



HAL
open science

**Separation rates for the detection of synchronization of
interacting point processes in a mean field frame.
Application to neuroscience.**

Josué Tchouanti, Eva Löcherbach, Patricia Reynaud-Bouret, Etienne Tanré

► **To cite this version:**

Josué Tchouanti, Eva Löcherbach, Patricia Reynaud-Bouret, Etienne Tanré. Separation rates for the detection of synchronization of interacting point processes in a mean field frame. Application to neuroscience.. 2024. hal-04653620

HAL Id: hal-04653620

<https://inria.hal.science/hal-04653620>

Preprint submitted on 19 Jul 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License

SEPARATION RATES FOR THE DETECTION OF SYNCHRONIZATION OF INTERACTING POINT PROCESSES IN A MEAN FIELD FRAME. APPLICATION TO NEUROSCIENCE.

JOSUÉ TCHOUANTI¹, ÉVA LÖCHERBACH², PATRICIA REYNAUD-BOURET³, AND ETIENNE TANRÉ⁴

ABSTRACT. Permutation tests have been proposed by Albert et al. (2015) to detect dependence between point processes, modeling in particular spike trains, that is the time occurrences of action potentials emitted by neurons. Our present work focuses on exhibiting a criterion on the separation rate to ensure that the Type II errors of these tests are controlled non asymptotically. This criterion is then discussed in two major models in neuroscience: the jittering Poisson model and Hawkes processes having M components interacting in a mean field frame and evolving in stationary regime. For both models, we obtain a lower bound of the size n of the sample necessary to detect the dependency between two neurons.

Keywords and phrases: *Permutation Test, Detection of synchronization, Hawkes Processes, Mean field interactions, Poisson Processes.*

MSC2020 subject classification: 62G10, 62M07, 62F40

1. INTRODUCTION

Neurons in the brain form a vast and dense directed network, in which each vertex (the neurons) emits action potentials that excite or inhibit downstream vertices. When recordings are performed, neurobiologists are usually able to track the spikes of a few neurons, that is, the times at which a neuron emits an action potential. The set of spikes of a given neuron is commonly called a *spike train*. The neurons that are recorded in a given region often show synchronization during specific periods depending on the cognitive process at play, even if they might be not physically directly connected. Here, *synchronization* means that the spikes of these neurons seem to appear "at the same time" during a given period. This phenomenon varies over time and usually marks cognitive events of importance, so that synchronization is usually thought a key element to understand the neural code, that is the way the brain encodes information [37, 35].

Let us formalize the problem in the case of two neurons emitting two spike trains, modeled by two point processes X^1 and X^2 . We reformulate the neuroscience problem of synchronization detection as a testing problem where one wants to reject the null hypothesis " X^1 is independent from X^2 " by counting the number of coincidences. Here by number of coincidences we mean the number of couples of spikes (one per neuron) that appear almost at the same time. Then the test rejects the null hypothesis if the number of coincidences is too large. Despite the fact that the synchronization is the result of the physical connectivity hidden behind, neurobiologists usually separate completely the problem of detecting synchronization from the problem of understanding how the network produces such phenomena. The above test is therefore usually applied by computing a reference distribution under the null hypothesis, which is done either by assuming that the spike trains have a known distribution (Poisson [38, 32, 21] or Integrate-and-Fire [36]) or by a bootstrap distribution [31, 2, 3]. In particular, it has been proved that permutation tests (based on permutation of i.i.d. trials) lead to very good results in practice because of their non asymptotic properties [2].

¹NEUROMOD INSTITUTE, UNIVERSITÉ CÔTE-D'AZUR

²SAMM, EA 4543 ET FR FP2M 2036 CNRS, UNIVERSITÉ PARIS 1 PANTHÉON SORBONNE

³UNIVERSITÉ CÔTE D'AZUR, CNRS, LJAD, FRANCE

⁴UNIVERSITÉ CÔTE D'AZUR, INRIA, CNRS, LJAD, FRANCE

E-mail addresses: Josue.Tchouanti-Fotso@univ-cotedazur.fr, josue.tchouanti_fotso@ens-lyon.fr, Eva.Locherbach@univ-paris1.fr, Patricia.Reynaud-Bouret@univ-cotedazur.fr, Etienne.Tanre@inria.fr.

In the present work, we investigate the non asymptotic separation rate of these permutation tests, focusing on two main examples in neuroscience. Our main goal is to understand what the fact that the test rejects the null hypothesis and detects synchronization reveals on the underlying link between the processes.

The first model we consider is the injection Poisson model which is a classic model of synchronization [10, 21]. For this model, we are in particular interested in the precise relationship between the minimal number of trials and the injection strength which are necessary to detect synchronization.

The second model is a homogeneous network of Hawkes processes. This case is of particular relevance since it is the first time that one tries to mathematically link the shape of the hidden network with the synchronization detection. Here we envision the case of a homogeneous network of exponential Hawkes processes of size M . The integer M represents the size of the relevant network at play, the two recorded neurons being just two neurons embedded in this network. We study more specifically the particular case of an exponential memory function.

A classical result by Delattre et al. [12] states that such a network has a mean-field limit when M tends to infinity, meaning that in the mean field limit, the two recorded neurons are independent, and therefore the corresponding spike trains satisfy the null hypothesis. Our goal is to precisely establish the link between M and the minimal number of trials n , to be able to detect synchronization. We show that n has to be larger than M^2 to detect the dependence between the two neurons. The implication for neuroscience is that if M is too large then one will never detect synchronization, despite the fact that the neurons are physically connected, because we will never have enough trials to do so. In this sense, we can do as if the neurons were independent, despite the fact that they are embedded in a common network.

Separation rates for general independence tests based on permutation have been recently investigated by Kim et al. [27] (see also [1] for the case of point processes). We derive from [27] a general bound with precise variances that improves the one of [1]. The main purpose of the remainder of the paper is to interpret these bounds in terms of our model parameters, within the two main examples we have in mind, coming from neuroscience.

If the injection Poisson model has been rarely investigated in the mathematical literature from a probabilistic point of view, probably because of its simplicity, stochastic systems of interacting particles in a mean field frame and their associated non-linear limit processes have been extensively studied during the last three or four decades. To build a statistical theory within this frame is the goal of a relatively new research direction (with the exception of [26]) that has been mainly explored with regards to estimation theory, and most often in the framework of interacting diffusion models. One can cite in particular the work of Della Maestra and Hoffmann [13, 14], which is devoted to an asymptotic estimation theory in the large population setting when N interacting diffusive particles evolve in a mean field frame and are observed continuously in time over a fixed time interval $[0, T]$, both in a parametric and in a non-parametric setting. See also [4] who deals with the parameter estimation of both drift and diffusion coefficient for discretely observed interacting diffusions, in the asymptotic where the step size tends to 0 and the number of particles to infinity. In [8], the authors have conducted a related study, yet in a semi parametric setting. Finally, in a series of two papers, Genon-Catalot and Laredo deal with the estimation problem in the limit model of a non-linear in the sense of McKean-Vlasov diffusion model, in the small noise regime [18, 19].

The statistical theory of systems of interacting point processes in a mean-field frame is less developed, but we can mention [5] devoted to the estimation of the underlying parameters in systems of interacting Hawkes processes, as well as [11] which proposes a study in a very particular framework where N point processes are observed during a fixed time interval, interacting on an Erdős-Renyi random graph, and one wants to recover the underlying parameters of the linear Hawkes process as well as the connection probability. To the best of our knowledge no test theory has been developed in either of the two frameworks (interacting diffusions or interacting point processes). Injection models in testing theory have been investigated slightly in [2, 3, 38] but mainly from a level (Type I error) point of view.

The goal of our paper is to understand in both situations (injection Poisson model and Hawkes processes in a mean-field frame) the *separation rate*. This separation rate is a measure of the

distance at which a certain distribution, \mathcal{P} – the joint distribution of the two spike trains X^1 and X^2 – has to be from the null hypothesis to be rejected by the test with large probability. The “measure” we use for this purpose is the difference between the expectation of the coincidence counts under \mathcal{P} and the expectation under $\mathcal{P}_1 \otimes \mathcal{P}_2$, where \mathcal{P}_1 and \mathcal{P}_2 are the marginals of \mathcal{P} . Bounds on these separation rates can in turn be interpreted as a minimal number of trials to detect a given dependence.

Our approach to obtain the separation rates is non asymptotic, in the line of [6, 7, 15, 16, 17]. The rates that we are getting are parametric in n , but the interesting feature is their dependence on the other important parameters of the problem (like M , the size of the hidden homogeneous network, or the strength of injection), and this is why the non asymptotic approach on the separation rates is important.

We treat the Poisson injection model in the most general way. However, concerning homogeneous networks of spike trains, we consider only a very particular case (namely exponential Hawkes processes). This is due to the fact that the general separation bounds we derive involve the computation of variances of U-statistics based on coincidences counts under \mathcal{P} , which involves the computation of cumulants of order 4 for the corresponding Hawkes process. At this point, we leveraged the result of Jovanović et al. [25] which gives an algorithmic way to compute these quantities. Related results are due to Privault [33] but in a more restricted framework, that can not be applied in our case. Most importantly, we show that all cumulants are of size $1/M$, and we study how these cumulants depend on the parameters of the exponential kernel and the window size of the coincidence count. We also give some more general results on cumulants, that might have a general interest per se: they are ready to use formulae for the cumulants of Hawkes processes, that are an application of the result in [25].

The rest of the article is structured as follows. In Section 2, we formulate our main results, that is the separation rates of the test in both settings (injection Poisson and homogeneous exponential Hawkes processes). In Section 3, we illustrate the sharpness of our bounds on simulations. In Section 4, we sketch some proofs and provide the ready-to-use formulae on Hawkes cumulants. Full proofs containing all computations are postponed to the appendix.

2. MAIN RESULTS

2.1. Presentation of the problem. Let us start by formulating the problem of interest. We denote by \mathcal{X} the set of finite increasing sequences of spiking times $0 \leq \tau_1 < \dots < \tau_\ell \leq T$ for $\ell \geq 0$ on a fixed time interval $[0, T]$. A point process X has random values in \mathcal{X} , and its associated point measure is denoted dX_t . We consider a probability distribution \mathcal{P} on \mathcal{X}^2 and denote by \mathcal{P}^1 and \mathcal{P}^2 its first and second marginals respectively. Given *i.i.d.* observations $\mathbb{X}_n := \{X_1 = (X_1^1, X_1^2), \dots, X_n = (X_n^1, X_n^2)\}$ from the same distribution \mathcal{P} , we are interested in the test of independence

$$(H_0) : \mathcal{P} = \mathcal{P}^1 \otimes \mathcal{P}^2 \quad \text{versus} \quad (H_1) : \mathcal{P} \neq \mathcal{P}^1 \otimes \mathcal{P}^2.$$

In terms of neuroscience, a typical observation $X_i = (X_i^1, X_i^2)$ corresponds to the spike-trains of two neurons, and a natural way to construct such a test consists of computing and comparing the number of coincidences under each hypothesis. We count those coincidences thanks to a symmetric function $\varphi(X_i^1, X_i^2)$. In this work and following the ideas of [3, 38], what we actually count is the number of realisations of points of both variables with a delay $\delta > 0$, $\delta < T$. More precisely, we put

$$(1) \quad \varphi(X_i^1, X_i^2) = \sum_{u \in X_i^1, v \in X_i^2} \mathbb{1}_{|u-v| \leq \delta}.$$

In this framework, the test aims to compare the number of observed coincidences with the number of coincidences that corresponds to an independent model, that is, that would be observed under (H_0) . It is based on the statistics defined by

$$(2) \quad \mathbf{C}_n(\mathbb{X}_n) = \frac{1}{n} \sum_{i=1}^n \varphi(X_i^1, X_i^2)$$

which corresponds to the empirical mean number of coincidences in the sample. The test works well whenever the distributions of this statistics under (H_1) and (H_0) are well separated. We denote by \mathbb{P} (and \mathbb{E}) the distribution (and expectation) with respect to \mathbb{X}_n and \mathbb{P}_\perp (and \mathbb{E}_\perp) the distribution (and expectation) with respect to an *i.i.d.* sample of size n with similar marginals as \mathbb{X}_n but with independent coordinates. Since \mathbb{P}_\perp , that is the distribution under (H_0) , is unknown in practice, we use a permutation approach.

2.2. The permutation test. Let us now introduce the permutation test and some of its most important properties.

Let \mathbb{S}_n be the set of permutations of the integers $\{1, \dots, n\}$. Let \mathbb{X}_n be a sample and $\pi \in \mathbb{S}_n$. The permuted sample \mathbb{X}_n^π is obtained by permutation of the second spike train in \mathbb{X}_n according to π . Precisely, we set

$$(3) \quad \mathbb{X}_n^\pi = \left\{ X_1^\pi = (X_1^1, X_{\pi(1)}^2), \dots, X_n^\pi = (X_n^1, X_{\pi(n)}^2) \right\}$$

and we consider the permuted statistics $\mathbf{C}_n(\mathbb{X}_n^\pi)$.

Given the sample \mathbb{X}_n , we denote by $q_{1-\alpha}(\mathbb{X}_n)$ the $(1-\alpha)$ -quantile of the uniform distribution $\frac{1}{n!} \sum_{\pi \in \mathbb{S}_n} \delta_{\mathbf{C}_n(\mathbb{X}_n^\pi)}$. Then the permutation test

$$\mathbb{1}_{\mathbf{C}_n(\mathbb{X}_n) > q_{1-\alpha}(\mathbb{X}_n)}$$

which rejects (H_0) when the coincidence count is too large is exactly of level α . This follows from the fact that under (H_0) , for any permutation $\pi \in \mathbb{S}_n$, the samples \mathbb{X}_n^π and \mathbb{X}_n have the same distribution. We refer the reader to [28, 30].

Now let us turn towards the control of the Type II error. The informal idea behind the permutation test is that, even if we are under (H_1) , if we permute the second coordinate, we can mimic \mathbb{P}_\perp by associating X_i^1 with X_j^2 for two different trials i and j , say for instance $i = 1$ and $j = 2$. One important quantity is therefore

$$f_\varphi(X_1, X_2) = \varphi(X_1^1, X_1^2) - \varphi(X_1^1, X_2^2).$$

We then introduce

$$(4) \quad \Delta_\varphi := \mathbb{E}[f_\varphi(X_1, X_2)] = \mathbb{E}[\varphi(X_1^1, X_1^2)] - \mathbb{E}_\perp[\varphi(X_1^1, X_1^2)]$$

as a measure of how far the current distribution \mathbb{P} is from (H_0) , that is far from \mathbb{P}_\perp . In the sequel the separation rates of the permutation test will therefore be expressed in terms of Δ_φ .

Kim et al. [27] showed a general bound on the separation rates of independence tests by permutation. By applying their result to our setting, one can show the following general lemma.

Lemma 1. *Let α and β be fixed constants in $(0, 1)$. Then there exists an absolute constant $C > 0$ such that if the number of trials n satisfies $n \geq 3/\sqrt{\alpha\beta}$ and if*

$$(5) \quad \Delta_\varphi \geq \frac{C}{\sqrt{n\alpha\beta}} \left\{ \sqrt{\mathbb{V}_\perp[f_\varphi(X_1, X_2)]} + \sqrt{\mathbb{V}[f_\varphi(X_1, X_2)]} \right\},$$

then the Type II error of the permutation test satisfies

$$\mathbb{P}(\mathbf{C}_n(\mathbb{X}_n) \leq q_{1-\alpha}(\mathbb{X}_n)) \leq \beta.$$

Remark 2. In practice, accessing the $n!$ permutations is computationally too intensive. However one can approximate $q_{1-\alpha}(\mathbb{X}_n)$ by a Monte-Carlo method. It can be proved that this does not affect the level of the test (see for instance [2] and the references therein). Also by slightly increasing β , one can show that this does not really affect the previous result either. However to keep the writing of the next results as simple as possible, we do not incorporate this case and focus on the ideal test.

Let us now interpret the above criterion within our two main examples, the injection Poisson case and the network of Hawkes processes in a mean-field frame.

2.3. Jittering injection Poisson model. In this section, we are interested in the following model introduced by Date et al. [10] (see also [21, 38]) under which the spike trains are modeled by Poisson processes.

Model 1. Consider three independent Poisson processes Z^1 , Y^1 and Y^2 on \mathbb{R} with respective intensities η^1 , λ^1 and λ^2 .

- The first spike train X^1 is a superposition of the points of Y^1 and Z^1 .
- The second spike train X^2 is the superposition of the points of Y^2 and of a random perturbation Z^2 of Z^1 . Precisely, we consider an i.i.d. sequence of real valued random variables $(\xi_i)_{i \in \mathbb{Z}}$ independent of anything else. Denote by $\dots < u_{k-1} < u_k < u_{k+1} < \dots$ the points of Z^1 , then the points of Z^2 are $\dots, u_{k-1} + \xi_{k-1}, u_k + \xi_k, u_{k+1} + \xi_{k+1}, \dots$.

We prove in Proposition 19 that Z^2 is still a Poisson process of rate η^1 such that the spike trains X^1 and X^2 correspond to the points of Poisson processes with intensities $\lambda^1 + \eta^1$ and $\lambda^2 + \eta^1$ respectively. For the convenience of the reader we give a proof of this fact in the appendix, section C.6.

The following result provides a lower bound for the separation rate between (H_0) and (H_1) , measured by Δ_φ . From this we can extract a minimal number of observations \mathbf{n}_{\min} to ensure a Type II error less than β .

Theorem 3. Let $\alpha, \beta \in (0, 1)$, $0 \leq \delta \leq T/2$ and assume that the observations obeys Model 1 with parameters $\lambda^1, \lambda^2, \eta^1$ on the time interval $[0, T]$. Then

$$(6) \quad \Delta_\varphi = \eta^1 \mathbb{E} [(T - |\xi|) \mathbb{1}_{|\xi| \leq \delta}].$$

Moreover there exists a constant $C > 0$ not depending on the model parameters, such that if $n \geq 3/\sqrt{\alpha\beta}$ and if

$$(7) \quad \Delta_\varphi \geq C \left[\sqrt{\frac{(1 + \lambda^m \delta)(\lambda^m)^2 \delta T}{n \alpha \beta}} + \frac{1 + \lambda^m \delta}{n \alpha \beta} \right],$$

with $\lambda^m = \max\{\lambda^1, \lambda^2\}$ and $\eta^1 \leq \lambda^m$, then the Type II error of the permutation test satisfies

$$\mathbb{P}(\mathbf{C}_n(\mathbb{X}_n) \leq q_{1-\alpha}(\mathbb{X}_n)) \leq \beta.$$

Moreover, there exists a constant $C' > 0$, not depending on the model parameters, such that the previous control of the Type II error is guaranteed as soon as

$$(8) \quad n \geq \mathbf{n}_{\min} := C' \max \left\{ \frac{1}{\sqrt{\alpha\beta}}, \frac{1 + \lambda^m \delta}{\alpha \beta \eta^1 T \mathbb{P}(|\xi| \leq \delta)} \left(1 + \frac{\lambda^m}{\eta^1} \frac{\lambda^m \delta}{\mathbb{P}(|\xi| \leq \delta)} \right) \right\}.$$

Note first that the measure of the distance between (H_0) and (H_1) , Δ_φ , only depends on the distribution of the perturbation ξ and the frequency of the injected synchronizations η^1 (see (6)). Next the bound (7) states that the separation rate diminishes in $n^{-1/2}$ as one would expect in this parametric situation. Moreover we also recover the dependency in \sqrt{T} that one would expect in this parametric Poissonian situation.

At this point the following remark is at order. Speaking of “trials” is a bit artificial here because of the memoryless property of the Poisson processes which implies that observing n independent copies of one Poisson process during a time interval of length T is equivalent to observing a single trajectory of the same Poisson process during a time interval of length nT . In other words, for a global duration of observation \mathcal{T} , any choice of n and T such that $\mathcal{T} = nT$ are essentially possible. Since we have renormalized our test statistics $\mathbf{C}_n(\mathbb{X}_n)$ by $1/n$, the number of trials, in (2), to take into account of this fact, one has to multiply (7) by n in order to obtain on its right hand side indeed a term involving $\sqrt{nT} = \sqrt{\mathcal{T}}$, as expected. We decided to work with the renormalized test statistics as it is commonly done.

Note also that to fix asymptotics, we work here with $T \geq 2\delta$ to ensure that in the middle of the interval $[0, T]$, we do not have boundary effects due to δ . Because of this convention, we have that

$$(9) \quad \eta^1 T \mathbb{P}(|\xi| \leq \delta) \geq \Delta_\varphi \geq \eta^1 \frac{T}{2} \mathbb{P}(|\xi| \leq \delta),$$

so that Δ_φ essentially behaves like $\mathbb{P}(|\xi| \leq \delta)$ and increases with δ . The right hand side of (7) increases with δ as well, at least as $\sqrt{\delta}$. Therefore, depending on the rate of growth of $\mathbb{P}(|\xi| \leq \delta)$ as a function of δ , one can have situations where the largest possible choice of δ is a good choice, or situations where a compromise need to be made.

We stress that (8) is essentially a lower bound on $\eta^1 n T = \eta^1 \mathcal{T}$, the average total number of injected common points in the data. So if the lower bounds are achieved, we are essentially saying that, for the independence test to detect synchronization, this number should be larger than

$$\frac{1 + \lambda^{\mathbf{m}} \delta}{\mathbb{P}(|\xi| \leq \delta)} \left(1 + \frac{(\lambda^{\mathbf{m}})^2 \delta}{\eta^1 \mathbb{P}(|\xi| \leq \delta)} \right).$$

In the asymptotic $\delta \rightarrow 0$, if there exists γ such that $\mathbb{P}(|\xi| \leq \delta) =_{\delta \rightarrow 0} \mathcal{O}(\delta^\gamma)$, we obtain that the minimal number of observations should be at least of order

$$\mathcal{O}(\delta^{-\gamma}) + \frac{(\lambda^{\mathbf{m}})^2}{\eta^1} \mathcal{O}(\delta^{1-2\gamma}).$$

Hence

- if $\gamma > 1/2$, it decreases with δ , and choosing δ large is a good choice. A typical example is when ξ is uniform on $[-D, D]$, for which the best choice would be $\delta = D$.
- if $\gamma < 1/2$, an equilibrium in δ can be found in $\mathcal{O}\left(\frac{\eta^1}{(\lambda^{\mathbf{m}})^2}\right)^{\frac{1}{1-\gamma}}$.

Therefore, one sees that, even in this very simple model, the choice of δ is in fact quite subtle and depends heavily on the distribution of the noise, that is not known in practice.

2.4. Hawkes model. Introduced for the first time by Hawkes [22], this model describes a stochastic dynamics on a network that is widely used nowadays to model the activity of a neuronal network (see also [9] for multivariate and non linear cases). We envision here a particular case where interactions in the (unobserved) network of size M are homogeneous, while we observe only two neurons embedded in the network.

Model 2. Let $\nu, a, b > 0$ and $\ell = a/b \in (0, 1)$. For $i \in \llbracket 1, M \rrbracket$, neuron i spikes at time t with intensity

$$(10) \quad \lambda^i(t) = \nu + \frac{a}{M} \sum_{j=1}^M \int_{(-\infty, t)} e^{-b(t-s)} dN_s^j,$$

where dN_t^j is the point measure associated to the spikes of neuron j . In this model, (X_1, X_2) is the observation of the first two coordinates of this M -multivariate Hawkes process in a stationary regime.

The hypothesis $\ell < 1$ and the exchangeability of the network allow to construct the stationary distribution of the previous model as follows.

- Simulate a univariate stationary Hawkes process N with intensity

$$(11) \quad \lambda(t) = \nu M + a \int_{(-\infty, t)} e^{-b(t-s)} dN_s.$$

This process is well defined since $\ell = a/b < 1$ is the integral of the interaction kernel of the univariate linear Hawkes process (see [23, 9]).

- Attribute at random, independently of anything else, each point T of N to the process N^j of neuron j with uniform probability on $\llbracket 1, M \rrbracket$.

Note then that $dN_t = \sum_{j=1}^M dN_t^j$. Note also that the univariate process N can be seen as a cluster process (see [23]) that is described in more details in Section 4. So one can see N as a superposition of clusters of points, each of them arriving according to a Poissonian rate νM .

For this model, one can show that

Theorem 4. *Let $\alpha, \beta \in (0, 1)$, $0 \leq \delta \leq T/2$ and assume that the observations obey Model 2 with parameters $\nu, a, b > 0$, and $\ell = a/b \in (0, 1)$ on the time interval $[0, T]$. Then*

$$(12) \quad \Delta_\varphi = \frac{\nu\delta(2-\ell)\ell}{M(1-\ell)^3} \left(1 + [b(1-\ell)(T-\delta) - 1] \frac{1 - e^{-b(1-\ell)\delta}}{b(1-\ell)\delta} \right).$$

Moreover there exists a constant $C > 0$ not depending on the model parameters, such that if $n \geq 3/\sqrt{\alpha\beta}$ and if

$$(13) \quad \Delta_\varphi \geq C \sqrt{\frac{\delta T}{n\alpha\beta}} \left(\frac{\nu + a/M}{1-\ell} \right) \sqrt{1 + \frac{\delta(\nu + a/M)}{1-\ell} \left[1 + \frac{\ell}{M(1-\ell)^2} \right]},$$

then the Type II error of the permutation test satisfies

$$\mathbb{P}(\mathbf{C}_n(\mathbb{X}_n) \leq q_{1-\alpha}(\mathbb{X}_n)) \leq \beta.$$

If in addition we assume that $b(1-\ell)T > 4$, then there exists a constant $C' > 0$ not depending on the model parameters, such that the previous control of the Type II error is guaranteed if $n \geq 3/\sqrt{\alpha\beta}$ and if

$$(14) \quad n \geq \frac{C'}{\alpha\beta} \frac{(M + a/\nu)^2(1-\ell)^4}{\ell^2 T} \left(\frac{\delta + \delta^2 \frac{(\nu+a/M)}{1-\ell} \left[1 + \frac{\ell}{M(1-\ell)^2} \right]}{(1 - e^{-b(1-\ell)\delta})^2} \right).$$

The expression of the distance between (H_0) and (H_1) in (12) is complicated and depends on all the parameters of the model. In particular, it involves the quantity $(1 - e^{-b(1-\ell)\delta})/(b(1-\ell)\delta)$ which is the mean total offspring of one single jump – that is, within one single cluster – occurring during a time interval of length δ . See Section 4 for more details.

By looking at (13), we see again, as for the jittering model, a parametric asymptotic in $n^{-1/2}$. We also see that the dependence is, as in the previous Poissonian case, in \sqrt{T} . If the dependency structure of the Hawkes process makes it less straightforward that one can cut the total duration of observation nT in as many trials as one wants, we still have that in a context of large T , much larger than the typical length of one cluster of N , T is proportional to the number of clusters and the clusters are independent, which provides an asymptotic for the variances of the coincidences count in \sqrt{T} as well.

Finally let us comment (14). Let us start by discussing the condition on $b(1-\ell)T > 4$: this compares T to the characteristic time $1/(b(1-\ell))$ of the cluster. Hence it ensures that one sees a certain number of clusters in one trial, making it possible to see interactions.

Next we see the clear asymptotics in M^2 for the number of trials. As announced in the introduction, this means that we need a number of trials larger than the square of the size of the network to detect the dependency. In particular, this leads us to the following interpretation in neuroscience: if we accept the null hypothesis on real data, and if the network is homogeneous as given by Model 2, we cannot distinguish the fact that the neurons are really independent from the fact that they are embedded in a network of order \sqrt{n} or larger, where n is the number of trials.

The factor ν also plays an important role. If ν tends to 0, the right hand side explodes: indeed if the rate of apparition of the clusters vanishes, we do not observe any cluster, hence no interaction.

In the same way, when $\ell = a/b$ tends to 0 it means that the interaction vanishes and again, the detection problem becomes more difficult to solve (more trials are needed).

Finally, the dependency in terms of δ is more complicated. When δ tends to 0 or infinity, the right hand side explodes, so there is a minimal value of the right hand side in δ that does not depend on T nor n or α and β . It seems however that a good choice would be of the order of $(b(1-\ell))^{-1}$, that is the characteristic time of the cluster.

3. SIMULATIONS

In this section, we perform numerical simulations in order to illustrate the separation rates proved for the permutation tests under Models 1 and 2. As explained in Section 2.2, the permutation test is based on estimating the $(1-\alpha)$ -quantile of $\mathbf{C}_n(\mathbb{X}_n^\pi)$, conditionally to \mathbb{X}_n where \mathbb{X}_n is the observed sequence of point processes and the permutation π is chosen uniformly. To reduce

computational time and following the idea of [2, 3], we use a Monte-Carlo method to approximate the bootstrap quantile, with only π^1, \dots, π^B i.i.d. permutations picked uniformly at random.

In all the sequel, we use $B = 5000$ and $\alpha = 0.05$.

Since the level of the test was studied in [2, 3], we do not provide results under the null hypothesis here, but we have checked on simulation that as expected, the permutation tests have a Type I error less than 0.05 in all cases.

3.1. The jittering Poisson model. We have simulated the jittering Poisson model with $\lambda_1 = \lambda_2 = 10\text{Hz}$ and the injection rate $\lambda = 1\text{Hz}$ on $[0, 2]$. The noise ξ is of three different shapes (i) Uniform on $[0, 0.1]$ (ii) of density $200 * (0.1 - x)$ on $[0, 0.1]$ (iii) of density $200x$ on $[0, 0.1]$.

Case (i) and (ii) have a density that is lower bounded around 0, whereas case (iii) implies a jitter that is seldom in a vicinity of 0. The resulting power of the permutation test is plotted in Figure 1.

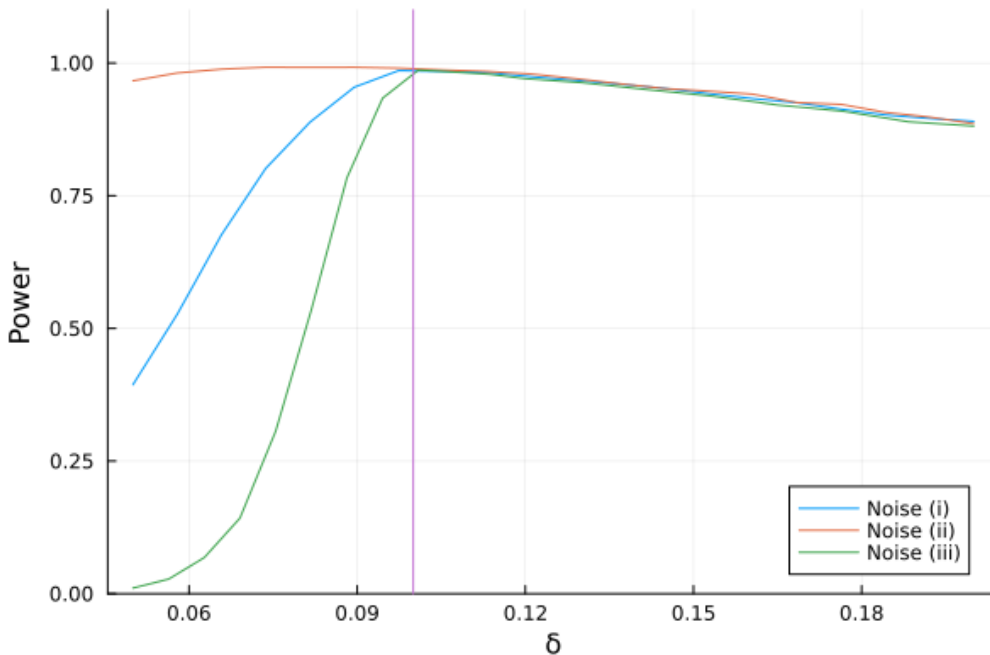


FIGURE 1 Influence of the noise and δ on the power of the test. We simulated $n = 200$ trials and we replicated this simulation $N = 10000$ times to estimate the power of the test. Noises (i),(ii) and (iii) are defined in Section 3.1. The parameter δ varies in a regular grid between 0.05 and 0.2. The vertical line indicates the boundary of support for all noises.

We see that there is a high sensitivity of the test as a function of δ , especially if the jittering noise does not charge a vicinity of 0 (Noise (iii)). The best choice of δ seems to depend on the noise, but a fair choice seems to be the one that will encompass most of the support of the jittering noise. After this (beyond 0.10), the power is decreasing slowly, indicating that there is no good reason to choose a large δ in all cases but that a slight overestimation of the support is not very problematic. The fact that there should be an optimal δ is also in accordance with the theoretical results of Theorem 3.

3.2. The Hawkes model. We simulated a homogeneous network of Hawkes processes of size M for various M between 10 and 30 and different a and b . We fixed $\nu = 1$ and we simulated all trials on $[0, 2]$ after a warming up time period of simulation of size 10 to reach stationarity.

In Figure 2, we see the evolution the power of the test as a function of n the number of trials. As expected the power rapidly deteriorates when M increases. With $M = 30$ and $n = 560$ we do not even reach a power of 80%.

