BOOTSTRAPPING THE CONDITIONAL COPULA*

MAREK OMELKA[1], NOËL VERAVERBEKE[2], IRÈNE GIJBELS[3]

[1] Department of Probability and Statistics, Faculty of Mathematics and Physics, Charles University Prague, Sokolovská 83, 186 75 Praha 8, Czech Republic.
[2] Center for Statistics, Hasselt University, Agoralaan -building D, B-3590 Diepenbeek, Belgium;
[3] Department of Mathematics and Leuven Statistics Research Center (LStat), Katholieke Universiteit Leuven, Celestijnenlaan 200B, B-3001 Leuven (Heverlee), Belgium;

DOI: 10.1016/j.jspi.2012.06.001

April 24, 2012

ABSTRACT. This paper is concerned with inference about the dependence or association between two random variables conditionally upon the given value of a covariate. A way to describe such a conditional dependence is via a conditional copula function. Nonparametric estimators for a conditional copula then lead to nonparametric estimates of conditional association measures such as a conditional Kendall's tau. The limiting distributions of nonparametric conditional copula estimators are rather involved. In this paper we propose a bootstrap procedure for approximating these distributions and their characteristics, and establish its consistency. We apply the proposed bootstrap procedure for constructing confidence intervals for conditional association measures, such as a conditional Blomqvist beta and a conditional Kendall's tau. The performances of the proposed methods are investigated via a simulation study involving a variety of models, ranging from models in which the dependence (weak or strong) on the covariate is only through the copula and not through the marginals, to models in which this dependence appears in both the copula and the marginal distributions. As a conclusion we provide practical recommendations for constructing bootstrap-based confidence intervals for the discussed conditional association measures.

Keywords and phrases: Asymptotic representation, bootstrap, empirical copula process, fixed design, random design, smoothing, weak convergence.

1. Introduction

Let $(Y_{11}, Y_{21}, X_1)^{\mathsf{T}}, \dots, (Y_{1n}, Y_{2n}, X_n)^{\mathsf{T}}$ be independent identically distributed three-dimensional vectors from the cumulative distribution function $H(y_1, y_2, x)$ of the random triple $(Y_1, Y_2, X)^{\mathsf{T}}$, where X is a covariate. The contributions in this paper are valid for the case of random design (X is a random variable) and fixed regular design (X is not random), with the design density satisfying some assumptions (see Section 2.4). Researchers are often interested in the dependence structure of the bivariate outcome $(Y_1, Y_2)^{\mathsf{T}}$ given a value of the covariate $X = x$. In an example concerning life expectancies of males and females, Veraverbeke et al. (2011) analyzed how the relationship between the two life expectancies changes with the logarithm of the under-five mortality rate of a country, which is often used as a characteristic for its development status. Another example is provided in Gijbels et al. (2011), where in a study on soil contamination it was investigated how the association between the amount of zinc in the soil and the soils microbial activity is influenced by the quantity of organic material present in the soil. We refer to these two papers and references therein for further background information on the study of conditional dependencies.

Let us denote the joint and marginal distribution functions of $(Y_1, Y_2)^{\mathsf{T}}$, conditionally upon $X = x$, as

$$H_x(y_1, y_2) = P(Y_1 \le y_1, Y_2 \le y_2 \,|\, X = x),$$
$$F_{1x}(y_1) = P(Y_1 \le y_1 \,|\, X = x), \quad F_{2x}(y_2) = P(Y_2 \le y_2 \,|\, X = x).$$

If $F_{1x}(\cdot)$ and $F_{2x}(\cdot)$ are continuous functions, then according to Sklar's theorem (see e.g. Nelsen (2006)) there exists a unique copula C_x which equals

$$C_x(u_1, u_2) = H_x(F_{1x}^{-1}(u_1), F_{2x}^{-1}(u_2)),$$

where $F_{1x}^{-1}(u) = \inf\{y : F_{1x}(y) \ge u\}$ is the conditional quantile function of Y_1 given $X = x$ and F_{2x}^{-1} is the conditional quantile function of Y_2 given $X = x$. The conditional copula C_x fully describes the conditional dependence structure of $(Y_1, Y_2)^{\mathsf{T}}$ given $X = x$.

Based on the observed data we have the following empirical estimator for $H_x(y_1, y_2)$:

$$H_{xh}(y_1, y_2) = \sum_{i=1}^{n} w_{ni}(x, h_n)\, \mathbb{I}\{Y_{1i} \le y_1, Y_{2i} \le y_2\}, \tag{1}$$

where $\{w_{ni}(x, h_n)\}$ is a sequence of weights that smooth over the covariate space and $h = h_n > 0$ is a bandwidth going to zero as the sample size increases. Gijbels et al. (2011) suggested the following empirical estimator of the copula C_x

$$C_{xh}(u_1, u_2) = H_{xh}\left(F_{1xh}^{-1}(u_1), F_{2xh}^{-1}(u_2)\right) \qquad \left(0 \le u_1, u_2 \le 1\right), \tag{2}$$

where F_{1xh} and F_{2xh} are corresponding marginal distribution functions of H_{xh}. Although this estimator seems to be quite natural Gijbels et al. (2011) illustrated that this estimator may suffer from severe bias if any of the marginal conditional distributions F_{1x}, F_{2x} change with x. In order to reduce that bias they suggested an alternative estimator constructed in the following way. First, put

$$(\tilde{U}_{1i}, \tilde{U}_{2i})^{\mathsf{T}} = (F_{1x_i g_1}(Y_{1i}), F_{2x_i g_2}(Y_{2i}))^{\mathsf{T}}, \quad i = 1, \dots, n, \tag{3}$$

where $g_1 = \{g_{1n}\} \searrow 0$ and $g_2 = \{g_{2n}\} \searrow 0$. Second, use the transformed observations $(\tilde{U}_{1i}, \tilde{U}_{2i})^{\mathsf{T}}$ in a similar way as the original observations and construct

$$\tilde{C}_{xh}(u_1, u_2) = \tilde{G}_{xh}\left(\tilde{G}_{1xh}^{-1}(u_1), \tilde{G}_{2xh}^{-1}(u_2)\right), \tag{4}$$

where

$$\tilde{G}_{xh}(u_1, u_2) = \sum_{i=1}^{n} w_{ni}(x, h_n)\, \mathbb{I}\{\tilde{U}_{1i} \le u_1, \tilde{U}_{2i} \le u_2\},$$

and $\tilde{G}_{1xh}$ and $\tilde{G}_{2xh}$ are its corresponding marginals.

Veraverbeke et al. (2011) studied the asymptotic properties of the estimators C_{xh} and $\tilde{C}_{xh}$ and showed that provided

$$h_n = O(n^{-1/5}), \quad \sqrt{n\, h_n}\, g_{jn}^2 = O(1), \quad \tfrac{h_n}{g_{jn}} = O(1), \quad n \min(h_n, g_{1n}, g_{2n}) \to \infty, \tag{5}$$

and some further regularity conditions are met the copula processes

$$\mathbb{C}_n = \sqrt{n\, h_n}(C_{xh} - C_x), \qquad \tilde{\mathbb{C}}_n = \sqrt{n\, h_n}(\tilde{C}_{xh} - C_x) \tag{6}$$

converge in distribution in the space of bounded functions on the unit square $[0,1]^2$, equipped with the supremum norm, to Gaussian processes with the same covariance structure but possibly different drift functions that correspond to the asymptotic biases of the estimators C_{xh} and $\tilde{C}_{xh}$. As the asymptotic limiting distributions are rather complicated, bootstrap procedures for estimating the sampling distribution of both C_{xh} and $\tilde{C}_{xh}$ are of practical interest. Our proposal, which is described in the next section, is inspired by the resampling procedure suggested in Aerts et al. (1994). For the unconditional situation (without a covariate), bootstrap approximations of the empirical copula process have been discussed in Fermanian et al. (2004) and Bücher and Dette (2010).

The paper is organised as follows. In Section 2 we describe the suggested resampling procedures and state the main results on their consistency. Section 3 gives applications of the bootstrap procedures to construct confidence intervals for conditional Blomqvist's beta and Kendall's tau. The proofs are given in the Appendix.

2. Bootstrap

In the following we will denote the observed values of the covariate X as $x_1, \dots, x_n$.

2.1. **Bootstrapping the process $\mathbb{C}_n$.** Let $g_b = \{g_{bn}\}$ be a bandwidth sequence that is asymptotically larger than $h = \{h_n\}$, more precisely assume that

$$h_n = O(n^{-1/5}), \qquad g_{bn} \to 0, \qquad n^{1-\delta}\, g_{bn}^5 \to +\infty, \quad \text{where } \delta > 0. \tag{7}$$

Substitute h_n with g_b in formula (1) and get $H_{x_i g_b}(y_1, y_2)$ – the estimate of the conditional joint distribution of $(Y_1, Y_2)^\mathsf{T}$ given $X = x_i$. We suggest to resample the process $\mathbb{C}_n$ as follows.
Algorithm:

Draw $(Y_{1i}^*, Y_{2i}^*)^\mathsf{T}$ from $H_{x_i g_b}(\cdot,\cdot)$; do this for $i = 1, \dots, n$.

Let H_{xh}^* be the empirical distribution function of the bootstrap sample, C_{xh}^* be the corresponding empirical copula and C_{xg_b} be the estimate of copula given by (2) with the bandwidth g_b. Denote the bootstrap process as $\mathbb{C}_n^* = \sqrt{n h_n}(C_{xh}^* - C_{xg_b})$.

2.2. **Bootstrapping the process $\tilde{\mathbb{C}}_n$.** Although assumption (5) implies that if $h_n \sim C n^{-1/5}$, then g_{1n} and g_{2n} used in (3) have to be of order $n^{-1/5}$ as well, it turns out that for the reason of bootstrapping we have to take bandwidths $g_{1b} = \{g_{1bn}\}$, $g_{2b} = \{g_{2bn}\}$ that are asymptotically larger than $O(n^{-1/5})$. Let us replace g_{1n} and g_{2n} with g_{1b} and g_{2b} in (3) and denote $(\tilde{U}_{1i}^b, \tilde{U}_{2i}^b)$ the corresponding transformed 'uniform' alike observations.

Further, similarly as in the bootstrap algorithm for C_{xh}, we need to introduce a bandwidth $g_b = \{g_{bn}\}$. In the following we will suppose that there exists $\delta_1 > 0$ and $\delta_2 > 0$ such that

$$\begin{aligned} &h_n = O(n^{-1/5}), \quad \tfrac{h_n \log^{1+\delta_1} n}{g_{jbn}} \to 0, \quad \forall_{\eta>0} \left(n\, h_n\, g_{jbn}^{4+\eta}\right) \to 0, \qquad \text{for } j = 1, 2, \\ &n^{1/5-\delta_2}\, g_{bn} \to +\infty, \quad g_{bn} \to 0. \end{aligned} \tag{8}$$

Finally, let us define

$$\tilde{G}_{z\, g_b}(u_1, u_2) = \sum_{i=1}^n w_{ni}(z, g_{bn})\, \mathbb{I}\{\tilde{U}_{1i}^b \le u_1, \tilde{U}_{2i}^b \le u_2\} \tag{9}$$

and $\tilde{C}_{xg_b}(u_1, u_2) = \tilde{G}_{xg_b}\left(\tilde{G}_{1xg_b}^{-1}(u_1), \tilde{G}_{2xg_b}^{-1}(u_2)\right)$ with $\tilde{G}_{1xg_b}$ and $\tilde{G}_{2xg_b}$ being the marginals of $\tilde{G}_{xg_b}$.
Algorithm:

Draw $(\tilde{U}_{1i}^*, \tilde{U}_{2i}^*)^\mathsf{T}$ from $\tilde{G}_{x_i g_b}(\cdot,\cdot)$; do this for $i = 1, \dots, n$.

Let $\tilde{C}_{xh}^*$ be a copula estimator given by (2) with the original observations (Y_{1i}, Y_{2i}) replaced by $(\tilde{U}_{1i}^*, \tilde{U}_{2i}^*)$. The bootstrap process is given by $\tilde{\mathbb{C}}_n^* = \sqrt{n h_n}(\tilde{C}_{xh}^* - \tilde{C}_{xg_b})$.

Remark 1. Sampling from H_{zg_b} (or $\tilde{G}_{zg_b}$) requires the weights $w_{ni}(x, g_b)$ to be non-negative and to sum up to 1. If this is not the case then the most straightforward way is to put negative weights to zero and then rescale all the weights so that they sum up to 1. Thanks to assumptions (**W3**) and (**W7**) given in Section 2.4 such a modification is 'asymptotically negligible' and the theoretical justification of the resampling procedures can use the original weights.

2.3. **Theoretical results.** Regularity conditions needed in the following theorems are given in Sections 2.4 and 2.5.

Theorem 1. *Assume* (7), (**W1**)–(**W7**), (**W'1**)–(**W'3**), (**W"1**)–(**W"5**), **(H)** *and* **(R2)**. *Then the bootstrap process* $\mathbb{C}_n^*$ *converges in bootstrap measure* P^* [P]*-almost surely to the same Gaussian process as the empirical process* $\mathbb{C}_n$.

Theorem 2. *Assume* (5), (8), (**W1**)–(**W7**), (**W'1**)–(**W'3**), (**W"1**)–(**W"6**), *and* **(R1)**–**(R4)**. *Then the bootstrap process* $\tilde{\mathbb{C}}_n^*$ *converges in bootstrap measure* P^* [P]*-almost surely to the same Gaussian process as the empirical process* $\tilde{\mathbb{C}}_n$.

The proofs of both theorems are rather technical and are postponed to the Appendix. The proofs mimic the proof of the weak convergence of the process $\mathbb{C}_n$ given in Veraverbeke et al. (2011). The difference is that in our situation we need to assure that things which are happening in the bootstrap probability P^* hold [P]-almost surely.

For $i = 1, \dots, n$ and $j = 1, 2$ put $U_{ji} = F_{jx_i}(Y_{ji})$ and let $\mathcal{C}_n$ stand either for the copula process $\mathbb{C}_n$ or $\tilde{\mathbb{C}}_n$. Veraverbeke et al. (2011) showed the following Bahadur type representations (uniform in (u_1, u_2))

$$\mathcal{C}_n(u_1, u_2) = \sqrt{n\,h_n} \sum_{i=1}^{n} w_{ni}(x; h_n)\, \xi_i(u_1, u_2) + o_P(1), \tag{10}$$

where

$$\begin{aligned} \xi_i(u_1, u_2) &= \mathbb{I}\{Y_{1i} \le F_{1x}^{-1}(u_1), Y_{2i} \le F_{2x}^{-1}(u_2)\} - C_x(u_1, u_2) \\ &\quad - C_x^{(1)}(u_1, u_2) \left[\mathbb{I}\{Y_{1i} \le F_{1x}^{-1}(u_1)\} - u_1\right] - C_x^{(2)}(u_1, u_2) \left[\mathbb{I}\{Y_{2i} \le F_{2x}^{-1}(u_2)\} - u_2\right], \end{aligned} \tag{11}$$

for $\mathcal{C}_n = \mathbb{C}_n$ and

$$\begin{aligned} \xi_i(u_1, u_2) &= \mathbb{I}\{U_{1i} \le u_1, U_{2i} \le u_2\} - C_x(u_1, u_2) \\ &\quad - C_x^{(1)}(u_1, u_2) \left[\mathbb{I}\{U_{1i} \le u_1\} - u_1\right] - C_x^{(2)}(u_1, u_2) \left[\mathbb{I}\{U_{2i} \le u_2\} - u_2\right], \end{aligned} \tag{12}$$

for $\mathcal{C}_n = \tilde{C}_n$; with $C_x^{(i)}(u_1, u_2) = \frac{\partial C_x(u_1,u_2)}{\partial u_i}$ for $i = 1, 2$.

In this paper we show analogous Bahadur representations

$$\mathcal{C}_n^*(u_1, u_2) = \sqrt{n\,h_n} \sum_{i=1}^{n} w_{ni}(x; h_n)\, \xi_i^*(u_1, u_2) + o_{P^*}(1), \qquad \text{[P]-a.s.}, \tag{13}$$

where

$$\begin{aligned} \xi_i^*(u_1, u_2) &= \mathbb{I}\{Y_{1i}^* \le F_{1xg_b}^{-1}(u_1), Y_{2i}^* \le F_{2xg_b}^{-1}(u_2)\} - C_{xg_b}(u_1, u_2) \\ &\quad - C_x^{(1)}(u_1, u_2) \left[\mathbb{I}\{Y_{1i}^* \le F_{1xg_b}^{-1}(u_1)\} - u_1\right] - C_x^{(2)}(u_1, u_2) \left[\mathbb{I}\{Y_{2i}^* \le F_{2xg_b}^{-1}(u_2)\} - u_2\right], \end{aligned} \tag{14}$$

for $\mathcal{C}_n^* = \mathbb{C}_n^*$ and

$$\begin{aligned}(15)\quad \xi_i^*(u_1,u_2) = \mathbb{I}\{\tilde{U}_{1i}^* \leq u_1, \tilde{U}_{2i}^* \leq u_2\} - \tilde{C}_{xg_b}(u_1,u_2) \\ - C_x^{(1)}(u_1,u_2)\Big[\mathbb{I}\{\tilde{U}_{1i}^* \leq u_1\} - u_1\Big] - C_x^{(2)}(u_1,u_2)\Big[\mathbb{I}\{\tilde{U}_{2i}^* \leq u_2\} - u_2\Big],\end{aligned}$$

for $\mathcal{C}_n^* = \tilde{C}_n^*$.

