Inference on local causality and tests of non-causality in time series

Taoufik Bouezmarni

Université de Sherbrooke
Centre interuniversitaire de recherche en economie quantitative (CIREQ)
e-mail: taoufik.bouezmarni@usherbrooke.ca

Félix Camirand Lemyre

Université de Sherbrooke
Centre de Recherche du Centre Hospitalier Universitaire de Sherbrooke
e-mail: felix.camirand.lemyre@usherbrooke.ca

Jean-François Quessy

Université du Québec à Trois-Rivières
e-mail: jean-francois.quessy@uqtr.ca

Abstract: The study of the causal relationships in a stochastic process $(Y_t, Z_t)_{t \in \mathbb{Z}}$ is a subject of a particular interest in finance and economy. A widely-used approach is to consider the notion of Granger causality, which in the case of first order Markovian processes is based on the joint distribution function of (Y_{t+1}, Z_t) given Y_t. The measures of Granger causality proposed so far are global in the sense that if the relationship between Y_{t+1} and Z_t changes with the value taken by Y_t, this may not be captured. To circumvent this limitation, this paper proposes *local* causality measures based on the conditional copula of (Y_{t+1}, Z_t) given $Y_t = x$. Exploiting results by [5] on the asymptotic behavior of two kernel-based conditional copula estimators under α-mixing, the asymptotic normality of nonparametric estimators of these local measures is deduced and asymptotically valid confidence intervals are built; tests of local non-causality are also developed. The suitability of the proposed methods is investigated with simulations and their usefulness is illustrated on the time series of Standard & Poor's 500 prices and trading volumes.

MSC 2010 subject classifications: Primary 62G99; secondary 62M99.
Keywords and phrases: α-mixing processes, bandwidth selection, conditional copula, Kendall and Spearman dependence measures, local linear kernel estimation, weak convergence.

Received May 2018.

1. Introduction

The concept of causality as originally introduced by [27] and [15] is helpful for studying the dynamic relationships in multivariate time series. This notion is defined in terms of predictability at horizon one of a random variable (or random vector) Y from its past and the past of another random variable (or vector) Z.

Specifically, assume that data are available for the process $(Y_t, Z_t)_{t \in \mathbb{Z}}$, and let $\mathbf{Y}_t$, $\mathbf{Z}_t$ be realizations, respectively, of Y and Z up to time t. According to [15], the causality from Z to Y *one period ahead* is defined as follows: Z is said to cause Y if $\mathbf{Z}_t$ can help predict Y_{t+1}, conditional on $\mathbf{Y}_t$.

Many works considered testing the null hypothesis of non-causality. For example, [4] investigated tests for multivariate ARMA models. Because the Granger non-causality is a form of conditional independence, tests can be deduced from standard conditional independence tests; see [9], for instance. Tests based on copulas have been proposed by [6]. When the hypothesis of non-causality is rejected, one may be interested in measuring the strength of this causal relationship. The first causality measures were proposed by [11] and [12] using the mean-squared forecast errors, and by [14] based on the Kullback–Leibler information. Causality measures for short and long run under parametric models were investigated by [23] and [7]. Mainly inspired by the fact that these measures suffer from model misspecification, nonparametric mesures where proposed by [24] and [16] using the Kullback–Leibler information and nonparametric density copula estimators. Recently, [28] investigated causality measures at multiple horizons for exchange rate and commodity prices.

It is worth mentioning that all of the above-cited papers focused on characterizing the global relationship between Y_{t+1} and $\mathbf{Z}_t$, conditional on $\mathbf{Y}_t$. However, if the nature of the link between Y_{t+1} and $\mathbf{Z}_t$ changes with the value taken by $\mathbf{Y}_t$, it may not be captured by global measures. A possible solution to this issue is to compute the partial correlation coefficient. However, such an approach implicitly assumes linear relationships and the measure depends on the marginal behavior. In other words, this strategy would suffer from the same well-documented drawbacks of the classical Pearson correlation coefficient.

To circumvent these limitations, this paper proposes marginal-free local causality indices for measuring the strength of the relationship in the pair $(Y_{t+1}, \mathbf{Z}_t)$ given a particular value taken by $\mathbf{Y}_t$. In order to simplify the presentation, a focus is put on Markovian models of order one where one considers the dependence structure of (Y_{t+1}, Z_t) given $Y_t = x \in \mathbb{R}$ as captured by its associated bivariate conditional copula. This approach allows for the definition of nonparametric measures of local causality that do not suffer from the drawbacks that arise when using partial correlations. Specifically, let $\{(Y_t, Z_t)\}_{t \in \mathbb{Z}}$ be a stationary process and define the local causality distribution function

$$H_x^{\mathrm{Z} \to \mathrm{Y}}(y, z) = \mathbb{P}\left(Y_{t+1} \leq y, Z_t \leq z | Y_t = x\right). \tag{1}$$

Assuming that the conditional marginal distributions $F_{1x}(y) = \mathbb{P}(Y_{t+1} \leq y | Y_t = x)$ and $F_{2x}(z) = \mathbb{P}(Z_t \leq z | Y_t = x)$ are continuous, Sklar's Theorem guarantees the existence of a unique copula $C_x^{\mathrm{Z} \to \mathrm{Y}} : [0, 1]^2 \to [0, 1]$ such that

$$H_x^{\mathrm{Z} \to \mathrm{Y}}(y, z) = C_x^{\mathrm{Z} \to \mathrm{Y}}\left\{F_{1x}(y), F_{2x}(z)\right\}.$$

The bivariate function $C_x^{\mathrm{Z} \to \mathrm{Y}}$ will be called the *local causality copula* and corresponds to the dependence structure of (Y_{t+1}, Z_t) given $Y_t = x$.

The first goal of this paper is to describe nonparametric estimators of $C_x^{Z \to Y}$ and to obtain their large-sample behavior under general α-mixing processes in the light of results by [5]. Then, local causality indices are defined for measuring the strength of the causal relationship in a bivariate time series and nonparametric estimators based on the empirical conditional copulas are proposed. The asymptotic normality of the latter is established using the functional delta method. These large-sample results are then exploited to develop inferential methods for local causality measures, including asymptotically valid confidence intervals and tests for the null hypothesis of local non-causality.

The article is organized as follows. In Section 2, two estimators of the local causality copula are described and their asymptotic behavior is obtained as an adaptation of results obtained by [5]. In Section 3, general local causality indices are defined and the large-sample behavior of nonparametric estimators is investigated. Section 4 addresses the issue of the estimation of the asymptotic variances and biases, as well as the optimal selection of the bandwidths. Section 5 develops interval estimates of the local causality measures, as well as tests for the null hypothesis of local non-causality. Section 6 investigates the sampling properties of the newly introduced methods when they are based on the Spearman and Kendall measures of association; their usefulness is illustrated in Section 7 on financial data. The assumptions under which the asymptotic results hold as well as the proofs are relegated to the appendix.

2. Estimation of the local causality copula

2.1. Two estimators

Consider a stationary process $\{(Y_t, Z_t)\}_{t \in \mathbb{Z}}$ that is α-mixing, *i.e.* its α-mixing coefficients $\alpha(r)$ are such that $\alpha(r) \to 0$ as $r \to \infty$. Recall that the latter are defined by $\alpha(0) = 1/2$ and for each lag $r \in \mathbb{N}$,

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

where $\mathcal{F}_a^b$ is the σ-field generated by $\{(Y_t, Z_t)\}_{a \le t \le b}$. In the sequel, one assumes a realization $(Y_1, Z_1), \ldots, (Y_{n+1}, Z_{n+1})$ of $\{(Y_t, Z_t)\}_{t \in \mathbb{Z}}$.

A first estimator of the local causality copula arises upon noting that $C_x^{Z \to Y}$ can be extracted from $H_x^{Z \to Y}$ via

$$C_x^{Z \to Y}(u, v) = H_x^{Z \to Y} \left\{ F_{1x}^{-1}(u), F_{2x}^{-1}(v) \right\}.$$

An estimator of the joint conditional distribution $H_x^{Z \to Y}$ defined in (1) will then provide a plug-in estimator of $C_x^{Z \to Y}$. In this paper, one considers estimators based on the local linear weights defined for each $i \in \{1, \ldots, n\}$ by

$$\omega_i(x, h) = \frac{1}{nh}\, K\left(\frac{Y_i - x}{h}\right) \frac{S_{n,2}(x) - \left(\frac{Y_i - x}{h}\right) S_{n,1}(x)}{S_{n,0}(x)\, S_{n,2}(x) - S_{n,1}(x)^2}, \tag{2}$$

where $h = h_n > 0$ is a bandwidth parameter that tends to zero as the sample size $n \to \infty$, K is a symmetric density and for $\ell \in \{0, 1, 2\}$,

$$S_{n,\ell}(x) = \frac{1}{nh} \sum_{i=1}^{n} \left(\frac{Y_i - x}{h} \right)^{\ell} K \left(\frac{Y_i - x}{h} \right).$$

The local linear estimator of $H_x^{Z \to Y}$ is then defined by

$$H_{xh}^{Z \to Y}(y, z) = \sum_{i=1}^{n} \omega_i(x, h) \, \mathbb{I}\left(Y_{i+1} \leq y, Z_i \leq z\right). \tag{3}$$

The reader is referred to [8] for details on how to derive this kind of estimators. A natural plug-in estimator of $C_x^{Z \to Y}$ is then given by

$$C_{xh}^{Z \to Y}(u, v) = H_{xh}^{Z \to Y}\left\{F_{1xh}^{-1}(u), F_{2xh}^{-1}(v)\right\}, \tag{4}$$

where F_{1xh}^{-1} and F_{2xh}^{-1} are the left-continuous generalized inverses of

$$F_{1xh}(y) = \sum_{i=1}^{n} \omega_i(x, h) \, \mathbb{I}\left(Y_{i+1} \leq y\right) \text{ and } F_{2xh}(z) = \sum_{i=1}^{n} \omega_i(x, h) \, \mathbb{I}\left(Z_i \leq z\right). \tag{5}$$

As noted by [26] and [13] in the i.i.d. case, the plug-in estimator $C_{xh}^{Z \to Y}$ may be severely biased, especially when the conditional marginal distributions strongly depend on the covariate. For that reason, a second estimator that aims at removing this possible effect of the covariate on the margins is proposed. To this end, let $h_1, h_2 > 0$ be bandwidth parameters that may differ from h, and define for each $i \in \{1, \ldots, n\}$ the pair of pseudo-observations

$$(\widehat{U}_{ih_1}, \widehat{V}_{ih_2}) = (F_{1Y_ih_1}(Y_{i+1}), F_{2Y_ih_2}(Z_i)), \tag{6}$$

where $F_{1Y_ih_1}$ and $F_{2Y_ih_2}$ follow from the definition in (5). Their associated empirical conditional joint distribution is then

$$G_{xh}^{Z \to Y}(y, z) = \sum_{i=1}^{n} \omega_i(x, h) \, \mathbb{I}\left(\widehat{U}_{ih_1} \leq y, \widehat{V}_{ih_2} \leq z\right).$$

An alternative estimator of $C_x^{Z \to Y}$ based on $G_{xh}^{Z \to Y}$ is

$$\widetilde{C}_{xh}^{Z \to Y}(u, v) = G_{xh}^{Z \to Y}\left\{G_{1xh}^{-1}(u), G_{2xh}^{-1}(v)\right\}, \tag{7}$$

where G_{1xh}^{-1} and G_{2xh}^{-1} are the left-continuous generalized inverses of

$$G_{1xh}(y) = \sum_{i=1}^{n} \omega_i(x, h) \, \mathbb{I}\left(\widehat{U}_{ih_1} \leq y\right) \text{ and } G_{2xh}(z) = \sum_{i=1}^{n} \omega_i(x, h) \, \mathbb{I}\left(\widehat{V}_{ih_2} \leq z\right).$$

Remark 2.1. *Although it will not be further treated in this paper, copula-based local causality could be extended beyond horizon one. In other words, one could consider the local causality copula that arises from the conditional distribution of (Y_{t+1}, Z_t) given $(Y_{t-k+1}, \ldots, Y_t) = \mathbf{x} \in \mathbb{R}^k$. The only difference would consists in using multivariate local linear weights $\omega_1(\mathbf{x}, h_1, \ldots, h_k), \ldots, \omega_n(\mathbf{x}, h_1, \ldots, h_k)$ instead of $\omega_1(x, h), \ldots, \omega_n(x, h)$. See [19], for instance, for more details on multivariate local linear weights.*

2.2. Weak convergence

In their paper, [5] obtained the weak convergence of two local linear estimators of conditional copulas from a realization of a general three-dimensional α-mixing process $(\epsilon_{1t}, \epsilon_{2t}, \epsilon_{3t})_{t \in \mathbb{Z}}$, where ϵ_{3t} is the conditioning variable, *i.e.* the covariate. The setup here is therefore a special case of [5] where $\epsilon_{1t} = Y_{t+1}$, $\epsilon_{2t} = Z_t$ and $\epsilon_{3t} = Y_t$. Before describing the results, let f_Y be the density associated to the stationary process $\{Y_t\}_{t \in \mathbb{Z}}$ and define

$$a_1 = \frac{1}{2} \int_{\mathbb{R}} z^2 K(z) \, \mathrm{d}z \quad \text{and} \quad a_2 = \int_{\mathbb{R}} \{K(z)\}^2 \, \mathrm{d}z.$$

Moreover, let $\alpha_x^{Z \to Y}$ be a Gaussian process on $[0,1]^2$ such that for κ defined in Assumption $(\mathcal{N})$ and $\ddot{H}_x^{Z \to Y} = \partial^2 H_x^{Z \to Y} / \partial^2 x$,

$$\mathrm{E}\left\{\alpha_x^{Z \to Y}(u,v)\right\} = \kappa \, a_1 \, \ddot{H}_x^{Z \to Y} \left\{F_{1x}^{-1}(u), F_{2x}^{-1}(v)\right\}.$$

Also, defining $\sigma_C^2(u, v, u', v') = C\left\{\min(u, u'), \min(v, v')\right\} - C(u, v) \, C(u', v')$ for an arbitrary copula C, the covariance function of α_x is

$$\mathrm{Cov}\left\{\alpha_x^{Z \to Y}(u,v), \alpha_x^{Z \to Y}(u', v')\right\} = \frac{a_2}{f_Y(x)} \, \sigma_{C_x^{Z \to Y}}^2(u, v, u', v').$$

As a special case of Proposition 2 in [5], and for a fixed value of $x \in \mathbb{R}$, one can then conclude that under Assumptions $(\mathcal{S})$, $(\mathcal{C})$, $(\mathcal{L})$, $(\mathcal{H})$ and $(\mathcal{N})$ described in Appendix A.1, the process $\mathbb{C}_{xh}^{Z \to Y} = \sqrt{nh} \left(C_{xh}^{Z \to Y} - C_x^{Z \to Y}\right)$ converges weakly in the space $\ell^\infty([0,1]^2)$ of bounded functions on $[0,1]^2$ to

$$\mathbb{C}_x^{Z \to Y}(u,v) = \alpha_x^{Z \to Y}(u,v) - C_x^{[1]}(u,v) \, \alpha_x^{Z \to Y}(u,1) - C_x^{[2]}(u,v) \, \alpha_x^{Z \to Y}(1,v),$$

where $C_x^{[1]}(u,v) = \partial C_x^{Z \to Y}(u,v)/\partial u$ and $C_x^{[2]}(u,v) = \partial C_x^{Z \to Y}(u,v)/\partial v$.

Before stating the result on the weak convergence of $\widetilde{\mathbb{C}}_{xh}^{Z \to Y} = \sqrt{nh}(\widetilde{C}_{xh}^{Z \to Y} - C_x^{Z \to Y})$, introduce the Gaussian process $\mathbb{G}_x^{Z \to Y}$ on $[0,1]^2$ such that for $\ddot{C}_x^{Z \to Y} = \partial^2 C_x^{Z \to Y}/\partial^2 x$,

$$\mathrm{E}\left\{\mathbb{G}_x^{Z \to Y}(u,v)\right\} = \kappa \, a_1 \, \ddot{C}_x^{Z \to Y}(u,v)$$

and whose covariance function is the same as that of α_x. Then one can invoke Proposition 3 of [5] and conclude that as long as Assumptions $(\mathcal{S})$, $(\mathcal{C})$, $(\mathcal{L})$, $(\mathcal{H}^\star)$ and $(\mathcal{N}^\star)$ in Appendix A.2 hold, and for a fixed value of $x \in \mathbb{R}$, $\widetilde{\mathbb{C}}_{xh}$ converges weakly in $\ell^\infty([0,1]^2)$ to

$$\widetilde{\mathbb{C}}_x^{Z \to Y}(u,v) = \mathbb{G}_x^{Z \to Y}(u,v) - C_x^{[1]}(u,v) \, \mathbb{G}_x^{Z \to Y}(u,1) - C_x^{[2]}(u,v) \, \mathbb{G}_x^{Z \to Y}(1,v).$$

As noted by [5] in the general case, the asymptotic behavior of the limit processes $\mathbb{C}_{xh}^{Z \to Y}$ and $\widetilde{\mathbb{C}}_{xh}^{Z \to Y}$ under α-mixing is the same as that under serial independence. In other words, their respective limits are the same as those identified by [26]. Thus, the impact of time-dependency is asymptotically negligible, somewhat as a consequence of Assumption $(\mathcal{S})$ on the α-mixing coefficients combined with

the use of a kernel function that smooths the covariate space in a shrinking neighborhood of x as n goes to infinity. Observe that when $C_{xh}^{Z \to Y}$ and $\widetilde{C}_{xh}^{Z \to Y}$ are computed with the same bandwidth h, the limit processes $\mathbb{C}_x^{Z \to Y}$ and $\widetilde{\mathbb{C}}_x^{Z \to Y}$ have the same covariance structure given implicitly by

$$\gamma_x(u, v, u', v') = \mathrm{Cov}\left\{\mathbb{C}_x^{Z \to Y}(u, v), \mathbb{C}_x^{Z \to Y}(u', v')\right\}.$$

Their bias functions are however different, since $\mathrm{E}\{\mathbb{C}_x(u, v)\} = \kappa\, a_1\, \mathbb{B}_x(u, v)$ and $\mathrm{E}\{\widetilde{\mathbb{C}}_x(u, v)\} = \kappa\, a_1\, \widetilde{\mathbb{B}}_x(u, v)$, where $\widetilde{\mathbb{B}}_x(u, v) = \ddot{C}_x^{Z \to Y}(u, v)$ and

$$\begin{aligned}
\mathbb{B}_x(u, v) &= \ddot{H}_x^{Z \to Y}\left\{F_{1x}^{-1}(u), F_{2x}^{-1}(v)\right\} \\
&\quad - \ddot{F}_{1x}^{Z \to Y}\left\{F_{1x}^{-1}(u)\right\} C_x^{[1]}(u, v) \\
&\quad - \ddot{F}_{2x}^{Z \to Y}\left\{F_{2x}^{-1}(v)\right\} C_x^{[2]}(u, v).
\end{aligned}$$

Hence, unlike for $C_{xh}^{Z \to Y}$, the asymptotic bias of $\widetilde{C}_{xh}^{Z \to Y}$ does not include the partial derivatives of F_{1x} and F_{2x} with respect to x.

