Review: Neuronal coding and spiking randomness

Lubomir Kostal[a)] and Petr Lansky
Institute of Physiology of the Czech Academy of Sciences, Videnska 1083, 14220 Prague 4,
Czech Republic
Jean-Pierre Rospars
UMR1272 UMPC-INRA-AgroParisTech "Physiologie de l'Insecte: Signalisation et Communication", INRA,
78026 Versailles Cedex, France

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. In this paper we review recent results on the notion of spiking randomness and discuss its properties with respect to the rate and temporal coding schemes. This method is compared with other widely used characteristics of spiking activity, namely the variability of interspike intervals and it is shown that randomness and variability provide two distinct views. We demonstrate, that estimation of spiking randomness from simulated and experimental data is capable of capturing characteristics that would otherwise be difficult to obtain with conventional methods.

Keywords: spike train, entropy, variability, randomness

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 and 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 and Kistler, 2002; Shadlen and Newsome, 1998; Stein, Gossen, and Jones, 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 and Lewis, 1966; Kass, Ventura, and Brown, 2005; Moore, Perkel, and Segundo, 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 and Bullock, 1968; Softky, 1995; Strong *et al.*, 1998). To answer this question, methods to compare different spike trains are needed first (Bhumbra, Inyushkin, and Dyball, 2004; Buracas and Albright, 1999; Nemenman, Bialek, and de Ruyter van Steveninck, 2004; Paninski, 2003; Rieke *et al.*, 1997; Victor and Purpura, 1997).

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, Schwartz, and Jessel, 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, Gossen, and Jones, 2005). The question whether 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 and Kistler, 2002; Shadlen and Newsome, 1994; Stein, Gossen, and Jones, 2005) and leads to the idea of temporal coding.

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 and Albright, 1999; Gerstner and Kistler, 2002; Rieke *et al.*, 1997). Possibility of information transmission by changes in ISIs serial correlation has been reported in crayfish interneurons (Sugano and Tsukada, 1978; Wiersma and Adams, 1950). For a classic overview of temporal coding see Perkel and Bullock (1968), for a more recent discussion see Abeles (1994); Rieke *et al.* (1997); Shadlen and Newsome (1994); Stein, Gossen, and Jones (2005); Theunissen and Miller (1995). It is worth denoting, that both rate and temporal coding schemes generally depend on the length of the observation window and therefore their precise separation may not be possible in general (Gerstner and Kistler, 2002). However, if the firing is stationary the rate does not depend on the observation window length and thus both coding schemes are well separated.

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

---

[a)]E-mail: kostal@biomed.cas.cz

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 review 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 of capturing 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 and Weaver, 1998), relation with other information-theoretic quantities can be established.

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, Perkel, and Segundo, 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 and Lewis, 1966; Landolt and Correia, 1978). Stationary neuronal firing is typically observed in the spontaneous activity, or under constant stimulus conditions (Gerstner and Kistler, 2002; Moore, Perkel, and Segundo, 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 and 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, Perkel, and Segundo, 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 and Lewis, 1966; Landolt and 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 and Lewis, 1966; Duchamp-Viret et al., 2005), which serves as a basis for some differential entropy estimators (see Appendix B). 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, Goldberg, and Smith, 1964; Moore, Perkel, and Segundo, 1966; Poggio and 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 $\mathrm{Var}(T)$ estimated by formulas (1) and (2). Such firing is called renewal process of ISIs (Cox and Lewis, 1966; Gerstner and 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 and Kistler, 2002; Landolt and Correia, 1978; Stein, 1967; Tuckwell, 1988). Sometimes, however, there might be a dependency structure between the observed ISIs (Chacron, Longtin, and Maler, 2001; Lansky and Rodriguez, 1999; Lindner, 2004; Longtin and Racicot, 1997; Ratnam and Nelson, 2000; Sakai, Funahashi, and Shinomoto, 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). 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 and Kistler, 2002; Moore, Perkel, and Segundo, 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 C for description of models employed in the figure.
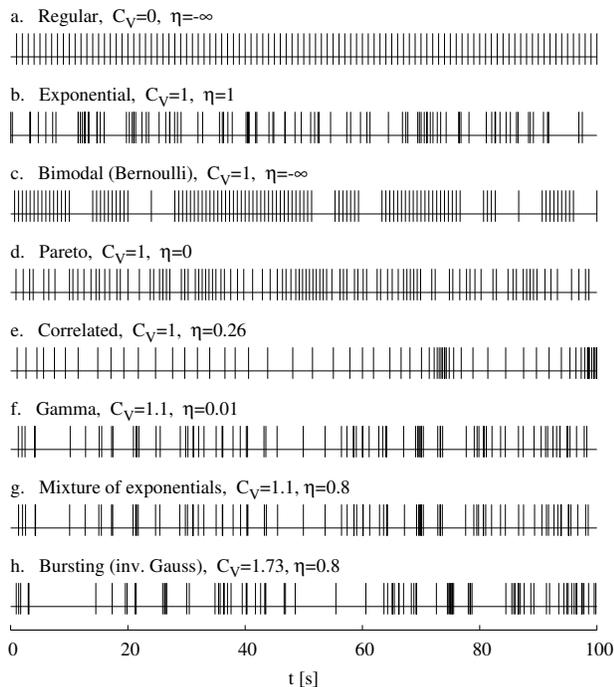
a. Regular, $C_V=0$, $\eta=-\infty$

