Higher-Order Cumulants Drive Neuronal Activity Patterns, Inducing UP-DOWN States in Neural Populations

Roman Baravalle [1,2] **and Fernando Montani** [1,2,*] (ID)

[1] Instituto de Física de La Plata (IFLP), Universidad Nacional de La Plata, CONICET CCT-La Plata, Diagonal 113 entre 63 y 64, La Plata, Buenos Aires 1900, Argentina; rbaravalle@fisica.unlp.edu.ar

[2] Departamento de Física, Facultad de Ciencias Exactas, UNLP Calle 49 y 115. C.C. 67, La Plata, Buenos Aires 1900, Argentina

[*] Correspondence: fmontani@gmail.com

check for updates

Abstract: A major challenge in neuroscience is to understand the role of the higher-order correlations structure of neuronal populations. The dichotomized Gaussian model (DG) generates spike trains by means of thresholding a multivariate Gaussian random variable. The DG inputs are Gaussian distributed, and thus have no interactions beyond the second order in their inputs; however, they can induce higher-order correlations in the outputs. We propose a combination of analytical and numerical techniques to estimate higher-order, above the second, cumulants of the firing probability distributions. Our findings show that a large amount of pairwise interactions in the inputs can induce the system into two possible regimes, one with low activity ("DOWN state") and another one with high activity ("UP state"), and the appearance of these states is due to a combination between the third- and fourth-order cumulant. This could be part of a mechanism that would help the neural code to upgrade specific information about the stimuli, motivating us to examine the behavior of the critical fluctuations through the Binder cumulant close to the critical point. We show, using the Binder cumulant, that higher-order correlations in the outputs generate a critical neural system that portrays a second-order phase transition.

Keywords: neuronal activity; collective dynamics; information processing; PDF evaluation; dynamic states; higher-order correlations; Fisher information and Shannon entropy; critical fluctuations and binder cumulant; high activity states; low activity states

1. Introduction

How synaptic input correlations from shared presynaptic neurons translate into membrane potential and spike-output correlations is still an open question [1–8]. Input correlations have been picked up in different brain regions at the level of the membrane potential when considering several neurons [9]. The output activity of a neuronal population depends typically on the input activity of a very large number of synaptic inputs. The interplay between inputs activity and the synchronized spiking in the outputs is not well known, and thus the structure underlying the correlated neuronal responses at population level is still not well deciphered [1–6,8].

The output spiking activity is usually associated with the outside world tasks like stimulus encoding, attention, cognition, perception and behavior [6,10–17]. It has been studied how the output spiking activity depends on the mean and variance of the inputs, but it remains unexplored whether these quantities might also affect higher-than-pairwise spike correlations in the outputs [1–3,18–20]. That is to say, it is still undisclosed what could be the exact amount of correlations carried by the

moments higher than two in the joint distribution of firing. Thus, it is not yet well understood whether a distribution of inputs without skewness and kurtosis can also generate higher-order cumulants in the joint distribution of firing, and how they might affect the neural networks dynamics [1–3,18–20]. In particular, how the input statistics affects the third and fourth output moments.

Using an appropriate theoretical modeling of the firing patterns that realistically emulate a neuronal population ensemble, and its spiking correlated activity, constitutes an important step as it can provide important insight on the diverse interaction between structures of the brain. Different statistical approaches have been proposed to investigate the correlated spiking activity in the brain [2–5,15–22]. In particular the DG approach delivers a formalism that allows us to successfully model the correlated structure of the spiking neuronal activity of the cortex [2,3]. The DG model simulates the spiking activity of a neuronal ensemble by generating multivariate Gaussian samples that characterize the correlated inputs to the neuronal population and thresholding them [2,3]. The DG approach generates dense higher-order statistics and can emulate higher-order correlations observed in the cortex [2,3,23]. The model generates, through a network with a multivariate distribution, with a given mean h and covariance Λ using a fixed threshold, certain correlations in the outputs [2,3]. These correlations are usually characterized by the output firing rate, μ, and the covariance, Σ [2,3,23]. However, it remains still unknown whether moments higher than two in the output joint distribution of firing could also shape the dynamics of the neural populations simulated using the DG model [1–3,18–20].

In this paper, we carefully investigate how pure pairwise correlations in the inputs can generate correlations higher than two in the outputs, by exactly quantifying them for different scenarios of the mean and variance of the inputs. That is to say, we compute the spike correlation coefficients up to fourth order considering a population of neurons receiving pairwise correlated inputs within the DG model. We propose a combination of analytical and numerical techniques to estimate the first, second, third and fourth moments of the probability distribution of firing. We first perform numerical estimations of the Amari's correlations coefficients up to fourth order for the neural simulations [8,21]. Secondly, we carry out an analytical and numerical expansion of the DG model to explore the impact of the estimation of the higher-than-second-order cumulants in the population dynamics. The main objective is to understand whether these higher-order components might provide further understanding of the neuronal dynamics when considering a population of finite size and in the asymptotic regime.

We examine the relationship between higher-order correlations and firing rate, and the behavior of the critical fluctuations in the outputs, through the fourth-order cumulant, considering different degrees of input correlations. Importantly, the dependence of the Binder cumulant as function of the correlation moments allows us to analyze higher-order fluctuations close to the critical point present in the model [24–27]. We investigate whether a large amount of pairwise correlations in the inputs may have any effect on quiescent ("DOWN state") and highly active state ("UP state") in the outputs [28–32]. In the next section we test the hypothesis that a large amount of pairwise correlations in the inputs affects higher-order cumulants of the joint distribution of firing and if this can induce "DOWN/UP states". Through the estimation of the fourth-order cumulant and the related Binder cumulant [24–27], which is independent of the number of neurons at the transition point, we show that the DG model exhibit a second-order phase transition. The connection between the correlations of spiking activity and the Binder cumulant links two fundamental features of the neural code and critical states. Consequently, this provides us a measurable methodology to evaluate detectable proof for higher-order correlations and to assess its role in the dynamical properties of the neuronal systems.

2. Characterization of the DG Model Considering Spike Correlations Higher than Two

The spiking activity of the network was described by a random variable $X \in \{0,1\}^n$ with output mean μ and output covariance matrix Σ. Here n is the number of neurons in the network.

In the DG model [22,33], an n-dimensional latent random variable U, described by a thresholded Gaussian distribution with mean h and covariance matrix Λ, is the cause of the spiking activity of the network. If we call Ω_i the threshold for the ith component of U, we can define the random variable X as follows: X takes the value 1 if and only if $U_i > \Omega_i$, and 0 otherwise. In the following we take $\Omega_i = 0 \,\forall i$. Thus, the input of the network consists of a distribution with pairs interaction, defined through the parameters h and Λ, with a threshold in 0. It is this threshold that causes the higher-order correlations in the output of the network, so the activity of the network cannot be fully described by the output mean and covariance, μ and Σ. The relevance of DG model on the investigation of higher-order correlations has been previously pointed out in Refs. [2,3,23]. The main idea here is to provide a quantification of the higher-order cumulants of the probability distributions of firing. More precisely we decompose these higher orders, putting special emphasis on third- and fourth-order interactions, as they are also important to describe the neural networks dynamics [15–17].

First, we want to model the first two moments of the variable X through the two first moments of the variable U. Without loss of generality, we put $\Lambda_{ii} = 1$, $i = 1, \ldots, n$. Next, we calculate the mean and the covariance of X by definition:

$$
\begin{aligned}
\mu_i = E[X_i] &= \sum_{l=0}^{1} l\, p\,(X_i = l) \\
&= p\,(X_i = 1) = P\,(U_i > 0) \\
&= 1 - P\,(U_i \leq 0) \\
&= 1 - \Phi\,(0) \\
&= \Phi\,(h_i),
\end{aligned}
\tag{1}
$$

where Φ is the cumulative function of the univariate Gaussian, and h_i is the ith component of the mean vector h. The last line comes from the fact that the distribution of U_i is symmetric around the mean h_i; it is not a general result for all the input distributions. For the covariance elements we can make the same reasoning:

$$
\begin{aligned}
\Sigma_{ij} = E\left[X_i X_j\right] - \mu_i \mu_j \\
&= \sum_{l=0,m=0}^{1} l\, m\, p\,(X_i = l, X_j = m) - \mu_i \mu_j \\
&= p\,(X_i = 1, X_j = 1) - \mu_i \mu_j \\
&= P\,(U_i > 0, U_j > 0) - \mu_i \mu_j \\
&= 1 - P\,(U_i \leq 0, U_j \leq 0) - \mu_i \mu_j \\
&= 1 - \Phi_2\,(0, 0, \Lambda_{ij}) - \mu_i \mu_j \\
&= \Phi_2\,(h_i, h_j, \Lambda_{ij}) - \mu_i \mu_j.
\end{aligned}
\tag{2}
$$

Here Φ_2 is the cumulative function of the bivariate Gaussian with unit variances and correlation coefficient Λ_{ij}. The previous line is again valid only for symmetric distributions.

