

On kernel-based estimation of conditional Kendall's tau: finite-distance bounds and asymptotic behavior

Alexis Derumigny and Jean-David Fermanian

CREST-ENSAE,

5, avenue Henry Le Chatelier

91764 Palaiseau cedex, France.

e-mail: alexis.derumigny@ensae.fr, jean-david.fermanian@ensae.fr

Abstract: We study nonparametric estimators of conditional Kendall's tau, a measure of concordance between two random variables given some covariates. We prove non-asymptotic pointwise and uniform bounds, that hold with high probabilities. We provide “direct proofs” of the consistency and the asymptotic law of conditional Kendall's tau. A simulation study evaluates the numerical performance of such nonparametric estimators.

Keywords and phrases: conditional dependence measures, kernel smoothing, conditional Kendall's tau.

MSC 2010 subject classifications: Primary 62H20; secondary 62G05, 62G08, 62G20..

1. Introduction

In the field of dependence modeling, it is common to work with dependence measures. Contrary to usual linear correlations, most of them have the advantage of being defined without any condition on moments, and of being invariant to changes in the underlying marginal distributions. Such summaries of information are very popular and can be explicitly written as functionals of the underlying copulas: Kendall's tau, Spearman's rho, Blomqvist's coefficient... See Nelsen [1] for an introduction. In particular, for more than a century (Spearman (1904), Kendall (1938)), Kendall's tau has become a popular dependence measure in $[-1, 1]$. It quantifies the positive or negative dependence between two random variables X_1 and X_2 . Denoting by $C_{1,2}$ the unique underlying copula of (X_1, X_2) that are assumed to be continuous, their Kendall's tau can be directly defined as

$$\begin{aligned} \tau_{1,2} &:= 4 \int_{[0,1]^2} C_{1,2}(u_1, u_2) C_{1,2}(du_1, du_2) - 1 \\ &= \mathbb{P}((X_{1,1} - X_{2,1})(X_{1,2} - X_{2,2}) > 0) - \mathbb{P}((X_{1,1} - X_{2,1})(X_{1,2} - X_{2,2}) < 0), \end{aligned} \quad (1)$$

where $(X_{i,1}, X_{i,2})_{i=1,2}$ are two independent versions of $\mathbf{X} := (X_1, X_2)$. This measure is then interpreted as the probability of observing a *concordant pair* minus the probability of observing a *discordant pair*. See [2] for an historical perspective on Kendall's tau. Its inference is discussed in many textbooks (see [3] or [4], e.g.). Its links with copulas and other dependence measures can be found in [1] or [5].

Similar dependence measure can be introduced in a conditional setup, when a p -dimensional covariate \mathbf{Z} is available. When hundreds of papers refer to Kendall's tau, only a few of them have considered conditional Kendall's tau (as defined below) until now. The goal is now to model the dependence between the two components X_1 and X_2 , given the vector of covariates \mathbf{Z} . Logically, we can invoke the conditional copula

$C_{1,2|\mathbf{Z}=\mathbf{z}}$ of (X_1, X_2) given $\mathbf{Z} = \mathbf{z}$ for any point $\mathbf{z} \in \mathbb{R}^p$ (see Patton [6, 7]), and the corresponding conditional Kendall's tau would be simply defined as

$$\begin{aligned} \tau_{1,2|\mathbf{Z}=\mathbf{z}} &:= 4 \int_{[0,1]^2} C_{1,2|\mathbf{Z}=\mathbf{z}}(u_1, u_2) C_{1,2|\mathbf{Z}=\mathbf{z}}(du_1, du_2) - 1 \\ &= \mathbb{P}((X_{1,1} - X_{2,1})(X_{1,2} - X_{2,2}) > 0 | \mathbf{Z}_1 = \mathbf{Z}_2 = \mathbf{z}) \\ &\quad - \mathbb{P}((X_{1,1} - X_{2,1})(X_{1,2} - X_{2,2}) < 0 | \mathbf{Z}_1 = \mathbf{Z}_2 = \mathbf{z}), \end{aligned}$$

where $(X_{i,1}, X_{i,2}, \mathbf{Z}_i)_{i=1,2}$ are two independent versions of (X_1, X_2, \mathbf{Z}) . As above, this is the probability of observing a *concordant pair* minus the probability of observing a *discordant pair*, conditionally on \mathbf{Z}_1 and \mathbf{Z}_2 being both equal to \mathbf{z} . Note that, as conditional copulas themselves, conditional Kendall's taus are invariant w.r.t. increasing transformations of the conditional margins X_1 and X_2 , given \mathbf{Z} . Of course, if \mathbf{Z} is independent of (X_1, X_2) then, for every $\mathbf{z} \in \mathbb{R}^p$, the conditional Kendall's tau $\tau_{1,2|\mathbf{Z}=\mathbf{z}}$ is equal to the (unconditional) Kendall's tau $\tau_{1,2}$.

Conditional Kendall's tau, and more generally conditional dependence measures, are of interest per se because they allow to summarize the evolution of the dependence between X_1 and X_2 , when the covariate \mathbf{Z} is changing. Surprisingly, their nonparametric estimates have been introduced in the literature only a few years ago ([8],[9],[10]) and their properties have not yet been fully studied in depth. Indeed, until now and to the best of our knowledge, the theoretical properties of nonparametric conditional Kendall's tau estimates have been obtained "in passing" in the literature, as a sub-product of the weak-convergence of conditional copula processes ([9]) or as intermediate quantities that will be "plugged-in" ([11]). Therefore, such properties have been stated under too demanding assumptions. In particular, some assumptions were related to the estimation of conditional margins, while this is not required because Kendall's tau are based on ranks. In this paper, we directly study nonparametric estimates $\hat{\tau}_{1,2|\mathbf{z}}$ without relying on the theory/inference of copulas. Therefore, we will state their main usual statistical properties: exponential bounds in probability, consistency, asymptotic normality.

Our $\tau_{1,2|\mathbf{Z}=\mathbf{z}}$ has not to be confused with the so-called "conditional Kendall's tau" in the case of truncated data ([12], [13]), in the case of semi-competing risk models ([14], [15]), or for other partial information schemes ([16], [17], among others). Indeed, particularly in biostatistics or reliability, the inference of dependence models under truncation/censoring can be led by considering some types of conditional Kendall's tau, given some algebraic relationships among the underlying random variables. This would induce conditioning by subsets. At the opposite, we will consider only pointwise conditioning events in this paper, under a nonparametric point-of-view. Nonetheless, such pointwise events can be found in the literature, but in some parametric or semi-parametric particular frameworks, as for the identifiability of frailty distributions in bivariate proportional models ([18], [19]). Other related papers are [20] or [21], that are dealing with extreme co-movements (bivariate extreme-value theory). There, the tail conditioning events of Kendall's tau have probabilities that go to zero with the sample size.

In Section 2, different kernel-based estimators of the conditional Kendall's tau are discussed. In Section 3, the theoretical properties of the latter estimators are proved, first with finite-distance bounds and then under an asymptotic point-of-view. A short simulation study is provided in Section 4. Proofs are postponed into the appendix.

2. Definition of several kernel-based estimators of $\tau_{1,2|\mathbf{z}}$

Let $(X_{i,1}, X_{i,2}, \mathbf{Z}_i)$, $i = 1, \dots, n$ be an i.i.d. sample distributed as (X_1, X_2, \mathbf{Z}) , and $n \geq 2$. Assuming continuous underlying distributions, there are several equivalent ways of defining the conditional Kendall's tau:

$$\begin{aligned}\tau_{1,2|\mathbf{Z}=\mathbf{z}} &= 4 \mathbb{P}(X_{1,1} > X_{2,1}, X_{1,2} > X_{2,2} | \mathbf{Z}_1 = \mathbf{Z}_2 = \mathbf{z}) - 1 \\ &= 1 - 4 \mathbb{P}(X_{1,1} > X_{2,1}, X_{1,2} < X_{2,2} | \mathbf{Z}_1 = \mathbf{Z}_2 = \mathbf{z}) \\ &= \mathbb{P}((X_{1,1} - X_{2,1})(X_{1,2} - X_{2,2}) > 0 | \mathbf{Z}_1 = \mathbf{Z}_2 = \mathbf{z}) \\ &\quad - \mathbb{P}((X_{1,1} - X_{2,1})(X_{1,2} - X_{2,2}) < 0 | \mathbf{Z}_1 = \mathbf{Z}_2 = \mathbf{z}).\end{aligned}$$

Motivated by each of the latter expressions, we introduce several kernel-based estimators of $\tau_{1,2|\mathbf{Z}=\mathbf{z}}$:

$$\begin{aligned}\hat{\tau}_{1,2|\mathbf{Z}=\mathbf{z}}^{(1)} &:= 4 \sum_{i=1}^n \sum_{j=1}^n w_{i,n}(\mathbf{z}) w_{j,n}(\mathbf{z}) \mathbb{1}\{X_{i,1} < X_{j,1}, X_{i,2} < X_{j,2}\} - 1, \\ \hat{\tau}_{1,2|\mathbf{Z}=\mathbf{z}}^{(2)} &:= \sum_{i=1}^n \sum_{j=1}^n w_{i,n}(\mathbf{z}) w_{j,n}(\mathbf{z}) \left(\mathbb{1}\{(X_{i,1} - X_{j,1})(X_{i,2} - X_{j,2}) > 0\} \right. \\ &\quad \left. - \mathbb{1}\{(X_{i,1} - X_{j,1})(X_{i,2} - X_{j,2}) < 0\} \right), \\ \hat{\tau}_{1,2|\mathbf{Z}=\mathbf{z}}^{(3)} &:= 1 - 4 \sum_{i=1}^n \sum_{j=1}^n w_{i,n}(\mathbf{z}) w_{j,n}(\mathbf{z}) \mathbb{1}\{X_{i,1} < X_{j,1}, X_{i,2} > X_{j,2}\},\end{aligned}$$

where $\mathbb{1}$ denotes the indicator function, $w_{i,n}$ is a sequence of weights given by

$$w_{i,n}(\mathbf{z}) = \frac{K_h(\mathbf{Z}_i - \mathbf{z})}{\sum_{j=1}^n K_h(\mathbf{Z}_j - \mathbf{z})}, \quad (2)$$

with $K_h(\cdot) := h^{-p} K(\cdot/h)$ for some kernel K on \mathbb{R}^p , and $h = h(n)$ denotes a usual bandwidth sequence that tends to zero when $n \rightarrow \infty$. In this paper, we have chosen usual Nadaraya-Watson weights. Obviously, there are alternatives (local linear, Priestley-Chao, Gasser-Müller, etc., weight), that would lead to different theoretical results.

The estimators $\hat{\tau}_{1,2|\mathbf{Z}=\mathbf{z}}^{(1)}$, $\hat{\tau}_{1,2|\mathbf{Z}=\mathbf{z}}^{(2)}$ and $\hat{\tau}_{1,2|\mathbf{Z}=\mathbf{z}}^{(3)}$ look similar, but they are nevertheless different, as shown in Proposition 1. These differences are due to the fact that all the $\hat{\tau}_{1,2|\mathbf{Z}=\mathbf{z}}^{(k)}$, $k = 1, 2, 3$ are affine transformations of a double-indexed sum, on every pair (i, j) , including the diagonal terms where $i = j$. The treatment of these diagonal terms is different for each of the three estimators defined above. Indeed, setting $s_n := \sum_{i=1}^n w_{i,n}^2(\mathbf{z})$, it can be easily proved that $\hat{\tau}_{1,2|\mathbf{Z}=\mathbf{z}}^{(1)}$ takes values in the interval $[-1, 1 - 2s_n]$, $\hat{\tau}_{1,2|\mathbf{Z}=\mathbf{z}}^{(2)}$ in $[-1 + s_n, 1 - s_n]$, and $\hat{\tau}_{1,2|\mathbf{Z}=\mathbf{z}}^{(3)}$ in $[-1 + 2s_n, 1]$. Moreover, there exists a direct relationship between these estimators, given by the following proposition.

Proposition 1. *Almost surely, $\hat{\tau}_{1,2|\mathbf{Z}=\mathbf{z}}^{(1)} + s_n = \hat{\tau}_{1,2|\mathbf{Z}=\mathbf{z}}^{(2)} = \hat{\tau}_{1,2|\mathbf{Z}=\mathbf{z}}^{(3)} - s_n$, where $s_n := \sum_{i=1}^n w_{i,n}^2(\mathbf{z})$.*

This proposition is proved in A.1. As a consequence, we can easily rescale the previous estimators so that the new estimator will take values in the whole interval $[-1, 1]$. This would yield

$$\tilde{\tau}_{1,2|\mathbf{Z}=\mathbf{z}} := \frac{\hat{\tau}_{1,2|\mathbf{Z}=\mathbf{z}}^{(1)}}{1 - s_n} + \frac{s_n}{1 - s_n} = \frac{\hat{\tau}_{1,2|\mathbf{Z}=\mathbf{z}}^{(2)}}{1 - s_n} = \frac{\hat{\tau}_{1,2|\mathbf{Z}=\mathbf{z}}^{(3)}}{1 - s_n} - \frac{s_n}{1 - s_n}.$$

Note that none of the latter estimators depends on any estimation of conditional marginal distributions. In other words, we only have to conveniently choose the weights $w_{i,n}$ to obtain an estimator of the conditional Kendall's tau. This is coherent with the fact that conditional Kendall's taus are invariant with respect to conditional marginal distributions. Moreover, note that, in the definition of our estimators, the inequalities are strict (there are no terms corresponding to the cases $i = j$). This is inline with the definition of (conditional) Kendall's tau itself through concordant/discordant pairs of observations.

The definition of $\hat{\tau}_{1,2|\mathbf{Z}=\mathbf{z}}^{(1)}$ can be motivated as follows. For $j = 1, 2$, let $\hat{F}_{j|\mathbf{Z}}(\cdot|\mathbf{Z} = \mathbf{z})$ be an estimator of the conditional cdf of X_j given $\mathbf{Z} = \mathbf{z}$. Then, a usual estimator of the conditional copula of X_1 and X_2 given $\mathbf{Z} = \mathbf{z}$ is

$$\hat{C}_{1,2|\mathbf{Z}}(u_1, u_2|\mathbf{Z} = \mathbf{z}) := \sum_{i=1}^n w_{i,n}(\mathbf{z}) \mathbb{1}\{\hat{F}_{1|\mathbf{Z}}(X_{i,1}|\mathbf{Z} = \mathbf{z}) \leq u_1, \hat{F}_{2|\mathbf{Z}}(X_{i,2}|\mathbf{Z} = \mathbf{z}) \leq u_2\}.$$

See [9] or [10], e.g. The latter estimator of the conditional copula can be plugged into (1) to define an estimator of the conditional Kendall's tau itself:

$$\begin{aligned} \hat{\tau}_{1,2|\mathbf{Z}=\mathbf{z}} &:= 4 \int \hat{C}_{1,2|\mathbf{Z}}(u_1, u_2|\mathbf{Z} = \mathbf{z}) \hat{C}_{1,2|\mathbf{Z}}(du_1, du_2|\mathbf{Z} = \mathbf{z}) - 1 \\ &= 4 \sum_{j=1}^n w_{j,n}(\mathbf{z}) \hat{C}_{1,2|\mathbf{Z}}(\hat{F}_{1|\mathbf{Z}}(X_{j,1}|\mathbf{Z} = \mathbf{z}), \hat{F}_{2|\mathbf{Z}}(X_{j,2}|\mathbf{Z} = \mathbf{z})|\mathbf{Z} = \mathbf{z}) - 1. \end{aligned} \quad (3)$$

Since the functions $\hat{F}_{j|\mathbf{Z}}(\cdot|\mathbf{Z} = \mathbf{z})$ are non-decreasing, this reduces to

$$\begin{aligned} \hat{\tau}_{1,2|\mathbf{Z}=\mathbf{z}} &= 4 \sum_{i=1}^n \sum_{j=1}^n w_{i,n}(\mathbf{z}) w_{j,n}(\mathbf{z}) \mathbb{1}\{X_{i,1} \leq X_{j,1}, X_{i,2} \leq X_{j,2}\} - 1 \\ &= 4 \sum_{i=1}^n \sum_{j=1}^n w_{i,n}(\mathbf{z}) w_{j,n}(\mathbf{z}) \mathbb{1}\{X_{i,1} < X_{j,1}, X_{i,2} < X_{j,2}\} - 1 + o_P(1) = \hat{\tau}_{1,2|\mathbf{Z}=\mathbf{z}}^{(1)} + o_P(1). \end{aligned}$$

Veraverbeke et al. [9], Subsection 3.2, introduced their estimator of $\tau_{1,2|\mathbf{z}}$ by (3). By the functional Delta-Method, they deduced its asymptotic normality as a sub-product of the weak convergence of the process $\sqrt{nh}(\hat{C}_{1,2|\mathbf{Z}}(\cdot, \cdot|z) - C_{1,2|\mathbf{Z}}(\cdot, \cdot|z))$ when \mathbf{Z} is univariate. In our case, we will obtain more and stronger theoretical properties of $\hat{\tau}_{1,2|\mathbf{Z}=\mathbf{z}}^{(1)}$ under weaker conditions by a more direct analysis based on ranks. In particular, we will not require any regularity condition on the conditional marginal distributions, contrary to [9]. Indeed, in the latter paper, it is required that $F_{j|\mathbf{Z}}(\cdot|\mathbf{Z} = \mathbf{z})$ has to be two times continuously differentiable (assumption $(\tilde{R}3)$) and its inverse has to be continuous (assumption $(R1)$). This is not satisfied for some simple univariate cdf as $F_j(t) = t\mathbb{1}(t \in [0, 1])/2 + \mathbb{1}(t \in (1, 2])/2 + t\mathbb{1}(t \in (2, 4])/4 + \mathbb{1}(t > 4)$, for instance. Note that we could justify $\hat{\tau}_{1,2|\mathbf{Z}=\mathbf{z}}^{(3)}$ in a similar way by considering conditional survival copulas.

Let us define g_1, g_2, g_3 by

$$\begin{aligned} g_1(\mathbf{X}_i, \mathbf{X}_j) &:= 4\mathbb{1}\{X_{i,1} < X_{j,1}, X_{i,2} < X_{j,2}\} - 1, \\ g_2(\mathbf{X}_i, \mathbf{X}_j) &:= \mathbb{1}\{(X_{i,1} - X_{j,1}) \times (X_{i,2} - X_{j,2}) > 0\} - \mathbb{1}\{(X_{i,1} - X_{j,1}) \times (X_{i,2} - X_{j,2}) < 0\}, \\ g_3(\mathbf{X}_i, \mathbf{X}_j) &:= 1 - 4\mathbb{1}\{X_{i,1} < X_{j,1}, X_{i,2} > X_{j,2}\}, \end{aligned}$$

where, for $i = 1, \dots, n$, we set $\mathbf{X}_i := (X_{i,1}, X_{i,2})$. Clearly, $\hat{\tau}_{1,2|\mathbf{Z}}^{(k)}$ is a smoothed estimator of $\mathbb{E}[g_k(\mathbf{X}_1, \mathbf{X}_2)|\mathbf{Z}_1 = \mathbf{Z}_2 = \mathbf{z}]$, $k = 1, 2, 3$.

Note that such dependence measures are of interest for the purpose of estimating (conditional or unconditional) copula models too. Indeed, several popular parametric families of copulas have a simple one-to-one mapping between their parameter and the associated Kendall's tau (or Spearman's rho): Gaussian, Student with a fixed degree of freedom, Clayton, Gumbel and Frank copulas, etc. Then, assume for instance that the conditional copula $C_{1,2|\mathbf{Z}=\mathbf{z}}$ belongs is a Gaussian copula with a parameter $\rho(\mathbf{z})$. Then, by estimating its conditional Kendall's tau $\tau_{1,2|\mathbf{Z}=\mathbf{z}}$, we get an estimate of the corresponding parameter $\rho(\mathbf{z})$, and finally of the conditional copula itself. See [22], e.g.

The choice of the bandwidth h could be done in a data-driven way, following the general conditional U-statistics framework detailed in Dony and Mason [23, Section 2]. Indeed, for any $k \in \{1, 2, 3\}$ and $\mathbf{z} \in \mathcal{Z}$, denote by $\hat{\tau}_{-(i,j),1,2|\mathbf{Z}=\mathbf{z}}^{(h,k)}$ the estimator $\hat{\tau}_{1,2|\mathbf{Z}=\mathbf{z}}^{(k)}$ that is made with the smoothing parameter h and our dataset, when the i -th and j -th observations have been removed. As a consequence, the random function $\hat{\tau}_{-(i,j),1,2|\mathbf{Z}=\cdot}^{(h,k)}$ is independent of $((\mathbf{X}_i, \mathbf{Z}_i), (\mathbf{X}_j, \mathbf{Z}_j))$. As usual with kernel methods, it would be tempting to propose h as the minimizer of the cross-validation criterion

$$CV_{DM}(h) := \frac{2}{n(n-1)} \sum_{i,j=1}^n \left(g_k(\mathbf{X}_i, \mathbf{X}_j) - \hat{\tau}_{-(i,j),1,2|\mathbf{Z}=(\mathbf{z}_i+\mathbf{z}_j)/2}^{(h,k)} \right)^2 K_h(\mathbf{Z}_i - \mathbf{Z}_j),$$

for $k = 1, 2, 3$ or for $\tilde{\tau}_{1,2|\mathbf{Z}=\cdot}$. The latter criterion would be a ‘‘naively localized’’ version of the usual cross-validation method. Unfortunately, we observe that the function $h \mapsto CV_{DM}(h)$ is most often decreasing in the range of realistic bandwidth values. If we remove the weight $K_h(\mathbf{Z}_i - \mathbf{Z}_j)$, then there is no reason why $g_k(\mathbf{X}_i, \mathbf{X}_j)$ should be equal to $\hat{\tau}_{-(i,j),1,2|\mathbf{Z}=(\mathbf{z}_i+\mathbf{z}_j)/2}^{(k)}$ (on average), and we are not interested in the prediction of concordance/discordance pairs for which the Z_i and Z_j are far apart. Therefore, a modification of this criteria is necessary. We propose to separate the choice of h for the terms $g_k(\mathbf{X}_i, \mathbf{X}_j) - \hat{\tau}_{-(i,j),1,2|\mathbf{Z}=(\mathbf{z}_i+\mathbf{z}_j)/2}^{(h,k)}$ and the selection of the ‘‘convenient pairs’’ of observations (i, j) . This leads to the new criterion

$$CV_{\tilde{h}}(h) := \frac{2}{n(n-1)} \sum_{i,j=1}^n \left(g_k(\mathbf{X}_i, \mathbf{X}_j) - \hat{\tau}_{-(i,j),1,2|\mathbf{Z}=(\mathbf{z}_i+\mathbf{z}_j)/2}^{(h,k)} \right)^2 \tilde{K}_{\tilde{h}}(\mathbf{Z}_i - \mathbf{Z}_j), \quad (4)$$

with a potentially different kernel \tilde{K} and a new fixed tuning parameter \tilde{h} . Even if more complex procedures are possible, we suggest to simply choose $\tilde{K}(\mathbf{z}) := \mathbb{1}\{|\mathbf{z}|_\infty \leq 1\}$ and to calibrate \tilde{h} so that only a fraction of the pairs (i, j) has non-zero weights. In practice, set \tilde{h} as the empirical quantile of $(\{|\mathbf{Z}_i - \mathbf{Z}_j|_\infty : 1 \leq i < j \neq n\})$ of order $2N_{pairs}/(n(n-1))$, where N_{pairs} is the number of pairs we want to keep.