The asymptotic tightness of the process $\mathbb{C}_n^*$ ($\tilde{\mathbb{C}}_n^*$) then follows by Step 1 of the proof in Appendix A (Appendix B). Further with the help of computations done in Step 2 of Appendix A (Appendix B) we can verify that the bias and covariance structure of process $\mathbb{C}_n^*$ ($\tilde{\mathbb{C}}_n^*$) asymptotically coincide with the bias and covariance structure of the process $\mathbb{C}_n$ ($\tilde{\mathbb{C}}_n$).

2.4. **Regularity conditions for weights.** Let $\{Z_n\}$ be a sequence of random variables and $\{a_n\}$ a sequence of constants. We will use the following notation.

$$Z_n = o_{a.s.}(a_n) \quad\Leftrightarrow\quad Z_n/a_n \to 0 \quad \text{a. s. as } n \to \infty,$$

$$Z_n = O_{a.s.}(a_n) \quad\Leftrightarrow\quad \exists C<\infty : P\Big[\sup_{m\geq n}|Z_m/a_m| \geq C\Big] \to 0 \quad \text{as } n\to\infty,$$

$$Z_n = O_e(a_n) \quad\Leftrightarrow\quad \exists C, C' < \infty\, \exists \eta > 0 : P\left[|Z_n/a_n| \geq C\right] \leq C\exp\left\{-\tfrac{n^\eta}{C'}\right\}.$$

Note that with the help of the Borel-Cantelli lemma $Z_n = O_e(a_n)$ implies $Z_n = O_{a.s.}(a_n)$. Further for a process $\{Z_n(z), z \in M_n\}$ (the set M_n is allowed to vary with n) we will write

$$Z_n(z) = O_e(a_n) \text{ on } M_n \quad\Leftrightarrow\quad \exists C, C' < \infty\, \exists \eta > 0 : \sup_{z\in M_n} P\left[|Z_n(z)/a_n| \geq C\right] \leq C\exp\left\{-\tfrac{n^\eta}{C'}\right\}.$$

Next by $w_{ni}'(z,h_n)$ and $w_{ni}''(z,h_n)$ we will understand the first and second derivative of the weight $w_{ni}(z,h_n)$ with respect to the variable z. Further put $I_x^{(n)} = \{i : w_{ni}(x, g_{bn}) \neq 0\}$ and $J_x^{(n)} = [\min_{i\in I_x^{(n)}} X_i, \max_{i\in I_x^{(n)}} X_i]$.

The following is a listing of assumptions on the system of weights $\{w_{ni};\, i = 1,\ldots,n\}$ in random design. The conditions for fixed design may be derived easily by replacing X_i by x_i and the symbols $o_{a.s.}, O_{a.s.}, O_e$ by o, O, O. Finally, a_n is a substitute for any of the sequences h_n, g_{1n}, g_{2n}, g_{bn}, g_{1bn}, g_{2bn}.

(W1) $\displaystyle \max_{1\leq i\leq n}|w_{ni}(x,a_n)| = o_{a.s.}\big(\tfrac{1}{\sqrt{n\,a_n}}\big)$, **(W2)** $\displaystyle \sum_{i=1}^n w_{ni}^2(z,a_n) = O_e\Big(\tfrac{1}{n\,a_n}\Big)$ on $J_x^{(n)}$,

(W3) $$\sup_{z\in J_x^{(n)}}\left|\sum_{i=1}^n w_{ni}(z,a_n) - 1\right| = o_{a.s.}\Big(\tfrac{1}{\sqrt{n\,h_n}}\Big)$$

(W4) $$\exists_{D_K<\infty}\,\forall_{a_n}\,\sup_{z\in J_x^{(n)}}\left|\sum_{i=1}^n w_{ni}(z,a_n)(X_i - z) - a_n^2\,D_K\right| = o_{a.s.}\left(a_n^2\right),$$

(W5) $$\exists_{E_K<\infty}\,\forall_{a_n}\,\sup_{z\in J_x^{(n)}}\left|\sum_{i=1}^n w_{ni}(z,a_n)(X_i - z)^2 - a_n^2\,E_K\right| = o_{a.s}\left(a_n^2\right),$$

(W6) $$\left(\max_{i \in I_x^{(n)}} X_i - \min_{i \in I_x^{(n)}} X_i \right) = o_{a.s.}(1),$$

(W7) $$\sup_{z \in J_x^{(n)}} \left| \sum_{i=1}^{n} w_{ni}(z, a_n)\, \mathbb{I}\{w_{ni}(z, a_n) < 0\} \right| = o_{a.s.}\big(\tfrac{1}{\sqrt{n\, h_n}}\big),$$

Assumptions on the first derivatives of weights $\{w'_{ni};\, i = 1, \dots, n\}$:

(W'1) $$\sup_{z \in J_x^{(n)}} \left| \sum_{i=1}^{n} w'_{ni}(z, g_{bn}) \right| = o_{a.s.}(1),$$ **(W'2)** $$\sum_{i=1}^{n} [w'_{ni}(z, g_{bn})]^2 = O_e\left(\tfrac{1}{n^{\delta}}\right) \text{ on } J_x^{(n)},$$

(W'3) $$\sup_{z \in J_x^{(n)}} \left| \sum_{i=1}^{n} w'_{ni}(z, g_{bn})(X_i - z) - 1 \right| = o_{a.s}(1)$$

Assumptions on the the second derivatives of weights $\{w''_{ni};\, i = 1, \dots, n\}$:

(W"1) $$\sup_{z \in J_x^{(n)}} \left| \sum_{i=1}^{n} w''_{ni}(z, g_{bn}) \right| = o_{a.s.}(1),$$ **(W"2)** $$\sum_{i=1}^{n} [w''_{ni}(z, g_{bn})]^2 = O_e\left(\tfrac{1}{n^{\delta}}\right) \text{ on } J_x^{(n)},$$

(W"3) $$\sup_{z \in J_x^{(n)}} \left| \sum_{i=1}^{n} w''_{ni}(z, g_{bn})(X_i - z) \right| = o_{a.s.}(1),$$

(W"4) $$\sup_{z \in J_x^{(n)}} \left| \sum_{i=1}^{n} w''_{ni}(z, g_{bn})(X_i - z)^2 - 1 \right| = o_{a.s}(1),$$

(W"5) $$\exists_{C, L < \infty}\ \exists_{\alpha > 0}\ \forall_{z_1, z_2 \in J_x^{(n)}} :\ \max_i \left| w''_{ni}(z_1, g_{bn}) - w''_{ni}(z_2, g_{bn}) \right| \le C\, g_{bn}^{-L}\, |z_1 - z_2|^{\alpha}.$$

(W"6) $$\sup_{z \in J_x^{(n)}} \sum_{i=1}^{n} |w''_{ni}(z, g_{bn})| = O_{a.s}\left(\tfrac{1}{g_{bn}^2}\right),$$

Discussion of the conditions. Although the long list of conditions is rather discouraging, it is quite informative, because it tells us what properties of the weights are used during our proof. We believe that in spite of its length this list is preferable over being stuck to a particular type of weights. And as we will see later, once the system of weights is chosen, we often end up with a few standard conditions about the distribution of the covariate X.

Assume for concreteness that a kernel density function K has a support on $[-1, 1]$, it is symmetric and three times continuously differentiable. Further suppose that $h_n = O(n^{-1/5})$ and $g_{bn} = O(n^{-1/7})$. First let us consider *Nadaraya-Watson weights* [NW] (see Nadaraya (1964) or Watson (1964)), which are defined as

$$w_{ni}(x, h_n) = \frac{K(\frac{X_i - x}{h_n})}{\sum_{j=1}^{n} K(\frac{X_j - x}{h_n})}, \qquad i = 1, \dots, n.$$

While assumptions (**W3**), (**W6**) and (**W"5**) hold immediately and assumption (**W1**) can be verified with the help of Bernstein exponential inequality together with the Borel-Cantelli lemma, the remaining assumptions require a uniform (in z) law of large numbers. After some calculations we find out that the key to the success of the verification of the conditions in the last group is in showing that for $j = 0, 1, 2$, $l = 0, 1, 2$ there exists a neighbourhood U_x of x such that

$$\sup_{z \in U_x} \left| \frac{1}{n\,h_n} \sum_{i=1}^{n} K^{(l)}\left(\tfrac{X_i - z}{h_n}\right)\left(\tfrac{X_i - z}{h_n}\right)^j - \mathrm{E}\,\tfrac{1}{h_n} K^{(l)}\left(\tfrac{X_1 - z}{h_n}\right)\left(\tfrac{X_1 - z}{h_n}\right)^j \right| = o_{a.s.}(1). \tag{17}$$

A very handy tool to verify (17) is e.g. Lemma 22 of Nolan and Pollard (1987). That lemma tells us that the family of functions

$$\mathcal{F} = \{x \mapsto K^{(l)}\left(\tfrac{x-z}{h}\right)\left(\tfrac{x-z}{h}\right)^j, z \in \mathbb{R}, h > 0\}$$

is a Vapnik-Chervonenkis class of functions provided the function $u \mapsto K^l(u)\,u^j$ is of bounded variation on $\mathbb{R}_+$. Then we can start the empirical process machinery (see e.g. van der Vaart and Wellner (1996), van de Geer (2000) or the proof of Lemma 3 of this paper) to arrive at an exponential inequality and the Borel-Cantelli lemma finishes the proof of (17).

Some further calculations show that for NW weights we end up with the following conditions on the distribution function F_X of the covariate X:

(F1) $f_X = F_X'$ is continuous and positive at point x,
(F2) $f_X' = F_X''$ is continuous in a neighbourhood of x.

Another system of weights, very commonly employed, is *local linear* [LL] system of weights (see e.g p. 20 of Fan and Gijbels (1996)), which is given by

$$w_{ni}(x, h_n) = \frac{\frac{1}{n\,h_n} K(\frac{X_i - x}{h_n})\left(S_{n,2} - \frac{X_i - x}{h_n} S_{n,1}\right)}{S_{n,0}\,S_{n,2} - S_{n,1}^2}, \qquad i = 1, \ldots, n, \tag{18}$$

where

$$S_{n,j} = \frac{1}{n\,h_n} \sum_{i=1}^{n} \left(\tfrac{X_i - x}{h_n}\right)^j K\left(\tfrac{X_i - x}{h_n}\right), \qquad j = 0, 1, 2.$$

The verification of the conditions (**W1**)–(**W"6**) is completely analogous to NW weights and the most difficult part is to verify (17). The only difference is that j in (17) may now take values from zero to four. The nice thing about LL weights is that thanks to $\sum_{i=1}^{n} w_{ni}(x, h_n)(X_i - x) = 0$ it is sufficient to assume only **(F1)**.

In a fixed regular design case (see e.g. Müller (1987)), there exists an absolutely continuous distribution function F_X (with associated density f_X) such that $x_i = F_X^{-1}\left(\frac{i}{n+1}\right)$. In this case the design points are ordered, that is $x_1 \leq x_2 \ldots \leq x_n$. In this setting *Gasser-Müller* [GM] weights (see Gasser and Müller (1979)) are quite popular. Consider fixed, but arbitrary values x_0 and x_{n+1} such that $x_0 < x_1$ and $x_{n+1} > x_n$. Then GM weights are defined as weights

$$w_{ni}(x, h_n) = \tfrac{1}{h_n} \int_{s_i}^{s_{i+1}} K(\tfrac{z-x}{h_n})\, dz, \quad \text{where } s_i = (x_i + x_{i-1})/2, \quad i = 1, \ldots, n. \tag{19}$$

To verify (**W1**)–(**W"6**) it is sufficient to assume **(F1)**, and that f'_X/f^3_X is continuous in a neighbourhood of x.

2.5. **Regularity conditions for the conditional distribution.** Assumptions about the joint conditional distribution of (Y_1, Y_2) given $X = z$, its marginals and the corresponding conditional copula.

- **(H)** $\dot{H}_z(u_1, u_2) = \frac{\partial}{\partial z} H_z(F^{-1}_{1x}(u_1), F^{-1}_{2x}(u_2))$, $\ddot{H}_z(u_1, u_2) = \frac{\partial^2}{\partial z^2} H_z(F^{-1}_{1x}(u_1), F^{-1}_{2x}(u_2))$ exist and are continuous as functions of (z, u_1, u_2), where z takes value in a neighbourhood of x;
- **(R1)** $\dot{C}_z(u_1, u_2) = \frac{\partial}{\partial z} C_z(u_1, u_2)$, $\ddot{C}_z(u_1, u_2) = \frac{\partial^2}{\partial z^2} C_z(u_1, u_2)$ exist and are continuous as functions of (z, u_1, u_2), where z takes value in a neighbourhood of x;
- **(R2)** For $j = 1, 2$, the j-th first-order partial derivative of C_z exists and is continuous on the set $U(x) \times \{(u_1, u_2) \in [0, 1]^2 : 0 < u_j < 1\}$, where $U(x)$ is a neighbourhood of the point x;
- **(R3)** For $j = 1, 2$: $F_{jz}(F^{-1}_{jz}(u))$, $\dot{F}_{jz}(F^{-1}_{jz}(u))$, $\ddot{F}_{jz}(F^{-1}_{jz}(u))$ are continuous as functions of (z, u) for z in a neighbourhood of x, where $\dot{F}_{jz}(y) = \frac{\partial}{\partial z} F_{jz}(y)$, $\ddot{F}_{jz}(y) = \frac{\partial^2}{\partial z^2} F_{jz}(y)$;
- **(R4)** For each $\varepsilon > 0$ there exists $C > 0$, $\eta > 0$ and a neighbourhood $U(x)$ of the point x such that

$$\max_{j=1,2} \sup_{z_1, z_2 \in U(x)} \sup_{u_1, u_2 \in [\varepsilon, 1-\varepsilon]} \left| \ddot{F}_{jz_1}(F^{-1}_{jz_2}(u_1)) - \ddot{F}_{jz_1}(F^{-1}_{jz_2}(u_2)) \right| \le C\, |u_1 - u_2|^{\eta}.$$

Note that the exclusion of the boundary points in Assumption **(R2)** is important, since this prevents commonly-used copulas (such as e.g. Clayton, Gumbel, normal and Student-t copulas) from being excluded. This was pointed out in the context of the (smooth) empirical copula process in Omelka et al. (2009), where it was shown that it is sufficient to assume the continuity of the partial derivatives on $[0, 1]^2 \setminus \{(0, 0), (0, 1), (1, 0), (1, 1)\}$. Recently, Segers (2012) further weakened this assumption to the continuity of the j-th first-order partial derivative of the copula on the set $\{(u_1, u_2) \in [0, 1]^2 : 0 < u_j < 1\}$. This weakening of the assumptions is in particular useful when considering copulas of more than two variables.

3. Applications

In the following we will be interested in a confidence interval for a quantity θ that can be expressed as a functional of a conditional copula C_x, thus formally we can write $\theta_x = T(C_x)$. Similarly put $\theta_{xh} = T(C_{xh})$, $\tilde{\theta}_{xh} = T(\tilde{C}_{xh})$, $\theta_{xg} = T(C_{xg})$, ... Further denote $\theta^{*(1)}_{xh}, \ldots, \theta^{*(B)}_{xh}$

or $\tilde{\theta}^{*(1)}_{xh}, \ldots, \tilde{\theta}^{*(B)}_{xh}$ the realizations of the bootstrap algorithm either of Section 2.1 or Section 2.2.

Basically, there are three type of confidence intervals we can construct.

`as`: Neglecting the possible bias we use the asymptotic representation given by (10) and either (11) or (12) to estimate the variance of the estimator $\hat{\theta}_h$ by $\hat{\sigma}^2$ and form a confidence interval based on a normal approximation as

$$[\theta_{xh} - z_{1-\alpha/2}\,\hat{\sigma},\, \theta_{xh} + z_{1-\alpha/2}\,\hat{\sigma}], \tag{20}$$

where z_α is an α-quantile of the standard normal distribution. Note that from (10) the asymptotic variance of the empirical copula process can be approximated by $nh_n \sum_{i=1}^n w_{ni}^2(x; h_n)\mathrm{Var}(\xi_i(u_1, u_2))$, where the variance-covariance structure for the latter quantities has been established in Veraverbeke et al. (2011) (see Corollary 1 therein).

