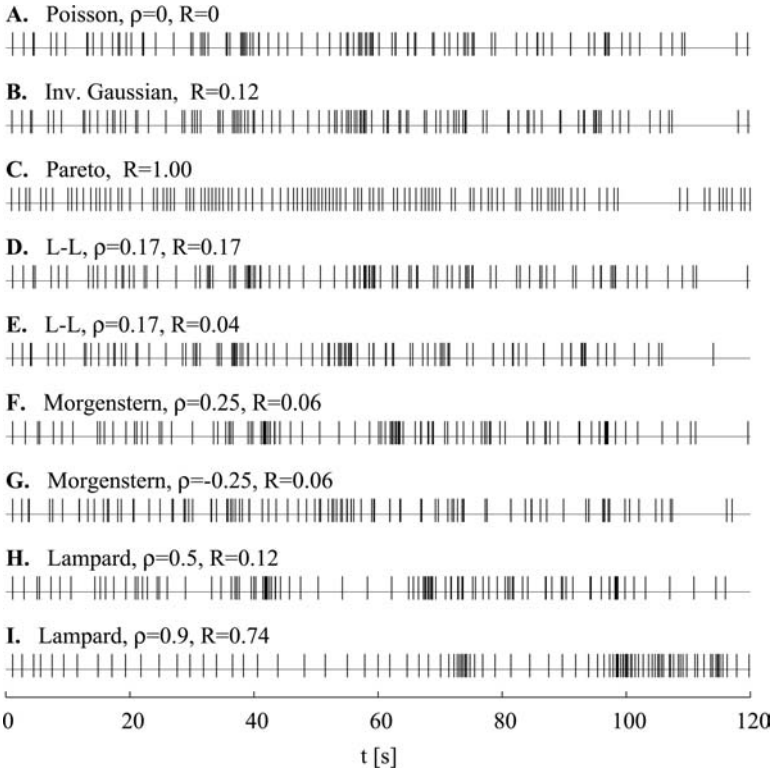


Figure 4. Simulated spike trains for renewal models with $CV = 1$ and Markov models with exponential marginal distributions in both cases $E(T) = 1$. The values of serial correlation ϱ and information rate R are given. (A) The Poisson process. The inverse Gaussian model (B) is similar to (A) but lacking the extremely short ISIs (see the probability density plot in Figure 1B). The Pareto model (C) differs strikingly from the Poisson process, with the main distinction being the "dead time" (approx. 0.58 s). The values of serial correlation ϱ in the cases of the Lawrance and Lewis (D + E), Morgenstern (F + G) and Lampard models (H + I) are relatively too small to produce apparent change in comparison with the Poisson process–the effect is only slightly more pronounced for negative ϱ (G) and for $\varrho = 0.5$ (H). The spike train generated according to Lampard model (I) with $\varrho = 0.9$ can be distinguished on the first sight.

Morgenstern model. The next model is constructed using the bivariate joint probability density function $f(x, y)$ first described by (Morgenstern 1956)

$$f(x, y) = a^2 e^{-2a(x+y)}[e^{a(x+y)} + 4\varrho(e^{ax} - 2)(e^{ay} - 2)], \tag{31}$$

with parameters $a > 0$ and $\varrho \in (-1/4, 1, 4)$. The function $f(x, y)$ is symmetric in its arguments $f(x, y) = f(y, x)$. The marginal distribution of this model is again exponential with parameter a. The serial correlation is equal to the parameter ϱ in Formula 31. The maximum serial correlation is again $|\varrho| = 1/4$, but contrary to the previous case of L–L model it can also be negative.

We use the same approach in applying Formula 26 as in the case of L–L model. The result was carried out numerically and is plotted for comparison with the L–L model in Figure 2. The behavior of the Morgenstern model appears simple compared to that of the L–L model. The information rate is symmetric $R(\varrho) = R(-\varrho)$ and $R(\varrho = 0) = 0$ implying that for zero serial correlation the ISIs are independent, as can be seen directly from Formula 31. An interesting observation is that for the same value of ϱ the rates $R(\varrho)$ of the Morgenstern

and L–L models differ and that for any $\varrho \in (0, 1/4)$ the rates of the Morgenstern model are smaller.

Figure 3C and D shows the $f(x, y)$ for both extreme values of serial correlation. The behavior of the model is simpler compared to the L–L model, there are no sharp regions of special interest and the probability density is smooth and symmetric around $x = y$. The positive correlation (C) is seen for ISIs smaller than 1 s, and then for large ISIs ($x \approx 2$ s) which are usually not followed by very short ones. The reverse statements hold for the negative serial correlation (D). Very short ISIs are not followed by comparatively short ones (the light region for small x, y) and very long ISIs ($x > 2$ s) are preferably followed by shorter ones ($y < 1$ s). Because the information rate is the same in both cases, we cannot distinguish between negative and positive serial correlation just by observing the value of R.

Neuronal firing that behaves according to the Morgenstern model is simulated in Figure 4F and G. Due to the small values of serial correlation only small differences can be seen compared to the Poisson process (A). This is confirmed by a very small value of the information rate $R \approx 0.06$.

Lampard model. The last example of the stationary Markov chain we employ as the ISI distribution model was first described by (Lampard 1968). It describes a counter system whose inputs are a pair of independent Poisson processes. The advantage of this model over the previous two lies in its possible interpretation from the neurophysiological point of view (Lawrance 1972).

The joint probability density function $f(x, y)$ is defined by

$$f(x, y) = \frac{(1/\varrho - 1)^{\xi}}{xy\Gamma(\xi)} \left(\frac{a\xi^2 \sqrt{xy\varrho}}{|\varrho - 1|} \right)^{1+\xi} \exp\left[\frac{a\xi(x + y)}{\varrho - 1} \right] I_{\xi-1}\left(\frac{2a\xi\sqrt{xy\varrho}}{1 - \varrho} \right), \tag{32}$$

where $\varrho \in (0, 1)$ is the first-order serial correlation, parameters $\xi > 0$, $a > 0$ and $I_\nu(z)$ is the modified Bessel function of the first kind (Abramowitz & Stegun 1972). The marginal distribution of the model is the gamma distribution (compare with equation (13)): $f(y) = (a\xi)^{\xi} y^{\xi-1} \exp(-a\xi y)/\Gamma(\xi)$, from which follows that parameter ξ is related to the CV by $CV^2 = 1/\xi$, and a describes the mean value of the ISI $E(T) = 1/a$. For $\xi = 1$, which implies $CV = 1$, the joint density $f(x, y)$ is reduced to the Downton bivariate exponential density (Downton 1970) with exponential marginal distributions:

$$f(x, y) = \frac{a^2}{1 - \varrho} \exp\left[\frac{a(x + y)}{\varrho - 1} \right] I_0\left(\frac{2a\sqrt{xy\varrho}}{1 - \varrho} \right). \tag{33}$$

To obtain the information rate R of the Lampard model we apply Formula 28, because the term R_1 is already given in the closed form by Formula 14. Therefore, the mutual information $I = I(X; Y)$ remains to be calculated. We parameterize the mutual information $I(\varrho)$ with the serial correlation (in agreement with the previous reasoning) and investigate its possible dependence on the CV through the remaining parameter ξ. The Bessel function $I_\nu(z)$ can be expressed in a simple form for two particular values of parameter ν (Abramowitz & Stegun 1972)

$$I_{1/2}(z) = \sqrt{\frac{2}{\pi}} \frac{\sinh z}{\sqrt{z}},$$

$$I_{-1/2}(z) = \sqrt{\frac{2}{\pi}} \frac{\cosh z}{\sqrt{z}}.$$

The solution for the first condition $\xi - 1 = 1/2$ is $CV = \sqrt{2/3} \approx 0.816$ and for the second $\xi - 1 = -1/2$ is $CV = \sqrt{2} \approx 1.414$. The joint density $f(x, y)$ of the Lampard model then reduces to much simpler forms

$$CV = \sqrt{\frac{2}{3}} : f(x,y) = -\frac{9a^2}{2\pi\sqrt{(1-\varrho)\varrho}} \exp\left[\frac{3}{2}\frac{a(x+y)}{\varrho-1}\right] \sinh\left(\frac{3a\sqrt{xy\varrho}}{\varrho-1}\right), \tag{34}$$

$$CV = \sqrt{2} : f(x,y) = \frac{a}{2\pi\sqrt{xy(1-\varrho)}} \exp\left[\frac{1}{2}\frac{a(x+y)}{\varrho-1}\right] \cosh\left(\frac{a\sqrt{xy\varrho}}{\varrho-1}\right). \tag{35}$$

The reduced forms given by Formulas 33 (for $CV = 1$), 34 and 35 make the analytical integration of at least some parts in $I(\varrho)$ possible. The remaining integration can thus be carried out numerically with much better precision than using the general and rather complicated form in Equation 32.

The resulting mutual information $I(\varrho)$ is plotted in Figure 5 for the three above mentioned values of CV. The total information rate R can be computed using Formula 28. For $CV = 1$ and $\varrho = 0$ the Lampard model reduces to the Poisson process, thus both $R_1 = 0$ and $I(X;Y) = 0$. For $CV = 1$ and $\varrho > 0$ holds $R_1 = 0$ and the mutual information corresponds to the total information rate R of the Lampard (=Downton) model. Similarly if $CV \neq 1$ and $\varrho = 0$, then $I(X;Y) = 0$ and the Lampard model reduces to the renewal process with gamma ISI distribution. Using Formula 28 for the term R_1 gives: $R_1(CV = \sqrt{2/3}) \approx 0.044$ and $R_1(CV = \sqrt{2}) \approx 0.216$. The shapes of the curves $I(\varrho)$ are very similar, starting from $I(\varrho = 0) = 0$ and continuing with monotonous increase $I(\varrho \to 1) \to \infty$. Mutual information increases dramatically for $\varrho > 0.8$, where an arbitrarily small change in ϱ has a strong effect on the information rate, while the changes in the serial correlation value for $\varrho < 0.5$ may be neglected. The dependence of $I(\varrho)$ on the value of CV is relatively small, nevertheless, it cannot be attributed to numerical errors. For CV more deviated from one the effect is slightly more pronounced. Again we have a situation where the serial correlation ϱ is not a sufficient measure of the true dependence between two random variables, contrary to $I(X;Y)$.

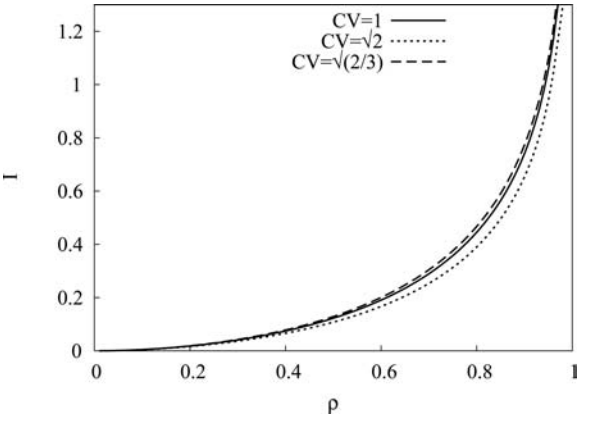


Figure 5. The mutual information $I = I(X;Y)$ of the Lampard model in dependence on serial correlation ϱ for three values of CV. The shapes of the curves are very similar but not identical, which means that ϱ does not measure the ISI dependence completely. A very sharp increase in I for $\varrho > 0.5$ suggests that low values of ϱ have only marginal effect, while a small increase in ϱ for $\varrho > 0.8$ can change R dramatically.

