Detection of dependence patterns with delay

Julien Chevallier, Thomas Laloë

Detection of dependence patterns with delay

Julien Chevallier*, Thomas Laloë

Université de Nice

Julien.CHEVALLIER@unice.fr

Summary

The Unitary Events (UE) method is a popular and efficient method used this last decade to detect dependence patterns of joint spike activity among simultaneously recorded neurons. The first introduced method is based on binned coincidence count (Grün, 1996) and can be applied on two or more simultaneously recorded neurons. This counting method is known to be subject to loss in synchrony detection (Grün *and others*, 1999). This defect has been corrected by the multiple shift coincidence count (Grün *and others*, 1999) for discrete time recordings of two simultaneously recorded neurons. This multiple shift coincidence count has recently been transposed in the continuous time framework (Tuleau-Malot *and others*, 2014) with the notion of delayed coincidence count (also for two neurons). The extension of this count to more than two neurons has not been investigated until the present work. First of all, we propose a generalization of the delayed coincidence count for more than two neurons. The point processes framework allows computations leading to a Gaussian approximation of the count for Poissonian spike trains. Since unknown parameters are involved in the approximation, a plug-in step is needed (where unknown parameters are replaced by estimated ones) and leads to a modification of the limit distribution. Finally the method takes the multiplicity of the tests into account via a Benjamini and Hochberg approach (Benjamini and Hochberg, 1995), to guarantee a prescribed control of the false discovery rate. We compare our new method and the UE method proposed in (Grün *and others*, 2002) over various simulations including changes in the underlying model. Furthermore our method is applied on real data.

Key words: Unitary Events - Coincidence pattern - Independence tests - Multiple testing - Poisson processes - Synchronization - Neuronal assemblies

*To whom correspondence should be addressed.

I. Introduction

The communication between neurons relies on their capacity to generate characteristic electric pulses called action potentials. These action potentials are usually assumed to be identical stereotyped events. Their maximum (called spike) can be considered as the relevant information. That is why the study of spike frequencies (firing rates) of neurons plays a key role in the comprehension of the information transmission in the brain (Abeles, 1982; Gerstein and Perkel, 1969; Shinomoto, 2010). One of the most important way to study theses spikes has been the recording of neurons activity via electrodes inserted in a laboratory animal's brain. Using spike sorting methods, these events are identified (associated to a neuron), and the measurements made by the electrodes end up in sequences of spikes (called spike trains).

In this article, the issue of detecting dependence patterns between simultaneously recorded spike trains is addressed. Despite the fact that some studies used to consider neurons as independent entities (Barlow, 1972), it is now well established that neurons can possibly coordinate their activity (Hebb, 1949; Palm, 1990; Sakurai, 1999; von der Malsburg, 1981). The understanding of this synchronization phenomenon (Singer, 1993) required the development of specific descriptive analysis methods of spike-timing over the last decades: cross-correlogram (Perkel *and others*, 1967), gravitational clustering (Gerstein *and others*, 1985) or joint peristimulus time histogram (JPSTH, Aertsen *and others* (1989)). In particular, Grün and collaborators developed one of the most popular and efficient method used this last decade: the Unitary Events (UE) analysis method (Grün, 1996) and the corresponding independence test, which detects where dependence lies by assessing p-values. A Unitary Event is a pattern that recurs more often than expected by chance. This method is based on a binned coincidence count that is unfortunately known to suffer a loss in synchrony detection. This flaw has been corrected by the multiple shift coincidence count (Grün *and others*, 1999). This method is used on discrete time processes. A new method (MTGAUE), based on a generalization of this count, the delayed coincidence count, has recently been proposed (Section 3.1 of Tuleau-Malot *and others* (2014) for two neurons). The results presented in this article are in the lineage of this newest method and is applied on continuous point processes which are random sets of points. Testing independence between real valued random variable is a well known problem. Various techniques have been developed, from the classical chi-square test to re-sampling methods for example. The interested reader may look at (Lehmann and Romano, 2005). Some of these methods and more general surrogate data methods have been applied on binned coincidence count, since the binned process transforms the spike train in vectors of finite dimension. However, the case of point processes that are not preprocessed needs other tools and remains to study. The binned method can indeed deal with several neurons (six simultaneously recorded neurons are analysed in (Grün *and others*, 2002). However, both of the improvements (Multiple Shift and MTGAUE) can only consider pairs of

neurons. Thus, our goal is to generalize the method introduced in (Tuleau-Malot *and others*, 2014) for more than two neurons. Following (Tuleau-Malot *and others*, 2014), spike trains are here modelled by point processes.

In Section II, we introduce the different notions of coincidence used through this article. In Section III, a test is established and the asymptotic control of its level is proved. In Section IV our test is confronted to the original UE method on simulated data and the accuracy of the Gaussian approximation is verified. In Section V the relevance of our method when our main theoretical assumptions are weakened is also empirically put on test. Section VI presents an illustration on real data. All the technical proofs are given in the Appendix.

II. Notions of coincidence and the classical UE methods

In order to detect synchronizations between the involved neurons, different notions of coincidence can be considered. Classically, there is a coincidence between neurons when they each emit a spike more or less simultaneously. This notion has already been used in UE methods (Grün *and others*, 2002) and is based on the following idea: a real dependency between $n \geqslant 2$ neurons should be characterized by an unusually large (or low) number of coincidence (Grammont and Riehle, 2003; Grün, 1996; Tuleau-Malot *and others*, 2014).

II.1 *Two notions of coincidence*

The UE method (see Grün (1996)) considers discretized spike trains at a resolution ℓ of typically 1 or 0.1 millisecond. Therefore, each trial consists in a set of n spike trains (one for each recorded neuron), each being represented by a sequence of 0 and 1 of length S. Since it is quite unlikely that two spikes occur at exactly the same time at this resolution ℓ, spike trains are binned and clipped at a coarser level. More precisely for a fixed bin size $\Delta = d\ell$, a new sequence of length S/d of 0 and 1 is associated to each spike train (1 if at least one spike occurs in the corresponding bin, 0 otherwise). For more precise informations on the binning procedure and the link with point processes we refer the interested reader to Tuleau-Malot *and others* (2014).

A constellation or pattern is a vector of size n of 0 and 1 (see Figure 1 or Grün *and others* (2002)). Of course, there are 2^n different constellations. The UE statistic associated to some constellation w consists in counting the number of occurrences of such w in the set of S/d vectors of size n

However, as shown in Figure 1, this method largely depends on the bin choice and it has been proved in Grün *and others* (1999) that this can lead in the case $n = 2$ to up to 60% of loss in detection when Δ is of the order of

the range of interaction.

Therefore, we propose a generalization to the case $n > 2$ of the notion of delayed coincidence count introduced in Tuleau-Malot *and others* (2014), which was already inspired by Grün *and others* (1999).

Because delayed coincidence count is based on non discretized data, constellations cannot be considered. However, it is always possible to associate to each constellation w a set $\mathcal{L}(w)$ of indices corresponding to the positions of the 1's (see Figure 1). In this respect, a dependence pattern refers either to a constellation w or to a set $\mathcal{L}$ of indices.

Considering $N_1, \ldots, N_n$, some point processes on $[a, b]$, and $\mathcal{L}$, a set of indices $i_1 < \cdots < i_L$, the *delayed coincidence count* $X_{\mathcal{L}}$ (of delay δ) over the neurons of subset $\mathcal{L}$ is given, for $\delta < \frac{b-a}{2}$, by

$$X_{\mathcal{L}} = X_{\mathcal{L}}(\delta) = \sum_{(x_1,\ldots,x_L)\in N_{i_1}\times\cdots\times N_{i_L}} \mathbf{1}_{\left|\max_{i\in\{1,\ldots,L\}} x_i - \min_{i\in\{1,\ldots,L\}} x_i\right| \leqslant \delta}.$$

The way coincidences are count can be explained in the following way (see Figure 1):

- Fix some duration parameter δ which is the equivalent of the bin size Δ,
- Count how many times each neuron in $\mathcal{L}$ spike almost at the same time, modulo the delay δ.

This is intuitively a good marker of the dependence between neurons because the influence of a neuron over others (whether exciting or inhibiting) results in the presence (or absence) of coincidence patterns, leading to the detection of synchronisation (or anti-synchronisation if a pattern occurs too few times).

That is a simple and meaningful way to count coincidences in a continuous manner. However, more general ways to count are possible and the results with respect to $X_{\mathcal{L}}$ can easily be transposed to more general counts (see the Appendix).

II.2 *Original UE method*

The final goal is to detect dependency between neurons. The idea is to compare two estimators of the expected coincidence count. The first one is the empirical mean $\bar{m}_w$ of the number of coincidence (i.e. for the UE method, the occurrences of a given constellation w) through M trials,

$$\bar{m}_w = \frac{1}{M}\sum_{k=1}^{M} m_w^{(k)},$$

where $m_w^{(k)}$ is the number of occurrences of w during the k^{th} trial. This estimator is consistent even with dependency between the spike trains. The second one is consistent under the independence hypothesis, and is given by

$$\hat{m}_{g,w} = \frac{S}{d} \prod_{l \in \mathcal{L}(w)} \hat{p}_l \prod_{k \notin \mathcal{L}(w)} (1 - \hat{p}_k), \tag{II.1}$$

where $\hat{p}_i$ is the empirical probability of finding a spike in a bin of neuron i.

This enables the construction of the test described in Grün *and others* (2002) and based on the comparison between the statistic $M\bar{m}_w$ and a quantile of the Poisson distribution $\mathcal{P}(M\hat{m}_{g,w})$ where M is the length of the sample. Most of the time only tests by upper values are computed (Grün, 1996; Grün *and others*, 2002). Following the study of Tuleau-Malot *and others* (2014), we have decided to focus on symmetric tests. The one based on the UE method rejects the independence hypothesis when $\bar{m}_w$ is too different from $\hat{m}_{g,w}$. More precisely, the symmetric independence test with significance level α is governed by the following rule: if

$$M\bar{m}_w \geqslant q_{1-\alpha/2} \quad \text{or} \quad M\bar{m}_w \leqslant q_{\alpha/2},$$

where q_x is the x-quantile of the Poisson distribution $\mathcal{P}(M\hat{m}_{g,w})$, then the independence hypothesis is rejected.

The UE method is applied under the hypothesis that the discrete processes modelling the spike trains of neurons are in fact Bernoulli processes. The equivalent in the "continuous" framework is the Poisson process (as it can be seen in Tuleau-Malot *and others* (2014)). This leads to a different estimator of the expected coincidence count and a different test which are defined properly in the next section.

III. Study of the delayed coincidence count

Once the notion of coincidence is defined with respect to continuous data, mathematical tools can be used to construct the desired independence test. The procedure is to compute the expectation and the variance of the variable $X_{\mathcal{L}}$. These computations classically implies a Gaussian approximation with respect to i.i.d trials. However, in order to be useful from a statistical point of view, the Gaussian approximation requires the knowledge of estimators of the expectation and the variance of $X_{\mathcal{L}}$. The next step is to replace these two parameters by corresponding estimates. This plug-in procedure is known to change the underlying distribution. As in Tuleau-Malot *and others* (2014), the delta method provides the exact nature of this change.

In the continuous framework, a sample is composed by M observations of $N_1, \cdots, N_n$ the point processes associated to spikes trains of n neurons on a window $[a, b]$. The goal is to answer the following question:

Given $\mathcal{L}$ a subset of $\{1,\dots,n\}$, are the processes $N_l,\ l\in\mathcal{L}$ independent?

To do this, a statistical test comparing the two hypotheses

$$\begin{cases}(\mathcal{H}_0) & \text{The processes } N_l,\ l\in\mathcal{L} \text{ are independent;}\\ (\mathcal{H}_1) & \text{The processes } N_l,\ l\in\mathcal{L} \text{ are not independent;}\end{cases}$$

is proposed.

In this section our test and its asymptotic relevance are introduced. First, let us present and discuss our main assumptions which are the same as in Tuleau-Malot *and others* (2014).

Assumption A1. *$N_1,\dots,N_n$ are Poisson processes.*

This assumption can be resumed to an assumption of independence of a point process with respect to itself over the time, as Bernoulli processes in discrete settings.

Assumption A2. *The Poisson processes $N_1,\dots,N_n$ are homogeneous on $[0,T]$*

Assumption A2 may also appear very restrictive. But once again Bernoulli processes considered in Grün *and others* (1999, 2002) have the same drawback. Moreover, if necessary, one can partition $[0,T]$ in smaller intervals on which A2 is satisfied. For more precise informations on Poisson processes we refer the interested reader to Kingman (1993).

This assumptions are necessary in this work in order to obtain an explicit form for the expected number of coincidences (and its variance). Note that there exist some trial-shuffling methods in the literature for which there is no need of a model on the data (Pipa *and others*, 2003; Pipa and Grün, 2003). However, they are based on binned coincidence count, and there is no equivalent with a delayed coincidence count, due to serious computational issues. Alternative works have also been done in the Bayesian paradigm (Archer *and others*, 2013).

