To appear in the *Journal of Nonparametric Statistics*
Vol. 00, No. 00, Month 20XX, 1–13

On the asymptotic behavior of two estimators of the conditional copula based on time series

T. Bouezmarni [a], F.C. Lemyre [a*], and J-F. Quessy [b]

[a]*Département de mathématiques, Université de Sherbrooke, Canada;* [b]*Département de mathématiques et d'informatique, Université du Québec à Trois-Rivières, Trois-Rivières, Canada*

(*Received 00 Month 20XX; accepted 00 Month 20XX*)

As defined by Patton (2006), the conditional copula of a random pair (Y_1, Y_2) given the value taken by some covariate $X \in \mathbb{R}$ is the function $C_x : [0,1]^2 \to [0,1]$ such that $\mathbb{P}(Y_1 \leq y_1, Y_2 \leq y_2) = C_x\{\mathbb{P}(Y_1 \leq y_1 | X = x), \mathbb{P}(Y_2 \leq y_2 | X = x)\}$. In this note, the weak convergence of two estimators of C_x proposed by Gijbels, Veraverbeke, and Omelka (2011) is established under strong-mixing. It is shown that under appropriate conditions on the weight functions and on the mixing coefficients, the limiting processes are the same as those obtained by Veraverbeke, Omelka, and Gijbels (2011) under the i.i.d. setting. The performance of these estimators in finite sample sizes is investigated.

Keywords: α-mixing processes; conditional copula; kernel estimation; weak convergence.

1. Introduction

Copulas have become a popular tool for modeling the dependence between the components of a random vector. The starting point of copula theory is Sklar's Theorem. In its classical formulation, this result ensures that for any random vector $\mathbf{Y} = (Y_1, \ldots, Y_d)$, there exists a function $C : [0,1]^d \to [0,1]$ called the copula of $\mathbf{Y}$ such that for all $\mathbf{y} = (y_1, \ldots, y_d) \in \mathbb{R}^d$,

$$\mathbb{P}(\mathbf{Y} \leq \mathbf{y}) = C\left\{\mathbb{P}(Y_1 \leq y_1), \ldots, \mathbb{P}(Y_d \leq y_d)\right\}.$$

When $Y_1, \ldots, Y_d$ are continuous, C is unique.

Recently, some works concentrated on capturing the influence of a covariate $X \in \mathbb{R}$ on the dependence structure of a random pair. A motivating example is given in Gijbels et al. (2011), where the relationship between the life expectancy of men (Y_1) and women (Y_2) with respect to the gross domestic product (X) is studied. Such an investigation relies

*Corresponding author. Email: Felix.Camirand.Lemyre@USherbrooke.ca

on an extension of Sklar's Theorem to the case of conditional dependence as initiated by Patton (2006). Formally, letting $H_x(y_1, y_2) = \mathbb{P}(Y_1 \leq y_1, Y_2 \leq y_2 | X = x)$, the dependence between Y_1 and Y_2 conditional on $X = x$ is characterized by the conditional copula $C_x : [0,1]^2 \to [0,1]$ such that for all $(y_1, y_2) \in \mathbb{R}^2$,

$$H_x(y_1, y_2) = C_x \left\{ \mathbb{P}(Y_1 \leq y_1 | X = x), \mathbb{P}(Y_2 \leq y_2 | X = x) \right\}. \tag{1}$$

Two nonparametric estimators of C_x were proposed by Gijbels et al. (2011) and their asymptotic behavior was formally investigated by Veraverbeke et al. (2011) in the i.i.d. case. The purpose of this note is to extend these large-sample results to the case of time series, since many contexts of applications involve serially dependent observations.

To this end, one adopts a very general framework where the stationary process $\{(Y_{1t}, Y_{2t}, X_t)\}_{t \in \mathbb{Z}}$ satisfies a strong mixing condition. In a certain sense, these results are versions of Bücher and Volgushev (2013) adapted to the context of conditional copulas. Specifically, let $\mathcal{F}_a^b$ be the σ-field generated by $\{(Y_{1t}, Y_{2t}, X_t)\}_{a \leq t \leq b}$ and define the α-mixing coefficients

$$\alpha(r) = \sup_{k \in \mathbb{Z}} \alpha \left(\mathcal{F}_{-\infty}^k, \mathcal{F}_{k+r}^\infty \right),$$

where

$$\alpha(\mathcal{A}, \mathcal{B}) = \sup_{(A,B) \in \mathcal{A} \times \mathcal{B}} \left| \mathbb{P}(A \cap B) - \mathbb{P}(A)\, \mathbb{P}(B) \right|.$$

The process $\{(Y_{1t}, Y_{2t}, X_t)\}_{t \in \mathbb{Z}}$ is said to be α-mixing, or strong mixing, if $\alpha(r) \to 0$ as $r \to \infty$. Several parametric time series models satisfy this strong mixing assumption, including ARMA and GARCH models under appropriate restrictions on the parameters involved. For more details, see Meitz and Saikkonen (2008), Doukhan (1994) and Carrasco and Chen (2002).

The remaining of the paper is organized as follows. Section 2 establishes the asymptotic behavior of a first estimator of the conditional copula and provides a sketch of the proof. Section 3 mimics Section 2 for a second estimator which aims at reducing the bias. Section 4 presents the results of a numerical study that investigates the performance of the two estimators when computed from serially dependent data. The assumptions needed for the theoretical results to hold are listed in an appendix and the detailed proofs of the main results are to be found in the Supplementary materials.

2. Investigation of a first estimator of C_x

2.1. *Description of the estimator*

Consider n realizations $(Y_{11}, Y_{21}, X_1), \ldots, (Y_{1n}, Y_{2n}, X_n)$ of a stationary process $\{(Y_{1t}, Y_{2t}, X_t)\}_{t\in\mathbb{Z}}$ that satisfies the strong mixing assumption. In that context, a first estimator of C_x arises naturally upon noting that

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

where $F_{1x}(y) = \mathbb{P}(Y_1 \leq y|X = x)$ and $F_{2x}(y) = \mathbb{P}(Y_2 \leq y|X = x)$. An estimator of H_x will then provide a plug-in estimation of C_x. Specifically, let

$$H_{xh}(y_1, y_2) = \sum_{i=1}^{n} w_{ni}(x, h)\, \mathbb{I}\left(Y_{1i} \leq y_1, Y_{2i} \leq y_2\right), \tag{3}$$

where $w_{n1}, \ldots, w_{nn}$ are nonnegative weight functions that smooth the covariate space and $h = h_n$ is a bandwidth parameter that typically depends on the sample size. Hereafter, it is assumed that the $w_{ni}(x, h)$'s sum to 1. From Equation (2), an estimator of C_x is given by

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

where $F_{1xh}(y) = \lim_{w\to\infty} H_{xh}(y, w)$ and $F_{2xh}(y) = \lim_{w\to\infty} H_{xh}(w, y)$ are the conditional empirical marginal distributions. Here and in the sequel, the inverse of a function is understood as its left-continuous generalized inverse.

2.2. *Weak convergence*

The aim of this subsection is to describe the large-sample behavior of the empirical process $\mathbb{C}_{xh} = \sqrt{nh}(C_{xh} - C_x)$ as a random element in the space $\ell^\infty([0, 1]^2)$ of bounded functions defined on $[0, 1]^2$. The first step toward this goal is to investigate the asymptotic behavior of $\mathbb{H}_{xh} = \sqrt{nh}(H_{xh} - H_x)$, where H_{xh} and H_x are defined respectively in Equation (3) and Equation (1).

Proposition 2.1 *Suppose that Assumptions $\mathcal{A}_1$–$\mathcal{A}_2$, W_1–W_5 and W_{11}–W_{13} are satisfied. If $nh \to \infty$ and $nh^5 \to K^2 < \infty$, then the empirical process $\mathbb{H}_{xh}$ converges weakly in the space $\ell^\infty(\mathbb{R}^2)$ to a Gaussian limit $\mathbb{H}_x$ such that*

$$\mathrm{E}\left\{\mathbb{H}_x(y_1, y_2)\right\} = K\left\{K_2\, \dot{H}_x(y_1, y_2) + K_3\, \ddot{H}_x(y_1, y_2)\right\}$$

and for $a \wedge b = \min(a, b)$,

$$\mathrm{Cov}\left\{\mathbb{H}_x(y_1, y_2), \mathbb{H}_x(y_1', y_2')\right\} = K_4 \left\{H_x(y_1 \wedge y_1', y_2 \wedge y_2') - H_x(y_1, y_2)\, H_x(y_1', y_2')\right\},$$

where the constants K_2–K_4 *are defined in Assumptions* W_2–W_4.

It is worth noting that the asymptotic covariance structure of $\mathbb{H}_{xh}$ under the strong mixing assumption is the same as that obtained by Veraverbeke et al. (2011) in the i.i.d. setting. In other words, the influence of time-dependency is asymptotically negligible here. An explanation is that the weight functions smooth the covariate space in a shrinking neighborhood of x as n goes to infinity. Note however that compared to the i.i.d. context, the additional assumptions W_{11}–W_{13} on the weight functions are needed in order to tackle moments of order six entailed by time-dependency.

Now the main result of this section can be established.

Proposition 2.2 *Suppose that the conditions in Proposition 2.1 are satisfied. Then, if Assumption* $\mathcal{A}_3$ *holds, the empirical process* $\mathbb{C}_{xh}$ *converges weakly in the space* $\ell^\infty([0,1]^2)$ *to a Gaussian limit* $\mathbb{C}_x$ *having representation*

$$\mathbb{C}_x(u_1, u_2) = \mathbb{B}_x(u_1, u_2) - C_x^{[1]}(u_1, u_2)\, \mathbb{B}_x(u_1, 1) - C_x^{[2]}(u_1, u_2)\, \mathbb{B}_x(1, u_2),$$

where $\mathbb{B}_x$ *is a Gaussian process on* $[0,1]^2$ *such that*

$$\mathrm{E}\left\{\mathbb{B}_x(u_1, u_2)\right\} = K\left[K_2\, \dot{C}_x(u_1, u_2) + K_3\, \ddot{C}_x\left\{F_{1x}^{-1}(u_1), F_{2x}^{-1}(u_2)\right\}\right]$$

and

$$\mathrm{Cov}\left\{\mathbb{B}_x(u_1, u_2), \mathbb{B}_x(u_1', u_2')\right\} = K_4 \left\{C_x(u_1 \wedge u_1', u_2 \wedge u_2') - C_x(u_1, u_2)\, C_x(u_1', u_2')\right\}.$$

The limiting representation of $\mathbb{C}_x$ in terms of $\mathbb{B}_x$ stated in Proposition 2.2 allows to compute the asymptotic bias function $\mathrm{E}\{\mathbb{C}_x(u_1, u_2)\}$, as well as the covariance function $\mathrm{Cov}\{\mathbb{C}_x(u_1, u_2), \mathbb{C}_x(u_1', u_2')\}$. These expressions are identical to those derived by Veraverbeke et al. (2011) in the i.i.d case.

Now sketches of the proofs of Proposition 2.1 and of Proposition 2.2 are provided in the next two subsections. The full arguments can be found in the Supplementary material section.

2.3. *Sketch of the proof of Proposition 2.1*

In the sequel, expectations of the form $\mathrm{E}\{f(Y_{1i}, Y_{2i}, X_i)\}$ are taken conditionally on X_i,

i.e.

$$\mathrm{E}\left\{f(Y_{1i}, Y_{2i}, X_i)\right\} = \int_{\mathbb{R}^2} f\left(y_1, y_2, X_i\right) \mathrm{d}H_{X_i}(y_1, y_2).$$

Since Assumptions W_2, W_3 and W_5 hold, one deduces from Veraverbeke et al. (2011) that $\sqrt{nh}\{\mathrm{E}(H_{xh}) - H_x\} = \mathrm{E}(\mathbb{H}_x) + o(1)$. Therefore, one only needs to show that the process

$$Z_{xn}(y_1, y_2) = \sqrt{nh}\left\{H_{xh}(y_1, y_2) - \mathrm{E}(H_{xh}(y_1, y_2))\right\}$$

is asymptotically gaussian and that its limiting covariance function matches that of $\mathbb{H}_x$. According for instance to Theorem 1.5.4 of van der Vaart and Wellner (1996), one needs to show that the finite-dimensional distributions of Z_{xn} are asymptotically jointly Normal and that Z_{xn} is asymptotically tight.

That the finite-dimensional distributions of Z_{xn} are asymptotically jointly Normal is established in the Supplementary material section, where the arguments in the proof of Theorem 27.4 of Billingsley (1968), which apply to serially dependent data, are adapted to the conditional setup.