A sample of neuronal firing behaving according to the Lampard model is shown in Figure 4 for $\varrho = 0.5$ (H) and $\varrho = 0.9$ (I). At such a high value of serial correlation (especially in Figure 4I) the pattern of spikes can be immediately distinguished from the other cases presented in the figure, though the marginal distribution (estimated, e.g., by histograms) of all the presented cases is exponential with equal parameters.

Experimental data

Formula 26 for the information rate R of the Markov chain can be expressed alternatively, substituting from the chain rule (Equation 4)

$$R = h(g) + h(X) - H(X, Y). \tag{36}$$

We see that the computation of R from experimental data for the renewal process and Markov chain is reduced to the estimation of entropy from one- and two-dimensional probability density functions. This makes equations 12 and 36 applicable in experimental data analysis as the problem of entropy from data estimation is well exploited in literature, see, e.g., Tsybakov & Meulen (1996); Beirlant et al. (1997) for an overview of available techniques.

For one dimension (the renewal process) the simple and well researched Vasicek's estimator (Vasicek 1976) gives reasonably good results on a wide range of data (Ebrahimi et al. 1992; Miller & Fisher III, 2003). Furthermore, our own experience with simulated data shows that for sample sizes $n = 500$ (the average size in the experimental data we used) the standard deviation is relatively small ($\sigma < 0.07$) and a possible positive bias with respect to true values is negligible. It is also preferable to avoid estimations based purely on histograms because the choice of binwidth affects the results greatly. The support of ISI distributions is always positive which makes the application of kernel estimators problematic due to possible overlapping into negative values.

In this section, we illustrate the use of information rate on experimental data in the case of renewal process. The data come from extracellular recordings made from olfactory receptor neurons of freely breathing and tracheotomized rats. Single-unit action potentials were recorded and more details on the data acquisition is described in (Duchamp-Viret et al. 2003). The sample sizes range from (circa) $n = 100$ to $n = 2000$ ISIs and all records have been tested for stationarity and ISI independence (the Wald–Wolfowitz test, serial correlation, periodogram).

Given the n ranked ISIs $\{t_{[1]} < t_{[2]} < \cdots < t_{[n]}\}$ we used the entropy estimator proposed by (Vasicek 1976)

$$h(\text{data}) = \frac{1}{n} \sum_{i=1}^{n} \ln \left[\frac{n}{2m} (t_{[i+m]} - t_{[i-m]}) \right]. \tag{37}$$

The positive integer parameter $m < n/2$ is set prior to computation and the two following conditions hold: $t_{[i-m]} = t_{[1]}$ for $(i - m) < 1$ and $x_{[i+m]} = x_{[n]}$ for $(i + m) > n$. The particular values of m corresponding to various values of n were determined by Ebrahimi et al. (1992).

The information rate R represents the average information gained per ISI and does not depend on $E(T)$, i.e., the firing rate. To include the effect offaster vs. slower neuronal firing, we examine the average distribution of information rate in time. We define the information rate flow η by

$$\eta = \frac{1}{\ln 2} \frac{R}{E(T)}. \tag{38}$$

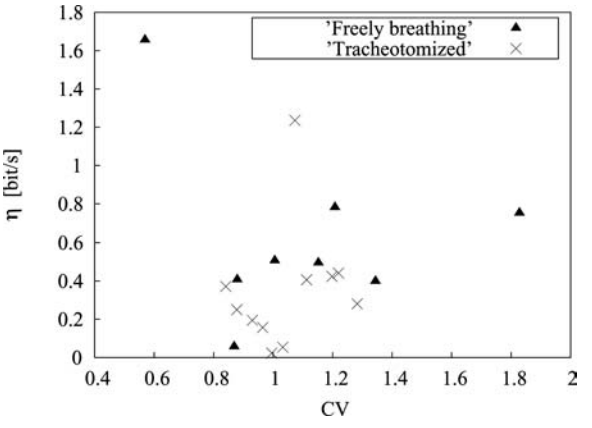


Figure 6. The information rate flow η in bits per second estimated from the experimental data in dependence on the CV. The stationary renewal activity of olfactory neurons in rats is compared for freely breathing (▲) (in the absence of any particular stimulus) and tracheotomized (×) animals. Except for two cases the η of tracheotomized animals is lower than in freely breathing ones. The results show that the activity in tracheotomized rats is closer to the state of null information (Poisson process) and that CV is not well related to the Poisson character of the process.

The factor $1/\ln 2$ is used to change the logarithm base in Equation 9 to 2, so the quantity η represents the average information gained (relative to the state of null information) in bits per second.

The estimated information rate flow η for the already mentioned two categories of data is plotted in Figure 6. We see that the η in the tracheotomized case (×) is in most of the cases lower than that of the freely breathing (▲). This is verified independently by other methods in Duchamp-Viret et al. (2005), where also some further inferences from the data are made. Due to the properties of R low values of η indicate that the firing is close to the state of null information, therefore justifying the hypothesis that spontaneous firing is not informative.

Conclusions

The information rate R based on the Kullback–Leibler distance between two ISI models was proposed for a general case of stationary neuronal activity. If the reference state maximizes the entropy of the ISI probability distribution and if the mean values of both distributions are equal, then R measures the information rate per ISI due to the temporal coding scheme. Beside the introduction of R, we proposed a related quantity, the information flow: $\eta = R/E(T)$, which measures the information gain per time unit. This quantity takes the firing rate of the neuron into account, thus even relatively small values of R must be taken into consideration when comparing the η of fast-firing neurons to the slower ones. The determination of R (and η) requires the computation (estimation) of differential entropy which makes the quantity applicable on suitable experimental or simulated data, as illustrated in the final section of this article. We showed that R is related to the mutual information and corresponds to the information due to specific stimulus.

We analyzed in detail several examples from two categories: the renewal process and the first-order Markov chain ISI models. The case of the renewal processes shows that even if neither spike frequency nor CV changes there still may be a gain of information. The chosen models behave differently for $CV > 1$, while for $CV \ll 1$ their information rates R are very

similar. On the other hand, the case of Markov chains indicates that if R is examined only due to the dependency among ISIs, small values of serial correlation ϱ imply small values of R. Moreover, the increase in the serial correlation of ISIs does not necessarily increase the information rate. Though the Markov chain ISI models discussed here are not resulting from realistic neuronal models, it is nevertheless clear that the relation of serial correlation to the information rate is not simple. As the entropy rates $\bar{h}$ were computed directly from the joint probability densities these Markov chain models serve well for testing purposes and further development of multi-dimensional entropy estimators. These estimators in turn may be used to estimate R in realistic models, where the joint probability density of ISIs is not available (including experimental data).

Theoretically, the information rate R may tend to infinity. This is due to the fact that a continuous random variable generally carries an infinite amount of information (van der Lubbe, 1997). Nevertheless this fact can be considered as merely formal and without consequences, in practice we are always working with finite precision on a finite time scale. Notably, the cases $CV \to 0$ or $\varrho \to \pm 1$ would require an infinite timing precision of the neuronal firing.

Acknowledgements

The authors wish to thank Patricia Duchamp-Viret and Michel Chaput for making their experimental data available. The authors are grateful to Nick Dorrel for linguistic correction. This work was supported by the Research project AV0Z 5011922, Center for Neuroscience LC554 and by the Academy of Sciences of the Czech Republic Grant (Information Society, 1ET400110401).

References

Abeles M. 1982. *Local cortical circuits: an electrophysiological study*, Berlin Heidelberg Springer-Verlag.

Abramowitz M, Stegun IA. 1972. Handbook of mathematical functions, New York, Dover.

Adrian ED. 1928. *The basis of sensation*, New York: W. W. Norton.

Beirlant J, Dudewicz EJ, Gyorfi L, van der Meulen EC. 1997. Nonparametric entropy estimation: an overview. Int J Math Stat Sci 6:17–39.

Bialek W, Rieke F, Van Steveninck RR, Warland D. 1991. Reading a neural code. Science 70:1629–1638.

Borst A, Theunissen FE. 1999. Information theory and neural coding. Nature Neuroscience 2:947–957.

Buracas GT, Albright TD. 1999. Gauging sensory representations in the brain. Trends Neurosci 22:303–309.

Chacron MJ, Longtin A, Maler L. 2001. Negative interspike interval correlations increase the neuronal capacity for encoding time-dependent stimuli. J Neurosci 21:5328–5343.

Chacron MJ, Longtin A, Maler L. 2003. The effects of spontaneous activity, background noise, and the stimulus ensemble on information transfer in neurons. Network: Comput Neural Syst 14:803–824.

Cover TM, Thomas JA. 1991. *Elements of information theory*. New York: John Wiley & sons Inc.

Cox DR, Lewis PAW. 1966. The statistical analysis of series of events. New York, John Wiley & Sons, Inc.

DeWeese MR, Meister M. 1999. How to measure the information gained from one symbol. Network: Comput Neural Syst 10:325–340.

Downton F. 1970. Bivariate exponential distributions in reliability theory. J Roy Statist Soc B 32:408–417.

Duchamp-Viret P, Duchamp A, Chaput MA. 2003. Single olfactory sensory neurons simultaneously integrate the components of an odour mixture. Eur J Neurosci 10:2690–2696.

Duchamp-Viret P, Chaput MA, Kostal L, Lánský P, Rospars J-P. 2005. Patterns of spontaneous activity in single rat olfactory receptor neurons are different in normally breathing and tracheotomized animals. J Neurobiology 65:97–114.

Ebrahimi N, Habibullah M, Soofi ES. 1992. Testing exponentiality based on Kullback-Leibler Information. J Statist Soc B 3:739–748.

Fuller MS, Looft F. 1984. An information theoretic analysis of cutaneous receptor responses. IEEE Trans on Biomed Eng 314:377–383.

Gerstein G, Mandelbrot B. 1964. Random walk models for the spike activity of a single neuron. Biophys J 4:41–68.

Gerstner W, Kistler W. 2002. *Spiking neuron models*. Cambridge: Cambridge University Press.
Kostal L, Lansky P. 2006. Similarity of interspike interval distributions and information gain in a stationary neuronal firing. Biol Cybern 94:157–167.
Lampard DG. 1968. A stochastic process whose successive intervals between events form a first order Markov chain. J Appl Probab 5:648–668.
Lánský P, Rodriguez R. 1999. Two-compartment stochastic model of a neuron. Physica D 132:267–286.
Lánský P, Rodriguez R, Sacerdote L. 2004. Mean instantaneous firing frequency is always higher than the firing rate. Neural Comput 16:477–489.
Lawrance AJ. 1972. Some models for stationary series of univariate events, In: *Stochastic point processes: statistical analysis, theory and applications*, editor P.A.W. Lewis, New York, pp. 99–256. John Wiley & Sons, Inc.
Lawrance A J, Lewis PAW. 1977. An exponential moving-average sequence and point process (EMA1). J Appl Probab 14:98–113.
Levine M.W. 1991. The distribution of the intervals between neural impulses in the maintained discharges of retinal ganglion cells. Biol Cybern 65:459–467.
Lindner B. 2004. Interspike interval statistics of neurons driven by colored noise. Phys Rev E 69:022901.
Longtin A, Racicot DM. 1999. Assesment of linear and non-linear correlations between neural firing events. In: *Nonlinear Dynamics and Time Series: Building a Bridge between the Natural and Statistical Sciences*, editor Cutler CD. and Kaplan DT. *Fields Institute Communications* 11:223–239.
van der Lubbe JCA. 1997. *Information theory*, Cambridge, Cambridge University Press.
Mandl G. 1992. Coding for stimulus velocity by temporal patterning of spike discharges in visual cells of cat superior colliculus. Vision Res 33:1451–1475.
Middlebrooks JC, Clock AB, Xu L, Green DM. 1994. A panoramic code for sound location by cortical neurons. Science 264:842–844.
Miller EG, Fisher III JW. 2003. ICA using space estimates of entropy. J Machine Learning Res 4:1271–1295.
Morgenstern D. 1956. Einfache beispiele zweidimensionaler vertailungen. Mitt Math Stat 8:234–235.
Perkel DH, Bullock TH. 1968. Neural coding: a report based on a NRP work session. Neurosci Res Program Bull 6:221–248.
Rieke F, Warland D, de Ruyter van Steveninck R R, Bialek W. 1997. Spikes: exploring the neural code. Cambridge: MIT Press.
Sakai Y, Funanashi S, Shinomoto S. 1999. Temporally correlated inputs to leaky integrate-and-fire models can reproduce spiking statistics of cortical neurons. Neural Networks 12:1181–1190.
Shannon C. 1948. A mathematical theory of communication. AT&T Tech J 27:379–423.
Tarantola A. 1994. *The inverse problem theory*. Amsterdam: Elsevier.
Tarantola A, Valette B. 1982. Inverse problems = Quest for information. J Geophys 50:159–170.
Theunissen F, Miller JP. 1995. Temporal encoding in nervous systems: a rigorous definition. J Comput Neurosci 2:149–162.
Tsybakov AB, van der Meulen EC. 1996. Root-*n* consistent estimators of entropy for densities with unbounded support. Scand J Statist 23:75–83.
Vasicek O. 1976. A test for normality based on sample entropy. J Statist Soc B 38:54–59.
Zador A. 1998. Impact of synaptic unreliability on the information transmitted by spiking neurons. J Neurophysiol 79:1219–1229.