b. Exponential, $C_V=1$, $\eta=1$

c. Bimodal (Bernoulli), $C_V=1$, $\eta=-\infty$

d. Pareto, $C_V=1$, $\eta=0$

e. Correlated, $C_V=1$, $\eta=0.26$

f. Gamma, $C_V=1.1$, $\eta=0.01$

g. Mixture of exponentials, $C_V=1.1$, $\eta=0.8$

h. Bursting (inv. Gauss), $C_V=1.73$, $\eta=0.8$

t [s]

Figure 1. Examples of different simulated spike trains. Mean interspike interval is $E(T) = 1$ 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.

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, $\mathrm{Var}(T)$, 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,

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

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, $\mathrm{Var}(\Theta) = C_V$, where

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

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 and Koch, 1993). Furthermore, the C_V of ISIs is related to the variability coding hypothesis (Perkel and Bullock, 1968). The coding characterized by C_V has been hypothetised to transmit information about light intensity in adapted cells of the horseshoe crab (Ratliff, Hartline, and Lange, 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 and Pritchard, 1964) and also in hippocampal place cells (Fenton, Lansky, and Olypher, 2002). The variability of ISIs generated by the leaky integrate-and-fire model (Burkitt, 2006; Gerstner and Kistler, 2002) was recently a topic for a very extensive discussion initiated by Softky and Koch (1993).

Nevertheless, $E(T)$ and C_V are not sufficient to describe all possible differences between spike trains (Fig. 1b–e, where $E(T) = 1$ 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. Additionally, 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.

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 and 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 $p_i = \mathrm{Prob}\{X = x_i\}$, the entropy $H(X)$ is defined as (Shannon and Weaver, 1998)

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

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 and Thomas, 1991). The entropy is a unique measure of randomness satisfying a set of intuitive conditions (Jaynes and Bretthorst, 2003; Shannon and 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 and Thomas, 1991). Therefore, the differential entropy $h(T)$ of the ISI probability density function $f(t)$ is defined as

$$h(T) = -\int_0^\infty f(t) \ln f(t)\, \mathrm{d}t. \qquad (8)$$

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, Longtin, and Maler, 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, Lansky, and Zucca, 2007)

$$\eta = h(T) - \ln E(T). \qquad (9)$$

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

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

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 and 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 and Thomas, 1991)

$$D(f, f_{exp}) = \int_0^\infty f(t) \ln \frac{f(t)}{f_{\exp}(t)}\, \mathrm{d}t \qquad (11)$$

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 and Lansky, 2007). It can be shown that η is related to $D(f, f_{\exp})$ by a simple formula (Kostal, Lansky, and Zucca, 2007)

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

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 and Lansky, 2006a), see Appendix A.

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 and Lansky, 2007). Second, while η in its general form accounts also for non-renewal spiking activity, $r(t)$ is used only in the renewal case (Moore, Perkel, and Segundo, 1966).

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, Lansky, and Zucca, 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 and 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, Lansky, and Zucca, 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 and 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 and 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 and 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, Warnke, and Timmer, 1998; Quiroga et al., 2000; Thakor and Tong, 2004).

Results on 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 and 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 and Lansky (2006b, 2007); Kostal, Lansky, and Zucca (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 and 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, Lansky, and Zucca, 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, Lansky, and Zucca, 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 and 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, Schimansky-Geier, and Longtin, 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, Lansky, and Zucca, 2007).

