Chinese Journal of Physiology 53(6): 454-464, 2010
DOI: 10.4077/CJP.2010.AMM031

Neuronal Jitter: Can We Measure the Spike Timing Dispersion Differently?

Lubomir Kostal[1] **and Petr Marsalek**[1, 2, 3, 4]

[1]*Institute of Physiology AS CR, v.v.i., Videnska 1083, CZ-142 20, Praha 4, Czech Republic*
[2]*Charles University of Prague, Department of Pathological Physiology, U nemocnice 5,*
CZ-128 53, Praha 2, Czech Republic
[3]*Czech Technical University of Prague, Faculty of Biomedical Engineering, Nam. Sitna 3105,*
CZ-272 01, Kladno, Czech Republic
and
[4]*Max Planck Institute for the Physics of Complex Systems, Nöthnitzer Str. 38,*
D-011 87, Dresden, Germany

Abstract

We propose a novel measure of statistical dispersion of a positive continuous random variable: the entropy-based dispersion (ED). We discuss the properties of ED and contrast them with the widely employed standard deviation (SD) measure. We show that the properties of SD and ED are different: while SD is a second moment characteristics measuring the dispersion relative to the mean value, ED measures an effective spread of the probability distribution and is more closely related to the notion of randomness of spiking activity. We apply both SD and ED to analyze the temporal precision of neuronal spiking activity of the perfect integrate-and-fire model, which is a plausible neural model under the assumption of high input synaptic activity. We show that SD and ED may give strikingly different results for some widely used models of presynaptic activity.

Key Words: perfect integrator neuronal model, standard deviation, entropy, spike timing jitter

Introduction

Generally, neurons communicate by employing 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 axonal excitable membrane. It is widely accepted, that information in neuronal systems is transferred by employing these spikes. 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 lengths of interspike intervals (ISIs) between two successive spikes in a spike train often vary, apparently randomly, both within and across trials (21, 42, 44).

There are two main hypotheses that attempt to describe the ways by which the spikes carry information: the frequency, or rate coding hypothesis and the temporal spike coding hypothesis (21, 26, 36, 45). In the rate coding scheme it is assumed, that the information sent along the axon is encoded in the number of spikes in a given time window (the number is the firing rate) (2, 21, 29). Temporal codes, on the other hand, employ those features of the spiking activity, that cannot be described by the firing rate, but can be described by spike firing relative to some other spike time. For example, a neural code based on a time to first spike after the stimulus onset, or characteristics based on the second and higher statistical moments of the ISI probability distribution, or precisely timed

---

Corresponding author: Dr. Lubomir Kostal, Institute of Physiology AS CR, v.v.i., Videnska 1083, CZ-142 20, Praha 4, Czech Republic. Tel: +420-24106-2276, Fax: +420-29644-2488, E-mail: kostal@biomed.cas.cz
Received: December 11, 2009; Revised: January 26, 2010; Accepted: January 27, 2010.
ISSN : 0304-4920. http://www.cps.org.tw

groups of spikes, temporal patterns, are all candidates for temporal codes (8, 21, 26, 36, 41). The rate coding scheme is therefore typically considered under stable time conditions. On the other hand, temporal codes are assumed to play a role whenever precise spike timing is important (32, 36, 41). Such situation occurs, for example, as one descends into the central nervous system from the peripheral auditory system in bats, where certain neurons fire with precise latency in response to the sound (14). Precise timing and coincidence detection are important for a variety of other neuronal systems (22, 33) and this was a subject for many theoretical studies, see (20, 32, 39) and references therein.

Traditionally, the spike-timing precision is described by employing the standard deviation (SD) of interspike interval or of a first peristimulus spike time (32, 39). The goal of this paper is to apply a different measure of statistical dispersion: the entropy-based dispersion measure (ED). Although SD is used ubiquitously and is almost synonymous to the measure of statistical dispersion, we show, that SD is not well suited to quantify some aspects of dispersion that are often expected intuitively. For example, SD does describe how far from the center are the values, and not rather how diverse the values are. The motivation behind ED is rooted in the measure of *randomness* (26), which is based on the information-theoretic quantities, such as entropy and Kullback-Leibler divergence (13). To illustrate and discuss the differences between ED and SD, we investigate the spiking activity of presynaptic neurons, described by several frequently used models of interspike interval probability density function: exponential, gamma, inverse Gaussian or bimodal log-normal. We show that SD and ED behave similarly for simple models of presynaptic activity, especially when the variation coeffcient C_V of input spike trains inter-spike intervals is not greater than one, $C_V \leq 1$. However, for bimodal lognormal or over-dispersed input spike trains with the coeffcient of variation larger than one, the results obtained by SD and ED differ strikingly. We also show, how the ED and SD might pinpoint some optimal regimes and parameter values of presynaptic spike trains.

Mathematical Framework

Measures of Statistical Dispersion

In order to describe and analyze neuronal firing, statistical methods and methods of probability theory and stochastic point processes are widely applied (15, 23, 35, 46). The probabilistic description of the spiking originates from the fact, that the positions of spikes cannot be predicted deterministically, only the probability that a spike occurs can be given (21). Let us assume, that the time to the first spike (or the inter-spike interval generally) is described by a continuous, positive random variable T. One of the most common probabilistic descriptors of T is the probability density function $f(t)$, defined so that $f(t)dt$ is the probability that spike occurs in an infinitesimally small time interval $[t; t+dt]$ (35). Probability density function is usually estimated from the data by means of histograms.

Generally, statistical dispersion is the *variability* or *spread* of the random variable T. The measure of statistical dispersion usually has the same physical units as T. There are many different measures of statistical dispersion, employed in different contexts, for example standard deviation, inter-quantile range or mean difference (11, 16, 24). By far, the most common measure is the standard deviation, σ, defined as

$$\sigma = \left[\int_0^\infty [t - E(T)]^2 f(t)dt \right]^{1/2}, \qquad [1]$$

where $E(T)$ is the mean value of T,

$$E(T) = \int_0^\infty t \; f(t)dt \, . \qquad [2]$$

Equivalently, the square of σ is the variance of T, $\sigma^2 = \mathrm{Var}(T)$. The relative measure of dispersion, with respect to the mean value, based on σ, is the coeffcient of variation, C_V, defined as

$$C_V = \frac{\sigma}{E(T)} \, . \qquad [3]$$

Besides the mean value, the C_V is one of the most frequently used characteristics of interspike intervals. The main advantage of the C_V as a characteristics of latency, or of interspike interval, as compared to σ is, that C_V is dimension-less and its value does not depend on the choice of units of T (e.g., seconds or milliseconds) and thus probability distributions with different means can be compared meaningfully (43). Furthermore, the observed C_V of interspike intervals is related to the variability coding hypothesis (10, 18, 36, 38).