To understand if the estimate of Theorem 4 in terms of number of trials necessary to achieve a certain Type II error is valid, at least in term of M , we decided to fix the Type II error $\beta = 0.05$ and find n^* the minimal number of trials to achieve such a small error. This is represented in Figure 3. We see a very good fit with the superposed curve, showing that indeed the number of necessary trials is growing as M^2 .

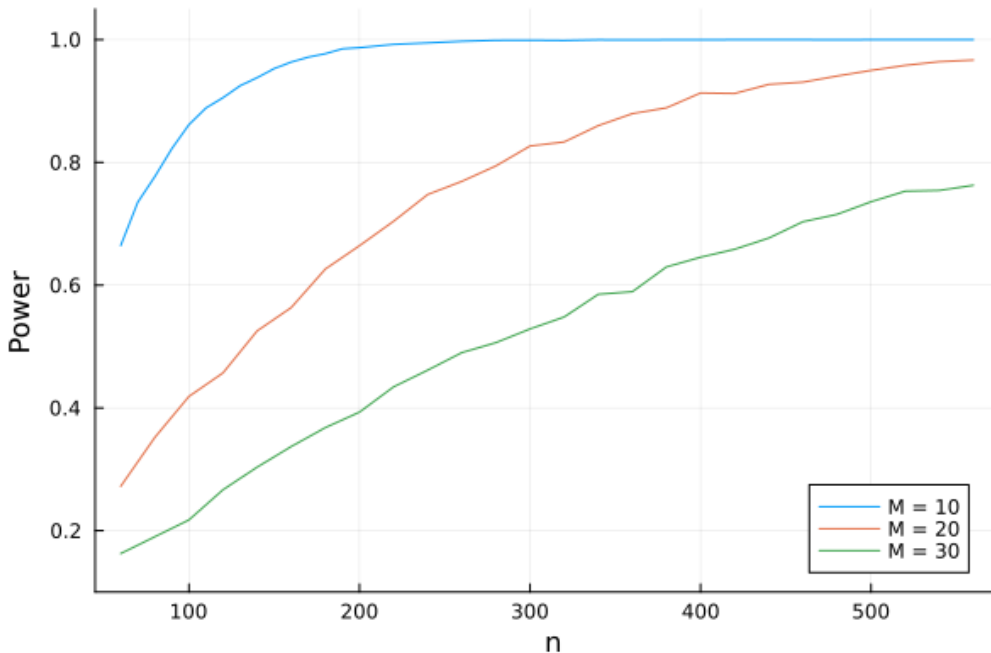


FIGURE 2 Influence of the number of trials n and of the size of the network M . The number of trials n varied in a regular grid between 60 and 550 and the point processes were the first two coordinates of an homogeneous network of size $M = 10, 20, 30$, with $a = 3$, $b = 4$ and $\delta = 0.1$. We simulated $N = 10000$ times the whole procedure to approximate the power of the test.

4. SKETCH OF PROOFS AND MOMENT FORMULA

The proofs of the above results are rather technical and postponed to the appendix. We give here some ideas of the proofs. Moreover we give exact formulas or bounds on the moments of the coincidence counts in both settings: the injection Poisson model and the homogeneous network of exponential Hawkes process.

4.1. On another formulation of the test. As already discussed in Section 2.2, the permutation test is based on the empirical mean number of coincidences $\mathbf{C}_n(\mathbb{X}_n)$ defined in (2). However the variance of this statistics can be very large while the correlation between the observed pairs of spike trains is not. This can be due to a large amount of independent spikes for example. Following [2, 3], we consider therefore the U-statistics

$$(15) \quad \mathbf{T}_n(\mathbb{X}_n) := \frac{1}{n(n-1)} \sum_{i,j=1}^n f_\varphi(X_i, X_j)$$

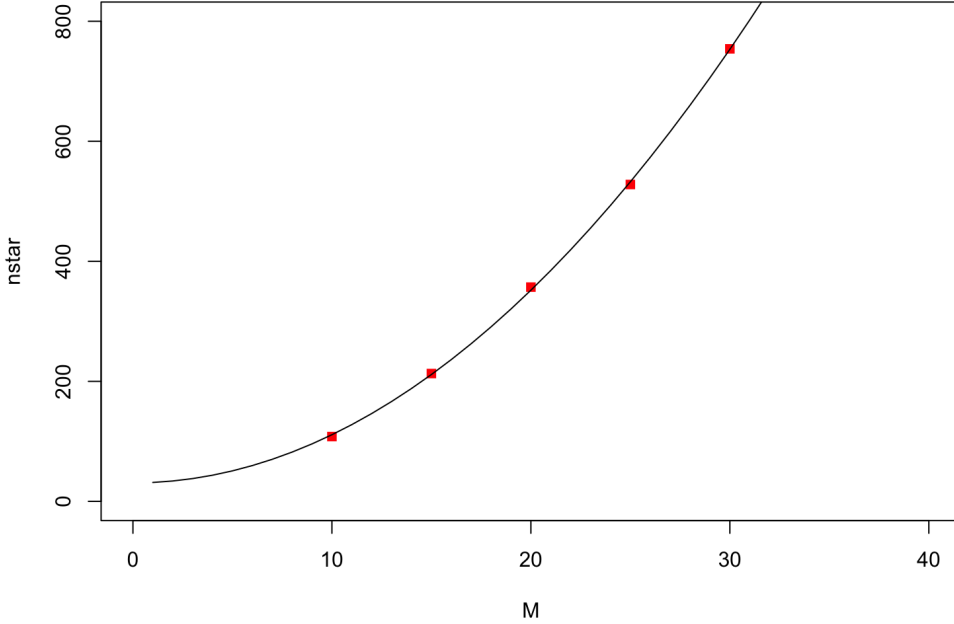


FIGURE 3 The number of necessary trials grows in M^2 . We plot with red squares n^* (see Section 3.2) as a function of the size of the network M . The simulation setting is similar to the one of Figure 2 except that $N = 10000$, $a = 10$, $b = 20$, $\delta = 0.1$. We fitted the curve by linear regression with respect to M^2 and superposed in black the curve $30 + 0.8M^2$.

where $f_\varphi(X_i, X_j) = \varphi(X_i^1, X_i^2) - \varphi(X_i^1, X_j^2)$. We have

$$(16) \quad \mathbb{E}[\mathbf{T}_n(\mathbb{X}_n)] = \mathbb{E}[\varphi(X_1^1, X_1^2)] - \mathbb{E}_\perp[\varphi(X_1^1, X_1^2)],$$

and so $\mathbf{T}_n(\mathbb{X}_n)$ is centered under the null hypothesis. As noted in [3], we have

$$(17) \quad \mathbf{T}_n(\mathbb{X}_n) = \frac{n}{n-1} \mathbf{C}_n(\mathbb{X}_n) - \frac{1}{n(n-1)} \sum_{i,j=1}^n \varphi(X_i^1, X_j^2).$$

Let us now denote by Π_n a random variable uniformly distributed on the set \mathbb{S}_n of permutations, independent of the sample \mathbb{X}_n . We deduce from (17) that for the permuted sample, we also have

$$(18) \quad \mathbf{T}_n(\mathbb{X}_n^{\Pi_n}) = \frac{n}{n-1} \mathbf{C}_n(\mathbb{X}_n^{\Pi_n}) - \frac{1}{n(n-1)} \sum_{i,j=1}^n \varphi(X_i^1, X_j^2).$$

It follows that the permutation tests based on $\mathbf{C}_n(\mathbb{X}_n)$ and $\mathbf{T}_n(\mathbb{X}_n)$ are exactly the same. In the sequel, we focus on this new formulation, use [27] and prove Lemma 1. The computations are quite technical and postponed to the appendix.

4.2. Proof of Theorem 3. This result follows from Lemma 1 and its proof consists of computing the means and variances that appear in this lemma under Model 1. Notice that having an upper bound of the variances is enough. The computations use the following intermediate result that can be useful (and even generalized) in other situations if needed.

Lemma 5. *Let us consider:*

- A Poisson process N on \mathbb{R} with intensity $\eta^1 > 0$.

- An i.i.d family $(\xi_k)_{k \in \mathbb{Z}}$ of real valued random variables that is independent of N .
- For any $l \geq 2$ the set of vectors with different coordinates

$$(19) \quad \Delta_l = \{(s_1, \dots, s_l) \in \mathbb{R}^l : s_i \neq s_j, \forall i \neq j\}.$$

Then for any deterministic and measurable function $F : \mathbb{R}^n \times \mathbb{R}^l \rightarrow \mathbb{R}$ such that

$$\mathbb{E} \left[\int_{\mathbb{R}^l} |F(s_1, \dots, s_l, \xi_1, \dots, \xi_l)| ds_1 \cdots ds_l \right] < \infty,$$

we have

$$(20) \quad \mathbb{E} \left[\int_{\Delta_l} F(s_1, \dots, s_l, \xi_{N_{s_1}}, \dots, \xi_{N_{s_l}}) dN_{s_1} \cdots dN_{s_l} \right] \\ = (\eta^1)^l \mathbb{E} \left[\int_{\mathbb{R}^l} F(s_1, \dots, s_l, \xi_1, \dots, \xi_l) ds_1 \cdots ds_l \right].$$

We give a short proof of this lemma in the appendix section C.6.

Remark 6. The counting process (N_t) has only jumps of size 1, i.e. $N_s = N_{s-} + 1$ for any jump time s . In (20), we should replace $\xi_{N_{s_1}}$ by $\xi_{N_{s_1-}+1}$ to avoid any doubt of measurability in the stochastic integral. However, we decide to keep this form to alleviate the notation.

Using this lemma, we split the proof of Theorem 3 in the following steps:

Step 1: We start with the proof of (6). It directly follows from the equality

$$(21) \quad \mathbb{E} [\varphi(X_1^1, X_1^2)] - \mathbb{E}_{\perp} [\varphi(X_1^1, X_1^2)] = \mathbb{E} \left[\int_0^T \mathbb{1}_{|\xi_{Z_1^1(u)}| \leq \delta} \mathbb{1}_{u + \xi_{Z_1^1(u)} \in [0, T]} dZ_1^1(u) \right],$$

where Z_1^1 is the Poisson process which links the spike trains X_1^1 and X_1^2 (see Model 1). Eq. (21) is obtained by distinguishing the jump times of both spike-trains that come from the same underlying jump time. It is then easy to verify that the right hand term in (21) is equal to the one of (6) (see Appendix B.1). In particular, we deduce the lower bound

$$(22) \quad \mathbb{E} [\varphi(X_1^1, X_1^2)] - \mathbb{E}_{\perp} [\varphi(X_1^1, X_1^2)] \geq \eta^1 (T - \delta) \mathbb{P}(|\xi| \leq \delta) \geq \frac{\eta^1 T}{2} \mathbb{P}(|\xi| \leq \delta),$$

since $\delta \leq T/2$.

Step 2: Setting $\lambda^{\mathbf{m}} = \max\{\lambda^1, \lambda^2\}$ and assuming that $\eta^1 \leq \lambda^{\mathbf{m}}$, we show that there exists a constant $C > 0$ independent of the parameters such that the variances of $f_{\varphi}(X_1, X_2)$ under the observed and the independent model satisfy the bounds

$$(23) \quad \mathbb{V}_{\perp} [f_{\varphi}(X_1, X_2)] \leq C(\lambda^{\mathbf{m}})^2 \delta T (1 + \lambda^{\mathbf{m}} \delta)$$

and

$$(24) \quad \mathbb{V} [f_{\varphi}(X_1, X_2)] \leq C(1 + \lambda^{\mathbf{m}} \delta) ((\lambda^{\mathbf{m}})^2 \delta T + \eta^1 T \mathbb{P}(|\xi| \leq \delta)).$$

We conclude the proof by replacing the mean, variances and the second order moment of $f_{\varphi}(X_1, X_2)$ in the inequality (5) of Lemma 1 by their bounds in (22), (23) and (24). The resulting inequality is solved according to the unknown n in order to get the lower bound (8) on the number of observations which ensures the control of the second kind error.

4.3. Proof of Theorem 4. As in the case of the jittering injection Poisson model, this result follows from Lemma 1, and its proof consists of computing the expectations and variances of $f_{\varphi}(X_1, X_2)$ for the observed and independent Hawkes models. These computations use the cumulant measures of the univariate Hawkes process N with intensity $\lambda(t)$ defined by (11) since this process allows to construct the entire network. We start by giving some important informations about cumulant measures and their computation which are mainly based on [25, 33].

About the cumulant measures of a univariate Hawkes process. Let N be a univariate Hawkes process with intensity

$$(25) \quad \lambda(t) = \mu + \int_{(-\infty, t)} h(t-s) dN_s$$

where the interaction function $h : \mathbb{R}_+ \rightarrow \mathbb{R}_+$ is such that $\|h\|_{L^1} < 1$. Let $l \geq 1$ be an integer, then (see [25, 33]) the l -order cumulant measure of N is the non negative measure defined by

$$(26) \quad k_l(dt_1, \dots, dt_l) = \sum_{\mathbf{p}} (|\mathbf{p}| - 1)! (-1)^{|\mathbf{p}|-1} \bigotimes_{B \in \mathbf{p}} \mathbb{E} \left[\bigotimes_{i \in B} dN_{t_i} \right]$$

where the sum goes over all partitions \mathbf{p} of the set $\{1, \dots, l\}$, $|\mathbf{p}|$ corresponds to the number of elements of this partition and the notation $\mathbb{E}[dN_{t_1} \dots dN_{t_n}]$ refers to the positive measure on \mathbb{R}^n defined by

$$\int \phi(t_1, \dots, t_n) \mathbb{E}[dN_{t_1} \dots dN_{t_n}] = \mathbb{E} \left\{ \int \phi(t_1, \dots, t_n) dN_{t_1} \dots dN_{t_n} \right\},$$

defined for all measurable test functions ϕ having compact support.

In the sequel we will mainly use the first and second order cumulants. The first order cumulant is the mean measure $k_1(dt_1) = \mathbb{E}[dN_{t_1}]$ and the second order cumulant is the covariance

$$k_2(dt_1, dt_2) = \mathbb{E}[dN_{t_1} dN_{t_2}] - \mathbb{E}[dN_{t_1}] \mathbb{E}[dN_{t_2}].$$

In particular, the second order cumulant is equivalent to the first order cumulant on the diagonal $\{t_1 = t_2\}$, that is

$$\mathbb{1}_{t_1=t_2} k_2(dt_1, dt_2) = \delta_{t_1}(dt_2) k_1(dt_1).$$

One well known property of the cumulant measure is the dual formula

$$(27) \quad \mathbb{E}[dN_{t_1} \dots dN_{t_l}] = \sum_{\mathbf{p}} \bigotimes_{B \in \mathbf{p}} k_{\#B}(dt_{i_1^B}, \dots, dt_{i_{\#B}^B})$$

where the notation $\#B$ refers to the cardinal of the subset B and where we suppose that $B = \{i_1^B, \dots, i_{\#B}^B\}$. Thus the positive measure $\mathbb{E}[dN_{t_1} \dots dN_{t_l}]$ is entirely determined by the knowledge of the cumulant measures up to order l . An efficient way to compute these cumulant measures has been proposed by Jovanović et al. [25] and consists in using the cluster representation of the Hawkes process (see [20, 23, 34]). Indeed, the Hawkes process N can be constructed as follows:

- (1) Simulate a realization of an homogeneous Poisson process with rate μ on $(-\infty, T]$. We call each point of this process an *immigrant*.
- (2) Each immigrant t_0 generates a cluster C_{t_0} . These clusters are mutually independent and each of them is constructed recursively:
 - (a) The first generation is an inhomogeneous Poisson process on $[t_0, T]$ with rate $r_1(t) = h(t - t_0), t \geq t_0$.
 - (b) Knowing the offspring generations $1, \dots, n$, every child s of generation n generates its own offspring of generation $n + 1$ as a realization of an inhomogeneous Poisson process on the time interval $[s, T]$ with rate $r_{n+1}(t) = h(t - s), t \geq s$.

The Hawkes process N corresponds to the superposition of all such clusters. The cumulant measure $k_l(dt_1, \dots, dt_l)$ expresses situations where t_1, \dots, t_l are indices of points belonging to the same cluster. We can pinpoint these points inside the whole genealogic tree of a given cluster described above. But to simplify, we only need to keep the nodes that are the *latest common ancestors* of these points, until the first immigrant from which the cluster is originated. So in the simplified tree, a node correspond to the *latest common ancestor* of its children, and the final leaves are the t_1, \dots, t_l . This structure can be put in a one-to-one map with dendrograms (see Murtagh [29]), where nodes represent sets that are the union of their children. We call the previously introduce simplified trees “dendrograms” as well.

Definition 7. A dendrogram is a labeled rooted tree with real-valued labels, where the terminal nodes’ labels are t_1, \dots, t_l , where the internal nodes are the latest common ancestor of their children. We add a final connection from the latest common ancestor of all the leaves to the

immigrant generating the whole cluster, which correspond to the root of the tree. We denote by $\mathbb{Q}(t_1, \dots, t_l)$ the set of dendrograms with terminal nodes t_1, \dots, t_l .

Figure 4 shows the unique dendrogram having two terminal nodes t_1 and t_2 . It expresses the fact that the jump times t_1 and t_2 are both caused by the jump at time u of their common ancestor, which was itself caused by an immigrant arriving at time x .

Notice that t_1 and t_2 are not necessarily of the same generation, that is, they are not necessarily brothers and/or sisters. We will also consider the situation when u and $t_1 \wedge t_2$ collapse to a single node, that is, when $t_1 \wedge t_2$ is an ancestor of $t_1 \vee t_2$.

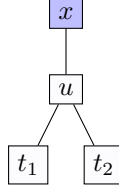


FIGURE 4 The unique dendrogram for $l = 2$.

In what follows we want to integrate with respect to the law of the time arrival of the common ancestors, that is, of the internal nodes (e.g. x and u in Figure 4). To do so, we assign measure valued weights to the non-terminal nodes and the roots of the dendrograms in $\mathbb{Q}(t_1, \dots, t_l)$. For each node, this weight corresponds to the intensity at which it appears given its children.

To do so, we first introduce the function given by

$$(28) \quad \Psi(t) = \sum_{m \geq 1} h^{*m}(t), t \geq 0,$$

where we define recursively for any $t \geq 0$,

$$(29) \quad \begin{cases} h^{*1}(t) = h(t), \\ h^{*(m+1)}(t) = \int_0^t h^{*m}(t-s)h(s)ds, m \geq 1. \end{cases}$$

The function $h^{*m}(t)$ is the rate of occurrence of offspring in the m -th generation from the root 0, thus $\Psi(t)$ is the rate of occurrence of offspring of 0 whatever their generation.

Given a dendrogram $\mathbf{q} \in \mathbb{Q}(t_1, \dots, t_l)$, its root x is weighted by the positive measure

$$(30) \quad W[u](x) = \mu(\delta_u(dx) + \Psi(u-x)dx),$$

where u is the only child of x in \mathbf{q} . In the cluster representation of the Hawkes process, the node u is the last common ancestor of t_1, \dots, t_l , and then Eq. (30) expresses the fact that u might be the root itself – appearing at rate μ – or an offspring of the root after a certain number of generations – appearing at rate $\mu\psi(\cdot - x)$.

Let us now consider a non-terminal node v of \mathbf{q} and introduce the following notations:

- $\text{ch}_{\mathbf{q}}(v)$ denotes the set of children of v in \mathbf{q} ,
- $\text{le}_{\mathbf{q}}(v)$ denotes the set of terminal nodes in $\text{ch}_{\mathbf{q}}(v)$.

The node v is weighted by the measure

$$(31) \quad W[\text{ch}_{\mathbf{q}}(v)](dv) = \left(\prod_{e \in \text{ch}_{\mathbf{q}}(v)} \Psi(e-v) \right) dv + \left(\prod_{\substack{e \in \text{ch}_{\mathbf{q}}(v) \\ \text{s.t } e \neq \min_{t \in \text{le}_{\mathbf{q}}(v)} t}} \Psi \left(e - \min_{t \in \text{le}_{\mathbf{q}}(v)} t \right) \right) \delta_{\min_{t \in \text{le}_{\mathbf{q}}(v)} t}(dv).$$

From the viewpoint of the cluster representation of the Hawkes process, the first term describes the emergence of offspring of v after at least one generation. The second term allows to consider the case where a terminal node (the smallest one) is the common ancestor of the others. This term is assumed to be null if $\text{le}_{\mathbf{q}}(v) = \emptyset$.

We now state the following result inspired by Jovanović et al. [25], that shows how to compute the cumulant measures.

Theorem 8. *The l -th order cumulant measure $k_l(dt_1, \dots, dt_l)$ of the Hawkes process N can be represented as follows.*

•

$$(32) \quad k_1(dt_1) = \frac{\mu}{1 - \|h\|_{L^1}} dt_1.$$

- *On the diagonals: if $1 \leq i, j \leq l$, then*

$$(33) \quad \mathbb{1}_{t_i=t_j} k_l(dt_1, \dots, dt_l) = \delta_{t_i}(dt_j) k_{l-1}(dt_1, \dots, dt_{j-1}, dt_{j+1}, \dots, dt_l).$$

- *Out of the diagonals (see Jovanović et al. [25]): we set*

$$(34) \quad \Delta_l = \{(t_1, \dots, t_l) \in \mathbb{R}^l : t_i \neq t_j, \forall i \neq j\}.$$

Then

$$(35) \quad \mathbb{1}_{\Delta_l} k_l(dt_1, \dots, dt_l) = \left(\sum_{\mathbf{q} \in \mathbb{Q}(t_1, \dots, t_l)} \int_{\mathbb{R}^{1+p}} \mathbb{W}[\text{ch}_{\mathbf{q}}(x)](dx) \mathbb{W}[\text{ch}_{\mathbf{q}}(v^1)](dv^1) \dots \mathbb{W}[\text{ch}_{\mathbf{q}}(v^p)](dv^p) \right) dt_1 \dots dt_l,$$

where $p = p(\mathbf{q})$ is the number of non-terminal nodes of \mathbf{q} , where v^1, \dots, v^p are the non-terminal nodes, and where x is its root.

In practice, Murtagh [29] showed how to compute the number of dendrograms with l terminal nodes, given $l \geq 1$. However listing the dendrograms is tedious when l increases. For example, according to Murtagh [29], there are: one dendrogram for $l = 2$, four for $l = 3$, twenty six for $l = 4$ and even more for l larger. But the set of dendrograms $\mathbb{Q}(t_1, \dots, t_l)$ can be decomposed in equivalence classes whose elements result from permutations of the positions of the terminal nodes t_1, \dots, t_l . Notice that the permutations that only swap the positions of terminal nodes with the same parents do not modify the dendrogram. Thus they only count as one. We have the following particular cases:

- In the case $l = 2$, we have only one dendrogram given in Figure 4. We deduce that the second order cumulant measure $k_2(dt_1, dt_2) = \mathbb{E}[dN_{t_1} dN_{t_2}] - \mathbb{E}[dN_{t_1}] \mathbb{E}[dN_{t_2}]$ satisfies

$$(36) \quad \mathbb{1}_{\Delta_2} k_2(dt_1, dt_2) = \int_{\mathbb{R}^2} \mathbb{W}[u](dx) \mathbb{W}[t_1, t_2](du) dt_1 dt_2.$$

- In the case $l = 3$, we have four dendrograms but only two equivalence classes described in Figure 5.



FIGURE 5 Equivalence classes of dendrograms for $l = 3$: (a) has only one permutation, (b) has three permutations..

It follows that the third order cumulant satisfies

$$(37) \quad \mathbb{1}_{\Delta_3} k_3(dt_1, dt_2, dt_3) = \left\{ \int_{\mathbb{R}^2} \mathbb{W}[u](dx) \mathbb{W}[t_1, t_2, t_3](du) \right. \\ \left. + \sum_{\sigma} \int_{\mathbb{R}^3} \mathbb{W}[u](dx) \mathbb{W}[v, t_{\sigma(3)}](du) \mathbb{W}[t_{\sigma(1)}, t_{\sigma(2)}](dv) \right\} dt_1 dt_2 dt_3$$

where the sum is over all the permutations of $\{1, 2, 3\}$ in the equivalence class that corresponds to the dendrogram **(b)** in Figure 5.

- In the case $l = 4$, there are 26 dendrograms, but only five equivalence classes that are described in Figure 6.

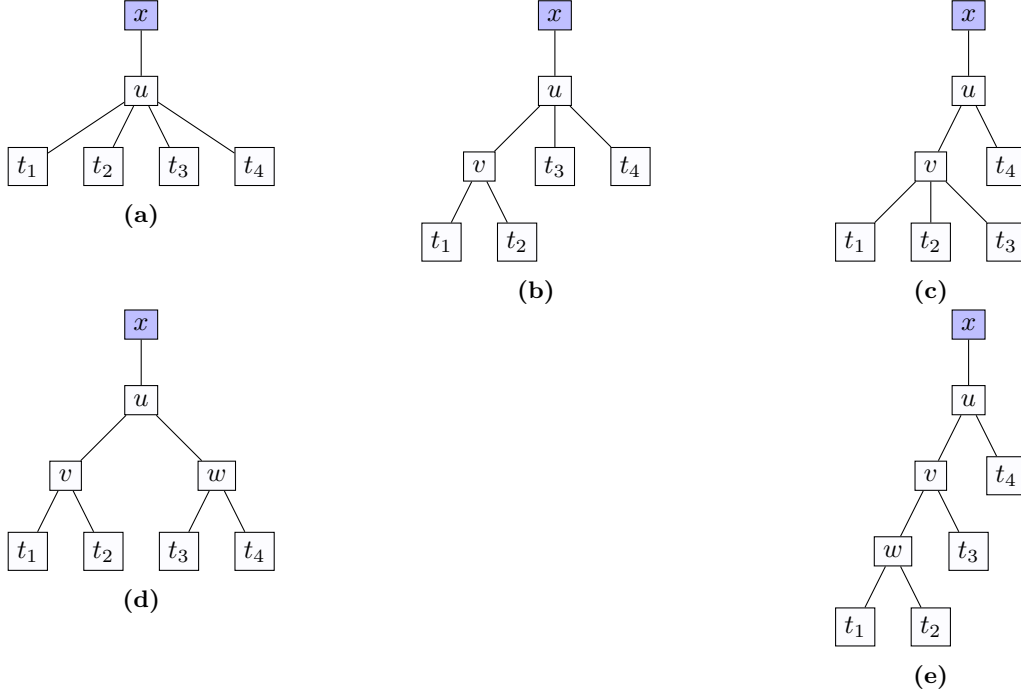


FIGURE 6 Equivalence classes of dendrograms for $l = 4$: **(a)** has one permutation, **(b)** has six permutations, **(c)** has four permutations, **(d)** has three permutations, **(e)** has twelve permutations.

We deduce that the fourth order cumulant measure satisfies

$$(38) \quad \mathbb{1}_{\Delta_4} k_4(dt_1, dt_2, dt_3, dt_4) = \left\{ \int_{\mathbb{R}^2} \mathbb{W}[u](dx) \mathbb{W}[t_1, t_2, t_3, t_4](du) \right. \\ + \sum_{\sigma} \int_{\mathbb{R}^3} \mathbb{W}[u](dx) \mathbb{W}[v, t_{\sigma(3)}, t_{\sigma(4)}](du) \mathbb{W}[t_{\sigma(1)}, t_{\sigma(2)}](dv) \\ + \sum_{\sigma} \int_{\mathbb{R}^3} \mathbb{W}[u](dx) \mathbb{W}[v, t_{\sigma(4)}](du) \mathbb{W}[t_{\sigma(1)}, t_{\sigma(2)}, t_{\sigma(3)}](dv) \\ + \sum_{\sigma} \int_{\mathbb{R}^4} \mathbb{W}[u](dx) \mathbb{W}[v, w](du) \mathbb{W}[t_{\sigma(1)}, t_{\sigma(2)}](dv) \mathbb{W}[t_{\sigma(3)}, t_{\sigma(4)}](dw) \\ \left. + \sum_{\sigma} \int_{\mathbb{R}^4} \mathbb{W}[u](dx) \mathbb{W}[v, t_{\sigma(4)}](du) \mathbb{W}[w, t_{\sigma(3)}](dv) \mathbb{W}[t_{\sigma(1)}, t_{\sigma(2)}](dw) \right\} dt_1 dt_2 dt_3 dt_4,$$

where the sums are taken over the permutations of the terminal nodes within the equivalence classes **(b)**, **(c)**, **(d)**, **(e)** in Figure 6 respectively.

This method can be extended to $l \geq 5$ in order to compute cumulant measures of higher order. In the particular case of our work and as we will see further, we only need bounds on the cumulant measures of the first four orders. We deduce the following result that will be very useful in the sequel.

Lemma 9. *Let us consider the univariate Hawkes process N with rate $\lambda(t)$ introduced in (25) above. We take $h(s) = ae^{-bs} \mathbb{1}_{s>0}$ where $0 < a < b$, and write $\ell = a/b < 1$. Then the first and second order cumulants of N satisfy*

$$(39) \quad k_1(dt_1) = \frac{\mu}{1-\ell} dt_1,$$

$$(40) \quad \begin{aligned} \mathbb{1}_{\Delta_2} k_2(dt_1, dt_2) &= \frac{a\mu}{1-\ell} \left(1 + \frac{1}{2} \frac{a}{b-a} \right) e^{-b(1-\ell)|t_1-t_2|} dt_1 dt_2 \\ &= \frac{a\mu}{1-\ell} \left(1 + \frac{1}{2} \frac{\ell}{1-\ell} \right) e^{-b(1-\ell)|t_1-t_2|} dt_1 dt_2. \end{aligned}$$

More generally, for any $m = 1, 2, 3, 4$, there exists $C_m > 0$ such that we have the bound

$$(41) \quad \mathbb{1}_{\Delta_m} k_m(dt_1, \dots, dt_m) \leq C_m \frac{a^{m-1} \mu}{(1-\ell)^m} e^{-b(1-\ell)(\max_i t_i - \min_i t_i)} dt_1 \dots dt_m,$$

with the convention $\Delta_1 = \mathbb{R}$.

In this result, the inequality $m_1 \leq m_2$ between two positive measures holds in the sense that for any non negative test function ϕ we have $\int \phi(t) m_1(dt) \leq \int \phi(t) m_2(dt)$.

Main steps of the proof of Theorem 4. Recall that we want to use the criterion given in Lemma 1 in order to control Type II error of the test $\mathbb{1}_{\mathbf{C}_n(\mathbb{X}_n) > q_{1-\alpha}(\mathbb{X}_n)}$ by β . For that purpose, we need to compute Δ_φ and the variances of $f_\varphi(X_1, X_2)$ in those models. We then split the proof as follows.

Step 1: We first show that the variances of $f_\varphi(X_1, X_2)$ under each model (the observed and the independent) can be approximated by its variance \mathbb{V}_0 under the mean-field limit model. Under this very model, the spike trains are independent Poisson processes with intensity $\frac{\nu}{1-\ell}$ (see Delattre et al. [12]). We prove in the appendix that

$$(42) \quad \mathbb{V}_0 = \frac{4\nu^2 \delta T}{(1-\ell)^2} \left\{ 1 + \frac{2\nu\delta}{1-\ell} - \frac{\delta}{2T} \left(1 + \frac{10}{3} \frac{\nu\delta}{1-\ell} \right) \right\}.$$

Using Lemma 9, we deduce that there exists a constant $C > 0$ independent of the parameters such that

$$(43) \quad 0 \leq \mathbb{V}_\perp [f_\varphi(X_1, X_2)] - \mathbb{V}_0 \leq C \frac{aT}{M} \frac{\nu^2 \delta^2}{(1-\ell)^3} \left\{ 1 + \frac{\nu + \frac{a}{M}}{b(1-\ell)^2} \right\}$$

and

$$(44) \quad 0 \leq \mathbb{V} [f_\varphi(X_1, X_2)] - \mathbb{V}_0 \leq C \frac{aT}{M} \frac{\nu\delta}{(1-\ell)^2} \left\{ 1 + \frac{(\nu + \frac{a}{M})\delta}{1-\ell} \left[1 + \frac{\nu + \frac{a}{M}}{b(1-\ell)^2} \right] \right\}.$$

As a consequence, the difference between the variances of $f_\varphi(X_1, X_2)$ under the observed and independent models is proportional to $1/M$.

Step 2: We prove in the Appendix C.2 that Δ_φ is equal to

$$(45) \quad \Delta_\varphi = \mathbb{E}[\varphi(X_1^1, X_1^2)] - \mathbb{E}_\perp[\varphi(X_1^1, X_1^2)] = \frac{1}{M^2} \int_0^T \int_0^T \mathbb{1}_{|u-v| \leq \delta} \mathbb{1}_{u \neq v} k_2(du, dv)$$

where $1/M^2$ is the probability that the jump times u, v correspond to the first and the second coordinates of an observation. Using (40) in Lemma 9, we deduce that

$$(46) \quad \Delta_\varphi = \frac{\nu\delta}{M} \frac{2-\ell}{(1-\ell)^3} \ell \left\{ (T-\delta) \frac{1-e^{-b(1-\ell)\delta}}{\delta} + 1 - \frac{1-e^{-b(1-\ell)\delta}}{b(1-\ell)\delta} \right\}.$$

Supposing that $b(1 - \ell)T > 4$, using moreover that $\delta \leq T/2$, we obtain the lower bound

$$(47) \quad \Delta_\varphi \geq C \frac{\nu \ell T}{M(1 - \ell)^3} (1 - e^{-b(1 - \ell)\delta}),$$

for some constant not depending on the model parameters.