3. Measuring local causality

3.1. Theoretical measures of local causality

Measuring the strength of the causal relationship from Z to Y in a stationary process $\{(Y_t, Z_t)\}_{t \in \mathbb{Z}}$ can be done from the information provided by the local causality copula. Specifically, consider a general functional $\Lambda : \ell^\infty([0, 1]^2) \to \mathbb{R}$ satisfying $\Lambda(\Pi) = 0$, $\Lambda(M) = 1$ and $\Lambda(W) = -1$, where $\Pi(u, v) = uv$, $M(u, v) = \min(u, v)$ and $W(u, v) = \max(u + v - 1, 0)$ are respectively the independence, perfect positive dependence and perfect negative dependence copulas. A measure of local causality from Z to Y at x based on Λ is then

$$\theta_{\Lambda,x}^{Z \to Y} = \Lambda\left(C_x^{Z \to Y}\right). \tag{8}$$

This measure has the desirable property of being marginal-free.

3.2. Nonparametric estimators

The estimation of the local causality index $\theta_{\Lambda,x}^{Z \to Y}$ defined in Equation (8) can be based on the empirical local causality copulas $C_{xh}^{Z \to Y}$ and $\widetilde{C}_{xh}^{Z \to Y}$. Specifically, two estimators of $\theta_{\Lambda,x}^{Z \to Y}$ are given by

$$\theta_{\Lambda,xh}^{Z \to Y} = \Lambda\left(C_{xh}^{Z \to Y}\right) \quad \text{and} \quad \widetilde{\theta}_{\Lambda,xh}^{Z \to Y} = \Lambda\left(\widetilde{C}_{xh}^{Z \to Y}\right).$$

The next result establishes the asymptotic normality of

$$\Theta_{\Lambda,xh}^{Z \to Y} = \sqrt{nh}\left(\theta_{\Lambda,xh}^{Z \to Y} - \theta_{\Lambda,x}^{Z \to Y}\right) \quad \text{and} \quad \widetilde{\Theta}_{\Lambda,xh}^{Z \to Y} = \sqrt{nh}\left(\widetilde{\theta}_{\Lambda,xh}^{Z \to Y} - \theta_{\Lambda,x}^{Z \to Y}\right).$$

Before stating it, note that in the sequel, the composition $\Lambda \circ \Lambda$ of a functional $\Lambda : \ell^\infty([0,1]^2) \to \mathbb{R}$ is taken such that for $\gamma \in \ell^\infty([0,1]^4)$,

$$\Lambda \circ \Lambda(\gamma) = \Lambda\left[\Lambda\left\{\gamma(\cdot, \cdot, u', v')\right\}\right],$$

i.e. the operator inside the brackets is computed with respect to the first two arguments of γ, and then Λ is computed again on $\Lambda\left\{\gamma(\cdot, \cdot, u', v')\right\} \in \ell^\infty([0,1]^2)$.

Proposition 3.1. *Suppose that the functional* $\Lambda : \ell^\infty([0,1]^2) \to \mathbb{R}$ *is Hadamard differentiable with derivative at* g *given by* Λ'_g.

(i) Under Assumptions (S), (C), (L), (H) and (N), $\Theta_{\Lambda,xh}^{Z \to Y}$ *converges in law to the Normal random variable* $\Theta_{\Lambda,x}^{Z \to Y} = \Lambda'_{C_x^{Z \to Y}}(\mathbb{C}_x^{Z \to Y})$ *having mean* $\mu_{\Lambda,x} = \kappa\, a_1\, \Lambda'_{C_x^{Z \to Y}}(\mathbb{B}_x)$ *and variance* $\sigma_{\Lambda,x}^2 = \Lambda'_{C_x^{Z \to Y}} \circ \Lambda'_{C_x^{Z \to Y}}(\gamma_x)$.

(ii) Under Assumptions (S), (C), (L), (H*) and (N*), $\widetilde{\Theta}_{\Lambda,xh}^{Z \to Y}$ *converges in law to the Normal random variable* $\widetilde{\Theta}_{\Lambda,x}^{Z \to Y} = \Lambda'_{C_x^{Z \to Y}}(\widetilde{\mathbb{C}}_x^{Z \to Y})$ *having mean* $\widetilde{\mu}_{\Lambda,x} = \kappa\, a_1\, \Lambda'_{C_x^{Z \to Y}}(\widetilde{\mathbb{B}}_x)$ *and variance* $\sigma_{\Lambda,x}^2 = \Lambda'_{C_x^{Z \to Y}} \circ \Lambda'_{C_x^{Z \to Y}}(\gamma_x)$.

3.3. The Kendall and Spearman measures of causality

It is well known, see [22] and [17] for details, that the popular Spearman's rho and Kendall's tau associated to a continuous random pair with unique copula C can be expressed as

$$\rho_S = 12 \int_{[0,1]^2} C(u,v)\, \mathrm{d}u\mathrm{d}v - 3 \quad \text{and} \quad \tau = 4 \int_{[0,1]^2} C(u,v)\, \mathrm{d}C(u,v) - 1.$$

Thus, these measures are marginal-free and can be expressed as functionals of C via $\rho_S = \Lambda_\rho(C)$ and $\tau = \Lambda_\tau(C)$, where for $\delta \in \ell^\infty([0,1]^2)$,

$$\Lambda_\rho(\delta) = 12 \int_{[0,1]^2} \delta(u,v)\, \mathrm{d}u\mathrm{d}v - 3 \quad \text{and} \quad \Lambda_\tau(\delta) = 4 \int_{[0,1]^2} \delta(u,v)\, \mathrm{d}\delta(u,v) - 1.$$

Local causality indices based on the Spearman and Kendall measures are then $\rho_x^{Z \to Y} = \Lambda_\rho(C_x^{Z \to Y})$ and $\tau_x^{Z \to Y} = \Lambda_\tau(C_x^{Z \to Y})$, respectively. Two possible estimators of $\rho_x^{Z \to Y}$ are then $\rho_{xh}^{Z \to Y} = \Lambda_\rho(C_{xh}^{Z \to Y})$ and $\widetilde{\rho}_{xh}^{Z \to Y} = \Lambda_\rho(\widetilde{C}_{xh}^{Z \to Y})$. By straightforward computations, one can derive the explicit expressions

$$\rho_{xh}^{Z \to Y} = 12 \sum_{i=1}^n \omega_i(x,h)\left\{1 - F_{1xh}(Y_{i+1})\right\}\left\{1 - F_{2xh}(Z_i)\right\} - 3,$$

$$\widetilde{\rho}_{xh}^{Z \to Y} = 12 \sum_{i=1}^n \omega_i(x,h)\left\{1 - G_{1xh_1}(\widehat{U}_{ih_1})\right\}\left\{1 - G_{2xh_2}(\widehat{V}_{ih_2})\right\} - 3.$$

Two estimators of $\tau_x^{Z \to Y}$ are given by $\tau_{xh}^{Z \to Y} = \Lambda_\tau(C_{xh}^{Z \to Y})$ and $\widetilde{\tau}_{xh}^{Z \to Y} = \Lambda_\tau(\widetilde{C}_{xh}^{Z \to Y})$, where it can be shown that

$$
\begin{aligned}
\tau_{xh}^{Z \to Y} &= 4 \sum_{i,j=1}^{n} \omega_i(x,h)\,\omega_j(x,h)\,\mathbb{I}\,(Y_{j+1} \leq Y_{i+1}, Z_j \leq Z_i) - 1, \\
\widetilde{\tau}_{xh}^{Z \to Y} &= 4 \sum_{i,j=1}^{n} \omega_i(x,h)\,\omega_j(x,h)\,\mathbb{I}\left(\widehat{U}_{jh_1} \leq \widehat{U}_{ih_1}, \widehat{V}_{jh_2} \leq \widehat{V}_{ih_2}\right) - 1.
\end{aligned}
$$

The conclusions of Proposition 3.1 apply to the Spearman and Kendall functionals Λ_ρ and Λ_τ since both are Hadamard differentiable with respective derivatives

$$
\left(\Lambda_\rho\right)'_g (\delta) = 12 \int_{[0,1]^2} \delta(u,v)\,\mathrm{d}u\,\mathrm{d}v \tag{9}
$$

and, from [26],

$$
\left(\Lambda_\tau\right)'_g (\delta) = 4 \int_{[0,1]^2} \left\{ \delta(u,v)\,\mathrm{d}g(u,v) + g(u,v)\,\mathrm{d}\delta(u,v) \right\}. \tag{10}
$$

4. Estimation of asymptotic biases/variances and bandwidth selection

In order to develop valid inference methods for a given local causality measure $\theta_{\Lambda,x}^{Z \to Y}$, the results stated in Proposition 3.1 cannot be used directly because the asymptotic variance $\sigma_{\Lambda,x}^2$ as well as the asymptotic biases $\mu_{\Lambda,x}$ and $\widetilde{\mu}_{\Lambda,x}$ must be estimated. Another aspect of importance is the choice of an appropriate bandwidth parameter with respect to some optimality criterion. These topics are treated in this section.

4.1. Estimation of the asymptotic variances

First observe that the process $\mathbb{C}_x^{Z \to Y}$ can be seen as the weak limit of

$$
\sqrt{nh} \sum_{i=1}^{n} \omega_i(x,h)\, L_{xi},
$$

where $\omega_i(x,h)$ is defined in (2) and for each $i \in \{1, \ldots, n\}$,

$$
\begin{aligned}
L_{xi}(u,v) &= \mathbb{I}\left\{ Y_{i+1} \leq F_{1x}^{-1}(u), Z_i \leq F_{2x}^{-1}(v) \right\} - C_x^{Z \to Y}(u,v) \\
&\quad - C_x^{[1]}(u,v)\left[\mathbb{I}\left\{ Y_{i+1} \leq F_{1x}^{-1}(u) \right\} - u \right] \\
&\quad - C_x^{[2]}(u,v)\left[\mathbb{I}\left\{ Z_i \leq F_{2x}^{-1}(v) - v \right\} \right].
\end{aligned}
$$

As a consequence, the random variable $\Theta_{\Lambda,x}^{Z \to Y} = \Lambda'_{C_x^{Z \to Y}}(\mathbb{C}_x^{Z \to Y})$ in Proposition 3.1 (i) arises as the limit in law of

$$
\lambda_{xh} = \Lambda'_{C_x^{Z \to Y}} \left\{ \sqrt{nh} \sum_{i=1}^{n} \omega_i(x,h)\, L_{xi} \right\} = \sqrt{nh} \sum_{i=1}^{n} \omega_i(x,h)\, \lambda_{xi},
$$

where $\lambda_{xi} = \Lambda'_{C_x^{Z \to Y}}(L_{xi})$; the last equality is a consequence of the linearity of Hadamard derivatives (see [25]). However, since the marginal conditional distributions F_{1x}, F_{2x} and the partial derivatives $C_x^{[1]}$, $C_x^{[2]}$ are unknown, one considers instead the version of L_{xi} given by

$$
\begin{aligned}
L_{xi}^{\star}(u, v) \;=\; & \mathbb{I}\left\{ Y_{i+1} \leq F_{1xh}^{-1}(u), Z_i \leq F_{2xh}^{-1}(v) \right\} \\
& - \widehat{C}_x^{[1]}(u, v) \, \mathbb{I}\left\{ Y_{i+1} \leq F_{1xh}^{-1}(u) \right\} \\
& - \widehat{C}_x^{[2]}(u, v) \, \mathbb{I}\left\{ Z_i \leq F_{2xh}^{-1}(v) \right\}.
\end{aligned}
$$

In the above expression, $\widehat{C}_x^{[1]}$ and $\widehat{C}_x^{[2]}$ are uniformly consistent estimators of the partial derivatives of $C_x^{Z \to Y}$ in the sense that for any $\varepsilon > 0$,

$$
\sup_{\substack{u \in [\varepsilon, 1-\varepsilon] \\ v \in [0,1]}} \left| \widehat{C}_x^{[1]}(u, v) - C_x^{[1]}(u, v) \right| \quad \text{and} \quad \sup_{\substack{v \in [\varepsilon, 1-\varepsilon] \\ u \in [0,1]}} \left| \widehat{C}_x^{[2]}(u, v) - C_x^{[2]}(u, v) \right| \quad (11)
$$

converge in probability to zero. A version of λ_{xh} is therefore given by

$$
\lambda_{xh}^{\star} = \sqrt{nh} \sum_{i=1}^{n} \omega_i(x, h) \, \lambda_{xi}^{\star},
$$

with $\lambda_{xi}^{\star} = \Lambda'_{C_x^{Z \to Y}}(L_{xi}^{\star})$. Based on the idea of [8] for the estimation of a conditional variance, the proposed estimator of $\sigma_{\Lambda, x}^2$ is given by

$$
\widehat{\sigma}_{\Lambda, x}^2 = \frac{a_2}{S_{n,0}(x)} \sum_{i=1}^{n} \omega_i(x, h) \left(\lambda_{xi}^{\star} - \frac{\lambda_{xh}^{\star}}{\sqrt{nh}} \right)^2.
$$

The consistency of $\widehat{\sigma}_{\Lambda, x}^2$ for the estimation of $\sigma_{\Lambda, x}^2$ is stated next.

Proposition 4.1. *Assume that Λ is Hadamard differentiable with derivative at g given by Λ'_g and that Assumptions $(\mathcal{S})$, $(\mathcal{C})$, $(\mathcal{L})$, $(\mathcal{H})$ and $(\mathcal{N})$ hold. Moreover, suppose that $\widehat{C}_x^{[1]}$ and $\widehat{C}_x^{[2]}$ satisfy (11) and that there exists a constant $D > 0$ such that as $n \to \infty$,*

$$
\mathbb{P}\left(\max_{j=1,2} \sup_{u,v \in [0,1]} \left| \widehat{C}_x^{[j]}(u, v) \right| > D \right) \to 0. \tag{12}
$$

Then $\widehat{\sigma}_{\Lambda, x}^2$ converges in probability to $\sigma_{\Lambda, x}^2$.

An alternative estimator of $\sigma_{\Lambda, x}^2$ can be based on the approximate version of the process $\widetilde{\mathbb{C}}_x^{Z \to Y}$ based on the pseudo-observations. Specifically, let

$$
\begin{aligned}
\widetilde{L}_{xi}^{\star}(u, v) \;=\; & \mathbb{I}\left\{ \widehat{U}_{ih_1} \leq \widetilde{G}_{1xh}^{-1}(u), \widehat{V}_{ih_2} \leq \widetilde{G}_{2xh}^{-1}(v) \right\} \\
& - \widehat{C}_x^{[1]}(u, v) \, \mathbb{I}\left\{ \widehat{U}_{ih_1} \leq \widetilde{G}_{1xh}^{-1}(u) \right\}
\end{aligned}
$$

$$-\widehat{C}_x^{[2]}(u,v)\,\mathbb{I}\left\{\widehat{V}_{ih_2}\leq\widetilde{G}_{2xh}^{-1}(v)\right\}$$

and define $\widetilde{\lambda}_{xi}^{\star}=\Lambda'_{C_x^{\mathrm{Z}\to\mathrm{Y}}}(\widetilde{L}_{xi}^{\star})$. Another estimator of $\sigma_{\Lambda,x}^2$ is therefore

$$\widehat{\widetilde{\sigma}}_{\Lambda,x}^2=\frac{a_2}{S_{n,0}(x)}\sum_{i=1}^{n}\omega_i(x,h)\left(\widetilde{\lambda}_{xi}^{\star}-\frac{\widetilde{\lambda}_{xh}^{\star}}{\sqrt{nh}}\right)^2,$$

where $\widetilde{\lambda}_{xh}^{\star}=\sqrt{nh}\sum_{i=1}^{n}\omega_i(x,h)\,\widetilde{\lambda}_{xi}^{\star}$. The consistency of $\widehat{\widetilde{\sigma}}_{\Lambda,x}^2$ for estimating $\sigma_{\Lambda,x}^2$ could be established using the same arguments as those leading to Proposition 4.1, except that $(\mathcal{H}^{\star})$ and $(\mathcal{N}^{\star})$ replace Assumptions $(\mathcal{H})$ and $(\mathcal{N})$.

