MEASURE OF THE REGULARITY OF EVENTS IN STOCHASTIC POINT PROCESSES, APPLICATION TO NEURON ACTIVITY ANALYSIS

D. Labarre, W. Meissner, T. Boraud

Université Bordeaux 2, CNRS, UMR 5227, 146 rue Léo-Saignat, 33076 Bordeaux Cedex, FRANCE

ABSTRACT

Numerous researches aim at understanding the high brain functions such as memory or decision making by analysing the activity of brain neurons. This activity corresponds to sequences of electrical potentials and thus can be viewed as a point processes. In this paper, we propose a method to measure the regularity level of event occurrences in point processes. Based on the analysis of the so-called density histogram, the proposed approach has the advantage of providing a decision to classify the process into one of the three following distinct classes: the "regular" processes, the "irregular" processes and the "bursting" processes. To illustrate the efficiency of the method, we first carry out a comparative study based on synthetic data. Then, the algorithm is tested in the framework of neurosciences for the classification of neurons according to their activity.

Index Terms**—** Point process, Poisson process, density histogram, goodness-of-fit, neuron classification.

1. INTRODUCTION

Analysing the temporal distribution of events is of interest in various frameworks such as queuing traffic in communication [7], data compression [6] or neurosciences [3] [5]. Indeed, neurons communicate by means of electrical potentials, the so-called spikes. A sequence of spikes, called a spike train, can be viewed as a stochastic point process and is usually assumed to be a Poisson process [2]. The information is embedded in the temporal arrangement of the spikes [2]. A neuron has one of the following three activities:

- Regular activity: the spikes are produced pseudo-periodically. Cf. figure 2.*a)*.
- Irregular activity: the spikes are produced randomly, according to a Poisson law [2]. Cf. figure 2.*b)*.
- Bursting activity: the neuron produces packs of spikes, the so-called bursts. Cf. figure 2.*c)*.

In this paper, we focus our attention on the classification issue. More particularly, we develop a method that allows to automatically make a decision about the activity that is displayed by such point processes.

Let us consider a stochastic point process in which each event is characterized by its time occurrence $t(k)$. The temporal arrangement of the events can be analyzed with various tools such as the number of events per second or the autocorrelogram of the observed point process for instance [2]. However, these techniques are not well-suited for the classification of point processes since they do not encode the regularity features on a reduced number of parameters. For this reason, we analyse the temporal arrangement of the events from the time intervals between two successive events. To this end, we define the Inter-Event Interval (*IEI*) process as follows:

$$IEI(k) = t(k) - t(k-1) \,. \qquad (1)$$

By counting the number of intervals $IEI(k)$ falling in time bins, we obtained the so-called histogram of the *IEI* process [2]. The *IEI* histogram corresponds to an estimation of the probability density function (PDF) related to the inter-event intervals. Then, it emphasizes the researched regularity features in the point process. As an example, an *IEI* histogram that fits a narrow band Gaussian PDF is typical of a regular appearance of events [2]. Thus, goodness-of-fit based methods exploiting the *IEI* histogram have been developed to characterize the distribution of event time occurrences [1] [5] [7]. However, such methods remain questionable. Indeed, choosing the number of bins to estimate the *IEI* histogram is a delicate issue and affects the shape of the obtained *IEI* histogram. Thereby, if the number of bins is not well tuned, the *IEI* histogram can exhibit similar shapes for different kind of activities [3].

As an alternative, we propose to take advantage of the so-called density histogram [3]. This latter provides a robust representation of the temporal distribution of events by taking into account the mean *IEIm* of the *IEI* defined as follows:

$$IEIm = \frac{1}{N}\sum_{k=1}^{N} IEI(k) \,. \qquad (2)$$

The density histogram $d(\lambda)$ is defined as the probability to have λ events in an interval of length $IEIm$, where λ is an integer. In practice, given the process $t(k)$, estimating the density histogram consists in first evaluating the number of events in each frame *i* of length $IEIm$ as follows:

$$card\left\{ t(k) \text{ such as } i \le \frac{t(k)}{IEIm} \le i+1 \right\}. \quad (3)$$

where $card\{.\}$ denotes the cardinal.

Then, for each integer $\lambda = 0,1,\ldots$, the density $d(\lambda)$ is estimated as the number of frames which contains λ events and normalized by the total number of frames.