In order to show that the covariance function of Z_{xn} converges to that of $\mathbb{H}_x$, one follows an idea similar to the one developed by Li and Racine (2007). Specifically, for $\mathbf{y} = (y_1, y_2)$ and $\mathbf{y}' = (y_1', y_2')$, one has

$$\mathrm{Cov}\left\{Z_{xn}(\mathbf{y}), Z_{xn}(\mathbf{y}')\right\} = \Lambda_{n1}(\mathbf{y}, \mathbf{y}') + \Lambda_{n2}(\mathbf{y}, \mathbf{y}') + \Lambda_{n3}(\mathbf{y}, \mathbf{y}'),$$

where for $\vartheta_i(\mathbf{y}) = \mathbb{I}(Y_{1i} \leq y_1, Y_{2i} \leq y_2) - H_{X_i}(y_1, y_2)$, $i \in \{1, \ldots, n\}$ and $L_{i,\ell}(\mathbf{y}, \mathbf{y}') = \mathrm{Cov}\{\vartheta_i(\mathbf{y}), \vartheta_{i+\ell}(\mathbf{y}')\}$,

$$\begin{aligned}
\Lambda_{n1}(\mathbf{y}, \mathbf{y}') &= nh \sum_{i=1}^{n} \left\{w_{ni}(x, h)\right\}^2 L_{i,0}(\mathbf{y}, \mathbf{y}'), \\
\Lambda_{n2}(\mathbf{y}, \mathbf{y}') &= \sum_{i=1}^{n} \sum_{\ell=1}^{\lfloor h^{-1/2} \rfloor} w_{ni}(x, h)\, w_{n,i+\ell}(x, h) \left\{L_{i,i+\ell}(\mathbf{y}, \mathbf{y}') + L_{i,i+\ell}(\mathbf{y}', \mathbf{y})\right\}, \\
\Lambda_{n3}(\mathbf{y}, \mathbf{y}') &= \sum_{i=1}^{n} \sum_{\ell=\lfloor h^{-1/2} \rfloor + 1}^{n} w_{ni}(x, h)\, w_{n,i+\ell}(x, h) \left\{L_{i,i+\ell}(\mathbf{y}, \mathbf{y}') + L_{i,i+\ell}(\mathbf{y}', \mathbf{y})\right\}.
\end{aligned}$$

Since $L_{i,0}(\mathbf{y}, \mathbf{y}') = H_{X_i}(y_1 \wedge y_1', y_2 \wedge y_2') - H_{X_i}(y_1, y_2)\, H_{X_i}(y_1, y_2')$, Assumption $\mathcal{A}_2$ ensures that $L_{i,0}(\mathbf{y}, \mathbf{y}') = H_x(y_1 \wedge y_1', y_2 \wedge y_2') - H_x(y_1, y_2)\, H_x(y_1, y_2') + o_{\mathbb{P}}(1)$ for any $i \in I_{nx}$. From Assumption W_4,

$$\lim_{n \to \infty} \Lambda_{n1}(\mathbf{y}, \mathbf{y}') = K_4 \left\{H_x(y_1 \wedge y_1', y_2 \wedge y_2') - H_x(y_1, y_2)\, H_x(y_1, y_2')\right\}.$$

Next,using the fact that $|L_{i,\ell}| \leq 1$ and from Assumption W_{11} in the special case when

$k = 1$ and $v_n = \lfloor h^{-1/2} \rfloor$, one obtains

$$\Lambda_{n2}(\mathbf{y}, \mathbf{y}') \leq 2nh\lfloor h^{-1/2} \rfloor \left\{ \max_{1 \leq \ell \leq \lfloor h^{-1/2} \rfloor} \sum_{i=1}^{n} w_{ni}(x,h)\, w_{n,i+\ell}(x,h) \right\} = O_{\mathbb{P}}(h^{-1/2}).$$

Finally, the strong mixing assumption entails that for all $\mathbf{y}, \mathbf{y}' \in \mathbb{R}^2$,

$$|\mathrm{Cov}\{\vartheta_i(\mathbf{y}), \vartheta_{i+\ell}(\mathbf{y}')\}| \leq \alpha(\ell).$$

From the fact that the weight functions sum to 1 and in view of Assumption W_1,

$$\Lambda_{n3}(\mathbf{y}, \mathbf{y}') \leq nh \left\{ \sum_{i=1}^{n} w_{ni}(x,h) \right\} \left\{ \max_{1 \leq i \leq n} w_{ni}(x,h) \right\} \sum_{\ell = \lfloor h^{-1/2} \rfloor + 1}^{n} \alpha(\ell) = o_{\mathbb{P}}(1),$$

where the last equality is a consequence of Assumption A_1 that ensures that $\alpha(\ell) = O(\ell^{-a})$ for some $a > 6$. One can finally conclude that

$$\lim_{n \to \infty} \mathrm{Cov}\left\{ Z_{xn}(\mathbf{y}), Z_{xn}(\mathbf{y}') \right\} = \mathrm{Cov}\left\{ \mathbb{H}_x(\mathbf{y}), \mathbb{H}_x(\mathbf{y}') \right\}.$$

Now in order to show the asymptotic tightness of Z_{xh}, consider for a fixed $x \in \mathbb{R}$ the semi-metric $\rho(\mathbf{y}, \mathbf{y}') = |F_{1x}(y_1) - F_{1x}(y_1')| + |F_{2x}(y_2) - F_{2x}(y_2')|$ and define for $\delta > 0$, $f : \mathbb{R}^2 \to \mathbb{R}$ bounded and $T \subseteq \mathbb{R}^2$,

$$\mathfrak{W}_\delta(f, T) = \sup_{\mathbf{y}, \mathbf{y}' \in T; \rho(\mathbf{y}, \mathbf{y}') < \delta} |f(\mathbf{y}) - f(\mathbf{y}')|.$$

The modulus of ρ-continuity of Z_{xn} is then given by $\mathfrak{W}_\cdot(Z_{xn}, \mathbb{R}^2)$. For a fixed $\mathbf{y} \in \mathbb{R}^2$, the random variable $Z_{xn}(\mathbf{y})$ is asymptotically tight in $\mathbb{R}$, so according to Theorem 1.5.7 of van der Vaart and Wellner (1996), the process Z_{xn} is asymptotically tight in $\ell^\infty([0,1]^2)$ if and only if for any $\delta > 0$, $\mathfrak{W}_\delta\left(Z_{xn}, \mathbb{R}^2\right)$ converges to zero in probability. To show that it is indeed the case, one proceeds similarly as in Theorem 3 of Bickel and Wichura (1971). Specifically, let $\kappa_\gamma = \lfloor (nh)^{1/2+\gamma} \rfloor$ for some $\gamma \in (0, 1/2)$ and define the rectangles

$$A_\gamma(i,j) = \left[F_{1x}^{-1}\left(\frac{i-1}{\kappa_\gamma}\right), F_{1x}^{-1}\left(\frac{i}{\kappa_\gamma}\right) \right] \times \left[F_{2x}^{-1}\left(\frac{j-1}{\kappa_\gamma}\right), F_{2x}^{-1}\left(\frac{j}{\kappa_\gamma}\right) \right].$$

The collection $A_\gamma(\cdot,\cdot)$ is a partition of $\mathbb{R}^2$ and the ρ-measure of each element is bounded by $2/\kappa_\gamma$. Now for an arbitrary nonempty rectangle $A \in \mathbb{R}^2$, let

$$\mathbb{H}_{xh}(A) = \sqrt{nh} \sum_{i=1}^{n} w_{ni}(x,h) \left[\mathbb{I}\left\{ (Y_{1i}, Y_{2i}) \in A \right\} - \nu_{X_i}(A) \right],$$

where $\nu_x(A) = \mathbb{P}\{(Y_{1i}, Y_{2i}) \in A | X_i = x\}$. The definition of the random function $\mathbb{H}_{xh}(A)$ is motivated by the following Lemma whose proof is to be found in the Supplementary

material section.

LEMMA 2.3 *Suppose that $\sqrt{nh}h^2 < \infty$ and Assumptions $\mathcal{A}_2$ and W_2–W_3 are satisfied. Then, for n sufficiently large, one has for any $\epsilon > 0$ that*

$$\mathbb{P}\left\{\mathfrak{W}_\delta\left(Z_{xn}, \mathbb{R}^2\right) \geq \epsilon\right\} \leq \mathbb{P}\left[\max_{1\leq i,j\leq \kappa_\gamma} |\mathbb{H}_{xh}\left\{A_\gamma(i,j)\right\}| \geq \epsilon\right].$$

Now let $\mu_x = \nu_x \otimes \lambda$, where λ is the ρ-measure of A. One then needs to find $\beta > 1$ and $C \in \mathbb{R}$ (that may depend on ϵ and β) such that

$$\mathbb{P}\left[|\mathbb{H}_{xh}\{A_\gamma(i,j)\}| \geq \epsilon\right] \leq C\left[\mu_x\left\{A_\gamma(i,j)\right\}\right]^\beta. \tag{4}$$

Such an inequality may be derived from the next lemma whose technical proof is to be found in the Supplementary material section.

LEMMA 2.4 *If Assumptions $\mathcal{A}_1$–$\mathcal{A}_2$, W_1 and W_{11}–W_{13} are satisfied, one can find a finite constant $\omega > 0$ such that for any rectangle $A \subseteq \mathbb{R}^2$,*

$$\mathrm{E}\left\{|\mathbb{H}_{xh}(A)|^6\right\} \leq \omega \sum_{k=1}^{3}\left\{\mu_x(A) + h^2\right\}^k (nh)^{-3+k}. \tag{5}$$

In view of equations (4) and (5), the Markov inequality entails

$$\mathbb{P}\left(|\mathbb{H}_{xh}\{A_\gamma(i,j)\}| \geq \epsilon\right) \leq \epsilon^{-6}\,\mathrm{E}\left\{\mathbb{H}_{xh}(A_\gamma(i,j))^6\right\} \leq \epsilon^{-6}\,K\,\mu_{x\gamma}^\beta a_{xh}(\gamma,\beta),$$

where $\mu_{x\gamma} = \mu_x(A_\gamma(i,j))$ and

$$\begin{aligned} a_{xh}(\gamma,\beta) &= \mu_{x\gamma}^{1-\beta}\left\{(nh)^{-2} + n^{-1} + h^4\right\} + \mu_{x\gamma}^{2-\beta}\left\{(nh)^{-1} + h^2\right\} \\ &\quad + \mu_{x\gamma}^{3-\beta} + \mu_{x\gamma}^{-\beta}\left\{n^{-2} + n^{-1}h^3 + h^6\right\}. \end{aligned} \tag{6}$$

From the definition of $\mu_{x\gamma}$, one has $(nh)^{-1-2\gamma} \leq \mu_{x\gamma} \leq (nh)^{-1/2-\gamma}$. Hence, a meticulous inspection of Equation (6) shows that for any small γ and any β close to 1, one has $a_{xh}(\gamma,\beta) < 1$ since h satisfies $nh^5 < \infty$. Hence, one can find a constant C (that depends on ϵ, β, γ and τ) such that Equation (4) is satisfied. The asymptotic ρ-equicontinuity follows, for instance, from an extension of Theorem 3 in Bickel and Wichura (1971).

2.4. *Proof of Proposition 2.2*

Let $\mathbb{D}$ be the space of bivariate distribution functions and define the mapping $\Lambda(H_x)(u_1,u_2) = H_x\{F_{1x}^{-1}(u_1), F_{2x}^{-1}(u_2)\}$. One can then write $\mathbb{C}_{xh} = \sqrt{nh}\{\Lambda(H_{xh}) -$

$\Lambda(H_x)\}$. From Bücher and Volgushev (2013), one can conclude in view of Assumption $\mathcal{A}_3$ that Λ is Hadamard differentiable with derivative at H_x given for $\widetilde{\Delta}(u_1,u_2) = \Delta\{F_{1x}^{-1}(u_1), F_{2x}^{-1}(u_2)\}$ by

$$\Lambda'_{H_x}(\Delta)(u_1,u_2) = \widetilde{\Delta}(u_1,u_2) - C_x^{[1]}(u_1,u_2)\,\widetilde{\Delta}(u_1,1) - C_x^{[2]}(u_1,u_2)\,\widetilde{\Delta}(1,u_2).$$

From the functional delta method, $\mathbb{C}_{xh}$ converges weakly to

$$\mathbb{C}_x = \Lambda'_{H_x}(\mathbb{H}_x) = \mathbb{B}_x(u_1,u_2) - C_x^{[1]}(u_1,u_2)\,\mathbb{B}_x(u_1,1) - C_x^{[2]}(u_1,u_2)\,\mathbb{B}_x(1,u_2),$$

where $\mathbb{B}_x(u_1,u_2) = \mathbb{H}_x\left\{F_{1x}^{-1}(u_1), F_{2x}^{-1}(u_2)\right\}$, which completes the proof.

3. Investigation of a second estimator of C_x

3.1. *Description of the estimator*

As noted by Gijbels et al. (2011), the estimator C_{xh} may be severely biased, especially when the marginal distributions strongly depend on the covariate. For that reason, they proposed a second estimator in order to reduce this effect of the covariate on the margins and hopefully obtain a smaller bias. To this end, define for each $i \in \{1,\ldots,n\}$ the *pseudo-uniformized* observations $(\widetilde{U}_{1i}, \widetilde{U}_{2i}) = (F_{1X_ih_1}(Y_{1i}), F_{2X_ih_2}(Y_{2i}))$, where h_1, h_2 are bandwidth parameters that may differ from h. Then, let

$$G_{xh}(v_1,v_2) = \sum_{i=1}^{n} w_{ni}(x,h)\,\mathbb{I}\left\{F_{1X_ih_1}(Y_{1i}) \le v_1, F_{2X_ih_2}(Y_{2i}) \le v_2\right\}$$

and note G_{1xh}, G_{2xh} the marginals of G_{xh}. An estimator of C_x is then

$$\widetilde{C}_{xh}(u_1,u_2) = G_{xh}\left\{G_{1xh}^{-1}(u_1), G_{2xh}^{-1}(u_2)\right\}.$$

3.2. *Weak convergence*

Let $\mathbb{G}_x$ be the Gaussian process of Proposition 2.1 when $H_x = C_x$. The weak convergence of $\widetilde{\mathbb{C}}_{xh} = \sqrt{nh}(\widetilde{C}_{xh} - C_x)$ is established next.

Proposition 3.1 *Suppose Assumptions $\mathcal{A}_1$, $\mathcal{A}_2^\star$, $\mathcal{A}_3$–$\mathcal{A}_4$ and W_1–W_{13} are satisfied. If $n\min(h_1,h_2) \to \infty$, $n\max(h_1^5,h_2^5) < \infty$ and $h/\min(h_1,h_2) < \infty$, then $\widetilde{\mathbb{C}}_{xh}$ converges weakly to a Gaussian limit $\widetilde{\mathbb{C}}_x$ having representation*

$$\widetilde{\mathbb{C}}_x(u_1,u_2) = \mathbb{G}_x(u_1,u_2) - C_x^{[1]}(u_1,u_2)\,\mathbb{G}_x(u_1,1) - C_x^{[2]}(u_1,u_2)\,\mathbb{G}_x(1,u_2).$$

Like $\mathbb{C}_{xh}$, the limit of $\widetilde{\mathbb{C}}_{xh}$ under strong mixing is the same as that obtained by Veraverbeke et al. (2011) in the i.i.d. case. In particular, the asymptotic bias and covariance function are the same as those find by these authors.