We conclude by replacing the variances in the criterion (5) of Lemma 1 by their upper bounds in the approximations (43)-(44). We solve the resulting inequality with respect to the unknown n and deduce the lower bound (14) on the number of observations that ensures the good control of the second kind error.

REFERENCES

- [1] M. Albert. Concentration inequalities for randomly permuted sums. In *High dimensional probability VIII—the Oaxaca volume*, volume 74 of *Progr. Probab.*, pages 341–383. Birkhäuser/Springer, Cham, 2019.
- [2] M. Albert, Y. Bouret, M. Fromont, and P. Reynaud-Bouret. Bootstrap and permutation tests of independence for point processes. *Ann. Statist.*, 43(6):2537–2564, 2015.
- [3] M. Albert, Y. Bouret, M. Fromont, and P. Reynaud-Bouret. Surrogate data methods based on a shuffling of the trials for synchrony detection: the centering issue. *Neural Comput.*, 28(11):2352–2392, 2016.
- [4] C. Amorino, A. Heidari, V. Pilipauskaitė, and M. Podolskij. Parameter estimation of discretely observed interacting particle systems. *Stochastic Process. Appl.*, 163:350–386, 2023.
- [5] E. Bacry, S. Gaïffas, I. Mastromatteo, and J.-F. Muzy. Mean-field inference of Hawkes point processes. *J. Phys. A*, 49(17):174006, 31, 2016.
- [6] Y. Baraud, S. Huet, and B. Laurent. Adaptive tests of linear hypotheses by model selection. *Ann. Statist.*, 31(1):225–251, 2003.
- [7] Y. Baraud, S. Huet, and B. Laurent. Adaptive tests of qualitative hypotheses. *ESAIM Probab. Stat.*, 7:147–159, 2003.
- [8] D. Belomestny, V. Pilipauskaitė, and M. Podolskij. Semiparametric estimation of McKean-Vlasov SDEs. *Ann. Inst. Henri Poincaré Probab. Stat.*, 59(1):79–96, 2023.
- [9] P. Brémaud and L. Massoulié. Stability of nonlinear Hawkes processes. *Ann. Probab.*, 24(3):1563–1588, 1996.
- [10] A. Date, E. Bienenstock, and S. Geman. On the temporal resolution of neural activity. Technical report, Division of Applied Mathematics, Brown University, 1998.
- [11] S. Delattre and N. Fournier. Statistical inference versus mean field limit for Hawkes processes. *Electron. J. Stat.*, 10(1):1223–1295, 2016.
- [12] S. Delattre, N. Fournier, and M. Hoffmann. Hawkes processes on large networks. *Ann. Appl. Probab.*, 26(1):216–261, 2016.
- [13] L. Della Maestra and M. Hoffmann. Nonparametric estimation for interacting particle systems: McKean-Vlasov models. *Probab. Theory Related Fields*, 182(1-2):551–613, 2022.
- [14] L. Della Maestra and M. Hoffmann. The LAN property for McKean-Vlasov models in a mean-field regime. *Stochastic Process. Appl.*, 155:109–146, 2023.
- [15] M. Fromont and B. Laurent. Adaptive goodness-of-fit tests in a density model. *Ann. Statist.*, 34(2):680–720, 2006.
- [16] M. Fromont, B. Laurent, and P. Reynaud-Bouret. Adaptive tests of homogeneity for a Poisson process. *Ann. Inst. Henri Poincaré Probab. Stat.*, 47(1):176–213, 2011.
- [17] M. Fromont, B. Laurent, and P. Reynaud-Bouret. The two-sample problem for Poisson processes: adaptive tests with a nonasymptotic wild bootstrap approach. *Ann. Statist.*, 41(3):1431–1461, 2013.
- [18] V. Genon-Catalot and C. Larédo. Asymptotic equivalence of nonparametric diffusion and Euler scheme experiments. *Ann. Statist.*, 42(3):1145–1165, 2014.
- [19] V. Genon-Catalot and C. Larédo. Probabilistic properties and parametric inference of small variance nonlinear self-stabilizing stochastic differential equations. *Stochastic Process. Appl.*, 142:513–548, 2021.
- [20] C. Graham. Regenerative properties of the linear Hawkes process with unbounded memory. *Ann. Appl. Probab.*, 31(6):2844–2863, 2021.
- [21] S. Grün, M. Diesmann, F. Grammont, A. Riehle, and A. Aertsen. Detecting unitary events without discretization of time. *J. Neurosci. Meth.*, 94(1):67–79, 1999.
- [22] A. G. Hawkes. Spectra of some self-exciting and mutually exciting point processes. *Biometrika*, 58:83–90, 1971.
- [23] A. G. Hawkes and D. Oakes. A cluster process representation of a self-exciting process. *J. Appl. Probability*, 11:493–503, 1974.
- [24] J. Jacod and A. N. Shiryaev. *Limit theorems for stochastic processes*, volume 288 of *Grundlehren der mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, second edition, 2003.
- [25] S. Jovanović, J. Hertz, and S. Rotter. Cumulants of hawkes point processes. *Phys. Rev. E*, 91:042802, Apr 2015.
- [26] R. A. Kasonga. Maximum likelihood theory for large interacting systems. *SIAM J. Appl. Math.*, 50(3):865–875, 1990.

- [27] I. Kim, S. Balakrishnan, and L. Wasserman. Minimax optimality of permutation tests. *Ann. Statist.*, 50(1):225–251, 2022.
- [28] E. L. Lehmann and J. P. Romano. *Testing statistical hypotheses*. Springer Texts in Statistics. Springer, Cham, fourth edition, [2021] ©2021.
- [29] F. Murtagh. Counting dendrograms: a survey. *Discrete Appl. Math.*, 7(2):191–199, 1984.
- [30] F. Pesarin and L. Salmaso. *Permutation Tests for Complex Data: Theory, Applications and Software*. Wiley, New-York, 2010.
- [31] G. Pipa and S. Grün. Non-parametric significance estimation of joint-spike events by shuffling and resampling. *Neurocomputing*, 52-54:31–37, 2003. Computational Neuroscience: Trends in Research 2003.
- [32] G. Pipa, S. Grün, and C. van Vreeswijk. Impact of spike train autostructure on probability distribution of joint spike events. *Neural Comput.*, 25(5):1123–1163, 2013.
- [33] N. Privault. Recursive computation of the Hawkes cumulants. *Statist. Probab. Lett.*, 177:Paper No. 109161, 11, 2021.
- [34] J. G. Rasmussen. Bayesian inference for Hawkes processes. *Methodol. Comput. Appl. Probab.*, 15(3):623–642, 2013.
- [35] A. Riehle, F. Grammont, M. Diesmann, and S. Grün. Dynamical changes and temporal precision of synchronized spiking activity in monkey motor cortex during movement preparation. *J. Physiol. (Paris)*, 94(5):569–582, 2000.
- [36] L. Sacerdote, M. Tamborrino, and C. Zucca. First passage times of two-dimensional correlated processes: Analytical results for the wiener process and a numerical method for diffusion processes. *Journal of Computational and Applied Mathematics*, 296:275–292, 2012.
- [37] W. Singer. Synchronization of cortical activity and its putative role in information processing and learning. *Annu. Rev. Physiol.*, 55(1):349–374, 1993.
- [38] C. Tuleau-Malot, A. Rouis, F. Grammont, and P. Reynaud-Bouret. Multiple tests based on a Gaussian approximation of the unitary events method with delayed coincidence count. *Neural Comput.*, 26(7):1408–1454, 2014.

APPENDIX A. CONTROL OF VARIANCES

This section is dedicated to the proof of Lemma 1. We start with a sketch of the proof and then detail the more technical parts in the next subsections.

A.1. Sketch of the Proof of Lemma 1. We easily show that given \mathbb{X}_n , the permuted statistics $\mathbf{T}_n(\mathbb{X}_n^{\Pi_n})$ is centered. Indeed, it follows from (15) that

$$\mathbb{E}[\mathbf{T}_n(\mathbb{X}_n^{\Pi_n})|\mathbb{X}_n] = \frac{1}{n(n-1)} \sum_{i \neq j} \mathbb{E}[\varphi(X_i^1, X_{\Pi_n(i)}^2) - \varphi(X_i^1, X_{\Pi_n(j)}^2)|\mathbb{X}_n] = 0,$$

since the laws of $(i, \Pi_n(i))$ and $(i, \Pi_n(j))$ are equal. Given this property, Kim et al. [27] showed that the type II error of the test $\mathbb{1}_{\mathbf{T}_n(\mathbb{X}_n) > q_{1-\alpha}(\mathbb{X}_n)}$ is controlled by β if

$$(48) \quad \mathbb{E}[\mathbf{T}_n(\mathbb{X}_n)] \geq \sqrt{\frac{2\mathbb{V}[\mathbf{T}_n(\mathbb{X}_n)]}{\beta}} + \sqrt{\frac{2\mathbb{E}[\mathbb{V}[\mathbf{T}_n(\mathbb{X}_n^{\Pi_n})|\mathbb{X}_n]]}{\alpha\beta}}.$$

Our aim is now to show that if the condition (5) is satisfied for a constant $C > 0$ and $n \geq 3/\sqrt{\alpha\beta}$, then the criterion (48) holds. We start with a presentation of the main steps of the proofs.

Step 1: First we prove that the variance of the test statistics $\mathbf{T}_n(\mathbb{X}_n)$ under the observed model is bounded by the one of $f_\varphi(X_1, X_2)$ through the inequality

$$(49) \quad \mathbb{V}[\mathbf{T}_n(\mathbb{X}_n)] \leq \frac{4}{n} \mathbb{V}[f_\varphi(X_1, X_2)].$$

Step 2: Using the fact that given \mathbb{X}_n the distribution of $\mathbf{T}_n(\mathbb{X}_n^{\Pi_n})$ is centered, we show that the expectation of the conditional variance of this permuted test statistics is mainly bounded by the variance of $f_\varphi(X_1, X_2)$ under independence with a rest. This bound is given as follows

$$(50) \quad \mathbb{E}\left(\mathbb{V}[\mathbf{T}_n(\mathbb{X}_n^{\Pi_n})|\mathbb{X}_n]\right) \leq \frac{2}{n} \left\{ \frac{2}{n} \mathbb{E}[|f_\varphi(X_1, X_2)|^2] + \frac{n-2}{n} \mathbb{V}_\perp[f_\varphi(X_1, X_2)] \right\}.$$

Step 3: We deduce from the bounds (49)-(50) that the right-hand term of (48) is upper bounded by

$$4\sqrt{\frac{2}{\alpha\beta}} \left\{ \sqrt{\frac{\mathbb{V}[f_\varphi(X_1, X_2)]}{n}} + \sqrt{\frac{\mathbb{V}_\perp[f_\varphi(X_1, X_2)]}{n}} \right\} + \frac{2}{n} \sqrt{\frac{2}{\alpha\beta}} \mathbb{E}[f_\varphi(X_1, X_2)].$$

Using (16) and $n \geq 3/\sqrt{\alpha\beta}$, we deduce a constant $C > 0$ for which the condition (5) implies that the criterion (48) holds, implying that the second kind error is bounded by β .

A.2. Proof of Inequality (49). This section is devoted to the proof of the bound

$$\mathbb{V}[\mathbf{T}_n(\mathbb{X}_n)] \leq \frac{4}{n} \mathbb{V}[f_\varphi(X_1, X_2)].$$

For that purpose, we deduce from (15) that

$$\mathbb{V}[\mathbf{T}_n(\mathbb{X}_n)] = \frac{1}{n^2(n-1)^2} \sum_{\substack{i \neq j \\ k \neq l}} \mathbb{E}[\{f_\varphi(X_i, X_j) - \mathbb{E}[f_\varphi(X_1, X_2)]\} \{f_\varphi(X_k, X_l) - \mathbb{E}[f_\varphi(X_1, X_2)]\}].$$

Thanks to the independence of the (X_i) , the terms in the previous sum with all different i, j, k, l are null. In addition, using Cauchy-Schwarz inequality,

$$\mathbb{E}[\{f_\varphi(X_i, X_j) - \mathbb{E}[f_\varphi(X_1, X_2)]\} \{f_\varphi(X_k, X_l) - \mathbb{E}[f_\varphi(X_1, X_2)]\}] \leq \mathbb{V}[f_\varphi(X_1, X_2)].$$

The number of non zero terms is at most

$$n^2(n-1)^2 - n(n-1)(n-2)(n-3) = n(n-1)(4n-6),$$

that ends the proof.

A.3. Proof of Ineq. (50). Since $\mathbf{T}_n(\mathbb{X}_n^{\Pi_n})$ is centered under the permutation law, its variance under this law is given by

$$(51) \quad \mathbb{V}[\mathbf{T}_n(\mathbb{X}_n^{\Pi_n}) | \mathbb{X}_n] = \frac{1}{n^2(n-1)^2} \sum_{\substack{i \neq j \\ k \neq l}} \mathbb{E} \left[f_\varphi(X_i^{\Pi_n}, X_j^{\Pi_n}) f_\varphi(X_k^{\Pi_n}, X_l^{\Pi_n}) \middle| \mathbb{X}_n \right].$$

Our aim is to compute precise bounds for the conditional expectations in this expression for $i \neq j$ and $k \neq l$. Let us first notice that

$$(52) \quad \mathbb{E} \left[f_\varphi(X_i^{\Pi_n}, X_j^{\Pi_n}) f_\varphi(X_k^{\Pi_n}, X_l^{\Pi_n}) \middle| \mathbb{X}_n \right] \\ = \sum_{\substack{i', j', k', l' \\ i' \neq j' \\ k' \neq l'}} \mathbb{P}(\Pi_n(i) = i', \Pi_n(j) = j', \Pi_n(k) = k', \Pi_n(l) = l') \\ \times \left\{ \varphi(X_i^1, X_{i'}^2) - \varphi(X_i^1, X_{j'}^2) \right\} \left\{ \varphi(X_k^1, X_{k'}^2) - \varphi(X_k^1, X_{l'}^2) \right\},$$

then it suffices to check the different possible choices of i, j, k, l and bound each of them thanks to the following lemma.

Lemma 10. *For any $i, i', j', k, k', l' \in \{1, \dots, n\}$ such that $i' \neq j'$ and $k' \neq l'$, we have the bound*

$$(53) \quad \mathbb{E} \left[\left\{ \varphi(X_i^1, X_{i'}^2) - \varphi(X_i^1, X_{j'}^2) \right\} \left\{ \varphi(X_k^1, X_{k'}^2) - \varphi(X_k^1, X_{l'}^2) \right\} \right] \\ \leq \begin{cases} \mathbb{E}[|f_\varphi(X_1, X_2)|^2] & \text{if } i \in \{i', j'\} \text{ and } k \in \{k', l'\}, \\ \mathbb{E}_\perp[|f_\varphi(X_1, X_2)|^2] & \text{if } i \notin \{i', j'\} \text{ and } k \notin \{k', l'\}, \\ \frac{1}{2} \{ \mathbb{E}[|f_\varphi(X_1, X_2)|^2] + \mathbb{E}_\perp[|f_\varphi(X_1, X_2)|^2] \} & \text{otherwise.} \end{cases}$$

This intermediate result directly follows from the inequality

$$\begin{aligned} \mathbb{E} \left[\left\{ \varphi(X_i^1, X_{i'}^2) - \varphi(X_i^1, X_{j'}^2) \right\} \left\{ \varphi(X_k^1, X_{k'}^2) - \varphi(X_k^1, X_{l'}^2) \right\} \right] \\ \leq \frac{1}{2} \left\{ \mathbb{E} \left[\left| \varphi(X_i^1, X_{i'}^2) - \varphi(X_i^1, X_{j'}^2) \right|^2 \right] + \mathbb{E} \left[\left| \varphi(X_k^1, X_{k'}^2) - \varphi(X_k^1, X_{l'}^2) \right|^2 \right] \right\} \end{aligned}$$

where we distinguished the cases $i \in \{i', j'\}$, $i \notin \{i', j'\}$, $k \in \{k', l'\}$ and $k \notin \{k', l'\}$. Returning to (52), we have seven cases depending on the possible values of the integers i, j, k, l :

Case 1 We assume that $i = k$ and $j = l$. In this case, the sum in (52) is such that $i' = k'$ and $j' = l'$. It follows that

$$\begin{aligned} \mathbb{E} \left[f_\varphi(X_i^{\Pi_n}, X_j^{\Pi_n}) f_\varphi(X_k^{\Pi_n}, X_l^{\Pi_n}) \middle| \mathbb{X}_n \right] &= \sum_{i' \neq j'} \mathbb{P}(\Pi_n(i) = i', \Pi_n(j) = j') \left\{ \varphi(X_i^1, X_{i'}^2) - \varphi(X_i^1, X_{j'}^2) \right\}^2 \\ &= \frac{1}{n(n-1)} \sum_{i' \neq j'} \left\{ \varphi(X_i^1, X_{i'}^2) - \varphi(X_i^1, X_{j'}^2) \right\}^2 \end{aligned}$$

and then thanks to (53) with $i = k, i' = k', j' = l'$,

$$\begin{aligned} \mathbb{E} \left[f_\varphi(X_i^{\Pi_n}, X_j^{\Pi_n}) f_\varphi(X_k^{\Pi_n}, X_l^{\Pi_n}) \right] &= \frac{1}{n(n-1)} \sum_{i' \neq j'} \mathbb{E} \left[\left\{ \varphi(X_i^1, X_{i'}^2) - \varphi(X_i^1, X_{j'}^2) \right\}^2 \right] \\ &\leq \frac{1}{n(n-1)} \sum_{i' \neq j'} \left\{ \mathbb{E} \left[\left| f_\varphi(X_1, X_2) \right|^2 \right] \mathbb{1}_{i=i' \text{ or } i=j'} + \mathbb{E}_\perp \left[\left| f_\varphi(X_1, X_2) \right|^2 \right] \mathbb{1}_{i \neq i' \text{ and } i \neq j'} \right\} \\ &= \frac{1}{n(n-1)} \left\{ 2(n-1) \mathbb{E} \left[\left| f_\varphi(X_1, X_2) \right|^2 \right] + (n-1)(n-2) \mathbb{E}_\perp \left[\left| f_\varphi(X_1, X_2) \right|^2 \right] \right\}, \end{aligned}$$

which finally gives

$$(54) \quad \mathbb{E} \left[f_\varphi(X_i^{\Pi_n}, X_j^{\Pi_n}) f_\varphi(X_k^{\Pi_n}, X_l^{\Pi_n}) \right] \leq \frac{2}{n} \mathbb{E} \left[\left| f_\varphi(X_1, X_2) \right|^2 \right] + \frac{n-2}{n} \mathbb{E}_\perp \left[\left| f_\varphi(X_1, X_2) \right|^2 \right].$$

Case 2 We now assume that $i = l$ and $j = k$. In this case, the sum in (52) is such that $i' = l'$ and $j' = k'$. We deduce that

$$\begin{aligned} \mathbb{E} \left[f_\varphi(X_i^{\Pi_n}, X_j^{\Pi_n}) f_\varphi(X_k^{\Pi_n}, X_l^{\Pi_n}) \middle| \mathbb{X}_n \right] &= \\ \sum_{i' \neq j'} \mathbb{P}(\Pi_n(i) = i', \Pi_n(j) = j') &\left\{ \varphi(X_i^1, X_{i'}^2) - \varphi(X_i^1, X_{j'}^2) \right\} \left\{ \varphi(X_j^1, X_{j'}^2) - \varphi(X_j^1, X_{i'}^2) \right\} \\ &= \frac{1}{n(n-1)} \sum_{i' \neq j'} \left\{ \varphi(X_i^1, X_{i'}^2) - \varphi(X_i^1, X_{j'}^2) \right\} \left\{ \varphi(X_j^1, X_{j'}^2) - \varphi(X_j^1, X_{i'}^2) \right\}, \end{aligned}$$

which implies thanks to (53) with $k = j, k' = j', l' = i'$ that

$$\begin{aligned}
 & \mathbb{E} \left[f_\varphi(X_i^{\Pi_n}, X_j^{\Pi_n}) f_\varphi(X_k^{\Pi_n}, X_l^{\Pi_n}) \right] \leq \frac{1}{n(n-1)} \times \\
 & \sum_{i' \neq j'} \left\{ \mathbb{E} [|f_\varphi(X_1, X_2)|^2] \mathbb{1}_{i \in \{i', j'\}} \mathbb{1}_{j \in \{i', j'\}} + \mathbb{E}_\perp [|f_\varphi(X_1, X_2)|^2] \mathbb{1}_{i \notin \{i', j'\}} \mathbb{1}_{j \notin \{i', j'\}} \right. \\
 & \left. + \frac{1}{2} \left(\mathbb{E} [|f_\varphi(X_1, X_2)|^2] + \mathbb{E}_\perp [|f_\varphi(X_1, X_2)|^2] \right) \times \right. \\
 & \quad \left. \times \left(1 - \mathbb{1}_{i \in \{i', j'\}} \mathbb{1}_{j \in \{i', j'\}} - \mathbb{1}_{i \notin \{i', j'\}} \mathbb{1}_{j \notin \{i', j'\}} \right) \right\} \\
 & = \frac{1}{n(n-1)} \left\{ 2\mathbb{E} [|f_\varphi(X_1, X_2)|^2] + (n-2)(n-3)\mathbb{E}_\perp [|f_\varphi(X_1, X_2)|^2] \right. \\
 & \left. + 2(n-2) \left(\mathbb{E} [|f_\varphi(X_1, X_2)|^2] + \mathbb{E}_\perp [|f_\varphi(X_1, X_2)|^2] \right) \right\}
 \end{aligned}$$

and finally gives

$$(55) \quad \mathbb{E} \left[f_\varphi(X_i^{\Pi_n}, X_j^{\Pi_n}) f_\varphi(X_k^{\Pi_n}, X_l^{\Pi_n}) \right] \leq \frac{2}{n} \mathbb{E} [|f_\varphi(X_1, X_2)|^2] + \frac{n-2}{n} \mathbb{E}_\perp [|f_\varphi(X_1, X_2)|^2].$$

Case 3 We assume that $i = k$ and $j \neq l$, then the sum in (52) is such that $i' = k'$ and $j' \neq l'$. We deduce that

$$\begin{aligned}
 & \mathbb{E} \left[f_\varphi(X_i^{\Pi_n}, X_j^{\Pi_n}) f_\varphi(X_k^{\Pi_n}, X_l^{\Pi_n}) \middle| \mathbb{X}_n \right] \\
 & = \sum_{\substack{i', j', l' \\ \text{different}}} \mathbb{P}(\Pi_n(i) = i', \Pi_n(j) = j', \Pi_n(l) = l') \\
 & \times \left\{ \varphi(X_i^1, X_{i'}^2) - \varphi(X_i^1, X_{j'}^2) \right\} \left\{ \varphi(X_i^1, X_{i'}^2) - \varphi(X_i^1, X_{l'}^2) \right\} \\
 & = \frac{1}{n(n-1)(n-2)} \sum_{\substack{i', j', l' \\ \text{different}}} \left\{ \varphi(X_i^1, X_{i'}^2) - \varphi(X_i^1, X_{j'}^2) \right\} \left\{ \varphi(X_i^1, X_{i'}^2) - \varphi(X_i^1, X_{l'}^2) \right\},
 \end{aligned}$$

and thanks to (53) with $k = i, k' = i'$, we have

$$\begin{aligned}
 & \mathbb{E} \left[f_\varphi(X_i^{\Pi_n}, X_j^{\Pi_n}) f_\varphi(X_k^{\Pi_n}, X_l^{\Pi_n}) \right] \\
 & \leq \frac{1}{n(n-1)(n-2)} \sum_{\substack{i', j', l' \\ \text{different}}} \left\{ \mathbb{E} [|f_\varphi(X_1, X_2)|^2] \mathbb{1}_{i \in \{i', j'\}} \mathbb{1}_{i \in \{i', l'\}} \right. \\
 & \left. + \mathbb{E}_\perp [|f_\varphi(X_1, X_2)|^2] \mathbb{1}_{i \notin \{i', j'\}} \mathbb{1}_{i \notin \{i', l'\}} + \frac{1}{2} \left(\mathbb{E} [|f_\varphi(X_1, X_2)|^2] + \mathbb{E}_\perp [|f_\varphi(X_1, X_2)|^2] \right) \right. \\
 & \quad \left. \times \left(1 - \mathbb{1}_{i \in \{i', j'\}} \mathbb{1}_{i \in \{i', l'\}} - \mathbb{1}_{i \notin \{i', j'\}} \mathbb{1}_{i \notin \{i', l'\}} \right) \right\} \\
 & = \frac{1}{n(n-1)(n-2)} \left\{ (n-1)(n-2)\mathbb{E} [|f_\varphi(X_1, X_2)|^2] \right. \\
 & \quad \left. + (n-1)(n-2)(n-3)\mathbb{E}_\perp [|f_\varphi(X_1, X_2)|^2] \right. \\
 & \quad \left. + (n-1)(n-2) \left(\mathbb{E} [|f_\varphi(X_1, X_2)|^2] + \mathbb{E}_\perp [|f_\varphi(X_1, X_2)|^2] \right) \right\},
 \end{aligned}$$

which implies

$$(56) \quad \mathbb{E} \left[f_\varphi(X_i^{\Pi_n}, X_j^{\Pi_n}) f_\varphi(X_k^{\Pi_n}, X_l^{\Pi_n}) \right] \leq \frac{2}{n} \mathbb{E} [|f_\varphi(X_1, X_2)|^2] + \frac{n-2}{n} \mathbb{E}_\perp [|f_\varphi(X_1, X_2)|^2].$$

Case 4 We assume that $i \neq k$ and $j = l$, then the sum in (52) is such that $i' \neq k'$ and $j' = l'$. It follows that

$$\begin{aligned}
& \mathbb{E} \left[f_\varphi(X_i^{\Pi_n}, X_j^{\Pi_n}) f_\varphi(X_k^{\Pi_n}, X_l^{\Pi_n}) \middle| \mathbb{X}_n \right] \\
&= \sum_{\substack{i', j', k' \\ \text{different}}} \mathbb{P}(\Pi_n(i) = i', \Pi_n(j) = j', \Pi_n(k) = k') \\
&\times \left\{ \varphi(X_i^1, X_{i'}^2) - \varphi(X_i^1, X_{j'}^2) \right\} \left\{ \varphi(X_k^1, X_{k'}^2) - \varphi(X_k^1, X_{j'}^2) \right\} \\
&= \frac{1}{n(n-1)(n-2)} \sum_{\substack{i', j', k' \\ \text{different}}} \left\{ \varphi(X_i^1, X_{i'}^2) - \varphi(X_i^1, X_{j'}^2) \right\} \left\{ \varphi(X_k^1, X_{k'}^2) - \varphi(X_k^1, X_{j'}^2) \right\},
\end{aligned}$$

and then

$$\begin{aligned}
(57) \quad & \mathbb{E} \left[f_\varphi(X_i^{\Pi_n}, X_j^{\Pi_n}) f_\varphi(X_k^{\Pi_n}, X_l^{\Pi_n}) \right] \\
&= \frac{1}{n(n-1)(n-2)} \sum_{\substack{i', j', k' \\ \text{different}}} \mathbb{E} \left[\left\{ \varphi(X_i^1, X_{i'}^2) - \varphi(X_i^1, X_{j'}^2) \right\} \left\{ \varphi(X_k^1, X_{k'}^2) - \varphi(X_k^1, X_{j'}^2) \right\} \right].
\end{aligned}$$

Case 5 We assume here that $i \neq l$ and $j = k$, then the sum in (52) is such that $i' \neq l'$ and $j' = k'$. Then,

$$\begin{aligned}
& \mathbb{E} \left[f_\varphi(X_i^{\Pi_n}, X_j^{\Pi_n}) f_\varphi(X_k^{\Pi_n}, X_l^{\Pi_n}) \middle| \mathbb{X}_n \right] \\
&= \sum_{\substack{i', j', l' \\ \text{different}}} \mathbb{P}(\Pi_n(i) = i', \Pi_n(j) = j', \Pi_n(l) = l') \\
&\times \left\{ \varphi(X_i^1, X_{i'}^2) - \varphi(X_i^1, X_{j'}^2) \right\} \left\{ \varphi(X_j^1, X_{j'}^2) - \varphi(X_j^1, X_{l'}^2) \right\} \\
&= \frac{1}{n(n-1)(n-2)} \sum_{\substack{i', j', l' \\ \text{different}}} \left\{ \varphi(X_i^1, X_{i'}^2) - \varphi(X_i^1, X_{j'}^2) \right\} \left\{ \varphi(X_j^1, X_{j'}^2) - \varphi(X_j^1, X_{l'}^2) \right\}
\end{aligned}$$

which finally gives

$$\begin{aligned}
(58) \quad & \mathbb{E} \left[f_\varphi(X_i^{\Pi_n}, X_j^{\Pi_n}) f_\varphi(X_k^{\Pi_n}, X_l^{\Pi_n}) \right] \\
&= \frac{1}{n(n-1)(n-2)} \sum_{\substack{i', j', l' \\ \text{different}}} \mathbb{E} \left[\left\{ \varphi(X_i^1, X_{i'}^2) - \varphi(X_i^1, X_{j'}^2) \right\} \left\{ \varphi(X_j^1, X_{j'}^2) - \varphi(X_j^1, X_{l'}^2) \right\} \right].
\end{aligned}$$

Remark 11. The sum of the expectations under **Case 4** given by (57) and **Case 5** given by (58) is null.

