

# Detecting changes in cross-sectional dependence in multivariate time series

Axel Bücher

Institut für Angewandte Mathematik

Universität Heidelberg

Im Neuenheimer Feld 294, 69120 Heidelberg, Germany

axel.buecher@rub.de

Ivan Kojadinovic, Tom Rohmer

Laboratoire de mathématiques et applications, UMR CNRS 5142

Université de Pau et des Pays de l'Adour

B.P. 1155, 64013 Pau Cedex, France

{ivan.kojadinovic,tom.rohmer}@univ-pau.fr

Johan Segers

Institut de statistique, biostatistique et sciences actuarielles

Université catholique de Louvain

Voie du Roman Pays 20, B-1348 Louvain-la-Neuve, Belgium

johan.segers@uclouvain.be

May 22, 2014

## Abstract

Classical and more recent tests for detecting distributional changes in multivariate time series often lack power against alternatives that involve changes in the cross-sectional dependence structure. To be able to detect such changes better, a test is introduced based on a recently studied variant of the sequential empirical copula process. In contrast to earlier attempts, ranks are computed with respect to relevant subsamples, with beneficial consequences for the sensitivity of the test. For the computation of p-values we propose a multiplier resampling scheme that takes the serial dependence into account. The large-sample theory for the test statistic and the resampling scheme is developed. The finite-sample performance of the procedure is assessed by Monte Carlo simulations. Two case studies involving time series of financial returns are presented as well.

# 1 Introduction

Given a sequence  $\mathbf{X}_1, \dots, \mathbf{X}_n$  of  $d$ -dimensional observations, change-point detection aims at testing

$$H_0 : \exists F \text{ such that } \mathbf{X}_1, \dots, \mathbf{X}_n \text{ have c.d.f. } F \quad (1.1)$$

against alternatives involving the nonconstancy of the c.d.f. Under  $H_0$  and the assumption that  $\mathbf{X}_1, \dots, \mathbf{X}_n$  have continuous marginal c.d.f.s  $F_1, \dots, F_d$ , we have from the work of Sklar (1959) that the common multivariate c.d.f.  $F$  can be written in a unique way as

$$F(\mathbf{x}) = C\{F_1(x_1), \dots, F_d(x_d)\}, \quad \mathbf{x} \in \mathbb{R}^d,$$

where the function  $C : [0, 1]^d \rightarrow [0, 1]$  is a *copula* and can be regarded as capturing the dependence between the components of  $\mathbf{X}_1, \dots, \mathbf{X}_n$ . It follows that  $H_0$  can be rewritten as  $H_{0,m} \cap H_{0,c}$ , where

$$H_{0,m} : \exists F_1, \dots, F_d \text{ such that } \mathbf{X}_1, \dots, \mathbf{X}_n \text{ have marginal c.d.f.s } F_1, \dots, F_d, \quad (1.2)$$

$$H_{0,c} : \exists C \text{ such that } \mathbf{X}_1, \dots, \mathbf{X}_n \text{ have copula } C. \quad (1.3)$$

Classical nonparametric tests for  $H_0$  are based on sequential empirical processes; see e.g. Bai (1994), Csörgő and Horváth (1997, Section 2.6) and Inoue (2001). For moderate sample sizes, however, such tests appear to have little power against alternative hypotheses that leave the margins unchanged but that involve a change in the copula, i.e., when  $H_{0,m} \cap (\neg H_{0,c})$  holds. Empirical evidence of the latter fact can be found in Holmes et al. (2013, Section 4). For that reason, nonparametric tests for change-point detection particularly sensitive to changes in the dependence structure are of practical interest.

Several tests designed to capture changes in cross-sectional dependence structure were proposed in the literature. Tests based on Kendall's tau were investigated by Gombay and Horváth (1999) (see also Gombay and Horváth, 2002), Quessy et al. (2013) and Dehling et al. (2013). Although these have good power when the copula changes in such a way that Kendall's tau changes as well, they are obviously useless when the copula changes but Kendall's tau does not change or only very little. Tests based on functionals of sequential empirical copula processes were considered in Rémillard (2010), Bücher and Ruppert (2013), van Kampen and Wied (2013) and Wied et al. (2013). However, the power of such tests is often disappointing; see Section 5 for some numerical evidence.

It is our aim to construct a new test for  $H_0$  that is more powerful than its predecessors against alternatives that involve a change in the copula. The test is based on sequential empirical copula processes as well, but the crucial difference lies in the computation of the ranks. Whereas in Rémillard (2010) and subsequent papers, ranks are always computed with respect to the full sample, we propose to compute the ranks with respect to the relevant subsamples; see Section 2 for details. The intuition is that in this way, the copulas of those subsamples are estimated more accurately, so that differences between copulas of disjoint subsamples are detected more quickly. The phenomenon is akin to the one observed in Genest and Segers (2010) that the empirical copula, which is based on

pseudo-observations, is often a better estimator of a copula than the empirical distribution function based on observations from the copula itself. For another illustration in the context of tail dependence functions, see Bücher (2013a).

The paper is organized as follows. The test statistic is presented in Section 2, and its asymptotic distribution under the null hypothesis is found in Section 3. Next, Section 4 contains a detailed description of the multiplier resampling scheme and its asymptotic validity under the null hypothesis. The results of a large-scale Monte Carlo simulation study are reported in Section 5, and two brief case studies are given in Section 6. Section 7 concludes. Proofs and details regarding the simulation study are deferred to the Appendices.

In the rest of the paper, the arrow ‘ $\rightsquigarrow$ ’ denotes weak convergence in the sense of Definition 1.3.3 in van der Vaart and Wellner (2000). Given a set  $T$ , let  $\ell^\infty(T)$  denote the space of all bounded real-valued functions on  $T$  equipped with the uniform metric.

## 2 Test statistic

We now describe our test statistic and highlight the difference with the one in Rémillard (2010) and Bücher and Ruppert (2013). Let  $\mathbf{X}_1, \dots, \mathbf{X}_n$  be random vectors. For integers  $1 \leq k \leq l \leq n$ , let  $C_{k:l}$  be the empirical copula of the sample  $\mathbf{X}_k, \dots, \mathbf{X}_l$ . Specifically,

$$C_{k:l}(\mathbf{u}) = \frac{1}{l-k+1} \sum_{i=k}^l \mathbf{1}(\hat{\mathbf{U}}_i^{k:l} \leq \mathbf{u}), \quad (2.1)$$

for  $\mathbf{u} \in [0, 1]^d$ , where

$$\hat{\mathbf{U}}_i^{k:l} = \frac{1}{l-k+1} (R_{i1}^{k:l}, \dots, R_{id}^{k:l}), \quad i \in \{k, \dots, l\}, \quad (2.2)$$

with  $R_{ij}^{k:l} = \sum_{t=k}^l \mathbf{1}(X_{tj} \leq X_{ij})$  the (maximal) rank of  $X_{ij}$  among  $X_{kj}, \dots, X_{lj}$ . (Because of serial dependence, there can be ties, even if the marginal distribution is continuous; think for instance of a moving maximum process.) An important point is that the ranks are computed within the subsample  $\mathbf{X}_k, \dots, \mathbf{X}_l$  and not within the whole sample  $\mathbf{X}_1, \dots, \mathbf{X}_n$ . As we continue, we adopt the convention that  $C_{k:l} = 0$  if  $k > l$ .

Write  $\Delta = \{(s, t) \in [0, 1]^2 : s \leq t\}$ . Let  $\lambda_n(s, t) = (\lfloor nt \rfloor - \lfloor ns \rfloor)/n$  for  $(s, t) \in \Delta$ . Our test statistic is based on the difference process,  $\mathbb{D}_n$ , defined by

$$\mathbb{D}_n(s, \mathbf{u}) = \sqrt{n} \lambda_n(0, s) \lambda_n(s, 1) \{C_{1:\lfloor ns \rfloor}(\mathbf{u}) - C_{\lfloor ns \rfloor + 1:n}(\mathbf{u})\} \quad (2.3)$$

for  $(s, \mathbf{u}) \in [0, 1]^{d+1}$ . For every  $s \in [0, 1]$ , it gives a weighted difference between the empirical copulas at  $\mathbf{u}$  of the first  $\lfloor ns \rfloor$  and the last  $n - \lfloor ns \rfloor$  points of the sample. Large absolute differences point in the direction of a change in the copula.

To aggregate over  $\mathbf{u}$ , we consider the Cramér–von Mises statistic

$$S_{n,k} = \int_{[0,1]^d} \{\mathbb{D}_n(k/n, \mathbf{u})\}^2 dC_{1:n}(\mathbf{u}), \quad k \in \{1, \dots, n-1\}.$$

The test statistic for detecting changes in cross-sectional dependence is then

$$S_n = \max_{1 \leq k \leq n-1} S_{n,k} = \sup_{s \in [0,1]} \int_{[0,1]^d} \{\mathbb{D}_n(s, \mathbf{u})\}^2 dC_{1:n}(\mathbf{u}). \quad (2.4)$$

Other aggregating functions can be thought of too, leading for instance to Kolmogorov–Smirnov and Kuiper statistics. In numerical experiments, the resulting tests were found to be less powerful than the one based on the Cramér–von Mises statistic and hence are not considered further in this paper.

The null hypothesis of a constant distribution is rejected when  $S_n$  is large. The p-values are determined by the null distribution of  $S_n$ , whose large-sample limit is derived in Section 3. To estimate the p-values from the data, a multiplier bootstrap method is proposed in Section 4.

Finally, if  $H_0$  is rejected, there could be one or several abrupt or smooth changes in the joint distribution. Moreover, the change(s) could concern one or more marginal distributions, the copula, or both. In the case where there is just a single (abrupt) change-point  $k^* \in \{1, \dots, n-1\}$ , one can for instance estimate it by

$$k_n^* = \arg \max_{1 \leq k \leq n-1} S_{n,k}. \quad (2.5)$$

We do not pursue the issue of single or multiple change-point estimation nor the diagnosis of the nature of the change-point.

Our test statistic  $S_n$  differs from the one considered in Rémillard (2010, Section 5.2) and Bücher and Ruppert (2013, Section 3.2) in the way the copulas of the subsamples  $\mathbf{X}_k, \dots, \mathbf{X}_l$  are estimated. Rather than the empirical copula  $C_{k:l}$ , these authors propose to use

$$C_{k:l,n}(\mathbf{u}) = \frac{1}{l-k+1} \sum_{i=k}^l \mathbf{1}(\hat{U}_i^{1:n} \leq \mathbf{u}), \quad \mathbf{u} \in [0,1]^d, \quad (2.6)$$

with the convention that  $C_{k:l,n} = 0$  if  $k > l$ . In comparison with  $C_{k:l}$  in (2.1), the ranks for the subsample  $\mathbf{X}_k, \dots, \mathbf{X}_l$  are computed relative to the complete sample  $\mathbf{X}_1, \dots, \mathbf{X}_n$ . The estimators  $C_{k:l,n}$  yield the difference process

$$\mathbb{D}_n^R(s, \mathbf{u}) = \sqrt{n} \lambda_n(0, s) \lambda_n(s, 1) \{C_{1:[ns],n}(\mathbf{u}) - C_{[ns]+1:n,n}(\mathbf{u})\} \quad (2.7)$$

for  $(s, \mathbf{u}) \in [0,1]^{d+1}$ . The process  $\mathbb{D}_n^R$  is to be compared with the process  $\mathbb{D}_n$  in (2.3). The difference lies in the use of  $C_{k:l,n}$  rather than  $C_{k:l}$ . From the process  $\mathbb{D}_n^R$ , one obtains the test statistic

$$S_n^R = \sup_{s \in [0,1]} \int_{[0,1]^d} \{\mathbb{D}_n^R(s, \mathbf{u})\}^2 dC_{1:n}(\mathbf{u}), \quad (2.8)$$

which is the analogue of  $S_n$  in (2.4).

In the Monte Carlo simulation experiments (Section 5), we will see that  $S_n$  is usually more powerful than  $S_n^R$  for detecting changes in the cross-sectional copula. Intuitively, the reason is that the empirical copula  $C_{k:l}$  in (2.1) is often a better copula estimator than  $C_{k:l,n}$  in (2.6). Note that  $C_{k:l}$  is not only the empirical copula of  $\mathbf{X}_k, \dots, \mathbf{X}_l$ , it

is also equal to the empirical copula of  $\hat{U}_k^{1:n}, \dots, \hat{U}_l^{1:n}$ , of which  $C_{k:l,n}$  is the empirical distribution function.

In Genest and Segers (2010), situations are identified where the empirical copula of an independent random sample drawn from a given bivariate copula has a lower asymptotic variance than the empirical distribution function of that sample. Of course, the situation here is different from the one in the cited paper: multivariate rather than bivariate, serial dependence rather than independence. But still, we suspect the same mechanisms to be active.

### 3 Large-sample distribution

The asymptotic distribution under  $H_0$  of our test statistic  $S_n$  in (2.4) can be obtained by writing it as a functional of the two-sided sequential empirical copula process studied in Bücher and Kojadinovic (2013). Let  $\mathbf{X}_1, \mathbf{X}_2, \dots$  be a strictly stationary  $d$ -variate time series with stationary c.d.f.  $F$  having continuous margins  $F_1, \dots, F_d$  and copula  $C$ . Recall  $C_{k:l}$  in (2.1) and  $\hat{U}_i^{k:l}$  in (2.2). The two-sided sequential empirical copula process,  $\mathbb{C}_n$ , is defined by

$$\mathbb{C}_n(s, t, \mathbf{u}) = \sqrt{n} \lambda_n(s, t) \{C_{\lfloor ns \rfloor + 1: \lfloor nt \rfloor}(\mathbf{u}) - C(\mathbf{u})\} \quad (3.1)$$

$$= \frac{1}{\sqrt{n}} \sum_{i=\lfloor ns \rfloor + 1}^{\lfloor nt \rfloor} \left\{ \mathbf{1}(\hat{U}_i^{\lfloor ns \rfloor + 1: \lfloor nt \rfloor} \leq \mathbf{u}) - C(\mathbf{u}) \right\}, \quad (3.2)$$

for  $(s, t, \mathbf{u}) \in \Delta \times [0, 1]^d$ . The link of  $\mathbb{C}_n$  to our test statistic  $S_n$  in (2.4) is that, under  $H_0$ , the difference process  $\mathbb{D}_n$  in (2.3) can be written as

$$\mathbb{D}_n(s, \mathbf{u}) = \lambda_n(s, 1) \mathbb{C}_n(0, s, \mathbf{u}) - \lambda_n(0, s) \mathbb{C}_n(s, 1, \mathbf{u}), \quad (3.3)$$

for  $(s, \mathbf{u}) \in [0, 1]^{d+1}$ .