4.2. Estimation of the asymptotic biases

The proposed estimator of the asymptotic bias $\mu_{\Lambda,x}$ follows an idea similar to that of [21] by considering a local polynomial regression approach. Specifically, for some integer $p\geq 2$ and some bandwidth $h_B>0$, consider

$$(\widehat{\beta}_{\Lambda,0},\ldots,\widehat{\beta}_{\Lambda,p})=\operatorname*{arg\,min}_{(\beta_0,\ldots,\beta_p)}\sum_{i=1}^{n}\left\{\lambda_{xi}^{\star}-\sum_{j=0}^{p}\beta_j\left(\frac{Y_i-x}{h_B}\right)^j\right\}^2 K\left(\frac{Y_i-x}{h_B}\right).$$

Then an estimator of $\Lambda'_{C_x^{\mathrm{Z}\to\mathrm{Y}}}(\mathbb{B}_x)$ is given by $\widehat{\beta}_{\Lambda,2}/h_B^2$, so that and estimator of $\mu_{\Lambda,x}=\kappa\,a_1\,\Lambda'_{C_x^{\mathrm{Z}\to\mathrm{Y}}}(\mathbb{B}_x)$ is

$$\widehat{\mu}_{\Lambda,x}=\sqrt{nh^5}\,a_1\,\frac{\widehat{\beta}_{\Lambda,2}}{h_B^2}\,.$$

Roughly speaking, the rationale for the above estimator is that the theory of local polynomial smoothing allows to deduce that $\widehat{\beta}_{\Lambda,j}$ consistently estimates the j-th derivative of the mapping

$$z\mapsto\mathrm{E}\left(\lambda_{ix}^{\star}|Y_i=z\right)\approx\Lambda'_{C_x^{\mathrm{Z}\to\mathrm{Y}}}\left\{\mathrm{E}(L_{xi}\mid Y_i=z)\right\}.$$

For any z in a neighbourhood of x, one then has

$$\frac{\partial^2}{\partial z^2}\,\Lambda'_{C_x^{\mathrm{Z}\to\mathrm{Y}}}\left\{\mathrm{E}(L_{xi}\mid Y_i=z)\right\}\approx h_B^2\,\Lambda'_{C_x^{\mathrm{Z}\to\mathrm{Y}}}\left\{\mathrm{E}(\mathbb{C}_x)\right\}=h_B^2\,\Lambda'_{C_x^{\mathrm{Z}\to\mathrm{Y}}}(\mathbb{B}_x).$$

The consistency of $\widehat{\mu}_{\Lambda,x}$ as an estimator of $\mu_{\Lambda,x}$ is established next.

Proposition 4.2. *Under the conditions of Proposition 4.1, and if in addition Assumptions $(\mathcal{H}_B)$ and $(\mathcal{N}_B)$ in Appendix A.3 are satisfied, then $\widehat{\mu}_{\Lambda,x}$ converges in probability to $\mu_{\Lambda,x}$.*

The estimation of the asymptotic bias $\widetilde{\mu}_{\Lambda,x}$ of the estimator $\widetilde{\theta}_{\Lambda,xh}^{\mathrm{Z}\to\mathrm{Y}}$ based on $\widetilde{C}_{xh}^{\mathrm{Z}\to\mathrm{Y}}$ can be done similarly by using $\widetilde{\lambda}_{xi}^{\star}$ instead of $\lambda_{xi}^{\star}$; the resulting estimator is noted $\widehat{\widetilde{\mu}}_{\Lambda,x}$ in the sequel. Its consistency as an estimator of $\widetilde{\mu}_{\Lambda,x}$

could be established similarly as in the proof of Proposition 4.2, again by replacing Assumptions $(\mathcal{H})$ and $(\mathcal{N})$ by Assumptions $(\mathcal{H}^\star)$ and $(\mathcal{N}^\star)$. Note that Assumption $(\mathcal{H}_B)$ is no longer required since the marginal distributions have been uniformized.

4.3. A note on the estimation of Hadamard derivatives

When the functional Λ is linear, its Hadamard derivative is free of g, *i.e.* $\Lambda'_g = \Lambda'$ for all g; this happens in particular for the Spearman functional, see Equation (9). Otherwise, Λ'_g needs to be estimated. In that case, if one can find an estimator $\widehat{\Lambda'_{C_x^{Z \to Y}}}$ such that for any $\delta \in \ell^\infty([0,1]^2)$,

$$\left| \widehat{\Lambda'_{C_x^{Z \to Y}}}(\delta) - \Lambda'_{C_x^{Z \to Y}}(\delta) \right| = o_{\mathbb{P}}(1),$$

then the conclusions of Proposition 4.1 and Proposition 4.2 remain valid. This situation occurs for the Kendall functional, for which a natural plug-in estimator of its derivative in the light of (10) is given by

$$\widehat{(\Lambda_\tau)'}_{C_x^{Z \to Y}}(\delta) = 4 \int_{[0,1]^2} \left\{ \delta(u,v) \, \mathrm{d}C_{xh}(u,v) + C_{xh}(u,v) \, \mathrm{d}\delta(u,v) \right\}.$$

When working with the second estimator $\widetilde{\tau}_{xh}^{Z \to Y}$, one uses $\widetilde{C}_{xh}$ instead of C_{xh} in the above expression.

4.4. Data-driven algorithms for the selection of optimal bandwidths

The computation of the estimator $\theta_{\Lambda,xh}^{Z \to Y}$ of the local causality index $\theta_{\Lambda,x}^{Z \to Y}$ requires the choice of the bandwidth h. From Proposition 3.1 (i), one deduces that $\theta_{\Lambda,xh}^{Z \to Y} - \theta_{\Lambda,x}^{Z \to Y}$ is approximately Normal with variance $\sigma_{\Lambda,x}^2/nh$ and mean

$$\frac{\kappa \, a_1 \, \Lambda'_{C_x^{Z \to Y}}(\mathbb{B}_x)}{\sqrt{nh}} \approx h^2 a_1 \Lambda'_{C_x^{Z \to Y}}(\mathbb{B}_x).$$

A strategy described for instance by [10] and [18] is to select the bandwidth that minimises the asymptotic mean-squared errors, which in the current context is

$$\mathrm{AMSE}_{\Lambda,x}(h) = h^4 a_1^2 \left(\Lambda'_{C_x^{Z \to Y}}(\mathbb{B}_x) \right)^2 + \frac{\sigma_{\Lambda,x}^2}{nh}.$$

It is then easy to see that the minimum of $\mathrm{AMSE}_{\Lambda,x}$ is achieved when

$$h = h_{\Lambda,x}^\star = \left\{ \frac{\sigma_{\Lambda,x}^2}{4a_1^2 (\Lambda'_{C_x^{Z \to Y}}(\mathbb{B}_x))^2} \right\}^{1/5} n^{-1/5}. \tag{13}$$

Since $\sigma_{\Lambda,x}^2$ and $\Lambda'_{C_x^{Z \to Y}}(\mathbb{B}_x)$ are unknown, an idea is to first estimate these quantities with the estimators introduced in subsection 4.1 and subsection 4.2, namely

$\widehat{\sigma}^2_{\Lambda,x}$ and $\widehat{\beta}_{\Lambda,2}/h_B^2$, by using pilot bandwidths and then plugging these values into the expression for $h^\star_{\Lambda,x}$ in (13). The procedure can then be repeated recursively until some convergence criterion is reached; see [10], for instance. Specifically, let *cst* be a constant whose value remains bounded as n increases and fix $p \geq 2$. Then, proceed as follows:

(1) Compute $\widehat{\sigma}^2_{\Lambda,x}$ based on the initial bandwidth $h_0 = cst \times n^{-1/5}$;
(2) Compute $\widehat{\beta}_{\Lambda,2}$ using the bandwidth $h_B = cst \times h_0 \times n^{1/5-1/(2p+3)}$;
(3) Compute $h^\star_{\Lambda,x}$ in Equation (13) based on $\widehat{\sigma}^2_{\Lambda,x}$ and $\widehat{\beta}_{\Lambda,2}$;
(4) Repeat Steps 1–3 with $h_0 = h^\star_{\Lambda,x}$ until a convergence criterion or a maximum number of iterations is reached.

It can readily be verified that the above algorithm ensures that h_0 and $h^\star_{\Lambda,x}$ satisfy Assumption $(\mathcal{N})$, while h_B satisfies Assumption $(\mathcal{N}_B)$. In practice, *cst* is set as the inter-quartile range of $Y_1, \ldots, Y_n$, as suggested for instance by [13].

The procedure is more involved for the estimator $\widetilde{\theta}^{Z \to Y}_{\Lambda,xh}$ since one needs to first select the bandwidths h_1, h_2 that are necessary for the computation of the pseudo-observations in Equation (6). To this end, one follows the approach suggested by [13] for the estimation of the marginal conditional distributions F_{1xh_1} and F_{2xh_2} in order to first select h_1 and h_2 in an optimal manner. Then, conditionally on these optimal choices $h_1^\star$ and $h_2^\star$, the pseudo-observations in (6) are computed and the previously described algorithm is run using the estimators of the asymptotic variance and bias of $\widetilde{\theta}^{Z \to Y}_{\Lambda,xh}$.

5. Inference for local causality

5.1. Confidence intervals

A confidence interval of level $1 - \alpha$ for $\theta^{Z \to Y}_{\Lambda,x}$ based on the asymptotic normality of $\theta^{Z \to Y}_{\Lambda,xh}$ established in Proposition 3.1 (i) is

$$\mathcal{CI}_{\alpha,\Lambda,x} = \left[\theta^{Z \to Y}_{\Lambda,xh} - \left(\frac{z_{\alpha/2}\,\widehat{\sigma}_{\Lambda,x} + \widehat{\mu}_{\Lambda,x}}{\sqrt{nh}} \right), \theta^{Z \to Y}_{\Lambda,xh} + \left(\frac{z_{\alpha/2}\,\widehat{\sigma}_{\Lambda,x} - \widehat{\mu}_{\Lambda,x}}{\sqrt{nh}} \right) \right],$$

where for Φ^{-1} being the inverse cumulative distribution function of the standard Normal, $z_\alpha = \Phi^{-1}(1-\alpha)$. Proposition 4.1 and Proposition 4.2 on the consistency of $\widehat{\sigma}_{\Lambda,x}$ and $\widehat{\mu}_{\Lambda,x}$ ensure that this interval is asymptotically of level $1 - \alpha$. The confidence interval is similar when $\theta^{Z \to Y}_{\Lambda,x}$ is estimated by $\widetilde{\theta}^{Z \to Y}_{\Lambda,xh}$, namely

$$\widetilde{\mathcal{CI}}_{\alpha,\Lambda,x} = \left[\widetilde{\theta}^{Z \to Y}_{\Lambda,xh} - \left(\frac{z_{\alpha/2}\,\widetilde{\widehat{\sigma}}_{\Lambda,x} + \widetilde{\widehat{\mu}}_{\Lambda,x}}{\sqrt{nh}} \right), \theta^{Z \to Y}_{\Lambda,xh} + \left(\frac{z_{\alpha/2}\,\widetilde{\widehat{\sigma}}_{\Lambda,x} - \widetilde{\widehat{\mu}}_{\Lambda,x}}{\sqrt{nh}} \right) \right].$$

The coverage probabilities of $\mathcal{CI}_{\alpha,\Lambda,x}$ and $\widetilde{\mathcal{CI}}_{\alpha,\Lambda,x}$ in the case of small and moderate sample sizes is investigated in Section 6.

5.2. Testing for local non-causality

Saying that there is no local causality relationship from Z to Y at x in the process $\{(Y_t, Z_t)\}_{t\in\mathbb{Z}}$ means that Y_{t+1} and Z_t are conditionally independent with respect to $Y_t = x$. In other words, $C_x^{Z\to Y}(u,v) = \Pi(u,v) = uv$, *i.e.* the local causality copula corresponds to the independence copula. In that case, the limit covariance structure of $\mathbb{C}_{xh}^{Z\to Y}$ and $\widetilde{\mathbb{C}}_{xh}^{Z\to Y}$ simplifies to $\gamma_\Pi(u,v,u',v') = \{\min(u,u') - uu'\}\{\min(v,v') - vv'\}$. The following result is a special case of Proposition 3.1 and is stated without proof.

Proposition 5.1. *Under the null hypothesis of local non-causality from Z to Y at x, one has under Assumptions ($\mathcal{S}$), ($\mathcal{C}$), ($\mathcal{L}$), ($\mathcal{H}$) and ($\mathcal{N}$) that $\sqrt{nh}\,\theta_{\Lambda,xh}^{Z\to Y}$ is asymptotically Normal with mean $\mu_{\Lambda,x}$ and variance*

$$\sigma_{\Lambda,\Pi,x}^2 = \frac{a_2}{f_Y(x)}\,\Lambda_\Pi' \circ \Lambda_\Pi'(\gamma_\Pi).$$

Similarly, as long as ($\mathcal{S}$), ($\mathcal{C}$), ($\mathcal{L}$), ($\mathcal{H}^\star$) and ($\mathcal{N}^\star$) hold, $\sqrt{nh}\,\widetilde{\theta}_{\Lambda,xh}^{Z\to Y}$ is asymptotically Normal with mean $\widetilde{\mu}_{\Lambda,x}$ and variance $\sigma_{\Lambda,\Pi,x}^2$.

Proposition 5.1 can be exploited to test the null hypothesis of local non-causality at x from Z to Y. In that case, the null and alternative hypotheses are

$$\mathbb{H}_0 : \theta_{\Lambda,x}^{Z\to Y} = 0 \quad \text{and} \quad \mathbb{H}_1 : \theta_{\Lambda,x}^{Z\to Y} \neq 0.$$

Since a_1 and $\Lambda_\Pi' \circ \Lambda_\Pi'(\gamma_\Pi)$ are known values in the expression for $\sigma_{\Lambda,\Pi,x}^2$, an estimator of the asymptotic variance under local non-causality is

$$\widehat{\sigma}_{\Lambda,\Pi,x}^2 = \frac{a_2}{S_{n,0}(x)}\,\Lambda_\Pi' \circ \Lambda_\Pi'(\gamma_\Pi).$$

Test based on the statistics $\theta_{\Lambda,xh}^{Z\to Y}$ and $\widetilde{\theta}_{\Lambda,xh}^{Z\to Y}$ will then reject the null hypothesis of local non-causality whenever

$$\left|\frac{\sqrt{nh}\,\theta_{\Lambda,xh}^{Z\to Y} - \widehat{\mu}_{\Lambda,x}}{\widehat{\sigma}_{\Lambda,\Pi,x}}\right| > z_{\alpha/2} \quad \text{or} \quad \left|\frac{\sqrt{nh}\,\widetilde{\theta}_{\Lambda,xh}^{Z\to Y} - \widehat{\widetilde{\mu}}_{\Lambda,x}}{\widehat{\sigma}_{\Lambda,\Pi,x}}\right| > z_{\alpha/2}.$$

For the Spearman measure of local causality, one has from (9) that

$$\Lambda_\Pi' \circ \Lambda_\Pi'(\gamma_\Pi) = 144 \int_{[0,1]^4} \gamma_\Pi(u,u',v,v')\,\mathrm{d}u\,\mathrm{d}v\,\mathrm{d}u'\,\mathrm{d}v' = 1,$$

while for the Kendall measure, (10) entails

$$\Lambda_\Pi' \circ \Lambda_\Pi'(\gamma_\Pi) = 64 \int_{[0,1]^4} \gamma_\Pi(u,u',v,v')\,\mathrm{d}u\,\mathrm{d}v\,\mathrm{d}u'\,\mathrm{d}v' = \frac{4}{9}.$$

These values correspond to the well-known asymptotic variances of Spearman's rho and Kendall's tau under independence.

6. Simulations study

6.1. Preliminaries

The goal of this section is to investigate the performance of the methodologies introduced in this work when local causality is measured using the Spearman and Kendall functionals. As suggested by [13], modified versions of Kendall's measure has been employed, namely

$$\tau_{xh}^{Z\to Y} = \frac{4}{1 - \sum_{i=1}^{n}\omega_i^2(x,h)} \sum_{i,j=1}^{n} \omega_i(x,h)\,\omega_j(x,h)\,\mathbb{I}\,(Y_{j+1} < Y_{i+1}, Z_j < Z_i) - 1,$$

and similarly for $\widetilde{\tau}_{xh}^{Z\to Y}$. These versions are asymptotically equivalent to those defined from the functional Λ_τ. The results that will be reported have been obtained using the local linear weights in (2) based on the Epanechnikov kernel

$$K(z) = \frac{3}{4}\left(1 - z^2\right)\mathbb{I}\left(|z| \le 1\right).$$

The latter satisfies Assumption $(\mathcal{L})$ and one can show that $a_1 = 1/10$ and $a_2 = 3/5$. The estimation of the partial derivatives $C_x^{[1]}$ and $C_x^{[2]}$, which is needed to estimate the asymptotic biases and variance, will be based on the finite-difference estimators

$$\widehat{C}_x^{[1]}(u,v) = \sqrt{nh}\left[C_{xh}^{Z\to Y}\left\{\min\left(u + \frac{1}{\sqrt{nh}}, 1\right), v\right\} - C_{xh}^{Z\to Y}(u,v)\right]$$

and

$$\widehat{C}_x^{[2]}(u,v) = \sqrt{nh}\left[C_{xh}^{Z\to Y}\left\{u, \min\left(v + \frac{1}{\sqrt{nh}}, 1\right)\right\} - C_{xh}^{Z\to Y}(u,v)\right].$$

The latter are uniformly consistent and fulfill condition (12) needed in Proposition 4.1 and Proposition 4.2. For the estimation of the biases, many experiments suggest that taking $p = 3$ provides the more accurate results. Finally note that the bandwidth selection used the convergence criterion $|h_0 - h_{\Lambda,x}^\star| < 5 \times 10^{-4}$.

