

Partial and average copulas and association measures

Irène Gijbels

*Department of Mathematics and Leuven Statistics Research Center (LStat),
KU Leuven, Belgium*

e-mail: Irene.Gijbels@wis.kuleuven.be

Marek Omelka

*Department of Probability and Statistics, Faculty of Mathematics and Physics,
Charles University in Prague, Czech Republic*

e-mail: omelka@karlin.mff.cuni.cz

and

Noël Veraverbeke

*Center for Statistics, Hasselt University, Belgium
and*

Unit for BMI, North-West University, Potchefstroom, South Africa

e-mail: noel.veraverbeke@uhasselt.be

Abstract: For a pair (Y_1, Y_2) of random variables there exist several measures of association that characterize the dependence between Y_1 and Y_2 by means of one single value. Classical examples are Pearson's correlation coefficient, Kendall's tau and Spearman's rho. For the situation where next to the pair (Y_1, Y_2) there is also a third variable X present, so-called partial association measures, such as a partial Pearson's correlation coefficient and a partial Kendall's tau, have been proposed in the 1940's. Following criticism on e.g. partial Kendall's tau, better alternatives to these original partial association measures appeared in the literature: the conditional association measures, e.g. conditional Kendall's tau, and conditional Spearman's rho. Both, unconditional and conditional association measures can be expressed in terms of copulas. Even in case the dependence structure between Y_1 and Y_2 is influenced by a third variable X , we still want to be able to summarize the level of dependence by one single number. In this paper we discuss two different ways to do so, leading to two relatively new concepts: the (new concept of) partial Kendall's tau, and the average Kendall's tau. We provide a unifying framework for the diversity of concepts: global (or unconditional) association measures, conditional association measures, and partial and average association measures. The main contribution is that we discuss estimation of the newly-defined concepts: the partial and average copulas and association measures, and establish theoretical results for the estimators. The various concepts of association measures are illustrated on a real data example.

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

Keywords and phrases: Average copula, conditional copula, empirical copula process, nonparametric estimation, partial copula, unconditional copula, smoothing, weak convergence.

Received October 2014.

Contents

1	Introduction	2422
2	Various concepts of association measures, defined in terms of copulas	2425
2.1	Unconditional (global) and conditional copulas and association measures	2425
2.2	Partial and average copulas and association measures	2426
2.3	Simplified pair-copula construction	2429
2.4	Summary: Various concepts of association measures	2429
3	Estimation of copulas and association measures	2430
3.1	Unconditional (global) Kendall's tau	2431
3.2	Conditional and average conditional Kendall's tau	2431
3.3	Partial Kendall's tau	2432
3.4	Some standard methods of adjustments	2433
3.4.1	Parametric location-scale model estimation of F_{1x} and F_{2x}	2433
3.4.2	Nonparametric location-scale model estimation of F_{1x} and F_{2x}	2433
3.4.3	General nonparametric estimation of F_{1x} and F_{2x}	2434
4	Application	2434
5	Theoretical results	2437
5.1	Asymptotic results for the nonparametric partial copula estimator	2438
5.1.1	Parametric location-scale adjustments	2439
5.1.2	Nonparametric location-scale adjustments	2440
5.1.3	General nonparametric adjustments	2440
5.2	Asymptotic results for the estimator of the partial Kendall's tau	2441
5.2.1	Parametric location-scale adjustments	2441
5.2.2	Nonparametric location-scale adjustments	2441
5.2.3	General nonparametric adjustments	2442
5.3	Asymptotic results for the estimator of the average conditional Kendall's tau	2442
6	Conclusion and Discussion	2443
Appendix A:	(Un)conditional, average and partial association measures	2444
A.1	Global or unconditional association measures	2444
A.2	On some <i>original</i> partial association measures	2445
A.3	Conditional association measures	2446
A.4	On average conditional and partial association measures	2447
Appendices B–F:	Assumptions, proofs and auxiliary results	2448
Appendix B:	Adjusting through parametric location-scale models	2449

Appendix C: Adjusting through nonparametric location-scale models . . .	2454
Appendix D: Adjusting through general nonparametric estimators of F_{1x} and F_{2x}	2458
Appendix E: Asymptotic result for the average conditional Kendall's tau	2458
Appendix F: Auxiliary results	2464
Acknowledgements	2472
References	2472

1. Introduction

Suppose we observe a random vector (Y_1, Y_2) . In statistics we often need to characterize the degree of dependence of Y_1 and Y_2 . The most standard (and probably also the oldest) measure of dependence is Pearson's correlation coefficient

$$\rho^{(P)}(Y_1, Y_2) = \frac{\text{cov}(Y_1, Y_2)}{\sqrt{\text{var}(Y_1) \text{var}(Y_2)}} = \frac{\mathbf{E}(Y_1 Y_2) - \mathbf{E}(Y_1) \mathbf{E}(Y_2)}{\sqrt{\text{var}(Y_1) \text{var}(Y_2)}}, \quad (1.1)$$

which proved to be useful in many situations. In particular, if (Y_1, Y_2) follows a bivariate normal distribution, then $\rho^{(P)}(Y_1, Y_2)$ completely characterizes the dependence structure of (Y_1, Y_2) . On the other hand, $\rho^{(P)}(Y_1, Y_2)$ can be of little use if the bivariate distribution of (Y_1, Y_2) is far away from being normal. Moreover, $\rho^{(P)}$ is even not defined if the distribution of (Y_1, Y_2) does not have finite and positive variances of Y_1 and Y_2 . That is why alternative measures of dependence have been introduced.

Among the most popular measures of dependence are Kendall's tau and Spearman's rho. See Section A.1 for a brief recall of their definitions and an overview of other commonly-used association measures.

The situation becomes more difficult when we observe a three-dimensional vector (Y_1, Y_2, X) and one is interested in the relationship between Y_1 and Y_2 when the effect of X is taken into consideration. A simple concept which has proved to be useful in many situations is that of the *original* partial Pearson's correlation coefficient given by

$$\rho_X^{(P)}(Y_1, Y_2) = \frac{\rho^{(P)}(Y_1, Y_2) - \rho^{(P)}(Y_1, X) \rho^{(P)}(Y_2, X)}{\sqrt{1 - (\rho^{(P)}(Y_1, X))^2} \sqrt{1 - (\rho^{(P)}(Y_2, X))^2}}. \quad (1.2)$$

In Section 2 we will introduce a new concept of partial association measures, not to be confused with this *original* partial correlation coefficient.

Similarly as for the global Pearson's correlation coefficient $\rho^{(P)}(Y_1, Y_2)$ defined in (1.1), the partial Pearson's correlation coefficient only completely characterizes the dependence structure of Y_1 and Y_2 taking into account X , if (Y_1, Y_2, X) has a trivariate normal distribution. Among the earliest attempts to get away from this normality assumption, is the concept of the *original* partial Kendall's tau (see (A.4)). However some criticism were formulated regarding this measure. See Section A.2 for a brief recall and some examples to illustrate the criticism.

A more comprehensive and detailed characterization of the dependence structure is provided by conditional measures of dependency/association, that measure the dependence structure in (Y_1, Y_2) conditionally upon an event formulated in terms of X . The simplest (and most common) conditional setting is to consider dependency in (Y_1, Y_2) given the event $X = x$ (i.e. a value x taken by the covariate). In Section A.3 we briefly recall such conditional association measures, and provide an example to illustrate their merits when compared to the *original* partial type of association measures. If one is in particular interested in high value settings of the covariate one might consider looking at the conditioning event that $X \geq x$ (or conversely $X \leq x$). This is for example often the case in economic (e.g. production frontier) or actuarial applications. In Section 4 an example of such a conditioning event is included. Although the presentation in this paper almost entirely focuses on conditioning upon the event $X = x$, the concepts and methodology apply to general events in terms of the covariate X .

Conditional association measures thus quantify clearly how the dependence structure between the two components in (Y_1, Y_2) changes in terms of (the event related to) X . Graphically such a conditional association measure is depicted as a function of X . Of interest is then to look into issues of average (or alternatively, for example, median) strength of dependence. Furthermore, one might want to quantify the differences in dependence structures within (Y_1, Y_2) and (V_1, V_2) , conditionally upon a similar event related to the same covariate X . Comparing the strengths of the two dependence structures, taken into account the behaviour of the common covariate, then translates into comparing two curves. A first approach to do so is to look into a kind of global (mean) behaviour of the curves.

In summary, the aim of the paper is to provide some insights in different ways to study such a global/mean behaviour of conditional dependencies. We discuss two approaches to do this, leading to the concepts of partial and average conditional association measures. A unifying framework for our study is provided by focusing on association measures that can be expressed as a functional of the copula function C (assumed to be unique), denoted as $\varphi(C)$. See Table 1 for some commonly-used association measures in this class. In case of the conditional dependence (given X) one has to deal with a conditional copula function C_X , leading to the corresponding conditional association measure $\varphi(C_X)$. A first approach towards a global/mean behaviour of the conditional dependencies is to take the average (with respect to X) of this conditional association measure, i.e. $E_X\{\varphi(C_X)\}$. We refer to this as the *average* conditional association measure. In a second approach one starts from the so-called partial copula, defined by $E_X\{C_X(\cdot, \cdot)\}$ and denoted by $\bar{C}(\cdot, \cdot)$; and then considers the corresponding association measure $\varphi(\bar{C})$. This is referred to as the partial association measure. Table 1 (third column) gives an overview of some average and partial association measures. We show that for most, but not all, conditional association measures these approaches coincide. An interesting case where they do not coincide is Kendall's tau. A second contribution of this paper consists of discussing estimation of the partial and average association measures. A crucial starting point for this is the estimation of C_X and \bar{C} ; that of C_X has been

TABLE 1
 Overview of some (un)conditional association measures and their average and partial versions. *Measures of tail dependence as introduced in Schmid and Schmidt (2007) (abbreviated as S.S. (2007) below)

Type	unconditional association measure in terms of C functional $\varphi(C)$	conditional, average, and partial versions	Facts
Spearman's rho	$\rho = 12 \iint C(u_1, u_2) du_1 du_2 - 3$	$\rho(x) = 12 \iint C_x(u_1, u_2) du_1 du_2 - 3$ $\rho^A = \mathbf{E}_X \rho(X)$ $\bar{\rho} = 12 \iint \bar{C}(u_1, u_2) du_1 du_2 - 3$	$\rho^A = \bar{\rho}$
Kendall's tau	$\tau = 4 \iint C(u_1, u_2) dC(u_1, u_2) - 1$	$\tau(x) = 4 \iint C_x(u_1, u_2) dC_x(u_1, u_2) - 1$ $\tau^A = \mathbf{E}_X \tau(X)$ $\bar{\tau} = 4 \iint \bar{C}(u_1, u_2) d\bar{C}(u_1, u_2) - 1$	in general: $\tau^A \neq \bar{\tau}$ Example A.3 (Appendix A)
Blomqvist's beta	$\beta = 4C(0.5, 0.5) - 1$	$\beta(x) = 4C_x(0.5, 0.5) - 1$ $\beta^A = \mathbf{E}_X \beta(X)$ $\bar{\beta} = 4\bar{C}(0.5, 0.5) - 1$	$\beta^A = \bar{\beta}$
Gini index	$\gamma = 4 \int C(u, 1-u) du - 4 \int [u - C(u, u)] du$	$\gamma(x) = 4 \int C_x(u, 1-u) du - 4 \int [u - C_x(u, u)] du$ $\gamma^A = \mathbf{E}_X \gamma(X)$ $\bar{\gamma} = 4 \int \bar{C}(u, 1-u) du - 4 \int [u - \bar{C}(u, u)] du$	$\gamma^A = \bar{\gamma}$
Tail coefficients	$\lambda_L = \lim_{t \rightarrow 0^+} \frac{C(t, t)}{t}$ $\lambda_U = \lim_{t \rightarrow 1^-} \frac{1 - 2t + C(t, t)}{1 - t}$	$\lambda_L(x) = \lim_{t \rightarrow 0^+} \frac{C_x(t, t)}{t}$ $\lambda_L^A = \mathbf{E}_X \lambda_L(X)$ $\bar{\lambda}_L = \lim_{t \rightarrow 0^+} \frac{\bar{C}(t, t)}{t}$ analogously	$\lambda_L^A = \bar{\lambda}_L$ Proposition A.1 (Appendix A) $\lambda_U^A = \bar{\lambda}_U$
S.S. (2007) Tail coefficients*	$\rho_L = \lim_{t \rightarrow 0^+} \frac{3}{t^3} \int_0^t \int_0^t C(u_1, u_2) du_1 du_2$ $\rho_U = \lim_{t \rightarrow 1^-} \frac{3}{(1-t)^3} \times \int_t^1 \int_t^1 [1 - u_1 - u_2 + C(u_1, u_2)] du_1 du_2$	$\rho_L(x) = \lim_{t \rightarrow 0^+} \frac{3}{t^3} \int_0^t \int_0^t C_x(u_1, u_2) du_1 du_2$ $\rho_L^A = \mathbf{E}_X \rho_L(X)$ $\bar{\rho}_L = \lim_{t \rightarrow 0^+} \frac{3}{t^3} \int_0^t \int_0^t \bar{C}(u_1, u_2) du_1 du_2$ analogously	$\rho_L^A = \bar{\rho}_L$ Proposition A.1 (Appendix A) $\rho_U^A = \bar{\rho}_U$

studied in recent literature, whereas that of \bar{C} is part of the contribution in this paper. For the most interesting case of partial and average Kendall's tau we also establish the asymptotic behaviour of the estimators.

The paper is further organized as follows. In Section 2 we discuss briefly the unifying framework for the various concepts of association measures (such as unconditional and conditional), and introduce the two fairly-new concepts of association measures. Section 3 discusses estimation of the various concepts, and for the proposed estimators of the average copula and the average and partial Kendall's tau, we establish the asymptotic properties in Section 5. The use of the various concepts of association measures is illustrated on a real data application in Section 4. Some conclusions and discussions are in Section 6. Appendix A provides a brief review on association measures, their specific drawbacks and merits. The proofs of the theoretical results, the assumptions under which these hold, as well as some needed auxiliary results are provided in Appendices B–F.

2. Various concepts of association measures, defined in terms of copulas

Many measures of association can be expressed as functionals of copulas, which link the marginal distributions into the joint distribution. This unifying framework, together with different conceptual notions of copulas, allows to provide a unified approach towards various concepts of association measures. In Section 2.1 we briefly review existing concepts, whereas in Section 2.2 we introduce new concepts, all in the same unifying framework.

2.1. Unconditional (global) and conditional copulas and association measures

To formalize the definition of a (unconditional) copula let $H(y_1, y_2)$ be the joint distribution function of the random vector (Y_1, Y_2) and denote by F_{Y_1} and F_{Y_2} the marginal distribution functions of Y_1 and Y_2 respectively. Then a copula C_{Y_1, Y_2} on $[0, 1]^2$ is a function such that

$$H(y_1, y_2) = C_{Y_1, Y_2}(F_{Y_1}(y_1), F_{Y_2}(y_2)), \quad (y_1, y_2) \in \mathbb{R}^2.$$

In case of continuous marginal distribution functions F_{Y_1} and F_{Y_2} , the copula function C_{Y_1, Y_2} is uniquely defined. See Nelsen (2006).

Many association measures can be expressed as specific functionals of C_{Y_1, Y_2} , say as $\varphi(C_{Y_1, Y_2})$ or, shortly, as $\varphi(C)$. For example, Kendall's tau and Spearman's rho, are given by

$$\begin{aligned} \tau(Y_1, Y_2) &= 4 \iint C_{Y_1, Y_2}(u_1, u_2) dC_{Y_1, Y_2}(u_1, u_2) - 1 \\ \rho^{(S)}(Y_1, Y_2) &= 12 \iint C_{Y_1, Y_2}(u_1, u_2) du_1 du_2 - 3. \end{aligned} \quad (2.1)$$

Table 1 lists other association measures indicating the specific functional $\varphi(\cdot)$. These include also the lower and upper tail coefficients (denoted with λ_L and λ_U) and other association measures focusing on tail behaviour, such as these

introduced by Schmid and Schmidt (2007). For a detailed study of association measures see Chapter 5 of Nelsen (2006).

In the literature so far, one has studied unconditional copulas as well as conditional copulas. The latter concept was introduced in Patton (2006), and serves to study the conditional dependence structure of Y_1 and Y_2 given $X = x$ (as the simplest conditioning event). Denote the joint and marginal distribution functions of (Y_1, Y_2) , conditionally upon $X = x$, as

$$H_x(y_1, y_2) = \mathbb{P}(Y_1 \leq y_1, Y_2 \leq y_2 \mid X = x),$$

$$F_{1x}(y_1) = \mathbb{P}(Y_1 \leq y_1 \mid X = x), \quad F_{2x}(y_2) = \mathbb{P}(Y_2 \leq y_2 \mid X = x).$$

If F_{1x} and F_{2x} are continuous, then according to Sklar's theorem (see e.g. Nelsen, 2006) there exists a unique copula C_x which links the conditional marginals into the conditional joint distribution

$$H_x(y_1, y_2) = C_x(F_{1x}(y_1), F_{2x}(y_2)). \quad (2.2)$$

The function C_x fully describes the conditional dependence structure of the bivariate vector (Y_1, Y_2) given $X = x$ and it is called a *conditional copula*.

As discussed in Gijbels, Veraverbeke and Omelka (2011) and Veraverbeke, Omelka and Gijbels (2011) the conditional measures of association that do not depend on the marginal distributions of Y_1 and Y_2 can be written as functionals of C_x . For instance the conditional Kendall's tau defined in (A.6) can be expressed as

$$\tau(x) = 4 \iint C_x(u_1, u_2) dC_x(u_1, u_2) - 1.$$

Similarly, the conditional Spearman's rho is given by

$$\rho^{(S)}(x) = 12 \iint C_x(u_1, u_2) du_1 du_2 - 3.$$

Other conditional association measures, including conditional tail coefficients are given in the first rows in each block of column 3 of Table 1.

2.2. Partial and average copulas and association measures

Conditional association measures (or more generally conditional copulas) are very useful when one wants to get a deeper insight into the dependence structure and how it changes with the covariate X . However, it still might be of interest to summarize/capture the strength of this dependence with one single number. Indeed, in case of two random vectors (Y_1, Y_2, X) and (V_1, V_2, X) , such a global number would allow us to make simple comparisons of the strengths of the dependencies between Y_1 and Y_2 on the one hand and these between V_1 and V_2 on the other hand, taking into account the covariate X .

We thus need one number (one copula) summarizing the dependence of Y_1 and Y_2 when adjusted for X . We now discuss two approaches to get to such a summarizing type of copula.

A first obvious idea is to average the conditional measures with respect to the distribution of X and to get the *average conditional copula*

$$C^A(u_1, u_2) = \mathbf{E}_X C_X(u_1, u_2), \tag{2.3}$$

and, for example, *average (conditional) Kendall's tau*

$$\tau^A = \mathbf{E}_X \tau(X),$$

or the *average (conditional) lower tail coefficient* (see Table 1)

$$\lambda_L^A = \mathbf{E}_X \lambda_L(X).$$

The concept of average conditional copula was first mentioned by Bergsma (2011, 2004).

Another way is to follow the original idea that a partial correlation coefficient is supposed to *measure correlation of Y_1 and Y_2 after removal of any part of the variation due to the influence of X* , (see e.g. p. 306 of Cramér, 1946). The most general way of removing the effect of X on Y_1 and on Y_2 is through their conditional (marginal) distribution functions, which results into $U_1 = F_{1X}(Y_1)$ and $U_2 = F_{2X}(Y_2)$. Note that neither U_1 nor U_2 depends on X any more, and both are uniformly distributed (due to the probability integral transform). Indeed, for example, for all $t \in [0, 1]$,

$$\begin{aligned} \mathbf{P}\{F_{1X}(Y_1) \leq t\} &= \int \mathbf{P}\{F_{1x}(Y_1) \leq t | X = x\} dF_X(x) \\ &= \int \mathbf{P}\{Y_1 \leq F_{1x}^{-1}(t) | X = x\} dF_X(x) \\ &= \int F_{1x}(F_{1x}^{-1}(t)) dF_X(x) = t, \end{aligned}$$

with F_X the cumulative distribution function of X . See also Song (2009) who exploits this transformation in the problem of testing for conditional independence.

So, after having removed the effect of X on the marginal distributions, the dependence structure of the transformed random variables is fully described by the copula function \tilde{C} corresponding to the pair (U_1, U_2) . We will call this the *partial copula*. See also Definition 3 of Bergsma (2004). As the marginals of (U_1, U_2) are already uniform, \tilde{C} coincides with the joint distribution function of (U_1, U_2) : the *partial copula* is defined by

$$\tilde{C}(u_1, u_2) = \mathbf{P}(U_1 \leq u_1, U_2 \leq u_2) \quad \text{with} \quad (U_1, U_2) = (F_{1X}(Y_1), F_{2X}(Y_2)). \tag{2.4}$$

Joe (2006) builds on partial correlations to generate random correlation matrices. The paper uses a vine decomposition to access the joint density of pairwise correlations. Bedford and Cooke (2002) introduced the concept of vines for dependent random variables. In Gaussian copulas commonly-used association measures such as Kendall's tau, Blomqvist's beta, Spearman's rho and Gini's

index can all be expressed in terms of Pearson’s correlation coefficient, whereas the upper and lower tail coefficients are zero. Kim et al. (2011) studied partial correlation assuming a Gaussian copula for \bar{C} .

The notions of average (conditional) copula and partial copula, defined in respectively (2.3) and (2.4), in fact coincide, as is stated and proved in Proposition 1.

Proposition 1. *For random variables Y_1 and Y_2 with continuous distribution functions, it holds that*

$$\bar{C}(u_1, u_2) = C^A(u_1, u_2) \quad \forall (u_1, u_2) \in [0, 1]^2.$$

Proof. This is straightforward since

$$\begin{aligned} \bar{C}(u_1, u_2) &= P(U_1 \leq u_1, U_2 \leq u_2) = \int P(U_1 \leq u_1, U_2 \leq u_2 \mid X = x) dF_X(x) \\ &= \int C_x(u_1, u_2) dF_X(x) = C^A(u_1, u_2). \quad \square \end{aligned}$$

Thus the copula of Y_1 and Y_2 after removal of the effect of X on the marginal distributions, coincides with the average conditional copula function. In other words, there are two ways of viewing the copula \bar{C} : it is the copula describing the dependence between Y_1 and Y_2 after removal of the effect of X ; but also the copula obtained after taking the expectation (with respect to the covariate X) of the conditional copula.

We can now think of considering association measures derived from the partial copula \bar{C} . We call these measures the *partial association measures*. For instance the partial Kendall’s tau is given by

$$\begin{aligned} \bar{\tau} = \tau(U_1, U_2) &= 2 P[(U_1 - U'_1)(U_2 - U'_2) > 0] - 1 \\ &= 4 \iint \bar{C}(u_1, u_2) d\bar{C}(u_1, u_2) - 1, \end{aligned} \tag{2.5}$$

where (U'_1, U'_2) is an independent copy of the random vector (U_1, U_2) , defined in (2.4). Similarly, the partial Spearman’s rho is defined by

$$\bar{\rho}^{(S)} = 12 \iint \bar{C}(u_1, u_2) du_1 du_2 - 3.$$

The partial measure $\bar{\tau}$ should not be confused with the *original* partial Kendall’s tau given in (A.4).