Several statistical models of non-renewal spiking activity described by first-order Markov chains (Cox and Lewis, 1966) were examined in Kostal and Lansky (2006a). Markov structure in experimental data is reported in literature (Ratnam and
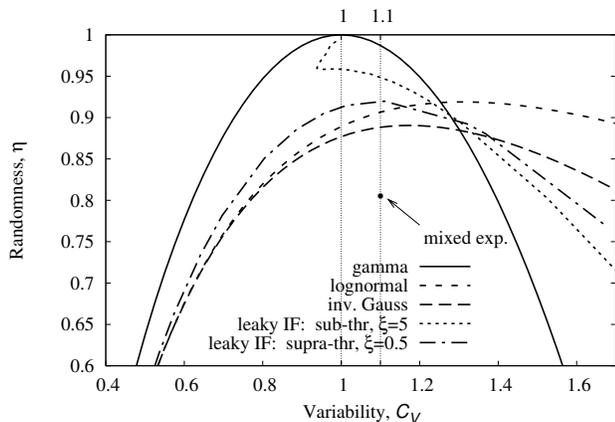
Figure 2. Randomness vs. variability for some widely used renewal models of neuronal activity. For $C_V = 1$ the gamma distribution becomes exponential and thus $\eta = 1$. The inverse Gaussian and lognormal models never become exponential their maximum randomness (minimal KL distance from the exponential 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.

Nelson, 2000), and even the first-order case makes the existence of certain (short) sequences of ISIs more probable than others (basic temporal pattern formation, Kostal and Lansky (2006a)). Although the examined models were not used in data analysis previously, the results show that the serial correlation coefficient (Cox and 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 and 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 and 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, $\text{Prob}\{T_n \le t_n | T_{n-1} = t_{n-1}, \ldots, T_1 = t_1\} = \text{Prob}\{T_n \le t_n | T_{n-1} = t_{n-1}\}$ (Cox and Lewis, 1966), does not hold in this case.

Results on 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, Inyushkin, and Dyball, 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 and Smith, 1965; Tuckwell, 1988)

$$f(t) = pae^{-ax} + (1-p)be^{-bx}, \qquad (13)$$

where $p \in (0, 1)$ and $a > 0$, $b > 0$, $a \ne b$. 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 and Lansky, 2006b). We compare two simulated spike trains with $E(T) = 1$ 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 in this case is $\eta = 0.99$. The parameters in formula (13) 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 (B-1) (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 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
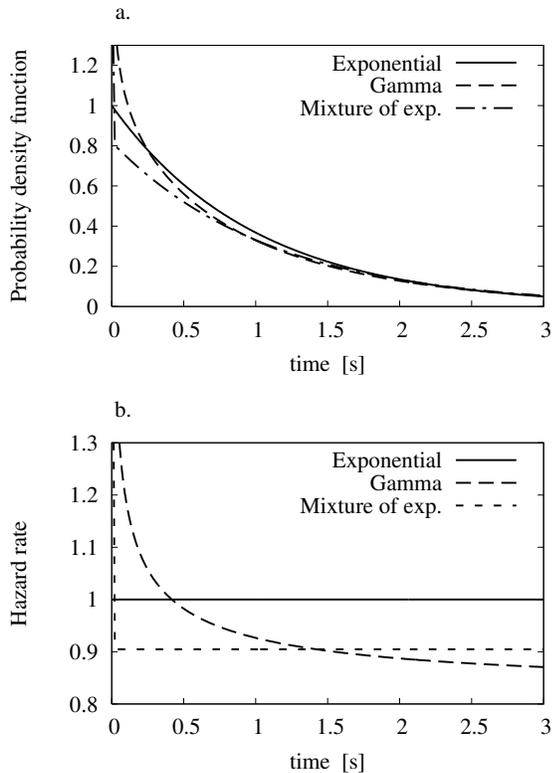
7

a.

b.

Figure 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.

$\eta = 0.8$ 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 and Lansky, 2006a).

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.

Acknowledgements

Appendix A: 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 and Lewis, 1966). The appropriate generalization of differential entropy $h(T)$ is given by the differential entropy rate $\bar{h}(f)$ (Cover and Thomas, 1991)

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

Equation (A-1) 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 and Thomas, 1991). Equation (A-1) then reads

$$\bar{h}(f) = -\int_0^\infty \int_0^\infty f(t_1, t_2) \ln f(t_2|t_1)\, \mathrm{d}t_1\, \mathrm{d}t_2, \qquad (A\text{-}2)$$

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

The joint probability density function $f(t_1, t_2, \ldots)$ 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 $f(t_1, \ldots, t_n)$ that governs the immediate response of a neuron (or a population of neurons). The definition (A-1) 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.