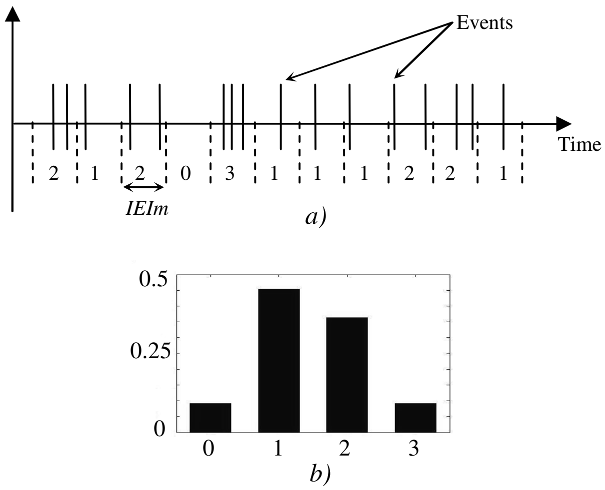

Figure 1: Example of event counting a) and the related density histogram b).

Let us consider the example in figure 1 *a)*. The point process is divided into eleven frames. One frame contains zero event, five frames contain one event, four frames contain two events and one frame contains three events. The resulting density histogram is illustrated in figure 1 *b)*.

The temporal distribution of events in stochastic point processes will lead to specific shapes for the density histogram. More particularly, let us consider the three kinds of activity:

- Regular activity: the probability to have one event in a frame is significantly higher than the other probabilities. The density histogram tends to be symmetric with mean equal to one.
- Irregular activity: the probability to obtain 0, 1 or 2 events in a frame will not be significantly different. The density histogram tends to fit a Poisson PDF with mean equal to one.
- Bursting activity: the probability to have 0 event in a frame is significantly higher than the other probabilities. The density histogram tends to fit a Poisson PDF with mean less than one.

Figure 2 provides examples of density histograms obtained from real spike trains.

In the neurosciences, Kaneoke et al. [3] have previously exploited the density histogram to detect bursts in spike trains. In their method, bursts are detected if the two following conditions are satisfied:

- The *IEI* histogram is sufficiently non symmetric, i.e. the activity is not regular. This condition is checked by comparing the skewness of the *IEI* histogram to an a priori threshold.
- The density histogram fits a Poisson PDF with mean equal to 1. A Chi square test is used to check this condition.

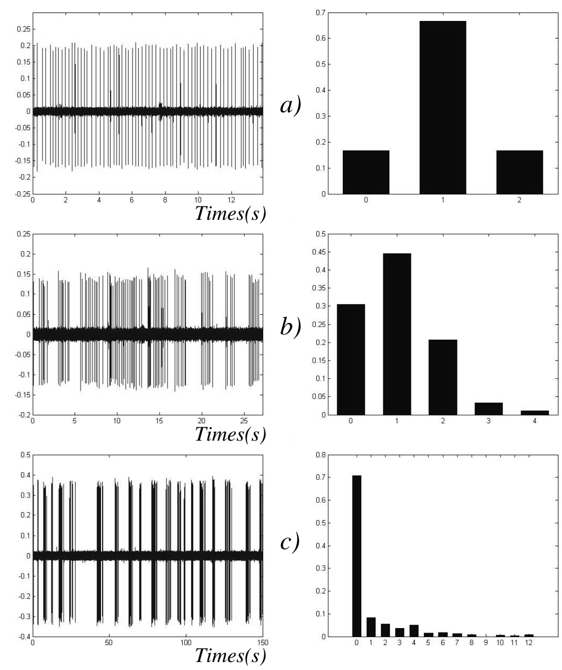

Figure 2: Examples of spike trains (left) and their corresponding density histograms (right) in different cases; a) regular activity, b) irregular activity, c) bursting activity. These signals have been recorded on sleeping rats in deep brain structures, the so-called basal ganglia.

However, such an approach does not allow to discriminate regular activities and irregular activities. In addition, it has the following drawbacks. Firstly, tuning the threshold for the skewness in the first condition is a key point of the method. Secondly, choosing the number of bins for the *IEI* histogram is a difficult task. Moreover, this choice affects the estimation of the skewness and thus has to be taken into account to adjust the threshold. Finally, it should be noted that the authors did not provide sufficient practical detail on the implementation of their method.

In this paper, we extend the approach of Kaneoke et al. [3]. The advantages of the proposed method are threefold. Firstly, no parameter such as the number of bins or a threshold has to be tuned. Secondly, no statistic such as the skewness has to be estimated, avoiding estimation and decision errors. Thirdly, while Kaneoke's method is oriented toward the detection of bursting activity, we propose a one-step based algorithm that provides a decision according to which the stochastic point process can be classified into three distinct classes: regular activity, irregular activity and bursting activity.

The paper is organized as follows: in section 2, we present the proposed classification method, in section 3, a comparative study on synthetic data is carried out and simulation results based on real data are provided.

