



Quantifying higher-order correlations in a neuronal pool



Lisandro Montangie and Fernando Montani*

IFLYSIB, CONICET & Universidad Nacional de La Plata, Calle 59-789, 1900 La Plata, Argentina

Departamento de Física, Facultad de Ciencias Exactas, UNLP Calle 49 y 115. C.C. 67 (1900), La Plata, Argentina

HIGHLIGHTS

- Population of neurons has shown significant amount of higher-order correlations.
- We account for beyond second order inputs correlations seen by each neuron.
- We obtain an exact analytical expression for the joint distribution of firing.
- This method allows us to characterize higher-order correlations in a neuronal pool.
- Input nonlinearities can enhance coding performance by neural populations.

ARTICLE INFO

Article history:

Received 3 July 2014

Available online 25 November 2014

Keywords:

Higher order correlations
Extended central limit theorem
Large neural ensemble
Information geometry

ABSTRACT

Recent experiments involving a relatively large population of neurons have shown a very significant amount of higher-order correlations. However, little is known of how these affect the integration and firing behavior of a population of neurons beyond the second order statistics. To investigate how higher-order inputs statistics can shape beyond pairwise spike correlations and affect information coding in the brain, we consider a neuronal pool where each neuron fires stochastically. We develop a simple mathematically tractable model that makes it feasible to account for higher-order spike correlations in a neuronal pool with highly interconnected common inputs beyond second order statistics. In our model, correlations between neurons appear from q -Gaussian inputs into threshold neurons. The approach constitutes the natural extension of the Dichotomized Gaussian model, where the inputs to the model are just Gaussian distributed and therefore have no input interactions beyond second order. We obtain an exact analytical expression for the joint distribution of firing, quantifying the degree of higher-order spike correlations, truly emphasizing the functional aspects of higher-order statistics, as we account for beyond second order inputs correlations seen by each neuron within the pool. We determine how higher-order correlations depend on the interaction structure of the input, showing that the joint distribution of firing is skewed as the parameter q increases inducing larger excursions of synchronized spikes. We show how input nonlinearities can shape higher-order correlations and enhance coding performance by neural populations.

© 2014 Elsevier B.V. All rights reserved.

1. Introduction

Neurons in the cortex receive 3000–10,000 synaptic inputs, 85% of which are excitatory. Nearly half of the excitatory inputs to any one neuron come from nearby neurons that fall within a cylinder of 100–200 μm radius, arranged as a column, sometimes termed a mini-column [1–4]. This suggests that cortical neurons receive abundant excitatory inputs and

* Corresponding author at: IFLYSIB, CONICET & Universidad Nacional de La Plata, Calle 59-789, 1900 La Plata, Argentina.

E-mail address: fmontani@gmail.com (F. Montani).

are embedded in a network of highly convergent signals. These networks have a recurrent nature, thus it is likely that those neurons receive similar inputs and emit spikes under similar conditions. This means that the conditions that lead to a response of any one neuron in the mini-column are likely to involve considerable activity from a large number of its inputs beyond second order statistics. It is therefore reasonable to expect that many spike inputs will arrive in synchrony within a very small time window.

The integration of features into gestalt entities [5–8] is one of the most important challenges in cognition. It has been proposed that correlated activity within the millisecond time range may be the signature of neuronal assembly formation. If this is the case, it may be essential in the context of multiple object encoding. According to the temporal binding hypothesis of von der Malsburg, cells belonging to the same assembly fire action potentials synchronously with a precision of a few milliseconds, and cells belonging to different assemblies fire asynchronously. This hypothesis requires cortical neurons to act as coincidence detectors [9–11]. In agreement with the latter, the major causes of correlated firing in neural networks are common presynaptic input. Behavior then stems from the emergent properties of a large set of neurons with overlapping neural circuits that share common dynamical inputs. A primary challenge in theoretical neuroscience is to gain further understanding of circuit dynamics incorporating the neuronal activity at a variety of spatial and temporal scales. Moreover, temporal dynamics and plasticity encode information about the outside world. Identification of relevant neural ensembles underlying cognitive behavior thus requires new modeling techniques and theoretical frameworks. Approaches that may help to link the multiple spatial, temporal, and organizational scales of neuronal assemblies could provide important insights into the emergent properties of the neural network, as they may lead to new discoveries concerning neural circuitry that could eventually shape the biophysical bases of behavior.

Information processing in the brain is usually encoded in the activity of large and highly interconnected neural populations. It has been proposed that synapses between neurons that fire synchronously are strengthened, forming cell assemblies and phase sequences. At short scales, it is expected that cell assemblies would affect information processing while at longer scales they could shape behavior and perception. Neuronal cells synchronize through correlated input, and spike synchronization between neurons emerges as a result of transient activity. Approaches using binary maximum entropy models at a pairwise level have been developed considering a very large number of neurons on short time scales [12–14]. These models can capture essential structures of the neural population activity, however, due to their pairwise nature their generality has been subject to debate [15–17]. In particular, E. Ohiorhenuan and J. D. Victor have shown the importance of triplets of spikes to characterize scale dependence in cortical networks [17,18]. That is to say, although models accounting for pairwise interactions have proved able to capture some of the most important features of population activity at the level of the retina [12,13], pairwise models are not enough to provide reliable descriptions of neural systems in general, as experiments considering a relatively large population of neurons have displayed a very significant amount of higher-order correlations ('HOCs') [15–19].

More specifically, neurophysiological research has shown that pairwise models fail to explain the responses of spatially localized triplets of cells [17–20], along with describing the activity of large neuronal populations responding to natural stimuli [19]. Deviations from the Maximum Entropy model indicate that HOCs have to be taken into account for modeling the population statistics [21–24]. Thus, the intricacy of the neurophysiological data highlights the need to develop a theoretical framework accounting for the statistical complexity of synchronous activity patterns. Pattern probabilities for the so-called Dichotomized Gaussian ('DG') model [20–24] were estimated using the cumulative distribution of multivariate Gaussians showing high precision fitting of the experimental data.

In this paper, we provide a simple mathematically tractable model able to account for HOCs in the joint firing distribution of a neuronal population. In our model, correlations between neurons arise from q -Gaussian inputs into threshold neurons. It is therefore an extension of the DG model proposed by Amari [21], where the inputs to the model are Gaussian distributed and therefore have no interactions beyond second order. Our current theoretical formalism relies on recent progress made on the Extended Central Limit Theorem ('ECLT'), and thus using mathematical tools of non-extensive statistical mechanics [25–33], we provide an approach that quantifies the degree of HOCs. We present the exact analytical solution of the joint distribution of firing including neural correlation patterns of all orders across a population. That is, we estimate by means of an analytically solvable model the amount of correlations of order higher than two in a neuronal pool through direct application of a q -Gaussian distribution of common synaptic inputs providing the expression of the joint distribution, Tsallis Relative Entropy and Fisher Information. We test the robustness of our approach using a set of simulated independent and correlated neurons. Using our model, we investigate different analytical solutions when considering three typical distributions: concentrated, widely spread, and bimodal. We study the emergent properties of the Fisher information in a large neural population, and show their impact on the efficiency of population coding. Our approach allows us to investigate how input nonlinearities can shape HOCs and improve information transmission. This could be a useful tool for understanding how groups of neurons could integrate into unique functional cell assemblies.

