



HAL
open science

Bootstrap and permutation tests of independence for point processes

Mélanide Albert, Yann Bouret, Magalie Fromont, Patricia Reynaud-Bouret

► **To cite this version:**

Mélanide Albert, Yann Bouret, Magalie Fromont, Patricia Reynaud-Bouret. Bootstrap and permutation tests of independence for point processes. 2014. hal-01001984v2

HAL Id: hal-01001984

<https://hal.science/hal-01001984v2>

Preprint submitted on 4 Aug 2014 (v2), last revised 7 Dec 2015 (v5)

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.

BOOTSTRAP AND PERMUTATION TESTS OF INDEPENDENCE FOR POINT PROCESSES

BY MÉLISANDE ALBERT^{*}, YANN BOURET[†], MAGALIE
FROMONT[‡], AND PATRICIA REYNAUD-BOURET^{*}

^{*} *Univ. Nice Sophia Antipolis, CNRS, LJAD, UMR 7351, 06100 Nice, France.* [†] *Univ. Nice Sophia Antipolis, CNRS, LPMC, UMR 7336, 06100 Nice, France.* [‡] *Univ. Européenne de Bretagne, CNRS, IRMAR, UMR 6625, 35043 Rennes Cedex, France.*

Abstract. Motivated by a neuroscience question about synchrony detection in spike train analysis, we deal with the independence testing problem for point processes. We introduce non-parametric test statistics, which are rescaled general U -statistics, whose corresponding critical values are constructed from bootstrap and randomisation/permutation approaches, making as few assumptions as possible on the underlying distribution of the point processes. We derive general consistency results for the bootstrap and for the permutation w.r.t. to Wasserstein's metric, which induce weak convergence as well as convergence of second order moments. The obtained bootstrap or permutation independence tests are thus proved to be asymptotically of the prescribed size, and to be consistent against any reasonable alternative, permutation independence tests having the further advantage to be exactly (that is non-asymptotically) of the prescribed level, even when Monte Carlo methods are used to approximate the randomised quantiles. A simulation study is performed to illustrate the derived theoretical results, and to compare the performance of our new tests with existing ones in the neuroscientific literature.

1. Introduction. Inspired by neuroscience problems, the present work is devoted to independence tests for point processes. The question of testing whether two random variables are independent is of course largely encountered in the statistical literature, as it is one of the central goals of data analysis. From the historical Pearson's (see [31, 32]) chi-square test of independence to the recent test of [19] using kernel methods in the spirit of statistical learning, many non-parametric independence tests have been developed for real valued random variables or even random vectors. Among them, of particular interest are the tests based on the randomisation/per-

MSC 2010 subject classifications: Primary 62M07, 62F40, 62E20, 60G55, 60F05; secondary 62P10

Keywords and phrases: Independence test, U -statistics, point processes, bootstrap, randomisation, permutation, neuroscience, spike train analysis

mutation principle introduced by Fisher [15], and covered thereafter in the series of papers by Pitman [36, 37], Scheffe [46], Hoeffding [24] for instance, or bootstrap approaches derived from Efron's [14] "naive" one. Two families of such permutation or bootstrap-based independence tests may be distinguished at least: the whole family of rank tests including the tests of Hotelling and Pabst [27], Kendall [28], Wolfowitz [53] or Hoeffding [23] on the one hand, the family of Kolmogorov-Smirnov type tests, like Blum, Kiefer, and Rosenblatt's [5], Romano's [42] or Van der Vaart and Wellner's [51] ones on the other hand. These tests are purely non-parametric that is they are completely free of the underlying distributions of the observed random variables or vectors. They are proved to achieve asymptotically the right desired size: the probability, under independence, that the independence hypothesis is rejected tends to a prescribed α in $]0, 1[$, as the size of the original samples of observations grows to infinity. Moreover, the tests based on permutation are known to be exactly (non-asymptotically meaning) of the desired level, that is the probability, under independence, that the independence hypothesis is rejected is smaller than the prescribed α , for any sample size. Some of these tests are proved to be consistent against many alternatives, such as Hoeffding's [23] one and the family of Kolmogorov-Smirnov type tests.

Detecting dependence is also a fundamental old point in the neuroscientific literature (see e.g. [18]). The neuroscience problem we were initially interested in consists in detecting interactions between occurrences of action potentials on two different neurons simultaneously recorded on n independent trials, as described in [20]. Each recorded set of time occurrences of action potentials for each neuron is usually referred to as a spike train, the spikes being the time occurrences themselves. It is commonly accepted that these spikes are one of the main components of the brain activity (see [48]). Therefore, when observing two spike trains coming from two different neurons, one of the main elementary problem is to assess whether these two spike trains are independent or not. Unfortunately, even if the real recordings of spike trains are discretized in time, due to the record resolution, and thus belong to finite dimensional spaces, the dimension of these spaces is so huge (from ten thousand up to one million) that it is neither realistic nor reasonable to model them by finite dimensional variables, and to apply usual independence tests. Several methods, such as the classical Unitary Events method (see [20] and the references therein), consists in binning the spike trains at first in order to deal with vectorial data with reduced dimension. But it has been shown that these dimension reduction methods involve an information loss of more than 60% in some cases, making this kind of preprocessing quite proscribed despite its simplicity of use.

Modelling the recordings of spike trains by point processes that are only almost surely finite without any a priori bound on the total number of points, and using, constructing if needed, independence tests specifically dedicated to such point processes then appear as realistic and reasonable solutions. Asymptotic independence tests have been introduced in [49] for point processes, but limited to homogeneous Poisson processes. Such a Poisson assumption is necessarily restrictive and even possibly inappropriate, considering spike train analysis (see [38]). Since the existence of any precise distribution for the point processes modelling spike train data is subject to broad debate, to construct model free independence tests for point processes is of utmost importance. To this end, bootstrap methods have been used in [35, 34] for binned data with relatively small dimension, without any proper mathematical justification.

Another field of applications, where detecting dependence between point processes is crucial, is, for instance, genomics, where point processes, such as Hawkes processes, may represent positions of motifs or Transcription Regulatory Elements on the DNA strand as described in [9].

We here propose to construct new non-parametric tests of independence between two point processes based on the observation of n independent copies, with as few assumptions as possible on the underlying distributions. Our test statistics are based on general U -statistics. The corresponding critical values are obtained from bootstrap or permutation approaches. Then the obtained tests are proved to be asymptotically (with respect to the sample size) of the desired size and consistent against many alternatives. They also have good performance in practice when the sample size is moderate to small, as is often the case in neuroscience or genomics for biological or economical reasons.

A huge number of papers deal with the bootstrap or permutation of U -statistics of i.i.d. real valued random variables or vectors. Among them, we can cite for instance [3, 7, 1, 12], or [29] devoted to the bootstrap in a general framework, [26] or [10] devoted to the permutation in various testing frameworks, and of course, the papers cited above (see [42, 51]), which are devoted to Kolmogorov-Smirnov type tests, and based on particular U -statistics in an independence testing framework.

To our knowledge, there is no previous work on the bootstrap or permutation of general U -statistics for i.i.d. pairs of point processes, as considered in the present work. The main difficulties lie here in the nature of the mathematical objects we handle, that is point processes and their associated point measures which are random measures on the one hand, in the general nature of the results we aim at on the other hand. The proofs of our results,

although inspired by Romano’s [41, 42] work and Hoeffding’s [24] precursor results on the permutation, are therefore more technical and complex on many aspects detailed in the sequel, all the more as we obtain the convergence of the conditional distribution of the permuted test statistics (see Theorem 4.1) under independence, but also under dependence. Such a result is, as far as we know, new even for more classical settings than point processes, thus partially solving a problem stated as open question in [51].

This paper is organized as follows.

We first present in Section 2 the testing problem, and introduce the main notation. Starting from existing works in neuroscience, we introduce our test statistics, based on general kernel-based U -statistics.

Section 3 is devoted to our bootstrap approach, from its description to the theoretical asymptotic properties of the corresponding independence tests, through general results about the consistency of the bootstrap for the considered U -statistics, expressed in terms of Wasserstein’s metric as in [3]. We thus state that our tests are asymptotically of the desired size, and that they are consistent against every reasonable alternative.

Section 4 is devoted to the permutation approach. After a description of this approach, we give general results about its consistency when the kernel of the U -statistic has a particular common form. These results are still expressed in terms of Wasserstein’s metric. Therefore the corresponding permutation independence tests satisfy the same asymptotic properties as the bootstrap ones, with the further advantage of being exactly (that is non-asymptotically) of the desired level, even when a Monte Carlo method is used to approximate the randomised quantiles.

As a comparison of the performance of our tests with existing ones in neuroscience, especially when the sample sizes are moderate or even small, a simulation study is presented in Section 5.

Finally notice that key proofs or sketches of proofs are postponed to the last section, while the complete proofs are given in a supplementary material.

2. From neuroscience interpretations to general test statistics.

2.1. *The testing problem.* We consider in this article finite point processes, defined on a probability space $(\Omega, \mathcal{A}, \mathbb{P})$ and observed on $[0, 1]$, that are random point processes on $[0, 1]$, for which the total number of points is almost surely finite (see [11] for instance). Typically, in a neuroscience framework, such finite point processes may represent spike trains recorded on a given finite interval of time, and rescaled so that their values may be assumed to belong to $[0, 1]$. The set \mathcal{X} of all their possible values is equipped with a

metric $d_{\mathcal{X}}$ that we introduce in (3.2). This metric, issued from the Skorohod topology, makes \mathcal{X} separable and allows to define accordingly borelian sets on \mathcal{X} and by extension on \mathcal{X}^2 through the product metric.

The point measure dN_x associated to an element x of \mathcal{X} is defined for all measurable real-valued function f by $\int_{[0,1]} f(u)dN_x(u) = \sum_{u \in x} f(u)$. In particular, the total number of points of x , denoted by $\#x$, is equal to $\int_{[0,1]} dN_x(u)$. Moreover, for a point process X defined on $(\Omega, \mathcal{A}, \mathbb{P})$ and observed on $[0, 1]$, $\int f(u)dN_X(u)$ becomes a real random variable, defined on the same probability space $(\Omega, \mathcal{A}, \mathbb{P})$.

A pair $X = (X^1, X^2)$ of finite point processes defined on $(\Omega, \mathcal{A}, \mathbb{P})$, observed on $[0, 1]$, has joint distribution P , with marginals P^1 and P^2 if $P(\mathcal{B}) = \mathbb{P}(X \in \mathcal{B})$, $P^1(\mathcal{B}^1) = \mathbb{P}(X^1 \in \mathcal{B}^1)$, and $P^2(\mathcal{B}^2) = \mathbb{P}(X^2 \in \mathcal{B}^2)$, for every borelian set \mathcal{B} of \mathcal{X}^2 , and all borelian sets $\mathcal{B}^1, \mathcal{B}^2$ of \mathcal{X} .

Given the observation of an i.i.d. sample $\mathbb{X}_n = (X_1, \dots, X_n)$ from the same distribution P as X , with $X_i = (X_i^1, X_i^2)$ for every $i = 1 \dots n$, we aim at testing (H_0) X^1 and X^2 are independent against (H_1) X^1 and X^2 are not independent, which can also be written as

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

2.2. *Independence test based on coincidences in neuroscience.* Considering that the i.i.d. sample $\mathbb{X}_n = (X_1, \dots, X_n)$ models pairs of rescaled spike trains issued from two distinct and simultaneously recorded neurons during n trials, the main dependence feature that needs to be detected between both neurons corresponds to synchronization in time, referred to as coincidences [20]. More precisely, neuroscientists expect to detect if such coincidences occur significantly, that is more than what may be due to chance. They speak in this case of a detected synchrony.

In [49], the notion of coincidence count between two point processes X^1 and X^2 with delay δ ($\delta > 0$) is defined by

$$(2.1) \quad \varphi_{\delta}^{coinc}(X^1, X^2) = \int_{[0,1]^2} \text{coinc}_{\delta}(u, v) dN_{X^1}(u) dN_{X^2}(v) = \sum_{u \in X^1, v \in X^2} \mathbb{1}_{\{|u-v| \leq \delta\}},$$

with $\text{coinc}_{\delta}(u, v) = \mathbb{1}_{\{|u-v| \leq \delta\}}$.

Under the assumption that both X^1 and X^2 are homogeneous Poisson processes, the independence test of [49] rejects (H_0) when a test statistic based on $\sum_{i=1}^n \varphi_{\delta}^{coinc}(X_i^1, X_i^2)$ is larger than a given critical value. This critical value is deduced from the asymptotic Gaussian distribution of the test statistic under (H_0) . The test is proved to be asymptotically of the desired size, but only under the homogeneous Poisson processes assumption.

However, it is now well-known that this assumption, as well as many other model assumptions, fails to be satisfied in practice for spike trains [38, 39].

2.3. General non-degenerate U -statistics as independence test statistics.

In the parametric homogeneous Poisson framework of [49], the expectation of $\varphi_\delta^{coinc}(X_i^1, X_i^2)$ has a simple expression as a function of δ and the intensities λ_1 and λ_2 of X^1 and X^2 . Since λ_1 and λ_2 can be easily estimated, an estimator of this expectation can thus be obtained using the plug-in principle, and subtracted from $\varphi_\delta^{coinc}(X_i^1, X_i^2)$ to lead to a test statistic with a centred asymptotic distribution under (H_0) .

In the present non-parametric framework where we want to make as few assumptions as possible on the point processes X^1 and X^2 , such a centring plug-in tool is not available, and we propose to use instead a self-centring trick, which amounts to considering

$$\sum_{i \neq i' \in \{1, \dots, n\}} (\varphi_\delta^{coinc}(X_i^1, X_i^2) - \varphi_\delta^{coinc}(X_i^1, X_{i'}^2)).$$

Furthermore, it is clear that the function φ_δ^{coinc} used in [49] suits the dependence feature the neuroscientists expect to detect in a spike train analysis. However, it is not necessarily the best choice for other kinds of dependence features to be detected in a general point processes analysis. So we introduce a more general interaction function defined by

$$(2.2) \quad \varphi_\delta^w(X^1, X^2) = \int_{[0,1]^2} w_\delta(u, v) dN_{X^1}(u) dN_{X^2}(v),$$

where w_δ is an integrable function possibly depending on a real-valued vector δ of parameters. Of course, the choice $w_\delta = coinc_\delta$ with $\delta > 0$ allows to get back the coincidence count function defined in (2.1), while the choice $w_{(j,k)}(u, v) = \psi_{j,k}(v - u)$, where $\psi_{j,k}$ is a rescaled (j) and translated (k) Haar mother wavelet, leads to the interaction function used in [45] for an interaction test in a specific Poisson framework. Even more generally, we could choose any integrable function φ , whose aim would be to detect particular dependence features on the pair (X^1, X^2) for instance, so we introduce

$$(2.3) \quad T_{n,\varphi} = \sum_{i \neq i' \in \{1, \dots, n\}} (\varphi(X_i^1, X_i^2) - \varphi(X_i^1, X_{i'}^2)).$$

Since $T_{n,\varphi}/(n(n-1))$ is an unbiased estimator of

$$\int \int \varphi(x^1, x^2) (dP(x^1, x^2) - dP^1(x^1) dP^2(x^2)),$$

it may be a reasonable independence test statistic, with expectation equal to 0 under (H_0) , without any assumption on the underlying point processes.

If X^1 and X^2 were finite dimensional variables with continuous distributions w.r.t. the Lebesgue measure, this test statistic would be closely related to generalized Kolmogorov-Smirnov tests of independence. For instance, the test statistics of Blum, Kiefer, and Rosenblatt [5], Romano [42], Van der Vaart and Wellner in [51] are equal to $n^{-3/2} \sup_{v^1 \in \mathcal{V}^1, v^2 \in \mathcal{V}^2} |T_{n, \varphi_{(v^1, v^2)}}|$, where, respectively:

- $\mathcal{V}^1 = \mathcal{V}^2 = \mathbb{R}$ and $\varphi_{(v^1, v^2)}(x^1, x^2) = \mathbb{1}_{\{-\infty, v^1\}}(x^1) \mathbb{1}_{\{-\infty, v^2\}}(x^2)$,
- \mathcal{V}^1 and \mathcal{V}^2 are countable V.-C. classes of subsets of \mathbb{R}^d ,
and $\varphi_{(v^1, v^2)}(x^1, x^2) = \mathbb{1}_{\{v^1\}}(x^1) \mathbb{1}_{\{v^2\}}(x^2)$,
- \mathcal{V}^1 and \mathcal{V}^2 are well-chosen classes of real-valued functions,
and $\varphi_{(v^1, v^2)}(x^1, x^2) = v^1(x^1)v^2(x^2)$.

Thus, up to our knowledge, the existing test statistics are based on functions φ of product type. However, as seen in Section 2.2, when dealing with point processes, natural functions φ are not of this type, as for instance φ_δ^{coinc} . This is why we consider more general test statistics.

By setting

$$(2.4) \quad h_\varphi(x, y) = \frac{1}{2} (\varphi(x^1, x^2) + \varphi(y^1, y^2) - \varphi(x^1, y^2) - \varphi(y^1, x^2)),$$

for every $x = (x^1, x^2)$, $y = (y^1, y^2)$ in \mathcal{X}^2 , $T_{n, \varphi}$ can also be written as $T_{n, \varphi} = \sum_{i \neq i' \in \{1, \dots, n\}} h_\varphi(X_i, X_{i'})$. Hence $T_{n, \varphi}/(n(n-1))$ is a classical U -statistic with a symmetric kernel h_φ , as introduced by Hoeffding [22].

Thus, we consider independence test statistics which are based on U -statistics of the form

$$(2.5) \quad U_{n, h}(\mathbb{X}_n) = \frac{1}{n(n-1)} \sum_{i \neq i' \in \{1, \dots, n\}} h(X_i, X_{i'}),$$

for some kernel h of the form h_φ given in (2.4), or even for some more general symmetric kernel $h : (\mathcal{X}^2)^2 \rightarrow \mathbb{R}$ such that $\mathbb{E}[h(X, X')] = 0$ for every X, X' i.i.d. with distribution P on \mathcal{X}^2 satisfying $(H_0) P = P^1 \otimes P^2$. Assuming that h is symmetric is however not restrictive since for any kernel h , $2U_{n, h} = U_{n, h_{sym}}$, where $h_{sym}(x, y) = h(x, y) + h(y, x)$ is symmetric.

Following the works of Romano [42] or Van der Vaart and Wellner [51], the tests we propose here are based on bootstrap and permutation approaches for the above general U -statistics. However, we focus in this article on non-degenerate U -statistics, discussing below what degeneracy means about the underlying processes.

2.4. *Non-degeneracy of the U -statistics under (H_0) .* Since under (H_0) $U_{n,h}(\mathbb{X}_n)$ is assumed to have zero mean, it is degenerate if and only if for X with distribution $P^1 \otimes P^2$ and $P^1 \otimes P^2$ -almost every x in \mathcal{X}^2 , $\mathbb{E}[h(x, X)] = 0$.

In the particular case where $h = h_\varphi$, with h_φ defined by (2.4), this amounts to state that for $X = (X^1, X^2)$ with distribution $P^1 \otimes P^2$ and for $P^1 \otimes P^2$ -almost every $x = (x^1, x^2)$,

$$\varphi(x^1, x^2) + \mathbb{E}[\varphi(X^1, X^2)] - \mathbb{E}[\varphi(x^1, X^2)] - \mathbb{E}[\varphi(X^1, x^2)] = 0.$$

This condition's fulfillment would however imply a very particular link between φ and the distribution of the bivariate point process X , which is unknown. When $\varphi = \varphi_\delta^w$ is given by (2.2), the condition can be rewritten as follows: for $P^1 \otimes P^2$ -almost every $x = (x^1, x^2)$,

$$(2.6) \quad \int w_\delta(u, v) dN_{x^1}(u) dN_{x^2}(v) + \mathbb{E} \left[\int w_\delta(u, v) dN_{X^1}(u) dN_{X^2}(v) \right] \\ - \mathbb{E} \left[\int w_\delta(u, v) dN_{x^1}(u) dN_{X^2}(v) \right] - \mathbb{E} \left[\int w_\delta(u, v) dN_{X^1}(u) dN_{x^2}(v) \right] = 0.$$

Therefore, if for instance $P^1(\{\emptyset\}) > 0$ and $P^2(\{\emptyset\}) > 0$, which is the case for Bernoulli processes (discretized processes), Poisson processes, Hawkes processes or even all processes with conditional intensities, this implies, by taking $x^1 = x^2 = \emptyset$, that

$$(2.7) \quad \mathbb{E} \left[\int w_\delta(u, v) dN_{X^1}(u) dN_{X^2}(v) \right] = 0.$$

If w_δ is nonnegative, then finally, $\int w_\delta(u, v) dN_{X^1}(u) dN_{X^2}(v) = 0$ a.s. This would not usually happen in practice. Indeed, for instance in spike train analysis, assuming that the processes may be empty is an obvious assumption as there often exist trials (usually short) where, just by chance, no spikes have been detected. Moreover, with $w_\delta = \text{coinc}_\delta$ (see (2.1)), δ is chosen large enough so that coincidences are observed in practice. Hence (2.7) is not satisfied, and the U -statistic is non-degenerate.

However notice that when $w_\delta(u, v) = \psi_{j,k}(v - u)$, where $\delta = (j, k)$ and $\psi_{j,k}$ is a rescaled and translated Haar mother wavelet as in [45], when considering specific Poisson processes, the condition (2.7) is fulfilled, since $\int \psi_{j,k}(u) du = 0$ is required to make the test statistic in [45] a reasonable test statistic. But even in this particular case, the U -statistic is still non-degenerate since the left-hand side of (2.6) has zero mean and positive variance, as soon as both Poisson processes have non zero intensities (see the supplementary material). It is therefore legitimate to assume $U_{n,h}(\mathbb{X}_n)$ non-degenerate in practice.

3. Bootstrap tests of independence. All along this section, $(X_n)_n$ denotes a sequence of i.i.d. pairs of point processes, with $X_i = (X_i^1, X_i^2)$ of distribution P on \mathcal{X}^2 , whose marginals are respectively P^1 and P^2 on \mathcal{X} . For every $n \geq 2$, let $\mathbb{X}_n = (X_1, \dots, X_n)$ and let $U_{n,h}(\mathbb{X}_n)$ be the U -statistic given by (2.5), with a measurable symmetric kernel h defined on $\mathcal{X}^2 \times \mathcal{X}^2$.

3.1. *Description of the bootstrap approach and informal view of the results.* Since the distribution of the test statistic $U_{n,h}(\mathbb{X}_n)$ is not free from the unknown underlying marginal distributions P^1 and P^2 under the null hypothesis, we turn to a classical bootstrap approach, whose aim is to mimic the distribution of the test statistic under (H_0) , for large, but also moderate or small sample sizes. As each $X_i = (X_i^1, X_i^2)$ is $P^1 \otimes P^2$ -distributed under (H_0) , the first and second coordinates of the elements of \mathbb{X}_n are resampled according to the corresponding marginal empirical distributions P_n^j given by

$$(3.1) \quad \text{for } j = 1, 2, \quad P_n^j = \frac{1}{n} \sum_{i=1}^n \delta_{X_i^j}.$$

The bootstrap sample $\mathbb{X}_n^* = (X_{n,1}^*, \dots, X_{n,n}^*)$, with $X_{n,i}^* = (X_{n,i}^{1*}, X_{n,i}^{2*})$, is then defined as an n i.i.d. sample from the distribution $P_n^1 \otimes P_n^2$.

We prove (see Theorem 3.1) that, under appropriate assumptions, the conditional distribution of $\sqrt{n}U_{n,h}(\mathbb{X}_n^*)$ given \mathbb{X}_n is asymptotically close to the one of $\sqrt{n}U_{n,h}(\mathbb{X}_n^\perp)$, where \mathbb{X}_n^\perp is an i.i.d sample from the distribution $P^1 \otimes P^2$. Therefore, using the quantiles of the distribution of $\sqrt{n}U_{n,h}(\mathbb{X}_n^*)$ given \mathbb{X}_n (or their Monte Carlo approximation) as critical values, we can build unilateral or bilateral tests of (H_0) against (H_1) with asymptotic size α . These tests are also consistent against some alternatives such that $\int h(x, x') dP(x) dP(x') \neq 0$ (see Theorem 3.2 and Proposition 3.6).

Before stating these results, and to avoid any confusion, we introduce below some notation.

- For any functional $Z : (\mathcal{X}^2)^n \rightarrow \mathbb{R}$, $\mathcal{L}(Z, Q)$ denotes the distribution of $Z(\mathbb{X}_n)$, where \mathbb{X}_n is an i.i.d. sample from the distribution Q on \mathcal{X}^2 .
- If the distribution $Q = Q(W)$ depends on a random variable W , $\mathcal{L}(Z, Q|W)$ is the conditional distribution of $Z(\mathbb{X}_n)$, \mathbb{X}_n being an i.i.d. sample from the distribution $Q = Q(W)$ on \mathcal{X}^2 , given W .
In particular, the conditional distribution of $\sqrt{n}U_{n,h}(\mathbb{X}_n^*)$ given \mathbb{X}_n is denoted by $\mathcal{L}(\sqrt{n}U_{n,h}, P_n^1 \otimes P_n^2 | \mathbb{X}_n)$.
- " Q -a.s. in $(X_n)_n$ " at the end of a statement means that the statement only depends on the sequence $(X_n)_n$, where the X_n 's are i.i.d with distribution Q , and that there exists an event \mathcal{C} only depending on $(X_n)_n$ such that $\mathbb{P}(\mathcal{C}) = 1$, on which the statement is true.

- " $Q_n \xrightarrow[n \rightarrow +\infty]{} Q$ " means that the sequence of distributions $(Q_n)_n$ converges towards Q in the weak sense, that is for any real valued, continuous and bounded function g , $\int g(z)dQ_n(z) \rightarrow_{n \rightarrow +\infty} \int g(z)dQ(z)$.
- Finally, as we often work conditionally on \mathbb{X}_n , we denote by $\mathbb{E}^*[\cdot]$ the conditional expectation given the sample \mathbb{X}_n .

3.2. *Main assumptions.* Since the random variables we deal with are not real-valued variables but point processes, our results need some assumptions whose interpretation may be difficult in this setting. We therefore describe and discuss them in the present section.

Centring assumptions. First notice that, under (H_0) , $U_{n,h}(\mathbb{X}_n)$ has zero mean if the kernel h satisfies the following assumption:

$$(\mathcal{A}_{Cent}) \quad \left| \begin{array}{l} \text{For } X_1 \text{ and } X_2, \text{ i.i.d. with distribution } P^1 \otimes P^2 \text{ on } \mathcal{X}^2, \\ \mathbb{E}[h(X_1, X_2)] = 0. \end{array} \right.$$

An empirical and resampled version of this assumption is also stated as:

$$(\mathcal{A}_{Cent}^*) \quad \left| \begin{array}{l} \text{For } x_1 = (x_1^1, x_1^2), \dots, x_n = (x_n^1, x_n^2) \text{ in } \mathcal{X}^2, \\ \sum_{i_1, i_2, i'_1, i'_2=1}^n h\left(\left(x_{i_1}^1, x_{i_2}^2\right), \left(x_{i'_1}^1, x_{i'_2}^2\right)\right) = 0. \end{array} \right.$$

Notice that both assumptions are fulfilled when h is of the form h_φ given by (2.4), but (\mathcal{A}_{Cent}^*) does not imply that h is of the form h_φ (see the supplementary material for a counterexample).

Moment assumptions. Since the considered metric is the \mathbb{L}^2 -Wasserstein one, moment assumptions are required. In particular, the variance of $U_{n,h}(\mathbb{X}_n)$ should exist, i.e.

$$(\mathcal{A}_{Mmt}) \quad \left| \begin{array}{l} \text{For } X_1 \text{ and } X_2, \text{ i.i.d. with distribution } P \text{ on } \mathcal{X}^2, \\ \mathbb{E}[h^2(X_1, X_2)] < +\infty. \end{array} \right.$$

A resampled version of this assumption is also needed:

$$(\mathcal{A}_{Mmt}^*) \quad \left| \begin{array}{l} \text{For } X_1, X_2, X_3, X_4 \text{ i.i.d. with distribution } P \text{ on } \mathcal{X}^2, \\ \text{and for } i_1, i_2, i'_1, i'_2 \text{ in } \{1, 2, 3, 4\}, \\ \mathbb{E}\left[h^2\left(\left(X_{i_1}^1, X_{i_2}^2\right), \left(X_{i'_1}^1, X_{i'_2}^2\right)\right)\right] < +\infty. \end{array} \right.$$

Notice that when (\mathcal{A}_{Mmt}^*) is satisfied, this implies that

- (\mathcal{A}_{Mmt}) is satisfied (taking $i_1 = i_2, i'_1 = i'_2$, and $i'_1 \neq i_1$),
- for $X \sim P$, $\mathbb{E}[h^2(X, X)] < +\infty$ (taking $i_1 = i_2 = i'_1 = i'_2$),
- for X_1, X_2 i.i.d with distribution $P^1 \otimes P^2$, $\mathbb{E}[h^2(X_1, X_2)] < +\infty$ (taking i_1, i_2, i'_1, i'_2 all different).

A sufficient condition for (\mathcal{A}_{Mmt}^*) and (\mathcal{A}_{Mmt}) to be satisfied is that there exist positive constants α_1, α_2, C such that for every $x = (x^1, x^2), y = (y^1, y^2)$ in \mathcal{X}^2 , $|h(x, y)| \leq C ((\#x^1)^{\alpha_1} + (\#y^1)^{\alpha_1}) ((\#x^2)^{\alpha_2} + (\#y^2)^{\alpha_2})$, with $\mathbb{E}[(\#X^1)^{4\alpha_1}] < +\infty$ and $\mathbb{E}[(\#X^2)^{4\alpha_2}] < +\infty$.

When h is of the form h_φ given by (2.4), a possible sufficient condition is that there exist some positive constants α_1, α_2 , and C such that for every x^1, x^2 in \mathcal{X} , $|\varphi(x^1, x^2)| \leq C(\#x^1)^{\alpha_1}(\#x^2)^{\alpha_2}$, with $\mathbb{E}[(\#X^1)^{4\alpha_1}] < +\infty$ and $\mathbb{E}[(\#X^2)^{4\alpha_2}] < +\infty$.

In particular, the coincidence count function φ_δ^{coinc} defined by (2.1) satisfies: for every x^1, x^2 in \mathcal{X} , $|\varphi_\delta^{coinc}(x^1, x^2)| \leq (\#x^1)(\#x^2)$. So, (\mathcal{A}_{Mmt}^*) and (\mathcal{A}_{Mmt}) are satisfied as soon as $\mathbb{E}[(\#X^1)^4] < +\infty$ and $\mathbb{E}[(\#X^2)^4] < +\infty$.

Such moment bounds for the total number of points of the processes are in fact satisfied by many kinds of point processes: discretized point processes at resolution $0 < r < 1$ (see [49] for a definition), which have at most $1/r$ points, Poisson processes, whose total number of points obeys a Poisson distribution having exponential moments of any order, and point processes with bounded conditional intensities, which can be constructed by thinning of homogeneous Poisson processes (see [30]). Similar moment bounds can also be obtained (see [21]) for linear stationary Hawkes processes with positive interaction functions that are classical models in spike train analysis (see e.g. [33, 49]). This finally may be extended to point processes whose conditional intensities are upper bounded by intensities of linear stationary Hawkes processes with positive interaction functions, by thinning arguments. This includes more general Hawkes processes (see [6]) and in particular Hawkes processes used to model inhibition in spike train analysis (see [21, 49, 40] or [39]).

