Computational Neuroscience

Modeling and analyzing higher-order correlations in non-Poissonian spike trains

Imke C.G. Reimer a, Benjamin Staude a, Werner Ehm b, Stefan Rotter a,∗

a Bernstein Center Freiburg and Faculty of Biology, Albert-Ludwig University, Freiburg, Germany
b Institute for Frontier Areas of Psychology and Mental Health, Freiburg, Germany

ARTICLE INFO

Article history:
Received 5 December 2011
Received in revised form 17 April 2012
Accepted 18 April 2012

Keywords:
Multiple single-unit spike trains
Surrogate data
Calibration of correlation measures

ABSTRACT

Measuring pairwise and higher-order spike correlations is crucial for studying their potential impact on neuronal information processing. In order to avoid misinterpretation of results, the tools used for data analysis need to be carefully calibrated with respect to their sensitivity and robustness. This, in turn, requires surrogate data with statistical properties common to experimental spike trains. Here, we present a novel method to generate correlated non-Poissonian spike trains and study the impact of single-neuron spike statistics on the inference of higher-order correlations. Our method to mimic cooperative neuronal spike activity allows the realization of a large variety of renewal processes with controlled higher-order correlation structure. Based on surrogate data obtained by this procedure we investigate the robustness of the recently proposed method empirical de-Poissonization (Ehm et al., 2007). It assumes Poissonian spiking, which is common also for many other estimation techniques. We observe that some degree of deviation from this assumption can generally be tolerated, that the results are more reliable for small analysis bins, and that the degree of misestimation depends on the detailed spike statistics. As a consequence of these findings we finally propose a strategy to assess the reliability of results for experimental data.

© 2012 Elsevier B.V. All rights reserved.

1. Introduction

Despite more than 50 years of experimental and theoretical research, the role of correlated spike activity for neural coding and cortical information processing remains highly debated. While earlier studies focused on pairwise spike correlations, more recently higher-order correlations have also attracted the interest of researchers. Various simulation studies have revealed that neurons are indeed very sensitive to the higher-order structure in their input (Diesmann et al., 1999; Bohte et al., 2000; Kuhn et al., 2003; Benucci et al., 2007). Advancements in multi-electrode recording techniques nowadays allow to investigate the spiking activity of three or more neurons simultaneously, and first findings about the incidence of higher-order correlations in neuronal data have been reported, along with speculations about their role for information processing (e.g. Shlens et al., 2006; Schneidman et al., 2006; Tang et al., 2008; Montani et al., 2009; Ohiorhenuan et al., 2010; Ganmor et al., 2011; Yu et al., 2011; Shimazaki et al., 2012, see also Macke et al., 2011 for possible explanations).

The measurement of neuronal correlations, and hence their interpretation, can be affected by various factors (Cohen and Kohn, 2011), in particular the brain state (Kohn et al., 2009), spike sorting errors (Bar-Gad et al., 2001; Pazienti and Grün, 2006) and systematic deviations from the (implicit) assumptions made by the employed data analysis method (Tetzlaff et al., 2008; Grün, 2009). By far the most common assumption is that measurements in subsequent time bins are independent of each other. For single-neuron spike trains, this effectively implies Bernoulli or Poisson statistics. However, recent studies emphasize that neuronal spike trains often deviate from this assumption. In particular, neurons can be more irregular or more regular than a Poisson process, and the degree of spiking irregularity varies with the brain area (Shinomoto et al., 2009; Maimon and Assad, 2009). For pairwise correlations, the influence of single-neuron spike statistics has been well investigated, and methods to account for model violations are available (see e.g. Tetzlaff et al., 2008; Grün, 2009; Louis et al., 2010a,b). In order to perform similar studies for estimators of higher-order correlations, proper surrogate data are inevitable. Currently, however, no method is available to generate non-Poissonian point processes with defined higher-order correlations.

In the first part of this paper (Section 2), we will therefore introduce a new method to simulate non-Poissonian spike trains with controlled higher-order correlations. It combines two procedures of generating point processes: First, Poisson processes with higher-order correlations can be obtained by injecting coincident spikes
(i.e. higher-order correlations) into a background of independently generated spikes (Kuhn et al., 2003; Ehm et al., 2007; Staude et al., 2010b,c). Second, a renewal process can be generated by appropriate deletion of events (“thinning”) of a Poisson process (Devroye, 1986). Applying the thinning procedure to each of the correlated Poisson processes then yields a population of non-Poissonian spike trains with higher-order correlations. We study the influence of thinning and devise a strategy to prescribe the higher-order structure of the resulting non-Poissonian population.

In the second part (Section 3), we investigate the influence of non-Poissonian spiking statistics on estimates of higher-order features. As a result, we propose an a posteriori strategy to judge the error of higher-order correlation estimates from experimental data. The method employed here to estimate higher-order correlations is empirical de-Poissonization (EDP; Ehm et al., 2007) which will also be briefly introduced. Based on the population spike count, EDP estimates a population average of the correlation structure, and a lower confidence limit for the maximal order of correlation present in the data. In contrast to other higher-order analysis methods (e.g. Martignon et al., 1995, 2000; Amari et al., 2003) its compact parametrization of higher-order correlations renders EDP applicable to large neuronal populations and sample sizes that are compatible with current experimental settings (see also Ehm et al., 2008). However, the underlying model assumes that neurons fire in a Poisson manner.

### 2. Generation of non-Poissonian processes with higher-order correlations

In order to introduce our new method to generate non-Poissonian processes with higher-order correlations some terminology and methodology of renewal theory is required.

#### 2.1. Renewal point process

##### 2.1.1. Definition

A renewal process is a point process where the times between two point-like events (i.e. the inter-spike intervals) are independent and identically distributed with probability density function (p.d.f.) \( f(x) \) (see e.g. Cox, 1962; Perkel et al., 1967).

Equivalently to this definition, a renewal process is completely determined by its hazard function. This function describes the conditional rate for the occurrence of an event (i.e. spike) with respect to the time \( x \) passed since the last event. Formally, the hazard function is defined as

\[
\phi(x) := \lim_{\Delta x \to 0} \frac{\mathbb{P}(x < X \leq x + \Delta x \mid X < x)}{\Delta x} = \frac{f(x)}{1 - F(x)},
\]

where \( X \) is the inter-event interval and \( F(x) := \mathbb{P}(X \leq x) = \int_0^x f(y)dy \) the cumulative distribution function.

Useful quantities to describe a renewal process include the intensity \( \lambda \) (i.e. the spike rate) and the coefficient of variation \( CV \) of the inter-spike interval distribution (i.e. the spiking irregularity). Let \( E[X] \) denote the expectation of \( X \) and \( \text{Var}[X] \) the variance of \( X \), then \( \lambda \) and \( CV \) are given by

\[
\lambda = \frac{1}{E[X]} \quad \text{and} \quad CV = \sqrt{\frac{\text{Var}[X]}{E[X]}}.
\]

##### 2.1.2. Examples

The most prominent example of a renewal process is the Poisson process with exponentially distributed inter-spike intervals and a constant hazard function. In neuroscience common non-Poissonian renewal processes are Poisson processes with dead time, gamma processes and lognormal processes. In particular, they are frequently used to fit experimentally observed inter-spike intervals and to mimic non-Poissonian spike activity in theoretical studies of neural processing (see e.g. Kuffer et al., 1957; Stein, 1965; Beyer et al., 1975; Burns and Webb, 1976; Tuckwell, 1988; Levine, 1991; McKeegan, 2002; Nawrot et al., 2008; Maimon and Assad, 2009; Minich et al., 2009; Grün, 2009; Ly and Tranchina, 2009; Deger et al., 2010, 2011; Rosenbaum and Josić, 2011). We briefly describe some of its main features and refer the reader to Appendix A for more details.

1. **2.1.2.1. Poisson process with dead time.** Poisson processes with dead time (PPD) are quite similar to Poisson processes but additionally they allow to account for the refractory period of a neuron. That is, the rate for the occurrence of a new spike after a spike is zero up to the dead time \( d \). This is depicted for the inter-spike interval distribution and for the corresponding hazard function in Fig. 1A and 1D (compare orange dashed line, Poisson process, with the solid lines, PPD). The spiking irregularity which can be mimicked by a PPD is restricted to a coefficient of variation \( CV \) smaller than one (cf. Appendix A).

2. **2.1.2.2. Gamma process.** In contrast, any coefficient of variation within the interval \([0, \infty]\) can be realized by gamma processes which have gamma-distributed inter-spike intervals. As for PPDs the Poisson process constitutes a special case for \( CV = 1 \). Fig. 1B and 1E shows that for \( CV > 1 \) small inter-spike intervals are very frequent, in contrast to neuronal spike activity. For \( CV < 1 \) small inter-spike intervals can occur but they are rare.

3. **2.1.2.3. Lognormal process.** As opposed to gamma processes the occurrence of small inter-spike intervals is very unlikely for processes with lognormal-distributed inter-spike intervals even for \( CV > 1 \). Accordingly, the hazard function of lognormal processes exhibits a different behavior (compare panels E and F in Fig. 1). Furthermore, a lognormal process with \( CV = 1 \) is not a Poisson process, as can be seen from the non-flat hazard function (Fig. 1F).