Note that $\bar{C} = C^A$ does not imply that the average conditional measures, obtained by averaging the conditional measures with respect to X , equal the partial measures. In general this holds true only when C_x does not depend on x (see also Section 2.3 below) or if the measure of association is a linear functional of the underlying copula. Thus while $\bar{\rho}^{(S)} = \rho^{(S),A} = E_X \rho^{(S)}(X)$, in general $\bar{\tau} \neq \tau^A$ as

$$\tau^A = E_X \tau(X) = 4 \int_{R_X} \iint_{[0,1]^2} C_x(u_1, u_2) dC_x(u_1, u_2) dF_X(x) - 1, \tag{2.6}$$

TABLE 2
Different concepts of copulas and association Kendall's measure

Concepts	notations		
	copula	population quantities association measure	nonparametric estimators
unconditional copula unconditional Kendall's tau	C_{Y_1, Y_2}	$\tau(Y_1, Y_2)$	$\tau_n(Y_1, Y_2)$
<i>original</i> partial Kendall's tau		$\bar{\tau}^K = \tau_X(Y_1, Y_2)$	$\bar{\tau}_n^K$
conditional copula conditional Kendall's tau	C_x	$\tau(x) = \tau(Y_1, Y_2 X = x)$	$\tau_n(x)$
partial copula partial Kendall's tau	\bar{C}	$\bar{\tau}$	$\bar{\tau}_n$
average (conditional) copula average (conditional) Kendall's tau	$C^A (= \bar{C})$	$\tau^A = E_X[\tau(X)]$	τ_n^A

and

$$\bar{\tau} = 4 \int_{R_X} \int_{R_X} \iint_{[0,1]^2} C_x(u_1, u_2) dC_{x'}(u_1, u_2) dF_X(x) dF_X(x') - 1.$$

See also Section A.4 in Appendix A for an example where $\bar{\tau} \neq \tau^A$.

Nevertheless, for many association measures the functional $\varphi(C)$ constitutes a linear functional in C , and hence the equality $E_X\{\varphi(C_X)\} = \varphi(\bar{C})$ is rather evident. For other association measures such as the upper and lower tail coefficients that involve limit expressions, Proposition A.1 establishes the coincidence.

2.3. Simplified pair-copula construction

Sometimes, it is reasonable to expect that the covariate X only affects the marginal distributions of Y_1 and Y_2 , but does not affect the dependence structure. This results in the conditional joint distribution of (Y_1, Y_2) given by

$$H_x(y_1, y_2) = C(F_{1x}(y_1), F_{2x}(y_2)). \tag{2.7}$$

This is also called the simplified pair-copula construction in the recent literature. See e.g. Hobæk Haff, Aas and Frigessi (2010) and Acar, Genest and Nešlehová (2012). Note that in model (2.7) the conditional copula C_x does not depend on x (i.e. $C_x = C$), which is in contrast to the general model (2.2). Hence, in this special setting, the conditional and the partial copula coincide ($C = \bar{C}$) and also all the three types of association measures (conditional, average conditional, partial) coincide, i.e. for the Kendall's type of association measures:

$$\tau(x) = \tau^A = \bar{\tau}.$$

2.4. Summary: Various concepts of association measures

Table 2 gives a summary of the different notions of copulas and corresponding association measures (focusing for ease of presentation only on Kendall's tau

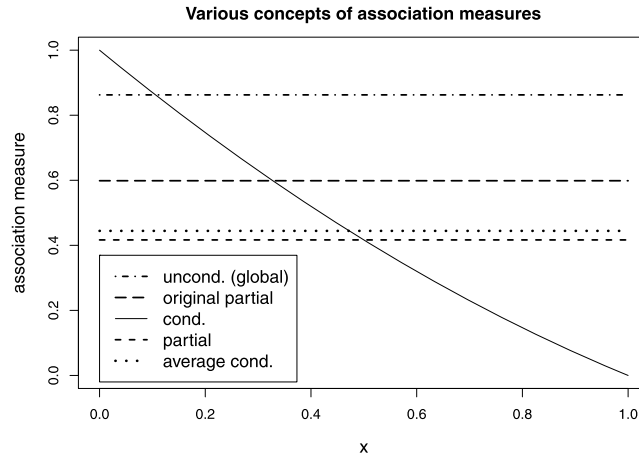


FIG 1. Example A.3: unconditional (global) Kendall's tau (horizontal dashed-dotted line), the original partial Kendall's tau (horizontal long-dashed line), the conditional Kendall's tau (solid line), the partial Kendall's tau (horizontal short-dashed line), and the average conditional Kendall's tau (horizontal dotted line).

types of association measures), with their respective notations. The entries in the last column will be discussed in Section 3.

In Figure 1 we depict the unconditional, *original* partial, the conditional, the partial and the average conditional Kendall's tau for Example A.3.

3. Estimation of copulas and association measures

Suppose we have independent random vectors $(Y_{11}, Y_{21}, X_1), \dots, (Y_{1n}, Y_{2n}, X_n)$, all having the same distribution as (Y_1, Y_2, X) . To illustrate the estimation of (un)conditional, average and partial association measures, we concentrate on estimation of different notions of Kendall's tau. The various notions of other association measures (Spearman's rho, Gini coefficient, Blomqvist's beta, upper and lower tail coefficients, ...) can be estimated analogously.

A crucial point is that all these different notions of association measures (unconditional or global, conditional, partial and average) can be expressed as functionals of the corresponding notion of copula (i.e. $\varphi(C_X)$ or $\varphi(\bar{C})$). Hence, plugging in an appropriate nonparametric estimator for the specific copula function (C_X or \bar{C}) into these expressions, leads directly to a nonparametric estimator for the specific notion of the association measure. Nonparametric estimation of unconditional and conditional copulas has been studied in the (recent) literature, whereas nonparametric estimation of the partial copula is largely unexplored. In Section 5 we discuss and study a nonparametric estimator for the partial copula. In this paper we focus on kernel type of estimation, for two reasons: (i) to be able to rely on results available in the literature on nonparametric estimation of a conditional copula; (ii) since all estimators have explicit forms this allows

to establish asymptotic results. Obviously alternative flexible estimation methods such as spline basis expansions (see e.g. Kauermann and Schellhase, 2014) and/or Bayesian methods (see e.g. Burda and Prokhorov, 2014) could also be applied.

In the sequel of this section, we immediately focus, for brevity and clarity, on the estimators of the association measures, resulting from the above plug-in step involving nonparametric estimation of the appropriate notion of copula.

3.1. Unconditional (global) Kendall's tau

Recall expression (2.1). Several nonparametric estimators for the unconditional copula C_{Y_1, Y_2} are available in the literature. The simplest estimator is the empirical copula function

$$C_n(u_1, u_2) = \frac{1}{n} \sum_{i=1}^n \mathbb{I}\{F_{1n}(Y_{1i}) \leq u_1, F_{2n}(Y_{2i}) \leq u_2\},$$

where F_{jn} , for $j = 1, 2$, is the empirical cumulative distribution function based on the observations Y_{j1}, \dots, Y_{jn} . This estimator was introduced and studied by Deheuvels (1979), and subsequently studied by Gänssler and Stute (1987), Fermaian, Radulović and Wegkamp (2004), Tsukahara (2005) and Segers (2012), among others. Kernel estimators for the unconditional copula are studied in Gijbels and Mielniczuk (1990), Chen and Huang (2007), and Omelka, Gijbels and Veraverbeke (2009), among others.

Substituting this estimator into the expression (2.1) leads to an estimator for $\tau(Y_1, Y_2)$. The obtained estimator is first order asymptotically equivalent with the estimator

$$\tau_n = \tau_n(Y_1, Y_2) = \frac{4}{n(n-1)} \sum_{i=1}^n \sum_{j=1}^n \mathbb{I}\{Y_{1i} < Y_{1j}, Y_{2i} < Y_{2j}\} - 1. \tag{3.1}$$

See Nelsen (2006). Subsequently, one can estimate the original partial Kendall's tau by Kendall (1942) simply via

$$\bar{\tau}_n^K = \frac{\tau_n(Y_1, Y_2) - \tau_n(Y_1, X) \tau_n(Y_2, X)}{\sqrt{1 - \tau_n^2(Y_1, X)} \sqrt{1 - \tau_n^2(Y_2, X)}}.$$

3.2. Conditional and average conditional Kendall's tau

Note that by mimicking formula (3.1) it would be possible to estimate the conditional Kendall's tau through the formula

$$\begin{aligned} \tau_n(x) &= \frac{4}{1 - \sum_{i=1}^n w_{ni}^2(x, h_n)} \\ &\times \sum_{i=1}^n \sum_{j=1}^n w_{ni}(x, h_n) w_{nj}(x, h_n) \mathbb{I}\{Y_{1i} < Y_{1j}, Y_{2i} < Y_{2j}\} - 1, \end{aligned} \tag{3.2}$$

where $\{w_{ni}(x, h_n)\}$ is a sequence of weights that smooth over the covariate space. But as discussed in Veraverbeke, Omelka and Gijbels (2011) it is better to replace the original observations (Y_{1i}, Y_{2i}) in formula (3.2) with observations that are already adjusted for the effect of the covariate X . For detailed discussions on nonparametric estimation of conditional copulas see Veraverbeke, Omelka and Gijbels (2011) and Gijbels, Veraverbeke and Omelka (2011), among others.

The method that is used to remove the effect of X on the marginal distributions of Y_1 and Y_2 depends on what can be assumed about this effect (see Section 3.4). In general let $G_j(y, x)$ stand for the transformation that removes the effect of X on Y_j . Generally, $G_j(y, x) = F_{jx}(y)$ does the job, but sometimes simpler functions (not requiring nonparametric estimation of F_{jx} , and hence the introduction of an additional smoothing parameter) are available. For instance in Example A.1 of Appendix A one can use $G_1(Y_1, X) = Y_1 - X$ and $G_2(Y_2, X) = Y_2 - X^2$. Further, let \hat{G}_{jn} stand for an estimate of the function G_j . Then the adjusted observations are given by

$$(\hat{Y}_{1i}^a, \hat{Y}_{2i}^a) = (\hat{G}_{1n}(Y_{1i}, X_i), \hat{G}_{2n}(Y_{2i}, X_i)), \quad i = 1, \dots, n. \quad (3.3)$$

The estimate of the conditional Kendall's tau is then given by

$$\begin{aligned} \tau_n(x) &= \frac{4}{1 - \sum_{i=1}^n w_{ni}^2(x, h_n)} \\ &\quad \times \sum_{i=1}^n \sum_{j=1}^n w_{ni}(x, h_n) w_{nj}(x, h_n) \mathbb{I}\{\hat{Y}_{1i}^a < \hat{Y}_{1j}^a, \hat{Y}_{2i}^a < \hat{Y}_{2j}^a\} - 1. \end{aligned} \quad (3.4)$$

An estimator of the *average* (conditional) Kendall's tau $\tau^A = \mathbb{E}_X \tau(X)$ is now simply

$$\tau_n^A = \frac{1}{n} \sum_{i=1}^n \tau_n(X_i), \quad (3.5)$$

where $\tau_n(X_i)$ is the estimate (3.4) evaluated at the point X_i .

3.3. Partial Kendall's tau

The population version of partial Kendall's tau was introduced in (2.5). With the help of the adjusted observations given by (3.3) one can estimate $\bar{\tau}$ by

$$\bar{\tau}_n = \frac{4}{n(n-1)} \sum_{i=1}^n \sum_{j=1}^n \mathbb{I}\{\hat{Y}_{1i}^a < \hat{Y}_{1j}^a, \hat{Y}_{2i}^a < \hat{Y}_{2j}^a\} - 1. \quad (3.6)$$

Note that while both $\bar{\tau}$ and τ^A are well defined and reasonable summaries of the dependence of Y_1 and Y_2 when adjusted for X , the advantage of $\bar{\tau}$ is that its estimator $\bar{\tau}_n$ given by (3.6) is less computationally intensive than τ_n^A . On the other hand we establish an asymptotic normality result for τ_n^A under more general assumptions than for $\bar{\tau}_n$ (see Section 5).

3.4. Some standard methods of adjustments

In this section we list some appealing methods for adjusting the observations for the effect of the covariate.

3.4.1. Parametric location-scale model estimation of F_{1x} and F_{2x}

Consider the following model

$$Y_1 = m_1(X, \beta_1) + \sigma_1(X, \gamma_1) \varepsilon_1, \quad Y_2 = m_2(X, \beta_2) + \sigma_2(X, \gamma_2) \varepsilon_2,$$

where $m_1, m_2, \sigma_1, \sigma_2$ are known functions, $\beta_1, \beta_2, \gamma_1, \gamma_2$ are unknown finite-dimensional parameters and ε_1 and ε_2 are independent of X with unknown distribution functions $F_{1\varepsilon}$ and $F_{2\varepsilon}$.

Note that the ‘ideal’ transformation function would be given by the function $G_j(y, x) = (y - m_j(x, \beta_j)) / \sigma_j(x, \gamma_j)$. The straightforward estimate of this function is given by $\hat{G}_{jn}(y, x) = (y - m_j(x, \hat{\beta}_{jn})) / \sigma_j(x, \hat{\gamma}_{jn})$, where $\hat{\beta}_{jn}$ and $\hat{\gamma}_{jn}$ are the estimates of the unknown parameters. The adjusted observations now coincide with the estimated residuals

$$\hat{Y}_{ji}^a = \hat{\varepsilon}_{ji} = \frac{Y_{ji} - m_j(X_i, \hat{\beta}_{jn})}{\sigma_j(X_i, \hat{\gamma}_{jn})}, \quad i = 1, \dots, n, \quad j = 1, 2.$$

3.4.2. Nonparametric location-scale model estimation of F_{1x} and F_{2x}

In this setting one assumes that the influence of the covariate on the marginal distributions is given by the model

$$Y_1 = m_1(X) + \sigma_1(X) \varepsilon_1, \quad Y_2 = m_2(X) + \sigma_2(X) \varepsilon_2,$$

where m_1, m_2, σ_1 and σ_2 are unknown functions and both ε_1 and ε_2 are independent of X with $E\varepsilon_1 = E\varepsilon_2 = 0$ and $\text{var}(\varepsilon_1) = \text{var}(\varepsilon_2) = 1$.

For simplicity of presentation we will consider only local linear regression estimates of these unknown functions ($j = 1, 2$):

$$\hat{m}_{jn}(t) = \sum_{i=1}^n w_{ni}(t, g_{jn}) Y_{ji}, \quad \hat{\sigma}_{jn}^2(t) = \sum_{i=1}^n w_{ni}(t, g_{jn}) (Y_{ji} - \hat{m}_{jn}(X_i))^2, \quad (3.7)$$

with the weights $w_{ni}(x, g_n)$ given by

$$w_{ni}(x, g_n) = \frac{\frac{1}{ng_n} k\left(\frac{x-X_i}{g_n}\right) (S_{n,2}(x) - \frac{x-X_i}{g_n} S_{n,1}(x))}{S_{n,0}(x) S_{n,2}(x) - S_{n,1}^2(x)}, \quad i = 1, \dots, n, \quad (3.8)$$

where

$$S_{n,j}(x) = \frac{1}{ng_n} \sum_{i=1}^n \left(\frac{x-X_i}{g_n}\right)^j k\left(\frac{x-X_i}{g_n}\right), \quad j = 0, 1, 2, \quad (3.9)$$

with $k(\cdot)$ a given kernel function.

The transformation function is now given by $G_j(y, x) = \frac{y - m_j(x)}{\sigma_j(x)}$ and its estimate by $\widehat{G}_{jn}(y, x) = \frac{y - \widehat{m}_{jn}(x)}{\widehat{\sigma}_{jn}(x)}$. The adjusted observations now coincide with the estimated residuals

$$\widehat{Y}_{ji}^a = \widehat{\varepsilon}_{ji} = \frac{Y_{ji} - \widehat{m}_{jn}(X_i)}{\widehat{\sigma}_{jn}(X_i)}, \quad i = 1, \dots, n, \quad j = 1, 2. \quad (3.10)$$

3.4.3. General nonparametric estimation of F_{1x} and F_{2x}

Sometimes, one has no idea about the influence of X on Y_1 and Y_2 . Then one uses the general transformation functions $G_j(y, x) = F_{jx}(y)$ that is estimated as

$$\widehat{G}_{jn}(y, x) = \widehat{F}_{jx}(y) = \sum_{i=1}^n w_{ni}(x, g_{jn}) \mathbb{I}\{Y_{ji} \leq y\}, \quad (3.11)$$

where $\{w_{ni}(x, g_{jn})\}$ is a sequence of local linear weights introduced in (3.8). The estimator in (3.11) is a standard kernel distribution function estimator. Other nonparametric estimators for a conditional distribution function can be used. For a recent contribution in this area, see e.g. Veraverbeke, Gijbels and Omelka (2014).

4. Application

As a practical illustration, the data on hydro-geochemical stream and sediment reconnaissance from Cook and Johnson (1981) are revisited. They consist of the observed log-concentrations of seven chemicals in 655 water samples collected near Grand Junction, Colorado. The data can be found e.g. as a data set called `uranium` in the R-package `copula` (Kojadinovic and Yan, 2010).

Following Acar, Genest and Nešlehová (2012) we first concentrate on Cobalt (Co), Scandium (Sc) and Titanium (Ti). The pairwise scatter plots are shown in Figure 2. Suppose we are interested in the relation of Cobalt (Y_1) and Scandium (Y_2) when Titanium (X) is taken into account. For exploration purpose, we fitted simple linear models (1m) $Y_j = \beta_{j1} + \beta_{j2} X + \varepsilon_j$, indicated in Figures 2 (b) and (c) with a dotted line. Similarly, nonparametric location models $Y_j = m_j(X) + \varepsilon_j$ were fitted with the help of the `locpol` R-package (Cabrera, 2012). The fits are indicated in Figures 2 (b) and (c) with a solid line (1p).

As the fits of the nonparametric mean functions in Figures 2 (b) and (c) are in reasonably good agreement with the simple linear fits for the majority of data points, we use the following methods of adjustments to estimate the partial Kendall's tau:

- `1m` Adjustment by simple linear regression models $Y_j = \beta_{j1} + \beta_{j2} X + \varepsilon_j$;
- `unif` Adjustment by nonparametric estimation of the conditional distribution functions F_{1x} and F_{2x} (see Section 3.4.3).

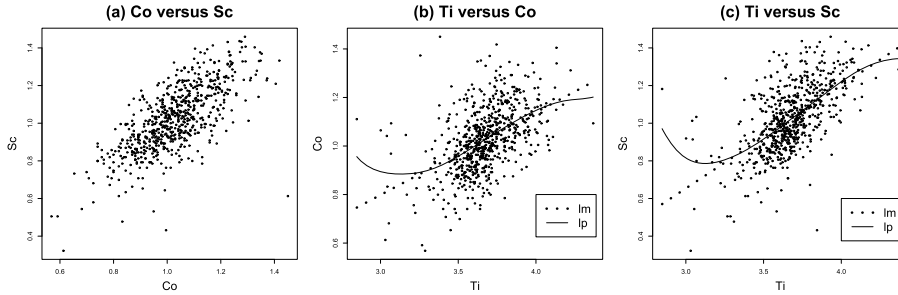


FIG 2. Pairwise scatter plots of the values of Cobalt (Co), Scandium (Sc) and Titanium (Ti).

TABLE 3

Estimated values of the unconditional (global), the original partial, the partial *lm* (using method *lm*), partial *unif* (using method *unif*), and the average conditional Kendall's tau for the indicated variables. Boldface notated values are plotted in Figure 3

estimated Kendall's tau	Co vs Sc (given Ti)	Cs vs K (given Ti)	Cs vs Sc (given Ti)
unconditional (global)	0.535	0.207	0.233
original partial	0.449	0.205	0.117
partial (<i>lm</i>)	0.406	0.201	0.070
partial (<i>unif</i>)	0.402	0.225	0.055
average conditional	0.391	0.204	0.066

The results for both methods, partial *lm* and partial *unif*, applied to the triplet $(Co, Sc, Ti) = (Y_1, Y_2, X)$ are quite comparable, as can be seen from Table 3. The table also lists the estimated value of the average conditional Kendall's tau defined in (2.6), and the sample version of *original* partial Kendall's tau $\bar{\tau}^K$, see (A.4), which is slightly higher than all the above values. Note that all these values are lower than the unconditional (global) Kendall's tau, defined in (A.2), of Co and Sc (so unadjusted for Ti) that is 0.535.

As explained in previous sections, although the marginals may be adjusted for the effect of the covariate, the dependence structure may change with the value of the covariate. Then the (original) partial, partial and average conditional measures provide different approaches to measuring average dependence over X . To quantify the effect of the covariate in more detail we present also the conditional Kendall's tau that measures the dependence of Co and Sc when Titanium is fixed to a given value. In Figure 3 we present the estimator of the conditional Kendall's tau, constructed from the observations adjusted (nonparametrically) for the values of the covariate, i.e. using $(\hat{Y}_{1i}^a, \hat{Y}_{2i}^a) = (\hat{F}_{1X_i}(Y_{1i}), \hat{F}_{2X_i}(Y_{2i}))$, where the observations are weighted according to the distance of X_i to the point of interest x , as in (3.4). For details about the construction of an estimator of the conditional Kendall's tau see Gijbels, Veraverbeke and Omelka (2011). The bandwidth used to construct the weights (for smoothing in the covariate direction) was fixed to 0.57 in order to have comparable results with Acar, Genest and Nešlehová (2012).

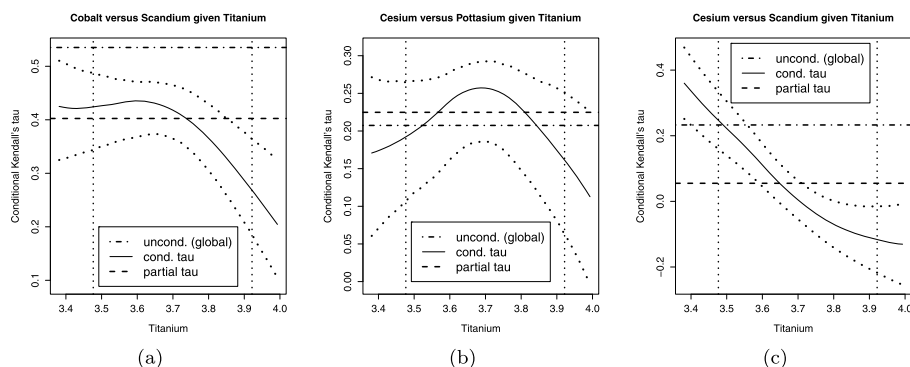


FIG 3. Estimated unconditional, conditional and partial Kendall's tau for (a) Cobalt (Co) and Scandium (Sc) given Titanium (Ti); (b) Cesium (Cs) and Pottasium (K) given Titanium (Ti); and (c) Cesium (Cs) and Scandium (Sc) given Titanium (Ti).