Summarizing, the connection between the input parameters h, Λ and the output parameters μ, Σ is given by the two equations:

$$
\mu_i = \Phi\,(h_i),
\tag{3}
$$

$$
\Sigma_{ij} = \Phi_2\,(h_i, h_j, \Lambda_{ij}) - \Phi\,(h_i)\,\Phi\,(h_j),
\tag{4}
$$

with the extra condition $\Lambda_{ii} = 1$. This pair of equations have a unique solution for any admissible moments μ and Σ, and can be solved numerically [2,3].

Up to this point we only measure the first and second central moment of the spiking distribution. In order to quantify the third and fourth central moments, we must have a similar expression for the moments in terms of the input distribution parameters. The third central moment can be estimated as (see Appendix A for further details):

$$
\begin{aligned}
\Sigma_{ijk} &= E\left[(X_i - \mu_i)(X_j - \mu_j)(X_k - \mu_k)\right] \\
&= E\left[X_i X_j X_k\right] - \mu_i E\left[X_j X_k\right] - \mu_j E\left[X_i X_k\right] \\
&\quad - \mu_k E\left[X_i X_j\right] + 2\mu_i \mu_j \mu_k.
\end{aligned}
\tag{5}
$$

The mean and the covariance of X can be described using Equations (3) and (4). The only term left to calculate is $E\left[X_i X_j X_k\right]$. Following the same reasoning as previously, we found that:

$$
\begin{aligned}
E\left[X_i X_j X_k\right] &= \sum_{l,m,n=0}^{1} l\, m\, n\, p\left(X_i = l, X_j = m, X_k = n\right) \\
&= p\left(X_i = 1, X_j = 1, X_k = 1\right) \\
&= P\left(U_i > 0, U_j > 0, U_k > 0\right) \\
&= 1 - \Phi_3\left(0, 0, 0, \Lambda_{ijk}\right) \\
&= \Phi_3\left(h_i, h_j, h_k, \Lambda_{ijk}\right),
\end{aligned}
\tag{6}
$$

where Φ_3 is the cumulative density function of a trivariate Gaussian with a 3×3 covariance matrix Λ_{ijk} corresponding to the random variables X_i, X_j and X_k.

For the fourth central moment, the following result is found (see Appendix A for further details):

$$
\begin{aligned}
\Sigma_{ijkl} &= E\left[(X_i - \mu_i)(X_j - \mu_j)(X_k - \mu_k)(X_l - \mu_l)\right] \\
&= E\left[X_i X_j X_k X_l\right] - \mu_i E\left[X_j X_k X_l\right] \\
&\quad - \mu_j E\left[X_i X_k X_l\right] - \mu_k E\left[X_i X_j X_l\right] \\
&\quad - \mu_l E\left[X_i X_j X_k\right] + \mu_i \mu_j E\left[X_k X_l\right] \\
&\quad + \mu_i \mu_k E\left[X_j X_l\right] + \mu_j \mu_k E\left[X_i X_l\right] \\
&\quad + \mu_i \mu_l E\left[X_j X_k\right] + \mu_j \mu_l E\left[X_i X_k\right] \\
&\quad + \mu_k \mu_l E\left[X_i X_j\right] - 3\mu_i \mu_j \mu_k \mu_l.
\end{aligned}
\tag{7}
$$

Again, the only term that must be calculated is $E\left[X_i X_j X_k X_l\right]$. As expected, this term is equal to $\Phi_4\left(h_i, h_j, h_k, h_l, \Lambda_{ijkl}\right)$, with Φ_4 the cumulative function of the four-dimensional Gaussian, and Λ_{ijkl} the 4×4 covariance matrix between the variables X_i, X_j, X_k and X_l.

We can see that the third and fourth central moments depends also on the lower-order moments, so we may have a combination of interactions between four, three and two neurons. In order to make a splitting of interactions between one, two, three or four neurons, we appeal to the Information Geometry tools, and make a coordinate system transformation, to the so-called theta coordinates [33]. In the next section we link the DG approach with the Information Geometry coordinates to estimate the effect of the correlations higher than two in the firing probability distributions.

2.1. Coordinate Transformations for the Pool of Neurons

On the whole we can symbolize the neuronal population activity by a binary vector $\mathbf{x} = (x_1, ..., x_N)$ in the space X of all binary vectors of length N, where $x_i = 0$ if neuron i is silent in some time window and $x_i = 1$ if it is firing one or more spikes. The probability $P(\mathbf{x})$ of observing a particular response can be portrayed using different coordinate systems. A useful way of characterizing the population activity distribution is by indicating the $2^N - 1$ individual probability values; these are

called the *p*-coordinates. The probability is identified by the $2^N - 1$ marginals; usually known as the η-coordinates [21]. Supplied $P(\mathbf{x}) \neq 0$ for a given $\mathbf{x}$, any such distribution can be extended in the so-called log-linear model, or θ-coordinates system [21,22]:

$$P(x) = \exp\left(\sum x_i + \theta_i + \sum_{i<j} x_i x_j \theta_{ij} + \right.$$
$$+ \sum_{i<j<k} x_i x_j x_k \theta_{ijk} + \cdots +$$
$$\left. + \sum_{i<\cdots<N} x_i \cdots x_N \theta_{i\cdots N} - \psi\right) \tag{8}$$

where the $2^N - 1$ different θ coefficients exclusively resolve the above distribution. This coordinate system was developed by Amari and peers to investigate the possibilities and interactions between neurons [21,22].

Allow us to consider that the neural population is a totally homogeneous pool; all of the parameters portraying single neuron properties do not depend upon the accurate character of each neurons, but instead on the amount of neurons being considered. Due to the symmetry all the θ at a given order i are the same and can be rewritten by θ_i. Within this framework, the probability of having exactly m neurons active at a certain time bin would be [8,20]:

$$P(m) \sim \binom{N}{m} \exp\left[\sum_{i=1}^{N} \binom{m}{i} \theta_i - \psi\right]. \tag{9}$$

That is, we assume that all the parameters characterizing single neuron properties and interactions between any group of neurons do not depend on the precise identity of the considered neurons, but rather on the number of neurons. Let us consider therefore a pool of N neurons where each unit has a membrane potential u_i subject to a joint Normal distribution. Given $U \sim N(h, \Lambda)$, where $\Lambda = \mathbb{I}(1 - \alpha) + \alpha \mathbf{1}_N \mathbf{1}_N^T$, then

$$u_i = h + \sqrt{1 - \alpha} v_i + \sqrt{\alpha} \varepsilon. \tag{10}$$

The variables v_i and ε are two independent random variables subject to the Gaussian distribution $N(0, 1)$ with mean 0 and variance 1. The input statistics are chosen such that the outputs X have mean μ and covariance Σ. Since here we focus on the case of homogeneous populations then $\mu_i = \mu$ and $\Sigma_{ij} = \sigma$, and we name the pairwise correlation coefficient as
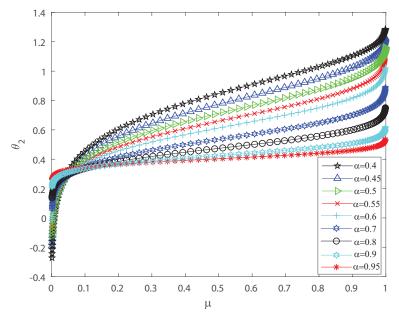
$$\rho = \frac{\sigma}{\mu(1 - \mu)}. \tag{11}$$

the skewness in Equation (5) as $\zeta = \Sigma_{ijk}$ and the kurtosis coefficient in Equation (7) as $\chi = \Sigma_{ijkl}$.