Available online at www.sciencedirect.com
ScienceDirect
BioSystems xxx (2006) xxx–xxx
Bio Systems
www.elsevier.com/locate/biosystems

Variability and randomness in stationary neuronal activity

Lubomir Kostal*, Petr Lánský[1]

Institute of Physiology, Academy of Sciences of Czech Republic, Videnska 1083, 142 20 Prague 4, Czech Republic

Received 29 November 2005; accepted 26 May 2006

Abstract

The patterns of neuronal activity can be different even if the mean firing rate is fixed. Investigating the variability of the firing may not be sufficient and we suggest to take into account the notion of randomness. The randomness is related to the entropy of the firing, which is bounded from above by the entropy of the Poisson process (given the mean interspike interval). Thus, we propose the Kullback–Leibler distance with respect to the Poisson process as a measure of randomness in a stationary neuronal activity. Under the condition of equal mean values the KL distance does not depend on the time scale and therefore can be compared to the coefficient of variation employed to measure the variability. Furthermore, this measure can be extended to account for correlated neuronal firing. Finally, we analyze the variability and randomness for three common ISI distributions in detail: gamma, lognormal and inverse Gaussian.

Keywords: Variability; Randomness; Entropy; Kullback–Leibler; Coefficient of variation

1. Introduction

The discharge activity of neurons is composed of the series of events called action potentials or spikes. It is generally accepted that the information in neuronal systems is transferred by the time series of spikes—the spike trains. There are two main hypotheses that attempt to classify possible ways in which the spike trains may carry information: the frequency (rate) codes and the temporal codes (Gerstner and Kistler, 2002; Theunissen and Miller, 1995). The classical results in early neuroscience (Adrian, 1928) show that the number of spikes per a time period (the firing rate) is related to the stimulus intensity, i.e., the firing rate increases with increasing stimulus intensity. The idea of temporal spike coding (Perkel and Bullock, 1968; Theunissen and Miller, 1995), on the other hand, employs the timing of the spikes or the particular ordering of interspike intervals. Whereas frequency code has quite specific meaning, the temporal code denotes all alternatives not classified as the former one. Therefore, the temporal coding involves on one hand precise patterns of spikes and on the other hand, for example, variability differences in the firing. Searching and comparing variability of different spike trains is a traditional tool in neuroscience studies. It holds for experimental as well as model spike trains and the most common way is by calculating the coefficient of variation (C_V) of interspike intervals (ISI).

The frequency codes in single neurons may carry information about both dynamic and stationary stimuli, see overview in Theunissen and Miller (1995). If the neuronal firing is stationary then the mean spike frequency (the inverse of the mean ISI, Lánský et al., 2004) carries the information from the frequency coding hypothesis point of view. The temporal coding, on the other hand, has been shown to occur almost exclusively under steady-state stimulus conditions (Fuller and Looft, 1984; Middlebrooks et al., 1994). Therefore, to classify

* Corresponding author. Tel.: +420 2 4106 2276; fax: +420 2 4106 2488.
[1] Tel.: +420 2 4106 2585; fax: +420 2 4106 2488.
E-mail address: kostal@biomed.cas.cz (L. Kostal), lansky@biomed.cas.cz (P. Lánský).

0303-2647/$ – see front matter © 2006 Published by Elsevier Ireland Ltd.
doi:10.1016/j.biosystems.2006.05.010

the spike trains solely from the temporal coding scheme point of view, one needs to describe differences between various stationary firing regimes with equal mean ISI.

The aim of this paper is to characterize the stationary neuronal firing. The method is based on a measure of randomness and we compare it with a measure of variability. The examples are restricted on the renewal spiking activity despite the fact that a general theory is available. The reason is that for the renewal model analytical results can be obtained, while there are no descriptions of correlated neuronal activity both realistic and mathematically suitable for our purpose.

2. Methods

2.1. Variability

Neuronal firing under stable conditions is often described as a renewal process of ISIs. In such a case the ISIs are mutually independent realizations of a positive random variable T and are fully characterized by the probability density function $f(t)$, where $f(t)\,\mathrm{d}t = \mathrm{Prob}(T \in [t, t+\mathrm{d}t))$ (Cox and Lewis, 1966). The renewal character of the ISIs implies stationarity of the neuronal activity. Often, though the neuronal firing is stationary, there is a dependency structure among the observed ISIs (Chacron et al., 2001; Longtin and Racicot, 1999). The dependence may arise due to the incomplete resetting of the membrane potential after the spike is emitted, which is experimentally observed especially in the distal parts of the neuron (Abeles, 1982). The other source of dependency may be a time structure in the input of the neuron. The successive ISIs $\{T_i\}$ are then statistically dependent, but due to the stationarity the expected value $E(T) = E(T_i)$ exists. The activity is fully described by the joint probability density function $f(t_1, t_2, \ldots)$, see e.g., Cox and Lewis (1966) for details.

The patterns of stationary neuronal activity may be strikingly different even if the mean firing rate, or equivalently the mean ISI, is fixed. The variability is probably the first issue to consider. It is often measured by employing the variance, $\mathrm{Var}(T)$, or the coefficient of variation, C_V, which relates variance to mean value, $C_\mathrm{V} = \sqrt{\mathrm{Var}(T)}/E(T)$. The main advantage of C_V over $\mathrm{Var}(T)$ – and the reason why it is used in data analysis – is that C_V does not depend on the 'scaling', $C_\mathrm{V}(aT) = C_\mathrm{V}(T)$. In other words, C_V is dimensionless and can be used to compare variability of spike trains with different mean ISI. For the Poisson process holds $C_\mathrm{V} = 1$ independently of $E(T)$. The differences in variability for several firing regimes are shown in Fig. 1A–C.

Even if the mean firing rate and variability of the neuronal firing are the same, the resulting spike trains may still have very different properties, compare Fig. 1A, D and E. The spike train in Fig. 1D is realized by the renewal process with Bernoulli distribution of ISIs and parameters chosen so that $C_\mathrm{V} = 1$ just as in the case of the Poisson process in Fig. 1A. In other words, though Poisson process implies $C_\mathrm{V} = 1$ the reverse implication does not hold. The example in Fig. 1E shows a spike train with $C_\mathrm{V} = 1$ again, but with properties significantly different from that of the Poisson firing (though the ISI densities of Fig. 1A and E are the same). The successive ISIs in Fig. 1E are not independent and the first-order serial correlation is $\rho = 0.86$. Thus, such a neuronal firing is not described by the renewal process. We may conclude by comparing spiking activities in Fig. 1A, D and E, that even though the variability is the same, the randomness of the firing can be different. We take this fact as an indication that the classification of stationary firing may be based also on different qualities than variability. In the following text we will precise the notion of randomness (or uncertainty) in neuronal activity.

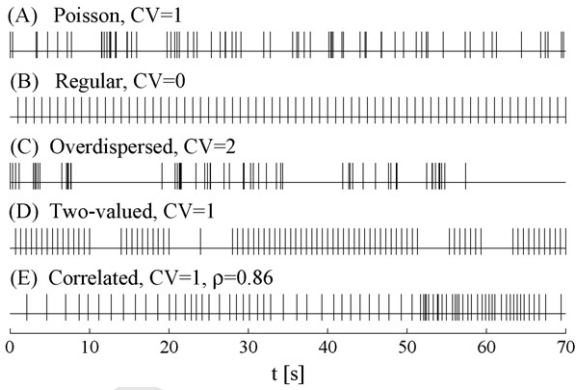


Fig. 1. Simulated spike trains illustrating the variability in stationary neuronal firing with $E(T) = 1$ s and different C_V. (A) Poisson process, $C_\mathrm{V} = 1$. (B) Regular firing, $C_\mathrm{V} = 0$. (C) Bursting neuronal activity (overdispersed firing), $C_\mathrm{V} > 1$. (D) Renewal process with Bernoulli distribution of ISIs, $C_\mathrm{V} = 1$. (E) activity 'derived' from the Poisson process (ISI densities are the same, thus $C_\mathrm{V} = 1$). The ISIs are ordered so that the serial correlation ρ between two successive ISIs is $\rho = 0.86$.

Next we introduce three "standard" renewal-process models of neuronal firing. These are gamma, inverse Gaussian and lognormal ISI distributions and we analyze their properties with respect to variability and randomness. The three mentioned distributions are fully determined by two parameters. We choose C_V as a parameter in order to employ variability directly. The C_V ranges from zero to infinity for all three mentioned models. We let the remaining parameter be the mean ISI, $\mu = E(T)$, which makes comparing distributions with equal $E(T)$ easier.

Gamma distribution is one of the most frequent statistical descriptors of ISIs (Hentall, 2000; Levine, 1991; McKeegan, 2002; Mandl, 1992; Reeke and Coop, 2004). Its probability density function is

$$f(t) = \left(\frac{1}{C_\mathrm{V}^2\mu}\right)^{1/C_\mathrm{V}^2} \Gamma(1/C_\mathrm{V}^2)\, t^{1/C_\mathrm{V}^2-1} \exp\left(-\frac{t}{C_\mathrm{V}^2\mu}\right), \quad (1)$$

where $\Gamma(z) = \int_0^\infty t^{z-1} \exp(-t)\,\mathrm{d}t$ is the gamma function. For $C_\mathrm{V} = 1$ it becomes exponential.

