Detection of synchronized firings in multivariate neural spike trains during motor tasks

Jérôme Bourien, Justin Sanchez, Jean-Jacques Bellanger, Fabrice Wendling, José Principe

Proceedings of the 29th Annual International
Conference of the IEEE EMBS
Cité Internationale, Lyon, France
August 23-26, 2007.

Detection of synchronized firings in multivariate neural spike trains during motor tasks

J. Bourien, J. C. Sanchez, *Member*, J. J. Bellanger, F. Wendling and J. C. Principe, *Fellow*

***Abstract*— This paper describes and compares two classical methods for the detection of neuron groups which exhibit synchronized firings in multivariate spike trains. These methods were compared on experimental and randomized data corresponding to the firing activity of 104 neurons located in motor, premotor, and parietal cortices in a monkey during movement tasks. Both methods exhibited high false positive rates in randomized data, but results showed that this rate can be advantageously reduced with a simple postprocessing. Otherwise, one method permitted to detect a significant number of synchronized groups of neurons related to the behavioral task.**

I. Introduction

Progress in the design of neural multielectrode recording techniques [1] has increased the need for analysis methods of multivariate neural spike trains of large dimension. The challenges of analyzing data of this type have been reviewed in [2]. In this paper, we compare two methods that are able to extract groups of neurons with synchronized firings. The first one was proposed in 1978 by Gerstein and colleagues [3] and the second was submitted by Grün and collaborators in 1994 [4].The classical way to deal with this problem is to extract groups for which the firing joint probability is higher than the expected firing joint probability when the neurons fire independently but with equal spike rates. In practice, the firing joint/marginal probabilities are estimated using a coincidence window and non-independence condition is dealt with using a statistical test.

Here we present these two methods in a common mathematical framework and to compare detection performances in experimental and randomized data in a Brain-Machine Interface behavioral

Manuscript received April 2, 2007.

J. Bourien, is with: CREFA-Université Montpellier1, 15 avenue Charles Flahault, 34094 Montpellier, Cedex 5 and with the Institute of Neurosciences of Montpellier, INSERM U583, 80 rue Augustin Fliche, 34091 Montpellier, France (phone: 334-6754-8691; fax: 334-6766-8123; e-mail: jerome.bourien@univ-montp1.fr)

J. C. Sanchez is with the Department of Pediatrics, Division of Neurology, University of Florida, P.O. Box 100296, JHMHC, Gainesville, FL 32610 USA (e-mail: jcs77@ufl.edu).

J.J. Belanger and F. Wendling, are with INSERM U642, LTSI, University of Rennes 1, Rennes, France (e-mail: fabrice.wendling@univ-rennes1.fr).

J. C. Principe is with the Department of Electrical and Computer Engineering, University of Florida, Gainesville, FL 32611 USA (e-mail: principe@cnel.ufl.edu).

II. Methods presentation

A. Multivariate spike train

A multivariate spike train of duration T may be represented by a multivariate signal $S(t)$, $t \in [0,T]$:

$$S(t) = \begin{pmatrix} \sum_{k=1}^{n_1} \delta(t-\tau_{1,k}) \\ \cdots \\ \sum_{k=1}^{n_i} \delta(t-\tau_{i,k}) \\ \cdots \\ \sum_{k=1}^{n_N} \delta(t-\tau_{N,k}) \end{pmatrix} \text{ with } \delta(t) = \begin{cases} 1 \textbf{ if } t = 0 \\ 0 \textbf{ if } t \neq 0 \end{cases}$$

where n_i and $\{\tau_{i,1},...,\tau_{i,n_i}\}$ for $i = 1,...,N$ denote respectively the count and the time occurrences of neuronal firings in the i^{th} neuron on the N observed neurons.

One classical way to randomize a multivariate spike train is to perform a random permutation of the interspike intervals in each spike train. This procedure is called data shuffling [5]. On the one hand, it guarantees that the i^{th} randomized and original spike trains have identical duration, spike counts, and interspike interval histograms. On the other hand, synchronized firings are broken in the randomized multivariate spike train.

B. Processing of a multivariate spike train

The two analysis methods are not directly applied to multivariate spike trains but on a binary matrix derived from these spike trains. The time axis $[0,T]$ is segmented into M adjacent coincidence time windows of duration D ($M = T/D$). Spikes from the same train that occurred into the same time window are grouped and transformed into a binary value equal to 1 if one or more spike(s) occurred during D and 0 otherwise. This transformation known as binning operation yields to a binary matrix B with N rows and M columns.