3. Theoretical results

3.1. Finite distance bounds

Hereafter, we will consider the behavior of conditional Kendall's tau estimates given $\mathbf{Z} = \mathbf{z}$ belongs to some fixed open subset \mathcal{Z} in \mathbb{R}^p . For the moment, let us state an instrumental result that is of interest per se. Let $\hat{f}_{\mathbf{Z}}(\mathbf{z}) := n^{-1} \sum_{j=1}^n K_h(\mathbf{Z}_j - \mathbf{z})$ be the usual kernel estimator of the density $f_{\mathbf{Z}}$ of the conditioning variable \mathbf{Z} . Note that the estimators $\hat{\tau}_{1,2|\mathbf{Z}=\mathbf{z}}^{(k)}$, $k = 1, \dots, 3$ are well-behaved only whenever $\hat{f}_{\mathbf{Z}}(\mathbf{z}) > 0$. Denote the joint density of (\mathbf{X}, \mathbf{Z}) by $f_{\mathbf{X},\mathbf{Z}}$. In our study, we need some usual conditions of regularity.

Assumption 3.1. The kernel K is bounded, and set $\|K\|_\infty =: C_K$. It is symmetrical and satisfies $\int K = 1$, $\int |K| < \infty$. This kernel is of order α for some integer $\alpha > 1$: for all $j = 1, \dots, \alpha - 1$ and every indices i_1, \dots, i_j in $\{1, \dots, p\}$, $\int K(\mathbf{u})u_{i_1} \dots u_{i_j} d\mathbf{u} = 0$. Moreover, $\mathbb{E}[K_h(\mathbf{Z} - \mathbf{z})] > 0$ for every $\mathbf{z} \in \mathcal{Z}$ and $h > 0$. Set $\tilde{K}(\cdot) := K^2(\cdot)/\int K^2$ and $\|\tilde{K}\|_\infty =: C_{\tilde{K}}$.

Assumption 3.2. $f_{\mathbf{Z}}$ is α -times continuously differentiable on \mathcal{Z} and there exists a constant $C_{K,\alpha} > 0$ s.t., for all $\mathbf{z} \in \mathcal{Z}$,

$$\int |K|(\mathbf{u}) \sum_{i_1, \dots, i_\alpha=1}^p |u_{i_1} \dots u_{i_\alpha}| \sup_{t \in [0,1]} \left| \frac{\partial^\alpha f_{\mathbf{Z}}}{\partial z_{i_1} \dots \partial z_{i_\alpha}}(\mathbf{z} + t\mathbf{u}) \right| d\mathbf{u} \leq C_{K,\alpha}.$$

Moreover, $C_{\tilde{K},2}$ denotes a similar constant replacing K by \tilde{K} and α by two.

Assumption 3.3. There exist two positive constants $f_{\mathbf{Z},\min}$ and $f_{\mathbf{Z},\max}$ such that, for every $\mathbf{z} \in \mathcal{Z}$, $f_{\mathbf{Z},\min} \leq f_{\mathbf{Z}}(\mathbf{z}) \leq f_{\mathbf{Z},\max}$.

Proposition 2. Under Assumptions 3.1-3.3 and if $C_{K,\alpha}h^\alpha/\alpha! < f_{\mathbf{Z},\min}$, for any $\mathbf{z} \in \mathcal{Z}$, the estimator $\hat{f}_{\mathbf{Z}}(\mathbf{z})$ is strictly positive with a probability larger than

$$1 - 2 \exp \left(-nh^p (f_{\mathbf{Z},\min} - C_{K,\alpha}h^\alpha/\alpha!)^2 / (2f_{\mathbf{Z},\max} \int K^2 + (2/3)C_K(f_{\mathbf{Z},\min} - C_{K,\alpha}h^\alpha/\alpha!)) \right).$$

The latter proposition is proved in A.2. It guarantees that our estimators $\hat{\tau}_{1,2|\mathbf{Z}}^{(k)}$, $k = 1, \dots, 3$, are well-behaved with a probability close to one. The next regularity assumption is necessary to explicitly control the bias of $\hat{\tau}_{1,2|\mathbf{Z}=\mathbf{z}}$.

Assumption 3.4. For every $\mathbf{x} \in \mathbb{R}^2$, $\mathbf{z} \mapsto f_{\mathbf{X},\mathbf{Z}}(\mathbf{x}, \mathbf{z})$ is differentiable on \mathcal{Z} almost everywhere up to the order α . For every $0 \leq k \leq \alpha$ and every $1 \leq i_1, \dots, i_\alpha \leq p$, let

$$\mathcal{H}_{k,\vec{i}}(\mathbf{u}, \mathbf{v}, \mathbf{x}_1, \mathbf{x}_2, \mathbf{z}) := \sup_{t \in [0,1]} \left| \frac{\partial^k f_{\mathbf{X},\mathbf{Z}}}{\partial z_{i_1} \dots \partial z_{i_k}}(\mathbf{x}_1, \mathbf{z} + t\mathbf{u}) \frac{\partial^{\alpha-k} f_{\mathbf{X},\mathbf{Z}}}{\partial z_{i_{k+1}} \dots \partial z_{i_\alpha}}(\mathbf{x}_2, \mathbf{z} + t\mathbf{v}) \right|,$$

denoting $\vec{i} = (i_1, \dots, i_\alpha)$. Assume that $\mathcal{H}_{k,\vec{i}}(\mathbf{u}, \mathbf{v}, \mathbf{x}_1, \mathbf{x}_2, \mathbf{z})$ is integrable and there exists a finite constant $C_{\mathbf{XZ},\alpha} > 0$ such that, for every $\mathbf{z} \in \mathcal{Z}$ and every $h < 1$,

$$\int |K|(\mathbf{u})|K|(\mathbf{v}) \sum_{k=0}^{\alpha} \binom{\alpha}{k} \sum_{i_1, \dots, i_\alpha=1}^p \mathcal{H}_{k,\vec{i}}(\mathbf{u}, \mathbf{v}, \mathbf{x}_1, \mathbf{x}_2, \mathbf{z}) |u_{i_1} \dots u_{i_k} v_{i_{k+1}} \dots v_{i_\alpha}| d\mathbf{u} d\mathbf{v} d\mathbf{x}_1 d\mathbf{x}_2$$

is less than $C_{\mathbf{XZ},\alpha}$.

The next three propositions state pointwise and uniform exponential inequalities for the estimators $\hat{\tau}_{1,2|\mathbf{Z}=\mathbf{z}}^{(k)}$, when $k = 1, 2, 3$. They are proved in A.3. We will denote $c_1 := c_3 := 4$ and $c_2 := 2$.

Proposition 3 (Exponential bound with explicit constants). Under Assumptions 3.1-3.4, for every $t > 0$ such that $C_{K,\alpha}h^\alpha/\alpha! + t \leq f_{\mathbf{Z},\min}/2$ and every $t' > 0$, if $C_{\tilde{K},2}h^2 < f_{\mathbf{Z}}(\mathbf{z})$, we have

$$\begin{aligned} & \mathbb{P} \left(\left| \hat{\tau}_{1,2|\mathbf{Z}=\mathbf{z}}^{(k)} - \tau_{1,2|\mathbf{Z}=\mathbf{z}} \right| > \frac{c_k}{f_{\mathbf{Z}}^2(\mathbf{z})} \left(\frac{C_{\mathbf{XZ},\alpha}h^\alpha}{\alpha!} + \frac{3f_{\mathbf{Z}}(\mathbf{z}) \int K^2}{2nh^p} + t' \right) \times \left(1 + \frac{16f_{\mathbf{Z}}^2(\mathbf{z})}{f_{\mathbf{Z},\min}^3} \left(\frac{C_{K,\alpha}h^\alpha}{\alpha!} + t \right) \right) \right) \\ & \leq 2 \exp \left(- \frac{nh^p t^2}{2f_{\mathbf{Z},\max} \int K^2 + (2/3)C_K t} \right) + 2 \exp \left(- \frac{(n-1)h^{2p} t'^2}{4f_{\mathbf{Z},\max}^2 (\int K^2)^2 + (8/3)C_K^2 t'} \right) \\ & + 2 \exp \left(- \frac{nh^p (f_{\mathbf{Z}}(\mathbf{z}) - C_{\tilde{K},2}h^2)^2}{8f_{\mathbf{Z},\max} \int \tilde{K}^2 + 4C_{\tilde{K}}(f_{\mathbf{Z}}(\mathbf{z}) - C_{\tilde{K},2}h^2)/3} \right), \end{aligned}$$

for any $\mathbf{z} \in \mathcal{Z}$ and every $k = 1, 2, 3$.

Alternatively, we can apply Theorem 1 in Major [24] instead of the Bernstein-type inequality that has been used in the proof of Proposition 3.

Proposition 4 (Alternative exponential bound without explicit constants). *Under Assumptions 3.1-3.4, for every $t > 0$ such that $C_{K,\alpha}h^\alpha/\alpha! + t \leq f_{\mathbf{z},\min}/2$ and every $t' > 0$ s.t. $t' \leq 2h^p(\int K^2)^3 f_{\mathbf{z},\max}^3/C_K^4$, there exist some universal constants C_2 and α_2 s.t.*

$$\begin{aligned} & \mathbb{P}\left(|\hat{\tau}_{1,2|\mathbf{Z}=\mathbf{z}}^{(k)} - \tau_{1,2|\mathbf{Z}=\mathbf{z}}| > \frac{c_k}{f_{\mathbf{z}}^2(\mathbf{z})} \left(\frac{C_{\mathbf{XZ},\alpha}h^\alpha}{\alpha!} + \frac{3f_{\mathbf{z}}(\mathbf{z}) \int K^2}{2nh^p} + t' \right) \times \left(1 + \frac{16f_{\mathbf{z}}^2(\mathbf{z})}{f_{\mathbf{z},\min}^3} \left(\frac{C_{K,\alpha}h^\alpha}{\alpha!} + t \right) \right)\right) \\ & \leq 2 \exp\left(-\frac{nh^p t^2}{2f_{\mathbf{z},\max} \int K^2 + (2/3)C_K t}\right) + 2 \exp\left(-\frac{nh^p(f_{\mathbf{z}}(\mathbf{z}) - C_{\tilde{K},2}h^2)^2}{8f_{\mathbf{z},\max} \int \tilde{K}^2 + 4C_{\tilde{K}}(f_{\mathbf{z}}(\mathbf{z}) - C_{\tilde{K},2}h^2)/3}\right) \\ & + 2 \exp\left(\frac{nh^p t^2}{32 \int K^2 (\int |K|)^2 f_{\mathbf{z},\max}^3 + 8C_K \int |K| f_{\mathbf{z},\max} t/3}\right) + C_2 \exp\left(-\frac{\alpha_2 nh^p t'}{8f_{\mathbf{z},\max} (\int K^2)}\right), \end{aligned}$$

for any $\mathbf{z} \in \mathcal{Z}$ and every $k = 1, 2, 3$, if $C_{\tilde{K},2}h^2 < f_{\mathbf{z}}(\mathbf{z})$ and $6h^p(\int |K|)^2 f_{\mathbf{z},\max} < \int K^2$.

Remark 5. In Propositions 2, 3 and 4, when the support of K is included in $[-c, c]^p$ for some $c > 0$, $f_{\mathbf{z},\max}$ can be replaced by a local bound $\sup_{\tilde{\mathbf{z}} \in \mathcal{V}(\mathbf{z}, \epsilon)} f_{\mathbf{z}}(\tilde{\mathbf{z}})$, denoting by $\mathcal{V}(\mathbf{z}, \epsilon)$ a closed ball of center \mathbf{z} and any radius $\epsilon > 0$, when $hc < \epsilon$.

As a corollary, the two latter result yield the weak consistency of $\hat{\tau}_{1,2|\mathbf{Z}=\mathbf{z}}^{(k)}$ for every $\mathbf{z} \in \mathcal{Z}$, when $nh^{2p} \rightarrow \infty$ (choose the constants t and $t' \sim h^p$ sufficiently small, in Proposition 4, e.g.).

It is possible to obtain uniform bounds, by slightly strengthening our assumptions. Note that this next result will be true if n is sufficiently large, when Proposition 4 was true for every n .

Assumption 3.5. *The kernel K is Lipschitz on $(\mathcal{Z}, \|\cdot\|_\infty)$, with a constant λ_K and \mathcal{Z} is a subset of an hypercube in \mathbb{R}^p whose volume is denoted by \mathcal{V} . Moreover, K and K^2 are regular in the sense of [25] or [26].*

Proposition 6 (Uniform exponential bound). *Under the assumptions 3.1-3.5, there exist some constants L_K and \bar{C}_K (resp. $L_{\tilde{K}}$ and $\bar{C}_{\tilde{K}}$) that depend only on the VC characteristics of K (resp. \tilde{K}), s.t., for every $\mu \in (0, 1)$ such that $\mu f_{\mathbf{z},\min} < C_{\mathbf{XZ},\alpha}h^\alpha/\alpha! + b_K \int K^2 f_{\mathbf{z},\max}/C_K$, if $f_{\mathbf{z},\max} < \tilde{C}_{\mathbf{XZ},2}h^2/2 + b_{\tilde{K}} \int \tilde{K}^2 f_{\mathbf{z},\max}/C_{\tilde{K}}$,*

$$\begin{aligned} & \mathbb{P}\left(\sup_{\mathbf{z} \in \mathcal{Z}} |\hat{\tau}_{1,2|\mathbf{Z}=\mathbf{z}}^{(k)} - \tau_{1,2|\mathbf{Z}=\mathbf{z}}| > \frac{c_k}{f_{\mathbf{z},\min}^2(1-\mu)^2} \left(\frac{C_{\mathbf{XZ},\alpha}h^\alpha}{\alpha!} + \frac{3f_{\mathbf{z},\max} \int K^2}{2nh^p} + t \right)\right) \\ & \leq L_K \exp\left(-C_{f,K}nh^p \left(\mu f_{\mathbf{z},\min} - \frac{C_{\mathbf{XZ},\alpha}h^\alpha}{\alpha!}\right)^2\right) \\ & + C_2 D \exp\left(-\frac{\alpha_2 nth^p}{8f_{\mathbf{z},\max} (\int K^2)}\right) + L_{\tilde{K}} \exp\left(-C_{f,\tilde{K}}nh^p (f_{\mathbf{z},\max} - \tilde{C}_{\mathbf{XZ},2}h^2)^2/4\right) \\ & + 2 \exp\left(-\frac{A_2 nh^p t^2 C_K^{-4}}{16^2 A_1^2 \int K^2 f_{\mathbf{z},\max}^3 (\int |K|)^2}\right) + 2 \exp\left(-\frac{A_2 nh^p t}{16 C_K^2 A_1}\right), \end{aligned}$$

for n sufficiently large, $k = 1, 2, 3$, and for every $t > 0$ s.t. $t \leq 2h^p(\int K^2)^3 f_{\mathbf{z},\max}^3/C_K^4$,

$$-16A_1 C_K^2 A_{\bar{g}} \int K^2 f_{\mathbf{z},\max}^3 \left(\int |K| \right)^2 \ln(h^p \int K^2 f_{\mathbf{z},\max}^3 \left(\int |K| \right)^2) < n^{1/2} h^{p/2} t, \text{ and}$$

$$nh^p t \geq \left(\int K^2 \right) f_{\mathbf{z},\max} M_2(p + \beta)^{3/2} \log\left(\frac{4C_K^2}{h^p f_{\mathbf{z},\max} \int K^2}\right), \quad \beta = \max\left(0, \frac{\log D}{\log n}\right), \quad D := \lceil \mathcal{V} \left(\frac{4C_K \lambda_K}{h} \right)^p \rceil,$$

for some universal constants $C_2, \alpha_2, M_2, A_1, A_2$ and a constant $A_{\bar{g}}$ that depends on K and $f_{\mathbf{z},\max}$.

We have denoted $C_{f,K} := \log(1 + b_K/(4L_K))/(L_K b_K f_{\mathbf{z},max} \int K^2)$, for any arbitrarily chosen constant $b_K \geq \overline{C}_K$. Similarly, $C_{f,\tilde{K}} := \log(1 + b_{\tilde{K}}/(4L_{\tilde{K}}))/(L_{\tilde{K}} b_{\tilde{K}} f_{\mathbf{z},max} \int \tilde{K}^2)$, $b_{\tilde{K}} \geq \overline{C}_{\tilde{K}}$.

3.2. Asymptotic behavior

The previous exponential inequalities are not optimal to prove usual asymptotic results. Indeed, they directly or indirectly rely on upper bounds of estimates, as in Hoeffding or Bernstein-type inequalities. In the case of kernel estimates, this implies the necessary condition $nh^{2p} \rightarrow \infty$, at least. By a direct approach, it is possible to state the consistency of $\hat{\tau}_{1,2|\mathbf{Z}=\mathbf{z}}^{(k)}$, $k = 1, 2, 3$, and then of $\tilde{\tau}_{1,2|\mathbf{Z}=\mathbf{z}}$, under the weaker condition $nh^p \rightarrow \infty$.

Proposition 7 (Consistency). *Under Assumption 3.1, if $nh_n^p \rightarrow \infty$, $\lim K(\mathbf{t})|\mathbf{t}|^p = 0$ when $|\mathbf{t}| \rightarrow \infty$, $f_{\mathbf{z}}$ and $\mathbf{z} \mapsto \tau_{1,2|\mathbf{Z}=\mathbf{z}}$ are continuous on \mathcal{Z} , then $\hat{\tau}_{1,2|\mathbf{Z}=\mathbf{z}}^{(k)}$ tends to $\tau_{1,2|\mathbf{Z}=\mathbf{z}}$ in probability, when $n \rightarrow \infty$ for any $k = 1, 2, 3$.*

This property is proved in A.6. Moreover, Proposition 6 does not allow to state the strong uniform consistency of $\hat{\tau}_{1,2|\mathbf{Z}=\mathbf{z}}^{(k)}$ because the threshold t has to be of order h^p at most. Here again, a direct approach is possible, nonetheless.

Proposition 8 (Uniform consistency). *Under Assumption 3.1, assume that $nh_n^{2p}/\log n \rightarrow \infty$, $\lim K(\mathbf{t})|\mathbf{t}|^p = 0$ when $|\mathbf{t}| \rightarrow \infty$, K is Lipschitz, $f_{\mathbf{z}}$ and $\mathbf{z} \mapsto \tau_{1,2|\mathbf{Z}=\mathbf{z}}$ are continuous on a bounded set \mathcal{Z} , and there exists a lower bound $f_{\mathbf{z},min}$ s.t. $f_{\mathbf{z},min} \leq f_{\mathbf{z}}(\mathbf{z})$ for any $\mathbf{z} \in \mathcal{Z}$. Then $\sup_{\mathbf{z} \in \mathcal{Z}} |\hat{\tau}_{1,2|\mathbf{Z}=\mathbf{z}}^{(k)} - \tau_{1,2|\mathbf{Z}=\mathbf{z}}| \rightarrow 0$ almost surely, when $n \rightarrow \infty$ for any $k = 1, 2, 3$.*

This property is proved in A.7. To derive the asymptotic law of this estimator, we will assume:

Assumption 3.6. (i) $nh_n^p \rightarrow \infty$ and $nh_n^{p+2\alpha} \rightarrow 0$; (ii) $K(\cdot)$ is compactly supported.

Proposition 9 (Joint asymptotic normality at different points). *Let $\mathbf{z}'_1, \dots, \mathbf{z}'_{n'}$ be fixed points in a set $\mathcal{Z} \subset \mathbb{R}^p$. Assume 3.1, 3.4, 3.6, that the \mathbf{z}'_i are distinct and that $f_{\mathbf{z}}$ and $\mathbf{z} \mapsto f_{\mathbf{x},\mathbf{z}}(\mathbf{x}, \mathbf{z})$ are continuous on \mathcal{Z} , for every \mathbf{x} . Then, as $n \rightarrow \infty$,*

$$(nh_n^p)^{1/2} (\hat{\tau}_{1,2|\mathbf{Z}=\mathbf{z}'_i} - \tau_{1,2|\mathbf{Z}=\mathbf{z}'_i})_{i=1,\dots,n'} \xrightarrow{D} \mathcal{N}(0, \mathbb{H}^{(k)}), \quad k = 1, 2, 3,$$

where $\hat{\tau}_{1,2|\mathbf{Z}=\mathbf{z}}$ denotes any of the estimators $\hat{\tau}_{1,2|\mathbf{Z}=\mathbf{z}}^{(k)}$, $k = 1, 2, 3$ or $\tilde{\tau}_{1,2|\mathbf{Z}=\mathbf{z}}$, and \mathbb{H} is the $n' \times n'$ diagonal real matrix defined by

$$[\mathbb{H}^{(k)}]_{i,j} = \frac{4 \int K^2 \mathbb{1}_{\{i=j\}}}{f_{\mathbf{z}}(\mathbf{z}'_i)} \{ \mathbb{E}[g_k(\mathbf{X}_1, \mathbf{X})g_k(\mathbf{X}_2, \mathbf{X})|\mathbf{Z} = \mathbf{Z}_1 = \mathbf{Z}_2 = \mathbf{z}'_i] - \tau_{1,2|\mathbf{Z}=\mathbf{z}'_i}^2 \},$$

for every $1 \leq i, j \leq n'$, and (\mathbf{X}, \mathbf{Z}) , $(\mathbf{X}_1, \mathbf{Z}_1)$, $(\mathbf{X}_2, \mathbf{Z}_2)$ are independent versions.