3.3. *Sketch of the proof of Proposition 3.1*

Consider a version of G_{xh} based on $(U_1, V_1, X_1), \ldots, (U_n, V_n, X_n)$, where $U_i = F_{1X_i}(Y_{1i})$ and $V_i = F_{2X_i}(Y_{2i})$, namely

$$\widetilde{G}_{xh}(u_1, u_2) = \sum_{i=1}^{n} w_{ni}(x, h)\, \mathbb{I}\left(U_i \leq u, V_i \leq v\right).$$

One can then write for Λ defined in the proof of Proposition 2.2 that

$$\widetilde{\mathbb{C}}_{xh} = \sqrt{nh}\left\{\Lambda(\widetilde{G}_{xh}) - C_x\right\} + \sqrt{nh}\left\{\Lambda(G_{xh}) - \Lambda(\widetilde{G}_{xh})\right\}.$$

The first summand is a special case of Proposition 2.1 with (Y_{1i}, Y_{2i}, X_i) replaced by (U_i, V_i, X_i). Since the conditional marginal distributions of (U_i, V_i) are uniform on $(0, 1)$, their joint conditional distribution is C_{X_i}. Since Assumptions $\mathcal{A}_1$, $\mathcal{A}_2^\star$, W_1–W_5 and W_{11}–W_{13} are satisfied, Proposition 2.1 ensures that $\sqrt{nh}\{\Lambda(\widetilde{G}_{xh}) - C_x\}$ converges weakly to $\Lambda'_{C_x}(\mathbb{G}_x) = \widetilde{\mathbb{C}}_x$.

It remains to show that $\sqrt{nh}\{\Lambda(G_{xh}) - \Lambda(\widetilde{G}_{xh})\}$ is asymptotically negligible. As pointed out by Veraverbeke et al. (2011), this is closely related to the asymptotic behavior of the processes

$$\widetilde{Z}_{1xn} = Z_{1xn} - \mathrm{E}(Z_{1xn}) \quad \text{and} \quad \widetilde{Z}_{2xn} = Z_{2xn} - \mathrm{E}(Z_{2xn}),$$

where for $j = 1, 2$ and $z_t = x + t\,C\,h$,

$$Z_{jxn}(t, u) = \sqrt{nh_j}\, F_{jz_t h_j}\left\{F_{jz_t}^{-1}(u_1)\right\}.$$

The key is the following lemma whose proof is deferred to the Supplementary material section.

Lemma 3.2 *Suppose that Assumptions $\mathcal{A}_1$, $\mathcal{A}_4$, W_1, W_6–W_{13} are satisfied. Then, as long as $nh_1^5 < \infty$ and $nh_2^5 < \infty$, the sequences $\widetilde{Z}_{1xn}$ and $\widetilde{Z}_{2xn}$ are asymptotically tight in $\ell^\infty([-1, 1] \times [0, 1])$.*

Finally, from similar arguments as those in Appendix B.2 of Veraverbeke et al. (2011), one obtains that $\sqrt{nh}\{\Lambda(G_{xh}) - \Lambda(\widetilde{G}_{xh})\} = o_{\mathbb{P}}(1)$. Hence,

$$\widetilde{\mathbb{C}}_{xh} = \sqrt{nh}\left\{\Lambda(\widetilde{G}_{xh}) - C_x\right\} + o_{\mathbb{P}}(1),$$

which completes the proof.

4. Sample behavior of the two conditional copula estimators

In order to evaluate the finite sample performance of the estimators C_{xh} and $\widetilde{C}_{xh}$, let $\mathbf{W}_t = (Y_{1t}, Y_{2t}, X_t)$ and consider for some $\theta \in (-1, 1)$ the autoregressive model $\mathbf{W}_t = \theta\, \mathbf{W}_{t-1} + (1-\theta^2)^{1/2} \boldsymbol{\varepsilon}_t$, where $(\boldsymbol{\varepsilon}_t)_{t\in\mathbb{Z}}$ is a i.i.d. process of innovations from the three-dimensional standard normal distribution with correlation matrix $R = (\rho_{ij})_{i,j=1}^3$. This model entails that $\mathbf{W}_t$ follows a standard Normal with correlation R. Then, the conditional distribution of (Y_{1t}, Y_{2t}) given $X_t = x$ is bivariate Normal with correlation coefficient

$$\rho_x = \frac{\rho_{12} - \rho_{13}\,\rho_{23}}{\sqrt{(1-\rho_{13}^2)(1-\rho_{23}^2)}}\,.$$

The conditional copula C_x is that case is therefore the Normal copula with parameter ρ_x; see Nelsen (2006) for more details on this model.

The performance of C_{xh} and $\widetilde{C}_{xh}$ under the above model has been evaluated in the light of the *average integrated squared bias* (AISB) and the *average integrated variance* (AIV) defined by

$$\begin{aligned}
\mathrm{AISB}(\widehat{C}) &= \int_{[0,1]^2} \left[\mathrm{E}\left\{\widehat{C}(u_1,u_2)\right\} - C_x(u_1,u_2)\right]^2 \mathrm{d}u_1\mathrm{d}u_2,\\
\mathrm{AIV}(\widehat{C}) &= \int_{[0,1]^2} \left[\left\{\widehat{C}(u_1,u_2)\right\}^2 - \left\{\mathrm{E}\left(\widehat{C}(u_1,u_2)\right)\right\}^2\right] \mathrm{d}u_1\mathrm{d}u_2.
\end{aligned}$$

The latter have been estimated from 1 000 replicates under each of the scenario considered; the results are reported in Table A1 for AISB and in Table A2 for AIV. All the simulations have been done using the Local–Linear weights defined for $\widetilde{X}_i = (X_i - x)/h$ by

$$w_{ni}(x,h) = K(\widetilde{X}_i)\left(\frac{S_{n,2} - \widetilde{X}_i\, S_{n,1}}{S_{n,0}\, S_{n,2} - S_{n,1}^2}\right),$$

where $K(y) = 35(1-y^2)^3\,\mathbb{I}(|y| \le 1)/32$ is the triweight function and $S_{n,\ell} = \sum_{i=1}^n \widetilde{X}_i^\ell\, K(\widetilde{X}_i)$, $\ell \in \{0,1,2\}$. Since negative weights can occur, they are taken to be zero in that case. By arguments similar as those in Li and Racine (2007), one can show that Assumptions W_1–W_{13} are satisfied whenever Assumption $\mathcal{A}_1$ on the alpha-mixing coefficients is satisfied.

From the entries in Table A1, one notes that $\widetilde{C}_{xh}$ outperforms C_{xh} in terms of AISB when $(\rho_{12}, \rho_{23}, \rho_{13}) \in \{(0.9, 0.8, 0.8), (-0.9, 0.8, -0.8)\}$; an explanation is the fact that $\mathrm{E}(\mathbb{C}_{xh})$ depends in general on F_{1x} and F_{2x}. This explanation is reinforced by the fact that their corresponding AISB are similar under the scenarios where $(\rho_{12}, \rho_{23}, \rho_{13}) \in$

$\{(0.8, 0.1, 0.1), (0.1, 0.1, 0.1)\}$, *i.e.* in cases where the influence on the marginal distributions is rather weak. Also observe that the integrated bias of $\widetilde{C}_{xh}$ stabilizes as the bandwidth parameter h takes large values; it is not the case for C_{xh}.

As can be seen in Table A2, the integrated variance is very similar for any values of θ. This is in accordance with the theoretical results that states that the estimators act, asymptotically, as in the i.i.d. case; in the model that was considered, it corresponds to $\theta = 0$. Generally speaking, C_{xh} do slightly better than $\widetilde{C}_{xh}$. Finally note that both AISB and AIV take smaller values when $n = 1\ 000$ compared to $n = 250$, as expected.

Acknowledgements

This research was supported in part by individual grants from the Natural Sciences and Engineering Research Council of Canada (NSERC) and by the Canadian Statistical Science Institute (CANSSI).

References

Bickel, P.J., and Wichura, M.J. (1971), 'Convergence criteria for multiparameter stochastic processes and some applications', *Ann. Math. Statist.*, 42, 1656–1670.

Billingsley, P. (1968), *Convergence of probability measures*, John Wiley & Sons, Inc., New York-London-Sydney.

Bücher, A., and Volgushev, S. (2013), 'Empirical and sequential empirical copula processes under serial dependence', *J. Multivariate Anal.*, 119, 61–70. http://dx.doi.org/10.1016/j.jmva.2013.04.003.

Carrasco, M., and Chen, X. (2002), 'Mixing and moment properties of various GARCH and stochastic volatility models', *Econometric Theory*, 18, 17–39. http://dx.doi.org/10.1017/S0266466602181023.

Doukhan, P. (1994), *Mixing*, *Lecture Notes in Statistics*, Vol. 85, Springer-Verlag, New York, http://dx.doi.org/10.1007/978-1-4612-2642-0. Properties and examples.

Gijbels, I., Veraverbeke, N., and Omelka, M. (2011), 'Conditional copulas, association measures and their applications', *Comput. Statist. Data Anal.*, 55, 1919–1932. http://dx.doi.org/10.1016/j.csda.2010.11.010.

Li, Q., and Racine, J.S. (2007), *Nonparametric econometrics*, Princeton University Press, Princeton, NJ. Theory and practice.

Meitz, M., and Saikkonen, P. (2008), 'Ergodicity, mixing, and existence of moments of a class of Markov models with applications to GARCH and ACD models', *Econometric Theory*, 24, 1291–1320. http://dx.doi.org/10.1017/S0266466608080511.

Nelsen, R.B. (2006), *An introduction to copulas*, 2nd ed., Springer Series in Statistics, New York: Springer.

Patton, A.J. (2006), 'Modelling asymmetric exchange rate dependence', *Internat. Econom. Rev.*, 47, 527–556. http://dx.doi.org/10.1111/j.1468-2354.2006.00387.x.

van der Vaart, A.W., and Wellner, J.A. (1996), *Weak convergence and empirical processes*, Springer Series in Statistics, New York: Springer-Verlag. With applications to statistics.

Veraverbeke, N., Omelka, M., and Gijbels, I. (2011), 'Estimation of a conditional copula and association measures', *Scand. J. Stat.*, 38, 766–780. http://dx.doi.org/10.1111/j.1467-9469.2011.00744.x.

Appendix A. Assumptions needed in Proposition 2.1, Proposition 2.2 and Proposition 3.1

A.1. *Distributional assumptions*

$\mathcal{A}_1$. The α-mixing coefficients of $\{(Y_{1t}, Y_{2t}, X_t)\}_{t\in\mathbb{Z}}$ are such that $\alpha(r) = O(r^{-a})$ for some $a > 6$.

$\mathcal{A}_2$. The functions $(w, y_1, y_2) \mapsto H_w(y_1, y_2)$, $\dot{H}_w = \partial H_w/\partial w$ and $\ddot{H}_w = \partial^2 H_w/\partial w^2$ exist and are uniformly continuous in $(w, y_1, y_2) \in J_x \times \mathbb{R}^2$, where J_x is an open neighborhood of x.

$\mathcal{A}_3$. The partial derivatives $C_x^{[1]}(u_1, u_2) = \partial\, C_x(u_1, u_2)/\partial u$ and $C_x^{[2]}(u_1, u_2) = \partial\, C_x(u_1, u_2)/\partial v$ exist and are continuous respectively on $(0, 1) \times [0, 1]$ and $[0, 1] \times (0, 1)$.

$\mathcal{A}_2^\star$. The functions $(w, u_1, u_2) \mapsto C_w(u_1, u_2)$, $\dot{C}_w = \partial C_w/\partial w$ and $\ddot{C}_w = \partial^2 C_w/\partial w^2$ exist and are uniformly continuous in $(w, u_1, u_2) \in J_x \times [0, 1]^2$, where J_x is an open neighborhood of x.

$\mathcal{A}_4$. For $j = 1, 2$, the functions $(w, u) \mapsto F_{jw}\{F_{jw}^{-1}(u_1)\}$, $\dot{F}_{jw}\{F_{jw}^{-1}(u_2)\}$ and $\ddot{F}_{jw}\{F_{jw}^{-1}(u_1)\}$ exist and are continuous in $(w, u) \in J_x \times [0, 1]$, where J_x is an open neighborhood of x.

A.2. *Assumptions on the weights*

W_1. $\max_{1\le i\le n} |w_{ni}(x, h)| = O_{\mathbb{P}}\left((nh)^{-1}\right)$;

W_2. $\sum_{i=1}^n w_{ni}(x, h)(X_i - x) = h^2 K_2 + o_{\mathbb{P}}\left((nh)^{-1/2}\right)$ for some $K_2 < \infty$;

W_3. $\sum_{i=1}^n w_{ni}(x, h)(X_i - x)^2/2 = h^2 K_3 + o_{\mathbb{P}}\left((nh)^{-1/2}\right)$ for some $K_3 < \infty$;

W_4. $nh \sum_{i=1}^n \{w_{ni}(x, h)\}^2 = K_4 + o_{\mathbb{P}}(1)$ for some $K_4 > 0$;

W_5. $\max_{i\in I_{nx}} X_i - \min_{i\in I_{nx}} X_i = o_{\mathbb{P}}(1)$, where $I_{nx} = \{j \in \{1, \ldots, n\} : w_{nj}(x, h) > 0\}$.

