December 7, 2020

ORDINAL PATTERN DEPENDENCE AS A MULTIVARIATE DEPENDENCE MEASURE

ANNIKA BETKEN, HEROLD DEHLING, INES MÜNKER, AND ALEXANDER SCHNURR

Abstract. In this article, we show that the recently introduced ordinal pattern dependence fits into the axiomatic framework of general multivariate dependence measures. Furthermore, we consider multivariate generalizations of established univariate dependence measures like Kendall's τ, Spearman's ρ and Pearson's correlation coefficient. Among these, only multivariate Kendall's τ proves to take the dynamical dependence of random vectors stemming from multidimensional time series into account. Consequently, the article focuses on a comparison of ordinal pattern dependence and multivariate Kendall's τ. To this end, limit theorems for multivariate Kendall's τ are established under the assumption of near epoch dependent, data-generating time series. We analyze how ordinal pattern dependence compares to multivariate Kendall's τ and Pearson's correlation coefficient on theoretical grounds. Additionally, a simulation study illustrates differences in the kind of dependencies that are revealed by multivariate Kendall's τ and ordinal pattern dependence.

1. Introduction

Recently, various attempts have been made to generalize classical dependence measures for one-dimensional random variables (like Pearson's correlation coefficient, Kendall's τ, Spearman's ρ) to a multivariate framework. The aim of these is to describe the degree of dependence between two random vectors with a single number. Roughly speaking one can separate the following two approaches: (I) In a first step, the main properties, which classical dependence measures between two random variables display, are extracted. In a second step, multivariate analogues of the dependence measures, which satisfy canonical generalizations of these properties in a multivariate framework, are defined. However, often a canonical interpretation of these measures is not at hand. (II) Given two time series, one wants to describe their co-movement.

Along these lines, the definition of ordinal pattern dependence (cf. Schnurr (2014)) follows the latter approach. Originally, axiomatic systems are disregarded by the notion of ordinal pattern dependence, which is naturally interpreted as the degree of co-monotonic behavior of two time series. Against the background of this approach, limit theorems have been proved in the time series setting (cf. Schnurr and Dehling (2017) for the SRD case and Betken et al. (2020) for the LRD case).

Both approaches in defining multivariate dependence measures have proved to be useful, but by now, they have been analyzed separately. In the present paper, we close the gap between the two. To this end, we recall the definition of ordinal pattern dependence in the subsequent section and show that it is a multivariate dependence measure according to the definition introduced in Grothe et al. (2014). In Section 3 we establish consistency and asymptotic normality for estimators of ordinal pattern dependence in the framework

Key words and phrases. ordinal patterns, time series, multivariate dependence, concordance ordering, limit theorems.

Research supported by DFG project SCHN 1231/3-1: *Ordinal-Pattern-Dependence: Grenzwertsätze und Strukturbrüche im langzeitabhängigen Fall mit Anwendungen in Hydrologie, Medizin und Finanzmathematik* and Collaborative Research Center SFB 823: *Statistical Modeling of Nonlinear Dynamic Processes* .

of i.i.d. random vectors. Section 4 deals with multivariate extensions of well-established univariate dependence measures. It turns out that multivariate Kendall's τ is the only one among these that captures the dynamical dependence between random vectors. Starting with approach (I) we prove limit theorems for an estimator of multivariate Kendall's τ in the time series context. In the last section, the different measures are compared from a theoretical point-of-view as well as by simulation studies.

2. Ordinal pattern dependence as a measure of multivariate dependence

If (X_i, Y_i), $i \geq 1$, denotes a stationary, bivariate process, we define, for any integers $i, h \geq 1$, the random vectors of consecutive observations

$$\mathbf{X}_i^{(h)} := (X_i, X_{i+1}, \ldots, X_{i+h}),$$
$$\mathbf{Y}_i^{(h)} := (Y_i, Y_{i+1}, \ldots, Y_{i+h}).$$

The goal of this paper is to consider the concept of ordinal pattern dependence as a multivariate measure of dependence between the random vectors $\mathbf{X}_i^{(h)}$ and $\mathbf{Y}_i^{(h)}$ stemming from a stationary, bivariate process (X_i, Y_i), $i \geq 1$, with continuous marginal distributions, and to compare it to established measures of dependence. Note that, by stationarity of the underlying process, the joint distribution of the vector $(\mathbf{X}_i^{(h)}, \mathbf{Y}_i^{(h)})$ does not depend on i. We will thus use the symbol $(\mathbf{X}^{(h)}, \mathbf{Y}^{(h)})$ for a generic random vector with the same joint distribution as any of the $(\mathbf{X}_i^{(h)}, \mathbf{Y}_i^{(h)})$ and we write $\mathbf{X}^{(h)} = (X_1, X_2, \ldots, X_{1+h})$, $\mathbf{Y}^{(h)} = (Y_1, Y_2, \ldots, Y_{1+h})$. Furthermore, note that it is common to count the number of increments h rather than the length of the vector, since ordinal patterns can be calculated by exclusively considering the increments of the time series.

2.1. Ordinal pattern dependence.

For $h \in \mathbb{N}$ let S_h denote the set of permutations of $\{0, \ldots, h\}$, which we write as $(h+1)$-tuples containing each of the numbers $0, \ldots, h$ exactly once. The ordinal pattern of order h refers to the permutation

$$\Pi(x_0, \ldots, x_h) = (\pi_0, \ldots, \pi_h) \in S_h$$

which satisfies

$$x_{\pi_0} \geq \ldots \geq x_{\pi_h}.$$

(cf. Bandt and Shiha (2007), Bandt and Pompe (2002)). In the present paper, we only consider continuous marginals. Allowing for non-continuous marginals would require the additional restriction $\pi_{j-1} > \pi_j$ if $x_{\pi_{j-1}} = x_{\pi_j}$ for $j \in \{1, ..., h\}$ (cf. Sinn and Keller (2011)).

Definition 2.1. We define the ordinal pattern dependence between two random vectors $\mathbf{X} := (X_1, \ldots, X_{h+1})$ and $\mathbf{Y} := (Y_1, \ldots, Y_{h+1})$ by

$$\begin{aligned}&\mathbf{OPD}_h(\mathbf{X}, \mathbf{Y})\\ &= \frac{P\left(\Pi\left(\mathbf{X}^{(h)}\right) = \Pi\left(\mathbf{Y}^{(h)}\right)\right) - \sum_{\pi \in S_h} P\left(\Pi(\mathbf{X}^{(h)}) = \pi\right) P\left(\Pi(\mathbf{Y}^{(h)}) = \pi\right)}{1 - \sum_{\pi \in S_h} P\left(\Pi(\mathbf{X}^{(h)}) = \pi\right) P\left(\Pi(\mathbf{Y}^{(h)}) = \pi\right)}.\end{aligned} \tag{1}$$

This definition of ordinal pattern dependence only takes positive dependence into account. Negative dependence can be included by analyzing the co-movement of X and

$-Y = (-Y_i)_{i\in\mathbb{N}}$. Typically, one is interested in measuring either positive *or* negative dependence. If one wants to consider both dependencies at the same time, a consideration of the quantity

$$\mathbf{OPD}_h(\mathbf{X}, \mathbf{Y})^+ - \mathbf{OPD}_h(\mathbf{X}, -\mathbf{Y})^+$$

where $a^+ := \max\{a, 0\}$ for every $a \in \mathbb{R}$, seems natural. In order to keep things less technical, we only consider the simpler measure (1). For recent developments in the theory of ordinal patterns cf. Mohr et al. (2020) and Bandt (2020).

2.2. **Axiomatic definition of multivariate dependence measures.** With the following definition, Grothe et al. (2014) establish an axiomatic theory for multivariate dependence measures between n-dimensional random vectors.

Definition 2.2. Let L_0 denote the space of random vectors with values in $\mathbb{R}^n$ on the common probability space $(\Omega, \mathcal{F}, P)$. We call a function $\mu : L_0 \times L_0 \longrightarrow \mathbb{R}$ an n-dimensional measure of dependence if

(1) it takes values in $[-1, 1]$,
(2) it is invariant with respect to simultaneous permutations of the components within two random vectors $\mathbf{X}$ and $\mathbf{Y}$,
(3) it is invariant with respect to increasing transformations of the components within two random vectors $\mathbf{X}$ and $\mathbf{Y}$,
(4) it is zero for two independent random vectors $\mathbf{X}$ and $\mathbf{Y}$,
(5) it respects concordance ordering, i.e., for two pairs of random vectors $\mathbf{X}$, $\mathbf{Y}$ and $\mathbf{X}^*$, $\mathbf{Y}^*$, it holds that

$$\begin{pmatrix}\mathbf{X}\\ \mathbf{Y}\end{pmatrix} \preccurlyeq_C \begin{pmatrix}\mathbf{X}^*\\ \mathbf{Y}^*\end{pmatrix} \Rightarrow \mu(\mathbf{X}, \mathbf{Y}) \le \mu(\mathbf{X}^*, \mathbf{Y}^*).$$

Here, $\preccurlyeq_C$ denotes concordance ordering, i.e.,

$$\begin{pmatrix}\mathbf{X}\\ \mathbf{Y}\end{pmatrix} \preccurlyeq_C \begin{pmatrix}\mathbf{X}^*\\ \mathbf{Y}^*\end{pmatrix} \text{ if and only if } F_{\binom{\mathbf{X}}{\mathbf{Y}}} \le F_{\binom{\mathbf{X}^*}{\mathbf{Y}^*}} \text{ and } \bar{F}_{\binom{\mathbf{X}}{\mathbf{Y}}} \le \bar{F}_{\binom{\mathbf{X}^*}{\mathbf{Y}^*}},$$

where $\le$ is meant pointwise and $\bar{F}$ denotes the survival function.

Theorem 2.3. *The ordinal pattern dependence* $\mathbf{OPD}_h$ *is an* $h+1$*-dimensional measure of dependence.*

Proof. The first four axioms in Definition 2.2 are easily verified, the fifth one is involved. We show that the fifth axiom is fulfilled for $\mathbf{OPD}_h$ with $h = 2$. For $h = 1$, the difficulties in the proof are not revealed, while for $h > 2$ the proof works analogously, but is notationally more complicated. Due to stationarity, it is enough to focus on the first three components of $\mathbf{X}$, $\mathbf{Y}$, $\mathbf{X}^*$, and $\mathbf{Y}^*$, i.e., without loss of generality we consider

$$\mathbf{X} = (X_1, X_2, X_3), \mathbf{Y} = (Y_1, Y_2, Y_3), \mathbf{X}^* = (X_1^*, X_2^*, X_3^*), \mathbf{Y}^* = (Y_1^*, Y_2^*, Y_3^*).$$

Moreover, we can restrict our considerations to $P(\Pi(\mathbf{X}) = \Pi(\mathbf{Y}))$ (the remaining summands of $\mathbf{OPD}_h(\mathbf{X}, \mathbf{Y})$ only relate to the distribution of $\mathbf{X}$ and $\mathbf{Y}$ separately). By Axiom 2 (invariance under permutation) it is, furthermore, enough to consider the monotone increasing pattern, that is, $\Pi(X_1, X_2, X_3) = (3, 2, 1)$. Let

$$\mathbf{X} \overset{\mathcal{D}}{=} \mathbf{X}^*,\ \mathbf{Y} \overset{\mathcal{D}}{=} \mathbf{Y}^*,\ \text{and} \begin{pmatrix}\mathbf{X}\\ \mathbf{Y}\end{pmatrix} \preccurlyeq_C \begin{pmatrix}\mathbf{X}^*\\ \mathbf{Y}^*\end{pmatrix}. \tag{2}$$

It is a well-known fact that (2) implies

$$\begin{pmatrix}\mathbf{X}\\ \mathbf{Y}\end{pmatrix}^I \preccurlyeq_C \begin{pmatrix}\mathbf{X}^*\\ \mathbf{Y}^*\end{pmatrix}^I$$

for all subvectors of variables with indices in $I \subseteq \{1,\dots,6\}$, i.e., removing dimensions does not influence which scenario has the larger dependence measure; see Joe (1990). We will make extensive use of this fact in what follows. Moreover, recall that we assume that all considered marginal distribution functions are continuous.