4. **2.1.3. Generation of a renewal process via thinning of a Poisson process**

An adaptation of the method proposed by Lewis and Shedler (1979) on generating non-stationary Poisson processes enables one to generate renewal processes via a thinning procedure (see Devroye, 1986, chapter VI.2.4). An intuition for its main idea can easily be obtained for Poisson processes with dead time. As has been mentioned earlier, Poisson processes with dead time are like Poisson processes except for the fact that given an event at time \( t \), no event can occur up to time \( t + d \), where \( d \) is the imposed dead time. Hence, one obtains such a process by first simulating a “source” Poisson process and then deleting all events which occur too early with respect to the previous event (see Fig. 2A). That is, one keeps an event only if the hazard function of the corresponding inter-event interval has a non-zero value. In doing so, the rate is reduced and therefore the initial rate of the source process has to be chosen sufficiently high.

The procedure generalizes to processes with more complicated hazard functions \( \phi(x) \) by deleting events with a probability proportional to the hazard function evaluated at the time \( x \) elapsed since the most recent event (see Fig. 2B). In order to obtain a certain rate in the “target” process after thinning, the rate of the source Poisson process has to be chosen at least as high as max \( \phi(x) \), the maximum of the hazard function of the target non-Poissonian process.

Let the list of spike times \( S = \{ t_1, t_2, \ldots \} \) and \( S' = \{ \hat{t}_1, \hat{t}_2, \ldots \} \) denote the spike trains of the source and the target process, respectively. Then the thinning algorithm is described by the following steps:
0. a) choose inter-event interval distribution of the target process with \( \lambda \) and CV according to Eq. (2)
b) determine \( \phi(x) \) via Eq. (1) and \( \max \phi(x) \)

1. a) set \( R_m \) such that \( R_m \geq \max \phi(x) \)
b) generate a Poisson process \( S \) with rate \( R_m \)

2. set \( t_{latest} \leftarrow t_1 \quad \% \ t_{latest} \) is latest renewal event time
   set \( S' \leftarrow \{ t_{latest} \} \)
   for \( i = 2: n \) do
      draw \( U \) with \( U \sim \xi(0, R_m) \)
      if \( U \leq \phi(t_i - t_{latest}) \) then
         \( t_{latest} \leftarrow t_i \)
         \( S' \leftarrow \{ S', t_{latest} \} \)
      end if
   end for

Note that the first event of the target process is a Poisson event.
In order to obtain a renewal process in “equilibrium” where all inter-spike intervals follow the desired distribution, one has to employ a “warm-up time” (Cox, 1962; Nawrot et al., 2007). Otherwise, strong onset transients may contaminate the spike count statistics (Muller et al., 2007; Deger et al., 2010).

Step 1. a) of the algorithm implies that the hazard function of the target process is bounded. Given this condition the thinning method allows to generate a renewal process with arbitrary hazard function. This could be an empirical one obtained from neuronal data, or a parametrized one. Examples for the latter are Poisson processes with dead time, gamma processes with a coefficient of variation smaller than one and lognormal processes (see Fig. 1, bottom, and Appendix A). The gamma process with CV > 1, however, cannot be simulated (see also discussion).

### 2.2. Correlated renewal point processes

The method described in the previous section enables one to generate single spike trains with any inter-spike interval distribution given its hazard function is bounded. We will now use this technique to simulate populations of renewal point processes with higher-order interactions. To that end, we define higher-order correlations and review how Poisson processes with these multi-neuron events can be obtained (cf. Ehm et al., 2007; Brette, 2009; Staude et al., 2010a,c).

#### 2.2.1. Higher-order correlations

Fig. 3C shows the spike trains \( \{ s_i(t) \} \) of \( N \) neurons. The colored ticks indicate times with simultaneous spikes in at least two neurons, where color encodes the number of neurons that spike synchronously. As can be seen in Fig. 3A via corresponding colors, we represent a pattern of exactly \( n \) coincident spikes by a process \( y_n(t) \) with intensity \( \nu_n \) (see B). In other words, the events of \( y_n(t) \) indicate all instants when such a multi-neuron event of order \( n \) (i.e. only a single spike in case of \( y_1(t) \) occurs).

![Thinning of a Poisson process (grey and red ticks at bottom of A, B) according to the hazard function \( \phi(x) \) (top) of the target renewal process (red ticks at bottom). The dashed line (top) indicates an upper bound \( R_m \) of the hazard function which is chosen as the rate of the source Poisson process. (A) Target process is a Poisson process with dead time. Events which occur within the dead time \( d \) are deleted. (B) Target process is a renewal process with bounded hazard function. One successively tests whether the interval \( x \) between a Poisson event and the latest renewal event is in accordance with the hazard function. That is, one draws a random number \( U \) with \( U \sim \xi(0, R_m) \) (indicated by crosses, top). If \( U \leq \phi(x) \) the event is kept and assigned to the target renewal process (marked red). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of the article.)](image)
The summed spike activity \( \sum_{n=1}^{N} s_i(t) \) of the population can then be represented as
\[
z(t) = \sum_{n=1}^{N} n y_n(t).
\] (3)

Thus, if \( v_n > 0 \) for some \( n \geq 2 \), there are groups of \( n \) neurons showing synchronized activity, which implies that the corresponding single neuron spike trains are correlated with each other. We therefore say “the neuronal spike activity exhibits correlations of order at least \( \xi \)” if the rate \( v_n \) is non-zero for at least one \( n \geq \xi \) (see also Stauder et al., 2010c, Theorem 1). Accordingly, the vector of component rates \((v_1, v_2, \ldots, v_N)\) will be referred to as the “correlation structure”. Moreover, we call its representation as a probability mass function \( f_k(k) = v_k / v_+ \) with \( v_+ = \sum_{n=1}^{N} v_n \) the “amplitude distribution”.  

2.2.2. Generation of Poisson processes with higher-order correlations

2.2.2.1. Compound Poisson process. If the component processes \((y_i(t))_n\) are independent stationary Poisson processes, the sum process \( z(t) \) is a compound Poisson process (CPP). This model can serve as the basis for the generation of various populations of spike trains. That is, given \( N \) independent component Poisson processes \( y_n(t) \) with intensity \( v_n \) \((n=1, \ldots, N)\), copying the event times of \( y_n(t) \) to \( n \) out of \( N \) neurons will result in a population with correlation structure \((v_1, v_2, \ldots, v_N)\). Here, many rules of assignment can be considered (see Stauder et al., 2010b, for more examples): for instance, if \( v_n = 0 \) for \( n \geq 2 \) and the events are assigned alternatively to one out of two neurons, this will result in two gamma processes with shape parameter \( \alpha = 2 \) (cf. e.g. Baker and Gerstein, 2000) and no coincident spikes.

2.2.2.2. Homogeneous Poisson population. If, however, event times are ascribed to each neuron independently one will end up with Poisson processes (Ehm et al., 2007; Brette, 2009; Stauder et al., 2010c). The simplest procedure is to assign the event times uniformly to the \( N \) neurons which gives a homogenous population. In more detail, let \( A \) be a random variable representing the correlation structure, i.e. having the probability mass function \( f_k(k) \) as defined in Section 2.2.1. Then, the single cell processes \( s_i(t) \) have the rate
\[
\lambda_i = \frac{E[A]}{N} v_i.
\] (4)

Moreover, with \( \lambda_c \) denoting the rate of coincident firing of any two cells,
\[
\rho := \frac{\lambda_c}{\lambda} = \frac{E[A^2] - E[A]}{E[A](N-1)}
\] (5)
is the pairwise coincidence probability, which in the case of Poisson processes coincides with Pearson’s count correlation coefficient (Kuhn et al., 2003; Tetzlaff et al., 2008). Note that Eqs. (4) and (5) depend only on the first two moments of the amplitude distribution \( f_k(k) \), and more details of it are not incorporated. Therefore, it is possible to construct Poisson populations with identical single and pairwise statistics that differ with regard to their higher-order correlation structure (Kuhn et al., 2003; Stauder et al., 2010c).

2.2.2.3. Heterogeneous Poisson population. Correlated Poisson processes with different rates and pairwise correlations as well as non-stationarities can be obtained by more complex rules according to which the event times of the component processes are independently assigned to the different neurons. Fig. 4 (right panel) shows an example of a heterogeneous population where the raster plot (bottom) visualizes four different subpopulations. These spike...
trains have been obtained by assigning the event times of $y_i(t)$ to the neurons with different probabilities according to their group membership. Moreover, only those higher-order events belonging to the same order of synchronized activity have been copied into one group. Note that the calculation of the mean pairwise coincidence probability for two spike trains of a non-homogeneous population needs to take into account the specific assignment probabilities (cf. Brette, 2009). Hence, for non-homogenous populations we will rather refer to $c_p$ given by the right-hand side of Eq. (5) as a measure of correlation strength (see Staude et al., 2010c).

### 2.2.3. Generation of non-Poissonian processes with higher-order correlations

We use the previous techniques to introduce a new method for the generation of non-Poissonian processes with higher-order correlations. First, we describe the main procedure and then we elaborate on how the higher-order correlation structure can be predefined.

#### 2.2.3.1. General procedure. Fig. 4C and D shows raster plots of non-Poissonian processes (red dots). These processes have been simulated by first generating a population of correlated Poisson processes (all spikes in the raster plot) as described in Section 2.2.2.

In doing so, the rate of each source Poisson process has been chosen as $R_m = \max_x \phi_{x,CV}(x)$, where $\phi_{x,CV}(x)$ denotes the hazard function of the target spike activity of one neuron with spike rate $\lambda$ and coefficient of variation CV. Then the thinning procedure of Section 2.1.3 was applied to each source process.