Case 6 We now assume that $i = l$ and $j \neq k$, then the sum in (52) is such that $i' = l'$ and $j' \neq k'$. We deduce that

$$\begin{aligned}
& \mathbb{E} \left[f_\varphi(X_i^{\Pi_n}, X_j^{\Pi_n}) f_\varphi(X_k^{\Pi_n}, X_l^{\Pi_n}) \middle| \mathbb{X}_n \right] \\
&= \sum_{\substack{i', j', k' \\ \text{different}}} \mathbb{P}(\Pi_n(i) = i', \Pi_n(j) = j', \Pi_n(k) = k') \\
&\times \left\{ \varphi(X_i^1, X_{i'}^2) - \varphi(X_i^1, X_{j'}^2) \right\} \left\{ \varphi(X_k^1, X_{k'}^2) - \varphi(X_k^1, X_{i'}^2) \right\} \\
&= \frac{1}{n(n-1)(n-2)} \sum_{\substack{i', j', k' \\ \text{different}}} \left\{ \varphi(X_i^1, X_{i'}^2) - \varphi(X_i^1, X_{j'}^2) \right\} \left\{ \varphi(X_k^1, X_{k'}^2) - \varphi(X_k^1, X_{i'}^2) \right\}
\end{aligned}$$

and then, thanks to (53) with $l' = i'$, we have

$$\begin{aligned}
& \mathbb{E} \left[f_\varphi(X_i^{\Pi_n}, X_j^{\Pi_n}) f_\varphi(X_k^{\Pi_n}, X_l^{\Pi_n}) \right] \\
&\leq \frac{1}{n(n-1)(n-2)} \sum_{\substack{i', j', k' \\ \text{different}}} \left\{ \mathbb{E} [|f_\varphi(X_1, X_2)|^2] \mathbb{1}_{i \in \{i', j'\}} \mathbb{1}_{k \in \{k', i'\}} \right. \\
&\quad \left. + \mathbb{E}_\perp [|f_\varphi(X_1, X_2)|^2] \mathbb{1}_{i \notin \{i', j'\}} \mathbb{1}_{k \notin \{k', i'\}} \right. \\
&\quad \left. + \frac{1}{2} \left(\mathbb{E} [|f_\varphi(X_1, X_2)|^2] + \mathbb{E}_\perp [|f_\varphi(X_1, X_2)|^2] \right) \right. \\
&\quad \left. \left(1 - \mathbb{1}_{i \in \{i', j'\}} \mathbb{1}_{k \in \{k', i'\}} - \mathbb{1}_{i \notin \{i', j'\}} \mathbb{1}_{k \notin \{k', i'\}} \right) \right\}.
\end{aligned}$$

In addition, we have

$$\sum_{\substack{i', j', k' \\ \text{different}}} \mathbb{1}_{i \in \{i', j'\}} \mathbb{1}_{k \in \{k', i'\}} = \sum_{\substack{i', j', k' \\ \text{different}}} (\mathbb{1}_{i'=i} \mathbb{1}_{k'=k} + \mathbb{1}_{j'=i} \mathbb{1}_{k'=k} + \mathbb{1}_{j'=i} \mathbb{1}_{i'=k}) = 3(n-2)$$

and

$$\begin{aligned}
& \sum_{\substack{i', j', k' \\ \text{different}}} \mathbb{1}_{i \notin \{i', j'\}} \mathbb{1}_{k \notin \{k', i'\}} = \sum_{\substack{i', j', k' \\ \text{different}}} \mathbb{1}_{i' \notin \{i, k\}} \mathbb{1}_{j' \neq i} \mathbb{1}_{k' \neq k} \\
&= \sum_{\substack{i', j', k' \\ \text{different}}} \mathbb{1}_{i' \notin \{i, k\}} \mathbb{1}_{j' \neq i} \mathbb{1}_{k' \neq k} (\mathbb{1}_{j'=k} \mathbb{1}_{k'=i} + \mathbb{1}_{j'=k} \mathbb{1}_{k' \neq i} + \mathbb{1}_{j' \neq k} \mathbb{1}_{k'=i} + \mathbb{1}_{j' \neq k} \mathbb{1}_{k' \neq i}) \\
&= \sum_{i'} \mathbb{1}_{i' \notin \{i, k\}} + \sum_{i' \neq k'} \mathbb{1}_{i', k' \notin \{i, k\}} + \sum_{i' \neq j'} \mathbb{1}_{i', j' \notin \{i, k\}} + \sum_{\substack{i', j', k' \\ \text{different}}} \mathbb{1}_{i', j', k' \notin \{i, k\}} \\
&= n-2 + 2(n-2)(n-3) + (n-2)(n-3)(n-4)
\end{aligned}$$

that gives

$$\sum_{\substack{i', j', k' \\ \text{different}}} \mathbb{1}_{i \notin \{i', j'\}} \mathbb{1}_{k \notin \{k', i'\}} = (n-2)(1 + (n-2)(n-3)).$$

It follows that

$$\begin{aligned}
& \mathbb{E} \left[f_\varphi(X_i^{\Pi_n}, X_j^{\Pi_n}) f_\varphi(X_k^{\Pi_n}, X_l^{\Pi_n}) \right] \\
& \leq \frac{1}{n(n-1)(n-2)} \left\{ 3(n-2) \mathbb{E} [|f_\varphi(X_1, X_2)|^2] \right. \\
& \quad \left. + (n-2)(1+(n-2)(n-3)) \mathbb{E}_\perp [|f_\varphi(X_1, X_2)|^2] \right. \\
& \quad \left. + (n-2)(2n-5) \left(\mathbb{E} [|f_\varphi(X_1, X_2)|^2] + \mathbb{E}_\perp [|f_\varphi(X_1, X_2)|^2] \right) \right\},
\end{aligned}$$

and then

$$(59) \quad \mathbb{E} \left[f_\varphi(X_i^{\Pi_n}, X_j^{\Pi_n}) f_\varphi(X_k^{\Pi_n}, X_l^{\Pi_n}) \right] \leq \frac{2}{n} \mathbb{E} [|f_\varphi(X_1, X_2)|^2] + \frac{n-2}{n} \mathbb{E}_\perp [|f_\varphi(X_1, X_2)|^2].$$

Case 7 We finally assume that $i \notin \{k, l\}$ and $j \notin \{k, l\}$, then the sum in (52) is such that $i' \notin \{k', l'\}$ and $j' \notin \{k', l'\}$. It follows that

$$\begin{aligned}
& \mathbb{E} \left[f_\varphi(X_i^{\Pi_n}, X_j^{\Pi_n}) f_\varphi(X_k^{\Pi_n}, X_l^{\Pi_n}) \middle| \mathbb{X}_n \right] \\
& = \sum_{\substack{i', j', k', l' \\ \text{different}}} \mathbb{P}(\Pi_n(i) = i', \Pi_n(j) = j', \Pi_n(k) = k', \Pi_n(l) = l') \\
& \times \left\{ \varphi(X_i^1, X_{i'}^2) - \varphi(X_i^1, X_{j'}^2) \right\} \left\{ \varphi(X_k^1, X_{k'}^2) - \varphi(X_k^1, X_{l'}^2) \right\} \\
& = \frac{1}{n(n-1)(n-2)(n-3)} \sum_{\substack{i', j', k', l' \\ \text{different}}} \left\{ \varphi(X_i^1, X_{i'}^2) - \varphi(X_i^1, X_{j'}^2) \right\} \\
& \qquad \qquad \qquad \left\{ \varphi(X_k^1, X_{k'}^2) - \varphi(X_k^1, X_{l'}^2) \right\}.
\end{aligned}$$

Distinguishing the cases $i' = i, j' = i$ and $i', j' \neq i$, we get

$$\begin{aligned}
 \mathbb{E} \left[f_\varphi(X_i^{\Pi_n}, X_j^{\Pi_n}) f_\varphi(X_k^{\Pi_n}, X_l^{\Pi_n}) \middle| \mathbb{X}_n \right] &= \frac{1}{n(n-1)(n-2)(n-3)} \times \\
 &\left[\sum_{\substack{i',j',k',l' \\ \text{different}}} \left\{ \varphi(X_i^1, X_{i'}^2) - \varphi(X_i^1, X_{j'}^2) \right\} \left\{ \varphi(X_k^1, X_{k'}^2) - \varphi(X_k^1, X_{l'}^2) \right\} \mathbb{1}_{i'=i} \right. \\
 &+ \sum_{\substack{i',j',k',l' \\ \text{different}}} \left\{ \varphi(X_i^1, X_{i'}^2) - \varphi(X_i^1, X_{j'}^2) \right\} \left\{ \varphi(X_k^1, X_{k'}^2) - \varphi(X_k^1, X_{l'}^2) \right\} \mathbb{1}_{j'=i} \\
 &+ \sum_{\substack{i',j',k',l' \\ \text{different}}} \left\{ \varphi(X_i^1, X_{i'}^2) - \varphi(X_i^1, X_{j'}^2) \right\} \left\{ \varphi(X_k^1, X_{k'}^2) - \varphi(X_k^1, X_{l'}^2) \right\} \mathbb{1}_{i',j' \neq i} \Big] \\
 &= \frac{1}{n(n-1)(n-2)(n-3)} \times \\
 &\left[\sum_{\substack{j',k',l' \\ \text{different} \\ j',k',l' \neq i}} \left\{ \varphi(X_i^1, X_{i'}^2) - \varphi(X_i^1, X_{j'}^2) \right\} \left\{ \varphi(X_k^1, X_{k'}^2) - \varphi(X_k^1, X_{l'}^2) \right\} \right. \\
 &+ \sum_{\substack{i',k',l' \\ \text{different} \\ i',k',l' \neq i}} \left\{ \varphi(X_i^1, X_{i'}^2) - \varphi(X_i^1, X_{j'}^2) \right\} \left\{ \varphi(X_k^1, X_{k'}^2) - \varphi(X_k^1, X_{l'}^2) \right\} \\
 &+ \sum_{\substack{i',j',k',l' \\ \text{different}}} \left\{ \varphi(X_i^1, X_{i'}^2) - \varphi(X_i^1, X_{j'}^2) \right\} \left\{ \varphi(X_k^1, X_{k'}^2) - \varphi(X_k^1, X_{l'}^2) \right\} \mathbb{1}_{i',j' \neq i} \Big].
 \end{aligned}$$

The two first sums compensates each other and we get

$$\begin{aligned}
 \mathbb{E} \left[f_\varphi(X_i^{\Pi_n}, X_j^{\Pi_n}) f_\varphi(X_k^{\Pi_n}, X_l^{\Pi_n}) \middle| \mathbb{X}_n \right] \\
 = \sum_{\substack{i',j',k',l' \\ \text{different}}} \left\{ \varphi(X_i^1, X_{i'}^2) - \varphi(X_i^1, X_{j'}^2) \right\} \left\{ \varphi(X_k^1, X_{k'}^2) - \varphi(X_k^1, X_{l'}^2) \right\} \mathbb{1}_{i',j' \neq i}.
 \end{aligned}$$

A similar reduction of this sum holds by distinguishing the cases: $i' = k, j' = k$ and $i', j' \neq k$; $k' = i, l' = i$ and $k', l' \neq i$; $k' = k, l' = k$ and $k', l' \neq k$. It follows that

$$\begin{aligned}
 \mathbb{E} \left[f_\varphi(X_i^{\Pi_n}, X_j^{\Pi_n}) f_\varphi(X_k^{\Pi_n}, X_l^{\Pi_n}) \middle| \mathbb{X}_n \right] &= \frac{1}{n(n-1)(n-2)(n-3)} \times \\
 &\sum_{\substack{i',j',k',l' \\ \text{different} \\ i',j',k',l' \notin \{i,k\}}} \left\{ \varphi(X_i^1, X_{i'}^2) - \varphi(X_i^1, X_{j'}^2) \right\} \left\{ \varphi(X_k^1, X_{k'}^2) - \varphi(X_k^1, X_{l'}^2) \right\},
 \end{aligned}$$

and we finally get

$$(60) \quad \mathbb{E} \left[f_\varphi(X_i^{\Pi_n}, X_j^{\Pi_n}) f_\varphi(X_k^{\Pi_n}, X_l^{\Pi_n}) \right] = 0.$$

We deduce from (54)–(60) and Remark 11 that

$$\begin{aligned} \mathbb{E} \left\{ \mathbb{V} [T_n(\mathbb{X}_n^{\Pi_n}) | \mathbb{X}_n] \right\} &= \frac{1}{n^2(n-1)^2} \\ &\sum_{\substack{i \neq j \\ k \neq l}} \left\{ \mathbb{E}^{\text{Case 1}} \left[f_\varphi(X_i^{\Pi_n}, X_j^{\Pi_n}) f_\varphi(X_k^{\Pi_n}, X_l^{\Pi_n}) \right] \mathbb{1}_{i=k, j=l} \right. \\ &+ \mathbb{E}^{\text{Case 2}} \left[f_\varphi(X_i^{\Pi_n}, X_j^{\Pi_n}) f_\varphi(X_k^{\Pi_n}, X_l^{\Pi_n}) \right] \mathbb{1}_{i=l, j=k} \\ &+ \mathbb{E}^{\text{Case 3}} \left[f_\varphi(X_i^{\Pi_n}, X_j^{\Pi_n}) f_\varphi(X_k^{\Pi_n}, X_l^{\Pi_n}) \right] \mathbb{1}_{i=k, j \neq l} \\ &\left. + \mathbb{E}^{\text{Case 6}} \left[f_\varphi(X_i^{\Pi_n}, X_j^{\Pi_n}) f_\varphi(X_k^{\Pi_n}, X_l^{\Pi_n}) \right] \mathbb{1}_{i=l, j \neq k} \right\}, \end{aligned}$$

and then

$$\begin{aligned} \mathbb{E} \left\{ \mathbb{V} [T_n(\mathbb{X}_n^{\Pi_n}) | \mathbb{X}_n] \right\} &\leq \\ &\frac{2}{n^2(n-1)^2} \left\{ n(n-1) \left[\frac{2}{n} \mathbb{E} [|f_\varphi(X_1, X_2)|^2] + \frac{n-2}{n} \mathbb{E}_\perp [|f_\varphi(X_1, X_2)|^2] \right] \right. \\ &+ n(n-1)(n-2) \left[\frac{2}{n} \mathbb{E} [|f_\varphi(X_1, X_2)|^2] + \frac{n-2}{n} \mathbb{E}_\perp [|f_\varphi(X_1, X_2)|^2] \right] \left. \right\} \\ &= \frac{2}{n} \left\{ \frac{2}{n} \mathbb{E} [|f_\varphi(X_1, X_2)|^2] + \frac{n-2}{n} \mathbb{E}_\perp [|f_\varphi(X_1, X_2)|^2] \right\}, \end{aligned}$$

which ends the proof.

A.4. Proof of Lemma 1. This proof consists in showing that the condition (48) – that allows to control the type II error by a given $\beta \in (0, 1)$ – is satisfied under the criterion of Lemma 1. For this purpose, let us first notice that from (49) and (50) proved in Appendix A.2 and A.3 respectively, we obtain

$$\begin{aligned} &\sqrt{\frac{2\mathbb{V}[\mathbf{T}_n(\mathbb{X}_n)]}{\beta}} + \sqrt{\frac{2\mathbb{E} \left[\mathbb{V} \left[\mathbf{T}_n(\mathbb{X}_n^{\Pi_n}) | \mathbb{X}_n \right] \right]}{\alpha\beta}} \\ &\leq 2\sqrt{\frac{2}{\beta}} \sqrt{\frac{\mathbb{V}[f_\varphi(X_1, X_2)]}{n}} + \frac{2}{\sqrt{\alpha\beta}} \sqrt{\frac{2}{n^2} \mathbb{E} [|f_\varphi(X_1, X_2)|^2] + \frac{n-2}{n^2} \mathbb{V}_\perp [f_\varphi(X_1, X_2)]} \\ &\leq 2\sqrt{\frac{2}{\alpha\beta}} \left(\sqrt{\alpha} + \frac{1}{\sqrt{n}} \right) \sqrt{\frac{\mathbb{V}[f_\varphi(X_1, X_2)]}{n}} + \frac{2}{\sqrt{\alpha\beta}} \sqrt{\frac{\mathbb{V}_\perp [f_\varphi(X_1, X_2)]}{n}} \\ &\quad + \frac{2}{n} \sqrt{\frac{2}{\alpha\beta}} \mathbb{E} [f_\varphi(X_1, X_2)], \end{aligned}$$

where we used the decomposition $\mathbb{E} [|f_\varphi(X_1, X_2)|^2] = \mathbb{V}[f_\varphi(X_1, X_2)] + (\mathbb{E}[f_\varphi(X_1, X_2)])^2$ and the inequality $\sqrt{a+b} \leq \sqrt{a} + \sqrt{b}$ for any $a, b \geq 0$. Since $\sqrt{\alpha} + 1/\sqrt{n} \leq 2$, it follows that

$$\begin{aligned} &\sqrt{\frac{2\mathbb{V}[\mathbf{T}_n(\mathbb{X}_n)]}{\beta}} + \sqrt{\frac{2\mathbb{E} \left[\mathbb{V} \left[\mathbf{T}_n(\mathbb{X}_n^{\Pi_n}) | \mathbb{X}_n \right] \right]}{\alpha\beta}} \\ &\leq 4\sqrt{\frac{2}{\alpha\beta}} \left\{ \sqrt{\frac{\mathbb{V}[f_\varphi(X_1, X_2)]}{n}} + \sqrt{\frac{\mathbb{V}_\perp [f_\varphi(X_1, X_2)]}{n}} \right\} + \frac{2}{n} \sqrt{\frac{2}{\alpha\beta}} \mathbb{E} [f_\varphi(X_1, X_2)]. \end{aligned}$$

In particular, if $n \geq 3/\sqrt{\alpha\beta}$, then

$$\begin{aligned} & \sqrt{\frac{2\mathbb{V}[\mathbf{T}_n(\mathbb{X}_n)]}{\beta}} + \sqrt{\frac{2\mathbb{E}\left[\mathbb{V}\left[\mathbf{T}_n(\mathbb{X}_n^{\Pi_n}) \mid \mathbb{X}_n\right]\right]}{\alpha\beta}} \\ & \leq 4\sqrt{\frac{2}{\alpha\beta}} \left\{ \sqrt{\frac{\mathbb{V}[f_\varphi(X_1, X_2)]}{n}} + \sqrt{\frac{\mathbb{V}_\perp[f_\varphi(X_1, X_2)]}{n}} \right\} + \frac{2\sqrt{2}}{3}\mathbb{E}[f_\varphi(X_1, X_2)]. \end{aligned}$$

We deduce that the condition (48) is satisfied if this upper bound is lower than $\mathbb{E}[\mathbf{T}_n(\mathbb{X}_n)] = \mathbb{E}[f_\varphi(X_1, X_2)]$, i.e. if

$$\mathbb{E}[f_\varphi(X_1, X_2)] \geq \frac{4}{1 - \frac{2\sqrt{2}}{3}} \sqrt{\frac{2}{\alpha\beta}} \left\{ \sqrt{\frac{\mathbb{V}[f_\varphi(X_1, X_2)]}{n}} + \sqrt{\frac{\mathbb{V}_\perp[f_\varphi(X_1, X_2)]}{n}} \right\}.$$

This corresponds to the condition (5) with $C = \frac{4\sqrt{2}}{1 - (2/3)\sqrt{2}}$.

APPENDIX B. THE JITTERING INJECTION MODEL

For the sake of simplicity, in what follows we use the same notation for a Poisson process and its sequence of jump times.

B.1. Proof of Eq. (21). We use the notations of Model 1. Notice that

$$(61) \quad \varphi(X^1, X^2) = \varphi(Y^1, Y^2) + \varphi(Y^1, Z^2) + \varphi(Z^1, Y^2) + \varphi(Z^1, Z^2).$$

Since Y^1, Y^2, Z^1 and $(\xi_k)_{k \in \mathbb{Z}}$ are independent, the expectations of $\varphi(Y^1, Y^2)$, $\varphi(Y^1, Z^2)$ and $\varphi(Z^1, Y^2)$ under the observed and the independent models do not change. In addition, we have

$$\begin{aligned} \varphi(Z^1, Z^2) &= \int_0^T \mathbb{1}_{|\xi_{Z_u^1}| \leq \delta} \mathbb{1}_{u + \xi_{Z_u^1} \in [0, T]} dZ_u^1 \\ &+ \int_{\Delta_2} \mathbb{1}_{|u_1 - u_2 - \xi_{Z_{u_2}^1}| \leq \delta} \mathbb{1}_{u_1 \in [0, T]} \mathbb{1}_{u_2 + \xi_{Z_{u_2}^1} \in [0, T]} dZ_{u_1}^1 dZ_{u_2}^1, \end{aligned}$$

where we recall that Δ_2 has been introduced in (19) above. Using Lemma 5 to calculate the expectation of the second integral, we get

$$\mathbb{E} \left[\int_{\Delta_2} \mathbb{1}_{|u_1 - u_2 - \xi_{Z_{u_2}^1}| \leq \delta} \mathbb{1}_{u_1 \in [0, T]} \mathbb{1}_{u_2 + \xi_{Z_{u_2}^1} \in [0, T]} dZ_{u_1}^1 dZ_{u_2}^1 \right] = \mathbb{E}_\perp [\varphi(Z^1, Z^2)].$$

Therefore,

$$\begin{aligned} \mathbb{E}[\varphi(X_1^1, X_1^2)] - \mathbb{E}_\perp[\varphi(X_1^1, X_1^2)] &= \mathbb{E} \left[\int_0^T \mathbb{1}_{|\xi_{Z_u^1}| \leq \delta} \mathbb{1}_{u + \xi_{Z_u^1} \in [0, T]} dZ_u^1 \right] \\ &= \eta^1 \mathbb{E} \left[\mathbb{1}_{|\xi| \leq \delta} \int_0^T \mathbb{1}_{u + \xi \in [0, T]} du \right] \\ (62) \quad &= \eta^1 \mathbb{E}[(T - |\xi|) \mathbb{1}_{|\xi| \leq \delta}]. \end{aligned}$$

The lower bound (22) follows by using the inequality $\delta \leq T/2$ which implies that $|\xi| \mathbb{1}_{|\xi| \leq \delta} \leq \frac{T}{2} \mathbb{1}_{|\xi| \leq \delta}$.

B.2. Proof of Eq. (23) and (24). Consider two counting processes N^1 and N^2 that almost surely do not jump together. Then we have

$$(63) \quad \varphi(N^1, N^2) = \int_0^T \left[N_{(u+\delta) \wedge T}^2 - N_{(u-\delta)_+}^2 \right] dN_u^1.$$

Remark 12. Notice that the integrand appearing in (63) is not predictable. Therefore we shall use this representation only in the case when N^1 and N^2 are independent.

We deduce the following result

Lemma 13. Let N^1, N^2, N^3 be three independent Poisson processes on $[0, T]$ with intensities μ^1, μ^2 and μ^3 respectively. We introduce the quantities

$$(64) \quad I_\delta^T = 2\delta T - \delta^2 \quad \text{and} \quad J_\delta^T = 4\delta^2 T - \frac{10}{3}\delta^3.$$

Then we have

$$(65) \quad \mathbb{E}[\varphi(N^1, N^2)] = \mu^1 \mu^2 I_\delta^T,$$

$$(66) \quad \mathbb{V}[\varphi(N^1, N^2)] = \mu^1 \mu^2 [(\mu^1 + \mu^2)J_\delta^T + I_\delta^T],$$

$$(67) \quad \mathbb{E}[\varphi(N^1, N^2)\varphi(N^1, N^3)] = \mu^1 \mu^2 \mu^3 J_\delta^T + (\mu^1)^2 \mu^2 \mu^3 (I_\delta^T)^2.$$

As a consequence, if $\mu^3 = \mu^2$, then we obtain

$$(68) \quad \mathbb{E} \left\{ [\varphi(N^1, N^2) - \varphi(N^1, N^3)]^2 \right\} = 2\mu^1 \mu^2 [I_\delta^T + \mu^1 J_\delta^T].$$

Remark 14. The quantities I_δ^T and J_δ^T correspond to the mean and variance of the conditional expectation $\mathbb{E}[\varphi(N^1, N^2)|N^2]$ in the particular case $\mu^1 = \mu^2 = 1$.

Proof. Step 1. We start with some preliminary remarks. Consider the measurable functions defined for $(u, v) \in \mathbb{R}^2$ by

$$(69) \quad w(u, v) = \mathbb{1}_{|u-v| \leq \delta} \mathbb{1}_{u \in [0, T]} \mathbb{1}_{v \in [0, T]},$$

$$(70) \quad \tilde{w}(u) := \int_{\mathbb{R}} w(u, v) dv = [(u + \delta) \wedge T - (u - \delta)_+] \mathbb{1}_{u \in [0, T]}.$$

Let N be a Poisson process with intensity 1. Then we have that

$$\int_0^T \mathbb{E}[N_{(u+\delta) \wedge T} - N_{(u-\delta)_+}] du = \int_{\mathbb{R}} \tilde{w}(u) du = I_\delta^T.$$

In particular, this implies that, since N^1 and N^2 are independent,

$$\mathbb{E}[\varphi(N^1, N^2)] = \mu^1 \int_0^T \mathbb{E}[N_{(u+\delta) \wedge T}^2 - N_{(u-\delta)_+}^2] du = \mu^1 \mu^2 \int_0^T \tilde{w}(u) du = \mu^1 \mu^2 I_\delta^T.$$

We shall also use throughout the remainder of this proof that

$$J_\delta^T = \int_0^T \tilde{w}(u)^2 du,$$

which can be easily checked.

Step 2. (63) implies

$$\mathbb{E}[\varphi^2(N^1, N^2)] = \mathbb{E} \left[\left(\int_0^T [N_{(u+\delta) \wedge T}^2 - N_{(u-\delta)_+}^2] dN_u^1 \right)^2 \right].$$

Conditioning on N^2 , it follows that

$$\begin{aligned} & \mathbb{E}[\varphi^2(N^1, N^2)] \\ &= \mu^1 \int_0^T \mathbb{E} \left\{ [N_{(u+\delta) \wedge T}^2 - N_{(u-\delta)_+}^2]^2 \right\} du + (\mu^1)^2 \mathbb{E} \left[\left(\int_0^T [N_{(u+\delta) \wedge T}^2 - N_{(u-\delta)_+}^2] du \right)^2 \right] \\ &= \mu^1 \int_0^T \{ (\mu^2)^2 \tilde{w}^2(u) + \mu^2 \tilde{w}(u) \} du \\ & \quad + (\mu^1)^2 \int_0^T \int_0^T \mathbb{E} \left\{ [N_{(u+\delta) \wedge T}^2 - N_{(u-\delta)_+}^2] [N_{(v+\delta) \wedge T}^2 - N_{(v-\delta)_+}^2] \right\} dudv. \end{aligned}$$

According to Step 0, the first expression appearing above can be rewritten as

$$\mu^1 (\mu^2)^2 J_\delta^T + \mu^1 \mu^2 I_\delta^T.$$

Moreover, studying the last expression appearing above and distinguishing the cases when the two intervals $[u - \delta, u + \delta]$, $[v - \delta, v + \delta]$ (intersected with $[0, T]$) do or do not overlap, we get

$$\begin{aligned} & (\mu^1)^2 \int_0^T \int_0^T \mathbb{E} \left\{ \left[N_{(u+\delta)\wedge T}^2 - N_{(u-\delta)_+}^2 \right] \left[N_{(v+\delta)\wedge T}^2 - N_{(v-\delta)_+}^2 \right] \right\} dudv \\ &= (\mu^1)^2 \int_0^T \int_0^T \left\{ (\mu^2)^2 \tilde{w}(u) \tilde{w}(v) + \mu^2 \int_0^T \mathbb{1}_{|v-s|\leq\delta} \mathbb{1}_{|u-s|\leq\delta} ds \right\} dudv \\ &= (\mu^1 \mu^2)^2 (I_\delta^T)^2 + (\mu^1)^2 \mu^2 J_\delta^T. \end{aligned}$$

Putting everything together, this implies that

$$\mathbb{E}[\varphi^2(N^1, N^2)] = \mu^1 \mu^2 [(\mu^1 + \mu^2) J_\delta^T + I_\delta^T + \mu^1 \mu^2 (I_\delta^T)^2].$$

Step 3. Using (63) again, we have

$$\begin{aligned} & \mathbb{E}[\varphi(N^1, N^2) \varphi(N^1, N^3)] \\ &= \mathbb{E} \left[\left(\int_0^T \left[N_{(u+\delta)\wedge T}^2 - N_{(u-\delta)_+}^2 \right] dN_u^1 \right) \left(\int_0^T \left[N_{(u+\delta)\wedge T}^3 - N_{(u-\delta)_+}^3 \right] dN_u^1 \right) \right], \end{aligned}$$

such that, conditioning on N^2 and N^3 ,

$$\begin{aligned} & \mathbb{E}[\varphi(N^1, N^2) \varphi(N^1, N^3)] \\ &= \mu^1 \int_0^T \mathbb{E} \left\{ \left[N_{(u+\delta)\wedge T}^2 - N_{(u-\delta)_+}^2 \right] \left[N_{(u+\delta)\wedge T}^3 - N_{(u-\delta)_+}^3 \right] \right\} du \\ &+ (\mu^1)^2 \mathbb{E} \left\{ \left(\int_0^T \left[N_{(u+\delta)\wedge T}^2 - N_{(u-\delta)_+}^2 \right] du \right) \left(\int_0^T \left[N_{(u+\delta)\wedge T}^3 - N_{(u-\delta)_+}^3 \right] du \right) \right\}. \end{aligned}$$

Since N_2 and N_3 are independent, we deduce that

$$\begin{aligned} & \mathbb{E}[\varphi(N^1, N^2) \varphi(N^1, N^3)] \\ &= \mu^1 \int_0^T \left(\mathbb{E} \left[N_{(u+\delta)\wedge T}^2 - N_{(u-\delta)_+}^2 \right] \right) \left(\mathbb{E} \left[N_{(u+\delta)\wedge T}^3 - N_{(u-\delta)_+}^3 \right] \right) du \\ &+ (\mu^1)^2 \mathbb{E} \left(\int_0^T \left[N_{(u+\delta)\wedge T}^2 - N_{(u-\delta)_+}^2 \right] du \right) \mathbb{E} \left(\int_0^T \left[N_{(u+\delta)\wedge T}^3 - N_{(u-\delta)_+}^3 \right] du \right) \\ &= \mu^1 \mu^2 \mu^3 J_\delta^T + (\mu^1)^2 \mu^2 \mu^3 (I_\delta^T)^2. \end{aligned}$$

Taking in particular $\mu^3 = \mu^2$, it follows that

$$\begin{aligned} & \mathbb{E} \left\{ [\varphi(N^1, N^2) - \varphi(N^1, N^3)]^2 \right\} = 2 \left\{ \mathbb{E}[\varphi^2(N^1, N^2)] - \mathbb{E}[\varphi(N^1, N^2) \varphi(N^1, N^3)] \right\} \\ &= 2\mu^1 \mu^2 [\mu^1 J_\delta^T + I_\delta^T]. \end{aligned}$$

□

Let us come back to the jittering injection Poisson model given by Model 1. Starting with the variance under the independent model, we have

$$\mathbb{V}_\perp [f_\varphi(X_1, X_2)] = \mathbb{E}_\perp \left[|\varphi(X_1^1, X_1^2) - \varphi(X_1^1, X_2^2)|^2 \right]$$

that implies thanks to (68) in Lemma 13 that

$$(71) \quad \mathbb{V}_\perp [f_\varphi(X_1, X_2)] = 2(\eta^1 + \lambda^1)(\eta^1 + \lambda^2) [I_\delta^T + (\eta^1 + \lambda^1) J_\delta^T].$$

The bound (23) follows thanks to the inequalities $\eta^1 \leq \lambda^{\mathbf{m}} = \max\{\lambda^1, \lambda^2\}$, $I_\delta^T \leq 2\delta T$ and $J_\delta^T \leq 4\delta^2 T$. Further, the decomposition (61) gives

$$(72) \quad f_\varphi(X_1, X_2) = \Delta^{(1)} + \Delta^{(2)} + \Delta^{(3)} + \Delta^{(4)}$$

where we set

$$\begin{aligned}\Delta^{(1)} &= \varphi(Y_1^1, Y_1^2) - \varphi(Y_1^1, Y_2^2), \\ \Delta^{(2)} &= \varphi(Y_1^1, Z_1^2) - \varphi(Y_1^1, Z_2^2), \\ \Delta^{(3)} &= \varphi(Z_1^1, Y_1^2) - \varphi(Z_1^1, Y_2^2), \\ \Delta^{(4)} &= \varphi(Z_1^1, Z_1^2) - \varphi(Z_1^1, Z_2^2).\end{aligned}$$

This implies that

$$(73) \quad \mathbb{E}[|f_\varphi(X_1, X_2)|^2] = \mathbb{E}[|\Delta^{(1)}|^2] + \mathbb{E}[|\Delta^{(2)}|^2] + \mathbb{E}[|\Delta^{(3)}|^2] + \mathbb{E}[|\Delta^{(4)}|^2] + 2 \sum_{i < j} \mathbb{E}[\Delta^{(i)} \Delta^{(j)}].$$

We deduce from (68) in Lemma 13 that

$$\begin{aligned}\mathbb{E}[|\Delta^{(1)}|^2] &= 2\lambda^1 \lambda^2 [\lambda^1 J_\delta^T + I_\delta^T] \\ \mathbb{E}[|\Delta^{(2)}|^2] &= 2\lambda^1 \eta^1 [\lambda^1 J_\delta^T + I_\delta^T] \\ \mathbb{E}[|\Delta^{(3)}|^2] &= 2\lambda^2 \eta^1 [\eta^1 J_\delta^T + I_\delta^T].\end{aligned}$$

In addition, conditioning on Y_1^1 gives

$$\mathbb{E}[\Delta^{(1)} \Delta^{(2)}] = 0,$$

and using (67) in Lemma 13, we obtain, due to the independence of Y^1 and Y^2 and of Z^1 and Y^2 ,

$$\begin{aligned}\mathbb{E}[\Delta^{(1)} \Delta^{(3)}] &= \mathbb{E}[(\varphi(Y_1^1, Y_1^2) - \varphi(Y_1^1, Y_2^2)) (\varphi(Z_1^1, Y_1^2) - \varphi(Z_1^1, Y_2^2))] \\ &= \mathbb{E}[\varphi(Y_1^1, Y_1^2) \varphi(Z_1^1, Y_1^2)] - 2\mathbb{E}[\varphi(Y_1^1, Y_1^2)] \mathbb{E}[\varphi(Z_1^1, Y_2^2)] + \mathbb{E}[\varphi(Y_1^1, Y_2^2) \varphi(Z_1^1, Y_2^2)] \\ &= \eta^1 \lambda^1 \lambda^2 J_\delta^T + \eta^1 \lambda^1 (\lambda^2)^2 (I_\delta^T)^2 - 2\eta^1 \lambda^1 (\lambda^2)^2 (I_\delta^T)^2 + \eta^1 \lambda^1 \lambda^2 J_\delta^T + \eta^1 \lambda^1 (\lambda^2)^2 (I_\delta^T)^2,\end{aligned}$$

which implies that

$$\mathbb{E}[\Delta^{(1)} \Delta^{(3)}] = 2\eta^1 \lambda^1 \lambda^2 J_\delta^T.$$

Further, it directly follows from the independence of (Y_1^1, Y_1^2, Y_2^2) and (Z_1^1, Z_1^2, Z_2^2) that

$$\mathbb{E}[\Delta^{(1)} \Delta^{(4)}] = 0.$$

We also have

$$\begin{aligned}\mathbb{E}[\Delta^{(2)} \Delta^{(3)}] &= \mathbb{E}[(\varphi(Y_1^1, Z_1^2) - \varphi(Y_1^1, Z_2^2)) (\varphi(Z_1^1, Y_1^2) - \varphi(Z_1^1, Y_2^2))] \\ &= \mathbb{E}[\varphi(Y_1^1, Z_1^2) (\varphi(Z_1^1, Y_1^2) - \varphi(Z_1^1, Y_2^2))] - \mathbb{E}[\varphi(Y_1^1, Z_2^2) (\varphi(Z_1^1, Y_1^2) - \varphi(Z_1^1, Y_2^2))].\end{aligned}$$

Conditioning on (Z_1^1, Z_1^2) in the first expectation and using the independence of (Y_1^1, Z_2^2) and (Z_1^1, Y_1^2, Y_2^2) to handle the second expression, we get

$$\mathbb{E}[\Delta^{(2)} \Delta^{(3)}] = 0.$$

Similarly, notice that

$$\mathbb{E}[\Delta^{(3)} \Delta^{(4)}] = \mathbb{E}[(\varphi(Z_1^1, Y_1^2) - \varphi(Z_1^1, Y_2^2)) (\varphi(Z_1^1, Z_1^2) - \varphi(Z_1^1, Z_2^2))].$$

By conditioning on (Z_1^1, Z_1^2, Z_2^2) , we get

$$\mathbb{E}[\Delta^{(3)} \Delta^{(4)}] = 0.$$

Then Eq. (73) becomes

$$(74) \quad \mathbb{E}[|f_\varphi(X_1, X_2)|^2] = 2[(\eta^1 + \lambda^1)(\eta^1 + \lambda^2) - (\eta^1)^2] I_\delta^T + 2[(\eta^1 + \lambda^1)^2(\eta^1 + \lambda^2) - (\eta^1)^2(\eta^1 + 2\lambda^1)] J_\delta^T + \mathbb{E}[|\Delta^{(4)}|^2] + 2\mathbb{E}[\Delta^{(2)} \Delta^{(4)}].$$

The computation of the expectations of $\Delta^{(2)} \Delta^{(4)}$ and $|\Delta^{(4)}|^2$ needs another representation of the number of coincidences, different of what we have used in (63), since now the two processes that are involved are not independent any more. Recalling the definition of w and of \tilde{w} above, we have:

Lemma 15.

$$(75) \quad \varphi(Z_1^1, Z_1^2) = \int_{\mathbb{R}} w(u, u + \xi_{Z_1^1(u^-)+1}) dZ_1^1(u) \\ + \int_{\Delta_2} w(u_1, u_2 + \xi_{Z_1^1(u_2^-)+1}) dZ_1^1(u_1) dZ_1^1(u_2).$$

In particular

$$(76) \quad \mathbb{E} [\varphi(Z_1^1, Z_1^2)] = \eta^1 \mathbb{E} [(T - |\xi|) \mathbb{1}_{|\xi| \leq \delta}] + (\eta^1)^2 I_\delta^T,$$

and

$$(77) \quad \mathbb{E} [\Delta^{(4)} | Z_1^1, Z_1^2] = \int_{\mathbb{R}} [w(u, u + \xi_{Z_1^1(u^-)+1}) - \eta^1 \tilde{w}(u)] dZ_1^1(u) \\ + \int_{\Delta_2} w(u_1, u_2 + \xi_{Z_1^1(u_2^-)+1}) dZ_1^1(u_1) dZ_1^1(u_2).$$

Proof. Equation (75) follows from the representation

$$\varphi(Z_1^1, Z_1^2) = \int_{\mathbb{R}^2} w(u_1, u_2 + \xi_{Z_1^1(u_2^-)+1}) dZ_1^1(u_1) dZ_1^1(u_2),$$

where we distinguished the case for which $u_1 = u_2$ from the case $u_1 \neq u_2$. Using Lemma 5 and the arguments of section B.1, the equation (76) follows immediately.