The inverse Gaussian distribution (Chhikhara and Folks, 1989) is often used to describe neural activity (Iynegar and

Please cite this article in press as: Lubomir, K., Petr L., Variability and randomness in stationary neuronal activity, BioSystems (2006), doi:10.1016/j.biosystems.2006.05.010

Liaom, 1997) and fitted to experimentally observed ISIs (Berger et al., 1990; Berger and Pribram, 1992; Gerstein and Mandelbrot, 1964; Levine, 1991). This distribution results from the Wiener process with positive drift (the depolarization has a linear trend to the threshold) and describes the spiking activity of non-leaky integrate-and-fire stochastic neuronal model (Ricciardi and Lánský, 2003). The probability density of the inverse Gaussian distribution can be expressed as

$$f(t) = \sqrt{\frac{\mu}{2\pi C_V^2 t^3}}\ \exp\left[-\frac{1}{2C_V^2\mu}\frac{(t-\mu)^2}{t}\right]. \quad (2)$$

The lognormal distribution of ISI, with some exceptions (Bershadskii et al., 2001), is rarely presented as a result of a neuronal model. However, it represents quite a common descriptor in ISI data analysis (Levine, 1991), e.g., a mixture of two lognormal distributions has been used recently (Bhumbra et al., 2004). It is given by the probability density function:

$$f(t) = \frac{1}{t\sqrt{2\pi \ln(1+C_V^2)}} \times \exp\left\{-\frac{1}{8}\frac{[\ln(1+C_V^2)+2\ln(t/\mu)]^2}{\ln(1+C_V^2)}\right\}. \quad (3)$$

Neither the inverse Gaussian nor the lognormal distribution is exponential for $C_V = 1$.

2.2. Randomness

The randomness of the renewal process with probability density function $f(t)$ can be judged by using the hazard rate $r(t)$:

$$r(t) = \frac{f(t)}{1-F(t)}, \quad (4)$$

where $F(t)$ is the cumulative distribution function $F(t) = \int_0^t f(z)\,dz$. The hazard rate determines the probability of spike occurrence in interval $[t, t+dt)$ under the condition that there was no firing in $[0, t)$. The most random firing is such that with elapsed time from the previous spike the probability of the next one does not change. It is well known that this holds for the renewal process with exponential distribution of ISIs. We denote the exponential probability density function as $g(t)$,

$$g(t) = \frac{1}{\mu}\,e^{-t/\mu}, \quad (5)$$

retaining the condition $E(T) = \mu$. The hazard rate for density (5) is then $r(t) = 1/\mu$.

Function $r(t)$ reflects the randomness of the renewal process but if we wish to relate it to the single value of C_V, we need to find a single-valued counterpart. The question how to measure the randomness of any renewal process with probability density $f(t)$ is answered by the concept of (differential) entropy, $h(f)$:

$$h(f) = -\int_0^\infty f(t)\ln f(t)\,dt. \quad (6)$$

The entropy $h(f)$ does not share the same properties and intuitive interpretation as the entropy H of a discrete probability mass function (Cover and Thomas, 1991). Namely, it can be negative and its value changes with a coordinate transform. Nevertheless, the most 'random' distribution is still the one that maximizes $h(f)$. The Poisson process thus represents the 'zero point' on the scale measuring the randomness of neuronal firing and we will relate it to any other stationary neuronal activity. It is not reasonable to choose regular spiking as the 'zero point' because the entropy h of the Dirac δ-distribution is $h = -\infty$ (Cover and Thomas, 1991). A measure $D(f, g)$ relating a renewal process with ISI probability density $f(t)$ to the Poisson process with the same mean value $\mu = E(T)$ is realized by the difference of the respective entropies:

$$D(f, g) = h(g) - h(f). \quad (7)$$

From Eqs. (5) and (6) follows that $h(g) = 1 + \ln\mu$ and then

$$D(f, g) = 1 + \ln\mu - h(f). \quad (8)$$

The proposed measure of randomness thus gives increasing values with decreasing randomness.

Formula (7) is related to the more general notion of Kullback–Leibler (KL) distance (relative entropy) of two probability density functions defined as (Cover and Thomas, 1991):

$$\mathrm{KL}(f, g) = \int_0^\infty f(t)\,\ln\frac{f(t)}{g(t)}\,dt. \quad (9)$$

Calculation shows immediately that if the mean values of $f(t)$ and $g(t)$ in formula (9) are the same and $g(t)$ is exponential then

$$\mathrm{KL}(f, g) = D(f, g). \quad (10)$$

Thus, the KL distance of a probability density function $f(t)$ from the exponential density under the condition of equal mean values can be used as a measure of randomness of a renewal neuronal activity.

Formula (10) can be extended to include any non-renewal stationary neuronal activity (see example in Fig. 1E). In such a case the activity is fully described by the joint probability density function $f(t_1, t_2, \ldots)$. The Kullback–Leibler distance per ISI then takes form:

$$\mathrm{KL}(f, g) = \lim_{n\to\infty}\frac{1}{n}\int_0^\infty\cdots\int_0^\infty f(t_1,\ldots,t_n) \times \ln\frac{f(t_1,\ldots,t_n)}{g(t_1,\ldots,t_n)}\,dt_1\cdots dt_n, \quad (11)$$

see Cover and Thomas (1991) for details. Formula (11) corresponds to the original definition (9) for the renewal process. Conditioning reduces entropy (Cover and Thomas, 1991) and thus the Poisson process maximizes entropy even in the generalized case. Letting $g(t_1, \ldots, t_n) = (1/\mu)^n \exp\left(-\sum_{i=1}^n t_i/\mu\right)$ in formula (11) and setting the mean values of $f(t_1, t_2, \ldots)$ and $g(t_1, t_2, \ldots)$ equal to $E(T)$ (this is possible because both activities are stationary) yields:

$$\mathrm{KL}(f, g) = h(g) - \bar{h}(f), \quad (12)$$

Please cite this article in press as: Lubomir, K., Petr L., Variability and randomness in stationary neuronal activity, BioSystems (2006), doi:10.1016/j.biosystems.2006.05.010

where

$$\bar{h}(f) = -\lim_{n\to\infty}\frac{1}{n}\int_0^\infty \cdots \int_0^\infty f(t_1,\ldots,t_n) \times \ln f(t_1,\ldots,t_n)\,\mathrm{d}t_1\cdots\mathrm{d}t_n. \quad (13)$$

Note that by taking the difference of two entropies in formula (7) instead of employing $h(f)$ directly, and by relating the result to the concept of the KL distance, several important issues are solved and some new properties emerge:

- $D(f, g)$ does not depend on coordinate transforms because in formula (11) both nominator and denominator are multiplied by the same factors.
- $D(f, g) \geq 0$ with equality if and only if $f(t)$ is exponential because $g(t)$ maximizes the entropy.
- $D(f, g)$ does not depend on $E(T)$ due to the invariance of KL distance to coordinate transforms. The logarithm of a time unit ($\ln E(T)$) "cancels out" as can be seen from the general formula (11).

3. Results

In this section we illustrate the application of formula (10) on the neuronal firing models given by Eqs. (1)–(3). Using formula (6), the entropy of gamma distribution (1) is

$$h(f) = \frac{1}{C_V^2} + \ln\left(\frac{1}{\mu C_V^2}\right) + \ln\,\Gamma\left(\frac{1}{C_V^2}\right) + \left(1-\frac{1}{C_V^2}\right)\Psi\left(\frac{1}{C_V^2}\right), \quad (14)$$

where $\Psi(z) = (\mathrm{d}/\mathrm{d}z)\ln\,\Gamma(z)$ is the digamma function. Combining Eqs. (8) and (14) we find the KL distance of the gamma distribution from the exponential one

$$\mathrm{KL}(C_V) = \ln\frac{\mathrm{e}}{C_V^2} - \ln\,\Gamma(1/C_V^2) + \frac{\Psi(1/C_V^2)-1}{C_V^2} - \Psi\left(\frac{1}{C_V^2}\right). \quad (15)$$

This result is illustrated in Fig. 2. Note that formula (15) does indeed not depend on $E(T)$ as mentioned before. The density f given by formula (1) is exponential for $C_V = 1$ and therefore $\mathrm{KL}(C_V = 1) = 0$. The KL distance tends to infinity for $C_V \to 0$ and $C_V \to \infty$. We can see from Fig. 2 that $\mathrm{KL}(C_V)$ increases rapidly for $C_V > 1$, especially if compared to the other models presented here. For $C_V < 0.25$ (approximately) the KL distances of gamma, lognormal and inverse Gaussian distributions become the same. The exponentiality of the gamma distribution for $C_V = 1$ and its difference from the Poisson process at $C_V = 2$ is illustrated using the hazard rates and probability density functions in Fig. 3.

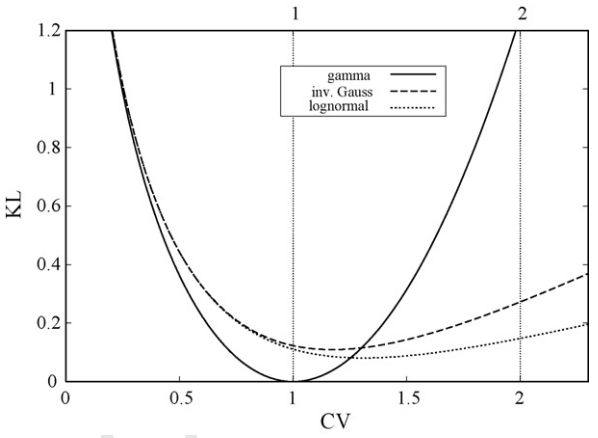


Fig. 2. The Kullback–Leibler (KL) distance as a function of C_V for three models of neuronal activity. The KL distance of the gamma distribution is zero for $C_V = 1$, implying that at this point it becomes exponential. Near $C_V = 1$ the values of KL distances are generally low. The minimum for the lognormal resp. inverse Gaussian distribution is located at $C_V \approx 1.31$, resp. $C_V \approx 1.17$. The distributions never become exponential. For $C_V \to 0$ and for $C_V \to \infty$ the KL distances tend to infinity. For C_V close zero the KL distances are initially the same. In general, low variability implies low randomness in the firing. On the other hand, the KL distances of the lognormal and inverse Gaussian grow very slowly with increasing C_V compared to the gamma distribution. This means that high variability may results in high as well as low randomness.

The identical approach as in the previous case reveals that the KL distance of the inverse Gaussian distribution (2) from the exponential one is

$$\mathrm{KL}(C_V) = \frac{1}{2}\ln\frac{\mathrm{e}}{2\pi C_V^2} + \frac{3\mathrm{e}^{1/C_V^2}}{\sqrt{2\pi C_V^2}}K_{1/2}^{(1,0)}\left(\frac{1}{C_V^2}\right), \quad (16)$$

where $K_\nu^{(1,0)}(z)$ is the derivative of the modified Bessel function of the second kind (Abramowitz and Stegun, 1972), $K_\nu^{(1,0)}(z) = (\partial/\partial\nu)K_\nu(z)$. The dependence is shown in Fig. 2. Due to the fact that the inverse Gaussian is never exponential, $\mathrm{KL}(C_V) > 0$. The minimum of $\mathrm{KL}(C_V)$ for the inverse Gaussian distribution is located at $C_V \approx 1.173$. We can see a difference compared to the gamma distribution. It has been already noted that the condition $C_V = 1$ does not imply exponentiality but in this case even the minimal distance is not located at $C_V = 1$, but at $C_V \approx 1.17$.