The estimated conditional Kendall's tau together with pointwise 95% confidence intervals is plotted for different values of Ti in Figure 3(a), using a solid line and dotted lines, respectively. The lower and the upper limits of the confidence intervals are derived by the bootstrap method presented in Omelka, Veraverbeke and Gijbels (2013). The estimates for the unconditional (global) Kendall's tau and the partial (via the method `unif`) Kendall's tau are indicated by the horizontal lines (dashed-dotted and dashed lines respectively). The range of Ti extends from the 5th to the 95th quantile of that variable. The 10th and the 90th quantiles of Ti are indicated by dotted vertical lines.

The dependence between Cobalt and Scandium clearly depends on the Titanium value, as shown by the estimated conditional Kendall's tau.

We also present similar results for log-concentrations of other chemicals. The considered additional triplets are (Cesium, Pottasium, Titanium) and (Cesium, Scandium, Titanium): $(Cs, K, Ti) = (Y_1, Y_2, X)$ and $(Cs, Sc, Ti) = (Y_1, Y_2, X)$. Figures 3(b) and (c) summarize the results. Of particular interest is to note that the unconditional Kendall's tau is for both pairs (Y_1, Y_2) around 0.2, whereas the partial Kendall's tau is for the pair (Cs, Sc) close to zero. In other words, the average strength of the dependence between the log-concentrations of Cesium and Scandium is far smaller than that between the log-concentrations of Cesium and Pottasium, when taking the log-concentration of Titanium into account. See also Table 3 for the estimated values of the other quantities. From this and other examples and simulations, we experienced that the values for the *original* partial Kendall's tau are often in between these of the unconditional Kendall's tau and the *newly-defined* partial Kendall's tau. See also Figure 3 and Table 3.

To illustrate further the use of other association measures and other conditioning settings, we provide in Figure 4(a) (respectively Figure 4(b)) the estimated upper (respectively lower) unconditional, conditional, partial and average tail coefficient of Schmid and Schmidt (2007) $(\rho_L, \rho_L(X), \bar{\rho}_L$ and ρ_L^A ; and similarly for the upper tail coefficients). Note that the average and partial co-

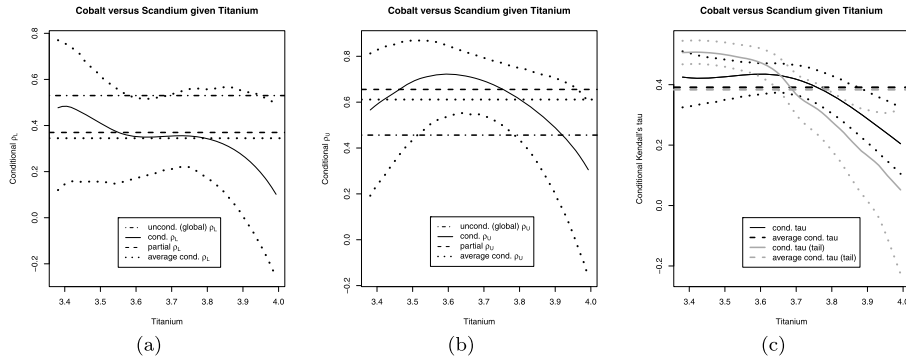


FIG 4. Dependence structures between Cobalt (Co) and Scandium (Sc) given Titanium (Ti). (a) & (b). Estimated unconditional, conditional, partial and average lower (ρ_L , $\rho_L(X)$, $\bar{\rho}_L$ and ρ_L^A) and upper tail (ρ_U , $\rho_U(X)$, $\bar{\rho}_U$ and ρ_U^A) coefficients; (c). Estimated conditional Kendall's tau (solid lines) and average Kendall's tau (horizontal dotted lines) for two different conditioning settings: $X = x$ (black lines) and $X \geq x$ (grey lines).

efficients are close (their population versions coincide as proved in Proposition A.1). Further, the tail dependence seems to be a bit higher in the upper tail than in the lower tail. Moreover the upper tail dependence reaches a maximum around a Ti-value of 3.6 and then the tail dependence weakens. Figure 4 (c) depicts the estimates for the conditional and average (conditional) Kendall's tau for two different conditioning settings: $X = x$ and $X \geq x$ (respectively the black and grey solid curves). Although the curves look quite different with a switching regime in dependence strength (around 3.65), their average values (the horizontal lines) are close to each other, meaning that on average the strength of the dependence between Cobalt and Scandium is comparable when either looking at a given Titanium value, or at Titanium values exceeding a given threshold. For all estimated conditional association measures in Figure 4 (a)–(c) we also plot 95% confidence intervals. For most confidence intervals the bootstrap procedure of Omelka, Veraverbeke and Gijbels (2013) was applied, but intervals for the conditional Kendall's tau, when conditioning on the event $X \geq x$, were constructed based on the asymptotic normality result for the conditional Kendall's tau.

5. Theoretical results

In this section we first discuss nonparametric estimation of a partial copula, defined in (2.4).

We need to transform the observed random variables Y_{1i} and Y_{2i} , $i = 1, \dots, n$ to be less (or not) dependent on X_i . The transformations are based on $G_j(y, x)$ (with $j = 1, 2$) using their estimates $\hat{G}_{jn}(y, x)$. Depending on whether the influence of X on Y_j ($j = 1, 2$) can be modelled by a parametric location-scale model (see Section 3.4.1), a nonparametric location-scale model (see Section 3.4.2), or

this influence is fully unknown as in Section 3.4.3, we use the following estimated transformations

- for parametric location-scale adjustments: $\widehat{G}_{jn}(y, x) = \frac{y - m_j(x, \widehat{\beta}_{jn})}{\sigma_j(x, \widehat{\gamma}_{jn})}$
- for nonparametric location-scale adjustments: $\widehat{G}_{jn}(y, x) = \frac{y - \widehat{m}_{jn}(x)}{\widehat{\sigma}_{jn}(x)}$
- for general nonparametric adjustments:

$$\widehat{G}_{jn}(y, x) = \widehat{F}_{jx}(y) = \sum_{i=1}^n w_{nj}(x, g_{jn}) \mathbb{I}\{Y_{ji} \leq y\}. \quad (5.1)$$

Based on the transformed observations $(\widehat{Y}_{11}^a, \widehat{Y}_{21}^a, X_1), \dots, (\widehat{Y}_{1n}^a, \widehat{Y}_{2n}^a, X_n)$, with

$$\widehat{Y}_{ji}^a = \widehat{\varepsilon}_{ji} = \widehat{G}_{jn}(Y_{ji}, X_i) \quad i = 1, \dots, n, \quad (5.2)$$

the nonparametric estimator of the partial copula in (2.4) is then given by

$$\bar{C}_n(u_1, u_2) = \frac{1}{n} \sum_{i=1}^n \mathbb{I}\left\{\widehat{\varepsilon}_{1i} \leq \widehat{F}_{1\widehat{\varepsilon}}^{-1}(u_1), \widehat{\varepsilon}_{2i} \leq \widehat{F}_{2\widehat{\varepsilon}}^{-1}(u_2)\right\}, \quad (5.3)$$

where, for $j = 1, 2$,

$$\widehat{F}_{j\widehat{\varepsilon}}(y) = \frac{1}{n} \sum_{i=1}^n \mathbb{I}\{\widehat{\varepsilon}_{ji} \leq y\},$$

is the estimate of the marginal distribution function $F_{j\varepsilon}(y) = \mathbb{P}(\varepsilon_j \leq y)$ of $\varepsilon_j = G_j(Y_j, X)$.

The estimator (5.3) was studied in Gijbels, Omelka and Veraverbeke (2015) but under the restrictive setting that the simplifying assumption holds, i.e. that only the marginal distribution functions are affected by the covariate X (see (2.7)).

In the next section we establish the asymptotic properties of the estimator defined in (5.3), in its general setting. These results are then the basis for proving asymptotic properties (see Section 5.2) of the estimator of the partial Kendall's tau, defined in (3.6). Finally, in Section 5.3, we provide asymptotic results for the estimator of the average conditional Kendall's tau, given in (3.5). For clarity of presentation, all assumptions are formulated in the Appendix. The theoretical results are presented according to the three major transformations considered in the adjustment/transformation step (5.2), as the asymptotic behaviour of the estimators depends on this step (and the degree of prior knowledge that it reflects).

5.1. Asymptotic results for the nonparametric partial copula estimator

Theorems 1, 2, and 3, establish asymptotic i.i.d. representations for the estimator (5.3) of the partial copula (2.4), when using the respective estimated transformations in (5.1).

In what follows let $U_{1i} = F_{1X_i}(Y_{1i})$, $U_{2i} = F_{2X_i}(Y_{2i})$ for $i = 1, \dots, n$.

5.1.1. Parametric location-scale adjustments

Theorem 1. Assume that the marginal distributions follow parametric location-scale models described in Section 3.4.1 and that (\mathbf{Cp}) , $(\beta\gamma)$, $(\mathbf{F1p})$, $(\mathbf{F2p})$, $(\mathbf{m}\sigma\mathbf{p})$ given in the Appendix hold. Then uniformly in $(u_1, u_2) \in [0, 1]^2$

$$\begin{aligned} & \sqrt{n} (\bar{C}_n(u_1, u_2) - \bar{C}(u_1, u_2)) \\ &= \frac{1}{\sqrt{n}} \sum_{i=1}^n \psi(U_{1i}, U_{2i}, u_1, u_2) + \sum_{j=1}^2 f_{j\varepsilon}(F_{j\varepsilon}^{-1}(u_j)) \mathbf{A}_j^\top(u_1, u_2) \sqrt{n} (\hat{\beta}_{jn} - \beta_j) \\ & \quad + \sum_{j=1}^2 f_{j\varepsilon}(F_{j\varepsilon}^{-1}(u_j)) F_{j\varepsilon}^{-1}(u_j) \mathbf{B}_j^\top(u_1, u_2) \sqrt{n} (\hat{\gamma}_{jn} - \gamma_j) + o_P(1), \end{aligned}$$

where

$$\begin{aligned} \psi(v_1, v_2, u_1, u_2) &= \mathbb{I}\{v_1 \leq u_1, v_2 \leq u_2\} - \bar{C}(u_1, u_2) \\ & \quad - \bar{C}^{(1)}(u_1, u_2) [\mathbb{I}\{v_1 \leq u_1\} - u_1] \\ & \quad - \bar{C}^{(2)}(u_1, u_2) [\mathbb{I}\{v_2 \leq u_2\} - u_2], \end{aligned} \tag{5.4}$$

$$\begin{aligned} \mathbf{A}_j(u_1, u_2) &= \mathbb{E}_X \left[C_X^{(j)}(u_1, u_2) \frac{\dot{m}_j(X, \beta_j)}{\sigma_j(X, \gamma_j)} \right] - \mathbb{E}_X \left[C_X^{(j)}(u_1, u_2) \right] \mathbb{E}_X \left[\frac{\dot{m}_j(X, \beta_j)}{\sigma_j(X, \gamma_j)} \right], \\ \mathbf{B}_j(u_1, u_2) &= \mathbb{E}_X \left[C_X^{(j)}(u_1, u_2) \frac{\dot{\sigma}_j(X, \gamma_j)}{\sigma_j(X, \gamma_j)} \right] - \mathbb{E}_X \left[C_X^{(j)}(u_1, u_2) \right] \mathbb{E}_X \left[\frac{\dot{\sigma}_j(X, \gamma_j)}{\sigma_j(X, \gamma_j)} \right], \end{aligned}$$

with

$$\dot{m}_j(X, \beta_j) = \frac{\partial m_j(X, \beta_j)}{\partial \beta_j} \quad \text{and} \quad \dot{\sigma}_j(X, \gamma_j) = \frac{\partial \sigma_j(X, \gamma_j)}{\partial \gamma_j}.$$

As mentioned in the introduction of Section 5, Theorem 1 can be viewed as an extension of the results presented in Gijbels, Omelka and Veraverbeke (2015). From this we thus can tell what happens if the pairwise simplifying assumption is wrongly assumed. The consequences for the estimator \bar{C}_n can be summarized as follows.

- \bar{C}_n still converges at \sqrt{n} -rate, but now \bar{C}_n estimates the partial copula function \bar{C} (and not C which is not even well defined now).
- The limiting structure of the estimator \bar{C}_n is more complicated due to the second and third term in the asymptotic representation of $\sqrt{n} (\bar{C}_n - \bar{C})$.

On the other hand if the pairwise simplifying assumption (2.7) really holds, then $\bar{C} \equiv C$ and both $\mathbf{A}_j(u_1, u_2)$ and $\mathbf{B}_j(u_1, u_2)$ vanish and thus also the second and third term in the asymptotic representation of $\sqrt{n} (\bar{C}_n - \bar{C})$. The latter then coincides with the results of Gijbels, Omelka and Veraverbeke (2015), where the asymptotic representation

$$\sqrt{n} (\bar{C}_n(u_1, u_2) - C(u_1, u_2)) = \frac{1}{\sqrt{n}} \sum_{i=1}^n \psi(U_{1i}, U_{2i}, u_1, u_2) + o_P(1)$$

is derived.

5.1.2. Nonparametric location-scale adjustments

Let $F_{j\varepsilon}$ be the distribution function of ε_j .

Theorem 2. *Assume that the marginal distributions follow nonparametric location-scale models described in Section 3.4.2 and that (\mathbf{Cn}) , (\mathbf{Bwn}) , $(\mathbf{F1n})$, $(\mathbf{F2n})$, (\mathbf{kn}) , $(\mathbf{m}\sigma)$ and (\mathbf{Xn}) given in the Appendix hold. Then uniformly in $(u_1, u_2) \in [0, 1]^2$*

$$\begin{aligned} & \sqrt{n} (\bar{C}_n(u_1, u_2) - \bar{C}(u_1, u_2)) \\ &= \frac{1}{\sqrt{n}} \sum_{i=1}^n \psi(U_{1i}, U_{2i}, u_1, u_2) \\ & \quad + \sum_{j=1}^2 f_{j\varepsilon}(F_{j\varepsilon}^{-1}(u_j)) \frac{1}{\sqrt{n}} \sum_{i=1}^n \phi_j(X_i, U_{ji}, u_1, u_2) + o_P(1), \end{aligned}$$

with ψ be given in (5.4) and where for $j = 1, 2$

$$\begin{aligned} \phi_j(x, v, u_1, u_2) &= \left\{ C_x^{(j)}(u_1, u_2) - \mathbb{E}_X \left[C_X^{(j)}(u_1, u_2) \right] \right\} \\ & \quad \times \left\{ F_{j\varepsilon}^{-1}(v) + \frac{F_{j\varepsilon}^{-1}(u_j)}{2} \left[(F_{j\varepsilon}^{-1}(v))^2 - 1 \right] \right\}. \end{aligned} \quad (5.5)$$

Analogously as Theorem 1 extends the result of Gijbels, Omelka and Veraverbeke (2015) for the parametric location-scale model adjustment, Theorem 2 does this for the nonparametric location-scale model adjustment. If the simplifying assumption (2.7) holds, then the functions ϕ_j given in (5.5) vanish and the result of Theorem 2 is in agreement with Gijbels, Omelka and Veraverbeke (2015). On the other hand if the simplifying assumption (2.7) does not hold, then \bar{C}_n still converges at \sqrt{n} -rate, but now it estimates \bar{C} and the limiting structure of the estimator is more involved.

5.1.3. General nonparametric adjustments

Theorem 3. *Suppose that assumptions (\mathbf{Bw}) , (\mathbf{F}) , (\mathbf{k}) and (\mathbf{X}) given in the Appendix are satisfied. Then the estimator \bar{C}_n is a consistent estimator of the copula C , that is*

$$\sup_{u_1, u_2} |\bar{C}_n(u_1, u_2) - \bar{C}(u_1, u_2)| = O_P(r_n), \quad (5.6)$$

where $r_n = \max \left\{ g_{1n}^2, g_{2n}^2, \sqrt{\frac{\log n}{n g_{n1}}}, \sqrt{\frac{\log n}{n g_{n2}}} \right\}$.

Note that this theorem also gives the same rate as the corresponding (more restrictive) theorem in Gijbels, Omelka and Veraverbeke (2015).

5.2. Asymptotic results for the estimator of the partial Kendall's tau

5.2.1. Parametric location-scale adjustments

Note that thanks to Hadamard differentiability of the functional $C \mapsto \iint C dC$ (tangentially to the set of functions that are continuous on $[0, 1]^2$) proved in Lemma 1 of Veraverbeke, Omelka and Gijbels (2011) one gets

$$\sqrt{n}(\bar{\tau}_n - \bar{\tau}) = 8 \iint \alpha_n(u_1, u_2) dC(u_1, u_2) + o_P(1), \tag{5.7}$$

where α_n stands for the asymptotic representation of $\sqrt{n}(\bar{C}_n - \bar{C})$ (see Theorem 1). Now (5.7) together with some further calculations yields the following result.

Theorem 4. *Assume that the marginal distributions follow parametric location-scale models described in Section 3.4.1 and that (\mathbf{Cp}) , $(\beta\gamma)$, $(\mathbf{F1p})$, $(\mathbf{F2p})$, $(\mathbf{m}\sigma\mathbf{p})$ are satisfied. Then*

$$\begin{aligned} \sqrt{n}(\bar{\tau}_n - \bar{\tau}) &= \frac{1}{\sqrt{n}} \sum_{i=1}^n [8\bar{C}(U_{1i}, U_{2i}) - 4U_{1i} - 4U_{2i} + 2 - 2\bar{\tau}] \\ &\quad + 8 \sum_{j=1}^2 \tilde{\mathbf{A}}_j^T \sqrt{n}(\hat{\beta}_{jn} - \beta_j) + 8 \sum_{j=1}^2 \tilde{\mathbf{B}}_j^T \sqrt{n}(\hat{\gamma}_{jn} - \gamma_j) \\ &\quad + o_P(1), \end{aligned} \tag{5.8}$$

where

$$\begin{aligned} \tilde{\mathbf{A}}_j &= \iint f_{j\varepsilon}(F_{j\varepsilon}^{-1}(u_j)) \mathbf{A}_j(u_1, u_2) d\bar{C}(u_1, u_2), \\ \tilde{\mathbf{B}}_j &= \iint f_{j\varepsilon}(F_{j\varepsilon}^{-1}(u_j)) F_{j\varepsilon}^{-1}(u_j) \mathbf{B}_j(u_1, u_2) d\bar{C}(u_1, u_2). \end{aligned}$$

Note that provided one has asymptotic representations for $\sqrt{n}(\hat{\beta}_{jn} - \beta_j)$ and $\sqrt{n}(\hat{\gamma}_{jn} - \gamma_j)$, one can with the help of (5.8) derive the asymptotic distribution of $\bar{\tau}_n$.

5.2.2. Nonparametric location-scale adjustments

With the help of (5.7) and similarly as in the previous section one can show the following i.i.d. representation of the estimator of the partial Kendall's tau.

Theorem 5. *Assume that the marginal distributions follow nonparametric location-scale models described in Section 3.4.2 and that (\mathbf{Cn}) , (\mathbf{Bwn}) , $(\mathbf{F1n})$, $(\mathbf{F2n})$, (\mathbf{kn}) , $(\mathbf{m}\sigma)$ and (\mathbf{Xn}) given in the Appendix hold. Then*

$$\sqrt{n}(\bar{\tau}_n - \bar{\tau}) = \frac{1}{\sqrt{n}} \sum_{i=1}^n [8\bar{C}(U_{1i}, U_{2i}) - 4U_{1i} - 4U_{2i} + 2 - 2\bar{\tau}]$$

$$\begin{aligned}
& + \frac{8}{\sqrt{n}} \sum_{j=1}^2 \sum_{i=1}^n \left\{ \tilde{\phi}_{1j}(X_i) F_{j\varepsilon}^{-1}(U_{ji}) + \tilde{\phi}_{2j}(X_i) \left[\frac{(F_{j\varepsilon}^{-1}(U_{ji}))^2}{2} - 1 \right] \right\} \\
& + o_P(1),
\end{aligned}$$

where

$$\begin{aligned}
\tilde{\phi}_{1j}(X_i) &= \iint f_{j\varepsilon}(F_{j\varepsilon}^{-1}(u_j)) \left\{ C_{X_i}^{(j)}(u_1, u_2) - \mathbb{E}_X \left[C_X^{(j)}(u_1, u_2) \right] \right\} d\bar{C}(u_1, u_2), \\
\tilde{\phi}_{2j}(X_i) &= \frac{1}{2} \iint f_{j\varepsilon}(F_{j\varepsilon}^{-1}(u_j)) F_{j\varepsilon}^{-1}(u_j) \\
&\quad \times \left\{ C_{X_i}^{(j)}(u_1, u_2) - \mathbb{E}_X \left[C_X^{(j)}(u_1, u_2) \right] \right\} d\bar{C}(u_1, u_2).
\end{aligned}$$

5.2.3. General nonparametric adjustments

Theorem 6. Suppose that assumptions **(Bw)**, **(F)**, **(k)** and **(X)** given in the Appendix are satisfied. Then $\bar{\tau}_n - \bar{\tau} = O_P(r_n)$, where r_n is given in (5.6).

5.3. Asymptotic results for the estimator of the average conditional Kendall's tau

Let $\tau_n^A = \frac{1}{n} \sum_{i=1}^n \tau_n(X_i)$ with $\tau_n(x)$ given by (3.2).

Theorem 7. Assume that **(kn)**, **(Xn)** and **(H)** given in the Appendix hold. Assume also that the bandwidth h_n satisfies the assumptions of a bandwidth stated in **(Bwn)**. Then

$$\sqrt{n} (\tau_n^A - \tau^A) = \frac{1}{\sqrt{n}} \sum_{i=1}^n \phi(X_i, U_{1i}, U_{2i}) + o_P(1),$$

where

$$\phi(x, u_1, u_2) = 2 [4 C_x(u_1, u_2) - 1 - \tau^A] - 4 [u_1 + u_2 - 1] - [\tau(x) - \tau^A]. \quad (5.9)$$

Note that the nice thing about τ_n^A when compared with $\bar{\tau}_n$ is that it is asymptotically normal without requiring that the marginal distributions follow either parametric or nonparametric location scale models. This might be surprising as for each $x \in R_X$ the estimator of the conditional Kendall's tau $\tau_n(x)$ converges typically at most at $n^{2/5}$ -rate (Veraverbeke, Omelka and Gijbels, 2011). But thanks to averaging of $\tau_n(X_i)$ this rate is improved to \sqrt{n} (see Akritas and Van Keilegom, 2001; Neumeyer and Van Keilegom, 2010, among others, for similar settings in nonparametric smoothing where averaging improves the rate of convergence).

Note that asymptotically we do not even need to bother about the adjustments of the marginals. At first sight this might be surprising in view of the previous results on conditional Kendall's tau estimation (Veraverbeke, Omelka

and Gijbels, 2011). This can be explained by assumption **(Bwn)** on the bandwidth which together with assumption **(H)** guarantees that for each $x \in R_X$ the conditional bias (given X_1, \dots, X_n) of the conditional Kendall's tau estimator $\tau_n(x)$ is of order $o_P(n^{-1/2})$ uniformly in x . The bias of τ_n^A is of the same order. To improve the finite sample properties we still recommend to pre-adjust the observations for the effect of X on their marginal distributions as described in Section 3.2. The asymptotic normality of the resulting estimator for the average conditional Kendall's tau has also been established by the authors (result not included here, for brevity).