This proposition is proved in A.8.

Remark 10. *The latter results will provide some simple tests of the constancy of the function $\mathbf{z} \mapsto \tau_{1,2|\mathbf{z}}$, and then of the constancy of the associated conditional copula itself. This would test the famous ‘‘simplifying assumption’’ ($\mathcal{H}_0 : C_{1,2|\mathbf{Z}=\mathbf{z}}$ does not depend on the choice of \mathbf{z}), a key assumption for vine modeling in particular: see [27] or [28] for a discussion, [29] for a review and a presentation of formal tests for this hypothesis.*

4. Simulation study

In this simulation study, we draw i.i.d. random samples $(X_{i,1}, X_{i,2}, Z_i)$, $i = 1, \dots, n$, with univariate explanatory variables ($p = 1$). We consider two settings, that correspond to bounded and/or unbounded explanatory variables respectively:

1. $\mathcal{Z} =]0, 1[$ and the law of Z is uniform on $]0, 1[$. Conditionally on $Z = z$, $X_1|Z = z$ and $X_2|Z = z$ both follow a Gaussian distribution $\mathcal{N}(z, 1)$. Their associated conditional copula is Gaussian and their conditional Kendall's tau is given by $\tau_{1,2|Z=z} = 2z - 1$.
2. $\mathcal{Z} = \mathbb{R}$ and the law of Z is $\mathcal{N}(0, 1)$. Conditionally on $Z = z$, $X_1|Z = z$ and $X_2|Z = z$ both follow a Gaussian distribution $\mathcal{N}(\Phi(z), 1)$, where $\Phi(\cdot)$ is the cdf of the \mathbf{Z} . Their associated conditional copula is Gaussian and their conditional Kendall's tau is given by $\tau_{1,2|Z=z} = 2\Phi(z) - 1$.

These simple frameworks allow us to compare the numerical properties of our different estimators in different parts of the space, in particular when Z is close to zero or one, i.e. when the conditional Kendall's tau is close to -1 or to 1 . We compute the different estimators $\hat{\tau}_{1,2|\mathbf{Z}=z}^{(k)}$ for $k = 1, 2, 3$, and the symmetrically rescaled version $\tilde{\tau}_{1,2|z}$. The bandwidth h is chosen as proportional to the usual "rule-of-thumb" for kernel density estimation, i.e. $h = \alpha_h \hat{\sigma}(Z) n^{-1/5}$ with $\alpha_h \in \{0.5, 0.75, 1, 1.5, 2\}$ and $n \in \{100, 500, 1000, 2000\}$. For each setting, we consider three local measures of goodness-of-fit: for a given z and for any Kendall's tau estimate (say $\hat{\tau}_{1,2|Z=z}$), let

- the (local) bias: $Bias(z) := \mathbb{E}[\hat{\tau}_{1,2|Z=z}] - \tau_{1,2|Z=z}$,
- the (local) standard deviation: $Sd(z) := \mathbb{E}\left[\left(\hat{\tau}_{1,2|Z=z} - \mathbb{E}[\hat{\tau}_{1,2|Z=z}]\right)^2\right]^{1/2}$,
- the (local) mean square-error: $MSE(z) := \mathbb{E}\left[\left(\hat{\tau}_{1,2|Z=z} - \tau_{1,2|Z=z}\right)^2\right]$.

We also consider their integrated version w.r.t the usual Lebesgue measure on the whole support of z , respectively denoted by $IBias$, ISd and $IMSE$. Some results concerning these integrated measures are given in Table 1 (resp. Table 2) for Setting 1 (resp. Setting 2), and for different choices of α_h and n . For the sake of effective calculations of these measures, all the theoretical previous expectations are replaced by their empirical counterparts based on 500 simulations.

For every n , the best results seem to be obtained with $\alpha_h = 1.5$ and the fourth (rescaled) estimator, particularly in terms of bias. This is not so surprising, because the estimators $\hat{\tau}^{(k)}$, $k = 1, 2, 3$, do not have the right support at a finite distance. Note that this comparative advantage of $\tilde{\tau}$ in terms of bias decreases with n , as expected. In terms of integrated variance, all the considered estimators behave more or less similarly, particularly when $n \geq 500$.

To illustrate our results for Setting 1 (resp. Setting 2), the functions $z \mapsto Bias(z)$, $Sd(z)$ and $MSE(z)$ have been plotted on Figures 1-2 (resp. Figures 3-4), both with our empirically optimal choice $\alpha_h = 1.5$. We can note that, considering the bias, the estimator $\tilde{\tau}$ behaves similarly as $\hat{\tau}^{(1)}$ when the true τ is close to -1 , and similarly as $\hat{\tau}^{(3)}$ when the true Kendall's tau is close to 1 . But globally, the best pointwise estimator is clearly obtained with the rescaled version $\tilde{\tau}_{1,2|\mathbf{Z}=\cdot}$, after a quick inspection of MSE levels, and even if the differences between our four estimators weaken for large sample sizes. The comparative advantage of $\tilde{\tau}_{1,2|z}$ more clearly appears with Setting 2 than with Setting 1. Indeed, in the former case, the support of \mathbf{Z} 's distribution is the whole line. Then $\hat{f}_{\mathbf{Z}}$ does not suffer any more from the boundary bias phenomenon,

contrary to what happened with Setting 1. As a consequence, the biases induced by the definitions of $\hat{\tau}_{1,2|z}^{(k)}$, $k = 1, 3$, appear more strikingly in Figure 3, for instance: when z is close to (-1) (resp. 1), the biases of $\hat{\tau}_{1,2|z}^{(1)}$ (resp. $\hat{\tau}_{1,2|z}^{(3)}$) and $\tilde{\tau}_{1,2|z}$ are close, when the bias $\hat{\tau}_{1,2|z}^{(3)}$ (resp. $\hat{\tau}_{1,2|z}^{(1)}$) is a lot larger. Since the squared biases are here significantly larger than the variances in the tails, $\tilde{\tau}_{1,2|z}$ provides the best estimator globally considering "both sides" together. But even in the center of \mathbf{Z} 's distribution, the latter estimator behaves very well.

In Setting 2 where there is no boundary problem, we also try to estimate the conditional Kendall's tau using our cross-validation criterion (4), with $N_{pairs} = 1000$. More precisely, denoting by h^{CV} the minimizer of the cross-validation criterion, we try different choices $h = \alpha_h \times h^{CV}$ with $\alpha_h \in \{0.5, 0.75, 1, 1.5, 2\}$. The results in terms of integrated bias, standard deviation and MSE are given in Table 3. We do not find any substantial improvements compared to the previous Table 2, where the bandwidth was chosen "roughly". In Table 4, we compare the average h^{CV} with the previous choice of h . The expectation of h^{CV} is always higher than the "rule-of-thumb" h^{ref} , but the difference between both decreases when the sample size n increases. The standard deviation of h^{CV} is quite high for low values of n , but decreases as a function of n . This may be seen as quite surprising given the fact that the number of pairs N_{pairs} used in the computation of the criterion stays constant. Nevertheless, when the sample size increases, the selected pairs are better in the sense that the differences $|Z_i - Z_j|$ can become smaller as more replications of Z_i are available.

References

- [1] R. Nelsen, An introduction to copulas, Springer Science & Business Media, 2007.
- [2] W. Kruskal, Ordinal measures of association, J. Amer. Statist. Ass. 53 (284) (1958) 814–861.
- [3] M. Hollander, D. Wolfe, Nonparametric Statistical Methods, Wiley, 1973.
- [4] E. Lehmann, Nonparametrics: Statistical Methods Based on Ranks., Holden-Day, 1975.
- [5] H. Joe, Multivariate models and multivariate dependence concepts, Chapman and Hall/CRC, 1997.
- [6] A. Patton, Estimation of multivariate models for time series of possibly different lengths, J. Appl. Econometrics 21 (2) (2006) 147–173.
- [7] A. Patton, Modelling asymmetric exchange rate dependence, Internat. Econom. Rev. 47 (2) (2006) 527–556.
- [8] I. Gijbels, N. Veraverbeke, M. Omelka, Conditional copulas, association measures and their applications, Comput. Statist. Data Anal. 55 (5) (2011) 1919–1932.
- [9] N. Veraverbeke, M. Omelka, I. Gijbels, Estimation of a conditional copula and association measures, Scand. J. Stat. 38 (4) (2011) 766–780.
- [10] J.-D. Fermanian, M. Wegkamp, Time-dependent copulas, J. Multivariate Anal. 110 (2012) 19–29.
- [11] J.-D. Fermanian, O. Lopez, Single-index copulas, J. Multivariate Anal. 165 (2018) 27–55.
- [12] W.-Y. Tsai, Testing the assumption of independence of truncation time and failure time, Biometrika 77 (1) (1990) 169–177.
- [13] E. C. Martin, R. A. Betensky, Testing quasi-independence of failure and truncation times via conditional kendall's tau, Journal of the American Statistical Association 100 (470) (2005) 484–492.
- [14] L. Lakhal, L.-P. Rivest, B. Abdous, Estimating survival and association in a semicompeting risks model, Biometrics 64 (1) (2008) 180–188.

		n = 100			n = 500			n = 1000			n = 2000		
		IBias	ISd	IMSE	IBias	ISd	IMSE	IBias	ISd	IMSE	IBias	ISd	IMSE
$\alpha_h = 0.5$	$\hat{\tau}_{1,2 Z=z}^{(1)}$	-133	197	66.5	-34.5	84.9	9.86	-18.2	61.6	4.85	-10.9	46	2.65
	$\hat{\tau}_{1,2 Z=z}^{(2)}$	-12.9	187	43.7	-4.08	84.4	8.58	-0.9	61.5	4.49	-1.07	46	2.53
	$\hat{\tau}_{1,2 Z=z}^{(3)}$	107	190	56.6	26.4	84.5	9.26	16.4	61.5	4.76	8.8	46	2.6
	$\tilde{\tau}_{1,2 Z=z}$	-0.91	213	48.2	-1.18	86.9	8.55	0.733	62.4	4.46	-0.149	46.4	2.5
$\alpha_h = 0.75$	$\hat{\tau}_{1,2 Z=z}^{(1)}$	-88	150	35.8	-26.3	68	6.32	-13.9	50.7	3.33	-7.98	37.6	1.8
	$\hat{\tau}_{1,2 Z=z}^{(2)}$	-10.4	145	26.3	-5.97	67.9	5.6	-2.33	50.6	3.12	-1.39	37.5	1.74
	$\hat{\tau}_{1,2 Z=z}^{(3)}$	67.2	146	30.6	14.3	67.9	5.75	9.2	50.6	3.19	5.2	37.5	1.76
	$\tilde{\tau}_{1,2 Z=z}$	-2.06	157	26.7	-3.99	69.2	5.49	-1.21	51.2	3.05	-0.76	37.8	1.69
$\alpha_h = 1$	$\hat{\tau}_{1,2 Z=z}^{(1)}$	-67.8	123	24.5	-19.2	58.7	4.8	-11	43.1	2.52	-6.34	33	1.44
	$\hat{\tau}_{1,2 Z=z}^{(2)}$	-9.99	121	19	-3.95	58.6	4.39	-2.35	43.1	2.39	-1.39	33	1.4
	$\hat{\tau}_{1,2 Z=z}^{(3)}$	47.8	122	20.9	11.3	58.7	4.47	6.34	43.1	2.41	3.57	33	1.41
	$\tilde{\tau}_{1,2 Z=z}$	-3.48	128	18.1	-2.34	59.5	4.18	-1.46	43.4	2.29	-0.897	33.2	1.35
$\alpha_h = 1.5$	$\hat{\tau}_{1,2 Z=z}^{(1)}$	-44.6	101	17.5	-15.9	50.4	4.12	-9.7	35.9	2.13	-5.52	27.6	1.28
	$\hat{\tau}_{1,2 Z=z}^{(2)}$	-5.81	100	14.9	-5.68	50.3	3.84	-3.84	35.9	2.02	-2.18	27.6	1.24
	$\hat{\tau}_{1,2 Z=z}^{(3)}$	33	101	15.5	4.58	50.3	3.77	2.01	35.9	1.99	1.15	27.6	1.23
	$\tilde{\tau}_{1,2 Z=z}$	-1.09	104	13.4	-4.55	50.8	3.57	-3.19	36.1	1.9	-1.83	27.7	1.18
$\alpha_h = 2$	$\hat{\tau}_{1,2 Z=z}^{(1)}$	-37.8	91.4	17.3	-11.8	43.8	4.14	-7.2	31.2	2.35	-5.97	23.7	1.43
	$\hat{\tau}_{1,2 Z=z}^{(2)}$	-8.03	91.4	15.4	-3.93	43.8	3.94	-2.75	31.2	2.28	-3.44	23.7	1.39
	$\hat{\tau}_{1,2 Z=z}^{(3)}$	21.7	91.7	15.4	3.91	43.8	3.87	1.7	31.2	2.24	-0.912	23.7	1.37
	$\tilde{\tau}_{1,2 Z=z}$	-4.5	94.2	13.5	-3.01	44.1	3.62	-2.24	31.3	2.12	-3.16	23.8	1.32

TABLE 1

Results of the simulation in Setting 1. All values have been multiplied by 1000. Bold values indicate optimal choices for the chosen measure of performance.

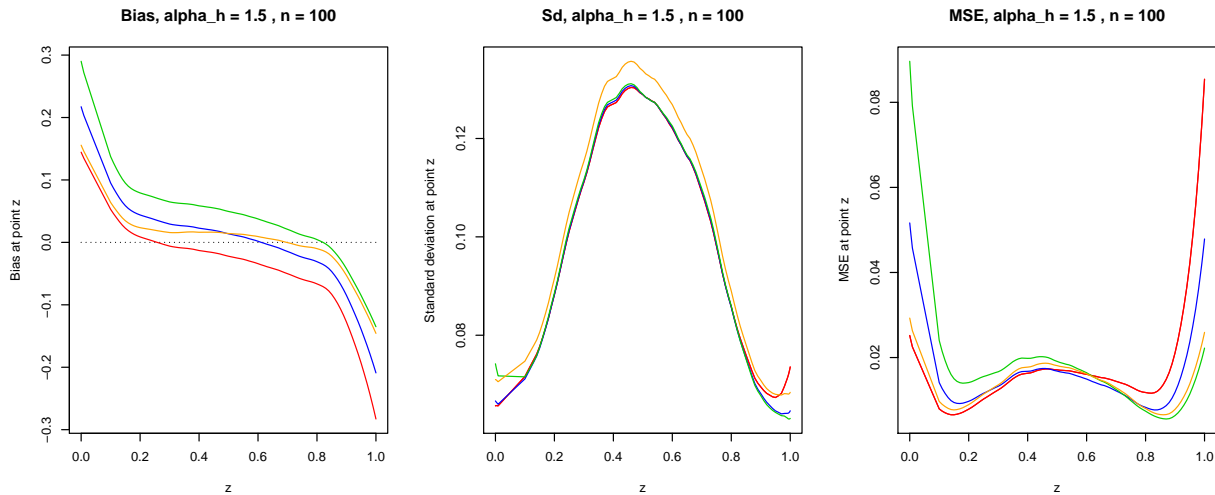


FIG 1. Local bias, standard deviation and MSE for the estimators $\hat{\tau}^{(1)}$ (red), $\hat{\tau}^{(2)}$ (blue), $\hat{\tau}^{(3)}$ (green), $\tilde{\tau}$ (orange), with $n = 100$ and $\alpha_h = 1.5$ in Setting 1. The dotted line on the first figure is the reference at 0.

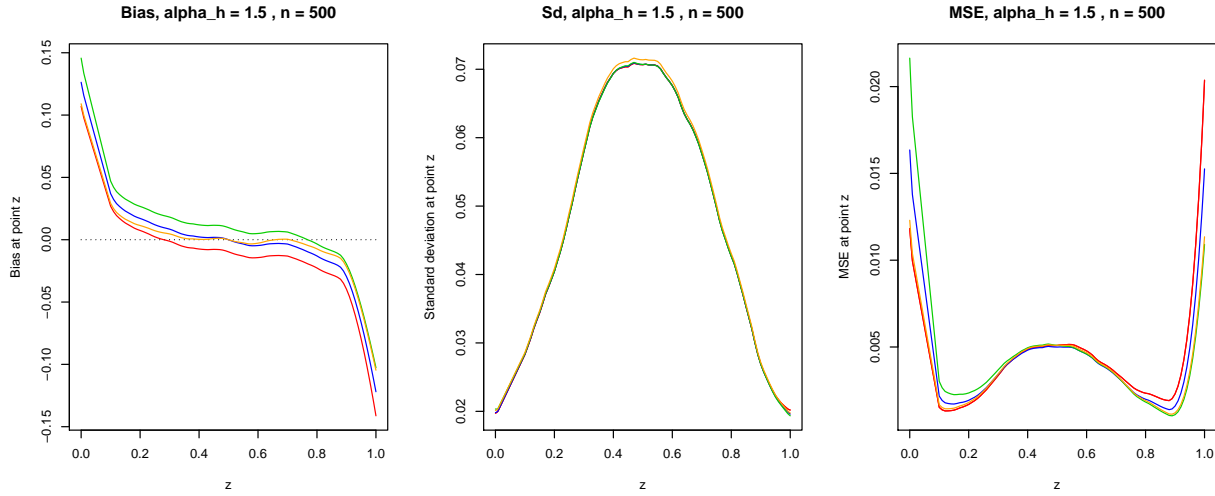


FIG 2. Local bias, standard deviation and MSE for the estimators $\hat{\tau}^{(1)}$ (red), $\hat{\tau}^{(2)}$ (blue), $\hat{\tau}^{(3)}$ (green), $\tilde{\tau}$ (orange), with $n = 500$ and $\alpha_h = 1.5$ in Setting 1. The dotted line on the first figure is the reference at 0.

		$n = 100$			$n = 500$			$n = 1000$			$n = 2000$		
		IBias	ISd	IMSE	IBias	ISd	IMSE	IBias	ISd	IMSE	IBias	ISd	IMSE
$\alpha_h = 0.5$	$\hat{\tau}_{1,2 Z=z}^{(1)}$	-207	227	180	-54.1	83.9	16.9	-29.6	55.3	5.81	-16.9	38.9	2.49
	$\hat{\tau}_{1,2 Z=z}^{(2)}$	1.15	207	97	0.845	80.5	10.8	0.557	54.4	4.35	0.145	38.6	2.04
	$\hat{\tau}_{1,2 Z=z}^{(3)}$	210	228	181	55.7	83.2	16.4	30.7	55.4	5.9	17.2	38.9	2.5
	$\hat{\tau}_{1,2 Z=z}^{(4)}$	1.4	225	51.9	0.987	81.4	6.86	0.456	55	3.22	0.175	38.9	1.66
$\alpha_h = 0.75$	$\hat{\tau}_{1,2 Z=z}^{(1)}$	-144	175	98.6	-33.3	60.6	7.5	-19.8	41.9	3.12	-10.6	30.5	1.42
	$\hat{\tau}_{1,2 Z=z}^{(2)}$	-2.33	163	56.2	1.73	59.4	5.56	-0.0619	41.7	2.51	0.665	30.4	1.24
	$\hat{\tau}_{1,2 Z=z}^{(3)}$	140	176	99.2	36.8	60.7	7.73	19.7	42.1	3.12	11.9	30.5	1.45
	$\hat{\tau}_{1,2 Z=z}^{(4)}$	-3.15	170	30.3	1.69	60.2	3.85	-0.093	42.1	1.95	0.645	30.5	1.05
$\alpha_h = 1$	$\hat{\tau}_{1,2 Z=z}^{(1)}$	-99.8	143	57.7	-24.9	50.9	5.06	-13.5	36.6	2.28	-6.92	26.6	1.09
	$\hat{\tau}_{1,2 Z=z}^{(2)}$	1.17	132	34.6	0.903	50.4	4.02	1.16	36.5	1.97	1.46	26.6	0.994
	$\hat{\tau}_{1,2 Z=z}^{(3)}$	102	139	54.4	26.7	51	5.13	15.8	36.6	2.33	9.83	26.6	1.11
	$\hat{\tau}_{1,2 Z=z}^{(4)}$	2.51	138	20.1	0.897	50.9	2.89	1.16	36.7	1.56	1.48	26.7	0.847
$\alpha_h = 1.5$	$\hat{\tau}_{1,2 Z=z}^{(1)}$	-59.1	104	28.1	-14.7	42.3	3.87	-7.56	29.7	1.86	-4.17	21.8	0.932
	$\hat{\tau}_{1,2 Z=z}^{(2)}$	4.34	99.7	21.4	2.05	42.1	3.48	2.07	29.6	1.75	1.35	21.8	0.899
	$\hat{\tau}_{1,2 Z=z}^{(3)}$	67.8	103	29.6	18.8	42.3	3.96	11.7	29.6	1.92	6.87	21.8	0.957
	$\hat{\tau}_{1,2 Z=z}^{(4)}$	3.34	103	13.4	2.08	42.5	2.6	2.08	29.7	1.39	1.35	21.8	0.755
$\alpha_h = 2$	$\hat{\tau}_{1,2 Z=z}^{(1)}$	-37.2	88.2	23.9	-9.57	38.2	4.6	-3.75	26.2	2.34	-1.09	19.8	1.32
	$\hat{\tau}_{1,2 Z=z}^{(2)}$	8.17	85.9	21.2	2.69	38	4.45	3.32	26.1	2.3	2.99	19.8	1.32
	$\hat{\tau}_{1,2 Z=z}^{(3)}$	53.5	87.4	25.3	14.9	38.1	4.74	10.4	26.2	2.41	7.08	19.8	1.36
	$\hat{\tau}_{1,2 Z=z}^{(4)}$	8.47	88.5	15	2.69	38.4	3.59	3.33	26.3	1.93	3	19.9	1.15

TABLE 2

Results of the simulation in Setting 2. All values have been multiplied by 1000. Bold values indicate optimal choices for the chosen measure of performance.

