



Letters

How spurious correlations affect a correlation-based measure of spike timing reliability

Jan A. Freund^{a,b,*}, Alexander Cerquera^c

^a Theoretical Physics/Complex Systems Research Group, ICBM, Carl von Ossietzky Universität Oldenburg, D-26111 Oldenburg, Germany

^b Research Center Neurosensory Science, Carl von Ossietzky Universität Oldenburg, D-26111 Oldenburg, Germany

^c Faculty for Electronic and Biomedical Engineering, Complex Systems Research Group, Antonio Nariño University, Bogotá, Colombia

ARTICLE INFO

Article history:

Received 8 April 2011

Received in revised form

30 August 2011

Accepted 3 October 2011

Communicated by A. Belatreche

Available online 11 November 2011

Keywords:

Stimulus reconstruction

Sensory neurons

Neuronal assemblies

Retinal ganglion cell population

ABSTRACT

The spike timing reliability of a neuron can be assessed via measuring the similarity of spike trains obtained in trials with repeated presentations of the same stimulus. Using a correlation-based measure of spike timing reliability we show that spurious correlations between independent Poisson spike trains can lead to a systematic misinterpretation to an extent that scales with the neural spike rate. Therefore, a correction is essential before comparing neurons with distinctly different spike rates. Such a comparison may, for instance, guide the choice of stimulus selective sensory neurons that are pooled for optimal stimulus reconstruction. We propose straightforward methods to abstract from these spurious correlations and demonstrate effects in an application to recorded spike trains of a retinal ganglion cell population.

© 2011 Elsevier B.V. All rights reserved.

1. Introduction

Neurons transmit information about a stimulus via spike trains. The specific way stimulus features shape these spike trains is the essence of a neural code. In theoretical approaches to neural coding [1–3] two dichotomies play a prominent role: rate code vs. spike timing code and single unit code vs. population code. Rate codes are robust against spike jitter and shuffling of spikes but a rather inefficient way of coding [4], in particular for short evaluation times. By contrast, the in principle high information capacity of spike timing codes is limited by the variability of spike trains observed in response to repeated presentations of the same stimulus. Spike jitter as well as additional or missing spikes are caused by neuronal noise and have the effect of limiting spike timing reliability.

The dichotomy of single unit vs. neural populations gains importance for fast coding strategies, i.e. when stimulus features have to be reconstructed within a short time span allowing at most one spike per neuron to be elicited. Under such circumstances a single unit rate estimate must necessarily be highly unreliable. On the contrary, a population rate is less prone to estimation errors because of mitigating ensemble statistics. However, many redundant neurons would be needed to suppress

statistical fluctuations sufficiently which makes this short-time population rate code highly inefficient [4]. An improvement is achieved by evaluating the timing (or just the rank-order [5]) of all those spikes across the population which are triggered by the stimulus. Admittedly, in a single trial situation stimulus-induced spikes cannot be easily separated from spontaneous spikes. However, in trials with repeated presentations of the same stimulus a significant similarity of spike trains across the trial ensemble can be seen as an indication of a stimulus selective neuron. It is evident that stimulus selectivity is beneficial for stimulus reconstruction. Therefore, the task to build optimal sub-populations means to separate stimulus selective neurons from non-selective ones. A constructive way to reach this goal is to assess stimulus selectivity of each neuron by measuring its reliability, i.e. the spike train similarity of the repeated trial ensemble, and then rank and threshold the population with respect to this measure (cf. the discussion in Section 5).

Several measures of spike train similarity exist, e.g. [6–11]. Paiva et al. [12] reported a comparison of binless spike train measures. Using measures without binning the time axis avoids boundary effects due to quantization of spike trains and is favorable for estimation. In the same publication three of these binless measures, Victor's and Purpura's spike train metrics [6], van Rossum's distance [7] and Schreiber et al.'s correlation measure, were put in a unifying perspective through a formulation via different kernel functions. Notwithstanding the fact that these methods are binless the kernel width introduces a

* Corresponding author at: Theoretical Physics/Complex Systems Research Group, ICBM, Carl von Ossietzky Universität Oldenburg, D-26111 Oldenburg, Germany.

E-mail address: jan.freund@uni-oldenburg.de (J.A. Freund).

smoothing parameter that controls the importance of spike timing precision.

Here we want to point out that in a comparative spike train reliability analysis of neurons with different rates the choice of this kernel width strongly influences the level of statistical significance which can be seen as a kind of bias. Statistical significance comes into play when assessing the reliability value observed for a specific neuron against the backdrop of spurious correlations occurring in reference ensembles of spike trains that are constructed in accordance with a null-hypothesis. In the case of comparatively short observation times that we have in mind the reference ensemble will be given by Poisson spike trains of a (constant) firing rate identical to the related estimate of the analyzed neuron. Generalizations to longer spike trains will be discussed in Section 6. Our presentation will be focused on Schreiber et al.'s correlation measure, however, the mentioned unifying perspective offered by the kernel function approach allows to anticipate similar statements for other spike train measures.

Starting from a binned version of the correlation measure (Section 2) we address the effect of spurious correlation arising in stationary Poisson spike train ensemble and combine numerical results (Section 3) with an analytical approximation (Section 4). A similar approach for the binless formulation is thinkable but will be left to future research. The essential point is that non-removal of the bias caused by spurious correlations can lead to an inappropriate assessment of neural populations and, consequently, to a non-optimal subpopulation. This will be illustrated by application to real data from a set of retinal ganglion cells (Section 5). Finally, we summarize and discuss implications and generalizations of our findings (Section 6).