Remark 1. Suppose that the pairwise simplifying assumption (2.7) holds. Then the function ϕ in (5.9) simplifies to

$$\phi(x, u_1, u_2) = 2 [4C(u_1, u_2) - 1 - \tau^A] - 4 [u_1 + u_2 - 1]$$

and in fact does not depend on x any more. This implies that the estimator of the average conditional Kendall's tau τ_n^A has the same asymptotic distribution as the oracle estimator

$$\tau_n^{A(or)} = \frac{1}{n(n-1)} \sum_{i=1}^n \sum_{j=1}^n \mathbb{I}\{U_{1i} < U_{1j}, U_{2i} < U_{2j}\}$$

based on unobserved $\{(U_{1i}, U_{2i}), i = 1, \dots, n\}$. Note that this asymptotic distribution then coincides also with the asymptotic distribution of the estimator of the partial Kendall's tau when either parametric or nonparametric location scale models are correctly used to remove the effects of the covariate on the marginal distributions (see Theorems 2 and 4).

6. Conclusion and Discussion

In this paper we focus on several conditional association measures describing the dependence between two response variables Y_1 and Y_2 , given that a third (covariate) variable X takes some value x . The common feature of all these measures is that they are copula-based, i.e. they can be described as functionals $\varphi(C_X)$ of the conditional copula C_X . This leads to two different ways of summarizing the level of dependence by a single number. The first is to consider $E_X(\varphi(C_X))$, leading to the average (conditional) association measures. The second is to calculate $\varphi(\bar{C})$ where \bar{C} is the so-called partial copula $\bar{C}(\cdot, \cdot) = E_X(C_X(\cdot, \cdot))$, resulting in partial association measures. We provide statistical inference for the corresponding estimators in the important case of the average and partial Kendall's tau.

Based on the obtained results on estimation of the average and partial Kendall's tau we reported on the following interesting findings. A first finding is that the nonparametric estimator of the partial Kendall's tau $\bar{\tau}_n$ (given in (3.6)) is easier to compute than the nonparametric estimator for the average (conditional) Kendall's tau τ_n^A (see (3.2), (3.4) and (3.5)). A second finding is that for τ_n^A we can establish an (asymptotic) i.i.d. representation for $\sqrt{n}(\tau_n^A - \tau^A)$

in a general setting (see Theorem 7) and hence for τ_n^A an asymptotic normality result is available under this general setting. For the partial Kendall's tau estimator $\bar{\tau}_n$ however we could only establish (asymptotic) i.i.d. representations for $\sqrt{n}(\bar{\tau}_n - \bar{\tau})$ under the more restrictive settings of parametric or nonparametric location-scale modelling of the conditional marginal distributions. Under the more general setting (not requiring such location-scales models to hold) we only achieve consistency of the estimator $\bar{\tau}_n$ at the nonparametric rate r_n (see Theorem 6). In conclusion, both estimators exhibit a specific advantage (for one of computational type and for the other estimator of theoretical type).

In practical examples the choice between the various association metrics depends, among others, on the considered research question, but also on the taste of the researcher. For example, if one is interested in dependence structures in the tails of joint distributions, then a study of tail coefficient type of association measures would be of primary interest.

The association measures of the Kendall's tau and Spearman's rho type are often used in economic and social statistics. A possible disadvantage however is that these concordance measures are not very sensitive to the dependence in the tails of the bivariate copula. Such additional information is crucial in bivariate extreme value theory and can be provided by the tail coefficient measures, see Table 1. Their definitions describe the limiting amount of dependence in the edges of the copula domain. The classical lower and upper tail dependence coefficients have some drawback since they only evaluate the copula at the diagonal sections. The association metrics of Schmid and Schmidt (2007) in Table 1 offers an alternative by averaging over all directions in the edge.

Appendix A: (Un)conditional, average and partial association measures

In this appendix we provide, in Sections A.1 and A.2, some background information on the *original* partial correlation measures, illustrate their drawback with an example, and illustrate the notion of conditional association measure (in Section A.3). Moreover, we provide some further insights in coincidence, or not, of the notions of average conditional association measure and partial association measure.

A.1. Global or unconditional association measures

Among the most popular unconditional association measures is Kendall's tau. Consider (Y'_1, Y'_2) an independent copy of (Y_1, Y_2) . Kendall's tau is then defined as the probability of concordance minus the probability of discordance between the couples (Y_1, Y_2) and (Y'_1, Y'_2) , i.e.

$$\tau(Y_1, Y_2) = \mathbb{P}[(Y_1 - Y'_1)(Y_2 - Y'_2) > 0] - \mathbb{P}[(Y_1 - Y'_1)(Y_2 - Y'_2) < 0] \quad (\text{A.1})$$

$$= 2\mathbb{P}[(Y_1 - Y'_1)(Y_2 - Y'_2) > 0] - 1, \quad (\text{A.2})$$

where the second equality holds when Y_1 and Y_2 are continuous random variables.

Another popular association measure is Spearman's rho defined as follows. Consider (Y'_1, Y'_2) and (Y''_1, Y''_2) two independent copies of (Y_1, Y_2) . Spearman's rho is defined as

$$\rho^{(S)}(Y_1, Y_2) = 3 \{ \mathbb{P}[(Y_1 - Y'_1)(Y_2 - Y''_2) > 0] - \mathbb{P}[(Y_1 - Y'_1)(Y_2 - Y''_2) < 0] \}.$$

Recall that, due to the probability integral transformation, it holds that: if Y_1 and Y_2 are continuous random variables with respective distribution functions F_{Y_1} and F_{Y_2} , then $F_{Y_1}(Y_1)$ and $F_{Y_2}(Y_2)$ are uniformly on $[0, 1]$ distributed random variables. In this particular case of continuous random variables Y_1 and Y_2 , Spearman's rho is equal to Pearson's correlation coefficient for the transformed random variables, $F_{Y_1}(Y_1)$ and $F_{Y_2}(Y_2)$, that is (see Nelsen, 2006)

$$\begin{aligned} \rho^{(S)}(Y_1, Y_2) = \rho^{(P)}(F_{Y_1}(Y_1), F_{Y_2}(Y_2)) &= \frac{\text{cov}(F_{Y_1}(Y_1), F_{Y_2}(Y_2))}{\sqrt{\text{var}(F_{Y_1}(Y_1)) \text{var}(F_{Y_2}(Y_2))}} \\ &= 12 \text{cov}(F_{Y_1}(Y_1), F_{Y_2}(Y_2)), \end{aligned}$$

since the variance of a uniform distribution on $[0, 1]$ equals $\frac{1}{12}$.

A.2. On some original partial association measures

Assuming a trivariate normal distribution for the triple (Y_1, Y_2, X) implies the following regression model structures:

$$Y_1 = \alpha_1 + \beta_1 X + \varepsilon_1, \quad Y_2 = \alpha_2 + \beta_2 X + \varepsilon_2, \tag{A.3}$$

where ε_1 and ε_2 are independent of X . The Pearson partial correlation coefficient $\rho_X^{(P)}(Y_1, Y_2)$ (see (1.2)) then measures the correlation of ε_1 and ε_2 . Note that model (A.3) implies that the dependence structure of the ' X adjusted' variables $Y_1 - \alpha_1 - \beta_1 X$ and $Y_2 - \alpha_2 - \beta_2 X$ does not depend on X any more, thus in this model the partial correlation coefficient coincides with the conditional correlation coefficient that measures the dependence of Y_1 and Y_2 given $X = x$.

Analogously as for the Pearson's correlation coefficient, researchers soon realized the need for alternatives to the partial Pearson's correlation coefficient that would not require the assumption of a trivariate normal distribution. Inspired by formula (1.2) for the partial Pearson's correlation coefficient, Kendall (1942) suggested to define the (*original*) partial Kendall's tau whose population version is denoted and given by

$$\bar{\tau}^K = \tau_X(Y_1, Y_2) = \frac{\tau(Y_1, Y_2) - \tau(Y_1, X) \tau(Y_2, X)}{\sqrt{1 - \tau^2(Y_1, X)} \sqrt{1 - \tau^2(Y_2, X)}}, \tag{A.4}$$

where $\tau(A, B)$ is the (global) Kendall's tau of the random variables A and B , see (A.1). See also Goodman (1959). While the obvious advantage of $\bar{\tau}^K$ is its

simplicity (to get the estimate it is sufficient to replace the pairwise Kendall's tau with their empirical versions), several criticism appeared in the literature, questioning whether $\bar{\tau}^K$ is a reasonable measure of the dependence of Y_1 and Y_2 when X is taken into consideration, see e.g. Korn (1984), Nelson and Yang (1988) and Gripenberg (1992). The difficulties of $\bar{\tau}^K$ are illustrated also by the following two examples (for other examples see the references given above).

Example A.1 Consider the models:

$$Y_1 = X + \varepsilon, \quad Y_2 = X^2 + \varepsilon,$$

where $X \sim \mathcal{N}(0, 1)$, $\varepsilon \sim \mathcal{N}(0, \sigma^2)$ and X and ε are independent. Note that $Y_1 - X = Y_2 - X^2$, thus Y_1 and Y_2 are perfectly dependent when (correctly) adjusted for the effect of X , but it can be checked via Monte Carlo simulations that $\bar{\tau}^K$ depends on the value of σ and even $\bar{\tau}^K \rightarrow 0$ as $\sigma \rightarrow 0$. Thus for small σ the coefficient $\bar{\tau}^K$ completely fails to measure the dependence of Y_1 and Y_2 when adjusted for X .

Example A.2 Suppose we have the models

$$Y_1 = 2 \exp\{X\} + \varepsilon_1, \quad Y_2 = 2 \exp\{X\} + \varepsilon_2, \quad (\text{A.5})$$

where X , ε_1 and ε_2 are independent and all with a uniform distribution on $[0, 1]$ (denoted by $\mathcal{U}[0, 1]$). Note that here $Y_1 - 2 \exp\{X\}$ and $Y_2 - 2 \exp\{X\}$ (that is Y_1 and Y_2 adjusted for X) are independent. But by Monte Carlo simulation one can find out that $\bar{\tau}^K \doteq -0.24$.

A.3. Conditional association measures

A more detailed characterization of the dependence structure can be gained with the help of conditional measures of dependency/association, that measure the dependency/association of (Y_1, Y_2) conditionally on the event that $X = x$. Let us consider for instance the conditional Kendall's tau (see Gijbels, Veraverbeke and Omelka, 2011) that is denoted and defined as

$$\tau(x) = \tau(Y_1, Y_2 | X = x) = 2 \mathbb{P}[(Y_1 - Y'_1)(Y_2 - Y'_2) > 0 | X = X' = x] - 1, \quad (\text{A.6})$$

where (Y'_1, Y'_2, X') is an independent copy of the random vector (Y_1, Y_2, X) and Y_1 and Y_2 are continuous random variables. See also (A.1) and (A.2). It is easy to see that in the first example (Example A.1) $\tau(x) = 1$ for all x , while in Example A.2 $\tau(x) = 0$ for all x . Examples A.1 and A.2 are thus simple in the sense that the conditional dependence structures do not change with x , i.e. are constant in x . Denote by Y_1^a the variable Y_1 adjusted for X , and similarly the variable Y_2^a , the variable Y_2 adjusted for X . Note that in the above two examples one can characterize the dependence of Y_1^a and Y_2^a by a single (global) association measure (respectively 1 and 0).

In more complex (and realistic) models, the dependency structure of Y_1^a and Y_2^a can still change with X , which prevents us from describing the dependency/association structure of Y_1^a and Y_2^a by one single real number. Such a more complex model is presented in the following example that modifies Example A.2.

Example A.3 Suppose that Y_1 and Y_2 follow (A.5), and that the marginal distributions of X , ε_1 and ε_2 are the same as in Example A.2. But now the conditional joint distribution of $(\varepsilon_1, \varepsilon_2)$ given $X = x$, with x in $[0, 1]$ and $(u_1, u_2) \in [0, 1]^2$, is given by

$$P(\varepsilon_1 \leq u_1, \varepsilon_2 \leq u_2 \mid X = x) = xu_1 u_2 + (1 - x) \min(u_1, u_2), \tag{A.7}$$

Note that ε_1 is independent of X , and thus ε_1 given $X = x$ is $\mathcal{U}[0, 1]$ for each value of $x \in [0, 1]$. The same holds for ε_2 given $X = x$. But X still influences the dependence structure of the random vector $(\varepsilon_1, \varepsilon_2)$. In this case, a straightforward calculation shows that the conditional Kendall's tau equals

$$\tau(x) = (1 - x)(1 - \frac{x}{3}).$$

This example just illustrates the usefulness of the concept of conditional Kendall's tau, or more generally, the concept of conditional association measures (see also Section 2).

A.4. On average conditional and partial association measures

We now look into the average conditional association measures, defined in general as $E_X \{ \varphi(X) \}$, and the partial association measures, defined as $\varphi(\bar{C})$. Table 1 in Section 1 indicates in the last column some interesting facts regarding equality (or not) of the two concepts. Example A.3 serves to illustrate that for Kendall's tau they in general do not coincide.

Example A.3 (continued) Note that in this example the conditional copula is given by the right-hand side of (A.7). Denote $C_1(u_1, u_2) = u_1 u_2$ (the independence copula) and $C_2(u_1, u_2) = \min(u_1, u_2)$ (the upper Fréchet-Hoeffding bound). Then one can write, with f_X denoting the density of X ,

$$\begin{aligned} \bar{\tau} &= 4 \int_{R_X} \int_{R_X} \iint_{[0,1]^2} [x C_1(u_1, u_2) + (1 - x) C_2(u_1, u_2)] \\ &\quad d[x' C_1(u_1, u_2) + (1 - x') C_2(u_1, u_2)] f_X(x) f_X(x') dx dx' - 1, \\ &= \frac{\tau_{11}}{4} + \frac{\tau_{12}}{4} + \frac{\tau_{21}}{4} + \frac{\tau_{22}}{4}, \end{aligned}$$

where $\tau_{ij} = 4 \iint_{[0,1]^2} C_i(u_1, u_2) dC_j(u_1, u_2) - 1$.

On the other hand

$$\begin{aligned} \mathbf{E}_X \tau(X) &= 4 \int_{R_X} \iint_{[0,1]^2} [x C_1(u_1, u_2) + (1-x) C_2(u_1, u_2)] \\ &\quad d[x C_1(u_1, u_2) + (1-x) C_2(u_1, u_2)] f_X(x) dx - 1, \\ &= \frac{\tau_{11}}{3} + \frac{\tau_{12}}{6} + \frac{\tau_{21}}{6} + \frac{\tau_{22}}{3}. \end{aligned}$$

Now it is easy to calculate that $\tau_{11} = 0$, $\tau_{12} = \tau_{21} = \frac{1}{3}$ and $\tau_{22} = 1$, which further implies

$$\bar{\tau} = \frac{1}{4}(0 + \frac{2}{3} + 1) = \frac{5}{12}, \quad \text{and} \quad \tau^A = \mathbf{E}_X \tau(X) = \frac{1}{6}(0 + \frac{2}{3} + 2) = \frac{4}{9}.$$

For the lower and upper tail coefficients, λ_L , λ_U , ρ_L and ρ_U in Table 1 the functional $\varphi(\cdot)$ involves a limit expression. It is fairly easy to show though that for this type of association measure, the concepts of average and partial tail coefficients coincide. Since the limits do not need to exist (see e.g. Larsson et al., 2011) we need to request existence.

Proposition A.1. *Suppose that all quantities below exist. Then*

$$\lambda_L^A = \bar{\lambda}_L, \quad \lambda_U^A = \bar{\lambda}_U, \quad \text{and} \quad \rho_L^A = \bar{\rho}_L, \quad \rho_U^A = \bar{\rho}_U.$$

Proof. Note that

$$\begin{aligned} \lambda_L^A = \mathbf{E}_X \lambda_L(X) &= \mathbf{E}_X \left[\lim_{t \rightarrow 0^+} \frac{C_X(t, t)}{t} \right] \\ &= \lim_{t \rightarrow 0^+} \mathbf{E}_X \left[\frac{C_X(t, t)}{t} \right] = \lim_{t \rightarrow 0^+} \frac{\bar{C}(t, t)}{t} = \bar{\lambda}_L, \end{aligned}$$

where the second equality is justified by applying Lebesgue’s dominated convergence theorem (allowing to interchange the integral and the limit). Indeed, by Lipschitz continuity of the copula C_x we have

$$|C_x(t, t)| = |C_x(t, t) - C_x(0, 0)| \leq 2|t - 0| \leq 2|t|,$$

and hence

$$\left| \frac{C_X(t, t)}{t} \right| \leq 2,$$

which is integrable with respect to $dF_X(x)$.

The other statements can be proven in a similar way. □

Appendices B–F: Assumptions, proofs and auxiliary results

Here we list the assumptions needed for each of the theoretical results, and provide proofs for these. The proofs of the asymptotic results for the nonparametric partial copula estimator are presented in Appendices B, C and D respectively, a division according to the three major transformations. The proof of Theorem 7, establishing the asymptotic representation for the estimator of the average conditional Kendalls’ tau, is given in Appendix E. Finally, some auxiliary results, needed in the proofs, can be found in Appendix F.

Appendix B: Adjusting through parametric location-scale models

In this appendix we prove Theorem 1.

Regularity assumptions

- (Cp) For $j = 1, 2$, the j -th first-order partial derivative of C_x exists and is continuous on the set $\{(u_1, u_2) \in [0, 1]^2 : 0 < u_j < 1\}$ uniformly in $x \in R_X$, where R_X is the support of f_X , the marginal density of X .
- (βγ) For $j = 1, 2$ the differences $\sqrt{n}(\hat{\beta}_{jn} - \beta_j)$ and $\sqrt{n}(\hat{\gamma}_{jn} - \gamma_j)$ converge in distribution.
- (F1p) The functions $F_{1\varepsilon}(y)$ and $F_{2\varepsilon}(y)$ are continuously differentiable.
- (F2p) For $j = 1, 2$ the functions $f_{j\varepsilon}(F_{j\varepsilon}^{-1}(u))$ and $f_{j\varepsilon}(F_{j\varepsilon}^{-1}(u)) F_{j\varepsilon}^{-1}(u)$ are continuous and tend to zero when $u \rightarrow 0_+$ or $u \rightarrow 1_-$, where $f_{j\varepsilon} = F_{j\varepsilon}'$.
- (mσp) For $j = 1, 2$ there exist open neighbourhoods $U(\beta_j)$ and $U(\gamma_j)$ of β_j and γ_j respectively, such that the partial derivatives of the functions $m_j(x, \cdot)$ of $\sigma_j(x, \cdot)$ with respect to their second arguments are uniformly continuous as functions of both arguments on $R_X \times U(\beta_j)$ and $R_X \times U(\gamma_j)$. Moreover, $\inf_{x \in R_X, \mathbf{c} \in U(\gamma_j)} \sigma_j(x, \mathbf{c}) > 0$.

Note that if the parameters $\beta_1, \beta_2, \gamma_1, \gamma_2$ were known, then one could estimate the unknown distribution function $F_{j\varepsilon}$ as

$$\hat{F}_{j\varepsilon}(y) = \frac{1}{n} \sum_{i=1}^n \mathbb{I}\{\varepsilon_{ji} \leq y\}, \quad j = 1, 2.$$

The estimator of \bar{C} then could be

$$\bar{C}_n^{(or)}(u_1, u_2) = \frac{1}{n} \sum_{i=1}^n \mathbb{I}\{\varepsilon_{1i} \leq \hat{F}_{1\varepsilon}^{-1}(u_1), \varepsilon_{2i} \leq \hat{F}_{2\varepsilon}^{-1}(u_2)\}, \quad (\text{B.1})$$

which can be viewed as a kind of ‘oracle’ estimator.

The proof of Theorem 1 follows by Proposition B.1 and by applying standard results on the asymptotic representation of the empirical copula process applied to $\bar{C}_n^{(or)}$ (see e.g. Gänssler and Stute, 1987; Fermanian, Radulović and Wegkamp, 2004; Tsukahara, 2005; Segers, 2012, among others).

Proposition B.1. *Assume that the assumptions of Theorem 1 are satisfied. Then*

$$\begin{aligned} & \sqrt{n} (\bar{C}_n(u_1, u_2) - \bar{C}_n^{(or)}(u_1, u_2)) \\ &= \sum_{j=1}^2 f_{j\varepsilon}(F_{j\varepsilon}^{-1}(u_j)) \mathbf{A}_j^\top(u_1, u_2) \sqrt{n} (\hat{\beta}_{jn} - \beta_j) \\ & \quad + \sum_{j=1}^2 f_{j\varepsilon}(F_{j\varepsilon}^{-1}(u_j)) F_{j\varepsilon}^{-1}(u_j) \mathbf{B}_j^\top(u_1, u_2) \sqrt{n} (\hat{\gamma}_{jn} - \gamma_j) + o_P(1), \quad (\text{B.2}) \end{aligned}$$

where $\bar{C}_n^{(or)}(u_1, u_2)$ is defined in (B.1).

Proof of Proposition B.1

The proof closely follows the proof of Theorem 4 in Gijbels, Omelka and Veraverbeke (2015).

Decomposition

Let us decompose the copula process $\sqrt{n}(\bar{C}_n - \bar{C}_n^{(or)})$ as

$$\sqrt{n}(\bar{C}_n - \bar{C}_n^{(or)}) = \bar{A}_n + \bar{B}_n + E_n, \tag{B.3}$$

where $\bar{A}_n = A_n - \mathbb{E} A_n$, with

$$A_n(u_1, u_2) = \frac{1}{\sqrt{n}} \sum_{i=1}^n \left[\mathbb{I}\{\hat{\varepsilon}_{1i} \leq \hat{F}_{1\hat{\varepsilon}}^{-1}(u_1), \hat{\varepsilon}_{2i} \leq \hat{F}_{2\hat{\varepsilon}}^{-1}(u_2)\} - \mathbb{I}\{\varepsilon_{1i} \leq \hat{F}_{1\varepsilon}^{-1}(u_1), \varepsilon_{2i} \leq \hat{F}_{2\varepsilon}^{-1}(u_2)\} \right],$$

and $\bar{B}_n = B_n - \mathbb{E} B_n$, with

$$B_n(u_1, u_2) = \frac{1}{\sqrt{n}} \sum_{i=1}^n \left[\mathbb{I}\{\varepsilon_{1i} \leq \hat{F}_{1\varepsilon}^{-1}(u_1), \varepsilon_{2i} \leq \hat{F}_{2\varepsilon}^{-1}(u_2)\} - \mathbb{I}\{\varepsilon_{1i} \leq \hat{F}_{1\varepsilon}^{-1}(u_1), \varepsilon_{2i} \leq \hat{F}_{2\varepsilon}^{-1}(u_2)\} \right].$$

Finally, E_n is given by

$$E_n(u_1, u_2) = \mathbb{E} A_n(u_1, u_2) + \mathbb{E} B_n(u_1, u_2) = \frac{1}{\sqrt{n}} \sum_{i=1}^n \left[\mathbb{E} \mathbb{I}\{\hat{\varepsilon}_{1i} \leq \hat{F}_{1\hat{\varepsilon}}^{-1}(u_1), \hat{\varepsilon}_{2i} \leq \hat{F}_{2\hat{\varepsilon}}^{-1}(u_2)\} - \mathbb{E} \mathbb{I}\{\varepsilon_{1i} \leq \hat{F}_{1\varepsilon}^{-1}(u_1), \varepsilon_{2i} \leq \hat{F}_{2\varepsilon}^{-1}(u_2)\} \right].$$

Completely analogously as in the proof of Theorem 4 in Gijbels, Omelka and Veraverbeke (2015) one can show that \bar{A}_n, \bar{B}_n are asymptotically negligible uniformly in (u_1, u_2) . Thus it remains to investigate the process E_n .

