Neural Networks 95 (2017) 57–65

Contents lists available at ScienceDirect

Neural Networks

journal homepage: www.elsevier.com/locate/neunet

Entropy factor for randomness quantification in neuronal data

K. Rajdl *, P. Lansky, L. Kostal

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

ARTICLE INFO

Article history:
Received 28 March 2017
Received in revised form 27 July 2017
Accepted 28 July 2017
Available online 15 August 2017

Keywords:
Variability measure
Shannon entropy
Renewal process
Fano factor

ABSTRACT

A novel measure of neural spike train randomness, an entropy factor, is proposed. It is based on the Shannon entropy of the number of spikes in a time window and can be seen as an analogy to the Fano factor. Theoretical properties of the new measure are studied for equilibrium renewal processes and further illustrated on gamma and inverse Gaussian probability distributions of interspike intervals. Finally, the entropy factor is evaluated from the experimental records of spontaneous activity in macaque primary visual cortex and compared to its theoretical behavior deduced for the renewal process models. Both theoretical and experimental results show substantial differences between the Fano and entropy factors. Rather paradoxically, an increase in the variability of spike count is often accompanied by an increase of its predictability, as evidenced by the entropy factor.

1. Introduction

One of the most important questions in neuroscience is how information is transferred in the nervous system. Some aspects of the mechanism are clear and generally accepted—information is transferred using spikes transmitted by neurons, coded only by times of their occurrence, not by their size or shape (Gerstner & Kistler, 2002; Rieke, Warl, de Ruyter van Steveninck, & Bialek, 1999). However, more specific details of the mechanism are not obvious, mainly the exact method of coding. The simplest, and most often assumed concept, is that coding is achieved through the spike rate (defined, for example, as the number of spikes fired in a unit of time), called the rate coding. It is a natural idea, as neuronal responses are usually strongly influenced by the presynaptic spiking rate. Nevertheless, the rate reflects only a small part of the character of spike trains and thus a possibility exists that a more complex coding is employed. The codes which assume some rate-independent behavior, given by the specific position of single spikes, are called temporal codes (temporal coding).

There are many possibilities how such a temporal code could operate, with one variant focusing on the variability in spike timing (variability coding). Under this concept, information is transferred not only by the rate, but also using the variability of spike trains. The crucial question while studying the variability is how to quantify it. The two of the most often used measures are the coefficient of variation (CV) and the Fano factor, see (Ditlevsen & Lansky, 2011; Rajdl & Lansky, 2014; Stevenson, 2016) and other references within these texts. The CV is defined as the standard deviation to mean of interspike intervals (ISIs) ratio and the Fano factor is defined as the variance to mean of number of spikes in a time window ratio. The number of experimental studies where these two measures have been applied is practically countless. The definitions of CV and the Fano factor also show two main ways how to see neuronal data—as ISIs or numbers of spikes in a time window of fixed length. A real neuron more likely registers the counts of spikes than the exact lengths of ISIs, nevertheless, both of these variants are useful and provide a different perspective in evaluation of experimental data. Another feature of these two variability measures, up to our knowledge never mentioned, is that they can be seen as variance (of ISIs or counts) related to a corresponding variance of a Poisson process with the same firing intensity. As a Poisson process has a unique position among the random processes used to model spike trains, it is an important property of these measures that increases their interpretability. Although we consider CV and Fano factor to be the most common variability measures, let us note that several others have been proposed and studied. Among others, we may mention CV2 proposed by Holt, Softky, Koch, and Douglas (1996), CVlog employed in Ruigrok, Hensbroek, and Simpson (2011) or coefficient of local variation, Lv, introduced by Shinomoto and his coworkers (Aoki, Takaguchi, Kobayashi, & Lambiotte, 2016; Shimokawa & Shinomoto, 2009; Shinomoto, Miura, & Koyama, 2005). For comparison of various measures in classification of neuronal discharge patterns, see Kumbhare and Baron (2015). Generally, as seen, there are various ways of quantifying the statistical heterogeneity of a given probability law, not only variance or entropy. Among others belong the Gini index, which measures the law's egalitarianism or its alternative—the Pietra index (Eliazar & Sokolov, 2010). This latter measure is especially useful in the case of asymmetric and skewed probability laws and its future applications on neuronal data remain open.

* Corresponding author.
E-mail address: rajdlk@mail.muni.cz (K. Rajdl).

http://dx.doi.org/10.1016/j.neunet.2017.07.016

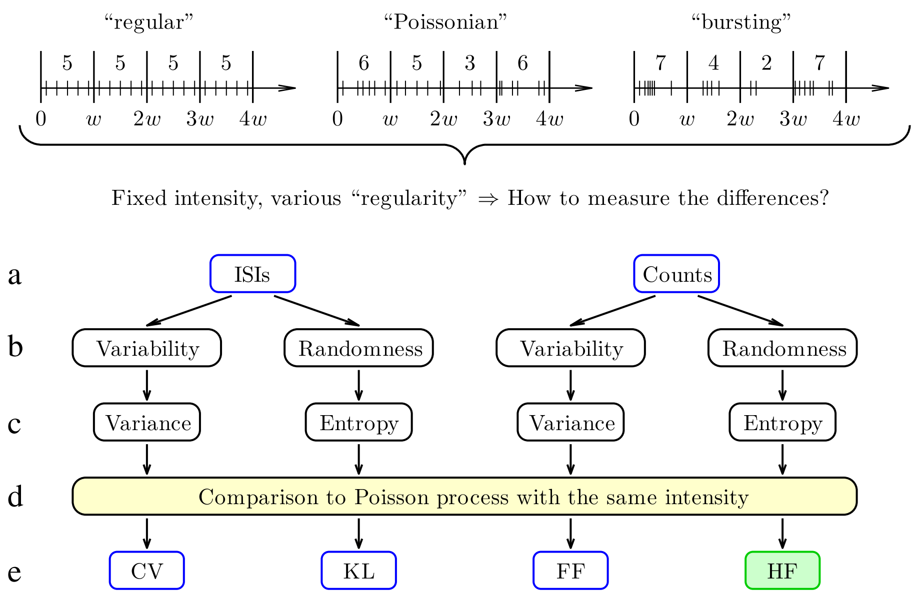

Fig. 1. An overview of selected measures of intensity independent behavior of spike trains. (a) Two main ways how to describe spike trains—using ISIs or counts of spikes in a time window of length w. (b) Two concepts how to understand variability and randomness. (c) Specific characteristics representing variability and randomness—variance and (Shannon) entropy. (d) Relating of given characteristics to a Poisson process of the same intensity. (e) Resulting measures—coefficient of variation (CV), Kullback–Leibler distance (KL), Fano factor (FF) and entropy factor (HF).

Although the variability is commonly used as a general term, there is a similar concept, but not equivalent one, which should be distinguished—the randomness or predictability. Both, the variability and randomness describe the character of spike trains which is not fully determined by the intensity, but there is a clear difference between them. It can be seen, for example, from the fact that even a variable process can be non-random. A very suitable quantity to measure randomness is Shannon entropy (Shannon & Weaver, 1998), which has been widely applied in neuroscience (among many others, Borst and Haag (2001); Chacron, Longtin, and Maler (2001); Ince et al. (2010); Kostal, Lansky, and Rospars (2007); McDonnell, Ikeda, and Manton (2011); Steuer et al. (2001); Strong, Koberle, de Ruyter van Steveninck, and Bialek (1998); Watters and Reeke (2014)). Some randomness measures based on entropy have been proposed in neural context and thoroughly studied (Kostal & Lansky, 2006; Kostal, Lansky, & Pokora, 2011; Kostal et al., 2007). Nevertheless, they focus only on ISIs, creating an alternative to CV. The most suitable way of measuring the randomness of ISIs appears to be through using the Kullback–Leibler (KL) distance of probability density of ISIs to density of ISIs in a Poisson process with the same intensity, thus to density of an exponential distribution.