Before focusing on the weak limit of the process  $\mathbb{D}_n$  under  $H_0$ , let us briefly recall the notion of *strongly mixing sequence*. For a sequence of  $d$ -dimensional random vectors  $(\mathbf{Y}_i)_{i \in \mathbb{Z}}$ , the  $\sigma$ -field generated by  $(\mathbf{Y}_i)_{a \leq i \leq b}$ ,  $a, b \in \mathbb{Z} \cup \{-\infty, +\infty\}$ , is denoted by  $\mathcal{F}_a^b$ . The strong mixing coefficients corresponding to the sequence  $(\mathbf{Y}_i)_{i \in \mathbb{Z}}$  are defined by

$$\alpha_r = \sup_{p \in \mathbb{Z}} \sup_{A \in \mathcal{F}_{-\infty}^p, B \in \mathcal{F}_{p+r}^{+\infty}} |P(A \cap B) - P(A)P(B)|$$

for positive integer  $r$ . The sequence  $(\mathbf{Y}_i)_{i \in \mathbb{Z}}$  is said to be *strongly mixing* if  $\alpha_r \rightarrow 0$  as  $r \rightarrow \infty$ .

The weak limit of the two-sided empirical copula process  $\mathbb{C}_n$  defined in (3.2) under strong mixing was established in Bücher and Kojadinovic (2013) under the following two conditions:

**Condition 3.1.** *With probability one, there are no ties in each of the  $d$  component series  $X_{1j}, X_{2j}, \dots$ , where  $j \in \{1, \dots, d\}$ .*

**Condition 3.2.** For any  $j \in \{1, \dots, d\}$ , the partial derivatives  $\dot{C}_j = \partial C / \partial u_j$  exist and are continuous on  $V_j = \{\mathbf{u} \in [0, 1]^d : u_j \in (0, 1)\}$ .

Condition 3.1 was considered in Bücher and Kojadinovic (2013) as continuity of the marginal distributions is *not* sufficient to guarantee the absence of ties when the observations are serially dependent (see e.g. Bücher and Segers, 2013, Example 4.2). One of the contributions of this work is to show that it actually can be dispensed with. Condition 3.2 was proposed in Segers (2012) and is nonrestrictive in the sense that it is necessary for the candidate weak limit of  $\mathbb{C}_n$  to exist pointwise and have continuous trajectories.

As we continue, for any  $j \in \{1, \dots, d\}$ , we define  $\dot{C}_j$  to be zero on the set  $\{\mathbf{u} \in [0, 1]^d : u_j \in \{0, 1\}\}$  (see also Segers, 2012; Bücher and Volgushev, 2013). Also, for any  $j \in \{1, \dots, d\}$  and any  $\mathbf{u} \in [0, 1]^d$ ,  $\mathbf{u}^{(j)}$  is the vector of  $[0, 1]^d$  defined by  $u_i^{(j)} = u_j$  if  $i = j$  and 1 otherwise.

The weak convergence of the process  $\mathbb{C}_n$  defined in (3.2) actually follows from that of the process

$$\mathbb{B}_n(s, t, \mathbf{u}) = \frac{1}{\sqrt{n}} \sum_{i=\lfloor ns \rfloor + 1}^{\lfloor nt \rfloor} \{\mathbf{1}(\mathbf{U}_i \leq \mathbf{u}) - C(\mathbf{u})\}, \quad (s, t, \mathbf{u}) \in \Delta \times [0, 1]^d, \quad (3.4)$$

where  $\mathbf{U}_1, \dots, \mathbf{U}_n$  is the unobservable sample obtained from  $\mathbf{X}_1, \dots, \mathbf{X}_n$  by the probability integral transforms  $U_{ij} = F_j(X_{ij})$ ,  $i \in \{1, \dots, n\}$ ,  $j \in \{1, \dots, d\}$ , and with the convention that  $\mathbb{B}_n(s, t, \cdot) = 0$  if  $\lfloor nt \rfloor - \lfloor ns \rfloor = 0$ .

If  $\mathbf{U}_1, \dots, \mathbf{U}_n$  is drawn from a strictly stationary sequence  $(\mathbf{U}_i)_{i \in \mathbb{Z}}$  whose strong mixing coefficients satisfy  $\alpha_r = O(r^{-a})$  with  $a > 1$ , we have from Bücher (2013b) that  $\mathbb{B}_n(0, \cdot, \cdot)$  converges weakly in  $\ell^\infty([0, 1]^{d+1})$  to a tight centered Gaussian process  $\mathbb{Z}_C$  with covariance function

$$\text{cov}\{\mathbb{Z}_C(s, \mathbf{u}), \mathbb{Z}_C(t, \mathbf{v})\} = \min(s, t) \sum_{k \in \mathbb{Z}} \text{cov}\{\mathbf{1}(\mathbf{U}_0 \leq \mathbf{u}), \mathbf{1}(\mathbf{U}_k \leq \mathbf{v})\}.$$

The latter is actually a consequence of Lemma 2 in Bücher (2013b) stating that  $\mathbb{B}_n(0, \cdot, \cdot)$  is asymptotically uniformly equicontinuous in probability, which in turn implies that  $\mathbb{Z}_C$  has continuous trajectories with probability one. As a consequence of the continuous mapping theorem,  $\mathbb{B}_n \rightsquigarrow \mathbb{B}_C$  in  $\ell^\infty(\Delta \times [0, 1]^d)$ , where

$$\mathbb{B}_C(s, t, \mathbf{u}) = \mathbb{Z}_C(t, \mathbf{u}) - \mathbb{Z}_C(s, \mathbf{u}), \quad (s, t, \mathbf{u}) \in \Delta \times [0, 1]^d. \quad (3.5)$$

The following result is a consequence of Theorem 3.4 of Bücher and Kojadinovic (2013) and the arguments used in the proof of Lemma A.2 of Bücher and Segers (2013). Its proof is given in Appendix A.

**Proposition 3.3.** Let  $\mathbf{X}_1, \dots, \mathbf{X}_n$  be drawn from a strictly stationary sequence  $(\mathbf{X}_i)_{i \in \mathbb{Z}}$  with continuous margins and whose strong mixing coefficients satisfy  $\alpha_r = O(r^{-a})$ ,  $a > 1$ . Then, provided Condition 3.2 holds,

$$\sup_{(s, t, \mathbf{u}) \in \Delta \times [0, 1]^d} \left| \mathbb{C}_n(s, t, \mathbf{u}) - \tilde{\mathbb{C}}_n(s, t, \mathbf{u}) \right| \xrightarrow{P} 0, \quad (3.6)$$

where

$$\tilde{\mathbb{C}}_n(s, t, \mathbf{u}) = \mathbb{B}_n(s, t, \mathbf{u}) - \sum_{j=1}^d \dot{C}_j(\mathbf{u}) \mathbb{B}_n(s, t, \mathbf{u}^{(j)}). \quad (3.7)$$

Consequently,  $\mathbb{C}_n \rightsquigarrow \mathbb{C}_C$  in  $\ell^\infty(\Delta \times [0, 1]^d)$ , where, for  $(s, t, \mathbf{u}) \in \Delta \times [0, 1]^d$ ,

$$\mathbb{C}_C(s, t, \mathbf{u}) = \mathbb{B}_C(s, t, \mathbf{u}) - \sum_{j=1}^d \dot{C}_j(\mathbf{u}) \mathbb{B}_C(s, t, \mathbf{u}^{(j)}). \quad (3.8)$$

In view of (3.3), the weak limit of  $\mathbb{D}_n$  under  $H_0$  is a mere corollary of Proposition 3.3 and the continuous mapping theorem.

**Corollary 3.4.** *Under the conditions of Proposition 3.3,  $\mathbb{D}_n \rightsquigarrow \mathbb{D}_C$  in  $\ell^\infty([0, 1]^{d+1})$ , where, for any  $(s, \mathbf{u}) \in [0, 1]^{d+1}$ ,*

$$\mathbb{D}_C(s, \mathbf{u}) = \mathbb{C}_C(0, s, \mathbf{u}) - s \mathbb{C}_C(0, 1, \mathbf{u}), \quad (3.9)$$

with  $\mathbb{C}_C$  defined in (3.8). As a consequence,

$$S_n \rightsquigarrow S = \sup_{s \in [0, 1]} \int_{[0, 1]^d} \{\mathbb{D}_C(s, \mathbf{u})\}^2 dC(\mathbf{u}). \quad (3.10)$$

The covariance function of  $\mathbb{D}_C$  can be expressed in terms of the one of  $\mathbb{C}_C$  by

$$\text{cov}\{\mathbb{D}_C(s, \mathbf{u}), \mathbb{D}_C(t, \mathbf{v})\} = \{\min(s, t) - st\} \text{cov}\{\mathbb{C}_C(0, 1, \mathbf{u}), \mathbb{C}_C(0, 1, \mathbf{v})\}.$$

## 4 Resampling

In order to compute p-values for  $S_n$  based on (3.10), we propose to use resampling methods. Tracing back the definition of  $S_n$  via  $\mathbb{D}_n$  to  $\mathbb{C}_n$  in (3.1) and using the approximation via  $\tilde{\mathbb{C}}_n$  in (3.7), we find that it suffices to construct a resampling scheme for  $\mathbb{B}_n$  defined in (3.4) and to estimate the first-order partial derivatives,  $\dot{C}_j$ , of  $C$ .

### 4.1 Multiplier sequences

In the case of i.i.d. observations, Scaillet (2005) proposed to use a *multiplier* approach in the spirit of van der Vaart and Wellner (2000, Chapter 2.9) to resample  $\mathbb{B}_n$ . When the first-order partial derivatives of  $C$  are estimated by finite-differencing as in Rémillard and Scaillet (2009), the resulting resampling scheme for  $\mathbb{C}_n$  is frequently referred to as a *multiplier bootstrap*. In a nonsequential setting based on independent observations, Bücher and Dette (2010) compared the finite-sample behavior of the various resampling techniques proposed in the literature and concluded that the multiplier bootstrap of Rémillard and Scaillet (2009) has, overall, the best finite-sample properties. This technique was revisited theoretically by Segers (2012) who showed its asymptotic validity under Condition 3.2. A sequential generalization of the latter result will be stated later in this section. In the case of independent observations, the multiplier bootstrap is based on *i.i.d. multiplier sequences*. We say that a sequence of random variables  $(\xi_{i,n})_{i \in \mathbb{Z}}$  is an *i.i.d. multiplier sequence* if:

(M0)  $(\xi_{i,n})_{i \in \mathbb{Z}}$  is i.i.d., independent of  $\mathbf{X}_1, \dots, \mathbf{X}_n$ , with distribution not changing with  $n$ , having mean 0, variance 1, and being such that  $\int_0^\infty \{P(|\xi_{0,n}| > x)\}^{1/2} dx < \infty$ .

Starting from the seminal work of Bühlmann (1993, Section 3.3), Bücher and Ruppert (2013) and Bücher and Kojadinovic (2013) have studied a *dependent* multiplier bootstrap for  $\mathbb{C}_n$  which extends the multiplier bootstrap of Rémillard and Scaillet (2009) to the sequential and strongly mixing setting. The key idea in Bühlmann (1993) is to replace i.i.d. multipliers by suitably serially dependent multipliers that will capture the serial dependence in the data. In the rest of the paper, we say that a sequence of random variables  $(\xi_{i,n})_{i \in \mathbb{Z}}$  is a *dependent multiplier sequence* if:

(M1) The sequence  $(\xi_{i,n})_{i \in \mathbb{Z}}$  is strictly stationary with  $E(\xi_{0,n}) = 0$ ,  $E(\xi_{0,n}^2) = 1$  and  $\sup_{n \geq 1} E(|\xi_{0,n}|^\nu) < \infty$  for all  $\nu \geq 1$ , and is independent of the available sample  $\mathbf{X}_1, \dots, \mathbf{X}_n$ .

(M2) There exists a sequence  $\ell_n \rightarrow \infty$  of strictly positive constants such that  $\ell_n = o(n)$  and the sequence  $(\xi_{i,n})_{i \in \mathbb{Z}}$  is  $\ell_n$ -dependent, i.e.,  $\xi_{i,n}$  is independent of  $\xi_{i+h,n}$  for all  $h > \ell_n$  and  $i \in \mathbb{N}$ .

(M3) There exists a function  $\varphi : \mathbb{R} \rightarrow [0, 1]$ , symmetric around 0, continuous at 0, satisfying  $\varphi(0) = 1$  and  $\varphi(x) = 0$  for all  $|x| > 1$  such that  $E(\xi_{0,n}\xi_{h,n}) = \varphi(h/\ell_n)$  for all  $h \in \mathbb{Z}$ .

Ways to generate dependent multiplier sequences are mentioned in Section 5 and Appendix C.

## 4.2 Computing p-values via resampling

Let  $M$  be a large integer and let  $(\xi_{i,n}^{(1)})_{i \in \mathbb{Z}}, \dots, (\xi_{i,n}^{(M)})_{i \in \mathbb{Z}}$  be  $M$  independent copies of the same multiplier sequence. We will define two multiplier resampling schemes for the process  $\mathbb{B}_n$  in (3.4). These will lead to two resampling schemes for the test statistic  $S_n$  in (2.4), on the basis of which approximate p-values can be computed.

Recall that  $C_{k:l}$  in (2.1) is the empirical copula of  $\mathbf{X}_k, \dots, \mathbf{X}_l$ , which is the empirical distribution of the vectors of rescaled ranks  $\hat{\mathbf{U}}_i^{k:l}$  in (2.2). For any  $m \in \{1, \dots, M\}$  and  $(s, t, \mathbf{u}) \in \Delta \times [0, 1]^d$ , let

$$\hat{\mathbb{B}}_n^{(m)}(s, t, \mathbf{u}) = \frac{1}{\sqrt{n}} \sum_{i=[ns]+1}^{[nt]} \xi_{i,n}^{(m)} \{ \mathbf{1}(\hat{\mathbf{U}}_i^{1:n} \leq \mathbf{u}) - C_{1:n}(\mathbf{u}) \}, \quad (4.1)$$

and

$$\begin{aligned} \check{\mathbb{B}}_n^{(m)}(s, t, \mathbf{u}) &= \frac{1}{\sqrt{n}} \sum_{i=[ns]+1}^{[nt]} \xi_{i,n}^{(m)} \{ \mathbf{1}(\hat{\mathbf{U}}_i^{[ns]+1:[nt]} \leq \mathbf{u}) - C_{[ns]+1:[nt]}(\mathbf{u}) \} \\ &= \frac{1}{\sqrt{n}} \sum_{i=[ns]+1}^{[nt]} (\xi_{i,n}^{(m)} - \bar{\xi}_{[ns]+1:[nt]}^{(m)}) \mathbf{1}(\hat{\mathbf{U}}_i^{[ns]+1:[nt]} \leq \mathbf{u}), \end{aligned} \quad (4.2)$$



where  $\bar{\xi}_{k:l}^{(m)}$  is the arithmetic mean of  $\xi_{i,n}^{(m)}$  for  $i \in \{k, \dots, l\}$ . By convention, the sums are zero if  $\lfloor ns \rfloor = \lfloor nt \rfloor$ . Note that the ranks are computed relative to the complete sample  $\mathbf{X}_1, \dots, \mathbf{X}_n$  for  $\hat{\mathbb{B}}_n^{(m)}(s, t, \cdot)$ , whereas they are computed relative to the subsample  $\mathbf{X}_{\lfloor ns \rfloor + 1}, \dots, \mathbf{X}_{\lfloor nt \rfloor}$  for  $\check{\mathbb{B}}_n^{(m)}(s, t, \cdot)$ .