The process E_n

Note that

$$E_n(u_1, u_2) = \sqrt{n} \mathbb{E}_X \left[C_X \left(F_{1X}(\hat{F}_{1X}^{-1}(u_1)), F_{2X}(\hat{F}_{2X}^{-1}(u_2)) \right) - C_X \left(F_{1\varepsilon}(\hat{F}_{1\varepsilon}^{-1}(u)), F_{2\varepsilon}(\hat{F}_{2\varepsilon}^{-1}(u)) \right) \right], \tag{B.4}$$

where

$$F_{jX}(y) = F_{j\varepsilon} \left(\frac{y - m_j(X, \beta_j)}{\sigma_j(X, \gamma_j)} \right) \quad \text{and} \quad \hat{F}_{jX}(y) = \hat{F}_{j\varepsilon} \left(\frac{y - m_j(X, \hat{\beta}_{jn})}{\sigma_j(X, \hat{\gamma}_{jn})} \right), \quad j = 1, 2.$$

Applying the mean value theorem yields

$$E_n(u_1, u_2) = E_X C_X^{(1)}(u_{1X}, u_{2X})Y_{1X}^{(n)}(u_1) + E_X C_X^{(2)}(u_{1X}, u_{2X})Y_{2X}^{(n)}(u_2), \quad (\text{B.5})$$

where for given $x \in R_X$

$$Y_{jx}^{(n)}(u) = \sqrt{n} \left[F_{jx}(\widehat{F}_{jx}^{-1}(u)) - F_{j\varepsilon}(\widehat{F}_{j\varepsilon}^{-1}(u)) \right], \quad j = 1, 2, \quad (\text{B.6})$$

and the point u_{jx} lies between the points $F_{jx}(\widehat{F}_{jx}^{-1}(u))$ and $F_{j\varepsilon}(\widehat{F}_{j\varepsilon}^{-1}(u))$.

Suppose for a moment that

$$\begin{aligned} Y_{jx}^{(n)}(u) &= f_{j\varepsilon}(F_{j\varepsilon}^{-1}(u))F_{j\varepsilon}^{-1}(u) \left(\frac{\dot{\sigma}_j(x, \gamma_j)}{\sigma_j(x, \gamma_j)} - E_X \left[\frac{\dot{\sigma}_j(X, \gamma_j)}{\sigma_j(X, \gamma_j)} \right] \right) \sqrt{n} (\widehat{\gamma}_{jn} - \gamma_j) \\ &\quad + f_{j\varepsilon}(F_{j\varepsilon}^{-1}(u)) \left[\dot{m}_j(x, \beta_j) - E_X \dot{m}_j(X, \beta_j) \right] \sqrt{n} (\widehat{\beta}_{jn} - \beta_j) \\ &\quad + o_P(1) \end{aligned} \quad (\text{B.7})$$

holds uniformly in $x \in R_X$ and $u \in [0, 1]$ (this is proved below, see pages 2451–2454). Then thanks to assumptions $(\beta\gamma)$ and $(\mathbf{F2p})$ for each $\varepsilon > 0$ one can find $\delta > 0$ such that

$$P \left(\sup_{x \in R_X} \sup_{u \in [0, \delta] \cup [1 - \delta, 1]} |Y_{jx}^{(n)}(u)| > \varepsilon \right) < \varepsilon \quad (\text{B.8})$$

for all sufficiently large n . This further implies that it is sufficient that we deal with the term $E_X C_X^{(1)}(u_{1X}, u_{2X})Y_{1X}^{(n)}(u_1)$ ($E_X C_X^{(2)}(u_{1X}, u_{2X})Y_{2X}^{(n)}(u_2)$) on the right-hand side of (B.5) when $u_1 \in [\delta, 1 - \delta]$ ($u_2 \in [\delta, 1 - \delta]$). As u_{jx} lies between $F_{jx}(\widehat{F}_{jx}^{-1}(u_j))$ and $F_{j\varepsilon}(\widehat{F}_{j\varepsilon}^{-1}(u_j))$, (B.22) and (B.23) (proved below, see page 2454) yield that $u_{jx} = u_j + o_P(1)$ uniformly in $(x, u_j) \in R_X \times [0, 1]$. This together with assumption (\mathbf{Cp}) implies that

$$\sup_{x \in R_X} \sup_{(u_1, u_2) \in [\delta, 1 - \delta] \times [0, 1]} |C_x^{(1)}(u_{1x}, u_{2x}) - C_x^{(1)}(u_1, u_2)| = o_P(1), \quad (\text{B.9})$$

$$\sup_{x \in R_X} \sup_{(u_1, u_2) \in [0, 1] \times [\delta, 1 - \delta]} |C_x^{(2)}(u_{1x}, u_{2x}) - C_x^{(2)}(u_1, u_2)| = o_P(1), \quad (\text{B.10})$$

Now combining (B.7), (B.8), (B.9) and (B.10) yields that the right-hand side of (B.5) can be approximated by the right-hand side of (B.2) uniformly in $(u_1, u_2) \in [0, 1]^2$, which finishes the proof of the theorem.

Proving (B.7)

Using the mean value theorem one can calculate that

$$\begin{aligned} F_{jx}(\widehat{F}_{jx}^{-1}(u)) &= F_{j\varepsilon} \left(\frac{\widehat{F}_{j\varepsilon}^{-1}(u)\sigma_j(x, \widehat{\gamma}_{jn}) + m_j(x, \widehat{\beta}_{jn}) - m_j(x, \beta_j)}{\sigma_j(x, \gamma_j)} \right) \\ &= F_{j\varepsilon}(\widehat{F}_{j\varepsilon}^{-1}(u)) + f_{j\varepsilon}(y_j^x(u))\widehat{F}_{j\varepsilon}^{-1}(u) \left(\frac{\sigma_j(x, \widehat{\gamma}_{jn})}{\sigma_j(x, \gamma_j)} - 1 \right) \end{aligned}$$

$$\begin{aligned}
& + f_{j\varepsilon}(y_j^x(u)) \frac{m_j(x, \hat{\beta}_{jn}) - m_j(x, \beta_j)}{\sigma_j(x, \gamma_j)} \\
& = F_{j\varepsilon}(\hat{F}_{j\varepsilon}^{-1}(u)) + f_{j\varepsilon}(y_j^x(u)) \hat{F}_{j\varepsilon}^{-1}(u) \frac{\dot{\sigma}_j(x, \gamma_j^*)}{\sigma_j(x, \gamma_j)} (\hat{\gamma}_{jn} - \gamma_j) \\
& + f_{j\varepsilon}(y_j^x(u)) \frac{\dot{m}_j(x, \beta_j^*)}{\sigma_j(x, \gamma_j)} (\hat{\beta}_{jn} - \beta_j), \tag{B.11}
\end{aligned}$$

where $y_j^x(u)$ lies between $\hat{F}_{j\varepsilon}^{-1}(u)$ and $\frac{\hat{F}_{j\varepsilon}^{-1}(u)\sigma_j(x, \hat{\gamma}_{jn}) + m_j(x, \hat{\beta}_{jn}) - m_j(x, \beta_j)}{\sigma_j(x, \gamma_j)}$, γ_j^* between $\hat{\gamma}_{jn}$ and γ_j , and finally β_j^* between $\hat{\beta}_{jn}$ and β_j .

First note that by assumptions $(\beta\gamma)$ and $(\mathbf{m}\sigma\mathbf{p})$

$$\dot{m}_j(x, \beta_j^*) = \dot{m}_j(x, \beta_j) + o_P(1), \quad \dot{\sigma}_j(x, \gamma_j^*) = \dot{\sigma}_j(x, \gamma_j) + o_P(1) \tag{B.12}$$

uniformly in x .

To proceed further with the analysis of the right-hand side of (B.11), one needs to analyse the term $F_{j\varepsilon}(\hat{F}_{j\varepsilon}^{-1}(u))$. To do so it is useful to investigate the process $\{\hat{F}_{j\varepsilon}(F_{j\varepsilon}^{-1}(u)), u \in [0, 1]\}$. In the same way as in the proof of Theorem 4 in Gijbels, Omelka and Veraverbeke (2015) one can show that uniformly in $u \in [0, 1]$

$$\hat{F}_{j\varepsilon}(F_{j\varepsilon}^{-1}(u)) - \hat{F}_{j\varepsilon}(F_{j\varepsilon}^{-1}(u)) - \mathbb{E} \hat{F}_{j\varepsilon}(F_{j\varepsilon}^{-1}(u)) + u = o_P\left(\frac{1}{\sqrt{n}}\right). \tag{B.13}$$

This together with

$$\begin{aligned}
\mathbb{E} \hat{F}_{j\varepsilon}(F_{j\varepsilon}^{-1}(u)) & = \mathbb{E} \mathbb{I}\left\{ \frac{\varepsilon_j \sigma_j(X, \gamma_j) + m_j(X, \beta_j) - m_j(X, \hat{\beta}_{jn})}{\sigma_j(X, \hat{\gamma}_{jn})} \leq F_{j\varepsilon}^{-1}(u) \right\} \\
& = \mathbb{E}_X F_{j\varepsilon} \left(\frac{F_{j\varepsilon}^{-1}(u) \sigma_j(X, \hat{\gamma}_{jn}) + m_j(X, \hat{\beta}_{jn}) - m_j(X, \beta_j)}{\sigma_j(X, \gamma_j)} \right)
\end{aligned}$$

further implies that uniformly in $u \in [0, 1]$

$$\begin{aligned}
\hat{F}_{j\varepsilon}(F_{j\varepsilon}^{-1}(u)) & = \hat{F}_{j\varepsilon}(F_{j\varepsilon}^{-1}(u)) + \mathbb{E}_X f_{j\varepsilon}(z_j^X(u)) F_{j\varepsilon}^{-1}(u) \left[\frac{\sigma_j(X, \hat{\gamma}_{jn})}{\sigma_j(X, \gamma_j)} - 1 \right] \\
& + \mathbb{E}_X f_{j\varepsilon}(z_j^X(u)) \left[\frac{m_j(X, \hat{\beta}_{jn}) - m_j(X, \beta_j)}{\sigma_j(X, \gamma_j)} \right] + o_P\left(\frac{1}{\sqrt{n}}\right),
\end{aligned}$$

where $z_j^x(u)$ lies between $F_{j\varepsilon}^{-1}(u)$ and $\frac{F_{j\varepsilon}^{-1}(u)\sigma_j(x, \hat{\gamma}_{jn}) + m_j(x, \hat{\beta}_{jn}) - m_j(x, \beta_j)}{\sigma_j(x, \gamma_j)}$. Now thanks to assumptions $(\beta\gamma)$, $(\mathbf{m}\sigma\mathbf{p})$ and Lemma F.4 (see Appendix F)

$$\begin{aligned}
f_{j\varepsilon}(z_j^x(u)) & = f_{j\varepsilon}(F_{j\varepsilon}^{-1}(u)) + o_P(1), \\
f_{j\varepsilon}(z_j^x(u)) F_{j\varepsilon}^{-1}(u) & = f_{j\varepsilon}(F_{j\varepsilon}^{-1}(u)) F_{j\varepsilon}^{-1}(u) + o_P(1),
\end{aligned}$$

uniformly in u and x , which further yields that

$$\begin{aligned}
\hat{F}_{j\varepsilon}(F_{j\varepsilon}^{-1}(u)) & = \hat{F}_{j\varepsilon}(F_{j\varepsilon}^{-1}(u)) + f_{j\varepsilon}(F_{j\varepsilon}^{-1}(u)) F_{j\varepsilon}^{-1}(u) \mathbb{E}_X \left[\frac{\dot{\sigma}_j(X, \gamma_j)}{\sigma_j(X, \gamma_j)} \right] (\hat{\gamma}_{jn} - \gamma_j) \\
& + f_{j\varepsilon}(F_{j\varepsilon}^{-1}(u)) \mathbb{E}_X \left[\frac{\dot{m}_j(X, \beta_j)}{\sigma_j(X, \gamma_j)} \right] (\hat{\beta}_{jn} - \beta_j) + o_P\left(\frac{1}{\sqrt{n}}\right). \tag{B.14}
\end{aligned}$$

Now, thanks to (B.14) and assumptions $(\beta\gamma)$, $(\mathbf{F2p})$ and $(\mathbf{m\sigma p})$ one gets that the process

$$\left\{ \sqrt{n} [\widehat{F}_{j\widehat{\varepsilon}}(z) - F_{j\varepsilon}(z)], z \in \mathbb{R} \right\}$$

converges in distribution to a limiting process \mathbb{F}_j . Further, \mathbb{F}_j satisfies that $\mathbb{P}(\mathbb{F}_j \in \mathcal{A}) = 1$, where \mathcal{A} is a set of functions on $[0, 1]$ such that each function $\alpha \in \mathcal{A}$ can be written as $\alpha(x) = h(F_{j\varepsilon}(x))$, where h is a continuous function on $[0, 1]$ that meets $h(0) = h(1) = 0$.

Now, we are ready to use the Hadamard-differentiability of the functional $\widehat{F} \rightarrow F \circ \widehat{F}^{-1}$ at the point $\widehat{F} = F$ (see Lemma A.2 of Omelka, Gijbels and Veraverbeke, 2009). Put $\alpha_n(z) = \sqrt{n}[\widehat{F}_{j\widehat{\varepsilon}}(z) - F_{j\varepsilon}(z)]$ and rewrite

$$\sqrt{n} [F_{j\varepsilon}(\widehat{F}_{j\widehat{\varepsilon}}^{-1}(u)) - u] = \sqrt{n} [F_{j\varepsilon}((F_{j\varepsilon} + \frac{\alpha_n}{\sqrt{n}})^{-1}(u)) - u]. \tag{B.15}$$

Now, one can use Theorem 3.9.4 of van der Vaart and Wellner (1996) together with the Hadamard differentiability, (B.14) and (B.15) to deduce that uniformly in $u \in [0, 1]$

$$\begin{aligned} \sqrt{n} [F_{j\varepsilon}(\widehat{F}_{j\widehat{\varepsilon}}^{-1}(u)) - u] &= -\alpha_n(F_{j\varepsilon}^{-1}(u)) + o_P(1) \\ &= -\sqrt{n} [\widehat{F}_{j\varepsilon}(F_{j\varepsilon}^{-1}(u)) - u] - \sqrt{n} f_{j\varepsilon}(F_{j\varepsilon}^{-1}(u)) F_{j\varepsilon}^{-1}(u) \mathbf{E}_X \left[\frac{\dot{\sigma}_j(X, \gamma_j)}{\sigma_j(X, \gamma_j)} \right] (\widehat{\gamma}_{jn} - \gamma_j) \\ &\quad - \sqrt{n} f_{j\varepsilon}(F_{j\varepsilon}^{-1}(u)) \mathbf{E}_X \left[\frac{\dot{m}_j(X, \beta_j)}{\sigma_j(X, \gamma_j)} \right] (\widehat{\beta}_{jn} - \beta_j) + o_P(1). \end{aligned} \tag{B.16}$$

Now, it remains to deal with the terms $f_{j\varepsilon}(y_j^x(u))$ and $f_{j\varepsilon}(y_j^x(u)) \widehat{F}_{j\widehat{\varepsilon}}^{-1}(u)$ on the right-hand side of (B.11). First note that (B.16) together with $(\mathbf{F2p})$ implies that uniformly in $u \in [0, 1]$

$$f_{j\varepsilon}(\widehat{F}_{j\widehat{\varepsilon}}^{-1}(u)) = f_{j\varepsilon}(F_{j\varepsilon}^{-1}[F_{j\varepsilon}(\widehat{F}_{j\widehat{\varepsilon}}^{-1}(u))]) = f_{j\varepsilon}(F_{j\varepsilon}^{-1}(u)) + o_P(1) \tag{B.17}$$

and analogously

$$f_{j\varepsilon}(\widehat{F}_{j\widehat{\varepsilon}}^{-1}(u)) \widehat{F}_{j\widehat{\varepsilon}}^{-1}(u) = f_{j\varepsilon}(F_{j\varepsilon}^{-1}(u)) F_{j\varepsilon}^{-1}(u) + o_P(1). \tag{B.18}$$

Further note that $y_j^x(u)$ introduced in (B.11) satisfies $y_j^x(u) = \widehat{F}_{j\widehat{\varepsilon}}^{-1}(u)(1 + a_n) + b_n$, where both a_n and b_n are of order $o_P(1)$ (uniformly in u and x). Thus it holds uniformly in u, x

$$\begin{aligned} F_{j\varepsilon}(y_j^x(u)) &= F_{j\varepsilon}(\widehat{F}_{j\widehat{\varepsilon}}^{-1}(u)(1 + a_n) + b_n) \\ &= F_{j\varepsilon}(\widehat{F}_{j\widehat{\varepsilon}}^{-1}(u)) + f_{j\varepsilon}(\widehat{F}_{j\widehat{\varepsilon}}^{-1}(u)(1 + a_n^*) + b_n^*)(a_n \widehat{F}_{j\widehat{\varepsilon}}^{-1}(u) + b_n) \\ &= F_{j\varepsilon}(\widehat{F}_{j\widehat{\varepsilon}}^{-1}(u)) \\ &\quad + \frac{a_n}{1+a_n^*} f_{j\varepsilon}(\widehat{F}_{j\widehat{\varepsilon}}^{-1}(u)(1 + a_n^*) + b_n^*)(\widehat{F}_{j\widehat{\varepsilon}}^{-1}(u)(1 + a_n^*) + b_n^*) + o_P(1) \\ &= F_{j\varepsilon}(\widehat{F}_{j\widehat{\varepsilon}}^{-1}(u)) + o_P(1), \end{aligned} \tag{B.19}$$

where a_n^* (b_n^*) lies between zero and a_n (b_n). Now, (B.19) together with (B.17) and (B.18) implies that (uniformly in u and x)

$$f_{j\varepsilon}(y_j^x(u)) = f_{j\varepsilon}(F_{j\varepsilon}^{-1}(u)) + o_P(1), \quad (\text{B.20})$$

$$f_{j\varepsilon}(y_j^x(u)) \widehat{F}_{j\varepsilon}^{-1}(u) = f_{j\varepsilon}(F_{j\varepsilon}^{-1}(u)) F_{j\varepsilon}^{-1}(u) + o_P(1). \quad (\text{B.21})$$

Finally using (B.12), (B.17), (B.18), (B.20), (B.21) and approximation (B.16) for $F_{j\varepsilon}(\widehat{F}_{j\varepsilon}^{-1}(u))$ in (B.11) yields that uniformly in $u \in [0, 1]$ and $x \in R_X$

$$\begin{aligned} & \sqrt{n} \left[F_{jx}(\widehat{F}_{jx}^{-1}(u)) - u \right] \\ &= -\sqrt{n} \left[\widehat{F}_{j\varepsilon}(F_{j\varepsilon}^{-1}(u)) - u \right] \\ & \quad + f_{j\varepsilon}(F_{j\varepsilon}^{-1}(u)) F_{j\varepsilon}^{-1}(u) \left[\frac{\dot{\sigma}_j(x, \gamma_j)}{\sigma_j(x, \gamma_j)} - \mathbb{E}_X \frac{\dot{\sigma}_j(X, \gamma_j)}{\sigma_j(X, \gamma_j)} \right] \sqrt{n} (\widehat{\gamma}_{jn} - \gamma_j) \\ & \quad + f_{j\varepsilon}(F_{j\varepsilon}^{-1}(u)) [\dot{m}_j(x, \beta_j) - \mathbb{E}_X \dot{m}_j(X, \beta_j)] \sqrt{n} (\widehat{\beta}_{jn} - \beta_j) \\ & \quad + o_P(1). \end{aligned} \quad (\text{B.22})$$

Analogously to (B.16) using the Hadamard differentiability of the functional $\tilde{F} \rightarrow F \circ \tilde{F}^{-1}$ but considering $\alpha_n(z) = \sqrt{n} [\widehat{F}_{j\varepsilon}(z) - F_{j\varepsilon}(z)]$ one can show that

$$\sqrt{n} \left[F_{j\varepsilon}(\widehat{F}_{j\varepsilon}^{-1}(u)) - u \right] = -\sqrt{n} \left[\widehat{F}_{j\varepsilon}(F_{j\varepsilon}^{-1}(u)) - u \right] + o_P(1), \quad (\text{B.23})$$

which together with (B.6) and (B.22) yields (B.7).

Appendix C: Adjusting through nonparametric location-scale models

In this appendix we list the assumptions needed for Theorem 2, and we provide a proof for this result.

Regularity assumptions

Let $F_{j\varepsilon}$ stand for the distribution function of ε_j .

- (Cn) The second-order partial derivatives $C_x^{(1,1)}$, $C_x^{(1,2)}$ and $C_x^{(2,2)}$ of the copula function C_x , with $C_x^{(i,j)}(u_1, u_2) = \frac{\partial^2 C_x(u_1, u_2)}{\partial u_i \partial u_j}$, satisfy

$$\begin{aligned} C_x^{(1,1)}(u_1, u_2) &= O\left(\frac{1}{u_1(1-u_1)}\right), \quad C_x^{(2,2)}(u_1, u_2) = O\left(\frac{1}{u_2(1-u_2)}\right), \\ C_x^{(1,2)}(u_1, u_2) &= O\left(\frac{1}{\sqrt{u_1 u_2 (1-u_1)(1-u_2)}}\right) \end{aligned}$$

uniformly in $x \in R_X$.

Further $C_x^{(j)}(u_1, u_2)$ viewed as a function of x is continuously differentiable in R_X uniformly in $(u_1, u_2) \in [0, 1]^2$ for $j = 1, 2$.

(Bwn) For $j = 1, 2$ and for some $\delta > 0$:

$$\frac{n g_{jn}^{3+\delta}}{\log n} \rightarrow \infty \quad \text{and} \quad n g_{jn}^4 \rightarrow 0.$$

- (F1n)** The functions $F_{1\varepsilon}(y)$ and $F_{2\varepsilon}(y)$ are twice continuously differentiable and $E \varepsilon_j^4 < \infty$ for $j = 1, 2$.
- (F2n)** For $j = 1, 2$ the functions $F'_{j\varepsilon}(F_{j\varepsilon}^{-1}(u))(1 + F_{j\varepsilon}^{-1}(u))$ are of order $O(u^\delta(1 - u)^\delta)$ for some $\delta \geq \frac{1}{3}$ and the functions $F''_{j\varepsilon}(y)(1 + y)^2$ are bounded.
- (kn)** The kernel k is twice continuously differentiable, symmetric with support $[-1, 1]$, decreasing on $[0, 1)$, and integrates out to one.
- (Xn)** The support R_X of X is a non-empty finite interval (a, b) . Suppose that $\inf_{x \in R_X} f_X(x) > 0$ and $f_X(x)$ is twice continuously differentiable in R_X .
- (mσ)** For $j = 1, 2$ the functions m_j and σ_j are twice continuously differentiable on the interior of R_X and $\inf_{x \in R_X} \sigma_j(x) > 0$.