2. Methodology

2.1. Higher order interactions in the pooled model

We represent the neuronal firing in a population of size N by a binary vector $\mathbf{x} = (x_1, \dots, x_N)$, where $x_i = 0$ if neuron i is silent in some time window ΔT and $x_i = 1$ if it is firing a spike. We consider the probability distribution of those binary

vectors, $\{P(\mathbf{x})\}$ that consists of 2^N probabilities

$$P(\mathbf{x}) = \text{Prob}\{x_1 = i_1, \dots, x_N = i_N\} = P_{i_1 \dots i_N} \tag{1}$$

subject to the normalization

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

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 [29,34]:

$$\begin{aligned} \eta_i &= E[x_i] = P\{x_i = 1\}, \quad i = 1, \dots, N \\ \eta_{ij} &= E[x_i x_j] = P\{x_i = x_j = 1\}, \quad i < j \\ &\vdots \\ \eta_{123 \dots N} &= E[x_1 \dots x_N] = P\{x_1 = x_2 = \dots = x_N = 1\}. \end{aligned}$$

These are called the η -coordinates [29]. Moreover, provided $P(\mathbf{x}) \neq 0$, any such distribution can be expanded as in Ref. [34]

$$P(\mathbf{x}) = \exp \left\{ \sum_{i=1}^N x_i \theta_i + \sum_{i < j} x_i x_j \theta_{ij} + \dots + x_1 \dots x_N \theta_{1 \dots N} - \psi \right\}, \tag{3}$$

where there are in total $2^N - 1$ different θ correlation coefficients that can be used to determine univocally the probability distribution. It is important to note that the estimation of all the parameters associated with HOCs suffers greatly from a combinatorial explosion [35].

It is said that we reach the ‘thermodynamic limit’ when the number of particles being considered reaches the limit $N \rightarrow \infty$ and the volume of the system also grows in proportion to the number of particles. The thermodynamic limit is asymptotically approximated in statistical mechanics using the Central Limit Theorem (‘CLT’) [36]. The CLT ensures that the probability distribution function of any measurable quantity is a normal Gaussian distribution, provided that a sufficiently large number of independent random variables with exactly the same mean and variance are being considered (see pages 324–330 [36]). Importantly, the CLT does not hold if correlations between random variables cannot be neglected.

In particular, the CLT has been used by Amari and colleagues [21] in order to obtain the DG model of correlated inputs to estimate the joint probability distribution of firing in a neuronal pool, considering the limit of a very large number of neurons. In their approach, pairwise correlations are quantified through the covariance $\langle u_i u_j \rangle$ of the weighted sum of Gaussian (due to the CLT [21]) inputs u_i and u_j of two given pairs of neurons ($i \neq j, i = 1 \dots N$ and $j = 1 \dots N$) [21]. That is $u_i = \sum_{j=1}^m w_{ij} - h$, where w_{ij} is the connection weight from the j th input to the i th neuron ($h = E[u_i]$ denotes the mean).

However, in recent publications by M Gell-Mann, C Tsallis, S Umarov, C Vignat, A Plastino (see: [26–28,32,33]), the CLT has been generalized for the case when a system with weakly or strongly correlated random variables is being considered. They have proved that if we gather a sufficiently large number of such systems together, the probability distribution will converge to a q -Gaussian distribution. This is in agreement with the theorems recently proved by Amari and Ohara [37], which allow the introduction of the q -geometrical structure to any arbitrary family of probability distributions, and guarantee that the family of all the probability distributions belongs to the q -exponential family of distributions.

We will use the ‘natural extension’ of the central limit theorem (ECLT) proposed in Ref. [26], which accounts for cases in which correlations between random variables are non-negligible. This results in so-called q -Gaussians (instead of Gaussians) as the probability density functions in the ECLT, as proved in Ref. [26]:

$$G_q(x) = \begin{cases} \left[1 + \frac{(1-q)(-x^2)}{2} \right]^{\frac{1}{1-q}} & \text{if } \left(1 + \frac{(1-q)(-x^2)}{2} \right) > 0, \\ 0 & \text{otherwise} \end{cases} \tag{4}$$

where q is a (problem-dependent) positive real index. Notice that in the limit of $q = 1$ a normal Gaussian distribution is recovered as $\lim_{N \rightarrow \infty} (1 + \frac{1}{N})^N = e$, which can be rewritten as $\lim_{q \rightarrow 1} (1 + (1-q))^{1/(1-q)} = e$. In other words, the CLT is recovered as $q \rightarrow 1$ [26–28,32,33].

Let us now consider, as in Refs. [21,35], the probability of exactly $k = N \cdot r$ (and thus $r = \frac{k}{N}$, with r being the population firing rate) neurons firing within a given time window ΔT across a population of N neurons. In the framework of the pooled model we have that

$$P_r \left[r = \frac{k}{N} \right] = P_r \{x_1 = x_2 = \dots = x_k = 1, x_{k+1} = \dots = x_N = 0\}, \tag{5}$$

where neuron i is subject to a weighted sum of inputs u_i , thus $x_i = 1$ if and only if $u_i > 0$ and $x_i = 0$ if $u_i \leq 0$. Following [35], the neuronal pool receives higher than pairwise-correlated common inputs s_1, s_2, \dots, s_M , and u_i is weighted by the common

inputs $u_i = \sum_{j=1}^M w_{ij}s_j - h$, where w_{ij} are randomly assigned connections weights. Considering that the u_i are subject to a q -Gaussian distribution (due to the ECLT [26–28,32,33]) $N_q(-h, 1)$, we define in analogy to Ref. [21] $u_i = \sqrt{1-\alpha} v_i + \sqrt{\alpha} \varepsilon - h$, for $i = 1, \dots, N$. We take $\alpha = E_q[u_i u_j]$ as a q -variance, $h = E_q[u_i]$ as the q -mean, and two independent q -Gaussian random variables v_i and ε subject to $N_q(0, 1)$ (see Ref. [38] for a detailed description of q -Gaussian random variables).

The joint firing distribution can therefore be estimated using the saddle point approximation as in Ref. [35]:

$$Q_q(r) = \sqrt{\frac{1}{r(1-r)|z_q''(\varepsilon_0)|}} \frac{1}{\sqrt{2\pi}} \exp \left[Nz_q(\varepsilon_0) - \frac{\varepsilon_0^2}{2} \right], \tag{6}$$

where $z_q(\varepsilon) = r \log\left(\frac{F_q(\varepsilon)}{r}\right) + (1-r) \log\left(\frac{1-F_q(\varepsilon)}{1-r}\right)$. Within the saddle point approximation [21,35,39]: $\varepsilon_0 = \arg \max_{\varepsilon \in \mathbb{R}} [z_q(\varepsilon)]$ and $\frac{dz(\varepsilon)}{d\varepsilon} = 0$. The solution is $\varepsilon_0 = F_q^{-1}(r)$, which implies $r = F_q(\varepsilon_0)$, where r goes between $[0, 1]$ and ε_0 is defined for all real numbers. Additionally, $\varepsilon_0 = F_q^{-1}(r)$ depends on the degree of correlation of the network architecture, which is quantified by q .

Notice that the functions $F_q(\varepsilon)$ and $z_q(\varepsilon)$ are defined as in Ref. [35]:

$$\begin{aligned} F_q(\varepsilon) &\equiv P_r(u > 0|\varepsilon) = P_r\left(u_i > \frac{h - \sqrt{\alpha}\varepsilon}{\sqrt{1-\alpha}}\right) \\ &= \frac{1}{\sqrt{2\pi}} \int_{\frac{h-\sqrt{\alpha}\varepsilon}{\sqrt{1-\alpha}}}^{\infty} \exp_q\left(-\frac{v^2}{2}\right) dv \end{aligned} \tag{7}$$

and

$$z_q(\varepsilon) = r \ln \left[\frac{F_q(\varepsilon)}{r} \right] + (1-r) \ln \left[\frac{1-F_q(\varepsilon)}{1-r} \right]. \tag{8}$$

After some algebra Eq. (7) reads as [35]:

$$F_q(\varepsilon) = \frac{1}{2\sqrt{2\pi}\sqrt{\frac{q-1}{2}}} B\left(\frac{1}{1+\xi_0(\varepsilon)}; \frac{1}{q-1}, \frac{1}{2}\right) \tag{9}$$

where $B(\cdot)$ is the Beta function and

$$\xi_0(\varepsilon) = \frac{(q-1)(h-\sqrt{\alpha}\varepsilon)^2}{2(1-\alpha)}. \tag{10}$$

If we consider the limit of the CLT framework ($q = 1$), Eq. (9) reduces to

$$F_{q=1}(\varepsilon) = \frac{1}{2} \text{Erfc}\left(\frac{1}{\sqrt{2}} \frac{h - \sqrt{\alpha}\varepsilon}{\sqrt{1-\alpha}}\right) \tag{11}$$

where $\text{Erfc}(x) = \frac{2}{\sqrt{\pi}} \int_x^{\infty} \exp(-t^2) dt$ denotes the complementary error function. However, if the effect of correlations of order higher than two is not negligible, then, according to the ECLT, q must be higher than 1 [35]. One can test for the presence of HOCs by measuring the distribution of activity in multi-unit recordings and fitting q , which represents the amount of higher-order correlations present in the distribution of firing. One can show by simple comparison how statistically different from the $q = 1$ case the measured distribution is [35].