Finally, the KL distance of lognormal distribution (3) from the exponential one is

$$\mathrm{KL}(C_V) = \frac{1}{2}\left[\ln\frac{C_V^2+1}{\ln(C_V^2+1)} + \ln\frac{\mathrm{e}}{2\pi}\right] \quad (17)$$

Please cite this article in press as: Lubomir, K., Petr L., Variability and randomness in stationary neuronal activity, BioSystems (2006), doi:10.1016/j.biosystems.2006.05.010

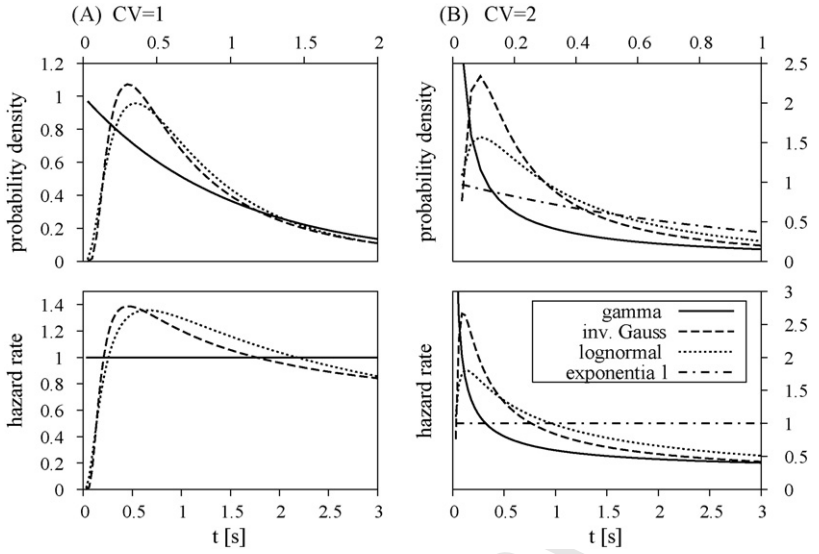


Fig. 3. Comparison of hazard rates and probability densities of the tested models with $E(T) = 1$ s for two values of C_V. (A) $C_V = 1$. The gamma distribution reduces to the exponential one. The hazard rates of inverse Gaussian and lognormal distributions are very similar in shape but none is constant. The lognormal and inverse Gaussian distributions can be hardly distinguished based on the hazard rates or densities and it is reflected in nearly equivalent values of the KL distance, see Fig. 2. (B) $C_V = 2$. The hazard rates are more different. In analogy with the KL distance, the lognormal is 'closest' to the exponential and gamma distribution is the most different. Similar (only less prominent) observation is yielded by comparing the probability density functions.

and the minimum is at $C_V = \sqrt{e - 1} \approx 1.311$. Again, the minimal possible deviation of lognormal distribution from exponential one is not at $C_V = 1$. It is interesting that for $C_V < 1$ (approximately) there is no difference in lognormal and inverse Gaussian distributions from the perspective of the KL distance. The equality in the KL distance, however, does not imply that these distributions are identical.

4. Conclusions

We demonstrated that variability in stationary neuronal firing is not sufficient to describe different firing regimes with equal mean ISI and that the notion of randomness brings an alternate point of view. The Kullback–Leibler (KL) distance was proposed as a measure of randomness with the exponential distribution being chosen as a template, because the exponential distribution is the most random one (maximizes entropy). Under the condition of equal mean values the KL distance resolves the problems of differential entropy and satisfies additional useful properties.

We concentrated mainly on the neuronal firing described by the renewal process and we analyzed three common two-parametric distributions using the proposed method: gamma, lognormal and inverse Gaussian.

The following inference can be made on the basis of the KL distance of ISI distributions:

1. The KL distances for all three investigated distributions is U-shaped with low values distributed around $C_V = 1$.
2. While small variability generally implies low randomness, high variability in the firing may result in both low as well as high randomness.
3. The same level of randomness in the firing can be obtained with different values of variability. Therefore, the notions of randomness and variability represent different descriptions of the neuronal activity.
4. It is well known that the lognormal and inverse Gaussian distributions never become exponential but in addition their minimal KL distances to this distribution are not located at $C_V = 1$.
5. For C_V increasing from zero (regular spiking) the KL distances of lognormal, inverse Gaussian and gamma distributions are initially the same. Then gamma branches off at $C_V \approx 0.25$ and the lognormal and inverse Gaussian depart at $C_V \approx 1$.
6. For lognormal and inverse Gaussian distributions the KL distance grows very slowly for $C_V > 1$, compared to the gamma distribution and their distances to the

Please cite this article in press as: Lubomir, K., Petr L., Variability and randomness in stationary neuronal activity, BioSystems (2006), doi:10.1016/j.biosystems.2006.05.010

exponential distribution are practically the same for $C_V = 1$ as for $C_V < 2$.

Acknowledgements

This work was supported by the Research project AV0Z 5011922, Center for Neuroscience LC554 and by Academy of Sciences of the Czech Republic Grant (Information Society, 1ET400110401).

References

Abeles, M., 1982. Local Cortical Circuits: An Electrophysiological Study. Springer-Verlag, Berlin, Heidelberg.

Abramowitz, M., Stegun, I.A., 1972. Handbook of mathematical functions. Dover, New York.

Adrian, E.D., 1928. The Basis of Sensation. W.W. Norton, New York.

Berger, D., Pribram, K., Wild, H., Bridges, C., 1990. An analysis of neural spike train distributions: determinants of the response of visual cortex neurons to changes in orientation and spatial frequency. Exp. Brain Res. 80, 129–134.

Berger, D., Pribram, K., 1992. The relationship between the gabor elementary function and a stochastic model of interspike interval distribution in the responses of visual cortex neurons. Biol. Cybern. 67, 191–194.

Bershadskii, A., Dremencov, E., Fukayama, D., Yadid, G., 2001. Probabilistic properties of neuron spiking time series obtained in vivo. Eur. Phys. J. B 24, 409–413.

Bhumbra, G.S., Inyushkin, A.N., Dyball, R.E.J., 2004. Assessment of spike activity in the supraoptic nucleus. J. Neuroendocrin. 16, 390–397.

Chacron, M.J., Longtin, A., Maler, L., 2001. Negative interspike interval correlations increase the neuronal capacity for encoding time-dependent stimuli. J. Neurosci. 21, 5328–5343.

Chhikhara, R.S., Folks, J.L., 1989. The Inverse Gaussian Distribution: Theory, Methodology and Applications. Marcel Dekker, Inc., New York.

Cover, T.M., Thomas, J.A., 1991. Elements of Information Theory. John Wiley & Sons, Inc., New York.

Cox, D.R., Lewis, P.A.W., 1966. The Statistical Analysis of Series of Events. John Wiley & Sons, Inc., New York.

Fuller, M.S., Looft, F., 1984. An information theoretic analysis of cutaneous receptor responses. IEEE Trans. Biomed. Eng. 314, 377–383.

Gerstein, G., Mandelbrot, B., 1964. Random walk models for the spike activity of a single neuron. Biophys. J. 4, 41–68.

Gerstner, W., Kistler, W., 2002. Spiking Neuron Models. Cambridge University Press, Cambridge.

Hentall, I.D., 2000. Interactions between brainstem and trigeminal neurons detected by cross-spectral analysis. Neuroscience 96, 601–610.

Iynegar, S., Liaom, Q., 1997. Modeling neural activity using the generalized inverse Gaussian distribution. Biol. Cybern. 77, 289–295.

Lánský, P., Rodriguez, R., Sacerdote, L., 2004. Mean instantaneous firing frequency is always higher than the firing rate. Neural Comput. 16, 477–489.

Levine, M.W., 1991. The distribution of the intervals between neural impulses in the maintained discharges of retinal ganglion cells. Biol. Cybern. 65, 459–467.

Longtin, A., Racicot, D.M., 1999. Assessment of linear and non-linear correlations between neural firing events. In: Cutler, C.D., Kaplan, D.T. (Eds.), Nonlinear Dynamics and Time Series: Building a Bridge between the Natural and Statistical Sciences, vol. 11. Fields Institute Communications, pp. 223–239.

Mandl, G., 1992. Coding for stimulus velocity by temporal patterning of spike discharges in visual cells of cat superior colliculus. Vision Res. 33, 1451–1475.

McKeegan, D.E.F., 2002. Spontaneous and odor evoked activity in single avian olfactory bulb neurons. Brain Res. 929, 48–58.

Middlebrooks, J.C., Clock, A.B., Xu, L., Green, D.M., 1994. A panoramic code for sound location by cortical neurons. Science 264, 842–844.

Perkel, D.H., Bullock, T.H., 1968. Neural coding: a report based on a NRP work session. Neurosci. Res. Program Bull. 6, 221–248.

Reeke, N.R., Coop, A.D., 2004. Estimating the temporal entropy of neuronal discharge. Neural Comp. 16, 941–970.

Ricciardi, L.M., Lánský, P., 2003. Diffusion models of neuron activity. In: Arbib, M.A. (Ed.), The Handbook of Brain Theory and Neural Networks, 2nd ed. MIT Press, Cambridge.

Theunissen, F., Miller, J.P., 1995. Temporal encoding in nervous systems: a rigorous definition. J. Comput. Neurosci. 2, 149–162.

Please cite this article in press as: Lubomir, K., Petr L., Variability and randomness in stationary neuronal activity, BioSystems (2006), doi:10.1016/j.biosystems.2006.05.010

Network: Computation in Neural Systems
Month?? 2007; ??(?): 1–13

informa
healthcare

Randomness and variability of the neuronal activity described by the Ornstein–Uhlenbeck model

LUBOMIR KOSTAL[1], PETR LANSKY[1], & CRISTINA ZUCCA[2]

[1]*Institute of Physiology, Academy of Sciences of Czech Republic, Videnska 1083, 142 20 Prague 4, Czech Republic and* [2]*Department of Mathematics, University of Torino, via Carlo Alberto 10, 10 123 Torino, Italy*

1 *(Received ■■ ■■ ■■; revised ■■ ■■ ■■; accepted ■■ ■■ ■■)*

Abstract

Normalized entropy as a measure of randomness is explored. It is employed to characterize those properties of neuronal firing that cannot be described by the first two statistical moments. We analyze randomness of firing of the Ornstein–Uhlenbeck (OU) neuronal model with respect either to the variability of interspike intervals (coefficient of variation) or the model parameters. A new form of the Siegert's equation for first-passage time of the OU process is given. The parametric space of the model is divided into two parts (sub-and supra-threshold) depending upon the neuron activity in the absence of noise. In the supra-threshold regime there are many similarities of the model with the Wiener process model. The sub-threshold behavior differs qualitatively both from the Wiener model and from the supra-threshold regime. For very low input the firing regularity increases (due to increase of noise) cannot be observed by employing the entropy, while it is clearly observable by employing the coefficient of variation. Finally, we introduce and quantify the converse effect of firing regularity decrease by employing the normalized entropy.

2 **Keywords:** ■■■

Introduction

Several approaches have been proposed in the literature to model the activity of a single neuron. Such models usually describe the evolution of the membrane potential and they can be either stochastic or deterministic (Tuckwell 1988;

Correspondence: Lubomir Kostal, Institute of Physiology, Academy of Sciences of Czech Republic, Videnska 1083, 142 20 Prague 4, Czech Republic. Tel: +420 2 4106 2276. Fax: +420 2 4106 2488. E-mail: kostal@biomed.cas.cz

ISSN 0954-898X print/ISSN 1361-6536 online
DOI: 10.1080/09548980701243134

Gerstner and Kistler 2002). The leaky integrate-and-fire concept is employed very often (for a recent review see Burkitt 2006). In one of the stochastic versions of this model the membrane potential evolution is described by the Ornstein–Uhlenbeck (OU) process bounded by a firing threshold. Another and simpler model, parallel to the OU process, is the perfect integrator described by a Wiener process with a positive drift. Due to its simplicity many results can be obtained in a closed form, while for the OU model only few analytical results are available.