From equation [1] we see, that σ measures essentially how off-centered is the distribution of T. The value of σ grows, as the probability of values close to $E(T)$ decreases. Furthermore, since the distance from the mean value, $(t - E(T))$, is squared in equation [1], it follows that σ is sensitive to outlying values. Most importantly, σ does not quantify how random, or unpredictable, are the latencies or interspike intervals described by random variable T.

Indeed, high value of σ (high variability) does not indicate that the probability distribution of T is close to the uniform distribution, in other words that the probabilities of different values that T can take are as evenly distributed as possible (26). Main motivation of this paper is to propose such measure of statistical dispersion, that would describe randomness of the probability distribution of T.

In order to proceed on, we extend our previous work done on measuring randomness of neuronal activity (26, 27), based on the information-theoretic concept of entropy (13). Informally, the entropy measures the choice of different values that T may take. The entropy $h(T)$ of the random variable T with probability density function $f(t)$ is defined as

$$h(T) = -\int_0^\infty f(t) \ln f(t) dt \,. \qquad [4]$$

However, $h(T)$ does not have the same properties and intuitive interpretation as the original Shannon's entropy of discrete random variables. Namely, the value of $h(T)$ may be positive or negative and changes with coordinate transforms. This way it depends on the choice of units of T, let it be seconds or milliseconds. Since the probability density function $f(t)$ has a physical dimension, it is a derivative of probability with respect to the variable t, $h(T)$ has physical dimension of the logarithm of the variable t, let it be the logarithm of second. Therefore $h(T)$ is not directly usable as a measure of statistical dispersion.

In order to obtain a properly behaved quantity, we propose a normalized entropy, η, as follows (26, 27). First, a new, dimensionless random variable Θ with $E(\Theta) = 1$ is obtained from T, by

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

Normalized entropy is then defined as the entropy of Θ, $h(\Theta)$. After a change of variables in integral [4] we obtain (27)

$$\eta \equiv h(\Theta) = h(T) - \ln E(T). \qquad [6]$$

Said in another way, formula [6] represents a unique decomposition of entropy $h(T)$ into a sum of 2 terms: $\ln E(T)$ and a dimensionless number $\eta(T)$. It follows from equation [6], that the exponential of $h(T)$ can be conveniently expressed as

$$\zeta \equiv \exp h(T) = E(T) \exp \eta. \qquad [7]$$

The definition of ζ is central to our paper and in the following we show that it can be interpreted as a dispersion measure and discuss its properties. The interpretation of ζ relies on the consequences of the asymptotic equipartition property theorem (AEP) (3, 13), which we briefly review in the Appendix. The main conclusion relevant for our purposes is that ζ is related to a volume of the typical set, $S_\delta^{(n)}$, associated with T. This is the set of almost all possible values that T can take (for the precise definition see the Appendix). For n independent realizations of T the following relation holds

$$\zeta \approx \sqrt[n]{\mathrm{vol}[S_\delta^{(n)}]} \,. \qquad [8]$$

The formula [8] means that ζ is the side length of a n-dimensional cube with the same volume as the asymptotic typical set. In other words, since we consider only sequences $T_1, \ldots, T_n$ of independent identically distributed random variables, asymptotically almost any observed sequence $\{\tau_1, \ldots, \tau_n\}$ comes from a limited subset of the whole support. The length of the support is ζ per observation (see Appendix for details). Thus, ζ can be considered as a type of measure of statistical dispersion of a random variable T. Small values of ζ mean that most of the probability is concentrated, while high values indicate spreaded distribution. However, the distribution spread measured by σ and ζ is different. The ζ increases as the probability of different values of T gets more and more uniform, which increases the volume of $S_\delta^{(n)}$. In other words, ζ measures how evenly is the probability distributed over the entire support of T. From this point of view, ζ is more appropriate if we wish to know how diverse the values of T are. Typically, the greatest difference between σ and ζ should be expected in multimodal distributions, as we later show for the bimodal lognormal model. Furthermore, 3 is directly related by equation [7] to the notion randomness, η, of spiking neuronal activity, see (26) for details.

Next we will derive one more useful relation for ζ. Let fexp(t) be a probability density function of an exponential distribution, $f_{\exp}(t) = \exp(-t/\lambda)/\lambda$, with the mean value equal to $E(T)$, $\lambda = E(T)$. Then formula [7] can be written as

$$\zeta = E(T) \exp[1 - D_{KL}(f \,\|\, f_{\exp})] \,, \qquad [9]$$

where $D_{\mathrm{KL}}(f \,\|\, g)$ is the Kullback-Leibler divergence (13),

$$D_{\mathrm{KL}}(f \,\|\, g) = \int_0^\infty f(t) \ln \frac{f(t)}{g(t)} dt \,. \qquad [10]$$

It is useful to scale the values of ζ with the base of natural logarithm, e, and define ζ_e in the following way

$$\zeta_e = \frac{\zeta}{e} = E(T)\exp[-D_{\mathrm{KL}}(f\,\|\,f_{\mathrm{exp}})]\,. \tag{11}$$

In analogy to the definition of C_V, equation [3], as a relative dispersion measure based on σ, it is straightforward to arrive to a relative dispersion measure based on ζ_e, here denoted simply as $\zeta_e/E(T)$,

$$\zeta_e/E(T) = \exp[-D_{\mathrm{KL}}(f\,\|\,f_{\mathrm{exp}})]\,. \tag{12}$$

By recalling that the entropy of exponential distribution is given by

$$h(T) = 1 + \ln E(T), \tag{13}$$

see reference (13), we also see, that ζ_e can be interpreted as a standard deviation, σ, of exponential distribution with the same entropy as the distribution in question, described by probability density function $f(t)$. This argument further supports ζ, or ζ_e, as a valid measure of statistical dispersion. We also note, that entropy, $h(T)$, is defined uniquely up to a multiplicative constant, or the logarithm base in equation [4], see (3) for details. However, from formula [27] follows, that the exponential function must be taken within the same base as the logarithm in the entropy definition. Therefore ζ is unique, independent on the logarithm base. Finally, the units of σ and ζ are the same, thus allowing for a direct comparison between these two quantities.