2.2. Synchronized Activity of the Outputs Considering a Neuronal Pool
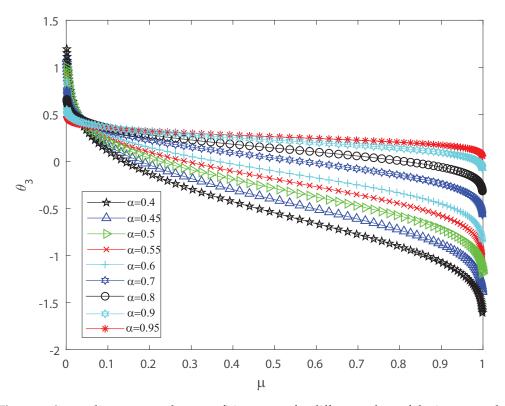
Several investigations pointed out the significance of higher-order correlations, as pairwise models fail to demonstrate the variability of the neuronal ensembles at a global level [8,14–18,34–36]. In the DG model, designed to produce spiking patterns thresholding a multivariate random Gaussian variable, spike correlations across neurons emerge from interactions in the inputs [2,3,33]. They are usually quantified through the mean firing rate as in Equation (3) and by second-order cumulant depicted in Equation (4) [2,3,33]. Next, let us investigate how input correlations may influence the spiking correlated activity of the outputs. Along the following lines we portray the DG model estimating the first-, second-, third- and fourth-order $\boldsymbol{\theta}$ coefficients of the Amari's expansion [21,22] when considering different input correlations $\boldsymbol{\alpha} = (\Lambda)_{ij}, i \neq j$. We consider a neuronal population of $n = 50$ neurons simulated using the DG model [2,3,33]. Figure 1 shows how $\boldsymbol{\theta_2}$ grows non-linearly as the mean firing

rate μ increases, for different values of the input correlations α. Figure 2 shows θ_3 versus μ that present the opposite behavior to Figure 1 as in this case the third order decreases non-linearly with the mean firing rate μ, for different values of the input correlations α. Please note that θ_3 versus μ takes larger negative values as μ increases. In addition, Figure 3 depicts the fourth-order correlation coordinate θ_4 for different values of the input correlations α, which takes negative values for very low firing rates, μ, but grows positively and non-linearly as μ becomes bigger. Overall this probabilistic model is imposing non-trivial constraints on population-level statistics as it appears that the third-order term helps to inhibit the correlation contribution to the global distribution of firing, because θ_3 is mostly a negative correlation coefficient (see Figure 2). We show using the Amari's formalism that the third term gives a solid restraint to the general DG system action, while the fourth-order term represents a sort of excitatory balance in the probability distribution. That is θ_4 is mostly a positive correlation coefficient (see Figure 3).
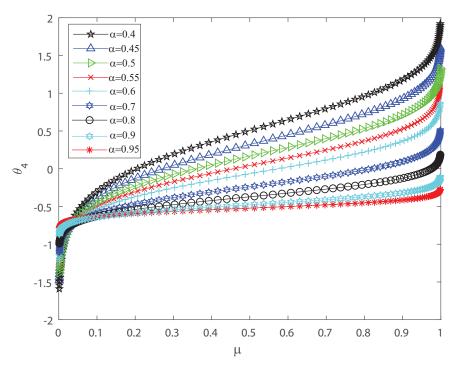
To detect the dynamical changes in the probability distributions we estimate Fisher information and Shannon entropy. Let us emphasize that Fisher information can be deciphered as an extent of the capacity to assess a parameter and as the amount of information that can be extracted from a set of measurements [37,38], while entropy is a measure of uncertainty [39]. We refer the reader to Appendix B for the detailed definitions. We show in Figures 4 and 5 that the effects of inputs correlations are not negligible in the output spiking activity when estimating the Shannon entropy and Fisher information, respectively. Both Shannon entropy and Fisher information depict an opposite behavior as the firing rate grows, for different values of the input correlations α. Notice that entropy becomes bigger as function of μ when the input correlation α is smaller. In contrast Fisher information is smaller as the firing rate μ becomes bigger for lower values of α. Thus, Fisher information is larger as the variance of the inputs grows. The previous results show that the effect of correlations in the inputs induces dynamical changes in the probability distributions that are significant on information transmission in the outputs. Moreover, as is well known, when describing the behavior of Fisher information close to a critical regime this quantity grows as the system is near a phase transition [40,41]. In the next sections we investigate the emergent properties of this system, the possible link between information theory and thermodynamic interpretations of critical behavior.
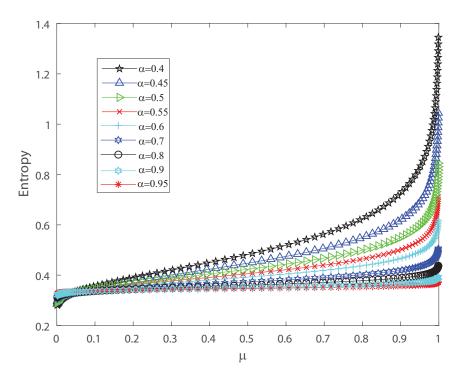

Figure 1. θ_2 coordinate versus the mean firing rate μ, for different values of the input correlations $\alpha = 0.4$, $\alpha = 0.45$, $\alpha = 0.5$, $\alpha = 0.5$, $\alpha = 0.55$, $\alpha = 0.6$, $\alpha = 0.7$, $\alpha = 0.8$, $\alpha = 0.9$ and $\alpha = 0.95$. θ_2 grows non-linearly as the firing rate μ increases. We consider $n = 50$.
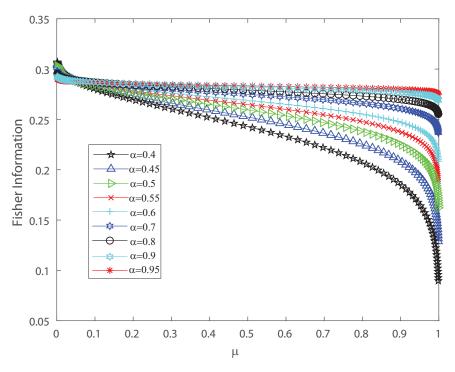
Figure 2. θ_3 coordinate versus the mean firing rate μ, for different values of the input correlations $\alpha = 0.4$, $\alpha = 0.45$, $\alpha = 0.5$, $\alpha = 0.5$, $\alpha = 0.55$, $\alpha = 0.6$, $\alpha = 0.7$, $\alpha = 0.8$, $\alpha = 0.9$ and $\alpha = 0.95$. θ_3 decreases non-linearly as the firing rate μ increases. We consider $n = 50$.

Figure 3. θ_4 coordinate versus the mean firing rate μ, for different values of the input correlations $\alpha = 0.4$, $\alpha = 0.45$, $\alpha = 0.5$, $\alpha = 0.5$, $\alpha = 0.55$, $\alpha = 0.6$, $\alpha = 0.7$, $\alpha = 0.8$, $\alpha = 0.9$ and $\alpha = 0.95$. θ_4 grows non-linearly as the firing rate μ increases. We consider $n = 50$.

Figure 4. Shannon entropy versus the mean firing rate μ, for different values of the input correlations $\alpha = 0.4$, $\alpha = 0.45$, $\alpha = 0.5$, $\alpha = 0.5$, $\alpha = 0.55$, $\alpha = 0.6$, $\alpha = 0.7$, $\alpha = 0.8$, $\alpha = 0.9$ and $\alpha = 0.95$. We consider $n = 50$.
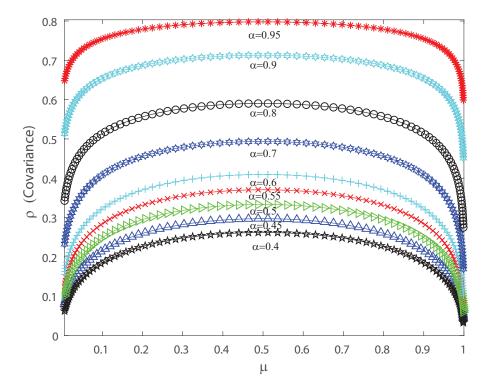

Figure 5. Fisher information versus the mean firing rate μ, for different values of the input correlations $\alpha = 0.4$, $\alpha = 0.45$, $\alpha = 0.5$, $\alpha = 0.5$, $\alpha = 0.55$, $\alpha = 0.6$, $\alpha = 0.7$, $\alpha = 0.8$, $\alpha = 0.9$ and $\alpha = 0.95$. We consider $n = 50$.

Despite the previous evidence showing non-negligible effects of correlations higher than two, we cannot interpret from the analysis depicted above whether they might reflect the

mechanisms responsible for tuning the network of stochastic neurons towards a critical boundary. Spiking correlations have proven to be important for the characterization of the DG model, which are characterized by the estimations of the cumulants depicted in Equations (3) and (4) [2,3,33]. This model has been broadly used to build quantitative forecasts on how takeoffs from pairwise models rely upon normal Gaussian like neuronal inputs and to create an ensemble of neurons with a spike trains with indicated mean and pairwise statistics [2,3,33]. However, we consider that the effect of the higher-order cumulants also must be taken into account [2,3,33], and the main idea here is to gain further understanding of the neuronal dynamics by extending this approach to higher orders. Thus, in the following section, we perform an extended analysis of the DG model to gain a broader picture of the role of higher correlations on the neural network dynamics.
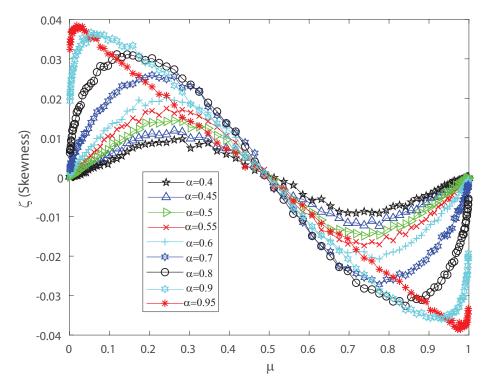
2.3. Higher-Order Cumulants and the Identification of a Second-Order Phase Transition