W_6. $\sup_{z\in J_x^{(n)}} \sum_{i=1}^n |w'_{ni}(z, h)| = O_{\mathbb{P}}(h^{-1})$, where $J_x^{(n)} = [\min_{i\in I_{nx}} Y_i, \max_{i\in I_{nx}} Y_i]$;

W_7. $\sup_{z\in J_x^{(n)}} \sum_{i=1}^n \{w'_{ni}(z, h)\}^2 = O_{\mathbb{P}}(n^{-1}h^{-3})$;

W_8. For some finite constant C,

$$\mathbb{P}\left(\sup_{z\in J_x^{(n)}} \max_{1\le i\le n} |w_{ni}(z, h)\, \mathbb{I}\left(|X_i - x| > Ch\right)| > 0\right) = o_{\mathbb{P}}(1);$$

W_9. There exists $D_K < \infty$ such that for all a_n,

$$\sup_{z \in J_x^{(n)}} \left| \sum_{i=1}^{n} w_{ni}(z, a_n)(X_i - z) - a_n^2 D_k \right| = o_{\mathbb{P}}(a_n^2);$$

W_{10}. There exists $E_K < \infty$ such that for all a_n,

$$\sup_{z \in J_x^{(n)}} \left| \sum_{i=1}^{n} w_{ni}(z, a_n)(X_i - z)^2 - a_n^2 E_k \right| = o_{\mathbb{P}}(a_n^2).$$

In order to establish moment inequalities of order r, one needs that for any integer $1 \leq k \leq 6$, any choice of $L_1, \ldots, L_k \in \mathbb{N}$ such that $L_1 + \cdots + L_k = 6$, and for some positive sequence υ_n satisfying $n - \upsilon_n \to \infty$:

W_{11}. $\displaystyle\sup_{z \in J_x} \max_{1 \leq \ell_2 < \cdots < \ell_k \leq \upsilon_n} \sum_{i=1}^{n-\ell_k} w_{ni}(z, h)^{L_1} \prod_{j=2}^{k} w_{n,i+\ell_j}(z, h)^{L_j} = O_{\mathbb{P}}\left(\frac{h^{k-1}}{(nh)^{r-1}} \right);$

W_{12}. $\displaystyle\sup_{z \in J_x} \sum_{i=1}^{n} w_{ni}(z, h)^{L_1} = O_{\mathbb{P}}\left((nh)^{-L_1+1} \right);$

W_{13}. $\displaystyle\sup_{z \in J_x} \max_{1 \leq \ell_2 < \cdots < \ell_k \leq \upsilon_n} \sum_{i=1}^{n-\ell_k} (X_i - z) w_{ni}(z, h)^{L_1} \prod_{j=2}^{k} w_{n,i+\ell_j}(z, h)^{L_j} = O_{\mathbb{P}}\left(\frac{h^{k+1}}{(nh)^{r-1}} \right).$

Table A1. Average integrated squared bias ($\times 10^4$) of C_{xh} and $\widetilde{C}_{xh}$, as estimated from 1 000 replicates of a first-order autoregressive Gaussian process. Upper panel: $n = 250$; bottom panel: $n = 1\,000$

		$(.9, .8, .8)$		$(-.9, .8, -.8)$		$(.8, .1, .1)$		$(.1, .1, .1)$	
θ	h	C_{xh}	$\widetilde{C}_{xh}$	C_{xh}	$\widetilde{C}_{xh}$	C_{xh}	$\widetilde{C}_{xh}$	C_{xh}	$\widetilde{C}_{xh}$
	0.5	0.934	1.359	1.461	1.058	1.192	1.253	0.959	0.974
	1.0	0.125	0.371	1.075	0.291	0.337	0.364	0.313	0.325
0.0	1.5	0.480	0.215	1.668	0.137	0.178	0.202	0.136	0.153
	2.0	1.165	0.154	2.501	0.107	0.111	0.127	0.060	0.069
	2.5	1.852	0.108	3.234	0.088	0.091	0.108	0.062	0.073
	0.5	0.968	1.381	1.418	1.100	1.254	1.334	1.097	1.115
	1.0	0.116	0.370	1.039	0.303	0.344	0.365	0.292	0.303
0.2	1.5	0.447	0.218	1.623	0.140	0.180	0.199	0.134	0.144
	2.0	1.149	0.152	2.459	0.110	0.112	0.132	0.063	0.075
	2.5	1.817	0.105	3.176	0.088	0.093	0.113	0.063	0.080
	0.5	1.044	1.411	1.619	1.263	1.343	1.473	1.336	1.361
	1.0	0.116	0.391	0.996	0.333	0.337	0.377	0.285	0.284
0.4	1.5	0.362	0.233	1.484	0.154	0.192	0.210	0.155	0.158
	2	1.011	0.152	2.261	0.115	0.118	0.138	0.076	0.089
	2.5	1.638	0.108	2.974	0.091	0.098	0.117	0.071	0.086
	0.5	0.026	0.085	0.147	0.061	0.068	0.074	0.046	0.046
	1.0	0.218	0.021	0.443	0.019	0.018	0.021	0.016	0.018
0.0	1.5	0.869	0.012	1.202	0.011	0.010	0.012	0.009	0.013
	2.0	1.646	0.008	2.025	0.008	0.006	0.008	0.004	0.007
	2.5	2.367	0.006	2.749	0.008	0.006	0.008	0.003	0.006
	0.5	0.027	0.082	0.147	0.062	0.074	0.079	0.056	0.058
	1.0	0.206	0.020	0.421	0.02	0.018	0.021	0.015	0.017
0.2	1.5	0.816	0.013	1.135	0.011	0.011	0.013	0.01	0.013
	2.0	1.612	0.008	1.957	0.007	0.006	0.008	0.004	0.006
	2.5	2.323	0.006	2.702	0.007	0.005	0.007	0.003	0.006
	0.5	0.035	0.088	0.153	0.077	0.081	0.086	0.074	0.074
	1.0	0.159	0.021	0.370	0.023	0.019	0.021	0.019	0.021
0.4	1.5	0.683	0.013	0.978	0.012	0.010	0.012	0.013	0.017
	2.0	1.447	0.007	1.786	0.009	0.006	0.008	0.005	0.008
	2.5	2.156	0.005	2.501	0.007	0.005	0.007	0.004	0.008

Table A2. Average integrated variance ($\times 10^4$) of C_{xh} and $\widetilde{C}_{xh}$, as estimated from 1 000 replicates of a first-order autoregressive Gaussian process. Upper panel: $n = 250$; bottom panel: $n = 1\ 000$

		$(.9, .8, .8)$		$(-.9, .8, -.8)$		$(.8, .1, .1)$		$(.1, .1, .1)$	
θ	h	C_{xh}	$\widetilde{C}_{xh}$	C_{xh}	$\widetilde{C}_{xh}$	C_{xh}	$\widetilde{C}_{xh}$	C_{xh}	$\widetilde{C}_{xh}$
	0.5	3.589	3.778	3.602	3.863	2.924	2.944	6.379	6.440
	1.0	1.649	1.882	1.636	1.894	1.412	1.428	3.287	3.300
0.0	1.5	1.052	1.337	1.087	1.382	0.992	0.995	2.466	2.459
	2.0	0.828	1.173	0.822	1.147	0.842	0.851	1.963	1.987
	2.5	0.671	1.016	0.688	1.054	0.745	0.752	1.850	1.845
	0.5	3.732	3.926	3.742	3.957	2.977	3.031	6.547	6.642
	1.0	1.716	1.967	1.735	1.968	1.473	1.477	3.407	3.394
0.2	1.5	1.115	1.360	1.094	1.401	1.011	1.014	2.495	2.482
	2.0	0.832	1.171	0.832	1.160	0.849	0.853	2.003	2.028
	2.5	0.673	1.026	0.698	1.060	0.757	0.769	1.869	1.882
	0.5	4.069	4.260	4.057	4.196	3.148	3.220	7.204	7.247
	1.0	1.851	2.061	1.853	2.059	1.528	1.544	3.653	3.667
0.4	1.5	1.210	1.484	1.198	1.487	1.077	1.085	2.749	2.755
	2.0	0.900	1.257	0.880	1.219	0.889	0.894	2.214	2.213
	2.5	0.712	1.084	0.741	1.106	0.796	0.810	2.012	2.020
	0.5	0.793	0.825	0.792	0.835	0.617	0.618	1.524	1.532
	1.0	0.385	0.434	0.389	0.444	0.316	0.316	0.799	0.796
0.0	1.5	0.257	0.324	0.246	0.315	0.233	0.233	0.588	0.592
	2.0	0.197	0.272	0.192	0.265	0.198	0.198	0.502	0.503
	2.5	0.162	0.255	0.164	0.249	0.183	0.185	0.453	0.455
	0.5	0.803	0.837	0.800	0.842	0.624	0.628	1.553	1.546
	1.0	0.397	0.443	0.408	0.466	0.319	0.320	0.787	0.789
0.2	1.5	0.263	0.332	0.253	0.325	0.232	0.234	0.598	0.602
	2.0	0.197	0.274	0.194	0.269	0.198	0.201	0.509	0.511
	2.5	0.168	0.256	0.164	0.253	0.185	0.188	0.469	0.470
	0.5	0.874	0.909	0.872	0.903	0.674	0.679	1.639	1.651
	1.0	0.432	0.478	0.437	0.499	0.338	0.337	0.842	0.844
0.4	1.5	0.283	0.357	0.278	0.354	0.245	0.246	0.641	0.644
	2.0	0.206	0.290	0.199	0.283	0.204	0.206	0.548	0.548
	2.5	0.178	0.269	0.175	0.269	0.196	0.197	0.512	0.515

Supplementary material for the paper *Estimation of a conditional copula based on time series*

1. Asymptotic normality of Z_{xn}

It will be shown that for any arbitrary $y, z \in \mathbb{R}$, the random variable $Z_{xn}(y,z)$ is asymptotically normal. The arguments can easily be adapted to show the joint weak convergence of $Z_{nx}(y_1,z_1),\ldots,Z_{nx}(y_K,z_K)$ by invoking the Cramér–Wold device.

To prove the asymptotic normality of $Z_{xn}(y,z)$, a blocking technique described for instance in Billingsley (1968) will be used. To this end, write $n = r_n(b_n+\ell_n)$ and assume without loss of generality that for for some $\delta,\epsilon > 0$, r_n, b_n and ℓ_n are integers such that $b_n \sim n^{1-\epsilon+\delta}$, $\ell_n \sim n^{1-\epsilon}$ and $\ell_n > h^{-1}$. Then, introduce for each $i \in \{1,\ldots,n\}$ the following sets:

$$\begin{aligned}\mathcal{U}_i &= \{j \in \mathbb{N} : (i-1)(b_n+\ell_n)+1 \le j \le (i-1)(b_n+\ell_n)+b_n\}\,;\\ \mathcal{W}_i &= \{j \in \mathbb{N} : i(b_n+\ell_n)+1-\ell_n \le j \le i(b_n+\ell_n)\}\,.\end{aligned}$$

Letting

$$U_{ni} = \sum_{j\in\mathcal{U}_i} w_{nj}(x,h)\,\vartheta_j(y,z) \quad \text{and} \quad W_{ni} = \sum_{j\in\mathcal{W}_i} w_{nj}(x,h)\,\vartheta_j(y,z),$$

one can write $Z_{xn}(y,z) = U_n + W_n$, where

$$U_n = \sqrt{nh}\sum_{i=1}^{r_n} U_{ni} \quad \text{and} \quad W_n = \sqrt{nh}\sum_{i=1}^{r_n} W_{ni}.$$

It will next be demonstrate that for an appropriate choice of ϵ and δ, the random variable W_n is asymptotically negligible while U_n is asymptotically normal.

Firstly, computation presented in Section 2 shows, that as long as $\ell_n > h^{-1}$, and since $\mathcal{A}_1$ holds, there exists a constant C_α that depends on the α-mixing coefficients and on the weight functions such that for sufficiently large n,

$$\sum_{i=1}^{r_n} \mathrm{E}\left(|W_{ni}|^4\right) \le C_\alpha n^{-3}h^{-2}\ell_n.$$

Hence, for any $\kappa > 0$, one has that

$$\begin{aligned}\mathbb{P}(|W_n| > \kappa) &\le \sum_{i=1}^{r_n}\mathbb{P}\left(|W_{ni}| > \frac{\kappa}{\sqrt{nh}\,r_n}\right)\\ &\le \frac{r_n^4(nh)^2}{\kappa^4}\sum_{i=1}^{r_n}\mathrm{E}\left(|W_{ni}|^4\right)\\ &\le \frac{C_\alpha\, r_n^4\,\ell_n}{\kappa^4\, n} \sim n^{3\epsilon-4\delta}.\end{aligned}$$

This last expression tends to zero as $n \to \infty$ whenever $3\epsilon < 4\delta$, which would entail $\mathbb{P}(|W_n| > \kappa) \to 0$.

Secondly, in order to deal with U_n, let $U'_{n1}, \ldots, U'_{nr_n}$ be independent random variables sharing the same conditional distribution as $U_{n1}, \ldots, U_{nr_n}$. Based on Billingsley (1968), p. 376, one can show that the respective characteristic functions of U_n and $U'_n = \sqrt{nh}\sum_{i=1}^{r_n} U'_{n,i}$ differ by at most $16\, r_n\, \alpha_{\ell_n}$. Since assumption $\mathcal{A}_1$ ensures $\alpha(\ell) \sim \ell^{-a}$ for some $a > 6$, one obtains that

$$16 r_n \alpha_{\ell_n} \sim n^{(a+1)\epsilon - a - \delta}.$$

Therefore $16\, r_n\, \alpha_{\ell_n} \to 0$ whenever $(a+1)\epsilon < a + \delta$. As a consequence, U_n and U'_n are asymptotically equivalent. Then, it suffices to show that U'_n is asymptotically normal to prove the asymptotic normality of U_n. To do this, it will next be establish that the random variable U_n satisfies the Lyapunov condition required for the central limit theorem. By a straightforward computation, $\mathrm{Var}(\sqrt{nh}\sum_{i=1}^{r_n} U'_{ni}) = K_4\, \sigma_x^2(y,z) + o(1)$. Moreover, it is shown in Section 2 that

$$\sum_{i=1}^{r_n} \mathrm{E}\left\{(U'_{ni})^4\right\} = O\left(\frac{b_n}{n^3 h^2}\right).$$

For any $\kappa > 0$, one then has

$$\begin{aligned}\sum_{i=1}^{r_n} \mathrm{E}\left[\sqrt{nh}\,(U'_{ni})^2\, \mathbb{I}\left\{\sqrt{nh}(U'_{ni})^2 > \kappa\right\}\right] &\leq \left(\frac{nh}{\kappa}\right)^2 \sum_{i=1}^{r_n} \mathrm{E}\left\{(U'_{ni})^4\right\} \\ &= O\left(\frac{b_n}{n}\right) \sim n^{\delta-\epsilon}.\end{aligned}$$

As long as $\epsilon > \delta$, this last expression is $o(1)$, leading to a Lyapunov ratio that converges to zero. Thus, if τ is such that $h \sim n^{-\tau}$, letting $\epsilon = 4/5 \min\{a/(a+1), 1-\tau\}$ and $\delta = 4/5\epsilon$ ensures the asymptotic normality of Z_{xh}.

