

# Variability Measures of Positive Random Variables

Lubomir Kostal\*, Petr Lansky, Ondrej Pokora

Department of Computational Neuroscience, Institute of Physiology, Academy of Sciences of the Czech Republic, Prague, Czech Republic

## Abstract

During the stationary part of neuronal spiking response, the stimulus can be encoded in the firing rate, but also in the statistical structure of the interspike intervals. We propose and discuss two information-based measures of statistical dispersion of the interspike interval distribution, the entropy-based dispersion and Fisher information-based dispersion. The measures are compared with the frequently used concept of standard deviation. It is shown, that standard deviation is not well suited to quantify some aspects of dispersion that are often expected intuitively, such as the degree of randomness. The proposed dispersion measures are not entirely independent, although each describes the interspike intervals from a different point of view. The new methods are applied to common models of neuronal firing and to both simulated and experimental data.

**Citation:** Kostal L, Lansky P, Pokora O (2011) Variability Measures of Positive Random Variables. PLoS ONE 6(7): e21998. doi:10.1371/journal.pone.0021998

**Editor:** Zheng Su, Genentech Inc., United States of America

**Received:** April 26, 2011; **Accepted:** June 9, 2011; **Published:** July 22, 2011

**Copyright:** © 2011 Kostal et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Funding:** This work was supported by AV0Z50110509, Centre for Neuroscience LC554, and the Grant Agency of the Czech Republic project P103/11/0282. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

**Competing Interests:** The authors have declared that no competing interests exist.

\* E-mail: kostal@biomed.cas.cz

## Introduction

One of the most fundamental problems in computational biology is the problem of neuronal coding, the question of how information is represented in neuronal signals [1,2]. The discharge activity of neurons is composed of series of events called action potentials (or spikes). It is widely accepted, that information in neuronal systems is transferred by employing these spikes. The shapes and durations of individual spikes are very similar, therefore it is generally assumed that the form of the action potential is not important in information transmission. When a stimulus is presented, the responding neuron usually produces a transient response followed by a sustained one, which is often treated as stationary in time [3]. The firing rate of the sustained part of the response depends on the stimulus, however, the stimulus can be also “encoded” in the statistical structure of the interspike intervals (ISI) by the *temporal coding* [1,4–6].

While the description of neuronal activity from the rate coding point of view is relatively straightforward [7], the temporal code allows infinite number of alternatives. Spike trains with equal firing rates may turn out to be different under various measures of their statistical structure beyond the firing rate. For example, even more than a half century ago, coefficient of variation ( $c_v$ ) of ISIs was reported to encode information about light intensity in adapted cells of the horseshoe crab [1,8]. Similarly, changes in the level of bursting activity, also characterized by  $c_v$ , are reported to be the proper code for edge detection in certain units of visual cortex [9]. In general, the bursting nature of neuronal firing is commonly described by  $c_v$  [10].

In order to describe and analyze the way information is represented in spike trains, methods for their mutual comparison are needed. Although the ISI probability density function (or histogram of data) usually provides a complete information, one needs quantitative methods [11–13], especially since a visual inspection of the density shape can be misleading. Here we restrict

our attention to the measures of the neuronal firing precision, e.g., of the the ISI distribution dispersion. We investigate the properties of the standard deviation, the entropy-based dispersion and the Fisher information-based dispersion. Although standard deviation is used ubiquitously and is almost synonymous to the “measure of statistical dispersion”, we show, that it is not well suited to quantify some aspects of spiking activity that are often expected intuitively [4,14]. We will show, that the diversity or randomness of ISIs is better described by entropy-based or Fisher information-based dispersions. The difference between entropy and Fisher information descriptions lies in the fact that the Fisher information describes how “smooth” is the distribution, while the entropy describes how “even” it is. The “smoothness” and “evenness” might be at first thought interchangeable, but we show that it is not the case.

The illustration of the proposed methods is provided on simple and frequently employed models of stationary neuronal activity, given by lognormal, gamma and inverse Gaussian distributions of ISIs. Finally, we apply the theory on experimental data obtained by recording the spontaneous activity of rat olfactory neurons [15].

## Methods

Statistical methods and methods of probability theory and stochastic point processes are widely applied to describe and to analyze neuronal firing [16–18]. The probabilistic description of spiking times results from the fact, that the positions of spikes cannot be predicted exactly, only the probability that the spike occurs is given [18]. Thus, under suitable conditions, the ISI or time-to-first spike after the stimulus onset can be described by a continuous positive random variable. We denote this random variable as  $T$ . Complete description of  $T$  is given by its probability density function  $f(t)$ , defined on  $[0, \infty)$ .

The (statistical) dispersion is a characteristics of “variability” or “spread” of the distribution of the random variable  $T$ . There are

different dispersion measures described in the literature and employed in different contexts, e.g., standard deviation, interquartile range [19], mean difference [20] or the coefficient of local variance [13]. The measures have the same physical units as  $T$ .

### Standard deviation

By far, the most common measure of dispersion is the standard deviation,  $\sigma$ , defined as the square root of the second central moment of the distribution. The corresponding *relative* dispersion measure is known as the coefficient of variation,  $c_v$ ,

$$c_v = \frac{\sigma}{E(T)}, \tag{1}$$

where  $E(T)$  is the mean value of  $T$ . Exponential distribution implies  $c_v = 1$ , however, this values of  $c_v$  may occur for other distributions as well.

### Entropy based dispersion

The randomness of a probability distribution can be defined as the measure of “choice” of possible outcomes. Bigger choice results, intuitively, in greater randomness. For discrete probability distributions such measure of randomness is provided by the Shannon entropy, which is known to be a unique, consistent with certain natural requirements [21]. The Shannon entropy is generally infinite for continuous variables, and therefore it cannot be used for our purposes. Formally, the notion of *differential entropy*,  $h(f)$ , of probability density function  $f(t)$ , is introduced as

$$h(f) = - \int_T f(t) \ln f(t) dt, \tag{2}$$

however, the value  $h(f)$  can be positive or negative and cannot be by itself used as a measure of randomness [22].

In order to obtain a properly behaving quantity, the entropy-based dispersion,  $\sigma_h$ , was proposed in [23],

$$\sigma_h = \exp[h(f) - 1]. \tag{3}$$

The interpretation of  $\sigma_h$  relies on the asymptotic equipartition property theorem and the entropy power concept [22]. Namely, since for the exponential probability density function  $f_{\text{exp}}(t) = 1/E(T) \exp[-t/E(T)]$  holds  $h(f_{\text{exp}}) = 1 + E(T)$  we see, that  $\sigma_h$  is the standard deviation of such exponential distribution, which satisfies  $h(f_{\text{exp}}) = h(f)$ . Informally, the value of  $\sigma_h$  is bigger for those random variables, which generate more diverse (or unpredictable) realizations.

Analogously to Eq. (1), we define the relative entropy-based dispersion coefficient,  $c_h$ , as

$$c_h = \frac{\sigma_h}{E(T)}. \tag{4}$$

Note, that Eq. (4) can be equivalently written as

$$c_h = \exp[-D_{\text{KL}}[f(t)||f_{\text{exp}}(t)]], \tag{5}$$

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

$$D_{\text{KL}}[f(t)||f_{\text{exp}}(t)] = \int_T f(t) \ln \frac{f(t)}{f_{\text{exp}}(t)} dt \tag{6}$$

is the Kullback-Leibler distance of the probability density  $f(t)$  from the exponential density with the same mean as  $f(t)$ . From Eq. (5) follows that  $c_h$  is essentially (up to the scale) equivalent to the measure of spiking randomness,  $\eta$ , proposed in [4], since  $c_h = e^{\eta-1}$ .

From the properties of the Kullback-Leibler distance in Eq. (5) follows, that the maximum value of  $c_h$  is  $c_h = 1$ , which occurs if and only if  $f(t)$  is exponential [22].

### Fisher information based dispersion

The Fisher information is a measure of the minimum error in estimating a parameter of a distribution. In a special case of the location parameter, the Fisher information  $J(f)$  does not depend on the parameter itself, and can be expressed directly as a functional of the density  $f(t)$  ([22], p.671),

$$J(f) = \int_0^\infty \left[ \frac{\partial \ln f(t)}{\partial t} \right]^2 f(t) dt. \tag{7}$$

We illustrate that the value of  $J(f)$  is small for smoothly-shaped probability densities. Any locally steep slope or the presence of modes in the shape of  $f(t)$  increases  $J(f)$  [24]. Due to the derivative in Eq. (7), certain regularity conditions are required on  $f(t)$ . In this paper we consider only the densities for which  $J(f)$  takes finite values. Further theoretical considerations are however beyond the scope of this paper.