Finally, we have

$$\mathbb{E} [\Delta^{(4)} | Z_1^1, Z_1^2] = \mathbb{E} [\varphi(Z_1^1, Z_1^2) - \varphi(Z_1^1, Z_2^2) | Z_1^1, Z_1^2] \\ = \varphi(Z_1^1, Z_1^2) - \mathbb{E} [\varphi(Z_1^1, Z_2^2) | Z_1^1] \\ = \varphi(Z_1^1, Z_1^2) - \eta^1 \int_{\mathbb{R}} \tilde{w}(u) dZ_1^1(u),$$

which implies (77) thanks to (75). □

Using (67) in Lemma 13 and equation (76), we get

$$\mathbb{E} [\Delta^{(2)} \Delta^{(4)}] = \mathbb{E} [(\varphi(Y_1^1, Z_1^2) - \varphi(Y_1^1, Z_2^2)) \Delta^{(4)}] \\ = \mathbb{E} [\varphi(Y_1^1, Z_1^2) \Delta^{(4)}] - \mathbb{E} [\varphi(Y_1^1, Z_2^2)] \mathbb{E} [\varphi(Z_1^1, Z_1^2)] \\ + \mathbb{E} [\varphi(Y_1^1, Z_2^2) \varphi(Z_1^1, Z_2^2)] \\ = \mathbb{E} [\varphi(Y_1^1, Z_1^2) \Delta^{(4)}] - \lambda^1 \eta^1 I_\delta^T ((\eta^1)^2 I_\delta^T + \eta^1 \mathbb{E} [(T - |\xi|) \mathbb{1}_{|\xi| \leq \delta}]) \\ + (\eta^1)^2 \lambda^1 J_\delta^T + (\eta^1)^3 \lambda^1 (I_\delta^T)^2$$

and then,

$$(78) \quad \mathbb{E} [\Delta^{(2)} \Delta^{(4)}] = \mathbb{E} [\varphi(Y_1^1, Z_1^2) \Delta^{(4)}] + (\eta^1)^2 \lambda^1 J_\delta^T - \lambda^1 (\eta^1)^2 I_\delta^T \mathbb{E} [(T - |\xi|) \mathbb{1}_{|\xi| \leq \delta}].$$

Conditioning on (Z_1^1, Z_1^2) and using (77), we obtain

$$\begin{aligned}
& \mathbb{E} \left[\varphi(Y_1^1, Z_1^2) \Delta^{(4)} \right] \\
&= \mathbb{E} \left\{ \mathbb{E} \left[\varphi(Y_1^1, Z_1^2) | Z_1^1 \right] \mathbb{E} \left[\Delta^{(4)} | Z_1^1, Z_1^2 \right] \right\} \\
&= \lambda^1 \mathbb{E} \left\{ \left(\int_{\mathbb{R}} \tilde{w}(u + \xi_{Z_1^1(u-)+1}) dZ_1^1(u) \right) \right. \\
&\quad \times \left(\int_{\mathbb{R}} \left[w(u, u + \xi_{Z_1^1(u-)+1}) - \eta^1 \tilde{w}(u) \right] dZ_1^1(u) \right. \\
&\quad \quad \left. \left. + \int_{\Delta_2} w(u_1, u_2 + \xi_{Z_1^1(u_2-)+1}) dZ_1^1(u_1) dZ_1^1(u_2) \right) \right\} \\
&= \lambda^1 \mathbb{E} \left\{ \int_{\mathbb{R}^2} \left[w(u_1, u_1 + \xi_{Z_1^1(u_1-)-}) - \eta^1 \tilde{w}(u_1) \right] \tilde{w}(u_2 + \xi_{Z_1^1(u_2-)-}) dZ_1^1(u_1) dZ_1^1(u_2) \right. \\
&\quad \left. + \int_{\mathbb{R}} \left(\int_{\Delta_2} w(u_1, u_2 + \xi_{Z_1^1(u_2-)+1}) \tilde{w}(u_3 + \xi_{Z_1^1(u_3-)+1}) dZ_1^1(u_1) dZ_1^1(u_2) \right) dZ_1^1(u_3) \right\},
\end{aligned}$$

which implies, by distinguishing common jump times,

$$\begin{aligned}
\mathbb{E} \left[\varphi(Y_1^1, Z_1^2) \Delta^{(4)} \right] &= \lambda^1 \mathbb{E} \left\{ \int_{\mathbb{R}} \left[w(u, u + \xi_{Z_1^1(u-)+1}) - \eta^1 \tilde{w}(u) \right] \tilde{w}(u + \xi_{Z_1^1(u-)+1}) dZ_1^1(u) \right. \\
&\quad + \int_{\Delta_2} \left[w(u_1, u_1 + \xi_{Z_1^1(u_1-)+1}) - \eta^1 \tilde{w}(u_1) \right] \tilde{w}(u_2 + \xi_{Z_1^1(u_2-)+1}) dZ_1^1(u_1) dZ_1^1(u_2) \\
&\quad + \int_{\Delta_2} w(u_1, u_2 + \xi_{Z_1^1(u_2-)+1}) \tilde{w}(u_1 + \xi_{Z_1^1(u_1-)+1}) dZ_1^1(u_1) dZ_1^1(u_2) \\
&\quad + \int_{\Delta_2} w(u_1, u_2 + \xi_{Z_1^1(u_2-)+1}) \tilde{w}(u_2 + \xi_{Z_1^1(u_2-)+1}) dZ_1^1(u_1) dZ_1^1(u_2) \\
&\quad \left. + \int_{\Delta_3} w(u_1, u_2 + \xi_{Z_1^1(u_2-)+1}) \tilde{w}(u_3 + \xi_{Z_1^1(u_3-)+1}) dZ_1^1(u_1) dZ_1^1(u_2) dZ_1^1(u_3) \right\}.
\end{aligned}$$

Then it follows from Lemma 5 that

$$\begin{aligned}
\mathbb{E} \left[\varphi(Y_1^1, Z_1^2) \Delta^{(4)} \right] &= \lambda^1 \mathbb{E} \left\{ \eta^1 \int_{\mathbb{R}} \left[w(u, u + \xi) - \eta^1 \tilde{w}(u) \right] \tilde{w}(u + \xi) du \right. \\
&\quad + (\eta^1)^2 \int_{\mathbb{R}^2} \left[w(u_1, u_1 + \xi_1) - \eta^1 \tilde{w}(u_1) \right] \tilde{w}(u_2 + \xi_2) du_1 du_2 \\
&\quad + (\eta^1)^2 \int_{\mathbb{R}^2} w(u_1, u_2 + \xi_1) \tilde{w}(u_1 + \xi_2) du_1 du_2 \\
&\quad + (\eta^1)^2 \int_{\mathbb{R}^2} w(u_1, u_2 + \xi) \tilde{w}(u_2 + \xi) du_1 du_2 \\
&\quad \left. + (\eta^1)^3 \int_{\mathbb{R}^3} w(u_1, u_2 + \xi_2) \tilde{w}(u_3 + \xi_3) du_1 du_2 du_3 \right\} \\
&= \lambda^1 \mathbb{E} \left\{ \eta^1 \int_{\mathbb{R}} \left[w(u, u + \xi) - \eta^1 \tilde{w}(u) \right] \tilde{w}(u + \xi) du \right. \\
&\quad + (\eta^1)^2 I_\delta^T \left[\int_{\mathbb{R}} w(u_1, u_1 + \xi_1) du_1 - \eta^1 I_\delta^T \right] \\
&\quad \left. + (\eta^1)^2 \int_{\mathbb{R}} \tilde{w}(u_1) \tilde{w}(u_1 + \xi_2) du_1 + (\eta^1)^2 J_\delta^T + (\eta^1)^3 (I_\delta^T)^2 \right\}
\end{aligned}$$

which finally gives

$$\begin{aligned} \mathbb{E} \left[\varphi(Y_1^1, Z_1^2) \Delta^{(4)} \right] &= \lambda^1 (\eta^1)^2 J_\delta^T + \lambda^1 (\eta^1)^2 I_\delta^T \mathbb{E} \left[(T - |\xi|) \mathbb{1}_{|\xi| \leq \delta} \right] \\ &\quad + \lambda^1 \eta^1 \mathbb{E} \left\{ \int_{\mathbb{R}} w(u, u + \xi) \tilde{w}(u + \xi) du \right\}. \end{aligned}$$

Using this equation in (78), we obtain

$$(79) \quad \mathbb{E} \left[\Delta^{(2)} \Delta^{(4)} \right] = 2\lambda^1 (\eta^1)^2 J_\delta^T + \lambda^1 \eta^1 \mathbb{E} \left\{ \int_{\mathbb{R}} w(u, u - \xi) \tilde{w}(u) du \right\}.$$

Let us now focus on the square moment of $\Delta^{(4)}$. We have the decomposition

$$(80) \quad \mathbb{E} \left[|\Delta^{(4)}|^2 \right] = \mathbb{E} \left[\left| \Delta^{(4)} - \mathbb{E}[\Delta^{(4)} | Z_1^1, Z_1^2] \right|^2 \right] + \mathbb{E} \left[\left| \mathbb{E} \left(\Delta^{(4)} | Z_1^1, Z_1^2 \right) \right|^2 \right].$$

Noticing that

$$\Delta^{(4)} - \mathbb{E}[\Delta^{(4)} | Z_1^1, Z_1^2] = \eta^1 \int_{\mathbb{R}} \tilde{w}(u) dZ_1^1(u) - \varphi(Z_1^1, Z_1^2),$$

it follows that

$$\begin{aligned} \mathbb{E} \left[\left| \Delta^{(4)} - \mathbb{E}[\Delta^{(4)} | Z_1^1, Z_1^2] \right|^2 \right] &= \mathbb{E} \left\{ \left| \int_{\mathbb{R}} \left[\eta^1 \tilde{w}(u) - \int_{\mathbb{R}} w(u, v) dZ_2^2(v) \right] dZ_1^1(u) \right|^2 \right\} \\ &= \eta^1 \mathbb{E} \left\{ \int_{\mathbb{R}} \left[\eta^1 \tilde{w}(u) - \int_{\mathbb{R}} w(u, v) dZ_2^2(v) \right]^2 du \right\} \\ &\quad + (\eta^1)^2 \mathbb{E} \left\{ \left| \eta^1 I_\delta^T - \int_{\mathbb{R}} \tilde{w}(u) dZ_2^2(v) \right|^2 \right\}. \end{aligned}$$

In the above expression, both squares involve a centered random variable. Developing the integral with respect to Z_2^2 in both cases, distinguishing once more common jumps, and finally using that $w^2(u, v) = w(u, v)$, we end up with

$$(81) \quad \mathbb{E} \left[\left| \Delta^{(4)} - \mathbb{E}[\Delta^{(4)} | Z_1^1, Z_1^2] \right|^2 \right] = (\eta^1)^2 I_\delta^T + (\eta^1)^3 J_\delta^T.$$

Further, we deduce from (77) that

$$\begin{aligned} &\mathbb{E} \left[\left| \mathbb{E}[\Delta^{(4)} | Z_1^1, Z_1^2] \right|^2 \right] \\ &= \mathbb{E} \left\{ \left| \int_{\mathbb{R}} \left[w(u, u + \xi_{Z_1^1(u^-)+1}) - \eta^1 \tilde{w}(u) \right] dZ_1^1(u) \right. \right. \\ &\quad \left. \left. + \int_{\Delta_2} w(u_1, u_2 + \xi_{Z_1^1(u_2^-)+1}) dZ_1^1(u_1) dZ_1^1(u_2) \right|^2 \right\} \\ &= \mathbb{E} \left\{ \left| \int_{\mathbb{R}} \left[w(u, u + \xi_{Z_1^1(u^-)+1}) - \eta^1 \tilde{w}(u) \right] dZ_1^1(u) \right|^2 \right. \\ &+ 2 \int_{\mathbb{R}} \left[\int_{\Delta_2} w(u_1, u_2 + \xi_{Z_1^1(u_2^-)+1}) \left[w(u_3, u_3 + \xi_{Z_1^1(u_3^-)+1}) - \eta^1 \tilde{w}(u_3) \right] dZ_1^1(u_1) \right. \\ &\quad \left. \left. dZ_1^1(u_2) \right] dZ_1^1(u_3) \right. \\ &\left. + \left| \int_{\Delta_2} w(u_1, u_2 + \xi_{Z_1^1(u_2^-)+1}) dZ_1^1(u_1) dZ_1^1(u_2) \right|^2 \right\}. \end{aligned}$$

By distinguishing the common jumps and using Lemma 5 and once more that $w^2 = w$, we obtain

$$\begin{aligned}
& \mathbb{E} \left\{ \left| \int_{\mathbb{R}} \left[w(u, u + \xi_{Z_1^1(u^-)+1}) - \eta^1 \tilde{w}(u) \right] dZ_1^1(u) \right|^2 \right\} \\
&= \mathbb{E} \left\{ \int_{\mathbb{R}} \left[w(u, u + \xi_{Z_1^1(u^-)+1}) - \eta^1 \tilde{w}(u) \right]^2 dZ_1^1(u) \right. \\
&+ \int_{\Delta_2} \left[w(u_1, u_1 + \xi_{Z_1^1(u_1^-)+1}) - \eta^1 \tilde{w}(u_1) \right] \left[w(u_2, u_2 + \xi_{Z_1^1(u_2^-)+1}) - \eta^1 \tilde{w}(u_2) \right] \\
&\qquad \qquad \qquad \left. dZ_1^1(u_1) dZ_1^1(u_2) \right\} \\
&= \mathbb{E} \left\{ \eta^1 \int_{\mathbb{R}} \left[w(u, u + \xi) - \eta^1 \tilde{w}(u) \right]^2 du \right. \\
&+ (\eta^1)^2 \int_{\mathbb{R}^2} \left[w(u_1, u_1 + \xi_1) - \eta^1 \tilde{w}(u_1) \right] \left[w(u_2, u_2 + \xi_2) - \eta^1 \tilde{w}(u_2) \right] du_1 du_2 \left. \right\} \\
&= \eta^1 \mathbb{E} \left\{ \int_{\mathbb{R}} w(u, u + \xi) du - 2\eta^1 \int_{\mathbb{R}} w(u, u + \xi) \tilde{w}(u) du + (\eta^1)^2 J_\delta^T \right\} \\
&+ (\eta^1)^2 \left\{ \left(\mathbb{E} \int_{\mathbb{R}} w(u, u + \xi) du \right)^2 - 2\eta^1 I_\delta^T \mathbb{E} \int_{\mathbb{R}} w(u, u + \xi) du + (\eta^1)^2 (I_\delta^T)^2 \right\}
\end{aligned}$$

and therefore

$$\begin{aligned}
(82) \quad & \mathbb{E} \left\{ \left| \int_{\mathbb{R}} \left[w(u, u + \xi_{Z_1^1(u^-)+1}) - \eta^1 \tilde{w}(u) \right] dZ_1^1(u) \right|^2 \right\} \\
&= (\eta^1)^3 J_\delta^T + (\eta^1)^4 (I_\delta^T)^2 + (\eta^1)^2 \left(\mathbb{E}[(T - |\xi|)_{|\xi| \leq \delta}] \right)^2 + \eta^1 (1 - 2(\eta^1)^2 I_\delta^T) \mathbb{E}[(T - |\xi|)_{|\xi| \leq \delta}] \\
&\qquad \qquad \qquad - 2(\eta^1)^2 \mathbb{E} \int_{\mathbb{R}} w(u, u + \xi) \tilde{w}(u) du.
\end{aligned}$$

Similarly,

$$\begin{aligned}
& \mathbb{E} \left\{ \int_{\mathbb{R}} \left[\int_{\Delta_2} w(u_1, u_2 + \xi_{Z_1^1(u_2^-)+1}) \left[w(u_3, u_3 + \xi_{Z_1^1(u_3^-)+1}) - \eta^1 \tilde{w}(u_3) \right] dZ_1^1(u_1) \right. \right. \\
&\qquad \qquad \qquad \left. \left. dZ_1^1(u_2) \right] dZ_1^1(u_3) \right\} \\
&= \mathbb{E} \left\{ \int_{\Delta_2} w(u_1, u_2 + \xi_{Z_1^1(u_2^-)+1}) \left[w(u_1, u_1 + \xi_{Z_1^1(u_1^-)+1}) - \eta^1 \tilde{w}(u_1) \right] dZ_1^1(u_1) dZ_1^1(u_2) \right. \\
&+ \int_{\Delta_2} w(u_1, u_2 + \xi_{Z_1^1(u_2^-)+1}) \left[w(u_2, u_2 + \xi_{Z_1^1(u_2^-)+1}) - \eta^1 \tilde{w}(u_2) \right] dZ_1^1(u_1) dZ_1^1(u_2) \\
&+ \int_{\Delta_3} w(u_1, u_2 + \xi_{Z_1^1(u_2^-)+1}) \left[w(u_3, u_3 + \xi_{Z_1^1(u_3^-)+1}) - \eta^1 \tilde{w}(u_3) \right] dZ_1^1(u_1) \\
&\qquad \qquad \qquad \left. \left. dZ_1^1(u_2) dZ_1^1(u_3) \right\} \\
&= \mathbb{E} \left\{ (\eta^1)^2 \int_{\mathbb{R}^2} w(u_1, u_2 + \xi_2) \left[w(u_1, u_1 + \xi_1) - \eta^1 \tilde{w}(u_1) \right] du_1 du_2 \right. \\
&+ (\eta^1)^2 \int_{\mathbb{R}^2} w(u_1, u_2 + \xi) \left[w(u_2, u_2 + \xi) - \eta^1 \tilde{w}(u_2) \right] du_1 du_2 \\
&+ (\eta^1)^3 \int_{\mathbb{R}^3} w(u_1, u_2 + \xi_2) \left[w(u_3, u_3 + \xi_3) - \eta^1 \tilde{w}(u_3) \right] du_1 du_2 du_3 \left. \right\},
\end{aligned}$$

from which we deduce that

$$\begin{aligned} & \mathbb{E} \left\{ \int_{\mathbb{R}} \int_{\Delta_2} w(u_1, u_2 + \xi_{Z_1^1(u_2^-)+1}) \left[w(u_3, u_3 + \xi_{Z_1^1(u_3^-)+1}) - \eta^1 \tilde{w}(u_3) \right] dZ_1^1(u_1) \right. \\ & \qquad \qquad \qquad \left. dZ_1^1(u_2) dZ_1^1(u_3) \right\} \\ &= (\eta^1)^2 \mathbb{E} \left\{ \int_{\mathbb{R}} \tilde{w}(u) w(u, u + \xi) du - \eta^1 J_\delta^T + \int_{\mathbb{R}} \tilde{w}(u + \xi) [w(u, u + \xi) - \eta^1 \tilde{w}(u)] du \right. \\ & \qquad \qquad \left. + \eta^1 I_\delta^T \left[\int_{\mathbb{R}} w(u, u + \xi) du - \eta^1 I_\delta^T \right] \right\} \end{aligned}$$

and then

$$\begin{aligned} (83) \quad & \mathbb{E} \left\{ \int_{\mathbb{R}} \int_{\Delta_2} w(u_1, u_2 + \xi_{Z_1^1(u_2^-)+1}) \left[w(u_3, u_3 + \xi_{Z_1^1(u_3^-)+1}) - \eta^1 \tilde{w}(u_3) \right] dZ_1^1(u_1) \right. \\ & \qquad \qquad \qquad \left. dZ_1^1(u_2) dZ_1^1(u_3) \right\} \\ &= -(\eta^1)^3 J_\delta^T - (\eta^1)^4 (I_\delta^T)^2 + (\eta^1)^3 I_\delta^T \mathbb{E} [(T - |\xi|) \mathbb{1}_{|\xi| \leq \delta}] \\ & \qquad \qquad \qquad + (\eta^1)^2 \mathbb{E} \int_{\mathbb{R}} \tilde{w}(u) [w(u, u + \xi) + w(u, u - \xi) - \eta^1 \tilde{w}(u + \xi)] du. \end{aligned}$$

The last term we need to compute is

$$\begin{aligned} & \mathbb{E} \left\{ \left| \int_{\Delta_2} w(u_1, u_2 + \xi_{Z_1^1(u_2^-)+1}) dZ_1^1(u_1) dZ_1^1(u_2) \right|^2 \right\} \\ &= \mathbb{E} \left\{ \int_{\substack{(u_1, u_2) \in \Delta_2 \\ (u_3, u_4) \in \Delta_2}} w(u_1, u_2 + \xi_{Z_1^1(u_2^-)+1}) w(u_3, u_4 + \xi_{Z_1^1(u_4^-)+1}) dZ_1^1(u_1) dZ_1^1(u_2) \right. \\ & \qquad \qquad \qquad \left. dZ_1^1(u_3) dZ_1^1(u_4) \right\}. \end{aligned}$$

By distinguishing again the cases $u_i = u_j$ and $u_i \neq u_j$ for $j = 1, \dots, 4$, we split this quantity in seven terms that can be explicitly computed.

The first term consists of taking the stochastic integral on the sub-domain $\{(u_1, u_2, u_3, u_4) : u_1 = u_3, u_2 = u_4\}$. The corresponding expectation is given by

$$\begin{aligned} K_1 &= \mathbb{E} \left\{ \int_{\Delta_2} w^2(u_1, u_2 + \xi_{Z_1^1(u_2^-)+1}) dZ_1^1(u_1) dZ_1^1(u_2) \right\} \\ &= (\eta^1)^2 \mathbb{E} \int_{\mathbb{R}^2} w^2(u_1, u_2 + \xi) du_1 du_2 = (\eta^1)^2 \mathbb{E} \int_{\mathbb{R}^2} w(u_1, u_2 + \xi) du_1 du_2 \\ &= (\eta^1)^2 I_\delta^T, \end{aligned}$$

since $w^2 = w$.

The second term consists of taking the stochastic integral on the sub-domain $\{(u_1, u_2, u_3, u_4) : u_1 = u_4, u_2 = u_3\}$. The corresponding expectation is given by

$$\begin{aligned} K_2 &= \mathbb{E} \left\{ \int_{\Delta_2} w(u_1, u_2 + \xi_{Z_1^1(u_2^-)+1}) w(u_2, u_1 + \xi_{Z_1^1(u_1^-)+1}) dZ_1^1(u_1) dZ_1^1(u_2) \right\} \\ &= (\eta^1)^2 \mathbb{E} \left\{ \int_{\mathbb{R}^2} w(u_1, u_2 + \xi_2) w(u_2, u_1 + \xi_1) du_1 du_2 \right\}. \end{aligned}$$

The third term consists of taking the stochastic integral on the sub-domain $\{(u_1, u_2, u_3, u_4) : u_1 = u_3, u_2 \neq u_4\}$. The corresponding expectation is given by

$$\begin{aligned} K_3 &= \mathbb{E} \left\{ \int_{\Delta_3} w(u_1, u_2 + \xi_{Z_1^1(u_2-)+1}) w(u_1, u_3 + \xi_{Z_1^1(u_3-)+1}) dZ_1^1(u_1) dZ_1^1(u_2) dZ_1^1(u_3) \right\} \\ &= (\eta^1)^3 \mathbb{E} \left\{ \int_{\mathbb{R}^3} w(u_1, u_2 + \xi_2) w(u_1, u_3 + \xi_3) du_1 du_2 du_3 \right\} \\ &= (\eta^1)^3 J_\delta^T. \end{aligned}$$

The fourth term consists of taking the stochastic integral on the sub-domain $\{(u_1, u_2, u_3, u_4) : u_1 = u_4, u_2 \neq u_3\}$. The corresponding expectation is given by

$$\begin{aligned} K_4 &= \mathbb{E} \left\{ \int_{\Delta_3} w(u_1, u_2 + \xi_{Z_1^1(u_2-)+1}) w(u_3, u_1 + \xi_{Z_1^1(u_1-)+1}) dZ_1^1(u_1) dZ_1^1(u_2) dZ_1^1(u_3) \right\} \\ &= (\eta^1)^3 \mathbb{E} \left\{ \int_{\mathbb{R}^3} w(u_1, u_2 + \xi_2) w(u_3, u_1 + \xi_1) du_1 du_2 du_3 \right\} \\ &= (\eta^1)^3 \mathbb{E} \int_{\mathbb{R}} \tilde{w}(u) \tilde{w}(u + \xi) du. \end{aligned}$$

The fifth term consists of taking the stochastic integral on the sub-domain $\{(u_1, u_2, u_3, u_4) : u_1 \neq u_3, u_2 = u_4\}$. The corresponding expectation K_5 is calculated in the same way the third term and it equals $K_5 = K_3 = (\eta^1)^3 J_\delta^T$.

The sixth term consists of taking the stochastic integral on the sub-domain $\{(u_1, u_2, u_3, u_4) : u_1 \neq u_4, u_2 = u_3\}$ and is treated in the same way as the fourth term, yielding the same expectation $K_6 = K_4$.

The seventh term consists of taking the stochastic integral on the sub-domain Δ_4 . The corresponding expectation is given by

$$\begin{aligned} K_7 &= \\ &\mathbb{E} \left\{ \int_{\Delta_4} w(u_1, u_2 + \xi_{Z_1^1(u_2-)+1}) w(u_3, u_4 + \xi_{Z_1^1(u_4-)+1}) dZ_1^1(u_1) dZ_1^1(u_2) dZ_1^1(u_3) dZ_1^1(u_4) \right\} \\ &= (\eta^1)^4 \mathbb{E} \left\{ \int_{\mathbb{R}^4} w(u_1, u_2 + \xi_2) w(u_3, u_4 + \xi_4) du_1 du_2 du_3 du_4 \right\} \\ &= (\eta^1)^4 (I_\delta^T)^4. \end{aligned}$$

By adding K_1, \dots, K_7 , we obtain

$$\begin{aligned} (84) \quad &\mathbb{E} \left\{ \left| \int_{\Delta_2} w(u_1, u_2 + \xi_{Z_1^1(u_2-)+1}) dZ_1^1(u_1) dZ_1^1(u_2) \right|^2 \right\} \\ &= (\eta^1)^2 I_\delta^T + 2(\eta^1)^3 J_\delta^T + (\eta^1)^4 (I_\delta^T)^2 + 2(\eta^1)^3 \mathbb{E} \int_{\mathbb{R}} \tilde{w}(u) \tilde{w}(u + \xi) du \\ &\quad + (\eta^1)^2 \mathbb{E} \int_{\mathbb{R}^2} w(u_1 + \xi_1, u_2) w(u_1, u_2 + \xi_2) du, \end{aligned}$$

and the sum (81) + (82) + 2 × (83) + (84) gives

$$\begin{aligned} (85) \quad &\mathbb{E} \left[|\Delta^{(4)}|^2 \right] \\ &= 2(\eta^1)^2 I_\delta^T + 2(\eta^1)^3 J_\delta^T + \eta^1 \mathbb{E} [(T - |\xi|) \mathbb{1}_{|\xi| \leq \delta}] + (\eta^1)^2 (\mathbb{E} [(T - |\xi|) \mathbb{1}_{|\xi| \leq \delta}])^2 \\ &\quad + 2(\eta^1)^2 \mathbb{E} \int_{\mathbb{R}} \tilde{w}(u) w(u, u - \xi) du + (\eta^1)^2 \mathbb{E} \int_{\mathbb{R}^2} w(u_1 + \xi_1, u_2) w(u_1, u_2 + \xi_2) du_1 du_2. \end{aligned}$$

Then using (79) and (85), we deduce from the decomposition (74) that

$$\begin{aligned}
 & \mathbb{V}[f_\varphi(X_1, X_2)] \\
 &= \mathbb{E}[|f_\varphi(X_1, X_2)|^2] - (\eta^1)^2 \left(\mathbb{E}[(T - |\xi|) \mathbb{1}_{|\xi| \leq \delta}] \right)^2 \\
 &= 2[(\eta^1 + \lambda^1)(\eta^1 + \lambda^2) - (\eta^1)^2] I_\delta^T + 2[(\eta^1 + \lambda^1)^2(\eta^1 + \lambda^2) - (\eta^1)^2(\eta^1 + 2\lambda^1)] J_\delta^T \\
 &\quad + 2(\eta^1)^2 I_\delta^T + 2(\eta^1)^3 J_\delta^T + \eta^1 \mathbb{E}[(T - |\xi|) \mathbb{1}_{|\xi| \leq \delta}] \\
 &\quad + 2(\eta^1)^2 \mathbb{E} \int_{\mathbb{R}} \tilde{w}(u) w(u, u - \xi) du + (\eta^1)^2 \mathbb{E} \int_{\mathbb{R}^2} w(u_1 + \xi_1, u_2) w(u_1, u_2 + \xi_2) du_1 du_2 \\
 &\quad + 4\lambda^1 (\eta^1)^2 J_\delta^T + 2\lambda^1 \eta^1 \mathbb{E} \left\{ \int_{\mathbb{R}} \tilde{w}(u) w(u, u - \xi) du \right\}.
 \end{aligned}$$

It follows that

$$\begin{aligned}
 \mathbb{V}[f_\varphi(X_1, X_2)] &= 2(\eta^1 + \lambda^1)(\eta^1 + \lambda^2) [I_\delta^T + (\eta^1 + \lambda^1) J_\delta^T] \\
 &\quad + \eta^1 \mathbb{E}[(T - |\xi|) \mathbb{1}_{|\xi| \leq \delta}] + 2\eta^1 (\eta^1 + \lambda^1) \mathbb{E} \left\{ \int_{\mathbb{R}} \tilde{w}(u) w(u, u - \xi) du \right\} \\
 &\quad + (\eta^1)^2 \mathbb{E} \left\{ \int_{\mathbb{R}^2} w(u_1, u_2) w(u_1 - \xi_1, u_2 + \xi_2) du_1 du_2 \right\},
 \end{aligned}$$

and then using the bounds $\eta^1 \leq \lambda^{\mathbf{m}} = \max\{\lambda^1, \lambda^2\}$, $I_\delta^T \leq 2\delta T$, $J_\delta^T \leq 4\delta^2 T$, $w(u, u - \xi) \leq \mathbb{1}_{|\xi| \leq \delta}$ and $w(u_1 - \xi_1, u_2 + \xi_2) \leq 1$, we obtain, for a constant C not depending on the model parameters,

$$\begin{aligned}
 \mathbb{V}[f_\varphi(X_1, X_2)] &\leq 2(\eta^1 + \lambda^1)(\eta^1 + \lambda^2) [I_\delta^T + (\eta^1 + \lambda^1) J_\delta^T] \\
 &\quad + \eta^1 T \mathbb{P}(|\xi| \leq \delta) + 2\eta^1 \lambda^{\mathbf{m}} I_\delta^T \mathbb{P}(|\xi| \leq \delta) + 3(\eta^1)^2 I_\delta^T \\
 &\leq C(1 + \lambda^{\mathbf{m}} \delta) ((\lambda^{\mathbf{m}})^2 \delta T + \eta^1 T \mathbb{P}(|\xi| \leq \delta)),
 \end{aligned}$$

which implies (24).