Evidently, the thinning affects not only the single neuron spike statistics but also the correlation structure in the population (see Fig. 4). As the spikes in the source spike trains are deleted randomly and independently, thinning reduces the order of higher-order events in a random fashion (for details, see discussion below). As a consequence, the random variable describing the post-thinning order of the multi-neuron event that had order $n$ in the source population is binomially distributed with parameters $n$ and $p$, where $p$ is the survival probability. For an arbitrary amplitude distribution, the above argument implies that the amplitude distribution after thinning $f_A$ is a superposition of binomial distributions

$$f_A(k) \propto \sum_n f_{A,\text{binomial}}(n) \cdot B(k|n, p),$$

where $f_{A,\text{binomial}}(n)$ is the amplitude distribution before thinning. Note that the binomial distribution assigns a non-zero probability to all $k \geq 0$, but within our model events of order zero $(k=0)$ are not accounted for. Hence, to obtain a well-defined probability mass function $f_A(k)$ the right-hand side in Eq. (6) has to be renormalized.

#### 2.2.3.2. Predefining the correlation structure. Not only firing rates and the coefficient of variation CV, but also the correlation structure can be controlled, as we will now explain with a simple example. Consider that we start with a two-point amplitude distribution $f_{A,\text{binomial}}$ that only allows for independent spikes and synchronized spikes of some order $\xi_{\text{syn}} > 1$ only, as illustrated in Fig. 4A (grey bars) and C (grey and red dots). Then

$$f_{A,\text{binomial}}(k) = \delta_{k,1} \cdot (1 - \eta) \cdot \delta_{\xi_{\text{syn}},k},$$

where $\delta_{ij}$ denotes the Kronecker delta ($\delta_{ij} = 1$ if $i=j$, otherwise $\delta_{ij} = 0$). The probability that a spike survives thinning is $\lambda/R_m$. Thus, following the above arguments (cf. Eq. (6)), the amplitude distribution after thinning is a superposition of binomial distributions with parameters $(1, \lambda/R_m)$ and $(\xi_{\text{syn}}, \lambda/R_m)$, which after normalization is

$$f_{A,\text{binomial}}(k) = \frac{\eta \cdot \lambda/R_m \cdot \delta_{k,1} + (1 - \eta) \cdot B(k|\xi_{\text{syn}}, \lambda/R_m)}{\eta \cdot \lambda/R_m + (1 - \eta) \cdot (1 - (1 - \lambda/R_m)^{\xi_{\text{syn}}})}.$$
Disregarding the independent spikes, the mean order of synchronized activity now is
\[ \xi = \xi_{\text{syn}} \cdot \frac{\lambda}{R_m}. \]

As can be seen in Fig. 4 (left top panel) for our above example the distribution \( f_{\text{binomial}}(k) \) (green bars) fits the empirical amplitude distribution after thinning \( f_{\text{syn}}(k) \) (red bars) quite well.

This binomial description enables one not only to predefined the rate \( \lambda \) and the coefficient of variation \( CV \) but also the correlation structure \( \nu = (\nu_1, \nu_2, \ldots, \nu_N) \) of the target processes. The location of the mean order of correlation \( \xi \) can be adjusted via Eq. (9).

In order to get a certain pairwise coincidence probability \( c_p \), one has to determine \( \eta \) and thereby the amplitude distribution of the source processes (see Eq. (7)). This can be achieved by calculating \( c_p \) via the right-hand side of Eq. (5) for the probability mass function \( f_{\text{binomial}}(k) \) and solving for \( \eta \). This yields
\[ \eta = \frac{\xi \cdot (\xi - \frac{\lambda}{R_m} - c_p(N - 1))}{\xi \cdot (\xi - \frac{\lambda}{R_m} - c_p(N - 1)) + c_p(N - 1) \cdot \frac{\lambda}{R_m}}. \]

It should be noted, however, that the binomial description is good only if the interactions of the source Poisson processes are not too strong. In spike trains which are dominated by high order spike pattern deletion of spikes tends to affect correlated neurons in a non-independent manner. In a nutshell, if the spikes belonging to a multi-neuron event survive in two neurons, the same synchronized hazard decides about which spikes of the source processes are kept or rejected next. So should the corresponding follow-up spikes again belong to a common multi-neuron event, then a non-flat hazard will imply the same bias in both neurons. That is, rejection is not independent across neurons any more. This also shows, why this lack of independence depends on both the degree of pairwise correlation (overall rate of patterns) and the degree of deviation from Poissonian firing (non-flat hazard). If not stated otherwise, in the following sections we refer to the regime where the binomial approximation holds.

2.2.3.3. Auto- and cross-correlation functions. The thinning procedure results in non-Poissonian spike trains with an auto-correlation function that is defined by the chosen hazard function (see Fig. 5A-B, grey lines). The cross-correlation function of two target processes has a delta-peak at time lag zero as depicted for Poisson processes with dead time in Fig. 5A-B (red lines). Its height is determined by \( \phi_2 \) and has an upper bound which depends on the correlation structure of the source population and the probability of spike deletion.

For instance, for the population considered in the previous section one obtains from the constraint that \( 0 \leq \eta \leq 1 \) together with Eq. (10) and Eq. (9)
\[ 0 \leq c_p \leq \frac{\xi - \lambda}{N - 1} = \frac{\xi_{\text{syn}} - 1}{N - 1} \cdot \frac{\lambda}{R_m}. \]

Thus, as \( \xi_{\text{syn}} \leq N \), the maximal pairwise coincidence probability is given by \( \lambda/R_m \). Hence, choosing \( R_m \) as small as possible is recommended to keep the range of possible values for \( c_p \) large.

More generally, the strongest correlation in the target processes is obtained by starting with two source processes with \( c_p = 1 \). In a homogeneous population the probability that a spike survives thinning is \( \lambda/R_m \) and thus the probability that the same spike time survives in both spike trains is \((\lambda/R_m)^2\). The pairwise coincidence probability of the target processes is given by the rate of coincident spikes (here: \((\lambda/R_m)^2 \cdot R_m\)) divided by the rate of a spike train (here: \( \lambda \)) which yields \( \lambda/R_m \).

Note that the effect of non-independent spike deletion during the thinning procedure (cf. discussion above) is also visible in the cross-correlation function of the target processes. Fig. 5A shows that the average cross-correlation histogram of two PPDs with \( CV = 0.9 \) exhibits a trough around lag 0. This feature is similar to the one in the corresponding autocorrelations. The amplitude of the former is, however, very small in comparison. The trough in the cross-correlation histogram is more prominent for processes deviating considerably from Poisson (compare Fig. 5A, CV = 0.9 and B, CV = 0.45) and the stronger the processes are correlated (not shown). Additionally, the height of the peak deviates more from the predefined correlation strength.

Our algorithm yields precise coincidences, however, correlations expanded in time can easily be obtained. By adding a random displacement to each spike time of the source Poisson processes one generates non-instantaneous correlations (cf. e.g. Bäuerle and Grübel, 2005; Brette, 2009), and this jitter is preserved under thinning. Fig. 5C, D show the average cross-correlation histograms (red lines) with the same single neuron spike statistics as in A, B but non-instantaneous coincidences. More precisely, the event times have been jittered according to a uniform distribution with support \([–10 \text{ ms}, 10 \text{ ms}]\) before the thinning procedure has been applied. This is reflected in the cross-correlation function by a triangular around lag 0 with a corresponding extent in time.

2.2.3.4. Correlation structure. In paragraph 2.2.3.2 we only explained how the correlation structure can be predefined when the source population exhibits synchronized activity of only order \( \xi_{\text{syn}} \). The ansatz outlined above can easily be extended to more complicated correlation structures. For instance, this can be achieved by describing the amplitude distribution with coincidence patterns of various orders \( \xi_{\text{syn}}^1, \xi_{\text{syn}}^2, \ldots \) before thinning as
\[ f_{\text{binomial}}(k) = \eta \cdot \delta_{1,k} + (1 - \eta) \cdot \sum_{i=1}^{N_{\text{syn}}} \alpha_i \delta_{\xi_{\text{syn}},k}. \]
Here, \( N_{\text{syn}} \) denotes the number of non-zero entries in the amplitude distribution for \( k \geq 1 \) and \( \omega_i \) determines the relative height of each peak with \( \sum \omega_i = 1 \).

Assuming independent deletion of spikes, the resulting amplitude distribution can be described as

\[
\begin{align*}
  f_{\text{binomial}}(k) &= \eta \cdot \lambda / R_m \cdot \delta_{1,k} + (1 - \eta) \cdot \sum_{i=1}^{N_{\text{syn}}} \omega_i \mathbb{E}[K] \xi_{i, \text{syn}, \lambda / R_m] \quad \text{for } k \leq N_{\text{syn}}/2, \\
  f_{\text{binomial}}(k) &= \eta \cdot \lambda / R_m + (1 - \eta) \cdot \sum_{k=1}^{N_{\text{syn}}} \sum_{i=1}^{N_{\text{syn}}} \omega_i \mathbb{E}[K] \xi_{i, \text{syn}, \lambda / R_m] \quad \text{for } k > N_{\text{syn}}/2.
\end{align*}
\]

Proceeding as above yields an expression for \( \eta \) which then, additionally to the choice of the weights \( \omega_i \), allows to redefine the distribution structure of the target non-Poissonian processes. An example with \( N_{\text{syn}} = 2 \) is employed in Section 3.3.2 (see also Appendix D).