Continuity of the kernel. The set \mathcal{X} of possible values of finite point processes defined on a probability space $(\Omega, \mathcal{A}, \mathbb{P})$ and observed on $[0, 1]$, can be embedded in the space \mathcal{D} of càdlàg functions on $[0, 1]$ through the identification $I : x \in \mathcal{X} \mapsto \left(t \mapsto \int_0^1 \mathbb{1}_{\{u \leq t\}} dN_x(u) \right) \in \mathcal{D}$. Considering the uniform Skorohod topology on \mathcal{D} (see [4]), associated with the metric $d_{\mathcal{D}}$ defined by

$$d_{\mathcal{D}}(f, f') = \inf \left\{ \varepsilon > 0 ; \exists \lambda \in \Lambda, \left\{ \begin{array}{l} \sup_{t \in [0,1]} |\lambda(t) - t| \leq \varepsilon, \\ \sup_{t \in [0,1]} |f(\lambda(t)) - f'(t)| \leq \varepsilon \end{array} \right\} \right\},$$

where Λ is the set of strictly increasing, continuous mappings of $[0, 1]$ onto itself, \mathcal{X} can be endowed with the topology induced by $d_{\mathcal{X}}$ defined on \mathcal{X} by

$$(3.2) \quad d_{\mathcal{X}}(x, x') = d_{\mathcal{D}}(I(x), I(x')) \quad \text{for every } x, x' \text{ in } \mathcal{X}.$$

Since $(\mathcal{D}, d_{\mathcal{D}})$ is a separable metric space, so are $(\mathcal{X}, d_{\mathcal{X}})$, $(\mathcal{X}^2, d_{\mathcal{X}^2})$, where $d_{\mathcal{X}^2}$ is the product metric defined from $d_{\mathcal{X}}$ (see [13, p 32]), and $(\mathcal{X}^2 \times \mathcal{X}^2, d)$,

where d , the product metric defined from $d_{\mathcal{X}^2}$, is given by

$$(3.3) \quad d((x, y), (x', y')) = \sup \left\{ \sup_{j=1,2} \left\{ d_{\mathcal{X}}(x^j, x'^j) \right\}, \sup_{j=1,2} \left\{ d_{\mathcal{X}}(y^j, y'^j) \right\} \right\},$$

for every $x = (x^1, x^2)$, $y = (y^1, y^2)$, $x' = (x'^1, x'^2)$, $y' = (y'^1, y'^2)$ in \mathcal{X}^2 .

The kernel h defining the U -statistic $U_{n,h}(\mathbb{X}_n)$ in (2.5) should satisfy:

$$(\mathcal{A}_{Cont}) \quad \left| \begin{array}{l} \text{There exists a subset } \mathcal{C} \text{ of } \mathcal{X}^2 \times \mathcal{X}^2, \text{ such that} \\ \text{(i) } h \text{ is continuous on } \mathcal{C} \text{ for the topology induced by } d, \\ \text{(ii) } (P^1 \otimes P^2)^{\otimes 2}(\mathcal{C}) = 1. \end{array} \right.$$

Here are some examples for which (\mathcal{A}_{Cont}) holds.

PROPOSITION 3.1. *Let $w_\delta : [0, 1]^2 \rightarrow \mathbb{R}$ be a continuous integrable function. Then the kernel $h_{\varphi_\delta^w}$ defined on $\mathcal{X}^2 \times \mathcal{X}^2$ by (2.2) and (2.4) is continuous w.r.t. the topology induced by d , defined by (3.3).*

The above result does not apply to $h_{\varphi_\delta^{coinc}}$ but the following one holds.

PROPOSITION 3.2. *The coincidence count kernel $h_{\varphi_\delta^{coinc}}$ defined on $\mathcal{X}^2 \times \mathcal{X}^2$ by (2.1) and (2.4) is continuous w.r.t. the topology induced by d , on*

$$(3.4) \quad \mathcal{C}_\delta = \left\{ ((x^1, x^2), (y^1, y^2)) \in \mathcal{X}^2 \times \mathcal{X}^2 ; \right. \\ \left. (\{x^1\} \cup \{y^1\}) \cap (\{x^2 \pm \delta\} \cup \{y^2 \pm \delta\}) = \emptyset \right\}.$$

As suggested in [49], when dealing with discretized point processes at resolution r , the right choice for δ is $kr + r/2$ for an integer k , so $(P^1 \otimes P^2)^{\otimes 2}(\mathcal{C}_\delta) = 1$, and $h_{\varphi_\delta^{coinc}}$ satisfies (\mathcal{A}_{Cont}) . Furthermore, when dealing with point processes with conditional intensities, so that they may be constructed by thinning Poisson processes, the probability $(P^1 \otimes P^2)^{\otimes 2}$ of \mathcal{C}_δ in (3.4) is larger than $\mathbb{P}(X \cap (X' \pm \delta) = \emptyset) = 1$, when X and X' are two independent Poisson processes. So $h_{\varphi_\delta^{coinc}}$ still satisfies (\mathcal{A}_{Cont}) .

3.3. Consistency of the bootstrap approach. As in the historical paper by Bickel et Freedman [3], the closeness between $\mathcal{L}(\sqrt{n}U_{n,h}, P_n^1 \otimes P_n^2 | \mathbb{X}_n)$ and $\mathcal{L}(\sqrt{n}U_{n,h}, P^1 \otimes P^2)$, both distributions on \mathbb{R} , is here measured via the classical \mathbb{L}^2 -Wasserstein's metric d_2 (also called Mallows' metric) defined by

$$(3.5) \quad d_2^2(Q, Q') = \inf \left\{ \mathbb{E}[(Z - Z')^2], (Z, Z') \text{ with marginals } Q \text{ and } Q' \right\},$$

for all the distributions Q, Q' with finite second order moments. Recall that convergence w.r.t. d_2 is equivalent to both weak convergence and convergence of second order moments.

The validity of the bootstrap approach described above for our independence tests is mainly due to the following consistency result.

THEOREM 3.1. *For every $n \geq 2$, let P_n^j for $j = 1, 2$ be the empirical marginal distributions defined by (3.1). Then, under (\mathcal{A}_{Cent}) , (\mathcal{A}_{Cent}^*) , (\mathcal{A}_{Mmt}^*) and (\mathcal{A}_{Cont}) ,*

$$d_2(\mathcal{L}(\sqrt{n}U_{n,h}, P_n^1 \otimes P_n^2 | \mathbb{X}_n), \mathcal{L}(\sqrt{n}U_{n,h}, P^1 \otimes P^2)) \xrightarrow[n \rightarrow +\infty]{} 0, \text{ } P\text{-a.s. in } (X_n)_n.$$

The proof follows similar arguments to the ones of [3] for the bootstrap of the mean, or to [12] and [29] for the bootstrap of U -statistics. But the original data being here finite point processes instead of real-valued random variables, the main difficult point is to transpose finite point processes in the separable Skorohod metric space $(\mathcal{D}, d_{\mathcal{D}})$, where weak convergence of sample probability distributions is available (see [52]). This theorem derives in fact from the following two propositions which may be useful in various frameworks. The first one states a non-asymptotic result, while the second one gives rather natural results of convergence.

PROPOSITION 3.3. *Under (\mathcal{A}_{Cent}) , (\mathcal{A}_{Cent}^*) , (\mathcal{A}_{Mmt}^*) , with the notation of Theorem 3.1, there exists an absolute constant $C > 0$ such that for $n \geq 2$,*

$$\begin{aligned} & d_2^2(\mathcal{L}(\sqrt{n}U_{n,h}, P_n^1 \otimes P_n^2 | \mathbb{X}_n), \mathcal{L}(\sqrt{n}U_{n,h}, P^1 \otimes P^2)) \\ & \leq C \inf \left\{ \mathbb{E}^* \left[(h(Y_{n,1}^*, Y_{n,2}^*) - h(Y_1, Y_2))^2 \right], Y_{n,1}^* \sim P_n^1 \otimes P_n^2, Y_1 \sim P^1 \otimes P^2, \right. \\ & \quad \left. \text{and } (Y_{n,2}^*, Y_2) \text{ is an independent copy of } (Y_{n,1}^*, Y_1) \right\}. \end{aligned}$$

PROPOSITION 3.4. *If $\mathbb{E}[|h(X_1, X_2)|] < +\infty$, then*

$$(3.6) \quad U_{n,h}(\mathbb{X}_n) \xrightarrow[n \rightarrow +\infty]{} \mathbb{E}[h(X_1, X_2)] = \int h(x, x') dP(x) dP(x'), \text{ } P\text{-a.s. in } (X_n)_n.$$

Under (\mathcal{A}_{Mmt}^) , one moreover obtains that P -a.s. in $(X_n)_n$,*

$$\frac{1}{n^4} \sum_{i,j,k,l=1}^n h^2((X_i^1, X_j^2), (X_k^1, X_l^2)) \xrightarrow[n \rightarrow +\infty]{} \mathbb{E}[h^2((X_1^1, X_2^2), (X_3^1, X_4^2))].$$

3.4. Convergence of cumulative distribution functions (c.d.f.) and quantiles. As usual, $\mathcal{N}(m, v)$ stands for the Gaussian distribution with mean m and variance v , $\Phi_{m,v}$ for its c.d.f. and $\Phi_{m,v}^{-1}$ for its quantile function. From the results of Rubin and Vitale [44] generalising Hoeffding's [22] Central Limit Theorem for non-degenerate U -statistics when the X_i 's are random vectors, we deduce the following Central Limit Theorem for $U_{n,h}(\mathbb{X}_n)$.

PROPOSITION 3.5. *Assume here that $P = P^1 \otimes P^2$, and that h satisfies (\mathcal{A}_{Cent}) and (\mathcal{A}_{Mmt}) , and let $\sigma_{h, P^1 \otimes P^2}^2$ be defined by*

$$(3.7) \quad \sigma_{h, P^1 \otimes P^2}^2 = 4 \text{Var}(\mathbb{E}[h(X_1, X_2) | X_1]).$$

If $U_{n,h}(\mathbb{X}_n)$ is non-degenerate (i.e. $\sigma_{h, P^1 \otimes P^2}^2 > 0$), then

$$d_2\left(\mathcal{L}\left(\sqrt{n}U_{n,h}, P^1 \otimes P^2\right), \mathcal{N}\left(0, \sigma_{h, P^1 \otimes P^2}^2\right)\right) \xrightarrow{n \rightarrow +\infty} 0.$$

Comment. The above asymptotic normality result may lead to a rather simple test, of the desired asymptotic size. Let

$$(3.8) \quad S_{n,h}(\mathbb{X}_n) = \frac{\sqrt{n}U_{n,h}(\mathbb{X}_n)}{\hat{\sigma}_h},$$

with

$$\hat{\sigma}_h^2 = \frac{4}{n(n-1)(n-2)} \sum_{i,j,k \in \{1, \dots, n\}, \#\{i,j,k\}=3} h(X_i, X_j)h(X_i, X_k).$$

By Slutsky's lemma and the law of large numbers for U -statistics of order 3, under (H_0) , $S_{n,h}(\mathbb{X}_n)$ converges in distribution to $\mathcal{N}(0, 1)$. The test rejecting (H_0) when $|S_{n,h}(\mathbb{X}_n)| \geq \Phi_{0,1}^{-1}(1 - \alpha/2)$ is therefore of asymptotic size α and consistent (i.e. of asymptotic power 1) against any alternative P such that $\mathbb{E}[h(X, X')] \neq 0$, for X, X' i.i.d. with distribution P and satisfying (\mathcal{A}_{Mmt}) . Such a purely asymptotic test may of course suffer from a lack of power when the sample size n is small or even moderate, which is typically the case for the application in neuroscience described in Section 2 for biological reasons (from few tens up to few hundreds at best). Though the bootstrap approach is mainly justified by asymptotic arguments, the simulation study presented in Section 5 shows its efficiency in a non-asymptotic context, with respect to this simpler test.

As Proposition 3.5 implies that the limit distribution of $\sqrt{n}U_{n,h}(\mathbb{X}_n)$ has a continuous c.d.f., the convergence of the conditional c.d.f. or quantiles of the considered bootstrap distributions holds.

COROLLARY 3.1. *For $n \geq 2$, with the notation of Theorem 3.1, let \mathbb{X}_n^* be a bootstrap sample, i.e. an i.i.d n -sample from the distribution $P_n^1 \otimes P_n^2$. Let \mathbb{X}_n^\perp be another i.i.d. n -sample from the distribution $P^1 \otimes P^2$ on \mathcal{X}^2 . Under the assumptions of Theorem 3.1, if $U_{n,h}(\mathbb{X}_n)$ is non-degenerate under (H_0) , then, P -a.s. in $(X_n)_n$,*

$$\sup_{z \in \mathbb{R}} \left| \mathbb{P}\left(\sqrt{n}U_{n,h}(\mathbb{X}_n^*) \leq z \mid \mathbb{X}_n\right) - \mathbb{P}\left(\sqrt{n}U_{n,h}(\mathbb{X}_n^\perp) \leq z\right) \right| \xrightarrow{n \rightarrow +\infty} 0.$$

If moreover, for η in $(0, 1)$, $q_{h,\eta,n}^*(\mathbb{X}_n)$ denotes the conditional η -quantile of $\sqrt{n}U_{n,h}(\mathbb{X}_n^*)$ given \mathbb{X}_n and $q_{h,\eta,n}^\perp$ denotes the η -quantile of $\sqrt{n}U_{n,h}(\mathbb{X}_n^\perp)$,

$$(3.9) \quad |q_{h,\eta,n}^*(\mathbb{X}_n) - q_{h,\eta,n}^\perp| \xrightarrow{n \rightarrow +\infty} 0, \text{ P-a.s. in } (X_n)_n.$$

3.5. *Asymptotic properties of the bootstrap tests.* Let α in $(0, 1)$ and let us now consider Γ the family of three sequences of tests $\Delta^+ = (\Delta_{h,\alpha,n}^+)_{n \geq 2}$, $\Delta^- = (\Delta_{h,\alpha,n}^-)_{n \geq 2}$, and $\Delta^{+/-} = (\Delta_{h,\alpha,n}^{+/-})_{n \geq 2}$, with critical values q where

$$(3.10) \quad q = [(q_{h,\alpha,n}^+, q_{h,\alpha,n}^-)]_{n \geq 2}, \text{ and for all } n \geq 2,$$

$$\begin{cases} \Delta_{h,\alpha,n}^+(\mathbb{X}_n) = \mathbb{1}_{\{\sqrt{n}U_{n,h}(\mathbb{X}_n) > q_{h,\alpha,n}^+(\mathbb{X}_n)\}} & \text{(test by upper values),} \\ \Delta_{h,\alpha,n}^-(\mathbb{X}_n) = \mathbb{1}_{\{\sqrt{n}U_{n,h}(\mathbb{X}_n) < q_{h,\alpha,n}^-(\mathbb{X}_n)\}} & \text{(test by lower values),} \\ \Delta_{h,\alpha,n}^{+/-}(\mathbb{X}_n) = \max(\Delta_{h,\alpha/2,n}^+(\mathbb{X}_n), \Delta_{h,\alpha/2,n}^-(\mathbb{X}_n)) & \text{(bilateral test),} \end{cases}$$

the last equation being naturally implicitly defined by the corresponding choices in $\alpha/2$. Of course, Γ , as well as Δ^+ , Δ^- and $\Delta^{+/-}$, depend on the choice of α , the desired level in $(0, 1)$ and the kernel h that are considered as fixed from the beginning of this section. They also more importantly depend on the critical values sequence q . However to keep the notation as simple and clear as possible, the corresponding indices are omitted.

Depending on the choice of q , the classical properties that can be expected to be satisfied by Γ are (\mathcal{P}_{size}) and $(\mathcal{P}_{consist.})$ defined by

$$\begin{array}{l} (\mathcal{P}_{size}) \quad \left| \begin{array}{l} \text{Each sequence } \Delta = (\Delta_n)_{n \geq 2} \text{ in } \Gamma \text{ is asymptotically of size } \alpha, \\ \text{i.e. if } P = P^1 \otimes P^2, \mathbb{P}(\Delta_n(\mathbb{X}_n) = 1) \rightarrow_{n \rightarrow +\infty} \alpha, \end{array} \right. \\ (\mathcal{P}_{consist.}) \quad \left| \begin{array}{l} \text{Each sequence } \Delta = (\Delta_n)_{n \geq 2} \text{ in } \Gamma \text{ is consistent,} \\ \text{i.e. } \mathbb{P}(\Delta_n(\mathbb{X}_n) = 1) \rightarrow_{n \rightarrow +\infty} 1, \text{ against every } P \text{ such that} \\ - \int h(x, x') dP(x) dP(x') > 0 \text{ if } \Delta = \Delta^+, \\ - \int h(x, x') dP(x) dP(x') < 0 \text{ if } \Delta = \Delta^-, \\ - \int h(x, x') dP(x) dP(x') \neq 0 \text{ if } \Delta = \Delta^{+/-}. \end{array} \right. \end{array}$$

Following Corollary 3.1, the bootstrap tests are associated to

$$(3.11) \quad q = [(q_{h,1-\alpha,n}^*, q_{h,\alpha,n}^*)]_{n \geq 2}.$$

THEOREM 3.2. *Let Γ^* be the family of tests defined by (3.10) and (3.11). If (\mathcal{A}_{Cent}) , (\mathcal{A}_{Cent}^*) , (\mathcal{A}_{Mmt}^*) and (\mathcal{A}_{Cont}) hold and if for all $n \geq 2$, $U_{n,h}(\mathbb{X}_n)$ is non-degenerate under (H_0) , then Γ^* satisfies both (\mathcal{P}_{size}) and $(\mathcal{P}_{consist.})$.*

Comments. In the particular case where $h = h_\varphi$ defined by (2.4),

$$\int h(x, x') dP(x) dP(x') = \int \varphi(x^1, x^2) [dP(x^1, x^2) - dP^1(x^1) dP^2(x^2)].$$

This means that under the assumptions of Theorem 3.2, the bilateral test in Γ^* is consistent against any alternative such that $\int \varphi(x^1, x^2) dP(x^1, x^2)$ differs from what is expected under (H_0) , i.e. $\int \varphi(x^1, x^2) dP^1(x^1) dP^2(x^2)$.

(i) When $\varphi = \varphi_\delta^{coinc}$ defined by (2.1), the assumptions of Theorem 3.2 are fulfilled for instance if X^1 and X^2 are discretized at resolution r , with $\delta = kr + r/2$ for some integer k , or if X^1 and X^2 have bounded conditional intensities, with δ large enough so that $\varphi_\delta^{coinc}(X^1, X^2)$ is not a.s. null. Theorem 3.2 means in such cases that the corresponding bilateral test is asymptotically of power 1 for any alternative P such that,

$$\int \mathbb{1}_{\{|v-u|\leq\delta\}} \mathbb{E} [dN_{X^1}(u) dN_{X^2}(v)] \neq \int \mathbb{1}_{\{|v-u|\leq\delta\}} \mathbb{E} [dN_{X^1}(u)] \mathbb{E} [dN_{X^2}(v)].$$

Note that no δ ensuring this condition can be found if heuristically, the repartition of the delays $|v - u|$ between points of X^1 and X^2 is the same under (H_0) and under (H_1) . For neuroscientists, it means that the cross-correlogram (histogram of the delays, that is classically represented as the first description of the data) does not show different behaviors in the dependent and independent cases. This would only occur if the dependence could not be measured in terms of delay between points. Though this is quite not likely to happen in spike train analysis, the question of the choice of δ remains an open question.

(ii) When $\varphi = \varphi_\delta^w$ defined by (2.2), Theorem 3.2 means that the corresponding bilateral test is consistent against any alternative such that

$$\beta_{w,\delta} = \int w_\delta(u, v) (\mathbb{E} [dN_{X^1}(u) dN_{X^2}(v)] - \mathbb{E} [dN_{X^1}(u)] \mathbb{E} [dN_{X^2}(v)]) \neq 0.$$

Under the specific Poisson assumptions of Sansonnet and Tuleau-Malot [45], if $w_\delta(u, v) = \psi_{j,k}(v - u)$ where $\delta = (j, k)$ and $\psi_{j,k}$ is a rescaled and translated Haar mother wavelet, $\beta_{w,\delta}$ is linked to the coefficient in the Haar basis of the so-called interaction function, which measures the dependence between both processes X^1 and X^2 . Working non asymptotically, one of the main result of [45] states, after reformulation in the present setting, that if $\beta_{w,\delta}$ is larger than an explicit lower bound, then the Type II error rate is less than a prescribed β in $(0, 1)$. Theorem 3.2 thus generalizes their result to a set-up with much less restrictive assumptions on the underlying stochastic models, but in an asymptotic way.

3.6. *Bootstrap tests with Monte Carlo approximation.* The above family of tests Γ^* involves an exact computation of the conditional quantiles $q_{h,\eta,n}^*(\mathbb{X}_n)$. Though such a computation is possible, it is not often reasonable

in practice even when the sample size n is moderate ($n \geq 15$), since computing $U_{n,h}(\mathbb{X}_n^*)$ itself may be complex from an algorithmic point of view, for particular choices of h . Therefore, the conditional quantiles $q_{h,\eta,n}^*(\mathbb{X}_n)$ are approximated, as usual, by a Monte Carlo method.

More precisely, for a chosen number $B \geq 1$ of iterations for the Monte Carlo method, let $(\mathbb{X}_n^{*1}, \dots, \mathbb{X}_n^{*B})$ be B independent bootstrap samples from \mathbb{X}_n , that is B i.i.d. random variables from the distribution $(P_n^1 \otimes P_n^2)^{\otimes n}$. For each \mathbf{b} in $\{1, \dots, B\}$, define $U^{*\mathbf{b}} = U_{n,h}(\mathbb{X}_n^{*\mathbf{b}})$. The order statistic associated with (U^{*1}, \dots, U^{*B}) is denoted as usual by $(U^{*(1)}, \dots, U^{*(B)})$. Then, for any sequence $(B_n)_{n \geq 2}$ of positive integers, the Monte Carlo bootstrap tests are associated to

$$(3.12) \quad q = \left[n^{1/2} \left(U^{*([\!(1-\alpha)B_n\!])}, U^{*([\!\alpha B_n\!] + 1)} \right) \right]_{n \geq 2}.$$

PROPOSITION 3.6. *Let Γ_{MC}^* be the family of tests defined by (3.10) combined with (3.12). If $B_n \rightarrow_{n \rightarrow +\infty} +\infty$, under the same assumptions as in Theorem 3.2, Γ_{MC}^* also satisfies both (\mathcal{P}_{size}) and $(\mathcal{P}_{consist.})$.*

4. Permutation tests of independence. All along this section, $(X_n)_n$ still denotes a sequence of i.i.d. pairs of point processes, with $X_i = (X_i^1, X_i^2)$ of distribution P on \mathcal{X}^2 , whose marginals are respectively P^1 and P^2 on \mathcal{X} . For every $n \geq 2$, $\mathbb{X}_n = (X_1, \dots, X_n)$, and $U_{n,h}(\mathbb{X}_n)$ is the U -statistic given by (2.5). We focus here on the particular case where h is of the form h_φ for some integrable function φ , as defined in (2.4).

4.1. *Description of the permutation approach and overview of the results.* The permutation approach we consider consists in randomly permuting the second coordinates of the observed pairs of point processes. More precisely, if π_n is a permutation of $\{1, \dots, n\}$, one defines

$$(4.1) \quad \mathbb{X}_n^{\pi_n} = (X_1^{\pi_n}, \dots, X_n^{\pi_n}) \text{ with } X_i^{\pi_n} = \left(X_i^1, X_{\pi_n(i)}^2 \right),$$

and for Π_n a random permutation taken uniformly at random in \mathfrak{S}_n independently of \mathbb{X}_n , the corresponding permuted sample is defined by $\mathbb{X}_n^{\Pi_n}$. Let P_n^* be the conditional distribution of $\mathbb{X}_n^{\Pi_n}$ given \mathbb{X}_n . Like for the bootstrap, the idea of the permutation principle is to mimic the distribution of the test statistic assuming that (H_0) is satisfied. Theorem 4.1 and Proposition 3.5 actually imply that under appropriate assumptions, the conditional distribution of $\sqrt{n}U_{n,h_\varphi}(\mathbb{X}_n^{\Pi_n})$ given \mathbb{X}_n is asymptotically close to the distribution of $\sqrt{n}U_{n,h_\varphi}(\mathbb{X}_n^\perp)$, where \mathbb{X}_n^\perp is an i.i.d. sample from the distribution $P^1 \otimes P^2$. Thus, we can propose new permutation tests of independence of the desired

asymptotic size, using the conditional quantiles of $\sqrt{n}U_{n,h_\varphi}(\mathbb{X}_n^{\Pi_n})$ given \mathbb{X}_n (or their Monte Carlo approximation) as critical values, which are consistent against any reasonable alternative. Following the statements of our bootstrap results in Section 3, we still express the closeness in distributions between the permuted and original statistics in terms of Wasserstein's metric and this even under (H_1) , which distinguishes, up to our knowledge, our results from previous ones in the permutation independence tests scene.

In addition, the main advantage of the permutation approach lies in the resulting tests being exactly of the desired level, even when the critical values are approximated by a Monte Carlo method. Such non-asymptotic results are proved thanks to [43, Lemma 1] and the following Proposition, at the heart of the permutation principle. Note that such results have no known counterpart for the bootstrap approach, despite very few non-asymptotic results in more classical settings [2, 17, 16].

PROPOSITION 4.1. *Let π_n be a deterministic permutation of $\{1, \dots, n\}$, and Π_n be a uniformly distributed random permutation of $\{1, \dots, n\}$, independent of the observed sample $\mathbb{X}_n = (X_1, \dots, X_n)$. If $P = P^1 \otimes P^2$, that is under (H_0) , then $\mathbb{X}_n^{\pi_n}$ and $\mathbb{X}_n^{\Pi_n}$ defined in (4.1) both have the same distribution as the original sample \mathbb{X}_n .*

4.2. Consistency of the permutation approach. Since we focus in this section on U -statistics based on symmetric kernels of the form h_φ , as defined in (2.4), the centring assumption (\mathcal{A}_{Cent}) is always satisfied, and we only need here the following moment assumption:

$(\mathcal{A}_{\varphi, Mmt})$ | For X with distribution P or $P^1 \otimes P^2$ on \mathcal{X}^2 , $\mathbb{E}[\varphi^4(X^1, X^2)] < \infty$.

THEOREM 4.1. *For all $n \geq 2$, let P_n^* be the conditional distribution of the permuted sample given \mathbb{X}_n . Under $(\mathcal{A}_{\varphi, Mmt})$, with the notation of Section 3,*

$$(4.2) \quad d_2 \left(\mathcal{L} \left(\sqrt{n}U_{n,h_\varphi}, P_n^* | \mathbb{X}_n \right), \mathcal{N} \left(0, \sigma_{h_\varphi, P^1 \otimes P^2}^2 \right) \right) \xrightarrow[n \rightarrow +\infty]{\mathbb{P}} 0,$$

where $\xrightarrow{\mathbb{P}}$ stands for the usual convergence in \mathbb{P} -probability.

Comments. Unlike the bootstrap approach, the convergence for Wasserstein's metric occurs in probability and not almost surely, but no continuity assumption for the kernel h_φ is used anymore, at the only price that the moment assumption is stronger than the one used for the bootstrap. The above result is one of the newest result presented here and its scope is well beyond the only generalisation to the point processes setting. Indeed, because it holds not only under (H_0) but also under (H_1) , it goes further than

existing ones such as the ones of Romano [42]. Note that the behavior under dependence of their permuted test statistic, closely related to ours, was also left as an open question by Van der Vaart and Wellner in [51]. The proof of this result is based on an asymptotic normality result for martingale difference arrays. A sketch of proof is given in Section 6, the complete longer version being given in the supplementary material.

From Theorem 4.1, we deduce the following Corollary, which, combined with Proposition 3.5, is a key point to prove that the conditional quantiles of $\mathcal{L}(\sqrt{n}U_{n,h_\varphi}, P_n^* | \mathbb{X}_n)$ can be used as critical values.

COROLLARY 4.1. *For every $n \geq 2$ and η in $(0, 1)$, let $q_{\varphi,\eta,n}^*(\mathbb{X}_n)$ denote the η -quantile of $\mathcal{L}(\sqrt{n}U_{n,h_\varphi}, P_n^* | \mathbb{X}_n)$, with the notation of Theorem 4.1. Then, under the assumptions of Theorem 4.1,*

$$q_{\varphi,\alpha,n}^*(\mathbb{X}_n) \xrightarrow[n \rightarrow +\infty]{\mathbb{P}} \Phi_{0, \sigma_{h_\varphi, P^1 \otimes P^2}^2}^{-1}(\eta).$$

4.3. Asymptotic properties of the permutation tests. With the notation of Corollary 4.1 and of (3.10), the permutation tests are associated to

$$(4.3) \quad q = [(q_{\varphi,1-\alpha,n}^*(\mathbb{X}_n), q_{\varphi,\alpha,n}^*(\mathbb{X}_n))]_{n \geq 2}.$$

THEOREM 4.2. *Let Γ^* be the family of tests defined by (3.10) and (4.3). If $(\mathcal{A}_{\varphi, Mmt})$ holds, and if for all $n \geq 2$, $U_{n,h_\varphi}(\mathbb{X}_n)$ is non-degenerate under (H_0) , then Γ^* satisfies both (\mathcal{P}_{size}) and $(\mathcal{P}_{consist.})$.*

4.4. Non-asymptotic properties of the permutation tests. As seen above, the main theoretical advantage of the permutation approach, as compared with the bootstrap one, lies on the result of Proposition 4.1, which allows in fact to prove that the tests also satisfy non-asymptotic properties. For a fixed integer $n \geq 2$, considering for instance the test $\Delta_{h_\varphi,\alpha,n}^+$ in Γ^* , one has indeed under (H_0) , by Proposition 4.1,

$$\begin{aligned} S &:= \mathbb{P}(\Delta_{h_\varphi,\alpha,n}^+(\mathbb{X}_n) = 1) = \mathbb{P}(\sqrt{n}U_{n,h_\varphi}(\mathbb{X}_n) > q_{\varphi,1-\alpha,n}^*(\mathbb{X}_n)) \\ &= \frac{1}{n!} \sum_{\pi_n \in \mathfrak{S}_n} \mathbb{P}(\sqrt{n}U_{n,h_\varphi}(\mathbb{X}_n^{\pi_n}) > q_{\varphi,1-\alpha,n}^*(\mathbb{X}_n^{\pi_n})). \end{aligned}$$

Since the distribution of $\sqrt{n}U_{n,h_\varphi}(\mathbb{X}_n^{\Pi_n \circ \pi_n})$ given $\mathbb{X}_n^{\pi_n}$ is the same as the one of $\sqrt{n}U_{n,h_\varphi}(\mathbb{X}_n^{\Pi_n})$ given \mathbb{X}_n , $q_{\varphi,1-\alpha,n}^*(\mathbb{X}_n^{\pi_n}) = q_{\varphi,1-\alpha,n}^*(\mathbb{X}_n)$. Hence,

$$\begin{aligned} S &= \sum_{\pi_n \in \mathfrak{S}_n} \mathbb{P}(\sqrt{n}U_{n,h_\varphi}(\mathbb{X}_n^{\Pi_n}) > q_{\varphi,1-\alpha,n}^*(\mathbb{X}_n) | \Pi_n = \pi_n) \mathbb{P}(\Pi_n = \pi_n) \\ &= \mathbb{P}(\sqrt{n}U_{n,h_\varphi}(\mathbb{X}_n^{\Pi_n}) > q_{\varphi,1-\alpha,n}^*(\mathbb{X}_n)) \\ &= \mathbb{E}[\mathbb{P}(\sqrt{n}U_{n,h_\varphi}(\mathbb{X}_n^{\Pi_n}) > q_{\varphi,1-\alpha,n}^*(\mathbb{X}_n) | \mathbb{X}_n)] \leq \alpha. \end{aligned}$$

The same arguments, adapted to any test in Γ^* , lead to the following result.

PROPOSITION 4.2. *Let Γ^* be the family of tests defined by (3.10) and (4.3). For each $n \geq 2$, each test Δ_n of a sequence $\Delta = (\Delta_n)_{n \geq 2}$ in Γ^* is exactly of level α , i.e. if $P = P^1 \otimes P^2$, $\mathbb{P}(\Delta_n(\mathbb{X}_n) = 1) \leq \alpha$.*

4.5. Permutation tests with Monte Carlo approximation. Every test in Γ^* involves an exact computation of the quantiles $q_{\varphi, \eta, n}^*(\mathbb{X}_n)$. Though such a computation is possible by sorting the $n!$ values of $\{U_{n, h_\varphi}(\mathbb{X}_n^{\pi_n})\}_{\pi_n \in \mathfrak{S}_n}$, as for the bootstrap approach, it does not often have a reasonable algorithmic complexity. Therefore, those quantiles are approximated as usual by a Monte Carlo method without changing their asymptotic properties. Moreover, it is interesting to point out that, even if they are based on a Monte Carlo approximation, these tests are still exactly of the desired level. This is proved using Proposition 4.1 and [43, Lemma 1].