To this end, we consider the Kullback–Leibler divergence, which constitutes an Information Theory quantity and is a measure of the ‘distance’ between two probability distributions:

$$D(P|P_0) = P(x) \log \frac{P(x)}{P_0(x)}. \tag{12}$$

A generalization of the Kullback–Leibler entropy, in the framework of the non-extensive thermodynamics, is the Tsallis relative entropy or q -relative entropy that is better suited to deal with non-Gaussian distributions. The relative entropy is a measure of the inefficiency of assuming that the distribution is P_0 when the true distribution is P [40–42]:

$$\begin{aligned} T_q(P, P_0) &\equiv \int P(x) \frac{[P(x)/P_0(x)]^{1-q} - 1}{1-q} dx \\ &= \frac{1}{q-1} \left[\int P(x) \left(\frac{P(x)}{P_0(x)}\right)^{q-1} dx - 1 \right] \\ &= \frac{1}{q-1} \left[\int P(x)^q P_0(x)^{1-q} dx - 1 \right]. \end{aligned} \tag{13}$$

A measure that is also particularly interesting in this context to test the applicability of Eq. (6), since it was used in Ref. [43] to give an information theoretic proof of the CLT, is the Fisher information I . This quantity [44,45] constitutes a measure of the gradient content of a distribution $P(x)$:

$$I(P) = \int \frac{|\vec{\nabla}P(x)|^2}{P(x)} dx, \quad (14)$$

and it is therefore quite sensitive even to small localized perturbations. The Fisher information can be variously interpreted as a scope of the ability to estimate a parameter, as the amount of information that can be extracted from a set of measurements, and also as a measure of the state of disorder of a system or phenomenon [45,46], its foremost property being the so-called Cramer–Rao bound. It is important to remark that the gradient operator significantly influences the contribution of minute local p -variations to the Fisher information value, so that the quantifier is referred as a ‘local’ one. Local sensitivity is useful in scenarios whose description needs to appeal to a notion of ‘order’ [47–49]. Fisher information is also useful for detecting dynamical changes in a probability density function (i.e. a sharper probability distribution function would tend to have higher Fisher information than a more widespread probability density function).

In the next section, we explicitly estimate the exact analytical expression of the probability distribution of firing extending the current formalism by means of our mathematically tractable model. We then investigate the analytical solutions of three most representative types of joint firing distributions: concentrated, widely spread, and bimodal. Then, we show the robustness of our method using two simulated set of neurons: the first one assuming independence across the cells and the second considering a set of correlated neurons as random binary vectors with specified correlations generated using the DG distribution [21–23]. Finally, we calculate the Tsallis relative entropy and Fisher information investigating the emergent properties of the system as the correlation degree q increases.

3. Results

Understanding brain functions requires interdisciplinary approaches involving many levels of study: from the molecular level through the cellular level (individual neurons), to the level of relatively small assemblies of neurons (for example, cortical columns); to the case of larger subsystems, as the one which subserves visual perception, and up to the level of large systems, including the cerebral cortex, the cerebellum and to nervous system as a whole. Theoretical neuroscience encompasses approaches ranging all levels, and together with mathematical modeling they are important tools for characterizing what nervous systems do, determining how they function and understanding why they operate in particular ways. Neurons interact through different dynamical pathways in the complex architecture of the brain connectivity. It may not be preposterous therefore to think of synchrony as a mechanism for reliable signal transmission that extracts HOCs as a gestalt rather than as an active binding mechanism that represents a secondary code to link a system of simple feature extraction. When considering high-dimensional data of large number of cells the joint distribution of firing accounts for the common overlapping inputs that neurons receive due to the network interactions. In the following, we will obtain the exact analytical expression of this distribution accounting for HOCs of all orders. In order to do so, let us first take the derivative of $F_q(\varepsilon)$ with respect to ε . Using the chain rule for the derivative, we can write

$$F'_q(\varepsilon) = \frac{1}{2\sqrt{2\pi}\sqrt{\frac{q-1}{2}}} \frac{dB\left(\frac{1}{1+\xi_0(\varepsilon)}; \frac{1}{q-1}, \frac{1}{2}\right)}{d\left(\frac{1}{1+\xi_0(\varepsilon)}\right)} \frac{d\left(\frac{1}{1+\xi_0(\varepsilon)}\right)}{d\varepsilon}. \quad (15)$$

Considering that

$$\frac{d}{dz} B(z; a, b) = (1-z)^{b-1} z^{a-1}, \quad (16)$$

then

$$\frac{dB\left(\frac{1}{1+\xi_0(\varepsilon)}; \frac{1}{q-1}, \frac{1}{2}\right)}{d\left(\frac{1}{1+\xi_0(\varepsilon)}\right)} = \left[1 - \frac{1}{1+\xi_0(\varepsilon)}\right]^{-\frac{1}{2}} \left[\frac{1}{1+\xi_0(\varepsilon)}\right]^{\left(\frac{1}{q-1}-1\right)} \quad (17)$$

and

$$\frac{d\left(\frac{1}{1+\xi_0(\varepsilon)}\right)}{d\varepsilon} = \frac{(q-1)\sqrt{\alpha} (h - \sqrt{\alpha}\varepsilon)}{(1-\alpha) [1+\xi_0(\varepsilon)]^2}. \quad (18)$$

Thus we can rewrite

$$F'_q(\varepsilon) = \frac{1}{\sqrt{2\pi}} \sqrt{\frac{\alpha}{1-\alpha}} [1+\xi_0(\varepsilon)]^{\left(\frac{q-2}{q-1}-\frac{3}{2}\right)}. \quad (19)$$

Taking the first derivative of $z_q(\varepsilon)$ with respect to ε ,

$$z'_q(\varepsilon) = \left\{ \frac{r}{F_q(\varepsilon)} - \frac{1-r}{1-F_q(\varepsilon)} \right\} \tag{20}$$

and the second derivative we obtain

$$z''_q(\varepsilon) = - \left\{ \frac{r}{F_q^2(\varepsilon)} + \frac{1-r}{[1-F_q(\varepsilon)]^2} \right\} [F'_q(\varepsilon)]^2 + \left\{ \frac{r}{F_q(\varepsilon)} - \frac{1-r}{1-F_q(\varepsilon)} \right\} F''_q(\varepsilon). \tag{21}$$

As we are working within the saddle-point approximation [21,35,39], one can define a parameter ε_0 that maximizes $z_q(\varepsilon)$, and thus [35],

$$\varepsilon = \varepsilon_0 \implies z'_q(\varepsilon_0) = 0 \wedge F_q(\varepsilon_0) = r. \tag{22}$$

Then,

$$\begin{aligned} z''_q(\varepsilon_0) &= - \left[\frac{1}{r} - \frac{1}{1-r} \right] [F'_q(\varepsilon_0)]^2 \\ &= - \frac{1}{r(1-r)} \frac{\alpha}{2\pi(1-\alpha)} [1 + \xi_0(\varepsilon_0)]^{2\left(\frac{q-2}{q-1}\right)-3}. \end{aligned} \tag{23}$$