The units of  $J(f)$  correspond to the inverse of the squared units of  $T$ , therefore we propose the Fisher information-based dispersion measure,  $\sigma_J$ , as

$$\sigma_J = \frac{1}{\sqrt{J(f)}}. \tag{8}$$

In analogy with Eqns. (1) and (4) we define the relative dispersion coefficient  $c_J$  as

$$c_J = \frac{\sigma_J}{E(T)}. \tag{9}$$

For exponential distribution holds  $c_J = 1$ , however, this value is not specific only for the case  $f(t) = f_{\text{exp}}(t)$ .

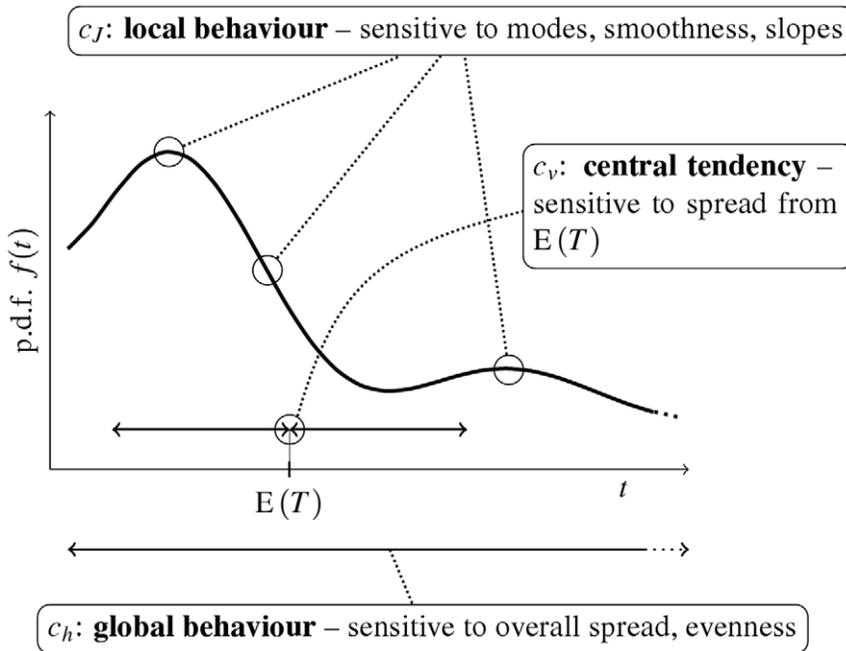
Just as  $c_h$  is related to the Kullback-Leibler distance by Eq. (5), we note that  $J(f)$  can be written as [25]

$$J(f) = \frac{\partial^2 D_{\text{KL}}[f(t-\Delta)||f(t)]}{\partial \Delta^2} \Big|_{\Delta=0}. \tag{10}$$

Although Eq. (10) is not suitable for evaluation of  $J(f)$ , it shows, that both  $c_h$  and  $c_J$  are connected on the fundamental level by the concept of the Kullback-Leibler distance.

### Basic properties of the proposed measures

Standard deviation (or  $c_v$ ) measures essentially how off-centered, with respect to  $E(T)$ , is the probability density of  $T$  and it is sensitive to outlying values. On the other hand,  $c_v$  does not quantify how random, or unpredictable, are the outcomes of  $T$ . Namely, high value of  $c_v$  (high variability) does not indicate that the possible values of  $T$  are distributed evenly [4]. On the other hand, the value of  $\sigma_h$  (and  $c_h$ ) quantifies how evenly is the probability distributed over  $[0, \infty)$ . The third measure,  $c_J$ , is sensitive to the modes and steepness of slopes of the density (due to



**Figure 1. Illustration of the main features of the studied measures.** Schematic example of a probability density function  $f(t)$  is shown. Although the evenness of the density (described by  $c_h$ ) and its smoothness (described by  $c_J$ ) are related, the sensitivity of  $c_J$  to modes and slopes enables it to differentiate shapes with otherwise equal  $c_v$  and  $c_h$ . doi:10.1371/journal.pone.0021998.g001

the dependence on the derivative of the probability density in Eq. (7)). Since multimodal densities can be more evenly spread than unimodal ones, the behavior of  $c_h$  cannot be generally deduced from  $c_J$  (and vice versa). The key features of the three considered dispersion measures are illustrated in Fig. 1.

A cartoon with typical density shapes resulting from a combination of  $c_v$ ,  $c_h$  and  $c_J$  values range is shown in Fig. 2. Very small value of  $c_v$  inevitably results in a density shapes concentrated around  $E(T)$ , and correspondingly small values of  $c_h$  and  $c_J$ . The intermediate,  $c_v \approx 1$ , and upper range of  $c_v$  offer more variable density shapes, where  $c_v$  and  $c_h$  are not sufficient for their classification and  $c_J$  can be employed for further description.

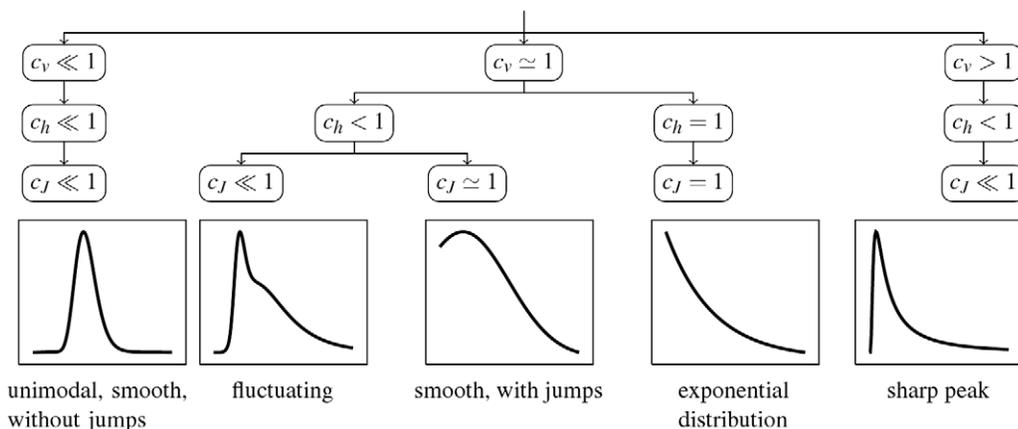
Note, that the number of possible scenarios is large and therefore Fig. 2 is not exhaustive.

**Results**

**Common distributions of interspike intervals**

We choose three widely employed statistical models of ISIs: gamma, inverse Gaussian and lognormal distributions, and analyze them by means of the three described dispersion coefficients  $c_v$ ,  $c_h$  and  $c_J$ .

Gamma distribution is one of the most frequent statistical descriptors of ISIs employed in analysis of experimental data



**Figure 2. Illustration of a classification tree of probability densities based on typical values of the dispersion measures.** Note, that not all combinations of values of  $c_v, c_h, c_J$  can appear. Selected identification signs or examples of corresponding distributions, which are typical but not necessarily comprehensive, are written below the corresponding illustrative plots of densities. doi:10.1371/journal.pone.0021998.g002

[15,26,27]. Its probability density function parametrized by shape parameter  $k$  and scale parameter  $\theta$  is

$$f(t) = \frac{t^{k-1} \exp\{-t/\theta\}}{\Gamma(k) \theta^k}, \tag{11}$$

where  $\Gamma(z)$  is the gamma function [28]. The mean value of the distribution is  $E(T) = k\theta$  and the coefficient of variation is equal to

$$c_v = 1/\sqrt{k}. \tag{12}$$

For  $c_v = 1$ , i.e.  $k = 1$ , the gamma distribution becomes exponential distribution. By parametrizing the density (11) by  $c_v$  and substituting it into Eqns. (4) and (9) we obtain the entropy-based and Fisher information-based dispersion coefficients as functions of  $c_v$ ,

$$c_h = c_v^2 \Gamma(c_v^{-2}) \exp\left\{\frac{1+(c_v^2-1)\Psi(c_v^{-2})}{c_v^2} - 1\right\}, \tag{13}$$

$$c_J = c_v \sqrt{1-2c_v^2} \quad \text{for } 0 < c_v < \frac{1}{\sqrt{2}}, \tag{14}$$

where  $\Psi(z) = \frac{d}{dz} \ln \Gamma(z)$  is the digamma function [28]. For details of the calculation of  $c_h$  and  $c_J$  see Supporting Information S1. Note, that the gamma density is not differentiable at  $t=0$  for  $c_v \geq 1/\sqrt{2}$ , thus  $c_J$  is evaluated only for  $0 < c_v < 1/\sqrt{2}$ .