More precisely, let $B \geq 1$ be a chosen number of iterations for the Monte Carlo method, and $(\Pi_n^1, \dots, \Pi_n^B)$ be a sample of i.i.d. random permutations uniformly distributed on \mathfrak{S}_n . For each \mathbf{b} in $\{1, \dots, B\}$, we introduce $U^{*, \mathbf{b}} = U_{n, h_\varphi}(\mathbb{X}_n^{\Pi_n^{\mathbf{b}}})$, and we denote by $U^{*, B+1}$ the statistic $U_{n, h_\varphi}(\mathbb{X}_n)$ computed on the original sample \mathbb{X}_n . The order statistic associated with $(U^{*, 1}, \dots, U^{*, B+1})$ is denoted as usual by $(U^{*, (1)}, \dots, U^{*, (B+1)})$. For a sequence $(B_n)_{n \geq 2}$ of positive integers, the Monte Carlo permuted tests are associated to

$$(4.4) \quad q = \left[n^{1/2} \left(U^{*, (\lceil (1-\alpha)(B_n+1) \rceil)}, U^{*, (\lfloor \alpha(B_n+1) \rfloor + 1)} \right) \right]_{n \geq 2}.$$

PROPOSITION 4.3. *Let Γ_{MC}^* be the family of tests defined by (3.10) and (4.4). Then for every $n \geq 2$, whatever the choice of B_n , each test Δ_n of a sequence $\Delta = (\Delta_n)_{n \geq 2}$ in Γ_{MC}^* is exactly of level α , i.e. if $P = P^1 \otimes P^2$, $\mathbb{P}(\Delta_n(\mathbb{X}_n) = 1) \leq \alpha$. Moreover, if $B_n \rightarrow_{n \rightarrow +\infty} +\infty$, under the same assumptions as in Theorem 4.2, Γ_{MC}^* satisfies both (\mathcal{P}_{size}) and $(\mathcal{P}_{consist.})$.*

5. Simulation study. In this section, we study our testing procedures from a practical point of view, by giving estimations of the size and the power for various underlying distributions that are coherent with real neuronal data. This allows to verify the usability of these new methods in practice, and to compare them to existing classical methods. A real data sets study and a more operational and complete method for neuroscientists derived from the present ones is the subject of an ongoing work. The programs have been optimised, parallelised in C++ and interfaced with R. The code is available at <https://code.google.com/p/neuro-stat/>.

5.1. *Presentation of the study.* All along the study, h is taken equal to $h_{\varphi_{\delta}^{coinc}}$ (see (2.4)), where φ_{δ}^{coinc} is defined in (2.1) and $\alpha = 0.05$. We only present the results for unilateral tests by upper values, but an analogous study has been performed for tests by lower values with similar results. Five different testing procedures are compared.

5.1.1. *Testing procedures.*

- (CLT) Test based on the Central Limit Theorem for U -statistics (see Proposition 3.5) which rejects (H_0) when the test statistic $S_{n,h}(\mathbb{X}_n)$ in (3.8) is larger than the $(1 - \alpha)$ -quantile of the standard normal distribution.
- (B) Monte Carlo bootstrap test by upper values of Γ_{MC}^* (see (3.10) and (3.12)).
- (P) Monte Carlo permutation test by upper values of Γ_{MC}^* (see (3.10) and (4.4)).
- (GA) Unilateral test by upper values introduced in [49, Definition 3] under the notation $\Delta_{GAUE}^+(\alpha)$, based on a Gaussian approximation of the total number of coincidences.
- (TS) Trial-shuffling test based on a Monte Carlo approximation of the p -value introduced in [34, equation (3)], but adapted to the present notion of coincidences. This test is the reference distribution-free method for neuroscientists. To be more precise, let $C(\mathbb{X}_n) = \sum_{i=1}^n \varphi_{\delta}^{coinc}(X_i^1, X_i^2)$, where $X_i = (X_i^1, X_i^2)$, be the total number of coincidences of $\mathbb{X}_n = (X_1, \dots, X_n)$. The trial-shuffling method consists in uniformly drawing with replacement n i.i.d. pairs of indices $\{(i^*(k), j^*(k))\}_{1 \leq k \leq n}$ in $\{(i, j), 1 \leq i \neq j \leq n\}$, and considering the associated TS -sample $\mathbb{X}_n^{TS} = \left(\left(X_{i^*(k)}^1, X_{j^*(k)}^2 \right) \right)_{1 \leq k \leq n}$. The Monte Carlo p -value is defined by $\alpha_B^{TS} = \frac{1}{B} \sum_{j=1}^B \mathbb{1}_{\{C(\mathbb{X}_n^{TS,j}) \geq C(\mathbb{X}_n)\}}$, where $(\mathbb{X}_n^{TS,1}, \dots, \mathbb{X}_n^{TS,B})$ are B independent TS -samples, and the test rejects (H_0) if $\alpha_B^{TS} \leq \alpha$.

The number B of steps in the Monte Carlo methods is taken equal to 10000.

5.1.2. *Simulated data.* Various types of point processes are simulated here to check the distribution-free character of our approaches and to investigate their limits. Of course, each of the considered point processes satisfies the moment assumptions on the number of points so that the theorems in this article can be applied. From now on and to be coherent with the neuroscience application which originally motivated this work, the point processes are simulated on a time slot $[a, b]$ of length 0.1.

Estimation of the size. The three data sets simulated under (H_0) consist of i.i.d. samples of pairs of independent point processes. For simplicity, both

processes have the same distribution, though this is not required.

Exp. A Homogeneous Poisson processes on $[a, b]$ with intensity $\lambda = 60$.

Exp. B Inhomogeneous Poisson processes with intensity $f_\lambda : t \in [a, b] \mapsto \lambda t$ and $\lambda = 60$.

Exp. C Hawkes Processes as detailed in [49] i.e. point processes with conditional intensity $\lambda(t) = \max\left(0, \mu - \int_a^t \nu \mathbf{1}_{\{]0, r]\}}(t-s) dN_X(s)\right)$, for t in $[a, b]$, with spontaneous intensity $\mu = 60$, refractory period $r = 0.001$, and $\nu > \mu$ such that for all point T in N_X and t in $]T, T+r]$, $\lambda(t) = 0$. This choice of ν prevents two points to occur at a distance less than the refractory period r to reflect typical neuronal behavior.

Study of the power. The three data sets simulated under (H_1) are such that the number of coincidences is larger than expected under (H_0) . The models (injection or Hawkes) are classical in neuroscience and already used in [49, 20].

Exp. D Homogeneous injection model. $X^1 = X_{ind}^1 \cup X_{com}$ and $X^2 = X_{ind}^2 \cup X_{com}$, with X_{ind}^1 and X_{ind}^2 being two independent homogeneous Poisson processes with intensity $\lambda_{ind} = 54$, X_{com} being a common homogeneous Poisson process with intensity $\lambda_{com} = 6$.

Exp. E Inhomogeneous injection model. As **Exp. D**, except that both X_{ind}^1 and X_{ind}^2 are independent inhomogeneous Poisson processes with intensity $f_{\lambda_{ind}}$ (see **Exp. B**) and $\lambda_{ind} = 54$.

Exp. F Dependent bivariate Hawkes processes. The coordinates X^1 and X^2 of a same pair respectively have the conditional intensities:

$$\lambda^1(t) = \max\left\{0, \mu - \int_a^t \nu \mathbf{1}_{\{]0, r]\}}(t-s) dN_{X^1}(s) + \int_a^t \eta \mathbf{1}_{\{]0, u]\}}(t-s) dN_{X^2}(s)\right\},$$

$$\lambda^2(t) = \max\left\{0, \mu - \int_a^t \nu \mathbf{1}_{\{]0, r]\}}(t-s) dN_{X^2}(s) + \int_a^t \eta \mathbf{1}_{\{]0, u]\}}(t-s) dN_{X^1}(s)\right\},$$

with the spontaneous intensity $\mu = 54$, the interaction intensity $\eta = 6$ in the period designated by $u = 0.005$ and the refractory period designated by $r = 0.001$ with $\nu \gg \mu + \eta u$ such that once again, $\lambda^j(t)$ is null on each $]T, T+r]$, T in N^j . We arbitrarily took $\nu = 50(2\mu + \eta)$.

5.2. Results.

Varying number of trials n . In Figure 1, the delay is fixed at $\delta = 0.01$ and the number n of trials varies in $\{10, 20, 50, 100\}$. Note that when the number of trials is too small ($n = 10$), the estimated variance in (CLT) is sometimes negative, therefore, the test cannot be implemented.

The left hand side of Figure 1 corresponds to estimated sizes. On the one hand, one can see in the case of homogeneous Poisson processes (**Exp. A**) and in the case of refractory Hawkes processes (**Exp. C**) that the methods

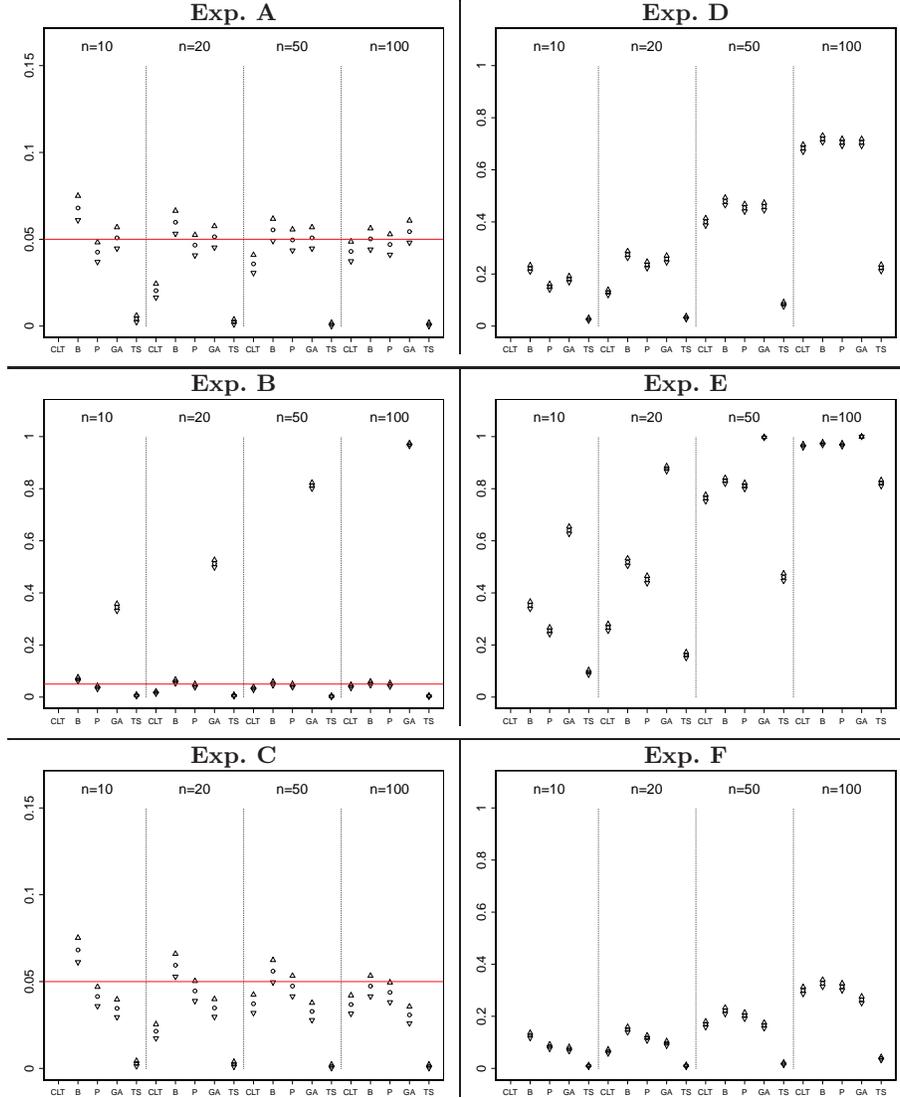


Figure 1: Estimated sizes and powers for different number of trials n , all the tests being performed with a level $\alpha = 0.05$. The circles represent the percentage of rejection on 5000 simulations for each different method, the triangles represent the corresponding endpoints of a 95% confidence interval. The corresponding experiments are described in Section 5.1.2.

(CLT), (B), (P) and (GA) are quite equivalent, even though the size, or Type I error rate, seems less controlled in the bootstrap approach (B) especially

for small numbers of trials. Yet, one can see the convergence of the size of the bootstrap test to α as the number of trials goes to infinity, which illustrates Proposition 3.6. On the other hand, in the case of inhomogeneous Poisson processes (**Exp. B**), one can see that the (GA) test has a huge Type I error rate and is thus inadequate here. Indeed it is based on the strong assumption that the data are homogeneous Poisson processes though they are in fact strongly nonstationary. The test tends thus to reject the independence null hypothesis even when the data are independent. Finally, in the three considered cases, the (TS) approach has a very small size, and is thus too conservative as one can see in the power study.

The right hand side of Figure 1 corresponds to estimated powers, which increase as n grows. This is in line with the consistency of the tests. Now, as it could be expected when looking at the sizes for the (TS) approach, the power is distinctly lower than the ones for the other methods, which confirms its conservative behavior. The other approaches are more similar in **Exp. D** or **Exp. F** though (B) clearly outperforms all tests. Note that in the inhomogeneous case (**Exp. E**), (GA) seems to have a better power, but this is at the price of an uncontrolled Type I error.

This part of the simulation study illustrates the convergences of the size and the power of the tests introduced here. The permutation approach seems to guarantee a better control of the size, whereas the bootstrap approach guarantees a better power, both approaches being effective for any considered kind of point processes and any sample size, unlike the (GA) test which has very restrictive assumptions. Moreover, the (CLT) test seems to have also satisfying results, but with a slower convergence than the (B) and (P) ones. This seems to illustrate that the conditional bootstrap and permutation distributions give better approximations of the original one under independence than a simple central limit theorem. This phenomenon is well-known as the second order accuracy of the bootstrap in more classical frameworks.

Varying delay δ . We now investigate the impact of the choice for the delay δ by making δ vary in $\{0.001, 0.005, 0.01, 0.02\}$ for a fixed number of trials $n = 50$. The results for the sizes being similar to the previous study, only the estimated powers are presented in Figure 2.

On the top row of Figure 2, the same process is injected in both coordinates: the coincidences are exact in the sense that they have no delay. Therefore, the best choice for the delay parameter δ is the smallest possible value: the obtained power is 1 for very small δ 's (e.g. $\delta = 0.001$) and then decreases as δ increases. On the contrary on the bottom row, it can be noticed that the highest power is for $\delta = 0.005$ which is the exact length of the interaction period u . Once again, the (TS) method performs poorly, as

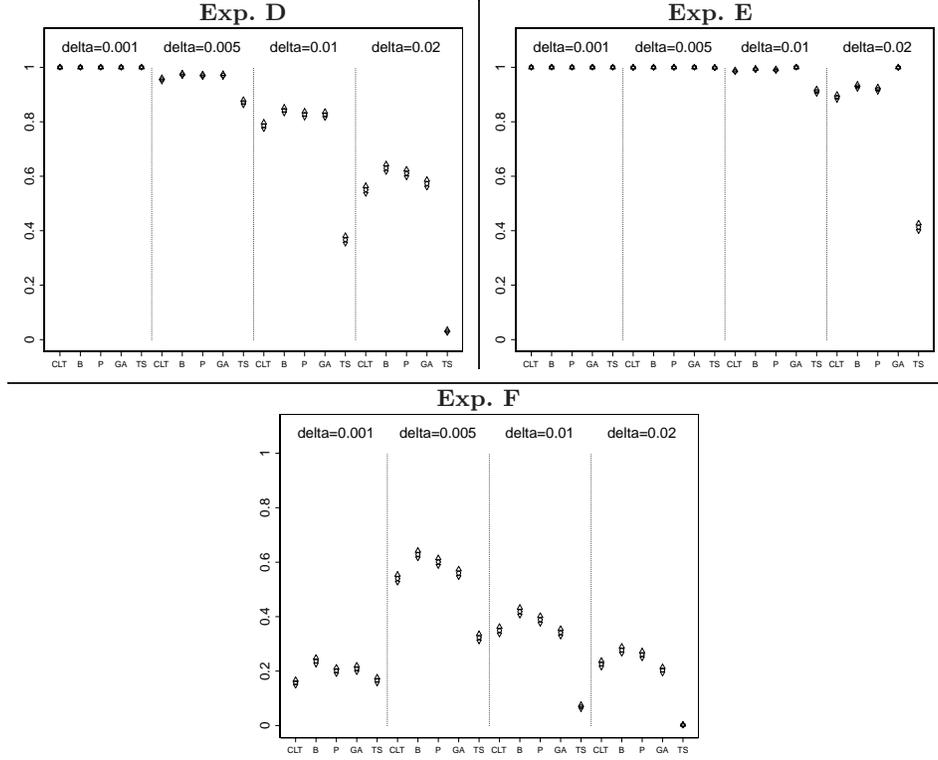


Figure 2: Estimated powers for different δ . Same convention as in Figure 1.

does the (CLT) method. The three other methods seem to be quite equivalent except in the inhomogeneous case (**Exp. E**) where the (GA) method has a power always equal to 1, but at the price of an uncontrolled Type I error.

To conclude, we wish to underline the importance of the choice of the delay parameter δ , which is an open theoretical question. Note that from a neuroscience point of view, this cannot be simply reduced to the study of $\sup_{\delta} |U_{n, h_{\varphi_{\delta}^{coinc}}}|$ since the values of δ for which the rejection takes place contain a very important information for biologists: the delay of interaction. This is the purpose of a future work.

6. Key proofs or sketches of proofs. All along this section, C and C' denote positive constants, that may vary from one line to another one.

6.1. *Proof of Theorem 3.1.* By Proposition 3.3, for all $n \geq 2$,

$$\begin{aligned} & d_2\left(\mathcal{L}\left(\sqrt{n}U_{n,h}, P_n^1 \otimes P_n^2 \mid \mathbb{X}_n\right), \mathcal{L}\left(\sqrt{n}U_{n,h}, P^1 \otimes P^2\right)\right) \\ & \leq C \inf_{\substack{(Y_{n,1}^*, Y_1), (Y_{n,2}^*, Y_2) \text{ i.i.d.} / \\ Y_{n,1}^*, Y_{n,2}^* \sim P_n^1 \otimes P_n^2, Y_1, Y_2 \sim P^1 \otimes P^2}} \mathbb{E}^* \left[\left(h(Y_{n,1}^*, Y_{n,2}^*) - h(Y_1, Y_2) \right)^2 \right]. \end{aligned}$$

Our goal is to construct, for almost all ω in Ω , a sequence of random variables $(\bar{Y}_{n,\omega,1}^*)_{n \geq 1}$ such that for every $n \geq 1$, $\bar{Y}_{n,\omega,1}^* \sim P_{n,\omega}^1 \otimes P_{n,\omega}^2$, where $P_{n,\omega}^j = n^{-1} \sum_{i=1}^n \delta_{X_i^j(\omega)}$ is the j th marginal empirical measure corresponding to the realisation $\mathbb{X}_n(\omega)$, a random variable $\bar{Y}_{\omega,1} \sim P^1 \otimes P^2$, and $\left\{ (\bar{Y}_{n,\omega,2}^*)_{n \geq 1}, \bar{Y}_{\omega,2} \right\}$ an independent copy of $\left\{ (\bar{Y}_{n,\omega,1}^*)_{n \geq 1}, \bar{Y}_{\omega,1} \right\}$ on some probability space $(\Omega'_\omega, \mathcal{A}'_\omega, \mathbb{P}'_\omega)$ depending on ω such that

$$(6.1) \quad \mathbb{E}'_\omega \left[\left(h(\bar{Y}_{n,\omega,1}^*, \bar{Y}_{n,\omega,2}^*) - h(\bar{Y}_{\omega,1}, \bar{Y}_{\omega,2}) \right)^2 \right] \xrightarrow{n \rightarrow +\infty} 0,$$

where \mathbb{E}'_ω denotes the expectation corresponding to \mathbb{P}'_ω . Then from (6.1), we can conclude by noting that, for almost all ω in Ω ,

$$\begin{aligned} & \inf_{\substack{(Y_{n,1}^*, Y_1), (Y_{n,2}^*, Y_2) \text{ i.i.d.} / \\ Y_{n,1}^*, Y_{n,2}^* \sim P_{n,\omega}^1 \otimes P_{n,\omega}^2, Y_1, Y_2 \sim P^1 \otimes P^2}} \mathbb{E}^* \left[\left(h(Y_{n,1}^*, Y_{n,2}^*) - h(Y_1, Y_2) \right)^2 \right] (\omega) \\ & \leq \mathbb{E}'_\omega \left[\left(h(\bar{Y}_{n,\omega,1}^*, \bar{Y}_{n,\omega,2}^*) - h(\bar{Y}_{\omega,1}, \bar{Y}_{\omega,2}) \right)^2 \right] \xrightarrow{n \rightarrow +\infty} 0. \end{aligned}$$

To prove (6.1), consider $(\Omega, \mathcal{A}, \mathbb{P})$ the probability space on which all the X_n 's are defined. In what follows, one can keep in mind that Ω represents the randomness in the original sequence $(X_n)_n$. Thus, a given ω in Ω represents a given realisation of $(X_n)_n$.

As a preliminary step, from Proposition 3.4, there exists some subset Ω_1 of Ω such that $\mathbb{P}(\Omega_1) = 1$ and for every ω in Ω_1 ,

$$(6.2) \quad \frac{1}{n^4} \sum_{i,j,k,l=1}^n h^2 \left((X_i^1(\omega), X_j^2(\omega)), (X_k^1(\omega), X_l^2(\omega)) \right) \xrightarrow{n \rightarrow +\infty} \mathbb{E} \left[h^2 \left((X_1^1, X_2^2), (X_3^1, X_4^2) \right) \right].$$

Applying Theorem 3 in [52], since $(\mathcal{X}, d_{\mathcal{X}})$ defined by (3.2) is separable, P -a.s. in $(X_n)_n$, $P_n^1 \xrightarrow{n \rightarrow +\infty} P^1$ and $P_n^2 \xrightarrow{n \rightarrow +\infty} P^2$. Hence there exists some subset Ω_2 of Ω such that $\mathbb{P}(\Omega_2) = 1$ and for every ω in Ω_2 ,

$$(6.3) \quad P_{n,\omega}^1 \otimes P_{n,\omega}^2 \xrightarrow{n \rightarrow +\infty} P^1 \otimes P^2,$$

Now, let us consider $\Omega_0 = \Omega_1 \cap \Omega_2$, and fix ω in Ω_0 .

Following the proof of Skorokhod's representation theorem in [13, Theorem 11.7.2 p. 415], since $(\mathcal{X}^2, d_{\mathcal{X}^2})$ is a separable space, it is possible to construct some probability space $(\Omega'_\omega, \mathcal{A}'_\omega, \mathbb{P}'_\omega)$, and some random variables $\bar{Y}_{n,\omega,1}^* : \Omega'_\omega \rightarrow \mathcal{X}^2$, $\bar{Y}_{n,\omega,2}^* : \Omega'_\omega \rightarrow \mathcal{X}^2$ with distribution $P_{n,\omega}^1 \otimes P_{n,\omega}^2$, and $\bar{Y}_{\omega,1} : \Omega'_\omega \rightarrow \mathcal{X}^2$, $\bar{Y}_{\omega,2} : \Omega'_\omega \rightarrow \mathcal{X}^2$ with distribution $P^1 \otimes P^2$ such that \mathbb{P}'_ω -a.s., $\bar{Y}_{n,\omega,1}^* \xrightarrow{n \rightarrow +\infty} \bar{Y}_{\omega,1}$ and $\bar{Y}_{n,\omega,2}^* \xrightarrow{n \rightarrow +\infty} \bar{Y}_{\omega,2}$, $\left\{ (\bar{Y}_{n,\omega,1}^*)_{n \geq 1}, \bar{Y}_{\omega,1} \right\}$ and $\left\{ (\bar{Y}_{n,\omega,2}^*)_{n \geq 1}, \bar{Y}_{\omega,2} \right\}$ being independent, so that w.r.t. the metric d (see (3.3)),

$$(6.4) \quad \mathbb{P}'_\omega\text{-a.s.}, \quad (\bar{Y}_{n,\omega,1}^*, \bar{Y}_{n,\omega,2}^*) \xrightarrow{n \rightarrow +\infty} (\bar{Y}_{\omega,1}, \bar{Y}_{\omega,2}).$$

But under (\mathcal{A}_{Cont}) , h is continuous on a set \mathcal{C} s. t. $\mathbb{P}'_\omega((\bar{Y}_{\omega,1}, \bar{Y}_{\omega,2}) \in \mathcal{C}) = (P^1 \otimes P^2)^{\otimes 2}(\mathcal{C}) = 1$, hence \mathbb{P}'_ω -a.s., $h(\bar{Y}_{n,\omega,1}^*, \bar{Y}_{n,\omega,2}^*) \xrightarrow{n \rightarrow +\infty} h(\bar{Y}_{\omega,1}, \bar{Y}_{\omega,2})$. As \mathbb{P}'_ω -a.s. convergence implies convergence in probability, to obtain (6.1), we only need to prove that the sequence $(h^2(\bar{Y}_{n,\omega,1}^*, \bar{Y}_{n,\omega,2}^*))_{n \geq 1}$ is uniformly integrable, according to Theorem 16.6 p. 165 of [47]. We therefore conclude since (6.2) is equivalent to

$$\begin{aligned} \mathbb{E}'_\omega[h^2(\bar{Y}_{n,\omega,1}^*, \bar{Y}_{n,\omega,2}^*)] &= \frac{1}{n^4} \sum_{i,j,k,l=1}^n h^2((X_i^1(\omega), X_j^2(\omega)), (X_k^1(\omega), X_l^2(\omega))) \\ &\xrightarrow{n \rightarrow +\infty} \mathbb{E}[h^2((X_1^1, X_2^2), (X_3^1, X_4^2))] = \mathbb{E}'_\omega[h^2(\bar{Y}_{\omega,1}, \bar{Y}_{\omega,2})]. \end{aligned}$$

6.2. *Proof of Corollary 3.1.* By Proposition 3.5, we have that

$$(6.5) \quad \mathcal{L}(\sqrt{n}U_{n,h}, P^1 \otimes P^2) \xrightarrow{n \rightarrow +\infty} \mathcal{N}(0, \sigma_{h, P^1 \otimes P^2}^2),$$

where $\mathcal{N}(0, \sigma_{h, P^1 \otimes P^2}^2)$ has a continuous c.d.f. Therefore, by [50, Lemma 2.11],

$$(6.6) \quad \sup_{z \in \mathbb{R}} \left| \mathbb{P}\left(\sqrt{n}U_{n,h}(\mathbb{X}_n^\perp) \leq z\right) - \Phi_{0, \sigma_{h, P^1 \otimes P^2}^2}(z) \right| \xrightarrow{n \rightarrow +\infty} 0.$$

Furthermore, since convergence w.r.t the d_2 distance implies weak convergence, Theorem 3.1 combined with (6.5) leads to

$$(6.7) \quad \mathcal{L}(\sqrt{n}U_{n,h}, P_n^1 \otimes P_n^2 | \mathbb{X}_n) \xrightarrow{n \rightarrow +\infty} \mathcal{N}(0, \sigma_{h, P^1 \otimes P^2}^2) \quad P\text{-a.s. in } (X_n)_n.$$

Hence,

$$(6.8) \quad \sup_{z \in \mathbb{R}} \left| \mathbb{P} \left(\sqrt{n} U_{n,h}(\mathbb{X}_n^*) \leq z | \mathbb{X}_n \right) - \Phi_{0, \sigma_{h, P^1 \otimes P^2}^2}^{-1}(z) \right| \xrightarrow{n \rightarrow +\infty} 0 \quad P\text{-a.s. in } (X_n)_n,$$

and the first part of the corollary is obtained. Then [50, Lemma 21.2] can then be applied to both (6.5) and (6.7), to obtain that on the event where (6.7) holds

$$(6.9) \quad q_{h, 1-\alpha, n}^*(\mathbb{X}_n) \xrightarrow{n \rightarrow +\infty} \Phi_{0, \sigma_{h, P^1 \otimes P^2}^2}^{-1}(1-\alpha) \quad P\text{-a.s. in } (X_n)_n,$$

and that $q_{h, \eta, n}^\perp$ also converges to $\Phi_{0, \sigma_{h, P^1 \otimes P^2}^2}^{-1}(\eta)$.

6.3. *Proof of Theorem 3.2.* Let us focus on the sequence of tests by upper values in Γ^* , the proof for the other tests being similar. Under (H_0) , from Proposition 3.5 and (6.9), by Slutsky's lemma, $(\sqrt{n} U_{n,h}(\mathbb{X}_n), q_{h, 1-\alpha, n}^*(\mathbb{X}_n))$ converges in distribution to $(Z, \Phi_{0, \sigma_{h, P^1 \otimes P^2}^2}^{-1}(1-\alpha))$, where $Z \sim \mathcal{N}(0, \sigma_{h, P^1 \otimes P^2}^2)$.

Therefore, under (H_0) , $\mathbb{P}(\sqrt{n} U_{n,h}(\mathbb{X}_n) > q_{h, 1-\alpha, n}^*(\mathbb{X}_n)) \rightarrow_{n \rightarrow +\infty} \alpha$.

Under (H_1) , by Proposition 3.4,

$$U_{n,h}(\mathbb{X}_n) \xrightarrow{n \rightarrow +\infty} \int h(x, x') dP(x) dP(x') > 0, \quad P\text{-a.s. in } (X_n)_n.$$

Furthermore, due to (6.9), $q_{h, 1-\alpha, n}^*(\mathbb{X}_n) / \sqrt{n} \rightarrow_{n \rightarrow +\infty} 0$ P -a.s. in $(X_n)_n$. Hence, $\mathbb{P}(\sqrt{n} U_{n,h}(\mathbb{X}_n) \leq q_{h, 1-\alpha, n}^*(\mathbb{X}_n)) \rightarrow_{n \rightarrow +\infty} 0$.