Remark 2. Note that assumption **(Bwn)** requires that $g_{jn} = o(n^{-1/4})$. As the optimal rate is usually $g_{jn} = O(n^{-1/5})$, assumption **(Bwn)** says that g_{jn} should converge to zero faster than at an optimal rate.

The proof of Theorem 2 follows from Proposition C.1 and standard results on the asymptotic representation of the empirical copula process applied to $\bar{C}_n^{(or)}$, that is given in (B.1).

Proposition C.1. *Assume that the assumptions of Theorem 2 are satisfied. Then uniformly in $(u_1, u_2) \in [0, 1]^2$*

$$\begin{aligned} & \sqrt{n} (\bar{C}_n(u_1, u_2) - \bar{C}_n^{(or)}(u_1, u_2)) \\ &= \sum_{j=1}^2 f_{j\varepsilon}(F_{j\varepsilon}^{-1}(u_j)) \frac{1}{\sqrt{n}} \sum_{i=1}^n \phi_j(X_i, U_{ji}, u_1, u_2) + o_P(1). \end{aligned} \tag{C.1}$$

Proof of Proposition C.1

As the proof is analogous to the proof of Proposition B.1 in Appendix B only the differences between these two proofs are briefly indicated.

In the same way as in the proof of Proposition B.1 one can use the decomposition (B.3) with the copula estimator \bar{C}_n and the distribution functions $\hat{F}_{j\varepsilon}$ based on the residuals given by (3.10).

Further, completely analogously as in the proof of Theorem 6 in Gijbels, Omelka and Veraverbeke (2015) one can show that \bar{A}_n, \bar{B}_n are asymptotically negligible uniformly in (u_1, u_2) . Thus it remains to investigate the process E_n .

Treatment of E_n

Note that E_n can be expressed as (B.4) with $F_{jx}(y)$ and $\widehat{F}_{jx}(y)$ given by

$$F_{jx}(y) = F_{j\varepsilon} \left(\frac{y - m_j(x)}{\sigma_j(x)} \right) \quad \text{and} \quad \widehat{F}_{jx}(y) = \widehat{F}_{j\varepsilon} \left(\frac{y - \widehat{m}_{jn}(x)}{\widehat{\sigma}_{jn}(x)} \right), \quad j = 1, 2. \quad (\text{C.2})$$

Now with the help of a second-order Taylor series expansion one gets

$$E_n(u_1, u_2) = \mathbb{E}_X C_X^{(1)}(u_1, u_2) Y_{1X}^{(n)}(u_1) + \mathbb{E}_X C_X^{(2)}(u_{1X}, u_{2X}) Y_{2X}^{(n)}(u_2) \quad (\text{C.3})$$

$$+ \frac{1}{2} \mathbb{E}_X \sum_{j=1}^2 \sum_{k=1}^2 C_X^{(j,k)}(u_{1X}, u_{2X}) Y_{jX}^{(n)}(u_j) Y_{kX}^{(n)}(u_k), \quad (\text{C.4})$$

where u_{jx} (for $j = 1, 2$) lies between the points $F_{jx}(\widehat{F}_{jx}^{-1}(u))$ and $F_{j\varepsilon}(\widehat{F}_{j\varepsilon}^{-1}(u))$ and $Y_{jx}^{(n)}$ is given in (B.6) (with $F_{jx}(y)$ and $\widehat{F}_{jx}(y)$ given by (C.2)).

Now completely analogously as in the proof of Theorem 6 of Gijbels, Omelka and Veraverbeke (2015) one can show that all the terms given in (C.4), are of order $o_P(\frac{1}{\sqrt{n}})$ uniformly in $(u_1, u_2) \in [0, 1]^2$. The only difference is that thanks to the results of Ojeda (2008) and assumption (Bwn) one gets

$$\sup_{x \in R_X} |m_j(x) - \widehat{m}_{jn}(x)| = o_P(n^{-a}), \quad \sup_{x \in R_X} |\sigma_j(x) - \widehat{\sigma}_{jn}(x)| = o_P(n^{-a}), \quad (\text{C.5})$$

with $a = \frac{1}{3}$ instead of $a = \frac{3}{8}$ as in (C5) of Gijbels, Omelka and Veraverbeke (2015). But this is compensated by requiring that $\delta > \frac{1}{3}$ in (F2n) (instead of $\delta > \frac{1}{4}$ as in (F2n) of Gijbels, Omelka and Veraverbeke (2015)) so that Lemma 3 of Gijbels, Omelka and Veraverbeke (2015) still holds true.

Thus it remains to investigate the right-hand side of (C.3). To proceed with the investigation of $Y_{jx}^{(n)}$ one needs to deal with the quantity $F_{jx}(\widehat{F}_{jx}^{-1}(u))$. By (F2n), (C.5) and a second-order Taylor series expansion one gets that uniformly in $u \in [0, 1]$ and $x \in R_X$

$$\begin{aligned} F_{jx}(\widehat{F}_{jx}^{-1}(u)) &= F_{j\varepsilon} \left(\frac{\widehat{F}_{j\varepsilon}^{-1}(u) \widehat{\sigma}_{jn}(x) + m_j(x) - \widehat{m}_{jn}(x)}{\sigma_j(x)} \right) \\ &= F_{j\varepsilon}(\widehat{F}_{j\varepsilon}^{-1}(u)) + f_{j\varepsilon}(\widehat{F}_{j\varepsilon}^{-1}(u)) \widehat{F}_{j\varepsilon}^{-1}(u) \left(\frac{\widehat{\sigma}_{jn}(x)}{\sigma_j(x)} - 1 \right) \\ &\quad + f_{j\varepsilon}(\widehat{F}_{j\varepsilon}^{-1}(u)) \frac{\widehat{m}_{jn}(x) - m_j(x)}{\sigma_j(x)} + o_P\left(\frac{1}{\sqrt{n}}\right). \end{aligned} \quad (\text{C.6})$$

To go on one needs to treat the quantity $F_{j\varepsilon}(\widehat{F}_{j\varepsilon}^{-1}(u))$. Analogously as in the proof of Proposition B.1 it is useful to investigate $\widehat{F}_{j\varepsilon}(F_{j\varepsilon}^{-1}(u))$. In the same way as in the proof of Theorem 6 in Gijbels, Omelka and Veraverbeke (2015) one can show that (B.13) holds. This together with

$$\mathbb{E} \widehat{F}_{j\varepsilon}(F_{j\varepsilon}^{-1}(u)) = \mathbb{E}_X F_{j\varepsilon} \left(\frac{F_{j\varepsilon}^{-1}(u) \widehat{\sigma}_j(X) + \widehat{m}_{jn}(X) - m_j(X)}{\sigma_j(X)} \right),$$

assumptions **(mσ)**, **(F1n)**, **(F2n)**, **(C.5)**, and Lemmas **F.2** and **F.3** (see Appendix F) yields that uniformly in $u \in [0, 1]$

$$\begin{aligned} \widehat{F}_{j\widehat{\varepsilon}}(F_{j\widehat{\varepsilon}}^{-1}(u)) &= \widehat{F}_{j\varepsilon}(F_{j\varepsilon}^{-1}(u)) + f_{j\varepsilon}(F_{j\varepsilon}^{-1}(u))F_{j\varepsilon}^{-1}(u) \mathbb{E}_X \left[\frac{\widehat{\sigma}_{jn}(X)}{\sigma_j(X)} - 1 \right] \\ &\quad + f_{j\varepsilon}(F_{j\varepsilon}^{-1}(u)) \mathbb{E}_X \left[\frac{\widehat{m}_{jn}(X) - m_j(X)}{\sigma_j(X)} \right] + o_P\left(\frac{1}{\sqrt{n}}\right) \\ &= \widehat{F}_{j\varepsilon}(F_{j\varepsilon}^{-1}(u)) + f_{j\varepsilon}(F_{j\varepsilon}^{-1}(u))F_{j\varepsilon}^{-1}(u) \frac{1}{n} \sum_{i=1}^n \frac{\varepsilon_{ji}^2 - 1}{2} \\ &\quad + f_{j\varepsilon}(F_{j\varepsilon}^{-1}(u)) \frac{1}{n} \sum_{i=1}^n \varepsilon_{ji} + o_P\left(\frac{1}{\sqrt{n}}\right). \end{aligned}$$

Now, analogously as in the proof of Proposition **B.1** one can use the Hadamard-differentiability of the functional $\tilde{F} \rightarrow F \circ \tilde{F}^{-1}$ at the point $\tilde{F} = F$ to deduce that uniformly in $u \in [0, 1]$

$$\begin{aligned} \sqrt{n} \left[F_{j\widehat{\varepsilon}}(\widehat{F}_{j\widehat{\varepsilon}}^{-1}(u)) - u \right] &= -\sqrt{n} \left[\widehat{F}_{j\varepsilon}(F_{j\varepsilon}^{-1}(u)) - u \right] \\ &\quad - f_{j\varepsilon}(F_{j\varepsilon}^{-1}(u))F_{j\varepsilon}^{-1}(u) \frac{1}{\sqrt{n}} \sum_{i=1}^n \frac{\varepsilon_{ji}^2 - 1}{2} \quad (\text{C.7}) \\ &\quad - f_{j\varepsilon}(F_{j\varepsilon}^{-1}(u)) \frac{1}{\sqrt{n}} \sum_{i=1}^n \varepsilon_{ji} + o_P(1). \end{aligned}$$

Note that **(C.7)** together with **(F2n)** also yields that **(B.17)** and **(B.18)** hold uniformly in $u \in [0, 1]$.

Now using **(B.17)** and **(B.18)** and approximation **(C.7)** for $F_{j\widehat{\varepsilon}}(\widehat{F}_{j\widehat{\varepsilon}}^{-1}(u))$ in **(C.6)** gives that uniformly in $u \in [0, 1]$ and $x \in R_X$

$$\begin{aligned} \sqrt{n} \left[F_{jx}(\widehat{F}_{jx}^{-1}(u)) - u \right] &= -\sqrt{n} \left[\widehat{F}_{j\varepsilon}(F_{j\varepsilon}^{-1}(u)) - u \right] \\ &\quad + f_{j\varepsilon}(F_{j\varepsilon}^{-1}(u))F_{j\varepsilon}^{-1}(u) \sqrt{n} \left[\frac{\widehat{\sigma}_{jn}(x)}{\sigma_j(x)} - 1 - \frac{1}{n} \sum_{i=1}^n \frac{\varepsilon_{ji}^2 - 1}{2} \right] \\ &\quad + f_{j\varepsilon}(F_{j\varepsilon}^{-1}(u)) \sqrt{n} \left[\frac{\widehat{m}_{jn}(x) - m_j(x)}{\sigma_j(x)} - \frac{1}{n} \sum_{i=1}^n \varepsilon_{ji} \right] \\ &\quad + o_P(1). \end{aligned}$$

Analogously one can show that **(B.23)** holds, which together with **(B.22)** yields that uniformly in $u \in [0, 1]$ and $x \in R_X$

$$\begin{aligned} Y_{jx}^{(n)}(u) &= f_{j\varepsilon}(F_{j\varepsilon}^{-1}(u))F_{j\varepsilon}^{-1}(u) \sqrt{n} \left[\frac{\widehat{\sigma}_{jn}(x)}{\sigma_j(x)} - 1 - \frac{1}{n} \sum_{i=1}^n \frac{\varepsilon_{ji}^2 - 1}{2} \right] \quad (\text{C.8}) \\ &\quad + f_{j\varepsilon}(F_{j\varepsilon}^{-1}(u)) \sqrt{n} \left[\frac{\widehat{m}_{jn}(x) - m_j(x)}{\sigma_j(x)} - \frac{1}{n} \sum_{i=1}^n \varepsilon_{ji} \right] + o_P(1). \end{aligned}$$

Now combining (C.8) together with Lemmas F.2 and F.3 yields that uniformly in (u_1, u_2)

$$\mathbb{E}_X C_X^{(j)}(u_1, u_2) Y_{jX}^{(n)}(u_j) = \frac{1}{\sqrt{n}} \sum_{i=1}^n \phi_j(X_i, \varepsilon_{ji}, u_1, u_2) + o_P(1),$$

where ϕ_j is defined in (5.5). Thus the right-hand side of (C.3) can be approximated by the right-hand side of (C.1) uniformly in $(u_1, u_2) \in [0, 1]^2$, which finishes the proof of the proposition.

Appendix D: Adjusting through general nonparametric estimators of F_{1x} and F_{2x}

The regularity assumptions needed in Theorem 3 are provided below.

Regularity assumptions

- (Bw) For $j = 1, 2$: $g_{jn} = o(1)$ and $n g_{jn} \rightarrow \infty$.
- (F) The second derivatives of the functions $F_{1x}(y)$ and $F_{2x}(y)$ with respect to x are bounded.
- (k) The kernel k is a symmetric and continuous function with support $[-1, 1]$.
- (X) The support R_X of X is a non-empty finite interval (a, b) .

The proof of Theorem 3 is completely analogous to the proof of Theorem 1 in Gijbels, Omelka and Veraverbeke (2015).

Appendix E: Asymptotic result for the average conditional Kendall's tau

We provide the technical details concerning the asymptotic result for the average conditional Kendall's tau, stated in Theorem 7.

- (H) The joint conditional distribution function $H_x(y_1, y_2)$ has continuous marginal conditional distribution functions F_{1x} and F_{2x} . Further $H_x(y_1, y_2)$ is twice continuously differentiable as a function of $x \in R_X$ uniformly in $(y_1, y_2) \in \mathbb{R}^2$.

Proof of Theorem 7

Analogously as in Lemma 4 of Gijbels, Omelka and Veraverbeke (2015) with the help of assumptions (Bwn), (Xn) and (kn) one can show that

$$\sup_{x \in R_X} \sum_{i=1}^n w_{ni}^2(x, h_n) = O_P\left(\frac{\sqrt{\log n}}{n h_n}\right) = o_P\left(\frac{1}{\sqrt{n}}\right).$$

This further yields that

$$\widehat{\tau}_n(x) = 4 \sum_{j=1}^n \sum_{k=1}^n w_{nj}(x, h_n) w_{nk}(x, h_n) \mathbb{I}\{Y_{1j} < Y_{1k}, Y_{2j} < Y_{2k}\} - 1 + o_P\left(\frac{1}{\sqrt{n}}\right)$$

uniformly in $x \in R_X$. So one can approximate τ_n^A as

$$\begin{aligned} \tau_n^A &= \frac{4}{n} \sum_{i=1}^n \sum_{j=1}^n \sum_{k=1}^n w_{nj}(X_i, h_n) w_{nk}(X_i, h_n) \mathbb{I}\{Y_{1j} < Y_{1k}, Y_{2j} < Y_{2k}\} - 1 \\ &\quad + o_P\left(\frac{1}{\sqrt{n}}\right). \end{aligned}$$

Further, let $\tau(x)$ stand for the conditional Kendall's tau at the point x and note that $\tau(x) = 4t(x) - 1$, where

$$t(x) = \iint H_x(y_1, y_2) dH_x(y_1, y_2).$$

Now, put

$$v(y_1, y_2, y'_1, y'_2, x) = [\mathbb{I}\{y_1 < y'_1, y_2 < y'_2\} - t(x)] \tag{E.1}$$

and note that one can decompose

$$\begin{aligned} \tau_n^A - \tau^A &= \frac{4}{n} \sum_{i=1}^n \sum_{j=1}^n \sum_{k=1}^n w_{nj}(X_i, h_n) w_{nk}(X_i, h_n) v(Y_{1j}, Y_{2j}, Y_{1k}, Y_{2k}, X_i) \\ &\quad + \frac{4}{n} \sum_{i=1}^n t(X_i) - 1 - \tau^A + o_P\left(\frac{1}{\sqrt{n}}\right). \end{aligned} \tag{E.2}$$

$$\equiv 4V_n + \frac{1}{n} \sum_{i=1}^n [\tau(X_i) - \tau^A] + o_P\left(\frac{1}{\sqrt{n}}\right), \tag{E.3}$$

where V_n is given by the first term (except for the factor 4) of the right-hand side of (E.2).

For $x \in R_X$ put

$$D_n(x) = S_{n,2}(x)S_{n,0}(x) - S_{n,1}^2(x), \tag{E.4}$$

where $S_{n,j}(x)$ is introduced in (3.9). Note that V_n can be rewritten as

$$V_n = \frac{1}{n} \sum_{i=1}^n [V_{1n}(X_i) + V_{2n}(X_i) + V_{3n}(X_i) + V_{4n}(X_i)], \tag{E.5}$$

where

$$V_{1n}(x) = \frac{S_{n,2}^2(x)}{D_n^2(x)n^2} \sum_{j=1}^n \sum_{k=1}^n \frac{1}{h_n^2} k\left(\frac{x-X_j}{h_n}\right) k\left(\frac{x-X_k}{h_n}\right) v(Y_{1j}, Y_{1k}, Y_{2j}, Y_{2k}, x),$$

$$\begin{aligned}
 V_{2n}(x) &= \frac{-S_{n,2}(x)S_{n,1}(x)}{D_n^2(x)n^2} \sum_{j=1}^n \sum_{k=1}^n \frac{x-X_k}{h_n^3} k\left(\frac{x-X_j}{h_n}\right) k\left(\frac{x-X_k}{h_n}\right) v(Y_{1j}, Y_{1k}, Y_{2j}, Y_{2k}, x), \\
 V_{3n}(x) &= \frac{-S_{n,2}(x)S_{n,1}(x)}{D_n^2(x)n^2} \sum_{j=1}^n \sum_{k=1}^n \frac{x-X_j}{h_n^2} k\left(\frac{x-X_j}{h_n}\right) k\left(\frac{x-X_k}{h_n}\right) v(Y_{1j}, Y_{1k}, Y_{2j}, Y_{2k}, x), \\
 V_{4n}(x) &= \frac{S_{n,1}^2(x)}{D_n^2(x)n^2} \sum_{j=1}^n \sum_{k=1}^n \frac{x-X_j}{h_n^2} \frac{x-X_k}{h_n^2} k\left(\frac{x-X_j}{h_n}\right) k\left(\frac{x-X_k}{h_n}\right) v(Y_{1j}, Y_{1k}, Y_{2j}, Y_{2k}, x),
 \end{aligned}$$

and the function $v(y_1, y_2, y'_1, y'_2, x)$ is defined in (E.1).

Below (see pages 2462–2463) we show that uniformly in $x \in R_X$

$$\begin{aligned}
 V_{1n}(x) &= \frac{1}{f_X^2(x)n^2} \sum_{j=1}^n \sum_{k=1}^n \frac{1}{h_n} k\left(\frac{x-X_j}{h_n}\right) \frac{1}{h_n} k\left(\frac{x-X_k}{h_n}\right) v(Y_{1j}, Y_{1k}, Y_{2j}, Y_{2k}, x) \\
 &\quad + o_P\left(\frac{1}{\sqrt{n}}\right)
 \end{aligned} \tag{E.6}$$

and

$$V_{jn}(x) = o_P\left(\frac{1}{\sqrt{n}}\right), \quad j = 2, 3, 4. \tag{E.7}$$

Combining (E.5), (E.6), (E.7) now yields

$$V_n = \frac{1}{n^3} \sum_{i=1}^n \sum_{j=1}^n \sum_{k=1}^n h(Y_{1i}, Y_{2i}, X_i, Y_{1j}, Y_{2j}, X_j, Y_{1k}, Y_{2k}, X_k) + o_P\left(\frac{1}{\sqrt{n}}\right),$$

where

$$\begin{aligned}
 h(y_1, y_2, x, y'_1, y'_2, x', y''_1, y''_2, x'') &= \frac{1}{f_X^2(x)} \frac{1}{h_n} k\left(\frac{x-x'}{h_n}\right) \frac{1}{h_n} k\left(\frac{x-x''}{h_n}\right) \\
 &\quad \times [\mathbb{I}\{y'_1 < y''_1, y'_2 < y''_2\} - t(x)].
 \end{aligned}$$

Note that h in fact does not depend on y_1 and y_2 . The reason why we put these arguments is to stress that V_n is a V -statistic (see e.g. Chapter 5.1.2 Serfling, 1980) with a kernel h of degree 3. Although the kernel is not symmetric, it could be easily symmetrized without affecting the quantity V_n . Further, as V -statistics and U -statistics are in terms of the \sqrt{n} -asymptotic distribution equivalent (Chapter 5.7.3 of Serfling, 1980), in what follows we can think of V_n as a U -statistic.

In the following the expectation \mathbf{E} is taken with respect to the distribution of the random vectors (Y'_1, Y'_2, X') and (Y''_1, Y''_2, X'') that are two independent copies of (Y_1, Y_2, X) .