In order to get to resampling versions of  $\tilde{\mathbb{C}}_n$  in (3.7), we need estimators of the first-order partial derivatives of  $C$ . A simple estimator based on  $\mathbf{X}_k, \dots, \mathbf{X}_l$  consists of finite differencing at a bandwidth of  $h \equiv h(k, l) = \min\{(l - k + 1)^{-1/2}, 1/2\}$ . Varying slightly upon the definition in Rémillard and Scaillet (2009) and following Kojadinovic et al. (2011a, Section 3), we put

$$\dot{C}_{j,k:l}(\mathbf{u}) = \frac{C_{k:l}(\mathbf{u} + h\mathbf{e}_j) - C_{k:l}(\mathbf{u} - h\mathbf{e}_j)}{\min(u_j + h, 1) - \max(u_j - h, 0)}$$

for  $\mathbf{u} \in [0, 1]^d$ , where  $\mathbf{e}_j$  is the  $j$ th canonical unit vector in  $\mathbb{R}^d$ . Note that if  $h \leq u_j \leq 1 - h$ , the denominator is just  $2h$ . The more general form of the denominator corrects for boundary effects ( $u_j$  close to 0 or 1). Proceeding for instance as in Kojadinovic et al. (2011a, proof of Proposition 2), we find that the previous estimator is uniformly bounded.

The resampling versions  $\hat{\mathbb{B}}_n^{(m)}$  and  $\check{\mathbb{B}}_n^{(m)}$  of  $\mathbb{B}_n$  then lead to the following resampling versions for  $\tilde{\mathbb{C}}_n$ : for  $(s, t, \mathbf{u}) \in \Delta \times [0, 1]^d$ ,

$$\begin{aligned} \hat{\mathbb{C}}_n^{(m)}(s, t, \mathbf{u}) &= \hat{\mathbb{B}}_n^{(m)}(s, t, \mathbf{u}) - \sum_{j=1}^d \dot{C}_{j,1:n}(\mathbf{u}) \hat{\mathbb{B}}_n^{(m)}(s, t, \mathbf{u}^{(j)}), \\ \check{\mathbb{C}}_n^{(m)}(s, t, \mathbf{u}) &= \check{\mathbb{B}}_n^{(m)}(s, t, \mathbf{u}) - \sum_{j=1}^d \dot{C}_{j,\lfloor ns \rfloor + 1:\lfloor nt \rfloor}(\mathbf{u}) \check{\mathbb{B}}_n^{(m)}(s, t, \mathbf{u}^{(j)}). \end{aligned} \quad (4.3)$$

Recall that  $\lambda_n(s, t) = (\lfloor nt \rfloor - \lfloor ns \rfloor)/n$ . The difference process  $\mathbb{D}_n$  is to be resampled by one of the following two methods:

$$\begin{aligned} \hat{\mathbb{D}}_n^{(m)}(s, \mathbf{u}) &= \lambda_n(s, 1) \hat{\mathbb{C}}_n^{(m)}(0, s, \mathbf{u}) - \lambda_n(0, s) \hat{\mathbb{C}}_n^{(m)}(s, 1, \mathbf{u}) \\ &= \hat{\mathbb{C}}_n^{(m)}(0, s, \mathbf{u}) - \lambda_n(0, s) \hat{\mathbb{C}}_n^{(m)}(0, 1, \mathbf{u}), \\ \check{\mathbb{D}}_n^{(m)}(s, \mathbf{u}) &= \lambda_n(s, 1) \check{\mathbb{C}}_n^{(m)}(0, s, \mathbf{u}) - \lambda_n(0, s) \check{\mathbb{C}}_n^{(m)}(s, 1, \mathbf{u}). \end{aligned}$$

For resampling the test statistic, one has the choice between

$$\hat{S}_n^{(m)} = \sup_{s \in [0, 1]} \int_{[0, 1]^d} \{\hat{\mathbb{D}}_n^{(m)}(s, \mathbf{u})\}^2 dC_{1:n}(\mathbf{u}), \quad (4.4)$$

$$\check{S}_n^{(m)} = \sup_{s \in [0, 1]} \int_{[0, 1]^d} \{\check{\mathbb{D}}_n^{(m)}(s, \mathbf{u})\}^2 dC_{1:n}(\mathbf{u}). \quad (4.5)$$

Finally, approximate p-values of the observed test statistic  $S_n$  can be computed via either

$$\frac{1}{M} \sum_{m=1}^M \mathbf{1}(\hat{S}_n^{(m)} \geq S_n) \quad \text{or} \quad \frac{1}{M} \sum_{m=1}^M \mathbf{1}(\check{S}_n^{(m)} \geq S_n). \quad (4.6)$$

The null hypothesis is rejected if the estimated p-value is smaller than the desired significance level.

By comparison, note that for the test statistic  $S_n^R$  in (2.8) based on the process  $\mathbb{D}_n^R$  in (2.7), an approximate p-value can be computed using the multiplier processes

$$\mathbb{D}_n^{R,(m)}(s, \mathbf{u}) = \hat{\mathbb{B}}_n^{(m)}(0, s, \mathbf{u}) - \lambda_n(0, s) \hat{\mathbb{B}}_n^{(m)}(0, 1, \mathbf{u}), \quad (4.7)$$

where  $\hat{\mathbb{B}}_n^{(m)}$  is defined in (4.1); see also Rémillard (2010, Section 5.2) and Bücher and Ruppert (2013, Section 3.2).

### 4.3 Asymptotic validity of the resampling scheme

We establish the asymptotic validity of the multiplier resampling schemes described above under the null hypothesis. First, we need to impose conditions on the data generating process  $\mathbf{X}_1, \dots, \mathbf{X}_n$  and the multiplier sequences  $(\xi_{i,n}^{(m)})_{i \in \mathbb{Z}}$  for  $m \in \{1, \dots, M\}$ .

**Condition 4.1.** *One of the following two conditions holds:*

- (i) *The random vectors  $\mathbf{X}_1, \dots, \mathbf{X}_n$  are i.i.d. and  $(\xi_{i,n}^{(1)})_{i \in \mathbb{Z}}, \dots, (\xi_{i,n}^{(M)})_{i \in \mathbb{Z}}$  are independent copies of a multiplier sequence satisfying (M0).*
- (ii) *The random vectors  $\mathbf{X}_1, \dots, \mathbf{X}_n$  are drawn from a strictly stationary sequence  $(\mathbf{X}_i)_{i \in \mathbb{Z}}$  whose strong mixing coefficients satisfy  $\alpha_r = O(r^{-a})$  for some  $a > 3 + 3d/2$ , and  $(\xi_{i,n}^{(1)})_{i \in \mathbb{Z}}, \dots, (\xi_{i,n}^{(M)})_{i \in \mathbb{Z}}$  are independent copies of a dependent multiplier sequence satisfying (M1)–(M3) with  $\ell_n = O(n^{1/2-\gamma})$  for some  $0 < \gamma < 1/2$ .*

*In both cases, the stationary distribution of  $\mathbf{X}_i$  has continuous margins and a copula  $C$  satisfying Condition 3.2.*

If the random vectors  $\mathbf{X}_1, \dots, \mathbf{X}_n$  are i.i.d., they can also be considered to be drawn from a strongly mixing, strictly stationary sequence. Hence, for the multiplier sequences  $(\xi_{i,n}^{(m)})_{i \in \mathbb{Z}}$ , one could either assume (M0) or (M1)–(M3): both should work. However, as discussed in Bücher and Kojadinovic (2013, Section 2), the use of dependent multipliers in the case of independent observations is likely to result in an efficiency loss. This is illustrated in the Monte Carlo simulations reported in Bücher and Ruppert (2013, Section 3) and carried out for the test based on the statistic  $S_n^R$  defined in (2.8) which is resampled using multiplier processes asymptotically equivalent to those given in (4.7): the use of dependent multipliers in the case of serially independent data usually results in a loss of power and in a slightly more conservative test. Thus, in finite samples, if there is no evidence against serial independence, it appears more sensible to work under (M0).

We can now state the asymptotic distributions of the multiplier resampling schemes under the null hypothesis of a constant distribution. We provide two propositions, one for the resampling scheme based on  $\hat{\mathbb{B}}_n^{(m)}$  in (4.1) and another one for the scheme based on  $\check{\mathbb{B}}_n^{(m)}$  in (4.2).

**Proposition 4.2.** *If Condition 4.1 holds, then*

$$\left(\mathbb{C}_n, \hat{\mathbb{C}}_n^{(1)}, \dots, \hat{\mathbb{C}}_n^{(M)}\right) \rightsquigarrow \left(\mathbb{C}_C, \mathbb{C}_C^{(1)}, \dots, \mathbb{C}_C^{(M)}\right)$$

in  $\{\ell^\infty(\Delta \times [0, 1]^d)\}^{M+1}$ , where  $\mathbb{C}_C$  is defined in (3.8), and  $\mathbb{C}_C^{(1)}, \dots, \mathbb{C}_C^{(M)}$  are independent copies of  $\mathbb{C}_C$ . As a consequence, also

$$\left(\mathbb{D}_n, \hat{\mathbb{D}}_n^{(1)}, \dots, \hat{\mathbb{D}}_n^{(M)}\right) \rightsquigarrow \left(\mathbb{D}_C, \mathbb{D}_C^{(1)}, \dots, \mathbb{D}_C^{(M)}\right),$$

in  $\{\ell^\infty([0, 1]^{(d+1)})\}^{M+1}$ , where  $\mathbb{D}_C$  is defined in (3.9) and  $\mathbb{D}_C^{(1)}, \dots, \mathbb{D}_C^{(M)}$  are independent copies of  $\mathbb{D}_C$ . Finally,

$$\left(S_n, \hat{S}_n^{(1)}, \dots, \hat{S}_n^{(M)}\right) \rightsquigarrow (S, S^{(1)}, \dots, S^{(M)})$$

where  $S$  is defined in (3.10) and  $S^{(1)}, \dots, S^{(M)}$  are independent copies of  $S$ .

Under Condition 4.1(i), the above result can be easily proved by starting from Theorem 1 of Holmes et al. (2013) and adapting the arguments used in Segers (2012, proof of Proposition 3.2). Under Conditions 4.1(ii) and 3.1, the result was obtained in Bücher and Kojadinovic (2013, Proposition 4.2). The additional arguments allowing to avoid Condition 3.1 will be given in the proof of the next result.

**Proposition 4.3.** *If Condition 4.1 holds, then the conclusions of Proposition 4.2 also hold with  $\hat{\mathbb{C}}_n^{(m)}$  replaced by  $\check{\mathbb{C}}_n^{(m)}$ ,  $\hat{\mathbb{D}}_n^{(m)}$  replaced by  $\check{\mathbb{D}}_n^{(m)}$ , and  $\hat{S}_n^{(m)}$  replaced by  $\check{S}_n^{(m)}$ .*

The proof of Proposition 4.3 is somewhat involved and is given in detail in Appendix B.

Combining the last claims of Propositions 4.2 and 4.3 with Proposition F.1 in Bücher and Kojadinovic (2013), we obtain that a test based on  $S_n$  whose p-value is computed using one of the two approaches in (4.6) will hold its level asymptotically as  $n \rightarrow \infty$  followed by  $M \rightarrow \infty$ .

## 5 Simulation study

Large-scale Monte Carlo experiments were carried out in order to study the finite-sample performance of the derived tests for detecting changes in cross-sectional dependence. The main questions addressed by the study are the following:

- (i) How well do the tests hold their size under the null hypothesis  $H_0$  in (1.1) of no change?
- (ii) What is the power of the tests against the alternative  $H_{1,c}$  of a single change in cross-sectional dependence at constant margins? Specifically, the alternative hypothesis is  $H_{1,c} \cap H_{0,m}$  with  $H_{0,m}$  in (1.2) and  $H_{1,c}$  defined by

$$H_{1,c} : \exists \text{ distinct } C_1 \text{ and } C_2, \text{ and } k^* \in \{1, \dots, n-1\} \text{ such that} \\ \mathbf{X}_1, \dots, \mathbf{X}_{k^*} \text{ have copula } C_1 \text{ and } \mathbf{X}_{k^*+1}, \dots, \mathbf{X}_n \text{ have copula } C_2. \quad (5.1)$$

- (iii) What happens if the change in distribution is only due to a change in the margins, the copula remaining constant? Specifically, the alternative hypothesis is  $H_{1,m} \cap H_{0,c}$  with  $H_{0,c}$  given in (1.3) and  $H_{1,m}$  defined by

$$\begin{aligned}
H_{1,m} : & \exists \text{ distinct } F_{1,1}, F_{1,2} \text{ as well as } F_2, \dots, F_d \text{ and } k_1^* \in \{1, \dots, n-1\} \\
& \text{such that } \mathbf{X}_1, \dots, \mathbf{X}_{k_1^*} \text{ have marginal c.d.f.s } F_{1,1}, F_2, \dots, F_d \\
& \text{and } \mathbf{X}_{k_1^*+1}, \dots, \mathbf{X}_n \text{ have marginal c.d.f.s } F_{1,2}, F_2, \dots, F_d.
\end{aligned} \tag{5.2}$$

In addition to the three questions above, many others can be formulated, involving other alternative hypotheses for instance. The problem is complex and there are countless ways of combining factors in the experimental design. In our study, the settings were chosen to represent a wide and hopefully representative variety of situations, in function of the three questions above. The main factors of our experiments are summarized below:

- Test statistics:
  - Our statistic  $S_n$  in (2.4) with p-values computed via resampling using  $\hat{S}_n^{(m)}$  or  $\check{S}_n^{(m)}$  in (4.4) and (4.5), respectively. As we continue, we shall simply talk about the test based on  $\hat{S}_n$  or  $\check{S}_n$ , respectively, to distinguish between these two situations.
  - The statistic  $S_n^R$  in (2.8) of Bücher and Ruppert (2013), with p-values computed according to the resampling method for  $\mathbb{D}_n^R$  in (4.7).
- Sample size:  $n \in \{50, 100, 200\}$ .
- Number of samples per setting: 1 000.
- Cross-sectional dimension:  $d \in \{2, 3\}$ .
- Significance level:  $\alpha = 5\%$ .
- Serial dependence: The data were generated either as being serially independent or via two time-series models, an autoregressive process and a multivariate version of the exponential autoregressive model considered in Auestad and Tjøstheim (1990) and Paparoditis and Politis (2001, Section 3.3). Independent standard normals were used as multipliers for independent observations, while for the serially dependent datasets, the dependent multiplier sequences were generated from initial independent standard normal sequences using the “moving average approach” proposed initially in Bühlmann (1993) and revisited in some detail in Bücher and Kojadinovic (2013, Section 6.1). The value of the bandwidth parameter  $\ell_n$  defined in Condition (M2) was chosen automatically using the approach described in Bücher and Kojadinovic (2013, Section 5). See Appendix D for details.
- Margins: in all but one setting, the margins were kept constant, i.e.,  $H_{0,m}$  in (1.2) was assumed. In one case (see Table 5), a break as in  $H_{1,m}$  in (5.2) was assumed, the marginal distribution of the first component changing from the  $N(0, 1)$  to the  $N(\mu, 1)$  distribution.
- Copulas: Clayton, Gumbel–Hougaard, Normal, Frank, with positive or negative (insofar possible) association, as well as asymmetric versions obtained via Khoudraji’s device (Khoudraji, 1995; Genest et al., 1998; Liebscher, 2008).
- Alternative hypotheses involving a single change-point occurring at time  $k^* = \lfloor nt \rfloor$  with  $t \in \{0.1, 0.25, 0.5, 0.75\}$ :
  - $H_{0,m} \cap H_{1,c}$  with a change of the parameter within a copula family.
  - $H_{0,m} \cap H_{1,c}$  with a change of the copula family at constant Kendall’s tau.