The DG model mimics population spiking action by producing multivariate Gaussian examples and thresholding them [2,3,33]. Dissimilar to most of the maximum entropy models [12,13,42], we can legitimately indicate the firing rates for the DG model so as to create cortical-like measurements [2,3,23]. It has been shown that correlations are important for the characterization of the DG model as it generates dense higher-order statistics and can replicate higher-order spike correlations [2,3,33]. It is our goal to show that a proper quantification of the third and fourth moment of the spike-count distribution is useful for characterizing and, in some cases, anticipating a change in the spiking behavior of the population. This means that the skewness and kurtosis could be important, not only in the theoretical case, but also in the characterization of neuronal ensembles in the cerebral cortex. Let us consider that the neural array is a totally homogeneous pool; all of the parameters depicting single neuron properties and interactions between any ensemble of neurons do not depend on the precise identity of the considered neurons, merely just on the amount of neurons considered. In the current simulations we also consider a population of **50** neurons within the DG model. In order to perform a proper characterization of the second, third and fourth moment of the spike-count distribution in the outputs we take several values for the covariance in the inputs. Figure 6 shows the pairwise correlation coefficient ρ versus the mean firing rate μ. This figure demonstrates that this model has a remarkable connection between pairwise correlations ρ and the firing rate μ, which is likely to be found in neural ensembles, and depicts that the pairwise correlation coefficient ρ increases monotonically with firing rate up to $\mu = $ **0.5**. Figure 6 depicts a maximum at $\mu = $ **0.5** and we have selected various values of the input correlations $\alpha = $ **0.4**, $\alpha = $ **0.45**, $\alpha = $ **0.5**, $\alpha = $ **0.55**, $\alpha = $ **0.6**, $\alpha = $ **0.7**, $\alpha = $ **0.8**, $\alpha = $ **0.9** and $\alpha = $ **0.95**. Figure 7 shows the skewness ζ as a function of the mean firing rate μ of the neuronal population, when considering several values of the input correlations $\alpha = $ **0.4**, $\alpha = $ **0.45**, $\alpha = $ **0.5**, $\alpha = $ **0.55**, $\alpha = $ **0.6**, $\alpha = $ **0.7**, $\alpha = $ **0.8**, $\alpha = $ **0.9** and $\alpha = $ **0.95**. Note from Figure 7 that the skewness changes its sign at $\mu = $ **0.5**, and presents a maximum for low firing rates and a minimum when considering very high firing rates.
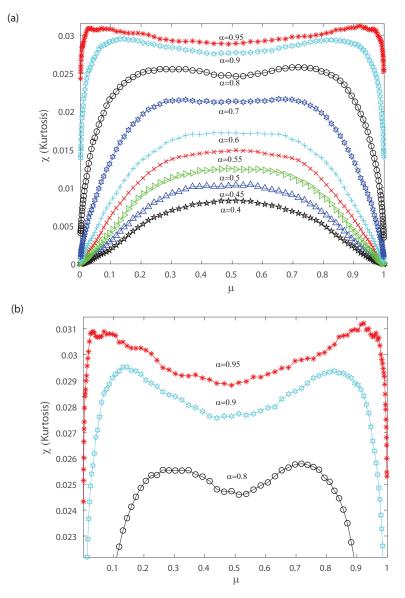
Figure 8a shows the kurtosis χ versus the mean firing rate for different values of the input correlations $\alpha = $ **0.4**, $\alpha = $ **0.45**, $\alpha = $ **0.5**, $\alpha = $ **0.55**, $\alpha = $ **0.6**, $\alpha = $ **0.7**, $\alpha = $ **0.8**, $\alpha = $ **0.9** and $\alpha = $ **0.95**. Notice from Figure 8a that there is a change in the shape of kurtosis for values of the input correlation α bigger than **0.7**. Please note that the kurtosis depicts a local minimum at $\mu = $ **0.5** when considering values of the input correlation α bigger than **0.7**. Bottom Figure 8b represents a zoom of top Figure 8a, and shows that as α becomes equal to **0.8**, **0.9** and **0.95** the kurtosis of the distribution depicts a bimodal behavior. Interestingly, below $\alpha = $ **0.7** the kurtosis is unimodal but for higher values the kurtosis show a bimodal behavior with the two peaks that indicates a broken-symmetry due to a induced phase transition, i.e., the fourth-order cumulant becomes highly suggestive that long-range fluctuations appear for a critical values of α between **0.8** and **0.95**. Below these critical values of α, the maximum of kurtosis occurs at $\mu = $ **0.5** for which the kurtosis shows a single peak. The previous results are similar for different mean h in the inputs.

Figure 6. Covariance ρ as function of the population mean firing rate μ for various values of the input correlations $\alpha = 0.4$, $\alpha = 0.45$, $\alpha = 0.5$, $\alpha = 0.5$, $\alpha = 0.55$, $\alpha = 0.6$, $\alpha = 0.7$, $\alpha = 0.8$, $\alpha = 0.9$ and $\alpha = 0.95$. We consider $n = 50$.

Figure 7. Skewness ζ as function of the population mean firing rate μ for various values of the input correlations $\alpha = 0.4$, $\alpha = 0.45$, $\alpha = 0.5$, $\alpha = 0.5$, $\alpha = 0.55$, $\alpha = 0.6$, $\alpha = 0.7$, $\alpha = 0.8$, $\alpha = 0.9$ and $\alpha = 0.95$. We consider $n = 50$.

(a)

(b)

Figure 8. Kurtosis χ as a function of the population mean firing rate μ. (**a**) We consider various values of the input correlations $\alpha = 0.4$, $\alpha = 0.45$, $\alpha = 0.5$, $\alpha = 0.5$, $\alpha = 0.55$, $\alpha = 0.6$, $\alpha = 0.7$, $\alpha = 0.8$, $\alpha = 0.9$ and $\alpha = 0.95$. (**b**) Zoom of the top figure for $\alpha = 0.8$, $\alpha = 0.9$ and $\alpha = 0.95$. We consider $n = 50$.

Let us note that for very low values of μ and pairwise interactions in the inputs bigger than $\alpha = 0.7$, Figures 7 and 8 depict a maximum with positive skewness and kurtosis. On the contrary taking a firing rate μ close to the maximum value and pairwise interactions in the inputs bigger than $\alpha = 0.7$ leads to a minimum (negative) skewness and positive (maximum) kurtosis in the outputs (see Figures 7 and 8). The bimodal behavior for the kurtosis seems to indicate that two states with maximum kurtosis that are accompanied with maximum or minimum skewness may coexist. This suggests that a phase transition could take place and that the third- and fourth-order cumulants might be of help to characterize it. Next, in order to gain further understanding of the nature of those cumulants we estimate them in the asymptotic limit.

The Asymptotic Behavior

To calculate the higher-order cumulants in the limit of an infinite number of neurons, let us consider a model of a homogeneous population with a number n of neurons [3]. Each neuron

has a firing rate μ and constant pairwise covariance ρ. We obtain samples X from the model by dichotomizing a latent Gaussian U with mean h, unit variances and pairwise covariance (and correlation coefficient) $\alpha > 0$ [2,3]. We decompose the Gaussian U into three parts: the constant corresponding to the mean, one multidimensional independent Gaussian V with zero mean and a univariate Gaussian distribution ε with zero mean and variance equal to **1**:

$$U \sim N(h, \Lambda) \text{ where } \Lambda = I_n(1 - \alpha) + \alpha 1_n 1_n^T,$$
$$U = h + \sqrt{1 - \alpha}V + \sqrt{\alpha}\varepsilon 1_n,$$
$$\text{where } V \sim N(0, I_n)$$
$$\text{and } \varepsilon \sim N(0, 1).$$

1_n is the unitary vector, formed by all ones. Such pool of neurons is fully characterized by the spike-count distribution of the population [2,3,33]. We name k as the spike count in the population, and $r = k/n$ the proportion of spiking neurons. The probability of having k spikes is $Q_{DG}(k) = P_{DG}(k)/\binom{n}{k}$ [2,3]. Hence, the population spike-count distribution is

$$P_{DG}(X = x) = \int_u P(X = x|u)p(u)dz$$
$$= \int_{-\infty}^{\infty} \phi_\alpha(\varepsilon)P(X = x|\varepsilon)d\varepsilon$$
$$= \int_{-\infty}^{\infty} \phi_\alpha(\varepsilon) \prod_{i=1}^n P(X_i = x|\varepsilon)d\varepsilon$$
$$= \int_{-\infty}^{\infty} \phi_\alpha(\varepsilon) \prod_{i=1}^n (1 - L(\varepsilon))^{1-x_i} L(\varepsilon)^{x_i}d\varepsilon$$
$$= \int_{-\infty}^{\infty} \phi_\alpha(\varepsilon) (1 - L(\varepsilon))^{n-k} L(\varepsilon)^k d\varepsilon$$
$$= \int_{-\infty}^{\infty} \phi_\alpha(\varepsilon) \exp\left(n W(s)\right)d\varepsilon$$
$$\text{where } W(\varepsilon) = r \log L(\varepsilon) + (1 - r) \log L(-\varepsilon)$$
$$\text{with } L(\varepsilon) = P(u_i > 0|\varepsilon)$$
$$= P\left(V_i > -\frac{\varepsilon + h}{\sqrt{1 - \alpha}}\right)$$
$$= \Phi\left(\frac{\varepsilon + h}{\sqrt{1 - \alpha}}\right).$$

Here $\phi_\alpha(\varepsilon)$ is the probability density function of a one-dimensional Gaussian with mean equal to **0** and variance α [2,3]. Φ is the cumulative function of a standard one-dimensional Gaussian.