2. Computation of $\sum_{k=1}^{r_n} \mathrm{E}\, W_{nk}^4$ and $\sum_{k=1}^{r_n} \mathrm{E}\, U'^4_{nk}$

Fix $y, z \in \mathbb{R}$ and for simplicity let ϑ_i stand for $\vartheta_i(y,z)$. For readability of the presentation, we write $\sum_{\mathcal{W}_k}$ when all the indices involved in the summation are in $\mathcal{W}_k$. The computations to be exposed next are valid provided assumptions $\mathcal{A}_1$, W_{11}–W_{12} are satisfied, as well as $\ell_n > h^{-1}$.

First decompose $W_{nk}^4 = \sum_{j=1}^5 W_{nk}^{(j)}$, where

$$W_{nk}^{(1)} = \sum_{\mathcal{W}_k} \vartheta_i^4 w_{ni}(x,h)^4, \qquad W_{nk}^{(4)} = 6 \sum_{\substack{\mathcal{W}_k \\ i_1 \neq i_2 \neq i_3}} \vartheta_{i_1}^2 \vartheta_{i_2} \vartheta_{i_3} w_{ni_1}(x,h)^2 w_{ni_2}(x,h) w_{ni_3}(x,h)$$

$$W_{nk}^{(2)} = 4 \sum_{\substack{\mathcal{W}_k \\ i_1 \neq i_2}} \vartheta_{i_1}^3 \vartheta_{i_2} w_{ni_1}(x,h)^3 w_{ni_2}(x,h), \qquad W_{nk}^{(5)} = \sum_{\substack{\mathcal{W}_k \\ i_1 \neq \ldots \neq i_4}} \prod_{k=1}^4 \vartheta_{i_k} w_{ni_k}(x,h)$$

$$W_{nk}^{(3)} = 6 \sum_{\substack{\mathcal{W}_k \\ i_1 \neq i_2}} \vartheta_{i_1}^2 \vartheta_{i_2}^2 w_{ni_1}(x,h)^2 w_{ni_2}(x,h)^2.$$

The goal is now to bound $\sum_{k=1}^{r_n} \mathrm{E}\, W_{nk}^{(j)}$ for each j. We begin with $j = 1$. In view of condition W_{12}, $\sum_{i=1}^n w_{ni}(x,h)^4 \leq K_{12} n^{-3} h^{-3}$ for some $K_{12} < \infty$. As $\ell_n > h^{-1}$, one

obtains that

$$\sum_{k=1}^{r_n} \mathrm{E}\, W_{nk}^{(1)} \leq K_{12} n^{-3} h^{-3} \leq K_{12} n^{-3} h^{-2} \ell_n.$$

To deal with $j = 2$, split $W_{nk}^{(2)}$ in $W_{nk}^{(2,<)}$ and $W_{nk}^{(2,>)}$ according to the cases $i_1 < i_2$ and $i_1 > i_2$. Next, as the random variables $\vartheta_1, \ldots, \vartheta_n$ satisfy a strong mixing assumption, it is useful to note that for any $1 \leq \ell \leq \ell_n - i$,

$$|\mathrm{E}(\vartheta_i^3 \vartheta_{i+\ell})| \leq \alpha(\ell) \leq 1$$

Letting π_h denote the integer part of $1/\sqrt{h}$, one writes

$$\begin{aligned} \sum_{k=1}^{r_n} \mathrm{E}\, W_{nk}^{(2,<)} \quad \leq \quad & \sum_{i=1}^{n} \sum_{l=1}^{\pi_h} \mathrm{E}\left(\vartheta_i^3 \vartheta_{i+\ell}\right) w_{ni}(x,h)^3 w_{n,i+\ell}(x,h) \\ & + \sum_{i=1}^{n} \sum_{l=\pi_h+1}^{l_n - i} \mathrm{E}\left(\vartheta_i^3 \vartheta_{i+\ell}\right) w_{ni}(x,h)^3 w_{n,i+\ell}(x,h). \end{aligned}$$

The first summand is bounded by

$$\pi_h \max_{1 \leq \ell \leq \pi_h} \sum_{i=1}^{n-\ell} w_{ni}(x,h)^3 w_{h,i+\ell}(x).$$

As W_{11} holds and $\ell_n > \pi_h$, the latter is $O(n^{-3} h^{-2} \ell_n)$. The second summand is bounded by

$$\left\{ \max_{1 \leq i \leq n} w_{ni}(x,h) \right\} \left\{ \sum_{i=1}^{n} w_{ni}(x,h)^3 \right\} \sum_{\ell=\pi_h+1}^{d} \alpha(\ell),$$

which can be bounded by

$$O(n^{-3} h^{-2} \ell_n) \times \sum_{\ell=\pi_h+1}^{d} \alpha(\ell)$$

since W_1 and W_{12} are satisfied together with $\ell_n > h^{-1}$. As assumption $\mathcal{A}_1$ holds, $\sum_{\ell=\pi_h+1}^{d} \alpha(\ell) \to 0$ and therefore the latter display is $o(n^{-3} h^{-2} \ell_n)$. As a consequence,

$$\sum_{k=1}^{r_n} \mathrm{E}\, W_{nk}^{(2,<)} = O(n^{-3} h^{-2} \ell_n).$$

Proceeding similarly for $W_{nk}^{(2,>)}$ yields

$$\sum_{k=1}^{r_n} \mathrm{E}\, W_{nk}^{(2)} = O(n^{-3} h^{-2} \ell_n)$$

In the case $j = 3$, one decomposes $W_{nk}^{(3)} = \overline{W}_{nk}^{(3)} + (W_{nk}^{(3)} - \overline{W}_{nk}^{(3)})$, where

$$\overline{W}_{nk}^{(3)} = 2 \sum_{i \in \mathcal{W}_k} \sum_{\ell=1}^{\ell_n - i} \mathrm{E}(\vartheta_i^2)\mathrm{E}(\vartheta_{i+\ell}^2) w_{ni}(x,h) w_{n,i+\ell}(x,h).$$

As $|\vartheta_i| \leq 1$, one deduces that

$$\sum_{k=1}^{r_n} \overline{W}_{nk}^{(3)} \quad \leq \quad \sum_{i=1}^{n} \sum_{\ell=1}^{\ell_n \wedge (n-i)} w_{ni}(x,h)^2 w_{n,i+\ell}(x,h)^2 = O(\ell_n n^{-3} h^{-2}),$$

where the last equality follows from assumption W_{11}. Next, since $|\mathrm{E}(\vartheta_i^2 \vartheta_{i+\ell}^2) - \mathrm{E}(\vartheta_i^2)\mathrm{E}(\vartheta_{i+\ell}^2)| \leq \alpha(\ell)$, one uses the same " splitting " strategy as previously to obtain that $\sum_{k=1}^{r_n}(W_{nk}^{(3)} - \overline{W}_{nk}^{(3)})$ is $O(n^{-3}h^{-2}\ell_n)$. We go directly to the case $j = 5$, as the case $j = 4$ can be dealt similarly but the computations are more involved in the former. Denote for each index $m = 1, 2, 3$ the sets $\mathcal{W}_{nk}^{(m)} = \{i_1 < ... < i_4 \in \mathcal{W}_{nk} : \max(g_1, g_2, g_3) \leq g_m\}$, where $g_m = i_{m+1} - i_m$ is the gap between two consecutive indices. The introduction of these sets of indices is justified by the fact that $\mathrm{E}\, W_{nk}^{(5)} \leq 4!(T_{nk}^{(1)} + T_{nk}^{(2)} + T_{nk}^{(3)})$, where

$$T_{nk}^{(m)} \quad = \quad \sum_{\mathcal{W}_{nk}^{(m)}} \mathrm{E}\left\{\prod_{\ell=1}^{4} \vartheta_{i_\ell} w_{ni_\ell}(x,h)\right\}.$$

As previously, one then re-decompose $T_{nk}^{(m)} = \overline{T}_{nk}^{(m)} + (T_{nk}^{(m)} - \overline{T}_{nk}^{(m)})$, where

$$\overline{T}_{nk}^{(m)} \quad = \quad \sum_{\mathcal{W}_{nk}^{(m)}} \mathrm{E}\left(\prod_{\ell=l}^{m} \vartheta_{i_l}\right) \mathrm{E}\left(\prod_{l'=m+1}^{4} \vartheta_{i_{l'}}\right) \prod_{\ell=1}^{4} w_{ni_\ell}(x,h).$$

As the expectation of ϑ_i is zero, both $\overline{T}_{nk}^{(1)}$ and $\overline{T}_{nk}^{(3)}$ are 0. Using the fact that $|\mathrm{E}(\vartheta_{i_1}\vartheta_{i_2}\vartheta_{i_3}\vartheta_{i_4}) - \mathrm{E}(\vartheta_{i_1}\vartheta_{i_2})\mathrm{E}(\vartheta_{i_3}\vartheta_{i_4})| \leq \alpha(g_2)$, one obtains that

$$\begin{aligned}\sum_{k=1}^{r_n} \overline{T}_{nk}^{(2)} \;&\leq\; \ell_n \left\{\max_{1 \leq \ell \leq \ell_n} \sum_{i=1}^{n-\ell} w_{ni}(x,h) w_{n,i+\ell}(x,h)\right\} \left\{\max_{1 \leq i \leq n} w_{ni}(x,h)\right\}^2 \sum_{g_2=1}^{\ell_n} \alpha(g_2) \\ &=\; O(\ell_n n^{-3} h^{-2}) \sum_{g_2=1}^{\ell_n} \alpha(g_2),\end{aligned}$$

where the last inequality follows from condition W_1 and W_{11}. As the assumption over the alpha mixing coefficients implies that $\sum_{g_2=1}^{\infty} \alpha(g_2) < \infty$, it follows that the latter equation is $O(\ell_n n^{-3} h^{-2})$. Finally, as

$$\left|\mathrm{E}(\prod_{\ell=1}^{4} \vartheta_{i_\ell}) - \mathrm{E}(\prod_{\ell=l}^{j} \vartheta_{i_l})\mathrm{E}(\prod_{l'=j+1}^{4} \vartheta_{i_{l'}})\right| \leq \alpha(g_j),$$

one has

$$\begin{aligned} T_{nk}^{(j)} - \overline{T}_{nk}^{(j)} &= [\sum_{\substack{\mathcal{W}_{nk}^{(j)} \\ g_j \le \pi_h}} + \sum_{\substack{\mathcal{W}_{nk}^{(j)} \\ g_j > \pi_h}}] \left\{ \mathrm{E}\left(\prod_{\ell=l}^{j} \vartheta_{i_l}\right) - \mathrm{E}\left(\prod_{\ell=l}^{j} \vartheta_{i_l}\right) \mathrm{E}\left(\prod_{l'=j+1}^{4} \vartheta_{i_{l'}}\right) \right\} \\ &\quad \times \prod_{\ell=1}^{4} w_{ni_\ell}(x,h) \\ &\le \sum_{\substack{\mathcal{W}_{nk}^{(j)} \\ g_j \le \pi_h}} \prod_{\ell=1}^{4} w_{ni_\ell}(x,h) + \sum_{\substack{\mathcal{W}_{nk}^{(j)} \\ g_j > \pi_h}} \alpha(g_j) \prod_{\ell=1}^{4} w_{ni_\ell}(x,h). \end{aligned}$$

In view of last equation, one directly obtains

$$\begin{aligned} \sum_{k=1}^{r_n} (T_{nk}^{(j)} - \overline{T}_{nk}^{(j)}) &\le \pi_h \left\{ \max_{1 \le h_1 < h_2 < h_3 < \pi_h} \sum_{i=1}^{n-h_3} w_{ni}(x,h) \prod_{k=1}^{3} w_{n,i+h_k}(x,h) \right\} \\ &\quad + \left\{ \max_{1 \le i \le n} w_{ni}(x,h) \right\}^3 \sum_{g_j = \pi_h + 1}^{\ell_n} (g_j+1)^2 \alpha(g_j) \end{aligned}$$

From assumption W_1 and W_{11}, the first summand is $O(n^{-3}h^{-1}\pi_h)$. Moreover, as the assumption $\mathcal{A}_1$ over the alpha-mixing coefficient implies the finiteness of $\sum_{g_j=1}^{\infty}(g_j+1)^2\alpha(g_j)$, one deduces that the previous equation is $o(n^{-3}h^{-2}\ell_n)$. Wrapping up all the terms, one obtains the desired result, i.e $\sum_{k=1}^{r_n} \mathrm{E}\, W_{nk}^4 = O(n^{-3}h^{-2}\ell_n)$. Similar arguments leads to $\sum_{k=1}^{r_n} \mathrm{E}\, U_{nk}'^4 = O(n^{-3}h^{-2}b_n)$.