- $H_{1,m} \cap H_{0,c}$ , i.e., a change of one of the *margins* rather than of the copula.
- For the serially dependent case, a change in the copula of the innovations, leading to a gradual change of the copula of the marginal distributions of the observables.

The experiments were carried out in the R statistical system (R Development Core Team, 2013) using the `copula` package (Hofert et al., 2013). To allow us to reuse previously written code, the rescaled ranks in (2.2) were computed by dividing the ranks by  $l - k + 2$  instead of  $l - k + 1$ . Because (4.4) only involves rescaled ranks computed from the entire sample, the test based on  $\hat{S}_n$  can be implemented to be substantially faster than the one based on  $\check{S}_n$  for larger sample sizes. The corresponding routines are available in the R package `npcp` (Kojadinovic, 2014).

For the sake of brevity, only a representative subset of the results is reported here. Specifically, the following tables are provided in Appendix D:

- Size of the tests under the null hypothesis  $H_0$ :
  - Table 1: Percentage of false rejections when data are serially independent.
  - Table 2: Percentage of false rejections when data are serially dependent.
- Power of the tests against specific alternatives:
  - Table 3: Power against  $H_{0,m} \cap H_{1,c}$  involving a change of the copula parameter within a copula family and at serial independence.
  - Table 4: Power against  $H_{0,m} \cap H_{1,c}$  involving a change of copula family at a constant value of Kendall’s tau and at serial independence.
  - Table 5: Power against  $H_{1,m} \cap H_{0,c}$  involving a change in one of the margins and at serial independence..
  - Table 6: Power against  $\neg H_0$  when data are serially dependent and the change occurs in the copula of the innovations.

Besides findings of a more anecdotal nature, the following conclusions may be drawn from the results:

- All tests hold their level reasonably well in the case of serial independence (Table 1), with minor fluctuations depending on sample size, test statistic, copula parameter and copula family.
- In case of serial dependence, the test based on  $\hat{S}_n$  is too conservative for the sample sizes under consideration (Table 2). In line with this observation, the test based on  $\check{S}_n$  appears to be more powerful than the one based on  $\hat{S}_n$  (Table 6).
- For alternative hypotheses involving a change in the *copula*, the tests based on  $\hat{S}_n$  and  $\check{S}_n$  have a higher power than  $S_n^R$  (Tables 3 and 4). When the copula changes in such a way that Kendall’s tau remains constant, the power of  $S_n^R$  is especially low. With respect to that last setting, note that distinguishing copulas on the basis of low amounts of data is known to be difficult (Genest et al., 2009; Kojadinovic et al., 2011b). The fact that the change-point is unknown makes the problem even harder.

- For alternative hypotheses involving a change in one of the *margins*, it is the test statistic  $S_n^R$  that is substantially more powerful than  $S_n$  (Table 5). The weak power of  $S_n$  can be explained by the fact that it is designed for detecting changes in the copula. Another tentative reading of the results is that the test based on  $S_n$ , regarded as a procedure for testing  $H_{0,c}$ , is relatively robust against small changes in one margin. In contrast, the test based on  $S_n^R$  behaves as an all-purpose test for the hypothesis of a constant distribution rather than as a test for a constant copula.

## 6 Case studies

As an illustration, we first applied the test based on  $\check{S}_n$  to bivariate financial data consisting of daily logreturns computed from the DAX and the Standard and Poor 500 indices. Following Dehling et al. (2013, Section 7), attention was restricted to the years 2006–2009. The corresponding closing quotes were obtained from <http://quote.yahoo.com> using the `get.hist.quote` function of the `tseries` R package (Trapletti and Hornik, 2013), which resulted in  $n = 993$  bivariate logreturns. Dependent multiplier sequences were generated as explained in Appendix C. An approximate p-value of 0.04 was obtained, providing some evidence against  $H_0$ . The conclusion is in line with the results reported in Dehling et al. (2013). Of course, as discussed earlier, it is only under the assumption that  $H_{0,m}$  in (1.2) holds that it would be fully justified to decide to reject  $H_{0,c}$  in (5.1) on the basis of the previous approximate p-value. The value of the change-point estimator  $k_n^*$  in (2.5) is 529, corresponding to February 22nd, 2008.

As a second illustration, we followed again Dehling et al. (2013) and considered  $n = 504$  bivariate logreturns computed from closing daily quotes of the Dow Jones Industrial Average and the Nasdaq Composite for the years 1987 and 1988. The former quotes, not being available on <http://quote.yahoo.com> anymore, were taken from the R package `QRM` (Pfaff and McNeil, 2013). This two-year period is of interest because it contains October 19th, 1987, known as “black Monday” (see Dehling et al., 2013, Figure 4). An approximate p-value of 0.59 was obtained. Hence, despite the extreme events that occurred during the period under consideration, the test based on  $\check{S}_n$  detects no evidence against  $H_0$  in the data, which is in line with the results reported in Dehling et al. (2013).

## 7 Conclusion

We have demonstrated that the sensitivity of rank-based tests for the null hypothesis of a constant distribution against changes in cross-sectional dependence can be improved if ranks are computed with respect to relevant subsamples. In this way, the test we propose achieves in many cases a higher power than the one proposed in Bücher and Ruppert (2013). The limit distribution of the test statistic under the null hypothesis is unwieldy, but approximate p-values can still be computed via a multiplier resampling scheme. To deal with potential serial dependence, we make use of dependent multiplier sequences, an idea going back to Bühlmann (1993) and revisited in Bücher and Kojadinovic (2013).

Here are some potential avenues for further research:

- Once the null hypothesis has been rejected, the nature of the nonstationary needs to be investigated further: is there a single change-point or is there more than one? Or maybe the change is gradual rather than sudden? And does the change concern the margins or the copula?
- Can one detect a change in the copula without the hypothesis that the margins are constant?
- The procedure is computationally intensive because the ranks have to be recomputed for every  $k \in \{1, \dots, n - 1\}$ . Efficient algorithms for reutilizing calculations from one value of  $k$  to the next one might speed up the computations.

## Acknowledgments

The authors are grateful to Mark Holmes, Jean-François Quessy and Martin Ruppert for fruitful discussions, and to an anonymous referee for pointing out that Condition 3.1 might be dispensed with.

The research by A. Bücher has been supported in parts by the Collaborative Research Center “Statistical modeling of nonlinear dynamic processes” (SFB 823, Project A7) of the German Research Foundation (DFG), which is gratefully acknowledged.

J. Segers gratefully acknowledges funding by contract “Projet d’Actions de Recherche Concertées” No. 12/17-045 of the “Communauté française de Belgique” and by IAP research network Grant P7/06 of the Belgian government (Belgian Science Policy).

## References

- B. Auestad and D. Tjøstheim. Identification of nonlinear time series: First order characterization and order determination. *Biometrika*, 77:669–687, 1990.
- J. Bai. Weak convergence of the sequential empirical processes of residuals in ARMA models. *The Annals of Statistics*, 22(4):2051–2061, 1994.
- A. Bücher. A note on nonparametric estimation of bivariate tail dependence. *Statistics and Risk Modeling*, page in press, 2013a. URL [http://www.ruhr-uni-bochum.de/imperia/md/content/mathematik3/publications/taildep\\_b](http://www.ruhr-uni-bochum.de/imperia/md/content/mathematik3/publications/taildep_b)
- A. Bücher. A note on weak convergence of the sequential multivariate empirical process under strong mixing. *Journal of Theoretical Probability*, page in press, 2013b.
- A. Bücher and H. Dette. A note on bootstrap approximations for the empirical copula process. *Statistics and Probability Letters*, 80(23–24):1925–1932, 2010.
- A. Bücher and I. Kojadinovic. A dependent multiplier bootstrap for the sequential empirical copula process under strong mixing. *arXiv:1306.3930*, 2013.

- A. Bücher and M. Ruppert. Consistent testing for a constant copula under strong mixing based on the tapered block multiplier technique. *Journal of Multivariate Analysis*, 116: 208–229, 2013.
- A. Bücher and J. Segers. Extreme value copula estimation based on block maxima of a multivariate stationary time series. *arXiv:1311.3060*, 2013.
- A. Bücher and S. Volgushev. Empirical and sequential empirical copula processes under serial dependence. *Journal of Multivariate Analysis*, 119:61–70, 2013.
- P. Bühlmann. *The blockwise bootstrap in time series and empirical processes*. PhD thesis, ETH Zürich, 1993. Diss. ETH No. 10354.
- M. Csörgő and L. Horváth. *Limit theorems in change-point analysis*. Wiley Series in Probability and Statistics. John Wiley & Sons, Chichester, UK, 1997.
- H. Dehling, D. Vogel, M. Wendler, and D. Wied. An efficient and robust test for a change-point in correlation. *arXiv:1203.4871*, 2013.
- C. Genest and J. Segers. On the covariance of the asymptotic empirical copula process. *Journal of Multivariate Analysis*, 101:1837–1845, 2010.
- C. Genest, K. Ghoudi, and L.-P. Rivest. Discussion of “Understanding relationships using copulas”, by E. Frees and E. Valdez. *North American Actuarial Journal*, 3:143–149, 1998.
- C. Genest, B. Rémillard, and D. Beaudoin. Goodness-of-fit tests for copulas: A review and a power study. *Insurance: Mathematics and Economics*, 44:199–213, 2009.
- E. Gombay and L. Horváth. Change-points and bootstrap. *Environmetrics*, 10(6), 1999.
- E. Gombay and L. Horváth. Rates of convergence for  $U$ -statistic processes and their bootstrapped versions. *Journal of Statistical Planning and Inference*, 102:247–272, 2002.
- M. Hofert, I. Kojadinovic, M. Mächler, and J. Yan. *copula: Multivariate dependence with copulas*, 2013. URL <http://CRAN.R-project.org/package=copula>. R package version 0.999-7.
- M. Holmes, I. Kojadinovic, and J-F. Quessy. Nonparametric tests for change-point detection à la Gombay and Horváth. *Journal of Multivariate Analysis*, 115:16–32, 2013.
- A. Inoue. Testing for distributional change in time series. *Econometric Theory*, 17(1): 156–187, 2001.
- A. Khoudraji. *Contributions à l’étude des copules et à la modélisation des valeurs extrêmes bivariées*. PhD thesis, Université Laval, Québec, Canada, 1995.
- I. Kojadinovic. *npcp: Some nonparametric tests for change-point detection in (multivariate) observations*, 2014. URL <http://CRAN.R-project.org/package=npcp>. R package version 0.0-1.



- I. Kojadinovic, J. Segers, and J. Yan. Large-sample tests of extreme-value dependence for multivariate copulas. *The Canadian Journal of Statistics*, 39(4):703–720, 2011a.
- I. Kojadinovic, J. Yan, and M. Holmes. Fast large-sample goodness-of-fit for copulas. *Statistica Sinica*, 21(2):841–871, 2011b.
- E. Liebscher. Construction of asymmetric multivariate copulas. *Journal of Multivariate Analysis*, 99:2234–2250, 2008.
- E. Paparoditis and D.N. Politis. Tapered block bootstrap. *Biometrika*, 88(4):1105–1119, 2001.
- B. Pfaff and A. McNeil. *QRM: Provides R-language Code to Examine Quantitative Risk Management Concepts*, 2013. URL <http://CRAN.R-project.org/package=QRM>. R package version 0.4-9.
- J.-F. Quessy, M. Saïd, and A.-C. Favre. Multivariate Kendall’s tau for change-point detection in copulas. *The Canadian Journal of Statistics*, 41:65–82, 2013.
- R Development Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, 2013. URL <http://www.R-project.org>. ISBN 3-900051-07-0.
- B. Rémillard. Goodness-of-fit tests for copulas of multivariate time series. *Social Science Research Network*, 1729982:1–32, 2010. URL <http://ssrn.com/abstract=1729982>.
- B. Rémillard and O. Scaillet. Testing for equality between two copulas. *Journal of Multivariate Analysis*, 100(3):377–386, 2009.
- J.P. Romano and M. Wolf. A more general central limit theorem for  $m$ -dependent random variables with unbounded  $m$ . *Statistics and Probability Letters*, 47:115–124, 2000.
- O. Scaillet. A Kolmogorov-Smirnov type test for positive quadrant dependence. *Canadian Journal of Statistics*, 33:415–427, 2005.
- J. Segers. Asymptotics of empirical copula processes under nonrestrictive smoothness assumptions. *Bernoulli*, 18:764–782, 2012.
- A. Sklar. Fonctions de répartition à  $n$  dimensions et leurs marges. *Publications de l’Institut de Statistique de l’Université de Paris*, 8:229–231, 1959.
- A. Trapletti and K. Hornik. *tseries: Time series analysis and computational finance*, 2013. URL <http://CRAN.R-project.org/package=tseries>. R package version 0.10-32.
- A.W. van der Vaart and J.A. Wellner. *Weak convergence and empirical processes*. Springer, New York, 2000. Second edition.
- M. van Kampen and D. Wied. A nonparametric constancy test for copulas under mixing conditions. *working paper*, 2013. URL <http://www.statistik.tu-dortmund.de/vankampen-eng.html>.

## A Proof of Proposition 3.3

Let us first introduce additional notation. For integers  $1 \leq k \leq l \leq n$ , let  $H_{k:l}$  denote the empirical c.d.f. of the unobservable sample  $\mathbf{U}_k, \dots, \mathbf{U}_l$  and let  $H_{k:l,j}$ , for  $j \in \{1, \dots, d\}$ , denote its margins. The empirical quantile functions are

$$H_{k:l,j}^{-1}(u) = \inf\{v \in [0, 1] : H_{k:l,j}(v) \geq u\}, \quad u \in [0, 1],$$

which are collected in a vector via

$$\mathbf{H}_{k:l}^{-1}(\mathbf{u}) = (H_{k:l,1}^{-1}(u_1), \dots, H_{k:l,d}^{-1}(u_d)), \quad \mathbf{u} \in [0, 1]^d.$$

By convention, the previously defined quantities are all taken equal to zero if  $k > l$ .

