

RESEARCH ARTICLE

# Interplay between Graph Topology and Correlations of Third Order in Spiking Neuronal Networks

Stojan Jovanović<sup>1,2\*</sup>, Stefan Rotter<sup>1</sup>

**1** Bernstein Center Freiburg & Faculty of Biology, University of Freiburg, Freiburg, Germany, **2** CB, CSC, KTH Royal Institute of Technology, Stockholm, Sweden

\* [stojan.jovanovic@ntnu.no](mailto:stojan.jovanovic@ntnu.no)



**OPEN ACCESS**

**Citation:** Jovanović S, Rotter S (2016) Interplay between Graph Topology and Correlations of Third Order in Spiking Neuronal Networks. *PLoS Comput Biol* 12(6): e1004963. doi:10.1371/journal.pcbi.1004963

**Editor:** Olaf Sporns, Indiana University, UNITED STATES

**Received:** January 26, 2016

**Accepted:** May 2, 2016

**Published:** June 6, 2016

**Copyright:** © 2016 Jovanović, Rotter. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Data Availability Statement:** Data are available on the Open Science Framework database at <https://osf.io/dm9g8/>.

**Funding:** This work was supported by the Erasmus Mundus Joint Doctoral program EuroSPIN, the German Federal Ministry of Education and Research (BFNT, grant 01GQ0830), and the German Research Foundation (BLBT, grant EXC 1086). The article processing charge was covered by the open access publication fund of the University of Freiburg. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

## Abstract

The study of processes evolving on networks has recently become a very popular research field, not only because of the rich mathematical theory that underpins it, but also because of its many possible applications, a number of them in the field of biology. Indeed, molecular signaling pathways, gene regulation, predator-prey interactions and the communication between neurons in the brain can be seen as examples of networks with complex dynamics. The properties of such dynamics depend largely on the topology of the underlying network graph. In this work, we want to answer the following question: Knowing network connectivity, what can be said about the level of third-order correlations that will characterize the network dynamics? We consider a linear point process as a model for pulse-coded, or spiking activity in a neuronal network. Using recent results from theory of such processes, we study third-order correlations between spike trains in such a system and explain which features of the network graph (i.e. which topological motifs) are responsible for their emergence. Comparing two different models of network topology—random networks of Erdős-Rényi type and networks with highly interconnected hubs—we find that, in random networks, the average measure of third-order correlations does not depend on the local connectivity properties, but rather on global parameters, such as the connection probability. This, however, ceases to be the case in networks with a geometric out-degree distribution, where topological specificities have a strong impact on average correlations.

## Author Summary

Many biological phenomena can be viewed as dynamical processes on a graph. Understanding coordinated activity of nodes in such a network is of some importance, as it helps to characterize the behavior of the complex system. Of course, the topology of a network plays a pivotal role in determining the level of coordination among its different vertices. In particular, correlations between triplets of events (here: action potentials generated by neurons) have recently garnered some interest in the theoretical neuroscience community. In this paper, we present a decomposition of an average measure of third-order

**Competing Interests:** The authors have declared that no competing interests exist.

coordinated activity of neurons in a spiking neuronal network in terms of the relevant topological motifs present in the underlying graph. We study different network topologies and show, in particular, that the presence of certain tree motifs in the synaptic connectivity graph greatly affects the strength of third-order correlations between spike trains of different neurons.

## Introduction

Analyzing networks of interacting elements has become the tool of choice in many areas of biology. In recent years, network models have been used to study the interactions between predator and prey [1], gene interactions [2] and neural network dynamics [3, 4]. A fundamental question in the study of complex networks is how the topology of the graph on which a dynamic process evolves influences its activity. A particularly interesting issue is the emergence of synchronized, or correlated patterns of events. While it is obvious that the presence or absence of such patterns of activity depends largely on how individual nodes in the network are connected, it is by no means a trivial task to explain exactly how this happens.

In theoretical neuroscience, the connection between network topology and correlated activity continues to be an important topic of study. Not only are correlations between neuronal spike trains believed to have an important function in information processing [5, 6] and coincidence detection [7], but they are also believed to be tied to expectation and attention (see [7] for details). In addition, it been shown that nerve cells can be extremely sensitive to synchronous input from large groups of neurons [8].

While there has been much work on elucidating the causes and effects of pairwise correlations between spike trains [3], it seems that correlations beyond second order also have a role to play in the brain. For example, it was indicated that a nonlinear neuron's firing rate profile depends on higher-order correlations between the presynaptic spikes [9]. Higher-order correlations have also been reported in the rat somatosensory cortex and the visual cortex of the behaving macaque [10]. Indeed, it has been suggested that these correlations are inherent properties of cortical dynamics in many species [11, 12]. As a result, neural data has recently been intensively investigated for signs of higher-order synchrony using classical means such as maximum entropy models [13–18]. In addition, new methods are being developed in order to shed more light on what seems to be a very important property of networks in the brain [19–21].

In this work, we study the relation between the topology (i.e. synaptic connectivity) and correlations of third order between neuronal spike trains. Our aim was to show how triplet correlations depend on topological motifs in a network with known connectivity. We hope our results can be used to facilitate thought experiments to relate hypothetical connectivity to third-order correlations by, for example, assuming specific network topologies and then computing how these assumptions affect the dynamics.

In the following text, the word “connection” is meant to be translated as “synapse”. While this might be a point of contention, in previous work, it was clearly shown that that a mapping between synaptically coupled spiking networks (e.g. comprising LIF neurons) and statistical, point process models, such as Hawkes process exist, with exactly the same underlying connectivity [22]. In addition, it has been demonstrated that synaptic connectivity can be reconstructed from simulated spike trains with very high fidelity, provided the network has a connectivity which is not too dense and not too sparse [23]. On the basis of these two results, we feel enough confidence to claim that in the Hawkes process network models considered

here “connections” in terms of coupling kernels can be safely identified with “synapses” in a spiking neuronal network.

However, we would also like to point out that knowing the true connectivity in an experimental setting is close to impossible. Indeed, the connectivity matrices, obtained by statistical inference methods applied to neural data are rarely more than a proxy for the actual “anatomical” connectivity. In other words, the existence of a statistical relationship between the firings of two neural cells (“correlation”) does generally not imply the existence of an actual synapse between them. In addition, the inference of connectivity from neural data is confounded by undersampling. One can typically only record from a tiny fraction of all neurons that constitute the network, while most of the population remains effectively hidden to the experimenter.

Similar work, pertaining to the influence of connectivity on correlations of second order has already been published [3, 24–26]. In it, the authors dissect the contribution of specific structural motifs to the emergence of pairwise correlations in a recurrent network of interconnected point processes, meant to represent neurons communicating via spikes. Interpreting known mathematical results [27] in an original fashion, they show how the influence of recurrent input can be disentangled to take into account not only effects of direct connections, but also indirect connectivity. However, no such result exists in the case of more complex patterns, stemming from correlations of higher order. With this paper, we aim to fill this gap.

Analogously to [3], we show that measures of third-order correlations (known in the statistical literature as “third-order joint cumulants”) are also heavily influenced by the presence of certain topological motifs in the network graph. We find that the motifs in question can be thought of representing “common input to triplets of neurons” and that, in graph theory terms, they represent rooted trees with three leaf nodes. Furthermore, we obtain an expansion of the joint third cumulants in terms of a sum of weights of all such subgraphs and show that, in a regular network (that is, a network with fixed in- and out-degrees), this expansion can be approximated by a formula that doesn’t depend on the specific adjacency matrix, but rather on global connectivity parameters, such as the connection probability  $p$ . In addition, our result extends to large random Erdős-Rényi type networks, as they are approximately regular when the number of nodes grows without bound. We find that the formula we derive is a useful approximation for quantifying the level of third-order correlations in networks with a narrow out-degree distribution. In addition, we look at networks of highly interconnected hubs and show that, in this case, the average joint third cumulant depends strongly on the details of the connectivity pattern.

## Methods

### The Hawkes process as a model of spiking neural networks

To study higher-order correlations in networks of spiking neurons with a fixed network topology, we apply a point process model introduced in [27, 28], which we will refer to as the “Hawkes process”. As the theory of Hawkes processes is rich and rather technical, we will only summarize the important definitions and equations needed to present our results. A more formal and thorough treatment of the model can be found in Hawkes’ original papers.

In what follows, we will use capital letters to denote matrices. Vectors will not be explicitly marked, as their identity will be clear from the context. Individual components of matrices and vectors are referred to by indices attached to the symbol. Furthermore, note that, from here onwards, the phrase “third-order correlations” should always be interpreted as referring to “third-order joint cumulants” (defined below).

Our spiking neuronal network consists of  $N$  neurons, of which  $N_E$  are excitatory and  $N_I$  are inhibitory. Spike trains of neuron  $i$ ,  $S_i(t) = \sum_n \delta(t - t_n^i)$ , are modeled as realizations of point

processes with time-dependent firing rates  $\Lambda_i(t)$ . In other words, we have

$$\Lambda_i(t) = \mathbb{E}[S_i(t)|S_j(t'), t' \leq t, 1 \leq j \leq N], \tag{1}$$

where  $\mathbb{E}[\cdot]$  is the (conditional) expectation operator. In the Hawkes process framework, the vector  $\Lambda(t)$  of instantaneous firing rates (conditional on  $S_i(t')$ , for  $t' \leq t$ ) is given by

$$\Lambda(t) = \mu + \int_{-\infty}^t G(t-t') \cdot S(t') dt' \equiv \mu + (G \star S)(t). \tag{2}$$

The vector  $\mu$  can be interpreted as the rate of spontaneous activity (due to constant external input) in the network. The neurons in the network would independently spike at rates, given by components of vector  $\mu$ , if there were no synaptic connections between neurons in the network.

Recurrent synaptic interaction in the network is governed by the matrix of interaction kernels  $G(t)$ , an  $N \times N$  matrix of causal functions  $g_{ij}(t)$ , describing the influence of a spike in neuron  $j$  imposed on the future rate of neuron  $i$ . Typically, this is a sparse matrix with most entries being zero, and only few of them being nonzero. In principle, all of the functions  $g_{ij}(t)$  can be different. However, for the sake of simplicity, we will assume that all source neurons in the excitatory subpopulation have interaction kernels equal to  $g_E(t)$  to contact their targets, and all inhibitory neurons have interaction kernels  $g_I(t)$ . Thus, the total synaptic weight of excitatory neurons equals  $g_E \equiv \int g_E(t) dt$  and is positive, i.e.  $g_E > 0$ . Similarly, for inhibitory neurons,  $g_I \equiv \int g_I(t) dt < 0$ .

The number  $g_E$  represents the expected number of extra spikes in the postsynaptic (target) neuron induced by a spike of the presynaptic (source) neuron. Analogously, for inhibitory neurons, the number  $g_I$  represents the expected reduction in the total number of spikes produced by the postsynaptic neuron.

The exact connectivity between neurons in the network is chosen randomly, according to various rules, as will be explained in the sections to follow.

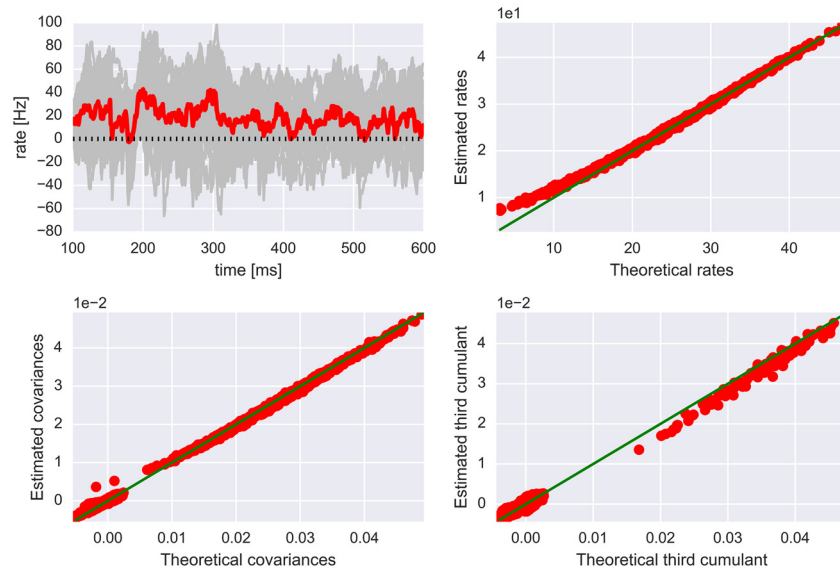
One important thing to note is that the Hawkes model only allows for pairwise interactions, and yet possesses correlations of all orders. Furthermore, the Hawkes process is a probabilistic spike generator and, as such, may exhibit a different behavior than an encoder with a deterministic threshold mechanism. It is, however, important to realize that real neurons that are embedded in a large network possess both stochastic and deterministic features. Another potential limitation of the Hawkes model is that it provides a good approximation when synapses are weak, but strong synapses may more thoroughly explore neuronal nonlinearities. Finally, the Hawkes process is formally correctly defined only for positive interaction kernels. Negative interactions may lead to a rate vector  $\Lambda(t)$  with negative entries, which is of course not a meaningful configuration. Thus, technically, one should use the rectified rate  $[\Lambda(t)]_+$  as a basis for spike generation in simulations. In the following, we will assume that the probability of having negative entries in the rate vector is negligibly low and will ignore the rectifying non-linearity. The goodness of this approximation is illustrated in Fig 1.