2.2.3.5. Heterogeneous populations. Thus far, we focused on populations of homogeneous non-Poissonian processes. However, one can also employ the thinning procedure to generate non-homogeneous populations. Fig. 4, right column, shows such an example. The population consists of four cell assemblies where neuronal spiking is correlated within each group, but not across groups (e.g., Berger et al., 2007). More precisely, four populations of lognormal processes (\( N = 30, 20, 15 \) and 35 neurons) have been independently generated with fixed firing rate \( \lambda = 1.5, 2.5, 12 \) and 0.5Hz, coefficient of variation \( CV = 1.5, 2.5, 2 \) and 1.2), pairwise coincidence probability \( \rho_{ij} = 0.03, 0.08, 0.12 \) and 0.005) and mean order of synchronized activity \( \xi = 2, 6, 7 \) and 3), resulting in a population that is heterogeneous in several respects. The amplitude distribution of the whole population (red bars, B) is well described by the superposition of the amplitude distributions of the four sub-populations (green bars, B). See also Section 3.3.3 where we study this example in more detail.

Another option to obtain a non-homogeneous population is to choose different hazard functions according to which the source Poisson processes are thinned. For instance, if one starts with a homogeneous Poisson population with rate \( R_m = \max_1^4 \xi_i \cdot \phi_{i, \text{CV}}(x) \) where \( \phi_{i, \text{CV}}(x) \) denotes the hazard function of the \( i \)-th target renewal process, then the amplitude distribution is well approximated (and hence controllable) by Eq. (8) by substituting \( \lambda \) with the mean rate \( \lambda \) (not shown). However, the pairwise coincidence probability \( \rho_{ij} \) will not be the same across all neurons in the population, if the deletion probabilities vary strongly across processes. Hence, populations with similar spike rates, but different types of inter-spike interval distributions can be better controlled by this approach than populations with different spike rates. By appropriately choosing the parameters \( R_m, E_{\text{syn}}, \) and \( \eta \) of the correlated source Poisson processes our method allows to simulate various kinds of surrogate data with certain statistics being fixed, while others are varied. For instance, it is possible to generate populations of spike trains with same rates and similar higher-order correlation structures, but different degrees of irregularity. We employ this unique feature of our method in the following chapter (see also Appendix C).

3. Impact of single neuron spike statistics on the estimation and inference of higher-order correlations

We study the impact of single neuron spike statistics on the estimation and inference of higher-order correlations using the recently proposed method of empirical de-Poissonization (EDP, Ehm et al., 2007). Here, we briefly review how the correlation structure \( \{\nu_1, \nu_2, \ldots, \nu_N\} \) and the maximal order of correlation \( \xi \) are estimated, but refer to Ehm et al. (2007) for a mathematical derivation and more details.

3.1. Empirical de-Poissonization

Let \( Z_t \) denote the population spike count, i.e., the number of spikes recorded from a neuronal population within the time interval \([t - 1 \cdot \lambda, t] \) of length \( \lambda \), as depicted in Fig. 3D. The distribution of these bin counts (Fig. 3E) evidently depends on the spike rates of the single cells, but also on their higher-order correlations, since correlations of any order put more weight on the tail of the distribution as compared to the case of independent firing. The distribution thus provides important information about the strength and the order of the correlations. In fact, EDP can reconstruct the entire unobservable correlation structure from that distribution (as indicated by the arrow in Fig. 3 from E to B).

In order to achieve this, EDP assumes the CPP as a model for the population spike activity. Practically, it is reasonable to think of the single-neuron spike trains as Poisson processes, although this is not a mathematical necessity. Apart from this property a large set of different populations can be described by this CPP model and, in particular, by the same correlation structure (cf. Section 2.2.2).

By employing the binned population spike activity one circumvents an apparent shortcoming of the underlying CPP model. A component process \( \nu_i(t) \) represents the exact coincidence of \( n \) spikes, whereas correlated firing of neurons does not show such a high precision. However, the value of the bin count \( Z_t \) will be approximately the same if these \( n \) spikes do not occur at the exact same time but with a jitter smaller than the bin size \( \lambda \).

3.1.1. Estimating the correlation structure

Within the framework of the CPP model the rate of synchronized activity of order \( n \) can be regarded as the ‘de-Poissonized’ version of the ‘Poissonized’ bin counts \( Z_t \):

\[
\nu_n = \frac{1}{2\pi} \int_{-\pi}^{\pi} h^{-1} \log \gamma_n(\theta) e^{-i n \theta} d\theta.
\]

Here, \( \gamma_n(\theta) = \mathbb{E}[e^{i n \theta}] \) denotes the characteristic function of the bin counts \( Z_t \) which uniquely determines, and is uniquely determined, by the distribution.

By substituting the unknown \( \gamma_n(\theta) \) in Eq. 14 with the empirical characteristic function

\[
\hat{\gamma}_n(\theta) = L^{-1} \sum_{l=1}^{L} e^{i n \theta}
\]

one obtains estimates \( \hat{\nu}_n \) of the correlation strengths \( \nu_n \) of order \( n = 1, 2, \ldots \). Note that the estimates \( \hat{\nu}_n \) can, in principle, take negative values. If the sample size \( L \) and the number of zero bins (i.e., bins with \( Z_t = 0 \)) are large, \( \hat{\nu}_n \) is a consistent estimate of \( \nu_n \) for each \( n \). In the present manuscript these conditions are satisfied, and we refer the reader to Ehm et al. (2007) for more details and some hints on corrective methods.

3.1.2. A lower bound for the maximal order of correlation

In addition to the direct rate estimates \( \hat{\nu}_n \), Ehm et al. (2007) proposed the maximal order of correlation \( \xi \) as a measure of interest (cf. also Staude et al., 2010b,c). The (unknown) true \( \xi \) as defined in Section 2.2.1 can formally be expressed as

\[
\xi = \max \left\{ m \mid \rho_m > 0 \right\}
\]

where \( \rho_m := \sum_{n=m}^{\infty} \nu_n. \)

Hence, an estimate for \( \xi \) can be obtained by successively estimating the rate tail \( \rho_m \) and testing whether it deviates significantly from 0. More precisely, one starts with \( m = 2 \) and increases \( m \) until
the null hypothesis $H_0^m : \rho_m = 0$ can not be rejected at a significance level $\alpha$. Let $p_m$ denote the $p$-value associated with such a hypothesis test, then by this procedure the estimated $\hat{\xi}$ is given as
\[
\hat{\xi} = \min \{ m \mid p_m > \alpha \} - 1.
\]

The $p$-value is determined in the following way: Under the CPP model $\rho_m$ can be represented as
\[
\rho_m = \frac{1}{2\pi} \int_{-\pi}^{\pi} h^{-1} \log \gamma_h(\theta) \frac{e^{-m\theta}}{1-e^{-\theta}} d\theta.
\]

Again plugging in $\hat{\gamma}_h(\theta)$ for $\gamma_h(\theta)$ in Eq. 18 yields an estimate $\hat{\rho}_m$ for sampled data. Under $H_0^m$, $\rho_m$ is approximately normally distributed with mean zero and variance $\sigma_m^2 = (\ln) \sum_{m,m}$ where
\[
\sum_{m,m} = \frac{1}{2\pi} \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \frac{1}{h} \left( \frac{\hat{\gamma}_h(\theta_1 + \theta_2) - 1}{\hat{\gamma}_h(\theta_1) \hat{\gamma}_h(\theta_2)} \right) \times e^{-m\theta_1} e^{-m\theta_2} \frac{1 - e^{-\theta_1}}{1 - e^{-\theta_2}} d\theta_1 d\theta_2.
\]

This expression, too, can be estimated by plugging in $\hat{\gamma}_h(\theta)$. An approximate $p$-value $p_m$ for testing $H_0^m : \rho_m = 0$ against $H_1^m : \rho_m > 0$ is thus given by
\[
p_m = 1 - \Phi \left( \frac{\hat{\rho}_m}{\hat{\sigma}_m} \right)
\]
where $\Phi$ denotes the cumulative distribution function of the standard normal distribution.

As shown in Appendix B, $\hat{\xi}$ as defined by Eq. (17) is, approximately, a lower confidence limit for $\xi$ at the level $\alpha$. This means that irrespective of the underlying spike rates, the probability that $\hat{\xi}$ is larger than the unknown parameter $\xi$ is at most $\alpha$.

3.2. Robustness of empirical de-Poissonization

The analysis of higher-order correlations with EDP requires choosing a bin size $h$ to determine the population spike count. Actually, this is the only adjustable parameter given a set of experimental data. Picking $h$ is hence the only possibility to affect results. Therefore, we study the robustness of EDP in dependence of $h$.