From the proof of Theorem 3.4 in Bücher and Kojadinovic (2013), we have that (3.6) holds with  $\mathbb{C}_n$  replaced by  $\mathbb{C}_n^{\text{alt}}$ , where

$$\mathbb{C}_n^{\text{alt}}(s, t, \mathbf{u}) = \frac{1}{\sqrt{n}} \sum_{i=\lfloor ns \rfloor + 1}^{\lfloor nt \rfloor} [\mathbf{1}\{\mathbf{U}_i \leq \mathbf{H}_{\lfloor ns \rfloor + 1 : \lfloor nt \rfloor}^{-1}(\mathbf{u})\} - C(\mathbf{u})], \quad (s, t, \mathbf{u}) \in \Delta \times [0, 1]^d.$$

To show (3.6), it remains therefore to prove that

$$\sup_{(s,t,\mathbf{u}) \in \Delta \times [0,1]^d} |\mathbb{C}_n^{\text{alt}}(s, t, \mathbf{u}) - \mathbb{C}_n(s, t, \mathbf{u})| \xrightarrow{\mathbb{P}} 0. \quad (\text{A.1})$$

To do so, we adapt the arguments used in Lemma A.2 of Bücher and Segers (2013). Fix  $1 \leq k \leq l \leq n$  and  $\mathbf{u} \in [0, 1]^d$ . For  $i \in \{k, \dots, l\}$ , the  $d$  components of  $\hat{\mathbf{U}}_i^{k:l}$  defined in (2.2) can be expressed as  $\hat{U}_{ij}^{k:l} = H_{k:l,j}(U_{ij})$ ,  $j \in \{1, \dots, d\}$ . Next, notice that

$$\begin{aligned} \mathbf{1}\{U_{ij} \leq H_{k:l,j}^{-1}(u_j)\} - \mathbf{1}\{\hat{U}_{ij}^{k:l} \leq u_j\} &= \mathbf{1}\{U_{ij} < H_{k:l,j}^{-1}(u_j)\} + \mathbf{1}\{U_{ij} = H_{k:l,j}^{-1}(u_j)\} \\ &\quad - \mathbf{1}\{\hat{U}_{ij}^{k:l} < u_j\} - \mathbf{1}\{\hat{U}_{ij}^{k:l} = u_j\} \\ &= \mathbf{1}\{U_{ij} = H_{k:l,j}^{-1}(u_j)\} - \mathbf{1}\{\hat{U}_{ij}^{k:l} = u_j\}, \end{aligned}$$

as  $x < H^{-1}(u)$  if and only  $H(x) < u$  for any distribution function  $H$ . Since  $\hat{U}_{ij}^{k:l} = H_{k:l,j}(U_{ij}) = u_j$  implies  $U_{ij} = H_{k:l,j}^{-1}(u_j)$ , we obtain that

$$0 \leq \mathbf{1}\{U_{ij} \leq H_{k:l,j}^{-1}(u_j)\} - \mathbf{1}\{\hat{U}_{ij}^{k:l} \leq u_j\} \leq \mathbf{1}\{U_{ij} = H_{k:l,j}^{-1}(u_j)\}.$$

Combining the previous inequality with the decomposition

$$\begin{aligned} &\mathbf{1}\{\mathbf{U}_i \leq \mathbf{H}_{k:l}^{-1}(\mathbf{u})\} - \mathbf{1}\{\hat{\mathbf{U}}_i^{k:l} \leq \mathbf{u}\} \\ &= \sum_{p=1}^d \left[ \prod_{1 \leq j \leq p} \mathbf{1}\{U_{ij} \leq H_{k:l,j}^{-1}(u_j)\} \prod_{p < j \leq d} \mathbf{1}\{\hat{U}_{ij}^{k:l} \leq u_j\} \right. \\ &\quad \left. - \prod_{1 \leq j \leq p-1} \mathbf{1}\{U_{ij} \leq H_{k:l,j}^{-1}(u_j)\} \prod_{p-1 < j \leq d} \mathbf{1}\{\hat{U}_{ij}^{k:l} \leq u_j\} \right], \end{aligned}$$

we obtain that

$$0 \leq \mathbf{1}\{\mathbf{U}_i \leq \mathbf{H}_{k:l}^{-1}(\mathbf{u})\} - \mathbf{1}(\hat{\mathbf{U}}_i^{k:l} \leq \mathbf{u}) \leq \sum_{j=1}^d \mathbf{1}\{U_{ij} = H_{k:l,j}^{-1}(u_j)\}. \quad (\text{A.2})$$

It follows that the supremum in (A.1) is smaller than

$$\sum_{j=1}^d \sup_{(s,t,\mathbf{u}) \in \Delta \times [0,1]} \frac{1}{\sqrt{n}} \sum_{i=[ns]+1}^{\lfloor nt \rfloor} \mathbf{1}\{U_{ij} = H_{[ns]+1:\lfloor nt \rfloor, j}^{-1}(u)\} \leq \sum_{j=1}^d \sup_{u \in [0,1]} \frac{1}{\sqrt{n}} \sum_{i=1}^n \mathbf{1}(U_{ij} = u).$$

Using the fact that  $\mathbf{1}(U_{ij} = u) \leq \mathbf{1}(U_{ij} \leq u) - \mathbf{1}(U_{ij} \leq u - 1/n)$ , the latter is smaller

$$d \sup_{\substack{\mathbf{u}, \mathbf{v} \in [0,1]^d \\ \|\mathbf{u} - \mathbf{v}\|_1 \leq n^{-1}}} |\mathbb{B}_n(0, 1, \mathbf{u}) - \mathbb{B}_n(0, 1, \mathbf{v})| + dn^{-1/2},$$

where  $\mathbb{B}_n$  is defined in (3.4). Using the asymptotic uniform equicontinuity in probability of  $\mathbb{B}_n$  established in Lemma 2 of Bücher (2013b), we finally obtain (A.1), which completes the proof.  $\blacksquare$

## B Proof of Proposition 4.3

We shall only prove the result in the case of strongly mixing observations, that is, when Condition 4.1(ii) is assumed. The proof is similar but simpler when Condition 4.1(i) is assumed instead.

It is sufficient to show the statement involving  $\check{\mathbb{C}}_n^{(m)}$ . The statements for  $\check{\mathbb{D}}_n^{(m)}$  and  $\check{\mathbb{S}}_n^{(m)}$  then follow from the continuous mapping theorem.

For any  $m \in \{1, \dots, M\}$  and  $(s, t, \mathbf{u}) \in \Delta \times [0, 1]^d$ , put

$$\mathbb{B}_n^{(m)}(s, t, \mathbf{u}) = \frac{1}{\sqrt{n}} \sum_{i=[ns]+1}^{\lfloor nt \rfloor} \xi_{i,n}^{(m)} \{\mathbf{1}(\mathbf{U}_i \leq \mathbf{u}) - C(\mathbf{u})\}, \quad (\text{B.1})$$

$$\mathbb{C}_n^{(m)}(s, t, \mathbf{u}) = \mathbb{B}_n^{(m)}(s, t, \mathbf{u}) - \sum_{j=1}^d \dot{C}_j(\mathbf{u}) \mathbb{B}_n^{(m)}(s, t, \mathbf{u}^{(j)}).$$

[Recall that  $\mathbf{u}^{(j)} = (1, \dots, 1, u_j, 1, \dots, 1) \in [0, 1]^d$ , with  $u_j$  appearing at the  $j$ -th coordinate.] From Theorem 2.1 in Bücher and Kojadinovic (2013), we have that

$$(\mathbb{B}_n, \mathbb{B}_n^{(1)}, \dots, \mathbb{B}_n^{(M)}) \rightsquigarrow (\mathbb{B}_C, \mathbb{B}_C^{(1)}, \dots, \mathbb{B}_C^{(M)})$$

in  $\{\ell^\infty(\Delta \times [0, 1]^d)\}^{M+1}$ , where  $\mathbb{B}_C^{(1)}, \dots, \mathbb{B}_C^{(M)}$  are independent copies of  $\mathbb{B}_C$  in (3.5), and thus, from the continuous mapping theorem and (3.6), we find that

$$(\mathbb{C}_n, \mathbb{C}_n^{(1)}, \dots, \mathbb{C}_n^{(M)}) \rightsquigarrow (\mathbb{C}_C, \mathbb{C}_C^{(1)}, \dots, \mathbb{C}_C^{(M)})$$

in  $\{\ell^\infty(\Delta \times [0, 1]^d)\}^{M+1}$ . It is therefore sufficient to show that

$$\sup_{(s,t,\mathbf{u}) \in \Delta \times [0,1]^d} |(\check{\mathbb{C}}_n^{(m)} - \mathbb{C}_n^{(m)})(s, t, \mathbf{u})| \xrightarrow{P} 0 \quad (\text{B.2})$$

for every  $m \in \{1, \dots, M\}$ . Below, we will show the following two assertions: first,

$$\sup_{(s,t,\mathbf{u}) \in \Delta \times [0,1]^d} |(\check{\mathbb{B}}_n^{(m)} - \mathbb{B}_n^{(m)})(s, t, \mathbf{u})| \xrightarrow{P} 0, \quad (\text{B.3})$$

and second, for every  $\delta \in (0, 1/2)$  and every  $\varepsilon \in (0, 1)$ ,

$$\sup_{\substack{\mathbf{u} \in [0,1]^d \\ \delta \leq u_j \leq 1-\delta}} \sup_{\substack{(s,t) \in [0,1]^2 \\ t-s \geq \varepsilon}} |\dot{C}_{j, [ns]+1: [nt]}(\mathbf{u}) - \dot{C}_j(\mathbf{u})| \xrightarrow{P} 0. \quad (\text{B.4})$$

In view of the structure of  $\check{\mathbb{C}}_n^{(m)}$  in (4.3), the assertions (B.3) and (B.4) imply (B.2), as we show next. Clearly,

$$\begin{aligned} & |(\check{\mathbb{C}}_n^{(m)} - \mathbb{C}_n^{(m)})(s, t, \mathbf{u})| \\ & \leq |(\check{\mathbb{B}}_n^{(m)} - \mathbb{B}_n^{(m)})(s, t, \mathbf{u})| + \sum_{j=1}^d |\dot{C}_{j, [ns]+1: [nt]}(\mathbf{u})| |(\check{\mathbb{B}}_n^{(m)} - \mathbb{B}_n^{(m)})(s, t, \mathbf{u}^{(j)})| \\ & \quad + \sum_{j=1}^d |\dot{C}_{j, [ns]+1: [nt]}(\mathbf{u}) - \dot{C}_j(\mathbf{u})| |\mathbb{B}_n^{(m)}(s, t, \mathbf{u}^{(j)})|. \end{aligned} \quad (\text{B.5})$$

Taking suprema over  $(s, t, \mathbf{u}) \in \Delta \times [0, 1]^d$ , the first and the second term on the right-hand side of (B.5) converge to zero in probability because of assertion (B.3) and uniform boundedness of  $\dot{C}_{j,k:l}$  (see Kojadinovic et al., 2011a, proof of Proposition 2). The third term on the right-hand side of (B.5) converges to zero in probability because of assertion (B.4) and the fact that  $(s, t, \mathbf{u}) \mapsto \mathbb{B}_n^{(m)}(s, t, \mathbf{u}^{(j)})$  vanishes as soon as  $s = t$  or  $u_j \in \{0, 1\}$ , and is asymptotically uniformly equicontinuous in probability as a consequence of Lemma A.3 in Bücher and Kojadinovic (2013).

It remains to show (B.3) and (B.4). The proof of the latter assertion is simplest and is given first.

*Proof of (B.4).* Observe that

$$C_{[ns]+1: [nt]}(\mathbf{u}) = C(\mathbf{u}) + \frac{1}{\sqrt{n} \lambda_n(s, t)} \mathbb{C}_n(s, t, \mathbf{u}).$$

Fix  $\delta \in (0, 1/2)$  and  $\varepsilon \in (0, 1)$ . Without loss of generality, assume that  $n$  is large enough so that the bandwidth  $h = h_n(s, t) = 1/\sqrt{[nt] - [ns]}$  is less than  $\delta$  whenever  $t - s \geq \varepsilon$ . Then, for  $t - s \geq \varepsilon$  and  $\mathbf{u} \in [0, 1]^d$  with  $\delta \leq u_j \leq 1 - \delta$ , we have

$$\begin{aligned} \dot{C}_{j, [ns]+1: [nt]}(\mathbf{u}) &= \frac{1}{2h} \{C(\mathbf{u} + h\mathbf{e}_j) - C(\mathbf{u} - h\mathbf{e}_j)\} \\ & \quad + \frac{1}{2h \sqrt{n} \lambda_n(s, t)} \{\mathbb{C}_n(s, t, \mathbf{u} + h\mathbf{e}_j) - \mathbb{C}_n(s, t, \mathbf{u} - h\mathbf{e}_j)\}. \end{aligned}$$

By the assumption of existence and continuity of  $\dot{C}_j$  on  $V_j$  (see Condition 3.2), and since  $0 \leq \dot{C}_j \leq 1$ , it follows from the mean-value theorem that

$$\sup_{\substack{\mathbf{u} \in [0,1]^d \\ \delta \leq u_j \leq 1-\delta}} \left| \frac{1}{2h} \{C(\mathbf{u} + h\mathbf{e}_j) - C(\mathbf{u} - h\mathbf{e}_j)\} - \dot{C}_j(\mathbf{u}) \right| \rightarrow 0, \quad h \rightarrow 0.$$

Using (3.6) and the fact that  $\mathbb{B}_n$  is asymptotically uniformly equicontinuous in probability, it can be verified that  $\mathbb{C}_n$  is asymptotically uniformly equicontinuous in probability as well. It follows that

$$\sup_{\substack{\mathbf{u} \in [0,1]^d \\ \delta \leq u_j \leq 1-\delta}} \sup_{\substack{(s,t) \in [0,1]^2 \\ t-s \geq \varepsilon}} \left| \mathbb{C}_n(s, t, \mathbf{u} + h\mathbf{e}_j) - \mathbb{C}_n(s, t, \mathbf{u} - h\mathbf{e}_j) \right| \xrightarrow{P} 0.$$

Finally,

$$\frac{1}{2h\sqrt{n}\lambda_n(s, t)} = \frac{1}{2\sqrt{\lambda_n(s, t)}} \leq \frac{1}{2\sqrt{\varepsilon - 1/n}}.$$

Combine the four previous displays to arrive at the desired conclusion.  $\blacksquare$

The proof of (B.3) is more complicated. Using the notation introduced in Appendix A, let us define the auto-centered version of the process  $\mathbb{B}_n^{(m)}$  in (B.1) as

$$\begin{aligned} \mathring{\mathbb{B}}_n^{(m)}(s, t, \mathbf{u}) &= \frac{1}{\sqrt{n}} \sum_{i=[ns]+1}^{[nt]} \xi_{i,n}^{(m)} \{ \mathbf{1}(\mathbf{U}_i \leq \mathbf{u}) - H_{[ns]+1:[nt]}(\mathbf{u}) \} \\ &= \frac{1}{\sqrt{n}} \sum_{i=[ns]+1}^{[nt]} (\xi_{i,n}^{(m)} - \bar{\xi}_{[ns]+1:[nt]}^{(m)}) \mathbf{1}(\mathbf{U}_i \leq \mathbf{u}), \end{aligned} \quad (\text{B.6})$$

with the usual convention that empty sums are zero.