At equilibrium, the expected firing rate vector of the Hawkes process,  $\mathbb{E}[\Lambda(t)]$ , no longer depends on time. We can compute the stationary rate vector, denoted  $\Lambda$ , as follows

$$\Lambda = \mu + \int_{-\infty}^{+\infty} \Lambda G(t-t') dt' = \mu + \Lambda \left( \int_{-\infty}^{+\infty} G(t) dt \right), \tag{3}$$

from which we obtain the stationary rate of the network as

$$\Lambda = (\mathbb{I} - G)^{-1} \mu, \tag{4}$$



**Fig 1. Hawkes process theory reproduces rates, correlations and third cumulants in a simulated network with Erdős-Rényi type random connectivity.** Network parameters are  $N = 1000$ ,  $N_E = 800$ ,  $N_I = 200$ ,  $p = 0.1$ ,  $g_E = 0.015$  and  $g_I = -0.075$ . Top left: Fluctuating firing rates of 50 randomly chosen neurons (gray traces) and their average (red line). The average rate only rarely goes below 0 (dashed line). Top right: Estimated temporal averages of firing rates scattered vs. rates predicted by Hawkes theory. The diagonal (green line) indicates a perfect match. Note that there is a slight discrepancy between theory and simulation for very low rates. Bottom left: Estimated integrated pairwise covariances (of all possible neuron pairs) scattered vs. integrated covariances predicted by Hawkes theory. Bottom right: Estimated integrated joint third cumulants (see the following sections for a definition) of a 100 randomly chosen neurons, scattered vs. integrated joint cumulants computed from Hawkes theory. The larger discrepancies are due to finite simulation time and a relatively small sample size.

doi:10.1371/journal.pcbi.1004963.g001

where we have used  $G$  as a shortcut for the matrix of integrated interaction kernels, i.e.  $G \equiv \int G(t) dt$  and  $\mathbb{I}$  denotes the  $N \times N$  unit matrix. A summary of symbols, used in the text can be found in [Table 1](#).

In what follows, we will also restrict ourselves to systems in which the spectral radius of the matrix  $G$  (the largest eigenvalue of  $G$ ), which we denote by  $\rho(G)$ , is less than 1. Indeed, this condition insures the existence of the matrix inverse in the rate [Eq 4](#). Furthermore, if  $\rho(G) > 1$ , it may happen that no stable equilibrium of the system exists and the spiking activity exhibits runaway solutions.

### Pairwise correlations in the Hawkes process framework

An important result, originally presented in Hawkes' original work [27, 28], was that the lagged cross-covariance of spike trains of different neurons can be analytically computed directly from the matrix of interaction kernels  $G(t)$ . More precisely, we can formally define the covariance density matrix, denoted by  $C(\tau)$ , as

$$C(\tau) = \mathbb{E}[S(t + \tau)S(t)^T] - \Lambda\Lambda^T. \tag{5}$$

As was discussed before, intuitively, the entry  $(i, j)$  in  $C(\tau)$  can be thought of as representing the probability that a spike of neuron  $j$  causes a spike of neuron  $i$  after time lag  $\tau$ , minus the probability that this happens by chance (which, assuming stationarity, equals  $\Lambda\Lambda^T$ ). As noted in

**Table 1. Symbols used in text (in order of appearance).**

Symbol	Description
$N$	total number of neurons
$N_E$	number of excitatory neurons
$N_I$	number of inhibitory neurons
$S(t)$	spike train vector
$\Lambda(t)$	conditional firing rate vector
$G(t), g_{ij}(t)$	matrix of interaction kernels, its components
$(g_E) g_E(t)$	(integrated) excitatory neuron interaction kernel
$(g_I) g_I(t)$	(integrated) inhibitory neuron interaction kernel
$\mu, \mu_i$	external input vector, external input to neuron $i$
$\Lambda$	stationary firing rate vector, $\Lambda = \mathbb{E}[\Lambda(t)]$
$G$	integrated matrix of interaction kernels
$C(\tau)$	covariance density matrix
$R(t)$	convolution power series of the matrix $G(t)$
$\kappa^{ijk}(t_1, t_2, t_3)$	third-order joint cumulant density of neurons $i, j$ and $k$
$N_i(T)$	spike count of neuron $i$ in a bin of size $T$
$\kappa^{ijk}(T)$	third cumulant of spike counts of neurons $i, j$ and $k$
$\Psi(t)$	$\Psi(t) = R(t) - \mathbb{I}\delta(t)$
$\bar{\kappa}^{ijk}$	integrated joint third cumulant of neurons $i, j$ and $k$
$B$	power series of matrix $G$
$C$	integrated covariance density matrix
$N_{\text{pop}}(T)$	population spike count in a bin of size $T$
$\bar{\kappa}_3$	average joint third cumulant
$p$	connection probability
$\mu^{(k)}$	average common input, shared by $k$ neurons
$\tilde{\kappa}_3$	quadratic approximation of the average third cumulant $\bar{\kappa}_3$

doi:10.1371/journal.pcbi.1004963.t001

[27, 28], it is possible to rewrite  $C(\tau)$  as

$$C(\tau) = D\delta(\tau) + C_0(\tau) - \Lambda\Lambda^T, \tag{6}$$

where  $D \equiv \text{diag}(\Lambda)$  is a diagonal matrix, with the entries of the rate vector  $\Lambda$  on the diagonal. Furthermore,  $C_0(\tau)$  denotes the continuous part of the covariance density matrix, which is the solution to the matrix convolution equation

$$C_0(\tau) = G(\tau)D + (G \star C_0)(\tau), \tau > 0, \tag{7}$$

where the convolution of two matrix functions  $F(t)$  and  $G(t)$  equals a matrix function  $H(t) \equiv (F \star G)(t)$  with

$$H_{ij}(t) = \int_{-\infty}^t F(t-s) \cdot G(s) ds = \sum_k \int_{-\infty}^t F_{ik}(t-s) G_{kj}(s) ds, \tag{8}$$

where  $\cdot$  denotes the usual product of two numerical matrices. An important result in [28] is that the Fourier transform of the covariance density matrix, i.e.  $\hat{C}(\omega) \equiv \int_{-\infty}^{+\infty} C(\tau) e^{-i\omega\tau} d\tau$  can be expressed in terms of the Fourier transform  $\hat{G}(\omega)$  of the matrix of interaction kernels



$G(t)$ . More precisely, we have

$$\hat{C}(\omega) = (\mathbb{I} - \hat{G}(\omega))^{-1} D (\mathbb{I} - \hat{G}^*(\omega))^{-1}, \tag{9}$$

where  $*$  denotes the conjugate transpose of a matrix.

Recently, it has been shown [29, 30] that, component-wise and in the time domain, the previous equation can be written as

$$C_{ij}(\tau) = \sum_{k=1}^N \Lambda_k \int_{-\infty}^{+\infty} R_{ik}(u) R_{jk}(u + \tau) du, \tag{10}$$

where  $\Lambda_k$  is the  $k$ -th component of the previously defined stationary rate vector, and the matrix  $R(t)$  is a function of  $G(t)$ . Namely, we have that  $R(t)$  is a “convolution power series” of  $G(t)$  or, more precisely,

$$R(t) = \sum_{n \geq 0} G^{*n}(t). \tag{11}$$

Here, the matrix  $G^{*n}(t)$  denotes the  $n$ -th convolution power of the interaction kernel  $G(t)$ , defined recursively by

$$G^{*0}(t) = \mathbb{I} \delta(t), \tag{12}$$

$$G^{*n}(t) = \int_{-\infty}^t G^{*(n-1)}(t-s) \cdot G(s) ds, \quad n \geq 1, \tag{13}$$

where  $\cdot$  again denotes a matrix product. We have the following heuristic interpretation of the matrix elements  $R_{ij}(t)$ :

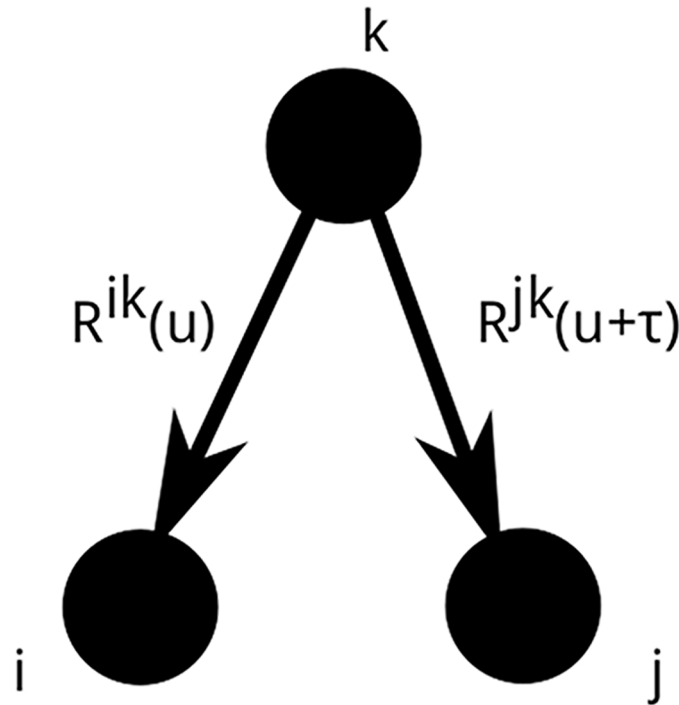
$$R_{ij}(t) dt \equiv P\{\text{spike of neuron } j \text{ at } 0 \text{ causes neuron } i \text{ to spike at } t\}. \tag{14}$$

This heuristic offer an interesting interpretation of Eq 10. Indeed, we can see the product  $\Lambda_k R_{ik}(u) R_{jk}(u + \tau) du$  as representing the probability that neuron  $k$ , spiking at its stationary rate  $\Lambda_k$ , causes neuron  $i$  to spike at  $u$  and neuron  $j$  at  $u + \tau$ . The covariance density  $C_{ij}(\tau)$  of neurons  $i$  and  $j$  at lag  $\tau$  is then nothing more than this probability, summed over all possible spikes times of neuron  $i$  (hence the integral w.r.t.  $u$ ) and over all possible “presynaptic” neurons  $k$ . Thus,  $C_{ij}(\tau)$  can be seen as a sum of all possible ways in which a neuron  $k$  can induce activity in neurons  $i$  and  $j$ , with spikes that are  $\tau$  apart.

Moreover, a simple graphical representation of  $C_{ij}(\tau)$  is now available. As was first shown in [3], the product  $\Lambda_k R_{ik}(u) R_{jk}(u + \tau) du$  can be represented as a rooted tree with leaves  $i$  and  $j$  (see Fig 2). Then, it can be shown that the lagged cross-covariance of spiking activity between neurons  $i$  and  $j$  is a sum of integral terms, each corresponding to a rooted tree with leaves  $i$  and  $j$  in the underlying network (for more details, see [3] and [29]).

### Higher order cumulants in analysis of network dynamics

We now move on to the problem of analyzing cumulants of higher order in networks of spiking neurons and introduce the tools necessary to do so. In statistics, a quantifier of third order correlations, analogous to the well-known covariance operator, is the third order joint cumulant, often denoted as  $\kappa_3[X, Y, Z]$ . It measures above-chance level third order dependence in the same way that covariance does for second order. It is defined, for random variables  $X, Y$  and  $Z$ ,



**Fig 2. Pictorial representation of terms contributing to the pairwise covariance density.** Each entry of  $C(\tau)$  is a weighted sum of integral terms, corresponding to rooted trees with 2 leaves,  $i$  and  $j$ .

doi:10.1371/journal.pcbi.1004963.g002

as (see [S3 Appendix](#) for a full derivation of the formula)

$$\kappa_3[X, Y, Z] = \mathbb{E}[XYZ] - \mathbb{E}[XY]\mathbb{E}[Z] - \mathbb{E}[XZ]\mathbb{E}[Y] - \mathbb{E}[YZ]\mathbb{E}[X] + 2\mathbb{E}[X]\mathbb{E}[Y]\mathbb{E}[Z]. \quad (15)$$

Let  $i, j$  and  $k$  be three distinct neurons in a recurrent neuronal network. Let further  $A = \{(i, t_1), (j, t_2), (k, t_3)\}$  denote a spike pattern, where neuron  $i$  spikes at time  $t_1$ , neuron  $j$  at  $t_2$  and neuron  $k$  at  $t_3$ . If we now plug in the variables  $S_i(t_1), S_j(t_2)$  and  $S_k(t_3)$  into [Eq 15](#) and denote

$$\kappa^{ijk}(t_1, t_2, t_3) \equiv \kappa_3[S_i(t_1), S_j(t_2), S_k(t_3)], \quad (16)$$

we see that the newly introduced function  $\kappa^{ijk}(t_1, t_2, t_3)$  measures the likelihood of the pattern  $A$  occurring not due to chance and not due to pairwise correlations.

Next, let  $N_i(T)$  represent the number of spikes of neuron  $i$  in a time bin of size  $T$ . Then, clearly,

$$N_i(T) = \int_0^T S_i(t) dt. \quad (17)$$

Now, using Fubini’s theorem, we find that

$$\kappa^{ijk}(T) \equiv \kappa_3[N_i(T), N_j(T), N_k(T)] = \int_0^T \int_0^T \int_0^T \kappa^{ijk}(t_1, t_2, t_3) dt_1 dt_2 dt_3. \quad (18)$$

In other words, while the function  $\kappa^{ijk}(t_1, t_2, t_3)$  encodes the probability of occurrence of a single pattern  $A$ , the “integrated cumulant”  $\kappa^{ijk}(T)$  (that is, the joint third cumulant of spike counts) measures the probability of the non-chance occurrence of any pattern of neurons  $i, j$  and  $k$  in a time bin of duration  $T$ . We will call the function  $\kappa^{ijk}(t_1, t_2, t_3)$  the (3rd order)



cumulant density, as one needs to integrate it in order to obtain the 3rd cumulant of spike counts, i.e.  $\kappa^{ijk}(T)$ .