For statistical analysis of coincidences, we considered that the column vectors of matrix B correspond to M independent realizations of a binary random vector $X = [x_1,...,x_N]^t$, $x_i \in \{0,1\}$ (the t exponent denotes the transpose operator). Then, the problem of extracting a group of co-activated neurons labeled by indices $\{w_1,...,w_n\} \subset \{1,...,N\}$, is equivalent to the problem of finding a group of coordinates in X conjointly equal to 1 (i.e. $\{x_{w_1} = 1,...,x_{w_n} = 1\}$) with an occurrence in B significantly higher than the expected occurrence when neurons fired in a independent way but with equal rates.

Two methodological approaches were proposed to solve the aforementioned problem. The first approach consists in searching groups of coordinates in X conjointly equal to 1 independently of the values of the other coordinates (i.e. non-exclusive activation of the considered group). The second approach is aimed at extracting groups of coordinates in X conjointly equal to 1 with all other coordinate values being equal to zero (i.e. exclusive activation of the considered group). In the following of the paper, the non-exclusive (resp. exclusive) occurrence of a neuron group $\{w_1,...,w_n\}$ will be referred as $n^{NE}(\{w_1,...,w_n\})$ (resp. $n^{E}(\{w_1,...,w_n\})$).

Gerstein and colleagues proposed an implementation of

TABLE I
TWO-WAY CONTINGENCY TABLE

	Activation of neuron group $\{w_1,...,w_n\}$	Not activation of neuron group $\{w_1,...,w_n\}$	
Activation of neuron $\{j\}$	a	b	$a+b$
Not-activation of neuron $\{j\}$	c	d	$c+d$
	$a+c$	$b+d$	M

With $\{j\}\in\{1,...,N\}-\{w_1,...,w_n\}$, $a=n^{NE}(\{w_1,...,w_n,j\})$, $b=n^{NE}(\{j\})-a$, $c=n^{NE}(\{w_1,...,w_n\})-a$, and $d=M-a-b-c$.

The M columns of matrix B can be partitioned in the above two-way contingency table. If the activation of the neuron group $\{w_1,...,w_n\}$ is independent of the activation of the neuron $\{j\}$, the indicator $I=\dfrac{M(ad\times cb)^2}{(a+b)(a+c)(c+d)(b+d)}$ may be assumed to have a χ^2 distribution with one DF for $M>40$. In this way, a co-activated group of neurons $\{w_1,...,w_n\}$ is detected if $I\geq I_\beta$ where I_β depends of the given confident probability β (for example $I_\beta=6.635$ for $\beta=99\%$).

the first approach (see [3] for details) in order to detect and identify "functional groups of neurons" in multivariate spike train. The iterative algorithm rejects independent groups of neurons using a χ^2 test (see Table I for details).

An implementation of the second approach, called Unitary Event analysis, was proposed in 1994 [4] (see [6] for implementation details) in order to detect episodes of synchronized neural activity in multivariate spike trains. This algorithm compares the observed occurrence $n^{E}(\{w_1,...,w_n\})$ and the expected occurrence $n^{E}_{\exp}(\{w_1,...,w_n\})$ when the neurons fire independently but with the same rates. The deviation is evaluated with the *joint-surprise* function $S(\Psi)=\log\dfrac{1-\Psi}{\Psi}$ with

$$\Psi(n^{E}(\{w_1,...,w_n\}\mid n^{E}_{\exp}(\{w_1,...,w_n\})=\sum_{k=n^{E}(\{w_1,...,w_n\})}^{+\infty}\frac{(n^{E}_{\exp}(\{w_1,...,w_n\}))^k}{k!}e^{-n^{E}_{\exp}(\{w_1,...,w_n\})}$$

With this transformation:

$S(\Psi)>0$ if $n^{E}(\{w_1,...,w_n\})>n^{E}_{\exp}(\{w_1,...,w_n\})$,

$S(\Psi)\approx 0$ if $n^{E}(\{w_1,...,w_n\})\approx n^{E}_{\exp}(\{w_1,...,w_n\})$, and

$S(\Psi)<0$ if $n^{E}(\{w_1,...,w_n\})<n^{E}_{\exp}(\{w_1,...,w_n\})$.

Then, a group of co-activated neurons $\{w_1,...,w_n\}$ is detected if $S(\Psi)\geq S_\beta$ where S_β depends of the given confident probability β (for example $S_\beta=2$ for $\beta=99\%$).