Defining

$$P^{x_1,y_1}(A) := P(A|X_1 = x_1, Y_1 = y_1),$$

considering $\Pi(X_1, X_2, X_3) = (3, 2, 1)$, and using disintegration twice yields

$$\begin{aligned}
&P\left(\Pi(X_1,X_2,X_3)=\Pi(Y_1,Y_2,Y_3)\right) = P\left(X_1 \le X_2 \le X_3, Y_1 \le Y_2 \le Y_3\right)\\
=&P\left(\{X_1 \le X_2\}\cap\{X_2 \le X_3\}\cap\{Y_1 \le Y_2\}\cap\{Y_2 \le Y_3\}\right)\\
=&\int_{\mathbb{R}^2} P^{x_1,y_1}(x_1 \le X_2 \le X_3, y_1 \le Y_2 \le Y_3)\, dP_{\binom{X_1}{Y_1}}(x_1,y_1)\\
=&\int_{\mathbb{R}^2} P^{x_1,y_1}(X_2 \le X_3, Y_2 \le Y_3 | x_1 \le X_2, y_1 \le Y_2)\cdot P^{x_1,y_1}(X_2 \ge x_1, Y_2 \ge y_1)\, dP_{\binom{X_1}{Y_1}}(x_1,y_1)\\
=&\int_{\mathbb{R}^2}\int_{[x_1,\infty[\times[y_1,\infty[} P^{x_1,y_1}(X_2 \le X_3, Y_2 \le Y_3 | X_2 = x_2, Y_2 = y_2)\, dP^{x_1,y_1}_{\binom{X_2}{Y_2}}(x_2,y_2)\\
&\qquad\qquad\times P^{x_1,y_1}(X_2 \ge x_1, Y_2 \ge y_1)\, dP_{\binom{X_1}{Y_1}}(x_1,y_1).
\end{aligned}$$

Since $\bar{F}_{\binom{X_3}{Y_3}}(x_2,y_2) = P^{x_1,y_1}(X_2 \le X_3, Y_2 \le Y_3 | x_1 \le X_2, y_1 \le Y_2)$, it follows that

$$\begin{aligned}
&P\left(\Pi(X_1,X_2,X_3)=\Pi(Y_1,Y_2,Y_3)\right)\\
&=\int_{\mathbb{R}^2}\int_{[x_1,\infty[\times[y_1,\infty[} \bar{F}_{\binom{X_3}{Y_3}}(x_2,y_2)\, dP^{x_1,y_1}_{\binom{X_2}{Y_2}}(x_2,y_2)\bar{F}_{\binom{X_2}{Y_2}}(x_1,y_1)\, dP_{\binom{X_1}{Y_1}}(x_1,y_1).
\end{aligned}$$

Due to (2), we deduce that

$$\int 1_{[a,\infty[\times[b,\infty[}(x,y)\, dP_{\binom{X_j}{Y_j}}(x,y) \le \int 1_{[a,\infty[\times[b,\infty[}(x,y)\, dP_{\binom{X_j^*}{Y_j^*}}(x,y)$$

for $a, b \in \mathbb{R}$ and $j = 1, 2, 3$. Since survival functions can be approximated by sums of indicator functions, the bounded convergence theorem yields

$$\int_{[x_1,\infty[\times[y_1,\infty[} \bar{F}_{\binom{X_3^*}{Y_3^*}}(x_2,y_2)\, dP^{x_1,y_1}_{\binom{X_2}{Y_2}}(x_2,y_2) \le \int_{[x_1,\infty[\times[y_1,\infty[} \bar{F}_{\binom{X_3^*}{Y_3^*}}(x_2,y_2)\, dP^{x_1,y_1}_{\binom{X_2^*}{Y_2^*}}(x_2,y_2).$$

Moreover, the function

$$H(x_1,x_2) := \int_{[x_1,\infty[\times[y_1,\infty[} \bar{F}_{\binom{X_3^*}{Y_3^*}}(x_2,y_2)\, dP^{x_1,y_1}_{\binom{X_2^*}{Y_2^*}}(x_2,y_2)$$

can be, up to scaling, considered as a survival function. Hence, we can use an approximation by sums of indicator functions for both, H and $\bar{F}_{\binom{X_2^*}{Y_2^*}}$. Most notably, the product of two functions having this property is of the same type, i.e., it can as well be approximated by sums of indicator functions.

For this reason, we finally arrive at

$$\begin{aligned}&P\left(\Pi(X_1,X_2,X_3)=\Pi(Y_1,Y_2,Y_3)\right)\\ \leq&\int_{\mathbb{R}^2}\int_{[x_1,\infty[\times[y_1,\infty[}\bar{F}_{\binom{X_3^*}{Y_3^*}}(x_2,y_2)\ dP^{x_1,y_1}_{\binom{X_2^*}{Y_2^*}}(x_2,y_2)\bar{F}_{\binom{X_2^*}{Y_2^*}}(x_1,y_1)\ dP_{\binom{X_1^*}{Y_1^*}}(x_1,y_1)\\ =&P(X_1^*\leq X_2^*\leq X_3^*, Y_1^*\leq Y_2^*\leq Y_3^*)=P\left(\Pi(X_1^*,X_2^*,X_3^*)=\Pi(Y_1^*,Y_2^*,Y_3^*)\right).\end{aligned}$$

□

3. Limit Theorems for Ordinal Pattern Dependence of i.i.d. Vectors

In Section 5, we compare ordinal pattern dependence to other concepts of multivariate dependence. These have been introduced and used for sequences of independent random vectors. In contrast to this, the definition of ordinal pattern dependence applies to random vectors stemming from multivariate time series. Nonetheless, ordinal pattern dependence can as well be applied to independent random vectors. Limit theorems that provide the asymptotic distribution of ordinal pattern dependence in this setting have not yet been established. We close this gap by the following considerations:

Let $(\mathbf{X}_i,\mathbf{Y}_i)$, $i\geq 1$, be independent copies of $(\mathbf{X}^{(h)},\mathbf{Y}^{(h)})$, and define

$$\begin{aligned}\hat{q}_{\mathbf{X},\pi,n} &:= \frac{1}{n}\sum_{i=1}^{n}1_{\{\Pi(\mathbf{X}_i)=\pi\}},\\ \hat{q}_{\mathbf{Y},\pi,n} &:= \frac{1}{n}\sum_{i=1}^{n-h}1_{\{\Pi(\mathbf{Y}_i)=\pi\}},\\ \hat{q}_{(\mathbf{X},\mathbf{Y}),n} &:= \frac{1}{n}\sum_{i=1}^{n-h}1_{\{\Pi(\mathbf{X}_i)=\Pi(\mathbf{Y}_i)\}},\end{aligned}$$

as well as the corresponding probabilities

$$\begin{aligned}q_{\mathbf{X}^{(h)},\pi} &:= P\left(\Pi(\mathbf{X}^{(h)})=\pi\right),\\ q_{\mathbf{Y}^{(h)},\pi} &:= P\left(\Pi(\mathbf{Y}^{(h)})=\pi\right),\\ q_{(\mathbf{X}^{(h)},\mathbf{Y}^{(h)})} &:= P\left(\Pi(\mathbf{X}^{(h)})=\Pi(\mathbf{Y}^{(h)})\right).\end{aligned}$$

According to the law of large numbers $\hat{q}_{\mathbf{X},\pi,n}$, $\hat{q}_{\mathbf{Y},\pi,n}$, and $\hat{q}_{(\mathbf{X},\mathbf{Y}),n}$ are strongly consistent estimators for these probabilities.

Proposition 3.1. *Let $(\mathbf{X}_i,\mathbf{Y}_i)$, $i\geq 1$, be independent copies of $(\mathbf{X}^{(h)},\mathbf{Y}^{(h)})$. Then, as $n\to\infty$,*

$$\begin{aligned}\hat{q}_{\mathbf{X},\pi,n} &\longrightarrow q_{\mathbf{X}^{(h)},\pi}\\ \hat{q}_{\mathbf{Y},\pi,n} &\longrightarrow q_{\mathbf{Y}^{(h)},\pi}\\ \hat{q}_{(\mathbf{X},\mathbf{Y}),n} &\longrightarrow q_{(\mathbf{X}^{(h)},\mathbf{Y}^{(h)})}\end{aligned}$$

almost surely.

The following theorem establishes asymptotic normality of ordinal pattern dependence of i.i.d. random vectors.

Theorem 3.2. *Let $(\mathbf{X}_i, \mathbf{Y}_i)$, $i \geq 1$, be independent copies of $(\mathbf{X}^{(h)}, \mathbf{Y}^{(h)})$. Then, as $n \to \infty$,*

$$\sqrt{n}\left(\frac{\hat{q}_{(\mathbf{X},\mathbf{Y}),n} - \sum_{\pi \in S_{h+1}} \hat{q}_{\mathbf{X},\pi,n}\hat{q}_{\mathbf{Y},\pi,n}}{1 - \sum_{\pi \in S_{h+1}} \hat{q}_{\mathbf{X}^{(h)},\pi,n}\hat{q}_{\mathbf{Y}^{(h)},\pi,n}} - \frac{q_{(\mathbf{X}^{(h)},\mathbf{Y}^{(h)})} - \sum_{\pi \in S_{h+1}} q_{\mathbf{X}^{(h)},\pi} q_{\mathbf{Y}^{(h)},\pi}}{1 - \sum_{\pi \in S_{h+1}} q_{\mathbf{X}^{(h)},\pi,n} q_{\mathbf{Y}^{(h)},\pi,n}}\right) \xrightarrow{\mathcal{D}} N(0, \sigma^2),$$

where the limit variance σ^2 is given by

$$\sigma^2 = \nabla f(q_{(\mathbf{X},\mathbf{Y})}, q_{\mathbf{X}}, q_{\mathbf{Y}})\, \Sigma \left(\nabla f(q_{(\mathbf{X},\mathbf{Y})}, q_{\mathbf{X}}, q_{\mathbf{Y}})\right)^t.$$

Here, the matrix Σ is defined as in Proposition 3.3 (see below), and ∇f is the gradient of the function $f : \mathbb{R} \times \mathbb{R}^{(h+1)!} \times \mathbb{R}^{(h+1)!} \to \mathbb{R}$, defined by

$$f(u, v, w) = \frac{u - v^t \cdot w}{1 - v^t \cdot w}.$$

The proof of Theorem 3.2 is based on the following proposition, which establishes the joint asymptotic normality of $\hat{q}_{\mathbf{X},\pi,n}$, $\hat{q}_{\mathbf{Y},\pi,n}$, and $\hat{q}_{(\mathbf{X},\mathbf{Y}),n}$. For this, we introduce the following notation:

$$\begin{aligned}
\hat{q}_{\mathbf{X},n} &:= \left(\hat{q}_{\mathbf{X},\pi,n}\right)_{\pi \in \mathcal{S}_{h+1}}, \\
\hat{q}_{\mathbf{Y},n} &:= \left(\hat{q}_{\mathbf{Y},\pi,n}\right)_{\pi \in \mathcal{S}_{h+1}}, \\
q_{\mathbf{X}} &:= \left(q_{\mathbf{X},\pi}\right)_{\pi \in \mathcal{S}_{h+1}}, \\
q_{\mathbf{Y}} &:= \left(q_{\mathbf{Y},\pi}\right)_{\pi \in \mathcal{S}_{h+1}}.
\end{aligned}$$

Proposition 3.3. *Under the same assumptions as in Theorem 3.2, we have*

$$\sqrt{n}\begin{pmatrix} \hat{q}_{(\mathbf{X},\mathbf{Y}),n} - q_{(\mathbf{X},\mathbf{Y})} \\ \hat{q}_{\mathbf{X},n} - q_{\mathbf{X}} \\ \hat{q}_{\mathbf{Y},n} - q_{\mathbf{Y}} \end{pmatrix} \xrightarrow{\mathcal{D}} N(0, \Sigma), \tag{3}$$

where Σ is the symmetric $(2(h+1)!+1) \times (2(h+1)!+1)$ matrix

$$\Sigma = \begin{pmatrix} \Sigma_{11} & \Sigma_{12} & \Sigma_{13} \\ \Sigma_{21} & \Sigma_{22} & \Sigma_{23} \\ \Sigma_{31} & \Sigma_{32} & \Sigma_{33} \end{pmatrix}$$

with

$$\begin{aligned}
&\Sigma_{11} = q_{(\mathbf{X},\mathbf{Y})}(1 - q_{(\mathbf{X},\mathbf{Y})}) \in \mathbb{R}, \\
&\Sigma_{12} = (\sigma_\pi(1,2))_{\pi \in \mathcal{S}_{h+1}} \in \mathbb{R}^{1 \times (h+1)!}, && \sigma_\pi(1,2) = P(\Pi(X) = \Pi(Y) = \pi) - q_{\mathbf{X},\mathbf{Y}} \cdot q_{\mathbf{X},\pi}, \\
&\Sigma_{13} = (\sigma_\pi(1,3))_{\pi \in \mathcal{S}_{h+1}} \in \mathbb{R}^{1 \times (h+1)!}, && \sigma_\pi(1,3) = P(\Pi(X) = \Pi(Y) = \pi) - q_{\mathbf{X},\mathbf{Y}} \cdot q_{\mathbf{Y},\pi}, \\
&\Sigma_{22} = (\sigma_{\pi,\pi'}(2,2))_{\pi,\pi' \in \mathcal{S}_{h+1}} \in \mathbb{R}^{(h+1)! \times (h+1)!}, && \sigma_{\pi,\pi'}(2,2) = \begin{cases} -q_{\mathbf{X},\pi} q_{\mathbf{X},\pi'} & \text{if } \pi \neq \pi' \\ q_{\mathbf{X},\pi}(1 - q_{\mathbf{X},\pi}) & \text{if } \pi = \pi', \end{cases} \\
&\Sigma_{23} = (\sigma_{\pi,\pi'}(2,3))_{\pi,\pi' \in \mathcal{S}_{h+1}} \in \mathbb{R}^{(h+1)! \times (h+1)!}, && \sigma_{\pi,\pi'}(2,3) = P(\Pi(\mathbf{X}) = \pi, \Pi(\mathbf{Y}) = \pi') - q_{\mathbf{X},\pi} q_{\mathbf{Y},\pi'}, \\
&\Sigma_{33} = (\sigma_{\pi,\pi'}(3,3))_{\pi,\pi' \in \mathcal{S}_{h+1}} \in \mathbb{R}^{(h+1)! \times (h+1)!}, && \sigma_{\pi,\pi'}(3,3) = \begin{cases} -q_{\mathbf{Y},\pi} q_{\mathbf{Y},\pi'} & \text{if } \pi \neq \pi' \\ q_{\mathbf{Y},\pi}(1 - q_{\mathbf{Y},\pi}) & \text{if } \pi = \pi'. \end{cases}
\end{aligned}$$

Due to symmetry of Σ, the remaining blocks are defined by $\Sigma_{21} = \Sigma_{12}^t$, $\Sigma_{31} = \Sigma_{13}^t$, $\Sigma_{32} = \Sigma_{23}^t$.