B.3. Proof of the lower bounds (7) and (8). At this stage, we have proved, combining all the inequalities on the various variances and Lemma 1, that there exists a constant $c > 0$ such that if

$$\Delta_\varphi \geq \frac{c}{\sqrt{n\alpha\beta}} \sqrt{(1 + \lambda^{\mathbf{m}} \delta) ((\lambda^{\mathbf{m}})^2 \delta T + \eta^1 T \mathbb{P}(|\xi| \leq \delta))},$$

then the Type II error is controlled. But since we can upper bound

$$\eta^1 T \mathbb{P}(|\xi| \leq \delta) \leq 2\Delta_\varphi$$

(recall (9)), it is sufficient to have

$$\Delta_\varphi \geq \frac{c}{\sqrt{n\alpha\beta}} \sqrt{(1 + \lambda^{\mathbf{m}} \delta) ((\lambda^{\mathbf{m}})^2 \delta T + 2\Delta_\varphi)}.$$

It remains to solve this in Δ_φ . With an additional upper bound, one can show that there exists another constant $C > 0$ such that the Type II error is controlled as soon as

$$\Delta_\varphi \geq C \left[\sqrt{\frac{(1 + \lambda^{\mathbf{m}} \delta) (\lambda^{\mathbf{m}})^2 \delta T}{n\alpha\beta}} + \frac{1 + \lambda^{\mathbf{m}} \delta}{n\alpha\beta} \right],$$

which concludes the proof of (7).

It remains to solve the second order polynomial in $x = 1/\sqrt{n}$ of (7) and to lower bound Δ_φ by $\eta^1 (T/2) \mathbb{P}(|\xi| \leq \delta)$, to obtain that (7) is implied by

$$n \geq C \frac{1 + \lambda^{\mathbf{m}} \delta}{\alpha\beta \eta^1 T \mathbb{P}(|\xi| \leq \delta)} \left[1 + \frac{(\lambda^{\mathbf{m}})^2 \delta}{\eta^1 \mathbb{P}(|\xi| \leq \delta)} \right].$$

Since all the reasoning is valid only if $n \geq 3/\sqrt{\alpha\beta}$, we obtain (8).

APPENDIX C. THE HAWKES MODEL

C.1. Proof of Lemma 9. The first step of the proof consists of computing the rate $\Psi(t)$ defined by (28). For the interaction function $h(t) = ae^{-bt} \mathbb{1}_{t>0}$, clearly for any $m \geq 1$,

$$h^{*m}(t) = \frac{a^m t^{m-1}}{(m-1)!} e^{-bt} \mathbb{1}_{t>0}$$

which implies that

$$(86) \quad \Psi(t) = ae^{-(b-a)t} \mathbb{1}_{t>0} = ae^{-b(1-\ell)t} \mathbb{1}_{t>0}.$$

We state the following intermediate result which will be used in our calculations later.

Lemma 16. *Let $\alpha_1 < \dots < \alpha_m$ be a sequence of real numbers with $m \geq 2$, then*

$$(87) \quad \int_{\mathbb{R}} \prod_{i=1}^m \Psi(\alpha_i - v) dv = \frac{a}{m(b-a)} \prod_{i=2}^m \Psi(\alpha_i - \alpha_1).$$

In addition, for any $\alpha \in \mathbb{R}$,

$$(88) \quad \int_{\mathbb{R}} \Psi(|\alpha - v|) \prod_{i=1}^m \Psi(\alpha_i - v) dv \leq \frac{a}{(m-1)(b-a)} \Psi(|\alpha - \alpha_1|) \prod_{i=2}^m \Psi(\alpha_i - \alpha_1).$$

Proof. We have

$$\begin{aligned} \int_{\mathbb{R}} \prod_{i=1}^m \Psi(\alpha_i - v) dv &= \int_{-\infty}^{\alpha_1} a^{m-1} \Psi\left(\sum_{i=1}^m \alpha_i - mv\right) dv \\ &= \frac{a^{m-1}}{m(b-a)} \Psi\left(\sum_{i=2}^m \alpha_i - (m-1)\alpha_1\right) \\ &= \frac{a^{m-1}}{m(b-a)} \Psi\left(\sum_{i=2}^m (\alpha_i - \alpha_1)\right) \end{aligned}$$

which is (87). Further, putting $\alpha_0 := \alpha$ to ease the notation, we have

$$\begin{aligned} \int_{\mathbb{R}} \Psi(|\alpha - v|) \prod_{i=1}^m \Psi(\alpha_i - v) dv \\ = \int_{-\infty}^{\alpha \wedge \alpha_1} \prod_{i=0}^m \Psi(\alpha_i - v) dv + \mathbb{1}_{\alpha < \alpha_1} \int_{\alpha}^{\alpha_1} \Psi(v - \alpha) \prod_{i=1}^m \Psi(\alpha_i - v) dv \end{aligned}$$

which in turn equals

$$\begin{aligned} a^m \int_{-\infty}^{\alpha \wedge \alpha_1} \Psi\left(\sum_{i=0}^m \alpha_i - (m+1)v\right) dv + a^m \mathbb{1}_{\alpha < \alpha_1} \int_{\alpha}^{\alpha_1} \Psi\left(\sum_{i=1}^m \alpha_i - \alpha - (m-1)v\right) dv \\ = \frac{a^m}{(m+1)(b-a)} \Psi\left(\sum_{i=0}^m \alpha_i - (m+1)(\alpha \wedge \alpha_1)\right) \\ + \frac{a^m \mathbb{1}_{\alpha < \alpha_1}}{(m-1)(b-a)} \left[\Psi\left(\sum_{i=1}^m \alpha_i - \alpha - (m-1)\alpha_1\right) - \Psi\left(\sum_{i=1}^m \alpha_i - m\alpha\right) \right]. \end{aligned}$$

By distinguishing the cases $\alpha \geq \alpha_1$ and $\alpha < \alpha_1$ in the first term, we obtain

$$\begin{aligned}
 & \int_{\mathbb{R}} \Psi(|\alpha - v|) \prod_{i=1}^m \Psi(\alpha_i - v) dv \\
 &= \frac{a^m \mathbb{1}_{\alpha \geq \alpha_1}}{(m+1)(b-a)} \Psi\left(\sum_{i=0}^m \alpha_i - (m+1)\alpha_1\right) \\
 &+ \frac{a^m \mathbb{1}_{\alpha < \alpha_1}}{(m-1)(b-a)} \left[\Psi\left(\sum_{i=1}^m \alpha_i - \alpha - (m-1)\alpha_1\right) - \frac{2}{m+1} \Psi\left(\sum_{i=1}^m \alpha_i - m\alpha\right) \right] \\
 &= \frac{a^m \mathbb{1}_{\alpha \geq \alpha_1}}{(m+1)(b-a)} \Psi\left(\alpha - \alpha_1 + \sum_{i=2}^m (\alpha_i - \alpha_1)\right) \\
 &+ \frac{a^m \mathbb{1}_{\alpha < \alpha_1}}{(m-1)(b-a)} \Psi\left(\alpha_1 - \alpha + \sum_{i=2}^m (\alpha_i - \alpha_1)\right) \left[1 - \frac{2}{m+1} e^{-(b-a)(m-1)(\alpha_1 - \alpha)} \right] \\
 &\leq \frac{a^m}{(m-1)(b-a)} \Psi\left(|\alpha - \alpha_1| + \sum_{i=2}^m (\alpha_i - \alpha_1)\right),
 \end{aligned}$$

implying (88). □

Let us now come back to the study of the first four cumulant measures.

According to (36), the second order cumulant measure satisfies

$$\begin{aligned}
 \mathbb{1}_{t_1 \neq t_2} k_2(dt_1, dt_2) &= \frac{\mu}{1 - \frac{a}{b}} \left(\int_{\mathbb{R}} \mathbb{W}[t_1, t_2](du) \right) dt_1 dt_2 \\
 &= \frac{\mu}{1 - \frac{a}{b}} \left(\Psi(|t_1 - t_2|) + \int_{\mathbb{R}} \Psi(t_1 - u) \Psi(t_2 - u) du \right) dt_1 dt_2 \\
 &= \frac{\mu}{1 - \frac{a}{b}} \left(1 + \frac{1}{2} \frac{a}{b-a} \right) \Psi(|t_1 - t_2|) dt_1 dt_2,
 \end{aligned}$$

where we have used (87) of Lemma 16.

The third and fourth order cumulant measures are more complicated. We need to compute the contribution corresponding to each equivalence class in Figures 5 and 6 respectively, and to take into account the corresponding permutations. Following (37), we set

$$S_3^{(1)}(dt_1, dt_2, dt_3) = \frac{\mu}{1 - \frac{a}{b}} \left(\int_{\mathbb{R}} \mathbb{W}[t_1, t_2, t_3](du) \right) dt_1 dt_2 dt_3$$

and

$$S_3^{(2,\sigma)}(dt_1, dt_2, dt_3) = \frac{\mu}{1 - \frac{a}{b}} \left(\int_{\mathbb{R}^2} \mathbb{W}[v, t_{\sigma(3)}](du) \mathbb{W}[t_{\sigma(1)}, t_{\sigma(2)}](dv) \right) dt_1 dt_2 dt_3$$

such that

$$(89) \quad \mathbb{1}_{\Delta_3} k_3 = S_3^{(1)} + \sum_{\sigma} S_3^{(2,\sigma)}.$$

We have

$$\begin{aligned}
 & \mathbb{1}_{t_1 < t_2 < t_3} S_3^{(1)}(dt_1, dt_2, dt_3) \\
 &= \frac{\mu}{1 - \frac{a}{b}} \mathbb{1}_{t_1 < t_2 < t_3} \left(\Psi(t_3 - t_1) \Psi(t_2 - t_1) + \int_{\mathbb{R}} \Psi(t_1 - u) \Psi(t_2 - u) \Psi(t_3 - u) du \right) dt_1 dt_2 dt_3 \\
 &= \frac{\mu}{1 - \frac{a}{b}} \mathbb{1}_{t_1 < t_2 < t_3} \left(1 + \frac{1}{3} \frac{a}{b-a} \right) \Psi(t_3 - t_1) \Psi(t_2 - t_1) dt_1 dt_2 dt_3,
 \end{aligned}$$

thanks to (87) in Lemma 16, and then by symmetry

$$(90) \quad S_3^{(1)}(dt_1, dt_2, dt_3) = a \frac{\mu}{1 - \frac{a}{b}} \left(1 + \frac{1}{3} \frac{a}{b-a} \right) \Psi\left(\sum_{i=1}^3 t_i - 3 \min_i t_i\right) dt_1 dt_2 dt_3.$$

Further, the measure associated to the equivalence class **(b)** in Figure 5 satisfies for a fixed permutation σ of $\{1, 2, 3\}$,

$$\begin{aligned}
S_3^{(2,\sigma)}(dt_1, dt_2, dt_3) &= \frac{\mu}{1 - \frac{a}{b}} \left(\int_{\mathbb{R}^2} \mathbb{W}[v, t_{\sigma(3)}](du) \mathbb{W}[t_{\sigma(1)}, t_{\sigma(2)}](dv) \right) dt_1 dt_2 dt_3 \\
&= \frac{\mu}{1 - \frac{a}{b}} \left(\int_{\mathbb{R}} \underbrace{\left\{ \int_{\mathbb{R}} \Psi(v-u) \Psi(t_{\sigma(3)}-u) du + \Psi(v-t_{\sigma(3)}) \right\}}_{= \frac{1}{2} \frac{a}{b-a} \Psi(|t_{\sigma(3)}-v|)} \right. \\
&\quad \left. \mathbb{W}[t_{\sigma(1)}, t_{\sigma(2)}](dv) \right) dt_1 dt_2 dt_3 \\
&\leq \frac{\mu}{1 - \frac{a}{b}} \left(1 + \frac{1}{2} \frac{a}{b-a} \right) \left\{ \int_{\mathbb{R}} \Psi(|v-t_{\sigma(3)}|) \mathbb{W}[t_{\sigma(1)}, t_{\sigma(2)}](dv) \right\} dt_1 dt_2 dt_3 \\
&= \frac{\mu}{1 - \frac{a}{b}} \left(1 + \frac{1}{2} \frac{a}{b-a} \right) \left\{ \Psi(|t_{\sigma(3)} - t_{\sigma(1)} \wedge t_{\sigma(2)}|) \Psi(|t_{\sigma(1)} - t_{\sigma(2)}|) \right. \\
&\quad \left. + \int_{\mathbb{R}} \Psi(|t_{\sigma(3)} - v|) \Psi(t_{\sigma(1)} - v) \Psi(t_{\sigma(2)} - v) dv \right\} dt_1 dt_2 dt_3.
\end{aligned}$$

This implies thanks to (88) in Lemma 16 that

$$\begin{aligned}
S_3^{(2,\sigma)}(dt_1, dt_2, dt_3) &\leq \\
&\frac{\mu}{1 - \frac{a}{b}} \left(1 + \frac{1}{2} \frac{a}{b-a} \right) \left(1 + \frac{a}{b-a} \right) \Psi(|t_{\sigma(3)} - t_{\sigma(1)} \wedge t_{\sigma(2)}|) \Psi(|t_{\sigma(1)} - t_{\sigma(2)}|) dt_1 dt_2 dt_3
\end{aligned}$$

and then

$$(91) \quad S_3^{(2,\sigma)}(dt_1, dt_2, dt_3) \leq \frac{a\mu}{(1 - \frac{a}{b})^3} \Psi(|t_{\sigma(3)} - t_{\sigma(1)} \wedge t_{\sigma(2)}| + |t_{\sigma(1)} - t_{\sigma(2)}|) dt_1 dt_2 dt_3.$$

Using (90) and (91), we deduce from (89) that

$$\begin{aligned}
&\mathbb{1}_{\Delta_\ell}(t_1, t_2, t_3) k_3(dt_1, dt_2, dt_3) \\
&\leq \frac{a\mu}{1 - \frac{a}{b}} \left\{ 1 + \frac{1}{3} \frac{a}{b-a} + \frac{3}{(1 - \frac{a}{b})^2} \right\} \Psi \left(\max_i t_i - \min_i t_i \right) dt_1 dt_2 dt_3 \\
&\leq \frac{4a\mu}{(1 - \frac{a}{b})^3} \Psi \left(\max_i t_i - \min_i t_i \right) dt_1 dt_2 dt_3,
\end{aligned}$$

where the factor 3 comes from the number of permutations associated to the equivalence class **(b)** in Figure 5. This implies (41) with $l = 3$.

We are now interested in the fourth order cumulant, and we introduce the measures associated to each equivalence class in Figure 6 that are defined from (38) by

$$\begin{aligned}
(92) \quad S_4^{(1)} &= \frac{\mu}{1 - \frac{a}{b}} \left(\int_{\mathbb{R}} \mathbb{W}[t_1, t_2, t_3, t_4](du) \right) dt_1 dt_2 dt_3 dt_4, \\
S_4^{(2,\sigma)} &= \frac{\mu}{1 - \frac{a}{b}} \left(\int_{\mathbb{R}^2} \mathbb{W}[v, t_{\sigma(3)}, t_{\sigma(4)}](du) \mathbb{W}[t_{\sigma(1)}, t_{\sigma(2)}](dv) \right) dt_1 dt_2 dt_3 dt_4, \\
S_4^{(3,\sigma)} &= \frac{\mu}{1 - \frac{a}{b}} \left(\int_{\mathbb{R}^2} \mathbb{W}[v, t_{\sigma(4)}](du) \mathbb{W}[t_{\sigma(1)}, t_{\sigma(2)}, t_{\sigma(3)}](dv) \right) dt_1 dt_2 dt_3 dt_4, \\
S_4^{(4,\sigma)} &= \frac{\mu}{1 - \frac{a}{b}} \left(\int_{\mathbb{R}^3} \mathbb{W}[v, w](du) \mathbb{W}[t_{\sigma(1)}, t_{\sigma(2)}](dv) \mathbb{W}[t_{\sigma(3)}, t_{\sigma(4)}](dw) \right) dt_1 dt_2 dt_3 dt_4, \\
S_4^{(5,\sigma)} &= \frac{\mu}{1 - \frac{a}{b}} \left(\int_{\mathbb{R}^3} \mathbb{W}[v, t_{\sigma(4)}](du) \mathbb{W}[w, t_{\sigma(3)}](dv) \mathbb{W}[t_{\sigma(1)}, t_{\sigma(2)}](dw) \right) dt_1 dt_2 dt_3 dt_4,
\end{aligned}$$

where to ease notation we wrote $S_4^{(k,\sigma)}$ instead of $S_4^{(k,\sigma)}(dt_1 dt_2 dt_3 dt_4)$, for each $1 \leq k \leq 5$.

Then

$$(93) \quad \mathbb{1}_{\Delta_4} k_4 = S_4^{(1)} + \sum_{\sigma} S_4^{(2,\sigma)} + \sum_{\sigma} S_4^{(3,\sigma)} + \sum_{\sigma} S_4^{(4,\sigma)} + \sum_{\sigma} S_4^{(5,\sigma)}.$$

Bound on $S_4^{(1)}$. It follows from the first equation in (92) and from Lemma 16 that

$$\begin{aligned} & \mathbb{1}_{t_1 < t_2 < t_3 < t_4} S_4^{(1)}(dt_1, dt_2, dt_3, dt_4) \\ &= \frac{\mu}{1 - \frac{a}{b}} \mathbb{1}_{t_1 < t_2 < t_3 < t_4} \left(\Psi(t_2 - t_1) \Psi(t_3 - t_1) \Psi(t_4 - t_1) \right. \\ &+ \left. \int_{\mathbb{R}} \Psi(t_1 - u) \Psi(t_2 - u) \Psi(t_3 - u) \Psi(t_4 - u) du \right) dt_1 dt_2 dt_3 dt_4 \\ &= a^2 \frac{\mu}{1 - \frac{a}{b}} \mathbb{1}_{t_1 < t_2 < t_3 < t_4} \left(1 + \frac{1}{4} \frac{a}{b - a} \right) \Psi(t_2 + t_3 + t_4 - 3t_1) dt_1 dt_2 dt_3 dt_4, \end{aligned}$$

which implies by symmetry that

$$(94) \quad S_4^{(1)}(dt_1, dt_2, dt_3, dt_4) = \frac{a^2 \mu}{1 - \frac{a}{b}} \left(1 + \frac{1}{4} \frac{a}{b - a} \right) \Psi \left(\sum_{i=1}^4 t_i - 4 \min_i t_i \right) dt_1 dt_2 dt_3 dt_4.$$

Bound on $S_4^{(2,\sigma)}$. We deduce from the second equation in (92) and from Lemma 16 that

$$\begin{aligned} S_4^{(2,\sigma)}(dt_1, dt_2, dt_3, dt_4) &= \frac{\mu}{1 - \frac{a}{b}} \\ & \left\{ \int_{\mathbb{R}} \left[\int_{\mathbb{R}} \underbrace{\Psi(v - u) \Psi(t_{\sigma(3)} - u) \Psi(t_{\sigma(4)} - u)}_{= \frac{a^2}{3(b-a)} \Psi(v + t_{\sigma(3)} + t_{\sigma(4)} - 3v \wedge t_{\sigma(3)} \wedge t_{\sigma(4)})} du + \Psi(v - t_{\sigma(3)} \wedge t_{\sigma(4)}) \Psi(|t_{\sigma(3)} - t_{\sigma(4)}|) \right] \right. \\ & \left. \mathbb{W}[t_{\sigma(1)}, t_{\sigma(2)}](dv) \right\} dt_1 dt_2 dt_3 dt_4 \end{aligned}$$

which is in turn upper bounded by

$$\begin{aligned} & \frac{a\mu}{1 - \frac{a}{b}} \left(1 + \frac{1}{3} \frac{a}{b - a} \right) \\ & \left\{ \int_{\mathbb{R}} \Psi(v + t_{\sigma(3)} + t_{\sigma(4)} - 3v \wedge t_{\sigma(3)} \wedge t_{\sigma(4)}) \mathbb{W}[t_{\sigma(1)}, t_{\sigma(2)}](dv) \right\} dt_1 dt_2 dt_3 dt_4 \end{aligned}$$

which equals

$$\begin{aligned} & \frac{a\mu}{1 - \frac{a}{b}} \left(1 + \frac{1}{3} \frac{a}{b - a} \right) \\ & \left\{ \Psi(t_{\sigma(1)} \wedge t_{\sigma(2)} + t_{\sigma(3)} + t_{\sigma(4)} - 3(t_{\sigma(1)} \wedge t_{\sigma(2)}) \wedge t_{\sigma(3)} \wedge t_{\sigma(4)}) \Psi(|t_{\sigma(1)} - t_{\sigma(2)}|) \right. \\ & \left. + \int_{\mathbb{R}} \Psi(v + t_{\sigma(3)} + t_{\sigma(4)} - 3v \wedge t_{\sigma(3)} \wedge t_{\sigma(4)}) \Psi(t_{\sigma(1)} - v) \Psi(t_{\sigma(2)} - v) dv \right\} dt_1 dt_2 dt_3 dt_4. \end{aligned}$$

Now,

$$\begin{aligned} & \int_{\mathbb{R}} \Psi(v + t_{\sigma(3)} + t_{\sigma(4)} - 3v \wedge t_{\sigma(3)} \wedge t_{\sigma(4)}) \Psi(t_{\sigma(1)} - v) \Psi(t_{\sigma(2)} - v) dv \\ &= \mathbb{1}_{t_{\sigma(3)} \wedge t_{\sigma(4)} > t_{\sigma(1)} \wedge t_{\sigma(2)}} \int_{\mathbb{R}} \Psi(t_{\sigma(3)} + t_{\sigma(4)} - 2v) \Psi(t_{\sigma(1)} - v) \Psi(t_{\sigma(2)} - v) dv \\ &+ \mathbb{1}_{t_{\sigma(3)} \wedge t_{\sigma(4)} \leq t_{\sigma(1)} \wedge t_{\sigma(2)}} \int_{t_{\sigma(3)} \wedge t_{\sigma(4)}}^{t_{\sigma(1)} \wedge t_{\sigma(2)}} \Psi(v + t_{\sigma(3)} \vee t_{\sigma(4)} - 2t_{\sigma(3)} \wedge t_{\sigma(4)}) \\ & \quad \Psi(t_{\sigma(1)} - v) \Psi(t_{\sigma(2)} - v) dv =: I_1 + I_2. \end{aligned}$$

Using Lemma 16, clearly

$$I_1 = \frac{a^2}{4(b-a)} \mathbb{1}_{t_{\sigma(3)} \wedge t_{\sigma(4)} > t_{\sigma(1)} \wedge t_{\sigma(2)}} \Psi(|t_{\sigma(1)} - t_{\sigma(2)}| + |t_{\sigma(3)} - t_{\sigma(4)}| + 2(t_{\sigma(3)} \wedge t_{\sigma(4)} - t_{\sigma(1)} \wedge t_{\sigma(2)})).$$

Moreover, on $\{t_{\sigma(3)} \wedge t_{\sigma(4)} \leq t_{\sigma(1)} \wedge t_{\sigma(2)}\}$,

$$\begin{aligned} I_2 &= \frac{a^2}{b-a} \\ &\left[\Psi(t_{\sigma(1)} \wedge t_{\sigma(2)} + |t_{\sigma(3)} - t_{\sigma(4)}| - t_{\sigma(3)} \wedge t_{\sigma(4)} + |t_{\sigma(1)} - t_{\sigma(2)}|) \right. \\ &\quad \left. - \Psi(|t_{\sigma(3)} - t_{\sigma(4)}| + (t_{\sigma(1)} - t_{\sigma(3)} \wedge t_{\sigma(4)}) + (t_{\sigma(2)} - t_{\sigma(3)} \wedge t_{\sigma(4)})) \right] \\ &= \frac{a^2}{b-a} \Psi(|t_{\sigma(1)} - t_{\sigma(2)}| + |t_{\sigma(3)} - t_{\sigma(4)}| + t_{\sigma(1)} \wedge t_{\sigma(2)} - t_{\sigma(3)} \wedge t_{\sigma(4)}) \\ &\quad \left[1 - e^{-(b-a)(t_{\sigma(1)} \wedge t_{\sigma(2)} - t_{\sigma(3)} \wedge t_{\sigma(4)})} \right], \end{aligned}$$

implying that

$$I_2 \leq \frac{a^2}{b-a} \Psi(|t_{\sigma(1)} - t_{\sigma(2)}| + |t_{\sigma(3)} - t_{\sigma(4)}| + t_{\sigma(1)} \wedge t_{\sigma(2)} - t_{\sigma(3)} \wedge t_{\sigma(4)}).$$

Putting things together, we obtain finally that

$$I_1 + I_2 \leq \frac{a^2}{b-a} \Psi(|t_{\sigma(3)} - t_{\sigma(4)}| + \beta(t_{\sigma(3)} \wedge t_{\sigma(4)} - t_{\sigma(1)} \wedge t_{\sigma(2)}) + |t_{\sigma(1)} - t_{\sigma(2)}|),$$

where we set $\beta(x) = |x| + x_+$ for any $x \in \mathbb{R}$. Thus,

$$\begin{aligned} S_4^{(2,\sigma)}(dt_1, dt_2, dt_3, dt_4) &\leq \frac{\mu}{1 - \frac{a}{b}} \left(1 + \frac{a}{b-a} \right) \\ &\quad \left(1 + \frac{a}{3(b-a)} \right) \Psi(|t_{\sigma(3)} - t_{\sigma(4)}|) \Psi \circ \beta(t_{\sigma(3)} \wedge t_{\sigma(4)} - t_{\sigma(1)} \wedge t_{\sigma(2)}) \Psi(|t_{\sigma(1)} - t_{\sigma(2)}|), \end{aligned}$$

and then

$$(95) \quad S_4^{(2,\sigma)}(dt_1, dt_2, dt_3, dt_4) \leq \frac{a^2 \mu}{(1 - \frac{a}{b})^3} \Psi(|t_{\sigma(3)} - t_{\sigma(4)}| + \beta(t_{\sigma(3)} \wedge t_{\sigma(4)} - t_{\sigma(1)} \wedge t_{\sigma(2)}) + |t_{\sigma(1)} - t_{\sigma(2)}|) dt_1 dt_2 dt_3 dt_4.$$

Bound on $S_4^{(3,\sigma)}$. We deduce from the third equation in (92) that, writing for short $dt_1^4 := dt_1 dt_2 dt_3 dt_4$,

$$\begin{aligned} S_4^{(3,\sigma)}(dt_1, dt_2, dt_3, dt_4) &= \frac{\mu}{1 - \frac{a}{b}} \\ &\left\{ \int_{\mathbb{R}} \left[\underbrace{\int_{\mathbb{R}} \Psi(v-u) \Psi(t_{\sigma(4)} - u) du + \Psi(v - t_{\sigma(4)})}_{= \frac{a}{2(b-a)} \Psi(|t_{\sigma(4)} - v|)} \right] \mathbb{W}[t_{\sigma(1)}, t_{\sigma(2)}, t_{\sigma(3)}](dv) \right\} dt_1^4 \\ &\leq \frac{\mu}{1 - \frac{a}{b}} \left(1 + \frac{1}{2} \frac{a}{b-a} \right) \left\{ \int_{\mathbb{R}} \Psi(|v - t_{\sigma(4)}|) \mathbb{W}[t_{\sigma(1)}, t_{\sigma(2)}, t_{\sigma(3)}](dv) \right\} dt_1^4 \\ &= \frac{\mu}{1 - \frac{a}{b}} \left(1 + \frac{1}{2} \frac{a}{b-a} \right) \left\{ a \Psi(|t_{\sigma(1)} \wedge t_{\sigma(2)} \wedge t_{\sigma(3)} - t_{\sigma(4)}|) \right. \\ &\quad \left. \Psi \left(\sum_{i=1}^3 t_{\sigma(i)} - 3t_{\sigma(1)} \wedge t_{\sigma(2)} \wedge t_{\sigma(3)} \right) \right. \\ &\quad \left. + \int_{\mathbb{R}} \Psi(|v - t_{\sigma(4)}|) \Psi(t_{\sigma(1)} - v) \Psi(t_{\sigma(2)} - v) \Psi(t_{\sigma(3)} - v) dv \right\} dt_1^4. \end{aligned}$$

Using (88) in Lemma 16, it follows that

$$S_4^{(3,\sigma)}(dt_1, dt_2, dt_3, dt_4) \leq \frac{a\mu}{1-\frac{a}{b}} \left(1 + \frac{1}{2} \frac{a}{b-a}\right)^2 \Psi(|t_{\sigma(1)} \wedge t_{\sigma(2)} \wedge t_{\sigma(3)} - t_{\sigma(4)}|) \\ \Psi\left(\sum_{i=1}^3 t_{\sigma(i)} - 3t_{\sigma(1)} \wedge t_{\sigma(2)} \wedge t_{\sigma(3)}\right) dt_1^4,$$

and then

$$(96) \quad S_4^{(3,\sigma)}(dt_1, dt_2, dt_3, dt_4) \\ \leq \frac{a^2\mu}{(1-\frac{a}{b})^3} \Psi\left(|t_{\sigma(4)} - t_{\sigma(1)} \wedge t_{\sigma(2)} \wedge t_{\sigma(3)}| + \sum_{i=1}^3 t_{\sigma(i)} - 3t_{\sigma(1)} \wedge t_{\sigma(2)} \wedge t_{\sigma(3)}\right) dt_1^4.$$

Bound on $S_4^{(4,\sigma)}$. We deduce from the fourth equation in (92) that

$$S_4^{(4,\sigma)}(dt_1, dt_2, dt_3, dt_4) \\ = \frac{\mu}{1-\frac{a}{b}} \left\{ \int_{\mathbb{R}^2} \left(\int_{\mathbb{R}} \Psi(v-u) \Psi(w-u) du \right) \mathbb{W}[t_{\sigma(1)}, t_{\sigma(2)}](dv) \mathbb{W}[t_{\sigma(3)}, t_{\sigma(4)}](dw) \right\} dt_1^4 \\ = \frac{a}{2b} \frac{\mu}{(1-\frac{a}{b})^2} \left\{ \int_{\mathbb{R}^2} \Psi(|v-w|) \mathbb{W}[t_{\sigma(1)}, t_{\sigma(2)}](dv) \mathbb{W}[t_{\sigma(3)}, t_{\sigma(4)}](dw) \right\} dt_1^4 \\ = \frac{a}{2b} \frac{\mu}{(1-\frac{a}{b})^2} \left\{ \int_{\mathbb{R}} \left[\Psi(|t_{\sigma(1)} \wedge t_{\sigma(2)} - w|) \Psi(|t_{\sigma(1)} - t_{\sigma(2)}|) \right. \right. \\ \left. \left. + \int_{\mathbb{R}} \Psi(|v-w|) \Psi(t_{\sigma(1)} - v) \Psi(t_{\sigma(2)} - v) dv \right] \mathbb{W}[t_{\sigma(3)}, t_{\sigma(4)}](dw) \right\} dt_1^4$$

Using (88) in Lemma 16, it follows that

$$S_4^{(4,\sigma)}(dt_1, dt_2, dt_3, dt_4) \\ \leq \frac{a}{2b} \frac{\mu}{(1-\frac{a}{b})^3} \Psi(|t_{\sigma(1)} - t_{\sigma(2)}|) \left\{ \int_{\mathbb{R}} \Psi(|t_{\sigma(1)} \wedge t_{\sigma(2)} - w|) \mathbb{W}[t_{\sigma(3)}, t_{\sigma(4)}](dw) \right\} dt_1^4 \\ \leq \frac{a}{2b} \frac{\mu}{(1-\frac{a}{b})^3} \Psi(|t_{\sigma(1)} - t_{\sigma(2)}|) \left\{ \Psi(|t_{\sigma(1)} \wedge t_{\sigma(2)} - t_{\sigma(3)} \wedge t_{\sigma(4)}|) \Psi(|t_{\sigma(3)} - t_{\sigma(4)}|) \right. \\ \left. + \int_{\mathbb{R}} \Psi(|t_{\sigma(1)} \wedge t_{\sigma(2)} - w|) \Psi(t_{\sigma(3)} - w) \Psi(t_{\sigma(4)} - w) dw \right\} dt_1^4$$

and then

$$(97) \quad S_4^{(4,\sigma)}(dt_1, dt_2, dt_3, dt_4) \\ \leq \frac{a^3}{2b} \frac{\mu}{(1-\frac{a}{b})^4} \Psi(|t_{\sigma(3)} - t_{\sigma(4)}| + |t_{\sigma(1)} \wedge t_{\sigma(2)} - t_{\sigma(3)} \wedge t_{\sigma(4)}| + |t_{\sigma(1)} - t_{\sigma(2)}|) dt_1^4.$$