Assuming stationarity, the density  $\kappa^{ijk}(t_1, t_2, t_3)$  can be written (with slight abuse of notation) as a function of only the (two) time lags between spike events at  $t_1, t_2$  and  $t_3$

$$\kappa^{ijk}(t_1, t_2, t_3) = \kappa^{ijk}(t_2 - t_1, t_3 - t_1) \equiv \kappa^{ijk}(\tau_1, \tau_2), \tag{19}$$

where we have defined  $\tau_1 = t_2 - t_1$  and  $\tau_2 = t_3 - t_1$ . In that case, we get (see [S1 Appendix](#))

$$\kappa^{ijk}(T) = \frac{\kappa_3[N_i(T), N_j(T), N_k(T)]}{T} = \int_{-T}^T \int_{-T}^T \kappa^{ijk}(\tau_1, \tau_2) d\tau_1 d\tau_2. \tag{20}$$

Thus, we obtain an alternative interpretation of  $\kappa^{ijk}(T)$ : It represents the third joint cumulant of spike counts of neurons  $i, j$  and  $k$  in a bin of size  $T$ , normalized by the bin size. As such, it is a quantity that can be easily computed from data, using unbiased estimators of higher-order cumulants, called  $k$ -statistics [31].

### Joint third cumulants in the Hawkes process framework

A recent result in the theory of Hawkes processes [29] shows that all 3rd order cumulant densities  $\kappa^{ijk}(t_1, t_2, t_3)$  can be computed, just as in the pairwise case, as sums of integral terms, each corresponding to a relevant topological motif (a subtree of the graph on which the process evolves), present in the underlying network. However, in the case of triplet correlations, the relevant rooted trees are somewhat more complicated (see [Fig 3](#)). Algebraically, we have

$$\begin{aligned} \kappa^{ijk}(t_1, t_2, t_3) &= \sum_{m=1}^N \Lambda_m \int_{-\infty}^{+\infty} R_{im}(t_1 - u) R_{jm}(t_2 - u) R_{km}(t_3 - u) du \\ &+ \sum_{m,n=1}^N \Lambda_n \int_{-\infty}^{+\infty} R_{in}(t_1 - u) \left( \int_{-\infty}^{+\infty} R_{jm}(t_2 - v) R_{km}(t_3 - v) \Psi_{mn}(v - u) dv \right) du \\ &+ \sum_{m,n=1}^N \Lambda_n \int_{-\infty}^{+\infty} R_{jn}(t_2 - u) \left( \int_{-\infty}^{+\infty} R_{im}(t_1 - v) R_{km}(t_3 - v) \Psi_{mn}(v - u) dv \right) du \\ &+ \sum_{m,n=1}^N \Lambda_n \int_{-\infty}^{+\infty} R_{kn}(t_3 - u) \left( \int_{-\infty}^{+\infty} R_{im}(t_1 - v) R_{jm}(t_2 - v) \Psi_{mn}(v - u) dv \right) du, \end{aligned} \tag{21}$$

where  $\Lambda_n$  (the stationary rate of neuron  $n$ ) and  $R_{ij}(t)$  (the rate change at time  $t$  in neuron  $i$  caused by a spike of neuron  $j$  at 0) have been defined previously, and

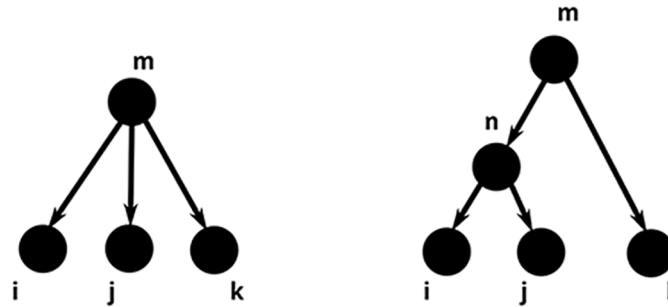
$$\Psi(t) = R(t) - \mathbb{I}\delta(t) = \sum_{n \geq 1} G^{*n}(t), \tag{22}$$

which, heuristically, simply means that

$$\Psi_{ij}(t) dt \equiv P\{\text{spike of neuron } j \text{ at } 0 \text{ causes neuron } i \neq j \text{ to spike at } t \neq 0\}. \tag{23}$$

Unfortunately, this formula is cumbersome, impractical and difficult to work with. However, a much more elegant expression is obtained if one considers the previously defined joint cumulants of spike counts,  $\kappa^{ijk}(T)$ . Formally, considering infinitely large time bins

$$\kappa^{ijk} \equiv \lim_{T \rightarrow +\infty} \kappa^{ijk}(T) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \kappa^{ijk}(\tau_1, \tau_2) d\tau_1 d\tau_2, \tag{24}$$



**Fig 3. Pictorial representation of terms contributing to  $\kappa^{ijk}(t_1, t_2, t_3)$ .** Each  $\kappa^{ijk}(t_1, t_2, t_3)$  is a weighted sum of integral terms, corresponding to rooted trees with leaves  $i, j$  and  $k$  (see Eq 21). The first term maps to the left tree, while the three remaining terms correspond to three possible ways in which three labeled leaves can be arranged into two groups to form the tree on the right. The first group would represent the daughter nodes of vertex  $m$ , and the second group would be a single child of the root node  $n$ .

doi:10.1371/journal.pcbi.1004963.g003

and letting  $B \equiv (\mathbb{I} - G)^{-1}$ , where  $G$  is the previously defined integrated matrix of interaction kernels, we have [29]

$$\begin{aligned} \kappa^{ijk} &= \sum_m \Lambda_m B_{im} B_{jm} B_{km} \\ &+ \sum_{m,n} \Lambda_n B_{im} B_{jm} (B_{mn} - \delta_{mn}) B_{kn} \\ &+ \sum_{m,n} \Lambda_n B_{jm} B_{km} (B_{mn} - \delta_{mn}) B_{in} \\ &+ \sum_{m,n} \Lambda_n B_{im} B_{km} (B_{mn} - \delta_{mn}) B_{jn}. \end{aligned} \tag{25}$$

This can be considered as a generalization of the pairwise correlation result from [3]. Indeed, if we let  $\omega = 0$  in Eq 9 and set  $C \equiv \hat{C}(0) = \int C(\tau) d\tau$ , we have

$$C_{ij} = BDB^* = \sum_{m=1}^N \Lambda_m B_{im} B_{jm}. \tag{26}$$

The problem, of course, is that the collection of all integrated cumulants  $\{\kappa^{ijk}\}_{i,j,k}$  represents a three-dimensional tensor, and as such cannot be represented in terms of a common matrix multiplication. For this reason, we must express  $\kappa^{ijk}$  as weighted sums and double sums of entries of the matrix  $B$  in formula 25.

### Populations cumulants as sums of joint cumulants of spike counts

Finally, let us touch upon the link between integrated covariances  $C_{ij}$ , cumulants  $\kappa^{ijk}$ , and moments of the population count distribution  $N_{\text{pop}}(T)$  which we define as the sum of activity of all neurons in the network

$$N_{\text{pop}}(T) \equiv \sum_{m=1}^N N_m(T). \tag{27}$$

From the general properties of cumulants [32], one can prove that

$$\lim_{T \rightarrow +\infty} \frac{\text{Var}[N_{\text{pop}}(T)]}{T} = \sum_{ij} C_{ij}. \quad (28)$$

In other words, the variance of the population activity is equal to the sum of all integrated covariances, normalized by bin size. Of course, this is only strictly true for infinitely large time bins, but we have found that Eq 28 is still a very good approximation whenever the size of bin  $T$  is much bigger than the temporal width of any entry in the matrix of interaction kernels  $G(t)$ .

Likewise, one can prove that

$$\lim_{T \rightarrow +\infty} \frac{\kappa_3[N_{\text{pop}}(T)]}{T} = \sum_{ij,k} \kappa^{ijk}. \quad (29)$$

Thus, the sums of all integrated cumulants of order 3 is equal to the third cumulant of population activity, normalized by bin size [31]. To understand why it is important to know the third cumulant  $\kappa_3[N_{\text{pop}}(T)]$  consider that, for a normally distributed random variable  $X$ , all cumulants of order 3 and higher are zero

$$X \sim \mathcal{N}(0, 1) \Rightarrow \kappa_n[X] = 0, \text{ for all } n \geq 3. \quad (30)$$

Therefore, in a sense, non-zero cumulants of order 3 and higher measure the departure from normality of the variable  $N_{\text{pop}}(T)$ . Furthermore, in statistics, a measure of skewness of the distribution of a random variable  $X$  is defined as the (scaled) third cumulant  $\kappa_3[X]$ . As the Gaussian distribution is symmetric about 0 (and thus  $\kappa_3[X] = 0$ ), any significant deviation of  $\kappa_3[N_{\text{pop}}(T)]$  indicates right (negative) or left (positive) skewness.

## Simulation and data analysis details

The simulation of linearly interacting point processes was conducted using the NEST simulator [33]. We simulated a network of 1000 neurons, of which 800 were excitatory and 200 inhibitory. The spikes of each neuron were generated according to a time-dependent rate function  $\Lambda(t)$ , defined by Eq 2. Negative values of  $\Lambda(t)$  were rectified to zero, resulting in no spike output. Neurons received external Poissonian drive with constant rate of 10 Hz. Incoming spikes induced an increment of amplitude 1.5 Hz and  $-7.5$  Hz for excitatory and inhibitory spikes, respectively, which decayed with a time constant of 10 ms. In the Hawkes process framework, this corresponds to an exponential interaction kernel with total integral  $g_E = 0.015$  and  $g_I = -0.075$ , respectively. The synaptic delay was set to 2 ms. The simulation time step was 0.1 ms. The total simulation time was 5000 s =  $5 \cdot 10^6$  ms.

Spike data from simulations were sampled in time bins of duration  $T = 100$  ms, producing  $5 \cdot 10^4$  bins. We found that the theoretical results concerning infinite sized bins are still largely valid when the bin size  $T$  is at least one order of magnitude larger than the interaction kernel time constant. The rates, pairwise covariances and joint third cumulants were estimated from a data matrix with 1000 rows (representing individual neurons) and 50000 columns (representing time bins) using  $k$ -statistics [31], which are known to be unbiased estimators of cumulants of any order. Note that pairwise covariances are nothing more than joint cumulants of second order.

## Results

### Weights of subtrees in the network determine the strength of triplet correlations

In this section we explain how recurrent connectivity affects joint third cumulants of triplets of neurons in a spiking neuronal network. As was mentioned before, the matrix of integrated interaction kernels  $G$  can be interpreted as an effective connectivity matrix, as each entry  $(i, j)$  represents the excess number of spikes in neuron  $i$ , caused by an individual spike in neuron  $j$ . With this in mind, let us now take a moment to develop a topological interpretation of Eq 25. Firstly, as  $\rho(G) < 1$  has been assumed, we have a power series expansion for the matrix  $B = (\mathbb{I} - G)^{-1}$ , namely  $B = \sum_n G^n$ . In order to develop intuition, we first consider what happens to Eq 26 when we plug the power series expansion of  $B$  into it (as was done in [3]). The formula for  $C_{ij}$  reads

$$C_{ij} = \sum_{m=1}^N \sum_{r=0}^{+\infty} \sum_{s=0}^{+\infty} \Lambda_m G_{im}^r G_{jm}^s. \tag{31}$$

We now interpret the matrix  $G^r$  in the sense of graph theory, i.e. as a matrix whose entry  $(i, j)$  corresponds to the sum of compound weights of all paths from node  $j$  to node  $i$  in exactly  $r$  steps. Indeed, a typical entry of matrix  $G^r$  equals

$$G_{ij}^r = \sum_{k_1, k_2, \dots, k_{r-1}} G_{ik_1} G_{k_1 k_2} \dots G_{k_{r-1} j}. \tag{32}$$

We observe that each of the summands in the above equation is the average number of excess spikes, caused by an individual length  $r$  chain of spiking events, originating in neuron  $j$ . The entry  $G_{ij}^r$  is then the sum over all such chains, i.e. over all possible intermediary neurons  $k_1, k_2, \dots, k_{r-1}$ . Thus, a procedure for computing  $C_{ij}$  would go as follows:

1. Pick a “root neuron”  $m$
2. Create a “spiking chain” from neuron  $m$  to neuron  $i$  that is  $r$  synaptic steps long
3. Create a “spiking chain” from neuron  $m$  to neuron  $j$  that is  $s$  neurons long
4. Compute the weight of the subtree defined in this way by multiplying together the weights of the branches (given by  $G_{im}^r$  and  $G_{jm}^s$ )
5. Multiply everything by the “weight of the root node”, which we can formally define to be  $\Lambda_m$

Note that  $r = 0$  ( $s = 0$ ) is a distinct possibility (as the first term in the power series expansion of  $B$  is  $G^0 \equiv I$ ). In that case, we identify neurons  $m$  and  $i$  ( $m$  and  $j$ ) and our “two-pronged tree” becomes a single branch with neuron  $i$  ( $j$ ) on top and neuron  $j$  ( $i$ ) on the bottom.

Our previous discussion shows that the integrated covariance density  $C_{ij}$  can be equivalently expressed as

$$C_{ij} = \sum_{T \in \mathcal{T}_{ij}^m} w(T), \tag{33}$$

where the sum goes over the set  $\mathcal{T}_{ij}^m$  of all rooted trees  $T$  with root  $m$ , containing nodes  $i$  and  $j$ . Here,  $w(T)$  denotes the weight of tree  $T$ , defined as the product of weights of all edges, contained in  $T$ , times the weight of the root  $m$ , defined as being equal to  $\Lambda_m$ .

Now, since, in the stationary case (see [S1 Appendix](#))

$$C_{ij} = \lim_{T \rightarrow +\infty} \frac{\text{cov}[N_i(T), N_j(T)]}{T}, \quad (34)$$

we have that, for infinitely large time bins, the probability (normalized by bin size) of the non-chance occurrence of ANY pattern of neurons  $i$  and  $j$  in a bin of size  $T$  can simply be computed as the sum of weights of ALL possible rooted trees with leaves  $i$  and  $j$ . Thus, in a nutshell, the only way pairwise interaction can arise between neurons  $i$  and  $j$  is through shared input by a neuron  $k$ , that can be arbitrarily far upstream from both  $i$  and  $j$ . This is the main result of [\[3\]](#).