For large population sizes n, we can use the saddle-point approximation in the last integral. After some calculations one can obtain the asymptotic continuous valued spike-count distribution $f(r)$, as in Ref. [2,3]:

$$f(r) = \sqrt{\frac{1 - \alpha}{\alpha}} \exp\left(\frac{h^2}{2 - 4\alpha}\right)$$
$$\exp\left(-\frac{(1 - 2\alpha)(\Phi^{-1}(r) - \frac{h\sqrt{1-\alpha}}{1-2\alpha})^2}{2\alpha}\right).$$

In the following we use $f(r)$ to estimate the asymptotic pairwise cumulant ρ, the skewness ζ and the kurtosis χ of the distribution.

Figure 9a–c depict the asymptotic estimations of the second-, third- and fourth-order cumulants, respectively. The asymptotic limit has a similar behavior as the previous DG simulation considering a finite number of neurons. The current findings emphasize the results of the previous section. In the following we investigate how higher-order correlations in the outputs might induce neural criticality. The partition-function has a discontinuity at $\alpha = 1/2$. Thus, the transition point is at $\alpha = 1/2$, independently of the mean h of the latent Gaussian. For values $\alpha < 1/2$ the distribution is unimodal, while for $\alpha > 1/2$ the distribution is bimodal. Next, let us consider the heat capacity and the fourth-order cumulant dependence with temperature. To do it so we derive, using the ideas in Ref. [2], the behavior of the specific heat and the Binder cumulant as a function of temperature. Given the model distribution $P(x)$, the extension at any temperature $T = 1/\beta$ is $P_\beta(x) = P(x)^\beta / Z_\beta$. Then, the spike-count distribution is given by $P_\beta(k) = n \exp\left(n(1-\beta)\eta(r)\right)\frac{P(k)^\beta}{Z_\beta}$, where $\eta(r) = -r\log(r) - (1-r)\log(1-r)$ [2,3].

It is useful to calculate the entropy rate $h_{DG} = \frac{1}{n}H(x)$ of the variable X. As shown in Ref. [2,3], for large populations, the result is:

$$h_{DG} \approx \int_0^1 f_\beta(r)\eta(r)\,dr,$$

where the subscript β indicates the dependence with the temperature. Taking into account that the specific heat is $c = Var\left[\log\left(P_\beta(x)\right)\right]/n$, and that we are in the limit of large n, the dominant term in c is [2,3]:

$$c \approx n \int_0^1 f_\beta(r)\left(\eta(r) - h_{DG}\right)^2 dr.$$

Extending the results in Refs. [2,3], we calculate the fourth-order cumulant. We depict here step by step the calculation. First, we define the fourth-order cumulant $v_4 = \sum_x P_\beta(x)\left(\log\left(P_\beta(x)\right) - E\left[\log\left(P_\beta(x)\right)\right]\right)^4$. Using the fact that the sum over x could be replaced by a sum over k (due to the fact that the neurons are identical), $\sum_x = \sum_k \binom{n}{k}$, and that $P_\beta(x) = P_\beta(k)/\binom{n}{k} \approx \frac{f_\beta(r)}{n}/\binom{n}{k}$, the fourth-order cumulant is:
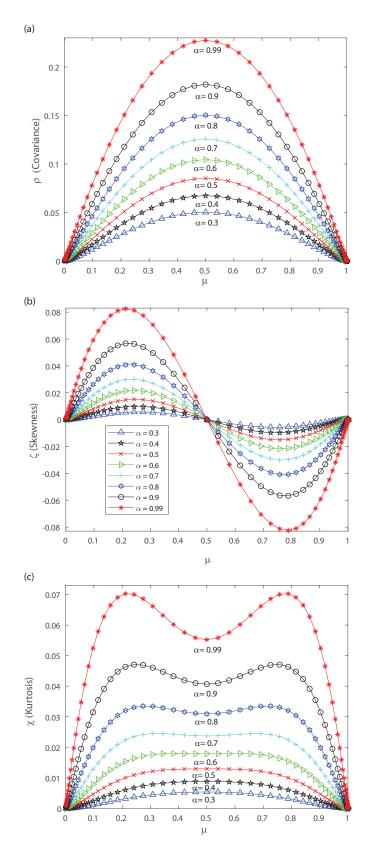
$$v_4 \approx \sum_k \frac{f_\beta(r)}{n}\left(\log\left(\frac{f_\beta(r)}{n}\right) - \log\binom{n}{k} + H_n\right)^4.$$

Defining $n\,\eta(r) = \binom{n}{k} = \binom{n}{nr}$, using the Stirling approximation for the binomial coefficient, and approximating $\frac{\sum_k}{n} \approx \int_0^1 dr$, we arrive at:

$$v_4 \approx \int_0^1 f_\beta(r)\left(\log f(r) - \log n - n\eta(r) + nh_{DG}\right)^4 dr$$

$$= \int_0^1 f_\beta(r)\left(\log\frac{f(r)}{n} - n\left(\eta(r) - h_{DG}\right)\right)^4 dr.$$

As in the limit of large populations the dominant term is the one which is linear with n inside the parentheses, thus the final result is:

$$v_4 = n^4 \int_0^1 f_\beta(r)\left(\eta(r) - h_{DG}\right)^4 dr. \tag{12}$$

Figure 9. Asymptotic estimations of the DG cumulants. We consider several values of the input correlations $\alpha = 0.3$, $\alpha = 0.4$, $\alpha = 0.5$, $\alpha = 0.6$, $\alpha = 0.7$, $\alpha = 0.8$, $\alpha = 0.9$ and $\alpha = 0.99$. (**a**) Covariance ρ versus the mean firing rate μ. (**b**) Skewness ζ versus the mean firing rate μ. (**c**) Kurtosis χ as a function of the population mean firing rate μ.

Considering that the second-order cumulant can be written as $v_2 = Var\left[\log\left(P_\beta(x)\right)\right] = n\,c$ and that the fourth-order cumulant is given by Equation (12), we define the Binder Cumulant as [24–27]:

$$B = 1 - \frac{v_4}{3v_2^2}.$$

Based on these estimations we can explore, as in Ref. [3], the scaling properties of the population in function of the temperature. In this case, the fourth-order moment is also important, because it allows us to calculate the Binder cumulant [24–27]. This cumulant is a robust way in which one can obtain the critical temperature of a system, finding the crossing points between this function for various population sizes [24–27]. In Figure 10 we depict the Binder cumulant as function of the temperature, and it is shown, as expected, that the critical point lies at $T = 1$. This cumulant is less biased in temperature than the specific heat, shown in Figure 11, which makes it a very good quantifier to perform finite size scaling analysis for a neuronal system depicting a critical phase transition. Thus, by computing the Binder critical cumulant which does not depend on the quantity of neurons at the transition point, we show that the system displays a second-order phase transition. Similar results are obtained for the Binder cumulant using the numerical estimation of the second- and fourth-order cumulant presented above. The current finding emphasize the relevance of estimating the higher-order terms, as the expanded moments of the proposed models capture the dynamics of the interconnected network in neural spiking through the estimation of the Binder cumulant. We show through the estimation of the higher-order cumulants the existence of a phase transition of second order. Our analysis reveals that the coding of the higher than second-order moments could dominate the brain dynamics as the stochastic restriction of the neurons leads them towards a critical boundary.
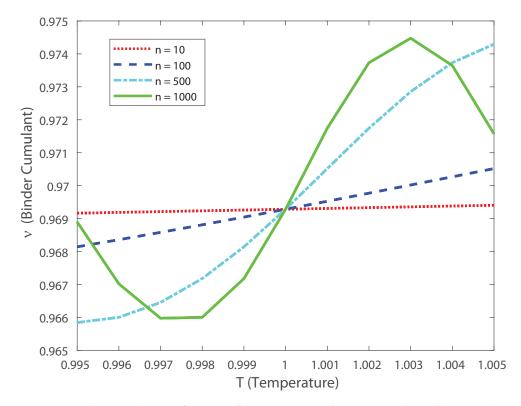

Figure 10. Binder cumulant as a function of the temperature for various values of the population size *n*. Notice that even for low values of populations (e.g., $n = 10$) the crossing point between the curves lies at the critical temperature $T = 1$.
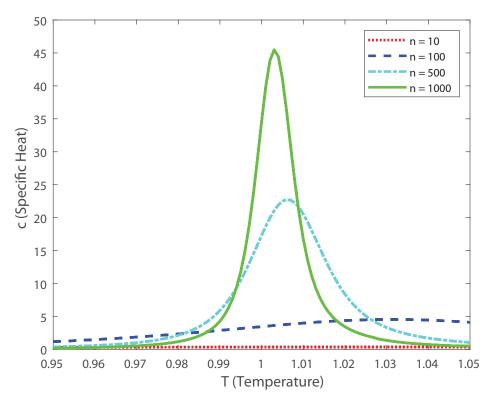
Figure 11. Specific heat as a function of the temperature for various values of the population size *n*.