The inverse Gaussian distribution is often used to describe neural activity and fitted to experimentally observed ISIs [26,29,30]. This distribution describes the spiking activity of a stochastic variant of the perfect integrate-and-fire neuronal model [18,31]. The probability density function of the inverse Gaussian distribution parametrized by its mean,  $\mu = E(T)$ , and scale parameter  $\sigma$  is

$$f(t) = \frac{1}{\sqrt{2\pi\sigma^2 t^3}} \exp\left\{-\frac{(t-\mu)^2}{2\sigma^2 \mu^2 t}\right\}. \tag{15}$$

The coefficient of variation is equal to

$$c_v = \sqrt{\mu\sigma^2} \tag{16}$$

and the other dispersion coefficients can be expressed as (see Supporting Information S1)

$$c_h = \sqrt{\frac{2\pi}{e}} c_v \exp\left\{-\frac{3 \exp(c_v^{-2}) K^{(1,0)}(-\frac{1}{2}, c_v^{-2})}{\sqrt{2\pi} c_v}\right\}, \tag{17}$$

$$c_J = \frac{\sqrt{2} c_v}{\sqrt{2+9c_v^2+21c_v^4+21c_v^6}}, \tag{18}$$

where  $K^{(1,0)}(v,z)$  is the derivative of the modified Bessel function of the second kind,  $K^{(1,0)}(v,z) = \frac{\partial}{\partial v} K(v,z)$  [28].

The lognormal distribution of ISIs, with some exceptions [32], is rarely presented as a result of a neuronal model. However, it represents a common descriptor in experimental data analysis [26,30]. The lognormal probability density function parametrized

by the mean,  $\mu$ , and standard deviation,  $\sigma$ , of variable  $\ln T$  is

$$f(t) = \frac{1}{\sqrt{2\pi\sigma^2} t} \exp\left\{-\frac{(\ln t - \mu)^2}{2\sigma^2}\right\}. \tag{19}$$

In this parametrization, the mean of the lognormal distribution is  $E(T) = \exp(\mu + \sigma^2/2)$  and the coefficient of variation is equal to

$$c_v = \sqrt{\exp(\sigma^2) - 1}. \tag{20}$$

The two other dispersion coefficients, expressed as functions of  $c_v$ , are (see Supporting Information S1)

$$c_h = \sqrt{\frac{2\pi}{e}} \sqrt{\frac{\ln(1+c_v^2)}{1+c_v^2}}, \tag{21}$$

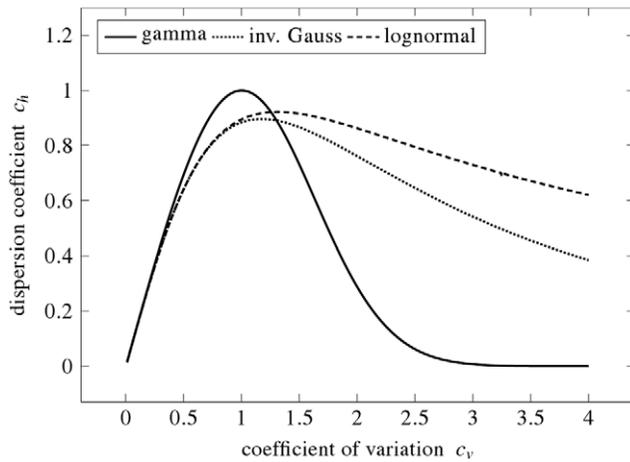
$$c_J = \sqrt{\frac{\ln(1+c_v^2)}{[1+c_v^2]^3 [1+\ln(1+c_v^2)]}}. \tag{22}$$

The dependence of  $c_h$  on  $c_v$  is shown in Fig. 3, the dependence of  $c_J$  on  $c_v$  is shown in Fig. 4, for all the three mentioned distributions. Obviously, the dependencies are not linear (even not monotonous) and thus neither  $c_h$  nor  $c_J$  is equivalent to  $c_v$ . Maxima of  $c_h$  and  $c_J$  occur for different  $c_v$  values, confirming that each of the proposed dispersion coefficients provides a different point of view. We see, that both  $c_h$  and  $c_J$  as functions of  $c_v$  show a “ $\cap$ ” shape with maxima around  $c_v \doteq 1$  (for  $c_h$ ) and around  $c_v \doteq 0.5$  (for  $c_J$ ). There is a reason why the maxima of  $c_h$  and  $c_J$  tend to occur at these values of  $c_v$ . It can be shown by the methods of variational calculus, that there exists a unique distribution maximizing  $c_h$ : the exponential distribution for which  $c_v = 1$ . Since some densities tend to resemble the exponential density if their  $c_v$  is close to 1, their maxima of  $c_h$  occur near this  $c_v$  value. Similarly, there exists a unique density maximizing  $c_J$ ; it is given in terms of the Airy functions with  $c_v = 0.44$ . Analogously, densities with  $c_v \approx 0.5$  may resemble this distribution and thus attain the maximum of  $c_J$  there. However, there exist distributions which does not attain the maximum of  $c_h$  around  $c_v = 1$  or the maximum of  $c_J$  around  $c_v = 0.5$ . Detailed mathematical treatment of the  $c_h$ - and  $c_J$ -maximizing distributions is beyond the scope of the manuscript and will be published elsewhere.

Note, that the plots of  $c_J$  against  $c_v$  appear like a scaled version of the plots of  $c_h$  against  $c_v$ , with the relative positions of the curves for each distribution preserved (to certain extent). In particular, while  $c_h$  of the lognormal is always greater than  $c_h$  of the inverse Gaussian, the ordering is reversed for the  $c_J$  for  $c_v > 2.2$ .

The dependence of  $c_J$  on  $c_h$  is plotted in Fig. 5. We observe, that  $c_h$  and  $c_J$  indeed do not describe the same qualities of the distribution, since a unique  $c_h$  value does not correspond to a unique  $c_J$  value (and vice versa). Except for the gamma distribution, the dependence between  $c_h$  and  $c_J$  forms a closed loop, where  $c_h = c_J = 0$  for both  $c_v \rightarrow 0$  and  $c_v \rightarrow \infty$ .

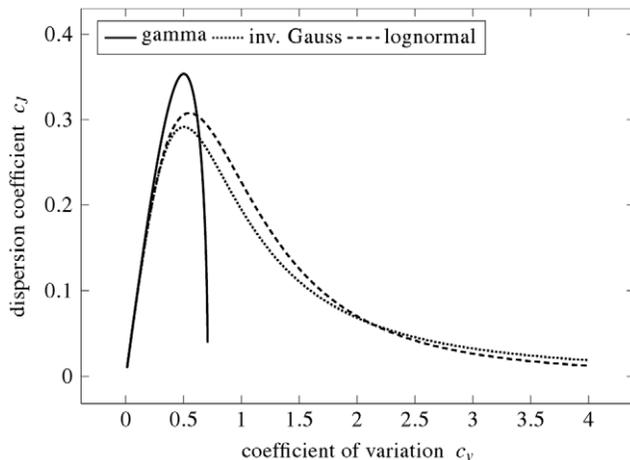
Additionally, just as  $c_h$  and  $c_J$  are related to  $c_v$  in Eqns. (13), (14), (17), (18), (21) and (22),  $c_h$  and  $c_J$  can also be related to higher statistical moments. For example, the skewness  $\gamma$  of the distribution is defined as the ratio of the third central moment and the third power of standard deviation. For gamma distribution holds  $\gamma = 2c_v$ , for inverse Gaussian  $\gamma = 3c_v$  and for the lognormal



**Figure 3. Entropy-based dispersion coefficient,  $c_h$ , in dependence on the coefficient of variation,  $c_v$ .** Three interspike interval models: gamma, inverse Gaussian and lognormal distribution, are employed. Both  $c_v$  and  $c_h$  describe “spread” of the interspike intervals, but from different points of view. Coefficient of variation,  $c_v$ , quantifies how off-centered is the mass of the probability density function, whereas  $c_h$  indicates how evenly is the mass distributed over all possible values. For all the shown distributions holds  $c_h = 0$  as  $c_v \rightarrow 0$  or  $c_v \rightarrow \infty$ . doi:10.1371/journal.pone.0021998.g003

$\gamma = c_v^3 + 3c_v$ . Thus the curves depicted in Fig. 3 and Fig. 4 would retain their unimodal shapes if plotted in dependence on  $\gamma$ .

