

**Copula-based dynamic models for  
multivariate time series**

B.R. Nasri,  
B.N. Rémillard

G-2018-102

December 2018

---

La collection *Les Cahiers du GERAD* est constituée des travaux de recherche menés par nos membres. La plupart de ces documents de travail a été soumis à des revues avec comité de révision. Lorsqu'un document est accepté et publié, le pdf original est retiré si c'est nécessaire et un lien vers l'article publié est ajouté.

**Citation suggérée:** B.R. Nasri, B. Rémillard (Décembre 2018). Copula-based dynamic models for multivariate time series, Rapport technique, Les Cahiers du GERAD G-2018-102, GERAD, HEC Montréal, Canada.

**Avant de citer ce rapport technique,** veuillez visiter notre site Web (<https://www.gerad.ca/fr/papers/G-2018-102>) afin de mettre à jour vos données de référence, s'il a été publié dans une revue scientifique.

The series *Les Cahiers du GERAD* consists of working papers carried out by our members. Most of these pre-prints have been submitted to peer-reviewed journals. When accepted and published, if necessary, the original pdf is removed and a link to the published article is added.

**Suggested citation:** B.R. Nasri, B. Rémillard (December 2018). Copula-based dynamic models for multivariate time series, Technical report, Les Cahiers du GERAD G-2018-102, GERAD, HEC Montréal, Canada.

**Before citing this technical report,** please visit our website (<https://www.gerad.ca/en/papers/G-2018-102>) to update your reference data, if it has been published in a scientific journal.

---

La publication de ces rapports de recherche est rendue possible grâce au soutien de HEC Montréal, Polytechnique Montréal, Université McGill, Université du Québec à Montréal, ainsi que du Fonds de recherche du Québec – Nature et technologies.

Dépôt légal – Bibliothèque et Archives nationales du Québec, 2018  
– Bibliothèque et Archives Canada, 2018

The publication of these research reports is made possible thanks to the support of HEC Montréal, Polytechnique Montréal, McGill University, Université du Québec à Montréal, as well as the Fonds de recherche du Québec – Nature et technologies.

Legal deposit – Bibliothèque et Archives nationales du Québec, 2018  
– Library and Archives Canada, 2018

---

GERAD HEC Montréal  
3000, chemin de la Côte-Sainte-Catherine  
Montréal (Québec) Canada H3T 2A7

Tél. : 514 340-6053  
Télec. : 514 340-5665  
info@gerad.ca  
www.gerad.ca

---



# Copula-based dynamic models for multivariate time series

Bouchra R. Nasri<sup>a</sup>

Bruno N. Rémillard<sup>b</sup>

<sup>a</sup> GERAD & Department of Mathematics and Statistics, McGill University, Montréal (Québec), Canada, H3A 0B9

<sup>b</sup> GERAD & Department of Decision Sciences, HEC Montréal, Montréal (Québec), Canada, H3T 2A7

bouchra.nasri@mail.mcgill.ca

bruno.remillard@hec.ca

December 2018

Les Cahiers du GERAD

G–2018–102

Copyright © 2018 GERAD, Nasri, Rémillard

Les textes publiés dans la série des rapports de recherche *Les Cahiers du GERAD* n'engagent que la responsabilité de leurs auteurs. Les auteurs conservent leur droit d'auteur et leurs droits moraux sur leurs publications et les utilisateurs s'engagent à reconnaître et respecter les exigences légales associées à ces droits. Ainsi, les utilisateurs:

- Peuvent télécharger et imprimer une copie de toute publication du portail public aux fins d'étude ou de recherche privée;
- Ne peuvent pas distribuer le matériel ou l'utiliser pour une activité à but lucratif ou pour un gain commercial;
- Peuvent distribuer gratuitement l'URL identifiant la publication.

Si vous pensez que ce document enfreint le droit d'auteur, contactez-nous en fournissant des détails. Nous supprimerons immédiatement l'accès au travail et enquêterons sur votre demande.

The authors are exclusively responsible for the content of their research papers published in the series *Les Cahiers du GERAD*. Copyright and moral rights for the publications are retained by the authors and the users must commit themselves to recognize and abide the legal requirements associated with these rights. Thus, users:

- May download and print one copy of any publication from the public portal for the purpose of private study or research;
- May not further distribute the material or use it for any profit-making activity or commercial gain;
- May freely distribute the URL identifying the publication.