The joint firing distribution can therefore be calculated as

$$Q_q(r) \simeq \sqrt{\frac{1-\alpha}{\alpha}} \left\{ \exp_q \left[-\frac{(h - \sqrt{\alpha}F_q^{-1}(r))^2}{2(1-\alpha)} \right] \right\}^{-\frac{(q+1)}{2}} \exp \left[-\frac{(F_q^{-1}(r))^2}{2} \right], \tag{24}$$

where

$$F_q^{-1}(r) = \frac{1}{\sqrt{\alpha}} \left\{ h - \sqrt{\frac{2(1-\alpha)}{q-1}} \left[B_{inverse}^{-1} \left(r; \frac{1}{q-1}, \frac{1}{2} \right) - 1 \right] \right\}. \tag{25}$$

Finally,

$$Q_q(r) \simeq \sqrt{\frac{1-\alpha}{\alpha}} \left[B_{inverse} \left(r; \frac{1}{q-1}, \frac{1}{2} \right) \right]^{-\frac{(q+1)}{2(q-1)}} \exp \left[-\frac{(F_q^{-1}(r))^2}{2} \right]. \tag{26}$$

Equivalently, the equation above can be expressed in terms of q -exponentials as

$$\begin{aligned} Q_q(r) &\simeq \sqrt{\frac{1-\alpha}{\alpha}} \left\{ \exp_q \left[B_{inverse}^{-1} \left(r; \frac{1}{q-1}, \frac{1}{2} \right) - 1 \right] \right\}^{-\frac{(q+1)}{2}} \\ &\cdot \exp \left[-\frac{1}{2\alpha} \left\{ h - \sqrt{\frac{2(1-\alpha)}{q-1}} \left[B_{inverse}^{-1} \left(r; \frac{1}{q-1}, \frac{1}{2} \right) - 1 \right] \right\}^2 \right]. \end{aligned} \tag{27}$$

In the limit of $q \rightarrow 1$,

$$\lim_{q \rightarrow 1} Q_q(r) = \sqrt{\frac{1-\alpha}{\alpha}} \lim_{q \rightarrow 1} \left\{ \exp_q \left[-\frac{(h - \sqrt{\alpha}F_q^{-1}(r))^2}{2(1-\alpha)} \right] \right\}^{-\frac{(q+1)}{2}} \cdot \exp \left[-\frac{(F_q^{-1}(r))^2}{2} \right], \tag{28}$$

considering

$$\lim_{q \rightarrow 1} F_q^{-1}(r) = F^{-1}(r) \tag{29}$$

and

$$\lim_{q \rightarrow 1} \exp_q(x) = \exp(x) \tag{30}$$

we obtain

$$\lim_{q \rightarrow 1} Q_q(r) = \sqrt{\frac{1-\alpha}{\alpha}} \exp \left[\frac{(h - \sqrt{\alpha}F^{-1}(r))^2}{2(1-\alpha)} - \frac{(F^{-1}(r))^2}{2} \right] = Q_1(r), \tag{31}$$

which is in agreement with the findings of Amari et al. [21].

Thus, by means of an analytical model we can test for the presence of HOCs by measuring the distribution of activity, and then fitting q . When the correlations in the common presynaptic input are just pairwise, q tends to 1 and the expression reduces to the standard Amari's Eq. (31) formula. However, when the correlations in a system are strong, q becomes greater than 1 to 'bias' the probabilities of certain microstates occurring, accounting therefore for HOCs in Eq. (27). That is, the parameter q , is therefore a way of characterizing the degree of correlation of the network architecture within $Q_q(r)$, and particularly how strong they are. Thus, estimating the mean h and the variance α of a distribution and the degree of HOCs q we can provide the joint probability distribution of firing $Q_q(r)$.

We compute the relative entropy using a Gaussian distribution as the reference probability density function. It is positive for the case in which a probability distribution is described by an encoding optimal distribution Q_q instead of the reference distribution Q_1 . In the limit of $q \rightarrow 1$, the Tsallis relative entropy becomes the Kullback–Leibler divergence. Both relative entropies are not symmetric, i.e., $T_q(Q_q, Q_1) \neq T_q(Q_1, Q_q)$. This measure allows us to quantitatively compare the extended activity distribution Q_q with the reference one (Q_1).

We calculate the analytical expression of the Fisher information for the extended joint probability distribution as

$$\begin{aligned}
 I(Q_q) &= \int_0^1 \frac{\left[\frac{\partial Q_q(r)}{\partial r} \right]^2}{Q_q(r)} dr \\
 &= \int_0^1 Q_q(r) \left\{ F_q^{-1}(r) + \frac{\sqrt{\alpha(q+1)}}{2(1-\alpha)} \frac{(h - \sqrt{\alpha} F_q^{-1}(r))}{\left[1 + \frac{(q-1)}{2(1-\alpha)} (h - \sqrt{\alpha} F_q^{-1}(r))^2 \right]} \right\}^2 \left[\frac{\partial F_q^{-1}(r)}{\partial r} \right]^2 dr.
 \end{aligned} \tag{32}$$

Considering that the derivative of the $B_{inverse}(z; a, b)$ is

$$\frac{\partial B_{inverse}(z; a, b)}{\partial z} = (1 - B_{inverse}(z; a, b))^{1-b} B_{inverse}(z; a, b)^{1-a} B(a, b), \tag{33}$$

we obtain, after some calculations,

$$\begin{aligned}
 I(Q_q) &= \int_0^1 \sqrt{\frac{1-\alpha}{\alpha}} \left(\frac{1-\alpha}{2\alpha(q-1)} \right) B\left(\frac{1}{q-1}, \frac{1}{2} \right) \left\{ \frac{\sqrt{1-\beta} [\beta]^{\left(\frac{q-2}{q-1}\right)}}{\beta^3 - \beta^4} \right\} \\
 &\cdot \left\{ \sqrt{\frac{(q+1)^2}{2(q-1)} \frac{\alpha}{(1-\alpha)} \beta(1-\beta)} + \frac{1}{\sqrt{\alpha}} \left\{ h - \sqrt{\frac{2(1-\alpha)}{q-1}} [\beta^{-1} - 1] \right\} \right\}^2 \\
 &\cdot \left\{ \exp_q [\beta^{-1} - 1] \right\}^{-\frac{(q+1)}{2}} \exp \left[-\frac{1}{2\alpha} \left\{ h - \sqrt{\frac{2(1-\alpha)}{q-1}} [\beta^{-1} - 1] \right\} \right] dr
 \end{aligned} \tag{34}$$

where $\beta = B_{inverse}\left(r; \frac{1}{q-1}, \frac{1}{2}\right)$.

In the following, we provide analytical estimates of how the joint probability behaves for different sets of h , α and q . We will consider three typical distributions as function of r : concentrated, widely spread, and bimodal. Fig. 1 shows a concentrated distribution. This concentration is caused by a very small amount of variance $\alpha = 0.01$ and such peak is moved towards a higher amount of synchronized neurons when q increases. Additionally, Fig. 2 shows a widely spread joint distribution considering $\alpha = 0.2$. Notice that when q is increased the probability of having a higher amount of synchronized spikes grows. As q grows the peak of the distribution moves to the right resulting in a larger number of synchronized neurons. Finally, Fig. 3 shows a bimodal distribution considering $\alpha = 0.8$. In such case the neurons fire synchronously at one time and are quiescent at other times when q is close to 1. By adding more correlations through q , the quiescent proportion of neurons become active leading to a larger amount of synchronized spikes at different times and at any correlation order. This could reflect the impact of noise correlation in the brain, and it provides us with an important feedback: small, perhaps undetectable, higher-order input correlations may have an important effect at the population level. Thus, our analytical expression allows us to assess how HOCs depend on the interaction structure of the input showing that the joint distribution of firing rates is more skewed as q becomes greater, inducing large excursions of synchronized spikes.