With our intuition primed by consideration of the simpler, pairwise correlation case, we are ready to tackle the computation of  $\kappa^{ijk}$ . Once again, plugging the power series expansion of matrix  $B$  into [Eq 25](#) yields

$$\kappa^{ijk} = \sum_{T \in \mathcal{T}_{ijk}^m} w(T), \quad (35)$$

where  $\mathcal{T}_{ijk}^m$  is the set of all rooted trees with root  $m$  containing nodes  $i, j, k$ , and  $w(\cdot)$  is the already defined weight function. As we have that (see [S1 Appendix](#))

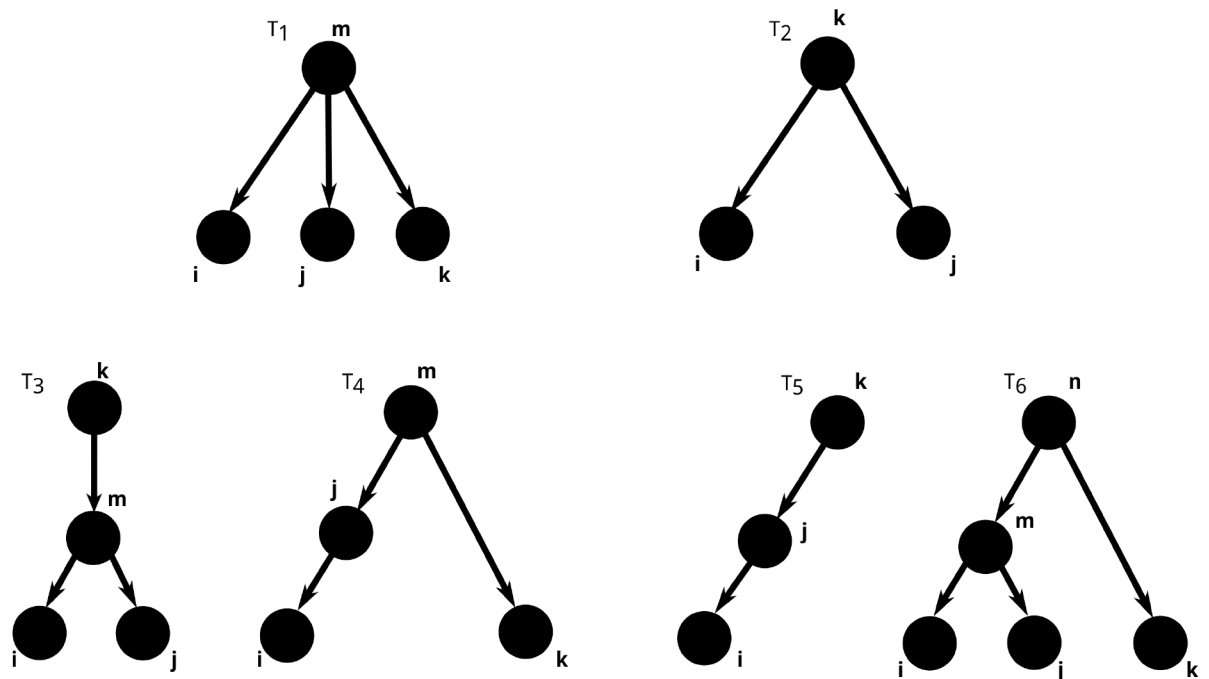
$$\kappa^{ijk} = \lim_{T \rightarrow +\infty} \frac{\kappa_3[N_i(T), N_j(T), N_k(T)]}{T}, \quad (36)$$

the interpretation of the “sum over trees” formula is analogous. In other words, for infinitely large time bins, the probability (normalized by bin size) of the non-chance occurrence of ANY pattern of neurons  $i, j$  and  $k$  in a bin of size  $T$  can simply be computed as the sum of weights of ALL possible rooted trees, containing nodes  $i, j$  and  $k$ . The only difference from the pairwise correlation case is that the topological motifs contributing to triplet correlations are different and more numerous.

What are the subtrees, contributing to  $\kappa^{ijk}$ ? We can get our first hint by comparing the [formula 25](#) and the trees in [Fig 3](#). Indeed, the first term in [Eq 25](#) corresponds to the left, “three-pronged” tree in [Fig 3](#)—in fact, it is the combined weight of all such structures found in the graph with adjacency matrix  $G$ , summed over all possible identities of the root node  $m$  and over all possible lengths of the tree branches terminating at  $i, j$  and  $k$ . However, as any of the three branches can also be of length 0, the left tree in [Fig 3](#) actually represents 4 different contributions to  $\kappa^{ijk}$ , one corresponding to the tree depicted, in which case all of the branches are of length at least 1, and three other “two-pronged” trees obtained by collapsing one of the three branches and identifying the node  $m$  with node  $i, j$  or  $k$  (see first row of [Fig 4](#)). Algebraically, this can also be seen by replacing one of the  $B$  matrices in the first row of [formula 25](#) by the identity matrix  $\mathbb{I}$ . Indeed, placing  $\mathbb{I}$  instead of  $B$  in any of the tree slots yields three possible contractions.

In the right tree in [Fig 3](#), each of the last three terms in [Eq 25](#) corresponds to one copy of it, the only difference among them being the label of the rightmost node. Indeed, the second term represents a tree in which the rightmost node is labeled  $k$ , for the third term the rightmost node is  $i$ , and for the last one it is  $j$ . Each of these terms contains three  $B$  matrices, and thus, each of these three terms will yield three additional trees whose weight will contribute to the overall sum, defining  $\kappa^{ijk}$  (see the second row of [Fig 4](#)). Like before, all of these are obtained by replacing one of the  $B$  matrices with the identity matrix  $\mathbb{I}$  and performing the corresponding summation.

Notice that the last three terms in [Eq 25](#) also depend on entries of the matrix  $B - \mathbb{I}$ . This signifies the fact that the link between nodes  $n$  and  $m$  in [Fig 3](#) can only “telescope out”, i.e. it



**Fig 4. The pictorial representation of all terms, contributing to  $\kappa^{ijk}$ .** The tree shapes depicted in this figure were obtained from those in Fig 3 by performing all possible contractions of branches (see text).

doi:10.1371/journal.pcbi.1004963.g004

cannot be contracted to 0 (indeed, it corresponds to the power series  $\sum_{n \geq 1} G^n$  in which the term of order 0 is not present). For reasons as to why this branch does not allow contractions, see [29].

To summarize, the six different tree shapes depicted in Fig 4 all contribute terms that, when summed up, yield  $\kappa^{ijk}$ . Likewise, as was mentioned previously, each branch, incident to each of the trees pictured, can have arbitrarily many intermediate nodes in between the two vertices shown.

### Average third cumulants in large random networks depend on the presence of a particular subtree

We are interested in computing the average third cumulant in the network, defined as

$$\frac{1}{N^3} \sum_{i,j,k} \kappa^{ijk}, \tag{37}$$

where  $\kappa^{ijk}$  represent the integrated joint third cumulants of neurons  $i, j$  and  $k$ , considered previously. From Eq 29, we have that the previous sum equals

$$\lim_{T \rightarrow +\infty} \frac{\kappa_3[N_{\text{pop}}(T)]}{TN^3}, \tag{38}$$

the third cumulant of population activity for an infinitely large time bin  $T$ , normalized by network size and bin width.

Note that the sum in Eq 37 goes over ALL indices  $i, j$  and  $k$ . Thus, we have three distinct cases:

1. The three indices are all distinct ( $i \neq j, j \neq k, k \neq i$ );

2. the three indices are all equal ( $i = j = k$ );
3. two of three indices are equal, with the third being distinct ( $i = j$  and  $j \neq k$ , or a permutation thereof).

The number of summands in the first case is equal to  $N(N - 1)(N - 2)$ , in the second case it is simply  $N$ , and in the third one it equals  $3N(N - 1)$ . Thus, we have

$$\bar{\kappa}_3 = \frac{1}{N^3} \left[ \sum_{i,j,k} \kappa^{ijk} + \sum_{i,j} \kappa^{ijj} + \sum_i \kappa^{iii} \right]. \tag{39}$$

In the limit of large networks, the first term becomes dominant, as

$$\lim_{N \rightarrow +\infty} \frac{3N(N - 1)}{N^3} = 0, \quad \lim_{N \rightarrow +\infty} \frac{N}{N^3} = 0, \quad \text{but} \quad \lim_{N \rightarrow +\infty} \frac{N(N - 1)(N - 2)}{N^3} = 1. \tag{40}$$

Therefore, in all calculations that follow, we will assume that  $i, j$  and  $k$  are all different

$$\bar{\kappa}_3 = \frac{1}{N^3} \sum_{i \neq j \neq k} \kappa^{ijk}. \tag{41}$$

Furthermore, we assume the following about the underlying network topology:

1. (Random network condition) Every node  $j$  has probability  $p$  of forming a connection with any of the other  $N - 1$  nodes.
2. (Generalized Dale’s law) To each node  $j$  we assign a *type*  $l \in L$  such that, for a fixed  $j$  we have  $\forall i, g_{ij} = g_l$ .

In other words, the probability of a directed connection between any pair of nodes is equal to  $p$ , and each node is of a *single type*  $l$  and as such, only makes outgoing connections of type  $l$ . Here,  $L$  denotes the set of *type labels*.

The derivations that follow can still be done under these general assumptions. Also, note that, even though the first assumption allows for random topologies, the results obtained in this section hold true for regular networks as well, as very large random networks are approximately regular. However, in the interest of concreteness, we will assume that  $L = \{E, I\}$ . In short, each node  $j$  can either be of type  $E$  (excitatory) or type  $I$  (inhibitory). Thus, for a given “excitatory” node  $j$ ,  $g_{ij}$  is either 0 (with probability  $1 - p$ ) or  $g_E$  (with probability  $p$ ), for every neuron  $i$ . Likewise if the neuron is inhibitory (in that case,  $g_{ij}$  equals  $g_I$ ).

We now compute the average input to a neuron, embedded in the network. First, we note that, mathematically, the total input to node  $i$  can be computed as  $\sum_j G_{ij}$ . Given our previous considerations, we have that the total input equals

$$p(N_E g_E + N_I g_I) = Np \left( \frac{N_E}{N} g_E + \frac{N_I}{N} g_I \right) \equiv N\mu^{in}, \tag{42}$$

where  $N_E$  and  $N_I$  are the numbers of excitatory and inhibitory neurons in the network, respectively. We have also  $\mu^{in}$  as  $p \left( \frac{N_E}{N} g_E + \frac{N_I}{N} g_I \right)$ , the average strength of the total input to a neuron. Now, if we set the external input  $\mu$  to 1, the stationary rate of neuron  $i$  can be seen to equal

$$\Lambda_i = \sum_j \left[ \sum_n G_{ij}^n \right] = \sum_j (\delta_{ij} + G_{ij} + G_{ij}^2 + \dots) = \frac{1}{1 - N\mu^{in}} \equiv \bar{\Lambda}. \tag{43}$$

Unsurprisingly, since the external input to all neurons is the same, the stationary rates are all



equal ( $\Lambda_i = \bar{\Lambda}, \forall i$ ). The computation of the average cumulant  $\bar{\kappa}_3$  can be done in much the same way (for details, see [S2 Appendix](#)). Note that, to simplify derivation, we assume that all neurons (irrespective of their type) have the same in-degree and out-degree. The final formula then reads

$$\begin{aligned} \bar{\kappa}_3 = & -\frac{\bar{\Lambda}}{N^3} \frac{N^4 p^2 \mu^{(3)}}{(1 - \mu^{(1)}N)^3} + \frac{3\bar{\Lambda}}{N^3} \frac{N^3 p \mu^{(2)}}{(1 - \mu^{(1)}N)^2} - \frac{3\bar{\Lambda}}{N^3} \frac{N^4 p \mu^{(1)} \mu^{(2)}}{(1 - \mu^{(1)}N)^3} \\ & - \frac{6\bar{\Lambda}}{N^3} \frac{N^4 p \mu^{(1)} \mu^{(2)}}{(1 - \mu^{(1)}N)^3} + \frac{6\bar{\Lambda}}{N^3} \frac{N^3 [\mu^{(1)}]^2}{(1 - \mu^{(1)}N)^2} + \frac{3\bar{\Lambda}}{N^3} \frac{N^5 p^3 [\mu^{(2)}]^2}{(1 - \mu^{(1)}N)^4}, \end{aligned} \tag{44}$$

where each term in the equation corresponds to one of the tree shapes in [Fig 4](#). We have chosen not to perform any simplifications in the formula, as we feel that this would obscure the correspondence each term has to its tree counterpart. Here, we have defined  $\mu^{(k)}$  as the average common input, shared by  $k$  neurons, equaling

$$\mu^{(k)} = p \left( \frac{N_E}{N} g_E^k + \frac{N_I}{N} g_I^k \right). \tag{45}$$

Note that in this formalism,  $\mu^{(1)}$  is the “average common input shared by one neuron”, equal to  $\mu^{in}$ , the average total input to a neuron. The precise nature of this relation between [formula 44](#) and the topology of specific trees is covered in [S2 Appendix](#). However, heuristically, the relationship is as follows

- The exponent of  $N$  in the formula counts the number of nodes of a particular tree
- The exponent of  $p$  is one less than the number of leaves of the tree
- The  $\mu^{(i)}$  terms each correspond to an internal node of the tree (that is, a node that is not a leaf). The number in parenthesis in the superscript denotes the out-degree of that particular internal node. Thus, for example,  $\mu^{(3)}$  indicates that the particular tree has an internal node with out-degree 3.
- The power  $k$  of the normalization factor  $\frac{1}{(1 - \mu^{(1)}N)^k}$  encodes the number of edges in the tree

[Eq 44](#) can be used as an approximation whenever the degree distribution of the network in question is narrow—formally, it is only exactly true for a regular network, in which all neurons have the same in- and out-degrees. For large random networks of the Erdős-Rényi type, this is true as the resulting Binomial distributions have a standard deviation that vanishes with increasing network size. The numerical efficacy of such an approximation can be found in the following section.