Quantities based on the KL distance can be seen as analogies to CV representing randomness instead of variance (Kostal, Lansky, & Pokora, 2013). It would be thus natural to use entropy to measure the randomness of spike-counts analogously to the Fano factor. We propose such a measure in this paper. It is defined as the ratio of the Shannon entropy of numbers of spikes in a time interval to the Shannon entropy of a Poisson counting process with the same intensity. Due to similarity to the Fano factor, we call the measure an entropy factor (HF). It creates a natural complement to existing measures. For an overview of selected measures of variability and randomness, see Fig. 1. The measures are classified into two groups according to whether they focus on ISIs or spike counts, the crucial difference between these two approaches being in the presence of an additional parameter (the observation window length) in the distribution of counts. Hence, the variability and randomness measures of spike-counts provide a more complete and extensive characterization of the underlying statistical spiking model, as reported in this paper. As Fig. 1 also shows, all the measures compare a characteristic of given spike train to a Poisson process, which is a simple, but efficient way how to improve interpretation and comparison of various experiments.

The aim of this paper is to explore properties of entropy factor and compare them with properties of Fano factor, to show that there are some fundamental differences between the measures. Behavior of the Fano factor has been studied in various papers, mostly its dependence on the length of the observation window or CV (Nawrot et al., 2008; Pipa, Gruen, & Vreeswijk, 2013; Rajdl & Lansky, 2014) or its statistical properties (Eden & Kramer, 2010). Here, the theoretical properties of the entropy factor are studied and explicit formulas describing its behavior are derived. Firstly, the equilibrium renewal process as a model of spike train is defined. Based on this model, the standard variability measures and their basic properties are summarized. Secondly, the new measure is defined and investigated in the next section. To illustrate its properties, two of the most often assumed models of the lengths of ISIs are used, gamma and inverse Gaussian distributions (Fisch, Schwalger, Lindner, Herz, & Benda, 2012; Koyama & Kostal, 2014; Lansky, Sacerdote, & Zucca, 2016; Nawrot et al., 2008; Omi & Shinomoto, 2011; Ostojic, 2011; Pipa et al., 2013; Shimokawa, Koyama, & Shinomoto, 2010). Finally, the entropy factor is compared to the Fano factor estimated from experimental data.

2. Spike train model

In a formal description, a spike train can be represented as a sequence of times of occurrence of the individual spikes, $X_1, \ldots, X_n$, $n \in \mathbb{N}$, and modeled using a (stochastic) point process. Probably the most often used are renewal processes, which assume that all the ISIs, $T_i = X_{i+1} - X_i$, $i = 1, \ldots, n-1$, are mutually independent and identically distributed random variables. This model is also used here, however, we are well aware that the true character of spike trains can be more complex (Avila-Akerberg & Chacron, 2011; Chacron et al., 2001; Farkhooi, Strube-Bloss, & Nawrot, 2009; Schwalger, Droste, & Lindner, 2015).

A renewal process is defined by a continuous positive random variable T, representing the lengths of ISIs, with a probability density function $f(t)$, cumulative distribution function $F(t)$, and mean $\mu = \mathrm{E}(T)$. To fully specify the renewal process, it is also necessary to state the relationship between the sequence of spikes and the zero time (start of the observation). From the perspective of an external observer, time zero is random with respect to the spikes (equilibrium renewal process). Then the time up to the first spike (T_0) has, generally, a different probability distribution than T. The only renewal process in which T is equivalent to T_0 is the Poisson process. In that case, the random variables T and T_0 are exponentially distributed.

As follows from the frequency coding approach, neurons are influenced by the number of spikes received in a time period. This point of view leads to a counting process $N(w)$, which denotes the number of spikes in an interval $(0, w]$, $w > 0$, where $N(w)$ is a discrete random variable with probability mass function

$$p_n(w) = P(N(w) = n),\ n = 0, 1, 2, \ldots . \tag{1}$$

The counting process is an alternative way of describing a renewal process using ISIs,

$$N(w) \geq n \Leftrightarrow T_0 + \cdots + T_{n-1} \leq w, \tag{2}$$

from which follows, (Jewell, 1960),

$$p_0(w) = \mathcal{L}^{-1}\left\{\frac{1}{s} - \frac{1-\tilde{f}(s)}{\mathrm{E}(T)s^2}\right\}(w), \tag{3}$$

$$p_n(w) = \mathcal{L}^{-1}\left\{\frac{(1-\tilde{f}(s))^2(\tilde{f}(s))^{n-1}}{s^2\mathrm{E}(T)}\right\}(w),\ n \in \mathbb{N}, \tag{4}$$

where $\mathcal{L}$ denotes the Laplace transform and $\tilde{f}(s) = \mathcal{L}\{f\}(s)$. Evaluation of formulas (3) and (4) is rather complicated, and even some basic characteristics of $N(w)$ are difficult to obtain. There is an explicit formula for the mean,

$$\mathrm{E}(N(w)) = \frac{w}{\mathrm{E}(T)} = \frac{w}{\mu}, \tag{5}$$

but no similar one exists for higher moments. For an approximation of the variance of $N(w)$, the following asymptotic relationship can be used, (Cox, 1962),

$$\mathrm{Var}(N(w)) \approx \frac{\mathrm{Var}(T)}{\mathrm{E}^3(T)}w + \frac{1}{2}\left(1 + \frac{\mathrm{Var}(T)}{\mathrm{E}^2(T)}\right)^2 - \frac{1}{3}\frac{\mathrm{E}(T^3)}{\mathrm{E}^3(T)}, \quad \text{for } w \to \infty. \tag{6}$$

Relationship (5) also yields a simple formula for the intensity of the process,

$$\lambda = \frac{1}{\mu}, \tag{7}$$

as for a general counting process the definition of intensity is

$$\lambda(w) = \frac{\mathrm{dE}(N(w))}{\mathrm{d}w}. \tag{8}$$

Note that probably the only non-trivial renewal process for which it is possible to express the probability mass function and many other characteristics of $N(w)$ by simple closed formulas is the Poisson process. Let us denote the corresponding counting process and random variable representing the ISIs as $N_P(w)$ and T_P. It holds that

$$\mathrm{E}(T_P) = \sqrt{\mathrm{Var}(T_P)} = 1/\lambda \tag{9}$$

and

$$\mathrm{E}(N_P(w)) = \mathrm{Var}(N_P(w)) = \lambda w. \tag{10}$$

3. Variability

The term "variability" is used to reflect the extent to which a distribution, under our scenario, of spike timing is stretched or squeezed. It is commonly related to the second statistical moment but also, for example, an interquartile range can be used. An important difference, related to the problemacy of stochastic point processes, is that the second statistical moments are not compared directly but scaled to a template which is represented by the Poisson process of the *same intensity*. Specifically, it is the coefficient of variation,

$$\mathrm{CV}^2 = \frac{\mathrm{Var}(T)}{\mathrm{Var}(T_P)}, \tag{11}$$

and Fano factor,

$$\mathrm{FF}(w) = \frac{\mathrm{Var}(N(w))}{\mathrm{Var}(N_P(w))},\ w > 0. \tag{12}$$

These measures are commonly defined with $\mathrm{E}^2(T)$ and $\mathrm{E}(N(w))$ instead of $\mathrm{Var}(T_P)$ and $\mathrm{Var}(N_P(w))$. However, due to relationships (5), (9) and (10), such definitions are equivalent and only formulas (11) and (12) fully clarify the role of CV and $\mathrm{FF}(w)$.