*Proof of (B.3).* Consider the decomposition

$$\begin{aligned} |(\check{\mathbb{B}}_n^{(m)} - \mathbb{B}_n^{(m)})(s, t, \mathbf{u})| &\leq |\check{\mathbb{B}}_n^{(m)}(s, t, \mathbf{u}) - \mathring{\mathbb{B}}_n^{(m)}(s, t, \mathbf{H}_{[ns]+1:[nt]}^{-1}(\mathbf{u}))| \\ &\quad + |\mathring{\mathbb{B}}_n^{(m)}(s, t, \mathbf{H}_{[ns]+1:[nt]}^{-1}(\mathbf{u})) - \mathbb{B}_n^{(m)}(s, t, \mathbf{H}_{[ns]+1:[nt]}^{-1}(\mathbf{u}))| \\ &\quad + |\mathbb{B}_n^{(m)}(s, t, \mathbf{H}_{[ns]+1:[nt]}^{-1}(\mathbf{u})) - \mathbb{B}_n^{(m)}(s, t, \mathbf{u})|. \end{aligned}$$

Write out the definitions of the processes  $\mathbb{B}_n^{(m)}$ ,  $\mathring{\mathbb{B}}_n^{(m)}$  and  $\check{\mathbb{B}}_n^{(m)}$  in (B.1), (B.6) and (4.2), respectively, and take suprema over  $(s, t, \mathbf{u}) \in \Delta \times [0, 1]^d$  to obtain

$$\begin{aligned} \sup_{(s,t,\mathbf{u}) \in \Delta \times [0,1]^d} |(\check{\mathbb{B}}_n^{(m)} - \mathbb{B}_n^{(m)})(s, t, \mathbf{u})| \\ \leq \sup_{(s,t,\mathbf{u}) \in \Delta \times [0,1]^d} |\check{\mathbb{B}}_n^{(m)}(s, t, \mathbf{u}) - \mathring{\mathbb{B}}_n^{(m)}(s, t, \mathbf{H}_{[ns]+1:[nt]}^{-1}(\mathbf{u}))| \end{aligned} \quad (\text{B.7})$$

$$+ \sup_{(s,t,\mathbf{u}) \in \Delta \times [0,1]^d} \left| \frac{1}{\sqrt{n}} \sum_{i=[ns]+1}^{[nt]} \xi_{i,n}^{(m)} \right| |H_{[ns]+1:[nt]}(\mathbf{u}) - C(\mathbf{u})| \quad (\text{B.8})$$

$$+ \sup_{(s,t,\mathbf{u}) \in \Delta \times [0,1]^d} |\mathbb{B}_n^{(m)}(s, t, \mathbf{H}_{[ns]+1:[nt]}^{-1}(\mathbf{u})) - \mathbb{B}_n^{(m)}(s, t, \mathbf{u})|. \quad (\text{B.9})$$

Each term requires a different treatment.

1- *The term (B.8):* Let  $\delta \in (0, 1/2)$ , to be specified later. We split the supremum into two parts, according to whether  $t - s$  is smaller or larger than  $a_n = n^{-1/2-\delta}$ :

$$A_{n,1} = \sup_{\substack{(s,t,\mathbf{u}) \in \Delta \times [0,1]^d \\ t-s \leq a_n}} \left| \frac{1}{\sqrt{n}} \sum_{i=[ns]+1}^{[nt]} \xi_{i,n}^{(m)} \right| |H_{[ns]+1:[nt]}(\mathbf{u}) - C(\mathbf{u})|,$$

$$A_{n,2} = \sup_{\substack{(s,t,\mathbf{u}) \in \Delta \times [0,1]^d \\ t-s \geq a_n}} \left| \frac{1}{\sqrt{n}} \sum_{i=[ns]+1}^{[nt]} \xi_{i,n}^{(m)} \right| |H_{[ns]+1:[nt]}(\mathbf{u}) - C(\mathbf{u})|.$$

We will show that both  $A_{n,1}$  and  $A_{n,2}$  converge to zero in probability.

(a) Since both  $H_{k:l}$  and  $C$  take values in  $[0, 1]$ , a crude bound for  $A_{n,1}$  is

$$A_{n,1} \leq \frac{1}{\sqrt{n}}(n a_n + 1) \max_{1 \leq i \leq n} |\xi_{i,n}^{(m)}| \leq 2 n^{-\delta} \max_{1 \leq i \leq n} |\xi_{i,n}^{(m)}|.$$

Using the fact that, from (M1), for any  $\nu \geq 1$ ,  $\sup_{n \geq 1} \mathbb{E}[|\xi_{1,n}^{(m)}|^\nu] < \infty$ , we have that, for every  $\alpha > 0$  and  $\nu \geq 1$  such that  $\nu > 1/\alpha$ ,

$$\mathbb{P}\left(\max_{1 \leq i \leq n} |\xi_{i,n}^{(m)}| \geq n^\alpha\right) \leq n \mathbb{P}(|\xi_{1,n}^{(m)}| \geq n^\alpha) \leq n^{1-\nu\alpha} \mathbb{E}[|\xi_{1,n}^{(m)}|^\nu] \rightarrow 0.$$

Apply the previous display with  $\alpha \in (0, \delta)$  to find that  $A_{n,1}$  converges to zero in probability.

(b) Recall  $\mathbb{B}_n$  in (3.4). Observe that

$$\mathbb{B}_n(s, t, \mathbf{u}) = \frac{[nt] - [ns]}{\sqrt{n}} \{H_{[ns]+1:[nt]}(\mathbf{u}) - C(\mathbf{u})\}.$$

We have

$$\begin{aligned} A_{n,2} &= \sup_{\substack{(s,t,\mathbf{u}) \in \Delta \times [0,1]^d \\ t-s \geq a_n}} \left| \frac{1}{[nt] - [ns]} \sum_{i=[ns]+1}^{[nt]} \xi_{i,n}^{(m)} \right| |\mathbb{B}_n(s, t, \mathbf{u})| \\ &\leq \sup_{\substack{[ns] < [nt] \\ [nt]+1 - [ns] \geq na_n}} \left| \frac{1}{[nt] - [ns]} \sum_{i=[ns]+1}^{[nt]} \xi_{i,n}^{(m)} \right| \sup_{(s,t,\mathbf{u}) \in \Delta \times [0,1]^d} |\mathbb{B}_n(s, t, \mathbf{u})| \\ &\leq \max_{\substack{1 \leq k \leq l \leq n \\ l-k \geq na_n}} \left| \frac{1}{l-k+1} \sum_{i=k}^l \xi_{i,n}^{(m)} \right| \sup_{(s,t,\mathbf{u}) \in \Delta \times [0,1]^d} |\mathbb{B}_n(s, t, \mathbf{u})|. \end{aligned}$$

By weak convergence  $\mathbb{B}_n \rightsquigarrow \mathbb{B}_C$  in  $\ell^\infty(\Delta \times [0, 1]^d)$ , the supremum at the end of the previous display is bounded in probability. Writing  $b_n = na_n$ , it is sufficient to show that

$$\max_{\substack{1 \leq k \leq l \leq n \\ l-k \geq b_n}} \left| \frac{1}{l-k+1} \sum_{i=k}^l \xi_{i,n}^{(m)} \right| \xrightarrow{\mathbb{P}} 0, \quad n \rightarrow \infty.$$

Fix  $\eta > 0$ . The probability that the previous maximum exceeds  $\eta$  is bounded by

$$\sum_{\substack{1 \leq k \leq l \leq n \\ l-k \geq b_n}} \mathbb{P} \left[ \left| \frac{1}{l-k+1} \sum_{i=k}^l \xi_{i,n}^{(m)} \right| > \eta \right]. \quad (\text{B.10})$$

Fix  $\nu \geq 2$ , to be specified later. By stationarity and Markov's inequality, the previous expression is bounded by

$$\sum_{\substack{1 \leq k \leq l \leq n \\ l-k \geq b_n}} \eta^{-\nu} (l-k+1)^{-\nu} \mathbb{E} \left[ \left| \sum_{i=1}^{l-k+1} \xi_{i,n}^{(m)} \right|^\nu \right] \leq \eta^{-\nu} n \sum_{b_n \leq r \leq n} r^{-\nu} \mathbb{E} \left[ \left| \sum_{i=1}^r \xi_{i,n}^{(m)} \right|^\nu \right].$$

Recall that the sequence  $(\xi_{i,n}^{(m)})_{i \in \mathbb{Z}}$  is  $\ell_n$ -dependent from (M2) and assume that  $n$  is sufficiently large so that  $n \geq 2\ell_n + b_n$ . Then, by Corollary A.1 in Romano and Wolf (2000), there exists a constant  $C_\nu$ , depending only on  $\nu$ , such that

$$\mathbb{E} \left[ \left| \sum_{i=1}^r \xi_{i,n}^{(m)} \right|^\nu \right] \leq C_\nu^\nu (4\ell_n r)^{\nu/2} \mathbb{E} [|\xi_{1,n}^{(m)}|^\nu].$$

Using the fact that, from (M1),  $\sup_{n \geq 1} \mathbb{E} [|\xi_{1,n}^{(m)}|^\nu] < \infty$ , and up to a multiplicative constant, the expression in (B.10) is bounded by

$$\begin{aligned} n \sum_{b_n \leq r \leq n} r^{-\nu} (\ell_n r)^{\nu/2} &\leq n^2 b_n^{-\nu/2} \ell_n^{\nu/2} \\ &= O(n^{2-(1/2-\delta)\nu/2+(1/2-\gamma)\nu/2}) \\ &= O(n^{2+(\delta-\gamma)\nu/2}). \end{aligned}$$

The right-hand side converges to zero if we choose  $\delta = \gamma/2$  and then  $\nu > 8/\gamma$ .

2- *The term (B.9):* We have to show that, for every  $\eta, \lambda > 0$ ,

$$\mathbb{P} \left[ \sup_{(s,t,\mathbf{u}) \in \Delta \times [0,1]^d} \left| \mathbb{B}_n^{(m)} \{s, t, \mathbf{H}_{[ns]+1:[nt]}^{-1}(\mathbf{u})\} - \mathbb{B}_n^{(m)}(s, t, \mathbf{u}) \right| > \lambda \right] \leq \eta,$$

for all sufficiently large  $n$ .

Fix  $\eta, \lambda > 0$ . Since (B.9) is smaller than  $2 \sup_{(s,t,\mathbf{u}) \in \Delta \times [0,1]^d} |\mathbb{B}_n^{(m)}(s, t, \mathbf{u})|$ , and using the fact that  $\mathbb{B}_n^{(m)}$  vanishes on the diagonal  $s = t$  and is asymptotically uniformly equicontinuous in probability, there exists  $\varepsilon \in (0, 1)$  such that, for all  $n$  sufficiently large,

$$\mathbb{P} \left[ \sup_{\substack{(s,t,\mathbf{u}) \in \Delta \times [0,1]^d \\ t-s < \varepsilon}} \left| \mathbb{B}_n^{(m)} \{s, t, \mathbf{H}_{[ns]+1:[nt]}^{-1}(\mathbf{u})\} - \mathbb{B}_n^{(m)}(s, t, \mathbf{u}) \right| > \lambda \right] \leq \eta/2.$$

Setting  $\zeta_n = \sup_{\substack{(s,t,\mathbf{u}) \in \Delta \times [0,1]^d \\ t-s \geq \varepsilon}} \| \mathbf{H}_{[ns]+1:[nt]}^{-1}(\mathbf{u}) - \mathbf{u} \|_1$ , we shall now show that, for all  $n$  sufficiently large,

$$A_n = \mathbb{P} \left\{ \sup_{\substack{(s,t,\mathbf{u},\mathbf{v}) \in \Delta \times [0,1]^{2d} \\ t-s \geq \varepsilon, \|\mathbf{u}-\mathbf{v}\|_1 \leq \zeta_n}} \left| \mathbb{B}_n^{(m)}(s,t,\mathbf{u}) - \mathbb{B}_n^{(m)}(s,t,\mathbf{v}) \right| > \lambda \right\} \leq \eta/2,$$

which will complete the proof. Using again the asymptotic uniformly equicontinuity in probability of  $\mathbb{B}_n^{(m)}$ , there exists  $\mu \in (0,1)$  such that, for all  $n$  sufficiently large,

$$A_{n,1} = \mathbb{P} \left\{ \sup_{\substack{(s,t,\mathbf{u},\mathbf{v}) \in \Delta \times [0,1]^{2d} \\ t-s \geq \varepsilon, \|\mathbf{u}-\mathbf{v}\|_1 \leq \mu}} \left| \mathbb{B}_n^{(m)}(s,t,\mathbf{u}) - \mathbb{B}_n^{(m)}(s,t,\mathbf{v}) \right| > \lambda \right\} \leq \eta/4.$$

We then bound  $A_n$  by  $A_{n,1} + A_{n,2}$ , where  $A_{n,2} = \mathbb{P}(\zeta_n > \mu)$ . From the weak convergence of  $\mathbb{B}_n$  to  $\mathbb{B}_C$  in  $\ell^\infty(\Delta \times [0,1]^d)$ , we have that

$$\begin{aligned} & \sup_{\substack{(s,t,\mathbf{u}) \in \Delta \times [0,1]^d \\ t-s \geq \varepsilon}} \left| H_{[ns]+1:[nt]}(\mathbf{u}) - C(\mathbf{u}) \right| \\ & \leq \sup_{(s,t,\mathbf{u}) \in \Delta \times [0,1]^d} \left| \mathbb{B}_n(s,t,\mathbf{u}) \right| \times n^{-1/2} \times \sup_{\substack{(s,t) \in \Delta \\ t-s \geq \varepsilon}} \{ \lambda_n(s,t) \}^{-1} \xrightarrow{\mathbb{P}} 0. \end{aligned}$$

Using the fact that  $\sup_{u \in [0,1]} |H_{[ns]+1:[nt],j}(u) - u| = \sup_{u \in [0,1]} |H_{[ns]+1:[nt],j}^{-1}(u) - u|$  for  $j \in \{1, \dots, d\}$  (for instance, by symmetry arguments on the graphs of  $H_{[ns]+1:[nt],j}$  and  $H_{[ns]+1:[nt],j}^{-1}$ ), we immediately obtain that  $\zeta_n \xrightarrow{\mathbb{P}} 0$ , which implies that, for all  $n$  sufficiently large,  $A_{n,2} \leq \eta/4$ , and thus that, for all  $n$  sufficiently large,  $A_n \leq \eta/2$ .

3- *The term (B.7):* For the following arguments, it is sufficient to assume that the sequence  $(\xi_{i,n}^{(m)})_{i \in \mathbb{Z}}$  appearing in  $\check{\mathbb{B}}_n^{(m)}$  and  $\hat{\mathbb{B}}_n^{(m)}$  satisfies only (M1) with  $\mathbb{E}\{\{\xi_{0,n}^{(m)}\}^2\} > 0$  not necessarily equal to one.