Fig. 6A shows the population spike count for a population of 15 correlated lognormal processes with coefficient of variation $CV = 0.65$ (dark blue) and 15 Poisson processes (light blue) with identical correlation structure, averaged over 100 realizations (see C for details on the generation of the surrogate data). For a small bin size of 2 ms (hatched bars) the distributions are almost indistinguishable. However, for a bigger bin size of 10 ms the distributions differ from each other (see e.g. $k = 0, 1$). Accordingly, applying EDP to these data sets yields similar results only for small bin sizes. The true correlation structure (red bars in Fig. 6B) is well estimated not only for the Poisson processes (light blue) but also for the lognormal processes (dark blue) given a bin size of $h = 2$ ms (hatched bars). While for a bin size of 10 ms (filled blue bars) the estimated correlation structure $\hat{\rho}_1, \hat{\rho}_2, \ldots, \hat{\rho}_N$ is on average correct for the Poisson processes, the processes violating the model assumptions induce a biased estimation. Correspondingly, for a small bin size the estimated maximal order of correlation for the lognormal processes (dark blue in Fig. 6C) lies on average only slightly below the one estimated for the Poisson processes (light blue). While in both cases $\hat{\xi}$ decreases with increasing bin size it appears that the results differ more for larger bins.

3.2.1. Estimated correlation structure

We quantify the degree to which non-Poissonian spiking impairs the estimation of the correlation structure in relation to the situation of Poissonian spiking by using the following estimation error
\[
\text{error} = \left( \sum_k k \cdot \hat{v}^{\text{non-PP}}_k - v^{\text{true}}_k \right) \left( \sum_k k \cdot \hat{v}^{\text{PP}}_k - v^{\text{true}}_k \right)
\]
where the angular brackets denote averaging over 100 realizations, $\hat{v}^{\text{non-PP}}_k$ is the rate of synchronous spiking of order $k$ estimated by EDP for (non-)Poisson processes, and $v^{\text{true}}_k$ is the actual rate in the considered data set. Note that our measure also accounts for the estimation error of EDP on Poissonian data: it measures the mean fraction of the misestimated overall rate (the population spike rate is $\sum_k k \cdot v_k$), or number of spikes, which is due to the deviation of the population spiking from the CPP model.

We determined the estimation error for populations of lognormal processes with different spiking irregularity, but the same correlation structure. For almost all values of $CV$ in the range [0.3, 3.5] we find a strong bin size dependence of the results (see Fig. 7A). However, the estimation error does not only depend on the bin size, but also on the CV. For very irregular spiking and spiking with a coefficient of variation close to 1 the error is considerable larger than for processes with CV close to 2 or to 0.3.

The dependence of the error on the coefficient of variation is different when the spike activity is not simulated as lognormal processes but as gamma processes (Fig. 7C) or Poisson processes with dead time (PPDs, 7D). Note that all remaining parameters regarding firing rate and correlation structure of the analyzed populations are identical. We conclude that the CV alone is not sufficient to measure the influence of non-Poissonian spiking on the estimation error.

While the estimation error has the same order of magnitude for all three types of point processes it is considerably reduced in Fig. 7B. In this example, the same correlation structure is realized by more neurons. Increasing the population size three-fold and, hence, decreasing the spike rate of each neuron three-fold, leads to a maximal error which is approximately three times smaller.

3.2.2. Lower bound for the maximal order of correlation

In order to investigate the effect of the bin size also on the maximal order of correlation $\xi$, we quantify its error by the average ratio for non-Poissonian and Poissonian populations, $\hat{\xi}^{\text{non-PP}} / \hat{\xi}^{\text{PP}}$. As can be seen in Fig. 8A, for our population of 15 correlated lognormal processes $\xi$ tends to be underestimated for $CV < 2$, and overestimated for $CV > 2$. While for small bin sizes the error is negligible, for larger bin sizes it can be substantial. A similar degree of underestimation is obtained for gamma processes (Fig. 8C) and PPDs (Fig. 8D). As observed before, the precise dependence of the error on the coefficient of variation varies with the point process model employed.

In line with our findings regarding the estimated correlation structure, the consequences of non-Poissonian firing for the maximal order of correlation are less pronounced when the population size is increased, without changing the correlation structure (compare Fig. 8A and 8B). Furthermore, the value of $CV$ marking the turning point between under- and overestimation of $\xi$ is shifted to the right for $N = 45$ as compared to $N = 15$.

In summary, we find that for small bin sizes both the correlation structure and the maximal order of correlation is well estimated also for non-Poissonian processes. However, the actual estimation error depends non-trivially on various factors (type of inter-spike interval distribution, coefficient of variation, population size and spike rate) and, hence, can hardly be predicted for correlated processes not considered here. Consequences and conclusions of our results are drawn in the following section and discussed in Section 4.2.
3.3. Validation of empirical de-Poissonization in neuronal data

As illustrated by our robustness studies, a deviation from the model assumptions can, under certain conditions, lead to a marked misestimation of higher-order correlations and hence, to a misinterpretation of experimental data. Therefore, a strategy is needed to either avoid the misestimation a priori, or, if not possible, to at least quantify its degree a posteriori.

Our findings suggest to choose the bin width of the population histogram small in order to get reliable results. However, there is

---

**Fig. 6.** Distribution of population spike count (A), corresponding (estimated) correlation structure (B) and inferred maximal order of correlation (C). (A) Population spike count of 15 lognormal processes with a coefficient of variation CV = 0.05 (non-PP, dark blue) and Poisson processes (PP, light blue) for different bin sizes h (hatched bars: 2 ms; filled bars: 10 ms) averaged over 100 realizations of 500 s each. Apart from the type of inter-spike interval distribution the populations have identical statistics, with each neuron having a spike rate of 6 Hz. (B) The underlying correlation structure is indicated by red bars, which corresponds to a pairwise coincidence probability of c_{ij} = 0.05. Blue bars show the estimates of these true rates obtained by EDP from the population spike count (same coding as in A). (C) The mean maximal order of correlation estimated by EDP is depicted in dependence of the bin size h for Poisson processes (light blue) and non-Poissonian processes (dark blue). The corresponding true correlation structure is depicted by red horizontal bars on the left. Error bars indicate standard deviation across independent trials. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of the article.)

**Fig. 7.** Error in estimating the correlation structure when model assumptions are not fulfilled. The bias due to the violation of the CPP/Poisson assumption is quantified by the error measure (Eq. (21)). This bias is depicted for different populations of renewal processes with identical correlation structures as illustrated in Fig. 6 (for details see C). Results are shown in dependence of the coefficient of variation (CV) for various bin widths h of the population histogram. The bin size ranges from 1 ms (dark red line) to 20 ms (light blue line). A: Population of N = 15 correlated lognormal processes with firing rate \( \lambda = 6 \text{ Hz} \) each. B: Population of N = 45 correlated lognormal processes with \( \lambda = 2 \text{ Hz} \). C: Population of N = 15 correlated gamma processes with \( \lambda = 6 \text{ Hz} \). Additionally, the order parameter \( \alpha \) of the gamma distribution (cf. A.2) corresponding to CV is indicated. D: Population of N = 15 correlated Poisson processes with dead time with \( \lambda = 6 \text{ Hz} \). Additionally, the dead time \( d \) of the inter-spike interval distribution (cf. A.1) corresponding to CV is marked. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of the article.)
Fig. 8. Mean ratio of the lower confidence bound on the maximal order of correlation of non-Poissonian processes (\(\hat{\xi}_{\text{non-PP}}\)) and Poisson processes (\(\hat{\xi}\)). Same populations as in Fig. 7 have been analysed. Results are shown in dependence of the coefficient of variation (CV) for various bin widths \(h\) of the population histogram. Order of panels and color coding as in Fig. 7. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of the article.)

a lower bound on the bin size to render the method insensitive to imprecise coincidences (see Section 3.1). Besides, we saw examples where bin sizes of 5 ms yielded reasonable results but also ones where bin sizes of even 20 ms were acceptable. Hence, it is difficult to decide a priori which bin size is appropriate to avoid a misestimation of higher-order correlations.

As a consequence, we need a possibility to assess the performance of EDP. In our robustness studies the performance of EDP could easily be checked because the underlying correlation structure was known. This is not the case in applications where only the experimental data is given, and we will circumvent this problem by the use of surrogate data. We will first propose the main procedure (Fig. 9A) before we illustrate its performance with an example (Fig. 9B and Fig. 9B).

3.3.1. General procedure

Validation of EDP in neuronal data proceeds in four steps. First, one determines the EDP estimate of the correlation

Fig. 9. Validation of empirical de-Poissonization in neuronal data. (A) Concept. Step I consists of the application of EDP to neuronal data to obtain an estimate \(\hat{v}\) (purple) for the true unknown correlation structure \(v^*(v_1, v_2, \ldots, v_n)\) (pink). Thereafter, one imitates the neuronal data (step II). This implies to fit a realizable (according to section 2.2.3) correlation structure \(\hat{v}^p\) to the estimated one \(\hat{v}\) and the estimation of various statistics of the experimental spike trains. Surrogate data with correlation structure \(\hat{v}^p\) are generated, and the higher-order correlations estimated \(\hat{\xi}\) (dark blue) (step III). The comparison of \(\hat{v}^p\) with \(\hat{\xi}\) gives an approximation of the relation between the unknown \(v\) and the estimated \(\hat{v}\) (dashed arrow, step IV). (B and C) Example, same color coding. Shown is the true correlation structure \(v\) (pink bars, B) of a population with four disjoint cell assemblies (see Fig. 4, right column, and D). The correlation structure estimated by EDP for a bin size of 5 ms from a 200 s sample is depicted in purple. For details on the fitted correlation structure \(v^p\) (yellow bars), see D and text. Surrogate data with true correlation structure \(v^*\) (red bars, C) have been simulated for 100 times. The average results by EDP with a bin size of 5 ms are depicted by the light blue bars for lognormal processes, and by light blue bars for the Poisson processes with identical correlation structure. Error bars denote standard deviation. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of the article.)
structure, \( \hat{v} = (\hat{v}_1, \hat{v}_2, \ldots) \), from the given data (step I, Fig. 9A, left). The task then is to find a way to generate surrogate data that mimic the basic statistical features of the original data, particularly the unknown correlation structure \( v \), as close as possible (step II, center). We consider the scenario where the experimental spike trains can be approximated by a renewal process with a bounded hazard function. Corresponding surrogate data can then be simulated by thinning correlated Poisson processes as described in Section 2.2.3. The correlation structure of the latter has to be chosen such that the correlation structure resulting after thinning, \( v_{sur} \), matches \( \hat{v} \) as well as possible (realized by the fit of \( f_{obs} \) in Fig. 9A, left, see Section 3.3.2 for more details). Finally, one determines the respective EDP estimate \( \hat{v}_{sur} \) for each surrogate data set generated according to that scheme (step III, Fig. 9A, right). On repeating this step many times, one obtains a Monte Carlo estimate of the distribution of \( v_{sur} \) (and likewise of the difference \( \hat{v}_{sur} - v_{sur} \), or other related statistics), which then serves as a substitute for the unknown distribution of \( \hat{v} \) one is actually interested in (step IV, dashed arrow in Fig. 9A). A means appropriate for assessing the bias of the EDP estimates under model misspecification would be the distribution of the random variable