Appendix B: 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 and 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, Habibullah, and Soofi, 1992; Esteban *et al.*, 2001; Miller and Fisher, 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

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

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, Habibullah, and Soofi (1992). The bias-correcting factor is

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

where $\Psi(z) = \frac{d}{dz} \ln \Gamma(z)$ is the digamma function (Abramowitz and Stegun, 1965). Our experience with simulated data shows, that for sample sizes $n \approx 500$ the error of estimation is relatively small (Kostal and Lansky, 2006b), the positive bias with respect to true values is not important for small samples (Ebrahimi, Habibullah, and Soofi, 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, Stögbauer, and Grassberger, 2004; Victor, 2002) is realized by the Kozachenko-Leonenko binless estimator (Kozachenko and 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

$$\hat{h} = \frac{d}{n} \sum_{i=1}^{n} \ln \lambda_i + \ln \left[\frac{(n-1)\sqrt{\pi^d}}{\Gamma(d/2+1)} \right] + \gamma, \qquad (B-3)$$

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 and 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.

Appendix C: 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 main 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

$$\text{Prob}(T = \theta_1) = 1 - \text{Prob}(T = \theta_2) = p, \qquad (C-1)$$

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

$$E(T) = p\theta_1 + (1-p)\theta_2, \qquad (C-2)$$
$$C_V = \frac{\sqrt{(1-p)p}|\theta_1 - \theta_2|}{(\theta_1 - \theta_2)p + \theta_2}. \qquad (C-3)$$

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

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

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

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)

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

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 and Stegun, 1965).

– ad f. Probability density function of the gamma distribution, parameterized by $\mu = E(T)$ and C_V is

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

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

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

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

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

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

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

Adrian, E. D., *Basis of Sensation* (W. W. Norton and Co., New York, 1928).

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

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

Berger, D., Pribram, K., Wild, H., and Bridges, C., "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 (1990).

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

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

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

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

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

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

Cover, T. M. and Thomas, J. A., *Elements of Information Theory* (John Wiley and Sons, Inc., New York, 1991).

Cox, D. R. and Lewis, P. A. W., *The statistical analysis of series of events* (Latimer Trend and Co. Ltd., Whistable, 1966).

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

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

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

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

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

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

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

Jaynes, E. T. and Bretthorst, G. L., *Probability Theory: The Logic of Science* (Cambridge University Press, Cambridge, 2003).

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

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

Kandel, E. R., Schwartz, J. H., and Jessel, T. M., *Principles of neural science* (Elsevier, New York, 1991).

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

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

Kostal, L. and Lansky, P., "Classification of stationary neuronal activity according to its information rate," Netw. Comput. Neural Syst. **17**, 193–210 (2006a).

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

Kostal, L. and Lansky, P., "Variability and randomness in stationary neuronal activity." BioSystems **89**, 44–49 (2007).

Kostal, L., Lansky, P., and Zucca, C., "Randomness and variability of the neuronal activity described by the Ornstein-Uhlenbeck model," Netw. Comput. Neural Syst. **18**, 63–75 (2007).

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

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

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

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

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

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

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

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

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

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

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

1475 (1992).

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

Miller, E. G. and Fisher, J. W., "Ica using spacings estimates of entropy," JMLR **4**, 1271–1295 (2003).

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

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

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

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

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

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

Ratliff, F., Hartline, H. K., and Lange, D., "Variability of interspike intervals in optic nerve fibers of limulus: Effect of light and dark adaptation," Proc. Natl. Acad. Sci. U.S.A. **60**, 464–469 (1968).

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

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

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

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

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

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

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

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

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

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

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

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

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

Sugano, N. and Tsukada, M., "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 (1978).

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

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

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

Tuckwell, H. C., *Introduction to Theoretical Neurobiology, Vol. 2* (Cambridge University Press, New York, 1988).

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

Victor, J. D., "Binless strategies for estimation of information from neural data," Phys. Rev. E **66**, 051903 (2002).

Victor, J. D. and Purpura, K. P., "Metric-space analysis of spike trains: theory, algorithms and application," Netw. Comput. Neural Syst. **8**, 127–164 (1997).

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

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