Ordinal pattern dependence as a multivariate dependence measure

Annika Betken [a], Herold Dehling [b], Ines Nüßgen [c], Alexander Schnurr [c],*

[a] *Faculty of Electrical Engineering, Mathematics and Computer Science, University of Twente, 7500 AE Enschede, The Netherlands*
[b] *Faculty of Mathematics, Ruhr-University Bochum, 44780 Bochum, Germany*
[c] *Department of Mathematics, Siegen University, 57072 Siegen, Germany*

ABSTRACT

In this article, we show that the recently introduced ordinal pattern dependence fits into the axiomatic framework of general multivariate dependence measures, i.e., measures of dependence between two multivariate random objects. 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 in this context. 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. This has to be separated from the branch of research where the dependence *within* one vector is described by a single number (see [13,14] and the references therein).

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 (see [15]) 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 (see [16] for the SRD case and [11] for the LRD case).

---

* Corresponding author.
    E-mail address: schnurr@mathematik.uni-siegen.de (A. Schnurr).

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 [7]. In Section 3, we establish consistency and asymptotic normality for estimators of ordinal pattern dependence in the framework 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

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

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 $\mathbf{X}^{(h)} = (X_1, \ldots, X_{1+h})$, $\mathbf{Y}^{(h)} = (Y_1, \ldots, Y_{1+h})$. 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. Moreover, here, and in the following, we consider vectors as column vectors. However, for the sake of readability and notational convenience, we omit the notation $\top$ indicating the transpose of vectors.

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 $x_{\pi_0} \geq \cdots \geq x_{\pi_h}$ (see [3,4]). 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 $j \in \{1, \ldots, h\}$ (see [17]).

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

$$\mathrm{OPD}_h(\mathbf{X}^{(h)}, \mathbf{Y}^{(h)}) = \frac{\Pr\left(\Pi\left(\mathbf{X}^{(h)}\right) = \Pi\left(\mathbf{Y}^{(h)}\right)\right) - \sum_{\pi \in S_h} \Pr\left(\Pi(\mathbf{X}^{(h)}) = \pi\right)\Pr\left(\Pi(\mathbf{Y}^{(h)}) = \pi\right)}{1 - \sum_{\pi \in S_h} \Pr\left(\Pi(\mathbf{X}^{(h)}) = \pi\right)\Pr\left(\Pi(\mathbf{Y}^{(h)}) = \pi\right)}. \tag{1}$$

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

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

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, see [2] and [9], and for a related approach to analyze dependence between dynamical systems, see [6].

2.2. Axiomatic definition of multivariate dependence measures

With the following definition, [7] establish an axiomatic theory for multivariate dependence measures between n-dimensional random vectors. This has been strongly inspired by the axiomatic framework of [14], who follow a copula-based approach to define and analyze multivariate dependence measures within one vector.

Definition 2. Let L_0 denote the space of random vectors with values in $\mathbb{R}^n$ on the common probability space $(\Omega, \mathcal{F}, \Pr)$. 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 monotonically increasing transformations of the components of the 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

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

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

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

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

Theorem 1. *The ordinal pattern dependence OPD_h is an $h + 1$-dimensional measure of dependence.*

The proof, which is a bit involved and makes use of multivariate distribution functions and survival functions, has been postponed to Section 6.

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 $(\mathbf{X}, \mathbf{Y})$, and define

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

as well as the corresponding probabilities

$$q_{\mathbf{X},\pi} := \Pr\left(\Pi(\mathbf{X}) = \pi\right), \ \ q_{\mathbf{Y},\pi} := \Pr\left(\Pi(\mathbf{Y}) = \pi\right), \ \ q_{(\mathbf{X},\mathbf{Y})} := \Pr\left(\Pi(\mathbf{X}) = \Pi(\mathbf{Y})\right).$$

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

$$\hat{q}_{\mathbf{X},\pi,n} \longrightarrow q_{\mathbf{X},\pi}, \ \ \hat{q}_{\mathbf{Y},\pi,n} \longrightarrow q_{\mathbf{Y},\pi}, \ \ \hat{q}_{(\mathbf{X},\mathbf{Y}),n} \longrightarrow q_{(\mathbf{X},\mathbf{Y})}$$

almost surely.

The following theorem establishes asymptotic normality of ordinal pattern dependence of i.i.d. random vectors. For this, we introduce the following notation:

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

Theorem 2. *Let $(\mathbf{X}_i, \mathbf{Y}_i)$, $i \geq 1$, be independent copies of $(\mathbf{X}, \mathbf{Y})$. Then, as $n \to \infty$,*

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

where the limit variance σ^2 is given by

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

Here, the matrix Σ is defined as in Proposition 2 (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

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

2a02989b9167c9d2

The proof of Theorem 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 $\hat{q}_{(\mathbf{X},\mathbf{Y}),n}$.

Proposition 2. *Under the same assumptions as in Theorem 2, we have*

$$\sqrt{n}\left(\begin{array}{c} \hat{q}_{(\mathbf{X},\mathbf{Y}),n} - q_{(\mathbf{X},\mathbf{Y})} \\ \hat{q}_{\mathbf{X},n} - q_{\mathbf{X}} \\ \hat{q}_{\mathbf{Y},n} - q_{\mathbf{Y}} \end{array}\right) \xrightarrow{\mathcal{D}} N(0,\Sigma),$$

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

$$\Sigma = \left(\begin{array}{ccc} \Sigma_{11} & \Sigma_{12} & \Sigma_{13} \\ \Sigma_{21} & \Sigma_{22} & \Sigma_{23} \\ \Sigma_{31} & \Sigma_{32} & \Sigma_{33} \end{array}\right)$$

with

$$\Sigma_{11} = q_{(\mathbf{X},\mathbf{Y})}(1 - q_{(\mathbf{X},\mathbf{Y})}) \in \mathbb{R},$$

$$\Sigma_{12} = (\sigma_\pi(1,2))_{\pi \in \mathcal{S}_h} \in \mathbb{R}^{1 \times (h+1)!}, \qquad \sigma_\pi(1,2) = \Pr(\Pi(X) = \Pi(Y) = \pi) - q_{\mathbf{X},\mathbf{Y}} \cdot q_{\mathbf{X},\pi},$$

$$\Sigma_{13} = (\sigma_\pi(1,3))_{\pi \in \mathcal{S}_h} \in \mathbb{R}^{1 \times (h+1)!}, \qquad \sigma_\pi(1,3) = \Pr(\Pi(X) = \Pi(Y) = \pi) - q_{\mathbf{X},\mathbf{Y}} \cdot q_{\mathbf{Y},\pi},$$

$$\Sigma_{22} = (\sigma_{\pi,\pi'}(2,2))_{\pi,\pi' \in \mathcal{S}_h} \in \mathbb{R}^{(h+1)! \times (h+1)!}, \qquad \sigma_{\pi,\pi'}(2,2) = \left\{\begin{array}{ll} -q_{\mathbf{X},\pi}q_{\mathbf{X},\pi'} & \text{if } \pi \neq \pi' \\ q_{\mathbf{X},\pi}(1 - q_{\mathbf{X},\pi}) & \text{if } \pi = \pi', \end{array}\right.$$

$$\Sigma_{23} = (\sigma_{\pi,\pi'}(2,3))_{\pi,\pi' \in \mathcal{S}_h} \in \mathbb{R}^{(h+1)! \times (h+1)!}, \qquad \sigma_{\pi,\pi'}(2,3) = \Pr(\Pi(\mathbf{X}) = \pi, \Pi(\mathbf{Y}) = \pi') - q_{\mathbf{X},\pi}q_{\mathbf{Y},\pi'},$$

$$\Sigma_{33} = (\sigma_{\pi,\pi'}(3,3))_{\pi,\pi' \in \mathcal{S}_h} \in \mathbb{R}^{(h+1)! \times (h+1)!}, \qquad \sigma_{\pi,\pi'}(3,3) = \left\{\begin{array}{ll} -q_{\mathbf{Y},\pi}q_{\mathbf{Y},\pi'} & \text{if } \pi \neq \pi' \\ q_{\mathbf{Y},\pi}(1 - q_{\mathbf{Y},\pi}) & \text{if } \pi = \pi'. \end{array}\right.$$

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

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

$$\xi_i = \left(1_{\{\Pi(\mathbf{X}_i) = \Pi(\mathbf{Y}_i)\}}, (1_{\{\Pi(\mathbf{X}_i) = \pi\}})_{\pi \in \mathcal{S}_h}, (1_{\{\Pi(\mathbf{Y}_i) = \pi\}})_{\pi \in \mathcal{S}_h}\right).$$

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

Proof of Theorem 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 2. In this way, we obtain

$$\sqrt{n}\left(f(\hat{q}_{(\mathbf{X},\mathbf{Y}),n}, \hat{q}_{\mathbf{X},n}, \hat{q}_{\mathbf{Y},n}) - f(q_{(\mathbf{X},\mathbf{Y})}, q_{\mathbf{X}}, q_{\mathbf{Y}})\right) \xrightarrow{\mathcal{D}} N(0, \nabla f(q_{(\mathbf{X},\mathbf{Y})}, q_{\mathbf{X}}, q_{\mathbf{Y}})\Sigma(\nabla f(q_{(\mathbf{X},\mathbf{Y})}, q_{\mathbf{X}}, q_{\mathbf{Y}})^\top).$$

This proves the statement of the theorem. $\square$

4. Ordinal pattern dependence in contrast to multivariate Kendall's τ

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 [12], and multivariate Kendall's τ as introduced in [7].

Definition 3. For two $(h+1)$-dimensional random vectors $\mathbf{X}, \mathbf{Y} \in L_2$ with invertible covariance matrices $\Sigma_{\mathbf{X}}$ and $\Sigma_{\mathbf{Y}}$ and cross-covariance matrix $\Sigma_{\mathbf{X},\mathbf{Y}}$, we define Pearson's correlation coefficient by

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

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

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

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 [7]. In that paper, the authors investigated the dependence between two multivariate random vectors. Therefore, for our purposes, it is appropriate to use it in the time series context. For a multivariate generalization of Kendall's τ *within* one random vector see [13].

