

# Distribution of interspike intervals estimated from multiple spike trains observed in a short time window

Zbyněk Pawlas\*

Department of Probability and Mathematical Statistics, Faculty of Mathematics and Physics, Charles University,  
Sokolovská 83, CZ-186 75 Prague 8, Czech Republic

Petr Lansky†

Institute of Physiology, Academy of Sciences of the Czech Republic, Vídeňská 1083, CZ-142 20 Prague 4, Czech Republic  
(Received 26 July 2010; published 18 January 2011)

Several nonparametric estimators of the probability distribution of interspike intervals are introduced. The methods are suitable for simultaneous spike trains observed in a time window of length comparable with the mean interspike interval. This reflects the situation in which a high number of input spike trains converge to a single cortical neuron that has to react in a relatively short time. The simulation study is performed to compare the estimators. For that purpose, several types of stationary point processes are considered as the models of neuronal activity. The methods permit one to estimate the distribution of interspike intervals even if practically none of them are observed. The Kaplan-Meier estimator seems to be the most flexible and reliable among all studied methods, but no direct conclusions as to how real neurons work can be deduced from it.

DOI: [10.1103/PhysRevE.83.011910](https://doi.org/10.1103/PhysRevE.83.011910)

PACS number(s): 87.10.Mn, 02.50.-r, 87.19.1l

## I. INTRODUCTION

Neurons generate action potentials (also called spikes or firings) in the process of information transmission. The statistical properties of the spikes, as they appear in time, have been intensively studied for more than 50 years using experimental as well as theoretical approaches. The reason is an endless effort to understand the code used in the neuronal information transmission. Many candidates of the code have been proposed, and nowadays the widely accepted concept is that there is not a unique code but many of them in coexistence [1–3]. The most common and most investigated one is the rate code [4]. It presumes that the information is encrypted in the frequency of the spikes. All the remaining codes are classified as the temporal codes [1,2,5–8] despite that this classification is rather vague. On one hand, there are codes based on exact timing of the spikes (e.g., latency or coincidence detectors), and on the other hand there are those that employ statistical features of spiking (e.g., variability or randomness), as it seems to be a certain fact that stochastic principles are behind the neuronal firing.

The firing rate is uniquely related (inversely proportional) to the mean interspike interval (ISI). Therefore, the next candidate for the rate code is that based on the variability of ISIs. It can be documented by the effort to characterize the neural data by the coefficient of variation (CV, the standard deviation divided by mean) of ISIs in the experimental as well as the modeling studies [9–12]. However, despite that the mean and variance tell us most of the information about a random variable, there are still some additional properties for which the complete probability distribution is required, for example, the higher moments of the distribution or its entropy [13,14]. The importance of the ISI probability distribution is thoroughly discussed and experimentally illustrated in Ref. [15]. Higher moments have been commonly calculated in

a neuronal context, and histogramming of ISIs to estimate their distribution has been the basic tool since when the ISIs started to be recorded [16–18]. To evaluate all these quantities (CV, higher moments, entropy, ISI distribution) usually requires long, single-trial data.

We have to stress that the character of information obtained in experiments in which the spikes are recorded, usually from a single neuron or from only a few ones, is fundamentally different from that which a neuron receives from the network of interconnected neurons. In experiments, a spike train is recorded for a relatively long period of time, and the properties of the firing are deduced. If the type of the investigated firing is transient, like in the stimulated activity, then the length of the record is replaced by repetitions assuming that these are statistically identical copies of the same firing behavior. In *in vivo* conditions, the neuron receives a large number of spike trains, up to several thousand, and the received information has to be extracted in short time intervals. This creates a discrepancy between what can be read from the experiments and how real neurons perform. The situation is well described in Ref. [3] from the rate code point of view. Three situations are distinguished there:

- (1) Rate as a spike count (average over time)
- (2) Rate as a spike density (average over several trials)
- (3) Rate as a population activity (average over several neurons)

and it is shown that the rate code can be alternatively investigated by all three strategies. However, we are not aware of any systematic comparison of these approaches for higher statistical characteristics. Masuda and Aihara in Ref. [19] investigated when trial averaging can be used as a representative of the network activity. The authors called the consistency between trial averaging and population averaging a *physiological ergodicity*. Nawrot *et al.* in Ref. [20] considered the relationship between single neuron spike train variability (measured by the CV) and variability across repeated trials (measured by the Fano factor of spike counts). In our recent paper [21] we showed how to treat the variability coding (CV estimation) under the population scenario.

\*pawlas@karlin.mff.cuni.cz

†lansky@biomed.cas.cz

The aim of the present paper is to compare several nonparametric estimators of the cumulative distribution function (cdf) of ISIs in the population activity. As mentioned, the estimation of the ISI distribution (density) is one of the most common tools in the statistical treatment of ISI data. However, we prefer to estimate the cdf rather than the probability density function (pdf). Both cdf and pdf carry the same information about the distribution. The histograms are perhaps so popular because it is easier to compare them visually with different pdfs than using the empirical cumulative distribution function (ecdf) for that purpose. However, the disadvantage of the histogram is the requirement of fixing the bins or the kernel properties. This results in the loss of some information that is not lost in the ecdf. Moreover, in our setting the estimation of the cdf seems to be more straightforward.

Most of the statistical procedures are based on assumptions idealizing the reality. The main simplifying assumption of the population rate code is the independence of spike trains generated by different neurons. Although this is hardly fully satisfied, it is generally accepted (see Ref. [5]), and we also have to take it into account. On the other hand, the population activity may vary rapidly and reflect changes in the external conditions nearly instantaneously [22,23]. Therefore, this population of neurons is observed for a relatively short period with length of the order of magnitude of the mean ISI, and consequently the assumption of stationarity is not restrictive.

The scenario with relatively short time observation windows was also considered in Ref. [24], where the effect on the time-rescaling method was investigated. The author speaks about short-trial spike train data. In such a situation, it is possible to derive a truncated version of the true ISI distribution, and Nawrot *et al.* in Ref. [20] propose its estimator. The main difficulty of this setting is the presence of length bias in the estimators because no ISIs are longer than the length of the window. Thus, we cannot construct the estimates for values exceeding the length of the window without further specific assumptions for the model. We have to deal with the censored data problem, and some solutions are offered in this paper.