A final thing to note about [Eq 44](#) is what happens when  $N \rightarrow +\infty$ . Firstly, note that, once we perform all possible cancellations of terms in [eq 44](#), we find, after rearranging

$$\bar{\kappa}_3 = 3p^3 [\mu^{(2)}]^2 N^2 \bar{\Lambda}^5 - (9p\mu^{(1)} \mu^{(2)} + p^2 \mu^{(3)}) N \bar{\Lambda}^4 + (3p\mu^{(2)} + 6[\mu^{(1)}]^2) \bar{\Lambda}^3. \tag{46}$$

Thus, in the limit of large networks, the most important term is the one corresponding to tree  $T_6$  in [Fig 4](#)

$$\tilde{\kappa}_3 \equiv \frac{3\bar{\Lambda}}{N^3} \frac{N^5 p^3 [\mu^{(2)}]^2}{(1 - \mu^{(1)}N)^4} = 3p^3 [\mu^{(2)}]^2 N^2 \bar{\Lambda}^5, \tag{47}$$

since we have  $1/(1 - \mu^{(1)}N) = \bar{\Lambda}$ . More precisely, we obtain the relation

$$\bar{\kappa}_3 = \tilde{\kappa}_3 + O(N) + O(1). \tag{48}$$

As is now evident, the contributions from all trees of this shape to  $\bar{\kappa}_3$  grows as a quadratic function of  $N$ . The reason for this is that, in large networks, the number of “more complicated” subgraphs grows faster than the number of simpler ones. To see why this is true, consider counting all possible trees with  $k$  nodes and  $k - 1$  edges in a random graph. Since each edge is generated independently, the number of such trees equals

$$\binom{N}{k} p^{k-1} (1 - p)^{\binom{k}{2} - k + 1}, \tag{49}$$

Thus, as long as  $k \leq \lfloor N/2 \rfloor$ , the number of tree structures with  $k$  nodes in a random graph of size  $N$  will increase with increasing  $k$ . This is, in a nutshell, why the most relevant contribution to  $\bar{\kappa}_3$  comes from the “most complicated” tree, i.e.  $T_6$ .

With the previous discussion in mind, one may expect that, for  $N \rightarrow +\infty$ , the quadratic term  $\tilde{\kappa}_3$  is a good approximation for  $\bar{\kappa}_3$ . Indeed, Fig 5 illustrates this. Thus, we are able to conclude that, in the limit of large networks, the dominating contribution to the average joint third cumulant  $\bar{\kappa}_3$  comes from the trees of topology  $T_6$  present in the network. One more thing to note is that the leading term  $\tilde{\kappa}_3$  is proportional to a power of the stationary rate  $\bar{\Lambda}$ . Let us briefly consider what happens to  $\bar{\Lambda}$  in very large networks, for  $N \rightarrow +\infty$ . We have

$$\bar{\Lambda} = \frac{1}{1 - Np \left( \frac{N_E}{N} g_E + \frac{N_I}{N} g_I \right)} \rightarrow 0, N \rightarrow +\infty, \tag{50}$$

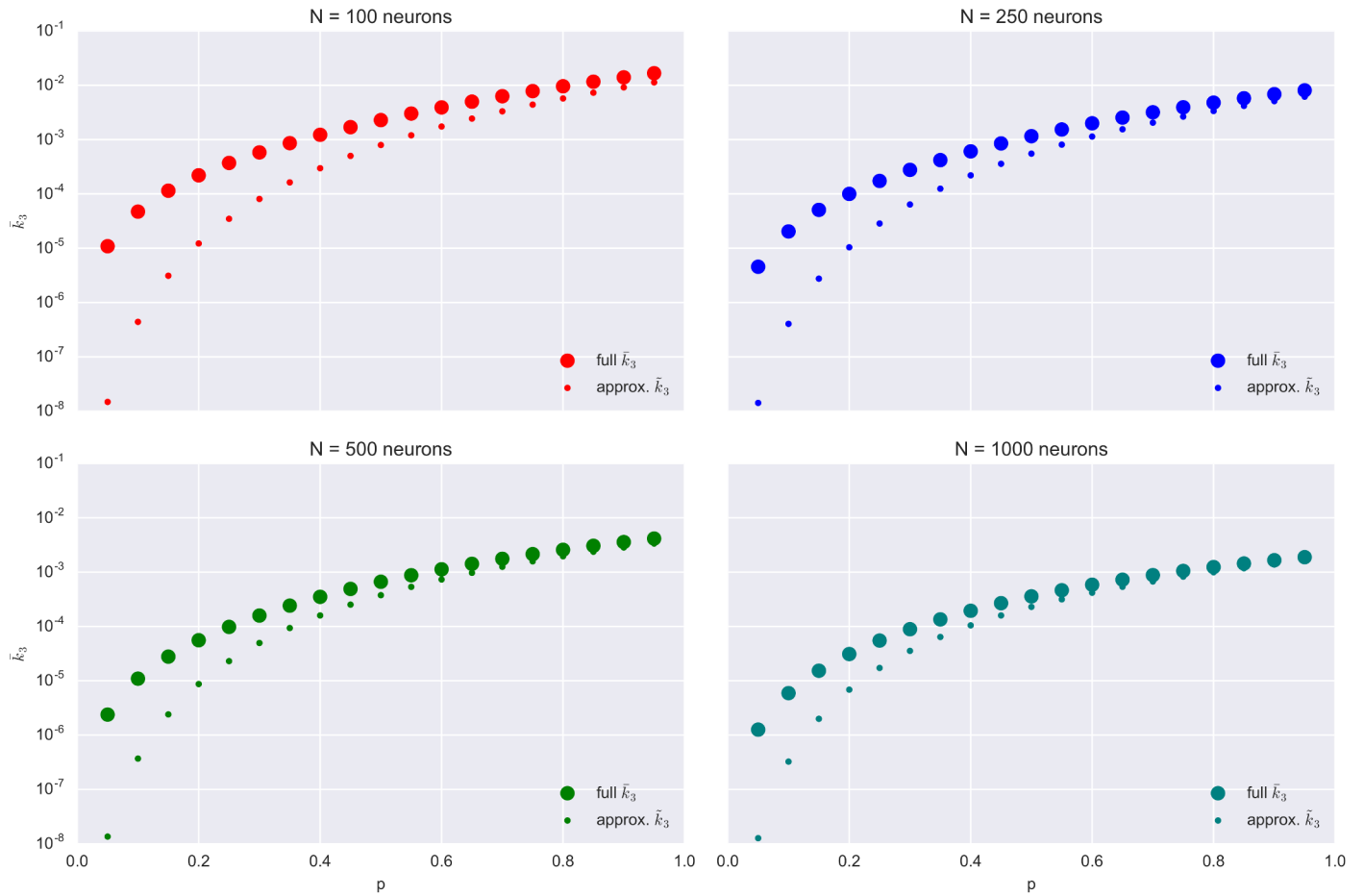
assuming we keep all other parameters fixed. As a result of this, the product  $N^2 \bar{\Lambda}^5$  in  $\tilde{\kappa}_3$ , will decay to zero with increasing network size. Thus, when the size of the network considered grows without bounds, two things happen:

- The leading term  $\tilde{\kappa}_3$  becomes a better approximation of the average third cumulant  $\bar{\kappa}_3$  but, at the same time,
- The average third cumulant  $\bar{\kappa}_3$  itself goes to 0.

The second point shouldn't be too surprising. Indeed, once we remember that  $\bar{\kappa}_3$  is proportional to the skewness of the population activity (defined as the sum of spike counts of all neurons in the network, in a bin of size  $T$ ), its asymptotic vanishing is a straightforward consequence of the Central Limit Theorem. As  $N$  increases, the population activity is behaving more and more like a Gaussian random variable and, as a consequence, its skewness inevitably decays to zero. This effect is reflected by the horizontal asymptote in Fig 5.

### The signs of tree motif contributions to the third cumulant in regular networks depend on their topological structure

In this section, we will analyze the contributions of terms, corresponding to tree shapes in Fig 4 with fixed branch length. More precisely, let us consider once again Eq 25, plugging in the power series expansion of matrix  $B$  and exchanging the order of summation over “branch length” (i.e. summation over powers of the  $G$  matrix) and summation “over nodes” (i.e.



**Fig 5. Efficacy of the quadratic approximation to  $\bar{\kappa}_3$  (Eq 47) for different network sizes  $N$ .** All four panels:  $\bar{\kappa}_3$  and its quadratic approximation,  $\tilde{\kappa}_3$ , plotted for different values of the connection probability  $\rho$ . Each panel corresponds to a network of a given size.

doi:10.1371/journal.pcbi.1004963.g005

summation over  $i, j$  and  $k$ , used to define  $\bar{\kappa}_3$ ), we get

$$\begin{aligned}
 \bar{\kappa}_3 &= \frac{\bar{\Lambda}}{N^3} \sum_{l_1, l_2, l_3} \left[ \sum_{i, j, k, m} G_{im}^{l_1} G_{jm}^{l_2} G_{km}^{l_3} \right] \\
 &+ \frac{3\bar{\Lambda}}{N^3} \sum_{l_1, l_2} \left[ \sum_{i, j, k} G_{ik}^{l_1} G_{jk}^{l_2} \right] \\
 &+ \frac{3\bar{\Lambda}}{N^3} \sum_{l_1, l_2, l_3} \left[ \sum_{i, j, k, m} G_{im}^{l_1} G_{jm}^{l_2} G_{mk}^{l_3} \right] \\
 &+ \frac{6\bar{\Lambda}}{N^3} \sum_{l_1, l_2, l_3} \left[ \sum_{i, j, k, n} G_{ij}^{l_1} G_{jn}^{l_2} G_{kn}^{l_3} \right] \\
 &+ \frac{6\bar{\Lambda}}{N^3} \sum_{l_1, l_2} \left[ \sum_{i, j, k} G_{ij}^{l_1} G_{jk}^{l_2} \right] \\
 &+ \frac{3\bar{\Lambda}}{N^3} \sum_{l_1, l_2, l_3, l_4} \left[ \sum_{i, j, k, m, n} G_{im}^{l_1} G_{jm}^{l_2} G_{mn}^{l_4} G_{kn}^{l_3} \right].
 \end{aligned} \tag{51}$$

The terms in the square brackets can be interpreted as the total weight of all relevant trees (see Fig 4) present in the network, with lengths of all branches fixed. Under the regularity assumption, i.e. if all neurons have the same in-degree and out-degree, it is straightforward to conclude that the “square bracket term” of a tree  $\mathcal{T}$  with  $n$  nodes and  $l$  leaves, embedded in a network of size  $N$ , can be computed as (see S2 Appendix)

$$N^n p^{l-1} \prod_v \mu^{(k_v)} [\mu^{(1)} N]^{l_1 + \dots + l_{n-1} - n + 1}, \tag{52}$$

where the product is over all internal nodes (i.e. nodes that are not leaves) of  $\mathcal{T}$  and  $k_v$  is the out-degree of node  $v$ . The numbers  $l_1, \dots, l_{n-1}$  encode the lengths of branches of  $\mathcal{T}$ , of which there are exactly  $n - 1$  in a tree with  $n$  nodes. In fact, it is this result that greatly simplifies the “summation over branch lengths” one needs to perform in order to obtain Eq 44.

Furthermore, from formula 52 we see that the only relevant characteristics of a tree  $\mathcal{T}$  that determine the weight of the contribution are the number of its nodes  $n$ , the number of its leaves  $l$  and the out-degrees of its internal nodes. Note that the root counts as an internal node here. Trees with a large total branch length contribute relatively little to  $\bar{\kappa}_3$ . Indeed, as

$$|\mu^{(1)} N| = |p(N_E g_E + N_I g_I)| < 1, \tag{53}$$

we have that, when the total length of all branches tends to infinity (i.e. when the sum  $s_n \equiv l_1 + \dots + l_{n-1}$  grows beyond all bounds), the corresponding term  $(\mu^{(1)} N)^{s_n}$  decays to zero.

Lastly, we consider the issue of determining the signs of various contributions to  $\bar{\kappa}_3$ . This can be done by once again closely analyzing formula 52. First, note that the common input terms  $\mu^{(k)}$  are positive for even and negative for odd  $k$ . Indeed, as we assume that underlying network is inhibition-dominated (that is, if we assume that the total input to a neuron is negative) we have, in mathematical terms that

$$N_E g_E + N_I g_I < 0 \quad \Leftrightarrow \quad g_I < -\frac{N_E}{N_I} g_E. \tag{54}$$

Thus,

$$\begin{aligned} \mu^{(2r+1)} &= p \left( \frac{N_E}{N} g_E^{2r+1} + \frac{N_I}{N} g_I^{2r+1} \right) \\ &< p \left( \frac{N_E}{N} g_E^{2r+1} + (-1)^{2r+1} \frac{N_E}{N} \left( \frac{N_E}{N_I} \right)^{2r} g_E^{2r+1} \right). \end{aligned}$$

Therefore,

$$\mu^{(2r+1)} < p \frac{N_E}{N} \left[ 1 - \left( \frac{N_E}{N_I} \right)^{2r} \right] g_E^{2r+1} < p \left[ 1 - \left( \frac{N_E}{N_I} \right)^{2r} \right] g_E^{2r+1} < 0, \tag{55}$$