Perfect Integrate-and-Fire Model

Integrate and fire neuronal model, perfect integrator (9, 21, 46), is one of the simplest neural models, yet it captures the integration property of neuron, especially for time scales shorter than the neuronal time constant, which can be measured electro-physiologically. In this paper we illustrate the effect of quantifying the dispersion of spike timing by employing different measures: σ and ζ. We consider a set of n presynaptic neurons. At time $t = 0$ there is a stimulus onset, and the time to the first spike of each of those n neurons is described by the same probability density function, $f_{\mathrm{in}}(t)$. The target neuron is described by the perfect integrator model. It fires after first k spikes ($k \le n$) from presynaptic neurons are received. Schematically, we write

$$\left.\begin{array}{c} 1-\text{st input}: f_{in}(t) \\ \vdots \\ n-\text{th input}: f_{in}(t) \end{array}\right\} \rightarrow \text{first (k) spikes} \rightarrow f_{\mathrm{out}}(t)\,. \tag{14}$$

The probability density function, $f_{\mathrm{out}}(t)$, of time to the first spike of the target neuron, perfect integrator, assuming that each input neuron fires only once, is given by the order statistics (17)

$$f_{\mathrm{out}}(t) = \frac{n!k}{k!(n-k)!}[F_{\mathrm{in}}(t)]^{k-1}[1-F_{\mathrm{in}}(t)]^{n-k} f_{\mathrm{in}}(t)\,, \tag{15}$$

where $F_{\mathrm{in}}(t)$ is the cumulative distribution function of the individual input random variables,

$$F_{\mathrm{in}}(t) = \int_0^t f_{\mathrm{in}}(z)dz\,. \tag{16}$$

Results

In this section we use several input timing densities $f_{\mathrm{in}}(t)$ in the perfect integrate-and-fire model to produce output densities $f_{\mathrm{out}}(t)$. We compare the statistical dispersion measures of these densities. The measures are: σ given by equation [1], C_V given by equation [3], ζ_e given by equation [11] and $\zeta_e/E(T)$ given by equation [12]. The individual presynaptic activities are given by the input probability density functions $f_{\mathrm{in}}(t)$. As the input densities we employ several widely used models of interspike interval distributions. The output densities are densities of the time to the first spike, or latency, of the perfect integrate-and-fire neural model. They are described by random variable T_{out} with probability density function $f_{\mathrm{out}}(t)$ defined by equation [15]. For selected input function $f_{\mathrm{in}}(t)$ we examine the dependence of the dispersion measures on k (the number of presynaptic spikes required for the perfect integrator neuron to spike) and n (the total number of presynaptic neurons). We distinguish two situations: a) fixed n and variable k, by which we study the effect of perfect integrator threshold (k) for a given intensity (n) of synaptic activity; b) fixed k and variable n, by which we study the effect of increasing synaptic input for a given threshold.

Gamma Model

Gamma distribution is one of the most frequent statistical descriptors of interspike intervals employed in experimental data analysis (31, 34, 37, 40). The probability density function of gamma distribution, parameterized by its mean value, μ, and coefficient of variation, C_V, instead by the usual pair of μ and σ, is

$$f_{in}(t) = \left(\frac{1}{C_V^2\mu}\right)^{1/C_V^2}\Gamma(1/C_V^2)t^{1/C_V^2-1}\exp\left(-\frac{t}{C_V^2\mu}\right), \tag{17}$$

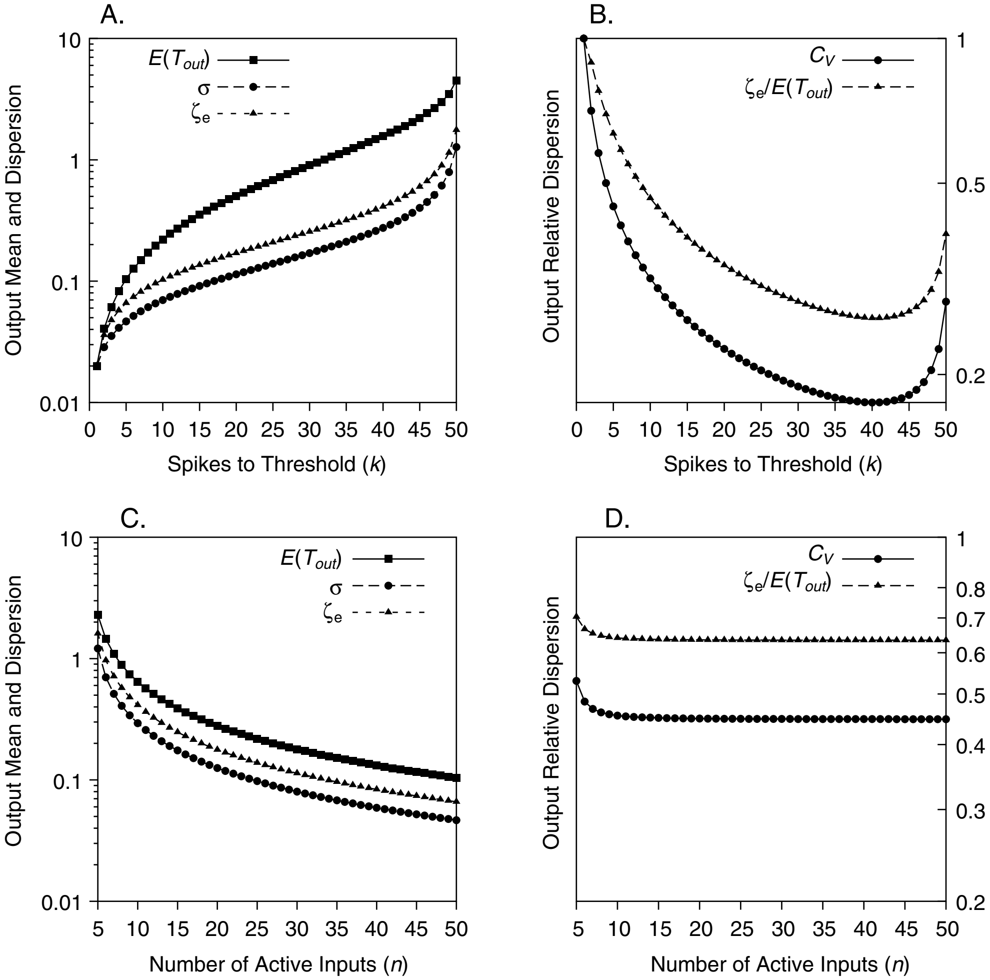

Fig. 1. In this and all subsequent figures, five characteristics of the output spike timing density $f_{\text{out}}(t)$ are shown: mean time $E(T_{\text{out}})$, standard deviation and entropy based dispersion as dimensional values (σ and $\zeta_e = \zeta/e$, panels A and C) and relative values (C_V and $\zeta_e/E(T_{\text{out}})$, panels B and D), in dependency on varying parameter k, panels A and B, and parameter n, panels C and D. Input model f_{in} at this figure has the exponential distribution with $E(T) = 1$ s.