Different distributions with equal  $c_v$  and different  $c_h$  (or  $c_J$ ) can be found, and vice versa; see Fig. 3 (or Fig. 4) for examples. Therefore, it cannot be said in general that  $c_h$ ,  $c_J$  are more informative than  $c_v$ . To provide an example in which  $c_J$  provides a different view over  $c_v$  and  $c_h$ , we consider the folded normal



**Figure 4. Fisher information-based dispersion coefficient,  $c_J$ , as a function of the coefficient of variation,  $c_v$ , for the same distributions as in Fig. 3.** The coefficient  $c_J$  grows as the average of squared derivative of the probability density function (see Eq. (7)) becomes smaller, that means as the distribution of the interspike intervals becomes more smooth. This confirms that “smoothness” and “evenness” of the distribution (compare with Fig. 3) are different notions, although there are qualitative similarities:  $c_J = 0$  for  $c_v \rightarrow 0$  for all shown distributions, and  $c_J = 0$  as  $c_v \rightarrow \infty$  for both lognormal and inverse Gaussian distributions. Note, that dispersion coefficient  $c_J$  for the gamma distribution can be calculated only for  $c_v < 1/\sqrt{2} \doteq 0.707$ . doi:10.1371/journal.pone.0021998.g004

probability density with parameters  $\alpha, \beta > 0$

$$f(t) = \sqrt{\frac{2}{\beta^2 \pi}} \left[ 1 + \operatorname{erf} \left( \frac{\alpha}{\sqrt{2}\beta} \right) \right]^{-1} \exp \left[ -\frac{(t-\alpha)^2}{2\beta^2} \right]. \quad (23)$$

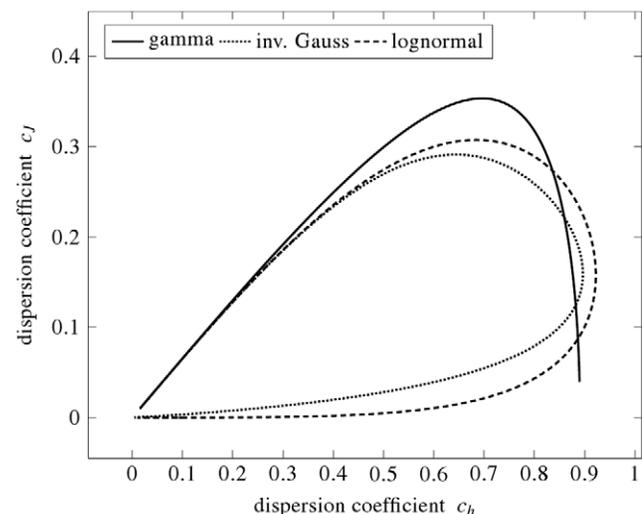
The shapes of the folded normal probability density function, Eq. (23), and gamma probability density function, Eq. (11), are compared in Fig. 6 for  $c_v = 0.69$ . Although their values of  $c_h$  are very similar, the values of  $c_J$  are very different. The reason lies mainly in the initial steep rise of the gamma density from zero.

**Simulated data**

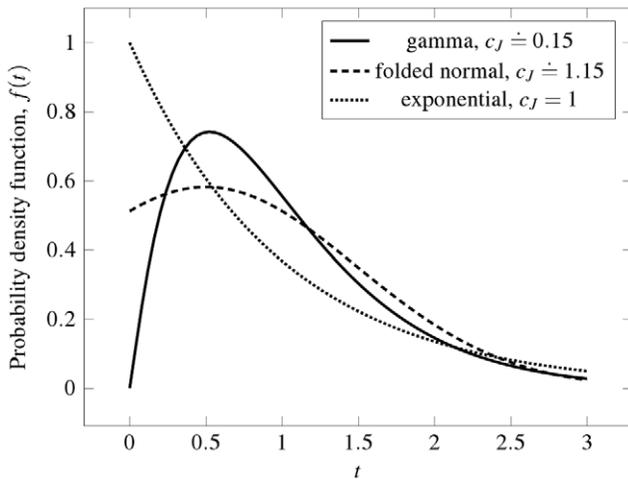
To illustrate the accuracy of the estimators  $\hat{c}_h$  and  $\hat{c}_J$  of dispersion coefficients  $c_h$  and  $c_J$ , we simulated spike trains with gamma, inverse Gaussian and lognormal distributions of ISIs by employing the R and STAR software packages [26,33]. In all the simulations the mean ISI was fixed to 1, while the coefficient of variation,  $c_v$ , varied from 0.05 to 4.00 in steps of 0.05. In other words, we generated random samples from the mentioned distributions with given parameters. The spike trains represented by sample point processes were constructed by using the generated values as the time intervals (ISIs) between successive events (spikes). Five thousand spike trains, each consisting of 100 ISIs, were simulated for each of the values of  $c_v$  and for each of the three distributions.

In the first study, the parameters of the distributions were estimated by the maximum likelihood method. For the gamma distribution (11) the maximum likelihood estimators  $\hat{k}$  and  $\hat{\theta}$  were found numerically (by minimizing the loglikelihood function). For the inverse Gaussian distribution (15) the maximum likelihood estimators were computed as

$$\hat{\mu} = \frac{1}{n} \sum_{i=1}^n t_i, \quad (24)$$



**Figure 5. The dispersion coefficients  $c_h$  and  $c_J$  for the same distributions as in Figs. 3 and 4.** The plot of dependencies between the two dispersion coefficients form closed curve for both inverse Gaussian and lognormal distribution. Starting from the origin and moving clockwise, the points on the loop correspond to the values of  $c_v$  growing from 0 to infinity. For gamma distribution,  $c_J$  is a common unimodal function of  $c_h$ . doi:10.1371/journal.pone.0021998.g005



**Figure 6. Comparison of probability density functions with  $E(T)=1$ , identical  $c_v$  and  $c_h$  but different  $c_J$ .** For the gamma distribution holds  $c_v=0.69$ ,  $c_h \doteq 0.88$  and  $c_J \doteq 0.15$ . For the folded normal distribution holds  $c_v=0.69$ ,  $c_h \doteq 0.91$  and  $c_J \doteq 1.15$ . The difference between these two distributions (from the  $c_J$  point of view) lies in the initial slope of the gamma probability density. For comparison, the exponential density ( $c_v=c_h=c_J=1$ ) is shown. doi:10.1371/journal.pone.0021998.g006

$$\hat{\sigma}^2 = \frac{1}{n} \sum_{i=1}^n \frac{1}{t_i} - \frac{1}{\hat{\mu}} \tag{25}$$

Similarly, for the lognormal distribution (19) of ISIs, maximum likelihood estimators of the parameters are

$$\hat{\mu} = \frac{1}{n} \sum_{i=1}^n \ln t_i, \tag{26}$$

$$\hat{\sigma}^2 = \frac{1}{n} \sum_{i=1}^n (\ln t_i - \hat{\mu})^2. \tag{27}$$

The values of coefficient of variation,  $c_v$ , were calculated by substitution of the maximum likelihood estimates into Eqns. (12), (16) or (20). Consequently, the other two dispersion coefficients,  $\hat{c}_h$  and  $\hat{c}_J$ , were computed by substitution of the estimated  $c_v$  into Eqns. (13) and (14) for the gamma distribution, into Eqns. (17) and (18) for the inverse Gaussian and into Eqns. (21) and (22) for the lognormal distribution.