Definition 4. For two $h+1$-dimensional random vectors $\mathbf{X}, \mathbf{Y}$, we define Kendall's τ by

$$\tau(\mathbf{X}, \mathbf{Y}) := \mathrm{Corr}\left(\mathbf{1}_{\left\{\mathbf{X} \leq \tilde{\mathbf{X}}\right\}}, \mathbf{1}_{\left\{\mathbf{Y} \leq \tilde{\mathbf{Y}}\right\}}\right),$$

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

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

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

$$\tau\left(\mathbf{X}^{(h)}, \mathbf{Y}^{(h)}\right) = \frac{\Pr\left(X_1 \leq 0, \ldots, X_{1+h} \leq 0, Y_1 \leq 0, \ldots, Y_{1+h} \leq 0\right) - \tilde{p}_{\mathbf{X}^{(h)}}\tilde{p}_{\mathbf{Y}^{(h)}}}{\sqrt{\tilde{p}_{\mathbf{X}^{(h)}}\left(1 - \tilde{p}_{\mathbf{X}^{(h)}}\right)\tilde{p}_{\mathbf{Y}^{(h)}}\left(1 - \tilde{p}_{\mathbf{Y}^{(h)}}\right)}},$$

where $\tilde{p}_{\mathbf{X}^{(h)}} = \Pr\left(X_1 \leq 0, \ldots, X_{1+h} \leq 0\right)$ *and* $\tilde{p}_{\mathbf{Y}^{(h)}} = \Pr\left(Y_1 \leq 0, \ldots, Y_{1+h} \leq 0\right)$.

Proof. Let $\left(\tilde{\mathbf{X}}^{(h)}, \tilde{\mathbf{Y}}^{(h)}\right)$ be an independent copy of $(\mathbf{X}^{(h)}, \mathbf{Y}^{(h)})$ with $\tilde{\mathbf{X}}^{(h)} = (\tilde{X}_1, \ldots, \tilde{X}_{1+h})$ and $\tilde{\mathbf{Y}}^{(h)} = (\tilde{Y}_1, \ldots, \tilde{Y}_{1+h})$. It then holds that

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

$$= \mathrm{Corr}\left(\mathbf{1}_{\left\{\mathbf{X}^{(h)} - \tilde{\mathbf{X}}^{(h)} \leq 0\right\}}, \mathbf{1}_{\left\{\mathbf{Y}^{(h)} - \tilde{\mathbf{Y}}^{(h)} \leq 0\right\}}\right)$$

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

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

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

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 $p_{\mathbf{X}^{(h)}} = \tilde{p}_{\mathbf{X}^{(h)}}$ and

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

$$= \frac{\Pr\left(X_1 \leq 0, \ldots, X_{1+h} \leq 0, Y_1 \leq 0, \ldots, Y_{1+h} \leq 0\right) - \tilde{p}_{\mathbf{X}^{(h)}}\tilde{p}_{\mathbf{Y}^{(h)}}}{\sqrt{\tilde{p}_{\mathbf{X}^{(h)}}\left(1 - \tilde{p}_{\mathbf{X}^{(h)}}\right)\tilde{p}_{\mathbf{Y}^{(h)}}\left(1 - \tilde{p}_{\mathbf{Y}^{(h)}}\right)}}. \quad \square$$

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

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

or

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

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

[7] 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 [5], 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

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

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

$$\psi(x, y, z) := \frac{z - xy}{\sqrt{x(1 - x)y(1 - y)}}. \tag{2}$$

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

$$\hat{p}_{\mathbf{X}^{(h)}, n} := \frac{1}{n(n-1)} \sum_{1 \leq i \neq j \leq n} \mathbf{1}_{\{\mathbf{X}_i^{(h)} \leq \mathbf{X}_j^{(h)}\}}, \ \hat{p}_{\mathbf{Y}^{(h)}, n} := \frac{1}{n(n-1)} \sum_{1 \leq i \neq j \leq n} \mathbf{1}_{\{\mathbf{Y}_i^{(h)} \leq \mathbf{Y}_j^{(h)}\}},$$

$$\hat{p}_{(\mathbf{X}^{(h)}, \mathbf{Y}^{(h)}), n} := \frac{1}{n(n-1)} \sum_{1 \leq i \neq j \leq n} \mathbf{1}_{\{\mathbf{X}_i^{(h)} \leq \mathbf{X}_j^{(h)}, \mathbf{Y}_i^{(h)} \leq \mathbf{Y}_j^{(h)}\}},$$

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

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

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 $\hat{\tau}_n(\mathbf{X}^{(h)}, \mathbf{Y}^{(h)})$ 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

$$f((x, y), (x', y')) = \frac{1}{2}\left(\mathbf{1}_{\{x \leq x'\}} + \mathbf{1}_{\{x \geq x'\}}\right), \ g((x, y), (x', y')) = \frac{1}{2}\left(\mathbf{1}_{\{y \leq y'\}} + \mathbf{1}_{\{y \geq y'\}}\right),$$

$$h((x, y), (x', y')) = \frac{1}{2}\left(\mathbf{1}_{\{x \leq x', y \leq y'\}} + \mathbf{1}_{\{x \geq x', y \geq y'\}}\right).$$

Note that the underlying random vectors $(\mathbf{X}_i^{(h)}, \mathbf{Y}_i^{(h)})$, $i \geq 1$, 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 [1].

Theorem 3. *Assume that (X_i, Y_i), $i \geq 1$, is a stationary ergodic process, and that $(\mathbf{X}^{(h)}, \mathbf{Y}^{(h)})$ has a continuous distribution. Then, as $n \to \infty$, we obtain almost surely*

$$\hat{p}_{\mathbf{X}^{(h)}, n} \longrightarrow p_{\mathbf{X}^{(h)}}, \ \hat{p}_{\mathbf{Y}^{(h)}, n} \longrightarrow p_{\mathbf{Y}^{(h)}}, \ \hat{p}_{(\mathbf{X}^{(h)}, \mathbf{Y}^{(h)}), n} \longrightarrow p_{(\mathbf{X}^{(h)}, \mathbf{Y}^{(h)})}.$$

Proof. We apply Theorem U from [1]. 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 [5]. This concept is a variation of the usual L_2-near-epoch dependence and does not require any moment assumptions.

Definition 5. (i) Given two sub-σ-fields $\mathcal{A}, \mathcal{B} \subset \mathcal{F}$, we define the absolute regularity coefficient

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

where the supremum is taken over all integers $m, n \geq 1$, 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, $i \geq 1$, 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), $i \geq 1$, is a stationary process, and if there exists a sequence $(a_k)_{k \geq 0}$ 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

$$\Pr(|X_0 - f_k(Z_{-k}, \ldots, Z_k)| \geq \epsilon) \leq a_k \Phi(\epsilon).$$

Proposition 3. *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:*

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

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

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

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

Remark 1. The first order term of the Hoeffding decomposition is given by

$$f_1(x, y) = Ef((x, y), (\tilde{\mathbf{X}}^{(h)}, \tilde{\mathbf{Y}}^{(h)})) - p_{\mathbf{X}^{(h)}} = \frac{1}{2}\left(\Pr(\mathbf{X}^{(h)} \leq x) + \Pr(\mathbf{X}^{(h)} \geq x)\right) - \Pr(\mathbf{X}^{(h)} \leq \tilde{\mathbf{X}}^{(h)})$$

$$= \frac{1}{2}\left(F(x) + \bar{F}(x)\right) - \Pr(\mathbf{X}^{(h)} \leq \tilde{\mathbf{X}}^{(h)}),$$

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

$$g_1(x, y) = \frac{1}{2}\left(G(x) + \bar{G}(x)\right) - \Pr(\mathbf{Y}^{(h)} \leq \tilde{\mathbf{Y}}^{(h)}),$$

$$h_1(x, y) = \frac{1}{2}\left(H(x, y) + \bar{H}(x, y)\right) - \Pr(\mathbf{X}^{(h)} \leq \tilde{\mathbf{X}}^{(h)}, \mathbf{Y}^{(h)} \leq \tilde{\mathbf{Y}}^{(h)}),$$

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

Proof of Proposition 3. This follows from Lemma D.6 of [5] noting that the variation condition is satisfied because the distribution of $\mathbf{Y}^{(h)} - \mathbf{X}^{(h)}$ has a bounded density. $\quad \square$

Theorem 4. *Under the same assumptions as in Proposition 3, 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 $\Sigma \in \mathbb{R}^{3 \times 3}$ *is the limit covariance matrix whose diagonal and off-diagonal entries are given, e.g., by*

$$\sigma_{11} = \text{Var}\left(f_1(\mathbf{X}_1^{(h)}, \mathbf{Y}_1^{(h)})\right) + 2\sum_{i=2}^\infty \text{Cov}\left(f_1(\mathbf{X}_1^{(h)}, \mathbf{Y}_1^{(h)}), f_1(\mathbf{X}_i^{(h)}, \mathbf{Y}_i^{(h)})\right),$$

$$\sigma_{12} = \text{Cov}\left(f_1(\mathbf{X}_1^{(h)}, \mathbf{Y}_1^{(h)}), g_1(\mathbf{X}_1^{(h)}, \mathbf{Y}_1^{(h)})\right) + \sum_{i=2}^{\infty} \text{Cov}\left(f_1(\mathbf{X}_1^{(h)}, \mathbf{Y}_1^{(h)}), g_1(\mathbf{X}_i^{(h)}, \mathbf{Y}_i^{(h)})\right)$$

$$+ \sum_{i=2}^{\infty} \text{Cov}\left(f_1(\mathbf{X}_i^{(h)}, \mathbf{Y}_i^{(h)}), g_1(\mathbf{X}_1^{(h)}, \mathbf{Y}_1^{(h)})\right).$$

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.,[18]. Now, the statement of the theorem follows from Proposition 3 together with an application of Slutsky's lemma. $\square$

Theorem 5. *Under the assumptions of Proposition 3, the estimator $\hat{\tau}_n(\mathbf{X}^{(h)}, \mathbf{Y}^{(h)})$ of $\tau(\mathbf{X}^{(h)}, \mathbf{Y}^{(h)})$ is consistent and asymptotically normal. More precisely, we obtain*

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

where ψ is defined by (2) and Σ is defined as in Theorem 4.