3. Discussion and Conclusions

The principal notable property of neural ensembles is that they are globally coupled. While it has been known for quite a while that neurons do not spike autonomously, early estimations of their cooperative contributions of functional connectivity were initially portrayed using pairwise correlations models (Ising-like) [12,13,42]. In particular the Ising model in the framework of the maximum entropy principle has been associated with a certain correlation between pairs of neurons in the retina [12,13,42]. As the technology made possible to simultaneously record the activity of hundred neurons in the cerebral cortex, be that as it may, it turned out to be progressively clear that pairwise models are inadequate to completely capture statistics in different data [8,14–17]. Moreover, it has been demonstrated that these intricate correlations are progressively adjusted by the stimulus [16] and that they can affect coding properties [8].

On the one hand it has been proposed that instead of directly estimating particular correlation parameters, it is possible to calculate the cumulant-based inference of higher-order correlations (CuBIC) to derive a lower bound in a given neuronal data set on the order of correlation [4,5]. On the other hand it has been shown that when modeling the collective behavior of biological systems, arisen from real data, the factual knowledge about the system is balanced close to an extremely uncommon point in their parameter space: a critical point [43]. Changes in triplet connections can altogether upgrade coding if those correlations depend on the stimuli, and it is through the skew that triplet correlations induce changes in population-wide distributions of neural responses [34]. Moreover, recent findings have shown the relevance of the kurtosis in the distribution of firing [20], and it has been also shown that the kurtosis is a very good quantifier for the description of the scaling properties of a neural population [44]. Higher correlations are indeed important for understanding neuronal avalanches [23] and even synchronous silence, or the co-inactivation of neurons, which is an omnipresent evidence of higher-order correlations [35]. Thus, assessing the accurate measure of pairwise, triplewise and quadruplewise correlations can be of great assistance to describe important functional properties of the cortex.

Cortical type models such as the DG have been used to generate spike trains with certain firing rates and degrees of correlation between action potentials [2,3,23,33]. For this purpose, it has been evaluated how different means and variances in the DG model inputs can affect the statistics of the neuronal firings in the outputs [2,3,23,33], and how statistical deformations of the Gaussian inputs changes the dynamics of the outputs [19,20,45]. In particular, a recent paper has shown the importance of maximum entropy models for analyzing the firing patterns of neuronal populations and the level of correlation between them [46]. In order to investigate the role of higher-order correlations in the neural code, we performed in this paper an extension of DG model proposed in Refs. [2,3] quantifying the contributions of the higher-order cumulants. Our current approach allows us to infer that, when considering strongly correlated inputs, a balance between skewness and kurtosis can induce jumps between "DOWN and UP states" in neuronal a population and this corresponds to a second-order phase transition, i.e., thought this approach we describe the basis of the dynamic mechanism by which these states transition occur.

In this paper, we show first using the Amari's formalism on the DG model that the third-order term provides a strong inhibition to the overall DG networks activity, while the fourth-order term account for excitatory contributions in the probability distribution providing insight that higher-order correlations should not be disregarded. Thereafter we extended the DG model [2,3,23,33] and to implement this we estimate the skewness and kurtosis of the simulated population while broadly changing the input parameters of the population, i.e., the inputs correlations of the Gaussian distribution. In the current study we take into account all correlations orders as it is important to consider that the neuronal systems work close a critical point to improve information transmission in the cortex [47,48]. The extended DG model permits us to emulate the activity of the neuronal cortex but it is through higher-order cumulants that the estimation of the Binder cumulant allows us to provide evidence showing that this realistic neuronal network exhibit a second-order phase transition. More importantly, we show using the extended DG model that higher-order correlations are the mechanisms responsible for tuning the neuronal network towards the critical state. Our findings show on one hand that high pairwise interactions in the inputs could be sufficient to generate a third- and fourth-order cumulant in which a positive skewness and kurtosis can induce the system to a "DOWN state" (quiescent) [28–32]. On the other hand high pairwise interactions in the inputs can produce third- and fourth-order cumulant in which a negative skewness and positive kurtosis can induce the system to an "UP state" (active) [28–32]. Thus, if the amount of pairwise correlations in the inputs is high enough there is always a nonzero probability that the system could pass from a quiescent to a highly active state, by inducing a change in the sign of the skewness and taking advantage of the bimodal shape of the kurtosis. Near the critical point the system can be maximally responsive, and this might be therefore a mechanism to optimally managing information about the external stimuli. Indeed, the reduced fourth-order cumulant (or Binder cumulant) allows us to show that a bimodal shape with the two peaks corresponds to the critical behavior of a second-order phase transition [24–27].

Numerous brain functions are optimized due to the fact that the brain operates between two coexisting phases of order and disorder [49]. The nervous system operates near a non-equilibrium phase transition where a large amount of spatial and temporal higher-order correlations are compatible. The brain keeps itself at the border of a whole instability urging a great number of adaptations at different scales which resemble a critical state. Criticality appears as a fundamental element of cortical boosts, and higher-order synchrony of firing neurons that are needed to investigate the emergent properties in spatial structure of the cortical activity and its reorganization become power laws at this transition point [23,49]. In the current paper we were able to investigate the collective dynamics of neuronal activity developing an extension of the DG model that quantifies the contributions of the higher-order cumulants, allowing us to investigate the scaling properties of the system. In addition, by combining different techniques that goes from Information Geometry to a DG extended approach we investigate higher interactions among a neuronal network. Our current study allows us to identify correlations of order higher than two as key factors that determine phase transitions in the DG model,

considering that at the level of the Amari's coordinates the triplet correlations gives an inhibitory-like restraint to the general DG system while the fourth-order term represents excitatory-like contributions depicting the non-locality of neural populations. However, it is the change of sign of skewness for high values of pairwise inputs correlations and the balance with a positive kurtosis what determines the possible excursions from quiescent to highly active states.

The DG model emulates dense higher-order statistics and can simulate higher-order correlations that can be recorded in vivo within specific areas of brain cortex [2,3,23]. In contrast with maximum entropy models [12,13,42], we can choose the firing rate in the DG model to emulate a cortical statistic patterns and that produces the same pairwise, triplewise and quadruplewise cumulant being quantified by ρ, ζ and χ, respectively. Using the DG simulation, we find a non-monotonic relationship between output higher-order correlations and firing rate, motivating us to investigate the behavior of the critical fluctuations by means of the Binder cumulant. Thus, we consider that delivering an appropriate framework for estimating the exact amount of correlations that pair, triple, and quadruple moments are carrying about the joint firing distribution in the DG model is of ultimate relevance for understanding further the emergent properties of cortical processing. Importantly, our discoveries demonstrates that a high amount of pairwise correlations in the inputs can yield to a third and a fourth cumulant that leads to a framework with to two different scenarios, one with very low firing rate ("DOWN state") and another with very high firing rate ("UP state"). Thus, the extended DG tools can be used to investigate further the complexity of different dynamic responses within neural populations of the cerebral cortex.