While the CV is a constant, $\mathrm{FF}(w)$ is a function of w, and therefore the Fano factor is sometimes defined as the limit

$$\mathrm{FF} = \lim_{w\to\infty} \mathrm{FF}(w), \tag{13}$$

with, (Cox, 1962),

$$\mathrm{FF} = \mathrm{CV}^2. \tag{14}$$

Next, the following formulas hold, (Jewell, 1960),

$$\lim_{w\to 0} \mathrm{FF}(w) = 1 \tag{15}$$

and, (Rajdl & Lansky, 2014),

$$\mathrm{FF}(w) = \frac{1}{w}\mathcal{L}^{-1}\left\{\frac{1+\mathcal{L}\{f\}(s)}{s^2[1-\mathcal{L}\{f\}(s)]}\right\}(w) - \frac{w}{\mathrm{E}(T)}, \tag{16}$$

where the latter can be used to numerically calculate $\mathrm{FF}(w)$ for any $w > 0$.

4. Predictability, the new measure and its properties

As already mentioned, we wish that predictability or randomness is distinguished from variability, as it reflects different features of the spike train. To measure randomness, it is suitable to use the Shannon entropy. For a discrete random variable $N(w)$ with probabilities $p_n(w)$, $n = 0, 1, \ldots$, it is defined as

$$\mathrm{H}(N(w)) = -\sum_{n=0}^{\infty} p_n(w) \ln p_n(w). \tag{17}$$

The larger the value of $\mathrm{H}(N(w))$, the more difficult it is to predict the value of $N(w)$, and so, the more information we obtain from a realization of $N(w)$ (the variable $N(w)$ is more random).

The randomness of T has been thoroughly studied (Kostal et al., 2011, 2013), however, we are not aware of any similar studies of $N(w)$. Therefore a randomness measure of $N(w)$ based on the Shannon entropy is defined and studied in this paper. It would be possible to directly use quantity (17). Nevertheless, for a better interpretation, it is suitable to also relate its value to a Poisson process such as in (11) and (12). We suggest to define the entropy factor as

$$\mathrm{HF}(w) = \frac{\mathrm{H}(N(w))}{\mathrm{H}(N_P(w))}, \tag{18}$$

where $N_P(w)$ denotes the Poisson process with the same intensity as $N(w)$.

Analogously to studies on the Fano factor, we are interested in the behavior of $\mathrm{HF}(w)$ in dependence on w and on properties of T, mainly the CV. Calculation of $\mathrm{HF}(w)$ requires calculation of the entropy of $N(w)$ and entropy of the Poisson process. The entropy of the counting Poisson process with intensity λ is

$$\mathrm{H}(N_P(w)) = \lambda w[1 - \ln(\lambda w)] + e^{-\lambda w} \sum_{n=0}^{\infty} \frac{(\lambda w)^n \ln(n!)}{n!}. \qquad (19)$$

Relationship (19) can be approximated using, (Evans & Boersma, 1988),

$$\mathrm{H}(N_P(w)) \approx \frac{1}{2}\ln(2\pi e \lambda w) - \frac{1}{12\lambda w} - \frac{1}{24(\lambda w)^2} - \frac{19}{360(\lambda w)^3}, \text{ for } \lambda w \to \infty. \qquad (20)$$

For a general renewal process with intensity λ, it is possible to numerically evaluate Eqs. (3), (4) and (17) or to use the asymptotic formula

$$\mathrm{H}(N(w)) \approx \frac{1}{2}\ln(2\pi e) + \frac{1}{2}\ln\left(\mathrm{CV}^2\lambda w + \frac{1}{2}\left(1+\mathrm{CV}^2\right)^2 - \frac{1}{3}\lambda^3 \mathrm{E}(T^3)\right), \text{ for } \lambda w \to \infty. \qquad (21)$$

To derive (21), we use the fact that probability distribution of $N(w)$ converges to a normal distribution with mean $\mathrm{E}(N) = \lambda w$ and variance $\sigma^2 = \mathrm{Var}(N(w))$ (Cox, 1962). Entropy of normal distribution is $\ln(2\pi e \sigma^2)/2$, which with formula (6) yields relationship (21).

Next, some limit properties of $\mathrm{HF}(w)$ are

$$\lim_{w\to 0} \mathrm{HF}(w) = 1, \qquad (22)$$

$$\lim_{w\to\infty} \mathrm{HF}(w) = 1 \qquad (23)$$

and, for large λw,

$$\lim_{\mathrm{CV}\to 0} \mathrm{HF}(w) \approx \frac{-q\ln(q) - (1-q)\ln(1-q)}{\frac{1}{2}\ln(2\pi e \lambda w)}, \qquad (24)$$

where $q = \lambda w - \lfloor \lambda w \rfloor$ and $\lfloor \lambda w \rfloor$ is the integer part of λw. Derivation of relationship (22) is shown in Appendix. Limit (23) can be calculated directly as limit from (21) and (20). For derivation of limit (24), let us first see the properties of an equilibrium renewal process for $\mathrm{CV} \to 0$. The process clearly consists of uniformly distributed time to the first spike ($T_0 \sim \mathrm{Unif}(0, \mu]$), and following constant ISIs of length μ. For any $w > 0$, there are thus only two possible numbers of spikes which can occur, $\lfloor \lambda w \rfloor$ spikes with probability $q = \lambda w - \lfloor \lambda w \rfloor$ or $\lfloor \lambda w \rfloor - 1$ spikes with probability $1 - q$. This yields the nominator of (24), the denominator is simply the first part of approximation (20).

According to relationship (23), $\mathrm{HF}(w)$ always converges to a value of one, which substantially differs to the behavior of the Fano factor. Another principal difference between $\mathrm{HF}(w)$ and $\mathrm{FF}(w)$ is that the entropy factor is bounded from above (for fixed w and μ), as there is a maximum value of entropy of $N(w)$, whereas the Fano factor can take arbitrarily large values. The maximum entropy among all the discrete distributions with mean λw has a geometric distribution (Cover & Thomas, 2006). Its probability mass function is

$$P(X(w) = n) = \left(1 - \frac{1}{1+\lambda w}\right)^n \frac{1}{1+\lambda w}, \; n = 0, 1, 2, \ldots \qquad (25)$$

and its entropy is

$$\mathrm{H}(X(w)) = (1+\lambda w)\ln(1+\lambda w) - \lambda w \ln(\lambda w). \qquad (26)$$

No renewal process has such a distribution of number of spikes, and thus no renewal process can reach the value (26). This fact can be easily seen through the limit behavior of $N(w)$ of any renewal process. As mentioned, it converges to the normal distribution for $\lambda w \to \infty$, however, the geometric distribution (25) does not. An upper bound of $\mathrm{HF}(w)$ is thus created by the ratio of $\mathrm{H}(X(w))$ to $\mathrm{H}(N_P(w))$. It holds that

$$\lim_{w\to\infty} \frac{\mathrm{H}(X(w))}{\mathrm{H}(N_P(w))} = 2, \qquad (27)$$

which is a direct consequence of relationships (20) and (26).