Let us now consider a population of simulated correlated binary spike trains using the model of Macke et al. [22,23] where the inputs are modeled by a correlated Gaussian with mean γ and covariance Λ , which are chosen such that the outputs \mathbf{x} have mean μ , covariance σ and a correlation coefficient $\rho = \sigma / [\mu(1 - \mu)]$ [22,23]. Here, a neuron is said to spike ($x_i = 1$) if its input is positive, and to be silent ($x_i = 0$) otherwise. By symmetry, all activity patterns with the same number of spikes are equally likely, and thus the model is fully specified by the distribution over spike counts $N_{sp} = \sum_i x_i$. The independent distribution of firing is generated by assuming that $P(\mathbf{x}) = \prod_i P(x_i = 1)$. Fig. 4 shows the joint probability distribution, considering the previous model with a correlation coefficient $\rho = 0.2852$ and the independent case $\rho = 0$. That is, we use the estimated q and the parameters α and h that give the best fit for Eq. (27). The optimization fitting criterion is the normalized

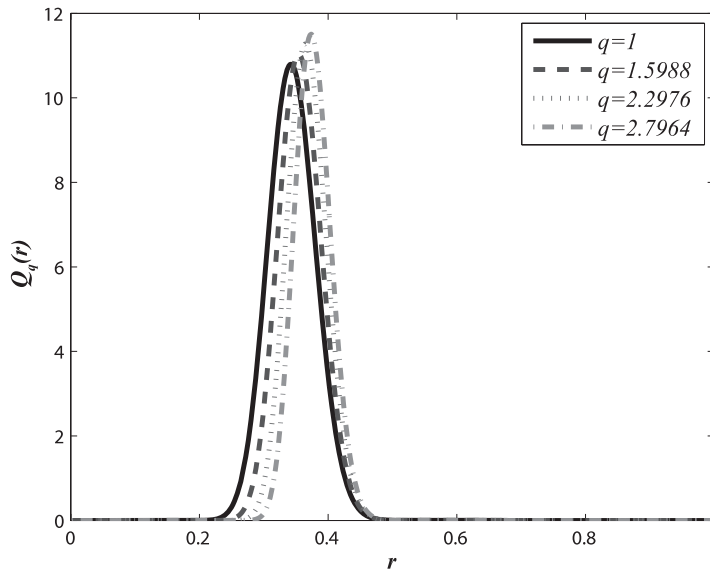


Fig. 1. Concentrated distribution of firing $Q_q(r)$ for various q , with fixed $\alpha = 0.01$ and $h = 0.4$. The peaks are moved towards a higher amount of synchronized neurons when q increases.

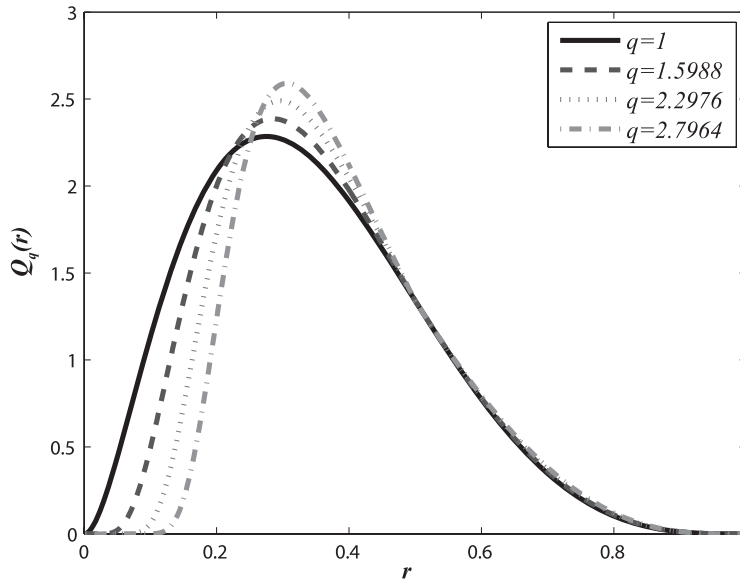


Fig. 2. Widely spread joint distribution of firing rates $Q_q(r)$ for various q , with fixed $\alpha = 0.2$ and $h = 0.4$. The peaks of the distribution moves to the right as q increases resulting in a larger number of synchronized neurons.

mean squared error (NMSE), and the default error value is lower than 0.05 (p -value < 0.05). We then fit the parameter q to find the best-fitting function $Q_q(r)$ in Eq. (27). The DG distribution with $\rho = 0.2852$ corresponds to a value of $q = 1.3788$, and if we test the hypothesis of absence of correlations with $\rho = 0$, the best fitting distribution $Q_q(r)$ corresponds to $q = 1$ with a quite small $\alpha = 0.019$. Thus, in this case, the pairwise model converges to the independent distribution.

Fig. 5 shows Tsallis relative entropy as a function of the correlation degree q . Notice that the function grows as the degree of correlation becomes higher, quantifying the inefficiency of incorrectly assuming that the distribution is the one that corresponds to $Q_1(r)$ when the true distribution is $Q_q(r)$. Meaning that the amount of information in a population cannot be computed without knowing the correlational structure. If our system is in a very ordered state and thus is represented by a very narrow probability density function, we have a maximal Fisher information measure. On the other hand, when the system under study lies in a very disordered state one gets an almost flat probability density function and Fisher information is closer to zero [50]. In the following, we take advantage of our proposed analytical solution for the joint probability distribution considering neural firing correlation of all orders to investigate how information behaves as noise correlation grows. Fig. 6 corresponds to a plot of Fisher information versus q . We show that it is possible to quantify the optimal amount of q

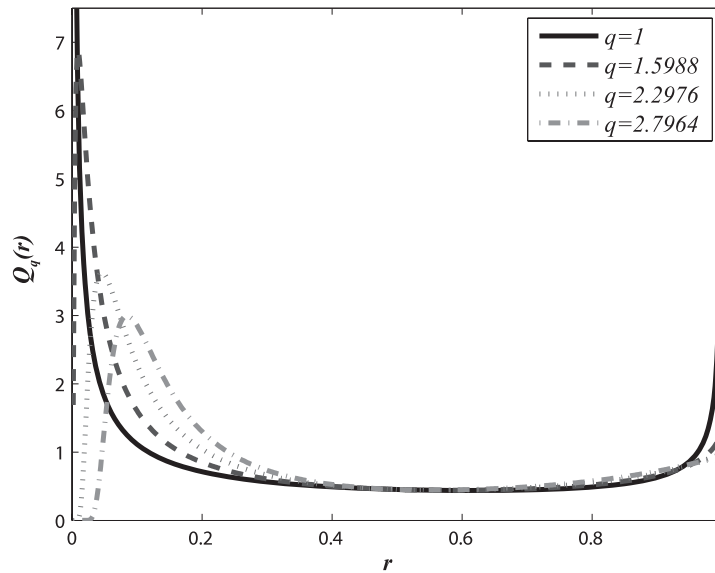


Fig. 3. Bimodal joint distribution $Q_q(r)$ for various q , with fixed $\alpha = 0.8$ and $h = 0.4$. By adding more correlations through q the quiescent proportion of neurons become active leading to a larger amount of synchronized neurons.