Bound on $S_4^{(5,\sigma)}$. We deduce from the fifth equation in (92) that

$$\begin{aligned}
& S_4^{(5,\sigma)}(dt_1, dt_2, dt_3, dt_4) \\
&= \frac{\mu}{1 - \frac{a}{b}} \left\{ \int_{\mathbb{R}^2} \left(\underbrace{\int_{\mathbb{R}} \Psi(v-u)\Psi(t_{\sigma(4)}-u)du + \Psi(v-t_{\sigma(4)})}_{= \frac{a}{2(b-a)}\Psi(|v-t_{\sigma(4)}|)} \right) \right. \\
&\quad \left. \mathbb{W}[w, t_{\sigma(3)}](dv)\mathbb{W}[t_{\sigma(1)}, t_{\sigma(2)}](dw) \right\} dt_1^4 \\
&\leq \frac{\mu}{1 - \frac{a}{b}} \left(1 + \frac{1}{2} \frac{a}{b-a} \right) \left\{ \int_{\mathbb{R}^2} \Psi(|v-t_{\sigma(4)}|)\mathbb{W}[w, t_{\sigma(3)}](dv)\mathbb{W}[t_{\sigma(1)}, t_{\sigma(2)}](dw) \right\} dt_1^4 \\
&= \frac{\mu}{1 - \frac{a}{b}} \left(1 + \frac{1}{2} \frac{a}{b-a} \right) \left\{ \int_{\mathbb{R}} \left[\int_{\mathbb{R}} \Psi(|v-t_{\sigma(4)}|)\Psi(w-v)\Psi(t_{\sigma(3)}-v)dv \right. \right. \\
&\quad \left. \left. + \Psi(|t_{\sigma(3)}-t_{\sigma(4)}|)\Psi(w-t_{\sigma(3)}) \right] \mathbb{W}[t_{\sigma(1)}, t_{\sigma(2)}](dw) \right\} dt_1^4.
\end{aligned}$$

Using (88) in Lemma 16, we deduce that

$$\begin{aligned}
& S_4^{(5,\sigma)}(dt_1, dt_2, dt_3, dt_4) \\
&\leq \frac{\mu}{1 - \frac{a}{b}} \left(1 + \frac{1}{2} \frac{a}{b-a} \right) \left\{ \int_{\mathbb{R}} \left[\frac{a}{b-a} \Psi(|t_{\sigma(4)}-w \wedge t_{\sigma(3)}|)\Psi(|w-t_{\sigma(3)}|) \right. \right. \\
&\quad \left. \left. + \Psi(|t_{\sigma(3)}-t_{\sigma(4)}|)\Psi(w-t_{\sigma(3)}) \right] \mathbb{W}[t_{\sigma(1)}, t_{\sigma(2)}](dw) \right\} dt_1^4 \\
&\leq \frac{\mu}{1 - \frac{a}{b}} \left(1 + \frac{1}{2} \frac{a}{b-a} \right) \left(1 + \frac{a}{b-a} \right) \\
&\quad \left\{ \int_{\mathbb{R}} \Psi(|t_{\sigma(4)}-w \wedge t_{\sigma(3)}|)\Psi(|w-t_{\sigma(3)}|)\mathbb{W}[t_{\sigma(1)}, t_{\sigma(2)}](dw) \right\} dt_1^4 \\
&\leq \frac{\mu}{(1 - \frac{a}{b})^3} \left\{ \Psi(|t_{\sigma(4)}-t_{\sigma(1)} \wedge t_{\sigma(2)} \wedge t_{\sigma(3)}|)\Psi(|t_{\sigma(1)} \wedge t_{\sigma(2)}-t_{\sigma(3)}|)\Psi(|t_{\sigma(1)}-t_{\sigma(2)}|) \right. \\
&\quad \left. + \int_{\mathbb{R}} \Psi(|t_{\sigma(4)}-w \wedge t_{\sigma(3)}|)\Psi(|w-t_{\sigma(3)}|)\Psi(t_{\sigma(1)}-w)\Psi(t_{\sigma(2)}-w)dw \right\} dt_1^4.
\end{aligned}$$

Moreover,

$$\begin{aligned}
& \int_{\mathbb{R}} \Psi(|t_{\sigma(4)}-w \wedge t_{\sigma(3)}|)\Psi(|w-t_{\sigma(3)}|)\Psi(t_{\sigma(1)}-w)\Psi(t_{\sigma(2)}-w)dw \\
&= \int_{\mathbb{R}} \Psi(|t_{\sigma(4)}-w|)\Psi(t_{\sigma(3)}-w)\Psi(t_{\sigma(1)}-w)\Psi(t_{\sigma(2)}-w)dw \\
&\quad + \mathbb{1}_{t_{\sigma(3)} < t_{\sigma(1)} \wedge t_{\sigma(2)}} \int_{t_{\sigma(3)}}^{t_{\sigma(1)} \wedge t_{\sigma(2)}} \Psi(w-t_{\sigma(3)})\Psi(t_{\sigma(1)}-w)\Psi(t_{\sigma(2)}-w)dw \\
&\leq \frac{a^2}{2(b-a)} \Psi(|t_{\sigma(4)}-t_{\sigma(1)} \wedge t_{\sigma(2)} \wedge t_{\sigma(3)}|)\Psi \left(\sum_{i=1}^3 t_{\sigma(i)} - 3t_{\sigma(1)} \wedge t_{\sigma(2)} \wedge t_{\sigma(3)} \right) \\
&\quad + \frac{a^2}{b-a} \mathbb{1}_{t_{\sigma(3)} < t_{\sigma(1)} \wedge t_{\sigma(2)}} \Psi(|t_{\sigma(3)}-t_{\sigma(4)}|) \left[\Psi(t_{\sigma(1)}+t_{\sigma(2)}-t_{\sigma(3)}-t_{\sigma(1)} \wedge t_{\sigma(2)}) \right. \\
&\quad \left. - \Psi(t_{\sigma(1)}+t_{\sigma(2)}-2t_{\sigma(3)}) \right]
\end{aligned}$$

thanks to (88) in Lemma 16. By distinguishing the cases $t_{\sigma(3)} \geq t_{\sigma(1)} \wedge t_{\sigma(2)}$ and the case $t_{\sigma(3)} < t_{\sigma(1)} \wedge t_{\sigma(2)}$ in the first term, we obtain

$$\begin{aligned}
 & \int_{\mathbb{R}} \Psi(|t_{\sigma(4)} - w \wedge t_{\sigma(3)}|) \Psi(|w - t_{\sigma(3)}|) \Psi(t_{\sigma(1)} - w) \Psi(t_{\sigma(2)} - w) dw \\
 & \leq \frac{a^2}{2(b-a)} \mathbb{1}_{t_{\sigma(3)} \geq t_{\sigma(1)} \wedge t_{\sigma(2)}} \Psi(|t_{\sigma(4)} - t_{\sigma(1)} \wedge t_{\sigma(2)}|) \Psi\left(\underbrace{\sum_{i=1}^3 t_{\sigma(i)} - 3t_{\sigma(1)} \wedge t_{\sigma(2)}}_{=|t_{\sigma(1)} - t_{\sigma(2)}| + t_{\sigma(3)} - t_{\sigma(1)} \wedge t_{\sigma(2)}}\right) \\
 & + \frac{a^2}{b-a} \mathbb{1}_{t_{\sigma(3)} < t_{\sigma(1)} \wedge t_{\sigma(2)}} \Psi(|t_{\sigma(4)} - t_{\sigma(3)}|) \Psi(\underbrace{t_{\sigma(1)} + t_{\sigma(2)} - t_{\sigma(3)} - t_{\sigma(1)} \wedge t_{\sigma(2)}}_{=|t_{\sigma(1)} - t_{\sigma(2)}| + t_{\sigma(1)} \wedge t_{\sigma(2)} - t_{\sigma(3)}}) \\
 & \leq \frac{a}{b-a} \Psi(|t_{\sigma(4)} - t_{\sigma(1)} \wedge t_{\sigma(2)} \wedge t_{\sigma(3)}|) \Psi(|t_{\sigma(1)} - t_{\sigma(2)}|) \Psi(|t_{\sigma(3)} - t_{\sigma(1)} \wedge t_{\sigma(2)}|).
 \end{aligned}$$

Then coming back to the bound on $S_4^{5,\sigma}$, we deduce that

$$\begin{aligned}
 (98) \quad S_4^{(5,\sigma)}(dt_1, dt_2, dt_3, dt_4) \\
 \leq \frac{a^2 \mu}{(1 - \frac{a}{b})^4} \Psi(|t_{\sigma(4)} - t_{\sigma(1)} \wedge t_{\sigma(2)} \wedge t_{\sigma(3)}| + |t_{\sigma(1)} \wedge t_{\sigma(2)} - t_{\sigma(3)}| + |t_{\sigma(1)} - t_{\sigma(2)}|) dt_1^4.
 \end{aligned}$$

Thanks to the inequalities (94), (95), (96), (97) and (98), we deduce from (93) that the fourth order cumulant measure satisfies the bound

$$\begin{aligned}
 & \mathbb{1}_{\Delta_4}(t_1, t_2, t_3, t_4) k_4(dt_1, dt_2 dt_3, dt_4) \\
 & \leq \frac{a^2 \mu}{1 - \frac{a}{b}} \left[1 + \frac{1}{4} \frac{a}{b-a} + 6 \times \frac{1}{(1 - \frac{a}{b})^2} + 4 \times \frac{1}{(1 - \frac{a}{b})^2} + 3 \times \frac{a}{2b} \frac{1}{(1 - \frac{a}{b})^3} + 12 \times \frac{1}{(1 - \frac{a}{b})^3} \right] \\
 & \quad \times \Psi\left(\max_i t_i - \min_i t_i\right) dt_1 dt_2 dt_3 dt_4 \\
 & \leq \frac{Ca^2 \mu}{(1 - \frac{a}{b})^4} \Psi\left(\max_i t_i - \min_i t_i\right) dt_1 dt_2 dt_3 dt_4,
 \end{aligned}$$

which implies (41) with $l = 4$.

C.2. Proof of Eq. (12). Let us recall that the entire network can be constructed using the univariate Hawkes process N with intensity

$$\lambda(t) = \nu M + a \int_{(-\infty, t)} e^{-b(t-s)} dN_s,$$

for all t .

We introduce an i.i.d family $\{J_i, i \in \mathbb{Z}\}$ of uniform random variables on $\llbracket 1, M \rrbracket$, independent of anything else. Each jump time u of the Hawkes process N corresponds to a jump time of the neuron with label $J_{N_{u-+1}}$.

It follows that under the observed model, we have

$$\begin{aligned}
 \varphi(X_1^1, X_1^2) &= \int_0^T \int_0^T \mathbb{1}_{|u-v| \leq \delta} dX_1^1(u) dX_1^2(v) \\
 &= \int_0^T \int_0^T \mathbb{1}_{|u-v| \leq \delta} \mathbb{1}_{J_{N_{u-+1}}=1} \mathbb{1}_{J_{N_{v-+1}}=2} dN_u dN_v,
 \end{aligned}$$

and then, conditioning by N , we deduce that

$$\begin{aligned}\mathbb{E}[\varphi(X_1^1, X_1^2)] &= \mathbb{E}\left[\int_0^T \int_0^T \mathbb{1}_{|u-v|\leq\delta} \mathbb{P}(J_1 = 1, J_2 = 2) \mathbb{1}_{u\neq v} dN_u dN_v\right] \\ &= \frac{1}{M^2} \mathbb{E}\left[\int_0^T \int_0^T \mathbb{1}_{|u-v|\leq\delta} \mathbb{1}_{u\neq v} dN_u dN_v\right] \\ &= \frac{1}{M^2} \int_0^T \int_0^T \mathbb{1}_{|u-v|\leq\delta} \mathbb{1}_{u\neq v} \mathbb{E}[dN_u dN_v].\end{aligned}$$

Using the same idea under the independent model, we have

$$\begin{aligned}\varphi(X_1^1, X_1^2) &= \int_0^T \int_0^T \mathbb{1}_{|u-v|\leq\delta} dX_1^1(u) dX_1^2(v) \\ &= \int_0^T \int_0^T \mathbb{1}_{|u-v|\leq\delta} \mathbb{1}_{J_{N_u-+1}=1} \mathbb{1}_{J_{\tilde{N}_v-+1}=2} dN_u d\tilde{N}_v,\end{aligned}$$

where \tilde{N} is an independent copy of N , independent of $\{J_i, i \in \mathbb{Z}\}$. Conditioning on (N, \tilde{N}) , it follows that

$$\begin{aligned}\mathbb{E}_{\perp}[\varphi(X_1^1, X_1^2)] &= \mathbb{E}\left[\int_0^T \int_0^T \mathbb{1}_{|u-v|\leq\delta} \mathbb{P}(J_1 = 1) \mathbb{P}(J_1 = 2) dN_u d\tilde{N}_v\right] \\ &= \frac{1}{M^2} \mathbb{E}\left[\int_0^T \int_0^T \mathbb{1}_{|u-v|\leq\delta} dN_u d\tilde{N}_v\right] \\ &= \frac{1}{M^2} \int_0^T \int_0^T \mathbb{1}_{|u-v|\leq\delta} \mathbb{E}[dN_u] \mathbb{E}[dN_v].\end{aligned}$$

We deduce that

$$\begin{aligned}\mathbb{E}[\varphi(X_1^1, X_1^2)] - \mathbb{E}_{\perp}[\varphi(X_1^1, X_1^2)] &= \frac{1}{M^2} \int_0^T \int_0^T \mathbb{1}_{|u-v|\leq\delta} \mathbb{1}_{u\neq v} (\mathbb{E}[dN_u dN_v] - \mathbb{E}[dN_u] \mathbb{E}[dN_v]) \\ &= \frac{1}{M^2} \int_0^T \int_0^T \mathbb{1}_{|u-v|\leq\delta} \mathbb{1}_{u\neq v} k_2(du, dv),\end{aligned}$$

implying (45).

Using (40) in Lemma 9, we obtain moreover

$$\begin{aligned}\mathbb{E}[\varphi(X_1^1, X_1^2)] - \mathbb{E}_{\perp}[\varphi(X_1^1, X_1^2)] &= \frac{\nu}{M} \frac{a}{1 - \frac{a}{b}} \left(1 + \frac{1}{2} \frac{a}{b-a}\right) \int_0^T \int_0^T e^{-(b-a)|u-v|} \mathbb{1}_{|u-v|\leq\delta} dudv \\ &= \frac{\nu}{M} \frac{2a}{1 - \frac{a}{b}} \left(1 + \frac{1}{2} \frac{a}{b-a}\right) \int_0^T \int_0^u e^{-(b-a)(u-v)} \mathbb{1}_{u-v\leq\delta} dvdu \\ &= \frac{2\nu}{M} \frac{a}{1 - \frac{a}{b}} \left(1 + \frac{1}{2} \frac{a}{b-a}\right) \left\{ \int_0^\delta \int_0^u e^{-(b-a)(u-v)} dvdu + \int_\delta^T \int_{u-\delta}^u e^{-(b-a)(u-v)} dvdu \right\} \\ &= \frac{2\nu}{M} \frac{a}{1 - \frac{a}{b}} \left(1 + \frac{1}{2} \frac{a}{b-a}\right) \frac{1}{b-a} \left\{ \int_0^\delta [1 - e^{-(b-a)u}] du + (T - \delta) [1 - e^{-(b-a)\delta}] \right\} \\ &= \frac{\nu}{M} \frac{2 - \frac{a}{b}}{(1 - \frac{a}{b})^3} \frac{a}{b} \left\{ \delta - \frac{1 - e^{-(b-a)\delta}}{b-a} + (T - \delta) [1 - e^{-(b-a)\delta}] \right\}.\end{aligned}$$

Factorizing δ and recalling that $\ell = a/b$, we can rewrite the above expression as

$$\begin{aligned}\Delta_\varphi &= \frac{\nu\delta(2-\ell)\ell}{M(1-\ell)^3} \left\{ 1 - \frac{1 - e^{-(b-a)\delta}}{(b-a)\delta} + (T-\delta) \frac{1 - e^{-(b-a)\delta}}{\delta} \right\} \\ &= \frac{\nu\delta(2-\ell)\ell}{M(1-\ell)^3} \left\{ 1 + [(T-\delta)(b-a) - 1] \frac{1 - e^{-(b-a)\delta}}{(b-a)\delta} \right\} \\ &= \frac{\nu\delta(2-\ell)\ell}{M(1-\ell)^3} \left\{ 1 + [(T-\delta)b(1-\ell) - 1] \frac{1 - e^{-b(1-\ell)\delta}}{b(1-\ell)\delta} \right\},\end{aligned}$$

which gives (12).

Finally, under the additional condition $b(1-\ell)T > 4$, and using moreover that $\delta \leq T/2$, we have $T - \delta \geq T/2$ and that $-1 > -b(1-\ell)T/4$, such that

$$b(1-\ell)(T-\delta) - 1 \geq \frac{1}{4}b(1-\ell)T.$$

Since $2 - \ell \geq 1$, this gives the lower bound

$$\Delta_\varphi \geq \frac{\nu\delta\ell}{M(1-\ell)^3} \frac{b(1-\ell)T}{4} \frac{1 - e^{-b(1-\ell)\delta}}{b(1-\ell)\delta} = \frac{1}{4} \frac{\nu\ell T}{M(1-\ell)^3} (1 - e^{-b(1-\ell)\delta}),$$

which is (47).

C.3. Proof of Eq. (43). Let us now focus on the variance of $f_\varphi(X_1, X_2)$ for the model of independent neurons. The proof of the inequality (43) follows from the following intermediate result.

Lemma 17. *Let us consider the positive measures*

$$\begin{aligned}\mathcal{R}_\perp(du, dv) &= 2k_1(du)k_1(dv), \\ \mathcal{G}_\perp(du, dv, dv') &= 4\mathbb{1}_{v \neq v'}k_1(du)k_2(dv, dv') + 2k_1(du)k_1(dv)k_1(dv'),\end{aligned}$$

and

$$\begin{aligned}\mathcal{Q}_\perp(du, du', dv, dv') &= 2\mathbb{1}_{v \neq v'}k_1(du)k_1(du')k_2(dv, dv') \\ &\quad + 2\mathbb{1}_{u \neq u'}\mathbb{1}_{v \neq v'}k_2(du, du')k_2(dv, dv').\end{aligned}$$

Then the variance of $f_\varphi(X_1, X_2)$ satisfies the decomposition

$$\begin{aligned}(99) \quad \mathbb{V}_\perp[f_\varphi(X_1, X_2)] &= \frac{1}{M^2} \int_{[0, T]^2} \mathbb{1}_{|u-v| \leq \delta} \mathcal{R}_\perp(du, dv) \\ &\quad + \frac{1}{M^3} \int_{[0, T]^3} \mathbb{1}_{|u-v| \leq \delta} \mathbb{1}_{|u-v'| \leq \delta} \mathcal{G}_\perp(du, dv, dv') \\ &\quad + \frac{1}{M^4} \int_{[0, T]^4} \mathbb{1}_{|u-v| \leq \delta} \mathbb{1}_{|u'-v'| \leq \delta} \mathcal{Q}_\perp(du, du', dv, dv').\end{aligned}$$

Proof. The neurons are *i.i.d* for this model, and we recall that

$$f_\varphi(X_1, X_2) = \int_{[0, T]^2} \mathbb{1}_{|u-v| \leq \delta} dX_1^1(u) dX_1^2(v) - \int_{[0, T]^2} \mathbb{1}_{|u-v| \leq \delta} dX_1^1(u) dX_2^2(v).$$

Then the random variable $f_\varphi(X_1, X_2)$ is centered and

$$\begin{aligned}
\mathbb{V}_\perp [f_\varphi(X_1, X_2)] &= \mathbb{E}_\perp \left[|f_\varphi(X_1, X_2)|^2 \right] \\
&= \mathbb{E}_\perp \left[\int_{[0, T]^4} \mathbb{1}_{|u-v| \leq \delta} \mathbb{1}_{|u'-v'| \leq \delta} dX_1^1(u) dX_1^1(u') dX_1^2(v) dX_1^2(v') \right. \\
&\quad + \int_{[0, T]^4} \mathbb{1}_{|u-v| \leq \delta} \mathbb{1}_{|u'-v'| \leq \delta} dX_1^1(u) dX_1^1(u') dX_2^2(v) dX_2^2(v') \\
&\quad \left. - 2 \int_{[0, T]^4} \mathbb{1}_{|u-v| \leq \delta} \mathbb{1}_{|u'-v'| \leq \delta} dX_1^1(u) dX_1^1(u') dX_1^2(v) dX_2^2(v') \right] \\
&= 2\mathbb{E}_\perp \left[\int_{[0, T]^4} \mathbb{1}_{|u-v| \leq \delta} \mathbb{1}_{|u'-v'| \leq \delta} dX_1^1(u) dX_1^1(u') dX_1^2(v) dX_1^2(v') \right. \\
&\quad \left. - \int_{[0, T]^4} \mathbb{1}_{|u-v| \leq \delta} \mathbb{1}_{|u'-v'| \leq \delta} dX_1^1(u) dX_1^1(u') dX_1^2(v) dX_2^2(v') \right],
\end{aligned}$$

since the stochastic integral $\int \mathbb{1}_{|u-v| \leq \delta} \mathbb{1}_{|u'-v'| \leq \delta} dX_1^1(u) dX_1^1(u') dX_1^2(v) dX_1^2(v')$ has under the independent model the same distribution as $\int \mathbb{1}_{|u-v| \leq \delta} \mathbb{1}_{|u'-v'| \leq \delta} dX_1^1(u) dX_1^1(u') dX_2^2(v) dX_2^2(v')$. By distinguishing simultaneous and non simultaneous jump times, it follows that

$$\begin{aligned}
\mathbb{V}_\perp [f_\varphi(X_1, X_2)] &= 2\mathbb{E}_\perp \left[\int_{[0, T]^3} \mathbb{1}_{|u-v| \leq \delta} \mathbb{1}_{|u-v'| \leq \delta} dX_1^1(u) dX_1^2(v) dX_1^2(v') \right. \\
&\quad \left. - \int_{[0, T]^3} \mathbb{1}_{|u-v| \leq \delta} \mathbb{1}_{|u-v'| \leq \delta} dX_1^1(u) dX_1^2(v) dX_2^2(v') \right] \\
&\quad + 2\mathbb{E}_\perp \left[\int_{[0, T]^4} \mathbb{1}_{|u-v| \leq \delta} \mathbb{1}_{|u'-v'| \leq \delta} \mathbb{1}_{u \neq u'} dX_1^1(u) dX_1^1(u') dX_1^2(v) dX_1^2(v') \right. \\
&\quad \left. - \int_{[0, T]^4} \mathbb{1}_{|u-v| \leq \delta} \mathbb{1}_{|u'-v'| \leq \delta} \mathbb{1}_{u \neq u'} dX_1^1(u) dX_1^1(u') dX_1^2(v) dX_2^2(v') \right],
\end{aligned}$$

and then

$$\begin{aligned}
\mathbb{V}_\perp [f_\varphi(X_1, X_2)] &= 2\mathbb{E}_\perp \left[\int_{[0, T]^2} \mathbb{1}_{|u-v| \leq \delta} dX_1^1(u) dX_1^2(v) \right. \\
&\quad + \int_{[0, T]^3} \mathbb{1}_{|u-v| \leq \delta} \mathbb{1}_{|u-v'| \leq \delta} \mathbb{1}_{v \neq v'} dX_1^1(u) dX_1^2(v) dX_1^2(v') \\
&\quad \left. - \int_{[0, T]^3} \mathbb{1}_{|u-v| \leq \delta} \mathbb{1}_{|u-v'| \leq \delta} dX_1^1(u) dX_1^2(v) dX_2^2(v') \right] \\
&\quad + 2\mathbb{E}_\perp \left[\int_{[0, T]^3} \mathbb{1}_{|u-v| \leq \delta} \mathbb{1}_{|u'-v| \leq \delta} \mathbb{1}_{u \neq u'} dX_1^1(u) dX_1^1(u') dX_1^2(v) \right. \\
&\quad + \int_{[0, T]^4} \mathbb{1}_{|u-v| \leq \delta} \mathbb{1}_{|u'-v'| \leq \delta} \mathbb{1}_{u \neq u'} \mathbb{1}_{v \neq v'} dX_1^1(u) dX_1^1(u') dX_1^2(v) dX_1^2(v') \\
&\quad \left. - \int_{[0, T]^4} \mathbb{1}_{|u-v| \leq \delta} \mathbb{1}_{|u'-v'| \leq \delta} \mathbb{1}_{u \neq u'} dX_1^1(u) dX_1^1(u') dX_1^2(v) dX_2^2(v') \right].
\end{aligned}$$

We deduce that

$$\begin{aligned} \mathbb{V}_{\perp} [f_{\varphi}(X_1, X_2)] &= 2 \left\{ \frac{1}{M^2} \int_{[0, T]^2} \mathbb{1}_{|u-v| \leq \delta} \mathbb{E}[dN_u] \mathbb{E}[dN_v] \right. \\ &\quad + \frac{1}{M^3} \int_{[0, T]^3} \mathbb{1}_{|u-v| \leq \delta} \mathbb{1}_{|u-v'| \leq \delta} \mathbb{1}_{v \neq v'} \mathbb{E}[dN_u] k_2(dv, dv') \\ &\quad + \frac{1}{M^3} \int_{[0, T]^3} \mathbb{1}_{|u-v| \leq \delta} \mathbb{1}_{|u'-v| \leq \delta} \mathbb{1}_{u \neq u'} \mathbb{E}[dN_v] \mathbb{E}[dN_u dN_{u'}] \\ &\quad \left. + \frac{1}{M^4} \int_{[0, T]^4} \mathbb{1}_{|u-v| \leq \delta} \mathbb{1}_{|u'-v'| \leq \delta} \mathbb{1}_{u \neq u'} \mathbb{1}_{v \neq v'} \mathbb{E}[dN_u dN_{u'}] k_2(dv, dv') \right\}. \end{aligned}$$

Using the fact that $\mathbb{E}[dN_u dN_v] = k_2(du, dv) + \mathbb{E}[dN_u] \mathbb{E}[dN_v]$ and $\mathbb{E}[dN_u] = k_1(du)$, it follows that the term in the third integral that is related to the second order cumulant measure is equal to the second integral. We obtain

$$\begin{aligned} \mathbb{V}_{\perp} [f_{\varphi}(X_1, X_2)] &= 2 \left\{ \frac{1}{M^2} \int_{[0, T]^2} \mathbb{1}_{|u-v| \leq \delta} k_1(du) k_1(dv) \right. \\ &\quad + \frac{2}{M^3} \int_{[0, T]^3} \mathbb{1}_{|u-v| \leq \delta} \mathbb{1}_{|u-v'| \leq \delta} \mathbb{1}_{v \neq v'} k_1(du) k_2(dv, dv') \\ &\quad + \frac{1}{M^3} \int_{[0, T]^3} \mathbb{1}_{|u-v| \leq \delta} \mathbb{1}_{|u'-v| \leq \delta} k_1(du) k_1(dv) k_1(du') \\ &\quad + \frac{1}{M^4} \int_{[0, T]^4} \mathbb{1}_{|u-v| \leq \delta} \mathbb{1}_{|u'-v'| \leq \delta} \mathbb{1}_{u \neq u'} \mathbb{1}_{v \neq v'} k_1(du) k_1(du') k_2(dv, dv') \\ &\quad \left. + \frac{1}{M^4} \int_{[0, T]^4} \mathbb{1}_{|u-v| \leq \delta} \mathbb{1}_{|u'-v'| \leq \delta} \mathbb{1}_{u \neq u'} \mathbb{1}_{v \neq v'} k_2(du, du') k_2(dv, dv') \right\}, \end{aligned}$$

ending the proof. \square

Proof of (42). Recall that in the mean-field limit, neurons 1 and 2 are independent and identically distributed, and their spike trains are Poisson processes with intensity $\frac{\nu}{1-\ell}$. The variance \mathbb{V}_0 of $f_{\varphi}(X_1, X_2)$ under this limit model can then be deduced from (68). \square

Coming back to the Hawkes model, we deduce from (99) in Lemma 17 and (39) in Lemma 9 that

$$\begin{aligned} \mathbb{V}_{\perp} [f_{\varphi}(X_1, X_2)] - \mathbb{V}_0 &= 2 \left\{ \frac{2}{M^3} \int_{[0, T]^3} \mathbb{1}_{|u-v| \leq \delta} \mathbb{1}_{|u-v'| \leq \delta} \mathbb{1}_{v \neq v'} k_1(du) k_2(dv, dv') \right. \\ &\quad + \frac{1}{M^4} \int_{[0, T]^4} \mathbb{1}_{|u-v| \leq \delta} \mathbb{1}_{|u'-v'| \leq \delta} \mathbb{1}_{v \neq v'} k_1(du) k_1(du') k_2(dv, dv') \\ &\quad \left. + \frac{1}{M^4} \int_{[0, T]^4} \mathbb{1}_{|u-v| \leq \delta} \mathbb{1}_{|u'-v'| \leq \delta} \mathbb{1}_{u \neq u'} \mathbb{1}_{v \neq v'} k_2(du, du') k_2(dv, dv') \right\}, \end{aligned}$$

which is non negative. Using the bound (41) in Lemma 9 for the second order cumulant, we obtain for a certain constant $C > 0$ (that can change from one line to another)

$$\begin{aligned}
& \mathbb{V}_{\perp} [f_{\varphi}(X_1, X_2)] - \mathbb{V}_0 \\
& \leq C \left\{ \frac{a}{M} \frac{\nu^2}{\left(1 - \frac{a}{b}\right)^3} \int_{[0, T]^3} e^{-(b-a)|v-v'|} \mathbb{1}_{|u-v| \leq \delta} \mathbb{1}_{|u-v'| \leq \delta} dudvdv' \right. \\
& \quad + \frac{a}{M} \frac{\nu^3}{\left(1 - \frac{a}{b}\right)^4} \int_{[0, T]^4} e^{-(b-a)|v-v'|} \mathbb{1}_{|u-v| \leq \delta} \mathbb{1}_{|u'-v'| \leq \delta} dudv' dv dv' \\
& \quad \left. + \frac{1}{M^2} \frac{a^2 \nu^2}{\left(1 - \frac{a}{b}\right)^4} \int_{[0, T]^4} e^{-(b-a)|u-u'|} e^{-(b-a)|v-v'|} \mathbb{1}_{|u-v| \leq \delta} \mathbb{1}_{|u'-v'| \leq \delta} dudv' dv dv' \right\} \\
& \leq C \left\{ \frac{a}{M} \frac{\nu^2}{\left(1 - \frac{a}{b}\right)^3} \int_{[0, T]^2} e^{-(b-a)|v-v'|} \underbrace{[(v \wedge v' + \delta) \wedge T - (v \vee v' - \delta)_+]_+}_{\leq (2\delta - |v-v'|)_+} dv dv' \right. \\
& \quad + \frac{a}{M} \frac{\nu^2}{\left(1 - \frac{a}{b}\right)^4} \left(\nu + \frac{a}{M} \right) \times \\
& \quad \left. \times \int_{[0, T]^2} e^{-(b-a)|v-v'|} \underbrace{[(v + \delta) \wedge T - (v - \delta)_+]_{\leq 2\delta}} \underbrace{[(v' + \delta) \wedge T - (v' - \delta)_+]_{\leq 2\delta}} dv dv' \right\}.
\end{aligned}$$

It follows that

$$\begin{aligned}
\mathbb{V}_{\perp} [f_{\varphi}(X_1, X_2)] - \mathbb{V}_0 & \leq C \left\{ \frac{a}{M} \frac{\nu^2}{\left(1 - \frac{a}{b}\right)^3} \int_0^T \int_0^T e^{-(b-a)|v-v'|} (2\delta - |v-v'|)_+ dv dv' \right. \\
& \quad \left. + \frac{a}{M} \frac{\nu^2}{\left(1 - \frac{a}{b}\right)^4} \left(\nu + \frac{a}{M} \right) \times 4\delta^2 \int_0^T \int_0^T e^{-(b-a)|v-v'|} dv dv' \right\} \\
& = C \left\{ \frac{a}{M} \frac{\nu^2}{\left(1 - \frac{a}{b}\right)^3} \times 2 \int_0^T \int_0^{v \wedge (2\delta)} e^{-(b-a)z} (2\delta - z) dz dv \right. \\
& \quad \left. + \frac{a}{M} \frac{\nu^2}{\left(1 - \frac{a}{b}\right)^4} \left(\nu + \frac{a}{M} \right) \times 4\delta^2 \int_0^T \int_0^T e^{-(b-a)|v-v'|} dv dv' \right\} \\
& \leq C \left\{ \frac{a}{M} \frac{\nu^2}{\left(1 - \frac{a}{b}\right)^3} \times 2 \int_0^T \int_0^{2\delta} e^{-(b-a)z} (2\delta) dz dv \right. \\
& \quad \left. + \frac{a}{M} \frac{\nu^2}{\left(1 - \frac{a}{b}\right)^4} \left(\nu + \frac{a}{M} \right) \times 4\delta^2 \int_0^T \int_0^T e^{-(b-a)|v-v'|} dv dv' \right\}.
\end{aligned}$$

Since

$$(2\delta) \int_0^T \int_0^{2\delta} e^{-(b-a)z} dz dv \leq 2\delta T \frac{1}{b-a} (1 - e^{-(b-a)2\delta}) \leq 4\delta^2 T,$$

where we have used that $1 - e^{-x} \leq x$, this implies that

$$\begin{aligned}
\mathbb{V}_{\perp} [f_{\varphi}(X_1, X_2)] - \mathbb{V}_0 & \leq C \left\{ \frac{a}{M} \frac{\nu^2}{\left(1 - \frac{a}{b}\right)^3} \delta^2 T \right. \\
& \quad \left. + \frac{a}{M} \frac{\nu^2}{\left(1 - \frac{a}{b}\right)^4} \left(\nu + \frac{a}{M} \right) \times \frac{\delta^2 T}{b-a} \left[1 - \frac{1 - e^{-(b-a)T}}{(b-a)T} \right] \right\}.
\end{aligned}$$

Using the inequality $1 - \frac{1-e^{-x}}{x} \leq 1 \wedge \frac{x}{2}$ for any $x > 0$, we obtain

$$\begin{aligned}
 & \mathbb{V}_{\perp} [f_{\varphi}(X_1, X_2)] - \mathbb{V}_0 \\
 & \leq C \left\{ \frac{a}{M} \frac{\nu^2}{(1-\frac{a}{b})^3} \delta^2 T + \frac{a}{M} \frac{\nu^2}{(1-\frac{a}{b})^4} \left(\nu + \frac{a}{M} \right) \times \delta^2 T \left(\frac{T}{2} \wedge \frac{1}{b-a} \right) \right\} \\
 & = C \frac{aT}{M} \frac{\nu^2 \delta^2}{(1-\frac{a}{b})^3} \left\{ 1 + \frac{\nu + \frac{a}{M}}{1-\frac{a}{b}} \left(\frac{T}{2} \wedge \frac{1}{b-a} \right) \right\} \\
 & \leq C \frac{aT}{M} \frac{\nu^2 \delta^2}{(1-\frac{a}{b})^3} \left\{ 1 + \frac{\nu + \frac{a}{M}}{b(1-\frac{a}{b})^2} \right\} = C \frac{aT}{M} \frac{\nu^2 \delta^2}{(1-\ell)^3} \left\{ 1 + \frac{\nu + \frac{a}{M}}{b(1-\ell)^2} \right\},
 \end{aligned}$$

which is (43) and ends the proof.