Currently, it is technically impossible to record simultaneously all the incoming spike trains to a target neuron. Therefore, any illustration of the presented methods with experimental data is not possible, and only simulations are available. We investigated several parametric models for ISI generation, the Poisson process, renewal process, and mixed Poisson process and verified the methods using them. The Poisson process is usually considered as too simple of a characterization of neuronal activity, and thus several extensions and generalizations are used. One of them assumes that the ISIs are generated by a renewal process [25] with a specific distribution of ISIs, for example, a gamma distribution [26–30] or an inverse Gaussian distribution [31–41]. These distributions reproduce features of real neurons (nonconstant failure rate, refractoriness) not reflected by the Poisson process. Extension in another direction retains the complete random character, but the homogeneity of the firing rate is not satisfied. This assumption leads to a doubly stochastic Poisson process (Cox process), which was studied theoretically in Refs. [42,43], and in the context of computational neuroscience it was used, for example, in Refs. [44–46].

## II. METHODS AND RESULTS

### A. Models

We define a point process  $N = \{X_i, i \in \mathbb{Z}\}$  as a random locally finite sequence of points (spikes) on the real line  $\mathbb{R}$ , assuming  $\dots < X_{-1} < X_0 \leq 0 < X_1 < X_2 < \dots$ ; i.e., no multiple points are allowed (only simple point processes are considered). The ISIs are denoted by  $T_i = X_{i+1} - X_i, i \in \mathbb{Z}$ , and  $N(t)$  is used for the number of spikes in the time interval  $[0, t]$ . Only stationary point processes are investigated; i.e., the distribution of  $N$  is always invariant under translation, and the mean number of points in a time interval is proportional to its length, i.e.,  $\mathbb{E}[N(t) - N(s)] = \lambda(t - s)$ ,  $s < t$ . The constant  $\lambda > 0$  is called the intensity of the point process  $N$ . The stationarity means that the time origin,  $t = 0$ , is unrelated to the ongoing activity. We would like to emphasize that there is almost surely no event at time 0. More details on the theory of point processes can be found, e.g., in Refs. [42,47,48].

Our aim is to find an estimator of the ISI distribution. For stationary renewal processes, ISIs are identically distributed, and we are interested in their common distribution. In the nonrenewal case we have to clarify what is meant by the ISI distribution. For this reason, it is convenient to introduce a marked point process  $N_m = \{(X_i, T_i)\}$ . The marks  $T_i$  depend on the point process  $N = \{X_i\}$ . In fact, we are in the situation when a point process is controlled by random marks (see Ref. [47], p. 133). It follows from stationarity that there exists a probability distribution  $Q$  (called a mark distribution) such that the mean number of points in the time interval  $(s, t]$  that have marks in a Borel set  $B$  is  $\lambda(t - s)Q(B)$ . In this context, our main interest is obtaining an estimate of the mark distribution  $Q$  or its corresponding cdf  $F(t) = Q((-\infty, t])$ . For a random variable  $T$  with cdf  $F$ , let  $\mu = \mathbb{E}(T)$  and  $\gamma = \text{CV}(T) = \sqrt{\text{Var}(T)}/\mathbb{E}(T)$  denote its mean and CV, respectively.

The methods introduced here are nonparametric and thus independent of the model. However, to check their performance, a model on which they are tested has to be selected. For such a purpose we choose the three following models: the Poisson process, renewal process, and mixed Poisson process. The first two correspond to the case of homogenous neuronal population, whereas the last one is a model of the nonhomogenous neuronal population. In all three cases it is clear how the cdf  $F(t)$  is defined.

#### 1. Homogenous neuronal population

**Poisson point process:** This is the simplest point process. The events occur completely at random with intensity  $\lambda > 0$ , and the ISIs  $T_i, i \in \mathbb{Z}$ , are independent copies of a positive random variable  $T$  that has an exponential distribution with mean  $1/\lambda$ ; i.e., its cdf is  $F(t) = \mathbb{P}(T \leq t) = 1 - e^{-\lambda t}, t \geq 0$ . For the Poisson process  $\mu = 1/\lambda$  and  $\gamma = 1$ .

The Poisson point process can be generalized in several directions. One of them leads to renewal processes that form an important class of point processes. The ISIs  $T_i$  are independent copies of a positive random variable  $T$  with cdf  $F(t)$ . In contrast to the Poisson process,  $T$  is not necessarily exponentially distributed, and the interval to the first spike,  $X_1$ , has a different distribution from  $T$ . It is well known that  $X_1$  (the so-called forward recurrence time) has a pdf

$f_1(t) = [1 - F(t)]/\mathbb{E}(T)$  (see Ref. [25]). We use two specific distributions for ISIs in this paper.

Renewal process with ISIs following a  $\Gamma(a, b)$  distribution: A gamma distribution forms a flexible and relatively simple class of distributions that is frequently used in neural modeling [26–30]. It has two parameters: a shape parameter  $a > 0$  and rate parameter  $b > 0$ . The pdf is

$$f(t) = \frac{b^a}{\Gamma(a)} e^{-bt} t^{a-1}, \quad t > 0,$$

the mean ISI is  $\mu = a/b$ , and the CV is  $\gamma = 1/\sqrt{a}$ .

Renewal process with ISIs having an inverse Gaussian distribution  $IG(m, 1/s)$ : The pdf is

$$f(t) = \frac{1}{\sqrt{2\pi s t^3}} \exp\left[-\frac{(t-m)^2}{2m^2 s t}\right], \quad t > 0,$$

the mean ISI is  $\mu = m$ , and the CV is  $\gamma = \sqrt{ms}$ . The popularity of inverse Gaussian distribution follows from the fact that it is related to the simplest neuronal model: the perfect integrate-and-fire neuron [31–41].

## 2. Nonhomogenous neuronal population

A different generalization of the Poisson process is a doubly stochastic Poisson process. The simplest type of the doubly stochastic Poisson process is obtained by considering a random variable  $\Lambda$  (called the driving intensity) such that conditionally on  $\Lambda = \lambda$ ,  $N$  is a Poisson process with intensity  $\lambda$ . Such a point process is referred to as a mixed Poisson process. The random driving intensity  $\Lambda$  cannot be observed. From a single realization, the mixed Poisson process cannot be distinguished from the Poisson process. For a population of neurons, this model means that the spike train of its  $k$ th neuron is the Poisson process with intensity  $\lambda^{(k)}$ , which is a realization of  $\Lambda$ . Then the ISIs,  $T^{(k)} = T \mid \lambda^{(k)}$ , conditionally in the  $k$ th neuron have an exponential distribution with parameter  $\lambda^{(k)}$ . The unconditional ISI for the whole population has the cdf

$$\begin{aligned} F(t) &= \mathbb{E}[\mathbb{P}(T \leq t \mid \Lambda)] \\ &= \mathbb{E}(1 - e^{-\Lambda t}) = 1 - \mathbb{E}(e^{-\Lambda t}), \quad t \geq 0. \end{aligned} \quad (1)$$

It is the mean cdf with respect to the distribution of the driving intensities.