see reference (28), where $\Gamma(z) = \int_0^\infty t^{z-1}\exp(-t)dt$ is the gamma function (1). For $C_V = 1$ the gamma distribution becomes exponential distribution, thus representing the canonical case of spiking being described by a homogeneous Poisson process. Due to its simplicity, the exponential model of presynaptic activity is the only one, for which the output characteristics can be expressed by means of tabulated functions. The mean value of time to the first spike, $E(T_{\text{out}})$, of the perfect integrator neuron then is

$$E(T_{\text{out}}) = \mu H_n - \mu H_{n-k}, \qquad [18]$$

where H_n is the n-th harmonic number (1). Standard deviation of T_{out} becomes

$$\sigma = \mu\sqrt{\psi'(1+n) + \psi'(1-k+n)}\,, \qquad [19]$$

where $\psi'(z)$ is the first derivative of the digamma function (1), and entropy based dispersion measure of T_{out} is

$$\zeta = \frac{\mu\Gamma(k)(n-k)!}{n!}\exp[(1-k)\psi(k) + n\psi(1+n) + (k-n-1)\psi(k+n-1)]\,, \qquad [20]$$

where $\Gamma(z)$ is the gamma function. Fig. 1 shows the comparison between different dispersion measures for the simple case of exponential presynaptic activity. Although the absolute numerical values differ, qualitatively there is no difference between measures based on standard deviation or entropy. We calculated the quantities for the presynaptic gamma model for the case $C_V < 1$, and detected again qualitatively similar behavior between the two measures, thus we do not show the corresponding figure here. Generally, we

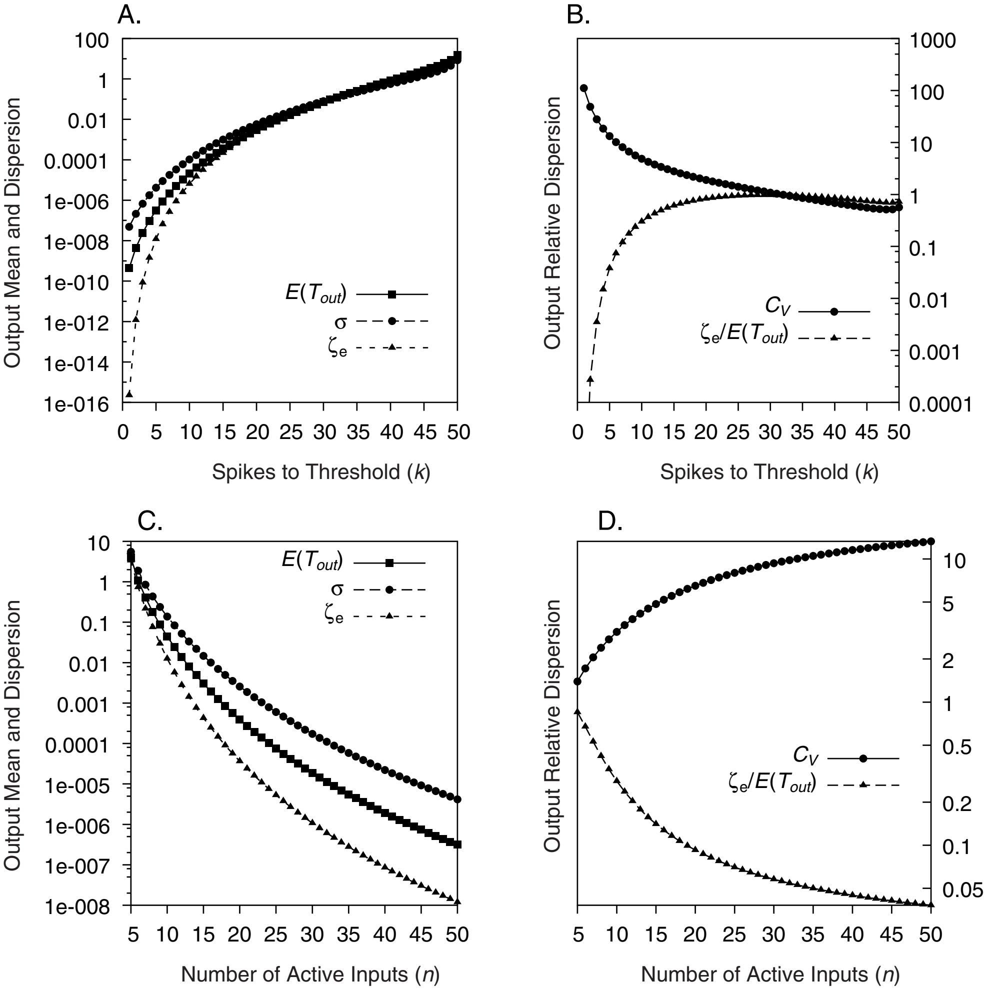

Fig. 2. Input distribution f_{in} is gamma distribution with $E(T) = 1$ s, $\sigma = 4$ s and $\zeta = 0.007$ s. With the rising number of synaptic inputs, the dimensional versions of the two measures are alike, yet the relative measures differ. In this and all other figures we can see that input distributions transformed by equation [15] exhibit distinguished output behaviors.

do not expect σ and ζ_e and their relative counterparts C_V and $\zeta_e/E(T)$, to give qualitatively different results for $C_V < 1$, whenever gamma, lognormal, inverse gaussian or the distribution resulting from the leaky integrate and fire model are employed, see (25, 27) for details. Fig. 2 shows the situation for $C_V > 1$ on the input. While qualitatively σ and ζ_e agree for the output latencies, the respective C_V and $\zeta_e/E(T)$ measures do not.

Inverse Gaussian Model

The inverse Gaussian distribution (12) is often used to describe neural activity and fitted to experimentally observed ISIs (4, 19, 31, 37). This distribution is the first threshold passage time of the neuronal membrane potential modeled by the Wiener process with positive drift. Simply, in this model the neuron depolarization has a linear trend to the threshold. This distribution describes the spiking activity of a stochastic variant of the perfect integrator: the non leaky integrate and fire stochastic neuronal model (30). The probability density of the inverse Gaussian distribution can be expressed as

$$f(t) = \sqrt{\frac{\mu}{2\pi C_V^2 t^3}}\exp\left[-\frac{1}{2C_V^2\mu}\frac{(t-\mu)^2}{t}\right]. \qquad [21]$$

In Fig. 3 we show the behavior of this model, which can give us an inference of how network of several layers of these neuronal models can behave, (32, 46).