These two methods will be referred respectively as Functional Groups Detection Method (FGDM) and Unitary Event Detection Method (UEDM) in the following of the paper. Methods are both controlled by two identical

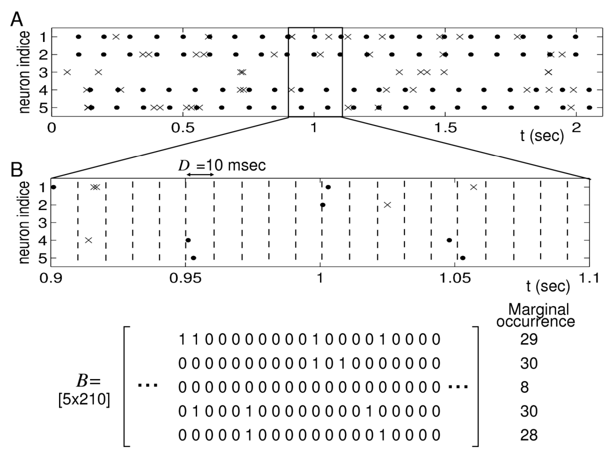

Fig. 1. Binning of a synthetic multivariate spike train. **A**-. Synchronized firings were introduced between neuron couples {1,2} and {4,5} with dithered regular spike trains of 10 spikes/s. The dither was such that the maximum interval-interspike duration between two synchronized firings can not exceed 6 ms whatever the synchronized group. The synchronized times between neurons 1-2 and 4-5 were in phase opposition in order to avoid overlapping. Additional independent firings were simulated using a Poisson process of 5 spikes/s in each spike train. **B**- A binning transformation was applied to the multivariate spike train for D =10 ms. This transformation yielded to a binary matrix B with 5 rows and 210 columns.

parameters: the coincidence window size D and the confident probability β.

III. DATA

A. Synthetic data

The Fig. 1-A shows a synthetic multivariate spike train introduced to benchmark the methods presented here. The simulated data exhibits synchronized (see black dots) and independent firings (see crosses) between five artificial neurons during an epoch of 2.1 s.

B. In Vivo Data

Simultaneously recorded spike trains (N=104 neurons from the motor, premotor, and parietal cortices) were

TABLE II
GROUPS DETECTION IN A SYNTHETIC MULTIVARIATE SPIKE TRAIN

	FGDM		UEDM		
Neuron groups	n^{NE}	I ($I_{99\%}=6.635$)	n^{E}	$n^{E}_{\exp}$	$S(\Psi)$ ($S_{99\%}=2$)
{1,2} [a,b]	15	**38.514 (*)**	12	2	**5.865 (*)**
{1,3}	3	3.922	0	-	-
{1,4}	3	0.427	2	2	0.165
{1,5}	3	0.260	3	2	0.321
{2,3}	3	3.660	0	-	-
{2,4}	4	0.026	0	-	-
{2,5}	4	0	1	2	0.805
{3,4}	1	0.022	0	-	-
{3,5}	0	-	0	-	-
{4,5} [a,b]	13	**27.260 (*)**	10	2	**4.332 (*)**
{1,2,3} [a,b]	3	**11.556 (*)**	2	0	**+∞ (*)**
{1,2,4}	1	0.766	0	-	-
{1,2,5}	0	-	0	-	-
{1,4,5}	0	-	0	-	-
{2,4,5} [b]	3	0.875	3	0	**+∞ (*)**
{1,2,3,4} [b]	1	0.902	1	0	**+∞ (*)**
{1,2,3,5}	0	-	0	-	-

Iterative progress of FGDM: At the first iteration, the non-exclusive occurrences of the ten pairs of neurons are estimated in the $[5\times210]$ matrix B defined in Fig. 1. The pairs $\{1,2\}$ and $\{4,5\}$ which exhibited indicators I higher to the threshold $I_{99\%}$ were kept. At the second iteration, the algorithm built 6 triplets and selected the triplet $\{1,2,3\}$. At the third iteration, two quadruplets were built but neither was selected consequently the algorithm stopped its progress.
Progress of UEDM: The algorithm determined the set of distinct columns in B with a least 2 active coordinates and occurring at least one time. Algorithm found a set of 8 columns occurrences varying from 1 to 12. Corresponding expected occurrences and joint-surprise values were then computed. Five groups exhibited joint-surprise values higher than the threshold $S_{99\%}$. Note that for a given group: $n^{NE}\geq n^{E}$. [a]group detected using FGDM, [b]group detected using UEDM, (*) detected group.