3. Proof of Lemma 2.3

First, define the product space $T_\gamma = T_\gamma^{(1)} \times T_\gamma^{(2)}$, where

$$\begin{aligned} T_\gamma^{(1)} &= \left\{ F_{1x}^{-1}(0), F_{1x}^{-1}\left(\frac{1}{\kappa_\gamma}\right), \ldots, F_{1x}^{-1}(1) \right\}, \\ T_\gamma^{(2)} &= \left\{ F_{2x}^{-1}(0), F_{2x}^{-1}\left(\frac{1}{\kappa_\gamma}\right), \ldots, F_{2x}^{-1}(1) \right\}. \end{aligned}$$

For $y \in \mathbb{R}$, define $\underline{y}_\gamma^{(1)} = \max\{\zeta \in I_\gamma^{(1)} : \zeta \le y\}$, $\overline{y}_\gamma^{(1)} = \min\{\zeta \in I_\gamma^{(1)} : \zeta \le y\}$, $\underline{y}_\gamma^{(2)} = \max\{\zeta \in I_\gamma^{(2)} : \zeta \le y\}$ and $\overline{y}_\gamma^{(2)} = \min\{\zeta \in I_\gamma^{(2)} : \zeta \le y\}$. With this notation,

$$F_{1x}\left(\overline{y}_\gamma^{(1)}\right) - F_{1x}\left(\underline{y}_\gamma^{(1)}\right) \le \frac{1}{\kappa_\gamma} \quad \text{and} \quad F_{2x}\left(\overline{z}_\gamma^{(2)}\right) - F_{2x}\left(\underline{z}_\gamma^{(2)}\right) \le \frac{1}{\kappa_\gamma}.$$

Now observe that for any $\omega = (y, z) \in \mathbb{R}^2$, one has for $\underline{\omega}_\gamma = (\underline{y}_\gamma^{(1)}, \underline{z}_\gamma^{(2)})$ that

$$\begin{aligned} Z_{xn}(\omega) - Z_{xn}(\underline{\omega}_\gamma) &\leq Z_{xn}(\overline{\omega}_\gamma) - Z_{xn}(\underline{\omega}_\gamma) \\ &\quad + 2\sqrt{nh}\sum_{i=1}^{n} w_{ni}(x,h)\left\{H_{X_i}(\overline{\omega}_\gamma) - H_{X_i}(\underline{\omega}_\gamma)\right\} \\ &= Z_{xn}(\overline{\omega}_\gamma) - Z_{xn}(\underline{\omega}_\gamma) + 2\sqrt{nh}\,\rho\left(\overline{\omega}_\gamma, \underline{\omega}_\gamma\right) + o(1). \end{aligned}$$

Since Assumption $\mathcal{A}_2$ holds, a Taylor expansion allows to write

$$\begin{aligned} &\sqrt{nh}\sum_{i=1}^{n} w_{ni}(x,h)\left\{H_{X_i}(\overline{\omega}_\gamma) - H_{X_i}(\underline{\omega}_\gamma)\right\} \\ &\quad = \sqrt{nh}\left\{H_x(\overline{\omega}_\gamma) - H_x(\underline{\omega}_\gamma)\right\} \\ &\qquad + \left\{\dot{H}_x(\overline{\omega}_\gamma) - \dot{H}_x(\underline{\omega}_\gamma)\right\}\sqrt{nh}\sum_{i=1}^{n} w_{ni}(x,h)(X_i - x) \\ &\qquad + \sqrt{nh}\sum_{i=1}^{n} w_{ni}(x,h)(X_i - x)^2\left\{\ddot{H}_{\zeta_i}(\overline{\omega}_\gamma) - \ddot{H}_{\zeta_i}(\underline{\omega}_\gamma)\right\}, \end{aligned} \tag{1}$$

where ζ_i lies between X_i and x. Now for any bivariate distribution function H with marginal distributions F_1 and F_2, one has for $\omega_1 = (y_1, z_1)$ and $\omega_2 = (y_2, z_2)$ that $|H(\omega_1) - H(\omega_2)| \leq |F_1(y_1) - F_1(y_2)| + |F_2(z_1) - F_2(z_2)|$. From Assumptions W_2–W_3 and $\mathcal{A}_2$, the right-hand side of equation (1) is bounded by $\sqrt{nh}\,\rho(\overline{\omega}_\gamma, \underline{\omega}_\gamma) + o(1)$. As a consequence, uniformly in $\omega \in \mathbb{R}^2$,

$$Z_{xn}(\omega) - Z_{xn}(\underline{\omega}_\gamma) \leq Z_{xn}(\overline{\omega}_\gamma) - Z_{xn}(\underline{\omega}_\gamma) + o(\sqrt{nh}h^2).$$

The negligibility of the remainder term $o(\sqrt{nh}h^2)$ is ensured by the fact that $\sqrt{nh}h^2 < \infty$. From similar arguments, one deduces

$$Z_{xn}(\underline{\omega}_\gamma) - Z_{xn}(\omega) \leq Z_{xn}(\overline{\omega}_\gamma) - Z_{xn}(\underline{\omega}_\gamma) + o(1).$$

Thus, for any $\omega_1, \omega_2 \in \mathbb{R}^2$,

$$\begin{aligned} |Z_{xn}(\omega_1) - Z_{xn}(\omega_2)| &\leq \left|Z_{xn}(\overline{\omega_1}_\gamma) - Z_{xn}(\underline{\omega_1}_\gamma)\right| \\ &\quad + \left|Z_{xn}(\overline{\omega_2}_\gamma) - Z_{xn}(\underline{\omega_2}_\gamma)\right| \\ &\quad + \left|Z_{xn}(\underline{\omega_1}_\gamma) - Z_{xn}(\underline{\omega_2}_\gamma)\right|. \end{aligned}$$

Since for n sufficiently large, $\rho(\omega_1, \omega_2) < \delta$ entails $\rho(\underline{\omega_1}_\gamma, \underline{\omega_2}_\gamma) < 2\delta$, it follows that $\mathfrak{W}_\delta(Z_{xn}, \mathbb{R}^2) \leq 3\,\mathfrak{W}_{2\delta}(Z_{xn}, T_\gamma)$. It remains to show that for any positive sequence δ_n that decreases to zero as $n \to \infty$ and for any $\epsilon > 0$, $\mathbb{P}\left(\mathfrak{W}_{\delta_n}\left(Z_{xn}, T_\gamma\right) > \epsilon\right)$ tends to zero. To this end, observe that $\mathfrak{W}_{\delta_n}\left(Z_{xn}, T_\gamma\right) = 0$ whenever $\delta_n < 2\kappa_\gamma^{-1}$, while $\mathfrak{W}_{\delta_n}\left(Z_{xn}, T_\gamma\right) \leq \mathfrak{W}_{2\kappa_\gamma}\left(Z_{xn}, T_\gamma\right)$ otherwise. One can then conclude that

$$\mathbb{P}\left(\mathfrak{W}_{\delta_n}\left(Z_{xn}, T_\gamma\right) \geq \epsilon\right) \leq \mathbb{P}\left(\max_{1 \leq i,j \leq \kappa_\gamma} \left|\mathbb{H}_{xh}\left\{A_\gamma(i,j)\right\}\right| \geq \epsilon\right).$$

4. Proof of Lemma 2.4

First note that since condition $\mathcal{A}_2$ holds, for any $i \in I_{nx}$,

$$\nu_{X_i}(A) = \nu_x(A) + \dot{\nu}_x(A)\{X_i - x\} + \frac{1}{2}\ddot{\nu}_{z_i}(A)\{X_i - x\}^2 \tag{2}$$

where z_i is between X_i and x. For simplicity let ϑ_i stand for $\vartheta_i(A)$ and ν_z for $\nu_z(A)$. Moreover, throughout this section, set $\nu_{xh} = \nu_x + h^2(\dot{\nu}_x + \ddot{\nu}_x)$. As a starting point, write $\mathcal{S}_j = \{\mathbf{L} = (L_1, \ldots, L_j) \in \{1, \ldots, 6\}^j : L_1 + \ldots + L_j = 6\}$ and notice that $|\mathbb{H}_{xh}(A)|^6 = (nh)^3 \sum_{j=1}^{6} \sum_{\mathbf{L}\in\mathcal{S}_j} T_{hx}^{(j)}(\mathbf{L})$, with

$$T_{hx}^{(j)}(\mathbf{L}) = \sum_{\substack{i_1\ldots i_j=1 \\ i_1\neq\ldots\neq i_j}}^{n} \prod_{k=1}^{j} \vartheta_{i_k}^{L_k} w_{ni_k}(x,h)^{L_k}.$$

The goal is now to bound each $T_{hx}^{(j)}(\mathbf{L})$, for $j = 1, \ldots, 6$. We begin with $j = 1$. In this case, $\mathcal{S}_1 = \{(6)\}$. One then obtains from equation (2) together with assumption $\mathcal{A}_2$ that

$$\begin{aligned} \mathrm{E}\, T_{hx}^{(1)}(1) \quad &\leq \textstyle\sum_{i=1}^{n} \nu_{X_i} w_{ni}(x,h)^6 = \nu_x \sum_{i=1}^{n} w_{ni}(x,h)^6 \\ &\quad + \dot{\nu}_x \textstyle\sum_{i=1}^{n} (X_i - x) w_{ni}(x,h)^6 + \frac{1}{2}\sum_{i=1}^{n} (X_i - x)^2 \ddot{\nu}_{z_i} w_{ni}(x,h)^6. \end{aligned}$$

Hence, in view of assumption W_{12}–W_{13} together with condition $\mathcal{A}_2$, the previous equation can be bounded by $\overline{\omega}_1(nh)^{-5}\{\nu_x + h^2(\dot{\nu}_x + \ddot{\nu}_x)\}$ for some $\overline{\omega}_1 > 0$. For $j = 2$, one first splits $T_{hx}^{(2)}(\mathbf{L})$ into $T_{hx}^{(2,<)}$ and $T_{hx}^{(2,>)}$ according to the cases $i_1 < i_2$ and $i_2 < i_1$. Starting with $T_{hx}^{(2,<)}$, one decomposes $T_{hx}^{(2,<)}(\mathbf{L}) = \overline{T}_{hx}^{(2)}(\mathbf{L}) + \{T_{hx}^{(2,<)}(\mathbf{L}) - \overline{T}_{hx}^{(2,<)}(\mathbf{L})\}$, where

$$\overline{T}_{hx}^{(2,<)}(\mathbf{L}) = \sum_{i_1<i_2} \mathrm{E}\,(\vartheta_{i_1}^{L_1})\mathrm{E}\,(\vartheta_{i_2}^{L_2}) w_{ni_1}(x,h)^{L_1} w_{ni_2}(x,h)^{L_2}.$$

Note that $\mathbf{L} \in \{(1,5),(5,1)\}$ implies $\overline{T}_{hx}^{(2,<)}(\mathbf{L}) = 0$. Otherwise, as

$$\overline{T}_{hx}^{(2,<)}(\mathbf{L}) \leq \sum_{i_1} \nu_{X_{i_1}} w_{ni_1}(x,h)^{L_1} \times \sum_{i_2} \nu_{X_{i_2}} w_{ni_2}(x,h)^{L_2},$$

one uses equation (2) together with assumptions W_{12}–W_{13} to deduce that

$$\overline{T}_{hx}^{(2,<)}(\mathbf{L}) \leq \omega_2 (nh)^{2-L_1-L_2}\{\nu_x + h^2(\dot{\nu}_x + \ddot{\nu}_x)\}^2 = \omega_2 (nh)^{2-L_1-L_2} \nu_{xh}^2 \tag{3}$$

for some $\omega_2 < \infty$. Next, let υ_n be the integer part of $\frac{n}{2}$. One then has

$$
\begin{aligned}
&T_{hx}^{(2,<)}(\mathbf{L}) - \overline{T}_{hx}^{(2,<)}(\mathbf{L}) \\
&= \sum_{\ell=1}^{\pi_h}\sum_{i=1}^{n-\ell}\left[\mathrm{E}(\vartheta_i^{L_1}\vartheta_{i+\ell}^{L_2}) - \mathbb{E}\left(\vartheta_i^{L_1}\right)\mathbb{E}\left(\vartheta_{i+\ell}^{L_2}\right)\right] w_{ni}(x,h)^{L_1} w_{n,i+\ell}(x)^{L_2} \\
&\quad + \sum_{\ell=\pi_h+1}^{n-1}\sum_{i=1}^{n-\ell}\left[\mathrm{E}(\vartheta_i^{L_1}\vartheta_{i+\ell}^{L_2}) - \mathbb{E}\left(\vartheta_i^{L_1}\right)\mathbb{E}\left(\vartheta_{i+\ell}^{L_2}\right)\right] w_{ni}(x,h)^{L_1} w_{n,i+\ell}(x)^{L_2} \\
&\leq \pi_h\left\{\max_{1\leq\ell\leq\pi_h}\sum_{i=1}^{n-\ell}\nu_{X_i} w_{ni}(x,h)^{L_1} w_{n,i+\ell}(x,h)^{L_2}\right\} \\
&\quad + \left\{\sum_{\ell=\pi_h+1}^{\upsilon_n}\alpha(\ell)\right\}\left\{\max_{\pi_h+1\leq\ell\leq\upsilon_n}\sum_{i=1}^{n-\ell} w_{ni}(x,h)^{L_1} w_{h,i+\ell}(x)^{L_2}\right\} \\
&\quad + \left\{\sum_{\ell=\upsilon_n+1}^{n-1}\alpha(\ell)\right\}\left\{\max_{\upsilon_n<\ell\leq n-1}\sum_{i=1}^{n-\ell} w_{ni}(x,h)^{L_1} w_{h,i+\ell}(x)^{L_2}\right\}.
\end{aligned}
$$

In view of equation (2), one uses assumption W_{11} and W_{13} to deduce that the first summand is $O(\pi_h h\nu_{xh}(nh)^{1-L_1-L_2}) = O(\nu_{xh}(nh)^{1-L_1-L_2})$. As assumption $\mathcal{A}_1$ holds, it follows that $\sum_{\ell=\gamma_n}^{n}\alpha(\ell) \sim O(\gamma_n^{-5})$ whenever $\gamma_n \to \infty$. Hence, from assumption W_{11}, the second summand is $O(h^6(nh)^{-L_1-L_2+1})$. Finally, as third summand is bounded by

$$
\left\{\sum_{\ell=\pi_h+1}^{\upsilon_n}\alpha(\ell)\right\}\left\{\sum_{i=1}^{n} w_{ni}(x,h)^{L_1}\right\}\left\{\max_{1\leq i\leq n} w_{ni}(x,h)\right\}^{L_2},
$$

one uses assumption W_1 and W_{12} to derive that

$$
T_{hx}^{(2,<)}(\mathbf{L}) - \overline{T}_{hx}^{(2,<)}(\mathbf{L}) \leq \omega_2'(nh)^{1-L_1-L_2}[\nu_{xh} + h^6] \tag{4}
$$

for some $\omega_2' > 0$. Wrapping last discussion, one concludes from equation (3) and (4) that

$$
T_{hx}^{(2,<)}(\mathbf{L}) \leq (nh)^{1-L_1-L_2}\left\{\omega_2(nh)\nu_{xh}^2 + \omega_2'(\nu_{xh} + h^6)\right\}.
$$