The neuronal firing in both these models is viewed as a first-passage time of the membrane potential through a threshold S. The spike itself is considered to be a point event after which the membrane potential is reset. The spiking activity of the OU and Wiener models is therefore described by the renewal process of interspike intervals (ISIs) T with probability density function $f(t)$. The densities $f(t)$ for different parameter values are mutually compared by calculating the moments of the distribution. The advantage of using the approach based on the moments lies in the possibility of relating statistical characteristics with some key concepts of neuronal coding or neuronal properties. For example, the mean value $E(T)$ describes the neuronal firing from the rate coding hypothesis point of view (Gerstner and Kistler 2002). The characteristics based on the variance can be related to the variability coding hypothesis (Perkel and Bullock 1968), and may also provide some information about the metabolic efficiency of the neuronal coding (Laughlin 2001). Variability, reflected by the coefficient of variation, of ISIs generated by the OU model was recently a topic for a very extensive discussion initiated by Softky and Koch (1993). On the other hand the estimates of higher moments cannot be reliably determined from samples of relatively small size as is normal with the neuronal data. Thus the attempts to use higher moments are not as frequent as using mean and C_V (Han et al. 1998; Lewis et al. 2001; Ruskin et al. 2002; Shinomoto et al. 2002). Besides the approach based on statistical moments, the information-theoretic quantities are employed in analysis of neuronal signal in order to provide an alternative point of view. Recently, for example, the concept of the Kullback–Leibler (KL) distance has been utilized in Johnson et al. (2001) and DeWeese and Meister (1999), although for a different purpose than in this article.

The aim of our article is to discuss the properties of different firing regimes of the OU model by employing statistical characteristics of the resulting spike trains with respect to the two input model parameters – signal μ and 'noise' σ^2. Besides the mentioned $E(T)$ and C_V, we define the normalized entropy (related to the KL distance), η, as a measure of randomness of the neuronal firing. Among these three characteristics ($E(T)$, C_V, η), we concentrate mainly on the differences between C_V and η, i.e., the differences that go beyond the first moment. We continue the work started in Kostal and Lansky (2007) and show, that the notions of variability and randomness describe different qualities of the neuronal firing and that these terms cannot be interchanged generally. We also note at this point, that the relation between η and some recently proposed information measures (DeWeese and Meister 1999; Chacron et al. 2001, 2003) has been exploited in Kostal and Lansky (2006b). Our current study describes similar phenomena to those obtained in Lindner et al. (2002), though from a different point of view. The approach used here permits a global comparison of the ISI probability densities generated by the model. Further on, working in the parameteric space of the original model gives an

opportunity to judge if the features observed in the standardized form can be realized by real neurons.

Theory and methods

Classification of neuronal firing

One of the most important characteristics of the neuronal firing (besides the firing rate) is its variability. The variability is often described by the coefficient of variation, C_V, which relates SD to the mean value, $C_V = \sqrt{Var(T)}/E(T)$. The C_V is dimensionless and does not depend on the linear scaling of the random variable, $C_V(aT) = C_V(T)$. In this sense the C_V does not depend on the actual $E(T)$ and both these numbers provide separate views on the spiking activity. Nevertheless, $E(T)$ and C_V cannot be used to distinguish between two probability distributions that differ in higher than second moments.

Instead of looking for classifications based on higher moments, we offer a conceptually different approach based on the randomness of the firing. The measure of randomness of a random variable T with probability density $f(t)$ is given by the entropy, $h(f)$,

$$h(f) = -\int_0^\infty f(t) \ln f(t) dt, \tag{1}$$

see Cover and Thomas (1991) for details. The entropy can be seen as measuring the 'choice' in neuronal firing, i.e., its value decreases as the possible ISI lengths are subject to more constraints. The maximum entropy on $[0, \infty)$ for a fixed $E(T)$ is realized by the exponential distribution, $h(f) = 1 + \ln E(T)$. The particular value of $h(f)$, however, generally depends on $E(T)$. In order to make the entropy independent of the linear scaling (in the same way as C_V), we transform the original random variable T to $\Theta = T/E(T)$. The 'new' variable Θ is dimensionless with mean $E(\Theta) = 1$, and we denote its entropy as 'normalized entropy' $\eta(f)$. The normalized entropy is related to the original entropy of the unscaled variable T as

$$\eta(f) = h(f) - \ln E(T) \tag{2}$$

If the firing is regular, i.e., $f(t) = \delta(t - t_0)$, then $\eta(f) = -\infty$. On the other hand, the value of η is maximized for the exponential probability density function and in that case $\eta(f) = 1$ which identifies maximum randomness of the firing.

Normalized entropy (2) can be related to the KL distance, $D(f, g)$, of a general probability density function $f(t)$ to the exponential density, $g(t) = \lambda e^{-\lambda t}$, with the same mean, i.e., with $\lambda = 1/E(T)$

$$D(f, g) = \int_0^\infty f(t) \ln \frac{f(t)}{g(t)} dt = 1 + \ln E(T) - h(f) \tag{3}$$

Combining formulas (2) and (3) yields

$$\eta(f) = 1 - D(f, g) \tag{4}$$

Note that for the exponential distribution $C_V = 1$ holds, but the reverse statement is not valid. On the other hand, it follows from the properties of the KL distance (Cover and Thomas 1991) that $\eta(f) = 1$ if and only if $f(t)$ is exponential density. Thus, for $C_V \neq 1$ it always holds $\eta(f) < 1$. Finally, we note, that the (normalized) entropy can be estimated directly from experimental data, see, e.g., Tsybakov and Meulen (1996), Beirlant et al. (1997); for an overview of different methods.

The models

We assume that the membrane depolarization of a neuron is described by a stochastic process $X = \{X_t;\ t \geq 0\}$ and a spike is elicited any time the process crosses a constant boundary S from below, under the condition that the process is reset after each spike. The time between two consecutive firings of the neuron is identified with the first-passage time of the stochastic process through a threshold S, $T = \inf\{t > 0 \,|\, X_t \geq S,\ X_0 = x_0 < S\}$.

The OU model describes the membrane depolarization by the stochastic process that fulfills the stochastic differential equation

$$dX_t = \left(-\frac{X_t}{\theta} + \mu\right)dt + \sigma dW_t, \quad X_0 = 0 \tag{5}$$

where $W = \{W_t;\ t \geq 0\}$ is a standard Wiener process, $\theta > 0$ is the membrane time constant, the constant μ characterizes the net-neuronal input and $\sigma > 0$ is a further constant related with the variability of the neuronal input. Commonly, the constant μ is denoted as the signal and σ^2 as the noise. However, such a distinction may appear to be confusing, because generally, there is dependence between μ and σ^2 (Hanson and Tuckwell 1983; Lansky and Sacerdote 2001) and large values of μ imply relatively low values of σ^2 (Ditlevsen and Lansky 2005). The choice $X_0 = 0$ implies that the resting and resetting potentials are set to zero. The parameters S and θ are the intrinsic parameters of the model while μ and σ^2 depend on the activity of other neurons in a network (Tuckwell and Richter 1978). The parameters of the process determine two firing regimes, depending on the behavior of the model in the absence of noise ($\sigma^2 = 0$). If $\mu\theta > S$ (supra-threshold regime) the neuron is active also in absence of noise and the firing is regular. If $\mu\theta < S$ (sub-threshold regime) the neuron is silent in absence of noise. The intermediate situation corresponds to the threshold regime $\mu\theta = S$.

The Wiener model can be obtained by taking the limit $\theta \to \infty$ in formula (5). The membrane potential is then described by the Wiener process with drift $Y = \{Y_t;\ t \geq 0\}$, that fulfills the stochastic differential equation

$$dY_t = \mu dt + \sigma dW_t, \quad Y_0 = 0. \tag{6}$$

Due to the properties of the model (6) an additional condition $\mu > 0$ has to be imposed, otherwise T is not a proper random variable, $\mathrm{Prob}(T = \infty) > 0$. From formula (6) follows that the membrane potential always approaches and crosses the threshold linearly for any σ^2. From this point of view, the Wiener model is always operating in the supra-threshold regime.

The ISI probability density of the Wiener model is known in a closed form, and is equivalent to an inverse Gaussian density (Chhikhara and Folks 1989),

$$f_W(t) = \frac{S}{\sqrt{2\pi\sigma^2 t^3}} \exp\left\{-\frac{(S-\mu t)^2}{2\sigma^2 t}\right\}. \tag{7}$$

The ISI probability density of the OU model is known in a closed form only in the threshold regime $\mu\theta = S$

$$f_{\mathrm{OU}}(t) = \frac{2S}{\sqrt{\pi\sigma^2 t^3}} \frac{\exp(2t/\theta)}{(\exp(2t/\theta)-1)^{3/2}} \exp\left\{-\frac{S^2}{\sigma^2\theta(\exp(2t/\theta)-1)}\right\}. \tag{8}$$

In the remaining cases, i.e., $\mu\theta \neq S$, numerical techniques have to be employed. The mean and C_V of the Wiener model follow from formula (7),

$$E(T) = \frac{S}{\mu}, \quad C_V = \frac{\sigma}{\sqrt{\mu S}}. \tag{9}$$

For the OU model the first two moments can be written in several ways in terms of integrals (Siegert 1951; Keilson and Ross 1975) or in terms of series (Tuckwell and Cope 1980; Ricciardi et al. 1999). Nevertheless these expressions are difficult to handle especially from a numerical point of view.

Numerical procedures

The method used throughout the article for the numerical evaluation of the ISI probability density function $f_{\mathrm{OU}}(t)$ is based on the integral equation with non-singular kernel derived in (Buonocore et al. 1987; Ricciardi et al. 1999). Furthermore, $E(T)$ can be evaluated using a closed form expression or calculated numerically from $f_{\mathrm{OU}}(t)$ together with C_V. Normalized entropy given by formula (2) is estimated numerically by an approximation of the involved integrals using a trapezoidal rule. We explicitly note that a reliable numerical integration is possible because the probability density function can be numerically evaluated with sufficient precision in a sufficiently dense set of points.

A simulation technique for the computation of the entropy does not give reliable results due to the estimation of the density function via histograms that are not smooth and depend on the binning. Kernel density estimators are also not useful here since, even if they give smooth densities they are bin-width dependent.

Results and discussion

Throughout the article, we set $S = 10\,\mathrm{mV}$ and $\theta = 10\,\mathrm{ms}$ to make our results biologically plausible and comparable with the previously published studies (Tuckwell and Richter 1978; Kandel and Schwartz 1985; Inoue et al. 1995; Lansky and Rospars 1995; Stevens and Zador 1998; La Camera et al. 2004). Consequently, the threshold value of neuronal input is $\mu = 1\,\mathrm{mV/ms}$ and any smaller value results the sub-threshold regime independently on the value of σ^2.

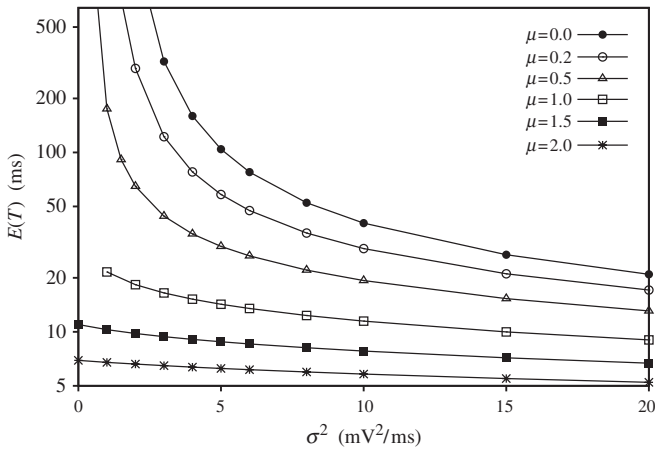


Figure 1. The dependence of the mean ISI upon the noise σ^2 for different values of μ for the OU model ($S = 10$ mV and $\theta = 10$ ms, μ is in mV/ms).