5. Examples of the entropy factor and comparison with the Fano factor

The above presented theoretical properties of $\mathrm{HF}(w)$ give a basic idea about its behavior. However, they do not yield the dependence on w or CV in a simple way. Therefore, it is illustrated for two commonly used distributions of T, gamma and inverse Gaussian (IG). Both distributions are parametrized using two independent parameters. For this purpose, it is suitable to use μ and CV. The density of the gamma distribution is

$$f_G(t) = \frac{t^{1/\mathrm{CV}^2-1}\exp[-t/(\mu\mathrm{CV}^2)]}{\Gamma(1/\mathrm{CV}^2)(\mu\mathrm{CV}^2)^{1/\mathrm{CV}^2}}, \quad t \geq 0, \qquad (28)$$

and the density of the IG distribution is

$$f_{IG}(t) = \sqrt{\frac{\mu}{2\pi\mathrm{CV}^2 t^3}}\exp[-(t-\mu)^2/(2\mathrm{CV}^2 t\mu)], \quad t \geq 0. \qquad (29)$$

Let us note that the gamma distribution becomes exponential for $\mathrm{CV} = 1$ and thus the corresponding renewal process is a Poisson process. On the contrary, IG distribution never attains the exponential distribution. Without loss of generality, in the following examples it is assumed that $\mu = 1$, this can be always achieved by using a suitable linear transformation of time. The time is thus shown in mean ISI units. The remaining free parameters are then CV and w.

To compare $\mathrm{HF}(w)$ with $\mathrm{FF}(w)$, the necessary values were numerically calculated using formulas (3), (4) and (17). In Fig. 2, the dependence on the length of the time window is illustrated. The common feature of both $\mathrm{HF}(w)$ and $\mathrm{FF}(w)$ is that they start at a value of one and the following behavior is influenced by the value of CV. In the presented cases, lower values of CV yield lower values of $\mathrm{FF}(w)$ and vice versa, however, this does not hold for $\mathrm{HF}(w)$. A large difference between the measures is in their limiting behavior for increasing w, where the Fano factor converges to CV^2, and the entropy factor returns to one. Nevertheless, the rate of convergence of $\mathrm{HF}(w)$ to the value of one is exceptionally slow.

A special property of the entropy factor is its upper bound. We can see that it is relatively low and thus the entropy of $N(w)$ cannot be much larger than the entropy of the Poisson process, at most double—according to relationship (27). Another feature is the sinusoid-like shape for a low CV. It is common for both of the measures, but it is more apparent for $\mathrm{HF}(w)$. The reason for this sensitivity of HF to small CV is the equilibrium property of the process we use to model the spike trains. Even for $\mathrm{CV} = 0$, when all the ISIs are same, the time to the first spike is random (due to the random start of the observation), but less than or equal to μ. It follows, that at times which are multiples of the mean ISI, we know exactly how many spikes there will be and that the entropy is zero. At other times, the number of spikes is still influenced by the random beginning and the entropy is thus positive.

For the sake of completeness, the range of w in Fig. 2(a), (b) is from 10^{-2} up to 10^8, while the mean ISI is equal to one. However, we may ask what range is interesting from a neural coding point of view. Typical experimental data contains a maximum of thousands of ISIs, the main reason being that current technology does not permit simultaneous observation of thousands of individual neurons and thus parallel observation is replaced by a single spike train. In

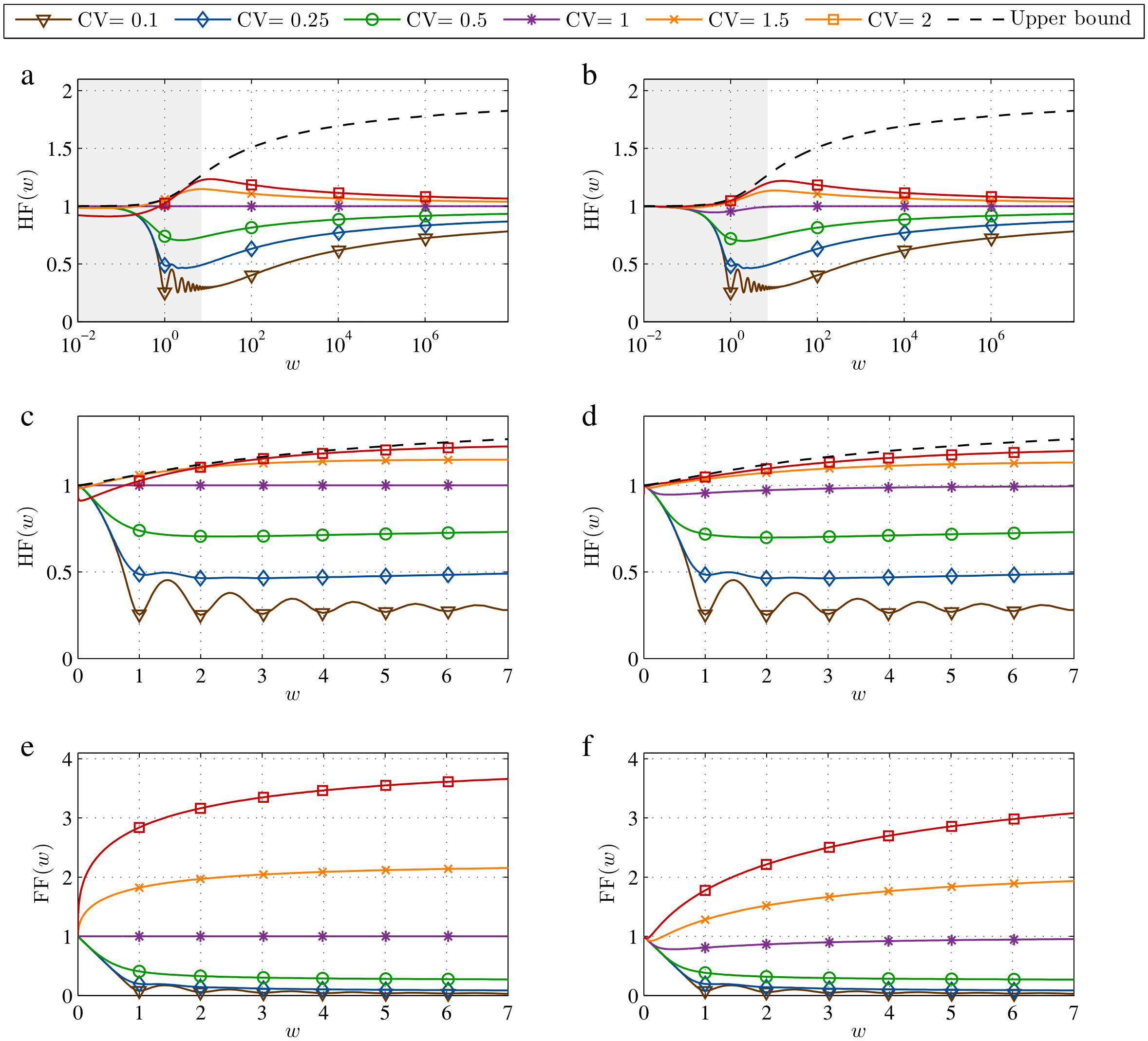

Fig. 2. Dependence of HF(w) (a, b, c, d) and FF(w) (e, f) on w for various values of CV. The ISIs follow gamma (a, c, e) and IG (b, d, f) probability distributions with mean ISI = 1. Window size w is on log-scale in panels (a) and (b), their gray part is shown with linear scale on panels (c) and (d). The values of CV are 0.1, 0.25, 0.5, 1, 1.5, 2 (corresponding to the individual full-line curves from bottom to top on the right sides of the panels). Dashed-lines: An upper boundary for HF(w), given by geometric distribution $X(w)$, (26), with parameter $\lambda = 1$.