III.1 *Asymptotical properties*

In order to build our independence test, we need to understand the behavior of the number of coincidence $X_{\mathcal{L}}$ under the independence hypothesis $\mathcal{H}_0$. In particular, the expected value and the variance of $X_{\mathcal{L}}$ must be computed. In a general point processes framework, these computations are impossible. This is why some restrictive assumptions are needed, such as A1, A2 and the independence of the processes, as done in the original UE method where independent Bernoulli processes have been considered.

Theorem III.1. *Let $\mathcal{L}$ and $X_{\mathcal{L}}$ be as defined previously. Suppose assumptions A1 and A2 and denote by $\lambda_1, \dots, \lambda_n$ the respective intensities of $N_1, \dots, N_n$. Under hypothesis $\mathcal{H}_0$, the expected value and the variance of the number of coincidences $X_{\mathcal{L}}$ are given by:*

$$m_0 := \mathbb{E}\left[X_{\mathcal{L}}\right] = \left(\prod_{l=1}^{L} \lambda_{i_l}\right) I(L,0)$$

and

$$Var(X_{\mathcal{L}}) = m_0 + \sum_{k=1}^{L-1} \left(\sum_{\substack{\mathcal{J} \subset \mathcal{L} \\ \#\mathcal{L}=k}} \prod_{j \in \mathcal{J}} \lambda_{i_j}^2 \prod_{l \notin \mathcal{J}} \lambda_{i_l} \right) I(L,k),$$

where the $I(L,k)$ are given by Proposition III.1 below.

The proof lies on the calculus of the moments of a sum over a Poisson Process and is given in Appendix VII.

The expressions of the integrals $I(L,k)$ are not trivial to obtain, but calculations can be made as showed in the following result.

Proposition III.1. *For $b > a \geqslant 0$ and $0 < \delta < b-a$, define for every $k \in \{0, \dots, L\}$*

$$I(L,k) = \int\limits_{[a,b]^{L-k}} \left(\int\limits_{[a,b]^k} \mathbf{1}_{\left| \max\limits_{i \in \{1,\dots,L\}} x_i - \min\limits_{i \in \{1,\dots,L\}} x_i \right| \leqslant \delta} dx_1 \dots dx_k \right)^2 dx_{k+1} \dots dx_L,$$

where the convention $\int\limits_{[a,b]^0} f(x) = f(x)$ is set. Then, for $L \geqslant 2$, and $k \in \{0, \dots, L-1\}$

- $I(L,L) = L^2 (b-a)^2 \delta^{2L-2} - 2L(L-1)(b-a)\delta^{2L-1} + (L-1)^2 \delta^{2L}$
- $I(L,k) = f(L,k)(b-a)\delta^{L+k-1} - h(L,k)\delta^{L+k}$

  where $f(L,k) = \dfrac{k(k+1) + L(L+1)}{L-k+1}$,

  and $h(L,k) = \dfrac{-k^3 + k^2(2+L) + k(5+2L-L^2) + L^3 + 2L^2 - L - 2}{(L-k+2)(L-k+1)}$

Remark Note that for $k < L$, $I(L,k)$ is of the order of δ^{L+k-1} (when $\delta \ll b-a$).

III.2 *Independence test*

Now that the behavior of $X_{\mathcal{L}}$ under $\mathcal{H}_0$ is known, the method to construct an independence test is quite clear. Suppose that M i.i.d. trials are given. Denote $N_i^{(k)}$ the spike train of neuron i during the k^{th} trial. The basic idea is to compare two estimates of the expectation of $X_{\mathcal{L}}$. The first one is the empirical mean of $X_{\mathcal{L}}$:

$$\bar{m}_{\mathcal{L}} = \frac{1}{M} \sum_{k=1}^{M} X_{\mathcal{L}}^{(k)}, \tag{III.2}$$

where $X_{\mathcal{L}}^{(k)}$ is the delayed coincidence count during the k^{th} trial. This estimate converges even if the processes are not independent. More precisely the following asymptotic result is given by the Central Limit Theorem

$$\sqrt{M}\left(\bar{m}_{\mathcal{L}} - \mathbb{E}\left[X_{\mathcal{L}}\right]\right) \underset{M\to\infty}{\overset{\mathcal{D}}{\longrightarrow}} \mathcal{N}\left(0, Var(X_{\mathcal{L}})\right),$$

where $\overset{\mathcal{D}}{\longrightarrow}$ denotes the convergence in distribution.

The second estimate is given by Theorem III.1. Indeed, under $\mathcal{H}_0$ the following equality holds

$$\mathbb{E}\left[X_{\mathcal{L}}\right] = m_0 = \left(\prod_{l=1}^{L} \lambda_{i_l}\right) I(L,0).$$

Then we only have to replace the spiking intensities λ_{i_l} by $\hat{\lambda}_i := \frac{1}{M(b-a)} \sum_{k=1}^{M} N_i^{(k)}\left([a,b]\right)$, where $N_i^{(k)}\left([a,b]\right)$ denotes the number of spikes in $[a,b]$ for neuron i during the k^{th} trial. Therefore, the following estimator is considered

$$\hat{m}_{0,\mathcal{L}} = \prod_{l=1}^{L} \hat{\lambda}_{i_l} . I\left(L,0\right) \tag{III.3}$$

So there are two estimates of $\mathbb{E}\left[X_{\mathcal{L}}\right]$ as for the UE method: $\bar{m}_{\mathcal{L}}$ who is always consistent and $\hat{m}_{0,\mathcal{L}}$ who is consistent only under $\mathcal{H}_0$. This leads to the following test: the independence assumption is rejected when the difference between $\bar{m}_{\mathcal{L}}$ and $\hat{m}_{0,\mathcal{L}}$ is large. More precisely, Theorem III.2 gives the asymptotic behavior of $\sqrt{M}\left(\bar{m}_{\mathcal{L}} - \hat{m}_{0,\mathcal{L}}\right)$ under $\mathcal{H}_0$.

Theorem III.2. *Under the notations and assumptions of Theorem III.1, and under $\mathcal{H}_0$, the following assumptions are true*

- *The following convergence in distribution holds:*

$$\sqrt{M}\left(\bar{m}_{\mathcal{L}} - \hat{m}_{0,\mathcal{L}}\right) \underset{M\to\infty}{\overset{\mathcal{D}}{\longrightarrow}} \mathcal{N}\left(0, \sigma^2\right),$$

  where $\mathcal{N}(\mu, s^2)$ denotes the gaussian distribution with mean μ, variance s^2 and

$$\sigma^2 = Var(X_{\mathcal{L}}) - (b-a)^{-1}\mathbb{E}\left[X_{\mathcal{L}}\right]^2 \left(\sum_{l=1}^{L} \lambda_{i_j}^{-1}\right).$$

- *Moreover, σ^2 can be estimated by*

$$\hat{\sigma}^2 = \hat{v}\left(X_{\mathcal{L}}\right) - (b-a)^{-1} I(L,L) \prod_{l=1}^{L} \hat{\lambda}_{i_l}^2 \left(\sum_{k=1}^{L} \hat{\lambda}_{i_k}^{-1}\right),$$

  where

$$\hat{v}(X_{\mathcal{L}}) = \hat{m}_{0,\mathcal{L}} + \sum_{k=1}^{L-1} \left(\sum_{\substack{\mathcal{J}\subset\mathcal{L} \\ \#\mathcal{L}=k}} \prod_{j\in\mathcal{J}} \hat{\lambda}_{i_j}^2 \prod_{l\notin\mathcal{J}} \hat{\lambda}_{i_l} \right) I(L,k),$$

  and

$$\sqrt{M} \frac{\left(\bar{m}_{\mathcal{L}} - \hat{m}_{0,\mathcal{L}}\right)}{\sqrt{\hat{\sigma}^2}} \overset{\mathcal{D}}{\to} \mathcal{N}\left(0,1\right).$$

The proof relies on a standard application of the delta method (Casella and Berger, 2002) and is given in Appendix B.

Note that the results obtained in Theorems III.1 and III.2 are true for more general delayed coincidence counts. A more general result and its proof are given in Appendix. However when one considers more general ways to count coincidences the integrals $I(L,k)$ are harder to compute.

The results obtained in Theorem III.2 allow us to easily build a test for detecting a dependency between neurons:

Definition III.1 (The GAUE test)**.** *For $\alpha \in]0,1[$, denote z_α the α-quantile of the standard Gaussian distribution $\mathcal{N}(0,1)$. Then the symmetric test which rejects $\mathcal{H}_0$ when $\bar{m}$ and $\hat{m}_{0,\mathcal{L}}$ are too different is defined by*

$$\left| \sqrt{M} \frac{(\bar{m}_{\mathcal{L}} - \hat{m}_{0,\mathcal{L}})}{\sqrt{\hat{\sigma}^2}} \right| \leqslant z_{1-\alpha/2}.$$

Note that once a subset is rejected by our test, one can determine if the dependency is rather excitatory or inhibitory according to the sign of $\bar{m}_{\mathcal{L}} - \hat{m}_{0,\mathcal{L}}$. If $\bar{m}_{\mathcal{L}} - \hat{m}_{0,\mathcal{L}} > 0$ (*respectively* < 0) then the dependency is rather excitatory (*respectively inhibitory*).

The following corollary is an immediate consequence of Theorem III.2.

Corollary III.1. *Under assumptions of Theorem III.2, the test presented in Definition III.1 is asymptotically of level α.*

IV. Illustration Study

In this section, an illustration of the previous theoretical results is given. In order to do that, Poisson processes are simulated. We choose a Framework $\mathbf{F}_1$ (size of the window, number of neurons, discharge rates) close to real data:

$\mathbf{F}_1$:
- Trial duration of $b-a$ is randomly selected (uniform distribution) between 0.2s and 0.4s;
- $n = 4$ neurons with different intensities randomly selected (uniform distribution) between 8 and 20 Hz;
- $\mathcal{L} = \{1,2,3,4\}$;
- Coincidence delay δ randomly selected (uniform distribution) between 0.015s and 0.025s.

IV.1 *Illustration of Theorem III.2*

To empirically validate the theoretical result on the level of our test we simulate independent Poisson processes with the previously provided parameters (Framework $\mathbf{F}_1$). Considering M independent trials of n point processes, the asymptotic of the delayed coincidence count is studied when M grows. To this aim, we use a Monte Carlo

method via a large number of simulations (1000) of M trials. For each iteration, we randomly select a new set of parameters (trial duration, intensities, coincidence delay). Finally, on each simulation, we compute our statistic $S_i = \sqrt{M}\frac{(\bar{m}_{\mathcal{L},i} - \hat{m}_{0,i})}{\sqrt{\hat{\sigma}_i^2}}$ (for i from 1 to 1000) and plot (Figure 2) the Kolmogorov distance $KS(F_{M,1000}, F)$ between the estimated distribution function over the 1000 repetition $F_{M,1000}$ and the standard Gaussian distribution function F:

$$KS(F_{M,1000}, F) = \sup_x |F_{M,1000}(x) - F(x)|.$$

As expected, the Kolmogorov distance decreases to 0 as M grows. Moreover, it seems reasonable to consider sample size of the order of 30 or larger since the KS distance does not decrease significantly after that.

Under $\mathcal{H}_0$, the theoretical results give that the p-values should be asymptotically distributed (in M) as the uniform distribution. Thus, the evolution (with respect to M) of the Kolmogorov distance between the empirical distribution function of the obtained p-values (with our test and the one given by the UE method) and the uniform distribution is plotted for symmetric tests (See Figure 3). As previously, it seems reasonable to consider sample size of the order of 30 or larger. Moreover, the distribution of the p-values given by the UE test does not converge to the uniform. In order to describe more precisely what happens, we plot in Figure 4 the sorted p-values in function of their normalized rank (for $M = 50$). Note that if a curve is below (respectively above) the diagonal, then the probability of rejecting independence is more (respectively less) important than it should be under $\mathcal{H}_0$. Our test seems to be too conservative except for very small p-values. The problem induced by this non conservativeness for very small p-values is detailed at the end of Section V. However, the empirical frequency of p-values lower than 0.05 is 5%. On the other side, the UE test rejects too many cases. For example, the UE test with level 5% rejects almost 10% of the cases.

IV.2 *Illustration of the Power of the test*

To evaluate the power of the test, we simulate a sample which is dependent and check how many times the test rejects $\mathcal{H}_0$. Note that unlike the level of the test, no theoretical information can be deduced from Theorem III.2, since we do not now the distribution of our statistic if we are not under $\mathcal{H}_0$.

To obtain dependent Poisson processes an injection model inspired by the one used in Grün *and others* (2002, 1999) or Tuleau-Malot *and others* (2014) is used. Consider independent homogeneous Poisson processes $\overline{N}_1, \ldots, \overline{N}_n$, drawn according to Framework $\mathbf{F}_1$. We simulate an other Poisson process (according to the same framework but independent from the previous ones) $N_{\{1,\ldots,n\}}$ which is injected for every neuron. Thus our se-

quence of dependent Poisson processes is given by

$$N_i = \overline{N}_i \cup N_{\{1,\dots,n\}}.$$

This new framework ($\mathbf{F}_1$ completed by the injection) is referred as Framework $\mathbf{F}_2$.