\[
d (v, \hat{v}) = \frac{\left( \sum_k k \cdot v_k - \hat{v}_k \right)}{\left( \sum_k k \cdot v_k \right)} .
\]

(22)

The closer it is concentrated near zero, the better the performance of EDP. Of course, this distribution is unknown in applications since \( v \) is unknown. However, using our surrogate data method we may estimate it by the empirical distribution of the quantities \( d(\hat{v}_{sur}, v_{sur}(j)) \) where \( v_{sur}(j) \) denotes the EDP estimate of the correlation structure obtained in the \( i \)-th simulation of a surrogate data set.

3.3.2. Example

An example is shown in Fig. 9.B and C). The pink bars (left) depict the correlation structure of an irregularly spiking population which consists of four disjoint cell assemblies as in Fig. 4 (right column).

3.3.2.1. Step I. The correlation structure \( \hat{v} \) has been estimated by EDP from a 200 s sample using a bin size of 5 ms (purple bars, 9B). It slightly deviates from the true correlation structure.

3.3.2.2. Step II. As in a real data scenario where this comparison is not possible, we validate the performance of EDP by the procedure outlined above. Initially, several parameters of the population in question have to be estimated: The number of neurons \( N \) and their firing characteristics, in particular their firing rates and inter-spike interval irregularity. Then we generate surrogate data with these parameters, including the estimated correlation structure \( \hat{v} \). However, as has been described in Section 2.2.3 it is not possible to generate non-Poissonian processes with arbitrary correlation structure by our procedure. Therefore, we have to approximate \( \hat{v} \) by a correlation structure \( v^{\text{fit}} \) which can be simulated. In our example, \( \hat{v} \) looks like a distribution with two peaks for \( v_2, \ldots, v_N \). Hence, we fitted the correlation structure \( v^{\text{fit}} \) (yellow bars) defined by Eq. (13) with \( N_{syn} = 2 \) to \( \hat{v} \) (see Appendix D for more details). This fit yields the parameters needed to construct the correlated Poisson processes, which then have to be thinned to obtain the non-Poissonian processes with correlation structure \( v^{\text{fit}} \). In doing so, we assumed that we had properly estimated the number of neurons and identified the inter-spike interval distribution as lognormal. However, we neglected the fact, that we had a heterogeneous population with respect to the correlation structure because EDP wouldn’t allow this inference from real data either. Furthermore, we simulated a homogenous population of lognormal processes with uniform rate \( \hat{\lambda} = (\hat{\lambda}_i) \) and coefficient of variation \( CV = (CV_i) \), \( 1 \leq i \leq N \), i.e. the mean spike statistics of all neurons.

3.3.2.3. Step III. The resulting surrogate data has a mean correlation structure \( v_{sur} \) (red bars in Fig. 9C). Since Eq. (8) is only an approximation, \( \hat{v}_{sur} \approx v_{sur} \). This surrogate data has been generated and analyzed with EDP for 100 times.

3.3.2.4. Step IV. For our particular example, a comparison of the true correlation structure \( v_{sur} \) (red bars) and the estimated one \( v_{sur} \) (dark blue bars) of the surrogate data suggests that the true correlation structure in the neuronal data is slightly misestimated, but that its overall shape is approximated quite well. As outlined above, the degree of misestimation can be quantified via the error measure \( d(v, \hat{v}) \) (Eq. (22)). Assuming consistency and normality, a 95% confidence interval for the deviation of \( \hat{v} \) from \( v \) (i.e. \( d(v, \hat{v}) \)) is approximately given by the mean of \( d(v_{sur}, \hat{v}_{sur}) \) plus/four times its standard deviation. This corresponds to [0.09, 0.33] in our example and hence contains the value of 0.16 which we obtained for the neuronal data set.

In order to get an idea how well EDP would have performed if model assumptions were fulfilled, one can additionally generate and analyze Poisson processes with a correlation structure \( v_{sur} \) identical to the one of the non-Poissonian processes. In our particular example, with \( d(v_{sur}, \hat{v}_{sur}) = 0.2 \pm 0.07 \) the results for the Poisson processes (light blue bars) are of a quality similar to the one for lognormal processes. Taken together with the rather large error bars of the rate estimates, this suggests that the misestimation in the neuronal data might be rather due to a small sample size, and less due to a deviation from model assumptions.

4. Discussion

We proposed a novel method to simulate non-Poissonian point processes with predefined higher-order correlations. We employed the new method to investigate the impact of single cell spiking statistics on the inference of higher-order correlations from empirical spike data. Furthermore, we outlined a strategy to assess the reliability of results obtained by empirical de-Poissonization on experimental data.

We will discuss the limitations and possible extensions of our approach to generate correlated spike trains and briefly compare it with methods that have recently been proposed in other contexts. Furthermore, we will contrast the results of our robustness study with those of others as far as possible, and explain the relevance and consequences of our findings for the analysis of biological spike trains.

4.1. Generation of non-Poissonian processes with higher-order correlations

4.1.1. Limitations and possible extensions

The method proposed here is based on the deletion of Poisson events. A bounded hazard function of the target process is required, such a property is shared by various renewal processes commonly employed in neuroscience. However, irregular gamma processes with a coefficient of variation of the ISI distribution larger than one represent a counter-example. Yet its overrepresentation of short inter-spike intervals is inconsistent with neuronal refractory periods and, hence, it is not first choice for mimicking irregular spiking in any case.

There are some constraints with respect to the across-neurons correlation structures that can be simulated employing the proposed method. First of all, we are restricted to positive correlations. That is, within our framework only surplus coincidences of spikes can be realized, but not a systematic lack of spikes, e.g. due to a direct inhibitory effect of one neuron on another one. This is due to the fact that based on their construction the Poisson processes can only be positively correlated (cf. Bäuerle and Grubel, 2005; Johnson
and Goodman, 2009), and these correlations are only weakened by thinning. While there is no way to overcome this constraint within the framework proposed here, the issue of imprecise coincidences can be addressed in our model by adding a random jitter to each spike time. In a similar way, elaborated shifting of the source processes against each other in time can yield precise firing patterns across neurons (Abeles et al., 1993; Prut et al., 1998). Moreover, our method does not allow to the realization of arbitrary amplitude distributions in combination with non-Poissonian processes which is possible for Poisson processes. For instance, we are not able to simulate a population of lognormal processes where synchronous spiking only occurs in a fixed number $k$ of neurons (i.e. $v_k = 0$ only if $k \in \{ 1, 2 \}$). Although the shape of amplitude distributions in real neuronal networks is unknown, it is quite unlikely that this scenario represents a biologically relevant example. Instead, an amplitude distribution with many nonzero entries is to be expected. And such a case can be realized by our method. Admissible amplitude distributions are superpositions of several binomial distributions. This allows a great variety of correlation structures to be modeled. Although the binomial approximation of the correlation structure holds only for weak interactions, recent reports about very small pairwise correlations (Ecker et al., 2010; Renart et al., 2010) suggest that our description fits biologically relevant parameter ranges.

We have outlined the generation of correlated spike trains with spike statistics which vary across neurons (see Section 2.2.3). However, a general approach to simulate a heterogeneous population of non-Poissonian processes with predefined correlation structure is expected to be rather difficult: As explained in Brette (2009), even the construction of correlated Poisson processes with arbitrarily defined single-neuron rates and pairwise correlations without any restrictions on the higher-order correlations can be cumbersome.

In this study, we focused on the generation of non-Poissonian processes in a stationary regime. However, neurons exhibit time-varying spike rates especially when they are driven by an external stimulus. Hence, generalizing our method to non-stationary scenarios is of major importance. While higher-order correlations in non-stationary spike trains have recently been considered for processes constituting a compound Poisson process (Staude et al., 2010b), here an extension is required for non-Poissonian processes. Two possible solutions are worth mentioning. The first idea is to generate correlated non-Poissonian processes in operational time, and to apply nonlinear time warping in a second step (Cox and Isham, 1980; Brown et al., 2002). If the correlation structure is defined by (exact) coincidences in operational time, they will be preserved if they are subject to the same time warping. Hence, non-Poissonian processes exhibiting the same non-stationarity can be equipped with a predefined higher-order correlation structure. The second ansatz is based on the fact that the thinning procedure applied to Poisson processes can be adapted straightforwardly to obtain non-stationary non-Poissonian processes. A non-stationary renewal process is determined by its hazard function $\phi(x, t)$, where $x$ denotes the time elapsed since the last event, and $t$ is the clock time (cf. Kass and Ventura, 2001). Thus, the generation of correlated non-stationary non-Poissonian processes is possible again via thinning of correlated Poisson processes by use of a 2-parameter hazard function. However, it remains to be explored how the correlation structure is affected by the thinning procedure.