Bimodal Lognormal Model

The lognormal distribution of interspike intervals, with some exceptions (5), is rarely presented as

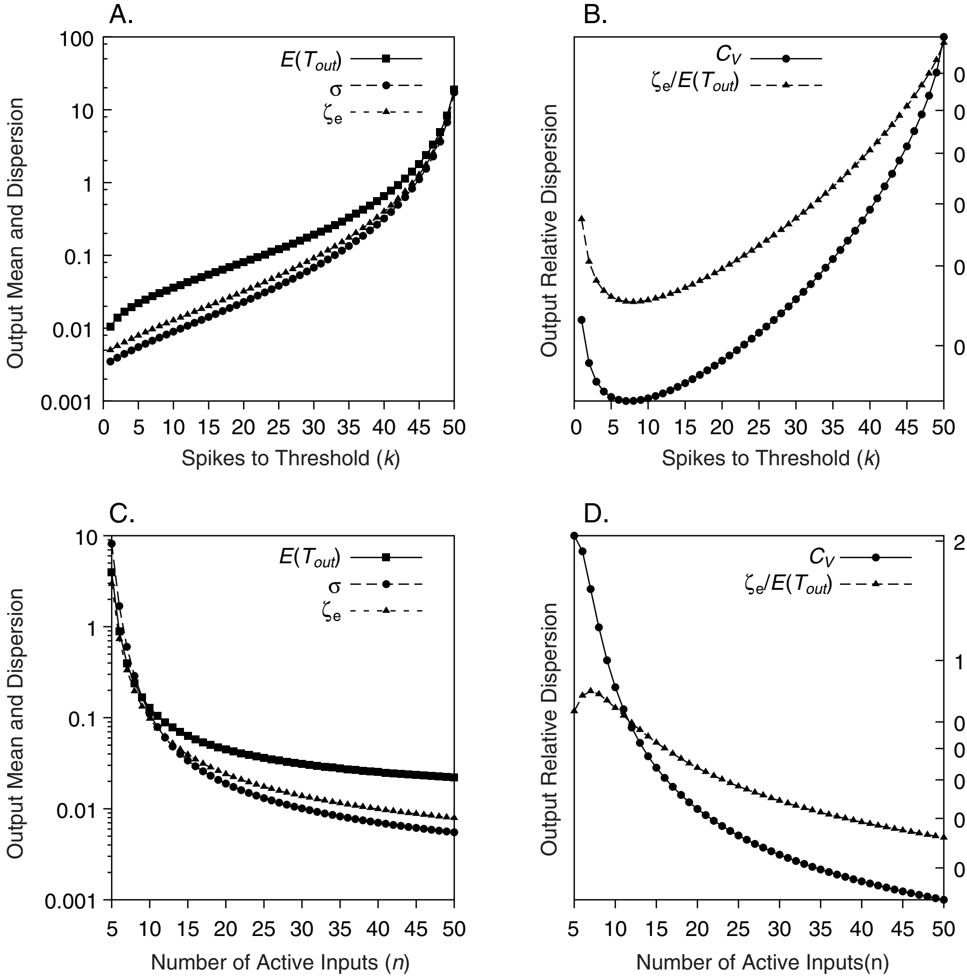

Fig. 3. Input model f_{in} is inverse Gaussian distribution with $E(T) = 1$ s, $\sigma = 4$ s and $\zeta = 0.39$ s. This particular distribution is not only frequently observed in experimental neuronal firing, but is also a natural description of output firing of a stochastic neuronal model. Note the non-monotonic behavior of the $\zeta_e/E(T_{\mathrm{out}})$, see the text for details.

a result of a neuronal model. However, it represents quite a common descriptor in experimental data analysis (31, 37). A mixture of two lognormal distributions has been used recently in identification of the ISI of the supraoptic nucleus activity (6, 7). The lognormal probability density function, parameterized by the mean value μ and coefficient of variation C_V, instead by the usual pair of μ and σ, is

$$f_{1n}(t; \mu, C_V) = \frac{1}{t\sqrt{2\pi \ln(1 + C_V^2)}} \exp \cdot \left\{ -\frac{1}{8} \frac{[\ln(1 + C_V^2) + 2\ln(t/\mu)]^2}{\ln(1 + C_V^2)} \right\}. \qquad [22]$$

The bimodal lognormal probability density function can be expressed as a mixture of two lognormal probability density functions (6) as

$$f_{\mathrm{ln}}^{(\mathrm{mix})}(t; \mu_1, \mu_2, C_{V1}, C_{V2}, p) = p\, f_{\mathrm{ln}}(t; \mu_1, C_{V1}) + (1 - p)\, f_{\mathrm{ln}}(t; \mu_2, C_{V2}), \qquad [23]$$

where $0 < p < 1$. The mean value 1 and standard deviation σ of the mixture can be calculated to be

$$\mu = p(\mu_1 - \mu_2) + \mu_2 \qquad [24]$$

and

$$\sigma = \sqrt{C_V^2\mu_2^2 - p^2(\mu_1 - \mu_2)^2 + p[(1 + C_{V1}^2)\mu_1^2 - 2\mu_1\mu_2 - (C_{V2}^2 - 1)\mu_2^2]}\,. \qquad [25]$$

Thus, the same values of μ and σ of the bimodal lognormal model can be achieved by employing different parameter values of the original lognormal distributions in the mixture [23]. The bimodality of the input distribution of latencies allows us to examine