6.4. *Proof of Proposition 3.6.* As above, let us focus on the sequence of tests by upper values in Γ_{MC}^* . Let $Z \sim \mathcal{N}(0, 1)$ and let us define for z in \mathbb{R} ,

$$F_{n, \mathbb{X}_n}^*(z) = \mathbb{P}(\sqrt{n} U_{n,h}(\mathbb{X}_n^*) \leq z | \mathbb{X}_n), \quad F_{n, \mathbb{X}_n}^{*B_n}(z) = \frac{1}{B_n} \sum_{\mathbf{b}=1}^{B_n} \mathbb{1}_{\{\sqrt{n} U_{n,h\varphi}(\mathbb{X}_n^{*\mathbf{b}}) \leq z\}}.$$

By the Dvoretzky-Kiefer-Wolfowitz inequality (see [50]), for $n \geq 2$ and $\varepsilon > 0$,

$$\begin{aligned} \mathbb{P} \left(\sup_{z \in \mathbb{R}} \left| F_{n, \mathbb{X}_n}^{*B_n}(z) - F_{n, \mathbb{X}_n}^*(z) \right| > \varepsilon \right) &= \mathbb{E} \left[\mathbb{P} \left(\sup_{z \in \mathbb{R}} \left| F_{n, \mathbb{X}_n}^{*B_n}(z) - F_{n, \mathbb{X}_n}^*(z) \right| > \varepsilon \mid \mathbb{X}_n \right) \right] \\ &\leq 2e^{-2B_n \varepsilon^2} \xrightarrow{n \rightarrow +\infty} 0, \end{aligned}$$

that is $\sup_{z \in \mathbb{R}} |F_{n, \mathbb{X}_n}^{*B_n}(z) - F_{n, \mathbb{X}_n}^*(z)| \xrightarrow[n \rightarrow +\infty]{\mathbb{P}} 0$. With (6.8), this leads to

$$(6.10) \quad \sup_{z \in \mathbb{R}} \left| F_{n, \mathbb{X}_n}^{*B_n}(z) - \Phi_{0, \sigma_{h, P^1 \otimes P^2}^2}^{-1}(z) \right| \xrightarrow[n \rightarrow +\infty]{\mathbb{P}} 0.$$

We finish the proof using similar arguments as in [50, Lemma 21.2], combined with a subsequence argument [13, Theorem 9.2.1]. Let ϕ_0 be an extraction. Then, by (6.10), there exists an extraction ϕ_1 , and some $\Omega_0 \subset \Omega$ such that $\mathbb{P}(\Omega_0) = 1$, and for every ω in Ω_0 ,

$$\sup_{z \in \mathbb{R}} \left| F_{\phi_1 \circ \phi_0(n), \mathbb{X}_{\phi_1 \circ \phi_0(n)}}^{*B_{\phi_1 \circ \phi_0(n)}}(\omega)(z) - \Phi_{0, \sigma_{h, P^1 \otimes P^2}^2}(z) \right| \xrightarrow{n \rightarrow +\infty} 0.$$

From now on, fix ω in Ω_0 . In particular, this fixes a realisation of \mathbb{X}_n , and a realisation of $(\mathbb{X}_n^{*1}, \dots, \mathbb{X}_n^{*B_n})$ and thus, $F_{n, \mathbb{X}_n}^{*B_n}(\omega)$ is deterministic. Hence, $F_{\phi_1 \circ \phi_0(n), \mathbb{X}_{\phi_1 \circ \phi_0(n)}}^{*B_{\phi_1 \circ \phi_0(n)}}(\omega)(Z) \xrightarrow[n \rightarrow +\infty]{a.s.} \Phi_{0, \sigma_{h, P^1 \otimes P^2}^2}(Z)$, and for η in $(0, 1)$,

$$\begin{aligned} \Phi_{0,1} \left(\left(F_{\phi_1 \circ \phi_0(n), \mathbb{X}_{\phi_1 \circ \phi_0(n)}}^{*B_{\phi_1 \circ \phi_0(n)}}(\omega) \right)^{-1}(\eta) \right) &= \mathbb{P} \left(F_{\phi_1 \circ \phi_0(n), \mathbb{X}_{\phi_1 \circ \phi_0(n)}}^{*B_{\phi_1 \circ \phi_0(n)}}(\omega)(Z) < \eta \right) \\ &\xrightarrow{n \rightarrow +\infty} \mathbb{P} \left(\Phi_{0, \sigma_{h, P^1 \otimes P^2}^2}(Z) < \eta \right) = \Phi_{0,1} \left(\left(\Phi_{0, \sigma_{h, P^1 \otimes P^2}^2} \right)^{-1}(\eta) \right). \end{aligned}$$

Finally, as $\Phi_{0,1}$ is a one-to-one function and $\Phi_{0,1}^{-1}$ is continuous,

$$\begin{aligned} (6.11) \quad &\sqrt{\phi_1 \circ \phi_0(n)} U^{*([\eta(B_{\phi_1 \circ \phi_0(n)})])}(\omega) \\ &= \left(F_{\phi_1 \circ \phi_0(n), \mathbb{X}_{\phi_1 \circ \phi_0(n)}}^{*B_{\phi_1 \circ \phi_0(n)}}(\omega) \right)^{-1}(\eta) \xrightarrow{n \rightarrow +\infty} \Phi_{0, \sigma_{h, P^1 \otimes P^2}^2}^{-1}(\eta), \end{aligned}$$

and this for all ω in Ω_0 , and any initial extraction ϕ_0 . Therefore, we obtain that $\sqrt{n} U^{*([\eta B_n])} \xrightarrow[n \rightarrow +\infty]{\mathbb{P}} \Phi_{0, \sigma_{h, P^1 \otimes P^2}^2}^{-1}(\eta)$. We conclude as for Theorem 3.2.

6.5. *Sketch of proof of Theorem 4.1.* We present here a sketch of the proof, whose complete version can be found in the supplementary material. Let d_{BL} denote the bounded Lipschitz metric, which metricises the weak convergence [13, Prop. 11.3.2 and Th. 11.3.3]. For Z_n any variable depending on \mathbb{X}_n and Π_n , $\mathcal{L}(Z_n | \mathbb{X}_n)$ denotes the conditional distribution of Z_n given \mathbb{X}_n and for any integrable function f , $\mathbb{E}_{P^1 \otimes P^2}[f] = \mathbb{E}[f(X_1^1, X_2^2)]$.

- The first step of the proof consists in decomposing $\sqrt{n} U_{n, h_\varphi}(\mathbb{X}_n^{\Pi_n})$ in

$$\sqrt{n} U_{n, h_\varphi}(\mathbb{X}_n^{\Pi_n}) = \frac{n}{n-1} \left(M_n^{\Pi_n}(\mathbb{X}_n) + \frac{R_n^{\Pi_n}(\mathbb{X}_n)}{\sqrt{n}} - \frac{T_n(\mathbb{X}_n)}{\sqrt{n}} \right),$$

where $M_n^{\Pi_n}(\mathbb{X}_n) = \frac{1}{\sqrt{n}} \sum_{i \neq j} \mathbb{1}_{\{\Pi_n(i)=j\}} C_{i,j}$,

$R_n^{\Pi_n}(\mathbb{X}_n) = \sum_{i=1}^n (\mathbb{1}_{\{\Pi_n(i)=i\}} - \frac{1}{n}) C_{i,i}$, and $T_n(\mathbb{X}_n) = \frac{1}{n} \sum_{i \neq j} C_{i,j}$, with

$C_{i,j} = \varphi(X_i^1, X_j^2) - \mathbb{E}[\varphi(X_i^1, X_j^2) | X_i^1] - \mathbb{E}[\varphi(X_i^1, X_j^2) | X_j^2] + \mathbb{E}_{P^1 \otimes P^2}[\varphi]$,

$X = (X^1, X^2)$ being P -distributed and independent of $(X_n)_n$. We then prove from Cauchy-Schwarz inequality that

$$\mathbb{E} \left[\left(\mathbb{E} \left[\frac{|R_n^{\Pi_n}(\mathbb{X}_n)|}{\sqrt{n}} \middle| \mathbb{X}_n \right] \right)^2 \right] \xrightarrow{n \rightarrow +\infty} 0 \text{ and } \mathbb{E} \left[\left(\frac{T_n(\mathbb{X}_n)}{\sqrt{n}} \right)^2 \right] \xrightarrow{n \rightarrow +\infty} 0,$$

therefore from Markov's inequality,

$$\mathbb{E} \left[\frac{|R_n^{\Pi_n}(\mathbb{X}_n)|}{\sqrt{n}} \middle| \mathbb{X}_n \right] \xrightarrow[n \rightarrow +\infty]{\mathbb{P}} 0 \text{ and } \frac{T_n(\mathbb{X}_n)}{\sqrt{n}} \xrightarrow[n \rightarrow +\infty]{\mathbb{P}} 0.$$

From the definition of d_{BL} , this allows us to derive that

$$(6.12) \quad d_{BL} \left(\mathcal{L}(\sqrt{n}U_{n,h_\varphi}(\mathbb{X}_n^{\Pi_n}) | \mathbb{X}_n), \mathcal{L} \left(\frac{n}{n-1} M_n^{\Pi_n}(\mathbb{X}_n) \middle| \mathbb{X}_n \right) \right) \xrightarrow[n \rightarrow +\infty]{\mathbb{P}} 0.$$

• The second, and most difficult, step of the proof consists in proving that

$$(6.13) \quad d_{BL} \left(\mathcal{L}(M_n^{\Pi_n}(\mathbb{X}_n) | \mathbb{X}_n), \mathcal{N} \left(0, \sigma_{h_\varphi, P^1 \otimes P^2}^2 \right) \right) \xrightarrow[n \rightarrow +\infty]{\mathbb{P}} 0.$$

Consider

$$(6.14) \quad Y_{n,i} = \frac{1}{\sqrt{n}} \sum_{j=1}^{i-1} (\mathbb{1}_{\{\Pi_n(i)=j\}} C_{i,j} + \mathbb{1}_{\{\Pi_n(j)=i\}} C_{j,i}),$$

and for Π'_n another uniformly distributed random permutation with values in \mathfrak{S}_n , independent of Π_n and \mathbb{X}_n , define accordingly $Y'_{n,i}$ by replacing Π_n by Π'_n in (6.14), so that $M_n^{\Pi_n}(\mathbb{X}_n) = \sum_{i=1}^n Y_{n,i}$ and similarly for $M_n^{\Pi'_n}(\mathbb{X}_n)$. Denoting $\mathcal{F}_{n,i} = \sigma(\Pi_n, \Pi'_n, X_1, X_2, \dots, X_i)$ for $n \geq i \geq 2$, we prove through technical computations that for a, b in \mathbb{R} , $(aY_{n,i} + bY'_{n,i}, \mathcal{F}_{n,i})_{2 \leq i \leq n}$ is a martingale difference array which satisfies the assumptions of the following result, commonly attributed to Brown [8].

THEOREM 6.1. *Let $(X_{n,k})_{k \in \{1, \dots, p_n\}, n \in \mathbb{N}^*}$ be a martingale difference array, i.e. such that there exists an array of σ -algebra $(\mathcal{F}_{n,k})_{k \in \{1, \dots, p_n\}, n \in \mathbb{N}^*}$ that is increasing w.r.t. k such that for all $k = 1, \dots, p_n$, $\mathbb{E}[X_{n,k} | \mathcal{F}_{n,k-1}] = 0$. Let $A_n = \sum_{k=1}^{p_n} \mathbb{E}[X_{n,k}^2 | \mathcal{F}_{n,k-1}]$, and assume that*

- $A_n \xrightarrow[n \rightarrow +\infty]{\mathbb{P}} \sigma^2 > 0$,
- $\forall \varepsilon > 0, \sum_{k=1}^{p_n} \mathbb{E}[X_{n,k}^2 \mathbb{1}_{\{|X_{n,k}| > \varepsilon\}}] \rightarrow_{n \rightarrow +\infty} 0$.

Then $Z_n = \sum_{k=1}^{p_n} X_{n,k}$ converges in distribution towards $\mathcal{N}(0, \sigma^2)$.

Thus, given a, b in \mathbb{R} , we obtain that

$$\mathcal{L} \left(aM_n^{\Pi_n}(\mathbb{X}_n) + bM_n^{\Pi'_n}(\mathbb{X}_n) \right) \xrightarrow[n \rightarrow +\infty]{} \mathcal{N} \left(0, (a^2 + b^2) \sigma_{h_\varphi, P^1 \otimes P^2}^2 \right),$$

which, according to the Cramér-Wold device, leads to Lemma 6.1 below.

LEMMA 6.1. *Considering the above notation,*

$$\mathcal{L} \left(\left(M_n^{\Pi_n}(\mathbb{X}_n), M_n^{\Pi'_n}(\mathbb{X}_n) \right)' \right) \xrightarrow[n \rightarrow +\infty]{} \mathcal{N}_2 \left(0, \begin{pmatrix} \sigma_{h_\varphi, P^1 \otimes P^2}^2 & 0 \\ 0 & \sigma_{h_\varphi, P^1 \otimes P^2}^2 \end{pmatrix} \right),$$

where $\mathcal{N}_2(M, V)$ denotes the 2-dimensional Gaussian distribution with mean vector M and variance-covariance matrix V .

From Lemma 6.1, we deduce that for every t in \mathbb{R} ,

$$\begin{cases} \mathbb{P} \left(M_n^{\Pi_n}(\mathbb{X}_n) \leq t \right) \xrightarrow[n \rightarrow +\infty]{} \Phi_{0, \sigma_{h_\varphi, P^1 \otimes P^2}^2}(t), \\ \mathbb{P} \left(M_n^{\Pi_n}(\mathbb{X}_n) \leq t, M_n^{\Pi'_n}(\mathbb{X}_n) \leq t \right) \xrightarrow[n \rightarrow +\infty]{} \Phi_{0, \sigma_{h_\varphi, P^1 \otimes P^2}^2}^2(t). \end{cases}$$

Using Chebychev's inequality, with the fact (see [13, Th. 9.2.1] for instance) that in a separable metric space, convergence in probability is metrizable, and therefore is equivalent to almost-sure convergence of a sub-sequence of any initial subsequence, we prove that this leads to (6.13), therefore,

$$d_{BL} \left(\mathcal{L} \left(\sqrt{n} U_{n, h_\varphi}(\mathbb{X}_n^{\Pi_n}) | \mathbb{X}_n \right), \mathcal{N} \left(0, \sigma_{h_\varphi, P^1 \otimes P^2}^2 \right) \right) \xrightarrow[n \rightarrow +\infty]{\mathbb{P}} 0.$$

• The third, and final, step of the proof consists in deriving, by direct computations and the strong law of large numbers of Hoeffding [25], the convergence

$$\mathbb{E} \left[\left(\sqrt{n} U_{n, h_\varphi}(\mathbb{X}_n^{\Pi_n}) \right)^2 | \mathbb{X}_n \right] \xrightarrow[n \rightarrow +\infty]{a.s.} \sigma_{h_\varphi, P^1 \otimes P^2}^2,$$

which ends the proof.

6.6. *Proof of Corollary 4.1.* Here, unlike the bootstrap approach, we only have in Theorem 4.1 a consistency result in probability. Thus, as for Proposition 3.6, we use an argument of sub-sequences. So let $\phi_0 : \mathbb{N} \rightarrow \mathbb{N}$ be an extraction defining a sub-sequence. By Theorem 4.1, there exists an extraction ϕ_1 such that P -a.s. in $(X_n)_n$,

$$(6.15) \quad \mathcal{L} \left(\sqrt{\phi_1 \circ \phi_0(n)} U_{\phi_1 \circ \phi_0(n), h_\varphi} \left(\mathbb{X}_{\phi_1 \circ \phi_0(n)}^{\Pi_{\phi_1 \circ \phi_0(n)}} \right) | \mathbb{X}_{\phi_1 \circ \phi_0(n)} \right) \xrightarrow[n \rightarrow +\infty]{} \mathcal{N} \left(0, \sigma_{h_\varphi, P^1 \otimes P^2}^2 \right).$$

In particular, applying [50, Lemma 21.2] on the event where the convergence is true, we obtain that for η in $(0, 1)$, $q_{\varphi, \eta, \phi_1 \circ \phi_0(n)}^*(\mathbb{X}_{\phi_1 \circ \phi_0(n)}) \xrightarrow[n \rightarrow +\infty]{a.s.} \Phi_{0, \sigma_{h_\varphi, P^1 \otimes P^2}^2}^{-1}(\eta)$, which ends the proof by [13, Theorem 9.2.1].

Acknowledgments. We thank F. Grammont for fruitful discussions. This work was granted access to the HPC and visualization resources of "Centre de Calcul Interactif" hosted by "Université Nice Sophia Antipolis". This research was partly supported by the french Agence Nationale de la Recherche (ANR 2011 BS01 010 01 projet Calibration) and by the PEPS BMI 2012-2013 *Estimation of dependence graphs for thalamo-cortical neurons and multivariate Hawkes processes*. The PhD grant of M. Albert is funded by the PACA french region.

References.

- [1] M. A. Arcones and E. Giné. On the bootstrap of U and V statistics. *Ann. Statist.*, 20(2):655–674, 1992.
- [2] S. Arlot, G. Blanchard, and E. Roquain. Some nonasymptotic results on resampling in high dimension, I: Confidence regions and II: Multiple tests. *Ann. Statist.*, 38(1):51–82, 83–99, 2010.
- [3] P. J. Bickel and D. A. Freedman. Some asymptotic theory for the bootstrap. *Ann. Statist.*, 9(6):1196–1217, 1981.
- [4] P. Billingsley. *Convergence of probability measures*. Wiley-Interscience, 2009.
- [5] J. R. Blum, J. Kiefer, and M. Rosenblatt. Distribution free tests of independence based on the sample distribution function. *Ann. Math. Statist.*, 32(2):485–498, 1961.
- [6] P. Brémaud and L. Massoulié. Stability of nonlinear Hawkes processes. *Ann. Probab.*, 24(3):1563–1588, 1996.
- [7] J. Bretagnolle. Lois limites du bootstrap de certaines fonctionnelles. *Ann. Inst. H. Poincaré Probab. Statist.*, 19(3):281–296, 1983.
- [8] B. M. Brown. Martingale central limit theorems. *Ann. Math. Statist.*, 42(1):59–66, 1971.
- [9] L. Carstensen, A. Sandelin, O. Winther, and N. R. Hansen. Multivariate Hawkes process models of the occurrence of regulatory elements and an analysis of the pilot ENCODE regions. *BMC Bioinformatics*, 11(456), 2010.
- [10] E. Chung and J. P. Romano. Exact and asymptotically robust permutation tests. *Ann. Statist.*, 41(2):484–507, 2013.
- [11] D. J. Daley and D. Vere-Jones. *An introduction to the theory of point processes. Vol. I. Probability and its Applications*.
- [12] H. Dehling and T. Mikosch. Random quadratic forms and the bootstrap for U-statistics. *J. Multivariate Anal.*, 51(2):392–413, 1994.
- [13] R. M. Dudley. *Real Analysis and Probability*. Cambridge University Press, 2002.
- [14] B. Efron. Bootstrap methods: another look at the jackknife. *Ann. Statist.*, 7(1):1–26, 1979.
- [15] R. A. Fisher. *The design of experiments*. 1935.
- [16] M. Fromont. Model selection by bootstrap penalization for classification. *Mach. Learn.*, 66(2-3):165–207, 2007.
- [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] G. L. Gerstein and D. H. Perkel. Simultaneous recorded trains of action potentials: analysis and functional interpretation. *Science*, 164:828–830, 1969.
- [19] A. Gretton and L. Györfi. Consistent nonparametric tests of independence. *J.M.L.R.*, 11:1391–1423, 2010.

- [20] S. Grün, M. Diesmann, and A. M. Aertsen. *Analysis of parallel spike trains*, chapter Unitary Events analysis. Springer Series in Computational Neuroscience, 2010.
- [21] N. R. Hansen, P. Reynaud-Bouret, and V. Rivoirard. Lasso and probabilistic inequalities for multivariate point processes. *To appear in Bernoulli*, 2014.
- [22] W. Hoeffding. A class of statistics with asymptotically normal distribution. *Ann. Math. Statist.*, 19(3):293–325, 1948.
- [23] W. Hoeffding. A non-parametric test of independence. *Ann. Math. Statist.*, 19(4):546–557, 1948.
- [24] W. Hoeffding. The large-sample power of tests based on permutation of the observations. *Ann. Math. Statist.*, 23(2):169–192, 1952.
- [25] W. Hoeffding. The strong law of large numbers for U-statistics. Institute of Statistics, Mimeograph Series No. 302, 1961.
- [26] L. Horváth and M. Hušková. Testing for changes using permutations of U-statistics. *J. Statist. Plann. Inference*, 128(2):351–371, 2005.
- [27] H. Hotelling and M. R. Pabst. Rank correlation and tests of significance involving no assumption of normality. *Ann. Math. Statist.*, 7(1):29–43, 1936.
- [28] M. G. Kendall. A new measure of rank correlation. *Biometrika*, 30(1–2):81–93, 1938.
- [29] A. Leucht and M. H. Neumann. Consistency of general bootstrap methods for degenerate U-type and V-type statistics. *J. Multivariate Anal.*, 100(8):1622–1633, 2009.
- [30] Y. Ogata. On Lewis’ simulation method for point processes. *IEEE Transactions on Information Theory*, 27(1):23–31, 1981.
- [31] K. Pearson. On the criterion that a given system of deviations from the probable in the case of a correlated system of variables is such that it can be reasonably supposed to have arisen from random sampling. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, 50(302):157–175, 1900.
- [32] K. Pearson. On the probability that two independent distributions of frequency are really samples from the same population. *Biometrika*, 8:250–254, 1911.
- [33] V. Pernice, B. Staude, S. Cardanobile, and S. Rotter. How structure determines correlations in neuronal networks. *PLoS Computational Biology*, 7(5):e1002059, 2012.
- [34] G. Pipa, M. Diesmann, and S. Grün. Significance of joint-spike events based on trial-shuffling by efficient combinatorial methods. *Complexity*, 8(4):1–8, 2003.
- [35] G. Pipa and S. Grün. Non-parametric significance estimation of joint-spike events by shuffling and resampling. *Neurocomputing*, 52–54:31–37, 2003.
- [36] E. J. Pitman. Significance tests which may be applied to samples from any populations. *Suppl. J. R. Stat. Soc.*, 4:117–130, 225–232, 1937.
- [37] E. J. Pitman. Significance tests which may be applied to samples from any populations. The analysis of variance test. *Biometrika*, 29:322–335, 1938.
- [38] C. Pouzat and A. Chaffiol. Automatic spike train analysis and report generation. an implementation with R, R2HTML and STAR. *J. Neurosci. Meth.*, 2009.
- [39] P. Reynaud-Bouret, V. Rivoirard, F. Grammont, and C. Tuleau-Malot. Goodness-of-fit tests and nonparametric adaptive estimation for spike train analysis. *J. Math. Neuro.*, 4(3), 2014.
- [40] P. Reynaud-Bouret, V. Rivoirard, and C. Tuleau-Malot. Inference of functional connectivity in neurosciences via Hawkes processes. In *1st IEEE Global Conference on Signal and Information Processing*, 2013, Austin Texas.
- [41] J. P. Romano. Bootstrap and randomization tests of some nonparametric hypotheses. Technical Report 270, Dept. Statistics, Stanford Univ., 1987.
- [42] J. P. Romano. Bootstrap and randomization tests of some nonparametric hypotheses. *Ann. Statist.*, 17(1):141–159, 1989.
- [43] J. P. Romano and M. Wolf. Exact and approximate stepdown methods for multiple

- hypothesis testing. *J. Amer. Statist. Assoc.*, 100(469):94–108, 2005.
- [44] H. Rubin and R. A. Vitale. Asymptotic distribution of symmetric statistics. *Ann. Statist.*, pages 165–170, 1980.
- [45] L. Sansonnet and C. Tuleau-Malot. A model of Poissonian interactions and detection of dependence. *to appear in Stat. Comput.*, 2014.
- [46] H. Scheffe. Statistical inference in the non-parametric case. *Ann. Math. Statist.*, 14(4):305–332, 1943.
- [47] R. L. Schilling. *Measures, integrals and martingales*. Cambridge Univ. Press, 2005.
- [48] W. Singer. Synchronization of cortical activity and its putative role in information processing and learning. *Annu. Rev. Physio.*, 55:349–374, 1993.
- [49] C. Tuleau-Malot, A. Rouis, F. Grammont, and P. Reynaud-Bouret. Multiple tests based on a Gaussian approximation of the Unitary Events method. *Neural Computation*, 26(7), 2014.
- [50] A. W. Van der Vaart. *Asymptotic statistics*. Cambridge series in Statistical and Probabilistic Mathematics, 1998.
- [51] A. W. Van der Vaart and J. A. Wellner. *Weak Convergence and Empirical Processes*. Springer, New York, 1996.
- [52] V. S. Varadarajan. On the convergence of sample probability distributions. *Sankhyà*, 19(1–2):23–26, 1958.
- [53] J. Wolfowitz. Additive partition functions and a class of statistical hypotheses. *Ann. Math. Statist.*, 13(3):247–279, 2042.

SUPPLEMENTARY MATERIAL

This Supplement consists of the proofs of Propositions 3.1, 3.2, 3.3, 3.4, 3.5, 4.1, 4.3, of Theorem 4.2, and of the complete proof of Theorem 4.1. It also contains some additional results concerning non-degeneracy and the empirical centring assumption.

APPENDIX A: COMPLETE PROOFS

A.1. Proof of Proposition 3.1. Consider $w_\delta : [0, 1]^2 \rightarrow \mathbb{R}$ a continuous function. For sake of simplicity, denote $w = w_\delta$ and $\varphi^w = \varphi_\delta^w$. Let us prove that $h = h_{\varphi^w}$ is continuous for the topology induced by d (see (3.3)). Recall that for $x_1 = (x_1^1, x_1^2)$ and $x_2 = (x_2^1, x_2^2)$ in (\mathcal{X}^2) ,

$$h_{\varphi^w}(x_1, x_2) = \varphi^w(x_1^1, x_1^2) + \varphi^w(x_2^1, x_2^2) - \varphi^w(x_1^1, x_2^2) - \varphi^w(x_2^1, x_1^2).$$

and that for z in \mathcal{X} , N_z denotes the counting function defined by

$$N_z(t) = \int \mathbb{1}_{\{u \leq t\}} dN_z(u).$$

The first step is to show that for each i, j in $\{1, 2\}$, the projection defined by

$$p_{i,j} : \left(\begin{array}{ccc} ((\mathcal{X}^2)^2, d) & \longrightarrow & (\mathcal{X}^2, d_{\mathcal{X}^2}) \\ ((x_1^1, x_1^2), (x_2^1, x_2^2)) & \longmapsto & (x_i^1, x_j^2) \end{array} \right),$$

is continuous. Let $\mathbf{x} = ((x_1^2, x_1^2), (x_2^1, x_2^2))$ and $\mathbf{x}' = ((x_1^1, x_1^2), (x_2^1, x_2^2))$ in $(\mathcal{X}^2)^2$. Then,

$$d_{\mathcal{X}^2}(p_{i,j}(\mathbf{x}), p_{i,j}(\mathbf{x}')) = d_{\mathcal{X}^2}((x_i^1, x_j^2), (x_i^1, x_j^2)) \leq d(\mathbf{x}, \mathbf{x}').$$

Hence, $p_{i,j}$ is 1-Lipschitz and therefore continuous.

The second step is to show that if w is continuous on $([0, 1]^2, \|\cdot\|_\infty)$, with $\|(u, v) - (u', v')\|_\infty = \max\{|u - u'|, |v - v'|\}$, then φ^w is also continuous. Let $\varepsilon > 0$. First notice that, w being continuous on the compact set $[0, 1]^2$, w is uniformly continuous. Thus one can find some η in $(0, 1)$ such that, for all $(u, v), (u', v')$ in $[0, 1]^2$,

$$(A.1) \quad \|(u, v) - (u', v')\|_\infty \leq \eta \quad \text{implies} \quad |w(u, v) - w(u', v')| \leq \varepsilon.$$

Consider such η .

Let $\{x_n\}_{n \geq 0}$ be a sequence in \mathcal{X}^2 such that $d_{\mathcal{X}^2}(x_n, x_0) \xrightarrow{n \rightarrow +\infty} 0$ and let us

show that $\varphi^w(x_n) \xrightarrow{n \rightarrow +\infty} \varphi^w(x_0)$. There exists n_0 in \mathbb{N} such that for each $n \geq n_0$, $d_{\mathcal{X}^2}(x_n, x_0) \leq \eta$. Then, for such n , by definition of $d_{\mathcal{X}^2}$, we have that $d_D(N_{x_n^1}, N_{x_0^1}) \leq \eta$ and $d_D(N_{x_n^2}, N_{x_0^2}) \leq \eta$. Thus, by definition of d_D ,

$$\exists \lambda_n^1 \in \Lambda / \begin{cases} \sup_{t \in [0,1]} |\lambda_n^1(t) - t| \leq \eta, & (1-i) \\ \sup_{t \in [0,1]} |N_{x_n^1}(t) - N_{x_0^1}(\lambda_n^1(t))| \leq \eta, & (1-ii) \end{cases}$$

$$\exists \lambda_n^2 \in \Lambda / \begin{cases} \sup_{t \in [0,1]} |\lambda_n^2(t) - t| \leq \eta, & (2-i) \\ \sup_{t \in [0,1]} |N_{x_n^2}(t) - N_{x_0^2}(\lambda_n^2(t))| \leq \eta. & (2-ii) \end{cases}$$

In particular, as η is chosen strictly smaller than 1 and as the $N_{x_n^j}$ for $n \geq 0$ and $j = 1, 2$ are counting processes with values in \mathbb{N} , (1-ii) implies that $\forall t \in [0, 1]$, $N_{x_n^1}(t) = N_{x_0^1}(\lambda_n^1(t))$ and thus,

$$u_0 \in x_0^1 \Leftrightarrow u_n = \lambda_n^1(u_0) \in x_n^1.$$

Similarly, (2-ii) implies that

$$v_0 \in x_0^2 \Leftrightarrow v_n = \lambda_n^2(v_0) \in x_n^2.$$

Therefore,

$$\begin{aligned} \varphi^w(x_n) &= \iint_{[0,1]^2} w(u, v) dN_{x_n^1}(u) dN_{x_n^2}(v) \\ &= \sum_{(u_n, v_n) \in x_n^1 \times x_n^2} w(u_n, v_n) \\ &= \sum_{(u_0, v_0) \in x_0^1 \times x_0^2} w(\lambda_n^1(u_0), \lambda_n^2(v_0)). \end{aligned}$$

Hence,

$$|\varphi^w(x_n) - \varphi^w(x_0)| \leq \sum_{(u_0, v_0) \in x_0^1 \times x_0^2} |w(\lambda_n^1(u_0), \lambda_n^2(v_0)) - w(u_0, v_0)|.$$

Yet, by (1-i) and (2-i), for each (u_0, v_0) in $x_0^1 \times x_0^2$, we have $\|(\lambda_n^1(u_0), \lambda_n^2(v_0)) - (u_0, v_0)\|_\infty \leq \eta$, and thus, applying (A.1), we obtain

$$|\varphi^w(x_n) - \varphi^w(x_0)| \leq \#x_0^1 \#x_0^2 \varepsilon,$$

and this for all $n \geq n_0$, which ends the proof of Proposition 3.1.

A.2. Proof of Proposition 3.2. Let us prove that $h = h_{\varphi_\delta^{coinc}}$ is continuous for the topology induced by the metric d (introduced in (3.3)) in any (x_0, y_0) in \mathcal{C}_δ satisfying

$$(\{x_0^1\} \cup \{y_0^1\}) \cap (\{x_0^2 \pm \delta\} \cup \{y_0^2 \pm \delta\}) = \emptyset.$$

Consider a sequence $\{(x_n, y_n)\}_{n \in \mathbb{N}}$ of elements of $\mathcal{X}^2 \times \mathcal{X}^2$, where $x_n = (x_n^1, x_n^2)$ and $y_n = (y_n^1, y_n^2)$ such that $d((x_n, y_n), (x_0, y_0)) \xrightarrow{n \rightarrow +\infty} 0$ and (x_0, y_0) belongs to \mathcal{C}_δ .

We want to show that $|h(x_n, y_n) - h(x_0, y_0)| \xrightarrow{n \rightarrow +\infty} 0$.