Denote the projections of h as

$$\begin{aligned}
 h^{\pi_1}(y_1, y_2, x) &= \mathbf{E} h(y_1, y_2, x, Y'_1, Y'_2, X', Y''_1, Y''_2, X''), \\
 h^{\pi_2}(y_1, y_2, x) &= \mathbf{E} h(Y'_1, Y'_2, X', y_1, y_2, x, Y''_1, Y''_2, X''), \\
 h^{\pi_3}(y_1, y_2, x) &= \mathbf{E} h(Y'_1, Y'_2, X', Y''_1, Y''_2, X'', y_1, y_2, x).
 \end{aligned}$$

Note that thanks to assumptions **(Bwn)**, **(kn)**, **(Xn)** and **(H)** it holds uniformly in $(y_1, y_2, x) \in \mathbb{R}^2 \times R_X$

$$\begin{aligned} h^{\pi_1}(y_1, y_2, x) &= \frac{1}{f_X^2(x)} \mathbb{E} \left\{ \frac{1}{h_n} k\left(\frac{x-X'}{h_n}\right) \frac{1}{h_n} k\left(\frac{x-X''}{h_n}\right) \right. \\ &\quad \left. \times [\mathbb{I}\{Y_1' < Y_1'', Y_2' < Y_2''\} - t(x)] \right\} \\ &= \frac{1}{f_X^2(x)} \mathbb{E} \left\{ \frac{1}{h_n} k\left(\frac{x-X'}{h_n}\right) [H_x(Y_1'', Y_2'') - t(x)] f_X(x) \right\} + o\left(\frac{1}{\sqrt{n}}\right) \\ &= o\left(\frac{1}{\sqrt{n}}\right). \end{aligned} \tag{E.8}$$

Analogously one can derive that it holds uniformly in (y_1, y_2, x)

$$\begin{aligned} h^{\pi_2}(y_1, y_2, x) &= \mathbb{E} \left\{ \frac{1}{f_X^2(X')} \frac{1}{h_n} k\left(\frac{X'-x}{h_n}\right) \frac{1}{h_n} k\left(\frac{X'-X''}{h_n}\right) \right. \\ &\quad \left. \times [\mathbb{I}\{y_1 < Y_1'', y_2 < Y_2''\} - t(X')] \right\} \\ &= \mathbb{E} \left\{ \frac{1}{f_X^2(X')} \frac{1}{h_n} k\left(\frac{X'-x}{h_n}\right) \frac{1}{h_n} k\left(\frac{X'-X''}{h_n}\right) \right. \\ &\quad \left. \times [1 - F_{1X''}(y_1) - F_{2X''}(y_2) + H_{X''}(y_1, y_2) - t(X')] \right\} \\ &= \mathbb{E}_{X'} \left\{ \frac{1}{f_X(X')} \frac{1}{h_n} k\left(\frac{X'-x}{h_n}\right) \right. \\ &\quad \left. \times [1 - F_{1X'}(y_1) - F_{2X'}(y_2) + H_{X'}(y_1, y_2) - t(X')] \right\} + o\left(\frac{1}{\sqrt{n}}\right) \\ &= 1 - F_{1x}(y_1) - F_{2x}(y_2) + H_x(y_1, y_2) - t(x) + o\left(\frac{1}{\sqrt{n}}\right), \end{aligned} \tag{E.9}$$

$$h^{\pi_3}(y_1, y_2, x) = H_x(y_1, y_2) - t(x) + o\left(\frac{1}{\sqrt{n}}\right), \tag{E.10}$$

and

$$\bar{h} = \mathbb{E} h(Y_1, Y_2, X, Y_1', Y_2', X', Y_1'', Y_2'', X'') = o\left(\frac{1}{\sqrt{n}}\right). \tag{E.11}$$

Now put

$$\begin{aligned} \tilde{h}(y_1, y_2, x, y_1', y_2', x') &= h(y_1, y_2, x, y_1', y_2', x', y_1'', y_2'', x'') \\ &\quad - h^{\pi_1}(y_1, y_2, x) - h^{\pi_2}(y_1', y_2', x') - h^{\pi_3}(y_1'', y_2'', x'') + 2\bar{h}. \end{aligned}$$

Using the notation introduced above one can decompose V_n as

$$\begin{aligned} V_n &= \frac{1}{n} \sum_{j=1}^n [h^{\pi_1}(Y_{1j}, Y_{2j}, X_j) + h^{\pi_2}(Y_{1j}, Y_{2j}, X_j) + h^{\pi_3}(Y_{1j}, Y_{2j}, X_j) - 2\bar{h}] \\ &\quad + \frac{1}{n(n-1)} \sum_{j=1}^n \sum_{k=1}^n \tilde{h}(Y_{1j}, Y_{2j}, X_j, Y_{1k}, Y_{2k}, X_k) + o_P\left(\frac{1}{\sqrt{n}}\right) \\ &\equiv V_n^\pi + R_n + o_P\left(\frac{1}{\sqrt{n}}\right), \end{aligned} \tag{E.12}$$

where V_n^π is the first term on the right-hand side of (E.12).

With the help of assumptions **(Bwn)**, **(kn)**, **(Xn)** and Lemma A of Section 5.1 of Serfling (1980) one gets that for sufficiently large constant K and for all sufficiently large $n \in \mathbb{N}$

$$\text{var}(\sqrt{n} R_n) \leq \frac{K}{n} \mathbb{E} [h(Y_1, Y_2, \dots, X'')]^2 = O\left(\frac{1}{nh_n^2}\right) = o(1),$$

which together with $\mathbb{E} R_n = 0$, Chebyshev's inequality implies that $R_n = o_P\left(\frac{1}{\sqrt{n}}\right)$.

Thus one can concentrate on V_n^π . Combining (E.8), (E.9), (E.10), (E.11) yields that

$$V_n^\pi = \frac{2}{n} \sum_{i=1}^n [C_{X_i}(U_{1i}, U_{2i}) - t(X_i)] + \frac{1}{n} \sum_{i=1}^n [1 - U_{1i} - U_{2i}] + o_P\left(\frac{1}{\sqrt{n}}\right),$$

which together with (E.3) implies

$$\begin{aligned} \tau_n^A - \tau^A &= \frac{2}{n} \sum_{i=1}^n [4C_{X_i}(U_{1i}, U_{2i}) - 1 - \tau(X_i)] + \frac{4}{n} \sum_{i=1}^n [1 - U_{1i} - U_{2i}] \\ &\quad + \frac{1}{n} \sum_{i=1}^n [\tau(X_i) - \tau^A] + o_P\left(\frac{1}{\sqrt{n}}\right), \\ &= \frac{2}{n} \sum_{i=1}^n [4C_{X_i}(U_{1i}, U_{2i}) - 1 - \tau^A] + \frac{4}{n} \sum_{i=1}^n [1 - U_{1i} - U_{2i}] \\ &\quad - \frac{1}{n} \sum_{i=1}^n [\tau(X_i) - \tau^A] + o_P\left(\frac{1}{\sqrt{n}}\right) \end{aligned}$$

and finishes the proof of the theorem.

Proof of statements (E.6) and (E.7)

First note that analogously as in Lemma 4 of Gijbels, Omelka and Veraverbeke (2015) using assumptions **(Bwn)**, **(kn)** and **(Xn)** one can show that uniformly in $x \in R_X$

$$\begin{aligned} S_{n,0}(x) &= f_X(x) + O_P(r_n), & S_{n,1}(x) &= O_P(r_n), \\ S_{n,2}(x) &= \mu_{2k} f_X(x) + O_P(r_n), \end{aligned} \tag{E.13}$$

where $r_n = \sqrt{\frac{\log n}{nh_n}}$ and $\mu_{2k} = \int u^2 k(u) du$.

Treatment of V_{1n} . (E.13) and the definition of $D_n(x)$ in (E.4) imply that uniformly in $x \in R_X$

$$\frac{S_{n,2}^2(x)}{D_n^2(x)} = \frac{1}{f_X^2(x)} + O_P\left(\sqrt{\frac{\log n}{nh_n}}\right). \tag{E.14}$$

Further by Lemma F.5 in Appendix F

$$\sup_{x \in R_X} |U_n(x) - \mathbb{E} U_n(x)| = o_P\left(\sqrt{\frac{h_n}{\log n}}\right), \tag{E.15}$$

where

$$U_n(x) = \frac{1}{n^2} \sum_{j=1}^n \sum_{k=1}^n \frac{1}{h_n} k\left(\frac{x-X_j}{h_n}\right) \frac{1}{h_n} k\left(\frac{x-X_k}{h_n}\right) \mathbb{I}\{Y_{1j} < Y_{1k}, Y_{2j} < Y_{2k}\}.$$

By a straightforward calculation uniformly in $x \in R_X$

$$\mathbb{E} U_n(x) = f_X^2(x)t(x) + O(h_n^2) \tag{E.16}$$

and

$$\begin{aligned} \frac{1}{n^2} \sum_{j=1}^n \sum_{k=1}^n \frac{1}{h_n} k\left(\frac{x-X_j}{h_n}\right) \frac{1}{h_n} k\left(\frac{x-X_k}{h_n}\right) t(x) &= S_{n,0}^2(x) t(x) \\ &= f_X^2(x)t(x) + O_P\left(\sqrt{\frac{\log n}{n h_n}}\right). \end{aligned} \tag{E.17}$$

Now combining (E.15), (E.16), (E.17) yields

$$\begin{aligned} \frac{1}{n^2} \sum_{j=1}^n \sum_{k=1}^n \frac{1}{h_n} k\left(\frac{x-X_j}{h_n}\right) \frac{1}{h_n} k\left(\frac{x-X_k}{h_n}\right) [\mathbb{I}\{Y_{1j} < Y_{1k}, Y_{2j} < Y_{2k}\} - t(x)] \\ = o_P\left(\sqrt{\frac{h_n}{\log n}}\right) \end{aligned}$$

uniformly in $x \in R_X$, which together with (E.14) yields (E.6).

Treatment of V_{jn} , $j = 2, 3, 4$. Let us start with $j = 2$. Analogously as before thanks to Lemma F.5 of Appendix F one gets (E.15) with $U_n(x)$ given by

$$U_n(x) = \frac{1}{n^2} \sum_{j=1}^n \sum_{k=1}^n \frac{1}{h_n} k\left(\frac{x-X_j}{h_n}\right) \frac{x-X_k}{h_n^2} k\left(\frac{x-X_k}{h_n}\right) \mathbb{I}\{Y_{1j} < Y_{1k}, Y_{2j} < Y_{2k}\}. \tag{E.18}$$

Further uniformly in $x \in R_X$

$$\mathbb{E} U_n(x) = O(h_n) \tag{E.19}$$

and with the help of (E.13)

$$\begin{aligned} \frac{1}{n^2} \sum_{j=1}^n \sum_{k=1}^n \frac{1}{h_n} k\left(\frac{x-X_j}{h_n}\right) \frac{x-X_k}{h_n^2} k\left(\frac{x-X_k}{h_n}\right) t(x) &= S_{n,0}(x) S_{n,1}(x) t(x) \\ &= O_P\left(\sqrt{\frac{\log n}{n h_n}}\right). \end{aligned} \tag{E.20}$$

Combining (E.15), (E.18), (E.19) and (E.20) yields (E.7) for $j = 2$.

Analogously one can prove (E.7) for $j = 3$ and 4 .

Appendix F: Auxiliary results

In this appendix we state auxiliary results that are used in the proofs of the main theorems and that can also be of independent interest. The results will be formulated with the help of assumptions defined in the previous appendices. To simplify the notation the index j is usually dropped.

The following lemma is a simple adaptation of Lemma 19.24 of van der Vaart (2000), and will be useful in the proofs of Lemmas F.2 and F.3.

Lemma F.1. *Suppose that $\mathbf{X}_1, \mathbf{X}_2, \dots$ are identically distributed random vectors with distribution P and \mathcal{H} is the set of real valued uniformly bounded functions defined on $S_{\mathbf{X}}$ that is P -Donsker. Further let g be a real valued function on $S_{\mathbf{X}}$ with a finite second moment and $\{h_n\}$ be a sequence of real valued functions on $S_{\mathbf{X}}$ such that*

$$\lim_{n \rightarrow \infty} P(h_n \in \mathcal{H}) = 1, \quad \sup_{\mathbf{x} \in S_{\mathbf{X}}} |h_n(\mathbf{x})| = o_P(1), \quad \text{and} \quad E g(\mathbf{X}_1) h_n(\mathbf{X}_1) = 0.$$

Then $\frac{1}{n} \sum_{i=1}^n g(\mathbf{X}_i) h_n(\mathbf{X}_i) = o_P\left(\frac{1}{\sqrt{n}}\right)$.

Proof. Thanks to assumption of the lemma one can suppose that $h_n \in \mathcal{H}$ and that h_n is bounded. Further by the permanence of Donsker property also the set of functions $\mathcal{F} = \{g h, h \in \mathcal{H}\}$ is Donsker. The proof now follows by applying Lemma 19.24 of van der Vaart (2000) with $f_n(\mathbf{x}) = g(\mathbf{x})h_n(\mathbf{x})$ and f_0 taken to be a zero function. \square

Further in the proofs of Lemmas F.2 and F.3 we make use of the bracketing numbers for sets of differentiable functions. Following the notation of Chapter 2.7 of van der Vaart and Wellner (1996) let $C_1^M(R_X)$ stand for a set of real valued functions defined on R_X that are Lipschitz of order 1, with the Lipschitz constant bounded with M , that is

$$\sup_{h \in C_1^M(R_X)} \sup_{x \in R_X} |h'(x)| \leq M.$$

By Corollary 2.7.2 of van der Vaart and Wellner (1996) there exists a constant K such that the logarithm of the bracketing number $N_{[\cdot]}(\varepsilon, C_1^M(R_X), L_2(P))$ of the set $C_1^M(R_X)$ (with $L_2(P)$ denoting the norm that is used) is bounded by

$$\log N_{[\cdot]}(\varepsilon, C_1^M(R_X), L_2(P)) \leq \frac{K}{\varepsilon} \tag{F.1}$$

for every $\varepsilon > 0$ and probability measure P on R_X . Note that (F.1) implies that $C_1^M(R_X)$ is Donsker (see e.g. Theorem 19.5 in van der Vaart, 2000).

Lemma F.2. *Suppose that our observations follow the nonparametric location scale models described at the beginning of Section 3.4.2 and $\hat{m}_n(X)$ be given by (3.7). Further let the assumptions (Bwn), (F1n), (kn), (Xn), and (m σ) hold and $b : R_X \rightarrow \mathbb{R}$ possesses a bounded derivative on R_X . Then*

$$E_X \left\{ b(X) (\hat{m}_n(X) - m(X)) \right\} = \frac{1}{n} \sum_{i=1}^n b(X_i) \sigma(X_i) \varepsilon_i + o_P\left(\frac{1}{\sqrt{n}}\right).$$

Proof.

$$\begin{aligned} & \mathbb{E}_X \left\{ b(X) (\widehat{m}_n(X) - m(X)) \right\} \\ &= \mathbb{E}_X \left\{ b(X) \sum_{i=1}^n w_{ni}(X, g_n) (Y_i - m(X)) \right\} \end{aligned} \tag{F.2}$$

$$\begin{aligned} &= \mathbb{E}_X \left\{ b(X) \sum_{i=1}^n w_{ni}(X, g_n) (m(X_i) - m(X)) \right\} \\ &\quad + \mathbb{E}_X \left\{ b(X) \sum_{i=1}^n w_{ni}(X, g_n) \sigma(X_i) \varepsilon_i \right\} \end{aligned} \tag{F.3}$$

$$\equiv A_n + B_n,$$

where A_n and B_n are the first and second terms on the right-hand side of (F.3) respectively.

Now to deal with A_n , which stands for the first term on the right-hand side of (F.3), one can use a second-order Taylor series expansion of $m(X_i)$ with respect to X_i at the point X which together with the properties of the local linear weights yields

$$A_n = \mathbb{E}_X \left\{ b(X) \sum_{i=1}^n w_{ni}(X, g_n) \frac{m''(X_i^*)}{2} (X - X_i)^2 \right\},$$

where X_i^* lies between X_i and X . Note that m'' is bounded and $w_{ni}(X, g_n) = 0$ for $|X_i - X| > g_n$. Further, following the arguments of Section 2.4.1 of Omelka, Veraverbeke and Gijbels (2013) one can easily show that $\sum_{i=1}^n |w_{ni}(x, g_n)| = 1 + O_P(1)$ uniformly in $x \in R_X$ and thus using **(mσ)** and **(Bwn)** one can bound

$$|A_n| \leq \mathbb{E}_X \left[|b(X)| \right] g_n^2 (1 + O_P(1)) = O_P(g_n^2) = o_P\left(\frac{1}{\sqrt{n}}\right). \tag{F.4}$$

Now, one can concentrate on B_n , which stands for the second term on the right-hand side of (F.3). As the local linear weights are given by (3.8), one needs to investigate the following two quantities

$$\begin{aligned} B_{1n} &= \mathbb{E}_X \left\{ b_{1n}(X) \sum_{i=1}^n \frac{1}{n g_n} k\left(\frac{X-X_i}{g_n}\right) \sigma(X_i) \varepsilon_i \right\}, \\ B_{2n} &= \mathbb{E}_X \left\{ b_{2n}(X) \sum_{i=1}^n \frac{1}{n g_n} \frac{X-X_i}{g_n} k\left(\frac{X-X_i}{g_n}\right) \sigma(X_i) \varepsilon_i \right\}, \end{aligned}$$

where

$$b_{1n}(x) = \frac{b(x) S_{n,2}(x)}{S_{n,0}(x) S_{n,2}(x) - S_{n,1}^2(x)} \quad \text{and} \quad b_{2n}(x) = \frac{b(x) S_{n,1}(x)}{S_{n,0}(x) S_{n,2}(x) - S_{n,1}^2(x)},$$

with $S_{n,j}(x)$ defined in (3.9).

Note that with the help of (E.13) and assumptions of the lemma about g one gets

$$b_{1n}(x) = \frac{b(x)}{f_X(x)} + O_P\left(\sqrt{\frac{\log n}{n g_n}}\right) \quad \text{and} \quad b_{2n}(x) = O_P\left(\sqrt{\frac{\log n}{n g_n}}\right). \quad (\text{F.5})$$

Further, analogously as in the proof of Lemma 4 of Gijbels, Omelka and Veraverbeke (2015) one can show

$$\sum_{i=1}^n \frac{(x-X_i)^l}{n g_n^{l+1}} k\left(\frac{x-X_i}{g_n}\right) \sigma(X_i) \varepsilon_i = O_P\left(\sqrt{\frac{\log n}{n g_n}}\right) \quad (\text{F.6})$$

for $l = 0, 1$ uniformly in $x \in R_X$. Now combining (Bwn), (F.5) and (F.6) with $l = 0$ yields that

$$B_{1n} = \mathbb{E}_X \left\{ \frac{b(X)}{f_X(X)} \sum_{i=1}^n \frac{1}{n g_n} k\left(\frac{X-X_i}{g_n}\right) \sigma(X_i) \varepsilon_i \right\} + o_P\left(\frac{1}{\sqrt{n}}\right), \quad (\text{F.7})$$

Analogously combining (Bwn), (F.5) and (F.6) with $l = 1$ yields $B_{2n} = o_P\left(\frac{1}{\sqrt{n}}\right)$. Thus it remains to treat only B_{1n} . Thanks to (F.7) this quantity can be approximated as

$$\begin{aligned} B_{1n} &= \frac{1}{n} \sum_{i=1}^n \mathbb{E}_X \left\{ \frac{b(X)}{f_X(X)} \frac{1}{g_n} k\left(\frac{X-X_i}{g_n}\right) \sigma(X_i) \varepsilon_i \right\} + o_P\left(\frac{1}{\sqrt{n}}\right), \\ &= \frac{1}{n} \sum_{i=1}^n \sigma(X_i) \varepsilon_i \int \frac{b(x)}{f_X(x)} \frac{1}{g_n} k\left(\frac{x-X_i}{g_n}\right) f_X(x) dx + o_P\left(\frac{1}{\sqrt{n}}\right), \\ &= \frac{1}{n} \sum_{i=1}^n \sigma(X_i) \varepsilon_i \int b(X_i + t g_n) k(t) dt + o_P\left(\frac{1}{\sqrt{n}}\right). \end{aligned} \quad (\text{F.8})$$

Let us introduce the set of functions on $R_X \times \mathbb{R}$

$$\mathcal{F} = \left\{ (x, e) \mapsto \int b(x + t g) k(t) dt \sigma(x) e, \quad g > 0 \right\}.$$

Then the first term on the right-hand side of (F.8) can be viewed as an empirical measure P_n of $(X_i, \varepsilon_i)_{i=1}^n$ that is indexed by the functions from \mathcal{F} and evaluated at the function $\tilde{f}_n(x, e) = \int b(x + t g_n) k(t) dt \sigma(x) e$. Thanks to the boundedness of the derivative of b the set of the functions

$$\mathcal{H} = \left\{ x \mapsto \int b(x + t g) k(t) dt, \quad g > 0 \right\}$$

is a subset of $C_1^M(R_X)$ for a sufficiently large M and thus Donsker. Thus Lemma F.1 with $\mathbf{X}_i = (X_i, \varepsilon_i)$ and $h_n(x, e) = \left(\int b(x + t g_n) k(t) dt - b(x)\right)$ and $g(x, e) = \sigma(x) e$ yields that

$$B_{1n} = \frac{1}{n} \sum_{i=1}^n b(X_i) \sigma(X_i) \varepsilon_i + o_P\left(\frac{1}{\sqrt{n}}\right),$$

which finishes the proof of the lemma. \square

Lemma F.3. *Suppose that our observations follow the nonparametric location scale models described at the beginning of Section 3.4.2 and $\hat{\sigma}_n(X)$ be given by (3.7). Further let the assumptions **(Bwn)**, **(F1n)**, **(kn)**, **(Xn)** and **(mσ)** hold and $b : R_X \rightarrow \mathbb{R}$ possesses a bounded derivative on R_X . Then*

$$\mathbb{E}_X \left\{ b(X) \left(\frac{\hat{\sigma}_n(X)}{\sigma(X)} - 1 \right) \right\} = \frac{1}{n} \sum_{i=1}^n \frac{b(X_i)}{2} (\varepsilon_i^2 - 1) + o_P\left(\frac{1}{\sqrt{n}}\right). \tag{F.9}$$

Proof. First, rewrite the left-hand side of (F.9) as

$$\begin{aligned} \mathbb{E}_X \left\{ b(X) \left(\frac{\hat{\sigma}_n(X)}{\sigma(X)} - 1 \right) \right\} &= \mathbb{E}_X \left\{ b(X) \frac{\hat{\sigma}_n^2(X) - \sigma^2(X)}{\sigma(X) [\hat{\sigma}_n(X) + \sigma(X)]} \right\} \\ &= \mathbb{E}_X \left\{ b_n(X) \left[\hat{\sigma}_n^2(X) - \sigma^2(X) \right] \right\}, \end{aligned} \tag{F.10}$$

where $b_n(x) = \frac{b(x)}{\sigma(x) [\hat{\sigma}_n(x) + \sigma(x)]}$. Now, the right-hand side of (F.10) equals

$$\begin{aligned} &\mathbb{E}_X \left\{ b_n(X) \left[\hat{\sigma}_n^2(X) - \sigma^2(X) \right] \right\} \\ &= \mathbb{E}_X \left\{ b_n(X) \sum_{i=1}^n w_{ni}(X, g_n) \left[(Y_i - \hat{m}_n(X_i))^2 - \sigma^2(X_i) + \sigma^2(X_i) - \sigma^2(X) \right] \right\} \\ &= \mathbb{E}_X \left\{ b_n(X) \sum_{i=1}^n w_{ni}(X, g_n) \left[(m(X_i) + \varepsilon_i \sigma(X_i) - \hat{m}_n(X_i))^2 - \sigma^2(X_i) \right] \right\} \\ &\quad + \mathbb{E}_X \left\{ b_n(X) \sum_{i=1}^n w_{ni}(X, g_n) \left[\sigma^2(X_i) - \sigma^2(X) \right] \right\} \\ &= \mathbb{E}_X \left\{ b_n(X) \sum_{i=1}^n w_{ni}(X, g_n) (m(X_i) - \hat{m}_n(X_i))^2 \right\} \end{aligned} \tag{F.11}$$

$$+ \mathbb{E}_X \left\{ b_n(X) \sum_{i=1}^n w_{ni}(X, g_n) \sigma^2(X_i) (\varepsilon_i^2 - 1) \right\} \tag{F.12}$$

$$+ 2 \mathbb{E}_X \left\{ b_n(X) \sum_{i=1}^n w_{ni}(X, g_n) (m(X_i) - \hat{m}_n(X_i)) \sigma(X_i) \varepsilon_i \right\} \tag{F.13}$$

$$+ \mathbb{E}_X \left\{ b_n(X) \sum_{i=1}^n w_{ni}(X, g_n) \left[\sigma^2(X_i) - \sigma^2(X) \right] \right\} \tag{F.14}$$

$$\equiv A_{1n} + A_{2n} + A_{3n} + A_{4n},$$

where A_{1n} , A_{2n} , A_{3n} and A_{4n} are given by (F.11), (F.12), (F.13) and (F.14) respectively.