The maximum value of σ^2 we allow is $\sigma^2 = 40\,\mathrm{mV}^2/\mathrm{ms}$ which is within the range of the estimated values from experimental data (Inoue et al. 1995).

Relations between model parameters and statistical characteristics of the generated ISIs

In order to obtain better insight into the behavior of the OU model we first review the dependence between the two input parameters μ and σ^2 and the two statistical characteristics C_V and $E(T)$ in the sub-threshold and supra-threshold regimes. The Siegert formula (Siegert 1951) for the mean first-passage time of the OU process is

$$E(T) = \sqrt{\frac{\pi\theta}{\sigma^2}} \int_{-\mu\theta}^{S-\mu\theta} \exp\left(\frac{z^2}{\theta\sigma^2}\right)\left[1 + \mathrm{erf}\left(\frac{z}{\sigma\sqrt{\theta}}\right)\right] dz. \tag{10}$$

Equation 10 can be written as

$$E(T) = \frac{(S-\mu\theta)^2}{\sigma^2}\,{}_2F_2\left(1,\ 1; \frac{3}{2}, 2; \frac{(S-\mu\theta)^2}{\sigma^2\theta}\right) - \frac{\mu^2\theta^2}{\sigma^2}\,{}_2F_2\left(1, 1; \frac{3}{2}, 2; \frac{\mu^2\theta}{\sigma^2}\right) + \frac{\pi\theta}{2}\left[\mathrm{erfi}\left(\frac{\mu\sqrt{\theta}}{\sigma}\right) + \mathrm{erfi}\left(\frac{S-\mu\theta}{\sigma\sqrt{\theta}}\right)\right], \tag{11}$$

where ${}_2F_2(a_1, a_2; b_1, b_2; z)$ is the generalized hypergeometric function (Abramowitz and Stegun 1972) and $\mathrm{erfi}(z) = \mathrm{erf}(iz)/i$ is the imaginary error function. This expression is particularly useful for numerical evaluation since the involved special functions can be implemented with sufficient precision.

As shown in Figure 1 the qualitative dependence of $E(T)$ on σ^2 is monotonous for all values of parameter μ. With decreasing σ^2 the $E(T)$ increases to infinity in threshold ($\mu = 1$) and sub-threshold ($\mu < 1$) regimes and to a constant in supra-threshold ($\mu > 1$) regime as expected intuitively. With increasing σ^2 the mean ISI always tends to zero, which is shown up to 5 ms.

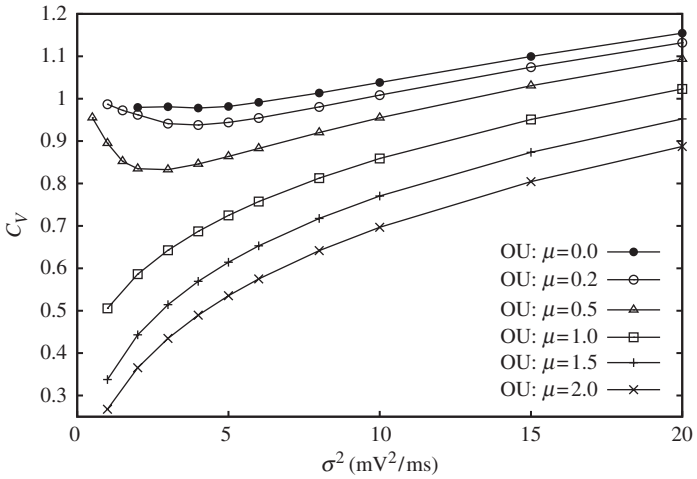


Figure 2. Coefficient of variation C_V for the OU model in dependency on the noise σ^2 with different values of the input parameter μ (in mV/ms).

The relation between C_V and σ^2 is shown in Figure 2. We can see again the different behavior between the two regimes in the OU model. In the supra-threshold regime as $\sigma^2 \to 0$ we get $C_V \to 0$ that means absence of variability, indeed the ISIs are almost constant. In the sub-threshold regime it holds $C_V \to 1$ for $\sigma^2 \to 0$. As expected, the density of T converges to an exponential density and this is illustrated by C_V close to one. With σ^2 increasing, we notice the local decrease of C_V. This effect is often denoted as the coherence resonance (Lindner et al. 2002).

In Figure 3 the dependence between $E(T)$ and C_V is illustrated. The dependency is always plotted for fixed μ and varying σ^2. In the supra-threshold regime as C_V increases $E(T)$ monotonically decreases. On the other hand, in the sub-threshold regime C_V does not determine $E(T)$ uniquely, i.e., it is impossible to determine $E(T)$ and σ^2 only from μ and C_V. However, it is possible to determine μ and C_V from $E(T)$ and σ^2 as Figure 1 shows. Comparison of Figures 1, 2 and 3 reveals, that knowledge of (almost) any two values in the quadruplet (μ, σ^2, $E(T)$, C_V) uniquely determines the remaining two. The only exception is that ($E(T)$, σ^2) cannot be determined from (μ, C_V) due to the coherence resonance effect (local C_V decrease) described later.

Randomness and moment characteristics of firing

We examine the behavior of randomness by employing the normalized entropy in dependency on the C_V. (Recall that the effect of different $E(T)$ values is removed for both of these measures.) For the Wiener model the normalized entropy can be written in terms of C_V as follows from Kostal and Lansky (2006a)

$$\eta(f_W) = \frac{1}{2} + \frac{1}{2}\ln(2\pi C_V^2) - \frac{3e^{1/C_V^2}}{\sqrt{2\pi C_V^2}} K^{(1,0)}_{(1/2)}\left(\frac{1}{C_V^2}\right), \tag{12}$$

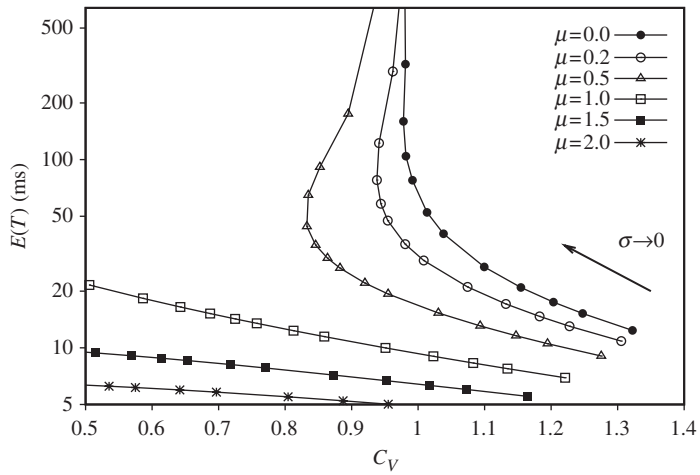


Figure 3. The mean ISI in dependency on C_V (along the curves from right to left σ^2 decreases) for different values of μ for the OU model ($S = 10$ mV and $\theta = 10$ ms, μ is in mV/ms). The maximum value of σ^2 is 40 mV^2/ms. Note the non-unique relation between $E(T)$ and C_V due to the coherence resonance effect (local decrease in firing variability with increasing input noise) in the sub-threshold regime.

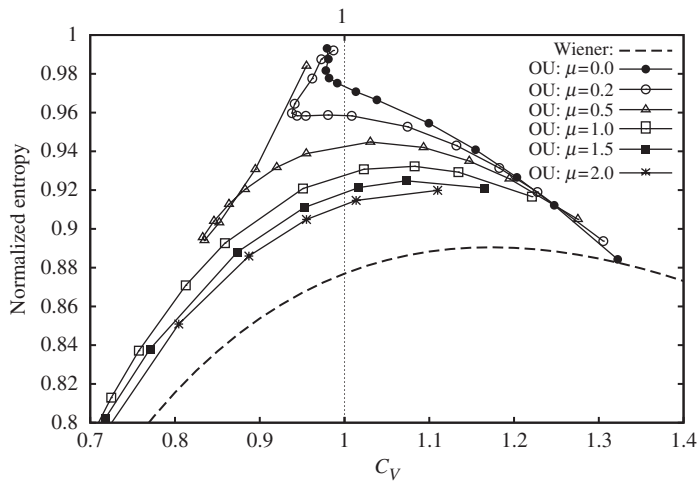


Figure 4. Normalized entropy in dependency on the coefficient of variation C_V with different values of the input parameter μ ($S = 10$ mV and $\theta = 10$ ms). The maximum value of σ^2 is 40 mV^2/ms. The dashed line represents the normalized entropy of the Wiener model.

where $K_\nu^{(1,0)}(z)$ is the derivative of the modified Bessel function of the second kind (Abramowitz and Stegun 1972). For the OU model only numerical procedure is available.

The results are shown in Figure 4 for the Wiener model and both sub and supra-threshold regimes of the OU model. We see that maximum randomness does not coincide with maximum variability. (Note, that there are distributions with

maximum variability and randomness coinciding, i.e., Pareto distribution, see Kostal and Lansky 2006b for details). Similarly, we see that $C_V = 1$ does not imply maximum randomness. For the OU model in the sub-threshold regime C_V does not determine the randomness (and thus the shape of the ISI probability density) uniquely if μ is fixed (see Figure 2). We can deduce from Figure 4 that there are always two different shapes of the probability density functions with C_V close to unity: one which is very close to the exponential density (maximum randomness) and the second one further away. With small input μ the C_V cannot be made deliberately small regardless of the noise amplitude. As the input in the sub-threshold regime increases the minimal value of C_V, $C_V^{(\min)}$, decreases. Note that around $C_V = 1.25$ the values of η are nearly the same for all sets of parameters of the sub-threshold OU model.

The effect of coherence resonance can be observed in both measures, C_V and η. However, it follows from the picture that the coherence resonance can be reliably observed in randomness only for $\mu > 0.2$. This fact further enhances the difference between variability and randomness, i.e, the increase of regularity (as measured by C_V) does not necessarily imply the decrease in randomness of the spike train. Namely, few sufficiently long ISIs in otherwise 'almost' regular spiking activity result in a high variability although the randomness may be low.

For the case of $\mu \geq 1$ (the threshold and supra-threshold regimes) the behavior of the model is less complicated and all C_V values are obtainable. Furthermore, it always holds $\eta \ll 1$, i.e., the ISI probability density is never close to the exponential distribution. The curves $\eta(f_{\mathrm{OU}})$ parameterized by the C_V in the supra-threshold regime have a very similar shape and are similar to the Wiener model as expected. For increasing μ the normalized entropy of the OU model converges to the entropy of the Wiener model and the shape of $\eta(f_{\mathrm{OU}})$ is less convex. This similarity holds only for small C_V. For $C_V > 1.4$ the randomness of the OU model is always lower than the randomness of the Wiener model.

Randomness and input parameters

In this section, we explore randomness in the Wiener and OU models with respect to the model parameters (μ, σ^2). The normalized entropy for the Wiener model is given in terms of (S, μ, σ^2) by combining formulas (9) and (12). For the OU model, we can write the normalized entropy only in the threshold regime $(\mu = 1)$ and we get an expression in terms of the parameters (S, θ, σ^2)

$$\eta(f_{\mathrm{OU}}) = \frac{1}{2} + \frac{3}{2}\left[\gamma + \ln\left(\frac{4S^2}{\sigma^2\theta}\right)\right] - \ln\left(\frac{2S}{\sqrt{\pi\sigma^2\theta^3}}\right) - \frac{2}{\theta}E(T) - \ln E(T), \qquad (13)$$

where $E(T)$ is given by formula (11) and $\gamma \approx 0.577$ is the Euler Gamma constant (Abramowitz and Stegun 1972). For the threshold regime, we checked the agreement between η given by formula (13) and its numerical estimation obtained directly from the probability density function.