`ab`: We estimate the variance from the bootstrap as

$$\hat{\sigma}_b^2 = \frac{1}{B-1}\sum_{b=1}^{B}(\theta^{*(b)}_{xh} - \bar{\theta}^*_{xh})^2, \qquad \text{where } \bar{\theta}^*_{xh} = \frac{1}{B}\sum_{b=1}^{B}\theta^{*(b)}_{xh}$$

and use the confidence interval given in (20). (`as-boot` or abbreviated `ab`)

`bo`: We use the bootstrap algorithm of Section 2.1 or 2.2 and construct the basic bootstrap confidence interval. (`boot` or abbreviated `bo`)

Our preliminary simulation experience has shown that estimating the bias of $\tilde{C}_{xh}$, i.e. the bias coming only from the influence of a covariate on the underlying copula, is rather difficult even in moderately large samples. In small sample settings it is usually much safer to neglect a possible bias, which means to estimate the bias by zero. That is why we include also the following 'sample-size adaptive' procedure.

`hybrid`: Note that $b(h) = \bar{\theta}^*_{xh} - \theta_{xg}$ is the bootstrap estimation of the bias. Use the bandwidth $h_2 = h/2$ and calculate also $b(h_2) = \bar{\theta}^*_{xh_2} - \theta_{xg}$. Based on the asymptotic considerations one would expect that

$$0 < b(h_2) < b(h) \qquad \text{or} \qquad 0 > b(h_2) > b(h). \tag{21}$$

But if (21) does not hold then asymptotic estimation of the bias has not probably kicked in yet. Thus we also include a hybrid procedure, which uses method `boot` if (21) holds and `as-boot` in the other case. (`hybrid` or abbreviated `hy`).

For practical implementation we have to specify several bandwidths. For the choice of h_n we used the plug-in bandwidth selection rule proposed in Gijbels et al. (2011). For determining the bandwidths g_{1n} and g_{2n} used to 'uniformize' the marginal distributions we used 'lokern' which is a library available for the R computing environment (see R Development Core Team (2009)) and which implements the ideas of bandwidth choice in nonparametric regression as

Model	mean functions	parameter ρ
1 / 2	$\mu_1(z) = 1 \qquad \mu_2(z) = 1$	1 / 5
3 / 4	$\mu_1(z) = 1 \qquad \mu_2(z) = \sin(z - x_0)$	1 / 5
5 / 6	$\mu_1(z) = \sin(z - x_0) \qquad \mu_2(z) = \sin(z - x_0)$	1 / 5
7 / 8	$\mu_1(z) = \cos(z - x_0) \qquad \mu_2(z) = \sin(z - x_0)$	1 / 5

TABLE 1. Simulation models.

introduced in Gasser et al. (1991) and Brockmann et al. (1993). Once, the bandwidths h_n, g_{1n} and g_{2n} are specified, the 'resampling' bandwidths g_{bn}, g_{1bn} and g_{2bn} are given by the equations

$$g_{bn} = 1.5\, h_n\, n^{1/10}, \qquad g_{1bn} = 0.25\, g_{1n}\, (\log n)^{1.01}, \qquad g_{2bn} = 0.25\, g_{2n}\, (\log n)^{1.01}. \tag{22}$$

While, the choice of g_{bn} was inspired by Section 3 of Härdle and Bowman (1988), the bandwidths g_{1bn} and g_{2bn} were simply chosen to satisfy (8), but to stay 'close' to g_{1n} and g_{2n}.

Note that the methods `as` and `as-boot` are asymptotically valid only when the asymptotic bias diminishes, that is generally when $h_n = o(n^{-1/5})$. Thus we will report the results not only for the value h_n given by the plug-in rule (say h_p), but also for $h_n = 0.5\, h_p$.

3.1. **Blomqvist's beta.** Let Y_1 and Y_2 be random variables with distribution functions F_1 and F_2 respectively. Blomqvist (1950) proposed and studied the following measure of concordance

$$\beta = \mathrm{P}\left[\left(Y_1 - F_1^{-1}(0.5)\right)\left(Y_2 - F_2^{-1}(0.5) > 0\right)\right] - \mathrm{P}\left[\left(Y_1 - F_1^{-1}(0.5)\right)\left(Y_2 - F_2^{-1}(0.5) < 0\right)\right],$$

which is often also called *the medial correlation coefficient.* Let C be the copula corresponding to Y_1 and Y_2. Then β can be expressed simply as $\beta = 4\, C(0.5, 0.5) - 1$ (see pp.182–183 of Nelsen (2006)). In the presence of a covariate we can consider Blomqvist's beta conditionally on $X = x$ and define it as $\beta_x = 4\, C_x(0.5, 0.5) - 1$. The considered estimators are thus

$$\hat{\theta}_h = \beta_{xh} = 4\, C_{xh}(0.5, 0.5) - 1 \quad \text{and} \quad \tilde{\theta}_h = \tilde{\beta}_{xh} = 4\, \tilde{C}_{xh}(0.5, 0.5) - 1.$$

To illustrate our main findings we report results for the following setup: the covariate is supposed to be standard normal and we are interested in the point $X = x_0 = 1$. The copula which joins the margins is a Frank copula with the parameter depending on the value of the covariate $X = z$ as $\theta(z) = 5 + \rho\, \sin(\frac{(z - x_0)\, \pi}{6})$. This results in Blomqvist's beta equal to 0.51 for $z = x_0$. The margins were taken normal with unit variances and mean functions $\mu_1(z)$ and $\mu_2(z)$. The eight models considered are given in Table 1.

Models 1 and 2 represent situations where the covariate does not influence the marginal distributions; in Models 3 and 4 only one of the marginals is affected; while in Models 5 and 6 both marginals are stochastically increasing with z; finally in Models 7 and 8 the marginals are affected in different directions. The two values of ρ represent the situations when there is a mild ($\rho = 1$) or strong effect ($\rho = 5$) of the covariate on the dependence structure.

Further, the nominal level of the confidence interval is 0.9, the considered sample sizes are $n = 200, 500, 1000$, the number of bootstrap samples is $B = 999$ and the number of generated samples is 1 000. Finally, we use LL weights introduced in Section 2.4 together with the triweight kernel $K(x) = \frac{35}{32}\,(1 - x^2)^3\,\mathbb{I}\{|x| \leq 1\}$.

The results on coverage and average lengths of the confidence intervals of the procedures `as`, `as-boot`, `boot` and `hybrid` that use either of the estimators C_{xh} and $\tilde{C}_{xh}$ are to be found in Tables 2 and 3. For the sake of brevity we omit the results for Models 5 and 6 as those are close to the results of Models 7 and 8. The main findings may be summarized as follows:

- The result for C_{xh} and $\tilde{C}_{xh}$ are completely comparable for Models 1 and 2 in which the marginal distributions are not influenced by the covariate.
- For Models 3 and 4 the confidence intervals based on C_{xh} do not work when the plug-in bandwidth h_p is used. The reason is that the bias of the estimator θ_{xh} is too high. But if we use the 'bias diminishing tactic' ($h_p/2$) we still get at least reasonable coverage at the price of an increased length of the confidence interval.
- The asymptotic method (`as`) suffers from undercoverage for the sample size $n = 200$.
- In all considered situations `as-boot` together with the estimator $\tilde{C}_{xh}$ work very well. There seems to be two reasons for that. First, the bootstrap slightly overestimates the true variance, and hence leads to conservative coverage probabilities. Second, the estimation of the bias of the estimator $\tilde{C}_{xh}$ is such a subtle problem that unless the sample size is very large ($n \geq 1000$), it is better to ignore the possible bias, that is to estimate it as zero.
- The simulation results not presented here suggest that we should start to worry about the bias of $\tilde{C}_{xh}$ for large sample sizes (say $n \geq 1000$) and where there is a danger of oversmoothing. Then we should switch either to `boot` or to the suggested `hybrid` which seems to work reasonably in all situations we encountered so far.

To get an insight why for small sizes it is better to simply ignore the bias let us consider the estimator β_{xh} and Model 2. Note that in this model, the covariate influences only the dependence structure. Figure 1 plots the difference $\beta_{xh} - \beta_x$ together with the bootstrap estimate of the bias, that is the Monte Carlo estimate of the quantity

$$\mathrm{E}^*\,\beta^*_{xh} - \beta_{xg}, \tag{23}$$

for sample sizes 500 and 1 000. The dotted lines (vertical as well as horizontal) represent the true finite sample bias of the estimator β_{xh}. The dashed line represents the bootstrap estimate of bias (23) averaged over 1 000 samples. We see that the 'averaged' bootstrap estimate of bias is in a very close agreement with the true bias. But we also see that in particular for the

			Model 1						Model 2					
n			200		500		1000		200		500		1000	
bandwidth			$h_p/2$	h_p	$h_p/2$	h_p	$h_p/2$	h_p	$h_p/2$	h_p	$h_p/2$	h_p	$h_p/2$	h_p
COVERAGE	C_{xh}	bo	0.904	0.889	0.895	0.894	0.902	0.893	0.915	0.903	0.891	0.895	0.917	0.889
		hy	0.925	0.920	0.917	0.930	0.926	0.918	0.922	0.923	0.923	0.926	0.929	0.898
		ab	0.925	0.931	0.918	0.933	0.930	0.923	0.923	0.928	0.923	0.925	0.927	0.909
		as	0.867	0.883	0.898	0.914	0.913	0.913	0.874	0.894	0.897	0.907	0.917	0.907
	$\tilde{C}_{xh}$	bo	0.881	0.890	0.897	0.884	0.904	0.897	0.899	0.885	0.892	0.885	0.910	0.893
		hy	0.904	0.912	0.916	0.899	0.922	0.916	0.910	0.904	0.906	0.893	0.918	0.904
		ab	0.911	0.925	0.919	0.907	0.927	0.924	0.916	0.917	0.912	0.896	0.919	0.896
		as	0.843	0.876	0.893	0.885	0.907	0.915	0.870	0.878	0.899	0.885	0.910	0.900
LENGTH	C_{xh}	bo	0.556	0.420	0.296	0.232	0.192	0.156	0.556	0.420	0.296	0.236	0.196	0.156
		hy	0.556	0.420	0.296	0.232	0.192	0.156	0.560	0.420	0.296	0.236	0.196	0.156
		ab	0.556	0.420	0.296	0.232	0.192	0.156	0.560	0.420	0.296	0.236	0.196	0.156
		as	0.468	0.360	0.268	0.216	0.184	0.148	0.476	0.376	0.276	0.228	0.188	0.156
	$\tilde{C}_{xh}$	bo	0.548	0.416	0.292	0.232	0.192	0.156	0.548	0.416	0.296	0.236	0.196	0.160
		hy	0.548	0.412	0.292	0.232	0.192	0.156	0.548	0.416	0.296	0.236	0.196	0.156
		ab	0.548	0.412	0.292	0.232	0.192	0.156	0.548	0.416	0.296	0.236	0.196	0.156
		as	0.468	0.360	0.268	0.216	0.184	0.148	0.476	0.376	0.276	0.228	0.188	0.160

			Model 3						Model 4					
n			200		500		1000		200		500		1000	
bandwidth			$h_p/2$	h_p	$h_p/2$	h_p	$h_p/2$	h_p	$h_p/2$	h_p	$h_p/2$	h_p	$h_p/2$	h_p
COVERAGE	C_{xh}	bo	0.904	0.885	0.900	0.808	0.885	0.703	0.907	0.872	0.899	0.808	0.878	0.697
		hy	0.910	0.864	0.884	0.763	0.864	0.664	0.896	0.846	0.882	0.758	0.864	0.671
		ab	0.912	0.864	0.886	0.758	0.863	0.609	0.898	0.846	0.880	0.710	0.860	0.558
		as	0.831	0.806	0.862	0.731	0.839	0.587	0.836	0.799	0.850	0.690	0.851	0.550
	$\tilde{C}_{xh}$	bo	0.894	0.872	0.891	0.877	0.916	0.896	0.889	0.887	0.898	0.883	0.904	0.905
		hy	0.907	0.903	0.902	0.891	0.929	0.913	0.915	0.909	0.908	0.892	0.916	0.902
		ab	0.913	0.921	0.907	0.905	0.934	0.923	0.919	0.925	0.911	0.895	0.919	0.895
		as	0.855	0.869	0.888	0.882	0.922	0.910	0.856	0.879	0.899	0.893	0.911	0.897
LENGTH	C_{xh}	bo	0.568	0.432	0.308	0.244	0.208	0.164	0.572	0.436	0.312	0.248	0.212	0.168
		hy	0.572	0.432	0.308	0.244	0.208	0.164	0.572	0.436	0.312	0.248	0.212	0.168
		ab	0.572	0.432	0.308	0.244	0.208	0.164	0.572	0.436	0.312	0.248	0.212	0.168
		as	0.480	0.380	0.284	0.232	0.196	0.160	0.488	0.392	0.288	0.240	0.200	0.168
	$\tilde{C}_{xh}$	bo	0.548	0.416	0.296	0.232	0.196	0.156	0.552	0.420	0.296	0.236	0.196	0.160
		hy	0.548	0.416	0.292	0.232	0.192	0.156	0.552	0.420	0.296	0.236	0.196	0.160
		ab	0.548	0.416	0.292	0.232	0.192	0.156	0.552	0.420	0.296	0.236	0.196	0.160
		as	0.468	0.360	0.268	0.216	0.184	0.148	0.476	0.376	0.276	0.228	0.188	0.160

TABLE 2. Coverage and average lengths of confidence intervals for the quantity β_x for Models 1, 2, 3 and 4. The nominal level is 0.90.

			Model 7						Model 8					
n			200		500		1000		200		500		1000	
bandwidth			$h_p/2$	h_p	$h_p/2$	h_p	$h_p/2$	h_p	$h_p/2$	h_p	$h_p/2$	h_p	$h_p/2$	h_p
COVERAGE	C_{xh}	bo	0.901	0.847	0.893	0.787	0.889	0.603	0.904	0.857	0.887	0.755	0.888	0.595
		hy	0.901	0.839	0.895	0.738	0.875	0.572	0.897	0.827	0.885	0.697	0.853	0.550
		ab	0.905	0.846	0.894	0.748	0.879	0.577	0.898	0.834	0.885	0.698	0.855	0.527
		as	0.843	0.777	0.857	0.673	0.844	0.503	0.843	0.775	0.859	0.646	0.830	0.480
	$\tilde{C}_{xh}$	bo	0.889	0.886	0.879	0.885	0.910	0.886	0.896	0.896	0.886	0.892	0.896	0.899
		hy	0.916	0.911	0.910	0.906	0.913	0.898	0.909	0.910	0.900	0.902	0.909	0.901
		ab	0.919	0.925	0.914	0.917	0.919	0.916	0.919	0.922	0.907	0.901	0.912	0.886
		as	0.861	0.879	0.881	0.894	0.903	0.898	0.859	0.882	0.887	0.890	0.905	0.884
LENGTH	C_{xh}	bo	0.568	0.432	0.308	0.244	0.212	0.168	0.572	0.436	0.312	0.248	0.216	0.172
		hy	0.568	0.432	0.308	0.244	0.208	0.168	0.572	0.436	0.312	0.248	0.216	0.172
		ab	0.568	0.432	0.308	0.244	0.208	0.168	0.572	0.436	0.312	0.248	0.216	0.172
		as	0.476	0.364	0.276	0.216	0.196	0.152	0.480	0.376	0.284	0.228	0.204	0.160
	$\tilde{C}_{xh}$	bo	0.548	0.416	0.296	0.232	0.192	0.156	0.552	0.420	0.296	0.236	0.196	0.156
		hy	0.548	0.416	0.296	0.232	0.192	0.156	0.552	0.420	0.296	0.236	0.196	0.156
		ab	0.548	0.416	0.296	0.232	0.192	0.156	0.552	0.420	0.296	0.236	0.196	0.156
		as	0.468	0.360	0.268	0.216	0.184	0.148	0.476	0.380	0.276	0.228	0.188	0.160

TABLE 3. Coverage and average lengths of confidence intervals for the quantity β_x for Models 7 and 8. The nominal level is 0.90.

sample size 500 the variance of the bootstrap estimate of bias is still rather high in comparison to the true bias. Further, the solid line given by the `lowess` smoother reveals that there seems to be a negative correlation between the quantities $\beta_{xh} - \beta_x$ and (23), which also negatively affects the centering (and thus coverage properties) of the basic bootstrap confidence intervals `boot` in not large samples. The same holds true also for the estimator $\tilde{\beta}_{xh}$ and bootstrap algorithm (2.2). Dealing with bias is even more difficult here because of the transformation of the marginals (3). Although according to the theoretical results of Veraverbeke et al. (2011) this transformation makes for instance the estimators β_{xh} of Model 2 and β_{xh} of Models 2, 4, 6 and 8 asymptotically equivalent in terms of asymptotic bias and variance (provided the same bandwidth h_n is used), the effect of the transformation on finite sample properties is not negligible unless the sample size is not very large.