Note that in the injection model used in Grün *and others* (1999), a small shift is applied before injection. In our Poissonian framework this shift cannot be performed in order to keep the Poissonian properties of the processes under H_1. Moreover, this injection model can only model excess of coincidences and not lack of coincidences. For a fixed level 0.05, Figure 5 illustrates the power of the two tests in function of M. Then Figure 6 represents the p-values in function of their normalized rank, for $M = 50$. The gap between the two tests in terms of test power may seem significant (around 20% for small sample sizes in favor of the original UE tests) but this is at the price of an uncontrolled first kind error.

V. Non-Poissonian framework

In this section, a more realistic framework than the Poisson one is considered. Indeed, it is interesting to see if our test is still reliable when the Poisson framework is not valid anymore. Our test is confronted to multivariate Hawkes processes, which can be simulated thanks to Ogata's Thinning method (Ogata, 1981) inspired by Lewis and Shedler (1979). The use of Hawkes processes in neurobiology was first introduced in Chornoboy *and others* (1988). With the development of simultaneous neuron recording there is a recent trend in favor of Hawkes processes in terms of modelling spike trains (Krumin *and others* (2010); Pernice *and others* (2011, 2012); Tuleau-Malot *and others* (2014)). In this model, interaction between two neurons can be easily and in a more realistic way inserted. This is one of the reasons of this trend.

A counting process N^j is characterized by its conditional intensity λ_t^j which is related with the local probability of finding a new point given the past. (Informally, the quantity $\lambda_t^j dt$ gives the probability that a new point on N^j appears in $[t, t+dt]$ given the past.) So let us define the conditional intensities of a multivariate Hawkes process.

The process $\left(N^i\right)_{i=1\dots n}$ is a multivariate Hawkes process if there exist some functions $(h_{ij})_{i,j=1\dots n}$ (called interaction functions) and some positive constants $(\mu_i)_{i=1\dots n}$ (spontaneous intensities) such that λ^j given by

$$\lambda_t^j = \max\left(0, \mu_j + \sum_{i=1}^{n} \int_{s<t} h_{ij}\left(t-s\right) N^i\left(ds\right)\right)$$

is the intensity of the point process N^j, where $N^i(ds)$ is the point measure associated to N^i that is $N^i(ds) = \sum_{T \in N^i} \delta_T$ where δ_T is the Dirac measure at point T.

The functions h_{ij} represent the influence of neuron i over neuron j in terms of spiking intensity. This influence can be exciting ($h \geqslant 0$) or inhibiting ($h \leqslant 0$)

Remark 1. This includes the case where the interaction function can be negative, which is possible thanks to the positive part.

2. The homogeneous Poisson process is a particular case of Hawkes processes (take null interaction functions).

For example, suppose that $h_{ij} = \alpha \mathbf{1}_{[0,x]}$. If $\alpha > 0$ (*respectively* $\alpha < 0$) then the apparition of a spike on N^i increases (*respectively decreases*) during a delay (namely x) the probability to have a spike on N^j: neuron i excites (*respectively inhibits*) neuron j. The processes (N^i) in this Hawkes model are independent if and only if $h_{ij} = 0$ for all $i \neq j$.

Note also that the self-interaction functions h_{jj} can model refractory periods, making the Hawkes model more realistic than Poisson processes, even in the independence case. In particular when $h_{jj} = -\mu_j \mathbf{1}_{[0,x]}$, all the other interaction functions being null, the n-dimensional process is composed by n independent Poisson processes with dead time x, modelling strict refractory periods of length x (Reimer *and others* (2012)).

All the following tests are computed according to the Framework $\mathbf{F}_3$ below:

- Trial duration of $b - a$ is randomly selected (uniform distribution) between 0.2 and 0.4;
- $n = 4$ neurons with spontaneous intensity $\mu_1, \ldots, \mu_4$ randomly selected (uniform distribution) between 8 and 20 Hz;
- Negative Auto interaction $h_{i,i} = -\mu_i \mathbf{1}_{[0,0.005s]}$;
- $\mathcal{L} = \{1, 2, 3, 4\}$;
- Coincidence delay δ randomly selected (uniform distribution) between 0.015 and 0.025.

$\Big\} \mathbf{F}_3$

To illustrate the level and power of the test in this new framework (where no theoretical result is proven), we use the same methodology as in the Poissonian framework.

V.1 *Illustration of the level*

As in the Poissonian framework (Section IV.1), Figure 7 shows the evolution of the KS distance between $F_{M,1000}$ and F and Figure 8 the evolution of the KS distance between the sorted p-values and the uniform. The conclusions appear to be similar: it seems reasonable to consider sample size of the order of 30 or larger and the U.E. sorted p-values are further away from the uniform distribution.

Finally, Figure 9 plays the same role than Figure 6 and presents the sorted p-values in function of their normalized rank (for $M = 50$). Our test appears to be even more conservative than in the Poissonnian framework

(for example, the empirical frequency of p-values lower than 0.05 is of the order of 1% instead of 5% in the Poissonian framework). On the other side, UE test still rejects too many cases. For example, the UE test with level 5% rejects almost 10% of the cases.

V.2 *Illustration of the Power of the test*

As said previously, it is more realistic to introduce dependency between Hawkes processes than Poisson processes. Still considering Framework $\mathbf{F}_3$, interaction functions $h_{i,j} = \alpha \mathbf{1}_{[0,0.005s]}$, α being randomly selected between 20 and 30 Hz, are added. More precisely, we add three interaction functions: $h_{3,1}$, $h_{4,2}$ and $h_{2,1}$ (see Figure 10). This new framework ($\mathbf{F}_3$ completed by the three interaction function) is referred as Framework $\mathbf{F}_4$.

As previously we provide first an illustration of the power of the two tests, associated to a level of 0.05, in function of M (Figure 11). Then Figure 12 represents the p-values in function of their normalized rank, for $M = 50$. The gap between the two tests in terms of test power is smaller in this case (around 10% for small sample sizes, in favor of the original UE tests) but still at the price of an uncontrolled first kind error.

V.3 *Multiple pattern test*

Rather than perform the test on the complete pattern $\{1, 2, 3, 4\}$, one can test all the eleven sub-pattern of two, three or four neurons. In multiple testing, the notion of test level is not relevant. The closest notion to the level of a test might be the *Family-Wise Error Rate* (FWER) which is the probability to wrongly reject at least one of the tests. This error rate can be controlled using Bonferroni's method but it is too restrictive, in particular when the number of tests involved is too large. One popular way to deal with multiple testing is the Benjamini-Hochberg procedure (Benjamini and Hochberg, 1995) which ensures a control of the *False Discovery Rate* (FDR). False discoveries cannot be avoided but it is not a problem if the ratio of F_p the number of false positives (detections) divided by R the total number of rejects is controlled. Therefore, the FWER and the FDR are mathematically defined by $\text{FWER} = \mathbb{P}(F_p > 0)$ and $\text{FDR} = \mathbb{E}\left[\frac{F_p}{R}\mathbf{1}_{R>0}\right]$.

Note that in the full independent case, the FWER and the FDR are equal. The following procedure, due to Benjamini and Hochberg ensures a small FDR:

1. Fix a level q ($q = 5\%$ for example);
2. Denote by $(P_1, \ldots, P_K)$ the p-values obtained for all considered tests;
3. Order them in increasing order and denote the increasing vector $(P_{(1)}, \ldots, P_{(K)})$;

4. Note k_0 the largest k such that $P_{(k)} \leqslant kq/K$;

5. Then, reject all the tests corresponding to p-values smaller than $P_{(k_0)}$.

The theoretical result of (Benjamini and Hochberg, 1995) ensures that if the p-values are upper bounded by a uniform distribution and independently distributed under the null hypothesis, then the procedure guarantees a FDR less than q. The main drawback of this procedure in our case is that one needs to compute p-values that are very small when K is large. For example, if $K \geqslant 50$ and $q = 0.05$, the upper bound given by kq/K can be smaller than one thousandth and as noted in Section IV.1 the empirical frequency of very small p-values is greater than expected and therefore the uniform upper bound of the p-values is not guaranteed in our case. However, only 11 tests are considered here and the procedure still returns reliable results.

We perform 1000 simulations and count how many times the tests reject the independence. The results, obtained for $M = 50$, are presented in Figure 13. They show that our test mostly detects the pattern $\{1, 2\}$, $\{1, 3\}$, $\{2, 4\}$, $\{1, 2, 3\}$, $\{1, 2, 4\}$ and $\{1, 2, 3, 4\}$. This is consistent with the considered framework ($\mathbf{F}_4$) since the real connection are between $\{1, 2\}$, $\{1, 3\}$ and $\{2, 4\}$. Moreover, the asymmetry in terms of detection between $\{1, 3\}$ and $\{2, 4\}$ can be explained by our configuration which excites more Neuron 1 than any other. More precisely, the auto-inhibitions imply strict refractory periods only for neurons 3 and 4 since they are not excited by an other neuron. This is not true for neurons 1 and 2: For example, neuron 1 is excited by neurons 3 and 2 whereas neuron 2 is excited by neuron 4 only.

The U.E. test essentially detects much less the pattern $\{1, 2, 3, 4\}$ and to a lesser extent $\{1, 2, 4\}$.

VI. Real data study

Our test being validated on simulations, our method can be now applied on real data and results in agreement with classical knowledge on those data are shown.

VI.1 *Description of the data*

The data set considered here is the same as in Tuleau-Malot *and others* (2014) and previous experimental studies Grammont and Riehle (2003); Riehle *and others* (2000, 2006). The following description of the experiment is copied from Section 4.1 of Tuleau-Malot *and others* (2014). These data were collected on a 5-year-old male Rhesus monkey who was trained to perform a delayed multidirectional pointing task. The animal sat in a primate chair in front of a vertical panel on which seven touch-sensitive light-emitting diodes were mounted, one in the center and

six placed equidistantly (60 degrees apart) on a circle around it. The monkey had to initiate a trial by touching and then holding with the left hand the central target. After a fix delay of 500ms, the preparatory signal (PS) was presented by illuminating one of the six peripheral targets in green. After a delay of either 600ms (with probability 0.3) or 1200ms (with probability 0.7), it turned red, serving as the response signal and pointing target. Signals recorded from up to seven microelectrodes (quartz insulated platinum-tungsten electrodes, impedance: 2-5MΩ at 1000Hz) were amplified and band-pass filtered from 300Hz to 10kHz. Using a window discriminator, spikes from only one single neuron per electrode were then isolated. Neuronal data along with behavioral events (occurrences of signals and performance of the animal) were stored on a PC for off-line analysis with a time resolution of 10kHz. The idea of the analysis is to detect some conspicuous patterns of coincident spike activity appearing during the response signal in the case of a long delay (1200ms). Therefore, we only consider trials where the response signal is indeed occurring after a long delay.

VI.2 *The test*

We dispose of recordings of four neurons (35 trials by neurons) and we consider two sub windows: one between 300ms and 500ms (i.e. before the preparatory signal), the other between 1100ms and 1300ms (i.e. around the expected signal). Our idea is that more synchronisation should be detected during the second window. Moreover, for each window, all eleven subsets (of at least two neurons) of the four considered neurons are tested. Thus we use the Benjamini-Hochberg procedure (presented in the previous section) for $K = 22$ tests to perform. Moreover, we took several values for the delay δ between 0.015s and 0.025s and the results remained stable.

The results are presented in Figure 14. Note that we saw in sections IV and V that our test is too conservative even for small number of trials. This ensure that the level of our test can be trusted. We see that synchronizations between the subsets $\{3, 4\}$ and $\{1, 3, 4\}$ appear in the second window. These results suggest that neurons 1, 3 and 4 belong to a neuronal assembly which is formed around the expected signal. This is in agreement with more quantitative results on those data (Grammont and Riehle, 2003; Tuleau-Malot *and others*, 2014).

VII. Conclusion

This paper generalizes the delayed coincidence count introduced in Tuleau-Malot *and others* (2014) to more than two neurons. This delayed coincidence count leads to an independence test for point processes which are commonly used to model spike trains.

Under the hypothesis that the point processes are homogeneous Poisson processes, the expectation and variance

of the delayed coincidence count can be computed (Theorem III.1), and then a test with prescribed level is built (Theorem III.2). A simulation study allows us to confirm our theoretical results and to state the empirical validity of our test with a relaxed Poisson assumption. Indeed, we considered Hawkes processes which are a more realistic model of spike trains. The simulation study gives good results, even for small sample size. This allows us to use our test on real data, in order to highlight the emergence of a neuronal assembly involved at some particular time of the experiment.