Proof. The proof follows directly from the multivariate central limit theorem applied to the partial sums of the $(2(h+1)!+1)$-dimensional i.i.d. random vectors

$$\xi_i = \left(1_{\{\Pi(\mathbf{X}_i)=\Pi(\mathbf{Y}_i)\}}, (1_{\{\Pi(\mathbf{X}_i)=\pi\}})_{\pi\in\mathcal{S}_{h+1}}, (1_{\{\Pi(\mathbf{Y}_i)=\pi\}})_{\pi\in\mathcal{S}_{h+1}}\right)^t$$

The limit covariance matrix is the covariance matrix of ξ_1, which is given by the formulae stated in the formulation of this proposition. □

Proof of Theorem 3.2. We apply the delta method to the function f, defined in the formulation of the theorem, together with the multivariate CLT established in Proposition 3.3. In this way, we obtain

$$\begin{aligned}&\sqrt{n}\left(f(\hat{q}_{(\mathbf{X},\mathbf{Y}),n}, \hat{q}_{\mathbf{X},n}, \hat{q}_{\mathbf{Y},n}) - f(q_{(\mathbf{X}^{(h)},\mathbf{Y}^{(h)})}, q_{\mathbf{X}^{(h)}}, q_{\mathbf{Y}^{(h)}})\right)\\ &\xrightarrow{\mathcal{D}} N(0, \nabla f(q_{(\mathbf{X}^{(h)},\mathbf{Y}^{(h)})}, q_{\mathbf{X}^{(h)}}, q_{\mathbf{Y}^{(h)}}) \cdot \Sigma \cdot (\nabla f(q_{(\mathbf{X}^{(h)},\mathbf{Y}^{(h)})}, q_{\mathbf{X}^{(h)}}, q_{\mathbf{Y}^{(h)}}))^t).\end{aligned}$$

This proves the statement of the theorem. □

4. Classical Measures of Multivariate Dependence in a Time Series Context

In this article, we are explicitly studying the dependence between random vectors stemming from stationary time series. In this regard, the main drawback of univariate dependence measures is that these do not incorporate cross-dependencies, which characterize the *dynamical* dependence between two random vectors. Univariate dependence measures focus on the dependence between X_i and Y_i, i.e., on the dependence at the same point in time. In contrast, ordinal pattern dependence captures the dynamics of time series.

In the following, we study two multivariate generalizations of univariate dependence measures, namely the multivariate extension of Pearson's correlation coefficient, established in Puccetti (2019), and multivariate Kendall's τ as introduced in Grothe et al. (2014).

Definition 4.1. For two random vectors $\mathbf{X}^{(h)}, \mathbf{Y}^{(h)} \in L_2\left(\mathbb{R}^{h+1}\right)$ with invertible correlation matrices $\Sigma_{\mathbf{X}^{(h)}}$ and $\Sigma_{\mathbf{Y}^{(h)}}$ and cross-covariance matrix $\Sigma_{\mathbf{X}^{(h)},\mathbf{Y}^{(h)}}$, we define Pearson's correlation coefficient by

$$\rho\left(\mathbf{X}^{(h)}, \mathbf{Y}^{(h)}\right) := \frac{\operatorname{tr}\left(\Sigma_{\mathbf{X}^{(h)},\mathbf{Y}^{(h)}}\right)}{\operatorname{tr}\left(\left(\Sigma_{\mathbf{X}^{(h)}}\Sigma_{\mathbf{Y}^{(h)}}\right)^{1/2}\right)},$$

where $A^{1/2}$ is the principal square root of the matrix A, such that $A^{1/2}A^{1/2} = A$.

For the multivariate generalization of Pearson's correlation coefficient, we obtain

$$\rho\left(\mathbf{X}^{(h)}, \mathbf{Y}^{(h)}\right) = \frac{\operatorname{tr}\left(\Sigma_{\mathbf{X}^{(h)},\mathbf{Y}^{(h)}}\right)}{\operatorname{tr}\left(\left(\Sigma_{\mathbf{X}^{(h)}}\Sigma_{\mathbf{Y}^{(h)}}\right)^{1/2}\right)} = \frac{\operatorname{Cov}(X_1, Y_1) + \ldots + \operatorname{Cov}(X_{1+h}, Y_{1+h})}{\operatorname{tr}\left(\left(\Sigma_{\mathbf{X}^{(h)}}\Sigma_{\mathbf{Y}^{(h)}}\right)^{1/2}\right)}.$$

As a result, the cross-correlations have no impact on the value of Pearson's correlation coefficient. Therefore, the multivariate Pearson's correlation coefficient does not seem to be appropriate for our approach. The same holds true for generalizations of Spearman's ρ, due to the close relationship between these concepts. We hence focus on the multivariate generalization of Kendall's τ:

4.1. **Multivariate Kendall's τ.** The definition of multivariate Kendall's τ that we consider in this section is taken from Grothe et al. (2014) .

Definition 4.2. For two random vectors $\mathbf{X}^{(h)}, \mathbf{Y}^{(h)}$, we define Kendall's τ by

$$\tau_h(\mathbf{X}^{(h)}, \mathbf{Y}^{(h)}) := \text{Corr}\left(\mathbf{1}_{\left\{\mathbf{X}^{(h)} \leq \tilde{\mathbf{X}}^{(h)}\right\}}, \mathbf{1}_{\left\{\mathbf{Y}^{(h)} \leq \tilde{\mathbf{Y}}^{(h)}\right\}}\right),$$

where $\left(\tilde{\mathbf{X}}^{(h)}, \tilde{\mathbf{Y}}^{(h)}\right)^t$ is an independent copy of $(\mathbf{X}^{(h)}, \mathbf{Y}^{(h)})^t$.

The following lemma establishes a representation of multivariate Kendall's τ in terms of the probabilities $p_{\mathbf{X}^{(h)}}$ and $p_{\mathbf{Y}^{(h)}}$ that enter in our definition of ordinal pattern dependence.

Lemma 4.3. *Let (X_i, Y_i), $i \geq 1$, denote a stationary Gaussian process and let $\mathbf{X}^{(h)} = (X_1, X_2, \ldots, X_{1+h})$ and $\mathbf{Y}^{(h)} = (Y_1, Y_2, \ldots, Y_{1+h})$. Then, we have*

$$\tau_h\left(\mathbf{X}^{(h)}, \mathbf{Y}^{(h)}\right) = \frac{P\left(X_1 \leq 0, \ldots, X_{1+h} \leq 0, Y_1 \leq 0, \ldots, Y_{1+h} \leq 0\right) - p_{\mathbf{X}^{(h)}} p_{\mathbf{Y}^{(h)}}}{\sqrt{p_{\mathbf{X}^{(h)}}\left(1 - p_{\mathbf{X}^{(h)}}\right) p_{\mathbf{Y}^{(h)}}\left(1 - p_{\mathbf{Y}^{(h)}}\right)}},$$

where $p_{\mathbf{X}^{(h)}} = P\left(X_1 \leq 0, \ldots, X_{1+h} \leq 0\right)$ and $p_{\mathbf{Y}^{(h)}} = P\left(Y_1 \leq 0, \ldots, Y_{1+h} \leq 0\right)$.

Proof. Let $\left(\tilde{\mathbf{X}}^{(h)}, \tilde{\mathbf{Y}}^{(h)}\right)^t$ be an independent copy of $(\mathbf{X}^{(h)}, \mathbf{Y}^{(h)})^t$ with $\tilde{\mathbf{X}}^{(h)} = (\tilde{X}_1, \tilde{X}_2, \ldots, \tilde{X}_{1+h})$ and $\tilde{\mathbf{Y}}^{(h)} = (\tilde{Y}_1', \tilde{Y}_2, \ldots, \tilde{Y}_{1+h})$. It then holds that

$$\begin{aligned}
&\tau_h\left(\mathbf{X}^{(h)}, \mathbf{Y}^{(h)}\right) \\
&= \text{Corr}\left(\mathbf{1}_{\left\{\mathbf{X}^{(h)} \leq \tilde{\mathbf{X}}^{(h)}\right\}}, \mathbf{1}_{\left\{\mathbf{Y}^{(h)} \leq \tilde{\mathbf{Y}}^{(h)}\right\}}\right) \\
&= \text{Corr}\left(\mathbf{1}_{\left\{\mathbf{X}^{(h)} - \tilde{\mathbf{X}}^{(h)} \leq 0\right\}}, \mathbf{1}_{\left\{\mathbf{Y}^{(h)} - \tilde{\mathbf{Y}}^{(h)} \leq 0\right\}}\right) \\
&= \frac{P\left(X_1 - \tilde{X}_1 \leq 0, \ldots, X_{1+h} - \tilde{X}_{1+h} \leq 0, Y_1 - \tilde{Y}_1 \leq 0, \ldots, Y_{1+h} - \tilde{Y}_{1+h} \leq 0\right) - p_{\mathbf{X}^{(h)}} p_{\mathbf{Y}^{(h)}}}{\sqrt{p_{\mathbf{X}^{(h)}}\left(1 - p_{\mathbf{X}^{(h)}}\right) p_{\mathbf{Y}^{(h)}}\left(1 - p_{\mathbf{Y}^{(h)}}\right)}}
\end{aligned}$$

with $p_{\mathbf{Y}^{(h)}} = P\left(Y_1 - \tilde{Y}_1 \leq 0, \ldots, Y_{1+h} - \tilde{Y}_{1+h} \leq 0\right)$ and $p_{\mathbf{X}^{(h)}} = P\left(X_1 - \tilde{X}_1 \leq 0, \ldots, X_{1+h} - \tilde{X}_{1+h} \leq 0\right)$.
Note that for centered Gaussian processes

$$\left(\mathbf{X}^{(h)} - \tilde{\mathbf{X}}^{(h)}, \mathbf{Y}^{(h)} - \tilde{\mathbf{Y}}^{(h)}\right) \stackrel{D}{=} \sqrt{2}\left(\mathbf{X}^{(h)}, \mathbf{Y}^{(h)}\right).$$

This explicitly implies that the cross-correlations within $\left(\mathbf{X}^{(h)} - \tilde{\mathbf{X}}^{(h)}, \mathbf{Y}^{(h)} - \tilde{\mathbf{Y}}^{(h)}\right)$ equal those within $\left(\mathbf{X}^{(h)}, \mathbf{Y}^{(h)}\right)$. Therefore, we have

$$\begin{aligned}
&\frac{P\left(X_1 - \tilde{X}_1 \leq 0, \ldots, X_{1+h} - \tilde{X}_{1+h} \leq 0, Y_1 - \tilde{Y}_1 \leq 0, \ldots, Y_{1+h} - \tilde{Y}_{1+h} \leq 0\right) - p_{\mathbf{X}^{(h)}} p_{\mathbf{Y}^{(h)}}}{\sqrt{p_{\mathbf{X}^{(h)}}\left(1 - p_{\mathbf{X}^{(h)}}\right) p_{\mathbf{Y}^{(h)}}\left(1 - p_{\mathbf{Y}^{(h)}}\right)}} \\
&= \frac{P\left(X_1 \leq 0, \ldots, X_{1+h} \leq 0, Y_1 \leq 0, \ldots, Y_{1+h} \leq 0\right) - p_{\mathbf{X}^{(h)}} p_{\mathbf{Y}^{(h)}}}{\sqrt{p_{\mathbf{X}^{(h)}}\left(1 - p_{\mathbf{X}^{(h)}}\right) p_{\mathbf{Y}^{(h)}}\left(1 - p_{\mathbf{Y}^{(h)}}\right)}}
\end{aligned}$$

with $p_{\mathbf{Y}^{(h)}} = P\left(Y_1 \leq 0, \ldots, Y_{1+h} \leq 0\right)$ and $p_{\mathbf{X}^{(h)}} = P\left(X_1 \leq 0, \ldots, X_{1+h} \leq 0\right)$. □

Although for $h \geq 2$ we cannot derive an analytic expression for

$$P\left(X_1 \leq 0, \ldots, X_{1+h} \leq 0\right), \ P\left(Y_1 \leq 0, \ldots, Y_{1+h} \leq 0\right),$$

or

$$P\left(X_1 \leq 0, \ldots, X_{1+h} \leq 0, Y_1 \leq 0, \ldots, Y_{1+h} \leq 0\right),$$

we know that these orthant probabilities of a multivariate Gaussian distribution are determined by the entries of the correlation matrices and by the entries of the cross-correlation matrix of $\mathbf{X}^{(h)}$ and $\mathbf{Y}^{(h)}$. In contrast to multivariate Pearson's correlation coefficient, multivariate Kendall's τ constitutes a multivariate dependence measure that takes the dynamical dependence of data stemming from time series into account.