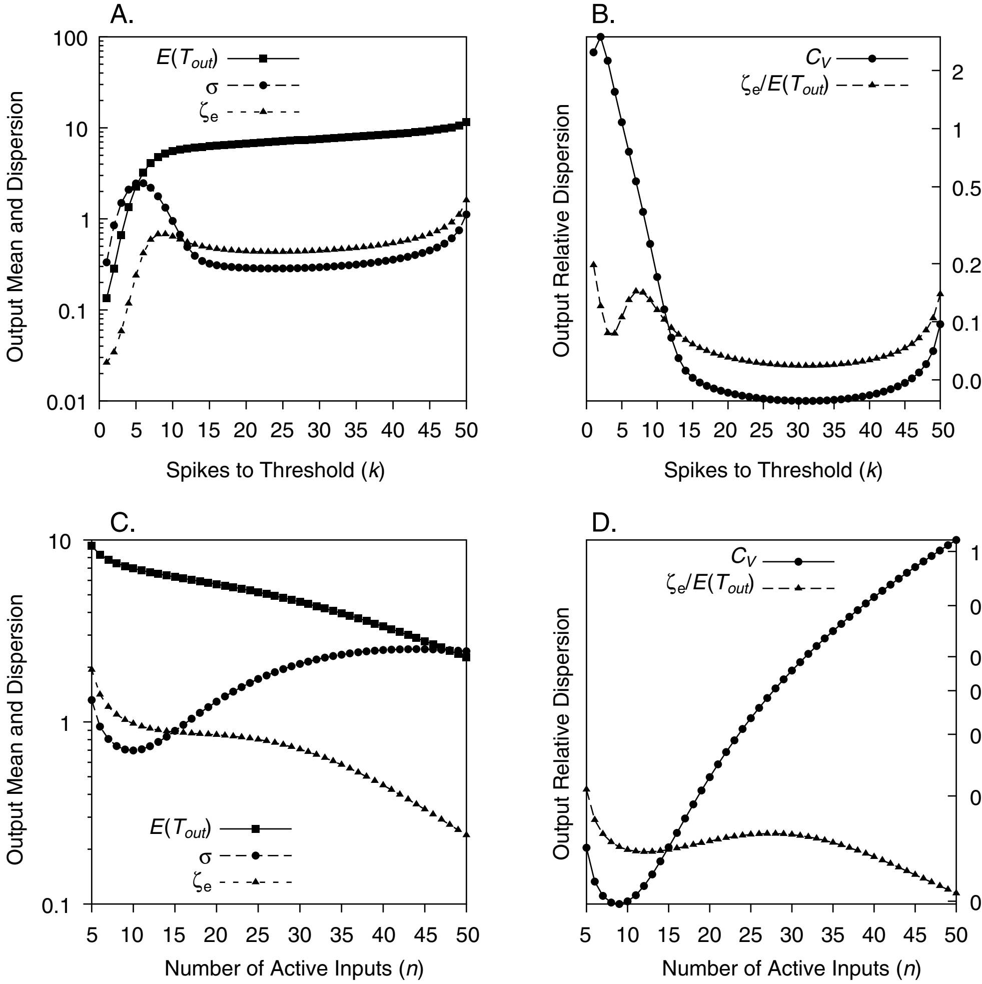

Fig. 4. Input model f_{in} is bimodal lognormal distribution, $E(T) = 6.8$ s, $\sigma = 2.65$ s and $\zeta = 2.08$ s. The bimodality of the distribution, frequently observed in data, yields higher C_V. This shows bimodality of critical neuronal regimes, which might be regarded as close to regimes optimal for neural coding. Here in panel C this can be seen for the σ curve, but not for the ζ_e curve.

striking differences between σ and ζ (and their relative counterparts), such as shown in Figs. 4 and 5.

Discussion

The aim of this study is to propose an alternative measure of spike timing variability, we call it entropy based dispersion and denote ζ. This measure reflects better some information theoretic properties of distributions, therefore it is useful, whenever we discuss possible computations and encodings performed by a single neuron. To demonstrate the properties of the new dispersion measure, we employ one of the simplest neuronal model, the (perfect) integrate and fire model.

What is the mechanism of firing in the perfect integrate and fire model? The variable corresponding to the membrane potential of the real neuron in the perfect integrate and fire model reaches the threshold without decay, as compared to the leaky integrate and fire model. Therefore with all parameters equal, perfect integrator will have higher firing rate and lower standard deviation, compared to the leaky integrate and fire. Also both models are equal when the membrane time constant τ value approaches infinity (9, 21). These are the reasons why results obtained in perfect integrator are representative for the class of simplified one point models, since the higher values of σ we get, more interesting values we get, possibly close to bimodal distributions.

When the amount of synaptic inputs bombarding the neuron is sufficiently high, $E(T_{\text{out}}) \ll E(T_{\text{in}})$, results obtained in both respective types of neuronal model are close. To capture the behavior of the synaptic transmission in other regimes, our systematic investigation of number of spikes needed to the threshold as one of the parameters gives the timing precision of the model when the number of synaptic input is low and therefore the leaky integrator and perfect integrator results may differ.

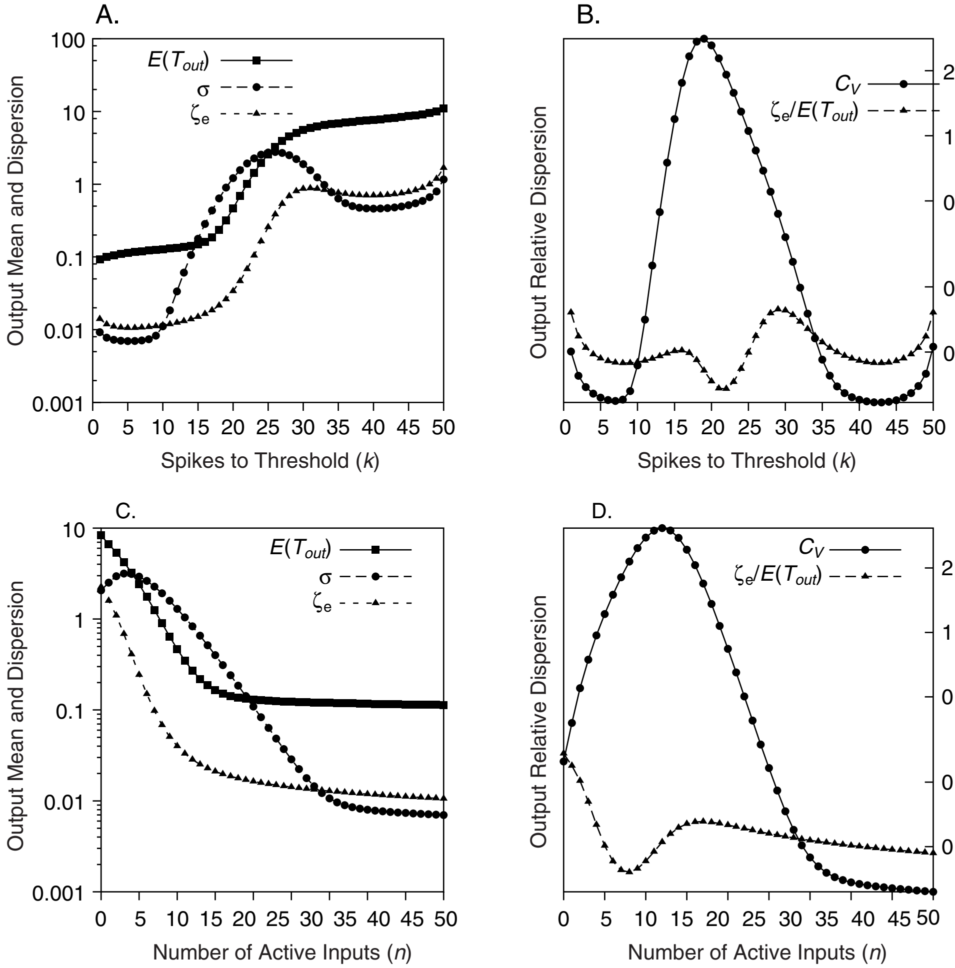

Fig. 5. Input model f_{in} is bimodal lognormal distribution, parameterized with $E(T) = 3.83$ s, $\sigma = 3.85$ s and $\zeta = 0.61$ s. The comparison with the previous figure clearly shows how different regimes of the same neural model with the same input distribution only with different scaling arrives to several critical points shown by local extremes.