3.2. **Kendall's tau.** While Blomqvist's beta is a simple functional of an underlying copula, Kendall's tau is a more complex one and its conditional version is given by

$$\tau_x = 4 \iint C_x \, dC_x - 1.$$

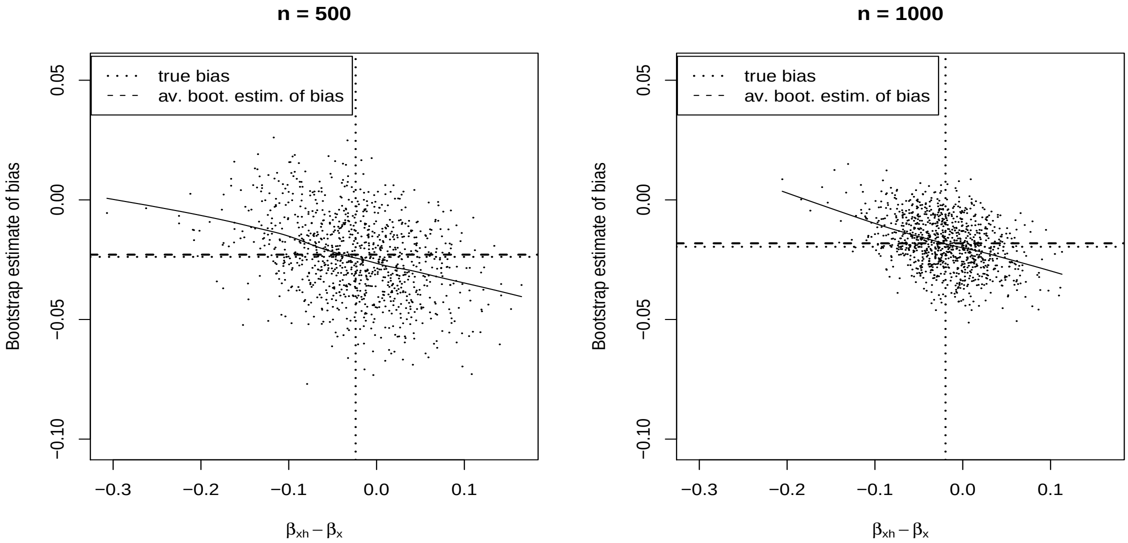

Figure 1. Scatterplot of $\beta_{xh} - \beta_x$ against the bootstrap estimator of the bias (23)

Gijbels et al. (2011) suggested the following estimator of τ_x

$$\tau_{xh} = \frac{4}{1 - \sum_{i=1}^{n} w_{ni}^2(x, h_n)} \sum_{i=1}^{n} \sum_{j=1}^{n} w_{ni}(x, h_n)\, w_{nj}(x, h_n)\, \mathbb{I}\{Y_{1i} < Y_{1j}, Y_{2i} < Y_{2j}\} - 1. \tag{24}$$

Further, Veraverbeke et al. (2011) showed that replacing the original observations (Y_{1i}, Y_{2i}) with the transformed 'uniform' alike observations $(\tilde{U}_{1i}, \tilde{U}_{2i})$ results in an estimator $\tilde{\tau}_{xh}$ that usually has better bias properties than the estimator τ_{xh}. Here we complement Gijbels et al. (2011), who compared the estimators τ_{xh} and $\tilde{\tau}_{xh}$ in terms of bias, variance and mean squared error, with results on the coverage and length of the corresponding confidence intervals.

The simulation design is the same as in Section 3.1, and we use the bandwidth h_p. As the general findings are similar as for Blomqvist's beta, in Table 4 we give only the results for Models 2, 4, 6 and 8, in which the covariate has a strong influence on the conditional copula.

4. Appendix – Proof of Theorem 1

As the proof is rather technical, we will present it only for fixed design. This may be justified in a similar way as in Veraverbeke et al. (2011). The technique of how to switch from random design to fixed design by conditioning on the values of the covariate is also presented in the proof of Lemma 5.

Similarly as in Veraverbeke et al. (2011) the expectations $\mathrm{E}^* f(Y_{i1}^*, Y_{i2}^*, F_{1xh}^*, F_{2xh}^*)$ with respect to bootstrap measure P^* have to be understood in a way that the functions F_{1xh}^*, F_{2xh}^* are fixed (nonrandom) and the expectation is computed only with respect to Y_{i1}^* and Y_{i2}^*.

			Model 2			Model 4			Model 6			Model 8		
bandwidth			200	500	1000	200	500	1000	200	500	1000	200	500	1000
COVERAGE	C_{xh}	bo	0.905	0.914	0.911	0.881	0.802	0.666	0.820	0.700	0.549	0.859	0.837	0.734
		hy	0.904	0.914	0.911	0.880	0.802	0.665	0.827	0.701	0.551	0.855	0.824	0.726
		ab	0.919	0.917	0.893	0.848	0.715	0.554	0.864	0.782	0.648	0.857	0.801	0.723
		as	0.906	0.902	0.876	0.830	0.695	0.513	0.866	0.763	0.642	0.839	0.786	0.697
	$\tilde{C}_{xh}$	bo	0.913	0.911	0.910	0.905	0.907	0.918	0.901	0.913	0.915	0.899	0.934	0.932
		hy	0.911	0.911	0.909	0.901	0.907	0.918	0.898	0.912	0.914	0.896	0.934	0.932
		ab	0.920	0.907	0.886	0.912	0.903	0.887	0.910	0.895	0.887	0.911	0.917	0.902
		as	0.905	0.889	0.872	0.901	0.885	0.874	0.899	0.883	0.875	0.908	0.913	0.887
LENGTH	C_{xh}	bo	0.230	0.136	0.094	0.246	0.148	0.104	0.214	0.127	0.089	0.259	0.158	0.114
		hy	0.230	0.136	0.094	0.246	0.148	0.104	0.214	0.127	0.089	0.259	0.158	0.114
		ab	0.230	0.136	0.094	0.245	0.147	0.103	0.214	0.127	0.089	0.259	0.158	0.113
		as	0.221	0.129	0.089	0.232	0.139	0.097	0.213	0.125	0.087	0.243	0.148	0.106
	$\tilde{C}_{xh}$	bo	0.232	0.137	0.095	0.233	0.137	0.095	0.232	0.138	0.095	0.233	0.137	0.095
		hy	0.232	0.137	0.095	0.233	0.137	0.095	0.232	0.138	0.095	0.233	0.137	0.095
		ab	0.232	0.137	0.095	0.232	0.137	0.095	0.232	0.137	0.095	0.233	0.137	0.095
		as	0.224	0.131	0.090	0.227	0.132	0.090	0.230	0.133	0.091	0.229	0.132	0.091

TABLE 4. Coverage and average lengths of confidence intervals for conditional Kendall's tau for Models 2, 4, 6 and 8. The nominal level is 0.90.

Formally

$$\mathrm{E}^* f(Y_{i1}^*, Y_{i2}^*, F_{1xh}^*, F_{2xh}^*) = \iint f(y_1, y_2, F_{1xh}^*, F_{2xh}^*)\, dH_{x_i g_b}(y_1, y_2),$$

whenever the integral on the right-hand side exists. A similar notation will be used for the expectation $\mathrm{E}\, f(Y_{i1}, Y_{i2}, \ldots)$ with respect to the original observed variables. Sometimes it will be even convenient to use the following convention

$$\mathrm{E}\, f(Y_{i1}, Y_{i2}, F_{1xg_b}, F_{2xg_b}, F_{1xh}^*, F_{2xh}^*) = \iint f(y_1, y_2, F_{1xg_b}, F_{2xg_b}, F_{1xh}^*, F_{2xh}^*)\, dH_{x_i}(y_1, y_2).$$

The reason for this notation is to simplify the presentation of the proof. This notation as well as the following decomposition apply the ideas of van der Vaart and Wellner (2007).

Similarly as in Veraverbeke et al. (2011) we can decompose

$$\sqrt{n\, h_n}\, (C_{xh}^* - C_{xg_b}) = A_n + B_n + C_n, \tag{A1}$$

where $A_n = D_n - \mathrm{E}^*\, D_n$, with

$$D_n(u_1, u_2) = \sqrt{n\, h_n} \sum_{i=1}^{n} w_{ni}(x; h_n) \Big[\mathbb{I}\{Y_{1i}^* \le F_{1xh}^{*-1}(u_1), Y_{2i}^* \le F_{2xh}^{*-1}(u_2)\} - \mathbb{I}\{Y_{1i}^* \le F_{1xg_b}^{-1}(u_1), Y_{2i}^* \le F_{2xg_b}^{-1}(u_2)\} \Big]. \tag{A2}$$

and

(A3)
$$B_n(u_1,u_2) = \sqrt{n\,h_n}\,\left[\sum_{i=1}^{n} w_{ni}(x;h_n)\,\mathbb{I}\{Y_{1i}^* \le F_{1xg_b}^{-1}(u_1), Y_{2i}^* \le F_{2xg_b}^{-1}(u_2)\} - C_{xg_b}(u_1,u_2)\right],$$

(A4)
$$\begin{aligned} C_n(u_1,u_2) &= \mathrm{E}^*\, D_n(u_1,u_2) \\ &= \sqrt{n\,h_n}\,\sum_{i=1}^{n} w_{ni}(x;h_n)\,\Big[H_{x_i g_b}\left(F_{1xh}^{*-1}(u_1), F_{2xh}^{*-1}(u_2)\right) - H_{x_i g_b}\Big(F_{1xg_b}^{-1}(u_1), F_{2xg_b}^{-1}(u_2)\Big)\Big]. \end{aligned}$$

The proof will be divided into two steps. In Step 1, we will show that the term A_n is asymptotically negligible uniformly in (u_1,u_2). In Step 2 we will find the asymptotic representation of the process C_n. This asymptotic representation together with B_n will give us the representation (13).

A1. Step 1 – Treatment of A_n. For $(u_1,u_2) \in [0,1]^2$ and G_1, G_2 nondecreasing functions from $\mathbb{R}$ to $[0,1]$ define the stochastic processes

$$Z_{ni}(u_1,u_2,G_1,G_2) = \sqrt{n\,h_n}\,w_{ni}(x,h_n)\,\mathbb{I}\{Y_{1i}^* \le G_1^{-1}(u_1), Y_{2i}^* \le G_2^{-1}(u_2)\},\quad i=1,\ldots,n,$$

and $Z_n = \sum_{i=1}^n Z_{ni}$. Equivalently, we can view the process Z_n as a process indexed by the family of functions from $\mathbb{R}^2$ to $\mathbb{R}$ given by

(A5)
$$\mathcal{F} = \big\{(w_1,w_2) \mapsto \mathbb{I}\{w_1 \le G_1^{-1}(u_1), w_2 \le G_2^{-1}(u_2)\}; \\ (u_1,u_2) \in [0,1]^2,\, G_1, G_2 : \mathbb{R} \to [0,1] \text{ nondecreasing}\big\}.$$

Thus each function $f \in \mathcal{F}$ may be formally identified by a quadruple (u_1,u_2,G_1,G_2). The introduction of the process Z_n is motivated by the fact that

$$D_n(u_1,u_2) = Z_n(f_n^*) - Z_n(f_n), \tag{A6}$$

where $f_n^* = (u_1,u_2,F_{1xh}^*,F_{2xh}^*)$ and $f_n = (u_1,u_2,F_{1xg_b},F_{2xg_b})$.

Finally, let us equip the index set $\mathcal{F}$ with a semimetric ρ defined as

$$\rho^2(f,f') = \Big|F_{1x}(G_1^{-1}(u_1)) - F_{1x}(G_1'^{-1}(u_1'))\Big| + \Big|F_{2x}(G_2^{-1}(u_2)) - F_{2x}(G_2'^{-1}(u_2'))\Big|$$

and note that the semimetric space $(\mathcal{F},\rho)$ is totally bounded.

Lemma 1. *Under conditions* (**W1**), (**W2**), (**W6**), *and* (**H1**) *the process* $\bar{Z}_n = Z_n - \mathrm{E}^*\, Z_n$ *indexed by* $(\mathcal{F},\rho)$ *is asymptotically uniformly ρ-equicontinuous in bootstrap probability* P^* [P]-*almost surely*

Proof. It is sufficient to verify that the following conditions of Theorem 2.11.1 of van der Vaart and Wellner (1996) hold [P]-almost surely.

$$\sum_{i=1}^{n} \mathrm{E}^* \, \|Z_{ni}\|_{\mathcal{F}}^2 \, \mathbb{I}\{\|Z_{ni}\|_{\mathcal{F}} > \eta\} \;\to\; 0, \qquad \text{for every } \eta > 0, \tag{A7}$$

$$\sup_{\rho(f,f')<\delta_n} \sum_{i=1}^{n} \mathrm{E}^* \, \big(Z_{ni}(f) - Z_{ni}(f')\big)^2 \;\to\; 0, \qquad \text{for every } \delta_n \downarrow 0, \tag{A8}$$

$$\int_0^{\delta_n} \sqrt{\log N(\varepsilon, \mathcal{F}, d_n)}\, d\varepsilon \;\overset{P^*}{\to}\; 0, \qquad \text{for every } \delta_n \downarrow 0, \tag{A9}$$

where $\|.\|_{\mathcal{F}}$ stands for the supremum over the set $\mathcal{F}$ and $N(\varepsilon, \mathcal{F}, d_n)$ is the corresponding covering number of the set $\mathcal{F}$ with a random semimetric d_n given by

$$d_n^2(f, f') = \sum_{i=1}^{n} \big[Z_{ni}(f) - Z_{ni}(f')\big]^2 .$$

The first condition (A7) is satisfied as (**W1**) immediately implies $\max_{i=1,\dots,n} \|Z_{ni}\|_{\mathcal{F}} = o(1)$. To verify (A8) use (**W2**) and estimate

$$\begin{aligned}
&\sup_{\rho(f,f')<\delta_n} \sum_{i=1}^{n} \mathrm{E}^* \big(Z_{ni}(f) - Z_{ni}(f')\big)^2 \\
&\quad \le \sup_{\rho(f,f')<\delta_n} 2\, n\, h_n \sum_{i=1}^{n} w_{ni}^2(x, h_n) \Big[\Big| F_{1ig_b}(G_1^{-1}(u_1)) - F_{1ig_b}(G_1'^{-1}(u_1')) \Big| \\
&\qquad\qquad + \Big| F_{2ig_b}(G_2^{-1}(u_2)) - F_{2ig_b}(G_2'^{-1}(u_2')) \Big| \Big] \\
&\quad \le O(1)(R_{n1} + R_{n2}),
\end{aligned} \tag{A10}$$

where (with the help of (**W6**) and (**H**))

$$\begin{aligned}
R_{n1} = \sup_{\rho(f,f')<\delta_n} \max_{i \in I_x^{(n)}} \Big[&|F_{1i}(G_1^{-1}(u_1)) - F_{1x}(G_1^{-1}(u_1))| + |F_{1x}(G_1^{-1}(u_1)) - F_{1x}(G_1'^{-1}(u_1'))| \\
&+ |F_{1x}(G_1'^{-1}(u_1')) - F_{1i}(G_1'^{-1}(u_1'))| + |F_{2i}(G_2^{-1}(u_2)) - F_{2x}(G_2^{-1}(u_2))| \\
&+ |F_{2x}(G_2^{-1}(u_2)) - F_{2x}(G_2'^{-1}(u_2'))| + |F_{2x}(G_2'^{-1}(u_2')) - F_{2i}(G_2'^{-1}(u_2'))| \Big] \\
= o(1) + \delta_n^2 = o(1)
\end{aligned} \tag{A11}$$

and

$$R_{n2} = 2 \max_{i \in I_x^{(n)}} \sup_y |F_{1ig_b}(y) - F_{1i}(y)| + 2 \max_{i \in I_x^{(n)}} \sup_y |F_{2ig_b}(y) - F_{2i}(y)| \tag{A12}$$

For $j = 1, 2$ we may bound

$$\text{(A13)} \quad \max_{i \in I_x^{(n)}} \sup_y |F_{jig_b}(y) - F_{ji}(y)| \\ \leq \max_{i \in I_x^{(n)}} \sup_y |F_{jig_b}(y) - \mathrm{E}\, F_{jig_b}(y)| + \max_{i \in I_x^{(n)}} \sup_y |\mathrm{E}\, F_{jig_b}(y) - F_{ji}(y)|.$$

By assumption (**W2**) and Remark 2 of Appendix C the first term on the right-hand side of equation (A13) converges to zero [P]-almost surely. The second term is of order $o(1)$ thanks to assumptions (**W6**) and **(R3)**. Combining (A10), (A11), (A12) and (A13) now proves (A8).