Since (x_0, y_0) is in \mathcal{C}_δ , for any t_0 in $\{x_0^2 \pm \delta\} \cup \{y_0^2 \pm \delta\}$, $t_0 \notin x_0^1$, which means that $N_{x_0^1}$ is continuous in t_0 and therefore constant in a neighborhood:

$$\exists \eta_{t_0} > 0 / \forall t \in [0, 1], \quad |t - t_0| \leq \eta_{t_0} \text{ implies } N_{x_0^1}(t) = N_{x_0^1}(t_0).$$

As $\{x_0^2 \pm \delta\} \cup \{y_0^2 \pm \delta\}$ is finite, $\eta_{x_0^1} = \min_{t_0 \in \{x_0^2 \pm \delta\} \cup \{y_0^2 \pm \delta\}} \eta_{t_0} > 0$ is well defined, and satisfies

$$\forall u \in \{x_0^2 \pm \delta\} \cup \{y_0^2 \pm \delta\}, \forall t \in [0, 1], \quad |t - u| \leq \eta_{x_0^1} \text{ implies } N_{x_0^1}(t) = N_{x_0^1}(u).$$

By the same argument using continuity of $N_{y_0^1}$ over $\{x_0^2 \pm \delta\} \cup \{y_0^2 \pm \delta\}$, one can find $\eta_{y_0^1} > 0$ such that

$$\forall u \in \{x_0^2 \pm \delta\} \cup \{y_0^2 \pm \delta\}, \forall t \in [0, 1], \quad |t - u| \leq \eta_{y_0^1} \text{ implies } N_{y_0^1}(t) = N_{y_0^1}(u).$$

Since $(x_0, y_0) \in \mathcal{C}_\delta \Leftrightarrow (\{x_0^2\} \cup \{y_0^2\}) \cap (\{x_0^1 \pm \delta\} \cup \{y_0^1 \pm \delta\}) = \emptyset$, one can construct $\eta_{x_0^2}$ and $\eta_{y_0^2}$ satisfying

$$\forall u \in \{x_0^1 \pm \delta\} \cup \{y_0^1 \pm \delta\}, \forall t \in [0, 1], \begin{cases} |t - u| \leq \eta_{x_0^2} \text{ implies } N_{x_0^2}(t) = N_{x_0^2}(u), \\ |t - u| \leq \eta_{y_0^2} \text{ implies } N_{y_0^2}(t) = N_{y_0^2}(u). \end{cases}$$

Finally, if $\eta = \min \{\eta_{x_0^1}, \eta_{y_0^1}, \eta_{x_0^2}, \eta_{y_0^2}\} > 0$,

$$(A.2) \quad \forall s \in \{x_0^2 \pm \delta\} \cup \{y_0^2 \pm \delta\}, \forall t \in [0, 1], |t - s| \leq \eta \text{ implies } \begin{cases} N_{x_0^1}(t) = N_{x_0^1}(s), \\ N_{y_0^1}(t) = N_{y_0^1}(s), \end{cases}$$

(A.3)

$$\forall s \in \{x_0^1 \pm \delta\} \cup \{y_0^1 \pm \delta\}, \forall t \in [0, 1], |t - s| \leq \eta \text{ implies } \begin{cases} N_{x_0^2}(t) = N_{x_0^2}(s), \\ N_{y_0^2}(t) = N_{y_0^2}(s). \end{cases}$$

As $d((x_n, y_n), (x_0, y_0)) \rightarrow_{n \rightarrow +\infty} 0$, there exists $n_0 \geq 0$ such that for $n \geq n_0$, $d((x_n, y_n), (x_0, y_0)) \leq \eta/4$. From the definition of d , we deduce that

$$\exists \lambda_n^1 \in \Lambda / \begin{cases} \sup_{t \in [0,1]} |\lambda_n^1(t) - t| \leq \frac{\eta}{4}, & (1-i) \\ \sup_{t \in [0,1]} |N_{x_n^1}(t) - N_{x_0^1}(\lambda_n^1(t))| \leq \frac{\eta}{4}, & (1-ii) \end{cases}$$

$$\exists \lambda_n^2 \in \Lambda / \begin{cases} \sup_{t \in [0,1]} |\lambda_n^2(t) - t| \leq \frac{\eta}{4}, & (2-i) \\ \sup_{t \in [0,1]} |N_{x_n^2}(t) - N_{x_0^2}(\lambda_n^2(t))| \leq \frac{\eta}{4}. & (2-ii) \end{cases}$$

Notice that similar results occur for y_n and y_0 , but there are not detailed here since we do not use them explicitly.

By definition of h ,

(A.4)

$$\begin{aligned} & h(x_n, y_n) - h(x_0, y_0) \\ &= \frac{1}{2} \iint \mathbb{1}_{\{|u-v| \leq \delta\}} \{dN_{x_n^1} dN_{x_n^2} + dN_{y_n^1} dN_{y_n^2} - dN_{x_n^1} dN_{y_n^2} - dN_{y_n^1} dN_{x_n^2}\} (u, v) \\ &\quad - \frac{1}{2} \iint \mathbb{1}_{\{|u-v| \leq \delta\}} \{dN_{x_0^1} dN_{x_0^2} + dN_{y_0^1} dN_{y_0^2} - dN_{x_0^1} dN_{y_0^2} - dN_{y_0^1} dN_{x_0^2}\} (u, v) \\ &= \frac{1}{2} \iint \mathbb{1}_{\{|u-v| \leq \delta\}} \left(dN_{x_n^1}(u) \left(dN_{x_n^2} - dN_{x_0^2} \right)(v) + dN_{y_n^1}(u) \left(dN_{y_n^2} - dN_{y_0^2} \right)(v) \right. \\ &\quad \left. - dN_{x_n^1}(u) \left(dN_{y_n^2} - dN_{y_0^2} \right)(v) - dN_{y_n^1}(u) \left(dN_{x_n^2} - dN_{x_0^2} \right)(v) \right. \\ &\quad \left. + \left(dN_{x_n^1} - dN_{x_0^1} \right)(u) dN_{x_0^2}(v) + \left(dN_{y_n^1} - dN_{y_0^1} \right)(u) dN_{y_0^2}(v) \right. \\ &\quad \left. - \left(dN_{x_n^1} - dN_{x_0^1} \right)(u) dN_{y_0^2}(v) + \left(dN_{y_n^1} - dN_{y_0^1} \right)(u) dN_{x_0^2}(v) \right). \end{aligned}$$

By symmetry of the problem, we just need to study the terms

$$A_n = \iint \mathbb{1}_{\{|u-v| \leq \delta\}} \left(dN_{x_n^1} - dN_{x_0^1} \right)(u) dN_{x_0^2}(v),$$

and

$$B_n = \iint \mathbb{1}_{\{|u-v| \leq \delta\}} dN_{x_n^1}(u) \left(dN_{x_n^2} - dN_{x_0^2} \right)(v).$$

Study of A_n .

$$\begin{aligned} A_n &= \iint \mathbb{1}_{\{|u-v| \leq \delta\}} \left(dN_{x_n^1} - dN_{x_0^1} \right)(u) dN_{x_0^2}(v) \\ &= \iint \mathbb{1}_{\{u \leq v + \delta\}} \left(dN_{x_n^1} - dN_{x_0^1} \right)(u) dN_{x_0^2}(v) \\ &\quad - \iint \mathbb{1}_{\{u < v - \delta\}} \left(dN_{x_n^1} - dN_{x_0^1} \right)(u) dN_{x_0^2}(v). \end{aligned}$$

We have that

$$\begin{aligned}
 & \left| \iint \mathbf{1}_{\{u \leq v + \delta\}} \left(dN_{x_n^1} - dN_{x_0^1} \right) (u) dN_{x_0^2}(v) \right| \\
 &= \left| \int \left(N_{x_n^1}(v + \delta) - N_{x_0^1}(v + \delta) \right) dN_{x_0^2}(v) \right| \\
 &\leq \sum_{T \in x_0^2} \left| N_{x_n^1}(T + \delta) - N_{x_0^1}(T + \delta) \right| \\
 &\leq \sum_{T \in x_0^2} \left| N_{x_n^1}(T + \delta) - N_{x_0^1}(\lambda_n^1(T + \delta)) \right| \\
 &+ \sum_{T \in x_0^2} \left| N_{x_0^1}(\lambda_n^1(T + \delta)) - N_{x_0^1}(T + \delta) \right|.
 \end{aligned}$$

Now, using the notation $N_{x_i^1}^-(t) = \int \mathbf{1}_{\{u < t\}} dN_{x_i^1}(u)$,

$$\left| \iint \mathbf{1}_{\{u < v - \delta\}} \left(dN_{x_n^1} - dN_{x_0^1} \right) (u) dN_{x_0^2}(v) \right| \leq \sum_{T \in x_0^2} \left| N_{x_n^1}^-(T - \delta) - N_{x_0^1}^-(T - \delta) \right|.$$

Therefore,

$$\begin{aligned}
 \text{(A.5)} \quad |A_n| &\leq \sum_{T \in x_0^2} \left(\left| N_{x_n^1}(T + \delta) - N_{x_0^1}(\lambda_n^1(T + \delta)) \right| \right. \\
 &+ \left| N_{x_0^1}(\lambda_n^1(T + \delta)) - N_{x_0^1}(T + \delta) \right| \\
 &\left. + \left| N_{x_n^1}^-(T - \delta) - N_{x_0^1}^-(T - \delta) \right| \right).
 \end{aligned}$$

Let us study individually each term in the sum.

Fix T in x_0^2 . By (1-ii),

$$\text{(A.6)} \quad \left| N_{x_n^1}(T + \delta) - N_{x_0^1}(\lambda_n^1(T + \delta)) \right| \leq \frac{\eta}{4} \leq \varepsilon.$$

From (1-i), one has $|\lambda_n^1(T + \delta) - (T + \delta)| \leq \frac{\eta}{2} \leq \eta$ which, with (A.2), implies

$$\text{(A.7)} \quad \left| N_{x_0^1}(\lambda_n^1(T + \delta)) - N_{x_0^1}(T + \delta) \right| = 0.$$

As $N_{x_n^1}^-(T - \delta) = \lim_{\substack{u \rightarrow T - \delta \\ u < T - \delta}} N_{x_n^1}(u)$, there exists u_T in $[T - \delta - \eta/4, T - \delta[$ such

that

$$\left| N_{x_n^1}^-(T - \delta) - N_{x_n^1}(u_T) \right| \leq \varepsilon,$$

so

$$(A.8) \left| N_{x_n^1}^-(T-\delta) - N_{x_0^1}^-(T-\delta) \right| \leq \varepsilon + \left| N_{x_n^1}(u_T) - N_{x_0^1}(\lambda_n^1(u_T)) \right| \\ + \left| N_{x_0^1}(\lambda_n^1(u_T)) - N_{x_0^1}^-(T-\delta) \right|.$$

From (1-ii), one has $\left| N_{x_n^1}(u_T) - N_{x_0^1}(\lambda_n^1(u_T)) \right| \leq \frac{\eta}{4} \leq \varepsilon$.

Then, by continuity of $N_{x_0^1}$ in $T-\delta$, first remark that $N_{x_0^1}^-(T-\delta) = N_{x_0^1}(T-\delta)$.

Moreover, by (1-i) and construction of u_T ,

$$\left| \lambda_n^1(u_T) - (T-\delta) \right| \leq \left| \lambda_n^1(u_T) - u_T \right| + \left| u_T - (T-\delta) \right| \leq \frac{\eta}{4} + \frac{\eta}{4} < \eta,$$

hence, using (A.2), $\left| N_{x_0^1}(\lambda_n^1(u_T)) - N_{x_0^1}^-(T-\delta) \right| = 0$. So finally, (A.8) gives

$$(A.9) \quad \left| N_{x_n^1}^-(T-\delta) - N_{x_0^1}^-(T-\delta) \right| \leq 2\varepsilon.$$

Combining (A.5), (A.6), (A.7), and (A.9), we obtain that for any $n \geq n_0$:

$$(A.10) \quad |A_n| \leq 3\varepsilon \#x_0^2.$$

Study of B_n . Recall that $B_n = \iint \mathbf{1}_{\{|u-v| \leq \delta\}} dN_{x_n^1}(u) \left(dN_{x_n^2} - dN_{x_0^2} \right) (v)$.

As for A_n , B_n is upper bounded by a sum of several terms, that we study separately.

$$B_n = \sum_{T \in x_n^1} \left(N_{x_n^2}(T+\delta) - N_{x_0^2}(T+\delta) \right) - \sum_{T \in x_n^1} \left(N_{x_n^2}^-(T-\delta) - N_{x_0^2}^-(T-\delta) \right).$$

So

$$(A.11) \quad B_n \leq |B_{n,1}| + |B_{n,2}| + |B_{n,3}| + |B_{n,4}|,$$

with

$$B_{n,1} = \sum_{T \in x_n^1} \left(N_{x_n^2}(T+\delta) - N_{x_0^2}(\lambda_n^2(T+\delta)) \right), \\ B_{n,2} = \sum_{T \in x_n^1} \left(N_{x_0^2}(\lambda_n^2(T+\delta)) - N_{x_0^2}(T+\delta) \right), \\ B_{n,3} = \sum_{T \in x_n^1} \left(N_{x_n^2}^-(T-\delta) - N_{x_0^2}^-(\lambda_n^2(T-\delta)) \right), \\ B_{n,4} = \sum_{T \in x_n^1} \left(N_{x_0^2}^-(\lambda_n^2(T-\delta)) + N_{x_0^2}^-(T-\delta) \right).$$

The control of B_n is quite similar to the one of A_n except that the sums are over T in x_n^1 instead of T in x_0^1 , which prevents us to use (A.3) and (A.2) directly.

Control of $B_{n,1}$. Due to (2-ii), $\left|N_{x_n^2}(T + \delta) - N_{x_0^2}(\lambda_n^2(T + \delta))\right| \leq \varepsilon$, so

$$(A.12) \quad |B_{n,1}| \leq \varepsilon \# x_n^1.$$

Control of $B_{n,2}$. One can easily see that

$$\begin{aligned} B_{n,2} &= \iint (\mathbb{1}_{\{v \leq \lambda_n^2(u+\delta)\}} - \mathbb{1}_{\{v \leq u+\delta\}}) dN_{x_0^2}(v) dN_{x_n^1}(u) \\ &= \iint \left[\left(1 - \mathbb{1}_{\{u < (\lambda_n^2)^{-1}(v) - \delta\}}\right) - \left(1 - \mathbb{1}_{\{u < v - \delta\}}\right) \right] dN_{x_0^2}(v) dN_{x_n^1}(u) \\ &= \sum_{T \in x_0^2} \left(N_{x_n^1}^-(T - \delta) - N_{x_n^1}^-((\lambda_n^2)^{-1}(T) - \delta) \right). \end{aligned}$$

Fix now T in x_0^2 .

$$\begin{aligned} \left| N_{x_n^1}^-(T - \delta) - N_{x_n^1}^-((\lambda_n^2)^{-1}(T) - \delta) \right| &\leq \left| N_{x_n^1}^-(T - \delta) - N_{x_0^1}(T - \delta) \right| \\ &\quad + \left| N_{x_0^1}(T - \delta) - N_{x_n^1}^-((\lambda_n^2)^{-1}(T) - \delta) \right|. \end{aligned}$$

As shown in (A.8), $\left| N_{x_n^1}^-(T - \delta) - N_{x_0^1}(T - \delta) \right| \leq 2\varepsilon$.

Furthermore, take v_T in $\left[(\lambda_n^2)^{-1}(T) - \delta - \eta/4, (\lambda_n^2)^{-1}(T) - \delta \right]$ such that

$$\left| N_{x_n^1}^-((\lambda_n^2)^{-1}(T) - \delta) - N_{x_n^1}(v_T) \right| \leq \varepsilon.$$

So,

$$\begin{aligned} \left| N_{x_0^1}(T - \delta) - N_{x_n^1}^-((\lambda_n^2)^{-1}(T) - \delta) \right| &\leq \varepsilon + \left| N_{x_n^1}(v_T) - N_{x_0^1}(\lambda_n^1(v_T)) \right| \\ &\quad + \left| N_{x_0^1}(\lambda_n^1(v_T)) - N_{x_0^1}(T - \delta) \right|. \end{aligned}$$

By construction of v_T and λ_n^1 (see (1-ii)), $\left| N_{x_n^1}(v_T) - N_{x_0^1}(\lambda_n^1(v_T)) \right| \leq \varepsilon$.

Because of (A.2) which is true as $\left| \lambda_n^1(v_T) - (T - \delta) \right| \leq \left| \lambda_n^1(v_T) - v_T \right| + \left| v_T - (T - \delta) \right| \leq \frac{\eta}{4} + \frac{\eta}{4} < \eta$ by (1-i), $\left| N_{x_0^1}(\lambda_n^1(v_T)) - N_{x_0^1}(T - \delta) \right| = 0$. Hence, $\left| N_{x_0^1}(T - \delta) - N_{x_n^1}^-((\lambda_n^2)^{-1}(T) - \delta) \right| \leq 2\varepsilon$. Finally,

$$\left| N_{x_n^1}^-(T - \delta) - N_{x_n^1}^-((\lambda_n^2)^{-1}(T) - \delta) \right| \leq 4\varepsilon,$$

and

$$(A.13) \quad |B_{n,2}| \leq 4\varepsilon \#x_0^2.$$

Control of $B_{n,3}$. First, for all T in x_n^1 , we find some $\nu_{n,T}$ in $]0, \eta/4]$ such that

$$\forall u \in [T - \delta - \nu_{n,T}, T - \delta[, \quad \left| N_{x_n^2}^-(T - \delta) - N_{x_n^2}(u) \right| \leq \varepsilon.$$

Setting $\nu_n = \min_{T \in x_n^1} \nu_{n,T}$,

$$\begin{aligned} |B_{n,3}| &\leq \sum_{T \in x_n^1} \left| N_{x_n^2}^-(T - \delta) - N_{x_n^2}(T - \delta - \nu_n) \right| \\ &\quad + \sum_{T \in x_n^1} \left| N_{x_n^2}(T - \delta - \nu_n) - N_{x_0^2}(\lambda_n^2(T - \delta - \nu_n)) \right| \\ &\quad + \left| \sum_{T \in x_n^1} \left(N_{x_0^2}(\lambda_n^2(T - \delta - \nu_n)) - N_{x_0^2}^-(\lambda_n^2(T - \delta)) \right) \right|. \end{aligned}$$

For each T in x_n^1 , $\left| N_{x_n^2}^-(T - \delta) - N_{x_n^2}(T - \delta - \nu_n) \right| \leq \varepsilon$ and

$\left| N_{x_n^2}(T - \delta - \nu_n) - N_{x_0^2}(\lambda_n^2(T - \delta - \nu_n)) \right| \leq \varepsilon$ by (2-ii). Therefore,

$$|B_{n,3}| \leq 2\varepsilon \#x_n^1 + \left| \sum_{T \in x_n^1} \left(N_{x_0^2}(\lambda_n^2(T - \delta - \nu_n)) - N_{x_0^2}^-(\lambda_n^2(T - \delta)) \right) \right|.$$

Now,

$$\begin{aligned} &\sum_{T \in x_n^1} \left(N_{x_0^2}(\lambda_n^2(T - \delta - \nu_n)) - N_{x_0^2}^-(\lambda_n^2(T - \delta)) \right) \\ &= \iint \mathbf{1}_{\{v \leq \lambda_n^2(u - \delta - \nu_n)\}} - \mathbf{1}_{\{v < \lambda_n^2(u - \delta)\}} dN_{X_n^1}(u) dN_{X_0^2}(v) \\ &= \sum_{T \in x_0^2} \left(N_{x_n^1} \left((\lambda_n^2)^{-1}(T) + \delta \right) - N_{x_n^1}^- \left((\lambda_n^2)^{-1}(T) + \delta + \nu_n \right) \right). \end{aligned}$$

For each T in x_0^2 ,

$$\begin{aligned} &\left| N_{x_n^1} \left((\lambda_n^2)^{-1}(T) + \delta \right) - N_{x_n^1}^- \left((\lambda_n^2)^{-1}(T) + \delta + \nu_n \right) \right| \\ &\leq \left| N_{x_n^1} \left((\lambda_n^2)^{-1}(T) + \delta \right) - N_{x_0^1} \left(\lambda_n^1 \left((\lambda_n^2)^{-1}(T) + \delta \right) \right) \right| \\ &\quad + \left| N_{x_0^1} \left(\lambda_n^1 \left((\lambda_n^2)^{-1}(T) + \delta \right) \right) - N_{x_0^1} \left(\lambda_n^1 \left((\lambda_n^2)^{-1}(T) + \delta + \nu_n \right) \right) \right| \\ &\quad + \left| N_{x_0^1} \left(\lambda_n^1 \left((\lambda_n^2)^{-1}(T) + \delta + \nu_n \right) \right) - N_{x_n^1}^- \left((\lambda_n^2)^{-1}(T) + \delta + \nu_n \right) \right| \\ &\leq 2\varepsilon + \left| N_{x_0^1} \left(\lambda_n^1 \left((\lambda_n^2)^{-1}(T) + \delta + \nu_n \right) \right) - N_{x_n^1}^- \left((\lambda_n^2)^{-1}(T) + \delta + \nu_n \right) \right|, \end{aligned}$$

where the last line comes from (1-ii), and (A.2).

We now find some w_T in $\left[(\lambda_n^2)^{-1}(T) + \delta + \nu_n - \eta/4, (\lambda_n^2)^{-1}(T) + \delta + \nu_n \right]$ such that

$$\left| N_{x_n^1}^- \left((\lambda_n^2)^{-1}(T) + \delta + \nu_n \right) - N_{x_n^1}(w_T) \right| \leq \varepsilon,$$

so

$$\begin{aligned} & \left| N_{x_0^1} \left(\lambda_n^1 \left((\lambda_n^2)^{-1}(T) + \delta + \nu_n \right) \right) - N_{x_n^1}^- \left((\lambda_n^2)^{-1}(T) + \delta + \nu_n \right) \right| \\ & \leq \left| N_{x_0^1} \left(\lambda_n^1 \left((\lambda_n^2)^{-1}(T) + \delta + \nu_n \right) \right) - N_{x_0^1} \left(\lambda_n^1(w_T) \right) \right| \\ & \quad + \left| N_{x_0^1} \left(\lambda_n^1(w_T) \right) - N_{x_n^1}(w_T) \right| + \varepsilon. \end{aligned}$$

From (1-ii), we deduce that $\left| N_{x_0^1} \left(\lambda_n^1(w_T) \right) - N_{x_n^1}(w_T) \right| \leq \varepsilon$. Due to (A.2), (1-i), and the construction of w_T ,

$$\left| \left(\lambda_n^1 \left((\lambda_n^2)^{-1}(T) + \delta + \nu_n \right) \right) - (T - \delta) \right| \leq \frac{3\eta}{4} < \eta,$$

and

$$\left| (\lambda_n^1(w_T)) - (T - \delta) \right| \leq \left| (\lambda_n^1(w_T) - w_T) \right| + |w_T - (T - \delta)| < \eta,$$

so $\left| N_{x_0^1} \left(\lambda_n^1 \left((\lambda_n^2)^{-1}(T) + \delta + \nu_n \right) \right) - N_{x_0^1} \left(\lambda_n^1(w_T) \right) \right| = 0$. As a consequence,

$$\left| N_{x_n^1} \left((\lambda_n^2)^{-1}(T) + \delta \right) - N_{x_n^1}^- \left((\lambda_n^2)^{-1}(T) + \delta + \nu_n \right) \right| \leq 4\varepsilon,$$

and

$$(A.14) \quad |B_{n,3}| \leq 2\varepsilon \#x_n^1 + 4\varepsilon \#x_0^2.$$

Control of $B_{n,4}$.

$$\begin{aligned} B_{n,4} &= \iint (\mathbb{1}_{\{v < \lambda_n^2(u-\delta)\}} - \mathbb{1}_{\{v < u-\delta\}}) dN_{x_0^2}(v) dN_{x_n^1}(u) \\ &= \sum_{T \in x_0^2} \left(N_{x_n^1}(T + \delta) - N_{x_n^1} \left((\lambda_n^2)^{-1}(T) + \delta \right) \right). \end{aligned}$$

Let us fix T in x_0^2 . We have

$$\begin{aligned} & \left| N_{x_n^1}(T + \delta) - N_{x_n^1} \left((\lambda_n^2)^{-1}(T) + \delta \right) \right| \\ & \leq \left| N_{x_n^1}(T + \delta) - N_{x_0^1} \left(\lambda_n^1(T + \delta) \right) \right| \\ & \quad + \left| N_{x_0^1} \left(\lambda_n^1(T + \delta) \right) - N_{x_0^1} \left(\lambda_n^1 \left((\lambda_n^2)^{-1}(T) + \delta \right) \right) \right| \\ & \quad + \left| N_{x_0^1} \left(\lambda_n^1 \left((\lambda_n^2)^{-1}(T) + \delta \right) \right) - N_{x_n^1} \left((\lambda_n^2)^{-1}(T) + \delta \right) \right|. \end{aligned}$$

The first and the last terms are upper bounded by ε due to (1-ii). Furthermore, since $N_{x_0^1} \left(\lambda_n^1 \left((\lambda_n^2)^{-1} (T) + \delta \right) \right) = N_{x_0^1} (T + \delta) = N_{x_0^1} (\lambda_n^1 (T + \delta))$ by applying (A.2) and using (1-i) and (2-i),

$$\left| N_{x_0^1} (\lambda_n^1 (T + \delta)) - N_{x_0^1} \left(\lambda_n^1 \left((\lambda_n^2)^{-1} (T) + \delta \right) \right) \right| = 0.$$

So finally,

$$(A.15) \quad |B_{n,4}| \leq 2\varepsilon \#x_0^2.$$

Combining (A.11), (A.12), (A.13), (A.14), and (A.15), we can conclude that

$$(A.16) \quad |B_n| \leq 3\varepsilon \#x_n^1 + 10\varepsilon \#x_0^2.$$

We now just remark that $(\#x_n^1)_{n \geq n_0}$ is bounded because it converges to $\#x_0^1$. Indeed, since $\#x_n^1 = N_{x_n^1}(1)$, $\#x_0^1 = N_{x_0^1}(1)$ and for every n , $\lambda_n^1(1) = 1$,

$$\begin{aligned} |\#x_n^1 - \#x_0^1| &= \left| N_{x_n^1}(1) - N_{x_0^1}(1) \right| \\ &= \left| N_{x_n^1}(1) - N_{x_0^1}(\lambda_n^1(1)) \right| \\ &\xrightarrow{n \rightarrow +\infty} 0. \end{aligned}$$

With (A.4), (A.10), and (A.16), this concludes the proof of Proposition 3.2.

A.3. Proof of Proposition 3.3. For some integer $n \geq 2$, let $(Y_{n,i}^*, Y_i)_{1 \leq i \leq n}$ be an i.i.d. sample such that for every $i = 1 \dots n$, $Y_{n,i}^* \sim P_n^1 \otimes P_n^2$, where the marginal empirical measures $P_n^j, j = 1, 2$ are associated to \mathbb{X}_n , $Y_i \sim P^1 \otimes P^2$, and such that, from the definition of Wasserstein's metric d_2 recalled in (3.5),

$$\begin{aligned} d_2^2(\mathcal{L}(\sqrt{n}U_n, P_n^1 \otimes P_n^2 | \mathbb{X}_n), \mathcal{L}(\sqrt{n}U_n, P^1 \otimes P^2)) \\ \leq \frac{1}{n(n-1)^2} \mathbb{E}^* \left[\left(\sum_{i \neq i'} (h(Y_{n,i}^*, Y_{n,i'}^*) - h(Y_i, Y_{i'})) \right)^2 \right], \end{aligned}$$

where the upper bound is finite under (\mathcal{A}_{Mmt}^*) .

Introducing for (i, i', j, j') in $\{1, 2, \dots, n\}^4$, and m in $\{2, 3, 4\}$,

$$\begin{aligned} \mathbb{E}_{(i, i', j, j')} &= \mathbb{E}^* \left[\left(h(Y_{n,i}^*, Y_{n,i'}^*) - h(Y_i, Y_{i'}) \right) \left(h(Y_{n,j}^*, Y_{n,j'}^*) - h(Y_j, Y_{j'}) \right) \right], \\ I_m &= \left\{ (i, i', j, j') \in \{1, 2, \dots, n\}^4; i \neq i', j \neq j', \# \{i, i', j, j'\} = m \right\}, \end{aligned}$$

where $\#\{i, i', j, j'\}$ denotes the number of different elements in $\{i, i', j, j'\}$, one has:

$$\begin{aligned} & \mathbb{E}^* \left[\left(\sum_{i \neq i'} h(Y_{n,i}^*, Y_{n,i'}^*) - h(Y_i, Y_{i'}) \right)^2 \right] \\ &= \sum_{(i, i', j, j') \in I_4} \mathbb{E}_{(i, i', j, j')} + \sum_{(i, i', j, j') \in I_3} \mathbb{E}_{(i, i', j, j')} + \sum_{(i, i', j, j') \in I_2} \mathbb{E}_{(i, i', j, j')}. \end{aligned}$$

Let us now upper bound each term of this sum separately.

If (i, i', j, j') is in I_4 , then by independence,

$$\begin{aligned} \mathbb{E}_{(i, i', j, j')} &= (\mathbb{E}^*[h(Y_{n,i}^*, Y_{n,i'}^*)] - \mathbb{E}[h(Y_i, Y_{i'})]) \times \\ & \quad (\mathbb{E}^*[h(Y_{n,j}^*, Y_{n,j'}^*)] - \mathbb{E}[h(Y_j, Y_{j'})]). \end{aligned}$$

Under (\mathcal{A}_{Cent}) and (\mathcal{A}_{Cent}^*) , $\mathbb{E}[h(Y_i, Y_{i'})] = \mathbb{E}^*[h(Y_{n,i}^*, Y_{n,i'}^*)] = 0$, so

$$\mathbb{E}_{(i, i', j, j')} = 0.$$

If (i, i', j, j') is in I_3 , by the Cauchy-Schwarz inequality,

$$\mathbb{E}_{(i, i', j, j')} \leq \mathbb{E}^* \left[\left(h(Y_{n,1}^*, Y_{n,2}^*) - h(Y_1, Y_2) \right)^2 \right].$$

If (i, i', j, j') is in I_2 , then $\mathbb{E}_{(i, i', j, j')} = \mathbb{E}^* \left[\left(h(Y_{n,1}^*, Y_{n,2}^*) - h(Y_1, Y_2) \right)^2 \right]$ is immediate.

But $\#I_3 = 4n(n-1)(n-2)$ and $\#I_2 = 2n(n-1)$, so

$$\begin{aligned} & d_n^2(\mathcal{L}(\sqrt{n}U_n, P_n^1 \otimes P_n^2 | \mathbb{X}_n), \mathcal{L}(\sqrt{n}U_n, P^1 \otimes P^2)) \\ & \leq 4\mathbb{E}^* \left[\left(h(Y_{n,1}^*, Y_{n,2}^*) - h(Y_1, Y_2) \right)^2 \right]. \end{aligned}$$

Since $(Y_{n,i}^*, Y_i)_{j=1,2}$ may be arbitrarily chosen, Proposition 3.3 follows.

A.4. Proof of Proposition 3.4. Let us first notice that (3.6) is a direct application of the strong law of large numbers for U -statistics, proved by Hoffding [25]. Next, let us notice that if for m in $\{1, \dots, 4\}$, $g_m(X_{i_1}, \dots, X_{i_m})$ denotes the sum $\sum_{(i,j,k,l) \in I_{\{i_1, \dots, i_m\}}} h^2((X_i^1, X_j^2), (X_k^1, X_l^2))$, where $I_{\{i_1, \dots, i_m\}}$ is the set $\{(i, j, k, l) \in \{i_1, \dots, i_m\}^4; \#\{i, j, k, l\} = m\}$, then

$$\frac{1}{n^4} \sum_{i,j,k,l=1}^n h^2((X_i^1, X_j^2), (X_k^1, X_l^2)) = \sum_{m=1}^4 \frac{1}{m!} \left(\frac{1}{n^4} \sum_{\substack{(i_1, \dots, i_m) \in \{1, \dots, n\}^m \\ i_1, \dots, i_m \text{ all different}}} g_m(X_{i_1}, \dots, X_{i_m}) \right).$$

Each of the four terms in the right hand side of the above decomposition being, up to a multiplicative factor, a classical U -statistic, and since under (\mathcal{A}_{Mmt}^*) , $\mathbb{E}[|g_m(X_{i_1}, \dots, X_{i_m})|] < +\infty$, we can now apply the strong law of large numbers for U -statistics again. Therefore P -a.s. in $(X_n)_n$,

$$\frac{1}{n(n-1)\dots(n-m+1)} \sum_{(i_1, \dots, i_m)} g_m(X_{i_1}, \dots, X_{i_m}) \xrightarrow[n \rightarrow +\infty]{} \mathbb{E}[g_m(X_1, \dots, X_m)].$$

In particular, P -a.s. in $(X_n)_n$, $n^{-4} \sum_{(i_1, \dots, i_m)} g_m(X_{i_1}, \dots, X_{i_m})$ converges towards 0 for m in $\{1, 2, 3\}$, and towards $\mathbb{E}[g_4(X_1, X_2, X_3, X_4)]$ for $m = 4$. Finally noticing that $\mathbb{E}[g_4(X_1, X_2, X_3, X_4)] = 4! \mathbb{E}[h^2((X_1^1, X_2^2), (X_3^1, X_4^2))]$ allows to conclude.

A.5. Proof of Proposition 3.5. Let $(X_n)_n$ be a sequence of i.i.d pairs of point processes with distribution $P^1 \otimes P^2$ on \mathcal{X}^2 . According to (\mathcal{A}_{Cent}) , for $i \neq j$, $\mathbb{E}[h(X_i, X_j)] = 0$. For a better readability, we set $\mathbb{E}[h|X_i] = \mathbb{E}[h(X_i, X)|X_i] = \mathbb{E}[h(X, X_i)|X_i]$ for some X with distribution $P^1 \otimes P^2$, and independent of X_i . By Hoeffding's decomposition for non-degenerate U -statistics, we obtain that

$$\sqrt{n}U_{n,h}(\mathbf{X}_n) = \frac{2}{\sqrt{n}(n-1)} (T_n + M_n),$$

where $T_n = \sum_{i < j} (\mathbb{E}[h|X_i] + \mathbb{E}[h|X_j])$, and $M_n = \sum_{i < j} g(X_i, X_j)$, with $g(X_i, X_j) = h(X_i, X_j) - \mathbb{E}[h|X_i] - \mathbb{E}[h|X_j]$.