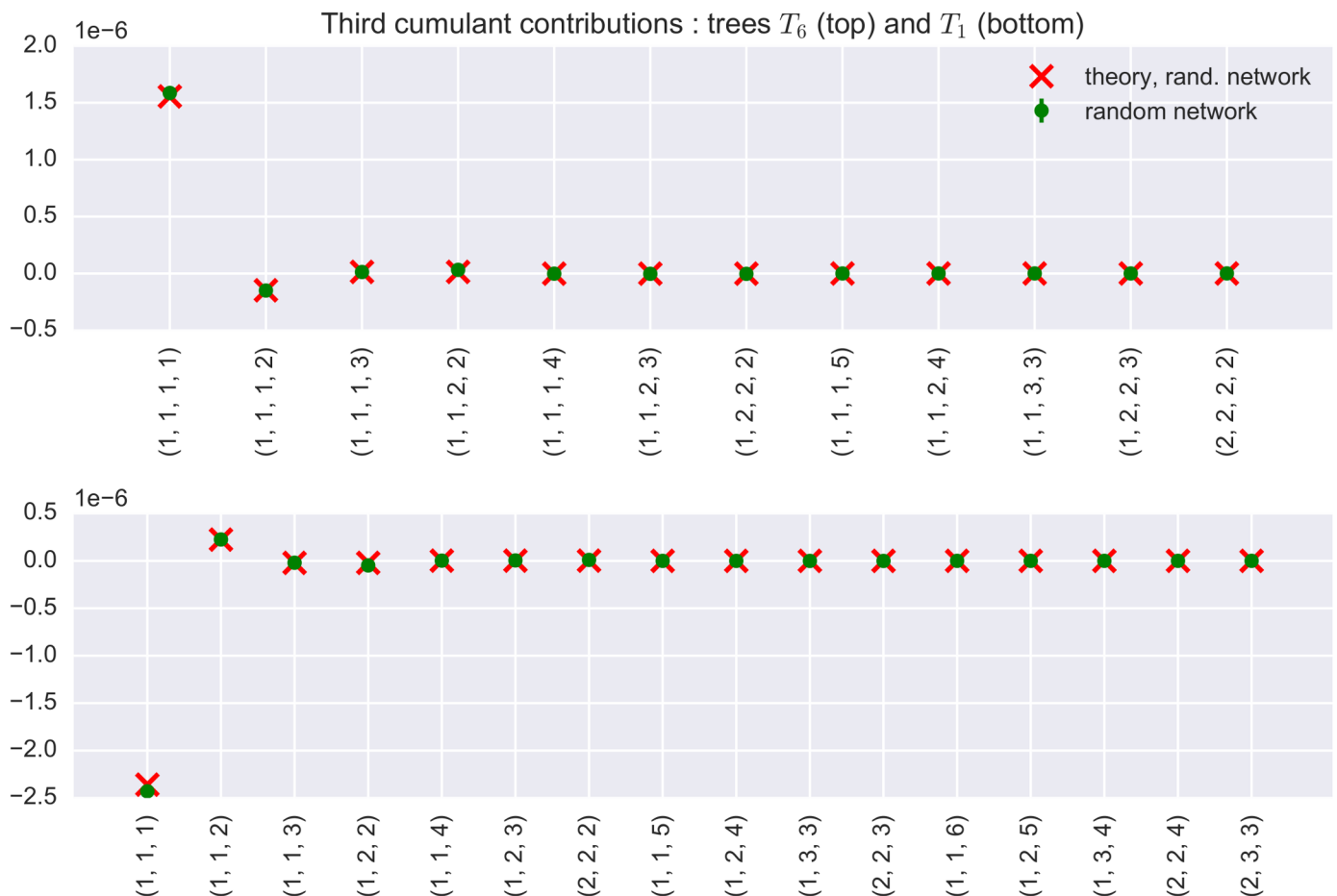
since  $g_E > 0$ ,  $N_E > N_I$  and  $0 \leq p \leq 1$ . In the same way, one can show that  $\mu^{(2r)} > 0$ . Therefore, the out-degree sequence of the internal nodes of the tree affects the sign of the corresponding contribution. If, for example, the tree has two internal nodes, with out-degrees 1 and 2, respectively, this will contribute an overall negative sign to the term. However, the out-degree sequence alone does not completely determine the sign of the contribution. Another factor is the parity of the total length of all branches, i.e. the sum  $s_n \equiv l_1 + \dots + l_{n-1}$ . To see why, note

that  $N\mu^{(1)} < 0$ , by our previous discussion, and likewise

$$[\mu^{(1)}N]^{s_n-n+1} \tag{56}$$

is either negative or positive, depending on whether  $s_n = 2r + 1$  or  $s_n = 2r$ . (Note that  $s_n \geq n - 1$ .)

To summarize, the resulting sign of the total contribution to the average third cumulant, of a specific tree with  $n$  nodes,  $l$  leaves, a given out-degree sequence and branch lengths depends on both the parity of the product of the internal node out-degrees and the parity of the total branch length. What this means in practice is that the presence of certain trees increases the overall level of third order correlation, while the existence of others can actually have the opposite effect. Whether the latter or the former is the case depends solely on the tree's topological structure, i.e. how the internal nodes branch and how many edges it contains. As an illustration, the signs and sizes of contributions of two sample trees in a recurrent random network are depicted in Fig 6. One can clearly see which trees increase third-order correlations in the network, and which trees actually decrease them.



**Fig 6. Contributions of some tree structures to the average third cumulant in a random network.** Top: Theoretical (narrow out-degree distribution approximation) and sample contributions to the average third cumulant of  $T_6$  (see Fig 4) tree topologies with fixed branch lengths. Ticks on the x-axis code for the lengths of the four branches of the tree. The ordering of the indices in the tick labels is done in a top-to-bottom and left-to-right fashion. More precisely, the first number corresponds to the length of the “leftmost” branch emanating from the root node, and so forth. The sample contributions were computed as averages of 3 independent realizations of a random network. Bottom: Theoretical and sample contributions to the average third cumulant of  $T_1$  tree topologies with fixed branch lengths.

doi:10.1371/journal.pcbi.1004963.g006

One last thing to note is how quickly the contributions, involving higher matrix powers of  $G$  (i.e. those trees with higher total branch length) decay to zero as the total branch length increases. This behavior is essentially governed by the spectral radius  $\rho(G)$  of the connectivity matrix. For example, in a large random network of both excitatory and inhibitory neurons, the spectrum consists of a single eigenvalue of size  $N\mu^{(1)} = Np\left(\frac{N_E}{N}g_E + \frac{N_I}{N}g_I\right)$  and a bulk spectrum, contained within a circle in the complex plane [34]. Its radius  $r$  is asymptotically given by

$$r^2 = Np(1-p)\left(\frac{N_E}{N}g_E^2 + \frac{N_I}{N}g_I^2\right). \tag{57}$$

While, as was already mentioned, the quantity  $N\mu^{(1)}$  corresponds to the total average input to a neuron, the radius  $r$  of the circle encompassing the bulk spectrum corresponds to the variance of this input. Thus, if the variance of the total input to a neuron in a random network is not too big ( $r < 1$ ), it will exhibit the aforementioned decay of contributions from trees with higher total branch lengths.

### Excitatory hubs in the network increase third-order cumulants

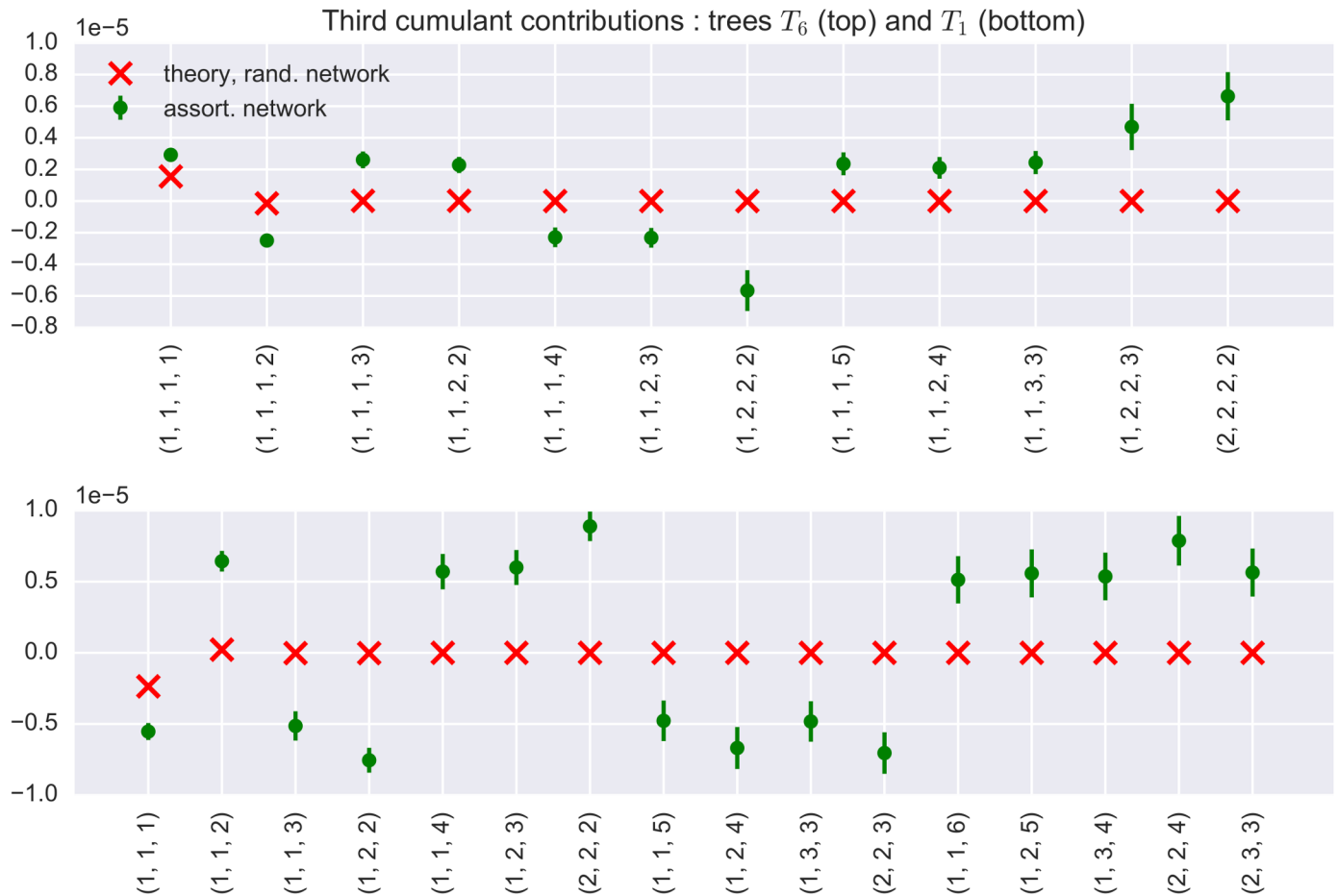
In the previous sections, we have demonstrated that the average third cumulant in networks with narrow degree distributions is determined by global parameters such as the number of neurons  $N$ , the connection probability  $p$ , and the average strength of input shared by  $k$  neurons,  $\mu^{(k)}$ . Of course, in networks with a wide degree distribution, the regular network approximation (which we used to derive the equation in [S2 Appendix](#)) is no longer valid. To demonstrate some of the new phenomena by simulation, we consider a network model with a geometric degree distribution, originally introduced in [3]. In short, the out-degrees  $k$  of excitatory and inhibitory neurons are chosen from a geometric distribution with parameter  $k_0$  (representing the mean out-degree) according to

$$P(k) = \left(1 - \frac{1}{k_0}\right)^{k-1} \frac{1}{k_0}. \tag{58}$$

This distribution exhibits a mean connection probability of  $1/k_0$  and a long tail. After the sampling of out-degrees, excitatory neurons are divided into “hubs” (out-degree  $k > k_0$ ) and “non-hubs” ( $k \leq k_0$ ). Postsynaptic neurons for non-hubs and inhibitory neurons are chosen randomly from the population consisting of all other neurons. However, for hub neurons, a fixed fraction  $f$  of all outgoing connections goes to other hubs. By varying  $f$  between 0 and 1, one can choose how densely connected the subnetwork of hubs will be. The “critical value” to keep in mind here is  $f_0 = 0.35$ . If  $f > f_0$ , hub neurons have a preference to connect to other hubs. Such a network is called “assortative”, otherwise it is called “disassortative”, see [3] for details. Similar networks have been studied in [35, 36]. The effect of the geometric out-degree distribution on the distribution of network motifs is depicted in [Fig 7](#).

If excitatory hubs preferentially connect to other hubs (for assortative networks), the number of relevant tree motifs with high total branch length grows in the network, and so does their combined strength. This is one major difference between assortative and random networks, and a reason why the contributions of longer trees in networks with hubs tend to be much larger than in Erdős-Rényi topologies. Of course, along the same lines, the number of “short” motifs (i.e. those with small total branch length) decreases (in comparison to their “longer” counterparts). This phenomenon is illustrated in [Fig 7](#).

This discrepancy can also be used to say something about the topology of the network that generated a given set of recorded spike data. Indeed, once the connection probability and third order correlations have been estimated (e.g. with the help of  $k$ -statistics), one could compare



**Fig 7. Contribution of tree motifs with longer total branch length increases in networks with excitatory hubs.** Top: Theoretical (narrow out-degree distribution approximation for random networks) and sample contributions (in non-regular networks) of the average third cumulant of  $T_6$  (see Fig 4) tree topologies with fixed branch lengths. Ticks on the x-axis code for the lengths of the branches of the tree. The ordering of the indices in the tick labels is done in a top-to-bottom and left-to-right fashion. The sample contributions were computed as averages of 3 independent realizations of an assortative network with a geometric out-degree distribution. Bottom: Theoretical and sample contributions to the average third cumulant of  $T_1$  tree topologies with fixed branch lengths.

doi:10.1371/journal.pcbi.1004963.g007

the regular network theory predictions with the third order cumulants obtained from data. A large disparity between the two could imply, for example, the presence of hubs and a wide in- and out-degree distribution in the network that generated the data.

## Discussion

In this work, we have studied connections between topology and measures of average third-order correlations in networks of spiking neurons. We have compared different connectivity rules with respect to their effect on the average joint third cumulant,  $\bar{\kappa}_3$ . Furthermore, we showed which topological motifs in the network contribute to the overall strength of third-order correlations. While our focus was on network models arising in neuroscience, we feel that the results presented here could as well be relevant in other fields, where correlations of higher-order play an important role.