a real neural network, in which synchrony (coincidence) of spiking plays a role, it can be expected that the range of interest for the time window is between fractions of the mean ISI up to several of its multiples. We can see that the extremes of HF(w) are reached in this range, which supports the importance of randomness evaluation. In Fig. 2(c), (d), the behavior of the entropy factor for the most important lengths of the observation window is shown in more detail.

In Figs. 3 and 4, the relationships between FF(w) and HF(w) are shown directly. In the first case, the values of CV of ISIs are fixed for the single curves and w is going continuously from 0 to ∞. In the second case, every curve corresponds to a fixed w while CV changes from 0.2 to 2. We can see that the relationship between HF and FF is not linear, often not even monotonic. Mainly, the dependence in Fig. 3 is relatively complicated, caused by the fact that the course of HF(w) is non-monotonic, as shown in Fig. 2. In Fig. 4, the curves are simpler, nevertheless, their order and intersections are not intuitive, which is again caused mainly by the non-monotonic dependency of the entropy factor and the length of the observation window.

Another question we were interested in is which CV maximizes HF(w) for each given model of ISIs and w. Such CVs in dependence on w are illustrated in Fig. 5(a). A larger w mostly requires a larger CV to maximize the entropy factor, however, the opposite holds for IG distribution and short windows. The individual maxima are always lower than the theoretical upper bound given by geometric distribution (26). In Fig. 5(b), the corresponding values of maximum HF(w) achieved by gamma and IG distributions in relation to the upper boundary are shown. We can see that for the presented lengths of the observation window ($w \leq 5$) the entropy factor of a renewal process can be relatively near to the upper boundary.

Finally, in Fig. 6 the accuracy of the approximation (21) is explored, using relative error,

$$\varrho(w) = 100\frac{\widetilde{\mathrm{HF}}(w) - \mathrm{HF}(w)}{\mathrm{HF}(w)}, \tag{30}$$

where $\widetilde{\mathrm{HF}}(w)$ is approximation of HF(w) using formula (21) to calculate the entropy. The precision is low mainly for small w and large CV. Very low values of CV are also problematic, as the approximation cannot reflect the sinusoid-like behavior of HF(w). However, for a standard range of CV (e.g., [0.5, 1.5]) the accuracy seems to be sufficient from w around several means of ISIs.

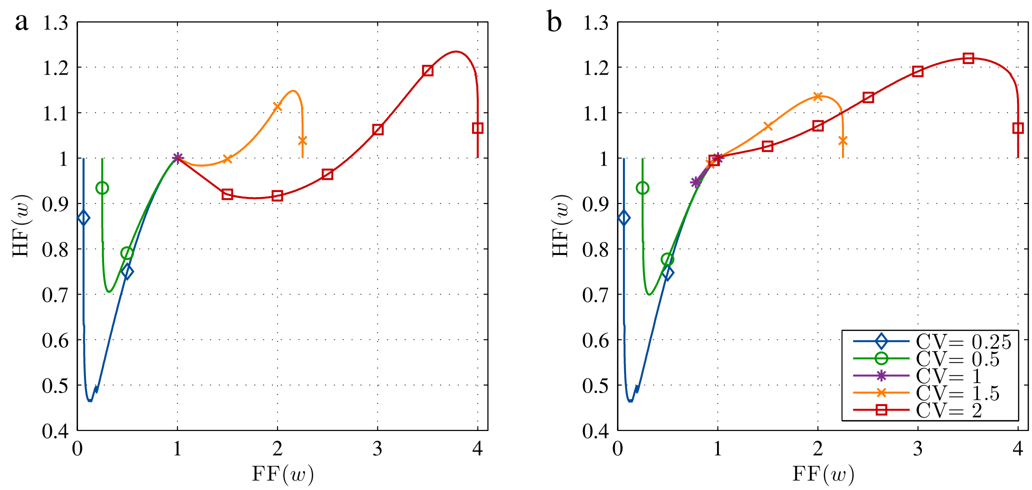

Fig. 3. The relation between FF(w) and HF(w) for various values of CV (0.25, 0.5, 1, 1.5, 2). The ISIs follow gamma (a) and IG (b) probability distributions with $\mu = 1$. The values of FF(w) and HF(w) are shown for $w \in (0, \infty)$.

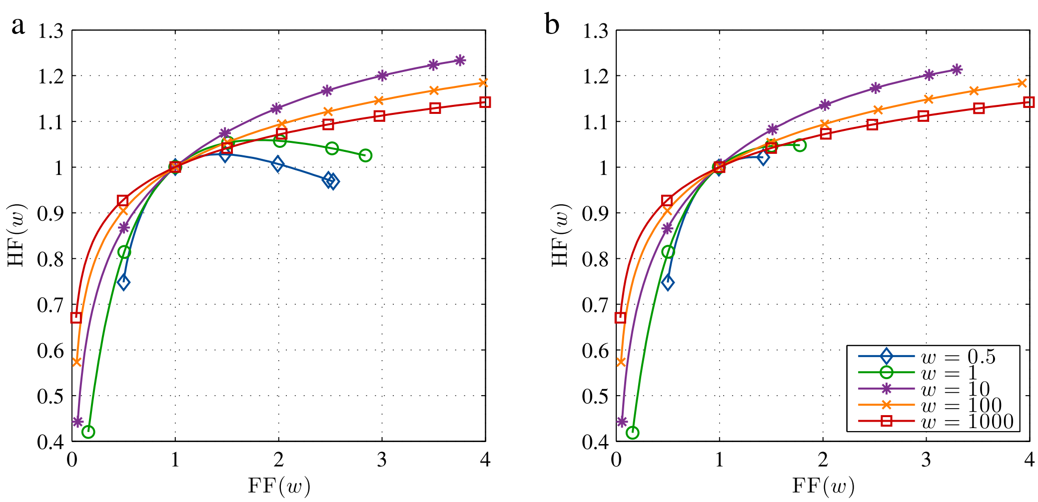

Fig. 4. The relation between FF(w) and HF(w) for various values of w (0.5, 1, 10, 100, 1000). The ISIs follow gamma (a) and IG (b) probability distributions with $\mu = 1$. The values of FF(w) and HF(w) are shown for CV $\in (0.2, 2)$.