Firstly, we have that $\mathbb{E}[M_n^2] = \sum_{i < j} \sum_{k < l} \mathbb{E}[g(X_i, X_j)g(X_k, X_l)]$. But if $\{i, j\} \cap \{k, l\} = \emptyset$, $i < j$, $k < l$, $\mathbb{E}[g(X_i, X_j)g(X_k, X_l)] = (\mathbb{E}[g(X_i, X_j)])^2 = 0$. If $\#\{i, j\} \cap \{k, l\} = 1$, with for instance $k = i$, $j \neq l$, $(i < j, i < l)$ (the other cases may be treated similarly), then

$$\mathbb{E}[g(X_i, X_j)g(X_i, X_l)] = \mathbb{E}[\mathbb{E}[g(X_i, X_j)|X_i] \mathbb{E}[g(X_i, X_l)|X_i]] = 0.$$

Therefore, $\mathbb{E}[M_n^2] = \sum_{i < j} \mathbb{E}[g^2(X_i, X_j)] = n(n-1)\mathbb{E}[g^2(X_1, X_2)]/2$, and since $\mathbb{E}[g^2(X_i, X_j)] < +\infty$, from Chebychev's inequality, we deduce that

$$(A.17) \quad \frac{2}{\sqrt{n}(n-1)} M_n \xrightarrow[n \rightarrow +\infty]{\mathbb{P}} 0.$$

Secondly, we have that $T_n = (n-1) \sum_{i=1}^n \mathbb{E}[h|X_i]$. Since the $\mathbb{E}[h|X_i]$'s are i.i.d, with $\mathbb{E}[\mathbb{E}[h|X_i]] = 0$ and $\text{Var}(\mathbb{E}[h|X_i]) = \sigma_{h, P^1 \otimes P^2}^2/4$, thanks to (\mathcal{A}_{Mmt}) , the Central Limit Theorem leads to

$$(A.18) \quad \frac{2}{\sqrt{n}(n-1)} T_n \xrightarrow[n \rightarrow +\infty]{\mathcal{L}} \mathcal{N}\left(0, \sigma_{h, P^1 \otimes P^2}^2\right).$$

Thus, combining (A.17) and (A.18), Slutsky's lemma ensures the convergence in distribution of $\sqrt{n}U_{n,h}(\mathbb{X}_n)$ towards $\mathcal{N}\left(0, \sigma_{h, P^1 \otimes P^2}^2\right)$.

Now, notice that

$$\mathbb{E}\left[(\sqrt{n}U_{n,h}(\mathbb{X}_n))^2\right] = \frac{1}{n(n-1)^2} \sum_{i \neq i'} \sum_{j \neq j'} \mathbb{E}\left[h(X_i, X_{i'})h(X_j, X_{j'})\right].$$

Let us consider all the cases where $i \neq i'$ and $j \neq j'$.

If $\#\{i, i', j, j'\} = 4$, $\mathbb{E}\left[h(X_i, X_{i'})h(X_j, X_{j'})\right] = 0$, by independence and (\mathcal{A}_{Cent}) .

If $\#\{i, i', j, j'\} = 3$, $\mathbb{E}\left[h(X_i, X_{i'})h(X_j, X_{j'})\right] = \sigma_{h, P^1 \otimes P^2}^2/4$, by symmetry of h .

If $\#\{i, i', j, j'\} = 2$, $\mathbb{E}\left[h(X_i, X_{i'})h(X_j, X_{j'})\right] = \mathbb{E}\left[(h(X_1, X_2))^2\right]$. Therefore,

$$\mathbb{E}\left[(\sqrt{n}U_{n,h}(\mathbb{X}_n))^2\right] = \frac{n-2}{n-1}\sigma_{h, P^1 \otimes P^2}^2 + \frac{2}{n-1}\mathbb{E}\left[(h(X_1, X_2))^2\right] \xrightarrow{n \rightarrow +\infty} \sigma_{h, P^1 \otimes P^2}^2,$$

which ends the proof of Proposition 3.5.

A.6. Proof of Proposition 4.1. As the σ -algebra considered on $(\mathcal{X}^2)^{\otimes n}$ is generated by the family (stable by intersection) containing all products $\prod_{i=1}^n (A_i^1 \times A_i^2)$, with A_i^1 and A_i^2 Borel sets of $(\mathcal{X}, d_{\mathcal{X}})$, we just need to verify that for all $A = \prod_{i=1}^n (A_i^1 \times A_i^2)$ of this form, $\mathbb{P}(\mathbb{X}_n^{\pi_n} \in A) = \mathbb{P}(\mathbb{X}_n \in A)$, and $\mathbb{P}(\mathbb{X}_n^{\Pi_n} \in A) = \mathbb{P}(\mathbb{X}_n \in A)$. We first have that under (H_0) ,

$$\begin{aligned} \mathbb{P}(\mathbb{X}_n^{\pi_n} \in A) &= \mathbb{P}\left(\bigcap_{i=1}^n \{X_i \in A_i^1 \times A_{\pi_n^{-1}(i)}^2\}\right) \\ &= \prod_{i=1}^n \left[P^1(A_i^1) P^2(A_{\pi_n^{-1}(i)}^2)\right] \\ &= \prod_{i=1}^n P^1(A_i^1) \prod_{k=1}^n P^2(A_k^2) \\ &= \prod_{i=1}^n [P^1(A_i^1) P^2(A_i^2)] \\ &= \mathbb{P}\left(\bigcap_{i=1}^n \{X_i \in A_i^1 \times A_i^2\}\right) \\ &= \mathbb{P}(\mathbb{X}_n \in A). \end{aligned}$$

Now consider a random permutation Π_n , uniformly distributed on \mathfrak{S}_n , and independent of \mathbb{X}_n . Then

$$\begin{aligned} \mathbb{P}(\mathbb{X}_n^{\Pi_n} \in A) &= \sum_{\pi_n \in \mathfrak{S}_n} \mathbb{P}(\mathbb{X}_n^{\pi_n} \in A) \mathbb{P}(\Pi_n = \pi_n) \\ &= \frac{1}{n!} \sum_{\pi_n \in \mathfrak{S}_n} \mathbb{P}(\mathbb{X}_n \in A) \\ &= \mathbb{P}(\mathbb{X}_n \in A), \end{aligned}$$

which ends the proof.

A.7. Complete proof of Theorem 4.1. Recall that d_{BL} denotes the bounded Lipschitz metric (which metrizes the weak convergence), defined by

$$d_{BL}(\mu, \nu) = \sup_{f \in BL, \|f\|_{BL} \leq 1} \left| \int_{\mathbb{R}} f \, (d\mu - d\nu) \right|,$$

where, as defined in [13], BL is the set of bounded Lipschitz function on \mathbb{R} , and $\|f\|_{BL} = \|f\|_{\infty} + \sup_{x \neq y} \frac{|f(x) - f(y)|}{|x - y|}$.

Recall that the proof consists of three steps presented in Section 6.5. We give below a complete proof for each of these steps.

First step: decomposition of $\sqrt{n}U_{n, h_\varphi}(\mathbb{X}_n^{\Pi_n})$. It is obvious that by the definition (2.4) of h_φ ,

$$(A.19) \quad U_{n, h_\varphi}(\mathbb{X}_n^{\Pi_n}) = \frac{1}{n-1} U_n^{\Pi_n},$$

where $U_n^{\Pi_n} = \sum_{i=1}^n \varphi(X_i^1, X_{\Pi_n(i)}^2) - \frac{1}{n} \sum_{i,j=1}^n \varphi(X_i^1, X_j^2)$, so,

$$\begin{aligned} U_n^{\Pi_n} &= \sum_{i=1}^n \varphi(X_i^1, X_{\Pi_n(i)}^2) - \frac{1}{n} \sum_{i,j=1}^n \mathbb{E}[\varphi(X_i^1, X_j^2) | X_i^1] \\ &\quad - \frac{1}{n} \sum_{i,j=1}^n \mathbb{E}[\varphi(X_i^1, X_j^2) | X_j^2] + \frac{1}{n} \sum_{i,j=1}^n \mathbb{E}[\varphi(X_i^1, X_j^2)] \\ &\quad - \frac{1}{n} \sum_{i,j=1}^n (\varphi(X_i^1, X_j^2) - \mathbb{E}[\varphi(X_i^1, X_j^2) | X_i^1] \\ &\quad - \mathbb{E}[\varphi(X_i^1, X_j^2) | X_j^2] + \mathbb{E}[\varphi(X_i^1, X_j^2)]). \end{aligned}$$

On the one hand, if $\mathbb{E}_P [f]$ and $\mathbb{E}_{P^1 \otimes P^2} [f]$ respectively denote $\mathbb{E} [f (X_1^1, X_1^2)]$, and $\mathbb{E} [f (X_1^1, X_2^2)]$, for any integrable function f , then

$$\begin{aligned} & \frac{1}{n} \sum_{i,j=1}^n \mathbb{E} [\varphi (X_i^1, X_j^2)] \\ &= \sum_{i,j=1}^n \mathbb{1}_{\{\Pi_n(i)=j\}} \mathbb{E} [\varphi (X_i^1, X_j^2)] - \sum_{i,j=1}^n \left(\mathbb{1}_{\{\Pi_n(i)=j\}} - \frac{1}{n} \right) \mathbb{E} [\varphi (X_i^1, X_j^2)] \\ &= \sum_{i,j=1}^n \mathbb{1}_{\{\Pi_n(i)=j\}} \mathbb{E} [\varphi (X_i^1, X_j^2)] - (\mathbb{E}_P [\varphi] - \mathbb{E}_{P^1 \otimes P^2} [\varphi]) \sum_{i=1}^n \left(\mathbb{1}_{\{\Pi_n(i)=i\}} - \frac{1}{n} \right). \end{aligned}$$

On the other hand,

$$\begin{aligned} \frac{1}{n} \sum_{i,j=1}^n \mathbb{E} [\varphi (X_i^1, X_j^2) | X_i^1] &= \sum_{i,j=1}^n \mathbb{1}_{\{\Pi_n(i)=j\}} \mathbb{E} [\varphi (X_i^1, X_j^2) | X_i^1] \\ &\quad - \sum_{i=1}^n \left(\mathbb{1}_{\{\Pi_n(i)=i\}} - \frac{1}{n} \right) (\mathbb{E} [\varphi (X_i^1, X_i^2) | X_i^1] - \mathbb{E} [\varphi (X_i^1, X^2) | X_i^1]), \end{aligned}$$

where $X = (X^1, X^2)$ is assumed to be P -distributed and independent of $(X_n)_n$, and in the same way,

$$\begin{aligned} \frac{1}{n} \sum_{i,j=1}^n \mathbb{E} [\varphi (X_i^1, X_j^2) | X_j^2] &= \sum_{i,j=1}^n \mathbb{1}_{\{\Pi_n(i)=j\}} \mathbb{E} [\varphi (X_i^1, X_j^2) | X_j^2] \\ &\quad - \sum_{j=1}^n \left(\mathbb{1}_{\{\Pi_n(j)=j\}} - \frac{1}{n} \right) (\mathbb{E} [\varphi (X_j^1, X_j^2) | X_j^2] - \mathbb{E} [\varphi (X^1, X_j^2) | X_j^2]). \end{aligned}$$

Therefore, $U_n^{\Pi_n}$ is equal to

$$\begin{aligned} & \sum_{i,j=1}^n \mathbb{1}_{\{\Pi_n(i)=j\}} \left(\varphi (X_i^1, X_j^2) - \mathbb{E} [\varphi (X_i^1, X_j^2) | X_i^1] \right. \\ & \quad \left. - \mathbb{E} [\varphi (X_i^1, X_j^2) | X_j^2] + \mathbb{E} [\varphi (X_i^1, X_j^2)] \right) \\ & + \sum_{i=1}^n \left(\mathbb{1}_{\{\Pi_n(i)=i\}} - \frac{1}{n} \right) \left(\mathbb{E} [\varphi (X_i^1, X_i^2) | X_i^1] + \mathbb{E} [\varphi (X_i^1, X_i^2) | X_i^2] \right. \\ & \quad \left. - \mathbb{E} [\varphi (X_i^1, X^2) | X_i^1] - \mathbb{E} [\varphi (X^1, X_i^2) | X_i^2] - \mathbb{E}_P [\varphi] + \mathbb{E}_{P^1 \otimes P^2} [\varphi] \right) \\ & - \frac{1}{n} \sum_{i,j=1}^n \left(\varphi (X_i^1, X_j^2) - \mathbb{E} [\varphi (X_i^1, X_j^2) | X_i^1] - \mathbb{E} [\varphi (X_i^1, X_j^2) | X_j^2] + \mathbb{E} [\varphi (X_i^1, X_j^2)] \right). \end{aligned}$$

As a consequence, setting

$$C_{i,j} = \varphi(X_i^1, X_j^2) - \mathbb{E}[\varphi(X_i^1, X^2) | X_i^1] - \mathbb{E}[\varphi(X^1, X_j^2) | X_j^2] + \mathbb{E}_{P^1 \otimes P^2}[\varphi],$$

where $X = (X^1, X^2)$ with distribution P is independent of \mathbb{X}_n ,

$$(A.20) \quad \sqrt{n}U_{n,h_\varphi}(\mathbb{X}_n^{\Pi_n}) = \frac{n}{n-1} \left(M_n^{\Pi_n}(\mathbb{X}_n) + \frac{R_n^{\Pi_n}(\mathbb{X}_n)}{\sqrt{n}} - \frac{T_n(\mathbb{X}_n)}{\sqrt{n}} \right),$$

with

$$\begin{aligned} M_n^{\Pi_n}(\mathbb{X}_n) &= \frac{1}{\sqrt{n}} \sum_{i \neq j} \mathbb{1}_{\{\Pi_n(i)=j\}} C_{i,j}, \\ R_n^{\Pi_n}(\mathbb{X}_n) &= \sum_{i=1}^n \left(\mathbb{1}_{\{\Pi_n(i)=i\}} - \frac{1}{n} \right) C_{i,i}, \\ T_n(\mathbb{X}_n) &= \frac{1}{n} \sum_{i \neq j} C_{i,j}. \end{aligned}$$

Let us now prove that

$$(A.21) \quad d_{BL} \left(\mathcal{L}(\sqrt{n}U_{n,h_\varphi}(\mathbb{X}_n^{\Pi_n}) | \mathbb{X}_n), \mathcal{L} \left(\frac{n}{n-1} M_n^{\Pi_n}(\mathbb{X}_n) \middle| \mathbb{X}_n \right) \right) \xrightarrow[n \rightarrow +\infty]{\mathbb{P}} 0.$$

To do this, first notice that for every function f in BL such that $\|f\|_{BL} \leq 1$,

$$\begin{aligned} \left| \mathbb{E} \left[f(\sqrt{n}U_{n,h_\varphi}(\mathbb{X}_n^{\Pi_n})) \middle| \mathbb{X}_n \right] - \mathbb{E} \left[f \left(\frac{n}{n-1} M_n^{\Pi_n}(\mathbb{X}_n) \right) \middle| \mathbb{X}_n \right] \right| \\ \leq \mathbb{E} \left[\left| \sqrt{n}U_{n,h_\varphi}(\mathbb{X}_n^{\Pi_n}) - \frac{n}{n-1} M_n^{\Pi_n}(\mathbb{X}_n) \right| \middle| \mathbb{X}_n \right] \\ \leq \frac{n}{n-1} \left(\mathbb{E} \left[\left| \frac{R_n^{\Pi_n}(\mathbb{X}_n)}{\sqrt{n}} \right| \middle| \mathbb{X}_n \right] + \frac{|T_n(\mathbb{X}_n)|}{\sqrt{n}} \right). \end{aligned}$$

Hence, taking the supremum over $\{f \in BL; \|f\|_{BL} \leq 1\}$,

$$(A.22) \quad d_{BL} \left(\mathcal{L}(\sqrt{n}U_{n,h_\varphi}(\mathbb{X}_n^{\Pi_n}) | \mathbb{X}_n), \mathcal{L} \left(\frac{n}{n-1} M_n^{\Pi_n}(\mathbb{X}_n) \middle| \mathbb{X}_n \right) \right) \\ \leq \frac{n}{n-1} \left(\mathbb{E} \left[\left| \frac{R_n^{\Pi_n}(\mathbb{X}_n)}{\sqrt{n}} \right| \middle| \mathbb{X}_n \right] + \frac{|T_n(\mathbb{X}_n)|}{\sqrt{n}} \right).$$

Moreover, on the one hand, since Π_n is independent of $(X_n)_n$, by Cauchy-Schwarz inequality,

$$\mathbb{E} \left[\left(\mathbb{E} \left[\frac{|R_n^{\Pi_n}(\mathbb{X}_n)|}{\sqrt{n}} \middle| \mathbb{X}_n \right] \right)^2 \right] \leq \frac{1}{n} \mathbb{E} \left[(R_n^{\Pi_n}(\mathbb{X}_n))^2 \right],$$

and

$$\begin{aligned} & \mathbb{E} \left[(R_n^{\Pi_n}(\mathbb{X}_n))^2 \right] \\ & \leq \sum_{i,j=1}^n \mathbb{E} \left[\left(\mathbb{1}_{\{\Pi_n(i)=i\}} - \frac{1}{n} \right) \left(\mathbb{1}_{\{\Pi_n(j)=j\}} - \frac{1}{n} \right) \right] \mathbb{E} [C_{i,i} C_{j,j}] \\ & \leq C (\mathbb{E}_P [\varphi^2] + \mathbb{E}_{P^1 \otimes P^2} [\varphi^2]) \sum_{i,j=1}^n \left(\mathbb{E} [\mathbb{1}_{\{\Pi_n(i)=i\}} \mathbb{1}_{\{\Pi_n(j)=j\}}] - \frac{1}{n^2} \right) \\ & \leq C (\mathbb{E}_P [\varphi^2] + \mathbb{E}_{P^1 \otimes P^2} [\varphi^2]) \left(\sum_{i=1}^n \left(\frac{1}{n} - \frac{1}{n^2} \right) + \sum_{i \neq j} \left(\frac{1}{n(n-1)} - \frac{1}{n^2} \right) \right) \\ & \leq C (\mathbb{E}_P [\varphi^2] + \mathbb{E}_{P^1 \otimes P^2} [\varphi^2]) < +\infty. \end{aligned}$$

Therefore, from Markov's inequality, we deduce that

$$\mathbb{E} \left[\frac{|R_n^{\Pi_n}(\mathbb{X}_n)|}{\sqrt{n}} \middle| \mathbb{X}_n \right] \xrightarrow[n \rightarrow +\infty]{\mathbb{P}} 0.$$

On the other hand,

$$\mathbb{E} \left[\left(\frac{T_n(\mathbb{X}_n)}{\sqrt{n}} \right)^2 \right] = \frac{1}{n^3} \sum_{i \neq j} \sum_{k \neq l} \mathbb{E} [C_{i,j} C_{k,l}].$$

Notice that for $i \neq j$, $\mathbb{E} [C_{i,j} | X_i] = \mathbb{E} [C_{i,j} | X_j] = 0$.

If $\# \{i, j, k, l\} = 4$, then $\mathbb{E} [C_{i,j} C_{k,l}] = (\mathbb{E} [C_{i,j}])^2 = 0$.

If i, j, l are all different, then

$$\begin{aligned} \mathbb{E} [C_{i,j} C_{i,l}] &= \mathbb{E} [\mathbb{E} [C_{i,j} C_{i,l} | X_i, X_l]] \\ &= \mathbb{E} [\mathbb{E} [C_{i,j} | X_i] C_{i,l}] \\ &= 0. \end{aligned}$$

In the same way, for i, j, k all different, then $\mathbb{E} [C_{i,j} C_{k,i}] = 0$.

If $i \neq j$,

$$(A.23) \quad \mathbb{E} [C_{i,j}^2] = \sigma_{h_\varphi, P^1 \otimes P^2}^2, \text{ and } \mathbb{E} [C_{i,j} C_{j,i}] \leq \sigma_{h_\varphi, P^1 \otimes P^2}^2,$$

by the Cauchy-Schwarz inequality. Combining the above computations, we obtain that

$$\mathbb{E} \left[\left(\frac{T_n(\mathbb{X}_n)}{\sqrt{n}} \right)^2 \right] \leq 2 \frac{n(n-1)}{n^3} \sigma_{h_\varphi, P^1 \otimes P^2}^2 \xrightarrow{n \rightarrow +\infty} 0,$$

and therefore,

$$\frac{T_n(\mathbb{X}_n)}{\sqrt{n}} \xrightarrow[n \rightarrow +\infty]{\mathbb{P}} 0.$$

Finally, from (A.22), we derive (A.21).

Second step: asymptotic normality of $M_n^{\Pi_n}(\mathbb{X}_n)$ given \mathbb{X}_n , in probability.

Recall that

$$\begin{aligned} M_n^{\Pi_n}(\mathbb{X}_n) &= \frac{1}{\sqrt{n}} \sum_{i \neq j} \mathbb{1}_{\{\Pi_n(i)=j\}} C_{i,j} \\ &= \frac{1}{\sqrt{n}} \sum_{i=2}^n \sum_{j=1}^{i-1} (\mathbb{1}_{\{\Pi_n(i)=j\}} C_{i,j} + \mathbb{1}_{\{\Pi_n(j)=i\}} C_{j,i}). \end{aligned}$$

Let Π'_n be another uniformly distributed random permutation with values in \mathfrak{S}_n , independent of Π_n and \mathbb{X}_n , and

$$\begin{aligned} M_n^{\Pi'_n}(\mathbb{X}_n) &= \frac{1}{\sqrt{n}} \sum_{i \neq j} \mathbb{1}_{\{\Pi'_n(i)=j\}} C_{i,j} \\ &= \frac{1}{\sqrt{n}} \sum_{i=2}^n \sum_{j=1}^{i-1} (\mathbb{1}_{\{\Pi'_n(i)=j\}} C_{i,j} + \mathbb{1}_{\{\Pi'_n(j)=i\}} C_{j,i}). \end{aligned}$$

Let us now recall the result of Lemma 6.1:

$$\mathcal{L} \left(\left(M_n^{\Pi_n}(\mathbb{X}_n), M_n^{\Pi'_n}(\mathbb{X}_n) \right)' \right) \xrightarrow[n \rightarrow +\infty]{} \mathcal{N}_2 \left(0, \begin{pmatrix} \sigma_{h_\varphi, P^1 \otimes P^2}^2 & 0 \\ 0 & \sigma_{h_\varphi, P^1 \otimes P^2}^2 \end{pmatrix} \right).$$

Proof of Lemma 6.1. According to the Cramér-Wold device, given a, b in \mathbb{R} , we aim at proving that

$$\mathcal{L} \left(a M_n^{\Pi_n}(\mathbb{X}_n) + b M_n^{\Pi'_n}(\mathbb{X}_n) \right) \xrightarrow[n \rightarrow +\infty]{} \mathcal{N} \left(0, (a^2 + b^2) \sigma_{h_\varphi, P^1 \otimes P^2}^2 \right).$$

In order to deal with simpler mathematical expressions, we introduce below some additional notation.

- For $n \geq i \geq 2$, $\mathcal{F}_{n,i} = \sigma(\Pi_n, \Pi'_n, X_1, X_2, \dots, X_i)$.

• Let

$$Y_{n,i} = \frac{1}{\sqrt{n}} \sum_{j=1}^{i-1} (\mathbb{1}_{\{\Pi_n(i)=j\}} C_{i,j} + \mathbb{1}_{\{\Pi_n(j)=i\}} C_{j,i}),$$

$$Y'_{n,i} = \frac{1}{\sqrt{n}} \sum_{j=1}^{i-1} (\mathbb{1}_{\{\Pi_n(i)=j\}} C_{i,j} + \mathbb{1}_{\{\Pi_n(j)=i\}} C_{j,i}),$$

so that $M_n^{\Pi_n}(\mathbb{X}_n) = \sum_{i=1}^n Y_{n,i}$ and $M_n^{\Pi'_n}(\mathbb{X}_n) = \sum_{i=1}^n Y'_{n,i}$.

Let us first prove that for a fixed integer $n \geq 2$, $(aY_{n,i} + bY'_{n,i}, \mathcal{F}_{n,i})_{2 \leq i \leq n}$ is a martingale difference array. Note that for $2 \leq i \leq n$,

$$\begin{aligned} \mathbb{E}[Y_{n,i} | \mathcal{F}_{n,i-1}] &= \frac{1}{\sqrt{n}} \sum_{j=1}^{i-1} \mathbb{E}[\mathbb{1}_{\{\Pi_n(i)=j\}} C_{i,j} + \mathbb{1}_{\{\Pi_n(j)=i\}} C_{j,i} | \mathcal{F}_{n,i-1}] \\ &= \frac{1}{\sqrt{n}} \sum_{j=1}^{i-1} (\mathbb{1}_{\{\Pi_n(i)=j\}} \mathbb{E}[C_{i,j} | X_j] + \mathbb{1}_{\{\Pi_n(j)=i\}} \mathbb{E}[C_{j,i} | X_j]) \\ &= 0. \end{aligned}$$

In the same way, we have that $\mathbb{E}[Y'_{n,i} | \mathcal{F}_{n,i-1}] = 0$, so $\mathbb{E}[aY_{n,i} + bY'_{n,i} | \mathcal{F}_{n,i-1}] = 0$. From Theorem 6.1, we thus deduce that if

$$\begin{aligned} (i) \quad & \sum_{i=2}^n \mathbb{E}[(aY_{n,i} + bY'_{n,i})^2 | \mathcal{F}_{n,i-1}] \xrightarrow[n \rightarrow +\infty]{\mathbb{P}} (a^2 + b^2) \sigma_{h_\varphi, P^1 \otimes P^2}^2, \\ (ii) \quad & \sum_{i=2}^n \mathbb{E}[(aY_{n,i} + bY'_{n,i})^2 \mathbb{1}_{\{|aY_{n,i} + bY'_{n,i}| > \varepsilon\}}] \xrightarrow[n \rightarrow +\infty]{} 0 \text{ for any } \varepsilon > 0, \end{aligned}$$

then

$$\mathcal{L} \left(aM_n^{\Pi_n}(\mathbb{X}_n) + bM_n^{\Pi'_n}(\mathbb{X}_n) \right) \xrightarrow[n \rightarrow +\infty]{} \mathcal{N} \left(0, (a^2 + b^2) \sigma_{h_\varphi, P^1 \otimes P^2}^2 \right).$$

Let us now check that both (i) and (ii) are satisfied.

Assumption (i). In all the following, only consider $n \geq 4$. Noticing that

$$\begin{aligned} (A.24) \quad & \sum_{i=2}^n \mathbb{E}[(aY_{n,i} + bY'_{n,i})^2 | \mathcal{F}_{n,i-1}] \\ &= (a^2 + b^2) \sum_{i=2}^n \mathbb{E}[Y_{n,i}^2 | \mathcal{F}_{n,i-1}] + 2ab \sum_{i=2}^n \mathbb{E}[Y_{n,i} Y'_{n,i} | \mathcal{F}_{n,i-1}], \end{aligned}$$

the proof of (i) can be decomposed into two points.

The first point consists in proving that

$$\sum_{i=2}^n \mathbb{E} [Y_{n,i}^2] \xrightarrow{n \rightarrow +\infty} \sigma_{h_\varphi, P^1 \otimes P^2}^2 \quad \text{and} \quad \text{Var} \left(\sum_{i=2}^n \mathbb{E} [Y_{n,i}^2 | \mathcal{F}_{n,i-1}] \right) \xrightarrow{n \rightarrow +\infty} 0,$$

which leads, thanks to Chebychev's inequality, to

$$\sum_{i=2}^n \mathbb{E} [Y_{n,i}^2 | \mathcal{F}_{n,i-1}] \xrightarrow[n \rightarrow +\infty]{\mathbb{P}} \sigma_{h_\varphi, P^1 \otimes P^2}^2.$$

The second point consists in proving that

$$\mathbb{E} \left[\left(\sum_{i=2}^n \mathbb{E} [Y_{n,i} Y'_{n,i} | \mathcal{F}_{n,i-1}] \right)^2 \right] \xrightarrow{n \rightarrow +\infty} 0,$$

so

$$\sum_{i=2}^n \mathbb{E} [Y_{n,i} Y'_{n,i} | \mathcal{F}_{n,i-1}] \xrightarrow[n \rightarrow +\infty]{\mathbb{P}} 0.$$

• First point. On the one hand,

$$\begin{aligned} \sum_{i=2}^n \mathbb{E} [Y_{n,i}^2] &= \frac{1}{n} \sum_{i=2}^n \sum_{j,k=1}^{i-1} \mathbb{E} [(\mathbb{1}_{\{\Pi_n(i)=j\}} C_{i,j} + \mathbb{1}_{\{\Pi_n(j)=i\}} C_{j,i}) \times \\ &\quad (\mathbb{1}_{\{\Pi_n(i)=k\}} C_{i,k} + \mathbb{1}_{\{\Pi_n(k)=i\}} C_{k,i})]. \end{aligned}$$

Furthermore, if $1 \leq j \neq k \leq i-1$,

$$\begin{aligned} &\mathbb{E} [(\mathbb{1}_{\{\Pi_n(i)=j\}} C_{i,j} + \mathbb{1}_{\{\Pi_n(j)=i\}} C_{j,i}) (\mathbb{1}_{\{\Pi_n(i)=k\}} C_{i,k} + \mathbb{1}_{\{\Pi_n(k)=i\}} C_{k,i})] \\ &= \mathbb{E} [\mathbb{E} [(\mathbb{1}_{\{\Pi_n(i)=j\}} C_{i,j} + \mathbb{1}_{\{\Pi_n(j)=i\}} C_{j,i}) (\mathbb{1}_{\{\Pi_n(i)=k\}} C_{i,k} + \mathbb{1}_{\{\Pi_n(k)=i\}} C_{k,i}) | X_i, X_j, \Pi_n]] \\ &= \mathbb{E} [(\mathbb{1}_{\{\Pi_n(i)=j\}} C_{i,j} + \mathbb{1}_{\{\Pi_n(j)=i\}} C_{j,i}) (\mathbb{1}_{\{\Pi_n(i)=k\}} \mathbb{E} [C_{i,k} | X_i] + \mathbb{1}_{\{\Pi_n(k)=i\}} \mathbb{E} [C_{k,i} | X_i])] \\ &= 0. \end{aligned}$$

Thus,

$$\begin{aligned}
 \sum_{i=2}^n \mathbb{E} [Y_{n,i}^2] &= \frac{1}{n} \sum_{i=2}^n \sum_{j=1}^{i-1} \mathbb{E} \left[\left(\mathbb{1}_{\{\Pi_n(i)=j\}} C_{i,j} + \mathbb{1}_{\{\Pi_n(j)=i\}} C_{j,i} \right)^2 \right] \\
 &= \frac{1}{n} \sum_{i=2}^n \sum_{j=1}^{i-1} \mathbb{E} \left[\mathbb{1}_{\{\Pi_n(i)=j\}} C_{i,j}^2 + \mathbb{1}_{\{\Pi_n(j)=i\}} C_{j,i}^2 + 2\mathbb{1}_{\{\Pi_n(i)=j\}} \mathbb{1}_{\{\Pi_n(j)=i\}} C_{i,j} C_{j,i} \right] \\
 &= \frac{1}{n} \sum_{i=2}^n \sum_{j=1}^{i-1} \left(\frac{2}{n} \mathbb{E} [C_{i,j}^2] + \frac{2}{n(n-1)} \mathbb{E} [C_{i,j} C_{j,i}] \right) \\
 &= \frac{2}{n^2} \sum_{i=2}^n (i-1) \left(\mathbb{E} [C_{1,2}^2] + \frac{1}{n-1} \mathbb{E} [C_{1,2} C_{2,1}] \right),
 \end{aligned}$$

so $\sum_{i=2}^n \mathbb{E} [Y_{n,i}^2] = \frac{n-1}{n} \mathbb{E} [C_{1,2}^2] + \frac{1}{n} \mathbb{E} [C_{1,2} C_{2,1}]$. From (A.23), we derive that

$$(A.25) \quad \sum_{i=2}^n \mathbb{E} [Y_{n,i}^2] \xrightarrow{n \rightarrow +\infty} \sigma_{h_\varphi, P^1 \otimes P^2}^2.$$

On the other hand, we have that

$$\begin{aligned}
 \mathbb{E} [Y_{n,i}^2 | \mathcal{F}_{n,i-1}] &= \frac{1}{n} \sum_{j=1}^{i-1} \mathbb{1}_{\{\Pi_n(i)=j\}} \mathbb{E} [C_{i,j}^2 | X_j] + \frac{1}{n} \sum_{j=1}^{i-1} \mathbb{1}_{\{\Pi_n(j)=i\}} \mathbb{E} [C_{j,i}^2 | X_j] \\
 &\quad + \frac{2}{n} \sum_{j=1}^{i-1} \mathbb{1}_{\{\Pi_n(i)=j\}} \mathbb{1}_{\{\Pi_n(j)=i\}} \mathbb{E} [C_{i,j} C_{j,i} | X_j] \\
 &\quad + \frac{2}{n} \sum_{1 \leq j \neq k \leq i-1} \mathbb{1}_{\{\Pi_n(i)=j\}} \mathbb{1}_{\{\Pi_n(k)=i\}} \mathbb{E} [C_{i,j} C_{k,i} | X_j, X_k].
 \end{aligned}$$

Then,

$$\sum_{i=2}^n \left(\mathbb{E} [Y_{n,i}^2 | \mathcal{F}_{n,i-1}] - \mathbb{E} [Y_{n,i}^2] \right) = A_{n,1} + A_{n,2} + 2A_{n,3} + 2A_{n,4},$$

with

$$\begin{aligned}
A_{n,1} &= \frac{1}{n} \sum_{1 \leq j < i \leq n} \left(\mathbf{1}_{\{\Pi_n(i)=j\}} \mathbb{E} [C_{i,j}^2 | X_j] - \frac{1}{n} \mathbb{E} [C_{i,j}^2] \right), \\
A_{n,2} &= \frac{1}{n} \sum_{1 \leq j < i \leq n} \left(\mathbf{1}_{\{\Pi_n(j)=i\}} \mathbb{E} [C_{j,i}^2 | X_j] - \frac{1}{n} \mathbb{E} [C_{i,j}^2] \right), \\
A_{n,3} &= \frac{1}{n} \sum_{1 \leq j < i \leq n} \left(\mathbf{1}_{\{\Pi_n(i)=j\}} \mathbf{1}_{\{\Pi_n(j)=i\}} \mathbb{E} [C_{i,j} C_{j,i} | X_j] - \frac{1}{n(n-1)} \mathbb{E} [C_{i,j} C_{j,i}] \right), \\
A_{n,4} &= \frac{1}{n} \sum_{1 \leq j \neq k < i \leq n} \left(\mathbf{1}_{\{\Pi_n(i)=j\}} \mathbf{1}_{\{\Pi_n(k)=i\}} \mathbb{E} [C_{i,j} C_{k,i} | X_j, X_k] \right).
\end{aligned}$$

Thus,

$$\begin{aligned}
\text{(A.26)} \quad \text{Var} \left(\sum_{i=2}^n (\mathbb{E} [Y_{n,i}^2 | \mathcal{F}_{n,i-1}]) \right) \\
\leq 4 \left(\mathbb{E} [A_{n,1}^2] + \mathbb{E} [A_{n,2}^2] + 4\mathbb{E} [A_{n,3}^2] + 4\mathbb{E} [A_{n,4}^2] \right).
\end{aligned}$$

Let us now control each term of the above right-hand side.