6.2. Accuracy of the local causality estimators

The performance of the nonparametric local causality measures $\rho_{xh}^{Z\to Y}$, $\widetilde{\rho}_{xh}^{Z\to Y}$, $\tau_{xh}^{Z\to Y}$ and $\widetilde{\tau}_{xh}^{Z\to Y}$ is studied here in the light of their bias and mean-squared errors. To this end, time series have been simulated from the stationary vector autoregressive model of order one given by

$$\begin{pmatrix} Y_t \\ Z_t \end{pmatrix} = \Sigma \begin{pmatrix} Y_{t-1} \\ Z_{t-1} \end{pmatrix} + \boldsymbol{\epsilon}_t,$$

where $\Sigma \in \mathbb{R}^{2\times 2}$ is such that $\Sigma_{11} = \Sigma_{22} = (\theta_2 - \theta_1\theta_3)/(1 - \theta_3^2)$ and $\Sigma_{21} = \Sigma_{12} = (\theta_1 - \theta_2\theta_3)/(1 - \theta_3^2)$; the process of innovations $(\boldsymbol{\epsilon}_t)_{t\in\mathbb{Z}}$ is i.i.d. from the mean-zero

bivariate Normal distribution with variance $\sigma_\epsilon^2 = 1 - \Sigma_{11}^2 - \Sigma_{21}^2 - 2\theta_3\Sigma_{11}\Sigma_{21}$ and correlation $\rho_\epsilon = \theta_3(1 - \Sigma_{11}^2 - \Sigma_{21}^2) - 2\Sigma_{11}\Sigma_{21}$. With this particular choice of parameters, $(Y_t, Z_t, Y_{t-1}, Z_{t-1})$ is centered Normal with some covariance matrix depending on $\theta_1, \theta_2, \theta_3$ in such a way that the local causality copula $C_x^{Z \to Y}$ is the Normal copula with parameter

$$\varrho = \frac{\theta_1 - \theta_2\,\theta_3}{\sqrt{1 - \theta_2^2}\sqrt{1 - \theta_3^2}}.$$

This model satisfies Assumptions $(\mathcal{S})$, $(\mathcal{H})$, $(\mathcal{H}^\star)$ and $(\mathcal{C})$. In Figure 1, the estimated bias and mean-squared error of $\rho_{xh}^{Z \to Y}$ and $\widetilde{\rho}_{xh}^{Z \to Y}$ when $x = 1/2$ are given as a function of the chosen bandwidth parameters $h = h_1 = h_2$; Figure 2 concerns Kendall's measure of local causality. These estimates are based on 1 000 replicates of the vector autoregressive process when $(\theta_1, \theta_2, \theta_3) \in \{(.25, .5, .5), (.482, .5, .5), (.485, .25, 0.6)\}$. The three variants of the model correspond respectively to $\tau_x^{Z \to Y} = 0$, $\tau_x^{Z \to Y} = .2$ and $\tau_x^{Z \to Y} = .4$, where the well-known relationship $\tau_x^{Z \to Y} = 2\sin^{-1}(\varrho)/\pi$ between Kendall's tau and the parameter of the Normal copula has been used. The mean value of the optimal bandwidth computed from the algorithm in subsection 4.4, taken over the replicates, are indicated on the curves.

A first look at Figure 1 indicates that $\widetilde{\rho}_{xh}^{Z \to Y}$ outperforms $\rho_{xh}^{Z \to Y}$ in terms of bias for all values of h and under the three scenarios. This is especially true under the scenarios when $\tau_x^{Z \to Y} \in \{.2, .4\}$, which might be explained by the strong influence of the conditional marginal distributions F_{1x} and F_{2x} on $\mathrm{E}(\rho_{xh}^{Z \to Y})$ for these models. The curves of the variance of $\rho_{xh}^{Z \to Y}$ and $\widetilde{\rho}_{xh}^{Z \to Y}$ as a function h (not presented here) are very similar, however, in accordance with the fact that they are asymptotically equal. Looking at the curves on the right panels, one sees that the data-driven selection rule for an optimal bandwidth is quite accurate in the sense that it minimizes the mean-squared error. All the above comments can also be made about $\tau_{xh}^{Z \to Y}$ and $\widetilde{\tau}_{xh}^{Z \to Y}$ upon looking at Figure 2.

6.3. Coverage probability of interval estimations

The aim of this subsection is to evaluate the finite-sample coverage probabilities of interval estimations of $\rho_x^{Z \to Y}$ and $\tau_x^{Z \to Y}$ as described in subsection 5.1. To this end, a general D-Vine structure for bivariate processes suggested by [2] will be used; this construction allows for the choice of the local causality copula $C_x^{Z \to Y}$; for the upcoming simulations, $C_x^{Z \to Y}$ is the Normal copula parameterized in such a way that $\tau_x^{Z \to Y} \in \{.0, .2, .4\}$. The algorithm of [2] necessitates to select a vector (C_1, C_2, C_3, C_4) of four copulas, where C_1 is the copula of (Y_t, Z_t), C_2 that of (Y_{t-1}, Y_t), C_3 that of the distribution of (Y_{t-1}, Z_t) given $Y_t = x$, and C_4 as that of (Z_t, Z_{t-1}) given $Y_{t-1} = x, Y_t = x'$. Here, $x = 1/2$.

The results in Table 1 concern the estimated coverage probabilities of 95% interval estimates based on $\rho_{xh}^{Z \to Y}$ and $\widetilde{\rho}_{xh}^{Z \to Y}$, while those in Table 2 concern $\tau_{xh}^{Z \to Y}$ and $\widetilde{\tau}_{xh}^{Z \to Y}$. The copulas (C_1, C_2, C_3, C_4) are Normal with Kendall's tau

$$(\tau_1, \tau_2, \tau_3, \tau_4) \in \{(.3, .1, .1, .1), (.05, .05, .05, .05), (.5, .3, .3, .05), (.5, .3, .75, .05)\}.$$
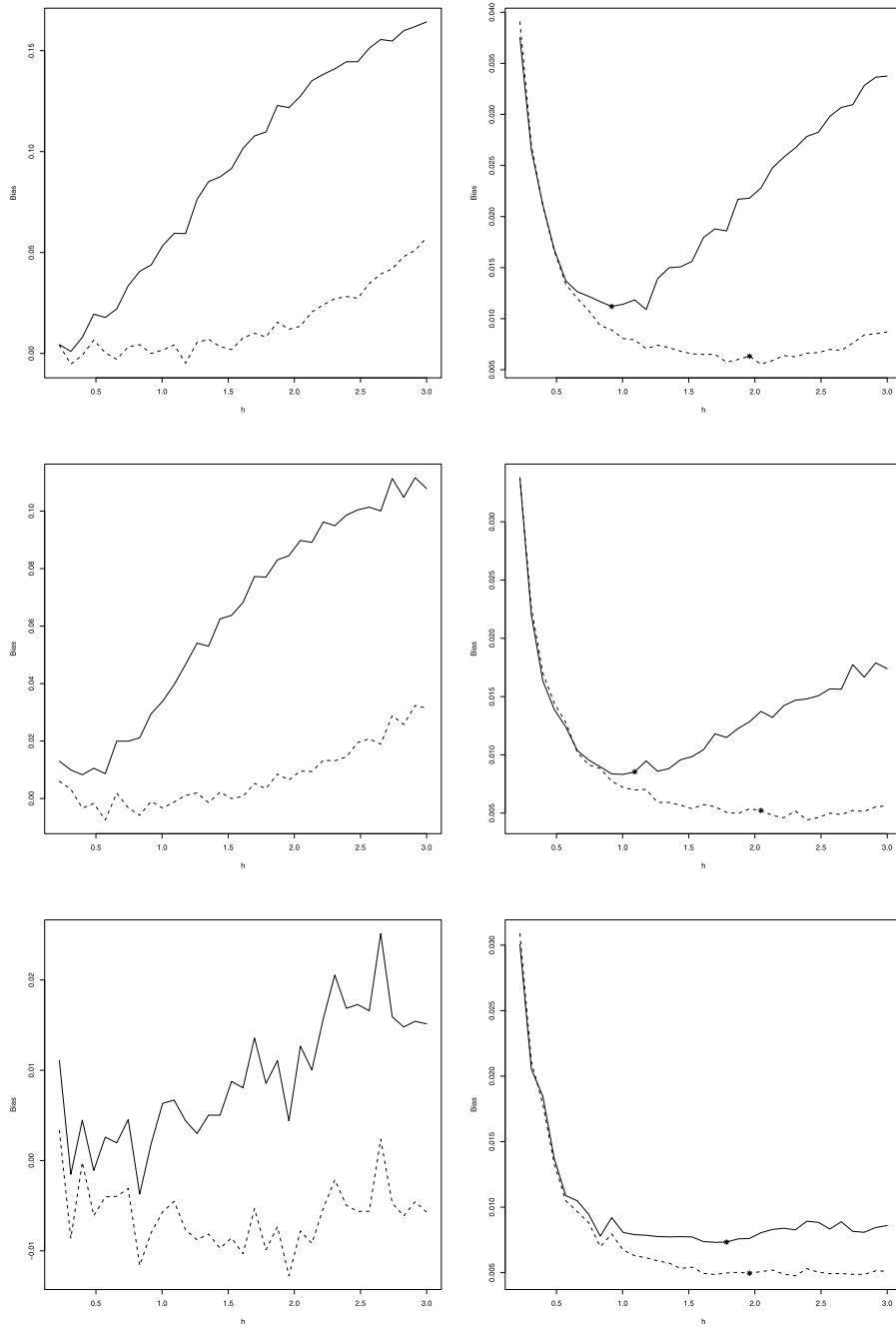
FIG 1. *Estimated bias (left) and mean-squared error (right) of $\rho_{xh}^{Z\to Y}$ (solid line) and $\tilde{\rho}_{xh}^{Z\to Y}$ (dashed line) when $n = 250$ as a function of the bandwidth parameter h under the vector autoregressive model such that $\tau_x^{Z\to Y} = .0$ (upper panels), $\tau_x^{Z\to Y} = .2$ (middle panels) and $\tau_x^{Z\to Y} = .4$ (bottom panels)*
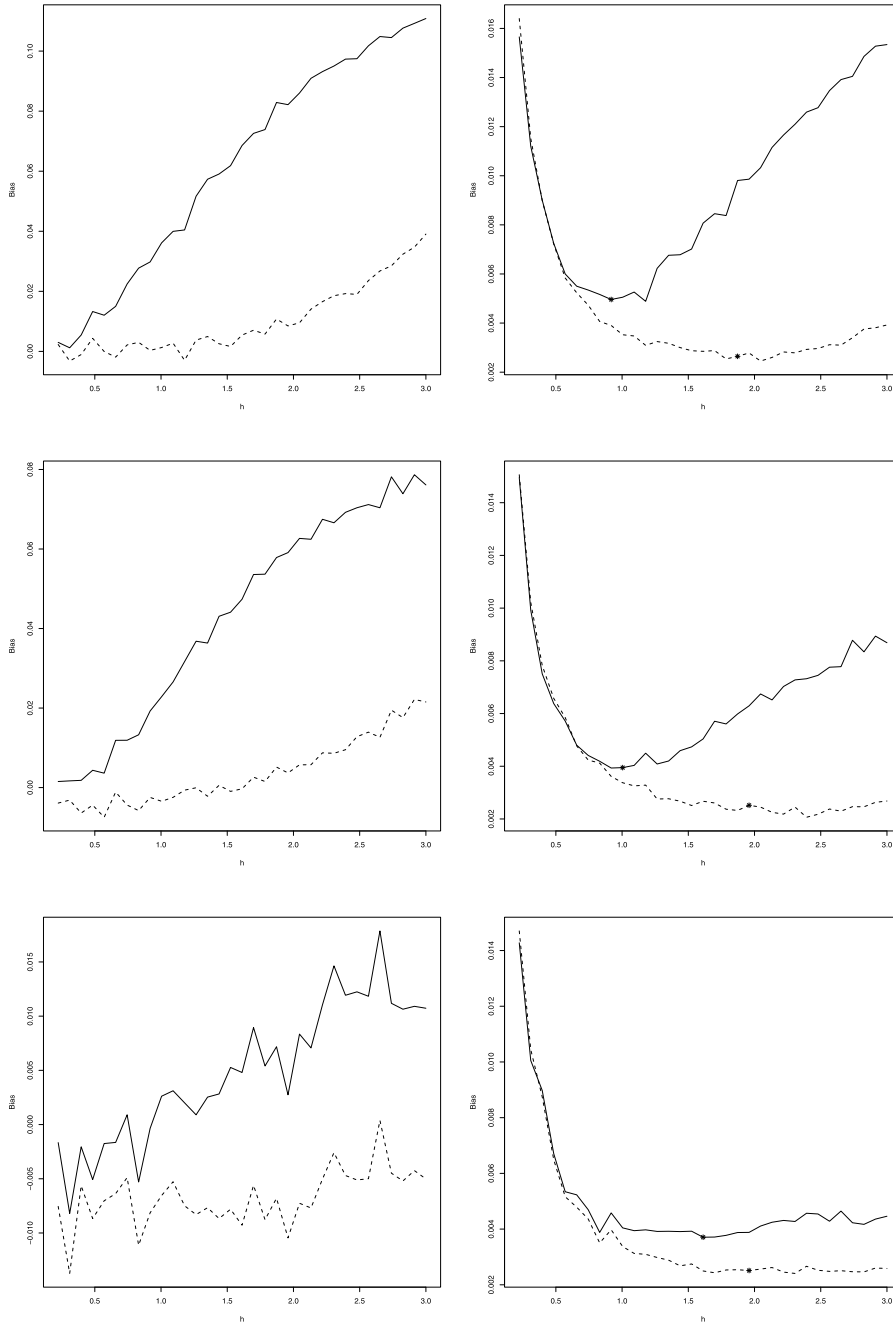
FIG 2. *Estimated bias (left) and mean-squared error (right) of $\tau_{xh}^{Z\to Y}$ (solid line) and $\widetilde{\tau}_{xh}^{Z\to Y}$ (dashed line) when $n = 250$ as a function of the bandwidth parameter h under the vector autoregressive model such that $\tau_{x}^{Z\to Y} = .0$ (upper panels), $\tau_{x}^{Z\to Y} = .2$ (middle panels) and $\tau_{x}^{Z\to Y} = .4$ (bottom panels)*

These scenarios generate various kinds of serial structures among the data. The results have been obtained by setting the bandwidth to $h^\star/2$ instead of the optimal bandwidth $h^\star$, since many numerical experiments suggested that it generally provides better results in terms of coverage probabilities. It is an indication that in the context of interval estimation, what matters the most is the bias; looking at Figures 1–2, the latter is minimized for h taken smaller than $h^\star$.

First of all, the coverage probabilities tend to be closer to their 95% confidence level as n increases; it is a simple consequence of the fact that the intervals are based on the asymptotic normality of the estimators. In Table 1, it can be seen that the estimated probabilities based on $\widetilde{\rho}_{xh}$ tend to be closer to their nominal level than those based on ρ_{xh} when $n \in \{250, 500\}$, while the results are similar when $n = 1000$. This is particularly true when $(\tau_1, \tau_2, \tau_3, \tau_4) = (.5, .3, .75, .05)$ and $\tau_x^{Z \to Y} \in \{.0, .2\}$. Similar observations can be drawn from Table 2.

6.4. Power of the tests of local non-causality

Consider testing the null hypothesis $\mathbb{H}_0$ of the local non-causality from Z to Y at x, *i.e.* the conditional independence between Y_t and Z_{t-1} given $Y_{t-1} = x$. To this end, one considers again the D-Vine structure for $(Y_t, Z_t)_{t \in \mathbb{Z}}$ described in Subsection 6.3. Here, the local causality copula $C_x^{Z \to Y}$ is taken to be either the Normal or the Clayton copula; the latter is defined by $C_\theta^{\mathrm{CL}}(u, v) = (u^{-\theta} + v^{-\theta} - 1)^{-1/\theta}$, $\theta > 0$. These two copulas have been parameterized in such a way that $\tau_x^{Z \to Y} = 2\{\Phi(x) - 1/2\}^2$ and x is chosen in order that $\tau_x^{Z \to Y} \in \{0, .1, .2\}$. Here again, (C_1, C_2, C_3, C_4) are Normal copulas parameterized such that

$$(\tau_1, \tau_2, \tau_3, \tau_4) \in \{(.3, .1, .1, .1), (.05, .3, .75, .05)\}.$$

The results on the estimated power of the tests based on $\rho_{xh}^{Z \to Y}$, $\widetilde{\rho}_{xh}^{Z \to Y}$, $\tau_{xh}^{Z \to Y}$ and $\widetilde{\tau}_{xh}^{Z \to Y}$ are in Table 3.

Under the null hypothesis of non-causality, *i.e.* when $\tau_x^{Z \to Y} = .0$, the four tests are slightly too liberal when $n = 250$; nevertheless, the estimated probabilities of rejection of $\mathbb{H}_0$ tend to the theoretical 5% nominal level. As expected, the ability of the four tests at rejecting the null hypothesis under departures from $\mathbb{H}_0$ increases with the sample size and is larger when $\tau_x^{Z \to Y} = .2$ compared to $\tau_x^{Z \to Y} = .1$. Interestingly, when $n \in \{500, 1000\}$, the tests based on $\widetilde{\rho}_{xh}$ and $\widetilde{\tau}_{xh}$ are clearly more powerful than those based on ρ_{xh} and τ_{xh} when $C_x^{Z \to Y}$ is the Clayton copula. Under a Normal causality copula, however, the four tests perform very similarly, except when $(\tau_1, \tau_2, \tau_3, \tau_4) = (.05, .3, .75, .05)$ and $\tau_x^{Y \to Z} = .1$, in which case the tests based on $\widetilde{\rho}_{xh}$ and $\widetilde{\tau}_{xh}$ are the best.