Mixed Poisson process with driving intensity  $\Lambda$  distributed according to the gamma distribution  $\Gamma(a, b)$ : The cdf is

$$F(t) = 1 - \left(\frac{b}{b+t}\right)^a$$

(a generalized Pareto distribution), the mean ISI is  $\mu = b/(a-1)$ , and the CV is  $\gamma = \sqrt{a/(a-2)}$ . We selected gamma for the distribution of the firing rates because it resembles experimentally obtained histograms of spike counts recorded from different types of neurons under different conditions, and, in addition, this choice results in a relatively simple form of  $F(t)$ .

The mixed Poisson process is a completely different type of model compared to the previous ones. It is a step toward modeling a neural population. The first three models consider the neural population as totally homogeneous, and spike trains are statistical copies of the same law. Here, despite that the individual spike trains are simple, there is an additional

assumption about the distribution of the spike rates  $\Lambda$  in the population.

## B. Estimation from a single neuron

We assume that the spikes form a realization of a stationary point process  $N$  observed in a time window  $[0, \Delta]$  of length  $\Delta > 0$ . The task corresponds to the situation when a single spike train is observed for a sufficiently long time to deduce the ISI distribution. The case of a single spike train is presented mainly as preparation for the situation when several spike trains are observed for a relatively short time (see the next subsection). We suppose that  $N(\Delta) > 0$ ; otherwise no reasonable estimator of cdf can be constructed.

A natural candidate for estimating  $F(t)$  is an *empirical cumulative distribution function*

$$\hat{F}^E(t) = \frac{1}{N(\Delta) - 1} \sum_{i=1}^{N(\Delta)-1} H(t - T_i) \quad (2)$$

for  $t \leq \Delta$ , provided that  $N(\Delta) \geq 2$ . Here  $H$  is the Heaviside step function; i.e.,  $H(x) = 1$  for  $x \geq 0$  and  $H(x) = 0$  for  $x < 0$ . However, not all information contained in the data is used in estimator (2), because the interval  $[X_{N(\Delta)}, \Delta]$  is not taken into account, although it gives additional information about  $T_{N(\Delta)} = X_{N(\Delta)+1} - X_{N(\Delta)}$ . We know that  $T_{N(\Delta)} > \Delta - X_{N(\Delta)} = B_\Delta$  (the backward recurrence time). Therefore, we define the modified estimator (if  $N(\Delta) \geq 2$ ) as

$$\hat{F}^{EM}(t) = \begin{cases} \frac{N(\Delta)-1}{N(\Delta)} \hat{F}^E(t), & t \leq B_\Delta, \\ \hat{F}^E(t), & t > B_\Delta \end{cases} \quad (3)$$

(see also Ref. [48], p. 313). If  $N(\Delta) = 1$ , we set  $\hat{F}^{EM}(t) = 0$  for  $t \leq B_\Delta$  and  $\hat{F}^{EM}(t) = 1$  otherwise. Alternatively to  $B_\Delta$ , we can consider time  $X_1$  to the first spike (the forward recurrence time). This also gives us additional information about  $T_0 = X_1 - X_0$  since we know that  $T_0 > X_1$ . We have to stress that it is not appropriate to use both backward and forward recurrence times simultaneously because it would introduce distinctive bias into the estimation due to the strong dependence of these variables. The knowledge of  $X_1$  and ISIs  $T_i$ ,  $i = 1, \dots, N(\Delta) - 1$ , uniquely determines  $B_\Delta = \Delta - (X_1 + \sum_{i=1}^{N(\Delta)-1} T_i)$ , and, conversely,  $X_1$  can be written as the function of  $B_\Delta$  and  $\{T_i\}$ . As also noted in Ref. [48], the shortcoming of estimators (2) and (3) is the length bias. Shorter intervals have a greater chance to be taken into account, and the intervals larger than  $\Delta$  are not observed at all.

The next possibility for estimation of  $F(t)$  is to use the so-called *reduced sample estimate*. In order not to have the problem with a bounded window we restrict our attention to the points in the interval  $[0, \Delta - t]$ . We define

$$\hat{F}^R(t) = \frac{\#\{T_i \leq t \mid X_i \in [0, \Delta - t]\}}{N(\Delta - t)} \quad (4)$$

for  $t \leq \Delta - X_1$ . Estimators of this type are often used in spatial statistics to eliminate edge effects caused by the bounded observation window [49]. Estimator (4) is not necessarily

monotone. Therefore, in order to get a monotone function, it can be modified:

$$\hat{F}^{\text{RM}}(t) = \sup_{s \leq t} \hat{F}^{\text{R}}(s), \quad t \geq 0. \quad (5)$$

Finally, the estimation can also be viewed as a censored data problem. Among ISIs  $T_1, \dots, T_{N(\Delta)-1}$  there are  $M$  distinct values that can be sorted, yielding  $T_{(1)} < T_{(2)} < \dots < T_{(M)}$ . In analogy with the classical theory of survival analysis, we propose the *Kaplan-Meier estimator* [if  $N(\Delta) \geq 2$ ]:

$$\hat{F}^{\text{KM}}(t) = 1 - \prod_{i: T_{(i)} \leq t} \left(1 - \frac{D_i}{S_i}\right), \quad (6)$$

where  $D_i = \#\{j = 1, \dots, N(\Delta) - 1 | T_j = T_{(i)}\}$  is the number of observed ISIs equal to  $T_{(i)}$ , and  $S_i = \sum_{j=1}^{N(\Delta)-1} H(T_j - T_{(i)}) + H(B_\Delta - T_{(i)})$  is the number of ISIs including a backward recurrence time greater or equal to  $T_{(i)}$ . If observed ISIs are all distinct, then  $M = N(\Delta) - 1$  and  $D_i = 1$  for all  $i$ . Since the censoring variables are not independent and depend on the survival times, the optimality of the Kaplan-Meier estimator is not preserved, but we may hope that it can compete with estimator (3). The Kaplan-Meier estimator is nondecreasing and piecewise constant but may not reach 1 (it is always less or equal to 1). In contrast to estimator (3), it does not have a jump at  $B_\Delta$ .