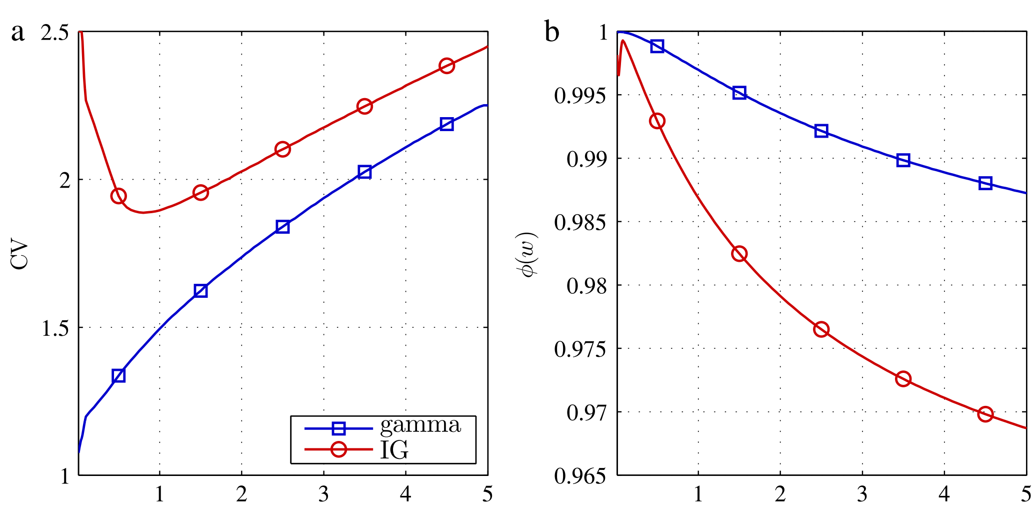

Fig. 5. (a) CV which maximizes HF(w) in dependence on w for gamma and IG distribution of T with $\mu = 1$. (b) Maximum of HF(w) to its upper boundary ratio in dependence on w, denoted as $\phi(w)$.

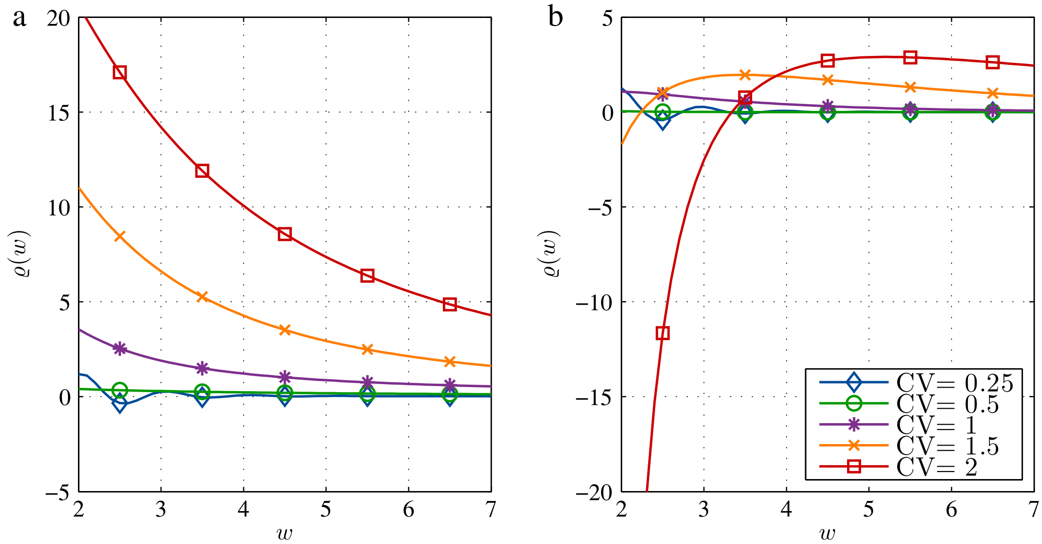

Fig. 6. The relative error $\varrho(w)$ [%] of approximation (21) for gamma (a) and IG (b) distribution of T in dependence on CV and w, $\mu = 1$.

6. Entropy factor evaluated from experimental data

In this section, we compare Fano and entropy factors evaluated from experiments. The data consists of 15 min records of spontaneous activity of neurons from primary visual cortex (V1) of the macaque monkey. The data were obtained from www.crcns.org www.crcns.org and for details see Cheng, Ping, and Chou (2014a, b). We initially analyzed 123 spike trains, each containing at least 200 spikes (most of them much more). The characteristics of the firing rates in the individual spike trains (calculated across the spike trains) were: mean = 427.2 ms, median = 193.6 ms, minimum = 22.6 ms, maximum = 2908.7 ms. To remove influence of the spike rate, the sequences of ISIs were linearly scaled to have the same (unit) mean. Although representing spontaneous activity the spike trains often showed some (rather minor) signs of trend. We performed a simple procedure (Cox & Lewis, 1966) to exclude these even slightly non-stationary spike trains. The procedure consisted of two steps. Firstly, the numbers of spikes in consecutive intervals of length 10 (in mean ISI units) were calculated. Secondly, significance of a linear regression in these time series was tested. If the linear trend was significant (at 0.05 significance level), the spike train was excluded. After this procedure 55 spike trains remained and were used in the following analysis.

The relationship of FF(w) to HF(w) was explored. Values of FF(w) and HF(w) were estimated for various values of w (from 0.5 to 15 mean ISI units). Before the estimation, it was necessary to transform the spike trains to spike counts in these time windows. The records were divided into segments of the required length and after each segment an interval of length 2 was discarded to reduce possible dependencies of the counts. Based on the counts, the entropy was estimated for every spike train by using formula (17) with the sample probabilities replacing the real ones. This simple method may lead to a negative bias of the entropy estimator (Schurmann, 2004). It is caused by the fact that very low probabilities are not observed in finite samples. To reduce the influence of this issue on the estimated value of HF(w), an analytically calculated entropy of the Poisson process in formula (18) was not used. It was analogously estimated from simulated data with a number of samples equal to the number of samples available in the individual records of the experimental data. The purpose was to have a similar bias in both parts of formula (18), which possibly reduced the bias of the ratio. The entropy of the Poisson process was generated 10 000 times and averaged to minimize the variance of the estimator.

The results are shown in Fig. 7. We can see that there is not a simple linear or at least monotonous relationship between the values of Fano and entropy factor. In addition, the shape of dependency strongly changes with the size of the counting window. For small w, initially, increasing values of FF(w) mostly implies larger values of HF(w), but in accordance with the theoretical results, the tendency turns over and further increase of FF(w) induces a decrease of HF(w). For large w, the latter tendency slows down to practically constant HF(w). From the point of view of a target neuron it means that higher variability does not automatically mean lower predictability. Often higher variability implies better predictability. This is rather counter-intuitive result which should be kept in mind when simply quantifying the spike trains by variability of the spike numbers only. The results confirm that the examined measures show different aspects of the spike trains.

7. Discussion and conclusions

We have defined and studied the entropy-based measure of randomness of the counting process $N(w)$ with the aim applying it to analysis of neural spike trains. It was related to other three commonly used quantities designed to measure the rate-independent behavior of spike trains—coefficient of variation, Fano factor and KL distance. As CV measures the variability of ISIs, the Fano factor reflects the variability of $N(w)$ and the KL distance the randomness of ISIs, the entropy factor is a natural complement to these measures, see (e) in Fig. 1. In fact, equally natural would be application of the KL distance between the observed counts and Poissonian counts. However, we stressed the analogy to the Fano factor due to its enormous popularity in neuronal as well as other studies.