Finally, (A9) may be verified in an analogous way as was done in Veraverbeke et al. (2011). □

Lemma 1 implies that for $\forall \varepsilon, \eta > 0$ $\exists \delta > 0$ such that [P]-almost surely

$$\text{(A14)} \qquad \limsup_{n \to \infty} \mathrm{P}^* \left[\sup_{\rho(f,f') < \delta} |\bar{Z}_n(f) - \bar{Z}_n(f')| > \varepsilon \right] < \eta.$$

Further by a triangular inequality, (**W1**) and (**W7**) for $j = 1, 2$

$$\text{(A15)} \quad \left| F_{jx}\left(F_{jxh}^{*-1}(u_j)\right) - F_{jx}\left(F_{jxg_b}^{-1}(u_j)\right)\right| \leq \\ \left| F_{jx}\left(F_{jxh}^{*-1}(u_j)\right) - F_{jxg_b}\left(F_{jxh}^{*-1}(u_j)\right)\right| + \left| F_{jxg_b}\left(F_{jxh}^{*-1}(u_j)\right) - F_{jxh}^{*}\left(F_{jxh}^{*-1}(u_j)\right)\right| \\ + \left| F_{jxh}^{*}\left(F_{jxh}^{*-1}(u_j)\right) - F_{jxg_b}\left(F_{jxg_b}^{-1}(u_j)\right)\right| + \left| F_{jxg_b}\left(F_{jxg_b}^{-1}(u_j)\right) - F_{jx}\left(F_{jxg_b}^{-1}(u_j)\right)\right| \\ \leq 2 \sup_y |F_{jxg_b}(y) - F_{jx}(y)| + \sup_y \left| F_{jxh}^{*}(y) - F_{jxg_b}(y)\right| + o(1).$$

It has been already proved that $\sup_y |F_{jxg_b}(y) - F_{jx}(y)| = o_{a.s.}(1)$ and in a similar way it can be proved $\sup_y \left| F_{jxh}^{*}(y) - F_{jxg_b}(y)\right| = o_{P^*}(1)$ [P]-almost surely.

Thus for $j = 1, 2$ uniformly in (u_1, u_2):

$$\left| F_{jx}\left(F_{jxh}^{*-1}(u_j)\right) - F_{jx}\left(F_{jxg_b}^{-1}(u_j)\right)\right| = o_{P^*}(1), \qquad \text{[P]-almost surely,}$$

which implies

$$\sup_{u_1,u_2} \rho^2 \left((u_1, u_2, F_{1xh}^{*}, F_{2xh}^{*}), (u_1, u_2, F_{1xg_b}, F_{2xg_b})\right) \\ \text{(A16)} \quad = \sup_{u_1} \left| F_{1x}\left(F_{1xh}^{*-1}(u_1)\right) - F_{1x}\left(F_{1xg_b}^{-1}(u_1)\right)\right| + \sup_{u_2} \left| F_{2x}\left(F_{2xh}^{*-1}(u_2)\right) - F_{2x}\left(F_{2xg_b}^{-1}(u_2)\right)\right|, \\ = o_{P^*}(1), \qquad \text{[P]-almost surely.}$$

Combining (A6), asymptotic ρ-equicontinuity (A14) and (A16) yields that

$$\text{(A17)} \quad \sup_{u_1,u_2} |A_n(u_1, u_2)| = \sup_{u_1,u_2} |D_n(u_1, u_2) - \mathrm{E}^*\, D_n(u_1, u_2)| = o_{P^*}(1), \quad \text{[P]-almost surely,}$$

which finishes the first step of the proof.

A2. Step 2 – Treatment of C_n. To simplify the notation for $j = 1, 2$ put $y^*_{jh} = F^{*-1}_{jxh}(u_j)$ and $y_{jg} = F^{-1}_{jxg_b}(u_j)$. With the help of (**W7**) and Taylor expansion (uniformly in (u_1, u_2))

$$\begin{aligned} C_n(u_1,u_2) &= \sqrt{n\,h_n}\sum_{i=1}^{n} w_{ni}(x;h_n)\;[H_{x_i g_b}(y^*_{1h},y^*_{2h}) - H_{x_i g_b}(y_{1g},y_{2g})] \\ &= C_{n1}(u_1,u_2) + C_{n2}(u_1,u_2) + C_{n3}(u_1,u_2), \end{aligned}$$

where

$$\begin{aligned} \text{(A18)}\; C_{n1}(u_1,u_2) &= \sqrt{n\,h_n}\sum_{i=1}^{n} w_{ni}(x;h_n)\,[H_{xg_b}(y^*_{1h},y^*_{2h}) - H_{xg_b}(y_{1g},y_{2g})]\,, \\ C_{n2}(u_1,u_2) &= \sqrt{n\,h_n}\sum_{i=1}^{n} w_{ni}(x;h_n)(x_i - x)\left[\dot{H}_{xg_b}(y^*_{1h},y^*_{2h}) - \dot{H}_{xg_b}(y_{1g},y_{2g})\right], \\ \text{(A19)}\; C_{n3}(u_1,u_2) &= \frac{1}{2}\sqrt{n\,h_n}\sum_{i=1}^{n} w_{ni}(x;h_n)(x_i - x)^2\left[\ddot{H}_{z_{hi}g_b}(y^*_{1h},y^*_{2h}) - \ddot{H}_{z_{gi}g_b}(y_{1g},y_{2g})\right], \end{aligned}$$

and z_{hi}, z_{gi} lie between x_i and x.

Part 1. Processes C_{n2} and C_{n3}. In the following we will show that the process C_{n3} is asymptotically negligible (the process C_{n2} can be handled in a similar way). To do so, we need to examine F^{*-1}_{1xh} and F^{*-1}_{2xh}.

Combining (**W2**), (**W3**), (**W6**), **(H1)** and Remark 2 of Appendix C yields that uniformly in y

$$\begin{aligned} \mathrm{E}^*\,F^*_{jxh}(y) &= \sum_{i=1}^{n} w_{ni}(x,h_n)F_{jx_i g_b}(y) = \sum_{i=1}^{n} w_{ni}(x,h_n)\mathrm{E}\,F_{jx_i g_b}(y) + o_{a.s.}(1) \\ \text{(A20)}\qquad &= F_{jx}(y) + o_{a.s.}(1), \qquad j = 1,2. \end{aligned}$$

Thus with the help of (A20), Lemma 3 of Appendix C, (**W2**) and (**W7**) we get that for every $\varepsilon > 0$ there exists a sufficiently large n such that

$$\begin{aligned} P^*\left[\sup_u\Big(F^{*-1}_{jxh}(u) - F^{-1}_{jx}(u+\varepsilon)\Big) \ge 0\right] &= P^*\left[\sup_u\Big(u - F^*_{jxh}(F^{-1}_{jx}(u+\varepsilon))\Big) \ge \tfrac{\varepsilon}{2}\right] \\ \le\; P^*\Bigg[\sup_u\Big(u - \mathrm{E}^*\,F^*_{jxh}(F^{-1}_{jx}(u+\varepsilon))\Big) &+ \\ \sup_u\Big(\mathrm{E}^*\,F^*_{jxh}(F^{-1}_{jx}(u+\varepsilon)) - F^*_{jxh}(F^{-1}_{jx}(u+\varepsilon))\Big) &\ge \tfrac{\varepsilon}{2}\Bigg] \\ \le\; P^*\left[\sup_y\big(\mathrm{E}^*\,F^*_{jxh}(y) - F^*_{jxh}(y)\big) \ge \tfrac{\varepsilon}{4}\right] &< \varepsilon. \qquad \text{[P]-a.s.} \end{aligned}$$

Similarly we can prove that the same holds true for the inequality $F^{*-1}_{jxh}(u) < F^{-1}_{jx}(u-\varepsilon)$. To summarize our conclusions, we have proved that for any $\varepsilon > 0$ with bootstrap probability P^*

going to one

$$F_{jx}^{-1}(u-\varepsilon) \leq F_{jxh}^{*-1}(u) \leq F_{jx}^{-1}(u+\varepsilon), \quad u \in [0,1], \qquad j=1,2, \qquad \text{[P]-a.s.} \tag{A21}$$

Analogously we can use Lemma 3 of Appendix C to prove that with probability P going to one

$$F_{jx}^{-1}(u-\varepsilon) \leq F_{jxg_b}^{-1}(u) \leq F_{jx}^{-1}(u+\varepsilon), \quad u \in [0,1], \qquad j=1,2. \tag{A22}$$

As the probability by which (A22) does not hold is going to zero at an exponential rate, we can use the Borel-Cantelli lemma to strengthen the result, such that (A22) holds [P]-a.s., as n tends to infinity.

With the help of (**W"2**) and (**W"5**) we can apply Lemma 4 of Appendix C with $d_{i,n}(z) = w''_{ni}(z, g_{bn})$ and I being a sufficiently small closed neighbourhood of the point x, which yields that uniformly in $(y_1, y_2) \in \mathbb{R}^2$ and $z \in I$

$$\ddot{H}_{zg_b}(y_1,y_2) = \mathrm{E}\,\ddot{H}_{zg_b}(y_1,y_2) + o_{a.s.}(1) = \sum_{i=1}^{n} w''_{ni}(z, g_{bn}) H_{x_i}(y_1,y_2) + o_{a.s.}(1). \tag{A23}$$

Using (**W"1**), (**W"3**), (**W"4**) and **(H)** we can further calculate (uniformly in (z, y_1, y_2))

$$\sum_{i=1}^{n} w''_{ni}(z, g_{bn}) H_{x_i}(y_1,y_2) o_{a.s.}(1). = \ddot{H}_z(y_1,y_2) + o(1). \tag{A24}$$

Definition of C_{n3} in (A19) together with (A23), (A24) yields

$$C_{n3}(u_1,u_2) = \frac{1}{2}\sqrt{n\,h_n}\sum_{i=1}^{n} w_{ni}(x;h_n)(x_i-x)^2\left[\ddot{H}_{z_{hi}}(y_{1h}^*, y_{2h}^*) - \ddot{H}_{z_{gi}}(y_{1g}, y_{2g})\right] + o_{a.s.}(1). \tag{A25}$$

Combining (A21), (A22), (A25), (**W5**) and **(H)** finally gives

$$\sup_{u_1,u_2} |C_{n3}(u_1,u_2)| = o_{P^*}(1), \qquad \text{[P]-a.s.}$$

Part 2. Process C_{n1}. Now, one can concentrate on the process C_{n1} given by (A18). First, note that with the help of (**W3**)

$$C_{n1}(u_1,u_2) = \sqrt{n\,h_n}\,\left[H_{xg_b}(y_{1h}^*, y_{2h}^*) - H_{xg_b}(y_{1g}, y_{2g})\right] + o(1).$$

For $j = 1,2$ and $u \in [0,1]$ put

$$\mathbb{Y}_{jn}(u) = \sqrt{n\,h_n}\left[u - F_{jxh}^*(F_{jxg_b}^{-1}(u))\right].$$

To finish the proof of Theorem 1 it remains to show that the process

$$Z_n(u_1,u_2) := C_{n1}(u_1,u_2) - C_x^{(1)}(u_1,u_2)\,\mathbb{Y}_{1n}(u_1) - C_x^{(2)}(u_1,u_2)\,\mathbb{Y}_{2n}(u_2), \tag{A26}$$

defined on $(u_1,u_2) \in [0,1]^2$ is asymptotically negligible in probability P^* [P]-almost surely.

First note that

$$\text{(A27)}\quad \mathbb{Y}_{jn}(u) = \sqrt{n\,h_n}\left[u - \mathrm{E}^*\, F^*_{jxh}(F^{-1}_{jxg_b}(u))\right] + \sqrt{n\,h_n}\left[\mathrm{E}^*\, F^*_{jxh}(F^{-1}_{jxg_b}(u)) - F^*_{jxh}(F^{-1}_{jxg_b}(u))\right].$$

Analogously as in Lemma 1 we can argue that the second term on the right-hand side of (A27) (viewed as a process in u) is uniformly asymptotic ρ-equicontinuous in probability P^* [P]-almost surely with $\rho(u,u') = |u-u'|$. For the first term on the right-hand side of (A27) we can use the similar reasoning as was used above for the treatment of the process C_{n3} and show that uniformly in u

$$\text{(A28)}\quad \sqrt{n\,h_n}\left[\mathrm{E}^*\, F^*_{jxh}(F^{-1}_{jxg_b}(u)) - u\right] = \sqrt{n\,h_n}\sum_{i=1}^{n} w_{ni}(x,h_n)(x_i-x)^2\ddot{F}_{jx}(F^{-1}_{jx}(u)) + o_{a.s.}(1).$$

Thus combining (A27) and (A28) implies

$$\text{(A29)}\qquad \sup_u |\mathbb{Y}_{jn}(u)| = O_{P^*}(1) \quad [\mathrm{P}]\text{ -a.s.}, \qquad j=1,2.$$

Moreover, thanks to **(R3)**, (A28) and the asymptotic ρ-continuity of the second term on the right-hand side of (A27), for each $\varepsilon > 0$ there exists $\delta_\varepsilon > 0$ such that for all sufficiently large n

$$\text{(A30)}\qquad \mathrm{P}^*\left(\max_{j=1,2}\ \sup_{u\in[0,\delta_\varepsilon]\cup[1-\delta_\varepsilon,1]} |\mathbb{Y}_{nj}(u)| > \varepsilon\right) \le \varepsilon \quad [\mathrm{P}]\text{ -a.s.}$$

In the following different representations of the processes $\mathbb{Y}_{1n}(u)$ and $\mathbb{Y}_{2n}(u)$ will be useful. For this reason note that $\sqrt{n\,h_n}(C^*_{xh}(u,1) - C_{xg_b}(u,1)) = o_{a.s.}(1)$ uniformly in u. Applying the decomposition (A1) and the asymptotic negligibility of the process A_n we get

$$\begin{aligned} C_n(u,1) &= -B_n(u,1) + o_{P^*}(1) = \sqrt{n\,h_n}\left[u - F^*_{1xh}(F^{-1}_{1xg_b}(u))\right] + o_{P^*}(1), \\ \text{(A31)}\qquad &= \mathbb{Y}_{1n}(u) + o_{P^*}(1), \quad [\mathrm{P}]\text{-a.s.} \end{aligned}$$

On the other hand decomposition of C_n given in (A18) and the asymptotic negligibility of the processes C_{n2} and C_{n3} yield

$$\text{(A32)}\quad C_n(u,1) = C_{n1}(u,1) + o_{P^*}(1) = \sqrt{n\,h_n}\left[F_{1xg_b}(F^{*-1}_{1xh}(u)) - u\right] + o_{P^*}(1)o_{P^*}(1), \quad [\mathrm{P}]\text{-a.s.}$$

Combining (A31) and (A32) gives

$$\text{(A33)}\qquad \mathbb{Y}_{1n}(u) = \sqrt{n\,h_n}\,[F_{1xg_b}(F^{*-1}_{1xh}(u)) - u] + o_{P^*}(1). \quad [\mathrm{P}]\text{-a.s.}$$

and in a completely analogous way we get

$$\text{(A34)}\qquad \mathbb{Y}_{2n}(u) = \sqrt{n\,h_n}\,[F_{2xg_b}(F^{*-1}_{2xh}(u)) - u] + o_{P^*}(1), \quad [\mathrm{P}]\text{-a.s.}$$

Focusing on the process Z_n note that with the help of (7), Lemma 3 in Appendix C, Condition (**W2**) and the Borel-Cantelli lemma

$$\text{(A35)}\quad \sqrt{nh_n}\,H_{xg_b}(y_1,y_2) = \sqrt{nh_n}\,\mathrm{E}\,H_{xg_b}(y_1,y_2) + o_{a.s.}(1), \quad \text{uniformly in } (y_1,y_2)\in\mathbb{R}^2,$$

which together with (A33) and (A34) implies

$$
\begin{aligned}
Z_n(u_1,u_2) = \sqrt{n\,h_n}\,\Big[&\mathrm{E}\,H_{xg_b}(y^*_{1h},y^*_{2h}) - \mathrm{E}\,H_{xg_b}(y_{1g},y_{2g}) \\
&- C^{(1)}_x(u_1,u_2)\,\mathrm{E}\,\left(F_{1xg_b}(y^*_{1h}) - F_{1xg_b}(y_{1g})\right) \\
&- C^{(2)}_x(u_1,u_2)\,\mathrm{E}\,\left(F_{2xg_b}(y^*_{2h}) - F_{2xg_b}(y_{2g})\right)\Big] + o_{a.s.}(1) \\
= \sqrt{n\,h_n}\,\sum_{i=1}^{n} w_{ni}(x,g_{bn})\Big[&C_{x_i}(F_{1x_i}(y^*_{1h}),F_{2x_i}(y^*_{2h})) - C_{x_i}(F_{1x_i}(y_{1g}),F_{2x_i}(y_{2g})) \\
&- C^{(1)}_x(u_1,u_2)\,(F_{1x_i}(y^*_{1h}) - F_{1x_i}(y_{1g})) \\
&- C^{(2)}_x(u_1,u_2)\,(F_{2x_i}(y^*_{2h}) - F_{2x_i}(y_{2g}))\Big] + o_{a.s.}(1)
\end{aligned}
$$

$$
\begin{aligned}
= \sqrt{n\,h_n}\,\sum_{i=1}^{n} w_{ni}(x,g_{bn})\,\Big[&\Big(C^{(1)}_{x_i}(u_{1i},u_{2i}) - C^{(1)}_x(u_1,u_2)\Big)\,(F_{1x_i}(y^*_{1h}) - F_{1x_i}(y_{1g})) \\
&+ \Big(C^{(2)}_{x_i}(u_{1i},u_{2i}) - C^{(2)}_x(u_1,u_2)\Big)\,(F_{2x_i}(y^*_{2h}) - F_{2x_i}(y_{2g}))\Big] + o_{a.s.}(1),
\end{aligned} \tag{A36}
$$

$$
= Z_{n1}(u_1,u_2) + Z_{n2}(u_1,u_2) + +o_{a.s.}(1), \tag{A37}
$$

where (u_{1i},u_{2i}) lies between $F_{1x_i}(y^*_{1h})$ and $F_{1x_i}(y_{1g})$.