Let us start with the term A_{1n} . Thanks to the results published in Ojeda (2008) we know that $\sup_{x \in R_X} |\hat{m}_n(x) - m(x)| = o_P(n^{-1/3})$, which yields

$$A_{1n} = o_P\left(\frac{1}{\sqrt{n}}\right). \tag{F.15}$$

Further, A_{2n} given by (F.12) can be rewritten as

$$A_{2n} = \frac{1}{n} \sum_{i=1}^n \tilde{b}_n(X_i) \sigma^2(X_i) (\varepsilon_i^2 - 1), \tag{F.16}$$

where

$$\tilde{b}_n(x) = \int b_{1n}(x + tg_n) k(t) f_X(x + tg_n) dt + \int b_{2n}(x + tg_n) t k(t) f_X(x + tg_n) dt,$$

with

$$b_{1n}(x) = \frac{b_n(x) S_{n,2}(x)}{S_{n,0}(x) S_{n,2}(x) - S_{n,1}^2(x)} \quad \text{and} \quad b_{2n}(x) = \frac{b_n(x) S_{n,1}(x)}{S_{n,0}(x) S_{n,2}(x) - S_{n,1}^2(x)}.$$

Thanks to Ojeda (2008) and our assumptions one gets $\sup_{x \in R_X} |\hat{\sigma}_n^2(x) - \sigma^2(x)| = o_P(n^{-1/3})$, which together with $(\mathbf{m}\sigma)$ yields that $\sup_{x \in R_X} |\hat{\sigma}_n(x) - \sigma(x)| = o_P(n^{-1/3})$, which gives us that uniformly in $x \in R_X$

$$b_{1n}(x) = \frac{b(x)}{2\sigma^2(x) f_X(x)} + o_P(n^{-1/3}), \quad b_{2n}(x) = O_P\left(\sqrt{\frac{\log n}{n g_n}}\right).$$

Analogously as in Lemma F.2 one can show that A_{2n} given by (F.16) can be approximated as

$$A_{2n} = \frac{1}{n} \sum_{i=1}^n \sigma^2(X_i) (\varepsilon_i^2 - 1) \int \frac{b(X_i + t g_n)}{2\sigma^2(X_i + t g_n)} k(t) dt + o_P\left(\frac{1}{\sqrt{n}}\right),$$

and further use Lemma F.1 to deduce that

$$\begin{aligned} A_{2n} &= \frac{1}{n} \sum_{i=1}^n \frac{b(X_i)}{2\sigma^2(X_i)} \sigma^2(X_i) (\varepsilon_i^2 - 1) + o_P\left(\frac{1}{\sqrt{n}}\right) \\ &= \frac{1}{n} \sum_{i=1}^n \frac{b(X_i)}{2} (\varepsilon_i^2 - 1) + o_P\left(\frac{1}{\sqrt{n}}\right). \end{aligned} \tag{F.17}$$

Further, similarly as in the proof of Lemma F.2 one can show that A_{3n} given by (F.13) can be approximated as

$$A_{3n} = \frac{1}{n} \sum_{i=1}^n (m(X_i) - \hat{m}_n(X_i)) \sigma(X_i) \varepsilon_i \int b(X_i + tg_n) k(t) dt + o_P\left(\frac{1}{\sqrt{n}}\right).$$

By the results of Ojeda (2008) one gets $\mathbb{P}(\hat{m}_n \in C_1^M(R_X)) \rightarrow 1$ and thus one can view A_{3n} as $P_n(f_n)$, where P_n is the empirical measure indexed by the set of functions

$$\mathcal{F} = \{(x, e) \mapsto r(x)\sigma(x)e; r \in C_1^M(R_X)\}$$

evaluated at $f_n(x, e) = (m(x) - \hat{m}_n(x))\sigma(x) \varepsilon_i \int b(x + tg_n) k(t) dt$. Further, $\hat{m}_n(x) = m(x) + o_P(1)$ uniformly in $x \in R_X$. Thus one can put $h_n(x, e) = m(x) -$

$\hat{m}_n(x)$ and $g(x, e) = \sigma(x)e$ in Lemma F.1 and get that $P_n(gh_n) = o_P(\frac{1}{\sqrt{n}})$, which further gives

$$A_{3n} = o_P\left(\frac{1}{\sqrt{n}}\right). \tag{F.18}$$

Finally similarly as in (F.4) in the proof of Lemma F.2 one can use (Bwn) and (mσ) to bound the quantity A_{4n} given by (F.14)

$$|A_{4n}| \leq E_X \left\{ |b_n(X)| \sum_{i=1}^n |w_{ni}(X, g_n)|(X_i - X)^2 \right\} \sup_{x \in R_x} \left| \frac{\partial^2 \sigma^2(x)}{\partial x^2} \right| = o_P\left(\frac{1}{\sqrt{n}}\right).$$

Now combining (F.15), (F.17) and (F.18) yields the statement of the lemma. □

Lemma F.4. *Let (F1p) and (F2p) be satisfied for a distribution function F (with the subscripts j and ε dropped). Then the functions*

$$g_u(s, t) = F^{-1}(u) f((1 + s)F^{-1}(u) + t), \quad h_u(s, t) = f((1 + s)F^{-1}(u) + t),$$

are continuous at the point $(0, 0)$ uniformly in $u \in [0, 1]$.

Proof. We prove the lemma only for g_u . The proof for h_u is analogous.

Note that by the assumptions the function $(|y| + 1)f(y)$ is bounded and

$$\lim_{|y| \rightarrow \infty} (|y| + 1)f(y) = 0.$$

Thus, for each $\varepsilon > 0$ there exists M such that $\sup_{|y| \geq M} (|y| + 1)f(y) < \varepsilon$. So there exists $u_0 \in (0, \frac{1}{2})$ such that for each $s, t \in (-\frac{1}{2}, \frac{1}{2})$:

$$\sup_{u \in I_{u_0}} |g_u(s, t)| \leq \sup_{u \in I_{u_0}} |F^{-1}(u) f(F^{-1}(u)(1 + s) + t)| < \varepsilon,$$

where $I_{u_0} = [0, u_0] \cup [1 - u_0, 1]$. Now, one can bound

$$\begin{aligned} & \sup_{u \in [0, 1]} |g_u(s, t) - g_u(0, 0)| \\ & \leq \sup_{u \in I_{u_0}} |g_u(s, t) - g_u(0, 0)| + \sup_{u \in [u_0, 1 - u_0]} |g_u(s, t) - g_u(0, 0)|. \end{aligned} \tag{F.19}$$

Now, the first term on the right-hand side of (F.19) can be bounded by 2ε and the second term can be made arbitrarily small for all s, t sufficiently close to zero. □

Lemma F.5. *For $l_1, l_2 \in \{0, 1\}$ define $U_n^{(l_1, l_2)}(x)$ as*

$$\begin{aligned} U_n^{(l_1, l_2)}(x) &= \frac{1}{n^2} \sum_{j=1}^n \sum_{k=1}^n \frac{(x - X_j)^{l_1}}{h_n^{l_1 + 1}} k\left(\frac{x - X_j}{h_n}\right) \\ & \quad \times \frac{(x - X_k)^{l_2}}{h_n^{l_2 + 1}} k\left(\frac{x - X_k}{h_n}\right) \mathbb{I}\{Y_{1j} < Y_{1k}, Y_{2j} < Y_{2k}\} \end{aligned}$$

and let (\mathbf{Xn}) , (\mathbf{kn}) , (\mathbf{Bwn}) and (\mathbf{H}) hold. Then

$$\sup_{x \in R_X} |U_n^{(l_1, l_2)}(x) - \mathbf{E} U_n^{(l_1, l_2)}(x)| = o_P \left(\sqrt{\frac{h_n^{1+\delta}}{\log n}} \right).$$

Proof. Note that $h_n^2 U_n^{(l_1, l_2)}$ can be viewed as U -process indexed by the set of functions \mathcal{F}_n that is for each $n \in \mathbb{N}$ a subset of

$$\begin{aligned} \mathcal{F} = \left\{ (y'_1, y'_2, x', y''_1, y''_2, x'') \mapsto \frac{(x-x')^{l_1}}{h^{l_1}} k\left(\frac{x-x'}{h}\right) \frac{(x-x'')^{l_2}}{h^{l_2}} k\left(\frac{x-x''}{h}\right) \right. \\ \left. \times \mathbb{I}\{y'_1 < y''_1, y'_2 < y''_2\}; x \in R_X, h > 0 \right\}. \end{aligned} \quad (\text{F.20})$$

Further it is easy to see that \mathcal{F} is a subset of a product of \mathcal{F}_1 and \mathcal{F}_2 , where

$$\begin{aligned} \mathcal{F}_1 = \left\{ (y'_1, y'_2, x', y''_1, y''_2, x'') \mapsto \frac{(x-x')^{l_1}}{h^{l_1}} k\left(\frac{x-x'}{h}\right) \mathbb{I}\{y'_1 < y''_1, y'_2 < y''_2\}; \right. \\ \left. x \in R_X, h > 0 \right\} \end{aligned} \quad (\text{F.21})$$

and

$$\begin{aligned} \mathcal{F}_2 = \left\{ (y'_1, y'_2, x', y''_1, y''_2, x'') \mapsto \frac{(x-x'')^{l_2}}{h^{l_2}} k\left(\frac{x-x''}{h}\right) \mathbb{I}\{y'_1 < y''_1, y'_2 < y''_2\}; \right. \\ \left. x \in R_X, h > 0 \right\}. \end{aligned} \quad (\text{F.22})$$

By Lemma 22 of Nolan and Pollard (1987) and Lemma 2.6.18 (vi) of van der Vaart and Wellner (1996) one can show that both \mathcal{F}_1 and \mathcal{F}_2 are VC-classes with bounded envelopes. Thus by Theorem 2.6.7 of van der Vaart and Wellner (1996) there exist universal constants A_1, B_1 and A_2, B_2 such that for each probability measure Q the covering numbers of \mathcal{F}_1 and \mathcal{F}_2 are bounded by

$$\forall \varepsilon > 0 \quad N(\varepsilon, \mathcal{F}_1, L_2(Q)) \leq A_1 \left(\frac{1}{\varepsilon}\right)^{B_1}, \quad N(\varepsilon, \mathcal{F}_2, L_2(Q)) \leq A_2 \left(\frac{1}{\varepsilon}\right)^{B_2}. \quad (\text{F.23})$$

Further, by the boundedness of \mathcal{F}_1 and \mathcal{F}_2 and (F.23) one can conclude that there exist universal constants A, B such that for each probability measure Q also

$$\forall \varepsilon > 0 \quad N(\varepsilon, \mathcal{F}, L_2(Q)) \leq A \left(\frac{1}{\varepsilon}\right)^B. \quad (\text{F.24})$$

Now by the Hoeffding decomposition one gets

$$U_n^{(l_1, l_2)}(x) - \mathbf{E} U_n^{(l_1, l_2)}(x) = U_{1n}^{(l_1, l_2)}(x) - \mathbf{E} U_{1n}^{(l_1, l_2)}(x) + U_{2n}^{(l_1, l_2)}(x),$$

where $U_{1n}^{(l_1, l_2)}(x)$ stands for the main (projection) term and $U_{2n}^{(l_1, l_2)}(x)$ for the remaining (degenerate) term. Thanks to (F.24) one can use Theorem 5.3.7 of De la Peña and Giné (1999) (with $k = 2$) to deduce that

$$\sup_{x \in R_X} |n h_n^2 U_{2n}^{(l_1, l_2)}(x)| = O_P(1),$$

which further together with **(Bwn)** implies that

$$\sup_{x \in R_X} |U_{2n}^{(l_1, l_2)}(x)| = O_P\left(\frac{1}{n h_n^2}\right) = o_P(h_n). \tag{F.25}$$

Thus it remains to deal with $U_{1n}^{(l_1, l_2)}(x)$. We consider only $l_1 = l_2 = 0$. The other cases can be treated analogously. With the help of **(Bwn)**, **(Xn)**, **(H)** and **(E.13)**

$$U_{1n}^{(0,0)}(x) = \frac{1}{n} \sum_{j=1}^n \frac{1}{h_n} k\left(\frac{x-X_j}{h_n}\right) \tag{F.26}$$

$$\begin{aligned} &\times \mathbb{E}_{X''} \frac{1}{h_n} k\left(\frac{x-X''}{h_n}\right) [2 H_{X''}(Y_{1j}, Y_{2j}) - F_{1X''}(Y_{1j}) - F_{2X''}(Y_{2j}) + 1] \\ &= \frac{1}{n} \sum_{j=1}^n \frac{1}{h_n} k\left(\frac{x-X_j}{h_n}\right) f_X(x) [2 H_x(Y_{1j}, Y_{2j}) - F_{1x}(Y_{1j}) - F_{2x}(Y_{2j}) + 1] \\ &\quad + o_P\left(\frac{1}{\sqrt{n}}\right) \end{aligned}$$

$$= \frac{1}{n} \sum_{j=1}^n \frac{1}{h_n} k\left(\frac{x-X_j}{h_n}\right) f_X(X_j) \tag{F.27}$$

$$\begin{aligned} &\times [2 H_{X_j}(Y_{1j}, Y_{2j}) - F_{1X_j}(Y_{1j}) - F_{2X_j}(Y_{2j}) + 1] + O_P(h_n) \\ &\equiv B_{1n}(x) + O_P(h_n), \end{aligned} \tag{F.28}$$

where for equation **(F.27)** we use a Taylor series expansion with respect to x at the point X_j , and $B_{1n}(x)$ stands for the first term on the right-hand side of **(F.27)**.

Note that $h_n B_{1n}(x)$ can be viewed as an empirical process indexed by a subset of the set of functions

$$\mathcal{G} = \left\{ (x', y'_1, y'_2) \mapsto k\left(\frac{x-x'}{h}\right) f_X(x') [2 H_{x'}(y'_1, y'_2) - F_{1x'}(y'_1) - F_{2x'}(y'_2) + 1], \right. \\ \left. x \in R_X, h > 0 \right\}.$$

With the help of Lemma 22 of Nolan and Pollard **(1987)** and Lemma 2.6.18 (vi) of van der Vaart and Wellner **(1996)** one can show that \mathcal{G} is a bounded VC-class. Thus for instance Theorem 2.5.2 of van der Vaart and Wellner **(1996)** yields that

$$\sup_{x \in R_X} |\sqrt{n} h_n (B_{1n}(x) - \mathbb{E} B_{1n}(x))| = O_P(1),$$

which further with assumption **(Bwn)** gives

$$\sup_{x \in R_X} |B_{1n}(x) - \mathbb{E} B_{1n}(x)| = O_P\left(\frac{1}{\sqrt{n} h_n^2}\right) = o_P\left(\sqrt{\frac{h_n^{1+\delta}}{\log n}}\right)$$

and thus together with **(F.25)** and **(F.28)** verifies the statement of the lemma. □

Acknowledgements

This research was supported by the IAP Research Network P7/06 of the Belgian State (Belgian Science Policy). The first author gratefully acknowledges support from the GOA/12/014 project of the Research Fund KU Leuven. The second author gratefully acknowledges support from the grant GACR 15-04774Y. The third author is an extraordinary professor at the North-West University, Potchefstroom, South Africa. The authors are grateful to an Associate Editor and two referees for their valuable comments that led to a considerable improvement of the paper.

References

- ACAR, E. F., GENEST, C. and NEŠLEHOVÁ, J. (2012). Beyond simplified pair-copula constructions. *J. Multivariate Anal.* **110** 74–90. [MR2927510](#)
- AKRITAS, M. G. and VAN KEILEGOM, I. (2001). Non-parametric Estimation of the Residual Distribution. *Scand. J. Statist.* **28** 549–567. [MR1858417](#)
- BEDFORD, T. and COOKE, R. M. (2002). Vines: a new graphical model for dependent random variables. *Ann. Statist.* **30** 1031–1068. [MR1926167](#)
- BERGSMA, W. P. (2004). Testing conditional independence for continuous random variables. EURANDOM-report.
- BERGSMA, W. P. (2011). Nonparametric testing of conditional independence by means of the partial copula. *Arxiv preprint*, [arXiv:1101.4607](#).
- BURDA, M. and PROKHOROV, A. (2014). Copula based factorization in Bayesian multivariate infinite mixture models. *J. Multivariate Anal.* **127** 200–213. [MR3188887](#)
- CABRERA, J. L. O. (2012). locpol: Kernel local polynomial regression R package version 0.6-0.
- CHEN, S. X. and HUANG, T. M. (2007). Nonparametric estimation of copula functions for dependence modelling. *Canad. J. of Statist.* **35** 265–282. [MR2393609](#)
- COOK, R. D. and JOHNSON, M. E. (1981). A family of distributions for modelling non-elliptically symmetric multivariate data. *J. R. Stat. Soc. Ser. B Stat. Methodol.* **43** 210–218. [MR0626768](#)
- CRAMÉR, H. (1946). *Mathematical methods of statistics*. Princeton University Press. [MR0016588](#)
- DE LA PEÑA, V. H. and GINÉ, E. (1999). *Decoupling: from dependence to independence*. Springer, New York. [MR1666908](#)
- DEHEUVELS, P. (1979). La fonction de dépendance empirique et ses propriétés. *Acad. Roy. Belg. Bull. Cl. Sci.* **65** 274–292. [MR0573609](#)
- FERMANIAN, J.-D., RADULOVIĆ, D. and WEGKAMP, M. (2004). Weak convergence of empirical copula processes. *Bernoulli* **10** 847–860. [MR2093613](#)
- GÄNSSLER, P. and STUTE, W. (1987). *Seminar on Empirical Processes*. Birkhäuser, Basel. [MR0902803](#)
- GIJBELS, I. and MIELNICZUK, J. (1990). Estimating the density of a copula function. *Comm. Statist. Theory Methods* **19** 445–464. [MR1067582](#)

- GIJBELS, I., OMELKA, M. and VERAVERBEKE, N. (2015). Estimation of a copula when a covariate affects only marginal distributions. *Scand. J. Statist.* To appear. DOI: [10.1111/sjos.12154](https://doi.org/10.1111/sjos.12154).
- GIJBELS, I., VERAVERBEKE, N. and OMELKA, M. (2011). Conditional copulas, association measures and their application. *Comput. Stat. Data An.* **55** 1919–1932. [MR2765054](#)
- GOODMAN, L. A. (1959). Partial tests for partial taus. *Biometrika* **46** 425–432. [MR0111097](#)
- GRIPENBERG, G. (1992). Confidence intervals for partial rank correlations. *J. Amer. Statist. Assoc.* **87** 546–551. [MR1173818](#)
- HOBÆK HAFF, I., AAS, K. and FRIGESSI, A. (2010). On the simplified pair-copula construction—Simply useful or too simplistic? *J. Multivariate Anal.* **101** 1296–1310. [MR2595309](#)
- JOE, H. (2006). Generating random correlation matrices based on partial correlations. *J. Multivariate Anal.* **97** 2177–2189. [MR2301633](#)
- KAUERMANN, G. and SCHELLHASE, C. (2014). Flexible pair-copula estimation in D-vines using bivariate penalized splines. *Statist. Comput.* **24** 1081–1100. [MR3253855](#)
- KENDALL, M. G. (1942). Partial Rank Correlation. *Biometrika* **32** 277–283. [MR0006660](#)
- KIM, J. M., JUNG, Y. S., CHOI, T. and SUNGUR, E. A. (2011). Partial correlation with copula modeling. *Comput. Stat. Data An.* **55** 1357–1366. [MR2741420](#)
- KOJADINOVIC, I. and YAN, J. (2010). Modeling multivariate distributions with continuous margins using the copula R package. *J. Statist. Software* **34** 1–20.
- KORN, E. L. (1984). The ranges of limiting values of some partial correlations under conditional independence. *Amer. Statist.* **38** 61–62.
- LARSSON, M., NEŠLEHOVÁ, J. et al. (2011). Extremal behavior of Archimedean copulas. *Adv. in Appl. Probab.* **43** 195–216. [MR2761154](#)
- NELSEN, R. B. (2006). *An Introduction to Copulas*. Springer, New York Second Edition. [MR2197664](#)
- NELSON, P. I. and YANG, S.-S. (1988). Some properties of Kendall's partial rank correlation coefficient. *Statist. Probab. Lett.* **6** 147–150. [MR0919365](#)
- NEUMEYER, N. and VAN KEILEGOM, I. (2010). Estimating the error distribution in nonparametric multiple regression with applications to model testing. *J. Multivariate Anal.* **101** 1067–1078. [MR2595293](#)
- NOLAN, D. and POLLARD, D. (1987). U-processes: Rates of Convergence. *Ann. Statist.* **15** 780–799. [MR0888439](#)
- OJEDA, J. (2008). Hölder continuity properties of the local polynomial estimator. *Pre-publicaciones del Seminario Matemático "García de Galdeano"*. Paper available at <http://www.unizar.es/galdeano/preprints/pre08.html>.
- OMELKA, M., GIJBELS, I. and VERAVERBEKE, N. (2009). Improved kernel estimation of copulas: Weak convergence and goodness-of-fit testing. *Ann. Statist.* **37** 3023–3058. [MR2541454](#)
- OMELKA, M., VERAVERBEKE, N. and GIJBELS, I. (2013). Bootstrapping the conditional copula. *J. Statist. Plann. Inference* **143** 1–23. [MR2969007](#)

- PATTON, J. A. (2006). Modeling asymmetric exchange rate dependence. *Internat. Econom. Rev.* **47** 527–556. [MR2216591](#)
- SCHMID, F. and SCHMIDT, R. (2007). Multivariate conditional versions of Spearman's rho and related measures of tail dependence. *J. Multivariate Anal.* **98** 1123–1140. [MR2326243](#)
- SEGERS, J. (2012). Weak convergence of empirical copula processes under non-restrictive smoothness assumptions. *Bernoulli* **18** 764–782. [MR2948900](#)
- SERFLING, R. J. (1980). *Approximation Theorems of Mathematical Statistics*. Wiley, New York. [MR0595165](#)
- SONG, K. (2009). Testing conditional independence via Rosenblatt transforms. *Ann. Statist.* **37** 4011–4045. [MR2572451](#)
- TSUKUHARA, H. (2005). Semiparametric estimation in copula models. *Canad. J. of Statist.* **33** 357–375. [MR2193980](#)
- VAN DER VAART, A. W. (2000). *Asymptotic Statistics*. Cambridge University Press, New York.
- VAN DER VAART, A. W. and WELLNER, J. A. (1996). *Weak Convergence and Empirical Processes*. Springer, New York. [MR1385671](#)
- VERAVERBEKE, N., GIJBELS, I. and OMELKA, M. (2014). Pre-adjusted non-parametric estimation of a conditional distribution function. *J. R. Stat. Soc. Ser. B Stat. Methodol.* **76** 399–438. [MR3164872](#)
- VERAVERBEKE, N., OMELKA, M. and GIJBELS, I. (2011). Estimation of a conditional copula and association measures. *Scand. J. Statist.* **38** 766–780. [MR2859749](#)