2. The correlation-based measure

We assume that N spike trains were recorded from a single neuron subjected to repeated presentations of the same stimulus. The method proposed by Schreiber et al. [8] proceeds as follows:

- A spike train, characterized by the sequence of M spike times t_1, \dots, t_M , is convoluted (smoothed) with a Gaussian filter of temporal width σ_c yielding the signal

$$s(t) = \sum_{i=1}^M \frac{1}{\sqrt{2\pi\sigma_c^2}} \exp\left(-\frac{(t-t_i)^2}{2\sigma_c^2}\right). \quad (1)$$

The parameter σ_c must be chosen by the experimenter and should match the estimated temporal timing precision of spikes. Therefore, it is characteristic for a specific (type of) neuron.

- Binning: the time-continuous signal $\{s(t)|0 \leq t \leq T\}$ is then converted to a vector $\vec{s} = (s_1, \dots, s_K)$ by choice of a bin width Δt and by the definition of the k -th vector component as the mean value

$$s_k = \frac{1}{\Delta t} \int_{(k-1)\Delta t}^{k\Delta t} s(t) dt \quad \text{for } k = 1, \dots, K. \quad (2)$$

Here the number K of vector components is related to the bin width Δt and the total observation time T by $K = \lfloor T/\Delta t \rfloor$ (where $\lfloor \dots \rfloor$ indicates the nearest smaller integer). Choosing $\Delta t < \sigma_c$ makes the integrand in (2) a mildly varying function. Hence, for any $0 \leq \theta \leq 1$, $s_k = s(\lfloor k-\theta \rfloor \Delta t)$ is an equally valid choice. In all of our calculations we fixed $\Delta t = 1$ ms, the shortest interspike interval reflecting absolute neuronal refractory time. Such a small bin width substantially reduces quantization effects, even more so, since binning is performed after smoothing.

- Performing the first two steps for each of the N spike trains yields a set of vectors $\{\vec{s}_1, \dots, \vec{s}_N\}$. Each of these vectors can be imagined as an arrow in a K -dimensional space. Reliable spike responses correspond to a bunch of N arrows pointing in the same direction. This however means that the set of angles between all possible pairs of arrows is scattered around zero or, equivalently, that the related average cosine is close to one. The mathematical expression for this average cosine is

$$R = \frac{2}{N(N-1)} \sum_{i=1}^{N-1} \sum_{j=i+1}^N \frac{\vec{s}_i \cdot \vec{s}_j}{|\vec{s}_i| |\vec{s}_j|}. \quad (3)$$

Notice that $0 \leq R \leq 1$ because all vector components are non-negative which means that the angle between an arbitrary pair of vectors \vec{s}_i and \vec{s}_j never exceeds the range $[-(\pi/2), \pi/2]$. We note that the sum in (3) is mathematically ill-defined if it contains a pair with $|\vec{s}_i| = 0$ or $|\vec{s}_j| = 0$ (or both). This reflects the impossibility to define an angle if (at least) one of the two vectors is the null vector. Since the numerator of (3) balances only non-vanishing segments of smoothed spike trains we choose to define

$$\frac{\vec{s}_i \cdot \vec{s}_j}{|\vec{s}_i| |\vec{s}_j|} = 0 \quad \text{if } |\vec{s}_i| = 0 \vee |\vec{s}_j| = 0. \quad (4)$$

As pointed out by Schreiber et al. [8] R accounts for the similarity of spike trains and, contrary to PSTH-based measures, is sensitive to slow variations in firing rate across cell trials.

3. Spurious correlations between independent Poisson spike trains

The use of the correlation-based measure (3) as a measure of spike timing reliability is justified whenever the quantified similarity of spike trains is beyond the accidental similarity of *statistically independent* spike trains. However, because of the non-negativity of the correlation-based measure spurious correlations reflecting such accidental similarities must necessarily be reflected by positive values. Therefore, these positive deviations from zero constitute a kind of bias that should be excluded from considerations of spike timing reliability.

To see how the bias statistics depends on the spike rate α of the neuron and on the choice of parameters we consider an ensemble of independent Poisson spike trains. These are constructed by concatenation of interspike-intervals τ_i chosen independently according to

$$\tau_i = \frac{\ln\left[\frac{1}{\xi_i}\right]}{\alpha}, \quad (5)$$

where ξ_i is a random number equidistributed in the unit interval. We thus generate a sequence of independent ISIs that are exponentially distributed, i.e.

$$p(\tau) = \alpha e^{-\alpha\tau}. \quad (6)$$

Both mean and standard deviation of the ISI distribution are given by α^{-1} which, therefore, is the characteristic time scale of the neuron. In Fig. 1 we show an ensemble of 10 independent spike trains that results from concatenation of such ISIs together with the signal traces resulting from the Gaussian filter.

When comparing the spike timing reliability of a population of neurons with vastly different spike rates one is left with two options:

- A: use of the correlation-based measure (3) with a *population specific* constant σ_c . This choice is reasonable if from physiological considerations the spike timing precision is identical for all neurons of the population, notwithstanding the fact that they might possess significantly varying spike rates. In case no such

Download English Version:

<https://daneshyari.com/en/article/410079>

Download Persian Version:

<https://daneshyari.com/article/410079>

[Daneshyari.com](https://daneshyari.com)