Convergence of $\mathbb{E} [A_{n,1}^2]$ and $\mathbb{E} [A_{n,2}^2]$.

$$\begin{aligned}
\mathbb{E} [A_{n,1}^2] &= \frac{1}{n^2} \sum_{1 \leq j < i \leq n} \sum_{1 \leq l < k \leq n} \left(\mathbb{E} [\mathbf{1}_{\{\Pi_n(i)=j\}} \mathbf{1}_{\{\Pi_n(k)=l\}}] \times \right. \\
&\quad \left. \mathbb{E} [\mathbb{E} [C_{i,j}^2 | X_j] \mathbb{E} [C_{k,l}^2 | X_l]] - \frac{1}{n^2} (\mathbb{E} [C_{k,l}^2])^2 \right).
\end{aligned}$$

Let us now consider all the cases where $1 \leq j < i \leq n$, and $1 \leq l < k \leq n$.

If $i = k$ and $j = l$, then

$$\mathbb{E} [\mathbf{1}_{\{\Pi_n(i)=j\}} \mathbf{1}_{\{\Pi_n(k)=l\}}] \mathbb{E} [\mathbb{E} [C_{i,j}^2 | X_j] \mathbb{E} [C_{k,l}^2 | X_l]] = \frac{1}{n} \mathbb{E} \left[(\mathbb{E} [C_{2,1}^2 | X_1])^2 \right].$$

If $i = k$ and $j \neq l$, or if $i \neq k$ and $j = l$, then

$$\mathbb{E} [\mathbf{1}_{\{\Pi_n(i)=j\}} \mathbf{1}_{\{\Pi_n(k)=l\}}] \mathbb{E} [\mathbb{E} [C_{i,j}^2 | X_j] \mathbb{E} [C_{k,l}^2 | X_l]] = 0.$$

If $i \neq k$ and $j \neq l$, then

$$\mathbb{E} [\mathbf{1}_{\{\Pi_n(i)=j\}} \mathbf{1}_{\{\Pi_n(k)=l\}}] \mathbb{E} [\mathbb{E} [C_{i,j}^2 | X_j] \mathbb{E} [C_{k,l}^2 | X_l]] = \frac{1}{n(n-1)} (\mathbb{E} [C_{2,1}^2])^2.$$

By combining these results, from (A.23) and under the assumption $(\mathcal{A}_{\varphi, Mmt})$, we obtain that

$$\begin{aligned} \mathbb{E} [A_{n,1}^2] &\leq \frac{n-1}{2n^2} \left(\mathbb{E} \left[(\mathbb{E} [C_{2,1}^2 | X_1])^2 \right] - \frac{\sigma_{h_{\varphi}, P^1 \otimes P^2}^4}{n} \right) \\ &\quad + C n^2 \left(\frac{1}{n(n-1)} - \frac{1}{n^2} \right) \sigma_{h_{\varphi}, P^1 \otimes P^2}^4 \xrightarrow{n \rightarrow +\infty} 0. \end{aligned}$$

One can prove in the same way that $\mathbb{E} [A_{n,2}^2] \xrightarrow{n \rightarrow +\infty} 0$.

Convergence of $\mathbb{E} [A_{n,3}^2]$. We easily prove that

$$\mathbb{E} [A_{n,3}^2] = \frac{1}{n^2} \sum_{1 \leq j < i \leq n} \sum_{1 \leq l < k \leq n} \kappa_{i,j,k,l} - \frac{1}{4n^2} (\mathbb{E} [C_{1,2} C_{2,1}])^2,$$

where

$$\begin{aligned} \kappa_{i,j,k,l} &= \mathbb{E} \left[\mathbb{1}_{\{\Pi_n(i)=j\}} \mathbb{1}_{\{\Pi_n(j)=i\}} \mathbb{1}_{\{\Pi_n(k)=l\}} \mathbb{1}_{\{\Pi_n(l)=k\}} \right] \times \\ &\quad \mathbb{E} [\mathbb{E} [C_{i,j} C_{j,i} | X_j] \mathbb{E} [C_{k,l} C_{l,k} | X_l]]. \end{aligned}$$

Let us again consider $\kappa_{i,j,k,l}$ in all the cases where $1 \leq j < i \leq n$, and $1 \leq l < k \leq n$. If $i = k$ and $j = l$, then

$$\kappa_{i,j,k,l} = \frac{1}{n(n-1)} \mathbb{E} \left[(\mathbb{E} [C_{2,1} C_{1,2} | X_1])^2 \right].$$

If $i = k$ and $j \neq l$, or if $i \neq k$ and $j = l$, then $\kappa_{i,j,k,l} = 0$.

If $i \neq k$ and $j \neq l$, then

$$\kappa_{i,j,k,l} = \frac{(\mathbb{E} [C_{1,2} C_{2,1}])^2}{n(n-1)(n-2)(n-3)}.$$

Thus, under $(\mathcal{A}_{\varphi, Mmt})$, we finally have that

$$\mathbb{E} [A_{n,3}^2] \leq \frac{1}{2n^2} \mathbb{E} \left[(\mathbb{E} [C_{1,2} C_{2,1} | X_1])^2 \right] + C \frac{n (\mathbb{E} [C_{1,2} C_{2,1}])^2}{(n-1)(n-2)(n-3)} \xrightarrow{n \rightarrow +\infty} 0.$$

Convergence of $\mathbb{E} [A_{n,4}^2]$.

$$\begin{aligned} \mathbb{E} [A_{n,4}^2] &= \frac{1}{n^2} \sum_{\substack{1 \leq j \neq k < i \leq n \\ 1 \leq p \neq q < l \leq n}} \left(\mathbb{E} \left[\mathbb{1}_{\{\Pi_n(i)=j\}} \mathbb{1}_{\{\Pi_n(k)=i\}} \mathbb{1}_{\{\Pi_n(l)=p\}} \mathbb{1}_{\{\Pi_n(q)=l\}} \right] \times \right. \\ &\quad \left. \mathbb{E} [\mathbb{E} [C_{i,j} C_{k,i} | X_j, X_k] \mathbb{E} [C_{l,p} C_{q,l} | X_p, X_q]] \right). \end{aligned}$$

Let us consider all the cases where $1 \leq j \neq k < i \leq n$, and $1 \leq p \neq q < l \leq n$. If $\#\{j, k, p, q\} \geq 3$, there exists at least one element in $\{j, k, p, q\}$, j for instance (the other cases are studied in the same way), which differs from the other ones. Then,

$$\begin{aligned} & \mathbb{E} [\mathbb{E} [C_{i,j} C_{k,i} | X_j, X_k] \mathbb{E} [C_{l,p} C_{q,l} | X_p, X_q]] \\ &= \mathbb{E} [\mathbb{E} [\mathbb{E} [C_{i,j} C_{k,i} | X_j, X_k] \mathbb{E} [C_{l,p} C_{q,l} | X_p, X_q] | X_k, X_p, X_q]] \\ &= \mathbb{E} [\mathbb{E} [C_{i,j} C_{k,i} | X_k] \mathbb{E} [C_{l,p} C_{q,l} | X_p, X_q]] \\ &= \mathbb{E} [\mathbb{E} [\mathbb{E} [C_{i,j} C_{k,i} | X_i, X_k] | X_k] \mathbb{E} [C_{l,p} C_{q,l} | X_p, X_q]] \\ &= \mathbb{E} [\mathbb{E} [C_{k,i} \mathbb{E} [C_{i,j} | X_i] | X_k] \mathbb{E} [C_{l,p} C_{q,l} | X_p, X_q]]. \end{aligned}$$

Since $\mathbb{E} [C_{i,j} | X_i] = 0$, this leads to

$$(A.27) \quad \mathbb{E} [\mathbb{E} [C_{i,j} C_{k,i} | X_j, X_k] \mathbb{E} [C_{l,p} C_{q,l} | X_p, X_q]] = 0.$$

If $j = p$, $k = q$, and $i = l$, then,

$$\mathbb{E} [\mathbb{1}_{\{\Pi_n(i)=j\}} \mathbb{1}_{\{\Pi_n(k)=i\}} \mathbb{1}_{\{\Pi_n(l)=p\}} \mathbb{1}_{\{\Pi_n(q)=l\}}] = \frac{1}{n(n-1)},$$

and

$$\begin{aligned} |\mathbb{E} [\mathbb{E} [C_{i,j} C_{k,i} | X_j, X_k] \mathbb{E} [C_{l,p} C_{q,l} | X_p, X_q]]| &= \mathbb{E} [(\mathbb{E} [C_{i,j} C_{k,i} | X_j, X_k])^2] \\ &= \mathbb{E} [(\mathbb{E} [C_{3,1} C_{2,3} | X_1, X_2])^2] \\ &< +\infty \quad \text{under } (\mathcal{A}_{\varphi, Mmt}). \end{aligned}$$

If $j = p$, $k = q$, and $i \neq l$, then $\mathbb{1}_{\{\Pi_n(k)=i\}} \mathbb{1}_{\{\Pi_n(q)=l\}} = 0$, so

$$\mathbb{E} [\mathbb{1}_{\{\Pi_n(i)=j\}} \mathbb{1}_{\{\Pi_n(k)=i\}} \mathbb{1}_{\{\Pi_n(l)=p\}} \mathbb{1}_{\{\Pi_n(q)=l\}}] = 0.$$

If $j = q$, $k = p$, and $i = l$, then $\mathbb{1}_{\{\Pi_n(i)=j\}} \mathbb{1}_{\{\Pi_n(l)=p\}} = 0$, so

$$\mathbb{E} [\mathbb{1}_{\{\Pi_n(i)=j\}} \mathbb{1}_{\{\Pi_n(k)=i\}} \mathbb{1}_{\{\Pi_n(l)=p\}} \mathbb{1}_{\{\Pi_n(q)=l\}}] = 0.$$

If $j = q$, $k = p$, and $i \neq l$, then

$$\mathbb{E} [\mathbb{1}_{\{\Pi_n(i)=j\}} \mathbb{1}_{\{\Pi_n(k)=i\}} \mathbb{1}_{\{\Pi_n(l)=p\}} \mathbb{1}_{\{\Pi_n(q)=l\}}] = \frac{(n-4)!}{n!},$$

and

$$\begin{aligned} |\mathbb{E} [\mathbb{E} [C_{i,j} C_{k,i} | X_j, X_k] \mathbb{E} [C_{l,p} C_{q,l} | X_p, X_q]]| &= |\mathbb{E} [\mathbb{E} [C_{i,j} C_{k,i} C_{l,k} C_{j,l} | X_j, X_k]]| \\ &\leq \mathbb{E} [C_{3,1} C_{2,3} C_{4,2} C_{1,4}] \\ &< +\infty \quad \text{under } (\mathcal{A}_{\varphi, Mmt}). \end{aligned}$$

By combining these results, we obtain that

$$\begin{aligned} \mathbb{E} [A_{n,4}^2] &\leq C \frac{n^3}{n^2} \mathbb{E} \left[\frac{(\mathbb{E} [C_{3,1} C_{2,3} | X_1, X_2])^2}{n(n-1)} \right] \\ &\quad + C' \frac{n^4 (n-4)!}{n^2 n!} \mathbb{E} [C_{3,1} C_{2,3} C_{4,2} C_{1,4}] \xrightarrow{n \rightarrow +\infty} 0. \end{aligned}$$

From (A.26), and the above results of convergence towards 0 for $\mathbb{E} [A_{n,1}^2]$, $\mathbb{E} [A_{n,2}^2]$, $\mathbb{E} [A_{n,3}^2]$, and $\mathbb{E} [A_{n,4}^2]$, we firstly derive that

$$\text{Var} \left(\sum_{i=2}^n (\mathbb{E} [Y_{n,i}^2 | \mathcal{F}_{n,i-1}]) \right) \xrightarrow{n \rightarrow +\infty} 0.$$

• Second point. Notice that

$$\mathbb{E} [Y_{n,i} Y'_{n,i} | \mathcal{F}_{n,i-1}] = B_{n,1} + B_{n,2} + B_{n,3} + B_{n,4},$$

with

$$\begin{aligned} B_{n,1} &= \frac{1}{n} \sum_{1 \leq j < i \leq n} \mathbf{1}_{\{\Pi_n(i)=j\}} \mathbf{1}_{\{\Pi'_n(i)=j\}} \mathbb{E} [C_{i,j}^2 | X_j], \\ B_{n,2} &= \frac{1}{n} \sum_{1 \leq j < i \leq n} \mathbf{1}_{\{\Pi_n(j)=i\}} \mathbf{1}_{\{\Pi'_n(j)=i\}} \mathbb{E} [C_{j,i}^2 | X_j], \\ B_{n,3} &= \frac{1}{n} \sum_{1 \leq j < i \leq n} (\mathbf{1}_{\{\Pi_n(i)=j\}} \mathbf{1}_{\{\Pi'_n(j)=i\}} + \mathbf{1}_{\{\Pi_n(j)=i\}} \mathbf{1}_{\{\Pi'_n(i)=j\}}) \mathbb{E} [C_{i,j} C_{j,i} | X_j], \end{aligned}$$

and

$$\begin{aligned} B_{n,4} &= \frac{1}{n} \sum_{1 \leq j \neq k < i \leq n} \left(\mathbf{1}_{\{\Pi_n(i)=j\}} \mathbf{1}_{\{\Pi'_n(i)=k\}} \mathbb{E} [C_{i,j} C_{i,k} | X_j, X_k] \right. \\ &\quad + \mathbf{1}_{\{\Pi_n(i)=j\}} \mathbf{1}_{\{\Pi'_n(k)=i\}} \mathbb{E} [C_{i,j} C_{k,i} | X_j, X_k] \\ &\quad + \mathbf{1}_{\{\Pi_n(j)=i\}} \mathbf{1}_{\{\Pi'_n(i)=k\}} \mathbb{E} [C_{j,i} C_{i,k} | X_j, X_k] \\ &\quad \left. + \mathbf{1}_{\{\Pi_n(j)=i\}} \mathbf{1}_{\{\Pi'_n(k)=i\}} \mathbb{E} [C_{j,i} C_{k,i} | X_j, X_k] \right). \end{aligned}$$

Thus,

$$\begin{aligned} \text{(A.28)} \quad \mathbb{E} \left[\left(\sum_{i=2}^n \mathbb{E} [Y_{n,i} Y'_{n,i} | \mathcal{F}_{n,i-1}] \right)^2 \right] \\ \leq 4 (\mathbb{E} [B_{n,1}^2] + \mathbb{E} [B_{n,2}^2] + \mathbb{E} [B_{n,3}^2] + \mathbb{E} [B_{n,4}^2]). \end{aligned}$$

Convergence of $\mathbb{E} [B_{n,1}^2]$ and $\mathbb{E} [B_{n,2}^2]$. It can be proved that

$$\mathbb{E} [B_{n,1}^2] \leq \frac{1}{n^3} \sum_{1 \leq j < i \leq n} \sum_{1 \leq l < k \leq n} \mathbb{E} [\mathbb{1}_{\{\Pi_n(i)=j\}} \mathbb{1}_{\{\Pi_n(k)=l\}}] \times \mathbb{E} [\mathbb{E} [C_{i,j}^2 | X_j] \mathbb{E} [C_{k,l}^2 | X_l]].$$

Then, with the same computations as for the convergence of $\mathbb{E} [A_{n,1}^2]$ above, we prove that

$$\mathbb{E} [B_{n,1}^2] \leq \frac{n-1}{2n^3} \mathbb{E} [(\mathbb{E} [C_{1,2}^2 | X_2])^2] + C \frac{\sigma_{h_\varphi, P^1 \otimes P^2}^4}{n-1} \xrightarrow{n \rightarrow +\infty} 0.$$

In the same way, we also prove that $\mathbb{E} [B_{n,2}^2] \xrightarrow{n \rightarrow +\infty} 0$.

Convergence of $\mathbb{E} [B_{n,3}^2]$. We also have that

$$\mathbb{E} [B_{n,3}^2] \leq \frac{4}{n^2} \sum_{1 \leq j < i \leq n} \sum_{1 \leq l < k \leq n} \mathbb{E} [\mathbb{1}_{\{\Pi_n(i)=j\}} \mathbb{1}_{\{\Pi_n(k)=l\}}] \times \mathbb{E} [\mathbb{1}_{\{\Pi'_n(j)=i\}} \mathbb{1}_{\{\Pi'_n(l)=k\}}] \mathbb{E} [\mathbb{E} [C_{i,j} C_{j,i} | X_j] \mathbb{E} [C_{k,l} C_{l,k} | X_l]].$$

Now, with similar computations as for the convergence of $\mathbb{E} [A_{n,1}^2]$ above again, we prove that

$$\mathbb{E} [B_{n,3}^2] \leq 2 \frac{n-1}{n^3} \mathbb{E} [(\mathbb{E} [C_{1,2} C_{2,1} | X_2])^2] + C \frac{(\mathbb{E} [C_{1,2} C_{2,1}])^2}{n-1} \xrightarrow{n \rightarrow +\infty} 0.$$

Convergence of $\mathbb{E} [B_{n,4}^2]$. Setting

$$B_{n,4,1} = \frac{1}{n} \sum_{1 \leq j \neq k < i \leq n} \mathbb{1}_{\{\Pi_n(i)=j\}} \mathbb{1}_{\{\Pi_n(i)=k\}} \mathbb{E} [C_{i,j} C_{i,k} | X_j, X_k],$$

$$B_{n,4,2} = \frac{1}{n} \sum_{1 \leq j \neq k < i \leq n} \mathbb{1}_{\{\Pi_n(i)=j\}} \mathbb{1}_{\{\Pi'_n(k)=i\}} \mathbb{E} [C_{i,j} C_{k,i} | X_j, X_k],$$

$$B_{n,4,3} = \frac{1}{n} \sum_{1 \leq j \neq k < i \leq n} \mathbb{1}_{\{\Pi_n(j)=i\}} \mathbb{1}_{\{\Pi_n(i)=k\}} \mathbb{E} [C_{j,i} C_{i,k} | X_j, X_k],$$

$$B_{n,4,4} = \frac{1}{n} \sum_{1 \leq j \neq k < i \leq n} \mathbb{1}_{\{\Pi_n(j)=i\}} \mathbb{1}_{\{\Pi'_n(k)=i\}} \mathbb{E} [C_{j,i} C_{k,i} | X_j, X_k],$$

then $B_{n,4} = B_{n,4,1} + B_{n,4,2} + B_{n,4,3} + B_{n,4,4}$ and in particular,

$$\mathbb{E} [B_{n,4}^2] \leq 4 (\mathbb{E} [B_{n,4,1}^2] + \mathbb{E} [B_{n,4,2}^2] + \mathbb{E} [B_{n,4,3}^2] + \mathbb{E} [B_{n,4,4}^2]).$$

Yet,

$$\begin{aligned} \mathbb{E} [B_{n,4,1}^2] &= \frac{1}{n^2} \sum_{1 \leq j \neq k < i \leq n} \sum_{1 \leq p \neq q < l \leq n} \mathbb{E} [\mathbb{1}_{\{\Pi_n(i)=j\}} \mathbb{1}_{\{\Pi_n(l)=p\}}] \times \\ &\quad \mathbb{E} [\mathbb{1}_{\{\Pi'_n(i)=k\}} \mathbb{1}_{\{\Pi'_n(l)=q\}}] \mathbb{E} [\mathbb{E} [C_{i,j} C_{i,k} | X_j, X_k] \mathbb{E} [C_{l,p} C_{l,q} | X_p, X_q]]]. \end{aligned}$$

Now, consider all the cases where $1 \leq j \neq k < i \leq n$, $1 \leq p \neq q < l \leq n$. If $\#\{j, k, p, q\} \geq 3$, using a similar argument as in (A.27), we obtain that

$$\mathbb{E} [\mathbb{E} [C_{i,j} C_{i,k} | X_j, X_k] \mathbb{E} [C_{l,p} C_{l,q} | X_p, X_q]] = 0.$$

If $j = p$, $k = q$, and $i = l$, then,

$$\mathbb{E} [\mathbb{1}_{\{\Pi_n(i)=j\}} \mathbb{1}_{\{\Pi_n(l)=p\}}] \mathbb{E} [\mathbb{1}_{\{\Pi'_n(i)=k\}} \mathbb{1}_{\{\Pi'_n(l)=q\}}] = \frac{1}{n^2},$$

and

$$\begin{aligned} |\mathbb{E} [\mathbb{E} [C_{i,j} C_{i,k} | X_j, X_k] \mathbb{E} [C_{l,p} C_{l,q} | X_p, X_q]]| &= \mathbb{E} \left[(\mathbb{E} [C_{i,j} C_{i,k} | X_j, X_k])^2 \right] \\ &= \mathbb{E} \left[(\mathbb{E} [C_{3,1} C_{3,2} | X_1, X_2])^2 \right] \\ &< +\infty \quad \text{under } (\mathcal{A}_\varphi, Mmt). \end{aligned}$$

If $j = p$, $k = q$, $i \neq l$, or if $j = q$, $k = p$, $i = l$, then $\mathbb{1}_{\{\Pi_n(i)=j\}} \mathbb{1}_{\{\Pi_n(l)=p\}}$ is equal to 0, so

$$\mathbb{E} [\mathbb{1}_{\{\Pi_n(i)=j\}} \mathbb{1}_{\{\Pi_n(l)=p\}}] \mathbb{E} [\mathbb{1}_{\{\Pi'_n(i)=k\}} \mathbb{1}_{\{\Pi'_n(l)=q\}}] = 0.$$

If $j = q$, $k = p$, and $i \neq l$, then

$$\mathbb{E} [\mathbb{1}_{\{\Pi_n(i)=j\}} \mathbb{1}_{\{\Pi_n(l)=p\}}] \mathbb{E} [\mathbb{1}_{\{\Pi'_n(i)=k\}} \mathbb{1}_{\{\Pi'_n(l)=q\}}] = \frac{1}{n^2(n-1)^2},$$

and

$$\begin{aligned} |\mathbb{E} [\mathbb{E} [C_{i,j} C_{i,k} | X_j, X_k] \mathbb{E} [C_{l,p} C_{l,q} | X_p, X_q]]| &= |\mathbb{E} [\mathbb{E} [C_{i,j} C_{i,k} C_{l,k} C_{l,j} | X_j, X_k]]| \\ &= \mathbb{E} [C_{3,1} C_{2,3} C_{4,2} C_{1,4}] \\ &< +\infty \quad \text{under } (\mathcal{A}_\varphi, Mmt). \end{aligned}$$

By combining these results, we obtain that

$$\mathbb{E} [B_{n,4,1}^2] \leq C \frac{\mathbb{E} \left[(\mathbb{E} [C_{3,1} C_{3,2} | X_1, X_2])^2 \right]}{n} + C' \frac{\mathbb{E} [C_{3,1} C_{2,3} C_{4,2} C_{1,4}]}{(n-1)^2} \xrightarrow{n \rightarrow +\infty} 0.$$

Following the same lines of proof, we furthermore obtain that $\mathbb{E} [B_{n,4,2}^2]$, $\mathbb{E} [B_{n,4,3}^2]$, and $\mathbb{E} [B_{n,4,4}^2]$ also converge towards 0. Hence, $\mathbb{E} [B_{n,4}^2] \xrightarrow{n \rightarrow +\infty} 0$. From (A.28), and the convergence towards 0 of $\mathbb{E} [B_{n,1}^2]$, $\mathbb{E} [B_{n,2}^2]$, $\mathbb{E} [B_{n,3}^2]$, and $\mathbb{E} [B_{n,4}^2]$, we derive that

$$\mathbb{E} \left[\left(\sum_{i=2}^n \mathbb{E} [Y_{n,i} Y'_{n,i} | \mathcal{F}_{n,i-1}] \right)^2 \right] \xrightarrow{n \rightarrow +\infty} 0,$$

which finally allows to conclude that assumption (i) is satisfied.

Assumption (ii). Given $\varepsilon > 0$, let us prove that

$$\sum_{i=2}^n \mathbb{E} \left[(aY_{n,i} + bY'_{n,i})^2 \mathbf{1}_{\{|aY_{n,i} + bY'_{n,i}| > \varepsilon\}} \right] \xrightarrow{n \rightarrow +\infty} 0.$$

$$\begin{aligned} \sum_{i=2}^n \mathbb{E} \left[(aY_{n,i} + bY'_{n,i})^2 \mathbf{1}_{\{|aY_{n,i} + bY'_{n,i}| > \varepsilon\}} \right] &\leq \frac{1}{\varepsilon^2} \sum_{i=2}^n \mathbb{E} \left[(aY_{n,i} + bY'_{n,i})^4 \right] \\ &\leq \frac{2^3}{\varepsilon^2} \sum_{i=2}^n \left(a^4 \mathbb{E} [Y_{n,i}^4] + b^4 \mathbb{E} [Y'_{n,i}{}^4] \right) \\ &\leq \frac{2^3(a^4 + b^4)}{\varepsilon^2} \sum_{i=2}^n \mathbb{E} [Y_{n,i}^4]. \end{aligned}$$

Since $Y_{n,i} = n^{-1/2} \left(\mathbf{1}_{\{\Pi_n(i) < i\}} C_{i, \Pi_n(i)} + \mathbf{1}_{\{\Pi_n^{-1}(i) < i\}} C_{\Pi_n^{-1}(i), i} \right)$,

$$\begin{aligned} \mathbb{E} [Y_{n,i}^4] &\leq \frac{2^3}{n^2} \mathbb{E} \left[\mathbf{1}_{\{\Pi_n(i) < i\}} C_{i, \Pi_n(i)}^4 + \mathbf{1}_{\{\Pi_n^{-1}(i) < i\}} C_{\Pi_n^{-1}(i), i}^4 \right] \\ &\leq \frac{2^3}{n^2} \sum_{j=1}^{i-1} \left(\mathbb{E} [\mathbf{1}_{\{\Pi_n(i)=j\}} C_{i,j}^4] + \mathbb{E} [\mathbf{1}_{\{\Pi_n^{-1}(i)=j\}} C_{j,i}^4] \right) \\ &\leq \frac{2^4}{n^2} \mathbb{E} [C_{1,2}^4]. \end{aligned}$$

We thus obtain that

$$\sum_{i=2}^n \mathbb{E} \left[(aY_{n,i} + bY'_{n,i})^2 \mathbf{1}_{\{|aY_{n,i} + bY'_{n,i}| > \varepsilon\}} \right] \leq \frac{2^7(a^4 + b^4)}{\varepsilon^2 n} \mathbb{E} [C_{1,2}^4],$$

where the right-hand side tends to 0 as soon as $\mathbb{E} [C_{1,2}^4] < +\infty$.

This last condition is ensured by $(\mathcal{A}_{\varphi, Mmt})$, which allows to confirm that

assumption (ii) is also checked, and that

$$\mathcal{L} \left(aM_n^{\Pi_n}(\mathbb{X}_n) + bM_n^{\Pi'_n}(\mathbb{X}_n) \right) \xrightarrow{n \rightarrow +\infty} \mathcal{N} \left(0, (a^2 + b^2) \sigma_{h_\varphi, P^1 \otimes P^2}^2 \right).$$

This ends the proof of Lemma 6.1.