4.2. **Estimation of multivariate Kendall's τ.** Grothe et al. (2014) consider an estimator for multivariate Kendall's τ based on independent vectors $(\mathbf{X}_i, \mathbf{Y}_i)$, $1 \leq i \leq n$. In our setup, we will define an empirical version of Kendall's τ based on the dependent vectors $(\mathbf{X}_i^{(h)}, \mathbf{Y}_i^{(h)})$, $1 \leq i \leq n$. For this, we will follow the ideas of Dehling et al. (2017), who considered estimation of the classical univariate Kendall's τ for bivariate time series under some mild dependence condition.

Given an independent copy $(\tilde{\mathbf{X}}^{(h)}, \tilde{\mathbf{Y}}^{(h)})$ of the vector $(\mathbf{X}^{(h)}, \mathbf{Y}^{(h)})$, we have

$$\begin{aligned} \tau_h(\mathbf{X}^{(h)}, \mathbf{Y}^{(h)}) &= \frac{p_{(\mathbf{X}^{(h)}, \mathbf{Y}^{(h)})} - p_{\mathbf{X}^{(h)}} p_{\mathbf{Y}^{(h)}}}{\sqrt{p_{\mathbf{X}^{(h)}}(1 - p_{\mathbf{X}^{(h)}}) p_{\mathbf{Y}^{(h)}}(1 - p_{\mathbf{Y}^{(h)}})}} \\ &= \psi(p_{\mathbf{X}^{(h)}}, p_{\mathbf{Y}^{(h)}}, p_{(\mathbf{X}^{(h)}, \mathbf{Y}^{(h)})}), \end{aligned}$$

where $p_{(\mathbf{X}^{(h)}, \mathbf{Y}^{(h)})} := P(\mathbf{X}^{(h)} \leq \tilde{\mathbf{X}}^{(h)}, \mathbf{Y}^{(h)} \leq \tilde{\mathbf{Y}}^{h})$, $p_{\mathbf{X}^{(h)}} := P(\mathbf{X}^{(h)} \leq \tilde{\mathbf{X}}(h))$, $p_{\mathbf{Y}^{(h)}} := P(\mathbf{Y}^{(h)} \leq \tilde{\mathbf{Y}}(h))$, and where $\psi : \mathbb{R}^3 \to \mathbb{R}$ is defined by

$$\psi(x, y, z) := \frac{z - x\, y}{x(1-x)y(1-y)}.$$

The probabilities $p_{\mathbf{X}^{(h)}}$, $p_{\mathbf{Y}^{(h)}}$, and $p_{(\mathbf{X}^{(h)}, \mathbf{Y}^{(h)})}$ can be estimated by their sample analogues, defined by

$$\begin{aligned} \hat{p}_{\mathbf{X}^{(h)},n} &:= \frac{2}{n(n-1)} \sum_{1 \leq i < j \leq n} 1_{\{\mathbf{X}_i^{(h)} \leq \mathbf{X}_j^{(h)}\}}, \\ \hat{p}_{\mathbf{Y}^{(h)},n} &:= \frac{2}{n(n-1)} \sum_{1 \leq i < j \leq n} 1_{\{\mathbf{Y}_i^{(h)} \leq \mathbf{Y}_j^{(h)}\}}, \\ \hat{p}_{(\mathbf{X}^{(h)}, \mathbf{Y}^{(h)}),n} &:= \frac{2}{n(n-1)} \sum_{1 \leq i < j \leq n} 1_{\{\mathbf{X}_i^{(h)} \leq \mathbf{X}_j^{(h)}, \mathbf{Y}_i^{(h)} \leq \mathbf{Y}_j^{(h)}\}}, \end{aligned}$$

where $\mathbf{X}_i^{(h)} = (X_i, X_{i+1}, \ldots, X_{i+h})$ and $\mathbf{Y}_i^{(h)} = (Y_i, Y_{i+1}, \ldots, Y_{i+h})$. The plug-in estimator for Kendall's τ is then given by

$$\hat{\tau}_n^{(h)}(\mathbf{X}, \mathbf{Y}) := \psi(\hat{p}_{\mathbf{X}^{(h)},n}, \hat{p}_{\mathbf{Y}^{(h)},n}, \hat{p}_{(\mathbf{X}^{(h)}, \mathbf{Y}^{(h)}),n}).$$

In what follows, we will derive the joint limit distribution of the random vector $(\hat{p}_{\mathbf{X}^{(h)},n}, \hat{p}_{\mathbf{Y}^{(h)},n}, \hat{p}_{(\mathbf{X}^{(h)}, \mathbf{Y}^{(h)}),n})$, and the limit distribution of $\hat{\tau}_n^{(h)}(\mathbf{X}, \mathbf{Y})$ by the delta method.

For this, observe that $\hat{p}_{\mathbf{X}^{(h)},n}$, $\hat{p}_{\mathbf{Y}^{(h)},n}$, and $\hat{p}_{(\mathbf{X}^{(h)},\mathbf{Y}^{(h)}),n}$ are U-statistics with symmetric kernels

$$\begin{aligned} f((x,y),(x',y')) &= \frac{1}{2}\left(1_{\{x\le x'\}} + 1_{\{x\ge x'\}}\right), \\ g((x,y),(x',y')) &= \frac{1}{2}\left(1_{\{y\le y'\}} + 1_{\{y\ge y'\}}\right), \\ h((x,y),(x',y')) &= \frac{1}{2}\left(1_{\{x\le x',y\le y'\}} + 1_{\{x\ge x',y\ge y'\}}\right). \end{aligned}$$

Note that the underlying random vectors $(\mathbf{X}_i^{(h)}, \mathbf{Y}_i^{(h)})$, $i \ge 1$, are dependent, so that standard U-statistics theory for independent data does not apply. However, we can apply an ergodic theorem for U-statistics, established in Aaronson et al. (1996).

Theorem 4.4. *Assume that (X_i, Y_i), $i \ge 1$, is a stationary ergodic process, and that $(\mathbf{X}^{(h)}, \mathbf{Y}^{(h)})$ has a continuous distribution. Then, as $n \to \infty$,*

$$\begin{aligned} \hat{p}_{\mathbf{X}^{(h)},n} &\longrightarrow p_{\mathbf{X}^{(h)}}, \\ \hat{p}_{\mathbf{Y}^{(h)},n} &\longrightarrow p_{\mathbf{Y}^{(h)}}, \\ \hat{p}_{(\mathbf{X}^{(h)},\mathbf{Y}^{(h)}),n} &\longrightarrow p_{(\mathbf{X}^{(h)},\mathbf{Y}^{(h)})}, \end{aligned}$$

almost surely.

Proof. We apply Theorem U from Aaronson et al. (1996). The kernels f, g, and h, are almost everywhere continuous, and thus condition (ii) of Theorem U holds. □

In order to establish asymptotic normality of these estimators, we have to make some assumptions assuring short-range dependence of the underlying process. We will use the concept of near epoch dependence in probability, introduced in Dehling et al. (2017). This concept is a variation of the usual L_2-near epoch dependence, and does not require any moment assumptions.

Definition 4.5. (i) Given two sub-σ-fields $\mathcal{A}, \mathcal{B} \subset \mathcal{F}$, we define the absolute regularity coefficient

$$\beta(\mathcal{A},\mathcal{B}) = \sup\{\sum_{i,j} |P(A_i \cap B_j) - P(A_i)P(B_j)|\},$$

where the supremum is taken over all integers $m, n \ge 1$, all partitions $A_1, \ldots, A_m \in \mathcal{A}$, and all partitions $B_1, \ldots, B_n \in \mathcal{B}$ of the sample space Ω.

(ii) For a stationary stochastic process Z_i, $i \in \mathbb{Z}$, we define the absolute regularity coefficients

$$\beta_k := \beta(\mathcal{F}_{-\infty}^0, \mathcal{F}_k^\infty),$$

where $\mathcal{F}_k^l$ denotes the σ-field generated by the random variables $X_k, \ldots, X_l$. The process $(Z_i)_{i\in\mathbb{Z}}$ is called absolutely regular if $\lim_{k\to\infty} \beta_k = 0$.

(iii) An $\mathbb{R}^d$-valued stochastic process X_i, $i \ge 1$, is called near epoch dependent in probability (in short P-NED) on the stationary process Z_i, $i \in \mathbb{Z}$, if (X_i, Z_i), $i \ge 1$, is a stationary process, and if there exists a sequence $(a_k)_{k\ge 0}$ of approximating constants with $\lim_{k\to\infty} a_k = 0$, a sequence of functions $f_k : \mathbb{R}^{2k+1} \to \mathbb{R}^d$, and a nonincreasing function $\Phi : (0,\infty) \to (0,\infty)$ such that

$$P(|X_0 - f_k(Z_{-k}, \ldots, Z_k)| \ge \epsilon) \le a_k \Phi(\epsilon).$$

Proposition 4.6. *Let $(\mathbf{X}_i, \mathbf{Y}_i)$, $i \geq 1$, be a stationary process that is P-NED on an absolutely regular process Z_k, $k \in \mathbb{Z}$, and assume that*

$$a_k \Phi(k^{-6}) = O(k^{-6(2+\delta)/\delta}) \quad \text{and} \quad \sum_{k=1}^{\infty} k \beta_k^{\delta/(2+\delta)} < \infty,$$

for some $\delta > 0$. Moreover, assume that $\mathbf{Y}^{(h)} - \mathbf{X}^{(h)}$ has a bounded density. Then, the following approximations hold

$$\begin{aligned} \sqrt{n}\left(\hat{p}_{\mathbf{X}^{(h)},n} - p_{\mathbf{X}^{(h)}}\right) &= \frac{2}{\sqrt{n}} \sum_{i=1}^{n} f_1(\mathbf{X}_i^{(h)}, \mathbf{Y}_i^{(h)}) + o_P(1), \\ \sqrt{n}\left(\hat{p}_{\mathbf{Y}^{(h)},n} - p_{\mathbf{Y}^{(h)}}\right) &= \frac{2}{\sqrt{n}} \sum_{i=1}^{n} g_1(\mathbf{X}_i^{(h)}, \mathbf{Y}_i^{(h)}) + o_P(1), \\ \sqrt{n}\left(\hat{p}_{(\mathbf{X}^{(h)},\mathbf{Y}^{(h)}),n} - p_{(\mathbf{X}^{(h)},\mathbf{Y}^{(h)})}\right) &= \frac{2}{\sqrt{n}} \sum_{i=1}^{n} h_1(\mathbf{X}_i^{(h)}, \mathbf{Y}_i^{(h)}) + o_P(1), \end{aligned}$$

where the functions f_1, g_1, and h_1 are the first order terms in the Hoeffding expansion of the kernels f, g, and h, respectively.

Remark 4.1. The first order term of the Hoeffding decomposition is, e.g., given by

$$\begin{aligned} f_1(x,y) &= Ef((x,y),(\tilde{\mathbf{X}}^{(h)}, \tilde{\mathbf{Y}}^{(h)})) - p_{\mathbf{X}^{(h)}} \\ &= \frac{1}{2}\left(P(\mathbf{X}^{(h)} \leq x) + P(\mathbf{X}^{(h)} \geq x)\right) - P(\mathbf{X}^{(h)} \leq \tilde{\mathbf{X}}^{(h)}) \\ &= \frac{1}{2}\left(F(x) + \bar{F}(x)\right) - P(\mathbf{X}^{(h)} \leq \tilde{\mathbf{X}}^{(h)}), \end{aligned}$$

where $F(x) := P(\mathbf{X}^{(h)} \leq x)$ and $\bar{F}(x) := P(\mathbf{X}^{(h)} \geq x)$. Similarly, we get

$$\begin{aligned} g_1(x,y) &= \frac{1}{2}\left(G(x) + \bar{G}(x)\right) - P(\mathbf{Y}^{(h)} \leq \tilde{\mathbf{Y}}^{(h)}), \\ h_1(x,y) &= \frac{1}{2}\left(H(x,y) + \bar{H}(x,y)\right) - P(\mathbf{X}^{(h)} \leq \tilde{\mathbf{X}}^{(h)}, \mathbf{Y}^{(h)} \leq \tilde{\mathbf{Y}}^{(h)}), \end{aligned}$$

where G, $\bar{G}$, H, and $\bar{H}$, are defined analogously to F and $\bar{F}$.