As a handy computational model of spiking neuronal activity, we have used the Hawkes point process [27, 28], which was originally introduced as a model of earthquake activity. It is



sufficiently rich in order to model interesting dependencies between various types of events (in our case, spikes of different neurons), but still simple enough to be tractable. Indeed, these are the exact properties that make Hawkes processes quite useful models in neuroscience. They have been employed in the analysis of pairwise correlations between spike trains [3, 37], modeling spike-timing dependent plasticity [38, 39], and, very recently, to model single unit activity recorded on a monkey during a sensory-motor task [40].

Using the Hawkes process theory, we have shown that a linear stochastic point process model can reproduce not only the event rates and pairwise correlations in networks (as was already shown in [3]), but also its third-order joint cumulants, which are statistical measures of correlations between groups of three nodes. These cumulants can be seen as a quantification of “non-Gaussian” properties of the total population activity observed in time bins of a given size.

The problem of quantifying higher-order correlations is of some importance in computational neuroscience. It has been suggested a long time ago [41, 42] that understanding the cooperative dynamics of populations of neurons would provide much needed insight into the neuron-level mechanisms of brain function. Indeed, there is now a large body of experimental evidence that supports the idea of computationally relevant correlations between neurons in a network [7, 43–45]. The evidence for coordinated activity of neuronal spike trains, however, mostly relies on the correlations between pairs of nerve cells [46–50]. Unfortunately, it is becoming increasingly clear that pairwise correlations cannot explain the intricate dynamics of neuronal populations [9, 12, 51, 52] and that higher-order moments of spiking activity need to be taken into account.

Traditionally in neuroscience, higher-order synchrony has been almost exclusively investigated with the help of classical tools borrowed from statistical physics such as maximum entropy models [13–18, 53]. In this approach, the quantifiers of higher-order coordination are the so-called “interaction parameters” of the binary exponential family. However, an alternative measure, commonly used in statistical literature, also exists—it is the joint cumulant. As already mentioned in [54, 55], cumulant correlations are not identical to the higher order exponential family parameters (for details, see [54]). In a sense, it can be said that non-zero cumulants indicate the presence of additive common input (a well-known model for correlated stochastic signals, see [56–58]), while the interaction parameters of maximum entropy models measure multiplicative interactions. The mathematical differences between the two types of dependence are currently under investigation [59–61]. As our neuronal network model each neuron “feels” only the linear sum of spiking activity of its presynaptic partners, in this work we have opted for quantifying synchrony using joint cumulants. Finally, it may be worthwhile to note that there are other ways of generating time structured correlations of higher order in computational models (see, for example, [9], but also [62]).

In addition, by generalizing the result in [3], we have found that integrated third-order correlations ( $\kappa^{ijk}$ ) also admit a representation in terms of sums of weights of certain topological sub-motifs in the network. While in the case of pairwise correlations between neurons these motifs were simple binary trees (see Fig 2), when dealing with third-order interactions the motifs become more complex (Fig 3) “trees with three leaves”, which are still manageable computationally. More precisely, it is the combined “strength” of all such trees containing a triplet of neurons that determine how often, on average, the activity of such a triplet exhibits coordinated spiking. Sadly, no concise matrix product formula is available for the whole third cumulant tensor  $\{\kappa^{ijk}\}_{i,j,k}$  and one has to resort to writing down equations for individual components, which still offer the possibility of efficient estimation. Indeed, computing the theoretical cumulants  $\kappa^{ijk}$  for (close to) regular networks is much less computationally intensive than estimating them from data via  $k$ -statistics and only relies on simple algebraic manipulations of the connectivity matrix  $G$ .

We have also studied analytically the average third-order cumulant  $\bar{\kappa}_3$ , derived from the sum of joint cumulants of all possible triplets of neurons in the network. We have shown that the value of  $\bar{\kappa}_3$  in random networks of Erdős-Rényi type does not depend on fine-scale topological structure and is instead a function of global network parameters, such as the network size  $N$ , the connection probability  $p$  and total common input to groups of neurons. Furthermore, we have shown that, in the limit of very large networks, the dominating contribution to  $\bar{\kappa}_3$  comes from the combined weight of all trees with a specific topology (which we denoted  $T_6$ , see Fig 4) present in the network. Thus, for large, random networks, it is tree-like connectivity motifs of this topology that affects the average third cumulant most.

We were able to show that the contributions of individual subtrees to the average joint cumulant depend on specific topological properties of the tree, such as its number of branches, number of nodes and, interestingly, the out-degrees of its internal nodes (nodes that are not leaves as they have a nonzero out-degree). Not surprisingly, in a stable network (whose connectivity matrix  $G$  has a spectral radius less than 1), the absolute contributions of trees with a large number of branches decays to 0 as the number of branches increases. However, the sign of the total contribution turns out to depend both on the parity of the sum of all internal node out-degrees and the parity of the total branch length. This, in principle, allows one to determine whether the presence of a particular sub-tree in a network will increase or decrease the third cumulant, and thus allow to compute the total size of third-order interactions.

Finally, we considered a case in which our regular network approximation fails, networks with interconnected hub neurons. Similar networks were already considered in [3]. Their main characteristic is a heavy-tailed out-degree distribution (in the case we considered, it was geometric). Such networks are, in a sense, the opposite of an Erdős-Rényi type random network. The presence of interconnected hubs increases the number of subtrees in the network with large total branch length and, consequently, their overall contribution to the average joint third cumulant. Thus, such networks illustrate nicely how “higher-order” motifs can, for certain networks, influence the overall third-order cumulant structure, which is not possible in networks with narrow out-degree distributions.

As far as the limitations of our approach are concerned, it is important to note that the linear theory of Hawkes processes which we resorted to [29] is strictly valid only for purely excitatory networks, as the instantaneous rate function is not allowed to become negative. For the case discussed here, this may happen, as the networks are inhibition-dominated. However, in accordance with what was already mentioned in [27], the theoretical results remain approximately valid for networks with negative interactions, as long as the probability of the rate being negative is small. Still, an interesting generalization of our model, and the results achieved with it, would be the case of multiplicative interaction [63]. More generally, a point process model in which a non-negative nonlinearity is applied to Eq 3 yields a necessarily positive rate for any choice of interaction kernels. The computational approach one would have to use in this case in order to study the effect of topology on higher-order correlations would be quite different, though, as almost no analytical results exist for such models [64, 65].

## Supporting Information

### S1 Appendix. Integrated covariances and joint third cumulants.

(PDF)

### S2 Appendix. Computation of $\bar{\kappa}_3$ .

(PDF)

**S3 Appendix. Derivation of the third-order joint cumulant formula.**  
(PDF)

**Acknowledgments**

The authors thank the reviewers, as well as John Hertz for their comments on this paper. S. J. acknowledges the hospitality of NORDITA.

**Author Contributions**

Conceived and designed the experiments: SJ SR. Performed the experiments: SJ. Analyzed the data: SJ. Wrote the paper: SJ SR.

**References**

1. Bascompte J, et al. Disentangling the web of life. *Science*. 2009; 325(5939):416. doi: [10.1126/science.1170749](https://doi.org/10.1126/science.1170749) PMID: [19628856](https://pubmed.ncbi.nlm.nih.gov/19628856/)
2. Reynaud-Bouret P, Schbath S, et al. Adaptive estimation for Hawkes processes; application to genome analysis. *The Annals of Statistics*. 2010; 38(5):2781–2822. doi: [10.1214/10-AOS806](https://doi.org/10.1214/10-AOS806)
3. Pernice V, Staude B, Cardanobile S, Rotter S. How structure determines correlations in neuronal networks. *PLoS Comput Biol*. 2011; 7(5):e1002059. doi: [10.1371/journal.pcbi.1002059](https://doi.org/10.1371/journal.pcbi.1002059) PMID: [21625580](https://pubmed.ncbi.nlm.nih.gov/21625580/)
4. Strogatz SH. Exploring complex networks. *Nature*. 2001; 410(6825):268–276. doi: [10.1038/35065725](https://doi.org/10.1038/35065725) PMID: [11258382](https://pubmed.ncbi.nlm.nih.gov/11258382/)
5. Shadlen MN, Newsome WT. The variable discharge of cortical neurons: implications for connectivity, computation, and information coding. *The Journal of neuroscience*. 1998; 18(10):3870–3896. PMID: [9570816](https://pubmed.ncbi.nlm.nih.gov/9570816/)
6. Averbeck BB, Latham PE, Pouget A. Neural correlations, population coding and computation. *Nature Reviews Neuroscience*. 2006; 7(5):358–366. doi: [10.1038/nrn1888](https://doi.org/10.1038/nrn1888) PMID: [16760916](https://pubmed.ncbi.nlm.nih.gov/16760916/)
7. Salinas E, Sejnowski TJ. Correlated neuronal activity and the flow of neural information. *Nature reviews neuroscience*. 2001; 2(8):539–550. doi: [10.1038/35086012](https://doi.org/10.1038/35086012) PMID: [11483997](https://pubmed.ncbi.nlm.nih.gov/11483997/)
8. Rossant C, Leijon S, Magnusson AK, Brette R. Sensitivity of noisy neurons to coincident inputs. *The Journal of Neuroscience*. 2011; 31(47):17193–17206. doi: [10.1523/JNEUROSCI.2482-11.2011](https://doi.org/10.1523/JNEUROSCI.2482-11.2011) PMID: [22114286](https://pubmed.ncbi.nlm.nih.gov/22114286/)
9. Kuhn A, Aertsen A, Rotter S. Higher-order statistics of input ensembles and the response of simple model neurons. *Neural Computation*. 2003; 15(1):67–101. doi: [10.1162/089976603321043702](https://doi.org/10.1162/089976603321043702) PMID: [12590820](https://pubmed.ncbi.nlm.nih.gov/12590820/)
10. Martignon L, Deco G, Laskey K, Diamond M, Freiwald W, Vaadia E. Neural coding: higher-order temporal patterns in the neurostatistics of cell assemblies. *Neural Computation*. 2000; 12(11):2621–2653. doi: [10.1162/089976600300014872](https://doi.org/10.1162/089976600300014872) PMID: [11110130](https://pubmed.ncbi.nlm.nih.gov/11110130/)
11. Yu S, Yang H, Nakahara H, Santos GS, Nikolić D, Plenž D. Higher-order interactions characterized in cortical activity. *The Journal of Neuroscience*. 2011; 31(48):17514–17526. doi: [10.1523/JNEUROSCI.3127-11.2011](https://doi.org/10.1523/JNEUROSCI.3127-11.2011) PMID: [22131413](https://pubmed.ncbi.nlm.nih.gov/22131413/)
12. Ohiorhenuan IE, Mechler F, Purpura KP, Schmid AM, Hu Q, Victor JD. Sparse coding and high-order correlations in fine-scale cortical networks. *Nature*. 2010; 466(7306):617–621. doi: [10.1038/nature09178](https://doi.org/10.1038/nature09178) PMID: [20601940](https://pubmed.ncbi.nlm.nih.gov/20601940/)
13. Tkačik G, Marre O, Amodei D, Schneidman E, Bialek W, Berry MJ II. Searching for collective behavior in a large network of sensory neurons. *PLoS Comput Biol*. 2014; 10(1):e1003408. doi: [10.1371/journal.pcbi.1003408](https://doi.org/10.1371/journal.pcbi.1003408) PMID: [24391485](https://pubmed.ncbi.nlm.nih.gov/24391485/)
14. Montani F, Ince RA, Senatore R, Arabzadeh E, Diamond ME, Panzeri S. The impact of high-order interactions on the rate of synchronous discharge and information transmission in somatosensory cortex. *Philosophical Transactions of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*. 2009; 367(1901):3297–3310. doi: [10.1098/rsta.2009.0082](https://doi.org/10.1098/rsta.2009.0082)
15. Köster U, Sohl-Dickstein J, Gray CM, Olshausen BA. Modeling higher-order correlations within cortical microcolumns. *PLoS Comput Biol*. 2014; 10(7):e1003684. doi: [10.1371/journal.pcbi.1003684](https://doi.org/10.1371/journal.pcbi.1003684) PMID: [24991969](https://pubmed.ncbi.nlm.nih.gov/24991969/)
16. Schneidman E, Berry MJ, Segev R, Bialek W. Weak pairwise correlations imply strongly correlated network states in a neural population. *Nature*. 2006; 440(7087):1007–1012. doi: [10.1038/nature04701](https://doi.org/10.1038/nature04701) PMID: [16625187](https://pubmed.ncbi.nlm.nih.gov/16625187/)