2. THE CLASSIFICATION METHOD

In this section, we present a simple method to classify stochastic point process according to their activity. The approach is based on the comparison of the density histogram $d(\lambda)$ to a reference density function $p_x(\lambda)$. Given the different possible shapes of the density histogram shown in figure 2, this reference PDF will be chosen for each activity hypothesis as follows:

- Regular hypothesis: The density histogram is expected to be symmetric with mean 1 and is thus compared to a Gaussian law $p_G(\lambda)$ with mean equal to 1 and variance equal to 0.5.
- Irregularity hypothesis: The density histogram is compared to a Poisson law $p_{P(1)}(\lambda)$ with mean equal to 1.
- Bursting hypothesis: The density histogram is compared to a Poisson law $p_{P(0.2)}(\lambda)$ with mean equal to 0.2.

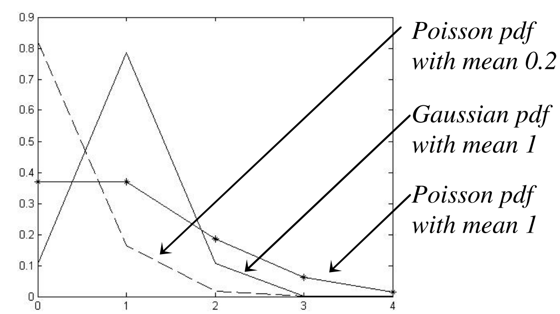

Figure 3: The three probability density functions used in the proposed method in the case of five bins

These three reference PDF are shown in figure 3. Various tests can be considered for the goodness-of-fit between $d(\lambda)$ and $p_x(\lambda)$, such as the Chi square, Kullback-Leibler or Kolmogorov-Smirnov tests [4]. However, the PDF are formed from a reduced number of samples. In addition, they can take values very close to zero. Thus, the Chi square and Kullback-Leibler tests are not considered in this paper. We propose the two following methods:

- Method 1: The estimated density histogram $d(\lambda)$ is compared to the reference PDF $p_x(\lambda)$ by means of a mathematical distance. The decision is made by choosing the hypothesis corresponding to the reference PDF that produces the smallest distance with $d(\lambda)$. In this paper, we test the three following distances: the 1-norm, 2-norm and infinity-norm distances.
- Method 2: The decision is made by comparing the cumulative distribution functions. As in method 1, the hypothesis corresponding to the reference PDF that produces the smallest distance is chosen. The same distances as in method 1 are tested. It should be noted that using the infinity-norm in that case can be related to the Kolmogorov-Smirnov test [4].

In the next section, we provide simulation results based on synthetic data and electrophysiological recordings.

3. SIMULATION RESULTS

3.1. Tests on synthetic data

We first test the method on synthetic data. The generated point processes have the following form:

$$t(k) = \sum_{i=1}^{k} s(i) . \quad (4)$$

The process $s(i)$ corresponds to the *IEI* process and is defined according to the desired activity:

- Regular activity:

In that case, the process $s(i)$ is simulated according to a Gaussian PDF with mean μ and variance σ^2. μ is set to 0.001, 0.01, 0.05, 0.08 or 0.1. σ^2 is set to 0.5, 1, 2, 5 or 10.

- Irregular activity:

The process $s(k)$ is generated as follows:

$$s(k) = P_{P(\mu)}(k) \times 100 \quad (5)$$

where $P_{P(\mu)}(k)$ is a Poisson distributed random number with mean μ.

μ is set to values varying from 1 to 3 with a step of 0.2.

- Bursting activity:

In that case, the inter-event intervals during a burst are chosen according to a Gaussian PDF with variance 1 and mean ranging from 1 to 10. In addition, the inter-burst intervals are chosen according to a Gaussian PDF with variance 100 and with mean ranging from 200 to 500.

The length of the process $t(k)$ varies from 20 to 500 events. In each activity case, 100 realizations are generated for each configuration of the parameters.

Table I: Percentage of detection on synthetic data using the 1-norm distance

		Estimated activity		
		Regular	*Irregular*	*Bursting*
Method 1	*Regular*	**99.08**	0.92	0
	Irregular	15.18	**84.80**	0.02
	Bursting	0	1.36	**98.64**
Method 2	*Regular*	**99.08**	0.92	0
	Irregular	16.05	**83.95**	0
	Bursting	0	43.25	**56.75**

Averaged percentages of decision are collected in table I, II and III. When using Method 1, results are similar with the three distances. When using Method 2, the 2-norm distance provides more reliable results. In addition, most of false decisions are made when analyzing an irregular activity. Indeed, in real cases, density histogram shapes corresponding to regular and irregular activities can be quite similar. Nevertheless, when using Method 1 with the 2-norm distance, the proposed approach allows to make a reliable decision on the kind of activity.