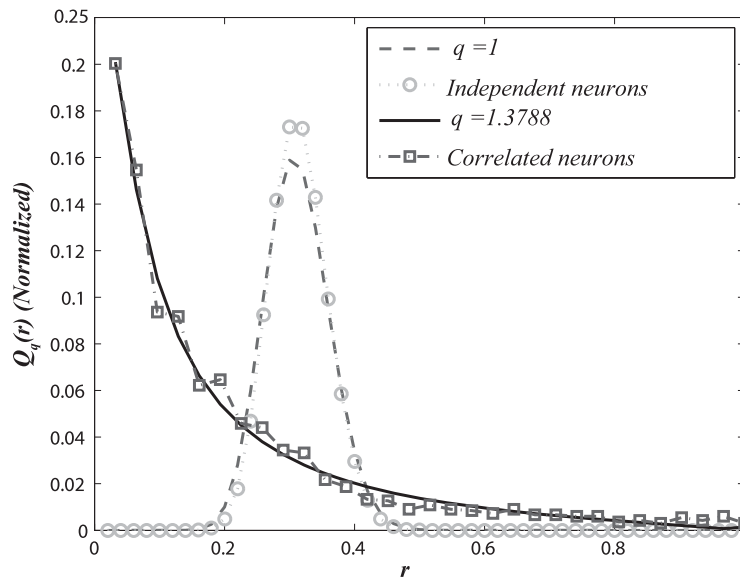


Fig. 4. Normalized joint firing distribution $Q_q(r)$, considering a model binary neural population patterns using simulated firing pattern based on a DG distribution. Correlated neurons correspond to the black squares joined by dashed lines: $\rho = 0.2852$ and $q = 1.3788$ ($\alpha = 0.48$); Independent case, gray circles joined by dashed lines, $\rho = 0$ and $q = 1$ ($\alpha = 0.019$).

that maximizes information. Coding performance could be improved by shaping the output via input HOCs and their non-linearities. Moreover, information saturates as the level of noise correlations increases. This is in agreement with the general intuition needed to understand how correlations may affect information in a population code: for positive correlations, information saturates as the number of correlated neurons increases and they have a large effect at the population level [51].

4. Discussion and conclusions

The brain is a dynamical system whose state variables encode information about the external world. Thus, the central assumption of theoretical neuroscience is that the brain computes; in short, computation equals coding plus dynamics. The detection of subtle changes in brain activity is therefore of importance to investigate the dynamics of functional interactions across neurons. Some neuroscientists attempt to characterize how these dynamical variables evolve with time. Others study the way that information is encoded in neural activity and diverse dynamical variables of the brain. However, building an

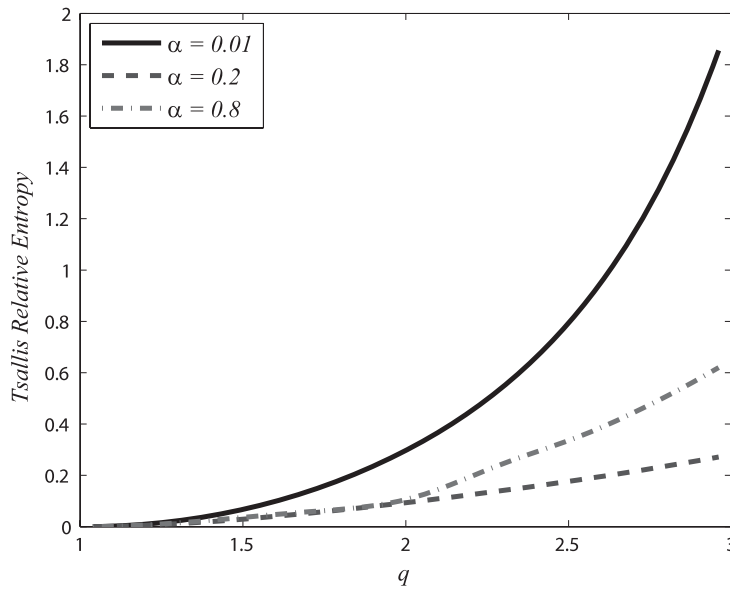


Fig. 5. Tsallis relative entropy of the Gaussian model $Q_1(r)$ and the extended model $Q_q(r)$ vs q , for various α and $h = 0.4$.

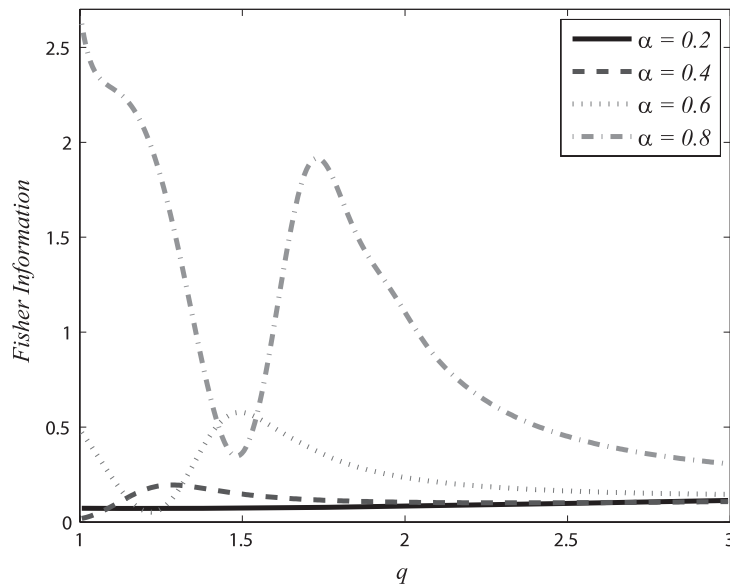


Fig. 6. Fisher information versus q . Coding performance can be improved by shaping the output via input higher-order correlations and their nonlinearities throughout q .

appropriate methodological approach to describe the collective activity of neural populations is a very difficult task. In particular, maximum entropy models have been used in the literature to investigate pairwise correlations in neural population activity [12,13]. But deviations from this model indicate that HOCs have to be taken into account for realistically modeling the population statistics [24].

Recent neurophysiological experiments with large neural populations show significant HOCs [17–20,24]. Statistical modeling on biologically inspired neural populations is important for inferring the function of complex neural circuits. Understanding how neurons process and integrate their input signals requires describing their cooperative action, instead of looking at individual elements. It has been proposed that HOCs allow a decoder to recover the stimulus presented to a neural population much faster than a decoder with access only to pairwise statistics. Several studies have revealed that neurons are indeed very sensitive to the higher-order structure in their input [52–55]. However, theoretical research have mostly focused on the second order statistical features of the input spike trains [21–23]. These properties substantially shape the output response of each neuron, but little is known of how HOCs affect the integration and firing behavior of a cell independently of the second order statistics.

The DG model based on the Amari [21] paper, has been extensively used by Macke et al. [22,23] to construct quantitative predictions on how departures from pairwise models depend on common Gaussian like neuronal inputs. They have shown that common pairwise inputs could be used to explain HOCs in neural population activity. Although the inputs to the model are Gaussian distributed and therefore have no interactions beyond second order, the nonlinear spiking thresholds may give rise to statistical interactions of all orders. Importantly, the DG approach has been developed within the CLT framework, which ensures that the probability distribution function of any measurable quantity is a normal Gaussian distribution, provided that a sufficiently large number of independent random variables with exactly the same mean and variance are being considered [36]. Importantly, the CLT does not hold if correlations between random variables cannot be neglected. More in detail, the CLT has been used within the DG model proposed by Amari and colleagues [21] to estimate the joint probability distribution of firing in a neuronal pool considering the limit of a very large number of neurons. Thus, in their approach, pairwise correlations are quantified through the covariance of the weighted sum of inputs of two given pairs of neurons [21], this sum being considered Gaussian due to the CLT.