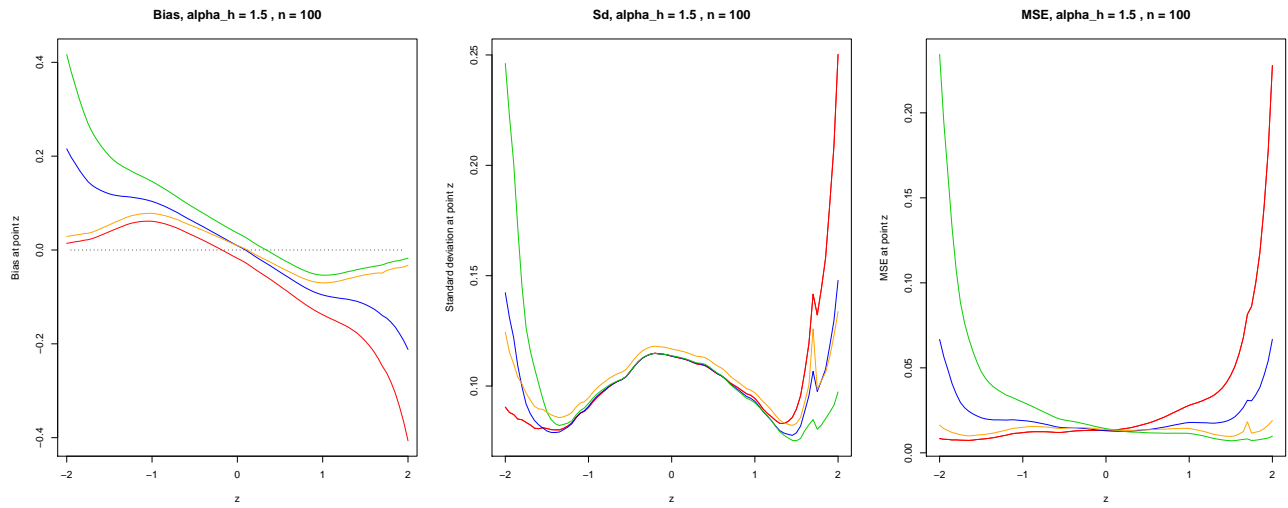


FIG 3. Local bias, standard deviation and MSE for the estimators $\hat{\tau}^{(1)}$ (red), $\hat{\tau}^{(2)}$ (blue), $\hat{\tau}^{(3)}$ (green), $\tilde{\tau}$ (orange), with $n = 100$ and $\alpha_h = 1.5$ in Setting 2. The dotted line on the first figure is the reference at 0.

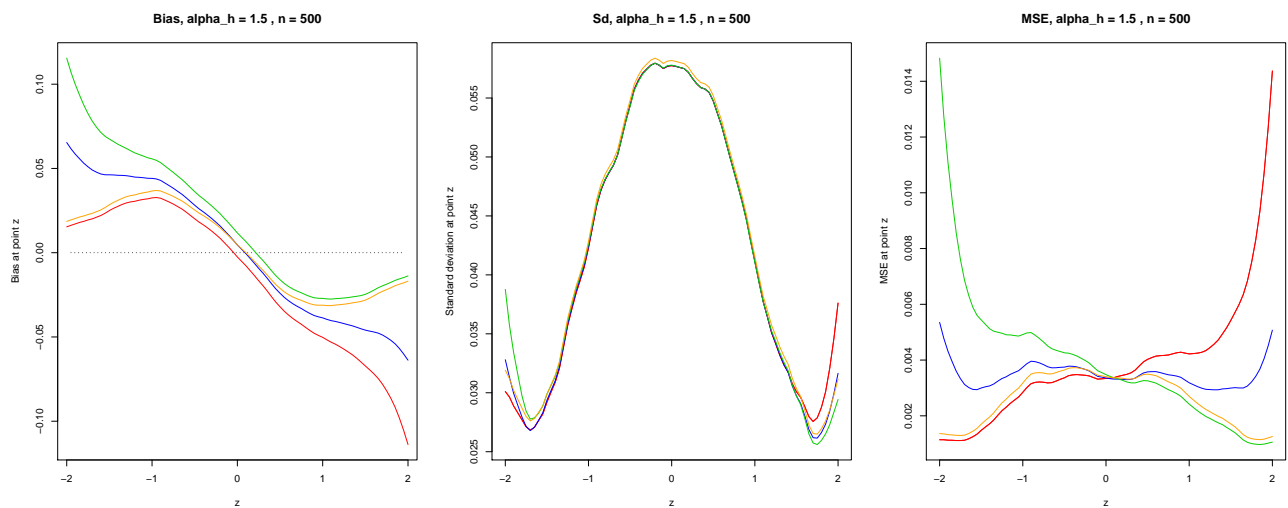


FIG 4. Local bias, standard deviation and MSE for the estimators $\hat{\tau}^{(1)}$ (red), $\hat{\tau}^{(2)}$ (blue), $\hat{\tau}^{(3)}$ (green), $\tilde{\tau}$ (orange), with $n = 500$ and $\alpha_h = 1.5$ in Setting 2. The dotted line on the first figure is the reference at 0.

		$n = 100$			$n = 500$			$n = 1000$			$n = 2000$		
		IBias	ISd	IMSE	IBias	ISd	IMSE	IBias	ISd	IMSE	IBias	ISd	IMSE
$\alpha_h = 0.5$	$\hat{\tau}_{1,2 Z=z}^{(1)}$	-111	154	66.2	-36.9	66.8	9.01	-22.4	48.2	4.06	-12.9	36.1	2.04
	$\hat{\tau}_{1,2 Z=z}^{(2)}$	0.0488	137	36.3	0.236	64.2	6.45	0.546	46.8	3.14	1.29	35.7	1.78
	$\hat{\tau}_{1,2 Z=z}^{(3)}$	111	151	60.6	37.4	66.3	8.88	23.5	47.2	4.07	15.5	36.2	2.18
	$\hat{\tau}_{1,2 Z=z}^{(4)}$	1.38	132	18.3	0.27	64.5	4.49	0.61	46.8	2.36	1.29	35.6	1.49
$\alpha_h = 0.75$	$\hat{\tau}_{1,2 Z=z}^{(1)}$	-67.4	117	35.7	-23.3	52.1	5.27	-13.9	37.8	2.4	-7.6	29	1.3
	$\hat{\tau}_{1,2 Z=z}^{(2)}$	4.32	108	23.5	0.809	50.7	4.21	1.03	37.2	2.07	1.78	28.8	1.21
	$\hat{\tau}_{1,2 Z=z}^{(3)}$	76.1	119	35.4	24.9	51.6	5.12	16	37.6	2.49	11.2	29.1	1.39
	$\hat{\tau}_{1,2 Z=z}^{(4)}$	4.98	106	13.3	0.86	51.6	3.13	1.03	37.5	1.63	1.81	28.9	1.02
$\alpha_h = 1$	$\hat{\tau}_{1,2 Z=z}^{(1)}$	-43	101	28	-15.8	45.7	4.44	-9.51	33.1	2.04	-4.68	25.1	1.07
	$\hat{\tau}_{1,2 Z=z}^{(2)}$	7.87	93.1	22.4	2.01	44.8	3.91	1.57	32.7	1.87	2.29	24.9	1.03
	$\hat{\tau}_{1,2 Z=z}^{(3)}$	58.8	97.6	27.2	19.8	45.3	4.41	12.7	32.9	2.1	9.27	25.1	1.14
	$\hat{\tau}_{1,2 Z=z}^{(4)}$	8.51	98	15.7	2.05	46	3.01	1.57	33.1	1.5	2.33	25.1	0.871
$\alpha_h = 1.5$	$\hat{\tau}_{1,2 Z=z}^{(1)}$	-16.1	95.6	41.7	-6.36	43	6.35	-4.04	30.6	2.87	-1.11	22.1	1.34
	$\hat{\tau}_{1,2 Z=z}^{(2)}$	14.9	92.6	40.4	5.08	42.6	6.2	3.17	30.4	2.83	3.47	22	1.34
	$\hat{\tau}_{1,2 Z=z}^{(3)}$	46	92.8	42.2	16.5	42.6	6.45	10.4	30.4	2.94	8.06	22.1	1.4
	$\hat{\tau}_{1,2 Z=z}^{(4)}$	15.6	100	35.2	5.11	44	5.31	3.17	31	2.45	3.5	22.4	1.17

TABLE 3

Results of the simulation in Setting 2 using $h = \alpha_h \times h^{CV}$ where h^{CV} has been chosen by cross-validation. All values have been multiplied by 1000. Bold values indicate optimal choices for the chosen measure of performance.

n	100	500	1000	2000
$\mathbb{E}[h^{CV}]$	0.77	0.43	0.34	0.27
$Sd[h^{CV}]$	0.17	0.091	0.060	0.057
$h^{ref} = n^{-1/5}$	0.40	0.29	0.25	0.22

TABLE 4

Expectation and standard deviation of the bandwidth selected by cross-validation as a function of the sample size n , and comparison with bandwidth h^{ref} chosen by the rule-of-thumb.

- [15] J.-J. Hsieh, W.-C. Huang, Nonparametric estimation and test of conditional Kendall's tau under semi-competing risks data and truncated data, *Journal of Applied Statistics* 42 (7) (2015) 1602–1616.
- [16] L. L. Chaieb, L.-P. Rivest, B. Abdous, Estimating survival under a dependent truncation, *Biometrika* 93 (3) (2006) 655–669.
- [17] Y.-J. Kim, Estimation of conditional Kendall's tau for bivariate interval censored data, *Communications for Statistical Applications and Methods* 22 (6) (2015) 599–604.
- [18] D. Oakes, Bivariate survival models induced by frailties, *Journal of the American Statistical Association* 84 (406) (1989) 487–493.
- [19] A. K. Manatunga, D. Oakes, A measure of association for bivariate frailty distributions, *Journal of Multivariate Analysis* 56 (1) (1996) 60–74.
- [20] A. V. Asimit, R. Gerrard, Y. Hou, L. Peng, Tail dependence measure for examining financial extreme co-movements, *Journal of Econometrics* 194 (2) (2016) 330–348.
- [21] A. Liu, Y. Hou, L. Peng, Interval estimation for a measure of tail dependence, *Insurance: Mathematics and Economics* 64 (2015) 294–305.
- [22] A. Sabeti, M. Wei, R. V. Craiu, Additive models for conditional copulas, *Stat* 3 (1) (2014) 300–312.
- [23] J. Dony, D. Mason, Uniform in bandwidth consistency of conditional u-statistics, *Bernoulli* 14 (4) (2008) 1108–1133.
- [24] P. Major, An estimate on the supremum of a nice class of stochastic integrals and u-statistics, *Probability Theory and Related Fields* 134 (3) (2006) 489–537.
- [25] E. Giné, A. Guillou, Rates of strong uniform consistency for multivariate kernel density estimators, *Annales de l'Institut Henri Poincaré (B) Probability and Statistics* 38 (6) (2002) 907–921.
- [26] U. Einmahl, D. Mason, Uniform in bandwidth consistency of kernel-type function estimators, *Ann. Statist.* 33 (3) (2005) 1380–1403.
- [27] E. F. Acar, C. Genest, J. Nešlehová, Beyond simplified pair-copula constructions, *Journal of Multivariate Analysis* 110 (2012) 74–90.
- [28] I. Hobæk Haff, K. Aas, A. Frigessi, On the simplified pair-copula construction—simply useful or too simplistic?, *J. Multivariate Anal.* 101 (2010) 1296–1310.
- [29] A. Derumigny, J.-D. Fermanian, About tests of the “simplifying” assumption for conditional copulas, *Depend. Model.* 5 (1) (2017) 154–197.
- [30] R. Serfling, *Approximation theorems of mathematical statistics*, John Wiley & Sons, 1980.
- [31] A. Rinaldo, L. Wasserman, et al., Generalized density clustering, *The Annals of Statistics* 38 (5) (2010) 2678–2722.
- [32] D. Bosq, J.-P. Lecoutre, *Théorie de l'estimation fonctionnelle*, Economica, 1987.
- [33] W. Stute, Conditional U-statistics, *Ann. Probab.* 19 (2) (1991) 812–825.

Appendix A: Proofs

For convenience, we recall Berk's (1970) inequality (see Theorem A in Serfling [30, p.201]). Note that, if $m = 1$, this reduces to Bernstein's inequality.

Lemma 11. *Let $m, n > 0$, $\mathbf{X}_1, \dots, \mathbf{X}_n$ i.i.d. random vectors with values in a measurable space \mathcal{X} and $g : \mathcal{X}^m \rightarrow [a, b]$ be a symmetric real bounded function. Set $\theta := \mathbb{E}[g(\mathbf{X}_1, \dots, \mathbf{X}_m)]$ and $\sigma^2 := \text{Var}[g(\mathbf{X}_1, \dots, \mathbf{X}_m)]$.*

Then, for any $t > 0$ and $n \geq m$,

$$\mathbb{P} \left(\binom{n}{m}^{-1} \sum_c g(\mathbf{X}_{i_1}, \dots, \mathbf{X}_{i_m}) - \theta \geq t \right) \leq \exp \left(- \frac{[n/m]t^2}{2\sigma^2 + (2/3)(b - \theta)t} \right),$$

where \sum_c denotes summation over all subgroups of m distinct integers (i_1, \dots, i_m) of $\{1, \dots, n\}$.

A.1. Proof of Proposition 1

Since there are no ties a.s.,

$$\begin{aligned} 1 + \hat{\tau}_{1,2|\mathbf{Z}=\mathbf{z}}^{(1)} &= 4 \sum_{i=1}^n \sum_{j=1}^n w_{i,n}(\mathbf{z}) w_{j,n}(\mathbf{z}) \left(\mathbb{1}\{X_{i,1} < X_{j,1}\} - \mathbb{1}\{X_{i,1} < X_{j,1}, X_{i,2} > X_{j,2}\} \right) \\ &= 4 \sum_{i=1}^n \sum_{j=1}^n w_{i,n}(\mathbf{z}) w_{j,n}(\mathbf{z}) \mathbb{1}\{X_{i,1} < X_{j,1}\} + \hat{\tau}_{1,2|\mathbf{Z}=\mathbf{z}}^{(3)} - 1. \end{aligned}$$

But

$$\begin{aligned} 1 &= \sum_{i=1}^n \sum_{j=1}^n w_{i,n}(\mathbf{z}) w_{j,n}(\mathbf{z}) = \sum_{i=1}^n \sum_{j=1}^n w_{i,n}(\mathbf{z}) w_{j,n}(\mathbf{z}) \left(\mathbb{1}\{X_{i,1} \leq X_{j,1}\} + \mathbb{1}\{X_{i,1} > X_{j,1}\} \right) \\ &= 2 \sum_{i=1}^n \sum_{j=1}^n w_{i,n}(\mathbf{z}) w_{j,n}(\mathbf{z}) \mathbb{1}\{X_{i,1} < X_{j,1}\} + \sum_{i=1}^n w_{i,n}^2, \end{aligned}$$

implying $1 + \hat{\tau}_{1,2|\mathbf{Z}=\mathbf{z}}^{(1)} = 2(1 - s_n) + \hat{\tau}_{1,2|\mathbf{Z}=\mathbf{z}}^{(3)} - 1$, and then $\hat{\tau}_{1,2|\mathbf{Z}=\mathbf{z}}^{(1)} = \hat{\tau}_{1,2|\mathbf{Z}=\mathbf{z}}^{(3)} - 2s_n$. Moreover,

$$\begin{aligned} \hat{\tau}_{1,2|\mathbf{Z}=\mathbf{z}}^{(2)} &= \sum_{i=1}^n \sum_{j=1}^n w_{i,n}(\mathbf{z}) w_{j,n}(\mathbf{z}) \left(\mathbb{1}\{X_{i,1} > X_{j,1}, X_{i,2} > X_{j,2}\} + \mathbb{1}\{X_{i,1} < X_{j,1}, X_{i,2} < X_{j,2}\} \right. \\ &\quad \left. - \mathbb{1}\{X_{i,1} > X_{j,1}, X_{i,2} < X_{j,2}\} - \mathbb{1}\{X_{i,1} < X_{j,1}, X_{i,2} > X_{j,2}\} \right) \\ &= 2 \sum_{i=1}^n \sum_{j=1}^n w_{i,n}(\mathbf{z}) w_{j,n}(\mathbf{z}) \left(\mathbb{1}\{X_{i,1} > X_{j,1}, X_{i,2} > X_{j,2}\} - \mathbb{1}\{X_{i,1} > X_{j,1}, X_{i,2} < X_{j,2}\} \right) \\ &= \frac{1}{2} (\hat{\tau}_{1,2|\mathbf{Z}=\mathbf{z}}^{(1)} + 1) + \frac{1}{2} (\hat{\tau}_{1,2|\mathbf{Z}=\mathbf{z}}^{(3)} - 1) = \frac{\hat{\tau}_{1,2|\mathbf{Z}=\mathbf{z}}^{(1)} + \hat{\tau}_{1,2|\mathbf{Z}=\mathbf{z}}^{(3)}}{2} = \hat{\tau}_{1,2|\mathbf{Z}=\mathbf{z}}^{(1)} + s_n = \hat{\tau}_{1,2|\mathbf{Z}=\mathbf{z}}^{(3)} - s_n. \quad \square \end{aligned}$$

A.2. Proof of Proposition 2

Lemma 12. Under Assumptions 3.1, 3.2 and 3.3, we have for any $t > 0$,

$$\mathbb{P} \left(|\hat{f}_{\mathbf{Z}}(\mathbf{z}) - f_{\mathbf{Z}}(\mathbf{z})| \geq \frac{C_{K,\alpha} h^\alpha}{\alpha!} + t \right) \leq 2 \exp \left(- \frac{nh^p t^2}{2f_{\mathbf{Z},\max} \int K^2 + (2/3)C_K t} \right).$$

This Lemma is proved below. If, for some $\epsilon > 0$, we have $C_{K,\alpha} h^\alpha / \alpha! + t \leq f_{\mathbf{Z},\min} - \epsilon$, then $\hat{f}_{\mathbf{Z}}(\mathbf{z}) \geq \epsilon > 0$ with a probability larger than $1 - 2 \exp \left(- nh^p t^2 / (2f_{\mathbf{Z},\max} \int K^2 + (2/3)C_K t) \right)$. So, we should choose the largest t as possible, which yields Proposition 2.

It remains to prove Lemma 12. Use the usual decomposition between a stochastic component and a bias: $\hat{f}_{\mathbf{Z}}(\mathbf{z}) - f_{\mathbf{Z}}(\mathbf{z}) = (\hat{f}_{\mathbf{Z}}(\mathbf{z}) - \mathbb{E}[\hat{f}_{\mathbf{Z}}(\mathbf{z})]) + (\mathbb{E}[\hat{f}_{\mathbf{Z}}(\mathbf{z})] - f_{\mathbf{Z}}(\mathbf{z}))$. We first bound the bias from above.

$$\mathbb{E}[\hat{f}_{\mathbf{Z}}(\mathbf{z})] - f_{\mathbf{Z}}(\mathbf{z}) = \int_{\mathbb{R}^p} K(\mathbf{u}) \left(f_{\mathbf{Z}}(\mathbf{z} + h\mathbf{u}) - f_{\mathbf{Z}}(\mathbf{z}) \right) d\mathbf{u}.$$

Set $\phi_{\mathbf{z},\mathbf{u}}(t) := f_{\mathbf{Z}}(\mathbf{z} + t h\mathbf{u})$ for $t \in [0, 1]$. This function has at least the same regularity as $f_{\mathbf{Z}}$, so it is α -differentiable. By a Taylor-Lagrange expansion, we get

$$\int_{\mathbb{R}^p} K(\mathbf{u}) \left(f_{\mathbf{Z}}(\mathbf{z} + h\mathbf{u}) - f_{\mathbf{Z}}(\mathbf{z}) \right) d\mathbf{u} = \int_{\mathbb{R}^p} K(\mathbf{u}) \left(\sum_{i=1}^{\alpha-1} \frac{1}{i!} \phi_{\mathbf{z},\mathbf{u}}^{(i)}(0) + \frac{1}{\alpha!} \phi_{\mathbf{z},\mathbf{u}}^{(\alpha)}(t_{\mathbf{z},\mathbf{u}}) \right) d\mathbf{u},$$

for some real number $t_{\mathbf{z},\mathbf{u}} \in (0, 1)$. By Assumption 3.1 and for every $i < \alpha$, $\int_{\mathbb{R}^p} K(\mathbf{u}) \phi_{\mathbf{z},\mathbf{u}}^{(i)}(0) d\mathbf{u} = 0$. Therefore,

$$\begin{aligned} \left| \mathbb{E}[\hat{f}_{\mathbf{Z}}(\mathbf{z})] - f_{\mathbf{Z}}(\mathbf{z}) \right| &= \left| \int_{\mathbb{R}^p} K(\mathbf{u}) \frac{1}{\alpha!} \phi_{\mathbf{z},\mathbf{u}}^{(\alpha)}(t_{\mathbf{z},\mathbf{u}}) d\mathbf{u} \right| \\ &= \frac{1}{\alpha!} \left| \int_{\mathbb{R}^p} K(\mathbf{u}) \sum_{i_1, \dots, i_\alpha=1}^p h^\alpha u_{i_1} \dots u_{i_\alpha} \frac{\partial^\alpha f_{\mathbf{Z}}}{\partial z_{i_1} \dots \partial z_{i_\alpha}}(\mathbf{z} + t_{\mathbf{z},\mathbf{u}} h\mathbf{u}) d\mathbf{u} \right| \leq \frac{C_{K,\alpha}}{\alpha!} h^\alpha. \end{aligned}$$

Second, the stochastic component may be written as

$$\hat{f}_{\mathbf{Z}}(\mathbf{z}) - \mathbb{E}[\hat{f}_{\mathbf{Z}}(\mathbf{z})] = n^{-1} \sum_{i=1}^n K_h(\mathbf{Z}_i - \mathbf{z}) - \mathbb{E} \left[n^{-1} \sum_{i=1}^n K_h(\mathbf{Z}_i - \mathbf{z}) \right] = n^{-1} \sum_{i=1}^n (g_{\mathbf{Z}}(\mathbf{Z}_i) - \mathbb{E}[g_{\mathbf{Z}}(\mathbf{Z}_i)]),$$

where $g(\mathbf{Z}_i) := K_h(\mathbf{Z}_i - \mathbf{z})$. Apply Lemma 11 with $m = 1$ and the latter $g(\mathbf{Z}_i)$. Here, we have $b = -a = h^{-p} C_K$, $\theta = \mathbb{E}[g(\mathbf{Z}_1)] \geq 0$ and $|\text{Var}[g(\mathbf{Z}_1)]| \leq h^{-p} f_{\mathbf{Z},\max} \int K^2$, and we get

$$\mathbb{P} \left(\left| \frac{1}{n} \sum_{i=1}^n K_h(\mathbf{Z}_i - \mathbf{z}) - \mathbb{E}[K_h(\mathbf{Z}_i - \mathbf{z})] \right| \geq t \right) \leq 2 \exp \left(- \frac{nt^2}{2h^{-p} f_{\mathbf{Z},\max} \int K^2 + (2/3)h^{-p} C_K t} \right). \square$$

A.3. Proof of Proposition 3

We show the result for $k = 1$. The two other cases can be proven in the same way.