However, we cannot achieve the full generalization of the MTGAUE method mainly because of the default of Gaussian approximation concerning extreme values of the test statistics. More precisely, very small p-values are not distributed as expected. In particular, as noted in Section IV, when the sample size (or number of trials M) is moderate, the present test returns too many very small p-values ($M = 50$). In Tuleau-Malot *and others* (2014), the MTGAUE method is applied simultaneously on 1900 sliding windows. In the present case, the total number of tests is even larger since, for each sliding window, there are $2^n - n - 1$ tests to perform, where n is the number of neurons. As said at the end of Section V this would lead to extremely small p-values. It could be therefore of interest to explore surrogate data method such as trial-shuffling (Pipa *and others*, 2003). A very recent work based on permutation approach for delayed coincidence count with $n = 2$ neurons (Albert *and others*, 2014) is a first step in this direction but needs to be generalized to more than 2 neurons.

Acknowledgments

We first of all want to thank Alexa Riehle, leader of the laboratory in which the data used in this article were previously collected, and Franck Grammont who collected these data. Finally we thank Patricia Reynaud-Bouret for fruitful discussions.

Conflict of Interest: None declared.

APPENDIX

As said in Section III.2, we prove more general results than Theorems III.1 and III.2. Considering $N_1, \dots, N_n$, some point processes on $[a,b]$, and $\mathcal{L}$, a set of indices $i_1 < \dots < i_L$ we prove the same results with any coincidence function $c(x_1, \dots, x_L)$ with value either 0 or 1 satisfying Definition A.1 below.

Definition A.1.

1. *A coincidence function is a function* $c : [a,b]^L \to \{0,1\}$ *which is symmetric.*

2. *Let* $(x_1, \dots, x_L) \in \prod_{l=1}^{L} N_{i_l}$ *be a L-uplet with a spiking time of every neuron of the subset* $\mathcal{L}$*. Say that* $(x_1, \dots, x_L)$ *is a coincidence if and only if* $c(x_1, \dots, x_L) = 1$*.*

3. *Given c a coincidence function we define* $X_{\mathcal{L}}$ *the number of coincidences on* $[a,b]$ *by:*

$$X_{\mathcal{L}} = \int_{[a,b]^L} c(x_1, \dots, x_L)\ dN_{i_1}(x_1) \dots dN_{i_L}(x_L)$$

where $dN_{i_1}, \dots, dN_{i_L}$ *are the point measures associated to* $N_{i_1}, \dots, N_{i_L}$*.*

4. *Define*

$$\forall k \in \{0, \dots, L\},\ I(L,k) = \int_{[a,b]^{L-k}} \left(\int_{[a,b]^k} c(x_1, \dots, x_L)\ dx_1 \dots dx_k \right)^2 dx_{k+1} \dots dx_L$$

A. Proof of Theorem III.1

Theorem A.1. *Under assumptions and Notations of Definition A.1, if* $N_1, \dots, N_n$ *are some independent homogeneous Poisson processes on* $[a,b]$ *with intensities* $\lambda_1, \dots, \lambda_n$*, the expected value and the variance of the number of coincidences* $X_{\mathcal{L}}$ *are given by:*

$$m_0 := \mathbb{E}[X_{\mathcal{L}}] = \left(\prod_{l=1}^{L} \lambda_{i_l} \right) I(L,0)$$

and

$$Var(X_{\mathcal{L}}) = m_0 + \sum_{k=1}^{L-1} \left(\sum_{\substack{\mathcal{J} \subset \mathcal{L} \\ \#\mathcal{L} = k}} \prod_{j \in \mathcal{J}} \lambda_{i_j}^2 \prod_{l \notin \mathcal{J}} \lambda_{i_l} \right) I(L,k).$$

Proof.

$$\mathbb{E}[X_{\mathcal{L}}] = \mathbb{E}\left[\int_{[a,b]^L} c(x_1, \dots, x_L)\ dN_{i_1}(x_1) \dots dN_{i_L}(x_L) \right]$$

Using the fact that $N_1, \cdots, N_n$ are independent homogeneous Poisson processes with respective intensities $\lambda_1, \cdots, \lambda_n$ one can prove (see Daley and Vere-Jones (2003)) that

$$\mathbb{E}[X_{\mathcal{L}}] = \left(\prod_{l=1}^{L} \lambda_{i_l} \right) \int_{[a,b]^L} c(x_1, \dots, x_L)\ dx_1 \dots dx_L$$

To compute the variance, first define $[a,b]^{(1)} = \left\{(x,y) \in [a,b]^2 \mid x = y\right\}$ and $[a,b]^{(2)} = [a,b]^2 \setminus [a,b]^{(1)}$.

Thanks to Fubini Theorem we have

$$\mathbb{E}\left[X_{\mathcal{L}}^2\right] = \mathbb{E}\left[\int\limits_{[a,b]^{2L}} c\left(x_1,\ldots,x_L\right) c\left(y_1,\ldots,y_L\right) \prod_{l=1}^{L} dN_{i_l}\left(x_l\right) dN_{i_l}\left(y_l\right)\right]$$

Let us see that $[a,b]^{2L} = \left([a,b]^2\right)^L$ and decompose the integral by piece taking $[a,b]^{(2)}$ or $[a,b]^{(1)}$ on each copy of $[a,b]^2$.

Denote $\Phi = \left\{\phi : \{1,\ldots,L\} \to \left\{[a,b]^{(1)}, [a,b]^{(2)}\right\}\right\}$ (used to formalize all the possible decompositions). We have

$$\mathbb{E}\left[X_{\mathcal{L}}^2\right] = \sum_{\phi\in\Phi} \mathbb{E}\left[\int\limits_{\prod_{k=1\ldots L}\phi(k)} c\left(x_1,\ldots,x_L\right) c\left(y_1,\ldots,y_L\right) \prod_{l=1}^{L} dN_{i_l}\left(x_l\right) dN_{i_l}\left(y_l\right)\right]$$

For $\phi \in \Phi$, denote $p = \text{Card}\left(\phi^{-1}\left([a,b]^{(1)}\right)\right)$. Using the symmetry of the coincidence function c, it is sufficient to compute when $\phi^{-1}\left([a,b]^{(1)}\right) = \{1,\ldots,p\}$. This leads by properties of the moment measure of Poisson processes (see Daley and Vere-Jones (2003) or Kingman (1993) in a more simplified framework):

$$\mathbb{E}\left[\int\limits_{\left([a,b]^{(1)}\right)^p}\int\limits_{\left([a,b]^{(2)}\right)^{L-p}} c\left(x_1,\ldots,x_L\right) c\left(y_1,\ldots,y_L\right) \prod_{l=1}^{L} dN_{i_l}\left(x_l\right) dN_{i_l}\left(y_l\right)\right]$$

$$= \prod_{l=1}^{p}\lambda_{i_l} \prod_{j=p+1}^{L}\lambda_{i_j}^2 \int\limits_{[a,b]^p}\left(\int\limits_{[a,b]^{2(L-p)}} c\left(t_1,\ldots,t_p,x_{p+1},\ldots,x_L\right) c\left(t_1,\ldots,t_p,y_{p+1},\ldots,y_L\right) \prod_{k=p+1}^{L} dx_k dy_k\right) dt_1\ldots dt_p$$

For fixed $(t_1,\ldots,t_p)$ one can apply Fubini Theorem to the inner integral which leads to:

$$\prod_{l=1}^{p}\lambda_{i_l} \prod_{j=p+1}^{L}\lambda_{i_j}^2 \int\limits_{[a,b]^p}\left(\int\limits_{[a,b]^{2(L-p)}} c\left(t_1,\ldots,t_p,t_{p+1},\ldots,t_J\right) dt_{p+1}\ldots dt_L\right)^2 dt_1\ldots dt_p$$

This expression is equal to $\prod_{l=1}^{p}\lambda_{i_l} \prod_{j=p+1}^{L}\lambda_{i_j}^2\, I\left(L, L-p\right)$. So we have:

$$\mathbb{E}\left[X_{\mathcal{L}}^2\right] = \mathbb{E}\left[X_{\mathcal{L}}\right]^2 + \sum_{k=0}^{L-1}\left(\sum_{\substack{\mathcal{J}\subset\mathcal{L}\\ \#\mathcal{J}=k}} \prod_{j\in\mathcal{J}}\lambda_{i_j}^2 \prod_{l\notin\mathcal{J}}\lambda_{i_l}\right) I(L,k)$$

where $k = \text{Card}\left(\phi^{-1}\left([a,b]^{(2)}\right)\right) = L-p$. More precisely, the term $\mathbb{E}\left[X_{\mathcal{L}}\right]^2$ corresponds to $\phi \equiv [a,b]^{(2)}$ $(k = L)$.

To finish the proof, it suffices to separate the case $k = 0$ in the summation, which leads to:

$$Var(X_{\mathcal{L}}) = \prod_{l=1}^{L}\lambda_{i_l} I(L,0) + \sum_{k=1}^{L-1}\left(\sum_{\substack{\mathcal{J}\subset\mathcal{L}\\ \#\mathcal{J}=k}} \prod_{j\in\mathcal{J}}\lambda_{i_j}^2 \prod_{l\notin\mathcal{J}}\lambda_{i_l}\right) I(L,k)$$

□

Theorem III.1 is a direct consequence of Theorem A.1 since the function $c_\delta : [a,b]^L \to \{0,1\}$ defined by

$$c_\delta(x_1,\dots,x_n) = \mathbf{1}_{\left|\max_{i\in\{1,\dots,L\}} x_i - \min_{i\in\{1,\dots,L\}} x_i\right| \leqslant \delta}, \ 0<\delta<\frac{b-a}{2}$$

satisfies Definition A.1.

B. Proof of Theorem III.2

Theorem B.1. *Under Notation and Assumptions of Theorem A.1, the two following assertion are valid:*

- *The following convergence in distribution holds:*

$$\sqrt{M}\left(\bar{m}_{\mathcal{L}} - \hat{m}_{0,\mathcal{L}}\right) \xrightarrow[M\to\infty]{\mathcal{D}} \mathcal{N}\left(0,\sigma^2\right),$$

  where

$$\sigma^2 = Var(X_{\mathcal{L}}) - (b-a)^{-1}\mathbb{E}\left[X_{\mathcal{L}}\right]^2\left(\sum_{l=1}^{L}\lambda_{i_j}^{-1}\right).$$

- *Moreover, σ^2 can be estimated by*

$$\hat{\sigma}^2 = \hat{v}\left(X_{\mathcal{L}}\right) - (b-a)^{-1}I(L,L)\prod_{l=1}^{L}\hat{\lambda}_{i_l}^2\left(\sum_{k=1}^{L}\hat{\lambda}_{i_k}^{-1}\right),$$

  where

$$\hat{v}(X_{\mathcal{L}}) = \hat{m}_{0,\mathcal{L}} + \sum_{k=1}^{L-1}\left(\sum_{\substack{\mathcal{J}\subset\mathcal{L}\\ \#\mathcal{L}=k}}\prod_{j\in\mathcal{J}}\hat{\lambda}_{i_j}^2\prod_{l\notin\mathcal{J}}\hat{\lambda}_{i_l}\right)I(L,k),$$

  and

$$\sqrt{M}\frac{(\bar{m}_{\mathcal{L}} - \hat{m}_{0,\mathcal{L}})}{\sqrt{\hat{\sigma}^2}} \xrightarrow{\mathcal{D}} \mathcal{N}\left(0,1\right).$$

Proof. An application of the Central Limit Theorem leads to:

$$\frac{1}{\sqrt{M}}\sum_{k=1}^{M}\left[\begin{pmatrix} X_{\mathcal{L}}^{(k)} \\ N_{i_1}^{(k)}\left([a,b]\right) \\ \vdots \\ N_{i_L}^{(k)}\left([a,b]\right)\end{pmatrix} - \begin{pmatrix} \mathbb{E}\left[X_{\mathcal{L}}\right] \\ \lambda_{i_1}(b-a) \\ \vdots \\ \lambda_{i_L}(b-a)\end{pmatrix}\right] \xrightarrow{\mathcal{D}} \mathcal{N}_{L+1}\left(0,\Gamma\right),$$

where Γ is the following covariance matrix:

$$\Gamma = \begin{pmatrix} Var(X_{\mathcal{L}}) & \mathbb{E}\left[X_{\mathcal{L}}\right] & \cdots & \mathbb{E}\left[X_{\mathcal{L}}\right] \\ \mathbb{E}\left[X_{\mathcal{L}}\right] & \lambda_{i_1}(b-a) & 0 & 0 \\ \vdots & 0 & \ddots & 0 \\ \mathbb{E}\left[X_{\mathcal{L}}\right] & 0 & 0 & \lambda_{i_L}(b-a)\end{pmatrix}$$