However, information in neural populations is often encoded in the activity of large, highly interconnected neural populations. But higher-order statistics can also be shaped and modulated by higher-order input statistics and their intrinsic circuit mechanisms [17–20,24]. In order to account for higher-order statistics within the neuronal inputs, we take advantage in this paper of recent mathematical progress on q -geometry in the asymptotic limit [26–28,32,33,37]. That is, we have considered that each neuron is subject to a weighted sum of inputs, and thus the neuronal population receives common inputs, which are q -Gaussian due to the ECLT [26–28,32,33]. We identify beyond pairwise inputs across neurons as the biophysical mechanisms that generate HOCs within the joint distribution. We provide the exact analytical expression of the joint probability distribution proposed in Ref. [35]. Hence, we quantify the extent of higher than pairwise spike correlations in pools of neural activity by means of an analytical solvable model showing that small, perhaps undetectable, higher-order input correlations may well have an important effect at the population level. The main advantage of our analytical methodology is that, by extending the theoretical framework of Amari et al. [21] to the ECLT, we are able to quantify the amount of correlations that come up from higher-order interconnectivity of the common overlapping neuronal inputs. In this manner, using our analytical expression and measuring the population activity distribution as a function of the normalized firing rate, we can then quantify the amount of correlation higher than two present in the neuronal data. Accounting for HOCs within the common neuronal inputs through the ECLT, allows us to characterize the degree of correlation in the joint distribution avoiding performing the comparison of how far the DG approach deviates from the maximum entropy pairwise estimations [21–23]. More importantly, our approach converges to the DG approach of Amari and collaborators when we consider the limit of the CLT framework ($q = 1$). Providing the exact analytical expression of a generalized joint distribution of firing rates within the ECLT allows us to show how HOCs can improve information coding performance in large neural populations. We can therefore infer the emergent properties of a neural population, showing how information saturates as the degree of correlation q grows.

The main basis of our formalism is that, when taking the limit of a very large number of neurons within the framework of the CLT as in Ref. [21], we are losing information about HOCs. Thus, in a novel theoretical approach, we take the limit of a very large number of neurons within the framework of the ECLT instead of the CLT. The inclusion of a deformation parameter q in the ECLT framework allows us to reproduce remarkably well the experimental distribution of firing and to avoid the sampling size problem of Eq. (3) due to the exponentially increasing number of parameters. We used the ECLT proposed in Ref. [26], which accounts for cases in which correlations between random variables are non negligible. The model we developed using an information geometric approach within the ECLT framework allows us to investigate how information might saturate as the degree of correlation increases at population level. Our current approach truly emphasizes the functional aspects of higher-order statistics, since we consider those inputs correlations seen by each neuron within the pool. We assessed how HOCs depend on the interaction structure of the input, showing that the (output) joint firing distribution is skewed inducing large excursions of synchronized spikes. We have shown through our formalism that coding performance could be improved by shaping the output via higher-order input correlations, and that common input nonlinearities can shape HOCs improving information transmission by neural populations.

Summing up, our approach could be of help to gain further insights into the role of HOCs in information transmission, and constitutes an important mathematical tool to characterize the dynamics of very large populations, ideally of up to several thousands of neurons, where the spike sorting methodologies are useless. Interestingly, as stated by the cell assembly hypothesis [56], neurons transiently synchronize in order to form elementary units of information processing. There are strong experimental evidences to support the idea that synchronous firing of the cell assemblies is related to development of memories and behavior [57–62]. It would be important therefore to apply our approach to evaluate the role of HOCs within the framework of the complementary Hebbian hypothesis that relevant information about concepts, percepts and behavior in general are encoded at the level of multiple assembly activations [56]. We believe that this will become a relevant analytical method for future research on the encoding capacity of large neuronal population and to assist neurophysiologists gain new insights into the neuronal interdependence in the cortex.

Acknowledgments

We acknowledge support by CONICET, Argentina. Research supported by PIP 0255/11 CONICET, Argentina (FM).