7. Illustration on financial data

The following illustration is based on the bivariate time series of the 1 512 daily observations taken between January 2010 and January 2016 for the compounded changes in prices (returns) and trading volume of the Standard and

TABLE 1
Coverage probabilities, as estimated from 1 000 replicates, of 95% confidence intervals for the Spearman local causality measures $\rho_x^{Z \to Y}$ under bivariate D-vine time series in which $C_x^{Z \to Y}$ is the Normal copula with Kendall's tau $\tau_x^{Z \to Y} \in \{.0, .2, .4\}$

$(\tau_1, \tau_2, \tau_3, \tau_4)$	$\tau_x^{Z \to Y}$	Estimator	$n = 250$	$n = 500$	$n = 1\ 000$
	.0	ρ_{xh}	93.2	92.2	93.0
		$\widetilde{\rho}_{xh}$	93.6	94.6	93.1
$(.30, .10, .10, .10)$	.2	ρ_{xh}	93.5	93.3	94.5
		$\widetilde{\rho}_{xh}$	94.3	93.0	95.1
	.4	ρ_{xh}	93.6	92.7	93.6
		$\widetilde{\rho}_{xh}$	93.4	93.9	93.0
	.0	ρ_{xh}	93.0	94.1	95.3
		$\widetilde{\rho}_{xh}$	93.7	93.6	93.2
$(.05, .05, .05, .05)$	.2	ρ_{xh}	93.5	93.4	93.6
		$\widetilde{\rho}_{xh}$	94.1	93.0	94.1
	.4	ρ_{xh}	93.8	93.4	94.1
		$\widetilde{\rho}_{xh}$	94.3	94.8	93.1
	.0	ρ_{xh}	93.9	93.0	93.6
		$\widetilde{\rho}_{xh}$	93.9	94.7	94.5
$(.50, .30, .30, .05)$	.2	ρ_{xh}	91.8	93.1	94.5
		$\widetilde{\rho}_{xh}$	94.1	93.8	94.4
	.4	ρ_{xh}	90.7	91.8	93.3
		$\widetilde{\rho}_{xh}$	92.5	93.6	93.2
	.0	ρ_{xh}	91.3	94.0	94.0
		$\widetilde{\rho}_{xh}$	94.7	94.2	93.4
$(.05, .30, .75, .05)$	.2	ρ_{xh}	90.3	92.1	93.8
		$\widetilde{\rho}_{xh}$	94.0	94.0	92.9
	.4	ρ_{xh}	88.1	89.8	89.5
		$\widetilde{\rho}_{xh}$	90.7	89.6	90.3

Poor's 500 (S&P500) Index. The relationship between these two indices has been extensively studied, both from a theoretical and from an empirical perspective. According to the tests of stationarity reported in [6], one will work instead with the first difference in logarithmic returns (Y) and with the first difference in logarithmic volume (Z). Put differently, Z and Y are respectively the log of the ratio of two consecutive recorded values of stock return and stock volume. Y_t and Z_t are therefore indicators of the growth rate of respectively stock return

TABLE 2

Coverage probabilities, as estimated from 1 000 replicates, of 95% confidence intervals for the Kendall local causality measures $\tau_x^{Z \to Y}$ under bivariate D-vine time series in which $C_x^{Z \to Y}$ is the Normal copula with Kendall's tau $\tau_x^{Z \to Y} \in \{.0, .2, .4\}$

$(\tau_1, \tau_2, \tau_3, \tau_4)$	$\tau_x^{Z \to Y}$	Estimator	$n = 250$	$n = 500$	$n = 1\,000$
	.0	τ_{xh}	93.7	92.4	93.2
		$\widetilde{\tau}_{xh}$	94.9	95.8	95.3
$(.30, .10, .10, .10)$	.2	τ_{xh}	94.2	93.6	94.3
		$\widetilde{\tau}_{xh}$	95.4	96.4	95.0
	.4	τ_{xh}	93.3	93.2	93.5
		$\widetilde{\tau}_{xh}$	94.8	95.6	94.8
	.0	τ_{xh}	93.0	94.0	95.7
		$\widetilde{\tau}_{xh}$	94.0	96.4	95.5
$(.05, .05, .05, .05)$	.2	τ_{xh}	94.0	93.3	93.6
		$\widetilde{\tau}_{xh}$	94.1	95.1	96.2
	.4	τ_{xh}	94.1	94.2	95.2
		$\widetilde{\tau}_{xh}$	94.7	94.4	95.0
	.0	τ_{xh}	94.1	93.5	94.0
		$\widetilde{\tau}_{xh}$	92.9	95.4	95.1
$(.50, .30, .30, .05)$	.2	τ_{xh}	92.4	94.4	94.9
		$\widetilde{\tau}_{xh}$	94.8	95.5	94.7
	.4	τ_{xh}	91.5	92.2	94.1
		$\widetilde{\tau}_{xh}$	95.3	93.8	95.1
	.0	τ_{xh}	92.5	94.1	94.4
		$\widetilde{\tau}_{xh}$	94.7	95.3	96.9
$(.50, .30, .75, .05)$	.2	τ_{xh}	91.8	92.7	94.4
		$\widetilde{\tau}_{xh}$	93.2	95.9	94.5
	.4	τ_{xh}	89.8	90.2	90.2
		$\widetilde{\tau}_{xh}$	92.0	92.4	93.5

and stock volume from period $t - 1$ to period t, and the upcoming conclusions will have to be interpreted accordingly.

The causality from Z to Y is then investigated from the sample (Y_2, Z_1, Y_1), ..., $(Y_{1511}, Z_{1510}, Y_{1510})$. For these data, the value of the partial correlation coefficient of (Y_{t+1}, Z_t) given Y_t is -0.024×10^{-4}, leading to the conclusion of a global non-causality $(p = .36)$. However, such a conclusion can be misleading when the relationship between Y_{t+1} and Z_t changes according to the value taken by Y_t. This is exactly what happens here. For example, if one considers the sub-

TABLE 3

Percentages of rejection of the null hypothesis of local non-causality, as estimated from 1 000 replicates, for the tests at level $\alpha = 0.05$ based on $\rho_{xh}^{Z\to Y}$, $\widetilde{\rho}_{xh}^{Z\to Y}$, $\tau_{xh}^{Z\to Y}$ and $\widetilde{\tau}_{xh}^{Z\to Y}$ under D-vine time series in which $C_x^{Z\to Y}$ is a Normal or Clayton copula; upper panel: $(\tau_1, \tau_2, \tau_3, \tau_4) = (.30, .10, .10, .10)$; lower panel: $(\tau_1, \tau_2, \tau_3, \tau_4) = (.50, .30, .75, .05)$

$\tau_x^{Z\to Y}$	Estimator	Normal copula			Clayton copula		
		$n = 250$	$n = 500$	$n = 1000$	$n = 250$	$n = 500$	$n = 1000$
.0	$\rho_{xh}^{Z\to Y}$	8.2	6.7	5.6	7.5	7.1	5.3
	$\widetilde{\rho}_{xh}^{Z\to Y}$	9.3	5.8	6.5	8.2	6.7	6.2
	$\tau_{xh}^{Z\to Y}$	7.3	6.4	5.6	7.1	7.0	5.4
	$\widetilde{\tau}_{xh}^{Z\to Y}$	8.8	5.9	6.4	7.3	6.5	6.0
.1	$\rho_{xh}^{Z\to Y}$	48.4	65.5	84.0	46.3	61.9	83.1
	$\widetilde{\rho}_{xh}^{Z\to Y}$	49.5	66.0	88.1	47.6	65.1	87.7
	$\tau_{xh}^{Z\to Y}$	45.9	63.1	82.7	44.5	60.6	81.5
	$\widetilde{\tau}_{xh}^{Z\to Y}$	46.2	65.2	87.1	46.1	64.2	87.0
.2	$\rho_{xh}^{Z\to Y}$	63.6	80.2	96.2	50.4	66.9	86.5
	$\widetilde{\rho}_{xh}^{Z\to Y}$	57.6	82.5	96.7	57.1	80.7	95.9
	$\tau_{xh}^{Z\to Y}$	59.5	79.3	95.7	48.0	65.2	85.7
	$\widetilde{\tau}_{xh}^{Z\to Y}$	52.8	81.7	96.0	54.7	79.0	95.3
.0	$\rho_{xh}^{Z\to Y}$	6.9	6.6	5.9	6.7	7.9	6.2
	$\widetilde{\rho}_{xh}^{Z\to Y}$	8.7	7.6	6.3	10.2	7.4	5.3
	$\tau_{xh}^{Z\to Y}$	6.5	6.6	5.9	5.8	7.6	5.9
	$\widetilde{\tau}_{xh}^{Z\to Y}$	8.4	7.1	6.2	9.2	7.1	5.5
.1	$\rho_{xh}^{Z\to Y}$	35.3	49.2	69.6	37.5	47.9	66.7
	$\widetilde{\rho}_{xh}^{Z\to Y}$	45.7	67.5	84.7	45.7	66.2	83.9
	$\tau_{xh}^{Z\to Y}$	32.3	47.7	67.9	34.1	44.9	64.8
	$\widetilde{\tau}_{xh}^{Z\to Y}$	43.5	66.8	84.2	44.0	64.8	83.3
.2	$\rho_{xh}^{Z\to Y}$	60.9	82.0	94.1	46.5	66.4	85.2
	$\widetilde{\rho}_{xh}^{Z\to Y}$	55.3	83.1	96.2	56.4	80.6	94.4
	$\tau_{xh}^{Z\to Y}$	57.6	80.3	93.7	42.4	64.1	84.1
	$\widetilde{\tau}_{xh}^{Z\to Y}$	52.6	81.3	96.2	52.2	79.5	94.2

sample for which $Y_t > 0$, then the partial correlation coefficient is 0.072, which this time is significantly different from zero ($p = .039$). On the other hand, the subsample for which $Y_t < 0$ leads to a partial correlation of -0.095, which also significantly departs from zero ($p < .01$).

In order to take into account the levels of Y_t, a solution is to rely on local causality indices as introduced in Section 3. Figure 3 reports the values of $\rho_{xh}^{Z\to Y}$, $\widetilde{\rho}_{xh}^{Z\to Y}$, $\tau_{xh}^{Z\to Y}$ and $\widetilde{\tau}_{xh}^{Z\to Y}$ as a function of x, together with 95% point-wise confidence intervals. The values of x that have been considered range between the 10-th and 90-th percentile of the Y. It is clear from these figures that the level of Y_t has an influence on the four local causality indices.

To interpret the recorded values at Figure 3, note that a positive value taken by an index under consideration (for e.g. at $x \approx 0.01$) suggests that, in cases
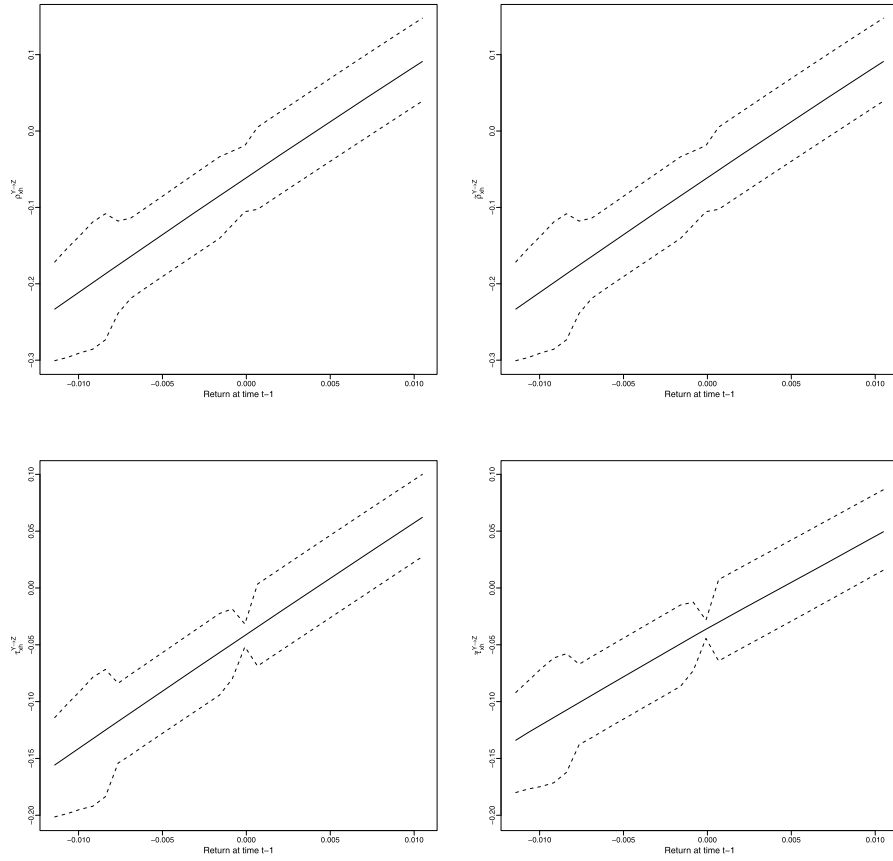
FIG 3. *Curves of $\rho_{xh}^{Z\rightarrow Y}$ (upper left), $\widetilde{\rho}_{xh}^{Z\rightarrow Y}$ (upper right), $\widetilde{\tau}_{xh}^{Z\rightarrow Y}$ (bottom left) and $\widetilde{\tau}_{xh}^{Y\rightarrow Z}$ (bottom right) as a function of x, together with the 95% point-wise confidence bands, for the Standard and Poor's 500 bivariate time series*

where the first difference in logarithmic volume at period t is given by x, then large (resp. small) values of the return's growth rate Z_t tend to be followed by large (resp. small) values of the volume's growth rate Y_{t+1}. Quite the opposite, when a local causality index is negative (for e.g. at $x \approx -0.01$), this conveys the idea that given $Y_t = x$, large (resp. small) values of Y_{t+1} generally occur after small (resp. large) values of Z_t.

From Figure 3, it can be seen that causal relationship between Y and Z appears to be stronger for negative Y_t, i.e. when the ratio of the volume at period t and the volume at period $t-1$ is less than 1, i.e. when there is a decrease in stock volumes from period $t-1$ to period t. This technically suggests that, in these cases, the return's growth rate at period t is more likely to influence the volume's growth rate at period $t+1$ compared to when Y_t is positive, i.e. compared to when there is an increase in stock volumes from period $t-1$ to period t.

TABLE 4

P-values of the tests of local non-causality based on $\rho_{xh}^{Z\to Y}$, $\widetilde{\rho}_{xh}^{Z\to Y}$, $\tau_{xh}^{Z\to Y}$ and $\widetilde{\tau}_{xh}^{Z\to Y}$ at selected values of x for the Standard and Poor's 500 bivariate time series

x	$\rho_{xh}^{Z\to Y}$	$\widetilde{\rho}_{xh}^{Z\to Y}$	$\tau_{xh}^{Z\to Y}$	$\widetilde{\tau}_{xh}^{Z\to Y}$
-0.0100	$< .001$	$< .001$	$< .001$	$< .001$
-0.0075	$< .001$	$< .001$	$< .001$	$< .001$
-0.0050	$< .001$	$< .001$	$< .001$	$< .001$
-0.0025	$< .001$	$.002$	$< .001$	$.002$
0.0000	$< .001$	$< .001$	$.005$	$.006$
0.0025	$.337$	$.376$	$.336$	$.373$
0.0050	$.514$	$.655$	$.526$	$.663$
0.0075	$.064$	$.144$	$.070$	$.150$
0.0010	$.001$	$.005$	$.001$	$.005$

To complement the above analysis, the tests for the null hypothesis of local non-causality have been performed for selected values of x; the results on the p-values of the tests are in Table 4. First note that the four tests are in agreement on the acceptance or rejection of $\mathbb{H}_0$ at the 5% level. Hence, for $x \in \{-.0100, -.0075, -.0050, -.0025, .0000, .0100\}$, the null hypothesis of non-causality is clearly rejected, while $\mathbb{H}_0$ is not rejected when $x \in \{.0025, .0050, .0075\}$. Further, an application of Bonferroni's correction to the tests based on $\rho_{xh}^{Z\to Y}$, $\widetilde{\rho}_{xh}^{Z\to Y}$, $\tau_{xh}^{Z\to Y}$ and $\widetilde{\tau}_{xh}^{Z\to Y}$ for testing at multiple values of x leads to the stronger statement that all of the reported p-values that were significant at the prescribed point-wise 5% level are also jointly significant at a family-wise error rate of 5%, except for the test based on $\widetilde{\tau}_{xh}^{Z\to Y}$ at $x = 0$, which is no longer regarded significant after the rectification for multiple testing. These results are in accordance with the conclusions of the tests of non-causality based on the partial correlation coefficient. This is also in agreement with the curves in Figure 3, where it can be observed that when $x < 0.005$, the 95% confidence bands all lie below zero, indicating a negative causal relationship; on the other side, these confidence bands are above zero for $x > 0.009$, suggesting a positive causal relationship.

Appendix A: Assumptions needed for the asymptotic results in Propositions 3.1, 4.1 and 4.2

In the sequel, let J_x be an open neighborhood of x.

A.1. Assumptions for the weak convergence of $C_{xh}^{\mathbf{Z}\to\mathbf{Y}}$ and $\theta_{\Lambda,xh}^{\mathbf{Z}\to\mathbf{Y}}$

($\mathcal{S}$) $\{(Y_{t+1}, Z_t, Y_t)\}_{t\in\mathbb{Z}}$ is stationary and its α-mixing coefficients satisfy $\alpha_r = O(r^{-a})$ for some $a > 6$. Also, one can find $M > 0$ such that the joint density $h_{\ell_1,\ldots,\ell_5}$ of $(Y_0, Y_{\ell_1}, \ldots, Y_{\ell_5})$ satisfies $h_{\ell_1,\ldots,\ell_5}(x_0, \ldots, x_{\ell_5}) \le M$ for all $x_0, \ldots, x_{\ell_5} \in J_x$..

($\mathcal{C}$) The first order partial derivatives $C_x^{[1]}(u,v) = \partial C_x(u,v)/\partial u$ and $C_x^{[2]}(u,v) = \partial C_x(u,v)/\partial v$ are continuous respectively on $(0,1)\times[0,1]$ and $[0,1]\times(0,1)$.

($\mathcal{L}$) The kernel K is positive, continuous, symmetric, vanishes outside of $[-1,1]$ and its second order derivative is bounded. Also, the density f_Y associated to $\{Y_t\}_{t\in\mathbb{Z}}$ exists, is twice differentiable and strictly positive at x.

($\mathcal{H}$) The following functions are uniformly continuous for $(w,y,z) \in J_x \times \mathbb{R}^2$:

$$\dot{H}_w(y,z) = \frac{\partial}{\partial w} H_w(y,z) \quad \text{and} \quad \ddot{H}_w(y,z) = \frac{\partial^2}{\partial w^2} H_w(y,z).$$

($\mathcal{N}$) As $n \to \infty$, $nh^5 \to \kappa^2 < \infty$ and $n^{8-\delta}h^{21} \to \infty$ for some $\delta > 0$.

A.2. Assumptions for the weak convergence of $\widetilde{C}_{xh}^{\mathbf{Z}\to\mathbf{Y}}$ and $\widetilde{\theta}_{\Lambda,xh}^{\mathbf{Z}\to\mathbf{Y}}$

($\mathcal{H}^\star$) For $j \in \{1,2\}$, the functions $F_{jw}\{F_{jw}^{-1}(u)\}$, $\dot{F}_{jw}\{F_{jw}^{-1}(u)\}$ and $\ddot{F}_{jw}\{F_{jw}^{-1}(u)\}$ are uniformly continuous for $(w,u) \in J_x\times[0,1]$. In addition, the following functions are uniformly continuous for $(w,u,v) \in J_x\times[0,1]^2$:

$$\dot{C}_w(u,v) = \frac{\partial}{\partial w} C_w(u,v) \quad \text{and} \quad \ddot{C}_w(u,v) = \frac{\partial^2}{\partial w^2} C_w(u,v).$$

($\mathcal{N}^\star$) As $n \to \infty$, $nh^5 \to \kappa^2 < \infty$, $n^{8-\delta}h^{21} \to \infty$ for some $\delta > 0$, $\max(nh_1^5, nh_2^5) < \infty$ and $\max(h/h_1, h/h_2) < \infty$.