### C. Estimation from multiple neurons

We consider the situation when  $n$  independent copies  $N_1, \dots, N_n$  of a stationary point process  $N$  are observed over the time interval  $[0, \Delta]$ . The points of  $N_k$  are denoted by  $X_i^{(k)}$  and the ISIs by  $T_i^{(k)}$ . An example is shown in Fig. 1, where 50 independent realizations of Poisson point process are shown. The time window of observation is equal to the mean ISI.

The *first* approach to the estimation from  $n$  parallel trains is to take the average of the separate estimators for each of the spike trains. In particular, we use

$$\bar{F}_n^{\text{EM}}(t) = \frac{1}{n} \sum_{k=1}^n \hat{F}_{N_k}^{\text{EM}}(t), \quad (7)$$

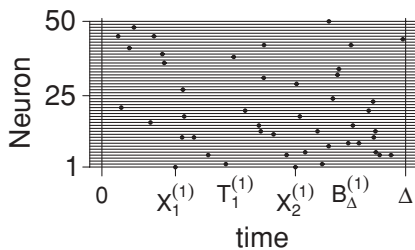


FIG. 1. Schematic illustration of the problem: 50 independent Poisson spike trains are observed in a time window of length  $\Delta = 1$ . Mean ISI is  $\mu = 1$ . The maximum number of spikes in one train is four (neuron no. 5), for 20 neurons there are no observed spikes in the interval  $[0, \Delta]$ , and for 18 neurons only one spike is observed. The firing times for  $k$ th neuron are denoted by  $X_i^{(k)}$ , and the first completely observed ISI is  $T_1^{(k)}$ . The time elapsed from the last observed spike to  $\Delta$  is denoted as  $B_\Delta^{(k)}$ .

where  $N_k$  means that estimator (3) is applied to the process  $N_k$ . Analogously, we also consider an estimator based on (2) and given by the formula

$$\bar{F}_n^{\text{E}}(t) = \frac{1}{n} \sum_{k=1}^n \hat{F}_{N_k}^{\text{E}}(t), \quad (8)$$

corresponding to the estimator given by (2).

The *second* possibility is to pool the information from all parallel observations together. The pooled reduced sample estimator and Kaplan-Meier estimator are obtained by analogues of (4) and (6) in which the numerators and denominators are replaced by the sums over all  $N_k$ ,

$$\hat{F}_n^{\text{R}}(t) = \frac{\sum_{k=1}^n \#\{T_i^{(k)} \leq t | X_i^{(k)} \in [0, \Delta - t]\}}{\sum_{k=1}^n N_k(\Delta - t)}$$

for  $t \leq \Delta - \min_{k=1, \dots, n} X_1^{(k)}$  and

$$\hat{F}_n^{\text{KM}}(t) = 1 - \prod_{T_{(i)} \leq t} \left(1 - \frac{\sum_{k=1}^n D_i^{(k)}}{\sum_{k=1}^n S_i^{(k)}}\right), \quad (9)$$

where  $T_{(1)} < T_{(2)} < \dots < T_{(M)}$  are the distinct ISIs from all replicated observations,  $D_i^{(k)} = \#\{j = 1, \dots, N_k(\Delta) - 1 | T_j^{(k)} = T_{(i)}\}$ , and  $S_i^{(k)} = \sum_{j=1}^{N_k(\Delta)-1} H(T_j^{(k)} - T_{(i)}) + H(B_\Delta^{(k)} - T_{(i)})$ . Again, we also consider a monotone version of the reduced sample estimator:

$$\hat{F}_n^{\text{RM}}(t) = \sup_{s \leq t} \hat{F}_n^{\text{R}}(s). \quad (10)$$

Despite the fact that there is no suitable pooled counterpart for the modified ecdf, it is possible to define the pooled version of the classical ecdf. Figure 2 shows the estimates of cdf of ISIs from the data presented in Fig. 1. We can see that neglecting the backward recurrence time deteriorates the estimate dramatically.

For a mixed Poisson process, it seems natural, as follows from (1), to estimate the distribution of spike rate  $\Lambda$  and insert it into the formula for  $F(t)$ . However, such an approach introduces bias into the estimation. Actually, we do not need to know the distribution of  $\Lambda$  in order to estimate  $\mathbb{E}(e^{-\Lambda t})$  directly. The quantity  $\mathbb{E}(e^{-\Lambda t})$  is closely related to the Laplace functional; for details about Laplace functionals of point

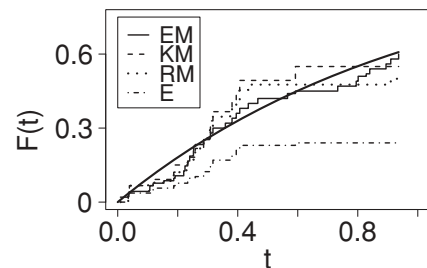


FIG. 2. The estimates of  $F(t)$  for the simulated data shown in Fig. 1. The estimators given by (7), (9), and (10) are compared with a true cdf (bold line). Moreover, we also show estimator (8), which does not use information given by backward recurrence times.



processes, see Refs. [42,48]. The Laplace functional of a point process  $N = \{X_i, i \in \mathbb{Z}\}$  is defined as

$$L_N(g) = \mathbb{E} \left\{ \exp \left[ - \sum_i g(X_i) \right] \right\},$$

where  $g$  is a nonnegative measurable function. The Laplace functional for a stationary Poisson process with intensity  $\lambda$  is

$$L_N(g) = \exp \left[ -\lambda \int_{-\infty}^{\infty} (1 - e^{-g(x)}) dx \right].$$

For a mixed Poisson process we have

$$L_N(g) = \mathbb{E} \left\{ \exp \left[ -\Lambda \int (1 - e^{-g(x)}) dx \right] \right\}.$$

In particular, for  $g(x) = -\log(1 - t/\Delta)$ ,  $x < \Delta$ ,

$$L_N(g) = \mathbb{E}(e^{-\Lambda t}).$$

Then the distribution function of ISI for the mixed Poisson process has the form

$$F(t) = 1 - \mathbb{E}(e^{-\Lambda t}) = 1 - L_N(g),$$

and we may estimate  $L_N(g)$  by using the empirical Laplace functional

$$\widehat{L_N}(g) = \frac{1}{n} \sum_{k=1}^n \exp \left[ - \sum_{i=1}^{N_k(\Delta)} g(X_i^{(k)}) \right].$$