4.1.2. Comparison with existing methods
Our approach to simulating correlated spike trains is weakly related to the one suggested by Baker and Gerstein (2000). They generated correlated non-Poissonian processes via non-probabilistic thinning of correlated Poisson processes. However, their procedure was restricted to gamma processes with integer-valued order parameters (this implies $CV < 1$), and they considered only two parallel processes. Here, we have adapted this idea to more than two processes (Staude et al., 2007, cf. for an extension of their ansatz to higher-order correlations). The method described here, however, employs independent probabilistic thinning, which can be applied to any renewal process with a bounded hazard function.

Within the last years various methods have been suggested to generate populations of correlated spike trains (e.g. Bohle et al., 2000; Kuhn et al., 2003; Niebur, 2007; Brette, 2009; Krumin and Shoham, 2009; Macke et al., 2009; Onken et al., 2009; Onken and Obermayer, 2009; Gutnisky and Josić, 2010; Lyamzin et al., 2010; Krumin et al., 2010). However, in most approaches higher-order correlations merely occur as a by-product. Sometimes, the resulting probability structure can be mathematically described as e.g. in Krumin and Shoham (2009). Only the copula based framework (e.g. Onken et al., 2009; Onken and Obermayer, 2009) allows the control of both second-order and higher-order interactions. In contrast to our work, they can also generate negative correlations. However, spike counts, rather than spike times, are targeted and, therefore, it is not clear what the corresponding spike trains would look like. In turn, it is to be expected that for a given type of spike trains defined in continuous time the counting properties strongly depend on the bin size used to represent them (Tetzlaff et al., 2008). We would like to stress that the different frameworks of correlated spike train modeling also adopt different concepts of higher-order correlations and, hence, the choice of any particular model must be carefully justified in view of its envisaged application (Staude et al., 2010a).

4.1.3. Application
Although neuronal spike activity is known to deviate from Poissonian statistics (see e.g. Kuffler et al., 1957; Beyer et al., 1975; Burns and Webb, 1976; Levine, 1991; Amarasingham et al., 2006; Nawrot et al., 2008; Minich et al., 2009; Maimon and Assad, 2009; Shionomoto et al., 2009), Poisson processes are nevertheless often employed in modeling and for analyzing experimental data. Recently, the need to consider non-Poissonian spiking to investigate neural computation has been pointed out (Câteau and Reyes, 2006; Lindner, 2006; Ly and Tranchina, 2009; Deger et al., 2011). Our proposed method enables the simulation of such spike trains. While we have focused here on higher-order correlations, our method also contributes to the available possibilities to generate pairs of correlated neurons within the renewal framework (see above) and hence allows to study e.g. the transfer of pairwise correlations in a richer and biologically more realistic environment (cf. Rosenbaum and Josić, 2011). Here, we used non-Poissonian processes with higher-order correlations to investigate the impact of non-Poissonian spiking on the estimation and inference of higher-order correlations by empirical de-Poissonization. Our point process model can, of course, also be employed for the analysis of the sensitivity and robustness of other methods, which aim to detect coordinated spike activity (as e.g. Grün et al., 2002a,b; Pipa et al., 2008; Staude et al., 2010b,c; Lopes-dos Santos et al., 2011).

4.2. Impact of single neuron spike statistics on the estimation and inference of higher-order correlations

4.2.1. Comparison to other results
Our robustness study showed that non-Poissonian spike trains only weakly distort the estimation of higher-order correlations when using very small bin sizes. This finding reflects the fact that, for very small bin sizes non-zero bin counts of a single spike train are sparse, and the corresponding count distribution is very close to a Poisson distribution. However, non-Poissonian spike trains may also imply serial correlations between consecutive spike counts if the bins are not chosen large enough, and this property can be more important for other analysis tools (see Roudi et al., 2009).
To our knowledge, there are currently no other studies which investigate the impact of single-neuron spike statistics on the estimation and inference of higher-order correlations. While the work by Onken et al. (2009) does treat higher-order correlations, their investigation of the importance of marginal spike statistics does not directly incorporate correlations. Rather, they considered the Shannon information of the single neuron spike count distribution. They found that violations of the Ising model assumption (i.e., a binomial distribution of the single-neuron spike count), which is frequently exploited in correlation analysis (e.g., Schneidman et al., 2006; Shlens et al., 2006), result in notable deviations of the estimates from the true entropy. In agreement with our findings, they concluded that it is important to use the proper single-neuron spike count distribution. Due to the different frameworks underlying their and our studies, it is difficult to compare the results in more detail.

The robustness study reported in Grün (2009) and the investigations performed in Tetzlaff et al. (2008) are more in line with our analysis, though restricted to pairs of spike trains. Unitary event (UE) analysis (Grün et al., 2002b,a;), which assumes independent Poissonian firing in its null-hypothesis, exhibits slightly different behavior for very irregular spiking where, in contrast to our results, no false positives occur. This is due to the fact that, contrary to EDP, UE operates on binned and clipped spike activity; i.e., spikes elicited by a neuron within one time bin are counted as one event irrespective of the actual number of spikes. Tetzlaff et al. (2008) revealed a qualitatively similar bin size dependence for Pearson’s count correlation coefficient as an estimator for the pairwise coincidence probability just as we found for our higher-order correlation measures: in the case of non-Poissonian spiking the estimates become worse for larger bin sizes.

An alternative to EDP for estimating the maximal order of correlation is a method called CuBIC (Cumulant Based Inference of higher-order Correlations; Staude et al., 2010b,c). Here we show no full analysis of CuBIC’s performance, but initial investigations revealed that CuBIC (Staude et al., 2010c) can be affected by non-Poissonian spiking in a qualitatively similar way as EDP.

4.2.2. Relevance and consequences of our findings

The parameter range, in which we investigated the impact of single-neuron spike statistics on the estimation of higher-order correlations, has been chosen to be as biologically reasonable as possible. That is, we matched the coefficient of variations to those reported by various studies (e.g., Sotnik and Koch, 1993; Nawrot et al., 2007, 2008; Maimon and Assad, 2009; Ponce-Alvarez et al., 2010) where we also considered ranges of much more irregular spiking (i.e., CV > 2). Spike rates have been set to rather small values due to recent reports on very low firing rates in both spontaneous and evoked activity (Lee et al., 2006; Hromádka et al., 2008; Haider et al., 2010; Wolfe et al., 2010).

Hence, care must be taken in interpreting measured higher-order correlations, especially when comparing results obtained under different stimulus conditions and from different cortical areas, where the neuronal spiking irregularity seems to vary (Maimon and Assad, 2009; Shinomoto et al., 2009). In order to keep the estimation bias as small as possible it is recommended that a small bin width is chosen. This is suggested by our results on bin size dependence, which is in line with the theory that the superposition of independent renewal processes locally constitutes a Poisson process (see e.g. chapter 6 in Cox (1962) and references therein, as well as e.g. Lindner (2006)). That is, one can always choose the bin size to be so small that the superposed spike count is approximately Poisson distributed, and the larger the bins chosen the more one deviates from this condition.

However, as we pointed out in Section 3.3, this can neither entirely solve all remaining problems, nor is there a need for this. Coordinated neuronal activity has only a certain degree of precision which has to be accounted for by a sufficiently large bin size. Besides, the degree of misestimation can be negligible even for rather large bin sizes, as we found for the examples of large sparsely spiking populations. Instead, we therefore suggest the use of our proposed method to evaluate the reliability of results by EDP for the analysis of some experimental data.

4.3. Validation of results by empirical de-Poissonization in neuronal data

While it is common to use surrogate data to judge the significance of correlations and spike patterns (see e.g. Gerstein (2004), Grün (2009), Louis et al. (2010b,a)) we had to adapt these approaches as they are based on the assumption that the null-hypothesis includes independent spiking. In contrast, by using EDP we either directly estimate the synchronized activity of various orders, or we infer a lower bound for the maximal order of correlation via successively testing null-hypotheses, assuming correlations only up to a certain maximal order. That is, considering independent processes is insufficient, and we outlined how correlations can be taken into account as well.

We proposed to characterize the performance of EDP for a parameter set matching the experimental data and concluded that our results for the neuronal data can be trusted to a similar degree. In doing so, we assume that the results by EDP change smoothly within the large parameter space of non-Poissonian firing, affirmed by our robustness studies. Furthermore, as in other approaches, the performance of our method depends on the reliability of the estimation of single-neuron spike statistics. However, as we showed by an example, the quality of the results can already be well judged based on population averaged spike rates and coefficient of variations.

We introduced a method to assess the estimated correlation structure. With the help of additional Poisson processes it can, of course, also be adapted for the evaluation of the inferred maximal order of correlation (cf. Section 3.2.2).