References

- [1] R. Desimone, Face-selective cells in the temporal cortex of monkeys, *J. Cogn. Neurosci.* 3 (1991) 1–8.
- [2] R.A. Andersen, Encoding of intention and spatial location in the posterior parietal cortex, *Cereb. Cortex* 5 (1995) 457–469.
- [3] P.S. Churchland, V.S. Ramachandran, T.J. Sejnowski, A critique of pure vision, in: C. Koch, J.L. Davis (Eds.), *Large-Scale Neuronal Theories of the Brain*, MIT Press, Cambridge, MA, 1994, pp. 23–60.
- [4] S.R. Cobb, E. Buhl, K. Halasy, O. Paulsen, P. Somogyi, Synchronization of neuronal activity in hippocampus by individual GABAergic interneurons, *Nature* 378 (1995) 75–78.
- [5] C. von der Malsburg, *The Correlation Theory of Brain Function*, in: Internal Report, vol. 81-2, Max-Planck-Institute for Biophysical Chemistry, Göttingen, 1981.
- [6] C. von der Malsburg, W. Schneider, A neural cocktail-party processor, *Biol. Cybernet.* 54 (1986) 29–40.
- [7] W. Singer, Search for coherence; a basic principle of cortical self organization, *Conc. Neurosci.* 1 (1) (1990) 1–26.
- [8] W. Singer, in: T.P. Hicks, S. Molotchnikoff, T. Ono. (Eds.), *Neuronal Representations, Assemblies and Temporal Coherence*, in: *Progress in Brain Research*, vol. 95, Elsevier, Amsterdam, 1993, pp. 461–474.
- [9] M. Abeles, Role of cortical neuron: integrator or coincidence detector, *Isr. J. Med. Sci.* 18 (1982) 83–92.
- [10] W. Softky, Sub-millisecond coincidence detection in active dendritic trees, *Neuroscience* 58 (1994) 13–41.
- [11] P. Koenig, A.K. Engel, W. Singer, Integrator or coincidence detector? The role of cortical neurons revisited, *Trends Neurosci.* 19 (1996) 130–137.
- [12] E. Schneidman, M.J. Berry, R. Segev, W. Bialek, Weak pairwise correlations imply strongly correlated network states in a neural population, *Nature* 440 (2006) 1007–1012.
- [13] J. Shlens, G.D. Field, J.L. Gauthier, M.I. Grivich, D. Petrusca, A. Sher, A.M. Litke, E.J. Chichilnisky, The structure of multi-neuron firing patterns in the primate retina, *J. Neurosci.* 26 (2006) 8254–8266.
- [14] J. Shlens, G.D. Field, J.L. Gauthier, M. Greschner, A. Sher, A. Litke, E. Chichilnisky, The structure of large-scale synchronized firing in primate retina, *J. Neurosci.* 29 (2009) 5022–5031.
- [15] P. Berens, M. Bethge, Near-maximum entropy models for binary neural representations of natural images, in: *Proceedings of the Twenty-First Annual Conference on Neural Information Processing Systems*, vol. 20, MIT Press, Cambridge, MA, 2008, pp. 97–104.
- [16] Y. Roudi, S. Nirenberg, P. Latham, Pairwise maximum entropy models for studying large biological systems: when they can work and when they can't, *PLoS Comput. Biol.* 5 (2009) e1000380.
- [17] I. Ohiorhenuan, J. Victor, Information-geometric measure of 3-neuron firing patterns characterizes scale-dependence in cortical networks, *J. Comput. Neurosci.* 30 (2010) 125–141.
- [18] I.E. Ohiorhenuan, F. Mechler, K.P. Purpura, A.M. Schmid, Q. Hu, J.D. Victor, Sparse coding and high-order correlations in fine-scale cortical networks, *Nature* 466 (2010) 617–621.
- [19] E. Ganmor, R. Segev, E. Schneidman, Sparse low-order interaction network underlies a highly correlated and learnable neural population code, *Proc. Nat. Acad. Sci.* 108 (2011) 9679–9684.
- [20] H. Shimazaki, S.I. Amari, E.N. Brown, S. Grün, State-space analysis of time-varying higher-order spike correlation for multiple neural spike train data, *PLoS Comput. Biol.* 8 (2012) e1002385.
- [21] S. Amari, H. Nakahara, S. Wu, Y. Sakai, Synchronous firing and higher-order interactions in neuron pool, *Neural Comput.* 15 (2003) 127–142.
- [22] J.H. Macke, P. Berens, A.S. Ecker, A.S. Tolias, M. Bethge, Generating spike trains with specified correlation coefficients, *Neural Comput.* 21 (2009) 397–423.
- [23] J.H. Macke, M. Opper, M. Bethge, Common input explains higher-order correlations and entropy in a simple model of neural population activity, *Phys. Rev. Lett.* 106 (2011) 208102.
- [24] S. Yu, H. Yang, H. Nakahara, G.S. Santos, D. Nikolic, D. Pleniz, Higher-order interactions characterized in cortical activity, *J. Neurosci.* 31 (2011) 17514–17526.
- [25] C. Tsallis, in: J.L. Morán-Lopez, J.M. Sánchez (Eds.), *New Trends in Magnetism, Magnetic Materials and their Applications*, Plenum, New York, 1994, p. 451.
- [26] C. Vignat, A. Plastino, Central limit theorem and deformed exponentials, *J. Phys. A: Math. Theor.* 40 (2007) F969–F978.
- [27] C. Vignat, A. Plastino, Estimation in a fluctuating medium and power-law distributions, *Phys. Lett. A* 360 (2007) 415–418.
- [28] C. Vignat, A. Plastino, Scale invariance and related properties of q -Gaussian systems, *Phys. Lett. A* 365 (2007) 370–375.
- [29] S. Amari, H. Nagaoka, *Methods of Information Geometry*, in: *Translations of Mathematical Monographs*, vol. 191, Oxford University Press, 2000.
- [30] C. Tsallis, M. Gell-Mann, Y. Sato, Asymptotically scale-invariant occupancy of phase space makes the entropy S_q extensive, *Proc. Nat. Acad. Sci.* 102 (2005) 15377–15382.
- [31] M.P.H. Stumpf, M.A. Porter, Critical truths about power laws, *Science* 335 (2012) 665–666.
- [32] S. Umarov, C. Tsallis, S. Steinberg, On a q -central limit theorem consistent with nonextensive statistical mechanics, *Milan J. Math.* 76 (2008) 307–328.
- [33] S. Umarov, C. Tsallis, M. Gell-Mann, S. Steinberg, Generalization of symmetric α -stable Lévy distributions for $q > 1$, *J. Math. Phys.* 51 (2010) 033502.
- [34] S. Nakahara, S. Amari, Information geometric measure for neural spikes, *Neural Comput.* 14 (2002) 2269–2316.
- [35] F. Montani, E. Phoka, M. Portesi, S.R. Schultz, Statistical modelling of higher-order correlations in pools of neural activity, *Physica A* 392 (2013) 3066–3086.
- [36] S.M. Ross, *A First Course in Probability*, second ed., Macmillan Publishing Company, 1976.
- [37] S. Amari, A. Ohara, Geometry of q -exponential family of probability distributions, *Entropy* 13 (2011) 1170–1185.
- [38] Generalized Box–Müller method for generating q -Gaussian random deviates, *IEEE Trans. Inform. Theory* 53 (2007) 4805–4810.
- [39] R. Butler, *Saddlepoint Approximations with Applications*, in: *Cambridge Series in Statistical and Probabilistic Mathematics*, Cambridge University Press, 2007.
- [40] J. Prehl, C. Essex, K.H. Hoffmann, Tsallis relative entropy and anomalous diffusion, *Entropy* 14 (2012) 701–716. <http://dx.doi.org/10.3390/e14040701>.
- [41] C. Tsallis, Generalized entropy-based criterion for consistent testing, *Phys. Rev. E* 58 (1998) 1442–1445.
- [42] L. Borland, A.R. Plastino, C. Tsallis, Information gain within nonextensive thermostatics, *J. Math. Phys.* 39 (1998) 6490. <http://dx.doi.org/10.1063/1.532660>.
- [43] Y. Linnik, An information-theoretic proof of the central limit theorem with the lindeberg condition, *Theory Probab. Appl.* 4 (1959) 288–299.
- [44] R.A. Fisher, On the mathematical foundations of theoretical statistics, *Philos. Trans. R. Soc. Lond. Ser. A* 222 (1922) 309–368.
- [45] B. Roy Frieden, *Science from Fisher Information: A Unification*, Cambridge University Press, Cambridge, 2004.
- [46] A.L. Mayer, C.W. Pawlowski, H. Cabezas, Fisher Information and dynamic regime changes in ecological systems, *Ecol. Modell.* 195 (2006) 72–82.
- [47] O.A. Rosso, L. De Micco, A. Plastino, H. Larrondo, Info-quantifiers' map-characterization revisited, *Physica A* 389 (2010) 249–262.
- [48] F. Olivares, A. Plastino, O.A. Rosso, Ambiguities in the Bandt–Pompe's methodology for local entropic quantifiers, *Physica A* 391 (2012) 2518–2526.
- [49] F. Olivares, A. Plastino, O.A. Rosso, Contrasting chaos with noise via local versus global information quantifiers, *Phys. Lett. A* 376 (2012) 1577–1583.
- [50] F. Montani, E.B. Deleglise, O.A. Rosso, Efficiency characterization of a large neuronal network: a causal information approach, *Physica A* 401 (2014) 58–70. <http://dx.doi.org/10.1016/j.physa.2013.12.053>. B.V. All rights reserved.
- [51] B.B. Averbeck, P.E. Latham, A. Pouget, Neural correlations, population coding and computation, *Nat. Rev. Neurosci.* 7 (2006) 358–366.
- [52] M. Diesmann, M.O. Gewaltig, A. Aertsen, Stable propagation of synchronous spiking in cortical neural networks, *Nature* 402 (1999) 529–533.
- [53] S.M. Bohde, H. Spekreijse, P.R. Roelfsema, The effects of pair-wise and higher-order correlations on the firing rate of a post-synaptic neuron, *Neural Comput.* 12 (2000) 153–179.
- [54] A. Kuhn, A. Aertsen, S. Rotter, Higher-order statistics of input ensembles and the response of simple model neurons, *Neural Comput.* 15 (2003) 67–101.
- [55] A. Benucci, P.F.M.J. Verschure, P. Koenig, Dynamical features of higher-order correlation events: impact on cortical cells, *Cogn. Neurodyn.* 1 (2007) 53–69.

- [56] D.O. Hebb, *The Organization of Behavior*, Wiley, New York, 1949.
- [57] M.A. Wilson, B.L. McNaughton, Reactivation of hippocampal ensemble memories during sleep, *Science* 265 (1994) 676–679. <http://dx.doi.org/10.1126/science.8036517>.
- [58] M. Stopfer, S. Bhagavan, B.H. Smith, G. Laurent, Impaired odour discrimination on desynchronization of odour-encoding neural assemblies, *Nature* 390 (1997) 70–74. <http://dx.doi.org/10.1038/36335>.
- [59] D. Robbe, S.M. Montgomery, A. Thome, P.E. Rueda-Orozco, B.L. Mcnaughton, G. Buzsaki, Cannabinoids reveal importance of spike timing coordination in hippocampal function, *Nat. Neurosci.* 9 (2006) 1526–1533. <http://dx.doi.org/10.1038/nn1801>.
- [60] A. Peyrache, M. Khamassi, K. Benchenane, S.I. Wiener, F.P. Battaglia, Replay of rule-learning related neural patterns in the prefrontal cortex during sleep, *Nat. Neurosci.* 12 (2009) 919–926. <http://dx.doi.org/10.1038/nn.2337>.
- [61] X. Liu, S. Ramirez, P.T. Pang, C.B. Puryear, A. Govindarajan, K. Deisseroth, S. Tonegawa, Optogenetic stimulation of a hippocampal engram activates fear memory recall, *Nature* 484 (2012) 381–385. <http://dx.doi.org/10.1038/nature11028>.
- [62] S. Ramirez, X. Liu, P.A. Lin, J. Suh, M. Pignatelli, R.L. Redondo, T.J. Ryan, S. Tonegawa, Creating a false memory in the hippocampus, *Science* 341 (2013) 387–391. <http://dx.doi.org/10.1126/science.1239073>.