The matrix is obtained using the fact that the (N_i) are independent and from the following computation:

$$\begin{aligned}
\mathbb{E}\left[X_{\mathcal{L}} N_{i_1}\left([a,b]\right)\right] &= \mathbb{E}\left[\int_{[a,b]^{L+1}} c\left(x_1,\ldots,x_L\right)\ dN_{i_1}\left(x_1\right)\ldots dN_{i_L}\left(x_L\right) dN_{i_1}\left(y\right)\right] \\
&= \mathbb{E}\left[\int_{[a,b]^{L-1}} \left(\int_{[a,b]^{(2)}} c\left(x_1,\ldots,x_L\right)\ dN_{i_1}\left(x_1\right) dN_{i_1}\left(y\right)\right) dN_{i_2}\left(x_2\right)\ldots dN_{i_L}\left(x_L\right)\right] \\
&\quad + \mathbb{E}\left[\int_{[a,b]^{L-1}} \left(\int_{[a,b]^{(1)}} c\left(x_1,\ldots,x_L\right)\ dN_{i_1}\left(x_1\right) dN_{i_1}\left(y\right)\right) dN_{i_2}\left(x_2\right)\ldots dN_{i_L}\left(x_L\right)\right] \\
&= \lambda_{i_1}\prod_{l=1}^{L}\lambda_{i_l}\int_{[a,b]^{L+1}} c\left(x_1,\ldots,x_L\right)\ dx_1\ldots dx_L dy + \prod_{l=1}^{L}\lambda_{i_l}.I(L,0) \\
&= \lambda_{i_1}(b-a)\left(\prod_{l=1}^{L}\lambda_{i_l}\right) I(L,0) + \left(\prod_{l=1}^{L}\lambda_{i_l}\right) I(L,0) \\
&= \mathbb{E}\left[N_{i_1}\left([a,b]\right)\right]\mathbb{E}\left[X_{\mathcal{L}}\right] + \mathbb{E}\left[X_{\mathcal{L}}\right]
\end{aligned}$$

Define $g(x,u_1,\ldots,u_L) = x - \prod_{l=1}^{L} u_l (b-a)^{-L} I(L,0)$ and remark that:

$$g\left(\frac{1}{M}\sum_{k=1}^{M} X_{\mathcal{L}}^{(k)}, \frac{1}{M}\sum_{k=1}^{M} N_{i_1}^{(k)}\left([a,b]\right),\ldots,\frac{1}{M}\sum_{k=1}^{M} N_{i_L}^{(k)}\left([a,b]\right)\right) = \bar{m}_{\mathcal{L}} - \hat{m}_{0,\mathcal{L}}$$

$$g\left(\mathbb{E}\left[X_{\mathcal{L}}\right], \lambda_{i_1}(b-a),\ldots,\lambda_{i_L}(b-a)\right) = 0$$

So we have

$$\begin{aligned}
\sqrt{M}\left(\bar{m}_{\mathcal{L}} - \hat{m}_{0,\mathcal{L}}\right) = \sqrt{M}\Bigg[&g\left(\frac{1}{M}\sum_{k=1}^{M} X_{\mathcal{L}}^{(k)}, \frac{1}{M}\sum_{k=1}^{M} N_{i_1}^{(k)}\left([a,b]\right),\ldots,\frac{1}{M}\sum_{k=1}^{M} N_{i_L}^{(k)}\left([a,b]\right)\right) \\
&- g\left(\mathbb{E}\left[X_{\mathcal{L}}\right], \lambda_{i_1}(b-a),\ldots,\lambda_{i_L}(b-a)\right)\Bigg]
\end{aligned}$$

And the delta method (Casella and Berger, 2002) gives the convergence in distribution:

$$\sqrt{M}\left(\bar{m}_{\mathcal{L}} - \hat{m}_{0,\mathcal{L}}\right) \xrightarrow{\mathcal{D}} \mathcal{N}\left(0, {}^{t}D\Gamma D\right),$$

where D is the gradient of g at the point $\left(\mathbb{E}\left[X_{\mathcal{L}}\right], \lambda_{i_1}(b-a),\ldots,\lambda_{i_L}(b-a)\right)$ i.e.

$$D = \begin{pmatrix} 1 \\ -\lambda_{i_1}^{-1}\mathbb{E}\left[X_{\mathcal{L}}\right](b-a)^{-1} \\ \vdots \\ -\lambda_{i_L}^{-1}\mathbb{E}\left[X_{\mathcal{L}}\right](b-a)^{-1} \end{pmatrix}$$

We have:

$${}^{t}D\Gamma D = {}^{t}D\begin{pmatrix} Var(X_{\mathcal{L}}) - (b-a)^{-1}\mathbb{E}\left[X_{\mathcal{L}}\right]^2\left(\sum_{l=1}^{L}\lambda_{i_l}^{-1}\right) \\ \mathbb{E}\left[X_{\mathcal{L}}\right] - \mathbb{E}\left[X_{\mathcal{L}}\right] \\ \vdots \\ \mathbb{E}\left[X_{\mathcal{L}}\right] - \mathbb{E}\left[X_{\mathcal{L}}\right] \end{pmatrix} = Var(X_{\mathcal{L}}) - (b-a)^{-1}\mathbb{E}\left[X_{\mathcal{L}}\right]^2\left(\sum_{l=1}^{L}\lambda_{i_l}^{-1}\right),$$

which proves the first part of the Theorem B.1.

To get the second part, it suffices to apply Slutsky lemma (Casella and Berger, 2002). □

Once again, Theorem III.2 is a direct consequence of Theorem B.1 since the function $c_\delta : [a,b]^L \to \{0,1\}$ defined by

$$c_\delta(x_1,\dots,x_n) = \mathbf{1}_{\left|\max\limits_{i\in\{1,\dots,L\}} x_i - \min\limits_{i\in\{1,\dots,L\}} x_i\right| \leqslant \delta},\ 0<\delta<\frac{b-a}{2}$$

satisfies Definition A.1.

C. Proof of Proposition III.1

To get simpler expressions we consider n instead of L, i.e., we compute

$$I(n,k) = \int\limits_{[a,b]^{n-k}} \left(\int\limits_{[a,b]^k} \mathbf{1}_{|\max(\vee x_i, \vee y_i) - \min(\wedge x_i, \wedge y_i)| \leqslant \delta}\, dx_1 \dots dx_k \right)^2 dx_{k+1}\dots dx_n$$

where $\wedge x_i = \min\{x_i, i \in \{1,\dots,k\}\}$ and $\vee x_i = \max\{x_i, i \in \{1,\dots,k\}\}$.

To calculate $I(n,k)$, we have to first compute the inner integral

$$\Sigma = \int\limits_{[a,b]^k} \mathbf{1}_{|\max(\vee x_i, \vee y_i) - \min(\wedge x_i, \wedge y_i)| \leqslant \delta}\, dx_1 \dots dx_k$$

In order to do that let us fix some $(y_1,\dots,y_{n-k}) \in [a,b]^{n-k}$, cut the integral into pieces according to the following cases:

1. $\wedge x_i > \wedge y_i$ and $\vee x_i > \vee y_i$. Denote the integral A.
2. $\wedge x_i < \wedge y_i$ and $\vee x_i < \vee y_i$. Denote the integral B.
3. $\wedge x_i > \wedge y_i$ and $\vee x_i < \vee y_i$. Denote the integral C.
4. $\wedge x_i < \wedge y_i$ and $\vee x_i > \vee y_i$. Denote the integral D.

Since we have partitioned $[a,b]^k$ up to a null measure set, we have $\Sigma = A + B + C + D$. The computation is summarized in the next Lemma.

Lemma C.1. $\forall k \in \{0,\dots,n-1\}$,

$$A = \mathbf{1}_{|\vee y_i - \wedge y_i| \leqslant \delta} \left[\left(\min(\delta, b - \wedge y_i)\right)^k - \left(\vee y_i - \wedge y_i\right)^k \right]$$

$$B = \mathbf{1}_{|\vee y_i - \wedge y_i| \leqslant \delta} \left[\left(\min(\delta, \vee y_i - a)\right)^k - \left(\vee y_i - \wedge y_i\right)^k \right]$$

$$C = \mathbf{1}_{|\vee y_i - \wedge y_i| \leqslant \delta} \left(\vee y_i - \wedge y_i\right)^k$$

$$\begin{aligned} D \;=\; & \mathbf{1}_{|\vee y_i - \wedge y_i| \leqslant \delta} \Big[\left(\vee y_i - \wedge y_i\right)^k - \left(\min(\delta, \vee y_i - a)\right)^k \\ & + k\left(\min(\wedge y_i, b-\delta) - \max\left(\vee y_i - \delta, a\right)\right)\delta^{k-1} + \left(\max(\delta, b - \wedge y_i)\right)^k - \left(b - \wedge y_i\right)^k \Big] \end{aligned}$$

and

$$\Sigma = \mathbf{1}_{|\vee y_i - \wedge y_i| \leqslant \delta} \left[(k+1)\,\delta^k + k\left(\min(\wedge y_i, b-\delta) - \max\left(\vee y_i, a+\delta\right)\right)\delta^{k-1} \right]$$

Proof of Lemma C.1 *Proof.* Let $k \in \{2, \dots, n-1\}$. To compute A, it is sufficient to only consider the case when $x_1 = \vee x_i$, provided a multiplication by k. So we have:

$$\begin{aligned}
A &= k \int_{x_1=\vee y_i}^{b} \left(\int_{[\wedge y_i, x_1]^{k-1}} \mathbf{1}_{|x_1 - \wedge y_i| \leqslant \delta} \, dx_2 \dots dx_k \right) dx_1 \\
&= k \mathbf{1}_{|\vee y_i - \wedge y_i| \leqslant \delta} \int_{x_1=\vee y_i}^{\min(\wedge y_i + \delta, b)} \left(\int_{[\wedge y_i, x_1]^{k-1}} 1 \, dx_2 \dots dx_k \right) dx_1 \\
&= k \mathbf{1}_{|\vee y_i - \wedge y_i| \leqslant \delta} \int_{x_1=\vee y_i}^{\min(\wedge y_i + \delta, b)} (x_1 - \wedge y_i)^{k-1} \, dx_1 \\
&= \mathbf{1}_{|\vee y_i - \wedge y_i| \leqslant \delta} \left[(\min(\wedge y_i + \delta, b) - \wedge y_i)^k - (\vee y_i - \wedge y_i)^k \right] \\
&= \mathbf{1}_{|\vee y_i - \wedge y_i| \leqslant \delta} \left[(\min(\delta, b - \wedge y_i))^k - (\vee y_i - \wedge y_i)^k \right]
\end{aligned}$$

To calculate B, let us do the same with $x_1 = \wedge x_i$.

$$\begin{aligned}
B &= k \int_{x_1=\max(\vee y_i - \delta, a)}^{\wedge y_i} (\vee y_i - x_1)^{k-1} \, dx_1 \\
&= \mathbf{1}_{|\vee y_i - \wedge y_i| \leqslant \delta} \left[(\vee y_i - \max(\vee y_i - \delta, a))^k - (\vee y_i - \wedge y_i)^k \right] \\
&= \mathbf{1}_{|\vee y_i - \wedge y_i| \leqslant \delta} \left[(\min(\delta, \vee y_i - a))^k - (\vee y_i - \wedge y_i)^k \right]
\end{aligned}$$

The computation of C is clear.

$$C = \int_{[\wedge y_i, \vee y_i]^k} \mathbf{1}_{|\vee y_i - \wedge y_i| \leqslant \delta} \, dx_1 \dots dx_k = \mathbf{1}_{|\vee y_i - \wedge y_i| \leqslant \delta} (\vee y_i - \wedge y_i)^k$$

To calculate D, it is sufficient to only consider the case when $x_1 = \wedge x_i$ and $x_2 = \vee x_i$, provided a multiplication by $k(k-1)$. So we have:

$$\begin{aligned}
D &= k(k-1) \int_{x_1=a}^{\wedge y_i} \int_{x_2=\vee y_i}^{b} \left(\int_{x_1}^{x_2} \mathbf{1}_{|x_2 - x_1| \leqslant \delta} \, dx_3 \dots dx_k \right) dx_2 dx_1 \\
&= k(k-1) \mathbf{1}_{|\vee y_i - \wedge y_i| \leqslant \delta} \int_{x_1=\max(\vee y_i - \delta, a)}^{\wedge y_i} \int_{x_2=\vee y_i}^{\min(x_1 + \delta, b)} (x_2 - x_1)^{k-2} \, dx_2 dx_1 \\
&= k \mathbf{1}_{|\vee y_i - \wedge y_i| \leqslant \delta} \int_{x_1=\max(\vee y_i - \delta, a)}^{\wedge y_i} (\min(x_1 + \delta, b) - x_1)^{k-1} - (\vee y_i - x_1)^{k-1} \, dx_1 \\
&= \mathbf{1}_{|\vee y_i - \wedge y_i| \leqslant \delta} \Big[(\vee y_i - \wedge y_i)^k - (\vee y_i - \max(\vee y_i - \delta, a))^k \\
&\quad + k (\min(\wedge y_i, b - \delta) - \max(\vee y_i - \delta, a)) \delta^{k-1} + (b - \min(\wedge y_i, b - \delta))^k - (b - \wedge y_i)^k \Big] \\
&= \mathbf{1}_{|\vee y_i - \wedge y_i| \leqslant \delta} \Big[(\vee y_i - \wedge y_i)^k - (\min(\delta, \vee y_i - a))^k \\
&\quad + k (\min(\wedge y_i, b - \delta) - \max(\vee y_i - \delta, a)) \delta^{k-1} + (\max(\delta, b - \wedge y_i))^k - (b - \wedge y_i)^k \Big]
\end{aligned}$$

To check the given expression of Σ it suffices to remark that

$$(\min(\delta, b - \wedge y_i))^k + (\max(\delta, b - \wedge y_i))^k = \delta^k + (b - \wedge y_i)^k$$

and

$$\max\left(\vee y_i - \delta, a\right) = \max\left(\vee y_i, a + \delta\right) - \delta$$

□

Remark We took $k \geqslant 2$ to avoid the division by 0 (divisions by k and $k-1$). However in these particular cases an easy computation shows that the general formula stands in these cases. $\Sigma = \mathbf{1}_{|\vee y_i - \wedge y_i| \leqslant \delta}$ when $k = 0$ and $\Sigma = \mathbf{1}_{|\vee y_i - \wedge y_i| \leqslant \delta} \left[\min(\wedge y_i + \delta, b) - \max\left(\vee y_i - \delta, a\right)\right]$ when $k = 1$.