collected in an owl monkey during a brain-machine interface experiment at Duke University, as described in [7]. Behaviorally, the monkey performed a motor task in which it was cued to reach food from a stationary position. The movement was repeated 72 times. The position of the hand was recorded in time synchronized with the neuronal data. The duration of each task was random but the movement sequence was the same: 1) the monkey's hand was at rest, 2) it reached for food, 3), food was carried to the mouth, and finally 4) the hand returned to the rest position. We manually segmented the neuronal recordings using the 3-D hand trajectories into only two behavioral states: hand at rest (i.e. preparation of movement) and hand moving (i.e. execution of movement). Time segments corresponding to repeated movement execution were concatenated and used for data analysis. The cumulative duration of execution of movement was equal to 60.303 s. In this way, we obtained a multivariate spike train composed of 104 single spike trains of duration 60.303 s.

C. *Randomized In Vivo Data*

Randomized experimental data were generated using the randomization procedure described in section II-A. Shuffled simulated data were composed of 104 independent spike trains with a duration of 60.303 s. In order to estimate the distribution of the number of extracted groups as a function of parameter *D*, a series of simulated data sets was generated by the repetition of this randomization procedure (100 times).

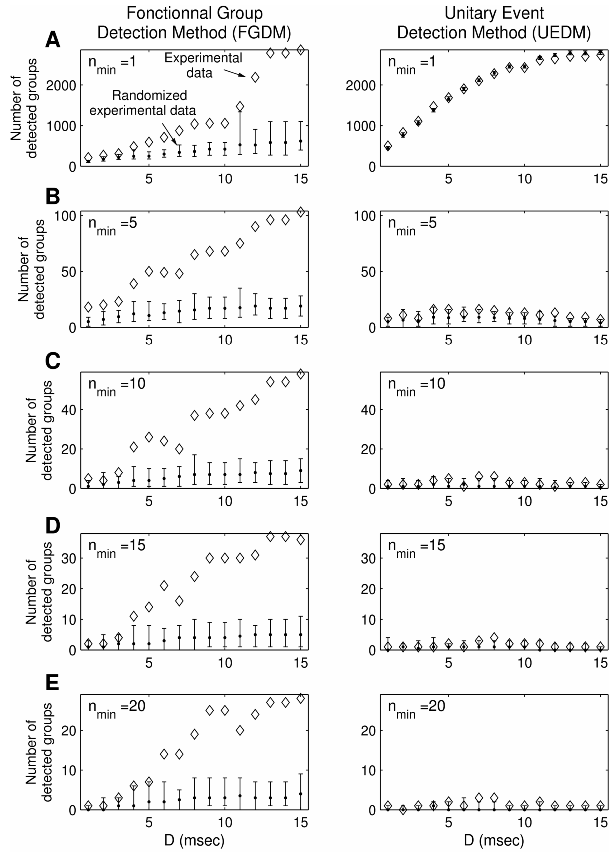

Fig. 2. Number of detected groups in experimental and randomized data. Detection results with FGDM and UEDM are respectively showed in left and right part and each row corresponds to a particular $n_{\min}$ value. For each method and for each D value, the detection was processed 1 time on experimental data and 1 time on each one of the 100 randomized data. The number of detected groups is then showed as diamond for experimental data and the distribution of detected groups in randomized data is show with an error bar (lower and upper borders correspond respectively to the min. and max. observed values and the dot corresponds to the median value).

IV. RESULTS

The Table II shows the list of detected groups for D =10 ms and $\beta=99\%$ in the synthetic multivariate presented in section III-A. The FGDM detected three groups (2 true positives + 1 false positive) whereas the UEDM extracted five groups (two true positives + 3 false positives). One way to select only the true positives among detected groups, is to select groups whose the occurrence is higher than a minimum occurrence $n_{\min}\in\{1,2,...\}$. For example, and for D =10 ms, $\beta=99\%$, and $n_{\min}$ =4, the two methods extract only the two true positives.

The two methods were applied on experimental data and randomized experimental data for fifteen values of the coincidence window D=1,2,…,15 ms, five values of the minimum occurrence $n_{\min}$=1,5,10,15,20, and for a same confident probability β=99%. For each method, each D value, and each $n_{\min}$ value, we superimposed the number of detected groups in experimental data (see diamond in Fig. 2) and the distribution of the number of extracted groups in the randomized experimental data (see error bars in Fig. 2).

Left and right plots of Fig. 2-A obtained for $n_{\min} = 1$ show the three following results. First, the number of extracted groups in randomized experimental data (i.e. false positive) is high whatever the method. For example, the number of false positives for D=15 ms in limited to 600 with FGDM and limited to 3000 for UEDM. Second, the number of detected groups increased when the D value increased especially for UEDM. Third, for $D > 4$ ms, the number of detected groups in experimental data is clearly higher than the number of groups detected in randomized experimental data with the functional group detection method whereas these numbers are similar with unitary event detection method.