A.3. Assumptions for the consistency of $\widehat{\mu}_{\Lambda,x}$ and $\widehat{\widetilde{\mu}}_{\Lambda,x}$

($\mathcal{H}_B$) Let $D_{1x}^{(j)}(y) = \partial^j F_{1x}(y)/\partial x^j$ and $D_{2x}^{(j)}(y) = \partial^j F_{2x}(y)/\partial x^j$. Then for each $j \in \{1,\ldots,p+1\}$, the mappings $u \mapsto D_{1x}^{(j)}\{F_{1x}^{-1}(u)\}$ and $u \mapsto D_{2x}^{(j)}\{F_{2x}^{-1}(u)\}$ are Lipchitz continuous in $u \in [0,1]$. In addition, $D_{1x}^{(p+1)}$ and $D_{2x}^{(p+1)}$ are uniformly bounded in $w \in J_x$.

($\mathcal{N}_B$) As $n \to \infty$, $h_B \to 0$, $nh_B^5 \to \infty$ and $nh_B^{2p+5} \to 0$.

Appendix B: Proofs of the main results

B.1. Proof of Proposition 3.1

(i) Since $\Theta_{\Lambda,xh}^{\mathbf{Z}\to\mathbf{Y}} = \sqrt{nh}\{\Lambda(C_{xh}^{\mathbf{Z}\to\mathbf{Y}}) - \Lambda(C_x^{\mathbf{Z}\to\mathbf{Y}})\}$ and because Λ is Hadamard differentiable, an application of the Functional delta method yields

$$\Theta_{\Lambda,xh}^{\mathbf{Z}\to\mathbf{Y}} = \Lambda'_{C_x^{\mathbf{Z}\to\mathbf{Y}}} \left\{ \sqrt{nh} \left(C_{xh}^{\mathbf{Z}\to\mathbf{Y}} - C_x^{\mathbf{Z}\to\mathbf{Y}} \right) \right\} + o_{\mathbb{P}}(1)$$

$$= \Lambda'_{C_x^{Z \to Y}} \left(\mathbb{C}_{xh}^{Z \to Y} \right) + o_{\mathbb{P}}(1).$$

Hence, $\Theta_{\Lambda,xh}^{Z \to Y}$ converges in distribution to $\Lambda'_{C_x^{Z \to Y}} \left(\mathbb{C}_x^{Z \to Y} \right)$. Since $\Lambda'_{C_x^{Z \to Y}}$ is a linear functional and $\mathbb{C}_x^{Z \to Y}$ is Gaussian, Lemma 3.9.8 in [25] entails that $\Theta_{\Lambda,xh}^{Z \to Y}$ converges in distribution to a normal random variable $\Theta_{\Lambda,x}^{Z \to Y}$ with mean

$$\mu_{\Lambda,x} = \mathrm{E} \left\{ \Lambda'_{C_x^{Z \to Y}} \left(\mathbb{C}_x^{Z \to Y} \right) \right\} = \Lambda'_{C_x^{Z \to Y}} \left\{ \mathrm{E} \left(\mathbb{C}_x^{Z \to Y} \right) \right\} = \kappa \, a_1 \, \Lambda'_{C_x^{Z \to Y}} (\mathbb{B}_x).$$

Next, upon noting that

$$
\begin{aligned}
\mu_{\Lambda,x}^2 &= \Lambda'_{C_x^{Z \to Y}} \left\{ \mathrm{E}(\mathbb{C}_x) \right\} \Lambda'_{C_x^{Z \to Y}} \left\{ \mathrm{E}(\mathbb{C}_x) \right\} \\
&= \Lambda'_{C_x^{Z \to Y}} \circ \Lambda'_{C_x^{Z \to Y}} \left\{ \mathrm{E} \left(\mathbb{C}_x(u,v) \right) \mathrm{E} \left(\mathbb{C}_x(u',v') \right) \right\},
\end{aligned}
$$

the variance of $\Theta_{\Lambda,x}^{Z \to Y}$ can be written

$$
\begin{aligned}
\sigma_{\Lambda,x}^2 &= \mathrm{Var} \left\{ \Lambda'_{C_x^{Z \to Y}} (\mathbb{C}_x^{Z \to Y}) \right\} \\
&= \mathrm{Cov} \left\{ \Lambda'_{C_x^{Z \to Y}} (\mathbb{C}_x^{Z \to Y}), \Lambda'_{C_x^{Z \to Y}} (\mathbb{C}_x^{Z \to Y}) \right\} \\
&= \mathrm{E} \left\{ \Lambda'_{C_x^{Z \to Y}} (\mathbb{C}_x^{Z \to Y}) \Lambda'_{C_x^{Z \to Y}} (\mathbb{C}_x^{Z \to Y}) \right\} - \mu_{\Lambda,x}^2 \\
&= \Lambda'_{C_x^{Z \to Y}} \circ \Lambda'_{C_x^{Z \to Y}} \left\{ \mathrm{E} \left(\mathbb{C}_x^{Z \to Y}(u,v) \, \mathbb{C}_x^{Z \to Y}(u',v') \right) \right\} - \mu_{\Lambda,x}^2 \\
&= \Lambda'_{C_x^{Z \to Y}} \circ \Lambda'_{C_x^{Z \to Y}} \left\{ \mathrm{E} \left(\mathbb{C}_x^{Z \to Y}(u,v) \, \mathbb{C}_x^{Z \to Y}(u',v') \right) \right\} \\
&\quad - \Lambda'_{C_x^{Z \to Y}} \circ \Lambda'_{C_x^{Z \to Y}} \left\{ \mathrm{E} \left(\mathbb{C}_x(u,v) \right) \mathrm{E} \left(\mathbb{C}_x(u',v') \right) \right\} \\
&= \Lambda'_{C_x^{Z \to Y}} \circ \Lambda'_{C_x^{Z \to Y}} \left\{ \mathrm{E} \left(\mathbb{C}_x^{Z \to Y}(u,v) \, \mathbb{C}_x^{Z \to Y}(u',v') \right) \right. \\
&\quad \left. - \mathrm{E} \left(\mathbb{C}_x(u,v) \right) \mathrm{E} \left(\mathbb{C}_x(u',v') \right) \right\} \\
&= \Lambda'_{C_x^{Z \to Y}} \circ \Lambda'_{C_x^{Z \to Y}} \left[\mathrm{Cov} \left\{ \mathbb{C}_x^{Z \to Y}(u,v), \mathbb{C}_x^{Z \to Y}(u',v') \right\} \right] \\
&= \Lambda'_{C_x^{Z \to Y}} \circ \Lambda'_{C_x^{Z \to Y}} (\gamma_x).
\end{aligned}
$$

(ii) By the same arguments as for $\Theta_{\Lambda,xh}^{Z \to Y}$, one can write

$$\widetilde{\Theta}_{\Lambda,xh}^{Z \to Y} = \Lambda'_{C_x^{Z \to Y}} \left(\widetilde{\mathbb{C}}_{xh}^{Z \to Y} \right) + o_{\mathbb{P}}(1) \rightsquigarrow \widetilde{\Theta}_{\Lambda,x}^{Z \to Y},$$

where $\widetilde{\Theta}_{\Lambda,x}^{Z \to Y}$ is a normal random variable with variance $\sigma_{\Lambda,x}^2$ and mean

$$\widetilde{\mu}_{\Lambda,x} = \mathrm{E} \left\{ \Lambda'_{C_x^{Z \to Y}} \left(\widetilde{\mathbb{C}}_x^{Z \to Y} \right) \right\} = \Lambda'_{C_x^{Z \to Y}} \left\{ \mathrm{E} \left(\widetilde{\mathbb{C}}_x^{Z \to Y} \right) \right\} = \kappa \, a_1 \, \Lambda'_{C_x^{Z \to Y}} (\widetilde{\mathbb{B}}_x).$$

B.2. Proof of Proposition 4.1

Consider the version of L_{xi} given by

$$L_{xi}^*(u,v) = \mathbb{I} \left\{ Y_{i+1} \le F_{1xh}^{-1}(u), Z_i \le F_{2xh}^{-1}(v) \right\}$$

$$- C_x^{[1]}(u,v)\,\mathbb{I}\left\{Y_{i+1}\le F_{1xh}^{-1}(u)\right\}$$
$$- C_x^{[2]}(u,v)\,\mathbb{I}\left\{Z_i\le F_{2xh}^{-1}(v)\right\}, \tag{14}$$

and set $\lambda_{xh}^* = \sqrt{nh}\sum_{i=1}^n \omega_i(x,h)\,\lambda_{xi}^*$ with $\lambda_{xi}^* = \Lambda'_{C_x^{Z\to Y}}(L_{xi}^*)$. The condition in (12) on the estimators $\widehat{C}_x^{[1]}$ and $\widehat{C}_x^{[2]}$ of the partial derivatives of $C_x^{Z\to Y}$ allows to write

$$\frac{\widehat{\sigma}_{\Lambda,x}^2}{a_2/S_{n,0}(x)} = \sum_{i=1}^n \omega_i(x,h)\left(\lambda_{xi}^* - \frac{\lambda_{xh}^*}{\sqrt{nh}}\right)^2 + o_{\mathbb{P}}(1).$$

Next, introducing the linear functional $\Upsilon : l^\infty([0,1]^2) \to l^\infty([0,1]^2)$ defined for any $\delta \in l^\infty([0,1]^2)$ as $\Upsilon(\delta)(u,v) = \delta(u,v) - C_x^{[1]}(u,v)\,\delta(u,1) - C_x^{[2]}(u,v)\,\delta(1,v)$, and letting

$$\zeta_{xi}(u,v) = \mathbb{I}\left\{Y_{i+1}\le F_{1xh}^{-1}(u), Z_i\le F_{2xh}^{-1}(v)\right\},$$

one can write $L_{xi}^*(u,v) = \Upsilon(\zeta_{xi})$. Since one has for $\widetilde{\Upsilon} = \Lambda'_{C_x^{Z\to Y}}\circ\Upsilon$ that $\lambda_{xi}^* = \widetilde{\Upsilon}(\zeta_{xi})$, it follows that

$$
\begin{aligned}
\frac{\widehat{\sigma}_{\Lambda,x}^2}{a_2/S_{n,0}(x)} &= \sum_{i=1}^n \omega_i(x,h)\left\{\widetilde{\Upsilon}(\zeta_{xi}) - \sum_{j=1}^n \omega_j(x,h)\,\widetilde{\Upsilon}(\zeta_{xj})\right\}^2 + o_{\mathbb{P}}(1)\\
&= \sum_{i=1}^n \omega_i(x,h)\left\{\widetilde{\Upsilon}\left(\zeta_{xi} - \sum_{j=1}^n \omega_j(x,h)\,\zeta_{xj}\right)\right\}^2 + o_{\mathbb{P}}(1)\\
&= \sum_{i=1}^n \omega_i(x,h)\,\widetilde{\Upsilon}\circ\widetilde{\Upsilon}\left\{\left(\zeta_{xi} - \sum_{j=1}^n \omega_j(x,h)\,\zeta_{xj}\right)\right.\\
&\qquad\qquad\qquad\left.\times\left(\zeta_{xi} - \sum_{j=1}^n \omega_j(x,h)\,\zeta_{xj}\right)\right\} + o_{\mathbb{P}}(1)\\
&= \widetilde{\Upsilon}\circ\widetilde{\Upsilon}\left\{\sum_{i=1}^n \omega_i(x,h)\left(\zeta_{xi} - \sum_{j=1}^n \omega_j(x,h)\,\zeta_{xj}\right)\right.\\
&\qquad\qquad\qquad\left.\times\left(\zeta_{xi} - \sum_{j=1}^n \omega_j(x,h)\,\zeta_{xj}\right)\right\} + o_{\mathbb{P}}(1).
\end{aligned}
$$

Now observe that for the term inside the brackets,

$$
\begin{aligned}
&\sum_{i=1}^n \omega_i(x,h)\left(\zeta_{xi} - \sum_{j=1}^n \omega_j(x,h)\,\zeta_{xj}\right)\left(\zeta_{xi} - \sum_{j=1}^n \omega_j(x,h)\,\zeta_{xj}\right)\\
&\qquad = \sum_{i=1}^n \omega_i(x,h)\,\zeta_{xi}(u,v)\,\zeta_{xi}(u',v')
\end{aligned}
$$

$$- \left\{ \sum_{i=1}^{n} \omega_i(x,h)\,\zeta_{xi}(u,v) \right\} \left\{ \sum_{j=1}^{n} \omega_j(x,h)\,\zeta_{xj}(u',v') \right\}$$

$$= \quad C_{xh}^{\mathrm{Z}\to\mathrm{Y}}\left\{\min(u,u'),\min(v,v')\right\} - C_{xh}^{\mathrm{Z}\to\mathrm{Y}}(u,v)\,C_{xh}^{\mathrm{Z}\to\mathrm{Y}}(u',v').$$

Since $C_{xh}^{\mathrm{Z}\to\mathrm{Y}}$ is consistent for $C_x^{\mathrm{Z}\to\mathrm{Y}}$, an application of the Continuous Mapping Theorem entails that as $n \to \infty$

$$\frac{\widehat{\sigma}_{\Lambda,x}^2}{a_2/S_{n,0}(x)}$$

$$= \widetilde{\Upsilon} \circ \widetilde{\Upsilon}\left\{ C_x^{\mathrm{Z}\to\mathrm{Y}}\left\{\min(u,u'),\min(v,v')\right\} - C_x^{\mathrm{Z}\to\mathrm{Y}}(u,v)\,C_x^{\mathrm{Z}\to\mathrm{Y}}(u',v')\right\} + o_{\mathbb{P}}(1)$$

$$= \Lambda'_{C_x^{\mathrm{Z}\to\mathrm{Y}}} \circ \Lambda'_{C_x^{\mathrm{Z}\to\mathrm{Y}}} \circ \Upsilon \circ \Upsilon \left\{ C_x^{\mathrm{Z}\to\mathrm{Y}}\left\{\min(u,u'),\min(v,v')\right\} \right.$$

$$\left. - C_x^{\mathrm{Z}\to\mathrm{Y}}(u,v)\,C_x^{\mathrm{Z}\to\mathrm{Y}}(u',v')\right\} + o_{\mathbb{P}}(1)$$

$$= \Lambda'_{C_x^{\mathrm{Z}\to\mathrm{Y}}} \circ \Lambda'_{C_x^{\mathrm{Z}\to\mathrm{Y}}} \left[\frac{f_Y(x)}{a_2}\,\mathrm{Cov}\left\{\mathbb{C}_x^{\mathrm{Z}\to\mathrm{Y}}(u,v), \mathbb{C}_x^{\mathrm{Z}\to\mathrm{Y}}(u',v')\right\} \right] + o_{\mathbb{P}}(1)$$

$$= \frac{f_Y(x)}{a_2}\,\sigma_{\Lambda,x}^2 + o_{\mathbb{P}}(1).$$

Using the fact that $S_{n,0}(x) = f_Y(x) + o_{\mathrm{a.s.}}(1)$, see for instance [20],

$$\widehat{\sigma}_{\Lambda,x}^2 = \frac{f_Y(x)}{S_{n,0}(x)}\,\sigma_{\Lambda,x}^2 + o_{\mathbb{P}}(1) \to \sigma_{\Lambda,x}^2.$$

B.3. Proof of Proposition 4.2

For $(u,v) \in [0,1]^2$, define the vectors $\mathbf{L}_x^\star(u,v)$ and $\mathbf{L}_x'(u,v)$ such that

$$\mathbf{L}_x^\star = (L_{x1}^\star, \ldots, L_{xn}^\star)^\top \quad \text{and} \quad \mathbf{L}_x' = (L_{x1}^*, \ldots, L_{xn}^*)^\top,$$

where L_{xh}^* is as in Equation (14). Also define the diagonal matrix $\mathbf{W}$ with i-th element $h_B^{-1} K\{(Y_i - x)/h_B\}$, as well as the matrix $\mathbf{X} \in \mathbb{R}^{n \times (p+1)}$ whose entries are given by

$$\mathbf{X}_{ij} = \left(\frac{Y_i - x}{h_B} \right)^{j-1}.$$

From p. 59 in [8], one has for $[A]_3$ being the third row of matrix A that

$$\widehat{\beta}_{2,\Lambda} = \left[\left(\mathbf{X}^\top \mathbf{W} \mathbf{X} \right)^{-1} \right]_3 \mathbf{X}^\top \mathbf{W} \left\{ \Lambda'_{C_x^{Y \to Z}}\left(\mathbf{L}_x^\star \right) \right\},$$

where it is understood that a functional applied to a vector is taken component-wise. As a primary step, consider

$$\widehat{\beta}_{2,\Lambda}' = \left[\left(\mathbf{X}^\top \mathbf{W} \mathbf{X} \right)^{-1} \right]_3 \mathbf{X}^\top \mathbf{W} \left\{ \Lambda'_{C_x^{Y \to Z}}\left(\mathbf{L}_x' \right) \right\}.$$

The following Lemma is helpful to derive the limiting behaviour of $\widehat{\beta}_{2,\Lambda}'$. In the sequel, let $\mathbf{Y}(y,z) = (\mathbb{I}(Y_2 \le y, Z_1 \le z), \ldots, \mathbb{I}(Y_{n+1} \le y, Z_n \le z))^\top$.