In the second study, the coefficient of variation was estimated by commonly used moment method as the ratio of the sample standard deviation,  $\hat{\sigma}$ , and the sample mean,  $\bar{T}$ ,

$$\hat{c}_v = \frac{\hat{\sigma}}{\bar{T}}, \tag{28}$$

for all the mentioned distributions. Both the entropy-based and Fisher information-based dispersion coefficients were then calculated by substitution of estimate (28) into the same equations for  $\hat{c}_h$  and  $\hat{c}_J$  as with maximum likelihood estimates, in accordance to the respective ISI distribution.

The accuracy of the estimates  $\hat{c}_h$  and  $\hat{c}_J$  was studied for both the types of estimators. The results are depicted in Fig. 7 for the

maximum likelihood estimates, and in Fig. 8 for the moment estimates. In both figures two measures of the accuracy of the estimators  $\hat{c}_h$  and  $\hat{c}_J$  are plotted against the true values of  $c_v$  (those used for simulation). The first is the bias of the estimate,  $b(c_h)$ , defined as

$$b(c_h) = \frac{1}{n} \sum_{i=1}^n (\hat{c}_{h,i} - c_h), \tag{29}$$

where  $n=5000$  is the number of simulated spike trains and  $\hat{c}_{h,i}$  is the value estimated from the  $i$ -th spike train. Analogous equation is used for evaluation of  $b(c_J)$ . The latter measure is the relative standard error,  $\varepsilon(c_h)$ , expressed as the ratio of the standard deviation and the mean value of the estimate,

$$\varepsilon(c_h) = \frac{1}{b(c_h) + c_h} \sqrt{\frac{1}{n-1} \sum_{i=1}^n \{ \hat{c}_{h,i} - [b(c_h) + c_h] \}^2}, \tag{30}$$

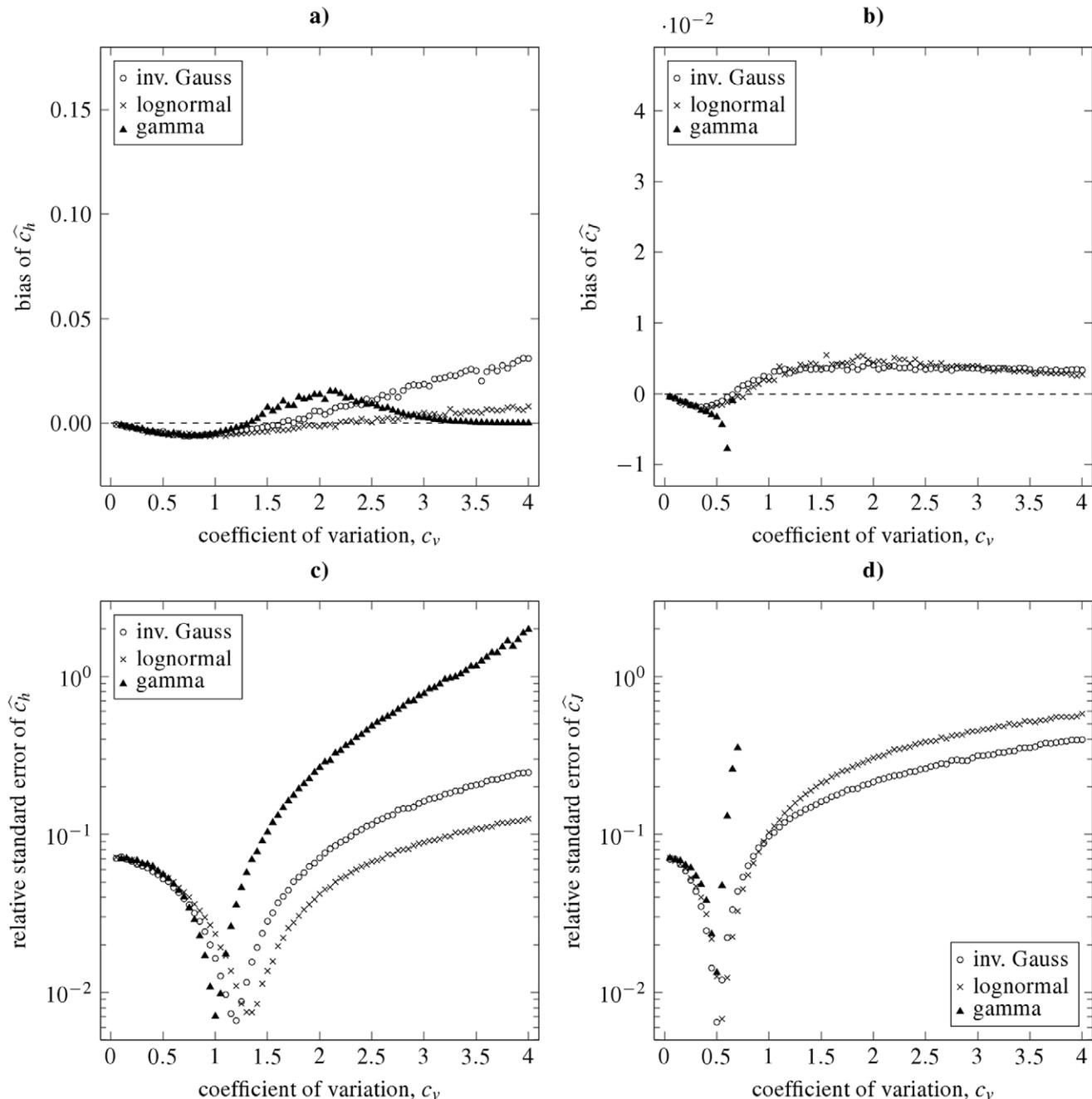
and analogously for  $\varepsilon(c_J)$ . This characteristics says how accurate the values of the estimator are when calculated from random sample of given  $c_v$ . The relative standard deviation with respect to the mean value is dimensionless and therefore it is suitable for comparisons of the quality of different estimators of  $c_h$  and  $c_J$ .

We observe qualitative similarities in the dependencies of both the bias and the relative standard error of the estimators  $\hat{c}_h$  and  $\hat{c}_J$ . In general, we see that the estimators are biased (see panels **a**, **b** in Figs. 7 and 8), but the values of bias of the moment estimators are approximately 10 times greater than the bias of the maximum likelihood estimators. For small values of  $c_v$  the dispersion coefficients are underestimated and the bias becomes positive as  $c_v$  grows. For gamma distribution, the bias of  $\hat{c}_h$  starts to decrease to zero after it attains its maxima for  $c_v \doteq 2$ , thus  $\hat{c}_h$  seems to be asymptotically unbiased estimator. On contrary, bias of  $\hat{c}_h$  for inverse Gaussian and lognormal distribution grows as  $c_v$  grows. There is also a difference between the maximum likelihood and moment estimator  $\hat{c}_h$ : in the maximum likelihood case the bias of  $\hat{c}_h$  for inverse Gaussian distribution is greater than for the lognormal, the difference seems to be negligible in the case of the moment estimator.

The bias of  $\hat{c}_J$  looks similar to the bias of  $\hat{c}_h$  for small  $c_v$ . But, in contrary to  $\hat{c}_h$ , the bias of  $\hat{c}_J$  starts to decrease slowly for large values of the coefficient of variation ( $c_v > 2$ ). This fact can be seen for both the inverse Gaussian and lognormal distribution. In the maximum likelihood case the bias of  $\hat{c}_J$  is almost the same for both these distributions. The bias of  $\hat{c}_J$  is greater for the lognormal than for inverse Gaussian distribution.

Focusing on the accuracy of the estimators (see panels **c**, **d** in Figs. 7 and 8), the shapes of the relative standard deviations of  $\hat{c}_h$  and  $\hat{c}_J$  are very similar, regardless of the ISI distribution and the method used for estimation. The relative standard deviations of  $\hat{c}_J$  look like scaled versions of analogous characteristics of  $\hat{c}_h$ . For  $c_v \rightarrow 0$  they starts at a value less than 0.1.