It is the maximum likelihood estimator of  $L_N(g)$ ; see Ref. [48]. Therefore, we define

$$\begin{aligned} \hat{F}_n^{\text{MP}}(t) &= 1 - \frac{1}{n} \sum_{k=1}^n \exp \left[ \sum_{i=1}^{N_k(\Delta)} \log \left( 1 - \frac{t}{\Delta} \right) \right] \\ &= 1 - \frac{1}{n} \sum_{k=1}^n \left( 1 - \frac{t}{\Delta} \right)^{N_k(\Delta)}. \end{aligned} \quad (11)$$

Since  $\mathbb{E}[\widehat{L_N}(g)] = L_N(g)$ , the estimator  $\hat{F}_n^{\text{MP}}(t)$  is unbiased. Note that in (11) we do not need the information about firing times, but the cdf of ISIs is estimated just from observed spike counts.

#### D. Estimation of tails

All estimators proposed in the previous section are defined only for  $t \leq \Delta$ , and now the aim is to extend them for  $t > \Delta$ . Since no information about the ISIs longer than  $\Delta$  is available, we estimate  $F(t)$  on the interval  $(\Delta, \infty)$  by extrapolating with the exponential tail. The exponential distribution is chosen because it resembles commonly used distributions of ISIs.

The mean ISI and the cdf of ISI are related by the identity

$$\mathbb{E}(T) = \int_0^{\infty} [1 - F(t)] dt. \quad (12)$$

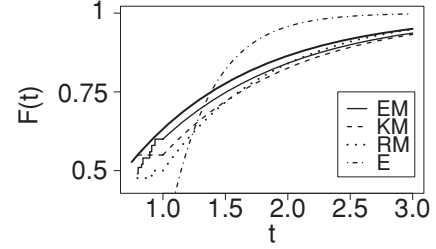


FIG. 3. The estimators  $\hat{F}_n^{\text{EM}}$ ,  $\hat{F}_n^{\text{KM}}$ ,  $\hat{F}_n^{\text{RM}}$ , and  $\hat{F}_n^{\text{E}}$ , based on the data shown in Fig. 1, extended to the interval  $[\Delta, \infty)$  using Eq. (15). The true cdf is depicted by the bold line.

The rate is reciprocal to  $\mathbb{E}(T)$  for renewal processes; thus we can estimate the mean ISI from the data by

$$\widehat{\mathbb{E}(T)} = \frac{n\Delta}{\sum_{i=1}^n N_i(\Delta)}.$$

We construct our estimator of the cdf so that the relation (12) holds also for the corresponding estimators:

$$\widehat{\mathbb{E}(T)} = \int_0^{\infty} [1 - \hat{F}_n(t)] dt. \quad (13)$$

If  $T$  has an exponential distribution with intensity  $\lambda$ , then

$$\begin{aligned} \mathbb{P}(T \leq t \mid T > \Delta) &= \frac{F(t) - F(\Delta)}{1 - F(\Delta)} \\ &= 1 - e^{-\lambda(t-\Delta)}, \quad t > \Delta; \end{aligned} \quad (14)$$

i.e., the distribution of  $T - \Delta$  conditional on  $T > \Delta$  is also exponential with intensity  $\lambda$  (memoryless property of exponential distribution). Using (14), we extend an estimate  $\hat{F}_n(t)$  defined for  $t \in [0, \Delta]$  to the interval  $(\Delta, \infty)$  as

$$\hat{F}_n(t) = 1 - [1 - \hat{F}_n(\Delta)]e^{-\hat{\lambda}(t-\Delta)}, \quad t > \Delta, \quad (15)$$

where  $\hat{\lambda}$  is the estimator of  $\lambda$ , determined so that condition (13) holds true. If  $\widehat{\mathbb{E}(T)} < \int_0^{\Delta} [1 - \hat{F}_n(t)] dt$ , condition (13) does not hold true, and therefore we define  $\hat{F}_n(t) = 1$  for  $t > \Delta$ . It is demonstrated in Fig. 3 how the method works for the situation with the simulated data given in Fig. 1.

### III. NUMERICAL RESULTS

The integrated square error is commonly used to judge the estimation efficiency. We normalize it to eliminate the influence of the limited observation window, and for the comparison of the estimators we use the *relative integrated square error* (RISE). It is defined on the interval  $(0, \delta)$ ,  $0 < \delta \leq \infty$ , as the integrated square error divided by  $F(\delta)^2$ :

$$R(\delta) = \frac{1}{F(\delta)^2} \int_0^{\delta} [\hat{F}_n(t) - F(t)]^2 dt.$$

Another possible way to measure the distance between the estimated and theoretical cdf is the maximum distance,  $\sup_{t \geq 0} |\hat{F}_n(t) - F(t)|$ , which is used in the Kolmogorov-Smirnov goodness-of-fit test. Since this quantity considers the maximal deviation only, it does not appear to be as good as those distances involving integration over the whole interval.

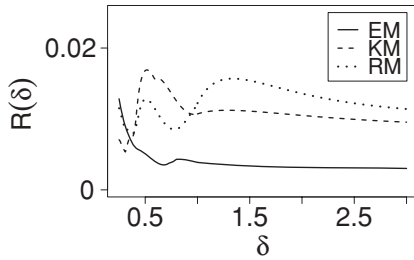


FIG. 4. The errors  $R(\delta)$  as the function of  $\delta$  for the estimators (7), (9), and (10), illustrated in Figs. 2 and 3 and evaluated from the data shown in Fig. 1. The values of  $R(\infty)$  are 0.0029, 0.0088, and 0.0103, respectively. We do not show  $R(\delta)$  for the ecdf given by (8) because the values are one order higher,  $R(\infty) = 0.0723$ .

The behavior of  $R(\delta)$  in dependency on  $\delta$  is demonstrated in Fig. 4 for the case of our simulated data ( $n = 50$ , Fig. 1). The errors  $R(\delta)$  are computed for the estimators (7), (9), and (10); see Figs. 2 and 3. Because of the fact that the data are from exponential distribution and the tail is estimated using the exponential distribution, the error remains practically constant for  $\delta$  larger than  $\Delta = 1$ .