Further (A29), (A33), (A34) and (A35) implies that for $j=1,2$

$$
\sup_{u_j}\left|\sqrt{n\,h_n}\,\sum_{i=1}^{n} w_{ni}(x,g_{bn})\,\big(F_{jx_i}(y^*_{jh}) - F_{jx_i}(y_{jg})\big)\right| = O_{P^*}(1), \quad [\mathrm{P}]\text{-a.s.} \tag{A38}
$$

Now, for given $\varepsilon > 0$ there exists $\delta > 0$ such that (A30) holds, which together with (A33), (A34) and (A35) yields that for sufficiently large n

$$
\mathrm{P}^*\left(\sup_{(u_1,u_2)\in([0,\delta_\varepsilon]\cup[1-\delta_\varepsilon,1])\times[0,1]} |\mathbb{Z}_{n1}(u_1,u_2)| > 2\,\varepsilon\right) \le 2\,\varepsilon \quad [\mathrm{P}]\text{ -a.s.} \tag{A39}
$$

Further combining (**W7**), (**R2**), (**R3**), (A36) and (A38) yields that

$$
\sup_{(u_1,u_2)\in[\delta_\varepsilon,1-\delta_\varepsilon]\times[0,1]} |Z_{n1}(u_1,u_2)| = o_{P*}(1) \qquad [\mathrm{P}]\text{-almost surely.} \tag{A40}
$$

Analogously one can treat the process Z_{n2}. Finally (A37), (A39) and (A40) imply that

$$
\sup_{(u_1,u_2)\in[0,1]^2} |Z_n(u_1,u_2)| = o_{P*}(1) \qquad [\mathrm{P}]\text{-almost surely,}
$$

which completes the proof.

5. Appendix – Proof of Theorem 2

The proof of Theorem 2 goes along the lines of the proof of Theorem 1 with (Y^*_{1i},Y^*_{2i}) being replaced with $(\tilde{U}^*_{1i},\tilde{U}^*_{2i})$, (Y_{1i},Y_{2i}) with $(\tilde{U}_{1i},\tilde{U}_{2i})$, $(\tilde{F}^*_{1x_ih},\tilde{F}^*_{2x_ih})$ with $(\tilde{G}^*_{1x_ih},\tilde{G}^*_{2x_ih})$, $(F_{1zg_b},F_{2zg_b},H_{zg_b})$, with $(\tilde{G}_{1zg_b},\tilde{G}_{2zg_b},\tilde{G}_{zg_b})$, ...

First, let us decompose

$$\sqrt{n\,h_n}\left(\tilde{C}^*_{xh} - \tilde{C}_{xg_b}\right) = \tilde{A}_n + \tilde{B}_n + \tilde{C}_n, \tag{B1}$$

where $\tilde{A}_n = \tilde{D}_n - \mathrm{E}^*\,\tilde{D}_n$, with

$$\tilde{D}_n(u_1,u_2) = \sqrt{n\,h_n}\sum_{i=1}^n w_{ni}(x;h_n)\,\big[\mathbb{I}\{\tilde{U}^*_{1i} \le \tilde{G}^{*-1}_{1xh}(u_1), U^*_{2i} \le \tilde{G}^{*-1}_{2xh}(u_2)\} - \mathbb{I}\{\tilde{U}^*_{1i} \le \tilde{G}^{-1}_{1xg_b}(u_1), \tilde{U}^*_{2i} \le \tilde{G}^{-1}_{2xg_b}(u_2)\}\big]. \tag{B2}$$

and

$$\tilde{B}_n(u_1,u_2) = \sqrt{n\,h_n}\left[\sum_{i=1}^n w_{ni}(x;h_n)\,\mathbb{I}\{\tilde{U}^*_{1i} \le \tilde{G}^{-1}_{1xg_b}(u_1), \tilde{U}^*_{2i} \le \tilde{G}^{-1}_{2xg_b}(u_2)\} - \tilde{C}_{xg_b}(u_1,u_2)\right]$$

$$\begin{aligned}\tilde{C}_n(u_1,u_2) &= \mathrm{E}^*\,\tilde{D}_n(u_1,u_2)\\ &= \sqrt{n\,h_n}\sum_{i=1}^n w_{ni}(x;h_n)\,\left[\tilde{G}_{x_ig_b}\left(\tilde{G}^{*-1}_{1xh}(u_1), \tilde{G}^{*-1}_{2xh}(u_2)\right) - \tilde{G}_{x_ig_b}\left(\tilde{G}^{-1}_{1xg_b}(u_1), \tilde{G}^{-1}_{2xg_b}(u_2)\right)\right].\end{aligned} \tag{B3}$$

As in the following we will make often use of Lemma 6, we need to investigate how close a pseudo-observation $\tilde{U}^b_{ji}$ is to an unobserved U_{ji}. For $j = 1, 2$ put

$$\mathcal{E}^{(n)}_{jg}(z,u) = F_{jzg_{jb}}(F^{-1}_{jz}(u)) - u, \qquad u \in [0,1], \tag{B4}$$

and note that $\mathcal{E}^{(n)}_{jg}(x_i, U_{ji}) = \tilde{U}^b_{ji} - U_{ji}$.

By Taylor expansion of $\mathrm{E}\,\mathcal{E}^{(n)}_{jg}(z,u)$, and assumptions (**W3**), (**W4**), (**W5**) and **(R3)** there exist a neighbourhood U_x of x and a constant C such that for all sufficiently large n

$$\sup_{u\in[0,1]}\sup_{z\in U_x}\left|\mathrm{E}\,\mathcal{E}^{(n)}_{jg}(z,u)\right| \le C\,g^2_{jbn} \tag{B5}$$

Thanks to the exponential inequality given by Lemma 3 of Appendix C and assumption (**W2**) there exist constants C_1, C_2, C_3 such for each $\delta > 0$

$$\begin{aligned}P&\left(\max_{i\in I^{(n)}_x}\sup_{u\in[0,1]}\left|\mathcal{E}^{(n)}_{jg}(x_i,u) - \mathrm{E}\,\mathcal{E}^{(n)}_{jg}(x_i,u)\right| > \tfrac{\log^{1/2+\delta} n}{\sqrt{n\,g_{jbn}}}\right)\\ &\le \sum_{i\in I^{(n)}_x} C_1 \exp\left\{-\tfrac{C_2\,\log^{1+\delta} n}{n\,g_{jbn}\sum_{k=1}^n w^2_{nk}(x_i,g_{jbn})}\right\} \le n\,C_2\,\exp\left\{-C_3\log^{1+\delta} n\right\} = O(n^{-2}).\end{aligned} \tag{B6}$$

The Borel-Cantelli lemma together with (B6) now implies that

$$\max_{i\in I^{(n)}_x}\sup_{u\in[0,1]}\left|\mathcal{E}^{(n)}_{jg}(x_i,u) - \mathrm{E}\,\mathcal{E}^{(n)}_{jg}(x_i,u)\right| = O_{a.s.}\left(\tfrac{\log^{1/2+\delta} n}{\sqrt{n\,g_{jbn}}}\right). \tag{B7}$$

Thus if one takes a sufficiently small δ in (B7), the assumption (8) together with (B5) and (B7) implies that

$$\max_{i \in I_x^{(n)}} |\tilde{U}_{ji}^b - U_{ji}| \le O_{a.s.}\left(g_{jbn}^2\right), \qquad j = 1, 2. \tag{B8}$$

B1. Step 1 – Treatment of $\tilde{A}_n$. The proof is analogous to Step 1 of Appendix A. Note that as the marginal distribution functions (G_{1x_i}, G_{2x_i}), corresponding to (F_{1x_i}, F_{2x_i}) in the proof Theorem 1, are uniform, the semimetric ρ is given directly by

$$\rho^2(f, f') = \left|G_1^{-1}(u_1) - G_1'^{-1}(u_1')\right| + \left|G_2^{-1}(u_2) - G_2'^{-1}(u_2')\right|.$$

The only difference in the proof of Lemma 1 is in (A12). The term R_{n2} is now given by

$$R_{n2} = 2 \max_{i \in I_x^{(n)}} \sup_u |\tilde{G}_{1ig_b}(u) - u| + 2 \max_{i \in I_x^{(n)}} \sup_u |\tilde{G}_{2ig_b}(u) - u|$$

which converges almost surely to zero by (**W2**), (B8) and Lemma 6.

Thus, to finish Step 1 of the proof of Theorem 2 it is sufficient to show that for $j = 1, 2$

$$\sup_u \left|\tilde{G}_{jxh}^{*-1}(u) - \tilde{G}_{jxg_b}^{-1}(u)\right| = o_{P^*}(1) \quad \text{[P]-a.s.}$$

In the same way as in (A15) we can bound

$$\left|\tilde{G}_{jxh}^{*-1}(u) - \tilde{G}_{jxg_b}^{-1}(u)\right| \le 2 \sup_u \left|\tilde{G}_{jxg_b}(u) - u\right| + \sup_u \left|\tilde{G}_{jxh}^{*}(u) - \tilde{G}_{jxg_b}(u)\right| + o(1) \tag{B9}$$

We have already argued that with the help of Lemma 6: $\sup_u \left|\tilde{G}_{jxg_b}(u) - u\right| = o_{a.s.}(1)$. Further, the second term on the right-hand side of (B9) can be bounded as

$$\sup_u \left|\tilde{G}_{jxh}^{*}(u) - \tilde{G}_{jxg_b}(u)\right| \le \sup_u \left|\tilde{G}_{jxh}^{*}(u) - \mathrm{E}^* \tilde{G}_{jxh}^{*}(u)\right| + \sup_u \left|\mathrm{E}^* \tilde{G}_{jxh}^{*}(u) - \tilde{G}_{jxg_b}(u)\right|. \tag{B10}$$

The first term on the right hand side of (B10) can be handled by Lemma 3 and assumption (**W2**). For the second term one can use assumption (**W3**) and Lemma 6 and show that

$$\begin{aligned} \sup_u \left|\mathrm{E}^* \tilde{G}_{jxh}^{*}(u) - \tilde{G}_{jxg_b}(u)\right| &= \sup_u \left|\sum_{i=1}^n w_{ni}(x, h_n)\, \tilde{G}_{jx_ig_b}(u) - \tilde{G}_{jxg_b}(u)\right| \\ &= \sup_u \left|\sum_{i=1}^n w_{ni}(x, h_n)\, u - u + o_{a.s.}(1)\right| = o_{a.s}(1). \end{aligned} \tag{B11}$$

B2. Step 2 – Treatment of $\tilde{C}_n$. To simplify the notation let us for $j = 1, 2$ denote $u_{jh}^* = \tilde{G}_{jxh}^{*-1}(u_j)$ and $u_{jg} = \tilde{G}_{jxg_b}^{-1}(u_j)$. With the help of Taylor expansion

$$\begin{aligned} \tilde{C}_n(u_1, u_2) &= \sqrt{n\, h_n} \sum_{i=1}^n w_{ni}(x; h_n) \left[\tilde{G}_{x_ig_b}(u_{1h}^*, u_{2h}^*) - \tilde{G}_{x_ig_b}(u_{1g}, u_{2g})\right] \\ &= \tilde{C}_{n1}(u_1, u_2) + \tilde{C}_{n2}(u_1, u_2) + \tilde{C}_{n3}(u_1, u_2), \end{aligned}$$

where

$$\text{(B12)}\quad \tilde{C}_{n1}(u_1,u_2) = \sqrt{n\,h_n}\sum_{i=1}^{n} w_{ni}(x;h_n)\left[\tilde{G}_{xg_b}(u_{1h}^*,u_{2h}^*) - \tilde{G}_{xg_b}(u_{1g},u_{2g})\right],$$

$$\tilde{C}_{n2}(u_1,u_2) = \sqrt{n\,h_n}\sum_{i=1}^{n} w_{ni}(x;h_n)(x_i-x)\left[\dot{\tilde{G}}_{xg_b}(u_{1h}^*,u_{2h}^*) - \dot{\tilde{G}}_{xg_b}(u_{1g},u_{2g})\right],$$

$$\text{(B13)}\quad \tilde{C}_{n3}(u_1,u_2) = \frac{1}{2}\sqrt{n\,h_n}\sum_{i=1}^{n} w_{ni}(x;h_n)(x_i-x)^2\left[\ddot{\tilde{G}}_{z_{hi}g_b}(u_{1h}^*,u_{2h}^*) - \ddot{\tilde{G}}_{z_{gi}g_b}(u_{1g},u_{2g})\right],$$

and z_{hi}, z_{gi} lie between x_i and x.

Part 1. Processes $\tilde{C}_{n2}$ *and* $\tilde{C}_{n3}$. To treat the process $\tilde{C}_{n3}$ (the process $\tilde{C}_{n2}$ may be treated analogously) we will first show that for $j = 1, 2$ uniformly in u

$$\text{(B14)}\qquad \tilde{G}_{jxh}^{*-1}(u) = u + o_{P^*}(1)\quad \text{[P]-a.s.};\qquad \text{(B15)}\qquad \tilde{G}_{jxg_b}^{-1}(u) = u + o_{a.s.}(1).$$

To prove equation (B14) let $\varepsilon > 0$ be given. Similarly as in (B11) we can show that $\mathrm{E}^*\,\tilde{G}_{jxh}(u) = u + o_{a.s.}(1)$ uniformly in u. Thus for all sufficiently large n with the help of the inequality (C1), (**W1**) and (**W2**)

$$\text{(B16)}\quad P^*\left[\sup_u\left(\tilde{G}_{jxh}^{*-1}(u) - u\right) > \varepsilon\right] \le P^*\left[\sup_u\left(u - \tilde{G}_{jxh}^*(u)\right) > \tfrac{\varepsilon}{2}\right]$$
$$\le P^*\left[\sup_u\left(\mathrm{E}^*\,\tilde{G}_{jxh}^*(u) - \tilde{G}_{jxh}^*(u)\right) > \tfrac{\varepsilon}{4}\right] \to 0.$$

In a similar way one can show that $P^*\left[\inf_u\left(\tilde{G}_{jxh}^{*-1}(u) - u\right) < -\varepsilon\right] \to 0$, which completes the proof of (B14).

The proof of (B15) follows from the inequality

$$\sup_u\left|\tilde{G}_{jxg_b}^{-1}(u) - u\right| \le \sup_u\left|\tilde{G}_{jxg_b}(u) - u\right| + o(1)$$

and Lemma 6.