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

5. Ordinal pattern dependence in contrast to other dependence measures

For independent vectors $(\mathbf{X}_i, \mathbf{Y}_i)$, $1 \le i \le n$, 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 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 4. *Let $\mathbf{X} = (X_1, X_2)$ and $\mathbf{Y} = (Y_1, Y_2)$ be two Gaussian random vectors satisfying $\text{E}(X_1) = \text{E}(X_2)$, $\text{E}(Y_1) = \text{E}(Y_2)$ and $\text{Var}(X_2 - X_1) \ne 0 \ne \text{Var}(Y_2 - Y_1)$. Then, it holds that*

$$\text{OPD}_1(\mathbf{X}, \mathbf{Y}) = \tau(X_2 - X_1, Y_2 - Y_1) = \frac{2}{\pi} \arcsin \text{Corr}(X_2 - X_1, Y_2 - Y_1). \tag{3}$$

Proof. By definition

$$\text{OPD}_1(\mathbf{X}, \mathbf{Y}) = \frac{\Pr(\Pi(X_1, X_2) = \Pi(Y_1, Y_2)) - \sum_{\pi \in S_1} \Pr(\Pi(X_1, X_2) = \pi)\Pr(\Pi(Y_1, Y_2) = \pi)}{1 - \sum_{\pi \in S_1} \Pr(\Pi(X_1, X_2) = \pi)\Pr(\Pi(Y_1, Y_2) = \pi)}.$$

Since $X_2 - X_1$ and $Y_2 - Y_1$ are both Gaussian random variables with mean zero and non-zero variance, we obtain

$$\Pr(X_1 < X_2) = \Pr(X_2 < X_1) = \Pr(Y_1 < Y_2) = \Pr(Y_2 < Y_1) = \frac{1}{2},$$

and thus $\Pr(\Pi(X_1, X_2) = \pi) = \Pr(\Pi(Y_1, Y_2) = \pi) = \frac{1}{2}$ for any $\pi \in S_1$. Hence, we obtain

$$\text{OPD}_1(\mathbf{X}, \mathbf{Y}) = 2\,\Pr(\Pi(X_1, X_2) = \Pi(Y_1, Y_2)) - 1.$$

Moreover, it holds that

$$\Pr(X_1 < X_2, Y_1 < Y_2) = \Pr(X_2 - X_1 > 0, Y_2 - Y_1 > 0) = \Pr(X_1 - X_2 > 0, Y_1 - Y_2 > 0) = \Pr(X_1 > X_2, Y_1 > Y_2),$$

and hence

$$\text{OPD}_1(\mathbf{X}, \mathbf{Y}) = 4\,\Pr(X_1 - X_2 > 0, Y_1 - Y_2 > 0) - 1.$$

From Lemma 1 with $h = 1$, we find $\tau(X_2 - X_1, Y_2 - Y_1) = 4 \Pr(X_2 - X_1 < 0, Y_2 - Y_1 < 0) - 1$, and thus $\mathrm{OPD}_1(\mathbf{X}, \mathbf{Y}) = \tau(X_2 - X_1, Y_2 - Y_1)$.

Finally, using the *orthant probabilities* formula for Gaussian random variables we obtain

$$\Pr(X_2 - X_1 < 0, Y_2 - Y_1 < 0) = \frac{1}{4} + \frac{1}{2\pi} \arcsin \mathrm{Corr}(X_2 - X_1, Y_2 - Y_1),$$

and thus $\mathrm{OPD}_1(\mathbf{X}, \mathbf{Y}) = \frac{2}{\pi} \arcsin \mathrm{Corr}(X_2 - X_1, Y_2 - Y_1)$. □

In the following we provide an example of marginally uncorrelated Gaussian random vectors with non-vanishing ordinal pattern dependence of order 1.

Definition 6. A stationary bivariate Gaussian process $W_i = (X_i, Y_i)$ is called AR(1) process, if there exist a matrix $A \in \mathbb{R}^{2 \times 2}$, and an i.i.d. $N(0, I_2)$-distributed Gaussian process $\xi_i = (\epsilon_i, \eta_i)$ such that the AR(1)-equation

$$W_i = AW_{i-1} + \xi_i \tag{4}$$

is satisfied.

Remark 2. Given a matrix $A \in \mathbb{R}^{2 \times 2}$, a stationary Gaussian AR(1)-process exists, if and only if all eigenvalues of A are strictly less than 1 in absolute value.

We will now consider the special example when the AR(1)-matrix is given by

$$A = \begin{pmatrix} a & b \\ b & -a \end{pmatrix}, \tag{5}$$

where $a^2 + b^2 < 1$. Thus, the AR(1)-equation for $W_i = (X_i, Y_i)$ takes the form

$$X_i = aX_{i-1} + bY_{i-1} + \epsilon_i, \quad Y_i = bX_{i-1} - aY_{i-1} + \eta_i.$$

In the next lemma, we state some properties of the processes $(X_i)_{i \geq 1}$ and $(Y_i)_{i \geq 1}$, and we give an explicit formula for their ordinal pattern dependence of order 1.

Lemma 2. *Consider the stationary bivariate Gaussian AR(1)-process $W_i = (X_i, Y_i)$, $i \geq 1$, satisfying (4) with matrix A given by (5) such that $a^2 + b^2 < 1$. Then, it holds that*

$$\mathrm{Cov}(X_1, Y_1) = 0, \ \mathrm{Var}(X_1) = \mathrm{Var}(Y_1) = \frac{1}{1 - a^2 - b^2}$$

$$\mathrm{OPD}_1(\mathbf{X}^{(1)}, \mathbf{Y}^{(1)}) = \frac{2}{\pi} \arcsin\left(-\frac{b}{\sqrt{1 - a^2}}\right) \tag{6}$$

Proof. The eigenvalues of A are $\lambda_{1,2} = \pm\sqrt{a^2 + b^2}$, and thus (4) has a unique stationary solution. Since the AR(1)-equation defines a Markov chain with state space $\mathbb{R}^2$, the joint distribution of $W_i = (X_i, Y_i)$ is uniquely characterized by the distributional fixed point equation

$$W \overset{\mathcal{D}}{=} AW + \xi,$$

where $\xi = (\epsilon, \eta)$ has a bivariate normal distribution with mean zero and covariance matrix I_2, and where ξ is independent of W. We will now show that $W \sim N(0, \sigma^2 I_2)$ satisfies this equation with

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

In order to prove this, we need to calculate the distribution of $AW + \xi$. Since the distribution is Gaussian, it suffices to calculate the variances and the covariance. We obtain

$$\mathrm{Cov}(aX + bY + \epsilon, bX - aY + \eta) = ab\sigma^2 - ab\sigma^2 = 0,$$

$$\mathrm{Var}(aX + bY + \epsilon) = a^2\sigma^2 + b^2\sigma^2 + 1 = \frac{a^2 + b^2}{1 - a^2 - b^2} + 1 = \frac{1}{1 - a^2 - b^2} = \sigma^2,$$

$$\mathrm{Var}(bX - aY + \epsilon) = b^2\sigma^2 + a^2\sigma^2 + 1 = \frac{a^2 + b^2}{1 - a^2 - b^2} + 1 = \frac{1}{1 - a^2 - b^2} = \sigma^2,$$

which shows that $AW + \xi$ has indeed the same distribution as W.

In order to determine the OPD$_1$ of the two processes, we need to calculate the correlation of the differences. The covariance of the increments is given by

$$\text{Cov}(X_2 - X_1, Y_2 - Y_1) = \text{Cov}((a-1)X_1 + bY_1 + \epsilon_2, bX_1 + (-a-1)Y_1 + \eta_2) = b(a-1)\sigma^2 - b(a+1)\sigma^2$$
$$= -2b\sigma^2 = \frac{-2b}{1-a^2-b^2}$$

and the variances of the increments are given by

$$\text{Var}(X_2 - X_1) = (a-1)^2\sigma^2 + b^2\sigma^2 + 1 = \frac{(a-1)^2 + b^2}{1-a^2-b^2} + 1 = \frac{2(1-a)}{1-a^2-b^2},$$
$$\text{Var}(Y_2 - Y_1) = b^2\sigma^2 + (a+1)^2\sigma^2 + 1 = \frac{b^2 + (a+1)^2}{1-a^2-b^2} + 1 = \frac{2(a+1)}{1-a^2-b^2}.$$

Thus, we obtain the following formula for the correlation of the increments:

$$\text{Corr}(X_2 - X_1, Y_2 - Y_1) = \frac{-2b}{\sqrt{4(1-a)(a+1)}} = -\frac{b}{\sqrt{1-a^2}}$$

Using the identity OPD$_1((X_1, X_2), (Y_1, Y_2)) = \frac{2}{\pi}\arcsin\text{Corr}(X_2 - X_1, Y_2 - Y_1)$, we finally obtain (6). □

Remark 3. (i) The special choice of the AR(1)-matrix A made in (5) assures that the two processes $(X_i)_{i\geq 1}$ and $(Y_i)_{i\geq 1}$ have identical marginals, and that X_i and Y_i are independent for each fixed i. In fact, one can show that the latter two properties only hold if A is either of the form (5) or of the form

$$A = \begin{pmatrix} a & b \\ -b & a \end{pmatrix}. \tag{7}$$

In this case, using similar calculations as above, one obtains OPD$_1((X_1, X_2), (Y_1, Y_2)) = 0$.

(ii) Lemma 2 provides an example of a Gaussian process for which Pearson's correlation of X_i and Y_i equals 0, i.e., the one-dimensional marginals are independent. However, the processes $(X_i)_{i\geq 1}$ and $(Y_i)_{i\geq 1}$ are not independent, as can be seen from the identity for OPD($\mathbf{X}^{(1)}, \mathbf{Y}^{(1)}$).