Consider the decomposition

$$\begin{aligned} \hat{\tau}_{1,2|\mathbf{Z}=\mathbf{z}} - \tau_{1,2|\mathbf{Z}=\mathbf{z}} &= 4 \sum_{1 \leq i,j \leq n} w_{i,n}(\mathbf{z}) w_{j,n}(\mathbf{z}) \mathbb{1}\{\mathbf{X}_i < \mathbf{X}_j\} - 4\mathbb{P}(\mathbf{X}_1 < \mathbf{X}_2 | \mathbf{Z}_1 = \mathbf{Z}_2 = \mathbf{z}) \\ &= \frac{4}{n^2 \hat{f}_{\mathbf{Z}}^2(\mathbf{z})} \sum_{1 \leq i,j \leq n} K_h(\mathbf{Z}_i - \mathbf{z}) K_h(\mathbf{Z}_j - \mathbf{z}) \left(\mathbb{1}\{\mathbf{X}_i < \mathbf{X}_j\} - \mathbb{P}(\mathbf{X}_1 < \mathbf{X}_2 | \mathbf{Z}_1 = \mathbf{Z}_2 = \mathbf{z}) \right) \\ &=: \frac{4}{\hat{f}_{\mathbf{Z}}^2(\mathbf{z})} \sum_{1 \leq i,j \leq n} S_{i,j}(\mathbf{z}). \end{aligned}$$

Therefore, for any positive numbers x and $\lambda(\mathbf{z})$, we have

$$\begin{aligned} \mathbb{P}(|\hat{\tau}_{1,2|\mathbf{Z}=\mathbf{z}} - \tau_{1,2|\mathbf{Z}=\mathbf{z}}| > x) &\leq \mathbb{P} \left(\frac{1}{\hat{f}_{\mathbf{Z}}^2(\mathbf{z})} > \frac{1 + \lambda(\mathbf{z})}{f_{\mathbf{Z}}^2(\mathbf{z})} \right) + \mathbb{P} \left(\frac{4(1 + \lambda(\mathbf{z}))}{f_{\mathbf{Z}}^2(\mathbf{z})} \times \left| \sum_{1 \leq i,j \leq n} S_{i,j}(\mathbf{z}) \right| > x \right) \\ &\leq \mathbb{P} \left(\left| \frac{1}{\hat{f}_{\mathbf{Z}}^2(\mathbf{z})} - \frac{1}{f_{\mathbf{Z}}^2(\mathbf{z})} \right| > \frac{\lambda(\mathbf{z})}{f_{\mathbf{Z}}^2(\mathbf{z})} \right) + \mathbb{P} \left(\frac{4(1 + \lambda(\mathbf{z}))}{f_{\mathbf{Z}}^2(\mathbf{z})} \times \left| \sum_{1 \leq i,j \leq n} S_{i,j}(\mathbf{z}) \right| > x \right). \end{aligned}$$

For any t s.t. $C_{K,\alpha}h^\alpha/\alpha! + t < f_{\mathbf{Z},\min}/2$, set $\lambda(\mathbf{z}) = 16f_{\mathbf{Z}}^2(\mathbf{z})(C_{K,\alpha}h^\alpha/\alpha! + t)/f_{\mathbf{Z},\min}^3$. This yields

$$\begin{aligned} \mathbb{P}\left(\left|\hat{\tau}_{1,2|\mathbf{Z}=\mathbf{z}} - \tau_{1,2|\mathbf{Z}=\mathbf{z}}\right| > x\right) &\leq \mathbb{P}\left(\left|\frac{1}{\hat{f}_{\mathbf{Z}}^2(\mathbf{z})} - \frac{1}{f_{\mathbf{Z}}^2(\mathbf{z})}\right| > \frac{16}{f_{\mathbf{Z},\min}^3}\left(\frac{C_{K,\alpha}h^\alpha}{\alpha!} + t\right)\right) \\ &+ \mathbb{P}\left(\left|\sum_{1 \leq i,j \leq n} S_{i,j}(\mathbf{z})\right| > \frac{f_{\mathbf{Z}}^2(\mathbf{z})x}{4(1 + \lambda(\mathbf{z}))}\right). \end{aligned}$$

By setting

$$x = \frac{4}{f_{\mathbf{Z}}^2(\mathbf{z})}\left(\frac{C_{\mathbf{XZ},\alpha}h^\alpha}{\alpha!} + \frac{3f_{\mathbf{Z}}(\mathbf{z})\int K^2}{2nh^p} + t'\right)\left(1 + \frac{16f_{\mathbf{Z}}^2(\mathbf{z})}{f_{\mathbf{Z},\min}^3}\left(\frac{C_{K,\alpha}h^\alpha}{\alpha!} + t\right)\right),$$

and applying the next two lemmas 13 and 14, we get the result. \square

Lemma 13. *Under Assumptions 3.1-3.3 and if $C_{K,\alpha}h^\alpha/\alpha! + t < f_{\mathbf{Z},\min}/2$ for some $t > 0$,*

$$\mathbb{P}\left(\left|\frac{1}{\hat{f}_{\mathbf{Z}}^2(\mathbf{z})} - \frac{1}{f_{\mathbf{Z}}^2(\mathbf{z})}\right| > \frac{16}{f_{\mathbf{Z},\min}^3}\left(\frac{C_{K,\alpha}h^\alpha}{\alpha!} + t\right)\right) \leq 2\exp\left(-\frac{nh^pt^2}{2f_{\mathbf{Z},\max}\int K^2 + (2/3)C_Kt}\right),$$

and $\hat{f}_{\mathbf{Z}}(\mathbf{z})$ is strictly positive on these events.

Proof : Applying the mean value inequality to the function $x \mapsto 1/x^2$, we get the inequality $\left|1/\hat{f}_{\mathbf{Z}}^2(\mathbf{z}) - 1/f_{\mathbf{Z}}^2(\mathbf{z})\right| \leq 2|\hat{f}_{\mathbf{Z}}(\mathbf{z}) - f_{\mathbf{Z}}(\mathbf{z})|/f_{\mathbf{Z}}^{*3}$, where $f_{\mathbf{Z}}^*$ lies between $\hat{f}_{\mathbf{Z}}(\mathbf{z})$ and $f_{\mathbf{Z}}(\mathbf{z})$. Denote by \mathcal{E} the event $\mathcal{E} := \{|\hat{f}_{\mathbf{Z}}(\mathbf{z}) - f_{\mathbf{Z}}(\mathbf{z})| \leq C_{K,\alpha}h^\alpha/\alpha! + t\}$. By Lemma 12, we obtain

$$\mathbb{P}(\mathcal{E}) \geq 1 - 2\exp\left(-\frac{nh^pt^2}{2f_{\mathbf{Z},\max}\int K^2 + (2/3)C_Kt}\right). \quad (5)$$

Therefore, on this event \mathcal{E} , $|\hat{f}_{\mathbf{Z}}(\mathbf{z}) - f_{\mathbf{Z}}(\mathbf{z})| \leq f_{\mathbf{Z},\min}/2$, so that $f_{\mathbf{Z},\min}/2 \leq \hat{f}_{\mathbf{Z}}(\mathbf{z})$. We have also $f_{\mathbf{Z},\min}/2 \leq f_{\mathbf{Z}}(\mathbf{z})$ and then $f_{\mathbf{Z},\min}/2 \leq f_{\mathbf{Z}}^*$. Combining the previous inequalities, we finally get

$$\left|\frac{1}{\hat{f}_{\mathbf{Z}}^2(\mathbf{z})} - \frac{1}{f_{\mathbf{Z}}^2(\mathbf{z})}\right| \leq \frac{16}{f_{\mathbf{Z},\min}^3}|\hat{f}_{\mathbf{Z}}(\mathbf{z}) - f_{\mathbf{Z}}(\mathbf{z})| \leq \frac{16}{f_{\mathbf{Z},\min}^3}\left(\frac{C_{K,\alpha}h^\alpha}{\alpha!} + t\right),$$

on \mathcal{E} . But since

$$\mathbb{P}\left(\left|\frac{1}{\hat{f}_{\mathbf{Z}}^2(\mathbf{z})} - \frac{1}{f_{\mathbf{Z}}^2(\mathbf{z})}\right| > \frac{16}{f_{\mathbf{Z},\min}^3}\left(\frac{C_{K,\alpha}h^\alpha}{\alpha!} + t\right)\right) \leq \mathbb{P}(\mathcal{E}^c),$$

we deduce the result. \square

Lemma 14. *Under Assumptions 3.1-3.4, if $C_{\tilde{K},2}h^2 < f_{\mathbf{z}}(\mathbf{z})$, we have for any $t > 0$*

$$\begin{aligned} \mathbb{P}\left(\left|\sum_{1 \leq i,j \leq n} S_{i,j}(\mathbf{z})\right| > \frac{C_{\mathbf{XZ},\alpha}h^\alpha}{\alpha!} + \frac{3f_{\mathbf{z}}(\mathbf{z})\int K^2}{2nh^p} + t\right) &\leq 2\exp\left(-\frac{(n-1)h^{2pt^2}}{4f_{\mathbf{Z},\max}^2(\int K^2)^2 + (8/3)C_{\tilde{K}}^2t}\right) \\ &+ 2\exp\left(-\frac{nh^p(f_{\mathbf{z}}(\mathbf{z}) - C_{\tilde{K},2}h^2)^2}{8f_{\mathbf{Z},\max}\int \tilde{K}^2 + 4C_{\tilde{K}}(f_{\mathbf{z}}(\mathbf{z}) - C_{\tilde{K},2}h^2)/3}\right). \end{aligned}$$

Proof : Note that $\sum_{1 \leq i,j \leq n} S_{i,j}(\mathbf{z}) = \sum_{1 \leq i \neq j \leq n} (S_{i,j}(\mathbf{z}) - \mathbb{E}[S_{i,j}(\mathbf{z})]) + n(n-1)\mathbb{E}[S_{1,2}(\mathbf{z})] + \sum_{i=1}^n S_{i,i}(\mathbf{z})$. The ‘‘diagonal term’’ $\sum_{i=1}^n S_{i,i}(\mathbf{z}) = -\mathbb{P}(\mathbf{X}_1 < \mathbf{X}_2 | \mathbf{Z}_1 = \mathbf{Z}_2 = \mathbf{z}) \sum_{i=1}^n K_h^2(\mathbf{Z}_i - \mathbf{z})/n^2$ is negative and negligible. It will be denoted by $-\Delta_n(\mathbf{z}) < 0$. Note that $\tilde{K}(\cdot) := K^2(\cdot)/\int K^2$ is a two-order kernel. Then,

$\tilde{f}_{\mathbf{z}}(\mathbf{z}) := \sum_{i=1}^n \tilde{K}_h(\mathbf{Z}_i - \mathbf{z})/n$ is a consistent estimator of $f_{\mathbf{z}}(\mathbf{z})$. Therefore, due to Lemma 12 and with obvious notations, we have for every $\varepsilon > 0$

$$\mathbb{P}\left(\left|\tilde{f}_{\mathbf{z}}(\mathbf{z}) - f_{\mathbf{z}}(\mathbf{z})\right| \geq \frac{C_{\tilde{K},2}h^2}{2} + \varepsilon\right) \leq 2 \exp\left(-\frac{nh^p\varepsilon^2}{2f_{\mathbf{z},max} \int \tilde{K}^2 + (2/3)C_{\tilde{K}}\varepsilon}\right).$$

This implies

$$\begin{aligned} \mathbb{P}\left(\left|\frac{\int K^2}{n^2h^p} \sum_{i=1}^n \tilde{K}_h(\mathbf{Z}_i - \mathbf{z}) - \frac{f_{\mathbf{z}}(\mathbf{z}) \int K^2}{nh^p}\right| \geq \left(\frac{\int K^2}{nh^p}\right)\left(\frac{C_{\tilde{K},2}h^2}{2} + \varepsilon\right)\right) \\ \leq 2 \exp\left(-\frac{nh^p\varepsilon^2}{2f_{\mathbf{z},max} \int \tilde{K}^2 + (2/3)C_{\tilde{K}}\varepsilon}\right). \end{aligned}$$

By choosing ε s.t. $C_{\tilde{K},2}h^2/2 + \varepsilon = f_{\mathbf{z}}(\mathbf{z})/2$, Δ_n will be smaller than $3f_{\mathbf{z}}(\mathbf{z}) \int K^2/(2nh^p)$ with a probability that is larger than

$$1 - 2 \exp\left(-\frac{nh^p\varepsilon^2}{2f_{\mathbf{z},max} \int \tilde{K}^2 + (2/3)C_{\tilde{K}}\varepsilon}\right). \quad (6)$$

Now, let us deal with the main term, that is decomposed as a stochastic component and a bias component. First, let us deal with the bias. Simple calculations provide, if $i \neq j$,

$$\begin{aligned} \mathbb{E}[S_{i,j}(\mathbf{z})] &= n^{-2} \mathbb{E}\left[K_h(\mathbf{Z}_i - \mathbf{z})K_h(\mathbf{Z}_j - \mathbf{z})\left(\mathbb{1}\{\mathbf{X}_i < \mathbf{X}_j\} - \mathbb{P}(\mathbf{X}_i < \mathbf{X}_j | \mathbf{Z}_i = \mathbf{Z}_j = \mathbf{z})\right)\right] \\ &= n^{-2} \int_{\mathbb{R}^{2p+2}} K_h(\mathbf{z}_1 - \mathbf{z})K_h(\mathbf{z}_2 - \mathbf{z})\left(\mathbb{1}\{\mathbf{x}_1 < \mathbf{x}_2\} - \mathbb{P}(\mathbf{X}_i < \mathbf{X}_j | \mathbf{Z}_i = \mathbf{Z}_j = \mathbf{z})\right) \\ &\quad \times f_{\mathbf{X},\mathbf{Z}}(\mathbf{x}_1, \mathbf{z}_1) f_{\mathbf{X},\mathbf{Z}}(\mathbf{x}_2, \mathbf{z}_2) d\mathbf{x}_1 d\mathbf{z}_1 d\mathbf{x}_2 d\mathbf{z}_2 \\ &= n^{-2} \int_{\mathbb{R}^{2p+2}} K(\mathbf{u})K(\mathbf{v})\left(\mathbb{1}\{\mathbf{x}_1 < \mathbf{x}_2\} - \mathbb{P}(\mathbf{X}_i < \mathbf{X}_j | \mathbf{Z}_i = \mathbf{Z}_j = \mathbf{z})\right) \\ &\quad \times \left(f_{\mathbf{X},\mathbf{Z}}(\mathbf{x}_1, \mathbf{z} + h\mathbf{u}) f_{\mathbf{X},\mathbf{Z}}(\mathbf{x}_2, \mathbf{z} + h\mathbf{v}) - f_{\mathbf{X},\mathbf{Z}}(\mathbf{x}_1, \mathbf{z}) f_{\mathbf{X},\mathbf{Z}}(\mathbf{x}_2, \mathbf{z})\right) d\mathbf{x}_1 d\mathbf{u} d\mathbf{x}_2 d\mathbf{v}, \end{aligned}$$

because, for every \mathbf{z} ,

$$0 = \int_{\mathbb{R}^4} \left(\mathbb{1}\{\mathbf{x}_1 < \mathbf{x}_2\} - \mathbb{P}(\mathbf{X}_1 < \mathbf{X}_2 | \mathbf{Z}_1 = \mathbf{Z}_2 = \mathbf{z})\right) f_{\mathbf{X},\mathbf{Z}}(\mathbf{x}_1, \mathbf{z}) f_{\mathbf{X},\mathbf{Z}}(\mathbf{x}_2, \mathbf{z}) d\mathbf{x}_1 d\mathbf{x}_2.$$

Apply the Taylor-Lagrange formula to the function $\phi_{\mathbf{x}_1, \mathbf{x}_2, \mathbf{u}, \mathbf{v}}(t) := f_{\mathbf{X},\mathbf{Z}}(\mathbf{x}_1, \mathbf{z} + t\mathbf{u}) f_{\mathbf{X},\mathbf{Z}}(\mathbf{x}_2, \mathbf{z} + t\mathbf{v})$. With obvious notation, this yields

$$\begin{aligned} \mathbb{E}[S_{i,j}(\mathbf{z})] &= n^{-2} \int K(\mathbf{u})K(\mathbf{v})\left(\mathbb{1}\{\mathbf{x}_1 < \mathbf{x}_2\} - \mathbb{P}(\mathbf{X}_i < \mathbf{X}_j | \mathbf{Z}_i = \mathbf{Z}_j = \mathbf{z})\right) \\ &\quad \times \left(\sum_{k=1}^{\alpha-1} \frac{1}{k!} \phi_{\mathbf{x}_1, \mathbf{x}_2, \mathbf{u}, \mathbf{v}}^{(k)}(0) + \frac{1}{\alpha!} \phi_{\mathbf{x}_1, \mathbf{x}_2, \mathbf{u}, \mathbf{v}}^{(\alpha)}(t_{\mathbf{x}_1, \mathbf{x}_2, \mathbf{u}, \mathbf{v}})\right) d\mathbf{x}_1 d\mathbf{u} d\mathbf{x}_2 d\mathbf{v} \\ &= \int \frac{K(\mathbf{u})K(\mathbf{v})}{n^2\alpha!} \left(\mathbb{1}\{\mathbf{x}_1 < \mathbf{x}_2\} - \mathbb{P}(\mathbf{X}_i < \mathbf{X}_j | \mathbf{Z}_i = \mathbf{Z}_j = \mathbf{z})\right) \phi_{\mathbf{x}_1, \mathbf{x}_2, \mathbf{u}, \mathbf{v}}^{(\alpha)}(t_{\mathbf{x}_1, \mathbf{x}_2, \mathbf{u}, \mathbf{v}}) d\mathbf{x}_1 d\mathbf{u} d\mathbf{x}_2 d\mathbf{v}. \end{aligned}$$

Since $\phi_{\mathbf{x}_1, \mathbf{x}_2, \mathbf{u}, \mathbf{v}}^{(\alpha)}(t)$ is equal to

$$\sum_{k=0}^{\alpha} \binom{\alpha}{k} \sum_{i_1, \dots, i_{\alpha-k}=1}^p h^{\alpha} u_{i_1} \dots u_{i_k} v_{i_{k+1}} \dots v_{i_{\alpha}} \frac{\partial^k f_{\mathbf{X},\mathbf{Z}}}{\partial z_{i_1} \dots \partial z_{i_k}}(\mathbf{x}_1, \mathbf{z} + t\mathbf{u}) \frac{\partial^{\alpha-k} f_{\mathbf{X},\mathbf{Z}}}{\partial z_{i_{k+1}} \dots \partial z_{i_{\alpha}}}(\mathbf{x}_2, \mathbf{z} + t\mathbf{v}),$$

using Assumption 3.4, we get

$$|\mathbb{E}[S_{1,2}(\mathbf{z})]| \leq C_{\mathbf{XZ},\alpha} h^\alpha / (n^2 \alpha!). \quad (7)$$

Second, the stochastic component will be bounded from above. Indeed,

$$\sum_{1 \leq i \neq j \leq n} (S_{i,j}(\mathbf{z}) - \mathbb{E}[S_{i,j}(\mathbf{z})]) = \frac{1}{n^2} \sum_{1 \leq i \neq j \leq n} g_{\mathbf{z}}((\mathbf{X}_i, \mathbf{Z}_i), (\mathbf{X}_j, \mathbf{Z}_j)),$$

with the function $g_{\mathbf{z}}$ defined by

$$\begin{aligned} g_{\mathbf{z}}((\mathbf{X}_i, \mathbf{Z}_i), (\mathbf{X}_j, \mathbf{Z}_j)) &:= K_h(\mathbf{Z}_i - \mathbf{z})K_h(\mathbf{Z}_j - \mathbf{z}) \left(\mathbb{1}\{\mathbf{X}_i < \mathbf{X}_j\} - \mathbb{P}(\mathbf{X}_i < \mathbf{X}_j | \mathbf{Z}_i = \mathbf{Z}_j = \mathbf{z}) \right) \\ &- \mathbb{E} \left[K_h(\mathbf{Z}_i - \mathbf{z})K_h(\mathbf{Z}_j - \mathbf{z}) \left(\mathbb{1}\{\mathbf{X}_i < \mathbf{X}_j\} - \mathbb{P}(\mathbf{X}_i < \mathbf{X}_j | \mathbf{Z}_i = \mathbf{Z}_j = \mathbf{z}) \right) \right]. \end{aligned}$$