When $n_{\min}$ increased, results presented in Fig. 2 B-E confirmed the preceding third point. Indeed, in the one hand, with functional group detection method and for $D > 4$ ms, the number of detected groups in experimental data was significantly higher that the numbers of extracted groups in randomized experimental data. On the other hand, with unitary event detection method, the number of extracted groups in experimental and in randomized data was similar. In addition, the number of detected groups decreased strongly both in experimental and randomized experimental data when $n_{\min}$ increased. For $n_{\min}$=5 in the case of EUDM and $n_{\min}$=10 in the case of FGDM, the mean number of false positives detection in randomized data was lower than 10.

V. Discussion

In this paper, we presented two standard methods ([3] and [4]) able to detect groups of neurons with synchronized firings in multivariate spike trains. For the first time, these methods were described in a same formal language. We consider that this description is interesting for the community because it gives a framework for a rapid implementation of the two approaches (especially for Gerstein's approach which was forgotten by the community). In addition, this presentation showed that these methods use the same parameters: a coincidence window value and a confident probability. This property permitted us to objectively compare these methods on experimental and randomized experimental data.

The detection results in synthetic and randomized experimental data show that the two methods detect a large number of false positives characterized by low occurrence. This observation confirms previous studies [8] which showed that this class of methods detects groups highly significant in a statistical view point but which appear only one or two time in the B matrix. In this paper, we show at the application of a minimum occurrence threshold on detected neurons groups decreased advantageously the false positive rate.

The comparison between the numbers of detected neurons groups in experimental and in randomized experimental data with the functional group detection method exhibited a significant excess of detected groups in experimental data whereas this result was not confirmed with unitary event detection method. This result revealed the two following findings. First, the functional group detection method, which works with non-exclusive activations, seems more sensitive in large multivariate spike train that unitary event detection method which works with exclusive activations. Second, the excess of co-activated groups observed in experimental data with functional group detection method has to be correlated with the behavior of the owl monkey. This interesting finding indicates that the hand movement of the owl monkey generates firing synchronizations between neurons distributed among motor, premotor, and parietal cortices. This second point has to be investigated in details in future works.

Acknowledgment

The multielectrode recordings utilized in this paper were collected in the Center for NeuroEngineering at Duke University. The authors extend their gratitude to the Center Director, Dr. Miguel Nicolelis, and Johan Wessberg for sharing the data and for valuable discussions. This work was partially supported by DARPA project # N66001-02-C-8022.

VI. References

[1] M. A. Nicolelis and S. Ribeiro, "Multielectrode recordings: the next steps," *Curr Opin Neurobiol*, vol. 12, pp. 602-6, 2002.

[2] E. N. Brown, R. E. Kass, and P. P. Mitra, "Multiple neural spike train data analysis: state-of-the-art and future challenges," *Nature Neuroscience*, vol. 7, pp. 456-61, 2004.

[3] G. L. Gerstein, D. H. Perkel, and K. N. Subramanian, "Identification of functionally related neural assemblies," *Brain Res*, vol. 140, pp. 43-62, 1978.

[4] S. Grun, A. Aertsen, M. Abeles, G. L. Gerstein, and G. Palm, "Behavior-related neuron group activity in the cortex.," presented at Proc. 17th Ann. Meeting European Neurosci. Assoc., 1994.

[5] D. H. Perkel, G. L. Gerstein, and G. P. Moore, "Neuronal spike trains and stochastic point processes. I. The single spike train," *Biophys J*, vol. 7, pp. 391-417, 1967.

[6] S. Grun, M. Diesmann, and A. Aertsen, "Unitary events in multiple single-neuron spiking activity: I. Detection and significance," *Neural Comput*, vol. 14, pp. 43-80, 2002.

[7] J. Wessberg, C. R. Stambaugh, J. D. Kralik, P. D. Beck, M. Laubach, J. K. Chapin, J. Kim, S. J. Biggs, M. A. Srinivasan, and M. A. Nicolelis, "Real-time prediction of hand trajectory by ensembles of cortical neurons in primates," *Nature*, vol. 408, pp. 361-5, 2000.

[8] A. Roy, P. N. Steinmetz, and E. Niebur, "Rate limitations of unitary event analysis," *Neural Comput*, vol. 12, pp. 2063-82, 2000.