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

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

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

$$A := \begin{pmatrix} a & b \\ b & -a \end{pmatrix}, \quad \xi_i := \begin{pmatrix} \epsilon_i \\ \eta_i \end{pmatrix}, \tag{8}$$

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 Cov(X_i, Y_i) = 0, but high ordinal pattern dependence. For the simulations summarized by the boxplots in Fig. 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.

Fig. 2 depicts one sample path of the single time series X_i, $i \in \{1, \ldots, 500\}$, and Y_i, $i \in \{1, \ldots, 500\}$, 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 Fig. 2 underline uncorrelatedness of the original processes and a high dependence of their increments.

5.2. The case $h = 2$

Let us recall that for the computation of the ordinal pattern dependence of order $h = 1$, the crucial quantity is Corr($X_2 - X_1, Y_2 - Y_1$) since, according to Proposition 4, 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 OPD$_1(\mathbf{X}^{(1)}, \mathbf{Y}^{(1)}) = 0$, but OPD$_2(\mathbf{X}^{(2)}, \mathbf{Y}^{(2)}) \neq 0$.

The AR(1)-process in Lemma 2 does not fulfill these conditions, since the restriction

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

implies $b = 0$. As a result, we obtain a process $W_i = (X_i, Y_i) = (aX_{i-1} + \xi_i, -aY_{i-1} + \eta_i)$, 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.
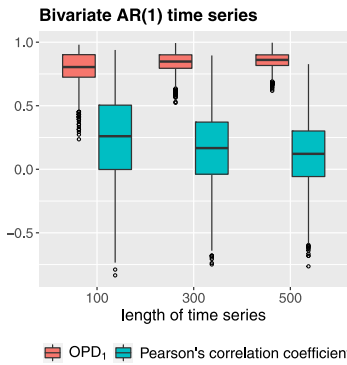
Fig. 1. Boxplots of Pearson's correlation coefficient and ordinal pattern dependence of order $h = 1$ based on 5000 repetitions of a bivariate AR(1)-process (X_i, Y_i), $i \in \{1, \ldots, n\}$, satisfying (8) with $a = 0.7$ and $b = -0.7$.
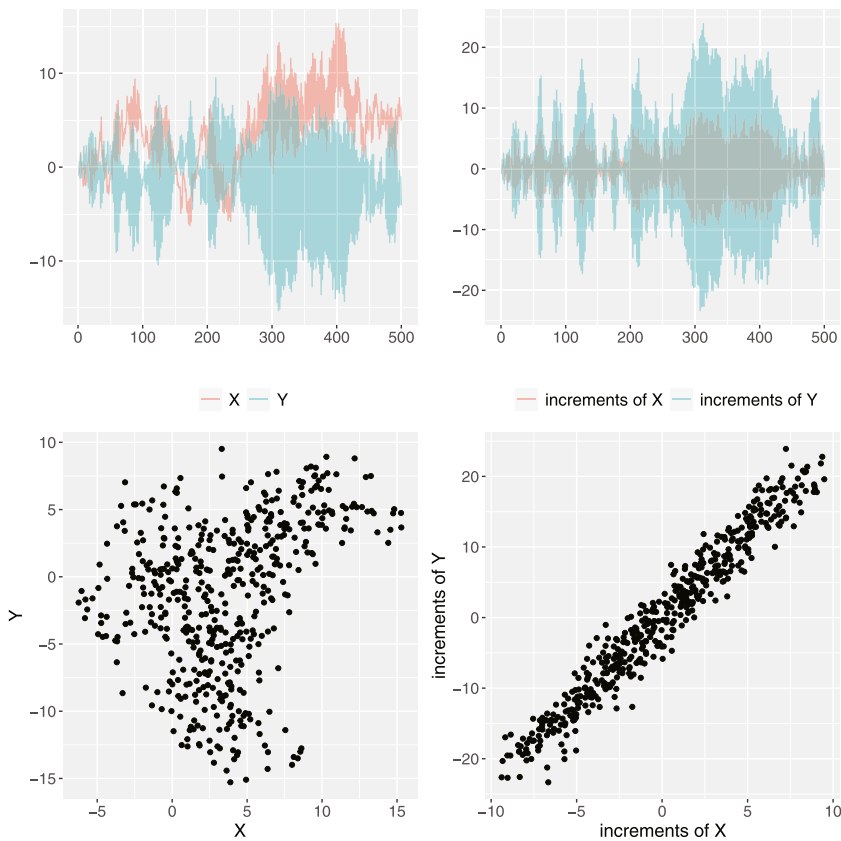

Fig. 2. Sample paths of X_i, $i \in \{1, \ldots, 500\}$, and Y_i, $i \in \{1, \ldots, 500\}$, their increments at lag 1, as well as corresponding scatterplots based on a bivariate AR(1)-process (X_i, Y_i), $i \in \{1, \ldots, n\}$, satisfying (8) with $a = 0.7$ and $b = -0.7$.

Following Remark 3, the choice of the matrix A in (7) yields Corr $(X_i, Y_i) = 0$ for $i \in \{1, 2\}$, and OPD$_1(\mathbf{X}^{(1)}, \mathbf{Y}^{(1)}) = 0$. This leads to the question whether this special construction of an AR(1)-process fulfills OPD$_2(\mathbf{X}^{(2)}, \mathbf{Y}^{(2)}) \neq 0$.

Lemma 3. *Consider the stationary bivariate Gaussian* AR(1)-*process* $W_i = (X_i, Y_i)$, $i \geq 1$, *satisfying* (4) *with matrix A given by* (7), *and where $a^2 + b^2 < 1$. Then, it holds that*

$$\text{Cov}(X_i, Y_i) = 0, \ \text{Var}(X_i) = \text{Var}(Y_i) = \sigma^2 = \frac{1}{1 - a^2 - b^2},$$

$$\text{OPD}_1((X_i, X_{i+1}), (Y_i, Y_{i+1})) = 0, \ \text{Corr}(X_2 - X_1, Y_3 - Y_2) = -b, \ \text{Corr}(X_3 - X_2, Y_2 - Y_1) = b.$$

Proof. The first three identities can be shown as in Lemma 2. Thus, it remains to show the latter two. It holds that

$$\mathrm{Var}(X_2 - X_1) = (a-1)^2\sigma^2 + b^2\sigma^2 + 1 = \frac{(a-1)^2 + b^2 + 1 - a^2 - b^2}{1 - a^2 - b^2} = 2(1-a)\sigma^2.$$

Analogously, we obtain

$$\mathrm{Var}(Y_3 - Y_2) = 2(1-a)\sigma^2.$$

Furthermore, it holds that

$$\mathrm{Cov}(Y_3 - Y_2, X_2 - X_1) = \mathbb{E}\,(Y_3 X_2) - \mathbb{E}\,(Y_2 X_2) - \mathbb{E}\,(Y_3 X_1) + \mathbb{E}\,(Y_2 X_1) = 2b(a-1)\sigma^2,$$

since

$$\mathbb{E}(Y_3 X_2) = -b\sigma^2, \ \mathbb{E}(Y_3 X_1) = -2ab\sigma^2.$$

Altogether, we arrive at

$$\mathrm{Corr}(Y_3 - Y_2, X_2 - X_1) = \frac{2ab - 2b}{2(1-a)} = -b.$$

$\mathrm{Corr}(Y_2 - Y_1, X_3 - X_2) = b$ is derived by similar calculations. $\square$

Lemma 3 provides an example of a bivariate process $(X_i, Y_i)_{i\geq 1}$ for which $\mathrm{Corr}(X_i, Y_i) = 0$ and $\mathrm{OPD}_1(\mathbf{X}^{(1)}, \mathbf{Y}^{(1)}) = 0$, but where the processes $(X_i)_{i\geq 1}$ and $(Y_i)_{i\geq 1}$ are nevertheless dependent. The fact that the increments $X_2 - X_1$ and $Y_3 - Y_2$ are dependent, leads us to conjecture that $\mathrm{OPD}_2(\mathbf{X}^{(2)}, \mathbf{Y}^{(2)}) \neq 0$, but we do not have a proof. In order to compute $\mathrm{OPD}_2(\mathbf{X}^{(2)}, \mathbf{Y}^{(2)})$, we would have to calculate

$$\mathrm{Pr}\,(\Pi(X_1, X_2, X_3) = \pi, \Pi(Y_1, Y_2, Y_3) = \pi)$$

for any $\pi \in \mathcal{S}_2$. This requires computations of orthant probabilities for 4-dimensional Gaussian vectors, e.g. for $\pi = (0, 1, 2)$ we obtain

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

To the best of our knowledge, there are no explicit formulas for these probabilities known.

In what follows, we present an example of a bivariate AR(2)-process $(X_i, Y_i)_{i\geq 1}$ for which $\mathrm{Corr}(X_i, Y_i) = 0$ and $\mathrm{OPD}_1(\mathbf{X}^{(1)}, \mathbf{Y}^{(1)}) = 0$, but where the processes $(X_i)_{i\geq 1}$ and $(Y_i)_{i\geq 1}$ are dependent. For this example, we show by means of a Monte Carlo simulation, that $\mathrm{OPD}_2(\mathbf{X}^{(2)}, \mathbf{Y}^{(2)}) \neq 0$.