Table II: Percentage of detection on synthetic data using the 2-norm distance

		Estimated activity		
		Regular	*Irregular*	*Bursting*
Method 1	*Regular*	**99.08**	0.92	0
	Irregular	15.38	**84.55**	0.07
	Bursting	0	1.36	**98.64**
Method 2	*Regular*	**97.65**	2.35	0
	Irregular	12.89	**86.84**	0.27
	Bursting	0	2.85	**97.15**

Table III: Percentage of detection on synthetic data using the infinity-norm distance

		Estimated activity		
		Regular	*Irregular*	*Bursting*
Method 1	*Regular*	**99.08**	0.92	0
	Irregular	14.98	**84.96**	0.06
	Bursting	0	4.66	**95.34**
Method 2	*Regular*	**99.82**	0.18	0
	Irregular	24.36	**75.64**	0
	Bursting	0	19.08	**80.92**

3.1. Tests on real data

Neurosciences constitute an active framework and numerous researches aim at understanding the high brain functions such as memory, learning or decision making by analysing the pattern of neuronal discharge. These studies are strongly interdisciplinary since they imply aspects of system modelling or the processing of electrophysiological signals recorded in the nervous system.

In this paper, we focus our attention on voluntary movement mechanisms and their related pathologies such as Parkinson's disease or dystonia. Dystonia is characterized by involuntary movements and prolonged muscle contractions, resulting in twisting body motions and abnormal postures. The pathophysiology of dystonia remains largely unknown.

Deep brain electrical stimulation (DBS) is a surgical operation which improves voluntary movement control of patients suffering from Parkinson's disease and dystonia. However, the effects of DBS are not completely understood. In this context, we have tested the proposed method to analyze neuronal activity of dystonic hamsters during DBS. Deep brain structures, the so-called basal ganglia that are implicated in the pathophysiology of dystonia, have been both stimulated and recorded. Ninety-six recorded signals have been analysed before and during DBS. The proposed approach has been applied to estimate the ratio of regular, irregular and bursting neurons. As shown in figure 4, DBS seems to disturb the regular neuron activity which becomes mainly irregular. It should be noted that Kaneoke's method only enables to distinguish regular and bursting neurons. Therefore, it can not be used for this analysis.

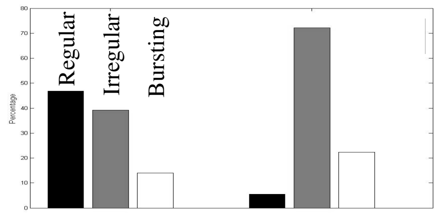

Figure 4: Classification of neurons in the context of deep brain stimulation

5. CONCLUSION

We have presented in this paper a method to analyze the temporal distribution of events in stochastic point processes. The approach is based on the estimation of the density histogram. This latter is compared to reference density probability functions. We have carried out a comparative study between methods using various mathematical distances. Finally, we have exercised the method in the context of deep brain electrical stimulation in an animal model of dystonia.

Acknowledgement: The authors would like to thank Abdelhamid Bennazouz and Jonathan Chetrit for providing the recordings used in figure 2.

REFERENCES

[1] J. H. Cocatre-Zielgien, F. Delcomyn, Identification of bursts in spike trains, Journal of Neuroscience, vol.41, n°1, pp.19-30, 1992.

[2] Gabbiani, F. and Koch, C. (1998). Principles of spike train analysis, In C. Koch and I. Segev, editors, Methods in Neuronal Modeling, chapter 9, pages 312-360. MIT press, 2nd edition.

[3] Y. Kaneoke, J.L. Vitek, Burst and oscillation as disparate neuronal properties, Journal of Neuroscience Method, vol. 68, pp.211-223, 1996.

[4] A. Papoulis, S. U. Pillai, Probability, Random Variables and Stochastic Processes, Fourth edition, Tata McGraw-Hill Edition, 2002.

[5] D. H. Perkel, G. L. Gerstein, G. P. Moore, Neuronal Spike Trains and Stochastic Point Process, I. The Single Spike Train, Biophysical Journal, vol. 7, n°4, pp. 391-418, 1967.

[6] I. Rubin, Information Rates and Data-Compression Schemes for Poisson Processes, IEEE Trans. on Information Theory, Vol. IT-20, n°2, pp.200-210, March 1974.

[7] P. Skelly, M. Schwartz, S. Dixit, A histogram-based model for video traffic behavior in an ATM multiplexer, IEEE/ACM Trans. Networking, vol. 1, no.4, pp. 446459, August 1993.