Physica A 392 (2013) 3066–3086

Contents lists available at SciVerse ScienceDirect

Physica A

journal homepage: www.elsevier.com/locate/physa

Statistical modelling of higher-order correlations in pools of neural activity

Fernando Montani[a,b,*], Elena Phoka[e,d], Mariela Portesi[a,c], Simon R. Schultz[e]

[a] Departamento de Física, Facultad de Ciencias Exactas, UNLP Calle 49 y 115. C.C. 67 (1900), La Plata, Argentina
[b] IFLYSIB, CONICET & Universidad Nacional de La Plata, Calle 59-789, 1900 La Plata, Argentina
[c] IFLP, CONICET & Universidad Nacional de La Plata, Calle 49 y 115, 1900 La Plata, Argentina
[d] Department of Mathematics, Imperial College London, South Kensington, London SW7 2AZ, United Kingdom
[e] Department of Bioengineering, Imperial College London, South Kensington, London SW7 2AZ, United Kingdom

HIGHLIGHTS

- We consider recordings from multiple neural units for a large number of neurons.
- We propose an approach within the framework of the extended central limit theorem.
- This allows us to characterise the extent of spike correlations higher than two.
- Infer dynamical properties of the network within pools of multiunit neural activity.
- Higher-order correlations needed to realise a widespread activity distribution.

ARTICLE INFO

Article history:
Received 20 November 2012
Received in revised form 4 February 2013
Available online 22 March 2013

Keywords:
Extended central limit theorem
Large neural ensemble
Multiunit neural activity
Information geometry

ABSTRACT

Simultaneous recordings from multiple neural units allow us to investigate the activity of very large neural ensembles. To understand how large ensembles of neurons process sensory information, it is necessary to develop suitable statistical models to describe the response variability of the recorded spike trains. Using the information geometry framework, it is possible to estimate higher-order correlations by assigning one interaction parameter to each degree of correlation, leading to a $(2^N - 1)$-dimensional model for a population with N neurons. However, this model suffers greatly from a combinatorial explosion, and the number of parameters to be estimated from the available sample size constitutes the main intractability reason of this approach. To quantify the extent of higher than pairwise spike correlations in pools of multiunit activity, we use an information-geometric approach within the framework of the extended central limit theorem considering all possible contributions from higher-order spike correlations. The identification of a deformation parameter allows us to provide a statistical characterisation of the amount of higher-order correlations in the case of a very large neural ensemble, significantly reducing the number of parameters, avoiding the sampling problem, and inferring the underlying dynamical properties of the network within pools of multiunit neural activity.

1. Introduction

To understand how sensory information is processed in the brain, we need to investigate how information is represented collectively by the activity of a population of neurons. There is a large body of evidence suggesting that pairwise correlations

* Corresponding author at: Departamento de Física, Facultad de Ciencias Exactas, UNLP Calle 49 y 115. C.C. 67, 1900 La Plata, Argentina. Fax: +54 221 4252006.
E-mail address: fmontani@gmail.com (F. Montani).

0378-4371/$ – see front matter
http://dx.doi.org/10.1016/j.physa.2013.03.012

are important for information representation or processing in retina [1,2], thalamus [3], cerebral [4–6] and cerebellar cortices [7,8]. However, there is also evidence that in at least some circumstances, pairwise correlations do not by themselves account for multineuronal firing patterns [9,10]; in such circumstances, triplet and higher-order interactions are important. The role of such higher-order interactions in information processing is still to be determined, although they are often interpreted as a signature of the formation of Hebbian cell assemblies [11].

Higher-order correlations may be important for neural coding even if they arise from random fluctuations. Amari and colleagues [12] have suggested that a widespread distribution of neuronal activity can generate higher-order stochastic interactions. In this case, pairwise or even triplet-wise correlations do not uniquely determine synchronised spiking in a population of neurons, and higher-order interactions across neurons cannot be disregarded. Thus, to gain a better understanding of how neural information is processed, we need to study whether higher-order interactions arise from cell assemblies or stochastic fluctuations. Information-geometric measures can be used to analyse neural firing patterns including correlations of all orders across the population of neurons [13–18,9,10].