Proof of Proposition 4.6. This follows from Lemma D.6 of Dehling et al. (2017), noting that the variation condition is satisfied because the distribution of $\mathbf{Y}^{(h)} - \mathbf{X}^{(h)}$ has a bounded density. □

Theorem 4.7. *Under the same assumptions as in Proposition 4.6, we have*

$$\sqrt{n}\begin{pmatrix} \hat{p}_{\mathbf{X}^{(h)},n} - p_{\mathbf{X}^{(h)}} \\ \hat{p}_{\mathbf{Y}^{(h)},n} - p_{\mathbf{Y}^{(h)}} \\ \hat{p}_{(\mathbf{X}^{(h)},\mathbf{Y}^{(h)}),n} - p_{(\mathbf{X}^{(h)},\mathbf{Y}^{(h)})} \end{pmatrix} \xrightarrow{\mathcal{D}} N(0, \Sigma),$$

where Σ is the limit covariance matrix, whose diagonal and off-diagonal entries are given, e.g., by

$$\begin{aligned}
\sigma_{11} &= \operatorname{Var}\left(f_1(\mathbf{X}_1^{(h)}, \mathbf{Y}_1^{(h)})\right) + 2\sum_{i=2}^{\infty} \operatorname{Cov}\left(f_1(\mathbf{X}_1^{(h)}, \mathbf{Y}_1^{(h)}), f_1(\mathbf{X}_i^{(h)}, \mathbf{Y}_i^{(h)})\right), \\
\sigma_{12} &= \operatorname{Cov}\left(f_1(\mathbf{X}_1^{(h)}, \mathbf{Y}_1^{(h)}), g_1(\mathbf{X}_1^{(h)}, \mathbf{Y}_1^{(h)})\right) \\
&\quad + \sum_{i=2}^{\infty} \operatorname{Cov}\left(f_1(\mathbf{X}_1^{(h)}, \mathbf{Y}_1^{(h)}), g_1(\mathbf{X}_i^{(h)}, \mathbf{Y}_i^{(h)})\right) \\
&\quad + \sum_{i=2}^{\infty} \operatorname{Cov}\left(f_1(\mathbf{X}_i^{(h)}, \mathbf{Y}_i^{(h)}), g_1(\mathbf{X}_1^{(h)}, \mathbf{Y}_1^{(h)})\right).
\end{aligned}$$

Proof. By the multivariate central limit theorem for partial sums of NED processes, we obtain

$$S_n := \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \left(f_1(\mathbf{X}_i^{(h)}, \mathbf{Y}_i^{(h)}), g_1(\mathbf{X}_i^{(h)}, \mathbf{Y}_i^{(h)}), h_1(\mathbf{X}_i^{(h)}, \mathbf{Y}_i^{(h)})\right) \xrightarrow{\mathcal{D}} N(0, \Sigma);$$

see, e.g. Wooldridge and White (1988). Now, the statement of the theorem follows from Proposition 4.6 together with an application of Slutsky's lemma. □

Theorem 4.8. *Under the assumptions of Proposition 4.6, the estimator $\hat{\tau}_n(\mathbf{X}^{(h)}, \mathbf{Y}^{(h)})$ of Kendall's τ $\tau_h(\mathbf{X}^{(h)}, \mathbf{Y}^{(h)})$ is consistent and asymptotically normal. More precisely, we obtain*

$$\sqrt{n}\left(\hat{\tau}_n(\mathbf{X}^{(h)}, \mathbf{Y}^{(h)}) - \tau_h(\mathbf{X}^{(h)}, \mathbf{Y}^{(h)})\right) \xrightarrow{\mathcal{D}} N(0, (\nabla\psi)^t \Sigma (\nabla\psi)),$$

where ψ and Σ are defined as above.

Proof. This follows from the previous theorem, together with the delta method applied to the function ψ. □

5. Ordinal Pattern Dependence in Contrast to Other Dependence Measures

For independent vectors $(\mathbf{X}_i, \mathbf{Y}_i)$, $1 \leq i \leq n$, all dependence measures considered in the previous sections make sense. Yet, for measuring dependence between two time series only ordinal pattern dependence and Kendall's τ seem to be reasonable choices of dependence measures. In this section, we point out what kind of dependencies are measured by ordinal pattern dependence and how ordinal pattern dependence compares to classical dependence measures such as Pearson's correlation coefficient and multivariate Kendall's τ.

5.1. **The case $h = 1$.** Axiom (4) in Definition 2.2 ensures that a multivariate dependence measure takes the value zero if the respective vectors are independent. In this regard, a natural question that arises when studying the dependence between two random vectors is whether the considered dependence measure may also differentiate between independent vectors and uncorrelated, but dependent, random vectors. In this section, we provide an answer to this question by giving examples of marginally uncorrelated Gaussian random vectors with non-vanishing ordinal pattern dependence. For this purpose, we initially characterize ordinal pattern dependence of order 1 for Gaussian random vectors.

Proposition 5.1. *Let* $\mathbf{X}^{(1)} := (X_1, X_2)$ *and* $\mathbf{Y}^{(1)} := (Y_1, Y_2)$ *be two Gaussian random vectors. Then, it holds that*

$$\mathbf{OPD}_1(\mathbf{X}^{(1)}, \mathbf{Y}^{(1)}) = \frac{2}{\pi} \arcsin \operatorname{Corr}(X_2 - X_1, Y_2 - Y_1).$$

Proof. Recall that

$$\begin{aligned} &\mathbf{OPD}_1(\mathbf{X}^{(1)}, \mathbf{Y}^{(1)}) \\ &= \frac{P\left(\Pi(X_1, X_2) = \Pi(Y_1, Y_2)\right) - \sum_{\pi \in S_1} P\left(\Pi(X_1, X_2) = \pi\right) P\left(\Pi(Y_1, Y_2) = \pi\right)}{1 - \sum_{\pi \in S_1} P\left(\Pi(X_1, X_2) = \pi\right) P\left(\Pi(Y_1, Y_2) = \pi\right)}. \end{aligned}$$

Moreover, for Gaussian observations, we have

$$P\left(X_i < X_{i+1}\right) = \frac{1}{2},$$

such that $\sum_{\pi \in S_1} P\left(\Pi(X_i, X_{i+1}) = \pi\right) P\left(\Pi(Y_i, Y_{i+1}) = \pi\right) = \frac{1}{2}$. Therefore, the ordinal pattern dependence reduces to

$$\mathbf{OPD}_1(\mathbf{X}^{(1)}, \mathbf{Y}^{(1)}) = 2\left(P\left(\Pi(X_1, X_2) = \Pi(Y_1, Y_2)\right) - \frac{1}{2}\right).$$

It holds that

$$\begin{aligned} P\left(X_1 \leq X_2, Y_1 \leq Y_2\right) &= P\left(0 \leq X_2 - X_1, 0 \leq Y_2 - Y_1\right) \\ &= \frac{1}{4} + \frac{1}{2\pi} \arcsin \operatorname{Corr}(X_2 - X_1, Y_2 - Y_1) \end{aligned}$$

according to the "orthant probabilities"-formula for Gaussian random variables.

All in all, it follows that

$$\begin{aligned} \mathbf{OPD}_1(\mathbf{X}^{(1)}, \mathbf{Y}^{(1)}) &= 2\left(P\left(X_1 \leq X_2, Y_1 \leq Y_2\right) + P\left(X_1 \geq X_2, Y_1 \geq Y_2\right) - \frac{1}{2}\right) \\ &= \frac{2}{\pi} \arcsin \operatorname{Corr}(X_2 - X_1, Y_2 - Y_1). \end{aligned}$$

□

The following Lemma gives an example of marginally uncorrelated Gaussian random vectors with non-vanishing ordinal pattern dependence of order 1.

Lemma 5.2. *Let* W_i*,* $i \geq 1$*, be a multivariate AR(1)-process defined by*

$$W_i := \begin{pmatrix} X_i \\ Y_i \end{pmatrix}, \ i \geq 1,$$

with $W_i = AW_{i-1} + \xi_i$*,*

$$A := \begin{pmatrix} a & b \\ c & d \end{pmatrix} \text{ and } \xi_i := \begin{pmatrix} \varepsilon_i \\ \eta_i \end{pmatrix},$$

where either '$a = d$ *and* $b = -c$*' or '*$a = -d$ *and* $b = c$*'. We assume that all the eigenvalues of* A *are less than 1 in absolute value and that* ξ_i*,* $i \geq 1$*, are bivariate Gaussian random vectors with covariance matrix* $\Sigma_\xi = I_2$ *(with* I_2 *denoting the identity matrix). Moreover, we* *assume that* $\operatorname{Cov}(X_1, Y_1) = 0$ *and* $\sigma^2 := \operatorname{Var}(X_1) = \operatorname{Var}(Y_1)$. *Then, for* $\mathbf{X}^{(1)} := (X_1, X_2)$ *and* $\mathbf{Y}^{(1)} := (Y_1, Y_2)$, *it holds that*

$$\mathbf{OPD}_1(\mathbf{X}^{(1)}, \mathbf{Y}^{(1)}) = \frac{2}{\pi} \arcsin\left(-\frac{b+c}{1-a^2-b^2}\right),$$

where $c^2 = b^2$ *and* $a^2 = d^2$.

Remark 5.1. The choice of the matrix A guarantees that $\operatorname{Cov}(X_2, Y_2) = 0$. With respect to the entries of A, we distinguish the following cases:

(1) If $b = -c$, we obtain $\operatorname{Cov}(X_2 - X_1, Y_2 - Y_1) = 0$.
(2) If $b = c$ and $d = -a$, it follows that

$$\operatorname{Corr}(X_2 - X_1, Y_2 - Y_1) = -\frac{b}{\sqrt{1-a^2}}.$$

In the second case, the correlation of the increments becomes big in absolute value for $a^2 + b^2$ close to 1. It approaches 1 if b is negative and -1 if b is positive.

Proof. According to Proposition 5.1, we have

$$\mathbf{OPD}_1(\mathbf{X}^{(1)}, \mathbf{Y}^{(1)}) = \frac{2}{\pi} \arcsin \operatorname{Corr}(X_2 - X_1, Y_2 - Y_1).$$

Since $\operatorname{Cov}(X_1, Y_1) = 0$, it follows that

$$\begin{aligned} \operatorname{Cov}(X_2 - X_1, Y_2 - Y_1) &= (c(a-1) + b(d-1))\,\sigma^2, \\ \operatorname{Var}(X_2 - X_1) &= \left((a-1)^2 + b^2\right)\sigma^2 + 1, \\ \operatorname{Var}(Y_2 - Y_1) &= \left(c^2 + (d-1)^2\right)\sigma^2 + 1. \end{aligned}$$

Therefore, it remains to compute σ^2.

Due to stationarity we have

$$\begin{aligned} \sigma^2 &= \operatorname{Var}(X_2) = \left(a^2 + b^2\right)\sigma^2 + 1, \\ \sigma^2 &= \operatorname{Var}(Y_2) = \left(c^2 + d^2\right)\sigma^2 + 1. \end{aligned}$$

It follows that

$$\begin{aligned} \sigma^2 &= (1-a^2)^{-1}\left(b^2\sigma^2 + 1\right), \\ \sigma^2 &= (1-d^2)^{-1}(c^2\sigma^2 + 1), \end{aligned}$$

such that

$$\sigma^2 = \left(1 - d^2 - \frac{b^2c^2}{1-a^2}\right)^{-1}\left(\frac{c^2}{1-a^2} + 1\right).$$

As a result, and since $\operatorname{Cov}(X_i, Y_i) = 0$,

$$0 = (ac + bd)\,\sigma^2, \text{ i.e., } d = -\frac{ac}{b}.$$

Therefore, we have $a^2 + b^2 = c^2 + d^2 = c^2(1 + \frac{a^2}{b^2})$.

It follows that $c^2 = b^2$ and $a^2 = d^2$ and, therefore,

$$\sigma^2 = (1 - a^2 - b^2)^{-1}.$$

All in all, we arrive at

$$\mathrm{Corr}(X_2 - X_1, Y_2 - Y_1) = -\frac{b+c}{1-a^2-b^2},$$

which proves

$$\mathbf{OPD}_1(\mathbf{X}^{(1)}, \mathbf{Y}^{(1)}) = \frac{2}{\pi} \arcsin\left(-\frac{b+c}{1-a^2-b^2}\right).$$

□

We illustrate our results by simulating a bivariate AR(1)-process

$$W_i := \begin{pmatrix} X_i \\ Y_i \end{pmatrix}, \ i = 1, \dots, 500,$$

with $W_i = AW_{i-1} + \xi_i$, where

$$A := \begin{pmatrix} a & b \\ b & -a \end{pmatrix} \text{ and } \xi_i := \begin{pmatrix} \varepsilon_i \\ \eta_i \end{pmatrix}, \tag{4}$$

ξ_i being a multivariate Gaussian random vector with covariance matrix $\Sigma_\xi = I_2$ (with I_2 denoting the identity matrix). We choose $a^2 + b^2 < 1$, but close to 1, in order to obtain $\mathrm{Cov}(X_i, Y_i) = 0$, but high ordinal pattern dependence. For the simulations summarized by the boxplots in Figure 1 we chose $a = 0.7$ and $b = -0.7$. Clearly the median of the boxplots that are based on the values of Pearson's correlation coefficient approaches zero, while the median of the boxplots that are based on the values of ordinal pattern dependence seem to converge to a value between 0.75 and 1.