Example 1. Let W_i, be a bivariate AR(2)-process defined by

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

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

$$A := \begin{pmatrix} a & b \\ b & -a \end{pmatrix} \text{ and } \xi_i := \begin{pmatrix} \varepsilon_i \\ \eta_i \end{pmatrix}, \tag{9}$$

ξ_i, bivariate Gaussian random vectors with covariance matrix $\Sigma_\xi = I_2$ (with I_2 denoting the identity matrix). 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 $\mathrm{OPD}_1(\mathbf{X}^{(1)}, \mathbf{Y}^{(1)}) = 0$ since

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

This construction of AR(2)-processes can be extended to AR(h) for $h \in \mathbb{N}$, if one wants to obtain $\mathrm{OPD}_h\,(\mathbf{X}^{(h)}, \mathbf{Y}^{(h)}) \neq 0$ but $\mathrm{OPD}_i\,(\mathbf{X}^{(i)}, \mathbf{Y}^{(i)}) = 0, i = 1, \ldots, h-1$ and $\mathrm{Corr}\,(X_1, Y_1) = 0$ by using h independent AR(1)-processes and couple them via

$$W_j = AW_{j-h} + \xi_j.$$

To illustrate the strong connection between a large correlation of the increments at lag 1 and $\mathrm{OPD}_2(\mathbf{X}^{(2)}, \mathbf{Y}^{(2)})$ we consider estimated values for $\mathrm{OPD}_2(\mathbf{X}^{(2)}, \mathbf{Y}^{(2)})$ based on simulations of a bivariate AR(2)-process that satisfies the above assumptions; see Fig. 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 value between 0.1 and 0.25.
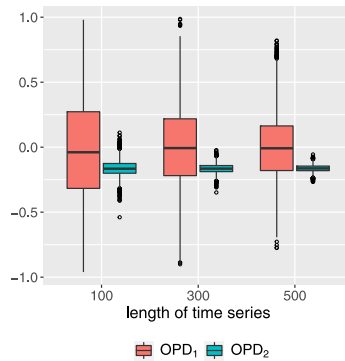
Fig. 3. Boxplots of ordinal pattern dependence of order $h = 1$ and $h = 2$ based on 5000 repetitions of a bivariate AR(2)-process (X_i, Y_i), $i \in \{1, \ldots, n\}$, satisfying (9) with $a = 0.01$ and $b = 0.98$.

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 $\mathbf{Y}^{(1)} = (Y_1, Y_2)$ satisfying $E(X_1) = E(X_2)$, $E(Y_1) = E(Y_2)$, and $\mathrm{Var}(X_2 - X_1) \neq 0 \neq \mathrm{Var}(Y_2 - Y_1)$, it holds that

$$\mathrm{OPD}_1(\mathbf{X}^{(1)}, \mathbf{Y}^{(1)}) = \tau\left(\tilde{X}_1, \tilde{Y}_1\right),$$

where $\tilde{X}_1 := X_2 - X_1$ and $\tilde{Y}_1 := Y_2 - Y_1$; see Proposition 4.

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

Proposition 5. *Let (X_i, Y_i), $i \geq 1$, be a bivariate, mean zero stationary Gaussian process. Then, it holds that*

$$\mathrm{OPD}_h\left(\mathbf{X}^{(h)}, \mathbf{Y}^{(h)}\right) = \frac{2\tau\left(\tilde{X}_1, \ldots, \tilde{X}_h, \tilde{Y}_1, \ldots, \tilde{Y}_h\right)\sqrt{\tilde{p}_{\tilde{\mathbf{X}}^{(h)}}\left(1 - \tilde{p}_{\tilde{\mathbf{X}}^{(h)}}\right)\tilde{p}_{\tilde{\mathbf{Y}}^{(h)}}\left(1 - \tilde{p}_{\tilde{\mathbf{Y}}^{(h)}}\right)}}{1 - \sum_{\pi \in S_h}\Pr\left(\Pi\left(\mathbf{X}^{(h)}\right) = \pi\right)\Pr\left(\Pi\left(\mathbf{Y}^{(h)}\right) = \pi\right)}$$
$$+ \sum_{\pi \in S_h \setminus T_h}\frac{\Pr\left(\Pi\left(\mathbf{X}^{(h)}\right) = \Pi\left(\mathbf{Y}^{(h)}\right) = \pi\right) - \Pr\left(\Pi\left(\mathbf{X}^{(h)}\right) = \pi\right)\Pr\left(\Pi\left(\mathbf{Y}^{(h)}\right) = \pi\right)}{1 - \sum_{\pi \in S_h}\Pr\left(\Pi\left(\mathbf{X}^{(h)}\right) = \pi\right)\Pr\left(\Pi\left(\mathbf{Y}^{(h)}\right) = \pi\right)},$$

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

$$\tilde{p}_{\tilde{\mathbf{X}}^{(h)}} = \Pr\left(\Pi\left(X_1, \ldots, X_{1+h}\right) = (0, \ldots, h)\right) = \Pr\left(\tilde{X}_1 \leq 0, \ldots, \tilde{X}_h \leq 0\right),$$

$$\tilde{p}_{\tilde{\mathbf{Y}}^{(h)}} = \Pr\left(\Pi\left(Y_1, \ldots, Y_{1+h}\right) = (0, \ldots, h)\right) = \Pr\left(\tilde{Y}_1 \leq 0, \ldots, \tilde{Y}_h \leq 0\right),$$

and

$$\mathrm{OPD}_h\left(\mathbf{X}^{(h)}, \mathbf{Y}^{(h)}\right) = \frac{\sum_{\pi \in S_h}\tau\left(X_{1+\pi_1} - X_{1+\pi_0}, \ldots, Y_{1+\pi_h} - Y_{1+\pi_{h-1}}\right)\sqrt{p_{\mathbf{X}^{(h)},\pi}\left(1 - p_{\mathbf{X}^{(h)},\pi}\right)p_{\mathbf{Y}^{(h)},\pi}\left(1 - p_{\mathbf{Y}^{(h)},\pi}\right)}}{1 - \sum_{\pi \in S_h}p_{\mathbf{X}^{(h)},\pi}p_{\mathbf{Y}^{(h)},\pi}}$$

with

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

Remark 4. It is a characteristic of Gaussian observations (X_i), $i \geq 1$, and (Y_i), $i \geq 1$, that the distribution of

$$\left(X_{1+\pi_1} - X_{1+\pi_0}, \ldots, X_{1+\pi_h} - X_{1+\pi_{h-1}}, Y_{1+\pi_1} - Y_{1+\pi_0}, \ldots, Y_{1+\pi_h} - Y_{1+\pi_{h-1}}\right)$$

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.

The proof of Proposition 5 has been postponed to Section 6. In order to illustrate the relation of multivariate Kendall's τ and ordinal pattern dependence, characterized through Proposition 5, we consider the case $h = 1$ in the following example:

Example 2. Let $\mathbf{X} = (X_1, X_2)$ and $\mathbf{Y} = (Y_1, Y_2)$ be Gaussian random vectors as in Proposition 4. Recall that

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

Moreover, if $\mathbf{X}$ and $\mathbf{Y}$ have standard normal marginal distributions, it holds that

$$\mathrm{Corr}(X_2 - X_1, Y_2 - Y_1) = \frac{2\mathbb{E}\,(X_1 Y_1) - \mathbb{E}\,(X_1 Y_2) - \mathbb{E}\,(Y_1 X_2)}{2 - 2\mathbb{E}\,(X_1 X_2)}.$$

In general, we know that $\tau(X, Y) = \frac{2}{\pi}\arcsin\left(\mathrm{Corr}(X, Y)\right)$ for a Gaussian random vector (X, Y), and hence $\mathrm{Corr}(X, Y) = \sin\left(\frac{\pi}{2}\tau(X, Y)\right)$.

As a result, we obtain

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

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

5.4. Simulation study

In this section, we compare the estimators for ordinal pattern dependence and multivariate Kendall's τ based on the vectors $\mathbf{X}_i^{(h)}, \mathbf{Y}_i^{(h)}$, generated by bivariate processes (X_i, Y_i), $i \geq 1$, in a simulation study. Fig. 4 corresponds to the case $h = 2$. 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 Fig. 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. For a systematic simulation study see Table A.1 in the appendix.

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

   $$\Sigma = \begin{pmatrix} 1 & 0 & 0 & \rho & \rho & \rho \\ 0 & 1 & 0 & \rho & \rho & \rho \\ 0 & 0 & 1 & \rho & \rho & \rho \\ \rho & \rho & \rho & 1 & 0 & 0 \\ \rho & \rho & \rho & 0 & 1 & 0 \\ \rho & \rho & \rho & 0 & 0 & 1 \end{pmatrix}. \tag{10}$$

   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 Fig. 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. For a systematic simulation study see Table A.2 in the appendix.

3. 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 one 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 Fig. 4, we chose $\rho = 0.5$.

   It is intriguing that ordinal pattern dependence does not detect the high correlation of the time series. The theoretical value of ordinal pattern dependence for $h = 1$ in this case is given by

   $$\mathrm{OPD}_1(\mathbf{X}^{(1)}, \mathbf{Y}^{(1)}) = \frac{2}{\pi}\arcsin\mathrm{Corr}(X_2 - X_1, Y_2 - Y_1) = \frac{2}{\pi}\arcsin\mathrm{Corr}(X_2 - X_1, X_3 - X_2).$$