The results relating randomness to μ and σ^2 are shown in Figure 5. For $\mu = 0$ the dependence $\eta(\sigma^2)$ is almost linear and decreases rather slowly. The nonlinearity is more pronounced for increasing values of μ. For $\mu < 1$ as $\sigma^2 \to 0$ the normalized entropy tends to 1, it means that the probability density function is

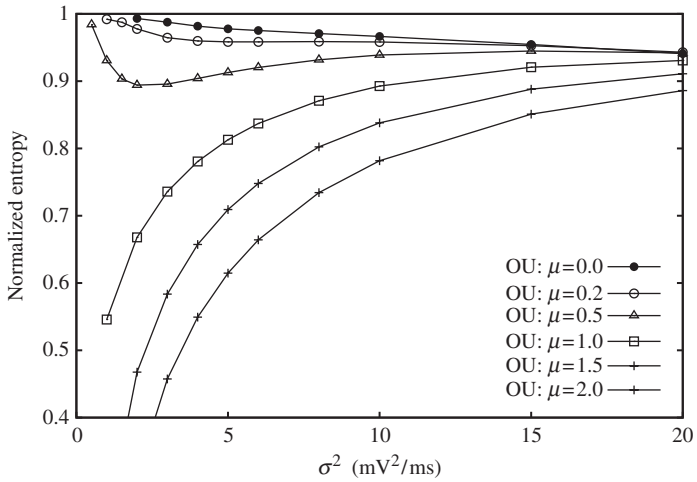


Figure 5. Normalized entropy as a function of the noise σ^2 for different values of parameter μ of the OU model ($S = 10$ mV and $\theta = 10$ ms). Note the similarity with Figure 2, however, the local decrease in C_V in the sub-threshold regime is not prominent for μ smaller than 0.2 mV/ms.

getting exponential. Then as σ^2 increases the randomness decreases to its local minimum (coherence resonance) and then increases again to its local maximum and it slowly decreases again. On the other hand, for $\mu \geq 1$ the firing gets more regular in the case of $\sigma^2 \to 0$ which is reflected by $\eta \to -\infty$. The local maxima of η occur for higher values of σ^2 with increasing μ.

Though Figure 5 looks very similar to Figure 2 there are substantial differences between the measures C_V and η. To demonstrate this, we plot the dependence of $E(T)$ on η in Figure 6 and compare it with a similar plot for C_V in Figure 3. The behavior of normalized entropy is more complex. The coherence resonance is present in the sub-threshold regime only for $\mu > 0.2$ as already mentioned. Moreover, the relation between $E(T)$ and η is non-unique even in the supra-threshold regime. For each $\mu > 0.2$, we observe a local increase in randomness (and irregularity of the firing) in dependence on $E(T)$ (or σ^2) even though C_V is monotonous in the supra-threshold regime.

Conclusions

The normalized entropy as a measure of randomness was introduced. We applied it to the neuronal activity described by the OU and Wiener models. In particular, we described the firing characteristics that go beyond the first statistical moment and analyzed the randomness of the ISIs with respect to the coefficient of variation and the mean ISI (the statistical approach) or with respect to the model parameters (the modeling approach).

The behavior of the OU model is qualitatively different in the sub- and supra-threshold regimes. In the sub-threshold regime C_V does not determine uniquely the

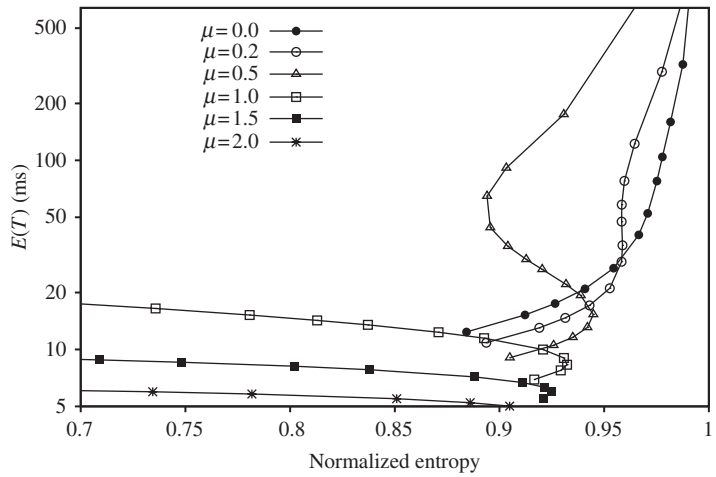


Figure 6. The mean ISI in dependency on the normalized entropy (compare with Figure 3). The maximum value of σ^2 is 40 mV^2/ms. Note the non-unique relationship in the supra-threshold regime and local increase of randomness present for regimes with μ greater than 0.2 mV/ms.

shape of the ISI probability density function even if the model parameters (except the input noise) are fixed. For both sub-and supra-threshold regimes, we identified such regions of C_V and model parameter values for which the randomness can be considered independent on the neuronal input. Finally, we noted that the local decrease in variability (with respect to the noise) for very low neuronal values input is not reliably observable by employing the normalized entropy of the firing, which demonstrates the key difference between randomness and variability. Moreover, we described the converse effect (the local decrease in regularity of the firing), which on the other hand cannot be observed by employing the notion of variability.

Acknowledgments

This work was supported by the Research project AV0Z50110509, Center for Neuroscience LC554, Academy of Sciences of the Czech Republic Grants (1ET400110401 and KJB100110701) and by MIUR project ''Mathematical methods to analyze the information content of interspike time series in small simulated or monitored neural networks.'' PRIN-Cofin 2005.

References

Abramowitz M, Stegun IA. 1972. Handbook of mathematical functions. New York: Dover.

Beirlant J, Dudewicz EJ, Gyorfi L, van der Meulen EC. 1997. Nonparametric entropy estimation: An overview. Int J Math Stat Sci 6:17–39.

Buonocore A, Nobile AG, Ricciardi LM. 1987. A new integral equation for the evaluation of FPT probability densities. Adv Appl Prob 19:784–800.

Burkitt AN. 2006. A review of the integrate-and-fire neuron model: I. Homogeneous synaptic input. Biol Cybern 95:1–19.

Chacron MJ, Longtin A, Maler L. 2001. Negative interspike interval correlations increase the neuronal capacity for encoding time-dependent stimuli. J Neurosci 21:5328–5343.

Chacron MJ, Longtin A, Maler L. 2003. The effects of spontaneous activity, background noise, and the stimulus ensemble on information transfer in neurons. Network: Comput Neural Syst 14:803–824.

Chhikhara RS, Folks JL. 1989. The inverse Gaussian distribution: Theory, methodology and applications. New York: Marcel Dekker, Inc.

Cover TM, Thomas JA. 1991. Elements of information theory. New York: John Wiley & Sons, Inc.

Cox DR, Lewis PAW. 1966. The statistical analysis of series of events. New York: John Wiley & Sons, Inc.

DeWeese MR, Meister M. 1999. How to measure the information gained from one symbol. Network: Comput Neural Syst 10:325–340.

Ditlevsen S, Lansky P. 2005. Estimation of the input parameters in the Ornstein-Uhlenbeck neuronal model. Phys Rev E 71:011907.

Gerstner W, Kistler WM. 2002. Spiking neuron models: Single neurons, populations, plasticity. ▮▮: Cambridge University Press.

Han YMY, Chan YS, Lo KS, Wong TM. 1998. Spontaneous activity and barosensitivity of the barosensitive neurons in the rostral ventrolateral medulla of hypertensive rats induced by transection of aortic depressor nerves. Brain Res 813:262–267.

Hanson FB, Tuckwell HC. 1983. Diffusion Approximation for neuronal activity including reversal potentials. J Theor Neurobiol 2:127–153.

Inoue J, Sato S, Ricciardi LM. 1995. On the parameter estimation for diffusion models of single neuron's activities. I. Application to spontaneous activities of mesencephalic reticular formation cells in sleep and waking states. Biol Cybern 73:209–221.

Johnson DH, Gruner CM, Baggerly K, Seshagiri C. 2001. Information-theoretic analysis of the neural code. J Comput Neurosci 10:47–69.

Kandel ER, Schwartz JH. 1985. Principles of neural science. 2nd ed. New York: Elsevier.

Keilson J, Ross HF. 1975. Passage time distributions for Gaussian Markov (Ornstein-Uhlenbeck) statistical processes. Selected Tables in Mathematical Statistics 3:233–328.

Kostal L, Lansky P. 2006a. Similarity of interspike interval distributions and information gain in a stationary neuronal firing. Biol Cybern 94:157–167.

Kostal L, Lansky P. 2006b. Classification of stationary neuronal activity according to its information rate. Network: Comput Neural Sys 17:193–210.

Kostal L, Lansky P. 2007. Variability and randomness in stationary neuronal activity, Biosystems, in print.

La Camera G, Rauch A, Luscher HR, Senn W, Fusi S. 2004. Minimal models of adapted neuronal response to in vivo-like input currents. Neural Comput 16:2101–2124.

Lansky P, Rospars J-P. 1995. Ornstein-Uhlenbeck neuronal model revisited. Biol Cybern 72:397–406.

Lansky P, Sacerdote L. 2001. The Ornstein-Uhlenbeck neuronal model with the signal-dependent noise. Phys Lett A 285:132–140.

Laughlin SB. 2001. Energy as a constraint on the coding and processing of sensory information. Curr Opin Neurobio 11:475480.

Lewis CD, Gebber GL, Larsen PD, Barman SM. 2001. Long-term correlations in the spike trains of medullary sympathetic neurons. J Neurophysiol 85:1614–1622.

Lindner B, Schimansky-Geier L, Longtin A. 2002. Maximizing spike train coherence or incoherence in the leaky integrate-and-fire neuron. Phys Rev E 66:031916 - 2.

Perkel DH, Bullock TH. 1968. Neural coding: A report based on a NRP work session. Neurosci Res Program Bull 6:221–248.

Ricciardi LM, Di Crescenzo A, Giorno V, Nobile AG. 1999. An outline of theoretical and algorithmic approaches to first passage time problems with applications to biological modeling. Math Jap 50:247–322.

Ruskin DN, Bergstrom DA, Walters JR. 2002. Nigrostriatal lesion and dopamine agonists affect firing patterns of rodent entopeduncular nucleus neurons. J Neurophysiol 88:487–496.

Siegert AJF. 1951. On the first passage time probability problem. Phys Rev 81:617–623.

Shinomoto S, Yutaka S, Hiroshi O. 2002. Recording site dependence of the neuronal spiking statistics. Biosystems 67:259–263.

Softky WR, Koch C. 1993. The highly irregular firing of cortical cells is inconsistent with temporal integration of random EPSP's. J Neurosci 13:334–530.

Stevens C, Zador A. 1998. Novel integrate-andfire -like model of repetitive firing in cortical neurons. In: Proceeding of the 5th Joint Symposium on Neural Computation 8:172–177.

Tsybakov AB, van der Meulen EC. 1996. Root-*n* consistent estimators of entropy for densities with unbounded support. Scand J Statist 23:75–83.

Tuckwell HC. 1988. Introduction to theoretical neurobiology. ▮▮: Cambridge University Press.

Tuckwell HC, Cope DK. 1980. Accuracy of neuronal interspike times calculated from a diffusion approximation. J Theor Biol 83:377–87.

Tuckwell HC, Richter W. 1978. Neuronal interspike time distributions and the estimation of neurophysiological and neuroanatomical parameters. J Theor Biol 71:167–183.