Lemma B.1. *Under Assumptions* $(\mathcal{L})$, $(\mathcal{N})$, $(\mathcal{S})$, $(\mathcal{H}_B)$ *and* $(\mathcal{N}_B)$,

$$\sup_{y,z\in\mathbb{R}} \left| \left[\left(\mathbf{X}^\top \mathbf{W} \mathbf{X} \right)^{-1} \right]_3 \mathbf{X}^\top \mathbf{W} \mathbf{Y}(y,z) - \frac{1}{2}\, h_B^2\, \ddot{H}_x(y,z) \right| = o_{\mathbb{P}}(h_B^2).$$

Invoking Lemma B.1, one obtains that

$$\sup_{(u,v)\in[0,1]^2} \left| \left[\left(\mathbf{X}^\top \mathbf{W} \mathbf{X} \right)^{-1} \right]_3 \mathbf{X}^\top \mathbf{W} \mathbf{Y} \left\{ F_{1xh}^{-1}(u), F_{2xh}^{-1}(v) \right\} \right.$$

$$\left. - \frac{h_B^2}{2} \ddot{H}_x \left\{ F_{1x}^{-1}(u), F_{2x}^{-1}(v) \right\} \right|$$

$$\leq \sup_{(y,z)\in\mathbb{R}^2} \left| \left[\left(\mathbf{X}^\top \mathbf{W} \mathbf{X} \right)^{-1} \right]_3 \mathbf{X}^\top \mathbf{W} \mathbf{Y}(y,z) - \frac{h_B^2}{2} \ddot{H}_x(y,z) \right|$$

$$+ \frac{h_B^2}{2} \sup_{(u,v)\in[0,1]^2} \left| \ddot{H}_x \left\{ F_{1xh}^{-1}(u), F_{2xh}^{-1}(v) \right\} - \ddot{H}_x \left\{ F_{1x}^{-1}(u), F_{2x}^{-1}(v) \right\} \right|$$

$$= o_{\mathbb{P}}(h_B^2),$$

where the last equality follows from the uniform continuity of $\ddot{H}_x$ and the fact that it can be deduced from Proposition 1 of [5] that $F_{kxh}^{-1} - F_{kx}^{-1} = o_{\mathbb{P}}(1)$. Therefore, since $\Lambda'_{C_x^{Y\to Z}}$ is a linear functional, the continuous mapping theorem implies

$$\left| \widehat{\beta}'_{2,\Lambda} - \frac{h_B^2}{2} \Lambda'_{C_x} \left(\Upsilon \left[\ddot{H}_x \{ F_{1x}^{-1}(\cdot), F_{2x}^{-1}(\cdot) \} \right] \right) \right| = o_{\mathbb{P}}(h_B^2). \tag{15}$$

The assumption on the estimators of the partial derivatives in (12) entails that $|\widehat{\beta}_{2,\Lambda} - \widehat{\beta}'_{2,\Lambda}| = o_{\mathbb{P}}(h_B^2)$. Now upon noting that

$$\mu_{\Lambda,x} = \kappa\, a_1\, \Lambda'_{C_x} \left(\Upsilon \left[\ddot{H}_x \{ F_{1x}^{-1}(\cdot), F_{2x}^{-1}(\cdot) \} \right] \right),$$

it follows that as $n \to \infty$,

$$\widehat{\mu}_{\Lambda,x} - \mu_{\Lambda,x} = \sqrt{nh^5}\, a_1 \left\{ \frac{\widehat{\beta}_{2,\Lambda}}{h_B^2} - \Lambda'_{C_x} \left(\Upsilon \left[\ddot{H}_x \{ F_{1x}^{-1}(\cdot), F_{2x}^{-1}(\cdot) \} \right] \right) \right\}$$

$$- \left(\sqrt{nh^5} - \kappa \right) a_1\, \Lambda'_{C_x} \left(\Upsilon \left[\ddot{H}_x \{ F_{1x}^{-1}(\cdot), F_{2x}^{-1}(\cdot) \} \right] \right)$$

$$= o_{\mathbb{P}}(1),$$

since Assumption $(\mathcal{N})$ implies $\sqrt{nh^5} \to \kappa$ as $n \to \infty$.

Appendix C: Proof of two lemmas

This section is devoted to the proof of Lemma B.1; the latter is used in order to establish Proposition 4.2. To this end, the following lemma will be invoked.

Lemma C.1. *Let* $\widetilde{Z}_{xn}^{(j)} = Z_{xn}^{(j)} - \mathrm{E}\, Z_{xn}^{(j)}$, *where for* $(y, z) \in \mathbb{R}^2$,

$$Z_{xn}^{(j)}(A) = \frac{1}{nh_B} \sum_{i=1}^{n} \left(\frac{Y_i - x}{h_B}\right)^j K\left(\frac{Y_i - x}{h_B}\right) \mathbb{I}\left(Y_{i+1} \leq y, Z_i \leq z\right).$$

Then under Assumptions $(\mathcal{L})$, $(\mathcal{N})$, $(\mathcal{S})$, $(\mathcal{H}_B)$ *and* $(\mathcal{N}_B)$, *the process* $\sqrt{nh_B}\, \widetilde{Z}_{xn}^{(j)}$ *is asymptotically tight.*

C.1. Proof of Lemma B.1

For any $y, z \in \mathbb{R}$, let

$$
\begin{aligned}
\left(\widehat{\beta}_0(y,z), \ldots, \widehat{\beta}_p(y,z)\right)^\top &= \left(\mathbf{X}^\top \mathbf{W} \mathbf{X}\right)^{-1} \mathbf{X}^\top \mathbf{W} \mathbf{Y}(y,z) \\
&= \left(\frac{1}{n} \mathbf{X}^\top \mathbf{W} \mathbf{X}\right)^{-1} \left(\frac{1}{n} \mathbf{X}^\top \mathbf{W} \mathbf{Y}(y,z)\right), \quad (16)
\end{aligned}
$$

so that

$$\left[\left(\mathbf{X}^\top \mathbf{W} \mathbf{X}\right)^{-1} \mathbf{X}^\top\right]_3 \mathbf{W} \mathbf{Y}(y,z) = \widehat{\beta}_2(y,z).$$

Observe that the (i,j)-th element of the matrix $\mathbf{X}^\top \mathbf{W} \mathbf{X}/n$ is $S_{n,i+j-2}(x)$. Also, from Corollary 1 of [20], one deduces that under Assumptions $(\mathcal{S})$ and $(\mathcal{L})$,

$$S_{n,j}(x) = f_Y(x) \int u^j K(u)\, \mathrm{d}u + o_{\mathrm{as}}(1).$$

Therefore, an application of the continuous mapping Theorem entails

$$\left(\frac{1}{n} \mathbf{X}^\top \mathbf{W} \mathbf{X}\right)^{-1} = \frac{1}{f_Y(x)} \mathbf{S}^{-1} + o_\mathbb{P}(1), \quad (17)$$

where $\mathbf{S} \in \mathbb{R}^{(p+1)\times(p+1)}$ is such that $\mathbf{S}_{ij} = \int u^{j+i-2} K(u)\, \mathrm{d}u$. Also note that the j-th element of $\mathbf{X}^\top \mathbf{W} \mathbf{Y}(y,z)/n$ is $Z_{xn}^{(j)}(y,z)$ and that from Lemma C.1 $Z_{xn}^{(j)} = \mathrm{E}\, Z_{xn}^{(j)} + O_\mathbb{P}(1/\sqrt{nh_B})$. Moreover, expanding H_w in a Taylor series of order p around $w = x$ yields, for all $|w - x| \leq h_B$,

$$H_w(y,z) = \sum_{k=0}^{p} \left\{\frac{\partial^k}{\partial x^k} H_x(y,z)\right\} \frac{(w-x)^k}{k!} + o(h_B^p).$$

As a consequence,

$$\mathrm{E}\left\{Z_{xn}^{(j)}(y,z)\right\}$$

$$= \mathrm{E}\left\{\left(\frac{Y_1 - x}{h_B}\right)^j K\left(\frac{Y_1 - x}{h_B}\right) H_{Y_1}(y,z)\right\}$$

$$
\begin{aligned}
&= o(h_B^p) + \sum_{k=0}^{p} \frac{h_B^k}{k!} \left\{ \frac{\partial^k}{\partial x^k} H_x(y,z) \right\} \mathrm{E} \left\{ \left(\frac{Y_1 - x}{h_B} \right)^{j+k} K \left(\frac{Y_1 - x}{h_B} \right) \right\} \\
&= o(h_B^p) + \sum_{k=0}^{p} \frac{h_B^k}{k!} \left\{ \frac{\partial^k}{\partial x^k} H_x(y,z) \right\} \left\{ f_Y(x) \int u^{j+k} K(u) \, \mathrm{d}u + o(1) \right\} \\
&= o(h_B^p) + \{ f_Y(x) \mathbf{S}_j + o(1) \} \\
&\qquad \times \left(\frac{h_B^0}{0!} \frac{\partial^0}{\partial x^0} H_x(y,z), \frac{h_B^1}{1!} \frac{\partial^1}{\partial x^1} H_x(y,z), \dots, \frac{h_B^p}{p!} \frac{\partial^p}{\partial x^p} H_x(y,z) \right)^{\top},
\end{aligned}
$$

where $\mathbf{S}_j$ is the $(j+1)$-th line of $\mathbf{S}$. Thus, uniformly in $(y,z) \in \mathbb{R}^2$,

$$
\begin{aligned}
\frac{1}{n} & \mathbf{X}^{\top} \mathbf{W} \mathbf{Y}(y,z) \\
&= \{ f_Y(x) \mathbf{S} + o(1) \} \\
&\qquad \times \left(\frac{h_B^0}{0!} \frac{\partial^0}{\partial x^0} H_x(y,z), \frac{h_B^1}{1!} \frac{\partial^1}{\partial x^1} H_x(y,z), \dots, \frac{h_B^p}{p!} \frac{\partial^p}{\partial x^p} H_x(y,z) \right)^{\top} \\
&\quad + o(h_B^p) + O_{\mathbb{P}} \left(\frac{1}{\sqrt{nh_B}} \right).
\end{aligned} \tag{18}
$$

Letting $I_{p+1} \in \mathbb{R}^{(p+1) \times (p+1)}$ be the identity matrix, one has combining (16), (17) and (18) that

$$
\begin{aligned}
\left(\widehat{\beta}_0(y,z), \dots, \widehat{\beta}_p(y,z) \right)^{\top} & \\
&= \{ I_{p+1} + o(1) \} \\
&\qquad \times \left(\frac{h_B^0}{0!} \frac{\partial^0}{\partial x^0} H_x(y,z), \frac{h_B^1}{1!} \frac{\partial^1}{\partial x^1} H_x(y,z), \dots, \frac{h_B^p}{p!} \frac{\partial^p}{\partial x^p} H_x(y,z) \right)^{\top} \\
&\quad + o(h_B^p) + O_{\mathbb{P}} \left(\frac{1}{\sqrt{nh_B}} \right).
\end{aligned}
$$

Since $1/\sqrt{nh_B} = o(h_B^2)$ (see Condition $(\mathcal{N}_B)$) and $p \geq 2$, the announced result follows.

C.2. Proof of Lemma C.1

For $u, v \in [0,1]$, consider the versions of $Z_{kxn}^{(j)}$ and $\widetilde{Z}_{kxn}^{(j)}$ given by

$$
\mathcal{U}_{xn}^{(j)}(u,v) = \sqrt{nh_B} Z_{xn}^{(j)} \left\{ F_{1x}^{-1}(u), F_{2x}^{-1}(v) \right\} \quad \text{and} \quad \widetilde{\mathcal{U}}_{xn}^{(j)} = \mathcal{U}_{xn}^{(j)} - \mathrm{E} \mathcal{U}_{xh}^{(j)}.
$$

Also, for $A = [u_1, u_2] \times [v_1, v_2] \subseteq [0,1]^2$ let

$$
\mathcal{U}_{xn}^{(j)}(A) = \mathcal{U}_{xn}^{(j)}(u_2, v_2) - \mathcal{U}_{xn}^{(j)}(u_1, v_2) - \mathcal{U}_{xn}^{(j)}(u_2, v_1) + \mathcal{U}_{xn}^{(j)}(u_1, v_1),
$$

and define $\widetilde{\mathcal{U}}_{xn}^{(j)}(A)$ similarly. From Theorem 1.5.7 of [25], the tightness of $\widetilde{\mathcal{U}}_{xn}^{(j)}$ will entail that of $\sqrt{nh_B}\,\widetilde{Z}_{xn}^{(j)}$. Now to prove that $\widetilde{\mathcal{U}}_{xn}^{(j)}$ is asymptotically tight in $l^\infty([0,1]^2)$, it suffices to show (see again Theorem 1.5.7 of [25]) that for any fixed $\mathbf{u} \in [0,1]^2$ the random variable $\widetilde{\mathcal{U}}_{xn}^{(j)}(\mathbf{u})$ is asymptotically tight in $\mathbb{R}$ and that for any $\epsilon > 0$,

$$\lim_{\delta \to 0} \lim_{n \to \infty} \mathbb{P} \left\{ \sup_{\substack{\mathbf{u},\mathbf{v}\in[0,1]^2 \\ |\mathbf{u}-\mathbf{v}|<\delta}} \left| \widetilde{\mathcal{U}}_{xn}^{(j)}(\mathbf{u}) - \widetilde{\mathcal{U}}_{xn}^{(j)}(\mathbf{v}) \right| > \epsilon \right\} = 0, \tag{19}$$

where it is understood that $|\cdot|$ is the Manhattan distance. First, for any fixed $\mathbf{u} \in [0,1]^2$, a consequence of Theorem 6 in [19] is that $\widetilde{\mathcal{U}}_{xn}^{(j)}(\mathbf{u})$ is asymptotically normal, and thus asymptotically tight in $\mathbb{R}$. It then remains to show that (19) holds. To this end, let $\kappa_\gamma = \lfloor (nh_B)^{1/2+\gamma} \rfloor$ for some $\gamma \in (0,1/2)$ and define the product space $T_{\kappa_\gamma} = I_{\kappa_\gamma} \times I_{\kappa_\gamma}$, where $I_{\kappa_\gamma} = \{0, 1/\kappa_\gamma, 2/\kappa_\gamma, \ldots, 1\}$. Next, for any $u \in [0,1]$, define $\underline{u}_{\kappa_\gamma} = \max\{\zeta \in I_{\kappa_\gamma} : \zeta \le u\}$ and $\overline{u}_{\kappa_\gamma} = \min\{\zeta \in I_{\kappa_\gamma} : \zeta > u\}$, and for any $\mathbf{u} = (u_1, u_2) \in [0,1]^2$, let $\underline{\mathbf{u}}_{\kappa_\gamma} = (\underline{u}_{1\kappa_\gamma}, \underline{u}_{2\kappa_\gamma})$ and $\overline{\mathbf{u}}_{\kappa_\gamma} = (\overline{u}_{1\kappa_\gamma}, \overline{u}_{2\kappa_\gamma})$. As under Assumption $(\mathcal{L})$ K is positive and vanishes outside of $[-1,1]$, one has for any $z \in \mathbb{R}$ that $|K(z)z^j| \le K(z)$. One can then write

$$\begin{aligned}
\left| \mathcal{U}_{xn}^{(j)}(\mathbf{u}) - \mathcal{U}_{xn}^{(j)}(\underline{\mathbf{u}}_{\kappa_\gamma}) \right| &\le \frac{1}{\sqrt{nh_B}} \sum_{i=1}^n K\left(\frac{Y_i - x}{h_B} \right) \\
&\quad \times \Big[\mathbb{I}\left\{ Y_{i+1} \le F_{1x}^{-1}(u_1), Z_i \le F_{2x}^{-1}(u_2) \right\} \\
&\qquad - \mathbb{I}\left\{ Y_{i+1} \le F_{1x}^{-1}(\underline{u}_{1\kappa_\gamma}), Z_i \le F_{2x}^{-1}(\underline{u}_{2\kappa_\gamma}) \right\} \Big] \\
&= \mathcal{U}_{xn}^{(0)}(\overline{\mathbf{u}}_{\kappa_\gamma}) - \mathcal{U}_{xn}^{(0)}(\underline{\mathbf{u}}_{\kappa_\gamma}).
\end{aligned}$$

Therefore,

$$\begin{aligned}
\left| \widetilde{\mathcal{U}}_{xn}^{(j)}(\mathbf{u}) - \widetilde{\mathcal{U}}_{xn}^{(j)}(\underline{\mathbf{u}}_{\kappa_\gamma}) \right| &\le \left| \widetilde{\mathcal{U}}_{xn}^{(0)}(\overline{\mathbf{u}}_{\kappa_\gamma}) - \widetilde{\mathcal{U}}_{xn}^{(0)}(\underline{\mathbf{u}}_{\kappa_\gamma}) \right| \\
&\quad + 2 \left| \mathrm{E}\left\{ \widetilde{\mathcal{U}}_{xn}^{(0)}(\overline{\mathbf{u}}_{\kappa_\gamma}) - \widetilde{\mathcal{U}}_{xn}^{(0)}(\underline{\mathbf{u}}_{\kappa_\gamma}) \right\} \right|. \tag{20}
\end{aligned}$$

In order to bound the second term on the right hand side of (20), one has in view of Assumption $(\mathcal{H}_B)$ that a Taylor expansion of order $p+1$ allows to write, uniformly in $y \in \mathbb{R}$,

$$\begin{aligned}
F_{kw}(y) &= F_{kx}(y) + \sum_{j=1}^{p+1} \left\{ \frac{\partial^j}{\partial x^j} F_{kx}(y) \right\} \frac{(w-x)^j}{j!} \\
&\quad + \left\{ \frac{\partial^{p+2}}{\partial x^{p+2}} F_{kx}(y) \Big|_{w=\xi_{w,x}} \right\} \frac{(w-x)^{p+2}}{(p+2)!},
\end{aligned}$$

where $\xi_{w,x}$ lies between w and x for any $w \in J_x$. Under Assumption $(\mathcal{H}_B)$, there exists $\eta > 0$ such that for $j \in \{1, \ldots, p+1\}$,

$$\left| \left\{ \frac{\partial^j}{\partial x^j} F_{kx}(y) \right\} - \left\{ \frac{\partial^j}{\partial x^j} F_{kx}(y') \right\} \right| \le \eta \left| F_{kx}(y) - F_{kx}(y') \right|$$

and

$$\sup_{w \in J_x} \sup_{y \in \mathbb{R}} \left| \frac{\partial^{p+2}}{\partial w^{p+2}} F_{kw}(y) \right| < \eta.$$