   Routine calculations yield the following formula for the ordinal pattern dependence of order 1 between an AR(1) process and the same process shifted by one time point:

   $$\mathrm{OPD}_1(\mathbf{X}^{(1)}, \mathbf{Y}^{(1)}) = \frac{2}{\pi}\arcsin\left(\frac{\rho - 1}{2}\right).$$
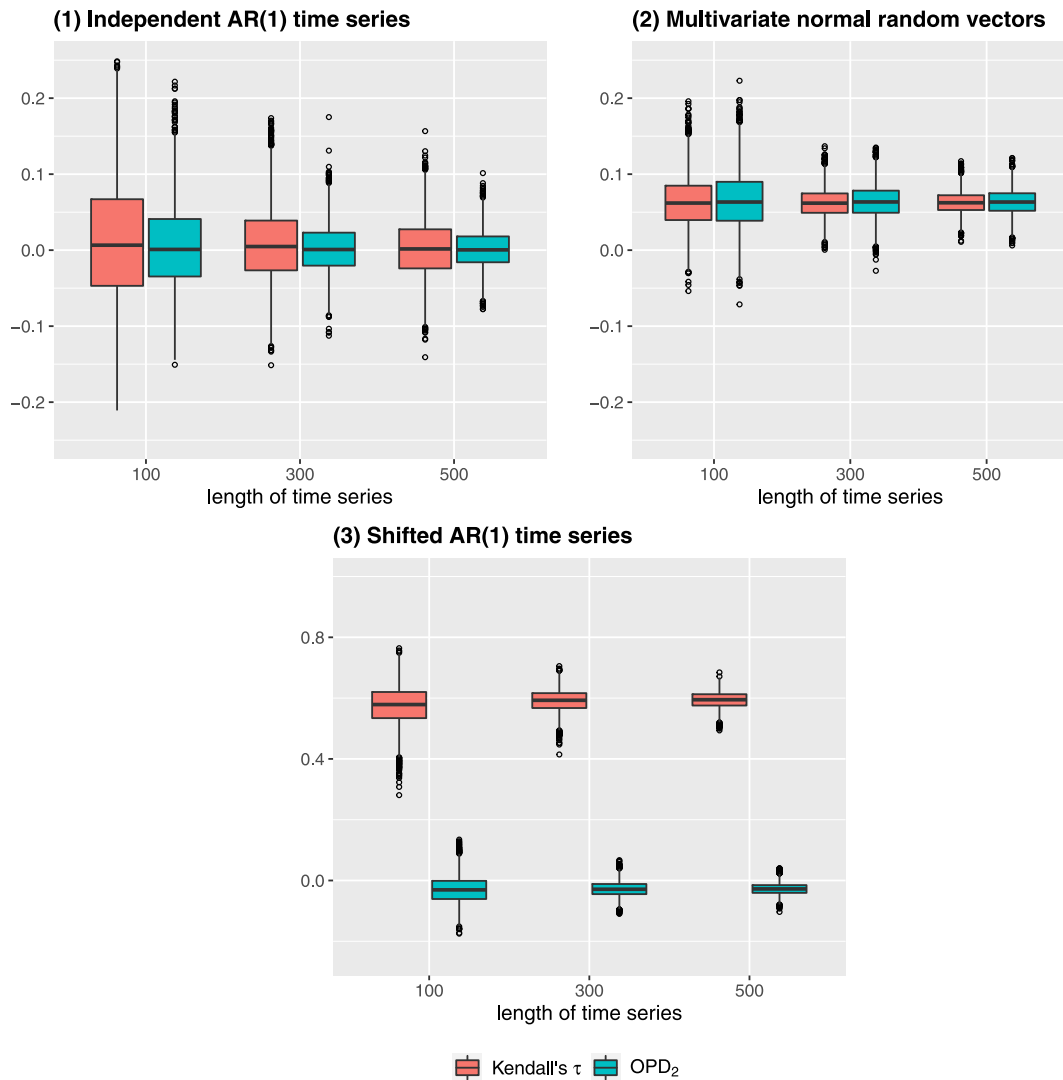
Fig. 4. Boxplots of ordinal pattern dependence of order $h = 2$ based on 5000 repetitions of a bivariate process (X_i, Y_i), $i \in \{1, \ldots, n\}$, corresponding to the situations (1)-(3).

As a result, $\mathrm{OPD}_1(\mathbf{X}^{(1)}, \mathbf{Y}^{(1)}) = -0.297, -0.161, -0.032$ for $\rho = 0.1, 0.5, 0.9$. These values coincide with the results of the corresponding systematic simulation study; see Table A.3 in the appendix.

In [16] on page 713, an approach is presented to solve the insensitivity of ordinal pattern dependence concerning time shifts. The authors introduced time shifted ordinal pattern dependence in order to investigate and compare time series that are known to have a similar behavior within a certain time deviation. These time series arise for example in the context of hydrology, if discharge data of a river is considered for two different locations. For a real-world data analysis see [10]. Using this approach, (time-shifted) ordinal pattern dependence of 1 is detected, since all patterns coincide if we reshift the second time series.

6. Proofs

Proof of Theorem 1. The first four axioms in Definition 2 are easily verified, the fifth one is involved. We show that the fifth axiom is fulfilled for 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 $\Pr(\Pi(\mathbf{X}) = \Pi(\mathbf{Y}))$ (the remaining summands of $\mathrm{OPD}_h(\mathbf{X}, \mathbf{Y})$ 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

$$\mathbf{X} \overset{\mathcal{D}}{=} \mathbf{X}^*, \ \mathbf{Y} \overset{\mathcal{D}}{=} \mathbf{Y}^*, \ \text{and} \ \begin{pmatrix} \mathbf{X} \\ \mathbf{Y} \end{pmatrix} \preccurlyeq_C \begin{pmatrix} \mathbf{X}^* \\ \mathbf{Y}^* \end{pmatrix}. \tag{11}$$

It is a well-known fact that (11) 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 $I \subseteq \{1, \ldots, 6\}$, i.e., removing dimensions does not influence which scenario has the larger dependence measure; see [8]. 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

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

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

$$\Pr(\Pi(X_1, X_2, X_3) = \Pi(Y_1, Y_2, Y_3)) = \Pr(X_1 \leq X_2 \leq X_3, Y_1 \leq Y_2 \leq Y_3)$$

$$= \Pr(\{X_1 \leq X_2\} \cap \{X_2 \leq X_3\} \cap \{Y_1 \leq Y_2\} \cap \{Y_2 \leq Y_3\}) = \int_{\mathbb{R}^2} \Pr^{x_1, y_1}(x_1 \leq X_2 \leq X_3, y_1 \leq Y_2 \leq Y_3) \, dP_{\binom{X_1}{Y_1}}(x_1, y_1)$$

$$= \int_{\mathbb{R}^2} \Pr^{x_1, y_1}(X_2 \leq X_3, Y_2 \leq Y_3 | x_1 \leq X_2, y_1 \leq Y_2) \cdot \Pr^{x_1, y_1}(X_2 \geq x_1, Y_2 \geq y_1) \, dP_{\binom{X_1}{Y_1}}(x_1, y_1)$$

$$= \int_{\mathbb{R}^2} \int_{[x_1, \infty[\times [y_1, \infty[} \Pr^{x_1, y_1}(X_2 \leq X_3, Y_2 \leq Y_3 | X_2 = x_2, Y_2 = y_2) \, dP_{\binom{X_2}{Y_2}}^{x_1, y_1}(x_2, y_2)$$

$$\times \Pr^{x_1, y_1}(X_2 \geq x_1, Y_2 \geq y_1) \, dP_{\binom{X_1}{Y_1}}(x_1, y_1).$$

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

$$\Pr(\Pi(X_1, X_2, X_3) = \Pi(Y_1, Y_2, Y_3)) = \int_{\mathbb{R}^2} \int_{[x_1, \infty[\times [y_1, \infty[} \bar{F}_{\binom{X_3}{Y_3}}(x_2, y_2) \, dP_{\binom{X_2}{Y_2}}^{x_1, y_1}(x_2, y_2) \bar{F}_{\binom{X_2}{Y_2}}(x_1, y_1) \, dP_{\binom{X_1}{Y_1}}(x_1, y_1).$$

Due to (11), we deduce that

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

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

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

Moreover, up to scaling, the function

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

can be 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.

Thus, we finally arrive at

$$\Pr(\Pi(X_1, X_2, X_3) = \Pi(Y_1, Y_2, Y_3) = (3, 2, 1))$$

$$\leq \int_{\mathbb{R}^2} \int_{[x_1, \infty[\times [y_1, \infty[} \bar{F}_{\binom{X_3^*}{Y_3^*}}(x_2, y_2) \, dP_{\binom{X_2^*}{Y_2^*}}^{x_1, y_1}(x_2, y_2) \bar{F}_{\binom{X_2^*}{Y_2^*}}(x_1, y_1) \, dP_{\binom{X_1^*}{Y_1^*}}(x_1, y_1)$$

$$= \Pr(X_1^* \leq X_2^* \leq X_3^*, Y_1^* \leq Y_2^* \leq Y_3^*) = \Pr\left(\Pi(X_1^*, X_2^*, X_3^*) = \Pi(Y_1^*, Y_2^*, Y_3^*) = (3, 2, 1)\right). \quad \square$$

Proof of Proposition 5. First, note that

$$\text{OPD}_h\left(\mathbf{X}^{(h)}, \mathbf{Y}^{(h)}\right) = \frac{\sum_{\pi \in T_h} \left\{\Pr\left(\Pi\left(\mathbf{X}^{(h)}\right) = \Pi\left(\mathbf{Y}^{(h)}\right) = \pi\right) - \Pr\left(\Pi\left(\mathbf{X}^{(h)}\right) = \pi\right)\Pr\left(\Pi\left(\mathbf{Y}^{(h)}\right) = \pi\right)\right\}}{1 - \sum_{\pi \in S_h} \Pr\left(\Pi\left(\mathbf{X}^{(h)}\right) = \pi\right)\Pr\left(\Pi\left(\mathbf{Y}^{(h)}\right) = \pi\right)}$$

$$+ \frac{\sum_{\pi \in S_h \setminus T_h} \left\{\Pr\left(\Pi\left(\mathbf{X}^{(h)}\right) = \Pi\left(\mathbf{Y}^{(h)}\right) = \pi\right) - \Pr\left(\Pi\left(\mathbf{X}^{(h)}\right) = \pi\right)\Pr\left(\Pi\left(\mathbf{Y}^{(h)}\right) = \pi\right)\right\}}{1 - \sum_{\pi \in S_h} \Pr\left(\Pi\left(\mathbf{X}^{(h)}\right) = \pi\right)\Pr\left(\Pi\left(\mathbf{Y}^{(h)}\right) = \pi\right)}.$$