For an illustration of how the estimates depend on the length of observation in the single-trial experiment, we consider a spike train observed in a window of length  $\Delta = 1$ . To see the reliability of the methods, we simulated 500 independent Poisson processes with intensity  $\lambda$ . Average RISEs (over 500 repetitions) of the estimators (3), (5), and (6) are shown in Fig. 5. Starting from 15 mean ISIs within the observation period ( $\lambda = 15$ ), there are no visible differences among the estimators. Obviously, below 10 mean ISIs the estimation errors dramatically increase.

In the remainder of this paper, we focus on the estimation from multiple spike trains and evaluate the estimators introduced in Sec. II C. The number of spike trains is always  $n = 400$ . For the simulation study we choose five different sets of parameters so that  $\mu$  is 0.25, 0.5, 1, 2, or 3. This choice determines how many complete ISIs fit the sampling window on average. The CV is  $\gamma = 1$  for the Poisson process. The value of  $\gamma$  has the substantial influence on the shape of the ISI densities. Thus, we choose  $\gamma = 0.5$  and  $\gamma = 1.5$  for renewal and non-Poisson processes. Note that the mixed Poisson process is a model for overdispersion, and therefore  $\gamma \geq 1$  in this case. We choose  $\gamma = 1.5$  for the mixed Poisson process.

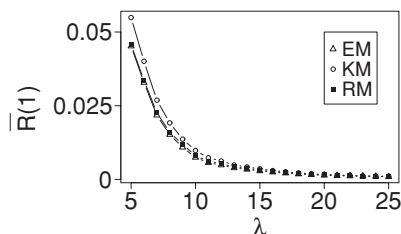


FIG. 5. The average  $R(1)$  (over 500 simulations) for the estimation of the cdf from a single time window of length  $\Delta = 1$  in the case of the Poisson process with intensity  $\lambda$ . Three different estimators are compared:  $\hat{F}_n^{\text{EM}}$  (triangles),  $\hat{F}_n^{\text{KM}}$  (circles), and  $\hat{F}_n^{\text{RM}}$  (squares). The errors decrease with increasing intensity  $\lambda$ .

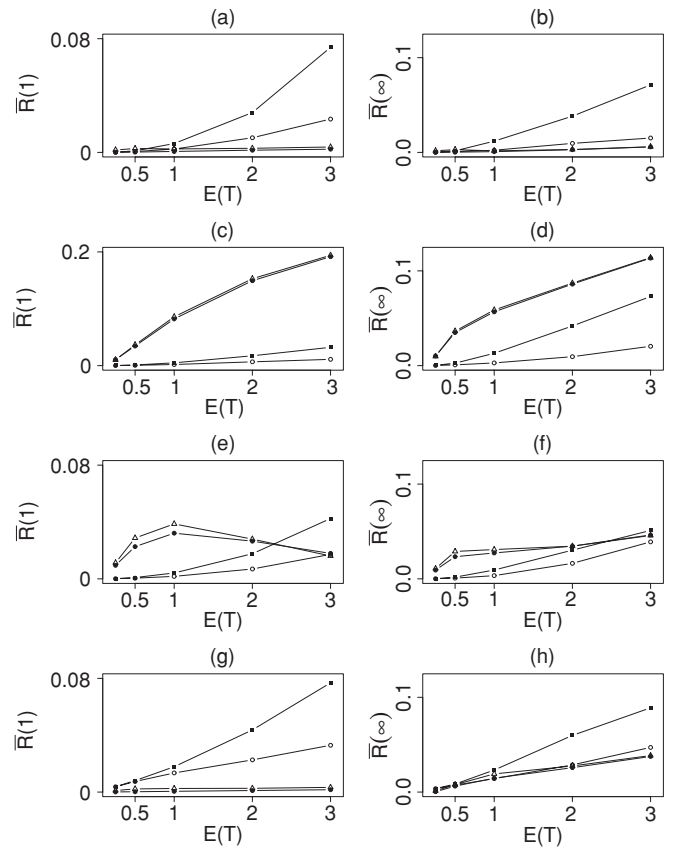


FIG. 6. The average RISEs  $R(1)$  (left column) and  $R(\infty)$  (right column) for four different estimators of the cdf computed from 500 realizations of 400 point processes observed in a window of length 1. Five different values of  $\mu$  are considered: 0.25, 0.5, 1, 2, and 3. For each  $\mu$  the average error is shown for the following estimators:  $\hat{F}_n^{\text{EM}}$  (triangles),  $\hat{F}_n^{\text{KM}}$  (circles),  $\hat{F}_n^{\text{RM}}$  (squares), and  $\hat{F}_n^{\text{MP}}$  (bullets). The results are depicted for (a)–(b) the Poisson process, (c)–(d) the renewal process with gamma distributed ISIs and  $\gamma = 1.5$ , (e)–(f) the renewal process with inverse Gaussian distributed ISIs and  $\gamma = 1.5$ , and (g)–(h) the mixed Poisson processes with  $\gamma = 1.5$ .

We have generated  $n = 400$  independent realizations of the selected point processes in the time window  $[0, 1]$ . From simulated data we estimate  $F(t)$  and compute RISEs with  $\delta = 1$  and  $\delta = \infty$ . This procedure is repeated 500 times for each choice of model parameters. The sample means of the computed errors are shown in Fig. 6 for the Poisson process ( $\gamma = 1$ ) and three different models with  $\gamma = 1.5$ . In Fig. 7 there are analogous results for two different models with  $\gamma = 0.5$ . The simulations and computations were carried out in R [50].

Independently of the estimation method, the errors increase with increasing mean ISI (causing fewer observations in the window). This result has been expected. The only exception can be seen in Fig. 6(e). This counterintuitive effect appears due to a larger bias of estimators  $\hat{F}_n^{\text{EM}}$  and  $\hat{F}_n^{\text{MP}}$  in the case of renewal processes. Furthermore, we can see that the errors are larger for smaller values of coefficient of variation. This seems to be a more complex phenomenon. The mean and CV are the two most common characteristics of the spiking activity. Thus, it is natural that the quality of the estimators depends on both