Let  $K > 0$  be a constant and let us first suppose that, for any  $n \geq 1$  and  $i \in \{1, \dots, n\}$ ,  $\xi_{i,n}^{(m)} \geq -K$ . With (A.2) in mind, the term (B.7) is smaller than  $A_{n,1} + A_{n,2}$ , where

$$A_{n,1} = \sup_{(s,t,\mathbf{u}) \in \Delta \times [0,1]^d} \frac{1}{\sqrt{n}} \sum_{i=[ns]+1}^{[nt]} (\xi_{i,n}^{(m)} + K) \left[ \mathbf{1}\{\mathbf{U}_i \leq \mathbf{H}_{[ns]+1:[nt]}^{-1}(\mathbf{u})\} - \mathbf{1}\{\hat{\mathbf{U}}_i^{[ns]+1:[nt]} \leq \mathbf{u}\} \right]$$

and

$$A_{n,2} = \sup_{(s,t,\mathbf{u}) \in \Delta \times [0,1]^d} \frac{K + \bar{\xi}_{[ns]+1:[nt]}^{(m)}}{\sqrt{n}} \sum_{i=[ns]+1}^{[nt]} \left[ \mathbf{1}\{\mathbf{U}_i \leq \mathbf{H}_{[ns]+1:[nt]}^{-1}(\mathbf{u})\} - \mathbf{1}\{\hat{\mathbf{U}}_i^{[ns]+1:[nt]} \leq \mathbf{u}\} \right].$$

Let us first show that  $A_{n,1} \xrightarrow{\mathbb{P}} 0$ . Plugging (A.2) into the expression of  $A_{n,1}$ , we bound  $A_{n,1}$  by  $A_{n,1,1} + \dots + A_{n,1,d}$ , where

$$A_{n,1,j} = \sup_{(s,t,\mathbf{u}) \in \Delta \times [0,1]^d} \frac{1}{\sqrt{n}} \sum_{i=[ns]+1}^{[nt]} (\xi_{i,n}^{(m)} + K) \mathbf{1}(U_{ij} = u).$$



To prove that  $A_{n,1} \xrightarrow{\mathbb{P}} 0$ , we shall now show that  $A_{n,1,j} \xrightarrow{\mathbb{P}} 0$  for all  $j \in \{1, \dots, d\}$ . Fix  $j \in \{1, \dots, d\}$ . Using the fact that  $\mathbf{1}(U_{ij} = u) \leq \mathbf{1}(U_{ij} \leq u) - \mathbf{1}(U_{ij} \leq u - 1/n)$ , we obtain that  $A_{n,1,j}$  is smaller than  $A'_{n,1,j} + A''_{n,1,j} + A'''_{n,1,j}$ , where

$$\begin{aligned} A'_{n,1,j} &= \sup_{(s,t,u) \in \Delta \times [0,1]} \left| \frac{1}{\sqrt{n}} \sum_{i=\lfloor ns \rfloor + 1}^{\lfloor nt \rfloor} (\xi_{i,n}^{(m)} - \bar{\xi}_{\lfloor ns \rfloor + 1 : \lfloor nt \rfloor}^{(m)}) \{ \mathbf{1}(U_{ij} \leq u) - \mathbf{1}(U_{ij} \leq u - 1/n) \} \right|, \\ A''_{n,1,j} &= K \sup_{(s,t,u) \in \Delta \times [0,1]} \frac{1}{\sqrt{n}} \sum_{i=\lfloor ns \rfloor + 1}^{\lfloor nt \rfloor} \{ \mathbf{1}(U_{ij} \leq u) - \mathbf{1}(U_{ij} \leq u - 1/n) \}, \\ A'''_{n,1,j} &= \sup_{(s,t,u) \in \Delta \times [0,1]} \frac{|\bar{\xi}_{\lfloor ns \rfloor + 1 : \lfloor nt \rfloor}^{(m)}|}{\sqrt{n}} \sum_{i=\lfloor ns \rfloor + 1}^{\lfloor nt \rfloor} \{ \mathbf{1}(U_{ij} \leq u) - \mathbf{1}(U_{ij} \leq u - 1/n) \}. \end{aligned}$$

From Lemma A.3 of Bücher and Kojadinovic (2013), we know that  $\mathbb{B}_n^{(m)}$  is asymptotically uniformly equicontinuous in probability under the weaker conditions on the sequence  $(\xi_{i,n}^{(m)})_{i \in \mathbb{Z}}$  considered above. The treatment of the term (B.8) carried out previously remains valid under these conditions and ensures that

$$\sup_{(s,t,\mathbf{u}) \in \Delta \times [0,1]^d} |\mathring{\mathbb{B}}_n^{(m)}(s, t, \mathbf{u}) - \mathbb{B}_n^{(m)}(s, t, \mathbf{u})| \xrightarrow{\mathbb{P}} 0,$$

which implies that  $\mathring{\mathbb{B}}_n^{(m)}$  is asymptotically uniformly equicontinuous in probability as well under the same weaker conditions on the sequence  $(\xi_{i,n}^{(m)})_{i \in \mathbb{Z}}$ . The latter immediately implies that  $A'_{n,1,j} \xrightarrow{\mathbb{P}} 0$ . For  $A''_{n,1,j}$ , we have

$$A''_{n,1,j} \leq K \sup_{\substack{\mathbf{u}, \mathbf{v} \in [0,1]^d \\ \|\mathbf{u} - \mathbf{v}\|_1 \leq n^{-1}}} |\mathbb{B}_n(0, 1, \mathbf{u}) - \mathbb{B}_n(0, 1, \mathbf{v})| + Kn^{-1/2} \xrightarrow{\mathbb{P}} 0,$$

by asymptotic uniform equicontinuity in probability of  $\mathbb{B}_n$ . The fact that  $A'''_{n,1,j} \xrightarrow{\mathbb{P}} 0$  can be shown by proceeding as for the term (B.8). Hence, we have that  $A_{n,1,j} \xrightarrow{\mathbb{P}} 0$ , which implies that  $A_{n,1} \xrightarrow{\mathbb{P}} 0$ .

The fact that  $A_{n,2} \xrightarrow{\mathbb{P}} 0$ , follows from (A.2) which implies that  $A_{n,2}$  is smaller than  $\sum_{j=1}^d (A''_{n,1,j} + A'''_{n,1,j})$ . This completes the proof under the condition  $\xi_{i,n}^{(m)} \geq -K$ .

To show that this condition is not necessary, we proceed as at the end of the proof of Lemma A.3 of Bücher and Kojadinovic (2013). Let  $Z_{i,n}^+ = \max(\xi_{i,n}^{(m)}, 0)$ ,  $Z_{i,n}^- = \max(-\xi_{i,n}^{(m)}, 0)$ ,  $K^+ = E(Z_{0,n}^+)$  and  $K^- = E(Z_{0,n}^-)$ . Furthermore, define  $\xi_{i,n}^{(m),+} = Z_{i,n}^+ - K^+$  and  $\xi_{i,n}^{(m),-} = Z_{i,n}^- - K^-$ . Then, using the fact that  $K^+ - K^- = 0$ , we can write

$$\xi_{i,n}^{(m)} = Z_{i,n}^+ - Z_{i,n}^- = Z_{i,n}^+ - K^+ - (Z_{i,n}^- - K^-) = \xi_{i,n}^{(m),+} - \xi_{i,n}^{(m),-}.$$

Let  $\mathbb{B}_n^{(m),+}$  and  $\mathbb{B}_n^{(m),-}$  be the analogues of  $\mathbb{B}_n^{(m)}$  defined from the sequences  $(\xi_{i,n}^{(m),+})_{i \in \mathbb{Z}}$  and  $(\xi_{i,n}^{(m),-})_{i \in \mathbb{Z}}$ , respectively, and similarly for  $\mathring{\mathbb{B}}_n^{(m),+}$  and  $\mathring{\mathbb{B}}_n^{(m),-}$ . The case treated

above yields

$$\begin{aligned} \sup_{(s,t,\mathbf{u}) \in \Delta \times [0,1]^d} \left| \check{\mathbb{B}}_n^{(m),+}(s,t,\mathbf{u}) - \mathring{\mathbb{B}}_n^{(m),+}(s,t, \mathbf{H}_{[ns]+1:[nt]}^{-1}(\mathbf{u})) \right| &\xrightarrow{P} 0, \\ \sup_{(s,t,\mathbf{u}) \in \Delta \times [0,1]^d} \left| \check{\mathbb{B}}_n^{(m),-}(s,t,\mathbf{u}) - \mathring{\mathbb{B}}_n^{(m),-}(s,t, \mathbf{H}_{[ns]+1:[nt]}^{-1}(\mathbf{u})) \right| &\xrightarrow{P} 0. \end{aligned}$$

The desired result finally follows from the fact that  $\mathbb{B}_n^{(m)} = \mathbb{B}_n^{(m),+} - \mathbb{B}_n^{(m),-}$  and  $\mathring{\mathbb{B}}_n^{(m)} = \mathring{\mathbb{B}}_n^{(m),+} - \mathring{\mathbb{B}}_n^{(m),-}$ . ■

## C On the set-up of the simulation experiments

For the numerical experiments involving serially dependent observations, we restricted ourselves to the bivariate case and only focused on the tests based on  $\check{S}_n$  and  $\hat{S}_n$ . Given a bivariate copula  $C$ , two models were used to generate serially dependent observations under  $H_0$  defined in (1.1).

- The first one is a simple autoregressive model of order one, AR(1). Let  $\mathbf{U}_i$ ,  $i \in \{-100, \dots, 0, \dots, n\}$ , be a bivariate i.i.d. sample from a copula  $C$ . Then, set  $\boldsymbol{\epsilon}_i = (\Phi^{-1}(U_{i1}), \Phi^{-1}(U_{i2}))$ , where  $\Phi$  is the c.d.f. of the standard normal distribution, and  $\mathbf{X}_{-100} = \boldsymbol{\epsilon}_{-100}$ . Finally, for any  $j \in \{1, 2\}$  and  $i \in \{-99, \dots, 0, \dots, n\}$ , compute recursively

$$X_{ij} = 0.5X_{i-1,j} + \epsilon_{ij}. \quad (\text{AR1})$$

- The second model is a bivariate version of the exponential autoregressive (EXPAR) model considered in Auestad and Tjøstheim (1990) and Paparoditis and Politis (2001, Section 3.3) (see also Bücher and Kojadinovic, 2013). The sample  $\mathbf{X}_1, \dots, \mathbf{X}_n$  is generated as previously with (AR1) replaced by

$$X_{ij} = \{0.8 - 1.1 \exp(-50X_{i-1,j}^2)\}X_{i-1,j} + 0.1\epsilon_{ij}. \quad (\text{EXPAR})$$

Data under  $(\neg H_0) \cap H_{0,m}$ , with  $H_{0,m}$  as in (1.2), were generated using the procedures described above except that the bivariate random vectors  $\mathbf{U}_i$ ,  $i \in \{-100, \dots, 0, \dots, n\}$  are independent such that  $\mathbf{U}_i$ ,  $i \in \{-100, \dots, 0, \dots, k^*\}$  are i.i.d. from a copula  $C_1$  and  $\mathbf{U}_i$ ,  $i \in \{k^* + 1, \dots, n\}$  are i.i.d. from a copula  $C_2$ , where  $C_1 \neq C_2$  and  $k^* = \lfloor nt \rfloor$  for some  $t \in (0, 1)$ . The resulting samples  $\mathbf{X}_1, \dots, \mathbf{X}_n$  are therefore not samples under  $H_{1,c} \cap H_{0,m}$  since the change in the dependence is gradual by (AR1) or (EXPAR). The copulas  $C_1$  and  $C_2$  were taken to be both either bivariate Clayton, Gumbel–Hougaard, Normal or Frank copulas such that  $C_1$  has a Kendall’s tau of 0.2 and  $C_2$  a Kendall’s tau of  $\tau \in \{0.4, 0.6\}$ . The parameter  $t$  defining  $k^*$  was chosen in  $\{0.25, 0.5\}$ .

The dependent multiplier sequences necessary to carry out the tests were generated using the “moving average approach” proposed initially in Bühlmann (1993, Section 6.2) and revisited in some detail in Bücher and Kojadinovic (2013, Section 6.1). A standard

normal sequence was used for the required initial i.i.d. sequence. The kernel function  $\kappa$  in that procedure was chosen to be the Parzen kernel defined by  $\kappa_P(x) = (1 - 6x^2 + 6|x|^3)\mathbf{1}(|x| \leq 1/2) + 2(1 - |x|)^3\mathbf{1}(1/2 < |x| \leq 1)$ ,  $x \in \mathbb{R}$ , which amounts to choosing the function  $\varphi$  in Condition (M3) as  $x \mapsto (\kappa_P \star \kappa_P)(2x)/(\kappa_P \star \kappa_P)(0)$ , where ‘ $\star$ ’ denotes the convolution operator. The value of the bandwidth parameter  $\ell_n$  defined in Condition (M2) was chosen using the procedure described in Bücher and Kojadinovic (2013, Section 5). Two choices for the “combining” function  $\psi$  in that procedure were considered: the median and the maximum. Both choices led to similar rejection rates. The results reported in Tables 2 and 6 below are those obtained with  $\psi = \text{maximum}$ .

## D Selected results of the simulation study

Tables 1 up to 6 provide partial results of the large-scale Monte Carlo simulation experiment described in Section 5. All the tests were carried out at the 5% level of significance.

Table 1: Percentage of rejection of  $H_0$  computed from 1000 random samples of size  $n \in \{50, 100, 200\}$  generated under  $H_0$ , where  $C$  is either the  $d$ -dimensional Clayton (Cl), the Gumbel–Hougaard (GH) or the normal (N) copula whose bivariate margins have a Kendall’s tau of  $\tau$ .