Recall that we aim at proving that

$$d_{BL} \left(\mathcal{L} \left(M_n^{\Pi_n}(\mathbb{X}_n) | \mathbb{X}_n \right), \mathcal{N} \left(0, \sigma_{h_\varphi, P^1 \otimes P^2}^2 \right) \right) \xrightarrow{n \rightarrow +\infty} 0.$$

From Lemma 6.1, we deduce that for every t in \mathbb{R} ,

$$\begin{cases} \mathbb{P} \left(M_n^{\Pi_n}(\mathbb{X}_n) \leq t \right) \xrightarrow{n \rightarrow +\infty} \Phi_{0, \sigma_{h_\varphi, P^1 \otimes P^2}^2}^2(t), \\ \mathbb{P} \left(M_n^{\Pi_n}(\mathbb{X}_n) \leq t, M_n^{\Pi'_n}(\mathbb{X}_n) \leq t \right) \xrightarrow{n \rightarrow +\infty} \Phi_{0, \sigma_{h_\varphi, P^1 \otimes P^2}^2}^2(t). \end{cases}$$

Setting $M_n = M_n^{\Pi_n}(\mathbb{X}_n)$ for the sake of simplicity, this leads to

$$(A.29) \quad \begin{cases} \mathbb{E} \left[\mathbb{E} \left[\mathbf{1}_{\{M_n \leq t\}} | \mathbb{X}_n \right] \right] \xrightarrow{n \rightarrow +\infty} \Phi_{0, \sigma_{h_\varphi, P^1 \otimes P^2}^2}^2(t), \\ \mathbb{E} \left[\left(\mathbb{E} \left[\mathbf{1}_{\{M_n \leq t\}} | \mathbb{X}_n \right] \right)^2 \right] \xrightarrow{n \rightarrow +\infty} \Phi_{0, \sigma_{h_\varphi, P^1 \otimes P^2}^2}^2(t). \end{cases}$$

In a separable metric space, convergence in probability is metrizable (see [13, Th. 9.2.1] for instance), therefore it is equivalent to almost-sure convergence of a sub-sequence of any initial subsequence. So, let us fix an initial extraction $\phi_0 : \mathbb{N} \rightarrow \mathbb{N}$, which defines a sub-sequence $(M_{\phi_0(n)})_{n \in \mathbb{N}}$ of $(M_n)_{n \in \mathbb{N}}$. Let us denote by $(q_m)_{m \in \mathbb{N}}$ a sequence such that $\{q_m, m \in \mathbb{N}\} = \mathbb{Q}$. For any m in \mathbb{N} , from (A.29), we derive that

$$\begin{cases} \mathbb{E} \left[\mathbb{E} \left[\mathbf{1}_{\{M_{\phi_0(n)} \leq q_m\}} | \mathbb{X}_{\phi_0(n)} \right] \right] \xrightarrow{n \rightarrow +\infty} \Phi_{0, \sigma_{h_\varphi, P^1 \otimes P^2}^2}^2(q_m), \\ \mathbb{E} \left[\left(\mathbb{E} \left[\mathbf{1}_{\{M_{\phi_0(n)} \leq q_m\}} | \mathbb{X}_{\phi_0(n)} \right] \right)^2 \right] \xrightarrow{n \rightarrow +\infty} \Phi_{0, \sigma_{h_\varphi, P^1 \otimes P^2}^2}^2(q_m), \end{cases}$$

which leads (by Chebychev's inequality) to

$$(A.30) \quad \mathbb{E} \left[\mathbf{1}_{\{M_{\phi_0(n)} \leq q_m\}} | \mathbb{X}_{\phi_0(n)} \right] \xrightarrow{n \rightarrow +\infty} \Phi_{0, \sigma_{h_\varphi, P^1 \otimes P^2}^2}^2(q_m).$$

Therefore, there exist an extraction ϕ_1 and a subset Ω_1 of Ω such that $\mathbb{P}(\Omega_1) = 1$, and for every ω in Ω_1 ,

$$\mathbb{E} \left[\mathbf{1}_{\{M_{\phi_1 \circ \phi(n)} \leq q_1\}} | \mathbb{X}_{\phi_1 \circ \phi(n)} \right] (\omega) \xrightarrow{n \rightarrow +\infty} \Phi_{0, \sigma_{h_\varphi, P^1 \otimes P^2}^2}^2(q_1).$$

Now, let $m \geq 1$ for which there exist an extraction ϕ_m and a subset Ω_m of Ω such that $\mathbb{P}(\Omega_m) = 1$, and for every $\omega \in \Omega_m$,

$$\mathbb{E} \left[\mathbb{1}_{\left\{M_{\phi_m \circ \phi_{m-1} \circ \dots \circ \phi_0(n)} \leq q_m\right\}} \left| \mathbb{X}_{\phi_m \circ \phi_{m-1} \circ \dots \circ \phi_0(n)} \right. \right] (\omega) \xrightarrow{n \rightarrow +\infty} \Phi_{0, \sigma_{h_\varphi, P^1 \otimes P^2}^2}^2(q_m).$$

Then, from (A.30), there also exist an extraction ϕ_{m+1} and a subset Ω_{m+1} of Ω such that $\mathbb{P}(\Omega_{m+1}) = 1$, and for every ω in Ω_{m+1} ,

$$\mathbb{E} \left[\mathbb{1}_{\left\{M_{\phi_{m+1} \circ \phi_m \circ \phi_{m-1} \circ \dots \circ \phi_0(n)} \leq q_{m+1}\right\}} \left| \mathbb{X}_{\phi_{m+1} \circ \phi_m \circ \dots \circ \phi_0(n)} \right. \right] (\omega) \xrightarrow{n \rightarrow +\infty} \Phi_{0, \sigma_{h_\varphi, P^1 \otimes P^2}^2}^2(q_{m+1}).$$

Setting $\tilde{\Omega} = \bigcap_{m \in \mathbb{N}} \Omega_m$ and for every n in \mathbb{N} , $\tilde{\phi}(n) = \phi_n \circ \dots \circ \phi_2 \circ \phi_1(n)$, then $\mathbb{P}(\tilde{\Omega}) = 1$. Moreover, for every ω in $\tilde{\Omega}$, m in \mathbb{N} ,

$$\mathbb{E} \left[\mathbb{1}_{\left\{M_{\tilde{\phi} \circ \phi_0(n)} \leq q_m\right\}} \left| \mathbb{X}_{\tilde{\phi} \circ \phi_0(n)} \right. \right] (\omega) \xrightarrow{n \rightarrow +\infty} \Phi_{0, \sigma_{h_\varphi, P^1 \otimes P^2}^2}^2(q_m).$$

Since $\Phi_{0, \sigma_{h_\varphi, P^1 \otimes P^2}^2}^2$ is a continuous distribution function, it can be proved that this follows

$$d_{BL} \left(\mathcal{L} \left(M_{\tilde{\phi} \circ \phi_0(n)} \left| \mathbb{X}_{\tilde{\phi} \circ \phi_0(n)} \right. \right), \mathcal{N} \left(0, \sigma_{h_\varphi, P^1 \otimes P^2}^2 \right) \right) \xrightarrow[n \rightarrow +\infty]{a.s.} 0.$$

To conclude, we actually proved that

$$d_{BL} \left(\mathcal{L} \left(M_n^{\Pi_n}(\mathbb{X}_n) \left| \mathbb{X}_n \right. \right), \mathcal{N} \left(0, \sigma_{h_\varphi, P^1 \otimes P^2}^2 \right) \right) \xrightarrow[n \rightarrow +\infty]{\mathbb{P}} 0,$$

which, combined with (A.21), leads to

$$d_{BL} \left(\mathcal{L} \left(\sqrt{n} U_{n, h_\varphi}(\mathbb{X}_n^{\Pi_n}) \left| \mathbb{X}_n \right. \right), \mathcal{N} \left(0, \sigma_{h_\varphi, P^1 \otimes P^2}^2 \right) \right) \xrightarrow[n \rightarrow +\infty]{\mathbb{P}} 0.$$

Third step: convergence of conditional second order moments. Recall that from (A.19), $U_{n, h_\varphi}(\mathbb{X}_n^{\Pi_n}) = \frac{1}{n-1} U_n^{\Pi_n}$, where

$$\begin{aligned} U_n^{\Pi_n} &= \sum_{i=1}^n \varphi \left(X_i^1, X_{\Pi_n(i)}^2 \right) - \frac{1}{n} \sum_{i,j=1}^n \varphi \left(X_i^1, X_j^2 \right) \\ &= \sum_{i,j=1}^n \left(\mathbb{1}_{\{\Pi_n(i)=j\}} - \frac{1}{n} \right) \varphi \left(X_i^1, X_j^2 \right). \end{aligned}$$

Therefore,

$$(A.31) \quad \mathbb{E} \left[(\sqrt{n}U_{n,h\varphi}(\mathbb{X}_n^{\Pi_n}))^2 \middle| \mathbb{X}_n \right] = \frac{n^2}{(n-1)^2} \left(\frac{1}{n} \mathbb{E} \left[(U_n^{\Pi_n})^2 \middle| \mathbb{X}_n \right] \right),$$

and if $C_{i,j,k,l} = (\mathbb{E} [\mathbb{1}_{\{\Pi_n(i)=j\}} \mathbb{1}_{\{\Pi_n(k)=l\}}] - \frac{1}{n^2}) \varphi(X_i^1, X_j^2) \varphi(X_k^1, X_l^2)$,

$$\frac{1}{n} \mathbb{E} \left[(U_n^{\Pi_n})^2 \middle| \mathbb{X}_n \right] = \frac{1}{n} \sum_{i,j=1}^n \sum_{k,l=1}^n C_{i,j,k,l}.$$

Firstly,

$$\frac{1}{n} \sum_{\substack{i,j,k,l \in \{1, \dots, n\} \\ \#\{i,j,k,l\}=4}} C_{i,j,k,l} = \frac{(n-2)(n-3)}{n^2} U_{n,1},$$

where

$$U_{n,1} = \frac{(n-4)!}{n!} \sum_{\substack{i,j,k,l \in \{1, \dots, n\} \\ \#\{i,j,k,l\}=4}} \varphi(X_i^1, X_j^2) \varphi(X_k^1, X_l^2)$$

is clearly a U -statistic of order 4. From the strong law of large numbers of Hoeffding [25], we thus have that

$$\frac{(n-2)(n-3)}{n^2} U_{n,1} \xrightarrow[n \rightarrow +\infty]{a.s.} (\mathbb{E} [\varphi(X_1^1, X_2^2)])^2.$$

Secondly,

$$\frac{1}{n} \sum_{\substack{i,j,k,l \in \{1, \dots, n\} \\ \#\{i,j,k,l\}=3 \\ i=j, i=l, j=k, \text{ or } k=l}} C_{i,j,k,l} = \frac{2(n-2)}{n^2} U_{n,2},$$

where

$$U_{n,2} = \frac{(n-3)!}{n!} \sum_{\substack{i,k,l \in \{1, \dots, n\} \\ \#\{i,k,l\}=3}} (\varphi(X_i^1, X_i^2) \varphi(X_k^1, X_l^2) + \varphi(X_i^1, X_l^2) \varphi(X_k^1, X_i^2))$$

is a U -statistic of order 3 which converges almost-surely, so

$$\frac{2(n-2)}{n^2} U_{n,2} \xrightarrow[n \rightarrow +\infty]{a.s.} 0.$$

Thirdly,

$$\frac{1}{n} \sum_{\substack{i,j,k,l \in \{1, \dots, n\} \\ \#\{i,j,k,l\}=3 \\ i=k, \text{ or } j=l}} C_{i,j,k,l} = -\frac{n(n-1)(n-2)}{n^3} U_{n,3},$$

where

$$U_{n,3} = \frac{(n-3)!}{n!} \sum_{\substack{i,k,l \in \{1,\dots,n\} \\ \#\{i,k,l\}=3}} (\varphi(X_i^1, X_k^2) \varphi(X_i^1, X_l^2) + \varphi(X_i^1, X_l^2) \varphi(X_k^1, X_l^2))$$

is a U -statistic of order 3. So,

$$-\frac{n(n-1)(n-2)}{n^3} U_{n,3} \xrightarrow[n \rightarrow +\infty]{a.s.} -\mathbb{E} \left[(\mathbb{E} [\varphi(X_1^1, X_2^2) | X_1])^2 \right] - \mathbb{E} \left[(\mathbb{E} [\varphi(X_1^1, X_2^2) | X_2])^2 \right].$$

Fourthly,

$$\frac{1}{n} \sum_{\substack{i,j,k,l \in \{1,\dots,n\} \\ \#\{i,j,k,l\}=2 \\ i=j=k, i=j=l, \\ i=k=l, \text{ or } j=k=l}} C_{i,j,k,l} = -\frac{2(n-1)}{n^2} U_{n,4},$$

where

$$U_{n,4} = \frac{1}{n(n-1)} \sum_{1 \leq i \neq j \leq n} (\varphi(X_i^1, X_i^2) \varphi(X_i^1, X_j^2) + \varphi(X_i^1, X_i^2) \varphi(X_j^1, X_i^2))$$

is a U -statistic of order 2, so

$$-\frac{2(n-1)}{n^2} U_{n,4} \xrightarrow[n \rightarrow +\infty]{a.s.} 0.$$

Fifthly,

$$\frac{1}{n} \sum_{\substack{i,j,k,l \in \{1,\dots,n\} \\ \#\{i,j,k,l\}=2 \\ i=j \neq k=l, \text{ or } i=l \neq j=k}} C_{i,j,k,l} = \frac{1}{n^2} U_{n,5},$$

where

$$U_{n,5} = \frac{1}{n(n-1)} \sum_{1 \leq i \neq j \leq n} (\varphi(X_i^1, X_i^2) \varphi(X_j^1, X_j^2) + \varphi(X_i^1, X_j^2) \varphi(X_j^1, X_i^2))$$

is a U -statistic of order 2, so

$$\frac{1}{n^2} U_{n,5} \xrightarrow[n \rightarrow +\infty]{a.s.} 0.$$

Sixthly,

$$\frac{1}{n} \sum_{\substack{i,j,k,l \in \{1,\dots,n\} \\ \#\{i,j,k,l\}=2 \\ i=k \neq j=l}} C_{i,j,k,l} = \frac{(n-1)^2}{n^2} U_{n,6},$$

where

$$U_{n,6} = \frac{1}{n(n-1)} \sum_{1 \leq i \neq j \leq n} \varphi^2(X_i^1, X_j^2)$$

is a U -statistic of order 2, so

$$\frac{(n-1)^2}{n^2} U_{n,6} \xrightarrow[n \rightarrow +\infty]{a.s.} \mathbb{E} [\varphi^2(X_1^1, X_2^2)].$$

Seventhly,

$$\frac{1}{n} \sum_{\substack{i,j,k,l \in \{1,\dots,n\} \\ \#\{i,j,k,l\}=1}} C_{i,j,k,l} = \frac{n-1}{n^3} \sum_{i=1}^n \varphi(X_i^1, X_i^2),$$

which almost-surely tends to 0 thanks to the strong law of large numbers.

By combining all these results, and the fact that

$$\begin{aligned} \sigma_{h_\varphi, P^1 \otimes P^2}^2 &= \mathbb{E} [\varphi^2(X_1^1, X_2^2)] + (\mathbb{E} [\varphi(X_1^1, X_2^2)])^2 \\ &\quad - \mathbb{E} \left[(\mathbb{E} [\varphi(X_1^1, X_2^2) | X_1])^2 \right] - \mathbb{E} \left[(\mathbb{E} [\varphi(X_1^1, X_2^2) | X_2])^2 \right], \end{aligned}$$

we finally obtain that

$$\frac{1}{n} \mathbb{E} \left[(U_n^{\Pi_n})^2 \middle| \mathbb{X}_n \right] \xrightarrow[n \rightarrow +\infty]{a.s.} \sigma_{h_\varphi, P^1 \otimes P^2}^2,$$

and from (A.31), we deduce that

$$\mathbb{E} \left[(\sqrt{n} U_{n, h_\varphi}(\mathbb{X}_n^{\Pi_n}))^2 \middle| \mathbb{X}_n \right] \xrightarrow[n \rightarrow +\infty]{a.s.} \sigma_{h_\varphi, P^1 \otimes P^2}^2.$$

Since $d_{BL} \left(\mathcal{L}(\sqrt{n} U_{n, h_\varphi}(\mathbb{X}_n^{\Pi_n}) | \mathbb{X}_n), \mathcal{N}(0, \sigma_{h_\varphi, P^1 \otimes P^2}^2) \right) \xrightarrow[n \rightarrow +\infty]{\mathbb{P}} 0$, this allows to conclude that

$$d_2 \left(\mathcal{L}(\sqrt{n} U_{n, h_\varphi}(\mathbb{X}_n^{\Pi_n}) | \mathbb{X}_n), \mathcal{N}(0, \sigma_{h_\varphi, P^1 \otimes P^2}^2) \right) \xrightarrow[n \rightarrow +\infty]{\mathbb{P}} 0.$$

A.8. Proof of Theorem 4.2. As in the proof of Theorem 3.2 and Proposition 3.6, we focus on the test by upper values. The proof of Theorem 4.2 is very similar to the one of Theorem 3.2, just replacing the argument of (6.9) by $q_{\varphi, 1-\alpha, n}^*(\mathbb{X}_n) \xrightarrow[n \rightarrow +\infty]{\mathbb{P}} \Phi_{0, \sigma_{h_\varphi, P^1 \otimes P^2}^2}^{-1}(1-\alpha)$, which is derived from Corollary 4.1.

A.9. Proof of Proposition 4.3. As above, we focus on the test by upper values. The proof is very similar to the one of Proposition 3.6. For the asymptotic part of the result, we first show, following the proof of [50, Lemma 21.2], that

$$(A.32) \quad \sqrt{n}U^{*,(\lceil(1-\alpha)(B_n+1)\rceil)} \xrightarrow[n \rightarrow +\infty]{\mathbb{P}} \Phi_{0,\sigma_{h,P^1 \otimes P^2}^2}^{-1}(1-\alpha),$$

then, we conclude using the same arguments as in the proof of Theorem 4.2.

Let F_{n,\mathbb{X}_n}^* be the c.d.f of $\mathcal{L}(\sqrt{n}U_{n,h,\varphi}, P_n^* | \mathbb{X}_n)$, and let us first show that

$$(A.33) \quad \sup_{z \in \mathbb{R}} \left| F_{n,\mathbb{X}_n}^*(z) - \Phi_{0,\sigma_{h,P^1 \otimes P^2}^2}(z) \right| \xrightarrow[n \rightarrow +\infty]{\mathbb{P}} 0.$$

As Theorem 4.1 provides only a convergence in probability, similar arguments of subsequences as in the proof of Corollary 4.1, have to be used. So, let ϕ_0 be an initial extraction and ϕ_1 be the extraction such that (6.15) is satisfied. As convergence in the d_{BL} metric is equivalent to a weak convergence (see [13, Proposition 11.3.3] for instance), and as the limit is continuous, by [50, Lemma 2.11] we obtain that

$$\sup_{z \in \mathbb{R}} \left| F_{\phi_1 \circ \phi_0(n), \mathbb{X}_{\phi_1 \circ \phi_0(n)}}^*(z) - \Phi_{0,\sigma_{h,P^1 \otimes P^2}^2}(z) \right| \xrightarrow[n \rightarrow +\infty]{a.s.} 0.$$

This being true for any initial extraction ϕ_0 , we obtain (A.33).

Let $F_{n,\mathbb{X}_n}^{*,B_n}$ denote the empirical c.d.f of $\mathcal{L}(\sqrt{n}U_{n,h,\varphi}, P_n^* | \mathbb{X}_n)$ associated with the sample $(\Pi_n^1, \dots, \Pi_n^{B_n})$, that is

$$\forall z \in \mathbb{R}, \quad F_{n,\mathbb{X}_n}^{*,B_n}(z) = \frac{1}{B_n} \sum_{b=1}^{B_n} \mathbb{1}_{\{\sqrt{n}U_{n,h,\varphi}(\mathbb{X}_n^{\Pi_n^b}) \leq z\}}.$$

Then, by the DKW inequality, we obtain as in the proof of Proposition 3.6,

$$(A.34) \quad \sup_{z \in \mathbb{R}} \left| F_{n,\mathbb{X}_n}^{*,B_n}(z) - F_{n,\mathbb{X}_n}^*(z) \right| \xrightarrow[n \rightarrow +\infty]{\mathbb{P}} 0.$$

Finally, let

$$G_{n,\mathbb{X}_n}^{*,B_n}(z) = \frac{1}{B_n+1} \sum_{b=1}^{B_n+1} \mathbb{1}_{\{\sqrt{n}U^{*,b} \leq z\}}.$$

Since $G_{n,\mathbb{X}_n}^{*,B_n}(z) = \frac{1}{B_n+1} \left(\mathbb{1}_{\{\sqrt{n}U_{n,h,\varphi}(\mathbb{X}_n) \leq z\}} + B_n F_{n,\mathbb{X}_n}^{*,B_n}(z) \right)$,

$$(A.35) \quad \sup_{z \in \mathbb{R}} \left| G_{n,\mathbb{X}_n}^{*,B_n}(z) - F_{n,\mathbb{X}_n}^{*,B_n}(z) \right| \leq \frac{2}{B_n+1} \xrightarrow[n \rightarrow +\infty]{} 0.$$

Combining (A.33), (A.34) and (A.35), we obtain that

$$\sup_{z \in \mathbb{R}} \left| G_{n, \mathbb{X}_n}^{*B_n}(z) - \Phi_{0, \sigma_{h, P^1 \otimes P^2}^2}(z) \right| \xrightarrow[n \rightarrow +\infty]{\mathbb{P}} 0.$$

Since

$$\sqrt{n} U^{*, (\lceil (1-\alpha)(B_n+1) \rceil)} = \left(G_{n, \mathbb{X}_n}^{*B_n} \right)^{-1} (1 - \alpha),$$

we obtain (A.32) and we conclude as in the proof of Proposition 3.6.

Now, let us prove the non asymptotic part of the result. Under (H_0) , from Proposition 4.1, we deduce that $\mathbb{X}_n^{\Pi_n^1}, \dots, \mathbb{X}_n^{\Pi_n^B}, \mathbb{X}_n$, and hence also $U^{*,1}, \dots, U^{*,B+1}$, are exchangeable real-valued random variables. Then,

$$\begin{aligned} & \mathbb{P} \left(U^{*,B+1} > U^{*, (\lceil (1-\alpha)(B+1) \rceil)} \right) \\ &= \mathbb{P} \left(\sum_{b=1}^{B+1} \mathbb{1}_{\{U^{*,b} < U^{*,B+1}\}} \geq (B+1) - \lfloor \alpha(B+1) \rfloor \right) \\ &= \mathbb{P} \left(\sum_{b=1}^{B+1} \mathbb{1}_{\{U^{*,b} \geq U^{*,B+1}\}} \leq \lfloor \alpha(B+1) \rfloor \right) \\ &= \mathbb{P} \left(\sum_{b=1}^{B+1} \mathbb{1}_{\{U^{*,b} \geq U^{*,B+1}\}} \leq \alpha(B+1) \right). \end{aligned}$$

By exchangeability of $U^{*,1}, \dots, U^{*,B+1}$, applying Lemma 1 in [43], whose proof is given in [2], we finally obtain that

$$\mathbb{P} \left(U^{*,B+1} > U^{*, (\lceil (1-\alpha)(B+1) \rceil)} \right) \leq \alpha,$$

which ends the proof.

APPENDIX B: ADDITIONAL RESULTS

B.1. About the non-degeneracy of the U -statistic. Define

$$\begin{aligned} Z(x) &= \int w(u, v) dN_{x^1}(u) dN_{x^2}(v) + \mathbb{E} \left[\int w(u, v) dN_{X^1}(u) dN_{X^2}(v) \right] \\ &\quad - \mathbb{E} \left[\int w(u, v) dN_{x^1}(u) dN_{X^2}(v) \right] - \mathbb{E} \left[\int w(u, v) dN_{X^1}(u) dN_{x^2}(v) \right]. \end{aligned}$$

Recall that in this case, the degeneration is equivalent to stating that for $X = (X^1, X^2)$ with distribution $P^1 \otimes P^2$, $Z(X)$ is a random variable

which is almost surely null under (H_0) . Since $\mathbb{E}[Z(X)] = 0$, $Z(X) = 0$ a.s. is equivalent to $\text{Var}(Z(X)) = 0$. Here we provide a computation of $\text{Var}(Z(X))$.

Let us introduce $dM_1^{[1]}(u)$ and $dM_2^{[1]}(v)$ the mean measures of respectively X^1 with distribution P^1 and X^2 of distribution P^2 [11, Chapter 5], then one can rewrite

$$\begin{aligned} Z(X) &= \int w(u, v) dN_{X^1}(u) dN_{X^2}(v) + \int w(u, v) dM_1^{[1]}(u) dM_2^{[1]}(v) \\ &\quad - \int w(u, v) dN_{X^1}(u) dM_2^{[1]}(v) - \int w(u, v) dM_1^{[1]}(u) dN_{X^2}(v). \end{aligned}$$

Therefore, $\mathbb{E}[Z(X)] = 0$, and

$$\begin{aligned} \text{Var}(Z(X)) &= \mathbb{E}[Z(X)^2] \\ &= \int_{[0,1]^4} w(u, v) w(s, t) \mathbb{E}[dN_{X^1}(u) dN_{X^1}(s)] \mathbb{E}[dN_{X^2}(v) dN_{X^2}(t)] \\ &\quad - \int_{[0,1]^4} w(u, v) w(s, t) \mathbb{E}[dN_{X^1}(u) dN_{X^1}(s)] dM_2^{[1]}(v) dM_2^{[1]}(t) \\ &\quad - \int_{[0,1]^4} w(u, v) w(s, t) dM_1^{[1]}(u) dM_1^{[1]}(s) \mathbb{E}[dN_{X^2}(v) dN_{X^2}(t)] \\ &\quad + \int_{[0,1]^4} w(u, v) w(s, t) dM_1^{[1]}(u) dM_1^{[1]}(s) dM_2^{[1]}(v) dM_2^{[1]}(t). \end{aligned}$$

By assuming that $\#X^1$ (resp $\#X^2$) has second order moment, (see also Section 3.2 for comment on this assumption), one can introduce the second factorial moment measure associated to X^1 (resp. X^2), and denoted $dM_1^{[2]}(u, s)$ (resp. $dM_2^{[2]}(v, t)$). Then straightforward computations show that

$$\begin{aligned} \text{Var}(Z(X)) &= \int_{[0,1]^2} w(u, v)^2 dM_1^{[1]}(u) dM_2^{[1]}(v) \\ &\quad + \int_{[0,1]^3} w(u, v) w(u, t) dM_1^{[1]}(u) \left(dM_2^{[2]}(v, t) - dM_2^{[1]}(v) dM_2^{[1]}(t) \right) \\ &\quad + \int_{[0,1]^3} w(u, v) w(s, v) \left(dM_1^{[2]}(u, s) - dM_1^{[1]}(u) dM_1^{[1]}(s) \right) dM_2^{[1]}(v) \\ &\quad + \int_{[0,1]^4} w(u, v) w(s, t) \left(dM_1^{[2]}(u, s) - dM_1^{[1]}(u) dM_1^{[1]}(s) \right) \times \\ &\quad \left(dM_2^{[2]}(v, t) - dM_2^{[1]}(v) dM_2^{[1]}(t) \right). \end{aligned}$$

In particular, for Poisson processes, $dM^{[2]}(u, s) = dM^{[1]}(u)dM^{[1]}(s)$ and

$$\text{Var}(Z(X)) = \int_{[0,1]^2} w(u, v)^2 dM_1^{[1]}(u)dM_2^{[1]}(v) > 0,$$

as soon as the Poisson processes have non zero intensities since for $j = 1, 2$, $dM_j^{[1]}(u) = \lambda_j(u)du$, with λ_j the intensity of X^j .

B.2. About the empirical centring assumption. Recall that

$$(\mathcal{A}_{Cent}^*) \quad \left\{ \begin{array}{l} \text{For } x_1 = (x_1^1, x_1^2), \dots, x_n = (x_n^1, x_n^2) \text{ in } \mathcal{X}^2, \\ \sum_{i_1, i_2, i'_1, i'_2=1}^n h\left(\left(x_{i_1}^1, x_{i_2}^2\right), \left(x_{i'_1}^1, x_{i'_2}^2\right)\right) = 0. \end{array} \right.$$

On the one hand, if $h = h_\varphi$, then for $n \geq 1$ and for $x_1 = (x_1^1, x_1^2), \dots, x_n = (x_n^1, x_n^2)$ in \mathcal{X}^2 ,

$$\begin{aligned} & \sum_{i, i', j, j'=1}^n h\left(\left(x_i^1, x_{i'}^2\right), \left(x_j^1, x_{j'}^2\right)\right) \\ &= \sum_{i, i', j, j'=1}^n \varphi\left(x_i^1, x_{i'}^2\right) + \varphi\left(x_j^1, x_{j'}^2\right) - \varphi\left(x_i^1, x_{j'}^2\right) - \varphi\left(x_j^1, x_{i'}^2\right) \\ &= n^2 \left(\sum_{i, i'=1}^n \varphi\left(x_i^1, x_{i'}^2\right) + \sum_{j, j'=1}^n \varphi\left(x_j^1, x_{j'}^2\right) - \sum_{i, j'=1}^n \varphi\left(x_i^1, x_{j'}^2\right) - \sum_{j, i'=1}^n \varphi\left(x_j^1, x_{i'}^2\right) \right) \\ &= 0. \end{aligned}$$

So (\mathcal{A}_{Cent}^*) is immediately satisfied if the kernel h is of the form h_φ .

On the other hand, (\mathcal{A}_{Cent}^*) does not imply that h is of the form h_φ . Indeed, consider

$$h\left(\left(x^1, x^2\right), \left(y^1, y^2\right)\right) = \#x^1 \cdot \#x^2 \cdot \#y^1 \cdot \#y^2 \left[(\#x^1 - \#y^1) (\#x^2 - \#y^2) \right].$$

- The kernel h is obviously symmetric.
- The kernel h satisfies (\mathcal{A}_{Cent}^*) . Let

$$f\left(x^1, y^1\right) = \#x^1 \cdot \#y^1 \left(\#x^1 - \#y^1\right).$$

First, notice that $f\left(x^1, x^1\right) = 0$ and $f\left(x^1, y^1\right) = -f\left(y^1, x^1\right)$. More-

over, $h((x^1, x^2), (y^1, y^2)) = f(x^1, y^1) f(x^2, y^2)$. Thus

$$\begin{aligned}
& \sum_{i,i',j,j'=1}^n h((x_i^1, x_{i'}^2), (x_j^1, x_{j'}^2)) \\
&= \sum_{i,i',j,j'=1}^n f(x_i^1, x_j^1) f(x_{i'}^2, x_{j'}^2) \\
&= \left(\sum_{i,j=1}^n f(x_i^1, x_j^1) \right) \left(\sum_{i',j'=1}^n f(x_{i'}^2, x_{j'}^2) \right) \\
&= \left(\sum_{i=1}^n \underbrace{f(x_i^1, x_i^1)}_0 + \sum_{1 \leq i < j \leq n} \underbrace{f(x_i^1, x_j^1) + f(x_j^1, x_i^1)}_0 \right) \left(\sum_{i',j'=1}^n f(x_{i'}^2, x_{j'}^2) \right) \\
&= 0,
\end{aligned}$$

and (\mathcal{A}_{Cent}^*) is satisfied by h .

- The kernel h cannot be written as an h_φ .

On the one hand, first notice that for any $\varphi : \mathcal{X}^2 \rightarrow \mathbb{R}$, the difference

$$\Delta_{h_\varphi} := h_\varphi((x^1, x^2), (y^1, y^2)) - h_\varphi((\tilde{x}^1, x^2), (y^1, y^2))$$

does not depend on y^1 . Indeed,

$$\begin{aligned}
\Delta_{h_\varphi} &= \varphi(x^1, x^2) + \varphi(y^1, y^2) - \varphi(x^1, y^2) - \varphi(y^1, x^2) \\
&\quad - \varphi(\tilde{x}^1, x^2) - \varphi(y^1, y^2) + \varphi(\tilde{x}^1, y^2) + \varphi(y^1, x^2) \\
&= \varphi(x^1, x^2) - \varphi(\tilde{x}^1, x^2) + \varphi(\tilde{x}^1, y^2) - \varphi(x^1, y^2).
\end{aligned}$$

And on the other hand, for the h introduced above, the difference Δ_h does depend on y^1 . Indeed

$$\begin{aligned}
\Delta_h &= h((x^1, x^2), (y^1, y^2)) - h((\tilde{x}^1, x^2), (y^1, y^2)) \\
&= \#x^2 \cdot \#y^1 \cdot \#y^2 (\#x^2 - \#y^2) \times \\
&\quad [\#x^1 \cdot (\#x^1 - \#y^1) - \#\tilde{x}^1 \cdot (\#\tilde{x}^1 - \#y^1)].
\end{aligned}$$

Thus, if for instance, $\#x^1 = \#y^2 = 1$ and $\#\tilde{x}^1 = \#x^2 = 2$, then

$$\Delta_h = 2\#y^1 [-3 + \#y^1],$$

which clearly depends on y^1 .

So finally, there does not exist any φ such that $h = h_\varphi$.

MÉLISANDE ALBERT
UNIV. NICE SOPHIA ANTIPOLIS
LJAD
06100 NICE, FRANCE
E-MAIL: Melisande.Albert@unice.fr
YANN BOURET
UNIV. NICE SOPHIA ANTIPOLIS
LPMC
06100 NICE, FRANCE
E-MAIL: Yann.Bouret@unice.fr

MAGALIE FROMONT
UNIV. RENNES 2
PLACE DU RECTEUR H. LE MOAL
35043 RENNES CEDEX, FRANCE
E-MAIL: magalie.fromont@univ-rennes2.fr
PATRICIA REYNAUD-BOURET
UNIV. NICE SOPHIA ANTIPOLIS
LJAD
06100 NICE, FRANCE
E-MAIL: Patricia.Reynaud-Bouret@unice.fr