Author Contributions: R.B. and F.M. contributed equally in the design of this research as well as in the writing of this paper. All authors have read and approved the final manuscript.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A. Estimations of the Third and Fourth Central Moment

Based on the fundamental properties of the expectation values [50] and performing some algebra to basic statistic one can obtain the third central moment as:

$$
\begin{aligned}
\Sigma_{ijk} &= E\left[(X_i - \mu_i)(X_j - \mu_j)(X_k - \mu_k)\right] \\
&= E\left[X_i X_j X_k - \mu_i X_j X_k - \mu_j X_i X_k - \mu_k X_i X_j \right. \\
&\quad \left. + \mu_i \mu_j X_k + \mu_j \mu_k X_i + \mu_i \mu_k X_j - \mu_i \mu_j \mu_k\right] \\
&= E\left[X_i X_j X_k\right] - \mu_i E\left[X_j X_k\right] - \mu_j E\left[X_i X_k\right] \\
&\quad - \mu_k E\left[X_i X_j\right] + \mu_i \mu_j E\left[X_k\right] + \mu_j \mu_k E\left[X_i\right] \\
&\quad + \mu_i \mu_k E\left[X_j\right] - \mu_i \mu_j \mu_k \\
&= E\left[X_i X_j X_k\right] - \mu_i E\left[X_j X_k\right] - \mu_j E\left[X_i X_k\right] \\
&\quad - \mu_k E\left[X_i X_j\right] + 2\mu_i \mu_j \mu_k.
\end{aligned}
\tag{A1}
$$

Similarly taking into account the definition of the fourth central moment [50], we can estimate:

$$
\begin{aligned}
\Sigma_{ijkl} \;=\; & E\left[\left(X_i - \mu_i\right)\left(X_j - \mu_j\right)\left(X_k - \mu_k\right)\left(X_l - \mu_l\right)\right] \\
=\; & E\left[X_i X_j X_k X_l\right] - \mu_i E\left[X_j X_k X_l\right] \\
& -\mu_j E\left[X_i X_k X_l\right] - \mu_k E\left[X_i X_j X_l\right] \\
& -\mu_l E\left[X_i X_j X_k\right] + \mu_i \mu_j E\left[X_k X_l\right] \\
& +\mu_i \mu_k E\left[X_j X_l\right] + \mu_j \mu_k E\left[X_i X_l\right] \\
& +\mu_i \mu_l E\left[X_j X_k\right] + \mu_j \mu_l E\left[X_i X_k\right] \\
& +\mu_k \mu_l E\left[X_i X_j\right] - E\left[X_i\right]\mu_j \mu_k \mu_l \\
& -\mu_i E\left[X_j\right]\mu_k \mu_l - \mu_i \mu_j E\left[X_k\right]\mu_l \\
& \;\; -\mu_i \mu_j \mu_k E\left[X_l\right] + \mu_i \mu_j \mu_k \mu_l \\
=\; & E\left[X_i X_j X_k X_l\right] - \mu_i E\left[X_j X_k X_l\right] \\
& -\mu_j E\left[X_i X_k X_l\right] - \mu_k E\left[X_i X_j X_l\right] \\
& -\mu_l E\left[X_i X_j X_k\right] + \mu_i \mu_j E\left[X_k X_l\right] \\
& +\mu_i \mu_k E\left[X_j X_l\right] + \mu_j \mu_k E\left[X_i X_l\right] \\
& +\mu_i \mu_l E\left[X_j X_k\right] + \mu_j \mu_l E\left[X_i X_k\right] \\
& +\mu_k \mu_l E\left[X_i X_j\right] - 3\mu_i \mu_j \mu_k \mu_l,
\end{aligned}
\tag{A2}
$$

which corresponds to the expression of Equation (7).

Appendix B. Shannon Entropy and Fisher Information

Shannon entropy [39] is usually understood as a quantifier of the lack of order or predictability, the gradual decline into disorder. Given a continuous probability distribution function $f(x)$ with $x \in \Delta \subset \mathbb{R}$ and $\int_\Delta f(x)\,dx = 1$, the Shannon entropy S [39] is defined as

$$
S[f] \;=\; -\int_\Delta f\,\ln(f)\,dx.
\tag{A3}
$$

It is considered a measure of uncertainty and a measure of "global character" that is not too sensitive to strong changes in the probability distribution taking place in a small region. Let us consider a time series $\mathcal{X}(t) \equiv \{x_t; t = 1, \cdots, M\}$, a set of size M from $\mathcal{X}$ which can be observed, related to a probability given by $P \equiv \{p_j; j = 1, \cdots, N\}$ with $\sum_{j=1}^{N} p_j = 1$ and N number of possible states of the system being considered, the Shannon entropy [39] is defined by

$$
S[P] = -\sum_{j=1}^{N} p_j \ln(p_j).
\tag{A4}
$$

This functional is equal to zero when we can predict with full certainty which of the possible outcomes j, whose probabilities are given by p_j, will actually take place. Our knowledge of the underlying process, described by the probability distribution, is maximal in this instance. In contrast, this knowledge is commonly minimal for a uniform distribution $P_e = \{p_j = 1/N, \forall j = 1, \cdots, N\}$.

Let us define *Fisher information* [37,38] which is a measure of the gradient content/size of the distribution f, so that is enough sensitive even for small local disturbances, as:

$$
F[f] \;=\; \int \frac{|\vec{\nabla} f(x)|^2}{f(x)}\,dx.
\tag{A5}
$$

When considering a discrete distribution, the Fisher information is estimated as [51]:

$$F[\psi] \;=\; 4 \int \left\{ \frac{d\psi}{dx} \right\}^2 dx \,.$$

Its discrete normalized version ($0 \leq F \leq 1$) is now

$$F[P] \;=\; F_0 \sum_{i=1}^{N-1} (\sqrt{p_{i+1}} - \sqrt{p_i})^2. \tag{A6}$$

where the amplitude of probability is $f(x) = \psi(x)^2$. Here the normalization constant F_0 reads

$$F_0 \;=\; \begin{cases} 1 & \text{if } p_{i*} = 1 \text{ for } i^* = 1 \text{ or } i^* = N \\ & \text{and } p_i = 0 \; \forall i \neq i^* \\ 1/2 & \text{otherwise.} \end{cases}$$

If our system is very organized and thus represented by a very narrow PDF, we have Shannon's entropy $S \sim 0$ and Fisher information measures $F \sim F_{max}$. When the system under study is in a very irregular state, it will get a PDF that is almost flat $S \sim S_{max}$, while $F \sim 0$. We named S_{max} and F_{max} to the maximum value for Shannon entropy and Fisher information, respectively. The general behavior of Fisher information is contrary to the one of the Shannon entropy [52].

References

1. De la Rocha, J.; Doiron, J.B.; Shea-Brown, B.E.; Josic, K.; Reyes, A. Correlation between neural spike trains increases with firing rate. *Nature* **2007**, *448*, 802–806. [CrossRef] [PubMed]
2. Macke, J.H.; Berens, P.; Ecker, A.S.; Tolias, A.S.; Bethge, M. Generating spike trains with specified correlation coefficients. *Neural Comput.* **2009**, *21*, 397–423. [CrossRef] [PubMed]
3. Macke, J.H.; Opper, M.; Bethge, M. Common input explains higher order correlations and entropy in a simple model of neural population activity. *Phys. Rev. Lett.* **2011**, 106, 208102. [CrossRef] [PubMed]
4. Staude, B.; Grün, S.; Rotter, S. Higher-order correlations in non-stationary parallel spike trains: Statistical modeling and inference. *Front. Comput. Neurosci.* **2010**, *4*, 16. [CrossRef]
5. Staude, B.; Rotter, S.; Grün, S.J. CuBIC: Cumulant based inference of higher-order correlations in massively parallel spike trains. *Comput Neurosci.* **2010**, *29*, 327. [CrossRef]
6. Montani, F.; Kohn, A.; Smith, M.A.; Schultz, S.R. The role of correlations in direction and contrast coding in the primary visual cortex. *J. Neurosci.* **2007**, *27*, 2338–2348. [CrossRef]
7. Montani, F.; Oliynyk, A.; Fadiga, L. Superlinear Summation of Information in Premotor Neurons Pairs. *Int. J. Neural Syst.* **2017**, *27*, 1650009. [CrossRef]
8. Montani, F.; Ince, R.; Senatore, R.; Arabzadeh, E.; Diamond, M.; Panzeri, S. The impact of high-order interactions on the rate of synchronous discharge and information transmission in somatosensory cortex. *Phil. Trans. R. Soc. A* **2009**, *367*, 3297–3310. [CrossRef]
9. Dayan, P.; Abbott, L.F. *Theoretical Neuroscience: Computational and Mathematical Modeling of Neural Systems*; MIT Press: Cambridge, UK, 2001.
10. Smith, M.A.; Jia, X.; Zandvakili, A.; Kohn, A. Laminar dependence of neuronal correlations in visual cortex. *J. Neurophysiol.* **2013**, *109*, 940–947. [CrossRef]
11. Nirenberg, S.; Latham, P.E. Decoding neuronal spike trains: How important are correlations?. *Proc. Natl. Acad. Sci. USA* **2003**, *100*, 7348–7553. [CrossRef]
12. Schneidman, E.; Berry, M.J.; Segev, R.; Bialek, W. Weak pairwise correlations imply strongly correlated network states in a neural population. *Nature* **2006**, *440*, 1007–1012. [CrossRef] [PubMed]
13. Shlens, J.; Field, G.D.; Gauthier, J.L.; Grivich, M.I.; Petrusca, D.; Sher, A.; Litke, A.M.; Chichilnisky, E.J. The structure of multi-neuron firing patterns in primate retina. *J. Neurosci.* **2006**, *26*, 8254–8266. [CrossRef]