$d$	$n$	$\tau$	Cl			GH			N		
			$\check{S}_n$	$\hat{S}_n$	$S_n^R$	$\check{S}_n$	$\hat{S}_n$	$S_n^R$	$\check{S}_n$	$\hat{S}_n$	$S_n^R$
2	50	0.00	6.2	4.0	4.6	5.4	2.9	4.5	7.3	3.4	4.8
		0.25	6.7	6.2	5.6	5.5	3.3	5.4	4.4	3.0	6.3
		0.50	5.6	7.9	6.0	4.4	3.3	4.6	4.4	5.3	4.9
		0.75	6.0	16.6	5.5	3.2	6.7	4.3	3.6	9.1	4.9
	100	0.00	4.9	3.5	5.3	5.2	4.1	5.5	4.3	2.8	5.5
		0.25	6.1	6.6	5.0	5.0	3.3	6.2	5.3	4.0	5.5
		0.50	4.4	9.3	5.9	3.7	2.8	5.7	3.1	3.4	5.3
		0.75	2.7	10.0	4.6	2.5	4.7	4.4	2.1	6.0	5.6
	200	0.00	4.0	3.5	5.2	4.3	4.0	5.2	5.4	4.9	4.3
		0.25	4.7	5.2	6.3	3.3	3.0	3.8	4.0	3.9	5.2
		0.50	5.1	8.5	4.9	3.2	2.3	4.5	4.0	4.7	4.8
		0.75	2.6	9.3	5.9	1.5	3.1	5.2	1.9	4.8	5.7
3	50	0.00	4.3	1.5	3.0	4.2	2.1	3.6	5.5	2.8	3.4
		0.25	6.3	5.0	5.1	5.5	1.0	5.1	5.3	3.0	4.3
		0.50	8.2	9.1	5.9	2.7	0.9	5.7	3.0	2.2	4.6
		0.75	2.0	2.9	6.9	0.5	0.4	6.3	1.1	1.3	4.1
	100	0.00	4.5	3.4	4.5	4.5	2.8	4.6	4.5	2.7	3.9
		0.25	5.0	5.1	5.4	4.2	2.6	4.4	5.4	3.5	4.5
		0.50	5.7	7.6	6.3	3.3	1.3	5.0	3.2	3.1	3.9
		0.75	2.5	4.9	5.0	1.0	1.0	5.2	0.8	1.6	5.5
	200	0.00	3.3	2.5	4.3	3.5	3.2	4.3	4.8	4.0	4.7
		0.25	6.6	7.1	5.5	5.0	3.3	4.5	4.8	4.1	5.0
		0.50	6.0	9.2	4.5	3.0	2.4	5.9	4.8	4.3	4.8
		0.75	2.9	6.4	6.4	0.7	0.9	3.8	1.3	2.2	4.9

Table 2: Percentage of rejection of  $H_0$  computed from 1000 samples of size  $n \in \{100, 200\}$  generated under  $H_0$  as explained in Appendix C, where  $C$  is either the bivariate Clayton (Cl), the Gumbel–Hougaard (GH), the normal (N) or the Frank (F) copula with a Kendall’s tau of  $\tau$ . The columns  $\widehat{\ell}_n^{opt}$  and std give the mean and the standard deviation of the values of  $\ell_n$  used for creating the dependent multiplier sequences.

$C$	$n$	$\tau$	AR1				EXPAR			
			$\widehat{\ell}_n^{opt}$	std	$\check{S}_n$	$\hat{S}_n$	$\widehat{\ell}_n^{opt}$	std	$\check{S}_n$	$\hat{S}_n$
Cl	100	0.00	14.2	8.5	4.2	0.7	16.7	9.5	5.1	0.6
		0.25	14.1	8.6	5.9	1.7	16.7	10.2	5.5	2.3
		0.50	14.0	10.3	4.5	2.9	16.4	10.7	6.1	3.2
		0.75	13.3	10.0	2.7	3.3	15.5	10.7	5.1	4.7
	200	0.00	16.7	8.0	5.1	2.6	20.9	9.5	4.5	1.7
		0.25	16.0	7.3	5.1	3.0	20.5	9.7	4.1	2.0
		0.50	15.8	7.8	2.6	2.5	19.8	9.9	3.5	2.8
		0.75	15.5	9.0	1.6	3.6	19.0	9.0	4.3	4.7
GH	100	0.00	14.5	9.7	4.6	0.9	16.9	8.2	4.4	0.6
		0.25	14.1	8.7	4.9	1.5	17.1	10.2	5.0	0.6
		0.50	14.0	9.5	3.9	1.2	15.8	9.5	4.0	0.3
		0.75	13.7	10.0	2.6	0.5	15.2	9.2	1.6	0.4
	200	0.00	16.8	7.9	4.3	1.8	21.5	10.1	3.6	1.5
		0.25	16.6	8.8	5.5	2.0	20.9	11.5	5.1	1.1
		0.50	15.9	7.4	3.7	1.6	20.1	10.9	2.9	0.7
		0.75	15.5	8.7	1.3	0.9	18.8	9.1	1.6	0.2
N	100	0.00	14.1	7.9	5.0	1.1	17.3	9.2	5.4	1.4
		0.25	13.5	8.1	5.9	1.4	17.3	10.8	5.0	1.1
		0.50	13.5	9.1	3.3	1.4	16.4	9.7	4.5	1.0
		0.75	12.9	7.9	1.7	1.7	15.7	10.8	2.7	1.1
	200	0.00	16.3	6.2	5.4	1.9	20.7	8.7	3.7	1.5
		0.25	16.0	7.1	4.2	2.4	20.9	8.9	5.0	1.5
		0.50	16.1	7.8	4.2	3.2	19.8	10.4	2.9	1.8
		0.75	15.4	7.5	0.9	1.4	19.3	10.7	0.8	0.5
F	100	0.00	13.8	7.8	5.8	1.5	17.4	9.7	6.4	0.8
		0.25	14.2	9.3	5.5	1.3	16.6	9.7	4.8	0.5
		0.50	13.9	9.0	3.3	1.7	16.8	11.4	3.6	0.5
		0.75	13.5	9.7	1.5	0.6	15.8	10.3	3.1	1.2
	200	0.00	16.8	7.2	4.2	2.3	20.9	9.4	4.3	1.4
		0.25	16.0	7.0	6.0	2.9	20.6	8.7	3.6	1.1
		0.50	16.1	8.2	3.1	1.5	20.3	10.4	3.2	1.2
		0.75	15.8	8.8	0.9	0.5	19.6	9.0	1.2	0.8

Table 3: Percentage of rejection of  $H_0$  computed from 1000 samples of size  $n \in \{50, 100, 200\}$  generated under  $H_{0,m} \cap H_{1,c}$ , where  $H_{1,c}$  is defined in (5.1),  $k^* = \lfloor nt \rfloor$ ,  $C_1$  and  $C_2$  are both either bivariate Clayton (Cl), Gumbel–Hougaard (GH) or normal (N) copulas such that  $C_1$  has a Kendall’s tau of 0.2 and  $C_2$  a Kendall’s tau of  $\tau$ .

$n$	$\tau$	$t$	Cl			GH			N		
			$\check{S}_n$	$\hat{S}_n$	$S_n^R$	$\check{S}_n$	$\hat{S}_n$	$S_n^R$	$\check{S}_n$	$\hat{S}_n$	$S_n^R$
50	0.4	0.10	7.5	8.1	5.7	6.1	3.7	4.3	6.1	4.8	4.3
		0.25	12.1	10.6	4.0	9.4	5.1	5.0	10.4	7.6	5.5
		0.50	18.0	16.1	6.3	12.1	7.8	4.8	12.3	8.4	5.8
	0.6	0.10	14.5	17.0	5.4	11.4	7.7	6.2	11.4	9.8	7.2
		0.25	35.5	34.4	7.4	29.9	21.4	6.4	31.3	21.4	7.1
		0.50	47.3	41.6	7.0	45.3	30.3	8.9	46.0	33.9	9.1
100	0.4	0.10	7.1	8.7	6.1	6.6	5.1	5.1	5.8	5.2	5.3
		0.25	18.8	19.9	5.2	16.9	13.2	5.8	14.9	12.5	6.1
		0.50	26.5	26.4	7.3	23.8	18.8	7.3	22.6	19.1	7.9
	0.6	0.10	21.5	25.1	5.8	16.7	12.0	5.1	17.5	16.9	6.1
		0.25	65.1	66.0	6.3	61.2	51.5	7.5	62.9	54.8	9.9
		0.50	82.1	81.6	14.7	78.8	69.7	14.9	79.1	73.7	11.8
200	0.4	0.10	11.1	13.8	5.8	8.3	8.1	5.5	9.5	9.8	5.0
		0.25	30.8	33.9	5.9	27.6	24.8	6.4	29.6	28.3	6.8
		0.50	47.1	48.6	9.0	45.8	41.4	8.7	47.1	46.1	9.4
	0.6	0.10	36.4	41.3	6.8	34.4	31.7	7.1	36.0	36.3	6.7
		0.25	92.6	93.2	12.3	91.4	88.9	16.7	91.3	90.2	12.0
		0.50	98.9	99.3	22.2	98.5	98.1	22.0	99.3	99.1	21.1

Table 4: Percentage of rejection of  $H_0$  computed from 1000 samples of size  $n \in \{100, 200\}$  generated under  $H_{0,m} \cap H_{1,c}$ , where  $H_{1,c}$  is defined in (5.1),  $k^* = \lfloor nt \rfloor$ ,  $C_1$  (resp.  $C_2$ ) is a  $d$ -dimensional Clayton (resp. Gumbel–Hougaard) copula whose bivariate margins have a Kendall’s tau of  $\tau$ .

$n$	$\tau$	$t$	$d = 2$			$d = 3$		
			$\check{S}_n$	$\hat{S}_n$	$S_n^R$	$\check{S}_n$	$\hat{S}_n$	$S_n^R$
100	0.25	0.25	5.7	4.1	5.3	5.3	3.1	4.4
		0.50	6.2	5.7	5.6	9.1	6.0	5.8
		0.75	6.6	6.3	3.5	5.8	5.4	4.9
	0.50	0.25	5.5	5.9	5.8	4.6	2.9	5.0
		0.50	10.5	12.2	5.1	15.1	15.1	6.8
		0.75	8.3	11.9	4.4	7.7	9.9	5.3
	0.75	0.25	4.0	7.6	5.1	2.5	1.9	4.3
		0.50	12.5	19.9	6.0	9.8	13.2	5.2
		0.75	8.2	16.4	6.5	4.8	6.7	3.8
200	0.25	0.25	5.8	5.3	6.1	5.8	3.7	5.8
		0.50	8.5	8.7	5.4	9.8	9.7	6.6
		0.75	6.4	6.9	5.5	9.0	9.1	5.3
	0.50	0.25	10.4	11.5	6.8	12.6	10.1	6.7
		0.50	30.9	37.5	5.3	44.0	45.8	7.1
		0.75	16.3	23.2	6.1	20.1	27.0	5.1
	0.75	0.25	11.3	18.0	4.9	15.6	16.7	7.3
		0.50	43.4	54.4	6.2	58.9	63.1	5.2
		0.75	21.3	36.4	4.6	23.6	36.0	6.8

Table 5: Rejection percentage of  $H_0$  computed from 1000 samples of size  $n \in \{50, 100, 200\}$  such that the  $\lfloor nt_1 \rfloor$  first observations of each sample are from a  $d$ -variate c.d.f. with normal copula and  $N(0, 1)$  margins (that is, from a multivariate standard normal c.d.f.), and the  $n - \lfloor nt_1 \rfloor$  last observations are from a  $d$ -variate c.d.f. with normal copula whose first margin is the  $N(\mu, 1)$  and whose  $d - 1$  remaining margins are the  $N(0, 1)$ . The bivariate margins of the normal copula have a Kendall's tau of  $\tau$ .

		$(\mu, t_1) =$		$(0.5, 0.25)$			$(0.5, 0.5)$			$(2, 0.25)$			$(2, 0.5)$		
$d$	$n$	$\tau$	$\check{S}_n$	$\hat{S}_n$	$S_n^R$	$\check{S}_n$	$\hat{S}_n$	$S_n^R$	$\check{S}_n$	$\hat{S}_n$	$S_n^R$	$\check{S}_n$	$\hat{S}_n$	$S_n^R$	
2	50	0.00	6.7	3.8	9.0	6.5	3.1	17.7	4.6	2.4	70.1	5.1	2.6	98.5	
		0.25	5.6	2.8	8.1	4.8	3.6	15.0	6.0	3.1	57.7	4.7	1.6	98.5	
		0.50	4.4	4.1	9.6	3.4	2.9	13.0	18.4	5.1	39.6	5.7	0.9	99.3	
	100	0.00	5.3	4.0	19.0	6.3	5.2	30.4	6.1	4.1	99.0	5.5	2.5	100.0	
		0.25	5.0	4.4	13.9	3.7	2.3	24.9	8.6	4.7	97.5	4.4	2.2	100.0	
		0.50	4.0	3.8	11.5	3.1	3.1	22.6	29.5	13.1	91.6	12.4	2.0	100.0	
	200	0.00	5.5	5.2	34.1	3.9	3.4	61.9	4.3	3.3	100.0	5.6	4.1	100.0	
		0.25	3.9	3.4	27.6	4.1	3.5	51.3	13.6	8.9	100.0	8.1	4.0	100.0	
		0.50	3.4	3.9	18.9	2.8	2.9	43.1	57.8	39.4	100.0	34.8	8.3	100.0	
3	50	0.00	4.9	1.6	5.0	4.5	1.7	10.0	4.8	1.9	36.5	6.0	2.3	79.6	
		0.25	5.0	2.7	6.9	5.0	2.6	9.9	6.9	2.6	24.8	5.1	2.6	87.4	
		0.50	3.7	2.6	5.5	3.3	1.3	8.9	9.4	3.5	17.4	3.0	0.7	94.2	
	100	0.00	4.5	2.3	11.3	4.2	2.5	18.6	4.8	3.0	87.5	3.6	2.0	99.6	
		0.25	4.9	3.1	10.7	5.1	3.5	14.8	6.7	3.8	67.9	5.9	4.2	99.9	
		0.50	2.8	2.0	7.3	3.0	2.4	13.7	16.9	8.6	60.6	6.0	1.3	100.0	
	200	0.00	3.0	2.4	20.1	4.5	4.0	37.3	5.3	3.3	100.0	4.3	3.2	100.0	
		0.25	4.8	4.1	15.3	5.6	4.4	30.6	11.9	8.1	99.2	7.4	5.2	100.0	
		0.50	4.9	3.7	11.8	3.8	3.2	24.9	41.0	30.9	99.2	23.9	7.8	100.0	

Table 6: Percentage of rejection of  $H_0$  computed from 1000 samples of size  $n \in \{100, 200\}$  generated under  $\neg H_0$  as explained in the second paragraph of Appendix C, where  $C_1$  and  $C_2$  are both bivariate Gumbel–Hougaard copulas such that  $C_1$  has a Kendall's tau of 0.2 and  $C_2$  a Kendall's tau of  $\tau$ . The columns  $\hat{\ell}_n^{opt}$  and std give the mean and the standard deviation of the values of  $\ell_n$  used for creating the dependent multiplier sequences.

		AR1						EXPAR			
$n$	$t$	$\tau$	$\hat{\ell}_n^{opt}$	std	$\check{S}_n$	$\hat{S}_n$	$\hat{\ell}_n^{opt}$	std	$\check{S}_n$	$\hat{S}_n$	
100	0.25	0.4	14.2	10.1	13.8	3.8	16.9	10.5	10.6	1.5	
		0.6	14.1	9.1	39.9	10.6	15.8	9.5	32.6	8.5	
	0.50	0.4	14.3	10.1	18.0	5.3	17.1	10.3	13.5	2.5	
		0.6	13.9	8.6	57.2	25.2	16.5	9.3	54.8	17.1	
200	0.25	0.4	16.5	8.6	16.4	8.6	20.6	9.8	14.6	4.5	
		0.6	16.4	8.0	71.2	46.0	19.6	9.0	63.1	32.5	
	0.50	0.4	16.6	7.4	31.9	19.0	20.7	8.8	25.8	12.4	
		0.6	16.3	7.1	89.8	75.5	20.6	9.1	80.9	62.5	