Multimodal data are important, because multimodal ISI reflects either that different regimes of neural firing are present (like regular firing and bursting), or they reflect the fine timing structure in the input, like the neural firing in the auditory nerve or in the visual cortical area MT, (32). Multimodal spike trains can be found not only in sensory systems like the previous two examples, but also in the supraoptic neurons interfacing the central and autonomous system in mammals, (6) and elsewhere.

We have following sets of parameters and distributions, where our results are of particular importance. First, interesting results are obtained, when ζ_e and σ attain local extremes, for example in Fig. 1 for $k = 40$ and analogously in Fig. 5 A for $k = 30$ and more than one local extreme, like in Fig. 4 and 5.

While σ measures how far from the central value $E(T)$ is the ISI distribution, ζ measures how evenly is the probability distributed over the entire supporting interval. From this point of view, ζ is more appropriate if we wish to know how diverse in a sense of transmitting meaningful information are the values that T can take. The ζ can also measure, how random the T values are. Both ζ and σ give similar results for simple models of presynaptic activity, especially if it holds $C_V < 1$ for presynaptic latencies. For bimodal probability density functions or over-dispersed presynaptic latencies with $C_V > 1$ the results obtained by ζ and σ differ.

Acknowledgments

This work was supported by the research projects MSM 0021620806 and MSM 6840770012 granted by the Ministry of Education, Youth and Sports of the Czech Republic, by Volkswagen Foundation grant I/83892, by research project AV0Z50110509 and by Centre for Neuroscience LC554.

Appendix: Asymptotic Equipartition Property

Analogously to the AEP for discrete random variables (3) we consider a sequence $T_1, T_2, \dots, T_n$ of independent identically distributed random variables with joint probability density function $f(t_1, t_2, \dots, t_n) = \Pi_{i=1}^{n} f(t)$. Then for a sequence of realizations $\{T_1 = \tau_1, \dots, T_n = \tau_n\}$ we have by the weak law of large numbers

$$-\frac{1}{n} \ln f(\tau_1, \dots, \tau_n) \to h(T)\,, \qquad [26]$$

where the convergence with increasing n is in probability and $h(T)$ is the entropy of the probability density $f(t)$. Said more precisely, for a fixed $\varepsilon > 0$ and $\delta > 0$ we can find n such, that for any sequence $\{\tau_1, \dots, \tau_n\}$ holds

$$\Pr\left[\left| -\frac{1}{n} \ln f(\tau_1, \dots, \tau_n) - h(T) \right| \le \delta\right] \ge 1 - \varepsilon\,. \qquad [27]$$

Expression [27] leads to the definition of typical set $S_\delta^{(n)}$, which contains sequences that satisfy

$$S_\delta^{(n)} = \left\{ \{\tau_1, \dots, \tau_n\} : \left| -\frac{1}{n} \ln f(\tau_1, \dots, \tau_n) - h(T) \right| \le \delta \right\}. \qquad [28]$$

Essentially, the "typicality" of $S_\delta^{(n)}$ means that a randomly drawn sequence of values $\{\tau_1, \dots, \tau_n\}$ probably belongs to $S_\delta^{(n)}$, and this probability can be made arbitrarily close to one for n sufficiently large.

By rewriting the definition [28] we have for any $\{\tau_1, \dots, \tau_n\} \in S_\delta^{(n)}$

$$\begin{aligned} \exp\,(-n(h(T) + \delta)) &\le f(\tau_1, \dots, \tau_n) \\ &\le \exp\,(-n(h(T) - \delta)). \end{aligned} \qquad [29]$$

We employ formulas [27] and [29] to calculate bounds on the volume of the typical set,

$$\mathrm{vol}[S_\delta^{(n)}] = \int_{S_\delta^{(n)}} dt_1 \dots dt_n\,. \qquad [30]$$

From expression [27] we have $\Pr[S_\delta^{(n)}] \ge 1 - \varepsilon$, in other words

$$\int_{S_\delta^{(n)}} f(t_1, \dots, t_n) dt_1 \dots dt_n \ge 1 - \varepsilon\,. \qquad [31]$$

By substituting the upper bound from inequality [29], formula [31] becomes

$$\exp(-n(h(T) - \delta)) \int_{S_\delta^{(n)}} dt_1 \dots dt_n \ge 1 - \varepsilon\,, \qquad [32]$$

since $\exp(-n(h(T) - \delta))$ is constant with respect to the integration. To employ the lower bound from inequality [29] we start from the fact that $S_\delta^{(n)}$ cannot exceed the support of $f(t_1, \dots, t_n)$ and therefore

$$\int_{S_\delta^{(n)}} f(t_1, \dots, t_n) dt_1 \dots dt_n \le 1\,, \qquad [33]$$

and by similar argument as before we come to

$$\exp(-n(h(T) + \delta)) \int_{S_\delta^{(n)}} dt_1 \dots dt_n \le 1\,. \qquad [34]$$

Combining formulas [30], [32] and [34] gives the following bounds on the volume of $S_\delta^{(n)}$,

$$\begin{aligned} (1 - \varepsilon)\, \exp(n(h(T) - \delta)) &\le \mathrm{vol}[S_\delta^{(n)}] \\ &\le \exp(n(h(T) + \delta)). \end{aligned} \qquad [35]$$

Finally, comparing formulas [7] and [35] yields formula [8].