The symmetrized version of g is $\tilde{g}_{i,j} = \left(g_{\mathbf{z}}((\mathbf{X}_i, \mathbf{Z}_i), (\mathbf{X}_j, \mathbf{Z}_j)) + g_{\mathbf{z}}((\mathbf{X}_j, \mathbf{Z}_j), (\mathbf{X}_i, \mathbf{Z}_i)) \right) / 2$. We can now apply Lemma 11 to the sum of the $\tilde{g}_{i,j}$. With its notation, $\theta = \mathbb{E}[\tilde{g}_{i,j}] = 0$. Moreover,

$$\begin{aligned} & \left| \text{Var} \left[g_{\mathbf{z}}((\mathbf{X}_i, \mathbf{Z}_i), (\mathbf{X}_j, \mathbf{Z}_j)) \right] \right| \\ & \leq \int K_h^2(\mathbf{z}_1 - \mathbf{z})K_h^2(\mathbf{z}_2 - \mathbf{z}) \left(\mathbb{1}\{\mathbf{x}_1 < \mathbf{x}_2\} - \mathbb{P}(\mathbf{X}_i < \mathbf{X}_j | \mathbf{Z}_i = \mathbf{Z}_j = \mathbf{z}) \right)^2 \\ & \quad \times f_{\mathbf{X},\mathbf{Z}}(\mathbf{x}_1, \mathbf{z}_1) f_{\mathbf{X},\mathbf{Z}}(\mathbf{x}_2, \mathbf{z}_2) d\mathbf{x}_1 d\mathbf{x}_2 d\mathbf{z}_1 d\mathbf{z}_2 \\ & \leq \int \frac{K^2(\mathbf{t}_1)K^2(\mathbf{t}_2)}{h^{2p}} f_{\mathbf{X},\mathbf{Z}}(\mathbf{x}_1, \mathbf{z} - h\mathbf{t}_1) f_{\mathbf{X},\mathbf{Z}}(\mathbf{x}_2, \mathbf{z} - h\mathbf{t}_2) d\mathbf{x}_1 d\mathbf{x}_2 d\mathbf{t}_1 d\mathbf{t}_2 \\ & \leq h^{-2p} f_{\mathbf{Z},\max}^2 \left(\int K^2 \right)^2, \end{aligned}$$

and the same upper bound applies for $\tilde{g}_{i,j}$ (invoke Cauchy-Schwarz inequality). Here, we choose $b = -a = 2C_K^2 h^{-2p}$. This yields

$$\mathbb{P} \left(\frac{2}{n(n-1)} \sum_{1 \leq i < j \leq n} \tilde{g}_{i,j} > t \right) \leq \exp \left(- \frac{[n/2]t^2}{2h^{-2p} f_{\mathbf{Z},\max}^2 (\int K^2)^2 + (4/3)C_K^2 h^{-2pt}} \right) \quad (8)$$

Then, for every $t > 0$, we obtain

$$\begin{aligned} & \mathbb{P} \left(\left| \sum_{1 \leq i \neq j \leq n} (S_{i,j}(\mathbf{z}) - \mathbb{E}[S_{i,j}(\mathbf{z})]) \right| \geq t \right) \leq \mathbb{P} \left(\frac{1}{n^2} \left| \sum_{1 \leq i \neq j \leq n} g_{\mathbf{z}}((\mathbf{X}_i, \mathbf{Z}_i), (\mathbf{X}_j, \mathbf{Z}_j)) \right| \geq t \right) \\ & \leq \mathbb{P} \left(\frac{(n-1)}{n} \times \frac{2}{n(n-1)} \left| \sum_{1 \leq i < j \leq n} \tilde{g}_{i,j} \right| \geq t \right) \leq 2 \exp \left(- \frac{[n/2]t^2}{2h^{-2p} f_{\mathbf{Z},\max}^2 (\int K^2)^2 + (4/3)C_K^2 h^{-2pt}} \right). \end{aligned}$$

The latter inequality, (6) and (7) yield the result. \square

A.4. Proof of Proposition 4

With the notations of the proof of Proposition 3, we get the following lemma, that straightforwardly implies the result.

Lemma 15. *Under the Assumptions and conditions of Proposition 4, we have*

$$\begin{aligned} \mathbb{P}\left(\left|\sum_{1 \leq i, j \leq n} S_{i,j}(\mathbf{z})\right| > \frac{C_{\mathbf{XZ},\alpha} h^\alpha}{\alpha!} + \frac{3f_{\mathbf{z}}(\mathbf{z}) \int K^2}{2nh^p} + t\right) &\leq C_2 \exp\left(-\frac{\alpha_2 nh^p t}{8f_{\mathbf{z},\max}(\int K^2)}\right) \\ &+ 2 \exp\left(-\frac{nh^p(f_{\mathbf{z}}(\mathbf{z}) - C_{\tilde{K},2} h^2)^2}{8f_{\mathbf{z},\max} \int \tilde{K}^2 + 4C_{\tilde{K}}(f_{\mathbf{z}}(\mathbf{z}) - C_{\tilde{K},2} h^2)/3}\right) \\ &+ 2 \exp\left(\frac{nh^p t^2}{32 \int K^2 (\int |K|)^2 f_{\mathbf{z},\max}^3 + 8C_K \int |K| f_{\mathbf{z},\max} t/3}\right). \end{aligned}$$

Proof : We lead exactly the same reasoning and the same notations as in Lemma 14, until (8). Now, with the same notations, introduce $\bar{g}_i := \mathbb{E}[\tilde{g}_{i,j} | \mathbf{X}_i, \mathbf{Z}_i]$ and consider $\xi_{i,j} := \tilde{g}_{i,j} - \bar{g}_i - \bar{g}_j$. Then, $\xi_{i,j}$ is a degenerate (symmetrical) U-statistics because $\mathbb{E}[\xi_{i,j} | \mathbf{X}_i, \mathbf{Z}_i] = \mathbb{E}[\xi_{i,j} | \mathbf{X}_j, \mathbf{Z}_j] = 0$, when $i \neq j$. Actually $\xi_{i,j} =: \xi_{\mathbf{z}}(\mathbf{X}_i, \mathbf{Z}_i, \mathbf{X}_j, \mathbf{Z}_j)$ for some function $\xi_{\mathbf{z}}$ and set

$$\ell_{\mathbf{z}} : (\mathbf{x}_1, \mathbf{z}_1, \mathbf{x}_2, \mathbf{z}_2) \mapsto \frac{h^{2p}}{4C_K^2} \xi_{\mathbf{z}}((\mathbf{x}_1, \mathbf{z}_1), (\mathbf{x}_2, \mathbf{z}_2)), \quad (9)$$

for a fixed \mathbf{z} and a fixed h . This yields $\|\ell_{\mathbf{z}}\|_{\infty} \leq 1$. By usual changes of variables, we get

$$\begin{aligned} &\int \ell_{\mathbf{z}}^2(\mathbf{x}_1, \mathbf{z}_1, \mathbf{x}_2, \mathbf{z}_2) f_{\mathbf{X},\mathbf{Z}}(\mathbf{x}_1, \mathbf{z}_1) f_{\mathbf{X},\mathbf{Z}}(\mathbf{x}_2, \mathbf{z}_2) d\mathbf{x}_1 d\mathbf{x}_2 d\mathbf{z}_1 d\mathbf{z}_2 \\ &\leq 3h^{2p} \frac{(\int K^2 f_{\mathbf{z},\max})^2}{(4C_K^2)^2} + 6h^{3p} \frac{\int K^2 f_{\mathbf{z},\max} (\int |K| f_{\mathbf{z},\max})^2}{(4C_K^2)^2} \leq \sigma^2, \text{ with} \\ &\sigma := h^p C_{\sigma}, \quad C_{\sigma} := \int K^2 f_{\mathbf{z},\max} / (2C_K^2), \end{aligned} \quad (10)$$

because $6h^p \int K^2 f_{\mathbf{z},\max} (\int |K| f_{\mathbf{z},\max})^2 \leq (\int K^2 f_{\mathbf{z},\max})^2$. With the notations of [24], this implies $D = 1$, $m = 1$ and L is arbitrarily small. Therefore, Theorem 2 in [24] yields

$$\mathbb{P}\left(\frac{1}{2n} \left| \sum_{i \neq j} \ell_{\mathbf{z}}(\mathbf{X}_i, \mathbf{Z}_i, \mathbf{X}_j, \mathbf{Z}_j) \right| > x\right) \leq C_2 \exp\left(-\frac{\alpha_2 x}{\sigma}\right), \quad (11)$$

for some universal constants C_2 and α_2 when $x \leq n\sigma^3$. By setting $t/2 = 4C_K^2 x / (nh^{2p})$ and applying Lemma 11, this provides

$$\begin{aligned} \mathbb{P}\left(\left|\sum_{1 \leq i \neq j \leq n} (S_{i,j}(\mathbf{z}) - \mathbb{E}[S_{i,j}(\mathbf{z})])\right| \geq t\right) &\leq \mathbb{P}\left(\frac{1}{n^2} \left| \sum_{1 \leq i \neq j \leq n} \xi_{ij} \right| \geq t/2\right) + \mathbb{P}\left(\frac{1}{n} \sum_{i=1}^n \bar{g}_i \geq t/4\right) \\ &\leq C_2 \exp\left(-\frac{\alpha_2 nth^p}{8f_{\mathbf{z},\max}(\int K^2)}\right) + 2 \exp\left(\frac{nh^p t^2}{32 \int K^2 (\int |K|)^2 f_{\mathbf{z},\max}^3 + 8/3 C_K \int |K| f_{\mathbf{z},\max} t}\right), \end{aligned}$$

when $t \leq 2h^p (\int K^2)^3 f_{\mathbf{z},\max}^3 / C_K^4$. This inequality, (6) and (7) conclude the proof. \square

A.5. Proof of Proposition 6

For $k = 1$, we follow the paths of the proof of Proposition 4. Since $\hat{\tau}_{1,2|\mathbf{Z}=\mathbf{z}} - \tau_{1,2|\mathbf{Z}=\mathbf{z}} = 4 \sum_{1 \leq i, j \leq n} S_{i,j}(\mathbf{z}) / \hat{f}_{\mathbf{z}}^2(\mathbf{z})$, we prove the result if we bound from above $1/\hat{f}_{\mathbf{z}}^2(\mathbf{z})$ and $|\sum_{1 \leq i, j \leq n} S_{i,j}(\mathbf{z})|$ uniformly w.r.t. $\mathbf{z} \in \mathcal{Z}$. To be

specific, for any positive constant $\mu < 1$, if $|\hat{f}_{\mathbf{Z}}(\mathbf{z}) - f_{\mathbf{Z}}(\mathbf{z})| \leq \mu f_{\mathbf{Z},\min}$, then $1/\hat{f}_{\mathbf{Z},\max}^2(\mathbf{z}) \leq f_{\mathbf{Z},\min}^{-2}(1 - \mu)^{-2}$. We deduce

$$\begin{aligned} \mathbb{P}(\sup_{\mathbf{z} \in \mathcal{Z}} |\hat{\tau}_{1,2|\mathbf{Z}=\mathbf{z}} - \tau_{1,2|\mathbf{Z}=\mathbf{z}}| > x) &\leq \mathbb{P}(\|\hat{f}_{\mathbf{Z}} - f_{\mathbf{Z}}\|_{\infty} > \mu f_{\mathbf{Z},\min}) \\ &+ \mathbb{P}\left(\frac{4}{f_{\mathbf{Z},\min}^2(1 - \mu)^2} \sup_{\mathbf{z} \in \mathcal{Z}} \left| \sum_{1 \leq i, j \leq n} S_{i,j}(\mathbf{z}) \right| > x\right). \end{aligned}$$

First invoke the uniform exponential inequality, as stated in [31], Proposition 9: for every $\varepsilon < b_K \int K^2 f_{\mathbf{Z},\max}/C_K$,

$$\mathbb{P}(\|\hat{f}_{\mathbf{Z}} - f_{\mathbf{Z}}\|_{\infty} > \varepsilon + \frac{C_{\mathbf{XZ},\alpha} h^{\alpha}}{\alpha!}) \leq \mathbb{P}(\|\hat{f}_{\mathbf{Z}} - \mathbb{E}[\hat{f}_{\mathbf{Z}}]\|_{\infty} > \varepsilon) \leq L_K \exp(-C_{f,K} n h^p \varepsilon^2), \quad (12)$$

for n sufficiently large. Then, apply Lemma 16, by setting (x, ε) so that

$$x = \frac{4}{f_{\mathbf{Z},\min}^2(1 - \mu)^2} \left(\frac{C_{\mathbf{XZ},\alpha} h^{\alpha}}{\alpha!} + \frac{3f_{\mathbf{Z},\max} \int K^2}{2nh^p} + t \right) \text{ and } \varepsilon + \frac{C_{\mathbf{XZ},\alpha} h^{\alpha}}{\alpha!} = \mu f_{\mathbf{Z},\min}. \quad \square$$

Lemma 16. *Under the assumptions and conditions of Proposition 6, we have*

$$\begin{aligned} &\mathbb{P}\left(\sup_{\mathbf{z} \in \mathcal{Z}} \left| \sum_{1 \leq i, j \leq n} S_{i,j}(\mathbf{z}) \right| > \frac{C_{\mathbf{XZ},\alpha} h^{\alpha}}{\alpha!} + \frac{3f_{\mathbf{Z},\max} \int K^2}{2nh^p} + t\right) \\ &\leq C_2 D \exp\left(-\frac{\alpha_2 n t h^p}{8f_{\mathbf{Z},\max}(\int K^2)}\right) + L_{\tilde{K}} \exp(-C_{f,\tilde{K}} n h^p (f_{\mathbf{Z},\max} - \tilde{C}_{\mathbf{XZ},2} h^2)^2 / 4) \\ &+ 2 \exp\left(-\frac{A_2 n h^p t^2 C_K^{-4}}{16^2 A_1^2 \int K^2 f_{\mathbf{Z},\max}^3 (\int |K|)^2}\right) + 2 \exp\left(-\frac{A_2 n h^p t}{16 C_K^2 A_1}\right). \end{aligned}$$

Proof : We will use the arguments and notations of the proof of Lemmas 14 and 15. We still invoke the decomposition $\sum_{1 \leq i, j \leq n} S_{i,j}(\mathbf{z}) = \sum_{1 \leq i \neq j \leq n} (S_{i,j}(\mathbf{z}) - \mathbb{E}[S_{i,j}(\mathbf{z})]) + n(n-1)\mathbb{E}[S_{1,2}(\mathbf{z})] + \sum_{i=1}^n S_{i,i}(\mathbf{z})$. First let us find a uniform bound for the ‘‘diagonal term’’ $\Delta_n(\mathbf{z}) = \sum_{i=1}^n S_{i,i}(\mathbf{z}) = \int K^2 \tilde{f}_{\mathbf{Z}}(\mathbf{z}) / (n h^p)$. As in (12), for every $\varepsilon < b_{\tilde{K}} \int \tilde{K}^2 f_{\mathbf{Z},\max} / C_{\tilde{K}}$,

$$\mathbb{P}(\|\tilde{f}_{\mathbf{Z}} - f_{\mathbf{Z}}\|_{\infty} > \varepsilon + \frac{\tilde{C}_{\mathbf{XZ},2} h^2}{2}) \leq L_{\tilde{K}} \exp(-C_{f,\tilde{K}} n h^p \varepsilon^2),$$

for n sufficiently large. This implies

$$\mathbb{P}\left(\sup_{\mathbf{z} \in \mathcal{Z}} \left| \frac{\int K^2}{n^2 h^p} \sum_{i=1}^n \tilde{K}_h(\mathbf{Z}_i - \mathbf{z}) - \frac{f_{\mathbf{Z}}(\mathbf{z}) \int K^2}{n h^p} \right| \geq \left(\frac{\int K^2}{n h^p} \right) \left(\varepsilon + \frac{\tilde{C}_{\mathbf{XZ},2} h^2}{2} \right) \right) \leq L_{\tilde{K}} \exp(-C_{f,\tilde{K}} n h^p \varepsilon^2).$$

Choose ε s.t. $\tilde{C}_{\mathbf{XZ},2} h^2 / 2 + \varepsilon = f_{\mathbf{Z},\max} / 2$. Then, $\sup_{\mathbf{z}} |\Delta_n(\mathbf{z})|$ will be smaller than $3f_{\mathbf{Z},\max} \int K^2 / (2n h^p)$ with a probability that is larger than

$$1 - L_{\tilde{K}} \exp(-C_{f,\tilde{K}} n h^p \varepsilon^2). \quad (13)$$

Moreover, it is easy to see that

$$\sup_{\mathbf{z} \in \mathcal{Z}} |\mathbb{E}[S_{1,2}(\mathbf{z})]| \leq C_{\mathbf{XZ},\alpha} h^{\alpha} / (n^2 \alpha!). \quad (14)$$

With the notations of Lemma 15's proof, the stochastic component is driven by

$$\begin{aligned} \sum_{1 \leq i \neq j \leq n} (S_{i,j}(\mathbf{z}) - \mathbb{E}[S_{i,j}(\mathbf{z})]) &= \frac{1}{n^2} \sum_{1 \leq i \neq j \leq n} g_{\mathbf{z}}((\mathbf{X}_i, \mathbf{Z}_i), (\mathbf{X}_j, \mathbf{Z}_j)) \\ &= \frac{1}{n^2} \sum_{1 \leq i \neq j \leq n} \tilde{g}_{i,j} = \frac{1}{n^2} \sum_{1 \leq i \neq j \leq n} \xi_{i,j} + \frac{2(n-1)}{n^2} \sum_{i=1}^n \bar{g}_i. \end{aligned}$$

Now apply Theorem 1 in [24], by recalling (9) and considering the family $\mathcal{F} := \{\ell_{\mathbf{z}}, \mathbf{z} \in \mathcal{Z}\}$, for a fixed bandwidth h . The constant σ has the same value as in (10). It is easy to check that the latter class of functions is L^2 dense (see [24]). Set $\varepsilon \in (0, 1)$. Since K is λ_K -Lipschitz, every function $\ell_{\mathbf{z}} \in \mathcal{F}$ can be approximated in L^2 by a function $\ell_{\mathbf{z}_j} \in \mathcal{F}$, for some $j \in \{1, \dots, m\}$ s.t. $\int |\ell_{\mathbf{z}} - \ell_{\mathbf{z}_j}|^2 d\nu \leq \varepsilon^2$, for any probability measure ν . Indeed, $\int |\ell_{\mathbf{z}} - \ell_{\mathbf{z}_j}|^2 d\nu \leq 64\lambda_K^2 \|\mathbf{z} - \mathbf{z}_j\|_{\infty}^2 C_K^2 h^{-2}$ that is less than ε^2 , if we cover \mathcal{Z} by a grid of m points (\mathbf{z}_j) in \mathcal{Z} s.t. $\|\mathbf{z} - \mathbf{z}_j\|_{\infty} \leq \varepsilon h / (8C_K \lambda_K) := \varepsilon \delta$. This can be done with $m \leq \varepsilon^{-p} \lceil \prod_{k=1}^p ((b_k - a_k) / \delta) \rceil = \varepsilon^{-p} \lceil \mathcal{V} \delta^{-p} \rceil$ points. Then, with the notations of [24], $L = p$ and $D = \mathcal{V}(8C_K \lambda_K / h)^p$. As above, this yields

$$\mathbb{P}\left(\sup_{\mathbf{z} \in \mathcal{Z}} \frac{1}{n^2} \left| \sum_{1 \leq i \neq j \leq n} \xi_{\mathbf{z}}(\mathbf{X}_i, \mathbf{Z}_i, \mathbf{X}_j, \mathbf{Z}_j) \right| > t\right) \leq C_2 D \exp\left(-\frac{\alpha_2 n h^p t}{8 f_{\mathbf{z}, \max}} \int K^2\right), \quad (15)$$

when $t \leq 2h^p (\int K^2)^3 f_{\mathbf{z}, \max}^3 / C_K^4$. It remains to bound $\mathbb{P}(\sup_{\mathbf{z} \in \mathcal{Z}} |n^{-1} \sum_{i=1}^n \bar{g}_i| > t/4)$. Consider the family of functions

$$\mathcal{F} := \{(\mathbf{x}_1, \mathbf{z}_1) \in \mathbb{R} \times \mathcal{Z} \mapsto \frac{h^p}{4C_K^2} \mathbb{E}[g_{\mathbf{z}}(\mathbf{x}_1, \mathbf{z}_1, \mathbf{X}, \mathbf{Z})], \mathbf{z} \in \mathcal{Z}\}.$$

This family of functions is bounded is one and its variance is less than $\bar{\sigma}^2 := h^p \int K^2 f_{\mathbf{z}, \max}^3 (\int |K|)^2$. Apply Propositions 9 and 10 in [11] that is coming from [26]: for some universal constants A_1 and A_2 , some constant $A_{\bar{g}}$ that depends on K and $f_{\mathbf{z}, \max}$ (see Proposition 1 in [26]) and for every $x > 0$,

$$\begin{aligned} \mathbb{P}\left(\sup_{\mathbf{z} \in \mathcal{Z}} \frac{h^p}{4C_K^2} \left| \sum_{i=1}^n \mathbb{E}[g_{\mathbf{z}}(\mathbf{X}_i, \mathbf{Z}_i, \mathbf{X}, \mathbf{Z}) | \mathbf{X}_i, \mathbf{Z}_i] \right| > A_1(x + A_{\bar{g}} n^{1/2} \bar{\sigma} \ln(1/\bar{\sigma}))\right) \\ \leq 2 \left(\exp\left(-\frac{A_2 x^2}{n \bar{\sigma}^2}\right) + \exp(-A_2 x) \right), \text{ or} \end{aligned}$$

$$\mathbb{P}\left(\sup_{\mathbf{z} \in \mathcal{Z}} \frac{1}{n} \left| \sum_{i=1}^n \bar{g}_i \right| > 4A_1 C_K^2 \left(x - \frac{A_{\bar{g}} \bar{\sigma}}{n^{1/2} h^p} \ln(\bar{\sigma})\right)\right) \leq 2 \exp\left(-\frac{A_2 n h^{2p} x^2}{\bar{\sigma}^2}\right) + 2 \exp(-A_2 n h^p x).$$

For any positive t s.t. $4A_1 C_K^2 (n-1) A_{\bar{g}} \bar{\sigma} \ln(1/\bar{\sigma}) < n^{3/2} h^p t / 8$, note that we can find a real $x > t h^p / (16 C_K^2 A_1)$.