We have studied the entropy factor under assumption that the spike trains correspond to a renewal process, which can be seen as a drawback as the real character of spike trains is probably far more complex. This simplification can be justified by the fact that neurons must "read" information relatively quickly—based on short time intervals during which the stationarity might be (approximately) satisfied. Nevertheless, it could be useful to study the behavior of HF even for some more realistic models, at least for serially dependent ISIs.

Another issue which should be explored in the future is estimation of the entropy factor. To use HF as a randomness measure for data, it has to be estimable with sufficient reliability but there

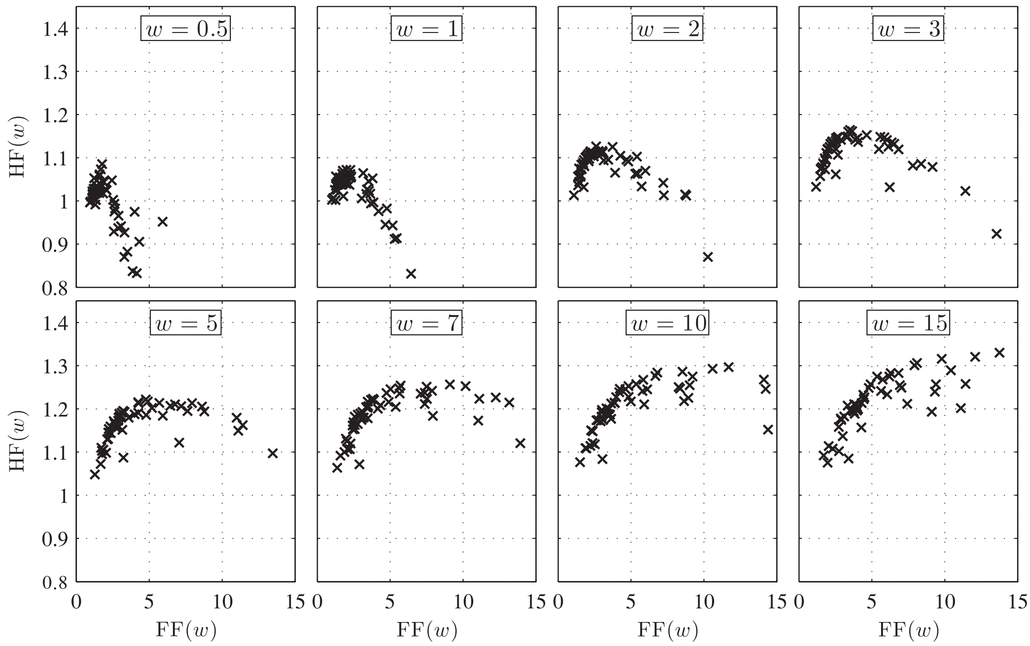

Fig. 7. The relationship of HF(w) to FF(w) in experimental data for the length of the observation window $w \in \{0.5, 1, 2, 3, 5, 7, 10, 15\}$, where w is in mean ISI units. The values were calculated individually for 55 spike trains.

is a problem of bias in estimation of entropy (Schurmann, 2004). In Section 6, we attempted to reduce this bias by creating a similar bias in the numerator and denominator of definition (18), however, the character of the bias and efficiency of this method should be theoretically examined.

We have focused mainly on behavior of the entropy factor in dependence on the length of the observation window w and on the relationship between HF(w) and FF(w). The main results can be summarized using the following points.

- For theoretical evaluation of the entropy factor from renewal models, except for large values of CV and low values of w, it appears possible to use formula (21). This fact can be utilized also for experimental data, under the condition of a careful statistical inference confirming the renewal character.
- The entropy factor can be larger than one—Poisson process does not maximize the entropy of $N(w)$ among the renewal processes. It contrasts the fact that the same process (exponentially distributed ISIs) maximizes the differential entropy of ISIs. Simultaneously, it is in agreement with the fact that the Fano factor can take values above (called over-dispersed) as well as below (under-dispersed) one.
- The extremes of the entropy factor in dependence on w occur for the observation window around several mean ISIs, which is the window that might be used by cortical or sensory neurons for information processing. At these windows the information about randomness is highest and it supports the possibility of randomness coding. Simultaneously, it may imply that slow synaptic processes are not involved in this way of coding. Of course, far detailed experimental investigation is required for any conclusion, which at this moment remains as a speculation only.
- One of the main differences of the entropy factor to the Fano factor is the existence of an upper bound. The Fano factor (variance) can reach arbitrarily large values, whereas the entropy of $N(w)$ is bounded from above by the entropy of geometric distribution and cannot be larger than double the entropy of a Poisson process with the same intensity. However, it seems that for a renewal process with properties corresponding to a realistic spike train, the values of the entropy factor cannot be much higher than one. In two investigated ISIs models, there is a maximum of the entropy factor which depends on CV. This maximum can be very close to the upper boundary for short observation window only.
- Both limits of HF(w) with respect to w (to 0 and ∞) are one. From the point of view of entropy, a renewal process starts to behave as a Poisson process and ends the same. However, the convergence for $w \to \infty$ is very slow.
- The relationship between the entropy factor and the Fano factor is not linear and often non-monotonic. At several occasions an increase in variability is accompanied with decreasing randomness.

All the above summarized results imply that the behavior of the entropy factor substantially differs from the behavior of the Fano factor. According to expectations, HF reflects a different aspect of the intensity independent character of spike trains (randomness) and thus can be seen as a complement to the commonly used measures.

Acknowledgments

This work was supported by the Institute of Physiology RVO:67985823 and by the Czech Science Foundation project 15-08066S.

Appendix. Derivation of formula (22)

For deriving formula (22) we assume regularity of the process, i.e., the probability of occurrence of more than one spike in $(0, w]$ is negligible for w small. The entropy of $N(w)$ then fully depends on the probability

$$p_1(w) = P(N(w) = 1). \tag{A.1}$$

This probability can be calculated from the distribution of the time up to the first spike, $f_0(w)$, which can be expressed (for equilibrium renewal processes) as, (Cox, 1962),

$$f_0(w) = \lambda(1 - F(w)). \tag{A.2}$$

Then,

$$p_1(w) = \int_0^w f_0(s)\,\mathrm{d}s = \lambda w - \int_0^w F(s)\,\mathrm{d}s. \tag{A.3}$$

Next, let us have two processes $N_a(w)$ and $N_b(w)$ with cumulative distribution functions $F_a(w)$ and $F_b(w)$, probabilities (A.1) $p_{1,a}(w)$ and $p_{1,b}(w)$ and the same intensity λ and calculate

$$\lim_{w\to 0} \frac{p_{1,a}(w)}{p_{1,b}(w)} = \lim_{w\to 0} \frac{\lambda w - \int_0^w F_a(s)\,\mathrm{d}s}{\lambda w - \int_0^w F_b(s)\,\mathrm{d}s} = \lim_{w\to 0} \frac{\lambda - F_a(w)}{\lambda - F_b(w)} = 1. \tag{A.4}$$

We have proved that in the limit for $w \to 0$ two equilibrium renewal processes with the same intensity have the same values of probability (A.1), thus the same entropy, which yields relationship (22).