References

1. Abramowitz, M. and Stegun, I. Handbook of Mathematical Functions, With Formulas, Graphs, and Mathematical Tables. New York: Dover, 1965.
2. Adrian, E. The basis of sensation. New York: W.W. Norton and Co., 1928.
3. Ash, R.B. Information Theory. New York: Dover, 1965.
4. 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.
5. Bershadskii, A., Dremencov, E., Fukayama, D. and Yadid, G. Probabilistic properties of neuron spiking time-series obtained *in vivo*. *Eur. Phys. J. B* 24: 409-413, 2001.
6. Bhumbra, G. and Dyball, R. Measuring spike coding in the rat supraoptic nucleus. *J. Physiol.* 555: 281-296, 2003.
7. Bhumbra, G., Inyushkin, A. and Dyball, R. Assessment of spike activity in the supraoptic nucleus. *J. Neuroendocrinol.* 16: 390-397, 2004.
8. Buracas, G. and Albright, T. Gauging sensory representations in the brain. *Trends Neurosci.* 22: 303-309, 1999.
9. Burkitt, A. A review of the integrate-and-fire neuron model: I. Homogeneous synaptic input. *Biol. Cybern.* 95: 1-19, 2006.
10. Burns, B. and Pritchard, R. Contrast discrimination by neurons in the cat's visual cerebral cortex. *J. Physiol.* 175: 445-463, 1964.
11. Chakravarty, S. Ethical social index numbers. New York: Springer-Verlag, 1990.
12. Chhikara, R. and Folks, J. The Inverse Gaussian Distribution: theory, methodology, and applications. New York: Marcel Dekker, 1989.
13. Cover, T. and Thomas, J. Elements of information theory. New York: Wiley, 1991.
14. Covey, E. and Casseday, J. The monaural nuclei of the lateral lemniscus in an echolocating bat: parallel pathways for analyzing temporal features of sound. *J. Neurosci.* 11: 3456, 1991.
15. Cox, D. and Lewis, P. The statistical analysis of series of events. London: Chapman and Hall, 1966.
16. Cramer, H. Mathematical methods of statistics. Princeton: Princeton University Press, 1946.
17. David, H. and Nagaraja, H. Order statistics. New Jersey: Wiley, 2003.
18. Fenton, A., Lansky, P. and Olypher, A. Properties of the extra-

positional signal in hippocampal place cell discharge derived from the overdispersion in location-specific firing. *Neuroscience* 111: 553-566, 2002.
19. Gerstein, G. and Mandelbrot, B. Random walk models for the spike activity of a single neuron. *Biophys. J.* 4: 41-68, 1964.
20. Gerstner, W., Kempter, R., van Hemmen, J. and Wagner, H. A neuronal learning rule for sub-millisecond temporal coding. *Nature* 383: 76-78, 1996.
21. Gerstner, W. and Kistler, W. Spiking Neuron Models: Single Neurons, Populations, Plasticity, Cambridge: Cambridge University Press, 2002.
22. Ivry, R. Cerebellar timing systems. *Int. Rev. Neurobiol.* 41: 555, 1997.
23. Kass, R.E., Ventura, V. and Brown, E.N. Statistical issues in the analysis of neuronal data. *J. Neurophysiol.* 94: 8-25, 2005.
24. Kendall, M. and Stuart, A. The advanced theory of statistics. Vol. 1: Distribution theory, London: Charles Griffin, 1977.
25. Kostal, L. and Lansky, P. Similarity of interspike interval distributions and information gain in a stationary neuronal firing. *Biol. Cybern.* 94: 157-167, 2006.
26. Kostal, L., Lansky, P. and Rospars, J.P. Review: Neuronal coding and spiking randomness. *Eur. J. Neurosci.* 26: 2693-2701, 2007.
27. Kostal, L., Lansky, P. and Zucca, C. Randomness and variability of the neuronal activity described by the Ornstein-Uhlenbeck model, Network. *Comp. Neur. Sys.* 18: 63-75, 2007.
28. Kostal, L. and Lansky, P. Variability and randomness in stationary neuronal activity. *Biosystems* 89: 44-49, 2007.
29. Lansky, P., Rodriguez, R. and Sacerdote, L. Mean instantaneous firing frequency is always higher than the firing rate. *Neural Comput.* 477-489, 2004.
30. Lansky, P. and Sato, S. The stochastic diffusion models of nerve membrane depolarization and interspike interval generation. *J. Peripher. Nerv. Syst.* 4: 27-42, 1999.
31. Levine, M. The distribution of the intervals between neural impulses in the maintained discharges of retinal ganglion cells. *Biol. Cybern.* 65: 459-467, 1991.
32. Marsalek, P., Koch, C. and Maunsell, J. On the relationship between synaptic input and spike output jitter in individual neurons. *Proc. Natl. Acad. Sci. U.S.A.* 94: 735-740, 1997.
33. Matell, M. and Meck, W. Neuropsychological mechanisms of interval timing behavior. *BioEssays* 22: 94-103, 2000.
34. McKeegan, D. Spontaneous and odour evoked activity in single avian olfactory bulb neurones. *Brain Res.* 929: 48-58, 2002.
35. Moore, G., Perkel, D. and Segundo, J. Statistical analysis and functional interpretation of neuronal spike data. *Annu. Rev. Physiol.* 28: 493-522, 1966.
36. Perkel, D. and Bullock, T. Neural coding. *Neurosci. Res. Prog. Sum.* 3: 405-527, 1968.
37. Pouzat, C. and Chaffiol, A. Automatic spike train analysis and report generation. An implementation with R, R2HTML and STAR. *J. Neurosci. Methods* 181: 119-144, 2009.
38. Ratliff, F., Hartline, H. 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.
39. Reed, M., Blum, J. and Mitchell, C. Precision of neural timing: Effects of convergence and time-windowing. *J. Comp. Neurosci.* 13: 35-47, 2002.
40. Reeke, G. and Coop, A. Estimating the temporal interval entropy of neuronal discharge. *Neural Comput.* 16: 941-970, 2004.
41. Rieke, F., Steveninck, R., Warland, D. and Bialek, W. Spikes: Exploring the Neural Code. Cambridge: MIT Press, 1997.
42. Shadlen, M. and Newsome, W. The variable discharge of cortical neurons: Implications for connectivity, computation, and information coding. *J. Neurosci.* 18: 3870-3896, 1998.
43. Softky, W. and Koch, C. The highly irregular firing of cortical cells is inconsistent with temporal integration of random Epsps. *J. Neurosci.* 13: 334-350, 1993.
44. Stein, R., Gossen, E. and Jones, K. Neuronal variability: noise or part of the signal? *Nat. Rev. Neurosci.* 6: 389-397, 2005.
45. Theunissen, F. and Miller, J. Temporal encoding in nervous systems: A rigorous definition. *J. Comp. Neurosci.* 2: 149-162, 1995.
46. Tuckwell, H. Introduction to Theoretical Neurobiology, Volume 2, New York: Cambridge University Press, 1988.