Then, we have

$$\mathbb{P}\left(\sup_{\mathbf{z} \in \mathcal{Z}} \frac{(n-1)}{n^2} \left| \sum_{i=1}^n \bar{g}_i \right| > \frac{t}{4}\right) \leq 2 \exp\left(-\frac{A_2 n h^p t^2 C_K^{-4}}{16^2 A_1^2 \int K^2 f_{\mathbf{z}, \max}^3 (\int |K|)^2}\right) + 2 \exp\left(-\frac{A_2 n h^p t}{16 C_K^2 A_1}\right). \quad (16)$$

Therefore, for such t , we obtain from (16) and (15) that

$$\begin{aligned} \mathbb{P}\left(\sup_{\mathbf{z} \in \mathcal{Z}} \left| \sum_{1 \leq i \neq j \leq n} (S_{i,j}(\mathbf{z}) - \mathbb{E}[S_{i,j}(\mathbf{z})]) \right| \geq t\right) &\leq C_2 D \exp\left(-\frac{\alpha_2 n h^p t}{8 (\int K^2) f_{\mathbf{z}, \max}}\right) \\ &+ 2 \exp\left(-\frac{A_2 n h^p t^2 C_K^{-4}}{15^2 A_1^2 \int K^2 f_{\mathbf{z}, \max}^3 (\int |K|)^2}\right) + 2 \exp\left(-\frac{A_2 n h^p t}{15 C_K^2 A_1}\right). \end{aligned}$$

for sufficiently large integers n . The latter inequality, (13) and (14) yield the exponential upper bound. \square

A.6. Proof of Proposition 7

Note that $\tau_{1,2|\mathbf{Z}=\mathbf{z}} = \mathbb{E}[g_k(\mathbf{X}_1, \mathbf{X}_2) | \mathbf{Z}_1 = \mathbf{z}, \mathbf{Z}_2 = \mathbf{z}]$ for every $k = 1, 2, 3$, and that our estimators with the weights (2) can be written as $\hat{\tau}_{1,2|\mathbf{Z}=\mathbf{z}}^{(k)} := U_n(g_k) / \{U_n(1) + \epsilon_n\}$, where

$$U_n(g) := \frac{1}{n(n-1)} \sum_{1 \leq i \neq j \leq n} g(\mathbf{X}_i, \mathbf{X}_j) \frac{K_h(\mathbf{z} - \mathbf{Z}_i) K_h(\mathbf{z} - \mathbf{Z}_j)}{\mathbb{E}[K_h(\mathbf{z} - \mathbf{Z})]^2} =: \frac{1}{n(n-1)} \sum_{1 \leq i \neq j \leq n} g_{i,j},$$

for any measurable bounded function g , with the residual diagonal term $\epsilon_n := \sum_{i=1}^n K_h^2(\mathbf{z} - \mathbf{Z}_i) / \{n(n-1) \mathbb{E}[K_h(\mathbf{z} - \mathbf{Z})]^2\}$. By Bochner's lemma (see Bosq and Lecoutre [32]), ϵ_n is $O_P((nh^p)^{-1})$, and it will be negligible compared to $U_n(1)$. Since the reasoning will be exactly the same for every estimator $\tau_{1,2|\mathbf{z}}^{(k)}$, i.e. for every function g_k , $k = 1, 2, 3$, we omit the sub-index k . Then, the functions g_k will be simply denoted by g .

The expectation of our U-statistics is

$$\begin{aligned} \mathbb{E}[U_n(g)] &:= \mathbb{E}[g(\mathbf{X}_1, \mathbf{X}_2) K_h(\mathbf{z} - \mathbf{Z}_1) K_h(\mathbf{z} - \mathbf{Z}_2)] / \mathbb{E}[K_h(\mathbf{z} - \mathbf{Z})]^2 \\ &= \int g(\mathbf{x}_1, \mathbf{x}_2) K(\mathbf{t}_1) K(\mathbf{t}_2) f_{\mathbf{X}, \mathbf{Z}}(\mathbf{x}_1, \mathbf{z} + h\mathbf{t}_1) f_{\mathbf{X}, \mathbf{Z}}(\mathbf{x}_2, \mathbf{z} + h\mathbf{t}_2) d\mathbf{x}_1 d\mathbf{x}_2 d\mathbf{t}_1 d\mathbf{t}_2 / \mathbb{E}[K_h(\mathbf{z} - \mathbf{Z})]^2 \\ &\rightarrow \frac{1}{f_{\mathbf{Z}}^2(\mathbf{z})} \int g(\mathbf{x}_1, \mathbf{x}_2) f_{\mathbf{X}, \mathbf{Z}}(\mathbf{x}_1, \mathbf{z}) f_{\mathbf{X}, \mathbf{Z}}(\mathbf{x}_2, \mathbf{z}) d\mathbf{x}_1 d\mathbf{x}_2 = \mathbb{E}[g(\mathbf{X}_1, \mathbf{X}_2) | \mathbf{Z}_1 = \mathbf{z}, \mathbf{Z}_2 = \mathbf{z}], \end{aligned}$$

applying Bochner's lemma to $\mathbf{z} \mapsto \int g(\mathbf{x}_1, \mathbf{x}_2) f_{\mathbf{X}|\mathbf{Z}=\mathbf{z}}(\mathbf{x}_1) f_{\mathbf{X}|\mathbf{Z}=\mathbf{z}}(\mathbf{x}_2) d\mathbf{x}_1 d\mathbf{x}_2 = \tau_{1,2|\mathbf{Z}=\mathbf{z}}$, that is a continuous function by assumption.

Set $\theta_n := \mathbb{E}[U_n(g)]$, $g^*(\mathbf{x}_1, \mathbf{x}_2) := (g(\mathbf{x}_1, \mathbf{x}_2) + g(\mathbf{x}_2, \mathbf{x}_1))/2$ and $g_{i,j}^* = (g_{i,j} + g_{j,i})/2$ for every (i, j) , $i \neq j$. Note that $U_n(g) = U_n(g^*)$. Since g^* is symmetrical, the Hájek projection $\hat{U}_n(g^*)$ of $U_n(g^*)$ satisfies $\hat{U}_n(g^*) := 2 \sum_{j=1}^n \mathbb{E}[g_{0,j}^* | \mathbf{X}_j, \mathbf{Z}_j] / n - \theta_n$. Note that $\mathbb{E}[\hat{U}_n(g^*)] = \theta_n = \tau_{1,2|\mathbf{Z}=\mathbf{z}} + o_P(1)$. Since $\text{Var}(\hat{U}_n(g^*)) = 4 \text{Var}(\mathbb{E}[g_{0,j}^* | \mathbf{X}_j, \mathbf{Z}_j]) / n = O((nh^p)^{-1})$, then $\hat{U}_n(g^*) = \theta_n + o_P(1) = \tau_{1,2|\mathbf{Z}=\mathbf{z}} + o_P(1)$.

Moreover, using the notation $\bar{g}_{i,j} := g_{i,j}^* - \mathbb{E}[g_{i,j}^* | \mathbf{X}_j, \mathbf{Z}_j] - \mathbb{E}[g_{i,j}^* | \mathbf{X}_i, \mathbf{Z}_i] + \theta_n$ for $1 \leq i \neq j \leq n$, we have $U_n(g^*) - \hat{U}_n(g^*) = \sum_{1 \leq i \neq j \leq n} \bar{g}_{i,j} / n(n-1)$. By usual U-statistics calculations, it can be easily checked that

$$\text{Var}(U_n(g^*) - \hat{U}_n(g^*)) = \frac{1}{n^2(n-1)^2} \sum_{1 \leq i_1 \neq j_1 \leq n} \sum_{1 \leq i_2 \neq j_2 \leq n} \mathbb{E}[\bar{g}_{i_1, j_1} \bar{g}_{i_2, j_2}] = O\left(\frac{1}{n^2 h^{2p}}\right).$$

Indeed, when all indices (i_1, i_2, j_1, j_2) are different, or when there is a single identity among them, $\mathbb{E}[\bar{g}_{i_1, j_1} \bar{g}_{i_2, j_2}]$ is zero. The first nonzero terms arise when there are two identities among the indices, i.e. $i_1 = i_2$ and $j_1 = j_2$ (or $i_1 = j_2$ and $j_1 = i_2$). In the latter case, we get an upper bound as $O((nh^p)^{-2})$ when $f_{\mathbf{Z}}$ is continuous at \mathbf{z} , by usual changes of variable techniques and Bochner's Lemma. Then, $U_n(g^*) = \hat{U}_n(g^*) + o_P(1) = \tau_{1,2|\mathbf{Z}=\mathbf{z}} + o_P(1)$. Note that $U_n(1) + \epsilon_n$ tends to one in probability (Bochner's lemma). As a consequence, $\hat{\tau}_{1,2|\mathbf{Z}=\mathbf{z}} = U_n(g^*) / (U_n(1) + \epsilon_n)$ tends to $\tau_{1,2|\mathbf{Z}=\mathbf{z}}/1$ by the continuous mapping theorem. \square

A.7. Proof of Proposition 8

Let us note that

$$\tau_{1,2|\mathbf{Z}=\mathbf{z}} = \mathbb{E}[g_k(\mathbf{X}_1, \mathbf{X}_2) | \mathbf{Z}_1 = \mathbf{z}, \mathbf{Z}_2 = \mathbf{z}] = \int g_k(\mathbf{x}_1, \mathbf{x}_2) f_{\mathbf{X}|\mathbf{Z}=\mathbf{z}}(\mathbf{x}_1) f_{\mathbf{X}|\mathbf{Z}=\mathbf{z}}(\mathbf{x}_2) d\mathbf{x}_1 d\mathbf{x}_2 = \frac{\phi_k(\mathbf{z})}{f_{\mathbf{Z}}^2(\mathbf{z})},$$

where $\phi_k(\mathbf{z}) := \int g_k(\mathbf{x}_1, \mathbf{x}_2) f_{\mathbf{X}, \mathbf{Z}}(\mathbf{x}_1, \mathbf{z}) f_{\mathbf{X}, \mathbf{Z}}(\mathbf{x}_2, \mathbf{z}) d\mathbf{x}_1 d\mathbf{x}_2$. Also write $\hat{\tau}_{1,2|\mathbf{Z}=\mathbf{z}}^{(k)} = \hat{\phi}_k(\mathbf{z}) / \hat{f}_{\mathbf{Z}}^2(\mathbf{z})$, where $\hat{\phi}_k(\mathbf{z}) := n^{-2} \sum_{i,j=1}^n K_h(\mathbf{Z}_i - \mathbf{z}) K_h(\mathbf{Z}_j - \mathbf{z}) g_k(\mathbf{X}_i, \mathbf{X}_j)$ and $\hat{f}_{\mathbf{Z}}(\mathbf{z}) := n^{-1} \sum_{i=1}^n K_h(\mathbf{Z}_i - \mathbf{z})$. Therefore, we have

$$\hat{\tau}_{1,2|\mathbf{Z}=\mathbf{z}}^{(k)} - \tau_{1,2|\mathbf{Z}=\mathbf{z}} = \frac{\hat{\phi}_k(\mathbf{z}) - \phi_k(\mathbf{z})}{\hat{f}_{\mathbf{Z}}^2(\mathbf{z})} - \tau_{1,2|\mathbf{Z}=\mathbf{z}} \frac{\hat{f}_{\mathbf{Z}}(\mathbf{z}) - f_{\mathbf{Z}}(\mathbf{z})}{\hat{f}_{\mathbf{Z}}^2(\mathbf{z})} \times (\hat{f}_{\mathbf{Z}}(\mathbf{z}) + f_{\mathbf{Z}}(\mathbf{z})).$$

By usual uniform consistency results (see for example Bosq and Lecoutre [32]), $\sup_{\mathbf{z} \in \mathcal{Z}} |\hat{f}_{\mathbf{Z}}(\mathbf{z}) - f_{\mathbf{Z}}(\mathbf{z})| \rightarrow 0$ almost surely, as $n \rightarrow \infty$. We deduce that

$$\min_{\mathbf{z} \in \mathcal{Z}} \hat{f}_{\mathbf{Z}}^2(\mathbf{z}) \geq f_{\mathbf{Z},\min}^2/2, \quad \text{and} \quad \max_{\mathbf{z} \in \mathcal{Z}} |\hat{f}_{\mathbf{Z}}(\mathbf{z}) + f_{\mathbf{Z}}(\mathbf{z})| \leq 2 \max_{\mathbf{z} \in \mathcal{Z}} f_{\mathbf{Z}}(\mathbf{z}) \quad \text{a.s.}$$

This means it is sufficient to prove the uniform strong consistency of $\hat{\phi}_k$ on \mathcal{Z} , to obtain that $\sup_{\mathbf{z} \in \mathcal{Z}} |\hat{\tau}_{1,2|\mathbf{Z}=\mathbf{z}}^{(k)} - \tau_{1,2|\mathbf{Z}=\mathbf{z}}^{(k)}|$ tends to zero a.s.

Note that, by Bochner's Lemma, $\sup_{\mathbf{z} \in \mathcal{Z}} |\mathbb{E}[\hat{\phi}_k(\mathbf{z})] - \phi_k(\mathbf{z})| \rightarrow 0$. Then, it remains to show that $\sup_{\mathbf{z} \in \mathcal{Z}} |\hat{\phi}_k(\mathbf{z}) - \mathbb{E}[\hat{\phi}_k(\mathbf{z})]| \rightarrow 0$ almost surely. Let $\rho_n > 0$ be such that we cover \mathcal{Z} by the union of l_n open balls $B(\mathbf{t}_l, \rho_n)$, where $\mathbf{t}_1, \dots, \mathbf{t}_{l_n} \in \mathbb{R}^p$ and $l_n \in \mathbb{N}^*$. Then

$$\sup_{\mathbf{z} \in \mathcal{Z}} |\hat{\phi}_k(\mathbf{z}) - \mathbb{E}[\hat{\phi}_k(\mathbf{z})]| \leq \sup_{l=1, \dots, l_n} |\hat{\phi}_k(\mathbf{t}_l) - \mathbb{E}[\hat{\phi}_k(\mathbf{t}_l)]| + A_n,$$

where $A_n := \sup_{l=1, \dots, l_n} \sup_{\mathbf{z} \in B(\mathbf{t}_l, \rho_n)} |\hat{\phi}_k(\mathbf{z}) - \hat{\phi}_k(\mathbf{t}_l) - (\mathbb{E}[\hat{\phi}_k(\mathbf{z})] - \mathbb{E}[\hat{\phi}_k(\mathbf{t}_l)])|$. For any index $l \in \{1, \dots, l_n\}$ and any $\mathbf{z} \in B(\mathbf{t}_l, \rho_n)$, a first-order expansion yields

$$\begin{aligned} & |\hat{\phi}_k(\mathbf{z}) - \hat{\phi}_k(\mathbf{t}_l) - (\mathbb{E}[\hat{\phi}_k(\mathbf{z})] - \mathbb{E}[\hat{\phi}_k(\mathbf{t}_l)])| \\ &= \left| \frac{1}{n(n-1)} \sum_{1 \leq i \neq j \leq n} g_k(\mathbf{X}_i, \mathbf{X}_j) K_h(\mathbf{z} - \mathbf{Z}_i) K_h(\mathbf{z} - \mathbf{Z}_j) \right. \\ & \quad \left. - \frac{1}{n(n-1)} \sum_{1 \leq i \neq j \leq n} g_k(\mathbf{X}_i, \mathbf{X}_j) K_h(\mathbf{t}_l - \mathbf{Z}_i) K_h(\mathbf{t}_l - \mathbf{Z}_j) \right. \\ & \quad \left. - \left(\mathbb{E}[g_k(\mathbf{X}_1, \mathbf{X}_2) K_h(\mathbf{z} - \mathbf{Z}_1) K_h(\mathbf{z} - \mathbf{Z}_2)] - \mathbb{E}[g_k(\mathbf{X}_i, \mathbf{X}_j) K_h(\mathbf{t}_l - \mathbf{Z}_i) K_h(\mathbf{t}_l - \mathbf{Z}_j)] \right) \right| \\ &\leq \frac{C_{Lip,K}}{h^{2p+1}} |\mathbf{z} - \mathbf{t}_l| \left(\mathbb{E}[|g_k(\mathbf{X}_1, \mathbf{X}_2)|] + \frac{1}{n(n-1)} \sum_{1 \leq i \neq j \leq n} |g_k(\mathbf{X}_i, \mathbf{X}_j)| \right) \\ &= O\left(\frac{\rho_n}{h^{2p+1}}\right) = o(1), \end{aligned}$$

for some constant $C_{Lip,K}$ and by choosing $\rho_n = o(h_n^{2p+1})$. Actually, we can cover \mathcal{Z} in such a way that $l_n = O(h_n^{-p(2p+1)})$. This is always possible because \mathcal{Z} is a bounded set in \mathbb{R}^p . The previous upper bound is uniform w.r.t. l and $\mathbf{z} \in B(\mathbf{t}_l, \rho_n)$, proving $A_n = o(1)$ everywhere.

Now, for every $l \leq l_n$, apply Equation (8) for every $\mathbf{z} = \mathbf{t}_l$. For any $t > 0$, this yields

$$\mathbb{P}\left(\frac{1}{n(n-1)} \left| \sum_{i \neq j} g^{(l)}((\mathbf{X}_i, \mathbf{Z}_i), (\mathbf{X}_j, \mathbf{Z}_j)) - \mathbb{E}\left[g^{(l)}((\mathbf{X}_1, \mathbf{Z}_1), (\mathbf{X}_2, \mathbf{Z}_2))\right] \right| > t\right) \leq \exp\left(-\frac{C_0 n h_n^{2p} t^2}{C_1 + C_2 t}\right),$$

for some positive constants C_0, C_1, C_2 , by setting

$$g^{(l)}((\mathbf{X}_i, \mathbf{Z}_i), (\mathbf{X}_j, \mathbf{Z}_j)) := g_k(\mathbf{X}_i, \mathbf{X}_j) K_h(\mathbf{t}_l - \mathbf{Z}_i) K_h(\mathbf{t}_l - \mathbf{Z}_j).$$

Therefore, we deduce

$$\mathbb{P} \left(\sup_{l=1, \dots, l_n} |\hat{\phi}_k(\mathbf{t}_l) - \mathbb{E}[\hat{\phi}_k(\mathbf{t}_l)]| \geq t \right) \leq C_4 h_n^{-p(2p+1)} \exp \left(- \frac{C_0 n h_n^{2p} t^2}{C_1 + C_2 t} \right),$$

for some constant C_4 . Finally, applying Borel-Cantelli lemma, $\sup_{\mathbf{z} \in \mathcal{Z}} |\hat{\phi}_k(\mathbf{z}) - \mathbb{E}[\hat{\phi}_k(\mathbf{z})]|$ tends to zero a.s., proving the result. \square

A.8. Proof of Proposition 9

By Markov's inequality, $\sum_{i=1}^n w_{i,n}^2(\mathbf{z}) = O_P((nh^p)^{-1})$ for any \mathbf{z} , that tends to zero. Then, by Slutsky's theorem, we get an asymptotic equivalence between the limiting laws of any $\hat{\tau}_{1,2|\mathbf{z}}^{(k)}$, $k = 1, 2, 3$, and of their linearly transformed versions $\tilde{\tau}_{1,2|\mathbf{z}}$. Thus, we will prove the asymptotic normality of $\hat{\tau}_{1,2|\mathbf{z}}^{(k)}$ for some index $k = 1, 2, 3$, simply denoted by $\hat{\tau}_{1,2|\mathbf{z}}$.

Let $g^*(\mathbf{x}_1, \mathbf{x}_2) := (g_k(\mathbf{x}_1, \mathbf{x}_2) + g_k(\mathbf{x}_2, \mathbf{x}_1))/2$ for some index $k = 1, 2, 3$ (that will be implicit in the proof). We now study the joint behavior of $(\hat{\tau}_{1,2|\mathbf{Z}=\mathbf{z}_i} - \tau_{1,2|\mathbf{Z}=\mathbf{z}_i})_{i=1, \dots, n'}$. We will extend Stute [33]'s approach, in the case of multivariate conditioning variable \mathbf{z} and studying the joint distribution of U-statistics at several conditioning points. As in the proof of Proposition 7, the estimator with the weights given by (2) can be rewritten as $\hat{\tau}_{1,2|\mathbf{Z}=\mathbf{z}_i} := U_{n,i}(g^*) / (U_{n,i}(1) + \epsilon_{n,i})$, where

$$U_{n,i}(g) := \frac{1}{n(n-1)\mathbb{E}[K_h(\mathbf{z}'_i - \mathbf{Z})]^2} \sum_{j_1, j_2=1, j_1 \neq j_2}^n g(\mathbf{X}_{j_1}, \mathbf{X}_{j_2}) K_h(\mathbf{z}'_i - \mathbf{Z}_{j_1}) K_h(\mathbf{z}'_i - \mathbf{Z}_{j_2}),$$

for any bounded measurable function $g : \mathbb{R}^4 \rightarrow \mathbb{R}$. Moreover, $\sup_{i=1, \dots, n'} |\epsilon_{n,i}| = O_P(n^{-1}h^{-p})$. By a limited expansion of $f_{\mathbf{X}, \mathbf{Z}}$ w.r.t. its second argument, and under Assumption 3.4, we easily check that $\mathbb{E}[U_{n,i}(g)] = \tau_{1,2|\mathbf{Z}=\mathbf{z}_i} + r_{n,i}$, where $|r_{n,i}| \leq C_0 h_n^\alpha / f_{\mathbf{Z}}^2(\mathbf{z}'_i)$, for some constant C_0 that is independent of i .