Further, with the help of (**W"2**), (**W"5**), (**W"6**) and (B8) we can apply Lemma 6 with $d_{i,n}(z) = w_{ni}''(z, g_{bn})$ and I being a sufficiently small closed neighbourhood of the point x, which yields that uniformly in $(u_1,u_2) \in [0,1]^2$ and $z \in I$

$$\text{(B17)}\qquad \ddot{\tilde{G}}_{zg_b}(u_1,u_2) = \sum_{i=1}^{n} w_{ni}''(z,g_{bn})\,C_{x_i}(u_1,u_2) + o_{a.s.}(1).$$

Using (**W"1**), (**W"3**), (**W"4**) and **(R1)** we can further calculate (uniformly in (z,u_1,u_2))

$$\text{(B18)}\qquad \sum_{i=1}^{n} w_{ni}''(z,g_{bn})\,C_{x_i}(y_1,y_2) = \ddot{C}_z(y_1,y_2) + o(1).$$

Definition of C_{n3} in (A19) together with (B17), (B18) yield

(B19)
$$\tilde{C}_{n3}(u_1,u_2) = \frac{1}{2}\sqrt{n\,h_n}\sum_{i=1}^{n} w_{ni}(x;h_n)(x_i-x)^2\left[\ddot{C}_{z_{hi}}(u_{1h}^*,u_{2h}^*) - \ddot{C}_{z_{gi}}(u_{1g},u_{2g})\right] + o_{a.s.}(1).$$

Combining (B14), (B15), (B19) with assumptions (**W5**) and **(R1)** finally gives

$$\sup_{u_1,u_2}|C_{n3}(u_1,u_2)| = o_{P^*}(1) \qquad \text{[P]-almost surely.}$$

Part 2. Process $\tilde{C}_{n1}$. First, note that with the help of (B12), (**W1**) and (**W3**)

$$\begin{aligned}\tilde{C}_{n1}(u_1,u_2) &= \sqrt{n\,h_n}\,\left[\tilde{G}_{xg_b}(u_{1h}^*,u_{2h}^*) - \tilde{G}_{xg_b}(u_{1g},u_{2g})\right] + o(1)\\ &= \sqrt{n\,h_n}\sum_{i=1}^{n} w_{ni}(x,g_{bn})\Big[\mathbb{I}\{\tilde{U}_{1i}^b \le u_{1h}^*, \tilde{U}_{2i}^b \le u_{2h}^*\}\\ &\qquad\qquad - \mathbb{I}\{\tilde{U}_{1i}^b \le u_{1g}, \tilde{U}_{2i}^b \le u_{2g}\}\Big] + o(1).\end{aligned}$$

For $j=1,2$ and $u\in[0,1]$ put

$$\mathbb{Y}_{jn}(u) = \sqrt{n\,h_n}\left[u - \tilde{G}_{jxh}^*(\tilde{G}_{jxg_b}^{-1}(u))\right].$$

To finish the proof of Theorem 2 we need to show that the process

(B20)
$$Z_n(u_1,u_2) := \tilde{C}_{n1}(u_1,u_2) - C_x^{(1)}(u_1,u_2)\,\mathbb{Y}_{1n}(u_1) - C_x^{(2)}(u_1,u_2)\,\mathbb{Y}_{2n}(u_2),$$

where $(u_1,u_2)\in[0,1]^2$, is asymptotically negligible in probability P^* [P]-almost surely.

Analogously as in Part 2 of the second step of the proof of Theorem 1 we can argue that (A29) and (A30) hold.

Further, putting either u_1 or u_2 to 1 in (B1) and the asymptotic negligibility of the processes $\tilde{A}_n$, $\tilde{C}_{n1}$ and $\tilde{C}_{n2}$ imply

(B21)
$$\mathbb{Y}_{jn}(u) = \sqrt{n\,h_n}\,[\tilde{G}_{jxg_b}(\tilde{G}_{jxh}^{*-1}(u)) - u] + o_{P^*}(1), \quad \text{[P]-a.s.}$$

Dealing with $\tilde{C}_{n1}$ is tricky as the transformed 'uniform' alike observations $(\tilde{U}_{1i},\tilde{U}_{2i})$ are involved. The following lemma will be useful.

Lemma 2. *For $j=1,\dots,n$ and $i=1,\dots,n$ put $R_{ji}^{(n)}(u) = \mathrm{E}\,\mathcal{E}_{jg}^{(n)}(x_i,u)$ and define*

$$\begin{aligned}V_n(u_1,u_2) = \sqrt{n\,h_n}\sum_{i=1}^{n} w_{ni}(x,g_{bn})\Big[&\mathbb{I}\{\tilde{U}_{1i}\le u_1, \tilde{U}_{2i}\le u_2\}\\ &- \mathbb{I}\{U_{1i} + R_{1i}^{(n)}(U_{1i}) \le u_1, U_{2i} + R_{2i}^{(n)}(U_{2i}) \le u_2\}\Big].\end{aligned}$$

Then

(B22)
$$\sup_{u_1,u_2}|V_n(u_1,u_2)| = o_{a.s.}(1).$$

Proof. First, note that for $j = 1, 2$ and $i = 1, \dots, n$

$$\tilde{U}_{ji} = U_{ji} + \tilde{U}_{ji} - U_{ji} = U_{ji} + \mathcal{E}_{jg}^{(n)}(x_i, U_{ji}) - R_{ji}^{(n)}(U_{ji}) + R_{ji}^{(n)}(U_{ji}). \tag{B23}$$

By (B7) one knows that for each $\delta > 0$

$$\max_j \max_{1 \le i \le n} \sup_u \left| \mathcal{E}_{jg}^{(n)}(x_i, u) - R_{ji}^{(n)}(u) \right| = O_{a.s.}(a_n), \quad \text{where } a_n = \frac{\log^{1/2+\delta} n}{\sqrt{n \, \min\{g_{1bn}, g_{2bn}\}}},$$

which together with (B23) imply that with $n \to \infty$ [P]-almost surely

$$\max_j \max_{1 \le i \le n} \left| \tilde{U}_{ji} - U_{ji} - R_{ji}^{(n)}(U_{ji}) \right| \le a_n.$$

Thus for sufficiently large n with the help of (**W7**) one can bound

$$|V_n(u_1, u_2)| \le V_{1n}(u_1) + V_{2n}(u_2) + o(1)$$

where

$$V_{jn}(u_j) = \sqrt{n\, h_n} \sum_{i=1}^n w_{ni}(x, g_{bn}) \Big[\mathbb{I}\{U_{ji} + R_{ji}^{(n)}(U_{ji}) \le u_j + a_n\} - \mathbb{I}\{U_{ji} + R_{ji}^{(n)}(U_{ji}) \le u_j - a_n\} \Big].$$

Assumption (**W2**) and Lemma 4 yield

$$V_{jn}(u_j) = \sqrt{n\, h_n} \sum_{i=1}^n w_{ni}(x, g_{bn}) \Big[\mathrm{P}\Big(U_{ji} + R_{ji}^{(n)}(U_{ji}) \le u_j + a_n \Big) - \mathrm{P}\Big(U_{ji} + R_{ji}^{(n)}(U_{ji}) \le u_j - a_n \Big) \Big] + o_{a.s.}(1),$$

which can be further with the help of (8), **(R4)** and (B5) bounded as

$$\begin{aligned} V_{jn}(u_j) &\le \sqrt{n\, h_n} \sum_{i=1}^n w_{ni}(x, g_{bn}) \Big[\mathrm{P}\Big(U_{ji} \le u_j + a_n - \inf_{w \in I_{u_j}^{(n)}} R_{ji}^{(n)}(w) \Big) \\ &\qquad - \mathrm{P}\Big(U_{ji} \le u_j - a_n - \sup_{w \in I_{u_j}^{(n)}} R_{ji}^{(n)}(w) \Big) \Big] + o_{a.s.}(1), \\ &\le 2 \sqrt{n\, h_n} \, \Big(a_n + C^{1+\eta} g_{jbn}^{2+2\eta} + C\, a_n^{\eta}\, g_{jbn}^2 \Big) + o_{a.s.}(1) = o_{a.s.}(1), \end{aligned}$$

where $I_{u_j}^{(n)} = \Big[u_j - a_n - C\, g_{jbn}^2,\, u_j + a_n + C\, g_{jbn}^2 \Big]$. □

Let us denote $G_{ji}^{(n)}$ the distribution function of the random variable $U_{ji} + R_{ji}^{(n)}(U_{ji})$. Note that by **(R3)**

$$\max_{j=1,2} \max_{i \in I_x^{(n)}} \sup_u \left| G_{ji}^{(n)}(u) - u \right| \le o(1). \tag{B24}$$

With the help of (A29), (B21), Lemma 2 and Lemma 4 we get that uniformly in u_j [P]-almost surely

$$
\begin{aligned}
O_{P^*}(1) &= \mathbb{Y}_{nj}(u_j) = \sqrt{n\,h_n}\,\Big[\tilde{G}_{jxg_b}(u^*_{jh}) - \tilde{G}_{jxg_b}(u_{jg})\Big] + o_{P^*}(1) \\
&= \sqrt{n\,h_n}\,\sum_{i=1}^{n} w_{ni}(x, g_{bn})\,\Big[\mathbb{I}\{U_{ji} + R^{(n)}_{ji}(U_{ji}) \le u^*_{jh}\} \\
&\qquad - \mathbb{I}\{U_{ji} + R^{(n)}_{ji}(U_{ji}) \le u_{jg}\}\Big] + o_{P^*}(1) \\
&= \sqrt{n\,h_n}\,\sum_{i=1}^{n} w_{ni}(x, g_{bn})\,\Big[G^{(n)}_{ji}\,(u^*_{jh}) - G^{(n)}_{ji}\,(u_{jg})\Big] + o_{P^*}(1).
\end{aligned} \tag{B25}
$$

Similarly with the help of Lemma 2 and Lemma 4 we get

$$
\begin{aligned}
\tilde{C}_{n1}(u_1,u_2) &= \sqrt{n\,h_n}\,\sum_{i=1}^{n} w_{ni}(x, g_b)\,\Big[\mathrm{P}\left(U_{1i} + R^{(n)}_{1i}(U_{1i}) \le u^*_{1h}, U_{2i} + R^{(n)}_{2i}(U_{ji}) \le u^*_{2h}\right) \\
&\qquad - \mathrm{P}\left(U_{1i} + R^{(n)}_{1i}(U_{1i}) \le u_{1g}, U_{2i} + R^{(n)}_{2i}(U_{ji}) \le u_{2g}\right)\Big] + o_{a.s.}(1).
\end{aligned}
$$

$$
\begin{aligned}
&= \sqrt{n\,h_n}\,\sum_{i=1}^{n} w_{ni}(x, g_b)\,\Big[C_{x_i}\left(G^{(n)}_{1i}(u^*_{1h}), G^{(n)}_{2i}(u^*_{2h})\right) \\
&\qquad - C_{x_i}\left(G^{(n)}_{1i}(u_{1g}), G^{(n)}_{2i}(u_{2g})\right)\Big] + o_{a.s.}(1).
\end{aligned} \tag{B26}
$$

Now with the help of (B14), (B15), (B20), (B24), and (B26) one can derive that uniformly in (u_1, u_2)

$$
\begin{aligned}
Z_n(u_1,u_2) = \sqrt{n\,h_n}\,\sum_{i=1}^{n} w_{ni}(x, g_{bn})\,&\Big[\left(C^{(1)}_{x_i}(u_{1i},u_{2i}) - C^{(1)}_{x}(u_1,u_2)\right)\left(G^{(n)}_{1i}(u^*_{1h}) - G^{(n)}_{1i}(u_{1g})\right) \\
&+ \left(C^{(2)}_{x_i}(u_{1i},u_{2i}) - C^{(2)}_{x}(u_1,u_2)\right)\left(G^{(n)}_{2i}(u^*_{2h}) - G^{(n)}_{2i}(u_{2g})\right)\Big] + o_{a.s.}(1),
\end{aligned} \tag{B27}
$$

where u_{ji} lies between $G^{(n)}_{ji}(u^*_{jh})$ and $G^{(n)}_{ji}(u_{jg})$.

Now, with the help of (**W7**), **(R2)** and (B25) one can use an analogous reasoning as at the end of the proof of Theorem 1 to show the asymptotic negligibility of the process Z_n which finishes the proof.

6. Appendix C – Auxiliary lemmas

The following technical lemmas are formulated for general p, but in this paper we will make use of them only for $p = 1, 2$. For brevity in the following we will refer to the book of van der Vaart and Wellner (1996) as VW (1996).

Lemma 3. *Let $\{d_{i,n}, i = 1, \ldots, n\}_{n=1}^{\infty}$ be a triangular array of constants and $\{\mathbf{Y}_i\}_{i=1}^{\infty}$ a sequence of p-dimensional independent random vectors with distribution functions $\{H_i\}_{i=1}^{\infty}$.*

Then there exist finite positive constants C_1, C_2 (which do not depend on the array $\{d_{i,n}\}$) such that for all $n \in \mathbb{N}$ and for all $\varepsilon > 0$

$$\text{(C1)} \qquad \mathrm{P}\left(\sup_{\mathbf{y}\in\mathbb{R}^p}\left|\sum_{i=1}^{n} d_{i,n}\,[\mathbb{I}\{\mathbf{Y}\leq\mathbf{y}\} - H_i(\mathbf{y})]\right| > \varepsilon\right) \leq C_1\, \exp\left\{-\tfrac{C_2\,\varepsilon^2}{|\mathbf{d}_n|_2^2}\right\},$$

where $|\mathbf{d}_n|_2^2 = \sum_{i=1}^n d_{i,n}^2$.

Proof. Let us introduce a family of functions from $\mathbb{R}^p \to \mathbb{R}$

$$\mathcal{F} = \{(z_1,\dots,z_p) \to \mathbb{I}\{z_1 \leq y_1,\dots,z_p \leq y_p\}, (y_1,\dots,y_p) \in \mathbb{R}^p\}$$

For $f_{\mathbf{y}} = \mathbb{I}\{\cdot \leq y_1,\dots,\cdot \leq y_p\} \in \mathcal{F}$ define a process

$$X_n(f_{\mathbf{y}}) = \sum_{i=1}^n d_{i,n}\,\mathbb{I}\{Y_{1i} \leq y_1,\dots,Y_{pi} \leq y_p\}.$$

Thus we need to prove the exponential inequality (C1) for $\sup_{f\in\mathcal{F}}|X_n(f) - \mathrm{E}\,X_n(f)|$. To prevent measurable difficulties notice that

$$\sup_{f\in\mathcal{F}}|X_n(f) - \mathrm{E}\,X_n(f)| = \sup_{f\in\mathcal{G}}|X_n(f) - \mathrm{E}\,X_n(f)|,$$

where

$$\mathcal{G} = \{(z_1,\dots,z_p) \to \mathbb{I}\{z_1 \leq y_1,\dots,z_p \leq y_p\}, (y_1,\dots,y_p) \in \mathbb{Q}^p\}$$

Thus the switch between $\mathcal{F}$ and $\mathcal{G}$ can be made whenever necessary to assure measurability (for details see VW (1996) p. 110).

Further put $\psi_2(x) = \exp\{x^2\} - 1$ and define the corresponding Orlicz norm

$$\|X\|_{\psi_2} = \inf\left\{C : \mathrm{E}\,\psi_2\left(\tfrac{|X|}{C}\right) \leq 1\right\}.$$

Applying the symmetrization inequality (see Lemma 2.3.1 of VW (1996)) yields

$$\mathrm{E}\left\|\sup_{f\in\mathcal{F}}|X_n(f) - \mathrm{E}\,X_n(f)|\right\|_{\psi_2} = 2\,\mathrm{E}\left\|\sup_{f\in\mathcal{F}}|X_n^s(f)|\right\|_{\psi_2},$$

where $X_n^s(f) = \sum_{i=1}^n \sigma_i\, d_{i,n}\,\mathbb{I}\{Y_{1i} \leq y_1, Y_{2i} \leq y_2\}$ and $\sigma_1,\dots,\sigma_n$ be Rademacher variables, that is they are independent with $P(\sigma_i = \pm 1) = \frac{1}{2}$.

Further for $f_{\mathbf{y}}, f_{\mathbf{z}} \in \mathcal{F}$ define a semimetric

$$\rho(f_{\mathbf{y}}, f_{\mathbf{z}})^2 = \sum_{i=1}^n d_{i,n}^2[f_{\mathbf{y}}(Y_{1i},Y_{2i}) - f_{\mathbf{z}}(Y_{1i},Y_{2i})]^2.$$

By Hoeffding inequality

$$\mathrm{P}\left(|X_n^s(f_{\mathbf{y}}) - X_n^s(f_{\mathbf{z}})| > \varepsilon\right) \leq 2\exp\left\{-\tfrac{\varepsilon^2}{2\,\rho(f_{\mathbf{y}},f_{\mathbf{z}})^2}\right\}.$$

Corollary 2.2.5 of VW (1996) together with the relation of packing and covering numbers imply the existence of a finite constant K such that

$$\mathrm{E}\,\Big\|\sup_{f\in\mathcal{F}}|X_n^s(f)|\Big\|_{\psi_2}\leq K\int_0^{|\mathbf{d}_n|_2}\sqrt{\log N(\tfrac{\varepsilon}{2},\rho)}\,d\varepsilon, \tag{C2}$$

where $N(\varepsilon,\rho)$ is a covering number of $\mathcal{F}$ equipped with a semimetric ρ. Define a measure Q on $\mathbb{R}^2$ by

$$Q=\sum_{i=1}^n \tfrac{d_{i,n}^2}{|\mathbf{d}_n|_2^2}\,\delta_{(Y_{1i},\ldots,Y_{pi})},\qquad \text{where } \delta \text{ is a Dirac measure.}$$

Then

$$\rho(f_{\mathbf{y}},f_{\mathbf{z}})^2=|\mathbf{d}_n|_2^2\,\mathrm{E}_Q(f_{\mathbf{y}}-f_{\mathbf{z}})^2=|\mathbf{d}_n|_2^2\,\|f_{\mathbf{y}}-f_{\mathbf{z}}\|_{L_2(Q)}^2,$$

which implies $N(\varepsilon,\rho)=N\Big(\tfrac{\varepsilon}{|\mathbf{d}_n|_2},\mathcal{F},L_2(Q)\Big)$.