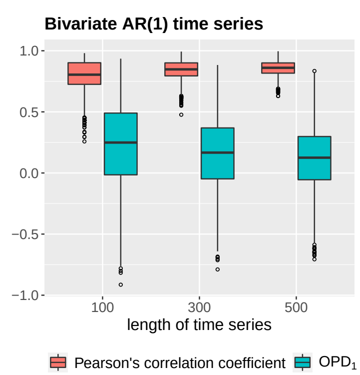

Figure 1. Boxplots of Pearson's correlation coefficient and ordinal pattern dependence of order $h = 1$ based on 5000 repetitions of a bivariate AR(1)-process (X_i, Y_i), $i = 1, \dots, n$, satisfying (4) with $a = 0.7$ and $b = -0.7$.

Figure 2 depicts one sample path of the single time series X_i, $i = 1, \dots, 500$, and Y_i, $i = 1, \dots, 500$, the corresponding increment processes, as well as scatterplots of the original observations and their increments. The scatterplots clearly indicate that, while the original observations are uncorrelated, the increment processes are positively correlated.

Moreover, the scatterplots of the two processes and their increments in Figure 2 underline uncorrelatedness of the original processes and a high dependence of their increments.

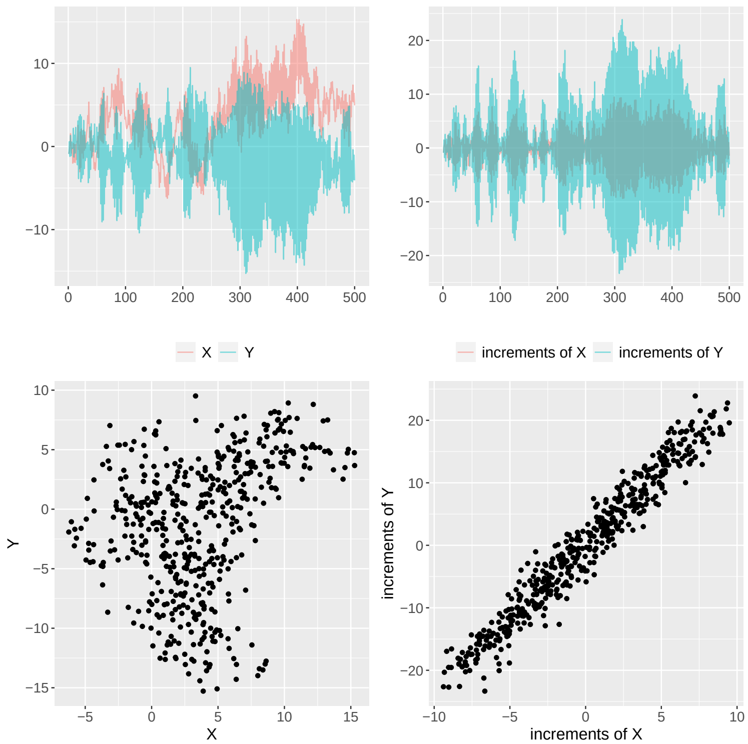

FIGURE 2. Sample paths of X_i, $i = 1, \ldots, 500$, and Y_i, $i = 1, \ldots, 500$, their increments at lag 1, as well as corresponding scatterplots based on a bivariate AR(1)-process (X_i, Y_i), $i = 1, \ldots, n$, satisfying (4) with $a = 0.7$ and $b = -0.7$.

5.2. **The case $h = 2$.** For the computation of the ordinal pattern dependence of order $h = 1$, the crucial quantity is $\mathrm{Corr}(X_2 - X_1, Y_2 - Y_1)$ since, according to Proposition 5.1, $\mathbf{OPD}_1(\mathbf{X}^{(1)}, \mathbf{Y}^{(1)})$ is just a monotone transformation of this correlation. It is, therefore, natural to wonder whether it is possible to construct a stationary, bivariate process $(X_i, Y_i)_{i\geq 1}$ with $\mathbf{OPD}_1(\mathbf{X}^{(1)}, \mathbf{Y}^{(1)}) = 0$, but $\mathbf{OPD}_2(\mathbf{X}^{(2)}, \mathbf{Y}^{(2)}) \neq 0$.

The AR(1)-process in Lemma 5.2 does not fulfill these conditions, since the restriction

$$\mathrm{Corr}(X_2 - X_1, Y_2 - Y_1) = -\frac{b}{\sqrt{1 - a^2}} = 0$$

implies $b = 0$. As a result, we obtain a process $W_i = (X_i, Y_i)^t = (aX_{i-1} + \xi_i, -aY_{i-1} + \eta_i)^t$, that does not incorporate any *dynamical* dependence between the processes X_i, $i \geq 1$, and Y_i, $i \geq 1$. The only dependence in this model exists within each component, yet, this does not have an impact on ordinal pattern dependence. Formally, these considerations correspond to the following lemma:

Lemma 5.3. *Let W_i, $i \geq 1$, be a multivariate AR(1)-process defined by*

$$W_i := \begin{pmatrix} X_i \\ Y_i \end{pmatrix}, \quad i \geq 1,$$

with $W_i = AW_{i-1} + \xi_i$*, where*

$$A := \begin{pmatrix} a & b \\ c & d \end{pmatrix} \text{ and } \xi_i := \begin{pmatrix} \varepsilon_i \\ \eta_i \end{pmatrix},$$

ξ_i*,* $i \geq 1$*, bivariate Gaussian random vectors with covariance matrix* $\Sigma_\xi = I_2$ *(with* I_2 *denoting the identity matrix). Moreover, we assume that* $\mathrm{Cov}(X_i, Y_i) = 0$ *for* $i = 1, 2$*,* $\mathbf{OPD}_1(\mathbf{X}^{(1)}, \mathbf{Y}^{(1)}) = 0$*, and* $\sigma^2 := \mathrm{Var}(X_1) = \mathrm{Var}(Y_1)$*. Then, for* $\mathbf{X}^{(2)} := (X_1, X_2, X_3)$ *and* $\mathbf{Y}^{(2)} := (Y_1, Y_2, Y_3)$*, it holds that*

$$\mathbf{OPD}_h(\mathbf{X}^{(h)}, \mathbf{Y}^{(h)}) = 0.$$

Proof. The assumptions $\mathrm{Cov}(X_1, Y_1) = \mathrm{Cov}(X_2, Y_2) = 0$ and $\mathbf{OPD}_1(\mathbf{X}^{(1)}, \mathbf{Y}^{(1)}) = 0$ imply that $b = c = 0$ and $a = d$. As a result, $\mathbf{X}^{(h)}$ is independent of $\mathbf{Y}^{(h)}$ for every $h \in \mathbb{N}$. Hence, we have $\mathbf{OPD}_h(\mathbf{X}^{(h)}, \mathbf{Y}^{(h)}) = 0$ for every $h \in \mathbb{N}$. $\square$

Due to Lemma 5.3, AR(1)-processes do not constitute a suitable model for bivariate, stationary Gaussian processes fulfilling $\mathrm{Corr}\,(X_i, Y_i) = 0$ for $i = 1, 2$, and $\mathbf{OPD}_1(\mathbf{X}^{(1)}, \mathbf{Y}^{(1)}) = 0$, but $\mathbf{OPD}_2(\mathbf{X}^{(2)}, \mathbf{Y}^{(2)}) \neq 0$.

Nonetheless, certain AR(2)-processes yield an example:

Example 5.1. Let W_i, $i \geq 1$, be a multivariate AR(2)-process defined by

$$W_i := \begin{pmatrix} X_i \\ Y_i \end{pmatrix}, \ i \geq 1,$$

where $W_i = AW_{i-2} + \xi_i$ with

$$A := \begin{pmatrix} a & b \\ b & -a \end{pmatrix} \text{ and } \xi_i := \begin{pmatrix} \varepsilon_i \\ \eta_i \end{pmatrix}, \tag{5}$$

ξ_i, $i \geq 1$, bivariate Gaussian random vectors with covariance matrix $\Sigma_\xi = I_2$ (with I_2 denoting the identity matrix) and $X_1 := \xi_1$, $Y_1 = \eta_1$, $X_2 := \xi_2$, $Y_2 = \eta_2$. Moreover, we assume that $\sigma^2 := \mathrm{Var}(X_1) = \mathrm{Var}(Y_1)$.

By definition it holds that $\mathrm{Cov}(X_1, Y_1) = \mathrm{Cov}(X_2, Y_2) = 0$. Moreover, we have $\mathbf{OPD}_1(\mathbf{X}^{(1)}, \mathbf{Y}^{(1)}) = 0$ since

$$\begin{aligned} \mathrm{Cov}(X_3 - X_2, Y_3 - Y_2) &= \mathbb{E}\left[(aX_1 + bY_1 + \xi_3 - X_2)\,(bX_1 - aY_1 + \eta_3 - Y_2)\right] \\ &= ab\sigma^2 - ba\sigma^2 = 0. \end{aligned}$$

In order to compute $\mathbf{OPD}_2(\mathbf{X}^{(2)}, \mathbf{Y}^{(h)})$, we have to calculate $P\left(\Pi(X_1, X_2, X_3) = \pi, = \Pi(Y_1, Y_2, Y_3) = \pi\right)$ for every $\pi \in S_2$. With $\pi = (0, 1, 2)$ it follows that

$$\begin{aligned} &P\left(\Pi(X_1, X_2, X_3) = \pi, = \Pi(Y_1, Y_2, Y_3) = \pi\right) \\ =&P\left(X_1 \leq X_2 \leq X_3, Y_1 \leq Y_2 \leq Y_3\right) \\ =&P\left(X_2 - X_1 \geq 0, X_3 - X_2 \geq 0, Y_2 - Y_1 \geq 0, Y_3 - Y_2 \geq 0\right). \end{aligned}$$

As a result, computing $\mathbf{OPD}_2(\mathbf{X}^{(2)}, \mathbf{Y}^{(2)})$ boils down to determining the orthant probabilities of a four-dimensional Gaussian vector. To our knowledge a closed expression for these probabilities is not at hand; see Abrahamson et al. (1964). For this reason, we consider estimated values for $\mathbf{OPD}_2(\mathbf{X}^{(2)}, \mathbf{Y}^{(2)})$ based on simulations of a bivariate AR(2)-process that satisfies the above assumptions; see Figure 3. The corresponding boxplots clearly indicate that, as the length of the time series increases, the ordinal pattern dependence of order $h = 1$ approaches zero, while the ordinal pattern dependence of order $h = 2$ converges to a a value between 0.1 and 0.25.

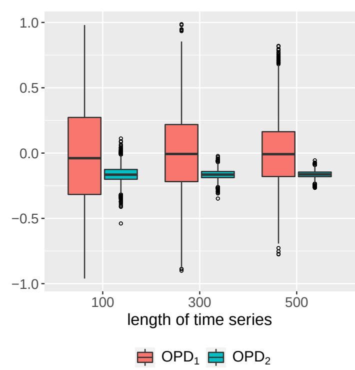

Figure 3. Boxplots of ordinal pattern dependence of order $h = 1$ and $h = 2$ based on 5000 repetitions of a bivariate AR(2)-process (X_i, Y_i), $i = 1, \ldots, n$, satisfying (5) with $a = 0.01$ and $b = 0.98$. Boxplots of $\mathbf{OPD}_1$ and $\mathbf{OPD}_2$.

5.3. **Ordinal pattern dependence in contrast to multivariate Kendall's τ.** Recall that for Gaussian random vectors $\mathbf{X}^{(1)} = (X_1, X_2)$ and $\mathbf{Y}^{(1)} = (Y_1, Y_2)$

$$\mathbf{OPD}_1(\mathbf{X}^{(1)}, \mathbf{Y}^{(1)}) = \frac{2}{\pi} \arcsin\left(\mathrm{Corr}(X_2 - X_1, Y_2 - Y_1)\right)$$

according to Proposition 5.1.

It follows that

$$\mathbf{OPD}_1(\mathbf{X}^{(1)}, \mathbf{Y}^{(1)}) = \frac{2}{\pi} \arcsin\left(\sin\left(\frac{\pi}{2}\tau_1\left(X_2 - X_1, Y_2 - Y_1\right)\right)\right) = \tau_1\left(\tilde{X}_1, \tilde{Y}_1\right),$$

where $\tilde{X}_1 := X_2 - X_1$ and $\tilde{Y}_1 := Y_2 - Y_1$; see Kruskal (1958). For this reason, in the case $h = 1$, ordinal pattern dependence is nothing but univariate Kendall's τ of the increment processes.