Now, we prove the joint asymptotic normality of $(U_{n,i}(g))_{i=1, \dots, n'}$. The Hájek projection $\hat{U}_{n,i}(g)$ of $U_{n,i}(g)$ satisfies $\hat{U}_{n,i}(g) := 2 \sum_{j=1}^n g_{n,i}(\mathbf{X}_j, \mathbf{Z}_j) / n - \theta_n$, where $\theta_n := \mathbb{E}[U_{n,i}(g)]$ and

$$g_{n,i}(\mathbf{x}, \mathbf{z}) := K_h(\mathbf{z}'_i - \mathbf{z}) \mathbb{E}[g(\mathbf{X}, \mathbf{x}) K_h(\mathbf{z}'_i - \mathbf{Z})] / \mathbb{E}[K_h(\mathbf{z}'_i - \mathbf{Z})]^2.$$

Lemma 17. *Under the assumptions of Proposition 9, for any measurable bounded function g ,*

$$(nh^p)^{1/2} \left(\hat{U}_{n,i}(g) - \mathbb{E}[U_{n,i}(g)] \right)_{i=1, \dots, n'} \xrightarrow{D} \mathcal{N}(0, M_\infty(g)), \text{ as } n \rightarrow \infty,$$

where, for $1 \leq i, j \leq n'$,

$$[M_\infty(g)]_{i,j} := \frac{4 \int K^2 \mathbb{1}_{\{\mathbf{z}'_i = \mathbf{z}'_j\}}}{f_{\mathbf{Z}}(\mathbf{z}'_i)} \int g(\mathbf{x}_1, \mathbf{x}) g(\mathbf{x}_2, \mathbf{x}) f_{\mathbf{X}|\mathbf{Z}=\mathbf{z}'_i}(\mathbf{x}) f_{\mathbf{X}|\mathbf{Z}=\mathbf{z}'_i}(\mathbf{x}_1) f_{\mathbf{X}|\mathbf{Z}=\mathbf{z}'_i}(\mathbf{x}_2) d\mathbf{x} d\mathbf{x}_1 d\mathbf{x}_2.$$

This lemma is proved in A.9. Similarly as in the proof of Lemma 2.2 in Stute [33], for every $i = 1, \dots, n'$ and every bounded symmetrical measurable function g , we have $(nh^p)^{1/2} \text{Var}[\hat{U}_{n,i}(g) - U_{n,i}(g)] = o(1)$, which implies

$$(nh^p)^{1/2} \left(U_{n,i}(g) - \mathbb{E}[U_{n,i}(g)] \right)_{i=1, \dots, n'} \xrightarrow{D} \mathcal{N}(0, M_\infty(g)), \text{ as } n \rightarrow \infty.$$

Considering two measurable bounded functions g_1 and g_2 , we have $U_{n,i}(c_1g_1 + c_2g_2) = c_1U_{n,i}(g_1) + c_2U_{n,i}(g_2)$ for every numbers c_1, c_2 . By the Cramér-Wold device, we check that

$$(nh^p)^{1/2} \left(\left(U_{n,i}(g_1) - \mathbb{E}[U_{n,i}(g_1)] \right)_{i=1, \dots, n'}, \left(U_{n,i}(g_2) - \mathbb{E}[U_{n,i}(g_2)] \right)_{i=1, \dots, n'} \right) \\ \xrightarrow{D} \mathcal{N} \left(0, \begin{bmatrix} M_\infty(g_1) & M_\infty(g_1, g_2) \\ M_\infty(g_1, g_2) & M_\infty(g_2) \end{bmatrix} \right),$$

as $n \rightarrow \infty$, where

$$[M_\infty(g_1, g_2)]_{i,j} := \frac{4 \int K^2 \mathbb{1}_{\{\mathbf{z}'_i = \mathbf{z}'_j\}}}{f_{\mathbf{Z}}(\mathbf{z}'_i)} \int g_1(\mathbf{x}_1, \mathbf{x}) g_2(\mathbf{x}_2, \mathbf{x}) f_{\mathbf{X}|\mathbf{Z}=\mathbf{z}'_i}(\mathbf{x}) f_{\mathbf{X}|\mathbf{Z}=\mathbf{z}'_i}(\mathbf{x}_1) f_{\mathbf{X}|\mathbf{Z}=\mathbf{z}'_i}(\mathbf{x}_2) d\mathbf{x} d\mathbf{x}_1 d\mathbf{x}_2.$$

Set $\tilde{\tau}_{1,2|\mathbf{Z}=\mathbf{z}'_i} := U_{n,i}(g^*) / U_{n,i}(1)$. Since $(nh_n^p)^{1/2}(\hat{\tau}_{1,2|\mathbf{Z}=\mathbf{z}'_i} - \tilde{\tau}_{1,2|\mathbf{Z}=\mathbf{z}'_i}) = O_P((nh_n^p)^{1/2}\epsilon_{n,i})$ is $o_P(1)$, it is sufficient to establish the asymptotic law of $(nh_n^p)^{1/2}(\tilde{\tau}_{1,2|\mathbf{Z}=\mathbf{z}'_i} - \tau_{1,2|\mathbf{Z}=\mathbf{z}'_i})$. Since $\mathbb{E}[U_{n,i}(1)] = 1 + o((nh^p)^{-1/2})$ and $\mathbb{E}[U_{n,i}(g^*)] = \tau_{1,2|\mathbf{Z}=\mathbf{z}'_i} + o((nh_n^p)^{-1/2})$, we get

$$(nh^p)^{1/2} \left(\left(U_{n,i}(g^*) - \tau_{1,2|\mathbf{Z}=\mathbf{z}'_i} \right)_{i=1, \dots, n'}, \left(U_{n,i}(1) - 1 \right)_{i=1, \dots, n'} \right) \\ \xrightarrow{D} \mathcal{N} \left(0, \begin{bmatrix} M_\infty(g^*) & M_\infty(g^*, 1) \\ M_\infty(g^*, 1) & M_\infty(1) \end{bmatrix} \right), \text{ as } n \rightarrow \infty.$$

Now apply the Delta-method with the function $\rho(\mathbf{x}, \mathbf{y}) := \mathbf{x}/\mathbf{y}$ where \mathbf{x} and \mathbf{y} are real-valued vectors of size n' and the division has to be understood component-wise. The Jacobian of ρ is given by the $n' \times 2n'$ matrix

$$J_\rho(\mathbf{x}, \mathbf{y}) = \left[\text{Diag}(y_1^{-1}, \dots, y_{n'}^{-1}), \text{Diag}(-x_1 y_1^{-2}, \dots, -x_{n'} y_{n'}^{-2}) \right],$$

where, for any vector v of size n' , $\text{Diag}(v)$ is the diagonal matrix whose diagonal elements are the v_i , with $i = 1, \dots, n'$. We deduce $(nh^p)^{1/2}(\tilde{\tau}_{1,2|\mathbf{Z}=\mathbf{z}'_i} - \tau_{1,2|\mathbf{Z}=\mathbf{z}'_i})_{i=1, \dots, n'} \xrightarrow{D} \mathcal{N}(0, \mathbb{H})$, as $n \rightarrow \infty$, setting

$$\mathbb{H} := J_\rho(\vec{\tau}, \mathbf{e}) \begin{bmatrix} M_\infty(g^*) & M_\infty(g^*, 1) \\ M_\infty(g^*, 1) & M_\infty(1) \end{bmatrix} J_\rho(\vec{\tau}, \mathbf{e})^T,$$

where $\vec{\tau} = (\tau_{1,2|\mathbf{Z}=\mathbf{z}'_i})_{i=1, \dots, n'}$ and \mathbf{e} is the vector of size n' whose all components are equal to 1. Thus, we have $J_\rho(\vec{\tau}, \mathbf{e}) = \left[\text{Id}_{n'}, -\text{Diag}(\vec{\tau}) \right]$, denoting by $\text{Id}_{n'}$ the identity matrix of size n' and by $\text{Diag}(\vec{\tau})$ the diagonal matrix of size n' whose diagonal elements are the $\tau_{1,2|\mathbf{z}'_i}$, for $i = 1, \dots, n'$. To be specific, we get

$$\mathbb{H} = M_\infty(g^*) - \text{Diag}(\vec{\tau})M_\infty(g^*, 1) - M_\infty(g^*, 1)\text{Diag}(\vec{\tau}) + \text{Diag}(\vec{\tau})M_\infty(1)\text{Diag}(\vec{\tau}).$$

For i, j in $\{1, \dots, n'\}$ and using the symmetry of the function g^* , we obtain

$$[M_\infty(g^*)]_{i,j} = 4 \int K^2 \mathbb{1}_{\{\mathbf{z}'_i = \mathbf{z}'_j\}} \mathbb{E}[g^*(\mathbf{X}_1, \mathbf{X})g^*(\mathbf{X}_2, \mathbf{X}) | \mathbf{Z} = \mathbf{Z}_1 = \mathbf{Z}_2 = \mathbf{z}'_i] / f_{\mathbf{Z}}(\mathbf{z}'_i), \\ [\text{Diag}(\vec{\tau})M_\infty(g^*, 1)]_{i,j} = 4\tau_{1,2|\mathbf{Z}=\mathbf{z}'_i} \int K^2 \mathbb{1}_{\{\mathbf{z}'_i = \mathbf{z}'_j\}} \mathbb{E}[g^*(\mathbf{X}_1, \mathbf{X}) | \mathbf{Z} = \mathbf{Z}_1 = \mathbf{z}'_i] / f_{\mathbf{Z}}(\mathbf{z}'_i) \\ = 4 \int K^2 \mathbb{1}_{\{\mathbf{z}'_i = \mathbf{z}'_j\}} \tau_{1,2|\mathbf{Z}=\mathbf{z}'_i}^2 / f_{\mathbf{Z}}(\mathbf{z}'_i) = [M_\infty(g^*, 1)\text{Diag}(\vec{\tau})]_{i,j} = [\text{Diag}(\vec{\tau})M_\infty(1)\text{Diag}(\vec{\tau})]_{i,j}.$$

As a consequence, we obtain

$$[\mathbb{H}]_{i,j} = \frac{4 \int K^2 \mathbb{1}_{\{\mathbf{z}'_i = \mathbf{z}'_j\}}}{f_{\mathbf{Z}}(\mathbf{z}'_i)} \left(\mathbb{E}[g^*(\mathbf{X}_1, \mathbf{X})g^*(\mathbf{X}_2, \mathbf{X}) | \mathbf{Z} = \mathbf{Z}_1 = \mathbf{Z}_2 = \mathbf{z}'_i] - \tau_{1,2|\mathbf{Z}=\mathbf{z}'_i}^2 \right). \quad \square$$

A.9. Proof of Lemma 17

Let us first evaluate the variance-covariance matrix $M_{n,n'} := [Cov(\hat{U}_{n,i}, \hat{U}_{n,j})]_{1 \leq i, j \leq n'}$. Note that $\mathbb{E}[g_{n,i}(\mathbf{X}_j, \mathbf{Z}_j)] = \mathbb{E}[\hat{U}_{n,i}] = \mathbb{E}[U_{n,i}(g)]$, and that

$$\left((nh^p)^{1/2} (\hat{U}_{n,i} - \mathbb{E}[U_{n,i}(g)]) \right)_{i=1, \dots, n'} = \frac{2h^{p/2}}{n^{1/2}} \sum_{j=1}^n (g_{n,i}(\mathbf{X}_j, \mathbf{Z}_j) - \mathbb{E}[U_{n,i}(g)])_{i=1, \dots, n'},$$

that is a sum of independent vectors. Thus, $Cov(\hat{U}_{n,i}, \hat{U}_{n,j}) = 4n^{-1} Cov(g_{n,i}(\mathbf{X}, \mathbf{Z}), g_{n,j}(\mathbf{X}, \mathbf{Z}))$, for every i, j in $\{1, \dots, n'\}$, and

$$\begin{aligned} & \mathbb{E}[g_{n,i}(\mathbf{X}, \mathbf{Z})g_{n,j}(\mathbf{X}, \mathbf{Z})] \\ &= \int K_h(\mathbf{z}'_i - \mathbf{z})K_h(\mathbf{z}'_j - \mathbf{z}) \frac{\mathbb{E}[g(\mathbf{X}, \mathbf{x})K_h(\mathbf{z}'_i - \mathbf{Z})]\mathbb{E}[g(\mathbf{X}, \mathbf{x})K_h(\mathbf{z}'_j - \mathbf{Z})]}{\mathbb{E}[K_h(\mathbf{z}'_i - \mathbf{Z})]^2\mathbb{E}[K_h(\mathbf{z}'_j - \mathbf{Z})]^2} f_{\mathbf{X}, \mathbf{Z}}(\mathbf{x}, \mathbf{z}) d\mathbf{x} d\mathbf{z} \\ &\sim \frac{1}{h^p f_{\mathbf{Z}}^2(\mathbf{z}'_i) f_{\mathbf{Z}}^2(\mathbf{z}'_j)} \int g(\mathbf{x}_1, \mathbf{x})g(\mathbf{x}_2, \mathbf{x})K_h(\mathbf{z}'_i - \mathbf{z})K_h(\mathbf{z}'_j - \mathbf{z})K_h(\mathbf{z}'_i - \mathbf{w}_1)K_h(\mathbf{z}'_j - \mathbf{w}_2) \\ &\times f_{\mathbf{X}, \mathbf{Z}}(\mathbf{x}, \mathbf{z})f_{\mathbf{X}, \mathbf{Z}}(\mathbf{x}_1, \mathbf{w}_1)f_{\mathbf{X}, \mathbf{Z}}(\mathbf{x}_2, \mathbf{w}_2) d\mathbf{x} d\mathbf{z} d\mathbf{x}_1 d\mathbf{w}_1 d\mathbf{x}_2 d\mathbf{w}_2 \\ &\sim \frac{1}{h^p f_{\mathbf{Z}}^2(\mathbf{z}'_i) f_{\mathbf{Z}}^2(\mathbf{z}'_j)} \int g(\mathbf{x}_1, \mathbf{x})g(\mathbf{x}_2, \mathbf{x})K(\mathbf{u}_1)K(\mathbf{u}_2)K(\mathbf{u})K\left(\frac{\mathbf{z}'_j - \mathbf{z}'_i}{h} + \mathbf{u}\right) f_{\mathbf{X}, \mathbf{Z}}(\mathbf{x}, \mathbf{z}'_i - h\mathbf{u}) \\ &\times f_{\mathbf{X}, \mathbf{Z}}(\mathbf{x}_1, \mathbf{z}'_i - h\mathbf{u}_1)f_{\mathbf{X}, \mathbf{Z}}(\mathbf{x}_2, \mathbf{z}'_j - h\mathbf{u}_2) d\mathbf{x} d\mathbf{u} d\mathbf{x}_1 d\mathbf{u}_1 d\mathbf{x}_2 d\mathbf{u}_2. \end{aligned}$$

If $i \neq j$ and K is compactly supported, the latter term is zero when n is sufficiently large, and $Cov(\hat{U}_{n,i}, \hat{U}_{n,j}) = -4n^{-1}\mathbb{E}[U_{n,i}]\mathbb{E}[U_{n,j}] \sim -4n^{-1}\tau_{1,2|\mathbf{Z}=\mathbf{z}'_i}\tau_{1,2|\mathbf{Z}=\mathbf{z}'_j} = o((nh^p)^{-1})$. Otherwise, $i = j$ and, as $\mathbb{E}[g_{n,i}(\mathbf{X}_1, \mathbf{Z}_1)] = O(1)$, we have

$$\begin{aligned} Var\left((g_{n,i}(\mathbf{X}, \mathbf{Z}))^2\right) &\sim \frac{1}{h^p f_{\mathbf{Z}}^4(\mathbf{z}'_i)} \int g(\mathbf{x}_1, \mathbf{x})g(\mathbf{x}_2, \mathbf{x})K(\mathbf{u}_1)K(\mathbf{u}_2)K^2(\mathbf{u})f_{\mathbf{X}, \mathbf{Z}}(\mathbf{x}, \mathbf{z}'_i - h\mathbf{u}) \\ &\times f_{\mathbf{X}, \mathbf{Z}}(\mathbf{x}_1, \mathbf{z}'_i - h\mathbf{u}_1)f_{\mathbf{X}, \mathbf{Z}}(\mathbf{x}_2, \mathbf{z}'_i - h\mathbf{u}_2) d\mathbf{x} d\mathbf{u} d\mathbf{x}_1 d\mathbf{u}_1 d\mathbf{x}_2 d\mathbf{u}_2 \\ &\sim \frac{\int K^2}{h^p f_{\mathbf{Z}}(\mathbf{z}'_i)} \int g(\mathbf{x}_1, \mathbf{x})g(\mathbf{x}_2, \mathbf{x})f_{\mathbf{X}|\mathbf{Z}=\mathbf{z}'_i}(\mathbf{x})f_{\mathbf{X}|\mathbf{Z}=\mathbf{z}'_i}(\mathbf{x}_1)f_{\mathbf{X}|\mathbf{Z}=\mathbf{z}'_i}(\mathbf{x}_2) d\mathbf{x} d\mathbf{x}_1 d\mathbf{x}_2, \end{aligned}$$

by Bochner's lemma. We have proved that, for every $i, j \in \{1, \dots, n'\}$,

$$nh^p [M_{n,n'}]_{i,j} \rightarrow \frac{4 \int K^2 \mathbb{1}_{\{\mathbf{z}'_i = \mathbf{z}'_j\}}}{f_{\mathbf{Z}}(\mathbf{z}'_i)} \int g(\mathbf{x}_1, \mathbf{x})g(\mathbf{x}_2, \mathbf{x})f_{\mathbf{X}|\mathbf{Z}=\mathbf{z}'_i}(\mathbf{x})f_{\mathbf{X}|\mathbf{Z}=\mathbf{z}'_i}(\mathbf{x}_1)f_{\mathbf{X}|\mathbf{Z}=\mathbf{z}'_i}(\mathbf{x}_2) d\mathbf{x} d\mathbf{x}_1 d\mathbf{x}_2,$$

as $n \rightarrow \infty$. Therefore, $nh^p M_{n,n'}$ tends to M_∞ .

We now verify Lyapunov's condition with third-order moments, so that the usual multivariate central limit theorem would apply. It is then sufficient to show that

$$\left(\frac{h^{p/2}}{n^{1/2}}\right)^3 \sum_{j=1}^n \mathbb{E}\left[|g_{n,i}(\mathbf{X}_j, \mathbf{Z}_j) - \mathbb{E}[U_{n,i}(g)]|^3\right] = o(1). \quad (17)$$

For any $j = 1, \dots, n$, we have

$$\begin{aligned} & \mathbb{E}\left[|g_{n,i}(\mathbf{X}_j, \mathbf{Z}_j) - \mathbb{E}[U_{n,i}(g)]|^3\right] \\ &\sim \int \left| \frac{1}{f_{\mathbf{Z}}^2(\mathbf{z}'_i)} \int g(\mathbf{x}_1, \mathbf{x})K_h(\mathbf{z}'_i - \mathbf{z}_1)K_h(\mathbf{z}'_i - \mathbf{z})f_{\mathbf{X}, \mathbf{Z}}(\mathbf{x}_1, \mathbf{z}_1) d\mathbf{x}_1 d\mathbf{z}_1 - \mathbb{E}[U_{n,i}(g)] \right|^3 f_{\mathbf{X}, \mathbf{Z}}(\mathbf{x}, \mathbf{z}) d\mathbf{x} d\mathbf{z}. \end{aligned}$$

By the change of variable $\mathbf{z}_1 = \mathbf{z}'_i - h\mathbf{t}_1$ and $\mathbf{z} = \mathbf{z}'_i - h\mathbf{t}$, we get

$$\begin{aligned} \mathbb{E}\left[|g_{n,i}(\mathbf{X}_j, \mathbf{Z}_j) - \mathbb{E}[U_{n,i}(g)]|^3\right] &\sim h^{-2p} \int \left| \frac{1}{f_{\mathbf{Z}}^2(\mathbf{z}'_i)} \int g(\mathbf{x}_1, \mathbf{x}) K(\mathbf{t}_1) K(\mathbf{t}) f_{\mathbf{X}, \mathbf{Z}}(\mathbf{x}_1, \mathbf{z}'_i - h\mathbf{t}_1) d\mathbf{x}_1 d\mathbf{t}_1 \right. \\ &\quad \left. - h^p \mathbb{E}[U_{n,i}(g)] \right|^3 f_{\mathbf{X}, \mathbf{Z}}(\mathbf{x}, \mathbf{z}'_i - h\mathbf{t}) d\mathbf{x} d\mathbf{t} = O(h^{-2p}), \end{aligned}$$

because of Bochner's lemma, under our assumptions. Therefore, we have obtained

$$\left(\frac{h^{p/2}}{n^{1/2}}\right)^3 \sum_{j=1}^n \mathbb{E}\left[|g_{n,i}(\mathbf{X}_j, \mathbf{Z}_j) - \mathbb{E}[U_{n,i}(g)]|^3\right] = O(h^{3p/2} n^{-3/2} n h^{-2p}) = O((nh^p)^{-1/2}) = o(1).$$

Therefore, we have checked Lyapunov's condition and the result follows. \square