C.4. Proof of Eq. (44). We are now interested in the variance of $f_{\varphi}(X_1, X_2)$ for the observed model, that is the one for which the neurons are in the same network. Then, the proof of the bound (44) follows from the following intermediate result.

Lemma 18. *Let us consider the positive measures defined by*

$$\begin{aligned}
 \mathcal{R}(du, dv) &= \mathbb{1}_{u \neq v} k_2(du, dv) + 2k_1(du)k_1(dv), \\
 \mathcal{G}(du, dv, dv') &= 2\mathbb{1}_{\Delta_3} \{k_3(du, dv, dv') + 2k_1(du)k_2(dv, dv') \\
 & \quad + k_1(dv)k_2(du, dv') + k_1(du)k_1(dv)k_1(dv')\}, \\
 \mathcal{Q}(du, du', dv, dv') &= \mathbb{1}_{\Delta_4} \{k_4(du, du', dv, dv') + 2k_1(du)k_3(du', dv, dv') \\
 & \quad + 3k_2(du, du')k_2(dv, dv') + 2k_1(du)k_1(du')k_2(dv, dv')\}.
 \end{aligned}$$

Then the variance of $f_{\varphi}(X_1, X_2)$ satisfies the decomposition

$$\begin{aligned}
 (100) \quad \mathbb{V}[f_{\varphi}(X_1, X_2)] &= \frac{1}{M^2} \int_{[0, T]^2} \mathbb{1}_{|u-v| \leq \delta} \mathcal{R}(du, dv) \\
 & \quad + \frac{1}{M^3} \int_{[0, T]^3} \mathbb{1}_{|u-v| \leq \delta} \mathbb{1}_{|u-v'| \leq \delta} \mathcal{G}(du, dv, dv') \\
 & \quad + \frac{1}{M^4} \int_{[0, T]^4} \mathbb{1}_{|u-v| \leq \delta} \mathbb{1}_{|u'-v'| \leq \delta} \mathcal{Q}(du, du', dv, dv').
 \end{aligned}$$

Proof. The proof of this result uses the same ideas as the one of Lemma 17 taking into account the fact that X_1^1 and X_1^2 belong to the same network under the observed model. The moment of order 2 of $f_{\varphi}(X_1, X_2)$ is then given by

$$\begin{aligned}
 \mathbb{E} \left[|f_{\varphi}(X_1, X_2)|^2 \right] &= \mathbb{E} \left[\int_{[0, T]^4} \mathbb{1}_{|u-v| \leq \delta} \mathbb{1}_{|u'-v'| \leq \delta} dX_1^1(u) dX_1^1(u') dX_1^2(v) dX_1^2(v') \right. \\
 & \quad + \int_{[0, T]^4} \mathbb{1}_{|u-v| \leq \delta} \mathbb{1}_{|u'-v'| \leq \delta} dX_1^1(u) dX_1^1(u') dX_2^2(v) dX_2^2(v') \\
 & \quad \left. - 2 \int_{[0, T]^4} \mathbb{1}_{|u-v| \leq \delta} \mathbb{1}_{|u'-v'| \leq \delta} dX_1^1(u) dX_1^1(u') dX_1^2(v) dX_2^2(v') \right],
 \end{aligned}$$

and by distinguishing simultaneous jumps, we get

$$\begin{aligned}
\mathbb{E} \left[|f_\varphi(X_1, X_2)|^2 \right] &= \mathbb{E} \left[\int_{[0, T]^2} \mathbb{1}_{|u-v| \leq \delta} dX_1^1(u) dX_1^2(v) \right. \\
&+ \int_{[0, T]^2} \mathbb{1}_{|u-v| \leq \delta} \mathbb{1}_{|u-v'| \leq \delta} \mathbb{1}_{v \neq v'} dX_1^1(u) dX_1^2(v) dX_1^2(v') \\
&+ \int_{[0, T]^2} \mathbb{1}_{|u-v| \leq \delta} dX_1^1(u) dX_2^2(v) \\
&+ \int_{[0, T]^3} \mathbb{1}_{|u-v| \leq \delta} \mathbb{1}_{|u-v'| \leq \delta} \mathbb{1}_{v \neq v'} dX_1^1(u) dX_2^2(v) dX_2^2(v') \\
&- 2 \int_{[0, T]^3} \mathbb{1}_{|u-v| \leq \delta} \mathbb{1}_{|u-v'| \leq \delta} dX_1^1(u) dX_1^2(v) dX_2^2(v') \\
&+ \int_{[0, T]^3} \mathbb{1}_{|u-v| \leq \delta} \mathbb{1}_{|u'-v| \leq \delta} \mathbb{1}_{u \neq u'} dX_1^1(u) dX_1^1(u') dX_1^2(v) \\
&+ \int_{[0, T]^4} \mathbb{1}_{|u-v| \leq \delta} \mathbb{1}_{|u'-v'| \leq \delta} \mathbb{1}_{u \neq u'} \mathbb{1}_{v \neq v'} dX_1^1(u) dX_1^1(u') dX_1^2(v) dX_1^2(v') \\
&+ \int_{[0, T]^3} \mathbb{1}_{|u-v| \leq \delta} \mathbb{1}_{|u'-v| \leq \delta} \mathbb{1}_{u \neq u'} dX_1^1(u) dX_1^1(u') dX_2^2(v) \\
&+ \int_{[0, T]^4} \mathbb{1}_{|u-v| \leq \delta} \mathbb{1}_{|u'-v'| \leq \delta} \mathbb{1}_{u \neq u'} \mathbb{1}_{v \neq v'} dX_1^1(u) dX_1^1(u') dX_2^2(v) dX_2^2(v') \\
&\left. - 2 \int_{[0, T]^4} \mathbb{1}_{|u-v| \leq \delta} \mathbb{1}_{|u'-v'| \leq \delta} \mathbb{1}_{u \neq u'} dX_1^1(u) dX_1^1(u') dX_1^2(v) dX_2^2(v') \right].
\end{aligned}$$

Since $\mathcal{L}(X_1^1, X_1^2) = \mathcal{L}(X_1^2, X_1^1)$ by exchangeability of the neurons, the random variables given by

$$\int \mathbb{1}_{|u-v| \leq \delta} \mathbb{1}_{|u-v'| \leq \delta} \mathbb{1}_{v \neq v'} dX_1^1(u) dX_1^2(v) dX_1^2(v')$$

and

$$\int \mathbb{1}_{|u-v| \leq \delta} \mathbb{1}_{|u'-v| \leq \delta} \mathbb{1}_{u \neq u'} dX_1^1(u) dX_1^1(u') dX_1^2(v)$$

have the same distribution. The same reason implies that the random variables

$$\begin{aligned}
&\int \mathbb{1}_{|u-v| \leq \delta} \mathbb{1}_{|u-v'| \leq \delta} \mathbb{1}_{v \neq v'} dX_1^1(u) dX_2^2(v) dX_2^2(v') \\
&\text{and} \quad \int \mathbb{1}_{|u-v| \leq \delta} \mathbb{1}_{|u'-v| \leq \delta} \mathbb{1}_{u \neq u'} dX_1^1(u) dX_1^1(u') dX_2^2(v)
\end{aligned}$$

have the same distribution. As a consequence, we obtain

$$\begin{aligned}
&\mathbb{E} \left[|f_\varphi(X_1, X_2)|^2 \right] \\
&= \mathbb{E} \left[\int \mathbb{1}_{|u-v| \leq \delta} dX_1^1(u) dX_1^2(v) + 2 \int \mathbb{1}_{|u-v| \leq \delta} \mathbb{1}_{|u-v'| \leq \delta} \mathbb{1}_{v \neq v'} dX_1^1(u) dX_1^2(v) dX_1^2(v') \right. \\
&+ \int \mathbb{1}_{|u-v| \leq \delta} dX_1^1(u) dX_2^2(v) + 2 \int \mathbb{1}_{|u-v| \leq \delta} \mathbb{1}_{|u-v'| \leq \delta} \mathbb{1}_{v \neq v'} dX_1^1(u) dX_2^2(v) dX_2^2(v') \\
&- 2 \int \mathbb{1}_{|u-v| \leq \delta} \mathbb{1}_{|u-v'| \leq \delta} dX_1^1(u) dX_1^2(v) dX_2^2(v') \\
&+ \int \mathbb{1}_{|u-v| \leq \delta} \mathbb{1}_{|u'-v| \leq \delta} \mathbb{1}_{u \neq u'} \mathbb{1}_{v \neq v'} dX_1^1(u) dX_1^1(u') dX_1^2(v) dX_1^2(v') \\
&+ \int \mathbb{1}_{|u-v| \leq \delta} \mathbb{1}_{|u'-v'| \leq \delta} \mathbb{1}_{u \neq u'} \mathbb{1}_{v \neq v'} dX_1^1(u) dX_1^1(u') dX_2^2(v) dX_2^2(v') \\
&\left. - 2 \int \mathbb{1}_{|u-v| \leq \delta} \mathbb{1}_{|u'-v'| \leq \delta} \mathbb{1}_{u \neq u'} dX_1^1(u) dX_1^1(u') dX_1^2(v) dX_2^2(v') \right].
\end{aligned}$$

Then using (45), that is,

$$\mathbb{E}[f_\varphi(X_1, X_2)] = \frac{1}{M^2} \int_{[0, T]^2} \mathbb{1}_{|u-v| \leq \delta} \mathbb{1}_{u \neq v} k_2(du, dv),$$

it follows that

$$\begin{aligned} \mathbb{V}[f_\varphi(X_1, X_2)] &= \frac{1}{M^2} \int_{[0, T]^2} \mathbb{1}_{|u-v| \leq \delta} \mathcal{R}(du, dv) \\ &\quad + \frac{1}{M^3} \int_{[0, T]^3} \mathbb{1}_{|u-v| \leq \delta} \mathbb{1}_{|u-v'| \leq \delta} \mathcal{G}(du, dv, dv') \\ &\quad + \frac{1}{M^4} \int_{[0, T]^4} \mathbb{1}_{|u-v| \leq \delta} \mathbb{1}_{|u'-v'| \leq \delta} \mathcal{Q}(du, du', dv, dv'), \end{aligned}$$

where we set

$$\begin{aligned} \mathcal{R}(du, dv) &= \mathbb{1}_{u \neq v} \mathbb{E}[dN_u dN_v] + \mathbb{E}[dN_u] \mathbb{E}[dN_v], \\ \mathcal{G}(du, dv, dv') &= 2 \mathbb{1}_{\Delta_3} \left\{ \mathbb{E}[dN_u dN_v dN_{v'}] + \mathbb{E}[dN_u] \mathbb{E}[dN_v dN_{v'}] \right. \\ &\quad \left. - \mathbb{E}[dN_{v'}] \mathbb{E}[dN_u dN_v] \right\}, \\ \mathcal{Q}(du, du', dv, dv') &= \mathbb{1}_{\Delta_4} \left\{ \mathbb{E}[dN_u dN_{u'} dN_v dN_{v'}] + \mathbb{E}[dN_u dN_{u'}] \mathbb{E}[dN_v dN_{v'}] \right. \\ &\quad \left. - 2 \mathbb{E}[dN_{v'}] \mathbb{E}[dN_u dN_{u'} dN_v] - k_2(du, dv) k_2(du', dv') \right\}. \end{aligned}$$

Since $k_1(du) = \mathbb{E}[dN_u]$ and $\mathbb{E}[dN_u dN_v] = k_2(du, dv) + k_1(du)k_1(dv)$, we get

$$\mathcal{R}(du, dv) = \mathbb{1}_{u \neq v} k_2(du, dv) + 2k_1(du)k_1(dv).$$

In addition, using (27),

$$\begin{aligned} \mathcal{G}(du, dv, dv') &= 2 \mathbb{1}_{\Delta_3} \left\{ k_3(du, dv, dv') + k_1(du)k_2(dv, dv') + k_1(dv)k_2(du, dv') \right. \\ &\quad + k_1(dv')k_2(du, dv) + k_1(du)k_1(dv)k_1(dv') \\ &\quad + k_1(du) \left(k_2(dv, dv') + k_1(dv)k_1(dv') \right) \\ &\quad \left. - k_1(dv') \left(k_2(du, dv) + k_1(du)k_1(dv) \right) \right\}, \end{aligned}$$

such that

$$\begin{aligned} \mathcal{G}(du, dv, dv') &= 2 \mathbb{1}_{\Delta_3} \left\{ k_3(du, dv, dv') + 2k_1(du)k_2(dv, dv') \right. \\ &\quad \left. + k_1(dv)k_2(du, dv') + k_1(du)k_1(dv)k_1(dv') \right\}. \end{aligned}$$

The following steps are devoted to obtain a simpler representation of the measure $\mathcal{Q}(du, du', dv, dv')$. Using (26), we get

$$\begin{aligned} \mathcal{Q}(du, du', dv, dv') &= \mathbb{1}_{\Delta_4} \left\{ \sum_{\substack{\mathbf{p} \text{ partition of} \\ \{u, u', v, v'\}}} \bigotimes_{B \in \mathbf{p}} k_{\#B} \left(\bigotimes_{u \in B} du \right) \right. \\ &\quad + \left(k_2(du, du') + k_1(du)k_1(du') \right) \left(k_2(dv, dv') + k_1(dv)k_1(dv') \right) \\ &\quad \left. - 2k_1(dv') \sum_{\substack{\mathbf{p} \text{ partition of} \\ \{u, u', v\}}} \bigotimes_{B \in \mathbf{p}} k_{\#B} \left(\bigotimes_{u \in B} du \right) - k_2(du, dv)k_2(du', dv') \right\}. \end{aligned}$$

We extract from the first sum the partitions that contain the singleton $\{v'\}$. The sum over this subset of partitions gives $k_1(dv') \sum_{\mathbf{p} \text{ partition of } \{u, u', v\}} \otimes_{B \in \mathbf{p}} k_{\#B}(\otimes_{u \in B} du)$, and then

$$\begin{aligned} \mathcal{Q}(du, du', dv, dv') &= \mathbb{1}_{\Delta_4} \left\{ \sum_{\substack{\mathbf{p} \text{ partition of } \{u, u', v, v'\} \\ \text{s.t. } \{v'\} \notin \mathbf{p}}} \otimes_{B \in \mathbf{p}} k_{\#B}(\otimes_{u \in B} du) + k_2(du, du')k_2(dv, dv') \right. \\ &+ k_2(du, du')k_1(dv)k_1(dv') + k_1(du)k_1(du')k_2(dv, dv') + k_1(du)k_1(du')k_1(dv)k_1(dv') \\ &\left. - k_1(dv') \sum_{\mathbf{p} \text{ partition of } \{u, u', v\}} \otimes_{B \in \mathbf{p}} k_{\#B}(\otimes_{u \in B} du) - k_2(du, dv)k_2(du', dv') \right\}. \end{aligned}$$

We extract the partition $\{\{u, v\}, \{u', v'\}\}$ from the first sum, and the partitions $\{\{u\}, \{u'\}, \{v\}\}$ and $\{\{u, u'\}, \{v\}\}$ from the second one. Then we obtain

$$(101) \quad \mathcal{Q}(du, du', dv, dv') = \mathbb{1}_{\Delta_4} \left\{ \sum_{\mathbf{p} \in \Pi^1(u, u', v, v')} \otimes_{B \in \mathbf{p}} k_{\#B}(\otimes_{u \in B} du) + k_2(du, du')k_2(dv, dv') \right. \\ \left. + k_1(du)k_1(du')k_2(dv, dv') - k_1(dv') \sum_{\mathbf{p} \in \Pi^2(u, u', v)} \otimes_{B \in \mathbf{p}} k_{\#B}(\otimes_{u \in B} du) \right\}$$

where

$$\begin{aligned} \Pi^1(u, u', v, v') &:= \{\mathbf{p} \text{ partition of } \{u, u', v, v'\} : \{v'\} \notin \mathbf{p} \text{ and } \mathbf{p} \neq \{\{u, v\}, \{u', v'\}\}\} \\ \Pi^2(u, u', v) &:= \{\mathbf{p} \text{ partition of } \{u, u', v\} : |\mathbf{p}| < 3 \text{ and } \mathbf{p} \neq \{\{u, u'\}, \{v\}\}\} \\ &= \left\{ \{\{u, u', v\}\}, \{\{u\}, \{u', v\}\}, \{\{u'\}, \{u, v\}\} \right\}. \end{aligned}$$

We now verify that the negative measure in the expression of $\mathcal{Q}(du, du', dv, dv')$ is compensated when we compute the integral $\int_{[0, T]^4} \mathbb{1}_{|u-v| \leq \delta} \mathbb{1}_{|u'-v'| \leq \delta} \mathbb{1}_{\Delta_4} k_1(dv')k_3(du, du', dv)$. This directly holds by symmetry. Indeed, the integral with the partition $\{\{u, u', v\}\}$ in the sum on $\Pi^2(u, u', v)$ in (101) that is

$$\int_{[0, T]^4} \mathbb{1}_{|u-v| \leq \delta} \mathbb{1}_{|u'-v'| \leq \delta} \mathbb{1}_{\Delta_4} k_1(dv')k_3(du, du', dv)$$

does not depend on the order of the variables in the measure $k_1(dv')k_3(du, du', dv)$. In particular, the value of the integral remains the same with what we get by using the partitions $\{\{u\}, \{u', v, v'\}\}$, $\{\{u'\}, \{u, v, v'\}\}$ and $\{\{v\}, \{u, u', v'\}\}$ in the sum on $\Pi^1(u, u', v, v')$ in (101). The same happens with:

- the values of the integral with the partition $\{\{u\}, \{u', v\}\}$ in the sum on $\Pi^2(u, u', v)$ and the partitions $\{\{u'\}, \{v\}, \{u, v'\}\}$, $\{\{u'\}, \{u\}, \{v, v'\}\}$ in the sum on $\Pi^1(u, u', v, v')$ in (101);
- the values of the integral with the partition $\{\{u'\}, \{u, v\}\}$ in the sum on $\Pi^2(u, u', v)$ and the partition $\{\{u\}, \{v\}, \{u', v'\}\}$ in the sum on $\Pi^1(u, u', v, v')$ in (101);
- the values of the integrals with the partitions $\{\{u, u'\}, \{v, v'\}\}$ and $\{\{u, v'\}, \{u', v\}\}$ in the sum on $\Pi^1(u, u', v, v')$ in (101).

We then deduce from (101) that

$$\begin{aligned} &\int_{[0, T]^4} \mathbb{1}_{|u-v| \leq \delta} \mathbb{1}_{|u'-v'| \leq \delta} \mathcal{Q}(du, du', dv, dv') \\ &= \int_{[0, T]^4} \mathbb{1}_{|u-v| \leq \delta} \mathbb{1}_{|u'-v'| \leq \delta} \mathbb{1}_{\Delta_4} \left\{ k_4(du, du', dv, dv') + 2k_1(du)k_3(du', dv, dv') \right. \\ &\quad \left. + 3k_2(du, du')k_2(dv, dv') + 2k_1(du)k_1(du')k_2(dv, dv') \right\}, \end{aligned}$$

that ends the proof. \square

From this Lemma and using the value of the variance \mathbb{V}_0 of $f_\varphi(X_1, X_2)$ under the mean-field limit given in (42), we deduce that

$$\begin{aligned} & \mathbb{V}[f_\varphi(X_1, X_2)] - \mathbb{V}_0 \\ &= \frac{1}{M^2} \int_{[0, T]^2} \mathbb{1}_{|u-v| \leq \delta} \mathbb{1}_{u \neq v} k_2(du, dv) \\ &+ \frac{1}{M^3} \int_{[0, T]^3} \mathbb{1}_{|u-v| \leq \delta} \mathbb{1}_{|u-v'| \leq \delta} \left(\mathcal{G}(du, dv, dv') - 2k_1(du)k_1(dv)k_1(dv') \right) \\ &+ \frac{1}{M^4} \int_{[0, T]^4} \mathbb{1}_{|u-v| \leq \delta} \mathbb{1}_{|u'-v'| \leq \delta} \mathcal{Q}(du, du', dv, dv'). \end{aligned}$$

This difference is positive. Using (41) in Lemma 9, it satisfies the following bound for a certain constant $C > 0$ (that can change from a line to another)

$$\begin{aligned} & \mathbb{V}[f_\varphi(X_1, X_2)] - \mathbb{V}_0 \\ & \leq C \left\{ \frac{a\nu}{M(1-\frac{a}{b})^2} \int_{[0, T]^2} \mathbb{1}_{|u-v| \leq \delta} dudv \right. \\ & + \frac{a\nu}{M(1-\frac{a}{b})^3} \int_{[0, T]^3} \mathbb{1}_{|u-v| \leq \delta} \mathbb{1}_{|u-v'| \leq \delta} \\ & \quad \times \left[\frac{a}{M} e^{-(b-a)[\max\{u, v, v'\} - \min\{u, v, v'\}]} + \nu \left(e^{-(b-a)|v-v'|} + e^{-(b-a)|u-v'|} \right) \right] dudvdv' \\ & + \frac{a\nu}{M(1-\frac{a}{b})^4} \int_{[0, T]^4} \mathbb{1}_{|u-v| \leq \delta} \mathbb{1}_{|u'-v'| \leq \delta} \left[\frac{a^2}{M^2} e^{-(b-a)[\max\{u, u', v, v'\} - \min\{u, u', v, v'\}]} \right. \\ & \quad \left. + \frac{a\nu}{M} \left(e^{-(b-a)[\max\{u', v, v'\} - \min\{u', v, v'\}]} + e^{-(b-a)(|u-u'| + |v-v'|)} \right) + \nu^2 e^{-(b-a)|v-v'|} \right] \\ & \quad \left. dudud'vdv' \right\}, \end{aligned}$$

and then

$$\begin{aligned} \mathbb{V}[f_\varphi(X_1, X_2)] - \mathbb{V}_0 & \leq C \left\{ \frac{a\nu\delta T}{M(1-\frac{a}{b})^2} + \frac{a\nu(\nu + \frac{a}{M})}{M(1-\frac{a}{b})^3} \underbrace{\int_{[0, T]^3} \mathbb{1}_{|u-v| \leq \delta} \mathbb{1}_{|u-v'| \leq \delta} dudvdv'}_{\leq 4\delta^2 T} \right. \\ & \quad \left. + \frac{a\nu(\nu + \frac{a}{M})^2}{M(1-\frac{a}{b})^4} \int_{[0, T]^4} e^{-(b-a)|v-v'|} \mathbb{1}_{|u-v| \leq \delta} \mathbb{1}_{|u'-v'| \leq \delta} dudud'vdv' \right\} \\ & \leq C \left\{ \frac{a\nu\delta T}{M(1-\frac{a}{b})^2} + \delta^2 T \frac{a\nu(\nu + \frac{a}{M})}{M(1-\frac{a}{b})^3} + \delta^2 \frac{a\nu(\nu + \frac{a}{M})^2}{M(1-\frac{a}{b})^4} \underbrace{\int_{[0, T]^2} e^{-(b-a)|v-v'|} dvdv'}_{\leq 2T/(b-a)} \right\} \\ & \leq C \frac{aT}{M} \frac{\nu\delta}{(1-\frac{a}{b})^2} \left\{ 1 + \frac{(\nu + \frac{a}{M})\delta}{1-\frac{a}{b}} \left[1 + \frac{b(\nu + \frac{a}{M})}{(b-a)^2} \right] \right\}. \end{aligned}$$

Replacing a/b by ℓ this implies (44) and ends the proof.

C.5. Proof of Theorem 4. Let us first notice that the variance \mathbb{V}_0 of $f_\varphi(X_1, X_2)$ under the mean-field limit (see (42)) satisfies the bound

$$\mathbb{V}_0 \leq \frac{4\nu^2\delta T}{(1-\ell)^2} \left(1 + \frac{2\nu\delta}{1-\ell} \right).$$

We deduce from (43) and (44) that there exists a constant $C > 0$ such that the criterion (5) in Lemma 1 which ensures the control of the type II error of the permutation test $\mathbb{1}_{\mathbf{C}_n(\mathbb{X}_n) > q_{1-\alpha}(\mathbb{X}_n)}$ by β holds if

$$\Delta_\varphi \geq \frac{C}{\sqrt{n\alpha\beta}} \sqrt{\frac{\nu^2\delta T}{(1-\ell)^2} \left(1 + \frac{2\nu\delta}{1-\ell}\right) + \frac{aT}{M} \frac{\nu\delta}{(1-\ell)^2} \left\{1 + \frac{(\nu + \frac{a}{M})\delta}{1-\ell} \left[1 + \frac{\nu + \frac{a}{M}}{b(1-\ell)^2}\right]\right\}}.$$

For the sake of simplicity, we denote by \mathcal{K} the term in the square root. Then we obtain

$$\begin{aligned} \mathcal{K} &= \frac{\nu\delta T}{(1-\ell)^2} \left\{ \nu \left(1 + \frac{2\nu\delta}{1-\ell}\right) + \frac{a}{M} \left\{1 + \frac{(\nu + \frac{a}{M})\delta}{1-\ell} \left[1 + \frac{\nu + \frac{a}{M}}{b(1-\ell)^2}\right]\right\} \right\} \\ &\leq 2 \frac{\nu(\nu + \frac{a}{M})\delta T}{(1-\ell)^2} \left\{1 + \frac{(\nu + \frac{a}{M})\delta}{1-\ell} + \frac{a(\nu + \frac{a}{M})\delta}{bM(1-\ell)^3}\right\} \\ &= 2 \frac{\nu(\nu + \frac{a}{M})\delta T}{(1-\ell)^2} \left\{1 + \frac{(\nu + \frac{a}{M})\delta}{1-\ell} \left[1 + \frac{\ell}{M(1-\ell)^2}\right]\right\}. \end{aligned}$$

Using that $\nu \leq \nu + a/M$, this implies the condition (13).

Finally, to prove (14), recall that (47) gives the lower bound

$$\Delta_\varphi \geq \frac{\nu\delta\ell}{M(1-\ell)^3} \frac{b(1-\ell)T}{4} \frac{1 - e^{-b(1-\ell)\delta}}{b(1-\ell)\delta} = \frac{1}{4} \frac{\nu\ell T}{M(1-\ell)^3} (1 - e^{-b(1-\ell)\delta}).$$

Therefore we only need n large enough such that the right hand side above is larger than the right hand side of (13), implying the bound given in (14). This concludes the proof.

C.6. Remaining proofs. We now give the proof of the fact that the process Z^2 introduced in Model 1 is still a Poisson process of intensity η^1 .

Proposition 19. *Let Z^1 be a Poisson process on \mathbb{R} with intensity η^1 . Denote by $\dots < u_{k-1} < u_k < u_{k+1} < \dots$ its points. Let moreover $(\xi_i)_{i \in \mathbb{Z}}$ be an i.i.d. sequence of real valued random variables, independent of Z . Then the point process Z^2 , having the points $u_k + \xi_k, k \in \mathbb{Z}$, is still a Poisson process with intensity η^1 .*

Proof. By construction, the intensity measure $\hat{\mu}$ of Z^2 is given by $\hat{\mu}(A) = \eta^1 \int_{\mathbb{R}} \mathbb{P}(t + \xi \in A) dt$, for any Borel subset A of \mathbb{R} . Using the change of variables $s = \xi + t$, we have

$$\int_{\mathbb{R}} \mathbb{P}(t + \xi \in A) dt = \mathbb{E} \int_{\mathbb{R}} \mathbb{1}_A(\xi + t) dt = \mathbb{E} \int_{\mathbb{R}} \mathbb{1}_A(s) ds = \int_A ds,$$

which implies that the intensity measure of Z^2 is still given by $\eta^1 dt$. A classical result (see e.g. Theorem II.4.5 of [24]) implies then that Z^2 is a Poisson process of intensity η^1 . \square

We close this section with an

Idea of the proof of Lemma 5. Let us start by verifying that the stochastic integral in (20) makes sense. We construct it by introducing an order between the instants s_1, \dots, s_l . Using the partition of $\Delta_l = \cup_{\pi \in \mathbb{S}_l} \Delta_l^\pi$, where for any $\pi \in \mathbb{S}_l$ we set $\Delta_l^\pi = \{(s_1, \dots, s_l) \in \mathbb{R}^l : s_{\pi(i)} < s_{\pi(j)} \text{ for any } i < j\}$, we have

$$\begin{aligned} &\int_{\Delta_l} F(s_1, \dots, s_l, \xi_{N_{s_1}}, \dots, \xi_{N_{s_l}}) dN_{s_1} \cdots dN_{s_l} \\ &= \sum_{\pi \in \mathbb{S}_l} \int_{\Delta_l^\pi} F(s_1, \dots, s_l, \xi_{N_{s_1}}, \dots, \xi_{N_{s_l}}) dN_{s_1} \cdots dN_{s_l} \\ &= \sum_{\pi \in \mathbb{S}_l} \int_{\mathbb{R}} \left\{ \int_{-\infty}^{s_{\pi(l)}^-} \cdots \left\{ \int_{-\infty}^{s_{\pi(2)}^-} F(s_1, \dots, s_l, \xi_{N_{s_1}}, \dots, \xi_{N_{s_l}}) dN_{s_{\pi(1)}} \right\} \cdots dN_{s_{\pi(l-1)}} \right\} dN_{s_{\pi(l)}}. \end{aligned}$$

Since the sequence $\{\xi_i, i \in \mathbb{Z}\}$ is independent of N , the most internal integral satisfies

$$\begin{aligned} & \int_{-\infty}^{s_{\pi(2)}^-} F(s_1, \dots, s_l, \xi_{N_{s_1}}, \dots, \xi_{N_{s_l}}) dN_{s_{\pi(1)}} \\ &= \left(\int_{-\infty}^{s_{\pi(2)}^-} F(s_1, \dots, s_l, \xi_{N_{s_1}}, x_2, \dots, x_l) dN_{s_{\pi(1)}} \right) \Big|_{x_2=\xi_{N_{s_2}}, \dots, x_l=\xi_{N_{s_l}}} . \end{aligned}$$

We compute the less internal integrals thanks to the same argument, and the expectation given by (20) directly follows by conditioning. \square

Acknowledgments. This work was supported by the French government, through CNRS, the UCA *Jedi* and 3IA Côte d’Azur Investissements d’Avenir managed by the National Research Agency (ANR-15 IDEX-01 and ANR-19-P3IA-0002), directly by the ANR project ChaMaNe (ANR-19-CE40-0024-02) and by the interdisciplinary Institute for Modeling in Neuroscience and Cognition (NeuroMod).