Identical arguments yields the same bound for $T_{hx}^{(2,>)}(\mathbf{L})$. Therefore, upon setting $\overline{\omega}_2 = 2(\omega_2 + \omega_2')$, one obtains:

$$
\begin{aligned}
T_{hx}^{(2)}(\mathbf{L}) &\leq \overline{\omega}_2\,(nh)^{1-L_1-L_2}\left\{(nh)\nu_{xh}^2 + \nu_{xh} + h^6\right\} \\
&= \overline{\omega}_2\,(nh)^{-5}\left\{(nh)\nu_{xh}^2 + \nu_{xh} + h^6\right\},
\end{aligned} \tag{5}
$$

where the last equality follows from the fact that $L_1 + L_2 = 6$. The case $j = 3$ calls for a special treatment. Denote for each $k = 1, 2$ the sets $\mathcal{T}_k = \{i_1 < i_2 < i_3 : \max(g_1, g_2) \leq g_k\}$, where $g_k = i_{k+1} - i_k$ is the gap between two consecutive indices. Next, let $\mathcal{P}$ be the set of all permutations σ over the indices $\{1, 2, 3\}$, i.e

$$
\mathcal{P} = \{\sigma : \{1,2,3\} \to \{1,2,3\} : \sigma(1) \neq \sigma(2) \neq \sigma(3)\}.
$$

The introduction of this notation is justified by the fact that $\mathrm{E}T_{hx}^{(3)}(\mathbf{L}) \leq \sum_{\sigma\in\mathcal{P}} W_{hx}^{(1,\sigma)}(\mathbf{L}) + W_{hx}^{(2,\sigma)}(\mathbf{L})$, where

$$W_{hx}^{(k,\sigma)}(\mathbf{L}) = \mathrm{E}\sum_{\mathcal{T}_k}\prod_{j=1}^{3} w_{ni_{\sigma(j)}}(x,h)^{L_{\sigma(j)}}\vartheta_{i_{\sigma(j)}}^{L_\sigma(j)}.$$

Note that the permutations play a similar role as the cases $i_1 < i_2$ and $i_2 < i_1$ required for the analysis of $T_{hx}^{(2)}$. However, as the treatment of $W_{hx}^{(k,\sigma)}$ is the same for all permutations, we consider only the case $(\sigma(1),\sigma(2),\sigma(3)) = (1,2,3)$ and we omit the σ in the notations.

Similarly as previously, one decomposes $W_{hx}^{(k)}(\mathbf{L}) = \overline{W}_{hx}^{(k)}(\mathbf{L}) + \{W_{hx}^{(k)}(\mathbf{L}) - \overline{W}_{hx}^{(k)}(\mathbf{L})\}$ with

$$\overline{W}_{hx}^{(k)}(\mathbf{L}) = \sum_{\mathcal{T}_k}\mathrm{E}\left(\prod_{j=1}^{k}\vartheta_{i_j}^{L_j}\right)\mathrm{E}\left(\prod_{j=k+1}^{3}\vartheta_{i_j}^{L_j}\right)\prod_{j=1}^{3} w_{ni_j}(x,h)^{L_j}.$$

We deal only with the term $\overline{W}_{hx}^{(1)}$, as the other case is identical. Note that $\overline{W}_{hx}^{(1)}(\mathbf{L}) = 0$ whenever $\mathbf{L} \in \{(1,b,5-b) : 1 \leq b \leq 4\}$. Otherwise, observe that

$$\overline{W}_{hx}^{(1)}(\mathbf{L}) = \sum_{g_1=1}^{n-1}\sum_{g_2=1}^{g_1}\mathrm{E}(\vartheta_{i_2}^{L_2}\vartheta_{i_3}^{L_3})w_{ni_2}(x,h)^{L_2}w_{ni_3}(x,h)^{L_3}\sum_{i_1=1}^{n-g_1}\mathrm{E}(\vartheta_{i_1}^{L_1})w_{ni_1}(x,h)^{L_1}.$$

Proceeding as previously, one uses equation (2) and assumption W_{12}-W_{13} to obtain:

$$\sum_{i_1=1}^{n-g_1}\mathrm{E}(\vartheta_{i_1}^{L_1})w_{ni_1}(x,h)^{L_1} \leq \omega_1(nh)^{-L_1+1}\nu_{xh}.$$

Moreover, one deduces from equation (5) that

$$\sum_{g_1=1}^{n-1}\sum_{g_2=1}^{g_1}\mathrm{E}(\vartheta_{i_2}^{L_2}\vartheta_{i_3}^{L_3})w_{ni_2}(x,h)^{L_2}w_{ni_3}(x,h)^{L_3} \leq \overline{\omega}_2(nh)^{1-L_2-L_3}\{(nh)\nu_{xh}^2 + \nu_{xh} + h^6\},$$

leading to

$$\overline{W}_{hx}^{(1)}(\mathbf{L}) \leq \omega_1\overline{\omega}_2(nh)^{-4}\{(nh)\nu_{xh}^3 + \nu_{xh}^2 + \nu_{xh}h^6\} \tag{6}$$

since the L_i's sum to 6. Finally,

$$
\begin{aligned}
&W_{hx}^{(1)}(\mathbf{L}) - \overline{W}_{hx}^{(1)}(\mathbf{L}) \\
&\quad = \sum_{\substack{\mathcal{T}_1 \\ g_1 \le \pi_h}} \left\{ \mathrm{E}\left(\prod_{j=1}^{3} \vartheta_{i_j}^{L_j}\right) - \mathrm{E}\left(\vartheta_{i_1}^{L_1}\right) \mathrm{E}\left(\vartheta_{i_2}^{L_2}\vartheta_{i_3}^{L_3}\right) \right\} \prod_{j=1}^{3} w_{ni_j}(x,h)^{L_j} \\
&\qquad + \sum_{\substack{\mathcal{T}_1 \\ g_1 > \pi_h}} \left\{ \mathrm{E}\left(\prod_{j=1}^{3} \vartheta_{i_j}^{L_j}\right) - \mathrm{E}\left(\vartheta_{i_1}^{L_1}\right) \mathrm{E}\left(\vartheta_{i_2}^{L_2}\vartheta_{i_3}^{L_3}\right) \right\} \prod_{j=1}^{3} w_{ni_j}(x,h)^{L_j} \\
&\quad \le \pi_h^2 \left\{ \max_{1 \le g_2 < g_1 \le \pi_h} \sum_{i_1=1}^{n-\ell} \nu_{X_{i_1}} \prod_{j=1}^{3} w_{ni_j}(x,h)^{L_j} \right\} \\
&\qquad + \sum_{\substack{\mathcal{T}_1 \\ \pi_h \le g_1 \le \upsilon_n}} \alpha(g_1) \prod_{j=1}^{3} w_{ni_j}(x,h)^{L_j} + \sum_{\substack{\mathcal{T}_1 \\ g_1 \ge \upsilon_n}} \alpha(g_1) \prod_{j=1}^{3} w_{ni_j}(x,h)^{L_j}.
\end{aligned}
$$

In view of equation (2), assumption W_{11} and W_{13}, the first summand is $O\{\nu_{xh}(nh)^{-5}\}$ since $\pi_h^2 h^2 \to 1$. The second summand is bounded by

$$
\begin{aligned}
\sum_{g_1=\pi_h}^{\upsilon_n} (g_1+1)\alpha(g_1) \max_{1<\ell_1<\ell_2<\pi_h} \sum_{i=1}^{n} w_{ni}(x,h)^{L_1} w_{n,i+\ell_1}(x,h)^{L_2} w_{n,i+\ell+2}(x,h)^{L_3} \\
+ \sum_{g_1=\upsilon_n}^{n} (g_1+1)\alpha(g_1) \times \upsilon_n \max_{1 \le i \le n} w_{ni}(x,h)^3.
\end{aligned}
$$

As condition $\mathcal{A}_1$ is satisfied, $\sum_{g_1=\gamma_n}^{n}(g_1+1)\alpha(g_1) = O(\gamma_n^{-4})$ provided $\gamma_n \to \infty$. Thus, from W_1 and W_{11}, one obtains that the previous equation is $O\{\pi_h^{-4}h^2(nh)^{-5} + \upsilon_n^{-3}(nh)^{-6}\} = O\{(nh)^{-5}h^6\}$. It follows that

$$
W_{hx}^{(1)}(\mathbf{L}) - \overline{W}_{hx}^{(1)}(\mathbf{L}) = O\left\{(nh)^{-5}(\nu_{xh} + h^6)\right\}. \tag{7}
$$

Hence, one deduces from equation (6) and (7) that there exist a constant $\kappa_3 > 0$ such that

$$
W_{hx}^{(1)}(\mathbf{L}) \le \omega_3 (nh)^{-5}\{(nh)^2\nu_{xh}^3 + (nh)\nu_{xh}^2 + (nh^7)\nu_x + h^6\}.
$$

For sufficiently large n, $nh^7 < 1$ since $nh^5 < \infty$ and $h \to 0$. Therefore the factor nh^7 in front of ν_x can be omitted. As the case $W_{hx}^{(2)}(\mathbf{T})$ is totally identical, one concludes that

$$
T_{xh}^{(3)}(\mathbf{L}) \le \overline{\omega}_3 (nh)^{-5}\{(nh)^2\nu_{xh}^3 + (nh)\nu_{xh}^2 + \nu_x + h^6\}
$$

for some constant $\overline{\omega}_3$. Similar but long computations yields to the same bound but a possibly different constant for the cases $j = 4, 5, 6$, i.e

$$
T_{xh}^{(j)}(\mathbf{L}) \le \overline{\omega}_j (nh)^{-5}\{(nh)^2\nu_{xh}^3 + (nh)\nu_{xh}^2 + \nu_x + h^6\} \qquad \overline{\omega}_j < \infty,\ j = 4, 5, 6.
$$

Collecting the bounds for each $T_{xh}^{(j)}(\mathbf{L})$ allows to conclude that there exist a global constant

$\omega < \infty$ such that

$$\mathrm{E}|\mathbb{H}_{xh}|^6 \leq \omega \sum_{k=1}^{3} \{\nu_x + h^2(\dot{\nu}_x + \ddot{\nu}_x)\}^k (nh)^{-3+k} + (nh)^{-2}h^6.$$

Since Assumption $\mathcal{A}_2$ holds, $\dot{\nu}_x(A)$ and $\ddot{\nu}_x(A)$ are uniformly bounded. Moreover, since $\nu_x(A) \leq \mu_x(A)$,

$$\mathrm{E}\left|\mathbb{H}_{xh}(A)\right|^6 \leq (\omega+1) \sum_{k=1}^{3} \left\{\mu_x(A) + h^2\right\}^k (nh)^{-3+k}.$$

5. Proof of Lemma 3.2

In order to ease readability, we simply write g for h_j. Moreover, as the cases $j = 1$ and $j = 2$ are identical, we drop the index j throughout the section. For any fixed $(t, u) \in \mathcal{I} = [-1, 1] \times [0, 1]$ the asymptotic normality of the random variable $\widetilde{Z}_{xn}(t, u)$ follows from similar arguments as in the proof of Proposition 2.1. This implies the asymptotic tightness of the random variable $\widetilde{Z}_{xn}(t, u)$ in $\mathbb{R}$. It remains to show the asymptotic tightness of the sequence $\widetilde{Z}_{xn}$ in $\ell^\infty(\mathcal{I})$. To this end, let $\rho(t, u, t', u') = |t - t'| + |u - u'|$ and for a bounded function $f : \mathcal{I} \to \mathbb{R}$ and a subset T of $\mathcal{I}$, define

$$\mathfrak{W}_\delta(f, T) = \sup_{\substack{(t,u),(t',u') \in T \\ \rho(t,u,t',u') < \delta}} |f(t, u) - f(t', u')|.$$

It will now be shonw that $\widetilde{Z}_{xn}$ is asymptotically ρ-equicontinuous i.e for any $\epsilon > 0$ there is a $\delta > 0$ such that

$$\lim_{n\to\infty} \mathbb{P}\left(\mathfrak{W}_\delta(\widetilde{Z}_{xn}, \mathcal{I}) > \epsilon\right) = 0.$$

For $\kappa_\gamma = \lfloor (ng)^{1/2+\gamma} \rfloor$, define grids $I_\gamma = \{0, \frac{1}{\kappa_\gamma}, \ldots, \frac{\kappa_\gamma - 1}{\kappa_\gamma}, 1\}$ and $J_\gamma = \{0, \pm\frac{1}{\kappa_\gamma}, \ldots, \pm\frac{\kappa_\gamma - 1}{\kappa_\gamma}, \pm 1\}$, where $\gamma \in (0, 1/2)$ is a grid parameter to be fixed later, and set $T_\gamma = J_\gamma \times I_\gamma$. For any $(t, u) \in [-1, 1] \times [0, 1]$, define $(\underline{t}_\gamma, \underline{u}_\gamma)$ and $(\overline{t}_\gamma, \overline{u}_\gamma)$ as in Section 3. Analogously to that section observe that

$$\begin{aligned}
&\widetilde{Z}_{xn}(t, u) - \widetilde{Z}_{xn}(\underline{t}_\gamma, \underline{u}_\gamma) \\
&\quad = \{\widetilde{Z}_{xn}(t, u) - \widetilde{Z}_{xn}(t, \underline{u}_\gamma)\} + \{\widetilde{Z}_{xn}(t, \underline{u}_\gamma) - \widetilde{Z}_{xn}(\underline{t}_\gamma, \underline{u}_\gamma)\} \\
&\quad \leq \{\widetilde{Z}_{xn}(t, \overline{u}_\gamma) - \widetilde{Z}_{xn}(t, \underline{u}_\gamma)\} + \{\widetilde{Z}_{xn}(t, \underline{u}_\gamma) - \widetilde{Z}_{xn}(\underline{t}_\gamma, \underline{u}_\gamma)\} \\
&\qquad\qquad + \mathrm{E}\left\{Z_{xn}(t, \overline{u}_\gamma) - Z_{xn}(t, \underline{u}_\gamma)\right\}. \qquad (8)
\end{aligned}$$

Starting with the last term of equation (8), using a Taylor expansion of F_{x_i} around z_t leads to:

$$\begin{aligned}
\mathrm{E}\left\{Z_{xn}(t,\overline{u}_\gamma) - Z_{xn}(t,\underline{u}_\gamma)\right\} &= \sqrt{ng}(\overline{u}_\gamma - \underline{u}_\gamma) \\
&+ \left[\dot{F}_{z_t}\{F_{z_t}^{-1}(\overline{u}_\gamma)\} - \dot{F}_{z_t}\{F_{z_t}^{-1}(\underline{u}_\gamma)\}\right]\sqrt{ng}\sum_{i=1}^{n} w_{ni}(g,z_t)(X_i - z_t) \\
&+ \frac{1}{2}\sqrt{ng}\sum_{i=1}^{n}\left[\ddot{F}_{r_{ti}}\{F_{z_t}^{-1}(\overline{u}_\gamma)\} - \ddot{F}_{r_{ti}}\{F_{z_t}^{-1}(\underline{u}_\gamma)\}\right] w_{ni}(g,z_t)(X_i - z_t)^2,
\end{aligned}$$

where r_{ti} lies between z_t and X_i. From assumptions W_9, W_{10} and $\mathcal{A}_4$, the previous equation is equal to

$$\sqrt{ng}(\overline{u}_\gamma - \underline{u}_\gamma) + o(1)O(\sqrt{ng}g^2) = o(1).$$

The last equality follows from the assumptions over the bandwidth parameters, ensuring that $\sqrt{ng}g^2 < \infty$, and the fact that the grid definition entails $\sqrt{ng}(\overline{u}_\gamma - \underline{u}_\gamma) = O\{(ng)^{-\gamma}\}$. This yields the negligibility of $\mathrm{E}\left\{Z_{xn}(t,\overline{u}_\gamma) - Z_{xn}(t,\underline{u}_\gamma)\right\}$.