4.4. Conclusions

We presented a novel method to generate non-Poissonian processes with defined higher-order correlations. Furthermore, we used data generated by this technique to perform a robustness study with empirical de-Poissonization (EDP, Ehm et al., 2007). The results emphasize the need to carefully calibrate estimators of higher-order correlations prior to their application. More precisely, the single neuron spike statistics represent crucial parameters which have to be taken into account when interpreting measurement for surrogate and experimental data. As a consequence of our findings, we ultimately proposed a method to assess the reliability of results obtained by EDP.

Acknowledgments

We thank A. Jasper, F. Siegfried and D. Suchanek for helpful comments on an earlier version of this manuscript. Supported by the German Ministry for Education and Research (BMBF grant 01GQ0420 to the BCCN Freiburg), the IGPP Freiburg and the German Research Foundation (DFG SFB 780).

Appendix A. Renewal processes and its hazard functions

Our proposed method to generate non-Poissonian processes with higher-order correlations requires that the hazard function of the renewal process is bounded (see Section 2.1.3). We describe
here the necessary details of the processes which we considered in our manuscript.

### A.1. Poisson process with dead time

The inter-spike interval distribution \( f_{\lambda,\theta}(x) \) and the hazard function \( \phi_{\lambda,\theta}(x) \) of a Poisson process with dead time read formally as

\[
f_{\lambda,\theta}(x) = \lambda e^{-\lambda(x-\theta)}I_{[\theta,\infty)}(x) \tag{A.1}
\]

and

\[
\phi_{\lambda,\theta}(x) = \lambda I_{[\theta,\infty)}(x), \tag{A.2}
\]

where \( I_A(x) \) equals 1 if \( x \in A \) and 0 otherwise. Hence, the maximum of the hazard function is \( \lambda \). The spike rate \( \lambda \) and the coefficient of variation \( CV \) relate to the parameters \( \lambda, \theta, d, m, k \) via:

\[
\lambda = \frac{\lambda}{CV} \quad \text{and} \quad d = \frac{1 - CV}{\lambda}. \tag{A.3}
\]

### A.2. Gamma processes

The inter-spike intervals follow a gamma distribution which is defined as

\[
f_{\alpha,\beta}(x) = \frac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\beta x}, \tag{A.4}
\]

where \( \Gamma \) denotes the gamma function. The parameters \( \alpha \) and \( \beta \) can be expressed in terms of the spike rate \( \lambda \) and the coefficient of variation \( CV \) via

\[
\alpha = \frac{1}{CV^2} \quad \text{and} \quad \beta = \frac{\lambda}{CV^2}. \tag{A.5}
\]

The hazard function of gamma processes decreases monotonically for \( CV > 1 \) where \( \phi_{\alpha,\beta}(x) \to \infty \) as \( x \to 0 \). In contrast, it increases monotonically for \( CV < 1 \) with \( \phi_{\alpha,\beta}(x) \to \beta \) as \( x \to 0 \) (Cox, 1962, see [p. 20]).

### A.3. Lognormal processes

The inter-spike interval distribution of lognormal processes reads as

\[
f_{\mu,\sigma}(x) = \frac{1}{x \sigma \sqrt{2\pi}} e^{-\frac{(\ln(x) - \mu)^2}{2\sigma^2}}, \tag{A.6}
\]

where

\[
\mu = -\log \frac{1}{2} \ln(CV^2 + 1) \quad \text{and} \quad \sigma = \sqrt{\log(CV^2 + 1)}. \tag{A.7}
\]

Its hazard function begins always at zero, rises to a maximum and then decreases to zero very slowly. This maximum does not have a closed analytical expression and, hence, needs to be determined numerically (Sweet, 1990, see [for more details]).

### Appendix B. Lower confidence bound for the maximal order of correlation \( \xi \)

Here we demonstrate that under the CPP model \( \hat{\xi} = \min \{ m \mid P_m > \alpha \} = 1 \) is an approximate level-\( \alpha \) lower confidence limit for the maximal correlation order \( \xi \) (cf. Section 3.1.2).

For simplicity we assume that the test statistic \( T_m \) is exactly standard normally distributed under the null-hypothesis \( H_0 \) (so that the p-value \( P_m \) from Eq. (20) is exactly uniformly distributed on the unit interval). It has to be shown that whatever the correlation structure \( \nu = (\nu_1, \nu_2, \ldots) \) underlying the CPP, one has

\[
P_\nu(\hat{\xi} \leq \xi) \geq 1 - \alpha.
\]

The first step is a basic fact about confidence regions. Let the observation \( x \) be distributed as \( P_\theta \) for some unknown parameter \( \theta \in \Theta \). Furthermore, let \( A(\theta_0) \) be the acceptance region of some level-\( \alpha \) test for \( H_0: \theta = \theta_0 \) versus \( H_1: \theta \neq \theta_0 \), for every \( \theta_0 \in \Theta \). Then \( C = \{ \theta \in \Theta \mid x \in A(\theta) \} \) is a level-\( \alpha \) confidence region, i.e.

\[
P_\theta(C \ni \theta) = 1 - \alpha \quad \text{for every } \theta \in \Theta \tag{see Section 3.5 in Lehmann, 1959}.
\]

Next, consider a parameter of interest \( \omega = \omega(\theta) \), that is, a real-valued function of \( \theta \). For every \( \omega_0 \) let \( A(\omega_0) \) denote the acceptance region of some level-\( \alpha \) test of the hypothesis \( H_0: \omega = \omega_0 \) versus \( H_1: \omega > \omega_0 \). Furthermore, define \( \omega_-(x) = \inf \{ \omega \mid x \in A(\omega) \} \).

Then

\[
P_\theta(\omega_-(x) \leq \omega(\theta)) \geq 1 - \alpha \quad \text{for every } \theta \in \Theta.
\]

Proof: For every \( \theta \in \Theta \) it holds that

\[
\omega_-(x) \leq \omega(\theta) \Leftrightarrow \exists \omega \leq \omega(\theta) \Leftrightarrow x \in A(\omega) \Leftrightarrow x \in \bigcup_{\omega \leq \omega(\theta)} A(\omega).
\]

Hence,

\[
P_\theta(\omega_-(x) \leq \omega(\theta)) = P_\theta(x \in \bigcup_{\omega \leq \omega(\theta)} A(\omega)) \geq P_\theta(x \in A(\omega(\theta))) = 1 - \alpha.
\]

Our present case corresponds to the special case where the parameters are \( \theta = \nu = (v_1, v_2, \ldots) \), \( \omega = \min \{ m \mid \sum_{k=0}^{m} v_k = 0 \} \).

Then \( \omega = \xi + 1 \), and in view of Definition (17) the proof is complete.

### Appendix C. Simulating populations of point processes with identical correlation structures

The non-Poissonian spike trains have been simulated with a warm-up time of \( m = 100 \). The parameters of the source Poisson processes have been chosen such that for all target populations the pairwise coincidence probability \( p_r \) was equal to 0.5 and \( \xi_{\text{peak}} \) of the binomial approximation of the correlation structure was equal to 3. That is, we chose \( \xi_{\text{peak}} = 15 \) and \( R_\text{m} = 5 \cdot \lambda \) where \( \lambda = 6, 2, 1 \) for the populations of size \( N = 15, 45, 90 \).

Note that this value for \( R_m \) lies below the maximum of the hazard function of gamma processes with \( \lambda = 6 \) only for \( CV < 0.4472 \) and hence more regular processes could not be generated.

In doing so, the resulting correlation structures are quite similar to each other. More precisely, with \( I_1(f_{1}(k), f_2(k)) = \sum_i f_1(k) - f_2(k) \) measuring the distance between two amplitude distributions, the correlation structures of processes with the same \( N \) and inter-spike interval distribution differed on average between \( I_1 = 0.0852 \) and \( I_1 = 0.0662 \). Thereby, the value did not vary much with the CV. The mean distance across all correlation structures was \( I_1 = 0.0666 \). In order to obtain comparable Poisson processes we determined the true correlation structure for the non-Poissonian spike trains in each of the 100 trials and generated Poisson process with the exact same number of spikes and higher-order correlations.

### Appendix D. Fitting the estimated amplitude distribution

In Section 3.3 the same parameter setting has been used with a simulation time of 200 s. Let \( f_{\text{est}}(k) = \hat{v}_k / \sum_k \hat{v}_k \) denote the amplitude distribution with correlation structure \( (\hat{v}_1, \hat{v}_2, \ldots) \) estimated via EDP. To this we fitted an amplitude distribution with two peaks

\[
f_{\text{bimodal}}(k) = \eta \cdot \frac{1}{R_m} \cdot \delta_{1,k} + \sum_{k=1}^{N} \eta_0 (1 - \eta) \delta(k) \delta_{\text{sym}} \cdot \frac{1}{R_m} \tag{D.1}
\]

More precisely, we solved numerically:

\[
(R_m, \omega_1, \omega_2, \xi_{\text{sym}}) = \arg \min \sum_k (f_{\text{bimodal}}(k) - f_{\text{est}}(k)) \tag{D.2}
\]

respecting the following settings and constraints.
where \( \lambda_i \) and \( C_i \) denote the firing rate and coefficient of variation of the \( i \)-th neuron, respectively, and \( \phi_i, \chi(x) \) the hazard function of a lognormal process with these averaged spike statistics. Property iii) is a combination of Eq. (5) and \( f_{\text{min}} \). The second inequality in iv) follows from \( \eta > 0 \).

References