A straightforward way to investigate the neural activity of a large population of neurons is to use binary maximum entropy models incorporating pairwise correlations on short timescales [1,2]. To estimate this model, one has to consider a sufficient amount of data to measure the mean activity of individual neurons and correlations in pairs of neurons. This allows us to estimate the functional connectivity in a population of neurons at pairwise level. However, if higher-order correlations are present in the data, and as the number of possible binary patterns grows exponentially with the number of neurons, we would need to use an appropriate mathematical approach to go beyond a pairwise modelling. This is in general a difficult problem, as sampling even third-order interactions can be difficult in a real neurophysiological setting [9,10].

However, under particular constraints, this sampling difficulty can be substantially ameliorated. An example of such a constraint is *pooling*, in which the identity of which neuron fires a spike in each pool is disregarded. Such a pooling process reflects the behaviour of a simple integrate-and-fire neuron model in reading out the activity of an ensemble of neurons. It also reflects a common measurement made in systems neuroscience: the recording of multiunit neural activity, without spike sorting. Whether or not such constraints permit a complete description of neural information processing is still a matter of debate [6,10], but they may allow substantial insight to be gained into the mechanisms of information processing in neural circuits. A method for the statistical quantification of correlations higher than two, in the representation of information by a neuronal pool, would be extremely useful, as it would allow the degree of higher-order correlations to be estimated from recordings of multiunit activity (MUA)—which can be performed in a much broader range of circumstances than "spike-sortable" recordings can.

In this paper, we use a pooling assumption to investigate the limit of a very large neural ensemble, within the framework of information geometry [12–20]. In particular, we take advantage of the recent mathematical developments in q-geometry (and q-information geometry) and the extended central limit theorem [21–26] to provide a statistical quantification of higher-order spike correlations. This approach allows us to identify a deformation parameter to characterise the extent of spike correlations higher than two in the limit of a very large neuronal ensemble. Our method accounts for the different regimes of firing within the probability distribution and provides a phenomenological description of the data, inferring the underlying role of noise correlations within pools of multiunit neural activity.

2. Methodology

2.1. Information geometry and the pooled model

We represent neuronal firing in a population of size N by a binary vector $X = (x_1, \ldots, x_N)$, where $x_i = 0$ if neuron i is silent in some time window ΔT and $x_i = 1$ if it is firing (see Fig. 1(a)). Then, for a given time window, we consider the probability distribution of binary vectors, $\{P(X)\}$. Any such probability distribution $\{P(X)\}$ is made up of 2^N probabilities

$$P(X) = Prob\{X_1 = i_1, \ldots, X_N = i_N\} = P_{i_1 \ldots i_N} \tag{1}$$

subject to the normalisation

$$\sum_{i_1,\ldots,i_N=0,1} P_{i_1\ldots i_N} = 1. \tag{2}$$

As proposed by Amari and co-workers (see, for instance, Refs. [20,15]), the set of all the probability distributions $\{P(X)\}$ forms a $(2^N - 1)$-dimensional manifold S_N. This approach uses the orthogonality of the natural and expectation parameters in the exponential family of distributions. It is also useful for analysing neural firing in a systematic manner based on information geometry. Any such probability distribution can be unequivocally determined using a *coordinate system*. One possible coordinate system is given by the set of $2^N - 1$ marginal probability values:

$$\tilde{\eta}_i = E[x_i] = P\{x_i = 1\}, \quad i = 1, \ldots, N \tag{3}$$

$$\tilde{\eta}_{ij} = E[x_i x_j] = P\{x_i = x_j = 1\}, \quad i < j \tag{4}$$

$$\vdots$$

$$\tilde{\eta}_{123\ldots N} = E[x_1 \ldots x_N] = P\{x_1 = x_2 = \cdots = x_N = 1\}. \tag{5}$$