In general, the following proposition establishes a relation between ordinal pattern dependence and multivariate Kendall's τ for Gaussian processes:

Proposition 5.4. *Let (X_i, Y_i), $i \geq 1$, be a bivariate, stationary Gaussian process. Then, it holds that*

$$\begin{aligned}&\mathbf{OPD}_h\left(\mathbf{X}^{(h)}, \mathbf{Y}^{(h)}\right)\\ &= \frac{2\tau_h\left(\tilde{X}_1, \ldots, \tilde{X}_h, \tilde{Y}_1, \ldots, \tilde{Y}_h\right)\sqrt{p_{\tilde{X}}\left(1 - p_{\tilde{X}}\right)p_{\tilde{Y}}\left(1 - p_{\tilde{Y}}\right)}}{1 - \sum_{\pi \in S_h} P\left(\Pi\left(\mathbf{X}^{(h)}\right) = \pi\right)P\left(\Pi\left(\mathbf{Y}^{(h)}\right) = \pi\right)}\\ &\quad + \sum_{\pi \in S_h \setminus T_h} \frac{P\left(\Pi\left(\mathbf{X}^{(h)}\right) = \Pi\left(\mathbf{Y}^{(h)}\right) = \pi\right) - P\left(\Pi\left(\mathbf{X}^{(h)}\right) = \pi\right)P\left(\Pi\left(\mathbf{Y}^{(h)}\right) = \pi\right)}{1 - \sum_{\pi \in S_h} P\left(\Pi\left(\mathbf{X}^{(h)}\right) = \pi\right)P\left(\Pi\left(\mathbf{Y}^{(h)}\right) = \pi\right)},\end{aligned}$$

where $T_h := \{(h, h-1, \ldots, 0), (0, 1, \ldots, h)\}$, $\tilde{X}_i = X_{i+1} - X_i$, $\tilde{Y}_i = Y_{i+1} - Y_i$,

$$p_{\tilde{X}} = P\left(\Pi\left(X_1, \ldots, X_{1+h}\right) = (0, 1, \ldots, h)\right) = P\left(\tilde{X}_1 \leq 0, \ldots, \tilde{X}_h \leq 0\right),$$

$$p_{\tilde{Y}} = P\left(\Pi\left(Y_1, \ldots, Y_{1+h}\right) = (0, 1, \ldots, h)\right) = P\left(\tilde{Y}_1 \leq 0, \ldots, \tilde{Y}_h \leq 0\right),$$

and

$$\begin{aligned}&\mathbf{OPD}_h\left(\mathbf{X}^{(h)}, \mathbf{Y}^{(h)}\right)\\ =&\frac{\sum_{\pi \in S_h} \tau_h\left(X_{1+\pi_2} - X_{1+\pi_1}, \ldots, Y_{1+\pi_{1+h}} - Y_{1+\pi_h}\right)\sqrt{p_{X,\pi}\left(1 - p_{X,\pi}\right) p_{Y,\pi}\left(p_{Y,\pi}\right)}}{1 - \sum_{\pi \in S_h} p_{X,\pi}, p_{Y,\pi}}\end{aligned}$$

with

$$p_{X,\pi} = P\left(\Pi\left(X_1, \ldots, X_{1+h}\right) = \pi\right),\ p_{Y,\pi} = P\left(\Pi\left(Y_1, \ldots, Y_{1+h}\right) = \pi\right).$$

Remark 5.2. It is a characteristic of Gaussian observations (X_i), $i \geq 1$, and (Y_i), $i \geq 1$, that the distribution of

$$\left(X_{1+\pi_2} - X_{1+\pi_1}, \ldots, X_{1+\pi_{1+h}} - X_{1+\pi_h}, Y_{1+\pi_2} - Y_{1+\pi_1}, \ldots, Y_{1+\pi_{1+h}} - Y_{1+\pi_h}\right)^t$$

is uniquely determined by the autocovariances and crosscovariances of $\mathbf{X}^{(h)}$ and $\mathbf{Y}^{(h)}$. For this reason, it is possible to express all the dependencies in the vector above by the two-dimensional marginals of a multivariate Gaussian distribution. However, since we do not have a closed expression for orthant probabilities of a multivariate Gaussian vector with more than 3 elements, it is not possible to constitute a closed form for ordinal pattern dependence in terms of Kendall's τ.

Proof. First, note that

$$\begin{aligned}&\mathbf{OPD}_h\left(\mathbf{X}^{(h)}, \mathbf{Y}^{(h)}\right)\\ =&\frac{\sum_{\pi \in T_h} P\left(\Pi\left(\mathbf{X}^{(h)}\right) = \Pi\left(\mathbf{Y}^{(h)}\right) = \pi\right) - P\left(\Pi\left(\mathbf{X}^{(h)}\right) = \pi\right) P\left(\Pi\left(\mathbf{Y}^{(h)}\right) = \pi\right)}{1 - \sum_{\pi \in S_h} P\left(\Pi\left(\mathbf{X}^{(h)}\right) = \pi\right) P\left(\Pi\left(\mathbf{Y}_m^{(h)}\right) = \pi\right)}\\ &+\frac{\sum_{\pi \in S_h \setminus T_h} P\left(\Pi\left(\mathbf{X}^{(h)}\right) = \Pi\left(\mathbf{Y}^{(h)}\right) = \pi\right) - P\left(\Pi\left(\mathbf{X}^{(h)}\right) = \pi\right) P\left(\Pi\left(\mathbf{Y}^{(h)}\right) = \pi\right)}{1 - \sum_{\pi \in S_h} P\left(\Pi\left(\mathbf{X}^{(h)}\right) = \pi\right) P\left(\Pi\left(\mathbf{Y}^{(h)}\right) = \pi\right)}.\end{aligned}$$

Focusing on the pattern $\pi = (0, 1, \ldots, h)$ in the first summand yields

$$\begin{aligned}&P\left(\Pi\left(\mathbf{X}^{(h)}\right) = \Pi\left(\mathbf{Y}^{(h)}\right) = (0, 1, \ldots, h)\right)\\ =&P\left(X_1 \geq X_2 \geq \ldots \geq X_{1+h}, Y_1 \geq Y_2 \geq \ldots \geq Y_{1+h}\right)\\ =&P\left(\tilde{X}_1 \leq 0, \ldots, \tilde{X}_h \leq 0, \ldots, \tilde{Y}_1 \leq 0, \ldots, \tilde{Y}_h \leq 0\right)\\ =&\tau_h\left(\tilde{X}_1, \ldots, \tilde{X}_h, \tilde{Y}_1, \ldots, \tilde{Y}_h\right)\sqrt{p_{\tilde{X}}\left(1 - p_{\tilde{X}}\right) p_{\tilde{Y}}\left(1 - p_{\tilde{Y}}\right)} + p_{\tilde{X}} p_{\tilde{Y}}.\end{aligned}$$

Due to symmetry of the multivariate normal distribution, we have $\left(\mathbf{X}^{(h)}, \mathbf{Y}^{(h)}\right)^t \overset{\mathcal{D}}{=} \left(-\mathbf{X}^{(h)}, -\mathbf{Y}^{(h)}\right)^t$. Therefore, it follows that

$$P\left(\Pi\left(\mathbf{X}^{(h)}\right) = \Pi\left(\mathbf{Y}^{(h)}\right) = (0, 1, \ldots, h)\right) = P\left(\Pi\left(\mathbf{X}^{(h)}\right) = \Pi\left(\mathbf{Y}^{(h)}\right) = (h, h-1, \ldots, 0)\right).$$

Finally, we obtain

$$\begin{aligned}\mathbf{OPD}_h\left(\mathbf{X}^{(h)},\mathbf{Y}^{(h)}\right) =& \frac{2\tau_h\left(\tilde{X}_1,\ldots,\tilde{X}_h,\tilde{Y}_1,\ldots,\tilde{Y}_h\right)\sqrt{p_{\tilde{X}}\left(1-p_{\tilde{X}}\right)p_{\tilde{Y}}\left(1-p_{\tilde{Y}}\right)}}{1-\sum_{\pi\in S_h}P\left(\Pi\left(\mathbf{X}^{(h)}\right)=\pi\right)P\left(\Pi\left(\mathbf{Y}^{(h)}\right)=\pi\right)} \\ &+ \frac{\sum_{\pi\in S_h\setminus T_h}P\left(\Pi\left(\mathbf{X}^{(h)}\right)=\Pi\left(\mathbf{Y}^{(h)}\right)=\pi\right)}{1-\sum_{\pi\in S_h}P\left(\Pi\left(\mathbf{X}^{(h)}\right)=\pi\right)P\left(\Pi\left(\mathbf{Y}^{(h)}\right)=\pi\right)} \\ &- \frac{\sum_{\pi\in S_h\setminus T_h}P\left(\Pi\left(\mathbf{X}^{(h)}\right)=\pi\right)P\left(\Pi\left(\mathbf{Y}^{(h)}\right)=\pi\right)}{1-\sum_{\pi\in S_h}P\left(\Pi\left(\mathbf{X}^{(h)}\right)=\pi\right)P\left(\Pi\left(\mathbf{Y}^{(h)}\right)=\pi\right)}.\end{aligned}$$

Let $\pi=(\pi_1,\ldots,\pi_{1+h})$ be a permutation in S_h. If $\Pi(X_1,\ldots,X_{1+h})=\pi$, it holds that $\left\{X_{1+\pi_1}\geq X_{1+\pi_2}\geq\ldots\geq X_{1+\pi_{1+h}}\right\}$. As a result, ordinal pattern dependence can be expressed by the following formula:

$$\begin{aligned}&\mathbf{OPD}_h\left(\mathbf{X}^{(h)},\mathbf{Y}^{(h)}\right) \\ =&\frac{\sum_{\pi\in S_h}P\left(\Pi\left(X_1,\ldots,X_{1+h}\right)=\Pi\left(Y_1,\ldots,Y_{1+h}\right)=\pi\right)-p_{X,\pi}p_{Y,\pi}}{1-\sum_{\pi\in S_h}P\left(\Pi\left(X_1,\ldots,X_{1+h}\right)=\pi\right)P\left(\Pi\left(Y_1,\ldots,Y_{1+h}\right)=\pi\right)} \\ =&\frac{\sum_{\pi\in S_h}P\left(X_{1+\pi_1}\geq X_{1+\pi_2}\geq\ldots\geq X_{1+\pi_{1+h}},Y_{1+\pi_1}\geq Y_{1+\pi_2}\geq\ldots\geq Y_{1+\pi_{1+h}}\right)-p_{X,\pi}p_{Y,\pi}}{1-\sum_{\pi\in S_h}p_{X,\pi},p_{Y,\pi}} \\ =&\frac{\sum_{\pi\in S_h}P\left(X_{1+\pi_2}-X_{1+\pi_1}\leq 0,\ldots,X_{1+\pi_{1+h}}-X_{1+\pi_h}\leq 0,\ldots,Y_{1+\pi_{1+h}}-Y_{1+\pi_h}\leq 0\right)-p_{X,\pi}p_{Y,\pi}}{1-\sum_{\pi\in S_h}p_{X,\pi},p_{Y,\pi}} \\ =&\frac{\sum_{\pi\in S_h}\tau_h\left(X_{1+\pi_2}-X_{1+\pi_1},\ldots,Y_{1+\pi_{1+h}}-Y_{1+\pi_h}\right)\sqrt{p_{X,\pi}\left(1-p_{X,\pi}\right)p_{Y,\pi}\left(p_{Y,\pi}\right)}}{1-\sum_{\pi\in S_h}p_{X,\pi},p_{Y,\pi}}\end{aligned}$$

with

$$\begin{aligned}p_{X,\pi}&=P\left(\Pi\left(X_1,\ldots,X_{1+h}\right)=\pi\right)\\ &=P\left(X_{1+\pi_1}\geq X_{1+\pi_2}\geq\ldots\geq X_{1+\pi_{1+h}}\right)\\ &=P\left(X_{1+\pi_2}-X_{1+\pi_1}\leq 0,\ldots,X_{1+\pi_{1+h}}-X_{1+\pi_h}\leq 0\right)\end{aligned}$$

and

$$\begin{aligned}p_{Y,\pi}&=P\left(\Pi\left(Y_1,\ldots,Y_{1+h}\right)=\pi\right)\\ &=P\left(Y_{1+\pi_1}\geq Y_{1+\pi_2}\geq\ldots\geq Y_{1+\pi_{1+h}}\right)\\ &=P\left(Y_{1+\pi_2}-Y_{1+\pi_1}\leq 0,\ldots,Y_{1+\pi_{1+h}}-Y_{1+\pi_h}\leq 0\right).\end{aligned}$$

□

In order to illustrate the relation of multivariate Kendall's τ and ordinal pattern dependence, characterized through Proposition 5.4, we consider the case $h=1$ in the following example:

Example 5.2. Recall that for Gaussian random vectors $\mathbf{X}^{(1)}=(X_1,X_2)$ and $\mathbf{Y}^{(1)}=(Y_1,Y_2)$, it holds that

$$\mathbf{OPD}_1(\mathbf{X}^{(1)},\mathbf{Y}^{(1)})=\tau_1\left(X_2-X_1,Y_2-Y_1\right)=\frac{2}{\pi}\arcsin\left(\mathrm{Corr}(X_2-X_1,Y_2-Y_1)\right)$$

according to Proposition 5.1. Moreover, if $\mathbf{X}^{(1)}$ and $\mathbf{Y}^{(1)}$ have standard normal marginal distributions, it holds that

$$\operatorname{Corr}(X_2 - X_1, Y_2 - Y_1) = \frac{2\mathbb{E}\left(X_1 Y_1\right) - \mathbb{E}\left(X_1 Y_2\right) - \mathbb{E}\left(Y_1 X_2\right)}{2 + 2\mathbb{E}\left(X_1 X_2\right)}.$$

In general, we know that $\tau_1(X, Y) = \frac{2}{\pi} \arcsin\left(\operatorname{Corr}(X, Y)\right)$ for a Gaussian random vector $(X, Y)^t$, and hence $\operatorname{Corr}(X, Y) = \sin\left(\frac{\pi}{2}\tau_1(X, Y)\right)$.
As a result, we obtain

$$\begin{aligned} &\mathbf{OPD}_1\left(\mathbf{X}_1^{(1)}, \mathbf{Y}_1^{(1)}\right) \\ &= \frac{2}{\pi}\arcsin\left(\frac{2\sin\left(\frac{\pi}{2}\tau_1(X_1, Y_1)\right) - \sin\left(\frac{\pi}{2}\tau_1(X_1, Y_2)\right) - \sin\left(\frac{\pi}{2}\tau_1(X_2, Y_1)\right)}{2 + 2\sin\left(\frac{\pi}{2}\tau_1(X_1, X_2)\right)}\right). \end{aligned}$$

Therefore, the ordinal pattern dependence of order 1 is determined by $\tau_1\left(X_1, Y_1\right)$, $\tau_1\left(X_1, Y_2\right)$, $\tau_1\left(X_2, Y_1\right)$, and $\tau_1\left(X_2, Y_2\right)$.

5.4. **Simulation Study.** In this section, we compare the estimators for ordinal pattern dependence and multivariate Kendall's τ based on the vectors $\mathbf{X}_i^{(2)} = (X_i, X_{i+1}, X_{i+2})$, $\mathbf{Y}_i^{(2)} = (Y_i, Y_{i+1}, Y_{i+2})$, $i = 1, \ldots, n-2$, generated by bivariate processes (X_i, Y_i), $i \geq 1$, in a simulation study. The following situations are considered:

(1) We simulate (X_i), $i \geq 1$, and (Y_i), $i \geq 1$, as two independent AR(1) time series with

$$X_i = \rho X_{i-1} + \varepsilon_i,\ Y_i = \rho Y_{i-1} + \eta_i,$$

where $|\rho| < 1$, and (ε_i), $i \geq 1$, and (η_i), $i \geq 1$, are two independent sequences of random variables, both i.i.d. standard normally distributed. In this case, the sequences (X_i), $i \geq 1$, and (Y_i), $i \geq 1$, are generated by the function `arima.sim` in `R`. For the simulations depicted in Figure 5.4, we chose $\rho = 0.5$. As expected, the values of both dependence measures vary around 0. Moreover, the boxplots become narrower as the sample size increases confirming consistency of the estimators. The boxplots that correspond to the estimate for Kendall's τ are wider. This indicates a faster convergence of the estimators for ordinal pattern dependence.

(2) We simulate (X_i), $i \geq 1$, and (Y_i), $i \geq 1$, as sequences of independent, multivariate normal random vectors with values in $\mathbb{R}^3$ and a joint normal distribution with expectation 0 and covariance matrix

$$\Sigma = \begin{pmatrix} 1 & 0 & 0 & \rho & \rho & \rho \\ 0 & 1 & 0 & \rho & \rho & \rho \\ 0 & 0 & 1 & \rho & \rho & \rho \\ \rho & \rho & \rho & 1 & 0 & 0 \\ \rho & \rho & \rho & 0 & 1 & 0 \\ \rho & \rho & \rho & 0 & 0 & 1 \end{pmatrix}.$$

The $\mathbb{R}^6$-valued random vectors (X_i, Y_i), $i \geq 1$, are generated by the function `rmvnorm` in `R`. For the simulations depicted in Figure 5.4, we chose $\rho = 0.2$. Note that the values of both dependence measures deviate from 0 thereby indicating a correlation between the two processes (X_i), $i \geq 1$, and (Y_i), $i \geq 1$. In fact, the boxplots of both estimators look very similar so that the rates of convergence seem to be comparable.

(3) We simulate (X_i, Y_i), $i \geq 1$, as a bivariate AR(1) process, i.e., we generate two independent, standard normally distributed random variables X_1 and Y_1 and

$$\begin{pmatrix} X_i \\ Y_i \end{pmatrix} = A \begin{pmatrix} X_{i-1} \\ Y_{i-1} \end{pmatrix} + \begin{pmatrix} \eta_i \\ \xi_i \end{pmatrix},$$

where

$$A = \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}$$

and (ε_i), $i \geq 1$, and (η_i), $i \geq 1$, are two independent sequences of random variables, both i.i.d. standard normally distributed. For the simulations depicted in Figure 5.4, we chose $\rho = 0.2$. We note that the median of the estimated values of Kendall's τ tends to 1 indicating strong dependence between the two time series, while the median of the estimated values of ordinal pattern dependence tends to a value between 0.3 and 0.4 indicating moderate dependence between the two time series. Moreover, the boxplots of the estimates for ordinal pattern dependence are narrower than those of the estimators for Kendall's τ. What catches the eye is that the interquartile distance of all boxplots does not change with increasing sample size. Furthermore, additional simulation results show that a change of the parameter ρ does not seem to affect the boxplots corresponding to the estimators of the ordinal pattern dependence. However, the values of the estimators for Kendall's τ clearly converge to 1, while their boxplots get narrower as ρ increases.

(4) We simulate (X_i), $i \geq 1$, as an AR(1) time series, while (Y_i), $i \geq 1$, corresponds to (X_i), $i \geq 1$, shifted by 1 time point. More precisely, we simulate

$$X_i = \rho X_{i-1} + \varepsilon_i,$$

where $|\rho| < 1$, and (ε_i), $i \geq 1$, is an i.i.d. standard normally distributed sequence of random variables, and we define $Y_i = X_{i+1}$. For the simulations depicted in Figure 5.4, we chose $\rho = 0.5$. These show that ordinal pattern dependence is sensitive to a shift in time in the sense that the values of the estimators for ordinal pattern dependence vary around 0. This does not hold for the estimator of Kendall's τ, as the corresponding values tend to be significantly higher than 0.5. In both cases, the boxplots are extremely narrow. This can be explained by the fact that (Y_i), $i \geq 1$, is fully determined by (X_i), $i \geq 1$.

5.5. **Conclusion.** We have shown that ordinal pattern dependence is a multivariate measure of dependence in an axiomatic sense. When applied to bivariate time series, it can be interpreted as a value describing the co-movement of the two component time series. In contrast to other dependence measures, it has thus been developed against a time series background. Univariate dependence measures do not carry enough information for an analysis of dependencies in a time series context while multivariate extensions of Pearson and Spearman proved not to be useful. For Gaussian observations, there is a close relationship between ordinal pattern dependence of order 1 and the multivariate version of Kendall's τ. However, simulations show that the values of Kendall's τ and ordinal pattern dependence differ in concrete situations. Since ordinal pattern dependence admits a canonical interpretation and limit theorems in the short and the long-range dependent framework are at hand, we suggest to use this measure in the context of time series analysis.

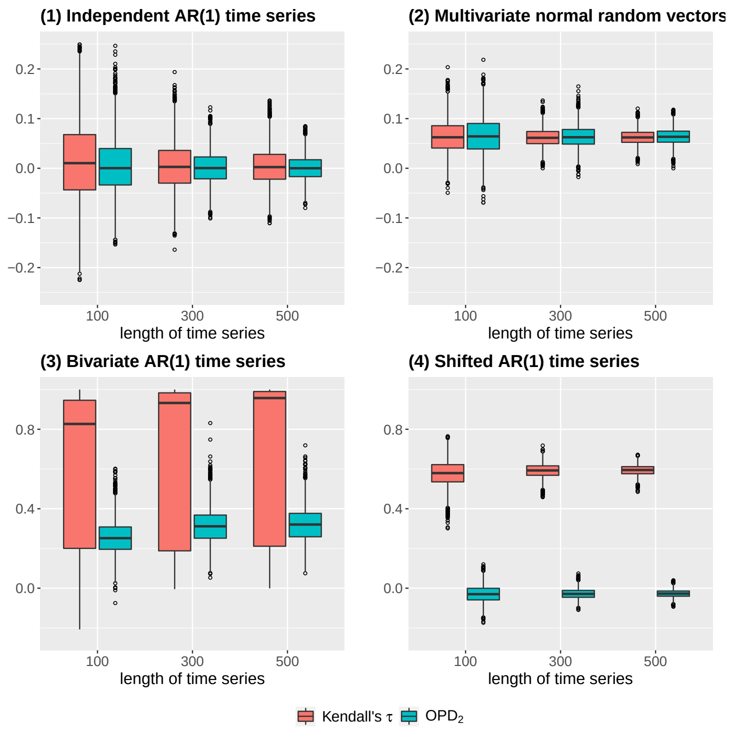

FIGURE 4. Boxplots of ordinal pattern dependence of order $h = 2$ based on 5000 repetitions of a bivariate process (X_i, Y_i), $i = 1, \ldots, n$, corresponding to the situations (1)-(4).

References

J. Aaronson, R. Burton, H. Dehling, D. Gilat, T. Hill, and B. Weiss. Strong laws for L- and U-statistics. *Transactions of the American Mathematical Society*, 348(7):2845 – 2866, 1996.

IG Abrahamson et al. Orthant probabilities for the quadrivariate normal distribution. *The Annals of Mathematical Statistics*, 35(4):1685 – 1703, 1964.

Christoph Bandt. Order patterns, their variation and change points in financial time series and brownian motion. *Statistical Papers*, pages 1–24, 2020.

Christoph Bandt and Bernd Pompe. Permutation Entropy: A Natural Complexity Measure for Time Series. *Physical review letters*, 88(17):174102–1 – 174102–4, 2002.

Christoph Bandt and Faten Shiha. Order Patterns in Time Series. *Journal of Time Series Analysis*, 28(5):646 – 665, 2007.

Annika Betken, Jannis Buchsteiner, Herold Dehling, Ines Münker, Alexander Schnurr, and Jeannette H. C. Woerner. Ordinal patterns in long-range dependent time series. *Scandinavian Journal of Statistics*, 2020.

Herold Dehling, Daniel Vogel, Martin Wendler, and Dominik Wied. Testing for changes in Kendall's tau. *Econometric Theory*, 33(6):1352 – 1386, 2017.

Oliver Grothe, Julius Schnieders, and Johan Segers. Measuring association and dependence between random vectors. *Journal of Multivariate Analysis*, 123:96 – 110, 2014.

Harry Joe. Multivariate concordance. *Journal of Multivariate Analysis*, 35(1):12 – 30, 1990.

William H. Kruskal. Ordinal measures of association. *Journal of the American Statistical Association*, 53(284):814 – 861, 1958.

Marisa Mohr, Florian Wilhelm, Mattis Hartwig, Ralf Möller, and Karsten Keller. New approaches in ordinal pattern representations for multivariate time series. In *FLAIRS Conference*, pages 124–129, 2020.

Giovanni Puccetti. Measuring Linear Correlation Between Random Vectors. *SSRN 3116066*, 2019.

Alexander Schnurr. An ordinal pattern approach to detect and to model leverage effects and dependence structures between financial time series. *Statistical Papers*, 55(4):919 – 931, 2014.

Alexander Schnurr and Herold Dehling. Testing for Structural Breaks via Ordinal Pattern Dependence. *Journal of the American Statistical Association*, 112(518):706 – 720, 2017.

Mathieu Sinn and Karsten Keller. Estimation of ordinal pattern probabilities in Gaussian processes with stationary increments. *Computational Statistics & Data Analysis*, 55(4): 1781 – 1790, 2011.

Jeffrey M. Wooldridge and Halbert White. Some invariance principles and central limit theorems for dependent heterogeneous processes. *Econometric Theory*, 4(2):210 – 230, 1988.

Fakultät für Mathematik, Ruhr-Universität Bochum, 44780 Bochum, Germany
Email address: annika.betken@rub.de

Fakultät für Mathematik, Ruhr-Universität Bochum, 44780 Bochum, Germany
Email address: herold.dehling@rub.de

Department Mathematik, Universität Siegen, 57068 Siegen, Germany
Email address: muenker@mathematik.uni-siegen.de

Department Mathematik, Universität Siegen, 57068 Siegen, Germany
Email address: schnurr@mathematik.uni-siegen.de