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

$$\Pr\left(\Pi\left(\mathbf{X}^{(h)}\right) = \Pi\left(\mathbf{Y}^{(h)}\right) = (0, \ldots, h)\right) = \Pr\left(X_1 \geq \cdots \geq X_{1+h}, Y_1 \geq \cdots \geq Y_{1+h}\right)$$

$$= \Pr\left(\tilde{X}_1 \leq 0, \ldots, \tilde{X}_h \leq 0, \tilde{Y}_1 \leq 0, \ldots, \tilde{Y}_h \leq 0\right)$$

$$= \tau\left(\tilde{X}_1, \ldots, \tilde{X}_h, \tilde{Y}_1, \ldots, \tilde{Y}_h\right)\sqrt{\tilde{p}_{\tilde{\mathbf{X}}^{(h)}}\left(1 - \tilde{p}_{\tilde{\mathbf{X}}^{(h)}}\right)\tilde{p}_{\tilde{\mathbf{Y}}^{(h)}}\left(1 - \tilde{p}_{\tilde{\mathbf{Y}}^{(h)}}\right)} + \tilde{p}_{\tilde{\mathbf{X}}^{(h)}}\tilde{p}_{\tilde{\mathbf{Y}}^{(h)}}.$$

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

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

Let $\pi = (\pi_0, \ldots, \pi_h)$ be a permutation of $(0, \ldots, h)$. If $\Pi(X_1, \ldots, X_{1+h}) = \pi$, it holds that $\left\{X_{1+\pi_0} \geq X_{1+\pi_1} \geq \cdots \geq X_{1+\pi_h}\right\}$. As a result, ordinal pattern dependence can be expressed by the following formula:

$$\text{OPD}_h\left(\mathbf{X}^{(h)}, \mathbf{Y}^{(h)}\right) = \frac{\sum_{\pi \in S_h} \Pr\left(\Pi(X_1, \ldots, X_{1+h}) = \Pi(Y_1, \ldots, Y_{1+h}) = \pi\right) - p_{\mathbf{X}^{(h)}, \pi} p_{\mathbf{Y}^{(h)}, \pi}}{1 - \sum_{\pi \in S_h} \Pr\left(\Pi(X_1, \ldots, X_{1+h}) = \pi\right)\Pr\left(\Pi(Y_1, \ldots, Y_{1+h}) = \pi\right)}$$

$$= \frac{\sum_{\pi \in S_h} \Pr\left(X_{1+\pi_0} \geq X_{1+\pi_1} \geq \cdots \geq X_{1+\pi_h}, Y_{1+\pi_0} \geq Y_{1+\pi_1} \geq \cdots \geq Y_{1+\pi_h}\right) - p_{\mathbf{X}^{(h)}, \pi} p_{\mathbf{Y}^{(h)}, \pi}}{1 - \sum_{\pi \in S_h} p_{\mathbf{X}^{(h)}, \pi} p_{\mathbf{Y}^{(h)}, \pi}}$$

$$= \frac{\sum_{\pi \in S_h} \Pr\left(X_{1+\pi_1} - X_{1+\pi_0} \leq 0, \ldots, X_{1+\pi_h} - X_{1+\pi_{h-1}} \leq 0, \ldots, Y_{1+\pi_h} - Y_{1+\pi_{h-1}} \leq 0\right) - p_{\mathbf{X}^{(h)}, \pi} p_{\mathbf{Y}^{(h)}, \pi}}{1 - \sum_{\pi \in S_h} p_{\mathbf{X}^{(h)}, \pi} p_{\mathbf{Y}^{(h)}, \pi}}$$

$$= \frac{\sum_{\pi \in S_h} \tau\left(X_{1+\pi_1} - X_{1+\pi_0}, \ldots, Y_{1+\pi_h} - Y_{1+\pi_{h-1}}\right)\sqrt{p_{\mathbf{X}^{(h)}, \pi}\left(1 - p_{\mathbf{X}^{(h)}, \pi}\right) p_{\mathbf{Y}^{(h)}, \pi}\left(1 - p_{\mathbf{Y}^{(h)}, \pi}\right)}}{1 - \sum_{\pi \in S_h} p_{\mathbf{X}^{(h)}, \pi} p_{\mathbf{Y}^{(h)}, \pi}}. \quad \square$$

7. Conclusion and outlook

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 equal moving windows. In contrast to other dependence measures, it has thus been developed against a time series background. To make ordinal pattern dependence comparable to other multivariate dependence measures, we adapted the definitions of the latter to the time series approach. We figured out that univariate dependence measures do not carry enough information for an analysis of dependencies between two random vectors. They do not take any dynamical dependence into account, given for example by the cross-correlations of the considered random vectors. The same holds true for the multivariate extension of Pearson's ρ. Hence, both measures are inappropriate in a time series context. For Gaussian observations, there is a close relationship between ordinal pattern dependence and the multivariate version of Kendall's τ. We proved that for $h = 1$ ordinal pattern dependence of two random vectors can be represented as Kendall's τ of the corresponding increment vectors. The provided simulations show that the values of Kendall's τ and ordinal pattern dependence of the same data set differ in concrete situations. This emphasizes that the two measures operate on different levels, i.e., one operates on the level of the original process, the other on the level of increments. Moreover, for this reason ordinal pattern dependence is insensitive with respect to time shifts: it only relies on the dependence between the corresponding increments. However, an extension of ordinal pattern dependence that is sensitive to dependence shifted in time is given in [16]. The authors of that article introduced time shifted ordinal pattern dependence that allows for time shifts between the moving windows of the two time series. Using this approach it is possible to detect a co-movement of the two time series that does not happen in the same moving window. Furthermore, it is an interesting topic for further research to compare ordinal pattern dependence to copula-based dependence measures. Originally, copulas were introduced to focus on the dependence *within* a multivariate random vector without taking the marginal distributions into account. If all univariate margins are continuous, the multivariate generalizations of Kendall's τ and multivariate Spearman's ρ in [7] only depend on the underlying copula. As these two measures are defined to measure dependence between two random vectors, this approach is a promising starting point to extend these ideas to a time series background. An investigation of ordinal pattern dependence with respect to the framework of copulas as well as a comparison to multivariate Spearman's ρ is ongoing research, but not in the scope of the present article. Since ordinal pattern dependence admits a canonical interpretation, and since limit theorems in the short and the long-range dependent framework are at hand, we suggest to use this measure to complement the classical time series analysis with an ordinal point of view.

Table A.1

We generate X_i, $1 \leq i \leq n$, and Y_i, $1 \leq i \leq n$, 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, $1 \leq i \leq n$, and η_i, $1 \leq i \leq n$, are two independent sequences of random variables, both i.i.d. standard normally distributed.

Method	n	h	$\rho = 0.1$		$\rho = 0.5$		$\rho = 0.9$	
			Mean (sd)	Median (IQR)	Mean (sd)	Median (IQR)	Mean (sd)	Median (IQR)
Kendall's τ	100	1	0.007 (0.062)	0.004 (0.086)	0.012 (0.087)	0.011 (0.121)	0.018 (0.168)	0.017 (0.238)
	300	1	0.003 (0.035)	0.003 (0.048)	0.004 (0.051)	0.004 (0.068)	0.008 (0.111)	0.007 (0.153)
	500	1	0.002 (0.027)	0.002 (0.036)	0.003 (0.040)	0.002 (0.053)	0.006 (0.088)	0.004 (0.119)
opd	100	1	0.009 (0.106)	0.003 (0.139)	0.011 (0.101)	0.016 (0.137)	0.008 (0.098)	0.003 (0.137)
	300	1	0.003 (0.062)	0.002 (0.085)	0.004 (0.059)	0.005 (0.083)	0.003 (0.058)	0.004 (0.075)
	500	1	0.002 (0.048)	0.004 (0.062)	0.003 (0.046)	0.003 (0.063)	0.001 (0.044)	-0.000 (0.061)
Kendall's τ	100	2	0.008 (0.053)	0.002 (0.070)	0.015 (0.083)	0.011 (0.114)	0.031 (0.170)	0.027 (0.239)
	300	2	0.003 (0.030)	0.002 (0.040)	0.005 (0.048)	0.004 (0.065)	0.013 (0.111)	0.011 (0.154)
	500	2	0.002 (0.023)	0.001 (0.031)	0.003 (0.038)	0.002 (0.050)	0.009 (0.089)	0.007 (0.120)
opd	100	2	0.003 (0.055)	-0.000 (0.073)	0.004 (0.055)	0.001 (0.077)	0.004 (0.056)	0.002 (0.075)
	300	2	0.001 (0.033)	-0.000 (0.044)	0.002 (0.032)	0.001 (0.044)	0.001 (0.033)	0.001 (0.044)
	500	2	0.000 (0.025)	0.000 (0.034)	0.001 (0.025)	0.001 (0.034)	0.000 (0.025)	-0.001 (0.034)
Kendall's τ	100	3	0.006 (0.043)	-0.002 (0.053)	0.015 (0.076)	0.007 (0.103)	0.040 (0.169)	0.031 (0.239)
	300	3	0.002 (0.024)	-0.000 (0.032)	0.005 (0.044)	0.002 (0.059)	0.016 (0.111)	0.014 (0.154)
	500	3	0.001 (0.019)	0.000 (0.024)	0.003 (0.035)	0.001 (0.047)	0.012 (0.089)	0.009 (0.120)
opd	100	3	0.001 (0.027)	-0.001 (0.035)	0.001 (0.027)	-0.002 (0.037)	0.002 (0.031)	-0.001 (0.041)
	300	3	0.001 (0.015)	-0.001 (0.021)	0.001 (0.016)	-0.000 (0.021)	0.001 (0.018)	-0.001 (0.024)
	500	3	-0.000 (0.012)	-0.001 (0.016)	0.000 (0.013)	-0.000 (0.017)	0.000 (0.014)	-0.000 (0.019)

Table A.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 (10).