Following of the proof of Proposition III.1

It remains to calculate $I(n,k) = \int\limits_{[a,b]^{n-k}} \Sigma\left(y_1, \ldots, y_{n-k}\right)^2 \, dy_1 \ldots dy_{n-k}$. In order to do that, cut the integral into pieces according to the following cases:

1. $\vee y_i \leqslant a + \delta$. In this case, $\Sigma = \delta^{k-1}\left[\delta + k\left(\wedge y_i - a\right)\right]$, and denote the integral Y.

2. $\wedge y_i \geqslant b - \delta$. In this case, $\Sigma = \delta^{k-1}\left[\delta + k\left(b - \vee y_i\right)\right]$, and denote the integral Z.

3. $\vee y_i \geqslant a + \delta$ and $\wedge y_i \leqslant b - \delta$. In this case, $\Sigma = \mathbf{1}_{|\vee y_i - \wedge y_i| \leqslant \delta} \delta^{k-1}\left[(k+1)\,\delta - k\left(\vee y_i - \wedge y_i\right)\right]$, and denote the integral W.

These three cases are distinct because $\delta < \frac{b-a}{2}$.

Lemma C.2. $\forall k \in \{0, \ldots, n-1\}$,

$$Y = Z = C\left(n,k\right)\delta^{n+k}$$

where

$$C\left(n,k\right) = \left(n-k\right)\frac{\left(k+1\right)^{n-k+2}}{k^{n-k}}\int_0^{\frac{k}{k+1}} t^{n-k-1}\left(1-t\right)^2 \, dt,$$

and

$$W = f\left(n,k\right)\left(b-a\right)\delta^{n+k-1} - \left[f\left(n,k\right) + g\left(n,k\right)\right]\delta^{n+k}$$

where

$$f\left(n,k\right) = \left(n-k\right)\left(k+1\right)^2 - 2\left(n-k-1\right)k\left(k+1\right) + \frac{\left(n-k\right)\left(n-k-1\right)}{\left(n-k+1\right)}k^2$$

and

$$g\left(n,k\right) = \left(k+1\right)^2 - 2\frac{\left(n-k-1\right)k\left(k+1\right)}{\left(n-k+1\right)} + \frac{\left(n-k\right)\left(n-k-1\right)}{\left(n-k+1\right)\left(n-k+2\right)}k^2$$

Proof of Lemma C.2 *Proof.* Let $k \in \{0, \dots, n-2\}$. To calculate Y, it is sufficient to only consider the case when $y_1 = \wedge y_i$, provided a multiplication by $(n-k)$. So we have:

$$\begin{aligned} Y &= \int\limits_{\vee y_i \leqslant a+\delta} \Sigma^2 \, dy_1 \dots dy_{n-k} \\ &= (n-k)\,\delta^{2k-2} \int\limits_{y_1=a}^{a+\delta} \left(\int\limits_{[y_1, a+\delta]^{n-k-1}} [\delta + k\,(y_1 - a)]^2 \, dy_2 \dots dy_{n-k} \right) dy_1 \\ &= (n-k)\,\delta^{2k-2} \int\limits_{y_1=a}^{a+\delta} (a+\delta-y_1)^{n-k-1} \, [\delta + k\,(y_1 - a)]^2 \, dy_1 \end{aligned}$$

Defining the variable $u = a + \delta - y_1$, leads to

$$\begin{aligned} Y &= (n-k)\,\delta^{2k-2} \int_0^{\delta} u^{n-k-1} \, [\delta + k\,(\delta - u)]^2 \, du \\ &= (n-k)\,\delta^{2k-2} \int_0^{\delta} u^{n-k-1} \, [(k+1)\,\delta - ku]^2 \, du \end{aligned}$$

And by defining the variable $t = \frac{ku}{(k+1)\delta}$, we have

$$\begin{aligned} Y &= (n-k)\,\delta^{2k-2} \int_0^{\frac{k}{k+1}} \left(\frac{(k+1)\,\delta t}{k} \right)^{n-k-1} (k+1)^2\,\delta^2\,(1-t)^2\,\frac{(k+1)\,\delta}{k} dt \\ &= (n-k)\,\delta^{n+k} \frac{(k+1)^{n-k+2}}{k^{n-k}} \int_0^{\frac{k}{k+1}} t^{n-k-1}\,(1-t)^2 \, dt \end{aligned}$$

The computation of Z can be done in the same way by inverting the roles of a and b on the one hand and of $\wedge y_i$ and $\vee y_i$ on the other hand. This leads to $Z = Y$. To calculate W, it is sufficient to only consider the case when $y_1 = \wedge y_i$ and $y_2 = \vee y_i$, provided a multiplication by $(n-k)\,(n-k-1)$. So we have:

$$\begin{aligned} W &= (n-k)\,(n-k-1)\,\delta^{2k-2} \\ &\quad \int\limits_{y_1=a}^{b-\delta} \int\limits_{y_2=\max(y_1, a+\delta)}^{b} \left(\int \mathbf{1}_{|y_2-y_1| \leqslant \delta} \, [(k+1)\,\delta - k\,(y_2 - y_1)]^2 \, dy_3 \dots dy_{n-k} \right) dy_2 dy_1 \\ &= (n-k)\,(n-k-1)\,\delta^{2k-2} \\ &\quad \int\limits_{y_1=a}^{b-\delta} \int\limits_{y_2=\max(y_1, a+\delta)}^{y_1+\delta} \left((y_2-y_1)^{n-k-2} \left[(k+1)^2\,\delta^2 - 2k\,(k+1)\,\delta\,(y_2-y_1) + k^2\,(y_2-y_1)^2 \right] \right) dy_2 dy_1 \end{aligned}$$

which leads to

$$\begin{aligned} W &= (n-k)\,\delta^{2k-2} \int_a^{b-\delta} (k+1)^2\,\delta^2 \left[\delta^{n-k-1} - (\max(y_1, a+\delta) - y_1)^{n-k-1} \right] dy_1 \\ &\quad - (n-k-1)\,\delta^{2k-2} \int_a^{b-\delta} 2k\,(k+1)\,\delta \left[\delta^{n-k} - (\max(y_1, a+\delta) - y_1)^{n-k} \right] dy_1 \\ &\quad + \frac{(n-k)\,(n-k-1)}{(n-k+1)} \delta^{2k-2} \int_a^{b-\delta} k^2 \left[\delta^{n-k+1} - (\max(y_1, a+\delta) - y_1)^{n-k+1} \right] dy_1 \end{aligned}$$

First, let us calculate W_1 the integral between $a+\delta$ and $b-\delta$.

$$W_1 = \delta^{2k-2} \int_{a+\delta}^{b-\delta} \left[(n-k)(k+1)^2 - 2(n-k-1)k(k+1) + \frac{(n-k)(n-k-1)}{(n-k+1)} k^2 \right] \delta^{n-k+1} \, dy_1$$
$$= (b-a-2\delta) f(n,k) \delta^{n+k-1}$$
$$= f(n,k)(b-a)\delta^{n+k-1} - 2f(n,k)\delta^{n+k}$$

where $f(n,k) = (n-k)(k+1)^2 - 2(n-k-1)k(k+1) + \frac{(n-k)(n-k-1)}{(n-k+1)} k^2$.

Second, let us calculate W_2 the integral between a and $a+\delta$.

$$W_2 = \delta^{2k-2} \left[\int_a^{a+\delta} f(n,k) \delta^{n-k+1} \, dy_1 \right.$$
$$- (n-k)(k+1)^2 \delta^2 \int_a^{a+\delta} (a+\delta-y_1)^{n-k-1} \, dy_1$$
$$+2(n-k-1)k(k+1)\delta \int_a^{a+\delta} (a+\delta-y_1)^{n-k} \, dy_1$$
$$\left. - \frac{(n-k)(n-k-1)}{(n-k+1)} k^2 \int_a^{a+\delta} (a+\delta-y_1)^{n-k+1} \, dy_1 \right]$$

Let us denote $g(n,k) = (k+1)^2 - 2\frac{(n-k-1)k(k+1)}{(n-k+1)} + \frac{(n-k)(n-k-1)}{(n-k+1)(n-k+2)} k^2$ which leads to

$$W_2 = f(n,k)\delta^{n+k} - g(n,k)\delta^{n+k}$$

So,

$$W = f(n,k)(b-a)\delta^{n+k-1} - [f(n,k) + g(n,k)]\delta^{n+k}$$

□

Remark We took $k \leqslant n-2$ to avoid division by 0 (divisions by $n-k$ and $n-k-1$). However, in the particular case where $k = n-1$, the computations of Y and Z are also valid because we only divide by $n-k=1$. Moreover the case "$\vee y_i \geqslant a+\delta$ and $\wedge y_i < a+\delta$" is impossible when $k = n-1$ because $\vee y_i = \wedge y_i$. An easy computation shows that

$$W_1 = n^2(b-a)\delta^{2n-2} - 2n^2\delta^{2n-1}$$

and that the general formula of W stands in this case.

Conclusion of the proof of Proposition III.1

It only remains to calculate

$$\int_0^{\frac{k}{k+1}} t^{n-k-1}(1-t)^2 \, dt = \left(\frac{k}{k+1}\right)^{n-k} \left[\frac{1}{n-k} - \frac{2k}{(k+1)(n-k+1)} + \frac{k^2}{(k+1)^2(n-k+2)} \right]$$

So we have the stated formula

$$I(n,k) = W + Y + Z = f(n,k)(b-a)\delta^{n+k-1} - [f(n,k) + g(n,k) - 2C(n,k)]\delta^{n+k}$$

In the stated result we just used the software Mathematica in order to simplify the expressions. This simplification leads to

$$f\left(n,k\right)=\frac{k\left(k+1\right)+n\left(n+1\right)}{n-k+1}$$

and

$$f\left(n,k\right)+g\left(n,k\right)-2C\left(n,k\right)=h\left(n,k\right)=\frac{-k^3+k^2(2+n)+k(5+2n-n^2)+n^3+2n^2-n-2}{(n-k+2)(n-k+1)}$$

References

Abeles, Moshe. (1982). Quantification, smoothing, and confidence limits for single-units' histograms. *Journal of Neuroscience Methods* **5**(4), 317–325.

Aertsen, A. M. H. J., Gerstein, G. L., Habib, M. K. and Palm, G. (1989). Dynamics of Neuronal Firing Correlation: Modulation of "Effective Connectivity". *Journal of Neurophysiology* **61**, 900–917.

Albert, Mélisande, Bouret, Yann, Fromont, Magalie and Reynaud-Bouret, Patricia. (2014). Bootstrap and permutation tests of independence for point processes. prepublication on HAL.

Archer, Evan W, Park, Il M and Pillow, Jonathan W. (2013). Bayesian entropy estimation for binary spike train data using parametric prior knowledge. In: *Advances in Neural Information Processing Systems*. pp. 1700–1708.

Barlow, H. B. (1972). Single Units and Sensation: A Neuron Doctrine for Perceptual Psychology? *Perception* **1**, 371–394.

Benjamini, Yoav and Hochberg, Yosef. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J. Roy. Statist. Soc. Ser. B* **57**(1), 289–300.

Casella, George and Berger, Roger. (2002). *Statistical Inference*. Duxbury.

Chornoboy, E.S., Schramm, L.P. and Karr, A.F. (1988). Maximum likelihood identification of neural point process systems. *Biological Cybernetics* **59**(4-5), 265–275.

Daley, D. J. and Vere-Jones, D. (2003). *An introduction to the theory of point processes. Vol. I*, Second edition., Probability and its Applications (New York). New York: Springer-Verlag. Elementary theory and methods.

Gerstein, George L, Aertsen, Ad M *and others*. (1985). Representation of cooperative firing activity among simultaneously recorded neurons. *J Neurophysiol* **54**(6), 1513–1528.

Gerstein, George L. and Perkel, Donald H. (1969). Simultaneously Recorded Trains of Action Potentials: Analysis and Functional Interpretation. *Science* **164**(3881), 828–830.