As a consequence, for h_B taken sufficiently small, one deduces that for any $y \leq y'$, $w \in [x - h_B, x + h_B]$ and $k \in \{1, 2\}$,

$$|F_{kw}(y) - F_{kw}(y')| \leq 4\eta \{F_{kx}(y) - F_{kx}(y')\} + \eta h_B^{p+2}. \tag{21}$$

Also, from Assumption $(\mathcal{L})$ and routine computations, one has for $n \to \infty$ that

$$\mathrm{E}\left\{ K\left(\frac{Y_1 - x}{h_B} \right) \right\} = \{1 + o(1)\} h_B f_Y(x) \leq 2 h_B f_Y(x). \tag{22}$$

Hence, Hoeffding's inequality combined with (21) and (22) yields

$$\left| \mathrm{E}\left\{ \widetilde{\mathcal{U}}_{xn}^{(0)}(\overline{\mathbf{u}}_{\kappa_\gamma}) - \widetilde{\mathcal{U}}_{xn}^{(0)}(\underline{\mathbf{u}}_{\kappa_\gamma}) \right\} \right|$$

$$= \sqrt{\frac{n}{h_B}} \left| \mathrm{E}\left[\left(\frac{Y_1 - x}{h_B} \right)^j K\left(\frac{Y_1 - x}{h_B} \right) \right. \right.$$

$$\left. \left. \times \left\{ H_{Y_1}\left(F_{1x}^{-1}(\overline{u}_{1\kappa_\gamma}), F_{2x}^{-1}(\underline{u}_{2\kappa_\gamma}) \right) - H_{Y_1}\left(F_{1x}^{-1}(\overline{v}_{1\kappa_\gamma}), F_{2x}^{-1}(\underline{v}_{2\kappa_\gamma}) \right) \right\} \right] \right|$$

$$\leq 2 \sqrt{\frac{n}{h_B}} \max_{k \in \{1,2\}} \mathrm{E}\left[K\left(\frac{Y_1 - x}{h_B} \right) \left\{ F_{kY_1} \circ F_{kx}^{-1}(\overline{u}_{k\kappa_\gamma}) - F_{kY_1} \circ F_{kx}^{-1}(\overline{v}_{k\kappa_\gamma}) \right\} \right]$$

$$\leq 16 \eta f_Y(x) \times \sqrt{nh_B} \left(\frac{1}{\kappa_\gamma} + h_B^{p+2} \right)$$

$$= o(1), \tag{23}$$

where the last equality is a consequence of the definition of κ_γ and Assumption $(\mathcal{N}_B)$. Hence, from (20) and (23),

$$\left| \widetilde{\mathcal{U}}_{xn}^{(j)}(\mathbf{u}) - \widetilde{\mathcal{U}}_{xn}^{(j)}(\mathbf{v}) \right| \leq \left| \widetilde{\mathcal{U}}_{xn}^{(0)}(\overline{\mathbf{u}}_{\kappa_\gamma}) - \widetilde{\mathcal{U}}_{xn}^{(0)}(\underline{\mathbf{u}}_{\kappa_\gamma}) \right| + \left| \widetilde{\mathcal{U}}_{xn}^{(0)}(\overline{\mathbf{v}}_{\kappa_\gamma}) - \widetilde{\mathcal{U}}_{xn}^{(0)}(\underline{\mathbf{v}}_{\kappa_\gamma}) \right|$$

$$+ \left| \widetilde{\mathcal{U}}_{xn}^{(j)}(\underline{\mathbf{u}}_{\kappa_\gamma}) - \widetilde{\mathcal{U}}_{xn}^{(j)}(\underline{\mathbf{v}}_{\kappa_\gamma}) \right| + o(1).$$

Since for $\delta > 2\kappa_\gamma$, $|\mathbf{u} - \mathbf{v}| < \delta$ implies $|\underline{\mathbf{u}}_{\kappa_\gamma} - \underline{\mathbf{v}}_{\kappa_\gamma}| < 2\delta$,

$$\sup_{\substack{\mathbf{u}, \mathbf{v}' \in [0,1]^2; \\ |\mathbf{u} - \mathbf{v}| < \delta}} \left| \widetilde{\mathcal{U}}_{xn}^{(j)}(\mathbf{u}) - \widetilde{\mathcal{U}}_{xn}^{(j)}(\mathbf{v}) \right| \leq 3 \max_{\ell \in \{1,2\}} \max_{\substack{\mathbf{u}, \mathbf{v}' \in T_{\kappa_\gamma}; \\ |\mathbf{u} - \mathbf{v}| < 2\delta}} \left| \widetilde{\mathcal{U}}_{xn}^{(\ell)}(\mathbf{u}) - \widetilde{\mathcal{U}}_{xn}^{(\ell)}(\mathbf{v}) \right| + o(1).$$

From there, Equation (19) will be established ift for any $\epsilon > 0$ and $\ell \in \{1, 2\}$,

$$\lim_{\delta \to 0} \lim_{n \to \infty} \mathbb{P}\left\{ \max_{\substack{\mathbf{u}, \mathbf{v}' \in T_{\kappa_\gamma}; \\ |\mathbf{u} - \mathbf{v}| < \delta}} \left| \widetilde{\mathcal{U}}_{xn}^{(\ell)}(\mathbf{u}) - \widetilde{\mathcal{U}}_{xn}^{(\ell)}(\mathbf{v}) \right| > \epsilon \right\} = 0.$$

To this end, Lemma 2 of [1] will be used (see also Theorem 3 and the remarks on page 1665 in [3]). The first step is to show that there exists $\eta^\star > 0$ and $\beta > 1$ such that for any $\epsilon > 0$ and n sufficiently large, it holds that for any rectangle $A \subseteq [0,1]^2$ whose corner points are all distinct and lie in T_{κ_γ},

$$\mathbb{P}\left\{\left|\widetilde{\mathcal{U}}_{xn}^{(j)}(A)\right| > \epsilon\right\} \leq \frac{\eta^\star}{\epsilon^4}\left\{\widetilde{\chi}_{h_B}(A)\right\}^\beta, \tag{24}$$

where $\widetilde{\chi}_{h_B}(A) = \mu(A) + \chi_{h_B}(A)$, with μ being Lebesgue's measure and

$$\chi_{h_B}(A) = \frac{1}{h_B}\,\mathrm{E}\left[\mathbb{I}\left\{(F_{1x}(Y_2), F_{2x}(Z_1)) \in A\right\} K\left(\frac{Y_1 - x}{h_B}\right)\right].$$

From the proof of Lemma 2 in [5] (see Equation (12) therein), one has under Assumptions $(\mathcal{S})$, $(\mathcal{L})$ and $(\mathcal{H})$ that there exists $\overline{\eta} > 0$ such that for any $A \subseteq [0,1]^2$ and n sufficiently large,

$$\mathrm{E}\left\{\left|\widetilde{\mathcal{U}}_{xn}^{(0)}(A)\right|^6\right\} \leq \overline{\eta}\left\{\frac{\chi_{h_B}(A) + h^{\frac{5a-30}{a}}}{n^2 h_B^2} + \frac{h_B^{\frac{a-6}{a}} + \chi_{h_B}(A)^{2-\frac{4}{a}}}{nh_B} + \chi_{h_B}(A)^{3-\frac{6}{a}}\right\}.$$

A close inspection of their arguments shows that the latter equation remains valid when $\widetilde{\mathcal{U}}_{xn}^{(0)}$ is replaced by $\widetilde{\mathcal{U}}_{xn}^{(j)}$. Since $\chi_{h_B}(A) \leq \widetilde{\chi}_{h_B}(A)$, one can conclude that for n sufficiently large and $\ell \in \{0, j\}$,

$$\mathrm{E}\left\{\left|\widetilde{\mathcal{U}}_{xn}^{(\ell)}(A)\right|^6\right\} \leq \overline{\eta}\left\{\frac{\widetilde{\chi}_{h_B}(A) + h_B^{\frac{5a-30}{a}}}{n^2 h_B^2} + \frac{h_B^{\frac{a-6}{a}} + \widetilde{\chi}_{h_B}(A)^{2-\frac{4}{a}}}{nh_B} + \widetilde{\chi}_{h_B}(A)^{3-\frac{6}{a}}\right\}$$

$$= \overline{\eta}\left\{\widetilde{\chi}_{h_B}(A)\right\}^\beta\left\{R_n^{(1)}(\beta, A) + R_n^{(2)}(\beta, A)\right\}, \tag{25}$$

where

$$R_n^{(1)}(\beta, A) = \frac{\{\widetilde{\chi}_{h_B}(A)\}^{1-\beta}}{(nh_B)^2} + \frac{\{\widetilde{\chi}_{h_B}(A)\}^{2-\frac{4}{a}-\beta}}{nh_B} + \chi_{h_B}(A)^{3-\frac{6}{a}-\beta}$$

and

$$R_n^{(2)}(\beta, A) = \{\widetilde{\chi}_{h_B}(A)\}^{-\beta}\left(\frac{h_B^{\frac{5a-30}{a}}}{n^2 h_B^2} + \frac{h_B^{\frac{a-6}{a}}}{nh_B}\right).$$

From the definition of μ and computations similar as those used for the derivation of (23), one can show that for n sufficiently large and any $A \subseteq [0,1]^2$ whose corner points are all distinct and lie in T_{κ_γ},

$$\frac{1}{\kappa_\gamma^2} \leq \mu(A) \leq \widetilde{\chi}_{h_B}(A) \leq 16\, f_Y(x). \tag{26}$$

Because $\kappa_\gamma = (nh_B)^{1/2+\gamma}$, one has from (26) and the fact that Assumption $(\mathcal{N}_B)$ ensures that $nh_B > 1$ when n is large, that for any $\beta \in (1, 1 + a^{-1})$,

$$R_n^{(1)}(\beta, A) \leq (nh_B)^{(1-\beta)(1+2\gamma)-2} + \{16 f_Y(x)\}^{1-\frac{5}{a}} + \{16 f_Y(x)\}^{2-\frac{7}{a}}.$$

Since $a > 6$, taking any $\gamma \in (0, 1/2)$ entails that for n sufficiently large,

$$R_n^{(1)}(\beta, A) \leq 1 + \{16 f_Y(x)\}^{1-\frac{5}{a}} + \{16 f_Y(x)\}^{2-\frac{7}{a}}. \tag{27}$$

Next, one obtains from the left hand side of (26) that

$$R_n^{(2)}(\beta, A) \leq (n h_B)^{(1+2\gamma)\beta-2} \, h_B^{\frac{5a-30}{a}} + (n h_B)^{(1+2\gamma)\beta-1} h_B^{\frac{a-6}{a}}.$$

Since Assumption $(\mathcal{N}_B)$ ensures that $n h_B^{2p+5} \to 0$ as $n \to \infty$, it follows that for n large enough, $n \leq h_B^{-2p-5}$. Therefore,

$$R_n^{(2)}(\beta, A) \leq h_B^{-(2p+4)\{(1+2\gamma)\beta-2\}+\frac{5a-30}{a}} + h_B^{-(2p+4)\{(1+2\gamma)\beta-1\}+\frac{a-6}{a}}.$$

For the first term on the right-hand side of the previous inequality, observe that as $a > 6$, taking $\gamma \in (0, 1/8)$ and $\beta \in (1, 1 + a^{-1})$ yields $(1 + 2\gamma)\beta < 2$; this implies that for n sufficiently large,

$$h_B^{-(2p+4)\{1+2\gamma)\beta-2\}} < 1.$$

For the second term, first choose γ, β such that

$$0 < \gamma < \frac{a-6}{8a(2p+4)} \quad \text{and} \quad 1 < \beta < \frac{a-6}{8a(2p+4)}.$$

This ensures that

$$(2p+4)\{(1+2\gamma)\beta - 1\} < \frac{a-6}{a},$$

and this finally entails

$$h_B^{-(2p+4)\{(1+2\gamma)\beta-1\}+\frac{a-6}{a}} < 1.$$

Hence, $R_n^{(2)}(\beta, A) \leq 2$. One can then invoke (25) and (27) in order to deduce that there exists $\beta > 1$ and $\eta' > 0$ such that for n taken sufficiently large,

$$\mathrm{E}\left\{ \left| \widetilde{\mathcal{U}}_{xn}^{(j)}(A) \right|^6 \right\} \leq \eta' \{\widetilde{\chi}_{h_B}(A)\}^\beta.$$

Equation (24) then follows from an application of the Markov inequality. Since $\widetilde{\chi}_{h_B}$ is a finite measure on $[0, 1]^2$, one uses Lemma 2 of [1] to conclude from similar computations as those presented at the end of the proof of Proposition 1 in [5] that the first term on the right hand side of (24) is asymptotically negligible as $\delta \to 0$. This in turn entails (19), so the proof is complete.

Acknowledgments

References

[1] BALACHEFF, S. and DUPONT, G. (1980). Normalite asymptotique des processus empiriques tronques et des processus de rang. *Lecture Notes in Mathematics* 19–45. MR0604019

[2] BEARE, B. K. and SEO, J. (2015). Vine copula specifications for stationary multivariate Markov chains. *J. Time Series Anal.* **36** 228–246. MR3316468

[3] BICKEL, P. J. and WICHURA, M. J. (1971). Convergence criteria for multiparameter stochastic processes and some applications. *The Annals of Mathematical Statistics* 1656–1670. MR0383482

[4] BOUDJELLABA, H., DUFOUR, J.-M. and ROY, R. (1992). Testing causality between two vectors in multivariate autoregressive moving average models. *J. Amer. Statist. Assoc.* **87** 1082–1090. MR1209566

[5] BOUEZMARNI, T., CAMIRAND LEMYRE, F. and QUESSY, J.-F. (2019). On the large-sample behavior of two estimators of the conditional copula under serially dependent data. *Metrika* **82** 823–841. MR4008664

[6] BOUEZMARNI, T., ROMBOUTS, J. V. K. and TAAMOUTI, A. (2012). Nonparametric copula-based test for conditional independence with applications to Granger causality. *J. Bus. Econom. Statist.* **30** 275–287. MR2934942

[7] DUFOUR, J. M. and TAAMOUTI, A. (2010). Short and long run causality measures: theory and inference. *Journal of Econometrics* **154** 42–58. MR2558950

[8] FAN, J. and GIJBELS, I. (1996). *Local polynomial modelling and its applications. Monographs on Statistics and Applied Probability* **66**. Boca Raton, Fla.: Chapman & Hall/CRC, 1996. MR1383587

[9] FLORENS, J.-P. and FOUGERE, D. (1996). Noncausality in continuous time. *Econometrica* **64** 1195–1212. MR1403234 (97b:60089)

[10] GASSER, T., KNEIP, A. and KÖHLER, W. (1991). A flexible and fast method for automatic smoothing. *Journal of the American Statistical Association* **86** 643–652. MR1147088

[11] GEWEKE, J. (1982). Measurement of linear dependence and feedback between multiple time series. *J. Amer. Statist. Assoc.* **77** 304–324. With discussion and with a reply by the author. MR0664676

[12] GEWEKE, J. F. (1984). Measures of conditional linear dependence and feedback between time series. *J. Amer. Statist. Assoc.* **79** 907–915. MR0770286

[13] GIJBELS, I., VERAVERBEKE, N. and OMELKA, M. (2011). Conditional copulas, association measures and their applications. *Comput. Statist. Data Anal.* **55** 1919–1932. MR2765054 (2012d:62180)

[14] GOURIÉROUX, C., MONFORT, A. and RENAULT, É. (1987). Kullback causality measures. *Annales d'Économie et de Statistique* **6/7** 369–410. MR0925061

[15] GRANGER, C. W. (1969). Investigating causal relations by econometric models and cross-spectral methods. *Econometrica: Journal of the Econometric Society* 424–438.

[16] Hu, M. and Liang, H. (2014). A copula approach to assessing Granger causality. *NeuroImage* **100** 125–134.

[17] Joe, H. (2015). *Dependence modeling with copulas. Monographs on Statistics and Applied Probability* **134**. CRC Press, Boca Raton, FL. MR3328438

[18] Li, Q. and Racine, J. S. (2007). *Nonparametric econometrics: Theory and practice.* Princeton University Press. MR2283034

[19] Masry, E. (1996a). Multivariate regression estimation—local polynomial fitting for time series. *Stochastic Process. Appl.* **65** 81–101. MR1422881

[20] Masry, E. (1996b). Multivariate local polynomial regression for time series: uniform strong consistency and rates. *Journal of Time Series Analysis* **17** 571. MR1424907

[21] Masry, E. and Fan, J. (1997). Local polynomial estimation of regression functions for mixing processes. *Scandinavian Journal of Statistics* **24** 165–179. MR1455865

[22] Nelsen, R. B. (1996). Nonparametric measures of multivariate association. In *Distributions with Fixed Marginals and Related Topics (Seattle, WA, 1993). IMS Lecture Notes Monogr. Ser.* **28** 223–232. Inst. Math. Statist., Hayward, CA. MR1485534 (98k:62104)

[23] Polasek, W. (1994). Temporal causality measures based on AIC. In *Proceedings of the First US/Japan Conference on the Frontiers of Statistical Modeling: an Informational Approach, Vol. 1 (Knoxville, TN, 1992)* 13, 159–168. Kluwer Acad. Publ., Dordrecht. MR1328874

[24] Taamouti, A., Bouezmarni, T. and El Ghouch, A. (2014). Nonparametric estimation and inference for conditional density based Granger causality measures. *J. Econometrics* **180** 251–264. MR3197796

[25] van der Vaart, A. W. and Wellner, J. A. (1996). *Weak convergence and empirical processes. Springer Series in Statistics.* Springer-Verlag, New York. With applications to statistics. MR1385671 (97g:60035)

[26] Veraverbeke, N., Omelka, M. and Gijbels, I. (2011). Estimation of a conditional copula and association measures. *Scand. J. Stat.* **38** 766–780. MR2859749

[27] Wiener, N. (1956). The theory of prediction. *Modern Mathematics for Engineers* **1** 125–139.

[28] Zhang, H. J., Dufour, J.-M. and Galbraith, J. W. (2016). Exchange rates and commodity prices: Measuring causality at multiple horizons. *Journal of Empirical Finance* **36** 100–120.