14. Roudi, Y.; Nirenberg, S.; Latham, P.E. Pairwise maximum entropy models for studying large biological systems: When they can work and when they can't. *PLoS Comput. Biol.* **2009**, *5*, e1000380. [CrossRef] [PubMed]

15. Ohiorhenuan, I.; Victor, J. Information-geometric measure of 3-neuron firing patterns characterizes scale-dependence in cortical networks. *J. Comput. Neurosci.* **2010**, *30*, 125–141. [CrossRef] [PubMed]

16. Ohiorhenuan, I.E.; Mechler, F.; Purpura, K.P.; Schmid, A.M.; Hu, Q.; Victor, J. Sparse coding and high-order correlations in fine-scale cortical networks. *Nature* **2010**, *466*, 617–621. [CrossRef] [PubMed]

17. Ganmor, E.; Segev, R.; Schneidman, E. Sparse low-order interaction network underlies a highly correlated and learnable neural population code. *Proc. Natl. Acad. Sci. USA* **2011**, *108*, 9679–9684. [CrossRef]

18. Montani, F.; Phoka, E.; Portesi, M.; Schultz, S.R. Statistical modelling of higher-order correlations in pools of neural activity. *Physica A* **2013**, *392*, 3066–3086. [CrossRef]

19. Montangie, L.; Montani, F. Higher-order correlations in common input shapes the output spiking activity of a neural population. *Physica A* **2017**, *471*, 845–861. [CrossRef]

20. Montangie, L.; Montani, F. Common inputs in subthreshold membrane potential: The role of quiescent states in neuronal activity. *Phys. Rev. E* **2018**, *97*, 060302(R). [CrossRef]

21. Amari, S.; Nagaoka, H. *Methods of Information Geometry*; Oxford University Press: Oxford, UK, 2000.

22. Nakahara, S.; Amari, S. Information geometric measure for neural spikes. *Neural Comput.* **2002**, *14*, 2269–2316. [CrossRef]

23. Yu, S.; Yang, H.; Nakahara, H.; Santos, G.S.; Nikolic, D.; Plenz, D. Higher-order interactions characterized in cortical activity. *J. Neurosci.* **2011**, *31*, 17514–17526. [CrossRef]

24. Binder, K. Finite size scaling analysis of Ising model block distribution functions. *Physik B* **1981**, *43*, 119–140. [CrossRef]

25. Binder, K. Critical Properties from Monte Carlo Coarse Graining and Renormalization. *Phys. Rev. Lett.* **1981**, *47*, 693. [CrossRef]

26. Albano, E.V.; Binder, K. Wetting transition in the two-dimensional Blume-Capel model: A Monte Carlo study. *Phys. Rev. E* **2012**, *85*, 061601. [CrossRef] [PubMed]

27. Albano, E.V.; Binder, K. Finite-Size Scaling Approach for Critical Wetting: Rationalization in Terms of a Bulk Transition with an Order Parameter Exponent Equal to Zero. *Phys. Rev. Lett.* **2012**, *109*, 036101. [CrossRef] [PubMed]

28. Cowan, R.L.; Wilson, C.J. Spontaneous firing patterns and axonal projections of single corticostriatal neurons in the rat medial agranular cortex. *J. Neurophysiol.* **1994**, *71*, 17–32. [CrossRef] [PubMed]

29. Wilson, C.J.; Kawaguchi, Y.J. The origins of two-state spontaneous membrane potential fluctuations of neostriatal spiny neurons. *J. Neurosci.* **1996**, *16*, 2397–2410. [CrossRef] [PubMed]

30. Contreras, D.; Timofeev, I.; Steriade, M.J. Mechanisms of long-lasting hyperpolarizations underlying slow sleep oscillations in cat corticothalamic networks. *J. Physiol.* **1996**, *494*, 251–264. [CrossRef] [PubMed]

31. Destexhe, A.; Rudolph, M.; Paré, D. The high-conductance state of neocortical neurons in vivo. *Nat. Rev. Neurosci.* **1996**, *20043*, 739–751.

32. Destexhe, A.; Hughes, S;W.; Rudolph, M.; Crunelli, V. Are corticothalamic 'up' states fragments of wakefulness?. *Trends Neurosci.* **2007**, *30*, 334–342. [CrossRef]

33. Amari, S.; Nakahara, H.; Wu, S.; Sakai, Y. Synchronous firing and higher order interactions in neuron pool. *Neural Comput.* **2003**, *15*, 127–142. [CrossRef] [PubMed]

34. Cayco-Gajic, N.A.; Zylberberg, J.; Shea-Brown, E. Triplet correlations among similarly tuned cells impact population coding. *Comput. Neurosci.* **2015**, *9*, 57. [CrossRef] [PubMed]

35. Shimazaki, H.; Sadeghi, K.; Ishikawa, T.; Ikegaya, Y.; Toyoizumi, T. Simultaneous silence organizes structured higher-order interactions in neural populations. *Sci. Rep.* **2015**, *5*, 9821. [CrossRef]

36. Berens, P.; Bethge, M. Near-maximum entropy models for binary neural representations of natural images. In Proceedings of the Twenty-First Annual Conference on Neural Information Processing Systems, Vancouver, BC, Canada, 8–10 December 2008; Volume 20, pp. 97–104.

37. Fisher, R.A. On the mathematical foundations of theoretical statistics. *Philos. Trans. R. Soc. Lond. Ser. A* **1922**, *222*, 309–368.

38. Roy Frieden, B. *Science from Fisher Information: A Unification*; Cambridge University Press: Cambridge, UK, 2004.

39.   Shannon, C.; Weaver, W. *The Mathematical Theory of Communication*; University of Illinois Press: Champaign, IL, USA, 1949.

40.   Prokopenko, M.; Lizier, J.T.; Obst, O.; Wang, X.R. Relating Fisher information to order parameters. *Phys. Rev. E* **2011**, *84*, 041116. [CrossRef]

41.   Prokopenko, M.; Barnett, L.; Harré, M.; Lizier, J.T.; Obst, O.; Wang, X.R. Fisher transfer entropy: Quantifying the gain in transient sensitivity. *Proc. R. Soc. A* **2015**, *471*, 20150610. [CrossRef]

42.   Tkacik, G.; Schneidman, E.; Berry, M.J., II; Bialek, W. Ising models for networks of real neurons. *arXiv* **2006**, arXiv:q-bio/0611072.

43.   Mora, T.; Bialek, W. Are Biological Systems Poised at Criticality? *J. Stat. Phys.* **2011**, *144*, 268–302. [CrossRef]

44.   Bradde, S.; Bialek, W. PCA meets RG. *J. Stat. Phys.* **2017**, *167*, 462–475. [CrossRef]

45.   Montangie, L.; Montani, F. Quantifying higher-order correlations in a neuronal pool. *Physica A* **2015**, *421*, 388–400. [CrossRef]

46.   Cayco-Gajic, N.A.; Zylberberg, J.; Shea-Brown, E. A Moment-Based Maximum Entropy Model for Fitting Higher-Order Interactions in Neural Data. *Entropy* **2018**, *20*, 489. [CrossRef]

47.   Scarpetta, S.; de Candia, A. Alternation of up and down states at a dynamical phase-transition of a neural network with spatiotemporal attractors. *Front. Syst. Neurosci.* **2014**, *8*, 88. [CrossRef] [PubMed]

48.   Brochini, L.; de Andrade Costa, A.; Abadi, M.; Roque, A.C.; Stolfi, J.; Kinouchi, O. Phase transitions and self-organized criticality in networks of stochastic spiking neurons. *Sci. Rep.* **2016**, *6*, 35831. [CrossRef] [PubMed]

49.   Ponce-Alvarez, A.; Jouary, A.; Privat, M.; Deco, G.; Sumbre, G. Whole-Brain Neuronal Activity Displays Crackling Noise Dynamics. *Neuron* **2018**, *100*, 1446–1459. [CrossRef]

50.   Casella, G.; Berger, R.L. *Statistical Inference*, 2nd ed.; Pacific Grove: Duxbury, MA, USA, 2002.

51.   Sanchez-Moreno, P.; Dehesa, J.S.; Yanez, R.J. Discrete densities and Fisher Information. In Proceedings of the 14th International Conference on Difference Equations and Applications, Estoril, Portugal, 19–23 October 2009; pp. 291–298.

52.   Pennini, F.; Plastino, A. Reciprocity relations between ordinary temperature and the Frieden-Soffer Fisher temperature. *Phys. Rev. E* **2005**, *71*, 047102. [CrossRef]