As  $c_v$  grows from zero, the relative standard deviation of the estimators decrease and attains its minima at around  $c_v \doteq 1.2$  (for  $\hat{c}_h$ ) and  $c_v \doteq 0.5$  (for  $\hat{c}_J$ ), respectively. It should be noted that these minima of relative standard deviations of  $\hat{c}_h$  and  $\hat{c}_J$  coincide with the maxima of  $c_h$  and  $c_J$  (compare with Figs. 3 and 4). In other words, the estimates  $\hat{c}_h$  and  $\hat{c}_J$  are most accurate for  $c_v$  values where  $c_h$  and  $c_J$  attain their theoretical maxima; but they are slightly negatively biased. For larger values of  $c_v$  the relative standard deviations of  $\hat{c}_h$  and  $\hat{c}_J$  are increasing functions of  $c_v$ . In addition, the values of relative standard deviations of the



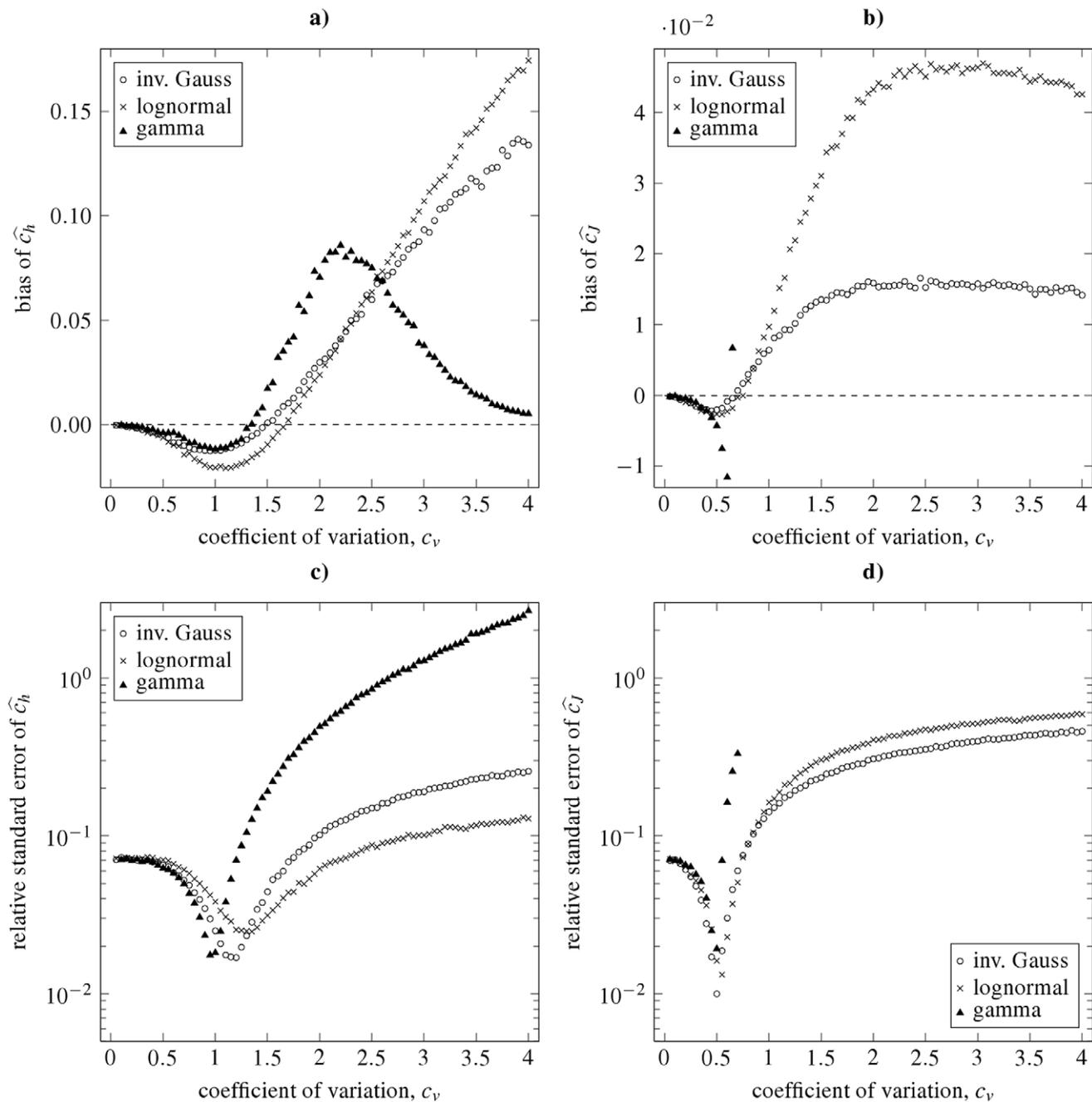
**Figure 7. Dispersion coefficients estimation by using the maximum likelihood method from simulated data.** Bias (panels **a, b**) and relative standard deviations (panels **c, d**) of the dispersion coefficients estimates  $\hat{c}_h$  (panels **a, c**) and  $\hat{c}_J$  (panels **b, d**), in dependence on the true value of the coefficient of variation,  $c_v$ , are shown. The depicted characteristics were estimated from simulated random samples drawn from inverse Gaussian (circles), lognormal (crosses) and gamma distribution (triangles). Coefficient  $\hat{c}_J$  for gamma distribution (panels **b, d**) can be computed for  $c_v < 0.707$  only.  
doi:10.1371/journal.pone.0021998.g007

estimators for large  $c_v$  values are ordered according to the ISI distribution. The order of the estimator accuracy (from high to low) is lognormal, inverse Gaussian and gamma distribution in the case of  $\hat{c}_h$ , and inverse Gaussian, lognormal and gamma distribution in the case of  $\hat{c}_J$ .

### Experimental data

In order to examine variability or irregularity of the ISIs in real neurons using the proposed dispersion coefficients, we apply the

measures on experimental data. The data come from extracellular recordings of olfactory receptor neurons of *freely breathing* and *tracheotomized* rats. Spontaneous, single-unit action potentials were recorded. The single unit nature of the recorded spikes was controlled. The experimental procedures and data analysis were published in [15], where complete details are given. The groups are not distinguishable on the basis of firing frequency only. For our purpose only samples with sufficient number of observations were chosen. Analyzed dataset consists of 6 records of ISIs from



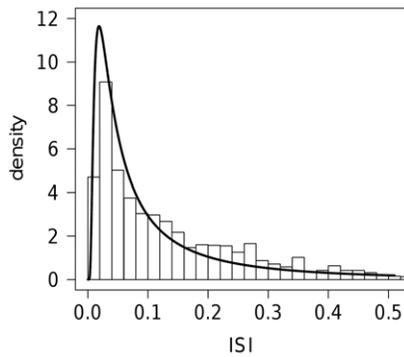
**Figure 8. Dispersion coefficients estimation by using the moment method.** The structure of the panels and the notation are equivalent to those in Fig. 7, except that the estimates  $\hat{c}_v$ , and consequently  $\hat{c}_h$  and  $\hat{c}_J$  were estimated by the moment method.  
doi:10.1371/journal.pone.0021998.g008

freely breathing rats and 11 records from tracheotomized rats. The sample sizes range from 150 to 1500 and all records were tested against nonstationarity.

All samples were fitted with inverse Gaussian distribution (15) as a commonly used distribution of ISI. The histogram of ISIs of typical record and fitted probability density function are depicted in Fig. 9. The mean,  $\mu$ , and the scale parameter  $\sigma$  were estimated by maximum likelihood method. The fit of the data to the inverse Gaussian distribution was checked by Kolmogorov-Smirnov test. The null hypothesis was not rejected on the 5% level in any sample. The dispersion coefficients  $c_v$ ,  $c_h$  and  $c_J$  were calculated

by substitution of the estimated parameters into Eqns. (16), (17) and (18).

The values of estimated dispersion coefficients are summarized and shown as box-and-whisker plots in Fig. 10. Generally, the two categories, *tracheotomized* and *freely breathing*, do not differ significantly in the medians of  $c_v$ ,  $c_h$  or  $c_J$ . Although the ranges of the values overlap in both categories, the values of the criteria seem to be relatively specific with respect to the *freely breathing* category. The difference between mean values are greater than between medians. However, we can observe that the *tracheotomized* category achieves higher values of  $c_v$  and lower values of both  $c_h$  and  $c_J$



**Figure 9. Histogram of interspike intervals from a typical record of the data.** The thick curve shows the shape of probability density function fitted by the maximum likelihood method. Estimated dispersion coefficients are  $c_v \doteq 1.59$ ,  $c_h \doteq 0.85$  and  $c_J \doteq 0.10$ . doi:10.1371/journal.pone.0021998.g009

than the *freely breathing* category. Taking into account the interquartile-range and the range between the whiskers, the Fisher information-based dispersion coefficient,  $c_J$ , seems to be the best of the three examined coefficients to distinguish the two categories for this data. Both groups of rats were compared by employing one-sided variant of the Mann-Whitney test to the three respective dispersion coefficients. However, due to the small sample sizes, no differences between the two groups were confirmed at 95% confidence level.