Grammont, F. and Riehle, Alexa. (2003). Spike synchronization and firing rate in a population of motor cortical neurons in relation to movement direction and reaction time. *Biological Cybernetics* **88**(5), 360–373.

Grün, Sonja. (1996). Unitary joint events in multiple neuron spiking activity: detection, significance, and interpretation [Ph.D. Thesis]. .

Grün, Sonja, Diesmann, Markus and Aertsen, Ad. (2002). Unitary Events in Multiple Single-Neuron Spiking Activity: I. Detection and Significance. *Neural Computation* **14**(1), 43–80.

Grün, Sonja, Diesmann, Markus, Grammont, Franck, Riehle, Alexa and Aertsen, Ad. (1999). Detecting unitary events without discretization of time. *Journal of neuroscience methods* **94**(1), 67–79.

Hebb, Donald O. (1949). *The Organization of Behavior: A Neuropsychological Theory*. New York: Wiley.

Kingman, J. F. C. (1993). *Poisson processes*, Volume 3, Oxford Studies in Probability. New York: The Clarendon Press Oxford University Press. Oxford Science Publications.

Krumin, Michael, Reutsky, Inna and Shoham, Shy. (2010). Correlation-based analysis and generation of multiple spike trains using Hawkes models with an exogenous input. *Frontiers in computational neuroscience* **4**.

Lehmann, E.L. and Romano, Joseph P. (2005). *Testing Statistical Hypotheses*, 3rd edition. Springer-Verlag New York Inc.

Lewis, P. A. W. and Shedler, G. S. (1979). Simulation of nonhomogeneous Poisson processes by thinning. *Naval Res. Logist. Quart.* **26**(3), 403–413.

Ogata, Yosihiko. (1981). On Lewis simulation method for point processes. *IEEE Transactions on Information Theory* **27**(1), 23–30.

Palm, G. (1990). Cell assemblies as a guideline for brain research. *Concepts in Neuroscience* **1**(1), 133–147.

Perkel, Donald H., Gerstein, George L. and Moore, George P. (1967). Neuronal Spike Trains and Stochastic Point Processes: II. Simultaneous Spike Trains. *Biophysical Journal* **7**(4), 419–440.

Pernice, Volker, Staude, Benjamin, Cardanobile, Stefano and Rotter, Stefan. (2011). How structure determines correlations in neuronal networks. *PLoS computational biology* **7**(5), e1002059.

Pernice, Volker, Staude, Benjamin, Cardanobile, Stefano and Rotter, Stefan. (2012). Recurrent interactions in spiking networks with arbitrary topology. *Physical Review E* **85**(3), 031916.

Pipa, Gordon, Diesmann, Markus and Grün, Sonja. (2003). Significance of joint-spike events based on trial-shuffling by efficient combinatorial methods. *Complexity* **8**(4), 79–86.

Pipa, Gordon and Grün, Sonja. (2003). Non-parametric significance estimation of joint-spike events by shuffling and resampling. *Neurocomputing* **52**, 31–37.

Reimer, Imke CG, Staude, Benjamin, Ehm, Werner and Rotter, Stefan. (2012). Modeling and analyzing higher-order correlations in non-Poissonian spike trains. *Journal of neuroscience methods* **208**(1), 18–33.

Riehle, Alexa, Grammont, Franck, Diesmann, Markus and Grün, Sonja. (2000). Dynamical changes and temporal precision of synchronized spiking activity in monkey motor cortex during movement preparation. *Journal of Physiology-Paris* **94**(5), 569–582.

Riehle, Alexa, Grammont, Franck and MacKay, William A. (2006). Cancellation of a planned movement in monkey motor cortex. *Neuroreport* **17**(3), 281–285.

Sakurai, Yoshio. (1999). How do cell assemblies encode information in the brain? *Neuroscience & Biobehavioral Reviews* **23**(6), 785–796.

Shinomoto, Shigeru. (2010). Estimating the Firing Rate. In: Grün, Sonja and Rotter, Stefan (editors), *Analysis of Parallel Spike Trains*, Volume 7, Springer Series in Computational Neuroscience. Springer US, pp. 21–35.

Singer, Wolf. (1993). Synchronization of Cortical Activity and its Putative Role in Information Processing and Learning. *Annual Review of Physiology* **55**, 349–374.

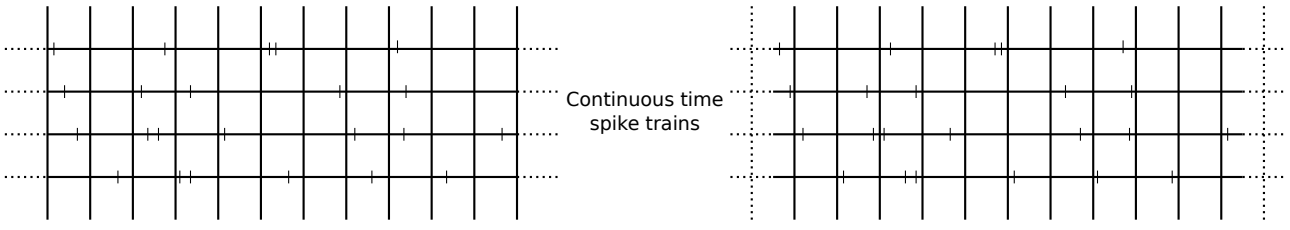
Tuleau-Malot, Christine, Rouis, Amel, Grammont, Franck and Reynaud-Bouret, Patricia. (2014). Multiple Tests Based on a Gaussian Approximation of the Unitary Events Method with delayed coincidence count. appearing in *Neural Computation* **26:7**.

von der Malsburg, Christoph. (1981). The Correlation Theory of Brain Function. *Internal Report* 81-2, Department of Neurobiology, Max-Planck-Institute for Biophysical Chemistry, Göttingen, Germany.

A : Simultaneously recorded neurons

Bin	1	2	3	...	S-1	S
Neuron 1	1	1	0	...	0	1
Neuron 2	0	1	0	...	1	1
Neuron 3	0	1	0	...	0	1
Neuron 4	0	0	1	...	1	1
$\mathcal{L}_w$	$\{1\}$	$\{1,2,3\}$	$\{4\}$	...	$\{2,4\}$	$\{1,2,3,4\}$

B : Discretization of spike trains



1	0	1	0	0	1	0	0	1	0	0
1	0	1	1	0	0	1	0	1	0	0
1	0	1	0	1	0	0	1	1	0	1
0	1	0	1	0	1	0	1	0	1	0

Discretized
spike trains

1	0	0	1	0	1	0	0	1	0	0	0
1	0	1	1	0	0	0	1	1	0	0	0
0	1	1	1	1	0	0	1	1	0	0	1
0	0	1	1	0	0	1	0	1	1	0	0

C : Delayed coincidence count

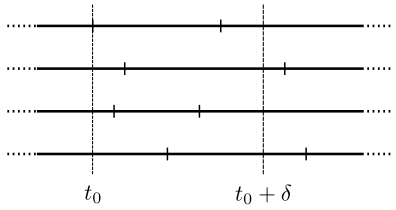


Figure 1. In **A**, 4 parallel binary processes of length S are displayed. At each time step, the constellation and its corresponding subset of $\{1,2,3,4\}$ are given. For instance, the constellation associated to the first bins is the vector $(1,0,0,0)$ and the corresponding subset is $\{1\}$. In **B**, illustration of the UE method with two different choices of bins of the same size (the results are different, for example the constellation full of 1s is present in the second case and not in the first one). In **C**, an illustration of the way delayed coincidence count is computed. The subset $\mathcal{L}$ being fixed, coincidences can be distinguished with respect to the minimal spike time of the coincidence. Indeed, it suffices to consider each spike time t_0 of every neuron of $\mathcal{L}$, and once t_0 is fixed, to count the spike times of the other neurons of $\mathcal{L}$ between t_0 and $t_0 + \delta$. For instance, in the figure, if $\mathcal{L} = \{1,2,3\}$, there are $2 \times 1 = 2$ coincidences with t_0 as the minimal spike time. If $\mathcal{L} = \{1,2,4\}$, there is $1 \times 1 = 1$ coincidence.

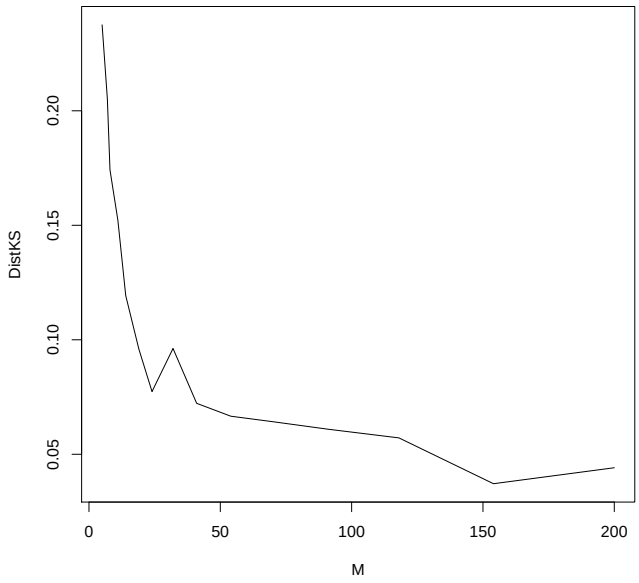


Figure 2. Under Framework $\mathbf{F}_1$. Evolution (under the independence assumption) of the Kolmogorov distance (in function of the number of trials) averaged on 1000 simulations between the estimated distribution of the test statistics and the standard Gaussian distribution function.

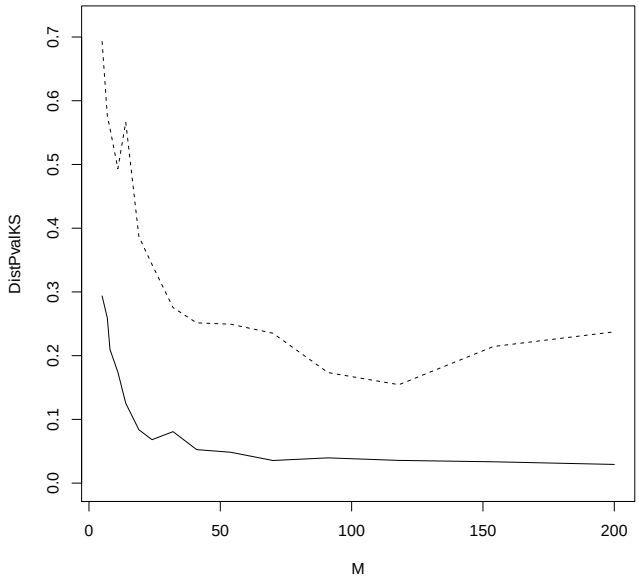


Figure 3. Under Framework $\mathbf{F}_1$. Evolution (under the independence assumption) of the Kolmogorov distance averaged on 1000 simulations between the repartition of the p-values and the uniform distribution with respect to the number of trials. The plain line stands for our test and the dashed line for the original UE one.

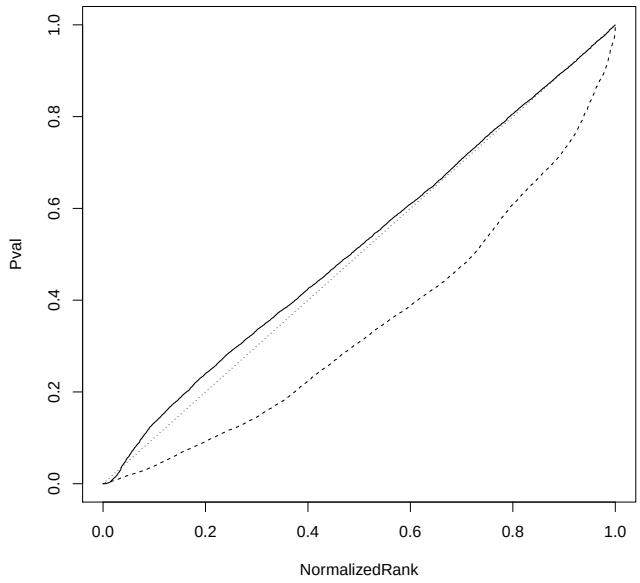


Figure 4. Under Framework $\mathbf{F}_1$. Graphs of the sorted 1000 p-values (under the independence assumption and for 50 trials) in function of their normalized rank under $\mathcal{H}_0$. The plain line stands for our test, the dashed line for the original UE one and the dotted line for the uniform distribution.