17. Shlens J, Field GD, Gauthier JL, Grivich MI, Petrusca D, Sher A, et al. The structure of multi-neuron firing patterns in primate retina. *The Journal of neuroscience*. 2006; 26(32):8254–8266. doi: [10.1523/JNEUROSCI.1282-06.2006](https://doi.org/10.1523/JNEUROSCI.1282-06.2006) PMID: [16899720](https://pubmed.ncbi.nlm.nih.gov/16899720/)
18. Pillow JW, Shlens J, Paninski L, Sher A, Litke AM, Chichilnisky E, et al. Spatio-temporal correlations and visual signalling in a complete neuronal population. *Nature*. 2008; 454(7207):995–999. doi: [10.1038/nature07140](https://doi.org/10.1038/nature07140) PMID: [18650810](https://pubmed.ncbi.nlm.nih.gov/18650810/)
19. Shimazaki H, Amari Si, Brown EN, Grün S. State-space analysis of time-varying higher-order spike correlation for multiple neural spike train data. *PLoS computational biology*. 2012; 8(3):e1002385. doi: [10.1371/journal.pcbi.1002385](https://doi.org/10.1371/journal.pcbi.1002385) PMID: [22412358](https://pubmed.ncbi.nlm.nih.gov/22412358/)
20. Hu Y, Trousdale J, Josić K, Shea-Brown E. Local paths to global coherence: cutting networks down to size. *Physical Review E*. 2014; 89(3):032802. doi: [10.1103/PhysRevE.89.032802](https://doi.org/10.1103/PhysRevE.89.032802)
21. Cayco-Gajic NA, Zylberberg J, Shea-Brown E. Triplet correlations among similarly tuned cells impact population coding. *Frontiers in Computational Neuroscience*. 2015; 9:57. doi: [10.3389/fncom.2015.00057](https://doi.org/10.3389/fncom.2015.00057) PMID: [26042024](https://pubmed.ncbi.nlm.nih.gov/26042024/)
22. Pernice V, Staude B, Cardanobile S, Rotter S. Recurrent interactions in spiking networks with arbitrary topology. *Physical review E*. 2012; 85(3):031916. doi: [10.1103/PhysRevE.85.031916](https://doi.org/10.1103/PhysRevE.85.031916)
23. Pernice V, Rotter S. Reconstruction of sparse connectivity in neural networks from spike train covariances. *Journal of Statistical Mechanics: Theory and Experiment*. 2013(03):P03008.
24. Rangan AV. Diagrammatic expansion of pulse-coupled network dynamics. *Physical review letters*. 2009; 102(15):158101. doi: [10.1103/PhysRevLett.102.158101](https://doi.org/10.1103/PhysRevLett.102.158101) PMID: [19518674](https://pubmed.ncbi.nlm.nih.gov/19518674/)
25. Rangan AV. Diagrammatic expansion of pulse-coupled network dynamics in terms of subnetworks. *Physical Review E*. 2009; 80(3):036101. doi: [10.1103/PhysRevE.80.036101](https://doi.org/10.1103/PhysRevE.80.036101)
26. Trousdale J, Hu Y, Shea-Brown E, Josic K. Impact of network structure and cellular response on spike time correlations. *PLoS Comput Biol*. 2012; 8(3):e1002408–e1002408. doi: [10.1371/journal.pcbi.1002408](https://doi.org/10.1371/journal.pcbi.1002408) PMID: [22457608](https://pubmed.ncbi.nlm.nih.gov/22457608/)
27. Hawkes AG. Spectra of some self-exciting and mutually exciting point processes. *Biometrika*. 1971; 58(1):83–90. doi: [10.1093/biomet/58.1.83](https://doi.org/10.1093/biomet/58.1.83)
28. Hawkes AG. Point Spectra of Some Mutually Exciting Point Processes. *Journal of the Royal Statistical Society Series B (Methodological)*. 33(3):pp. 438–443.
29. Jovanović S, Hertz J, Rotter S. Cumulants of Hawkes point processes. *Phys Rev E*. 2015 Apr; 91:042802. doi: [10.1103/PhysRevE.91.042802](https://doi.org/10.1103/PhysRevE.91.042802)
30. Bacry E, Muzy JF. Second order statistics characterization of Hawkes processes and non-parametric estimation. *arXiv preprint arXiv:14010903*. 2014;.
31. McCullagh P. *Tensor methods in statistics*. vol. 161. Chapman and Hall London; 1987.
32. Lukacs E. *Characteristics functions*. Griffin, London. 1970;.
33. Diesmann M, Gewaltig MO. NEST: An environment for neural systems simulations. *Forschung und wissenschaftliches Rechnen, Beiträge zum Heinz-Billing-Preis*. 2001; 58:43–70.
34. Rajan K, Abbott L. Eigenvalue spectra of random matrices for neural networks. *Physical review letters*. 2006; 97(18):188104. doi: [10.1103/PhysRevLett.97.188104](https://doi.org/10.1103/PhysRevLett.97.188104) PMID: [17155583](https://pubmed.ncbi.nlm.nih.gov/17155583/)
35. Roxin A, Hakim V, Brunel N. The statistics of repeating patterns of cortical activity can be reproduced by a model network of stochastic binary neurons. *The Journal of neuroscience*. 2008; 28(42):10734–10745. doi: [10.1523/JNEUROSCI.1016-08.2008](https://doi.org/10.1523/JNEUROSCI.1016-08.2008) PMID: [18923048](https://pubmed.ncbi.nlm.nih.gov/18923048/)
36. Rubinov M, Sporns O. Complex network measures of brain connectivity: uses and interpretations. *Neuroimage*. 2010; 52(3):1059–1069. doi: [10.1016/j.neuroimage.2009.10.003](https://doi.org/10.1016/j.neuroimage.2009.10.003) PMID: [19819337](https://pubmed.ncbi.nlm.nih.gov/19819337/)
37. Krumin M, Reutsky I, Shoham S. Correlation-based analysis and generation of multiple spike trains using Hawkes models with an exogenous input. *Frontiers in Computational Neuroscience*. 2010; 4(147). Available from: [http://www.frontiersin.org/computational\\_neuroscience/10.3389/fncom.2010.00147/abstract](http://www.frontiersin.org/computational_neuroscience/10.3389/fncom.2010.00147/abstract). doi: [10.3389/fncom.2010.00147](https://doi.org/10.3389/fncom.2010.00147) PMID: [21151360](https://pubmed.ncbi.nlm.nih.gov/21151360/)
38. Gilson M, Burkitt AN, Grayden DB, Thomas DA, van Hemmen JL. Emergence of network structure due to spike-timing-dependent plasticity in recurrent neuronal networks. I. Input selectivity–strengthening correlated input pathways. *Biological cybernetics*. 2009; 101(2):81–102. doi: [10.1007/s00422-009-0319-4](https://doi.org/10.1007/s00422-009-0319-4) PMID: [19536560](https://pubmed.ncbi.nlm.nih.gov/19536560/)
39. Kempter R, Gerstner W, Van Hemmen JL. Hebbian learning and spiking neurons. *Physical Review E*. 1999; 59(4):4498. doi: [10.1103/PhysRevE.59.4498](https://doi.org/10.1103/PhysRevE.59.4498)
40. Reynaud-Bouret P, Rivoirard V, Grammont F, Tuleau-Malot C. Goodness-of-fit tests and nonparametric adaptive estimation for spike train analysis. *Journal of mathematical neuroscience*. 2014; 4(3). doi: [10.1186/2190-8567-4-3](https://doi.org/10.1186/2190-8567-4-3) PMID: [24742008](https://pubmed.ncbi.nlm.nih.gov/24742008/)

41. Hebb DO. *The Organization of Behavior: A Neuropsychological Theory*. New ed ed. New York: Wiley;1949
42. Gerstein G, Bedenbaugh P, Aertsen AM. Neuronal assemblies. *Biomedical Engineering, IEEE Transactions on*. 1989; 36(1):4–14. doi: [10.1109/10.16444](https://doi.org/10.1109/10.16444)
43. Lestienne R. Spike timing, synchronization and information processing on the sensory side of the central nervous system. *Progress in neurobiology*. 2001; 65(6):545–591. doi: [10.1016/S0301-0082\(01\)00019-3](https://doi.org/10.1016/S0301-0082(01)00019-3) PMID: [11728644](https://pubmed.ncbi.nlm.nih.gov/11728644/)
44. Womelsdorf T, Fries P. The role of neuronal synchronization in selective attention. *Current opinion in neurobiology*. 2007; 17(2):154–160. doi: [10.1016/j.conb.2007.02.002](https://doi.org/10.1016/j.conb.2007.02.002) PMID: [17306527](https://pubmed.ncbi.nlm.nih.gov/17306527/)
45. Kohn A, Zandvakili A, Smith MA. Correlations and brain states: from electrophysiology to functional imaging. *Current opinion in neurobiology*. 2009; 19(4):434–438. doi: [10.1016/j.conb.2009.06.007](https://doi.org/10.1016/j.conb.2009.06.007) PMID: [19608406](https://pubmed.ncbi.nlm.nih.gov/19608406/)
46. Gray CM, Singer W. Stimulus-specific neuronal oscillations in orientation columns of cat visual cortex. *Proceedings of the National Academy of Sciences*. 1989 Mar; 86(5):1698–1702. doi: [10.1073/pnas.86.5.1698](https://doi.org/10.1073/pnas.86.5.1698)
47. Vaadia E, Haalman I, Abeles M, Bergman H, Prut Y, Slovin H, et al. Dynamics of neuronal interactions in monkey cortex in relation to behavioural events. *Nature*. 1995; 373(6514):515–518. doi: [10.1038/373515a0](https://doi.org/10.1038/373515a0) PMID: [7845462](https://pubmed.ncbi.nlm.nih.gov/7845462/)
48. Riehle A, Grün S, Diesmann M, Aertsen A. Spike synchronization and rate modulation differentially involved in motor cortical function. *Science*. 1997; 278(5345):1950–1953. doi: [10.1126/science.278.5345.1950](https://doi.org/10.1126/science.278.5345.1950) PMID: [9395398](https://pubmed.ncbi.nlm.nih.gov/9395398/)
49. Bair W, Zohary E, Newsome WT. Correlated firing in macaque visual area MT: time scales and relationship to behavior. *The journal of Neuroscience*. 2001; 21(5):1676–1697. PMID: [11222658](https://pubmed.ncbi.nlm.nih.gov/11222658/)
50. Kohn A, Smith MA. Stimulus dependence of neuronal correlation in primary visual cortex of the macaque. *The Journal of neuroscience*. 2005; 25(14):3661–3673. doi: [10.1523/JNEUROSCI.5106-04.2005](https://doi.org/10.1523/JNEUROSCI.5106-04.2005) PMID: [15814797](https://pubmed.ncbi.nlm.nih.gov/15814797/)
51. Martignon L, Von Hase H, Grün S, Aertsen A, Palm G. Detecting higher-order interactions among the spiking events in a group of neurons. *Biological cybernetics*. 1995; 73(1):69–81. doi: [10.1007/BF00199057](https://doi.org/10.1007/BF00199057) PMID: [7654851](https://pubmed.ncbi.nlm.nih.gov/7654851/)
52. Bohtë SM, Spekreijse H, Roelfsema PR. The effects of pair-wise and higher-order correlations on the firing rate of a postsynaptic neuron. *Neural Computation*. 2000; 12(1):153–179. doi: [10.1162/089976600300015934](https://doi.org/10.1162/089976600300015934) PMID: [10636937](https://pubmed.ncbi.nlm.nih.gov/10636937/)
53. Nakahara H, Amari Si. Information-geometric measure for neural spikes. *Neural Computation*. 2002; 14(10):2269–2316. doi: [10.1162/08997660260293238](https://doi.org/10.1162/08997660260293238) PMID: [12396564](https://pubmed.ncbi.nlm.nih.gov/12396564/)
54. Staude B, Grün S, Rotter S. Higher-order correlations and cumulants. In: *Analysis of parallel spike trains*. Springer; 2010. p. 253–280.
55. Staude B, Rotter S, Grün S. CuBIC: cumulant based inference of higher-order correlations in massively parallel spike trains. *Journal of Computational Neuroscience*. 2010; 29(1–2):327–350. doi: [10.1007/s10827-009-0195-x](https://doi.org/10.1007/s10827-009-0195-x) PMID: [19862611](https://pubmed.ncbi.nlm.nih.gov/19862611/)
56. De La Rocha J, Doiron B, Shea-Brown E, Josić K, Reyes A. Correlation between neural spike trains increases with firing rate. *Nature*. 2007; 448(7155):802–806. doi: [10.1038/nature06028](https://doi.org/10.1038/nature06028) PMID: [17700699](https://pubmed.ncbi.nlm.nih.gov/17700699/)
57. Staude B, Rotter S, Grün S. Can spike coordination be differentiated from rate covariation? *Neural Computation*. 2008; 20(8):1973–1999. doi: [10.1162/neco.2008.06-07-550](https://doi.org/10.1162/neco.2008.06-07-550) PMID: [18336077](https://pubmed.ncbi.nlm.nih.gov/18336077/)
58. Tetzlaff T, Rotter S, Stark E, Abeles M, Aertsen A, Diesmann M. Dependence of neuronal correlations on filter characteristics and marginal spike train statistics. *Neural computation*. 2008; 20(9):2133–2184. doi: [10.1162/neco.2008.05-07-525](https://doi.org/10.1162/neco.2008.05-07-525) PMID: [18439140](https://pubmed.ncbi.nlm.nih.gov/18439140/)
59. Darroch J, Speed T. Additive and multiplicative models and interactions. *The Annals of Statistics*. 1983; p. 724–738. doi: [10.1214/aos/1176346240](https://doi.org/10.1214/aos/1176346240)
60. Streitberg B. Lancaster interactions revisited. *The Annals of Statistics*. 1990;p. 1878–1885. doi: [10.1214/aos/1176347885](https://doi.org/10.1214/aos/1176347885)
61. Ip EH, Wang YJ, Yeh Yn. Structural decompositions of multivariate distributions with applications in moment and cumulant. *Journal of multivariate analysis*. 2004; 89(1):119–134. doi: [10.1016/j.jmva.2003.09.001](https://doi.org/10.1016/j.jmva.2003.09.001)
62. Trousdale J, Hu Y, Shea-Brown E, Josic K. A generative spike train model with time-structured higher order correlations. *Frontiers in Computational Neuroscience*. 2013; 7(84). Available from: [http://www.frontiersin.org/computational\\_neuroscience/10.3389/fncom.2013.00084/abstract](http://www.frontiersin.org/computational_neuroscience/10.3389/fncom.2013.00084/abstract). doi: [10.3389/fncom.2013.00084](https://doi.org/10.3389/fncom.2013.00084) PMID: [23908626](https://pubmed.ncbi.nlm.nih.gov/23908626/)

63. Cardanobile S, Rotter S. Multiplicatively interacting point processes and applications to neural modeling. *Journal of computational neuroscience*. 2010; 28(2):267–284. doi: [10.1007/s10827-009-0204-0](https://doi.org/10.1007/s10827-009-0204-0) PMID: [20052525](https://pubmed.ncbi.nlm.nih.gov/20052525/)
64. Bremaud P, Massoulié L. Stability of Nonlinear Hawkes Processes. *The Annals of Probability*. 1996; 24(3):pp. 1563–1588. doi: [10.1214/aop/1065725193](https://doi.org/10.1214/aop/1065725193)
65. Zhu L. Central Limit Theorem for Nonlinear Hawkes Processes. *Journal of Applied Probability*. 2013; (50):760–771. doi: [10.1239/jap/1378401234](https://doi.org/10.1239/jap/1378401234)