Moreover, obtained scatterplots of pairs of the dispersion coefficients  $c_h$  and  $c_J$  are shown in Fig. 11. The two categories of rats are best distinguishable in panel **c**), for the *tracheotomized* category having lower values of  $c_h$  together with lower values of  $c_J$  than the latter one. Note also the positions of the points in panel **c**), which confirm that there can be two different  $c_J$  values corresponding to unique  $c_h$  value.

## Discussion

In recent years, information-based measures of signal regularity or randomness have gained significant popularity in various

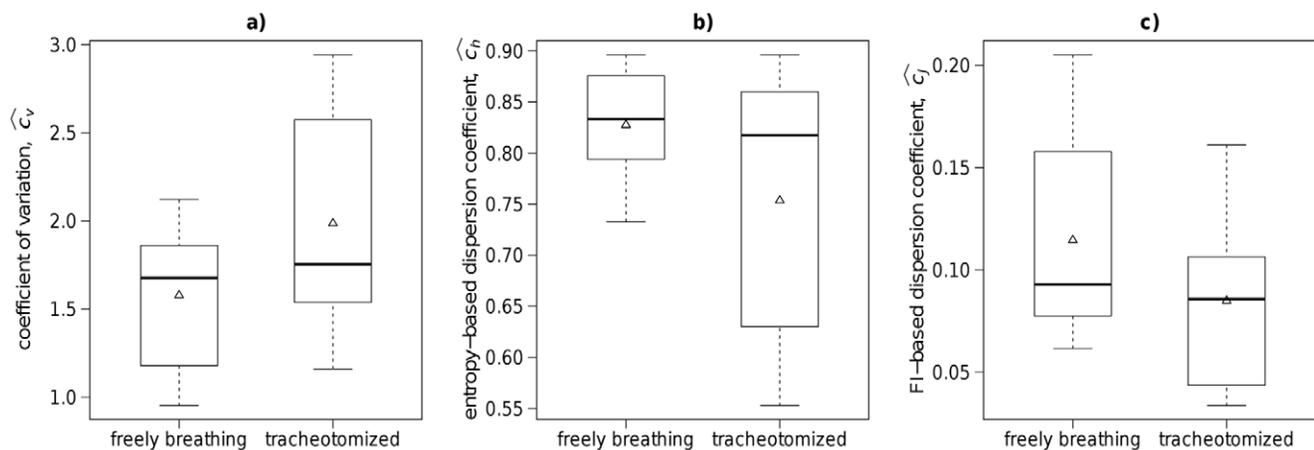
branches of science [24,34–37]. In this paper, we constructed dispersion-like quantities based on these information measures and applied them to the description of neuronal ISI distributions. In particular, we continued the effort initiated in [4,23] by taking into account a variant of Fisher information, which has been employed also in different contexts [24,38–41].

We are motivated by the difference between frequently mixed up notions of ISI variability and randomness, which, however, represent two different concepts. Consider, for example, a spike train consisting of “long” and “short” ISIs with no serial correlations. By adding “medium” length ISIs we do not increase the spiking variability, contrary to what expected intuitively, but decrease it. On the other hand, since the count of ISI of different lengths increases, the spiking randomness is increased. Furthermore, even if conventional analysis of two spike trains reveals no difference, the spike trains may still differ in their randomness and the difference is tractable with relatively limited amount of data [4].

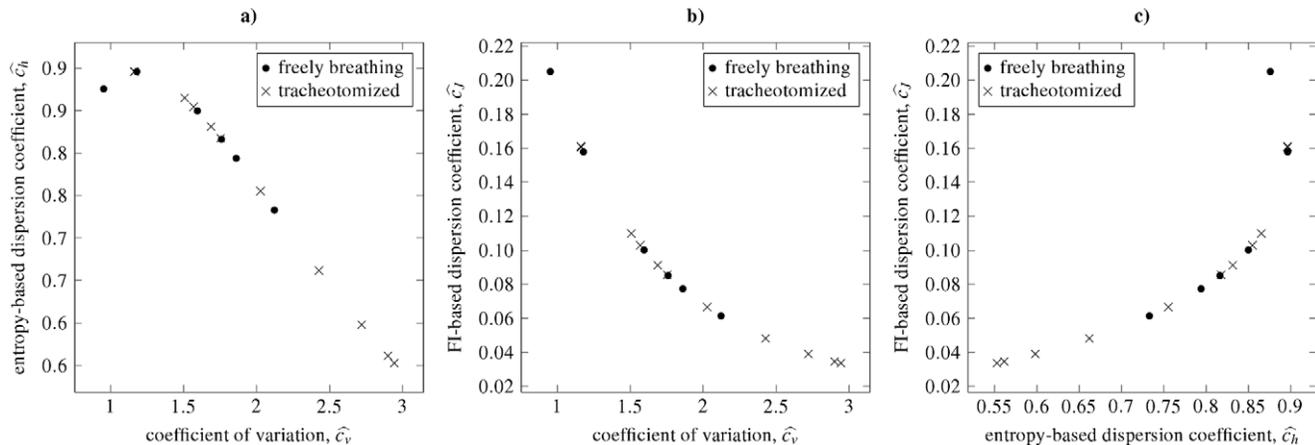
Additionally, by considering the Fisher information-based dispersion coefficient,  $c_J$ , we show that ISI randomness (increasing with diversity of the ISI lengths) and probability density “smoothness” are related, but still different notions. For example, all of the tested distributions are “maximally smooth” for  $c_v \doteq 0.5$  and “maximally even” (maximum ISI randomness) for  $c_v \doteq 1$ .

The statistical properties of the parametric estimations of  $c_v$  and of  $c_h$  and  $c_J$  consequently, are illustrated on simulated data. The results show that the accuracy of the dispersion coefficients depends on the distribution. However, similar property can be found: estimated values of  $c_h$  as well as of  $c_J$  become accurate at the point of maxima of these dispersion coefficients, regardless on the used ISI distribution. It is shown that the ISI distribution as well as the method used for estimation of the parameters from the sample highly influence the bias of the estimators  $\hat{c}_h$  and  $\hat{c}_J$ .

In this paper, we used the parametrical estimates of  $c_v, c_h, c_J$  for both simulated and experimental data analyses. Specific parametric family of distributions was assumed and only the parameters were estimated. On the other hand, it is natural to ask for the non-parametric versions of the estimators. The non-parametric estimate of  $c_v$  is simply calculated by using the first two sample moments. Recently [42], discussed disadvantages of this estimator,



**Figure 10. Box-and-whisker plots of estimated dispersion coefficients.** The coefficients  $c_v$  (panel **a**),  $c_h$  (panel **b**), and  $c_J$  (panel **c**) estimated from the experimental data (spontaneously active rat olfactory neurons) are shown for two categories: freely breathing and tracheotomized rats. The lower and upper sides of the boxes denotes the first ( $Q_1$ ) and third ( $Q_3$ ) quartile, thick lines inside the boxes are medians, triangles denote mean values. The whiskers show the lowest and greatest data value between  $Q_1 - 1.5 \times IQR$  and  $Q_3 + 1.5 \times IQR$  (where  $IQR = Q_3 - Q_1$  is interquartile range). doi:10.1371/journal.pone.0021998.g010



**Figure 11. Dependencies between different dispersion coefficients for the same experimental data as in Fig. 10.** The data was fitted by inverse Gaussian distribution and from maximum likelihood estimators of its parameters the dispersion coefficients were computed. Two categories of data are distinguished: tracheotomized and freely breathing rats. There are small differences in the two categories, as quantified by all the three coefficients. See Figs. 3, 4 and 5 for inverse Gaussian distribution to see the complete curves of the dependencies. doi:10.1371/journal.pone.0021998.g011

stressors out its bias. Non-parametric estimates of the entropy are known [43,44], and we found  $c_h$  can be estimated reliably. As regards the non-parametric estimate of  $c_J$ , approaches based either on spline interpolation of the empirical cumulative distribution function [45] or on specialized kernel-based method for the estimation of the probability density function [46] can be used. Nevertheless, the estimation of the Fisher information-based dispersion coefficient  $c_J$  is a complex task. Preliminary results of our work in progress are promising.