References

Aoki, T., Takaguchi, T., Kobayashi, R., & Lambiotte, R. (2016). Input–output relationship in social communications characterized by spike train analysis. *Physical Review E*, *94*, 042313.

Avila-Akerberg, O., & Chacron, M. J. (2011). Nonrenewal spike train statistics: causes and functional consequences on neural coding. *Experimental Brain Research*, *210*, 353–371.

Borst, A., & Haag, J. (2001). Effects of mean firing on neural information rate. *Journal of Computational Neuroscience*, *10*, 213–221.

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

Cheng, C. J. C., Ping, F. C., & Chou, P. H. (2014a). Multi-electrode recordings of ongoing activity and responses to parametric stimuli in macaque V1. CRCNS.org. http://dx.doi.org/10.6080/K0J1012K.

Cheng, C. J. C., Ping, F. C., & Chou, P. H. (2014b). Tuning dissimilarity explains short distance decline of spontaneous spike correlation in macaque V1. *Vision Research*, *96*, 113–132.

Cover, T. M., & Thomas, J. A. (2006). *Elements of information theory*. New Jersey: John Wiley & Sons.

Cox, D. R. (1962). *Renewal theory*. London: Methuen.

Cox, D. R., & Lewis, P. A. W. (1966). *The statistical analysis of series of events*. London: Methuen.

Ditlevsen, S., & Lansky, P. (2011). Firing variability is higher than deduced from the empirical coefficient of variation. *Neural Computation*, *23*, 1944–1966.

Eden, U. T., & Kramer, M. A. (2010). Drawing inferences from Fano factor calculations. *Journal of Neuroscience Methods*, *190*, 149–152.

Eliazar, I. I., & Sokolov, I. M. (2010). Measuring statistical heterogeneity: The Pietra index. *Physica A*, *389*, 117–125.

Evans, R. J., & Boersma, J. (1988). The entropy of a Poisson-distribution. *SIAM Review*, *30*, 314–317.

Farkhooi, F., Strube-Bloss, M. F., & Nawrot, M. P. (2009). Serial correlation in neural spike trains: Experimental evidence, stochastic modeling, and single neuron variability. *Physical Review E*, *79*, 021905.

Fisch, K., Schwalger, T., Lindner, B., Herz, A. V., & Benda, J. (2012). Channel noise from both slow adaptation currents and fast currents is required to explain spike-response variability in a sensory neuron. *Journal of Neuroscience*, *32*, 17332–17344.

Gerstner, W., & Kistler, W. M. (2002). *Spiking neuron models*. Cambridge: Cambridge University Press.

Holt, G. R., Softky, W. R., Koch, C., & Douglas, R. J. (1996). Comparison of discharge variability in vitro and in vivo in cat visual cortex neurons. *Journal of Neurophysiology*, *75*, 1806–1814.

Ince, R. A. A., Senatore, R., Arabzadeh, E., Montani, F., Diamond, M. E., & Panzeri, S. (2010). Information-theoretic methods for studying population codes. *Neural Networks*, *23*, 713–727.

Jewell, W. S. (1960). The properties of recurrent-event processes. *Operations Research*, *8*, 446–472.

Kostal, L., & Lansky, P. (2006). Similarity of interspike interval distributions and information gain. *Biological Cybernetics*, *94*, 157–167.

Kostal, L., Lansky, P., & Pokora, O. (2011). Variability measures of positive random variables. *Plos One*, *6*, e21998.

Kostal, L., Lansky, P., & Pokora, O. (2013). Measures of statistical dispersion based on Shannon and Fisher information concepts. *Information Sciences*, *235*, 214–223.

Kostal, L., Lansky, P., & Rospars, J. P. (2007). Neuronal coding and spiking randomness. *European Journal of Neuroscience*, *26*, 2693–2701.

Koyama, S., & Kostal, L. (2014). The effect of interspike interval statistics on the information gain under the rate coding hypothesis. *Mathematical Biosciences and Engineering*, *11*, 63–80.

Kumbhare, D., & Baron, M. S. (2015). A novel tri-component scheme for classifying neuronal discharge patterns. *Journal of Neuroscience Methods*, *239*, 148–161.

Lansky, P., Sacerdote, L., & Zucca, C. (2016). The Gamma renewal process as an output of the diffusion leaky integrate-and-fire neuronal model. *Biological Cybernetics*, *110*, 193–200.

McDonnell, M. D., Ikeda, S., & Manton, J. H. (2011). An introductory review of information theory in the context of computational neuroscience. *Biological Cybernetics*, *105*, 55–70.

Nawrot, M. P., Bouscein, C., Molina, V. R., Riehle, A., Aertsen, A., & Rotter, S. (2008). Measurement of variability dynamics in cortical spike trains. *Journal of Neuroscience Methods*, *169*, 374–390.

Omi, T., & Shinomoto, S. (2011). Optimizing time histograms for non-poissonian spike trains. *Neural Computation*, *23*, 3125–3144.

Ostojic, S. (2011). Interspike interval distributions of spiking neurons driven by fluctuating inputs. *Journal of Neurophysiology*, *106*, 361–373.

Pipa, G., Gruen, S., & van Vreeswijk, C. (2013). Impact of spike train autostructure on probability distribution of joint spike events. *Neural Computation*, *25*, 1123–1163.

Rajdl, K., & Lansky, P. (2014). Fano factor estimation. *Mathematical Biosciences and Engineering*, *11*, 105–123.

Rieke, F., Warl, D., de Ruyter van Steveninck, R., & Bialek, W. (1999). *Spikes: exploring the neural code*. Cambridge: MIT Press.

Ruigrok, T. H. J., Hensbroek, R. A., & Simpson, J. I. (2011). Spontaneous activity signatures of morphologically identified interneurons in the vestibulocerebellum. *Journal of Neuroscience*, *31*, 712–724.

Schurmann, T. (2004). Bias analysis in entropy estimation. *Journal of Physics A: Mathematical and General*, *37*, 295–301.

Schwalger, T., Droste, F., & Lindner, B. (2015). Statistical structure of neural spiking under non-Poissonian or other non-white stimulation. *Journal of Computational Neuroscience*, *39*, 29–51.

Shannon, C. E., & Weaver, W. (1998). *The mathematical theory of communication*. Illinois: University of Illinois Press.

Shimokawa, T., Koyama, S., & Shinomoto, S. (2010). A characterization of the time-rescaled gamma process as a model for spike trains. *Journal of Computational Neuroscience*, *29*, 183–191.

Shimokawa, T., & Shinomoto, S. (2009). Estimating instantaneous irregularity of neuronal firing. *Neural Computation*, *21*, 1931–1951.

Shinomoto, S., Miura, K., & Koyama, S. (2005). A measure of local variation of interspike intervals. *BioSystems*, *79*, 67–72.

Steuer, R., Ebeling, W., Russell, D. F., Bahar, S., Neiman, A., & Moss, F. (2001). Entropy and local uncertainty of data from sensory neurons. *Physical Review E*, *64*, 061911.

Stevenson, I. H. (2016). Flexible models for spike count data with both over- and under- dispersion. *Journal of Computational Neuroscience*, *41*, 29–43.

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

Watters, N., & Reeke, G. N. (2014). Neuronal spike train entropy estimation by history clustering. *Neural Computation*, *26*, 1840–1872.