As $\mathcal{F}$ is a VC-class of functions with envelope $F=1$, then by Theorem 2.6.7 of VW (1996) there exist finite constants C_0,C_1 (independent of the measure Q) such that

$$N(\varepsilon,\mathcal{F},L_2(Q))\leq C_0\left(\tfrac{1}{\varepsilon}\right)^{C_1},\qquad 0<\varepsilon<1.$$

Thus for $\varepsilon<|\mathbf{d}_n|_2$ we get

$$N(\tfrac{\varepsilon}{2},\rho)=N\Big(\tfrac{\varepsilon}{2\,|\mathbf{d}_n|_2},\mathcal{F},L_2(Q)\Big)\leq C_0\left(\tfrac{|\mathbf{d}_n|_2}{2\,\varepsilon}\right)^{C_1}. \tag{C3}$$

Combining (C2) with (C3) yields

$$\begin{aligned}\mathrm{E}\,\Big\|\sup_{f\in\mathcal{F}}|X_n^s(f)|\Big\|_{\psi_2}&\leq K\int_0^{|\mathbf{d}_n|_2}\sqrt{\log\Big(C_0\left(\tfrac{|\mathbf{d}_n|_2}{2\,\varepsilon}\right)^{C_1}\Big)}\,d\varepsilon\\&=K|\mathbf{d}_n|_2\int_0^1\sqrt{\log\Big(C_0\left(\tfrac{1}{2\,\varepsilon}\right)^{C_1}\Big)}\,d\varepsilon=O(|\mathbf{d}_n|_2),\end{aligned}$$

which together with the tail inequality $P(|X|>x)\leq\frac{1}{\psi_2(x/\|X\|_{\psi_2})}$ finishes the proof of the lemma. ☐

Lemma 4. *Let $\{d_{i,n}(x),i=1,\ldots,n\}_{n=1}^\infty$ be a triangular array of functions on a finite interval I and $\{\mathbf{Y}_i\}_{i=1}^\infty$ be a sequence of p-dimensional independent random vectors with distribution functions $\{H_i\}_{i=1}^\infty$. Suppose there exist finite positive constants C_0,L,α such that for all $x,x'\in I$ and all $n\in\mathbb{N}$*

$$\sum_{i=1}^n|d_{i,n}(x)-d_{i,n}(x')|\leq C_0\,n^L|x-x'|^\alpha. \tag{C4}$$

Further suppose that there exists a finite positive $C_3>0$ and $\delta>0$ such that for all $n\in\mathbb{N}$

$$\sup_{x\in I}\sum_{i=1}^n d_{i,n}^2(x)\leq\tfrac{C_3}{n^\delta}. \tag{C5}$$

Then as $n \to \infty$

$$\sup_{x\in I}\sup_{y_1,\dots,y_p}\left|\sum_{i=1}^n d_{i,n}(x)\,[\mathbb{I}\{Y_{1i}\le y_1,\dots,Y_{pi}\le y_p\}-H_i(y_1,\dots,y_p)]\right| = o_{a.s}(1). \tag{C6}$$

Proof. Without loss of generality we may suppose the interval I to be a unit interval $[0,1]$. Further for each n define a grid of points $\mathbf{X}_n = \{0 = x_{0,n} < x_{1,n} < \dots. < x_{m_n,n} = 1\}$ such that

$$\max_i |x_{i,n} - x_{i-1,n}| \le \tfrac{1}{[C_0 n^{L+1}]^{1/\alpha}} \qquad \text{and} \qquad \min_i |x_{i,n} - x_{i-1,n}| \ge \tfrac{1}{2\,[C_0 n^{L+1}]^{1/\alpha}},$$

where the constants C_0, L are taken from assumption (C4).

Let us denote π_n the projection from I to $\mathbf{X}_n$ which maps $x \in I$ to its closest left neighbour in $\mathbf{X}_n$. Then

$$\begin{aligned}\sup_{x\in I}\sup_{y_1,\dots,y_p}&\left|\sum_{i=1}^n d_{i,n}(x)\,[\mathbb{I}\{Y_{1i}\le y_1,\dots,Y_{pi}\le y_p\}-H_i(y_1,\dots,y_p)]\right|\\ &\le \sup_{x\in \mathbf{X}_n}\sup_{y_1,\dots,y_p}\left|\sum_{i=1}^n d_{i,n}(x)\,[\mathbb{I}\{Y_{1i}\le y_1,\dots,Y_{pi}\le y_p\}-H_i(y_1,\dots,y_p)]\right| + \frac{2}{n}.\end{aligned} \tag{C7}$$

Put $\mathbf{d}_n(x) = \sum_{i=1}^n d_{i,n}^2(x)$. Now with the help of Lemma 3 and assumption (C5)

$$\begin{aligned}\mathrm{P}&\left(\sup_{x\in \mathbf{X}_n}\sup_{y_1,\dots,y_p}\left|\sum_{i=1}^n d_{i,n}(x)\,[\mathbb{I}\{Y_{1i}\le y_1,\dots,Y_{pi}\le y_p\}-H_i(y_1,\dots,y_p)]\right| > \varepsilon\right)\\ &\le C_1 \sum_{i=1}^{m_n} \exp\left\{-\tfrac{C_2\,\varepsilon^2}{|\mathbf{d}_n(x_{i,n})|_2^2}\right\} \le 2\,C_1\,[C_0\,n^{L+1}]^{1/\alpha}\exp\left\{-\tfrac{C_2\,\varepsilon^2 n^{\delta}}{C_3}\right\},\end{aligned} \tag{C8}$$

which together with Borel-Cantelli lemma implies the almost sure convergence of the first term on left-hand side of (C7) and finishes the proof of the lemma. □

Remark 2. Let $\{I^{(n)}\}$ be a sequence of finite subsets of the interval I in Lemma 4. Then from the proof of Lemma 4 we can deduce that

$$\max_{x\in I^{(n)}}\sup_{y_1,\dots,y_p}\left|\sum_{i=1}^n d_{i,n}(x)\,[\mathbb{I}\{Y_{1i}\le y_1,\dots,Y_{pi}\le y_p\}-H_i(y_1,\dots,y_p)]\right| = o_{a.s}(1). \tag{C9}$$

provided that the condition (C4) is replaced with $\mu(I^{(n)}) \le C_0\,n^L$, where μ is a counting measure.

In random design, the following lemma is useful.

Lemma 5. *Let* $\{d_{i,n}(x), i = 1,\dots,n\}_{n=1}^{\infty}$ *be a triangular array of random functions on a finite interval* I*. Then Lemma 4 (as well as Remark 2) still holds provided that the assumption* (C5) *is reformulated in the following way: There exist finite positive constants* C_3, C_4 *and* $\delta > 0$ *such that for all* $n \in \mathbb{N}$

$$\text{(C10)} \qquad \sup_{x\in I} \mathrm{P}\left(\sum_{i=1}^{n} d_{i,n}^2(x) \geq \tfrac{C_3}{n^\delta}\right) \leq C_4 \exp\left\{-\tfrac{n^\delta}{C_2}\right\}.$$

Proof. Let us denote the event $A_n(x) = \left[\sum_{i=1}^n d_{i,n}^2(x) \leq \frac{C_3}{n^\delta}\right]$.

Conditioning on the values of $d_{i,n}(x)$, the proof goes along the lines of proof of Lemma 4. The difference is in (C8). Now, with the help of assumption (C10) we can estimate

$$\begin{aligned}
&\mathrm{P}\left(\sup_{x\in\mathbf{X}_n}\sup_{y_1,\dots,y_p}\left|\sum_{i=1}^{n} d_{i,n}(x)\left[\mathbb{I}\{Y_{1i}\leq y_1,\dots,Y_{pi}\leq y_p\} - H_i(y_1,\dots,y_p)\right]\right| > \varepsilon\right)\\
&\quad\leq \sum_{i=1}^{m_n} \mathrm{E}\left[C_1 \exp\left\{-\tfrac{C_2\,\varepsilon^2}{|\mathbf{d}_n(x_{i,n})|_2^2}\right\}\wedge 1\right]\\
&\quad\leq \sum_{i=1}^{m_n} \left[\mathrm{E}\, C_1 \exp\left\{-\tfrac{C_2\,\varepsilon^2}{|\mathbf{d}_n(x_{i,n})|_2^2}\right\}\mathbb{I}_{A_n(x)} + P(A_n^c(x))\right]\\
&\quad\leq 2\,n^{L+1}\left[C_0\, C_1 \exp\left\{-\tfrac{C_2\,\varepsilon^2 n^\delta}{C_3}\right\} + C_4 \exp\left\{-\tfrac{n^\delta}{C_3}\right\}\right]
\end{aligned}$$

and the proof is finished by applying the Borel-Cantelli lemma. □

The following lemma is useful for proving the bootstrap of the estimator $\tilde{C}_{xh}$. Let $\{\mathbf{U}_i\}_{i=1}^\infty$ be a sequence of p-dimensional independent random vectors with cumulative distribution functions $\{C_i\}_{i=1}^\infty$, whose marginals are uniform. Further consider a triangular array of p-dimensional random vectors $\{\tilde{\mathbf{U}}_i^{(n)} = (\tilde{U}_{1i}^{(n)}, \dots, \tilde{U}_{pi}^{(n)})^\mathsf{T}, i = 1, \dots, n\}_{n=1}^\infty$.

Lemma 6. *Suppose there exists a sequence of positive constants going to zero* a_n *such that*

$$\text{(C11)} \qquad \max_{j=1,\dots,p}\ \max_{i=1,\dots,n} |\tilde{U}_{ji}^{(n)} - U_{ji}| \leq O_{a.s.}(a_n).$$

Then Lemma 4 (Remark (2)*, Lemma 5) holds true if* $\{\mathbf{Y}_i\}$ *is replaced with the triangular array* $\{\tilde{\mathbf{U}}_i^{(n)}\}$ *and* $\{H_i\}$ *with* $\{C_i\}$*, provided*

$$a_n \sup_{x\in I}\sum_{i=1}^{n} |d_{i,n}(x)| = o(1) \qquad \left(a_n \sup_{x\in I}\sum_{i=1}^{n} |d_{i,n}(x)| = o_{a.s.}(1)\right).$$

Proof. With the assumptions of the lemma and the help of Lemma 4 for nonrandom $\{d_{i,n}(x)\}$ or Lemma 5 for random $\{d_{i,n}(x)\}$ we get

$$\sup_{x\in I}\sup_{u_1,\dots,u_p}\left|\sum_{i=1}^{n} d_{i,n}(x)\left[\mathbb{I}\{\tilde{U}_{1i}\leq u_1,\dots,\tilde{U}_{pi}\leq u_p\} - C_i(u_1,\dots,u_p)\right]\right|$$

$$\leq \sup_{x\in I}\sup_{u_1,\dots,u_p}\left|\sum_{i=1}^n d_{i,n}(x)\left[\mathbb{I}\{\tilde{U}_{1i}\leq u_1,\dots,\tilde{U}_{pi}\leq u_p\}-\mathbb{I}\{U_{1i}\leq u_1,\dots,U_{pi}\leq u_p\}\right]\right|$$
$$+\sup_{x\in I}\sup_{u_1,\dots,u_p}\left|\sum_{i=1}^n d_{i,n}(x)\left[\mathbb{I}\{U_{1i}\leq u_1,\dots,U_{pi}\leq u_p\}-C_i(u_1,\dots,u_p)\right]\right|$$
$$\leq \sup_{x\in I}\sup_{u_1,\dots,u_p}\left|\sum_{i=1}^n |d_{i,n}(x)|\left[\mathbb{I}\{U_{1i}\leq u_1+a_n,\dots,U_{pi}\leq u_p+a_n\}\right.\right.$$
$$\left.\left.-\mathbb{I}\{U_{1i}\leq u_1-a_n,\dots,U_{pi}\leq u_p-a_n\}\right]\right|+o_{a.s}(1)$$
$$\leq \sup_{x\in I}\sup_{u_1,\dots,u_p}\left|\sum_{i=1}^n |d_{i,n}(x)|\left[C_i(u_1+a_n,\dots,u_p+a_n)-C_i(u_1-a_n,\dots,u_p-a_n)\right]\right|+o_{a.s.}(1)$$
$$\leq 2\,p\,a_n\sup_{x\in I}\sum_{i=1}^n |d_{i,n}(x)|+o_{a.s.}(1)=o_{a.s.}(1).$$

□

Acknowledgement. The authors are grateful to the Editors and referees for their very valuable comments which led to a considerable improvement of the manuscript. This work was supported by the IAP Research Networks P6/03 and P7/13 of the Belgian State (Belgian Science Policy). This work was started while Marek Omelka was a postdoctoral researcher at the Katholieke Universiteit Leuven and the Universiteit Hasselt within the IAP Research Network. The work of the first author was supported by the grant GACR P201/11/P290. The second author acknowledges support from research grant MTM 2008-03129 of the Spanish Ministerio de Ciencia e Innovacion. The third author gratefully acknowledges support from the projects GOA/07/04 and GOA/12/014 of the Research Fund KULeuven, as well as support from the FWO-project G.0328.08N of the Flemish Science Foundation.

References

Aerts, M., Janssen, P., and Veraverbeke, N. (1994). Bootstrapping regression quantiles. *J. Nonparametr. Stat.*, 4:1–20.

Blomqvist, N. (1950). On a measure of dependence between two random variables. *Ann. Math. Statist.*, 21:593–600.

Brockmann, M., Gasser, T., and Hermann, E. (1993). Locally adaptive bandwidth choice for kernel regression estimators. *J. Amer. Statist. Assoc.*, 88:1302–1309.

Bücher, A. and Dette, H. (2010). A note on bootstrap approximations for the empirical copula process. *Statist. Probab. Lett.*, 80:1925–1932.

Fan, J. and Gijbels, I. (1996). *Local Polynomial Modelling and Its Applications.* Chapman & Hall.

Fermanian, J.-D., Radulovič, D., and Wegkamp, M. (2004). Weak convergence of empirical copula processes. *Bernoulli*, 10:847–860.

Gasser, T., Kneip, A., and Kohler, W. (1991). A flexible and fast method for automatic smoothing. *J. Amer. Statist. Assoc.*, 86:643–652.

Gasser, T. and Müller, H.-G. (1979). Kernel estimates of regression functions. In Gasser, T. and Rosenblatt, M., editors, *Lecture Notes in Mathematics 757*, pages 23–68. Springer, New York.

Gijbels, I., Veraverbeke, N., and Omelka, M. (2011). Conditional copulas, association measures and their application. *Comput. Stat. Data An.*, 55:1919–1932.

Härdle, W. and Bowman, A. W. (1988). Bootstrapping in nonparametric regression: Local adaptive smoothing and confidence bands. *J. Amer. Statist. Assoc.*, 83:102–110.

Müller, H.-G. (1987). Weighted local regression and kernel methods for nonparametric curve fitting. *J. Amer. Statist. Assoc.*, 82:231–238.

Nadaraya, E. A. (1964). On estimating regression. *Theory Probab. Appl.*, 9:141–142.

Nelsen, R. B. (2006). *An Introduction to Copulas.* Springer, New York. Second Edition.

Nolan, D. and Pollard, D. (1987). U-processes: Rates of convergence. *Ann. Statist.*, 15:780–799.

Omelka, M., Gijbels, I., and Veraverbeke, N. (2009). Improved kernel estimation of copulas: Weak convergence and goodness-of-fit testing. *Ann. Statist.*, 37(5B):3023–3058.

R Development Core Team (2009). *R: A Language and Environment for Statistical Computing.* R Foundation for Statistical Computing, Vienna, Austria.

Segers, J. (2012). Weak convergence of empirical copula processes under nonrestrictive smoothness assumptions. *Bernoulli*, 18:764–782.

van de Geer, S. A. (2000). *Empirical Processes in M-Estimation.* Cambridge University Press, New York.

van der Vaart, A. W. and Wellner, J. A. (1996). *Weak Convergence and Empirical Processes.* Springer, New York.

van der Vaart, A. W. and Wellner, J. A. (2007). Empirical processes indexed by estimated functions. In Cator, E. A., Jongbloed, G., Kraaikamp, C., Lopuhaä, H. P., and Wellner, J. A., editors, *IMS Lecture Notes Monograph Series 2007: Asymptotics: Particles, Processes and Inverse Problems*, volume 55, pages 234–252. Institute of Mathematical Statistics.

Veraverbeke, N., Omelka, M., and Gijbels, I. (2011). Estimation of a conditional copula and association measures. *Scand. J. Statist.*, 38:766–780.

Watson, G. S. (1964). Smooth regression analysis. *Sankhyā A*, 26:359–372.