The coefficients were also evaluated from the experimental data, spontaneous action potentials of olfactory receptor neurons in tracheotomized and freely breathing rats. Assuming the inverse Gaussian model, the three estimated dispersion coefficients quantify small differences in the two categories. Taking into account their variability,  $c_J$  seems to be the best measure for distinguishing the categories. Other approach use the pairs of

coefficients  $c_v$ ,  $c_h$  and  $c_J$  to discriminate between the groups. For the analyzed data, the pair of values  $c_h$  and  $c_J$  seems to be the most effective choice.

## Supporting Information

**Supporting Information S1** Detailed calculation of the coefficients  $c_h$  and  $c_J$  for the gamma, inverse Gaussian and lognormal distributions. (PDF)

## Author Contributions

Conceived and designed the experiments: LK PL OP. Analyzed the data: LK OP. Contributed reagents/materials/analysis tools: LK PL OP. Wrote the paper: LK OP.

## References

- Perkel DH, Bullock TH (1968) Neural coding. *Neurosci Res Prog Sum* 3: 405–527.
- Stein R, Gossen E, Jones K (2005) Neuronal variability: noise or part of the signal? *Nat Rev Neurosci* 6: 389–397.
- Johnson DH, Glantz RM (2004) When does interval coding occur? *Neurocomputing* 59: 13–18.
- Kostal L, Lansky P, Rospars JP (2007) Review: Neuronal coding and spiking randomness. *Eur J Neurosci* 26: 2693–2701.
- Nemenman I, Lewen GD, Bialek W, de Ruyter van Steveninck RR (2008) Neural coding of natural stimuli: Information at sub-millisecond resolution. *PLoS Comput Biol* 4: e1000025.
- Theunissen F, Miller JP (1995) Temporal encoding in nervous systems: A rigorous definition. *J Comput Neurosci* 2: 149–162.
- Lansky P, Rodriguez R, Sacerdote L (2004) Mean instantaneous firing frequency is always higher than the firing rate. *Neural Comput* 16: 477–489.
- Ratliff F, Hartline HK, Lange D (1968) Variability of interspike intervals in optic nerve fibers of limulus: Effect of light and dark adaptation. *Proc Natl Acad Sci USA* 60: 464–469.
- Burns BD, Pritchard R (1964) Contrast discrimination by neurons in the cat's visual cerebral cortex. *J Physiol* 175: 445–463.
- Fenton AA, Lansky P, Olypher AV (2002) Properties of the extra-positional signal in hippocampal place cell discharge derived from the overdispersion in location-specific firing. *Neuroscience* 111: 553–566.
- Buracas GT, Albright TD (1999) Gauging sensory representations in the brain. *Trends Neurosci* 22: 303–309.
- Victor JD, Purpura KP (1997) Metric-space analysis of spike trains: theory, algorithms and application. *Network: Comput Neural Syst* 8: 127–164.
- Shinomoto S, Kim H, Shimokawa T, Matsuno N, Funahashi S, et al. (2009) Relating neuronal firing patterns to functional differentiation of cerebral cortex. *PLoS Comput Biol* 5: e1000433.
- Kostal L, Lansky P, Zucca C (2007) Randomness and variability of the neuronal activity described by the Ornstein-Uhlenbeck model. *Network: Comput Neural Syst* 18: 63–75.
- Duchamp-Viret P, Kostal L, Chaput M, Lansky P, Rospars JP (2005) Patterns of spontaneous activity in single rat olfactory receptor neurons are different in normally breathing and tracheotomized animals. *J Neurobiol* 65: 97–114.
- Kass RE, Ventura V, Brown EN (2005) Statistical issues in the analysis of neuronal data. *J Neurophysiol* 94: 8–25.
- Moore GP, Perkel DH, Segundo JP (1966) Statistical analysis and functional interpretation of neuronal spike data. *Annu Rev Physiol* 28: 493–522.
- Tuckwell HC (1988) *Introduction to Theoretical Neurobiology, Volume 2*. New York: Cambridge University Press.
- Kendall M, Stuart A, Ord JK (1977) *The advanced theory of statistics. Vol. 1: Distribution theory*. London: Charles Griffin.
- Chakravarty SR (1990) *Ethical social index numbers*. New York: Springer-Verlag.
- Ash RB (1965) *Information Theory*. New York: Dover.
- Cover TM, Thomas JA (1991) *Elements of Information Theory*. New York: John Wiley and Sons, Inc.
- Kostal L, Marsalek P (2010) Neuronal jitter: can we measure the spike timing dispersion differently? *Chin J Physiol* 53: 454–464.
- Frieden BR (1998) *Physics from Fisher information: a unification*. New York: Cambridge University Press.
- Kullback S (1968) *Information theory and statistics*. New York: Dover.
- Pouzat C, Chaffiol A (2009) Automatic Spike Train Analysis and Report Generation. An Implementation with R, R2HTML and STAR. *J Neurosci Methods* 181: 119–144.
- Reeke GN, Coop AD (2004) Estimating the temporal interval entropy of neuronal discharge. *Neural Comput* 16: 941–970.

28. Abramowitz M, Stegun IA (1965) Handbook of Mathematical Functions, With Formulas, Graphs, and Mathematical Tables. New York: Dover.
29. Berger D, Pribram K, Wild H, Bridges C (1990) An analysis of neural spike-train distributions: determinants of the response of visual cortex neurons to changes in orientation and spatial frequency. *Exp Brain Res* 80: 129–134.
30. Levine MW (1991) The distribution of the intervals between neural impulses in the maintained discharges of retinal ganglion cells. *Biol Cybern* 65: 459–467.
31. Lansky P, Sato S (1999) The stochastic diffusion models of nerve membrane depolarization and interspike interval generation. *J Peripher Nerv Syst* 4: 27–42.
32. Bershanskii A, Dremencov E, Fukayama D, Yadid G (2001) Probabilistic properties of neuron spiking time-series obtained in vivo. *Eur Phys J B* 24: 409–413.
33. R Development Core Team (2009) R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing.
34. Bercher JF, Vignat C (2009) On minimum Fisher information distributions with restricted support and fixed variance. *Inform Sciences* 179: 3832–3842.
35. Berger AL, Della Pietra VJ, Della Pietra SA (1996) A maximum entropy approach to natural language processing. *Comput Linguist* 22: 39–71.
36. Della Pietra SA, Della Pietra VJ, Lafferty J (1997) Inducing features of random fields. *IEEE Trans on Pattern Anal and Machine Int* 19: 380–393.
37. Di Crescenzo A, Longobardi M (2002) Entropy-based measure of uncertainty in past lifetime distributions. *J Appl Probab* 39: 434–440.
38. Frieden BR (1990) Fisher information, disorder, and the equilibrium distributions of physics. *Phys Rev A* 41: 4265–4276.
39. Telesca L, Lapenna V, Lovullo M (2005) Fisher information measure of geoelectrical signals. *Physica A* 351: 637–644.
40. Vignat C, Bercher JF (2003) Analysis of signals in the Fisher–Shannon information plane. *Phys Lett A* 312: 27–33.
41. Zivojnovic V (1993) A robust accuracy improvement method for blind identification using higher order statistics. In: *IEEE Internat. Conf. Acous. Speech. Signal Proc.*, 1993. ICASSP-93, volume 4, pp 516–519.
42. Ditlevsen S, Lansky P (2011) Firing variability is higher than deduced from the empirical coefficient of variation. *Neural Comput*, in press.
43. Vasicek O (1976) A test for normality based on sample entropy. *J Roy Stat Soc B* 38: 54–59.
44. Tsybakov AB, van der Meulen EC (1994) Root-n consistent estimators of entropy for densities with unbounded support. *Scand J Statist* 23: 75–83.
45. Huber PJ (1974) Fisher information and spline interpolation. *Ann Stat* 2: 1029–1033.
46. Good IJ, Gaskins RA (1971) Nonparametric roughness penalties for probability densities. *Biometrika* 58: 255–277.