Next we deal with the term $\widetilde{Z}_{xn}(t,\underline{u}_\gamma) - \widetilde{Z}_{xn}(\underline{t}_\gamma,\underline{u}_\gamma)$ in equation (8). Denote $\mathbb{F}_{zg} = \sqrt{ng}(F_{zg} - \mathrm{E}\,F_{zg})$ and notice that $\mathbb{F}_{z_tg}\{F_{z_t}^{-1}(u)\} = \widetilde{Z}_{xn}(t,u)$. Therefore, one writes

$$\begin{aligned}
\widetilde{Z}_{xn}(t,\underline{u}_\gamma) - \widetilde{Z}_{xn}(\underline{t}_\gamma,\underline{u}_\gamma) &= \left[\mathbb{F}_{z_tg}\left\{F_{z_t}^{-1}(\underline{u}_\gamma)\right\} - \mathbb{F}_{z_tg}\left\{F_{z_{\underline{t}_\gamma}}^{-1}(\underline{u}_\gamma)\right\}\right] \\
&\quad + \left[\mathbb{F}_{z_tg}\left\{F_{z_{\underline{t}_\gamma}}^{-1}(\underline{u}_\gamma)\right\} - \mathbb{F}_{jz_{\underline{t}_\gamma}g}\left\{F_{z_{\underline{t}_\gamma}}^{-1}(\underline{u}_\gamma)\right\}\right].
\end{aligned}$$

In view of assumption W_6 and the fact that $z_t - z_{\underline{t}_\gamma} = Ch(t - \underline{t}_\gamma)$, for any $y \in \mathbb{R}$:

$$\begin{aligned}
&\sqrt{ng}|F_{z_tg}(y) - F_{jz_{\underline{t}_\gamma}g}(y)| \\
&\quad = \sqrt{ng}|\sum_{i=1}^{n}\mathbb{I}(Z_i \le y)\left\{w_{ni}(g,z_t) - w_{ni}(g,z_{\underline{t}_\gamma})\right\}| \\
&\quad \le \sqrt{ng}\sup_{z\in I_x}\sum_{i=1}^{n}|w'_{ni}(z,g)| \times h(t - \underline{t}_\gamma) = O\left\{(ng)^{-\gamma}g^{-1}h\right\}.
\end{aligned}$$

Since $h/g < \infty$ the latter is o(1) uniformly in y. From similar arguments one deduces that $\sup_{y,t}|\mathbb{F}_{z_tg}(y) - \mathbb{F}_{z_{\underline{t}_\gamma}g}(y)| = o(1)$. It follows that

$$\begin{aligned}
\widetilde{Z}_{xn}(t,\underline{u}_\gamma) - \widetilde{Z}_{xn}(\underline{t}_\gamma,\underline{u}_\gamma) &= \mathbb{F}_{z_tg}\{F_{z_t}^{-1}(\underline{u}_\gamma)\} - \mathbb{F}_{z_tg}\{F_{z_{\underline{t}_\gamma}}^{-1}(\underline{u}_\gamma)\} + o(1) \\
&= \mathbb{F}_{z_{\underline{t}_\gamma}g}\{F_{z_t}^{-1}(\underline{u}_\gamma)\} - \mathbb{F}_{z_{\underline{t}_\gamma}g}\{F_{z_{\underline{t}_\gamma}}^{-1}(\underline{u}_\gamma)\} + o(1).
\end{aligned}$$

Using the same strategy with the first term of equation (8), one deduces that

$$\widetilde{Z}_{xn}(t,\overline{u}_\gamma) - \widetilde{Z}_{xn}(t,\underline{u}_\gamma) = \mathbb{F}_{z_{\underline{t}_\gamma}g}\{F_{z_t}^{-1}(\overline{u}_\gamma)\} - \mathbb{F}_{z_{\underline{t}_\gamma}g}\{F_{z_{\underline{t}_\gamma}}^{-1}(\underline{u}_\gamma)\} + o(1).$$

As assumption $\mathcal{A}_4$ implies that the function $z \mapsto F_z^{-1}$ is continuous in a neighborhood of x, one deduces that

$$\begin{aligned}
&\left|\mathbb{F}_{z_{\underline{t}_\gamma}} g\left\{F_{z_t}^{-1}(\overline{u}_\gamma)\right\} - \mathbb{F}_{z_{\underline{t}_\gamma}} g\left\{F_{z_{\underline{t}_\gamma}}^{-1}(\underline{u}_\gamma)\right\}\right| \\
&\quad \leq \left|\mathbb{F}_{z_{\underline{t}_\gamma}} g\left\{F_{z_{\overline{t}_\gamma}}^{-1}(\overline{u}_\gamma)\right\} - \mathbb{F}_{z_{\underline{t}_\gamma}} g\left\{F_{z_{\underline{t}_\gamma}}^{-1}(\overline{u}_\gamma)\right\}\right| \\
&\qquad + \left|\mathbb{F}_{z_{\underline{t}_\gamma}} g\left\{F_{z_{\underline{t}_\gamma}}^{-1}(\overline{u}_\gamma)\right\} - \mathbb{F}_{z_{\underline{t}_\gamma}} g\left\{F_{z_{\underline{t}_\gamma}}^{-1}(\underline{u}_\gamma)\right\}\right| + o(1).
\end{aligned}$$

In view of last discussion and of decomposition (8), one concludes that

$$\begin{aligned}
&\sup_{(t,u)\in\mathcal{I}} \left|\widetilde{Z}_{xn}(t,u) - \widetilde{Z}_{xn}(\underline{t}_\gamma, \underline{u}_\gamma)\right| \\
&\quad \leq 2 \sup_{(t,u)\in T_n} \left|\mathbb{F}_{z_t} g\left\{F_{z_t}^{-1}(\overline{u}_\gamma)\right\} - \mathbb{F}_{z_t} g\left\{F_{z_t}^{-1}(\underline{u}_\gamma)\right\}\right| \\
&\qquad + 2 \sup_{(t,u)\in T_n} \left|\mathbb{F}_{z_{\underline{t}_\gamma}} g\left\{F_{z_{\overline{t}_\gamma}}^{-1}(u)\right\} - \mathbb{F}_{z_{\underline{t}_\gamma}} g\left\{F_{z_{\underline{t}_\gamma}}^{-1}(u)\right\}\right| + o(1).
\end{aligned}$$

For $t_k = \frac{k}{\kappa_\gamma}$, denote

$$A_\gamma(i,t) = [F_{z_t}^{-1}(\frac{i-1}{\kappa_\gamma}), F_{jz_t}^{-1}(\frac{i}{\kappa_\gamma})] \text{ and } B_\gamma(i,k) = [F_{z_{t_k}}^{-1}(\frac{i}{\kappa_\gamma}), F_{z_{t_{k+1}}}^{-1}(\frac{i}{\kappa_\gamma})].$$

Moreover, write $\mathcal{G}_\gamma = \{0, 1, \ldots, \kappa_\gamma^{-1}\} \times \{0, \pm 1, \ldots, \pm\kappa_\gamma^{-1}\}$. Then from similar arguments as in the end of Section 3, for sufficiently large n:

$$\mathfrak{W}_\delta(\widetilde{Z}_{xn}, \mathcal{I}) \leq 6 \max_{(i,k)\in\mathcal{G}_\gamma} |\mathbb{F}_{z_{t_k}} g\{A_\gamma(i,t_k)\}| + 6 \max_{(i,k)\in\mathcal{G}_\gamma} |\mathbb{F}_{z_{t_k}} g\{B_\gamma(i,t_k)\}|.$$

For any interval $A = [a,b] \subset \mathbb{R}$, denote $\nu_z(A) = F_z(b) - F_z(a)$. On one hand, $\nu_{z_{t_k}}\{A_\gamma(i,t_k)\} = (ng)^{-1/2-\gamma}$. On the other hand,

$$\begin{aligned}
\nu_{z_{t_k}}\{B_\gamma(i,k)\} &= \left|u - F_{z_{t_k}}\left\{F_{z_{t_{k+1}}}^{-1}(u)\right\}\right| \\
&= \left|\dot{F}_{z_{t_{k+1}}}\left\{F_{z_{t_{k+1}}}^{-1}(u)\right\}(z_{t_k} - z_{t_{k+1}}) \right. \\
&\qquad \left. + \frac{1}{2}\ddot{F}_{z_{t^\star}}\left\{F_{z_{t_{k+1}}}^{-1}(u)\right\}(z_{t_k} - z_{t_{k+1}})^2\right|,
\end{aligned}$$

where $t^\star \in [t_k, t_{k+1}]$. Since $z_{t_k} - z_{t_{k+1}} = h(ng)^{1/2+\gamma}$, the $\nu_{z_{t_k}}$-measure of the set $B_\gamma(i,k)$ is smaller than the $\nu_{z_{t_k}}$-measure of the set $A_\gamma(i,t_k)$. One then argues that for n sufficiently large, for any $(i,k) \in \mathcal{G}_\gamma$, either $B_\gamma(i,k) \subset A_\gamma(i-1,t_k)$ or $B_\gamma(i,k) \subset A_\gamma(i,t_k)$. Thus for any $\epsilon > 0$:

$$\begin{aligned}
\mathbb{P}\left\{\mathfrak{W}_\delta\left(\widetilde{Z}_{xn}, \mathcal{I}\right) \geq \epsilon\right\} &\leq \mathbb{P}\left[\max_{(i,k)\in\mathcal{G}_\gamma} \left|\mathbb{F}_{z_{t_k}} g\left\{A_\gamma(i,t_k)\right\}\right| \geq \frac{\epsilon}{12}\right] \\
&\leq \sum_{(i,k)\in\mathcal{G}_\gamma} \mathbb{P}\left[\left|\mathbb{F}_{z_{t_k}} g\left\{A_\gamma(i,t_k)\right\}\right| \geq \frac{\epsilon}{12}\right] \\
&\leq \frac{(ng)^{1+2\gamma}}{\epsilon^6}\left(\max_{(i,k)\in\mathcal{G}_\gamma} \mathrm{E}\left[\mathbb{F}_{z_{t_k}} g\left\{A_\gamma(i,t_k)\right\}\right]^6\right),
\end{aligned}$$

where the last line follows from the use Markov inequality. As the assumptions of Lemma 2.4 are satisfied, identical computations as in section 4 with ν_x being replace with $\nu_{z_{t_k}}$ enables to find a constant $\omega < \infty$ such that for any interval $A \in [0,1]$:

$$\mathrm{E}\{\mathbb{F}_{z_{t_k}} g(A)^6\} \leq \omega \sum_{k=1}^{3} \{\nu_{z_{t_k}}(A) + g^2\}^k (ng)^{-3+k} + (ng)^{-2} g^6.$$

Since $\nu_{z_{t_k}}\{A_\gamma(i,t_k)\} = (ng)^{-1/2-\gamma}$ and $\sqrt{ng}g^2 < \infty$, it follow that for sufficiently large n, $g^2 > (ng)^{-1/2-\gamma}$. The previous equation is therefore bounded by $8\omega(n^{-2} + n^{-1}g^2 + g^6)$. In follows that

$$\begin{aligned} \mathbb{P}\left\{\mathfrak{W}_\delta(\widetilde{Z}_{xn}, \mathcal{I}) \geq \epsilon\right\} &\leq \frac{8\kappa}{\epsilon^6}(ng)^{1+2\gamma}(n^{-2} + n^{-1}g^2 + g^6) \\ &= \frac{8\kappa}{\epsilon^6}(n^{-1+2\gamma}g^{1+2\gamma} + n^{2\gamma}g^{3+2\gamma} + n^{1+2\gamma}g^{7+2\gamma}) \end{aligned}$$

Since $\sqrt{ng}g^2 < \infty$ implies $g = O(n^{-\tau})$ with $\tau \geq 1/5$, it follows that the latter is $o(1)$ upon taking $\gamma \in (0, r_\tau)$ with $r_\tau = \min\{1/2, \frac{3\tau}{2(1-\tau)}, \frac{7\tau-1}{2(1-\tau)}\}$. The lemma is therefore proven.

References

Billingsley, P. (1968), *Convergence of probability measures*, John Wiley & Sons, Inc., New York-London-Sydney.