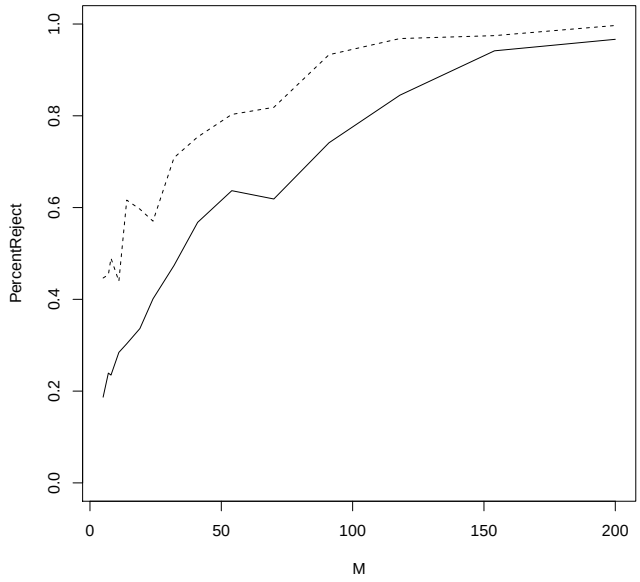


Figure 5. Under Framework $\mathbf{F}_2$. Illustration of the power of the test, for a level 0.05. The curves represent the evolution, with respect to the number of trials, of the percentage (averaged on 1000 simulations) of the rejection of the independence assumption when there is a dependence structure (induced by an injection model) between neurons. The plain line stands for our test and the dashed line for the original UE one.

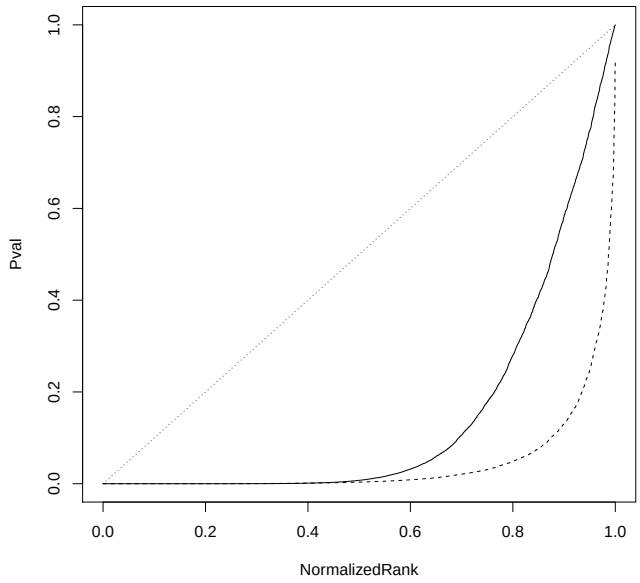


Figure 6. Under Framework $\mathbf{F}_2$. Graphs of the sorted 1000 p-values for dependent Poisson processes (50 trials). The plain line stands for our test, the dashed line for the original UE one and the dotted line for the uniform distribution.

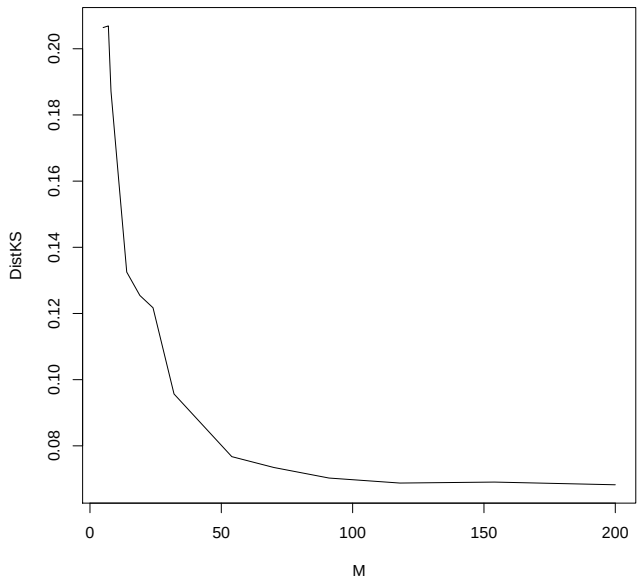


Figure 7. Under Framework $\mathbf{F}_3$. Evolution (under the independence assumption) of the Kolmogorov distance (in function of the number of trials) averaged on 1000 simulations between the estimated distribution of the test statistics and the standard Gaussian distribution function.

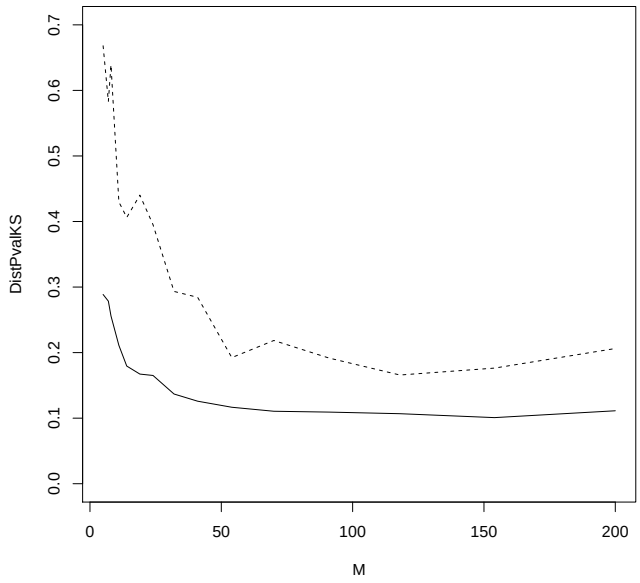


Figure 8. Under Framework $\mathbf{F}_1$. Evolution (under the independence assumption) of the Kolmogorov distance averaged on 1000 simulations between the repartition of the p-values and the uniform distribution with respect to the number of trials. The plain line stands for our test and the dashed line for the original UE one.

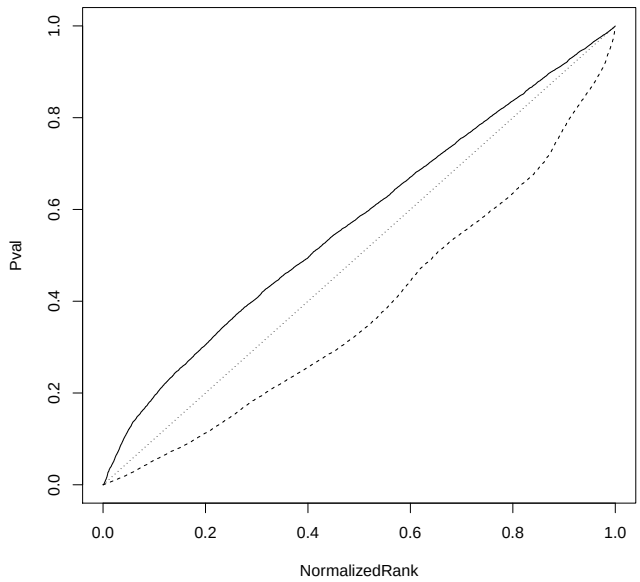


Figure 9. Under Framework $\mathbf{F}_3$. Graphs of the sorted 1000 p-values (under the independence assumption and for 50 trials) in function of their normalized rank under $\mathcal{H}_0$. The plain line stands for our test, the dashed line for the original UE one and the dotted line for the uniform distribution.

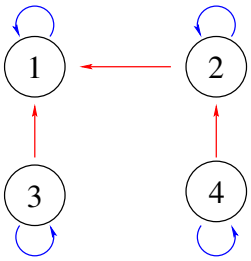


Figure 10. Local independence graph. An arrow means a non null interaction function. Blue arrow means inhibition and red arrow means excitation.

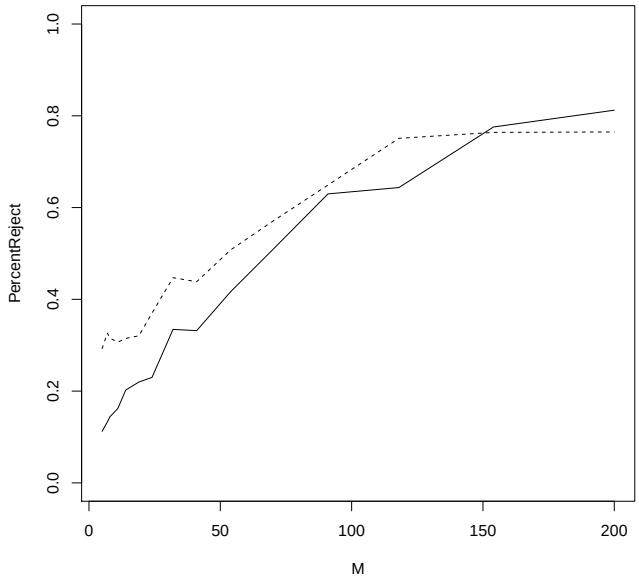


Figure 11. Under Framework $\mathbf{F}_4$. Illustration of the power of the test, for a level 0.05. The curves represent the evolution, with respect to the number of trials, of the percentage (averaged on 1000 simulations) of the rejection of the independence assumption when there is a dependence structure (presented in Figure 10) between neurons. The plain line stands for our test and the dashed line for the original UE one.

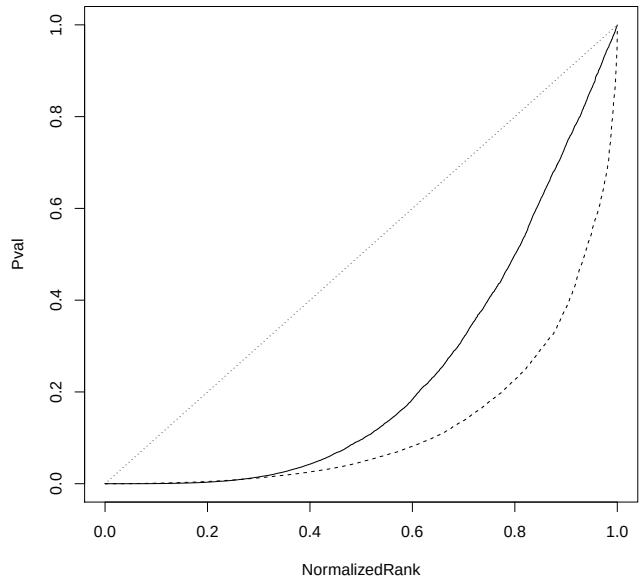


Figure 12. Under Framework $\mathbf{F}_4$. Graphs of the sorted 1000 p-values for dependent Hawkes processes (see Figure 10), for 50 trials. The plain line stands for our test, the dashed line for the original UE one and the dotted line for the uniform distribution.

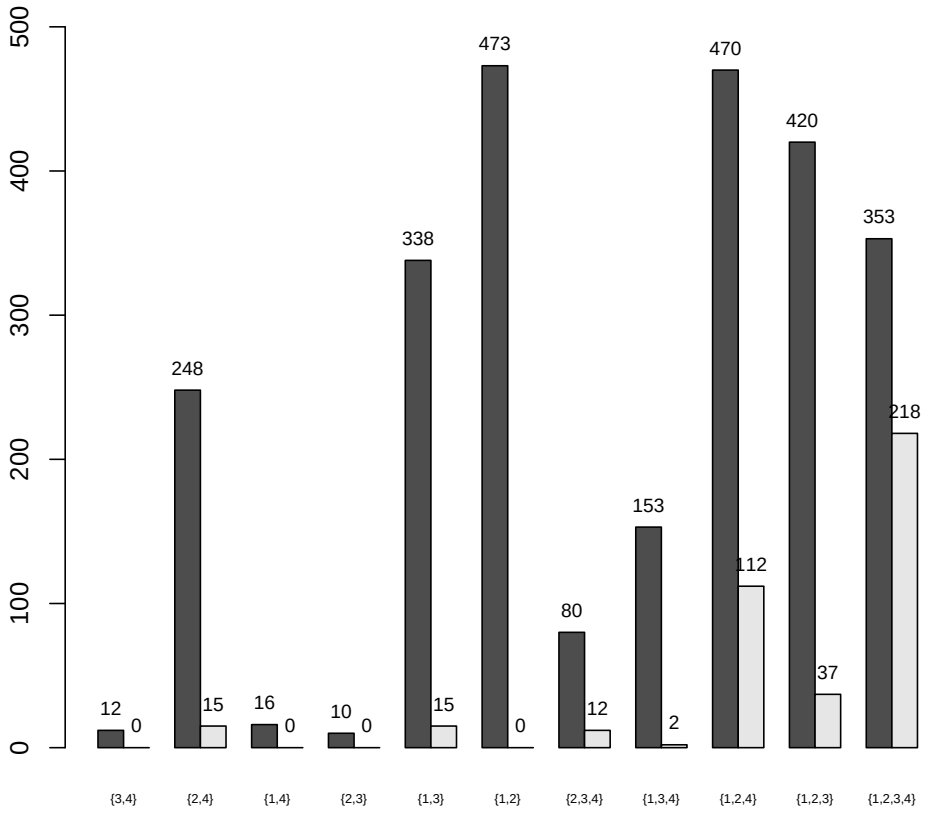


Figure 13. Under Framework $\mathbf{F}_4$. Number of dependence detection (among 1000 simulations) for each pattern. Grey for our test, white for the original UE method.

Figure 14. Evolution of the synchronization between neurons. The lines indicate the subset for which our test detects dependence. Here we detect an excess of coincidences between neurons $\{1, 3, 4\}$ and $\{3, 4\}$

□