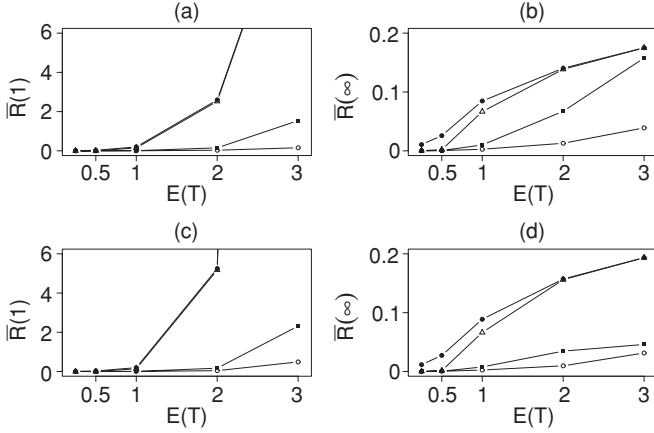


FIG. 7. The average RISEs  $R(1)$  (left column) and  $R(\infty)$  (right column) for four different estimators of the cdf computed from 500 realizations of 400 point processes observed in a window of length 1. Five different values of  $\mu$  are considered: 0.25, 0.5, 1, 2, and 3. For each  $\mu$  the average error is shown for the following estimators:  $\hat{F}_n^{\text{EM}}$  (triangles),  $\hat{F}_n^{\text{KM}}$  (circles),  $\hat{F}_n^{\text{RM}}$  (squares), and  $\hat{F}_n^{\text{MP}}$  (bullets). The results are depicted for (a)–(b) the renewal process with gamma distributed ISIs and  $\gamma = 0.5$ , (c)–(d) the renewal process with inverse Gaussian distributed ISIs and  $\gamma = 0.5$ .

of them simultaneously, and finally both the shape and position of the ISI distribution determine how well it can be estimated from observations obtained in the limited size window. An alternative characteristic of the data, probably closely related to the estimators' performance, would depend on the value of  $F(t)$  at  $\Delta$ ,  $F(\Delta)$ , which is not investigated in this paper.

For the case  $E(T) = 3$  and  $\text{CV}(T) = 0.5$  there are no spikes or just one observed in majority of trials because  $F(1) \doteq 0.0465$  for gamma-distributed ISIs and  $F(1) \doteq 0.0162$  for inverse Gaussian-distributed ISIs. These small values of  $F(1)$  entail large errors  $R(1)$ , in particular for  $\hat{F}_n^{\text{EM}}$  and  $\hat{F}_n^{\text{MP}}$ ; see Fig. 7. While  $R(1)$  tells us how precisely we estimate  $F(t)$  for  $t \in [0, \Delta]$ ,  $R(\infty)$  contains information about the error caused by estimating the tail using methods from Sec. IID. Since we are using exponential distribution to estimate the tail, the differences between  $R(1)$  and  $R(\infty)$  are less pronounced in the case of the Poisson process and more apparent for other processes, in particular for estimators  $\hat{F}_n^{\text{EM}}$  and  $\hat{F}_n^{\text{MP}}$ .

In the case of a Poisson or mixed Poisson process [Fig. 6(a)–6(b) and 6(g)–6(h)],  $\hat{F}_n^{\text{MP}}$  as well as  $\hat{F}_n^{\text{EM}}$  have the best performance. However, they are much worse when applied to the data formed by the renewal process [Figs. 6(c)–6(f) and 7]. In this case the estimators  $\hat{F}_n^{\text{KM}}$  and  $\hat{F}_n^{\text{RM}}$  are preferable. The Kaplan-Meier estimator is more efficient and has the smallest deviation from the true cdf. Moreover, the Kaplan-Meier estimator is quite reliable also for the case of the Poisson process. Therefore, it seems that the Kaplan-Meier estimator offers a good compromise.

Finally, let us shortly show the role of the number of spike trains. Figure 8 illustrates the results of a simulation study for  $n$  Poisson processes with  $\lambda = 1$  observed in a window of length  $\Delta = 1$ . Since the Poisson process is a special case of mixed Poisson process, the estimator  $\hat{F}_n^{\text{MP}}$  gives the best results. As expected, the errors of the estimates are decreasing with an increasing number of spike trains.

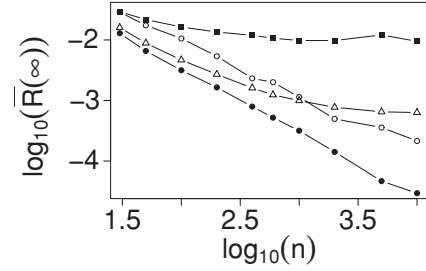


FIG. 8. The errors  $R(\infty)$  as the function of the number of records in the case of the Poisson process with intensity  $\lambda = 1$  observed in a window of length  $\Delta = 1$ . The errors are obtained as averages over 500 simulations. For better visualization, a log-log plot is used. Four estimators of  $F(t)$  are compared:  $\hat{F}_n^{\text{EM}}$  (triangles),  $\hat{F}_n^{\text{KM}}$  (circles),  $\hat{F}_n^{\text{RM}}$  (squares), and  $\hat{F}_n^{\text{MP}}$  (bullets).

#### IV. CONCLUDING REMARKS

We have considered several nonparametric estimators of the cdf for ISIs. At first, we introduced the estimators for data obtained in a single spike train experiment. The methods were remarkably different only if just a few ISIs were available. These methods were introduced in order to proceed to the case of multiple spike trains. In that situation, two basic approaches were applied: averaging the individual estimators for each spike train and pooling the information from all spike trains to evaluate the estimators. The simulations indicated that the estimators based on the pooled data were better than those calculated as the means of the separate estimators.

The estimator  $\hat{F}_n^{\text{MP}}$  was constructed under the assumption of a mixed Poisson process, and it has much lower quality if this assumption is violated. Other estimators were more robust. They worked reasonably even if there were no or very few ISIs observed, provided that the number of spike trains was sufficiently large. A reduced sample estimator discarded a lot of information given by the data when estimating  $F(t)$  for larger  $t$ . A modified ecdf had poor quality for non-Poisson renewal processes. In conclusion, in regard to the accuracy, among all the tested estimators we recommend the Kaplan-Meier estimator  $\hat{F}_n^{\text{KM}}$ , which gave satisfactory results for all considered models.

Another important point is the biological plausibility of the estimators. It was not our aim to speculate how real neurons could evaluate the distribution of interspike intervals. However, we can conclude that for a postsynaptic neuron it would be easier to implement  $\hat{F}_n^{\text{MP}}$  because it requires only the spike counts, and the exact firing times are not needed. In order to evaluate all other estimators, a neuron would have to use the exact timing of spikes.