Method	n	h	$\rho = 0.1$		$\rho = 0.2$		$\rho = 0.3$	
			Mean (sd)	Median (IQR)	Mean (sd)	Median (IQR)	Mean (sd)	Median (IQR)
Kendall's τ	100	1	0.045 (0.034)	0.045 (0.047)	0.091 (0.037)	0.091 (0.050)	0.141 (0.039)	0.141 (0.053)
	300	1	0.045 (0.020)	0.045 (0.026)	0.091 (0.021)	0.091 (0.028)	0.142 (0.022)	0.142 (0.030)
	500	1	0.045 (0.015)	0.044 (0.021)	0.091 (0.016)	0.091 (0.023)	0.142 (0.017)	0.142 (0.023)
opd	100	1	0.066 (0.064)	0.066 (0.081)	0.129 (0.063)	0.130 (0.086)	0.196 (0.066)	0.198 (0.087)
	300	1	0.065 (0.036)	0.064 (0.048)	0.128 (0.037)	0.129 (0.051)	0.195 (0.038)	0.195 (0.053)
	500	1	0.065 (0.029)	0.065 (0.038)	0.128 (0.028)	0.128 (0.039)	0.195 (0.030)	0.194 (0.040)
Kendall's τ	100	2	0.031 (0.030)	0.029 (0.041)	0.063 (0.033)	0.062 (0.045)	0.100 (0.036)	0.100 (0.050)
	300	2	0.030 (0.017)	0.029 (0.023)	0.062 (0.019)	0.062 (0.026)	0.100 (0.021)	0.100 (0.029)
	500	2	0.030 (0.013)	0.030 (0.018)	0.063 (0.015)	0.062 (0.019)	0.100 (0.016)	0.100 (0.022)
opd	100	2	0.031 (0.034)	0.031 (0.047)	0.065 (0.037)	0.063 (0.051)	0.101 (0.040)	0.099 (0.054)
	300	2	0.030 (0.020)	0.030 (0.027)	0.064 (0.022)	0.063 (0.029)	0.102 (0.024)	0.101 (0.031)
	500	2	0.030 (0.016)	0.030 (0.021)	0.064 (0.017)	0.063 (0.023)	0.102 (0.018)	0.102 (0.025)
Kendall's τ	100	3	0.019 (0.024)	0.017 (0.032)	0.041 (0.029)	0.039 (0.038)	0.067 (0.033)	0.064 (0.043)
	300	3	0.019 (0.014)	0.018 (0.019)	0.042 (0.017)	0.041 (0.022)	0.068 (0.019)	0.067 (0.026)
	500	3	0.019 (0.011)	0.018 (0.015)	0.041 (0.013)	0.041 (0.018)	0.069 (0.015)	0.068 (0.020)
opd	100	3	0.011 (0.017)	0.010 (0.023)	0.024 (0.020)	0.023 (0.028)	0.041 (0.023)	0.040 (0.032)
	300	3	0.011 (0.010)	0.011 (0.014)	0.024 (0.012)	0.024 (0.015)	0.041 (0.013)	0.040 (0.018)
	500	3	0.011 (0.008)	0.011 (0.010)	0.024 (0.009)	0.024 (0.012)	0.041 (0.010)	0.041 (0.014)

Acknowledgments

Appendix. Simulation study

In Table A.1 to A.3, we empirically compare multivariate Kendall's τ and ordinal pattern dependence by simulation studies for different settings.

Table A.3

We generate X_i, $1 \leq i \leq n$, from an AR(1) time series with parameter ρ, while Y_i, $1 \leq i \leq n$, corresponds to X_i, $1 \leq i \leq n$, shifted by one time point, i.e., $Y_i = X_{i+1}$.

Method	n	h	$\rho = 0.1$		$\rho = 0.5$		$\rho = 0.9$	
			Mean (sd)	Median (IQR)	Mean (sd)	Median (IQR)	Mean (sd)	Median (IQR)
Kendall's τ	100	1	0.354 (0.059)	0.354 (0.080)	0.509 (0.058)	0.512 (0.077)	0.738 (0.055)	0.744 (0.072)
	300	1	0.360 (0.034)	0.361 (0.046)	0.521 (0.033)	0.522 (0.044)	0.777 (0.030)	0.779 (0.039)
	500	1	0.361 (0.026)	0.361 (0.036)	0.524 (0.025)	0.525 (0.033)	0.784 (0.023)	0.786 (0.031)
opd	100	1	−0.282 (0.083)	−0.279 (0.112)	−0.152 (0.090)	−0.157 (0.118)	−0.026 (0.097)	−0.023 (0.131)
	300	1	−0.292 (0.049)	−0.292 (0.066)	−0.159 (0.053)	−0.159 (0.071)	−0.030 (0.057)	−0.028 (0.074)
	500	1	−0.295 (0.038)	−0.295 (0.050)	−0.158 (0.041)	−0.157 (0.054)	−0.030 (0.044)	−0.029 (0.060)
Kendall's τ	100	2	0.437 (0.070)	0.438 (0.095)	0.575 (0.065)	0.579 (0.086)	0.783 (0.055)	0.791 (0.071)
	300	2	0.450 (0.038)	0.450 (0.052)	0.591 (0.036)	0.593 (0.049)	0.816 (0.027)	0.817 (0.036)
	500	2	0.453 (0.030)	0.454 (0.041)	0.594 (0.027)	0.595 (0.037)	0.823 (0.021)	0.824 (0.028)
opd	100	2	−0.088 (0.037)	−0.088 (0.049)	−0.030 (0.044)	−0.031 (0.060)	0.050 (0.052)	0.049 (0.070)
	300	2	−0.088 (0.021)	−0.088 (0.028)	−0.028 (0.025)	−0.029 (0.033)	0.052 (0.030)	0.051 (0.041)
	500	2	−0.087 (0.016)	−0.087 (0.022)	−0.028 (0.019)	−0.028 (0.025)	0.052 (0.024)	0.052 (0.032)
Kendall's τ	100	3	0.462 (0.083)	0.465 (0.113)	0.602 (0.075)	0.609 (0.101)	0.804 (0.058)	0.814 (0.073)
	300	3	0.485 (0.047)	0.485 (0.062)	0.624 (0.040)	0.625 (0.053)	0.837 (0.027)	0.840 (0.035)
	500	3	0.489 (0.035)	0.489 (0.047)	0.629 (0.030)	0.630 (0.040)	0.844 (0.020)	0.845 (0.026)
opd	100	3	−0.029 (0.016)	−0.030 (0.022)	−0.008 (0.024)	−0.010 (0.033)	0.040 (0.038)	0.037 (0.050)
	300	3	−0.025 (0.010)	−0.025 (0.013)	−0.003 (0.014)	−0.004 (0.019)	0.045 (0.022)	0.045 (0.030)
	500	3	−0.024 (0.007)	−0.024 (0.010)	−0.002 (0.011)	−0.002 (0.015)	0.046 (0.017)	0.046 (0.023)

References

[1] J. Aaronson, R. Burton, H. Dehling, D. Gilat, T. Hill, B. Weiss, Strong laws for L- and U-statistics, Trans. Amer. Math. Soc. 348 (7) (1996) 2845–2866.

[2] C. Bandt, Order patterns, their variation and change points in financial time series and Brownian motion, Statist. Papers 61 (2020) 1565–1588.

[3] C. Bandt, B. Pompe, Permutation entropy: A natural complexity measure for time series, Phys. Rev. Lett. 88 (17) (2002) 174102–1 – 4.

[4] C. Bandt, F. Shiha, Order patterns in time series, J. Time Series Anal. 28 (5) (2007) 646–665.

[5] H. Dehling, D. Vogel, M. Wendler, D. Wied, Testing for changes in Kendall's tau, Econom. Theory 33 (6) (2017) 1352–1386.

[6] I. Echegoyen, V. Vera-Ávila, R. Sevilla-Escoboza, J.H. Martinez, J.M. Buldu, Ordinal synchronization: Using ordinal patterns to capture interdependencies between time series, Chaos Solitons Fractals 119 (2019) 8–18.

[7] O. Grothe, J. Schnieders, J. Segers, Measuring association and dependence between random vectors, J. Multivariate Anal. 123 (2014) 96–110.

[8] H. Joe, Multivariate concordance, J. Multivariate Anal. 35 (1) (1990) 12–30.

[9] M. Mohr, F. Wilhelm, M. Hartwig, R. Möller, K. Keller, New approaches in ordinal pattern representations for multivariate time series, in: Proceedings of the 33rd Florida Artificial Intelligence Research Society Conference (FLAIRS-33), 2020, pp. 124–129.

[10] I. Nüßgen, Ordinal Pattern Analysis: Limit Theorems for Multivariate Long-Range Dependent Gaussian Time Series and a Comparison to Multivariate Dependence Measures, (Ph.D. thesis), University of Siegen, 2021.

[11] I. Nüßgen, A. Schnurr, Ordinal pattern dependence in the context of long-range dependence, Entropy (23) (2021) 670.

[12] G. Puccetti, Measuring linear correlation between random vectors, 2019, http://dx.doi.org/10.2139/Ssrn.3116066, Available at SSRN: https://ssrn.com/abstract=3116066.

[13] J.-F. Quessy, M. Saïd, A.-C. Favre, Multivariate Kendall's tau for change-point detection in copulas, Canad. J. Statist. 41 (1) (2013) 65–82.

[14] F. Schmid, R. Schmidt, T. Blumentritt, S. Gaißer, M. Ruppert, Copula-based measures of multivariate association, in: P. Jaworski, F. Durante, W.K. Härdle, T. Rychlik (Eds.), Copula Theory and Its Applications, in: Lecture Notes in Statistics, vol. 198, Springer, Berlin, Heidelberg, 2010, pp. 209–236.

[15] A. Schnurr, An ordinal pattern approach to detect and to model leverage effects and dependence structures between financial time series, Statist. Papers 55 (4) (2014) 919–931.

[16] A. Schnurr, H. Dehling, Testing for structural breaks via ordinal pattern dependence, J. Amer. Statist. Assoc. 112 (518) (2017) 706–720.

[17] M. Sinn, K. Keller, Estimation of ordinal pattern probabilities in Gaussian processes with stationary increments, Comput. Statist. Data Anal. 55 (4) (2011) 1781–1790.

[18] J.M. Wooldridge, H. White, Some invariance principles and central limit theorems for dependent heterogeneous processes, Econom. Theory 4 (2) (1988) 210–230.