#### ACKNOWLEDGMENTS

The authors thank an anonymous referee for his or her valuable comments. Z.P. was supported by the Czech Ministry of Education, project MSM 0021620839, and by the Grant Agency of the Academy of Sciences of the Czech Republic, project IAA 101120604. P.L. was supported by the Center for Neurosciences LC554, Grant No. AV0Z50110509, and by the Grant Agency of the Czech Republic, project P103/11/0282.

- [1] D. H. Perkel and T. H. Bullock, *Neurosci. Res. Program Bull.* **6**, 221 (1968).
- [2] F. Theunissen and J. P. Miller, *J. Comput. Neurosci.* **2**, 149 (1995).
- [3] W. Gerstner and W. M. Kistler, *Spiking Neuron Models* (Cambridge University Press, New York, 2002).
- [4] E. D. Adrian, *The Basis of Sensation* (W. W. Norton, New York, 1928).
- [5] N. Masuda and K. Aihara, *Phys. Rev. Lett.* **88**, 248101 (2002).
- [6] F. Rieke, D. Warland, R. de Ruyter van Steveninck, and W. Bialek, *Spikes: Exploring the Neural Code* (MIT Press, Cambridge, MA, 1997).
- [7] S. Shinomoto and S. Koyama, *Stat. Med.* **26**, 4032 (2007).
- [8] R. B. Stein, E. R. Gossen, and K. E. Jones, *Nat. Rev. Neurosci.* **6**, 389 (2005).
- [9] L. Kostal and P. Lansky, *Biol. Cybern.* **94**, 157 (2006).
- [10] T. Shimokawa and S. Shinomoto, *Neural Comput.* **21**, 1931 (2009).
- [11] H. C. Tuckwell, *Math. Biosci.* **207**, 246 (2007).
- [12] R. D. Vilela and B. Lindner, *J. Theor. Biol.* **257**, 90 (2009).
- [13] L. Kostal, P. Lansky, and J.-P. Rospars, *Eur. J. Neurosci.* **26**, 2693 (2007).
- [14] D. N. Ruskin, D. A. Bergstrom, and J. R. Walters, *J. Neurophysiol.* **88**, 487 (2002).
- [15] B. N. Lundstrom and A. L. Fairhall, *J. Neurosci.* **26**, 9030 (2006).
- [16] A. Longtin and D. M. Racicot, *Biosystems* **40**, 111 (1997).
- [17] H. Shimazaki and S. Shinomoto, *Neural Comput.* **19**, 1503 (2007).
- [18] E. Urdapilleta and I. Samengo, *Biol. Cybern.* **101**, 265 (2009).
- [19] N. Masuda and K. Aihara, *Neural Comput.* **15**, 1341 (2003).
- [20] M. P. Nawrot, C. Boucsein, V. R. Molina, A. Riehle, A. Aertsen, and S. Rotter, *J. Neurosci. Meth.* **169**, 374 (2008).
- [21] Z. Pawlas, L. B. Klebanov, M. Prokop, and P. Lansky, *Neural Comput.* **20**, 1325 (2008).
- [22] N. Brunel, F. S. Chance, N. Fourcaud, and L. F. Abbott, *Phys. Rev. Lett.* **86**, 2186 (2001).
- [23] W. Gerstner, *Neural Comput.* **12**, 43 (2000).
- [24] M. C. Wiener, *Neural Comput.* **15**, 2565 (2003).
- [25] D. R. Cox and P. A. W. Lewis, *The Statistical Analysis of Series of Events* (Methuen, London, 1966).
- [26] M. Badoual, M. Rudolph, Z. Piwkowska, A. Destexhe, and T. Bal, *Neurocomputing* **65–66**, 493 (2005).
- [27] K. Fujiwara and K. Aihara, *Eur. Phys. J. B* **68**, 283 (2009).
- [28] S. Ikeda and J. H. Manton, *Neural Comput.* **21**, 1714 (2009).
- [29] D. E. McKeegan, *Brain Res.* **929**, 48 (2002).
- [30] K. Miura, M. Okada, and S. Amari, *Neural Comput.* **18**, 2359 (2006).
- [31] D. H. Berger and K. H. Pribram, *Biol. Cybern.* **67**, 191 (1992).
- [32] A. R. Bulsara, T. C. Elston, C. R. Doering, S. B. Lowen, and K. Lindenberg, *Phys. Rev. E* **53**, 3958 (1996).
- [33] A. R. Bulsara, S. B. Lowen, and C. D. Rees, *Phys. Rev. E* **49**, 4989 (1994).
- [34] A. Buonocore, A. D. Crescenzo, and E. D. Nardo, *Biosystems* **67**, 27 (2002).
- [35] A. Burkitt, *Biol. Cybern.* **95**, 1 (2006).
- [36] G. L. Gerstein and B. Mandelbrot, *Biophys. J.* **4**, 41 (1964).
- [37] S. Iyengar and Q. Liao, *Biol. Cybern.* **77**, 289 (1997).
- [38] K. Kang and S. Amari, *Neural Comput.* **20**, 1411 (2008).
- [39] H. E. Plesser and T. Geisel, *Phys. Rev. E* **63**, 031916 (2001).
- [40] T. Shimokawa, K. Pakdaman, and S. Sato, *Phys. Rev. E* **60**, R33 (1999).
- [41] E. Urdapilleta and I. Samengo, *Phys. Rev. E* **80**, 011915 (2009).
- [42] D. J. Daley and D. Vere-Jones, *An Introduction to the Theory of Point Processes*, 2nd ed. (Springer, Berlin, 2005).
- [43] J. Grandell, *Doubly Stochastic Poisson Processes*, Lect. Notes in Math. 529 (Springer, Berlin, 1976).
- [44] D. H. Johnson, *J. Comput. Neurosci.* **3**, 275 (1996).
- [45] P. Lansky and J. Vaillant, *Biosystems* **58**, 27 (2000).
- [46] S. B. Lowen and M. C. Teich, *Phys. Rev. A* **43**, 4192 (1991).
- [47] D. R. Cox and V. Isham, *Point Processes* (Chapman and Hall, London, 1980).
- [48] A. F. Karr, *Point Processes and Their Statistical Inference*, 2nd ed. (Marcel Dekker, New York, 1991).
- [49] A. Baddeley and R. D. Gill, *Ann. Statist.* **25**, 263 (1997).
- [50] R Development Core Team, *R: A Language and Environment for Statistical Computing* (R Foundation for Statistical Computing, Vienna, Austria, 2010), [<http://www.R-project.org>].