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

**Abstract:** In this paper, we propose an intuitive way to couple several dynamic time series models even when there are no innovations. This extends previous work for modeling dependence between innovations of stochastic volatility models. We consider time-independent and time-dependent copula models and we study the asymptotic behavior of some empirical processes constructed from pseudo-observations, as well as the behavior of pseudo-maximum likelihood estimators of the associated copula parameters. The results show that even if the univariate dynamic models depend on unknown parameters, the limiting behavior of many processes of interest do not depend on the estimation errors. One can easily perform tests of change point on the full distribution, the margins or the copula, as if the parameters of the dynamic models were known. This is also true for some interesting parametric models of time-dependent copulas. This interesting property makes it possible to construct consistent tests of specification for the dependence models, without having to consider the dynamic time series models. Monte Carlo simulations are used to demonstrate the power of the proposed goodness-of-fit test for finite samples. An example of application with financial data is given.

**Keywords:** Goodness-of-fit, time series, copulas, dynamic models, generalized error models

# 1 Introduction

Considering the dependence between economic or financial variables is an important issue, as exemplified by the 2008 financial crisis, where the risk of contagion was underestimated, mainly because of overly simplified dependence models. Therefore more adequate models must be developed to capture the dependence between multiple time series. As advocated by Embrechts et al. (2002), one cannot rely on Pearson's correlation as a measure of dependence, unless the series are jointly Gaussian, which is rarely the case in practice. Copulas are a much more flexible tool for dependence modeling. As explained by Patton (2009, 2012), there are typically two ways to exploit copulas in time series modeling. Copulas can be used either to model the dependence between successive values of a univariate time series, or can be used to model the conditional dependence of a random vector, given some information about its past, thereby leading to time-varying copulas.

In most papers where copula-based models for multivariate time series are proposed, individual series are typically modeled first, and a copula is used to capture the dependence between serially independent innovations; see, e.g., van den Goorbergh et al. (2005), Chen and Fan (2006b), Patton (2006), Jondeau and Rockinger (2006). However, not all dynamic models have innovations. Take for example univariate time series modeled by hidden Markov models. One aim of this paper is to propose a natural way to model dependence between any dynamic time series, even those without innovations, by using time-dependant copulas. We also propose a way to estimate the parameters and validate the model. To fix ideas, consider a multivariate time series  $\mathbf{X}_t = (X_{1t}, \dots, X_{dt})$ ,  $t \in \{1, \dots, n\}$ . We suppose that each univariate time series is a "generalized error model" (Du, 2016), meaning that for any  $j \in \{1, \dots, d\}$ ,  $\varepsilon_{jt} = G_{\boldsymbol{\theta}, jt}(X_{jt})$  are i.i.d. with continuous distribution function  $F_j$  and density  $f_j$ , for some  $\mathcal{F}_{j, t-1}$ -measurable transformation  $G_{\boldsymbol{\theta}, jt}$ ,  $\boldsymbol{\theta} \in \mathcal{O}$ . These dynamic models contain the innovation models, in particular all stochastic volatility models, Hidden Markov Models (HMM), and the models in Bai (2003) as particular cases. Note that linking together generalized errors have been proposed implicitly in Patton (2004) who considered the bivariate case where  $F_1$  and  $F_2$  are uniform cdf and  $\mathcal{F}_{1,t} = \mathcal{F}_{2,t}$ . However, only stochastic volatility models were studied.

To model the dependence between the time series, consider  $\mathcal{F}_t \supset \bigvee_{j=1}^d \mathcal{F}_{jt}$ , where  $\bigvee_{j=1}^d \mathcal{F}_{jt}$  is the smallest sigma-algebra containing  $\mathcal{F}_{1t}, \dots, \mathcal{F}_{dt}$ , and assume that for any  $j \in \{1, \dots, d\}$ ,

$$P(\varepsilon_{jt} \leq x_j | \mathcal{F}_{t-1}) = F_j(x_j), \quad x_j \in \mathbb{R}. \quad (1)$$

Condition (1) means that the  $\varepsilon_{jt}$  are i.i.d. with conditional distribution  $F_j$  given  $\mathcal{F}_{t-1}$ . It then follows from Patton (2004)[Theorem 1] or Fermanian and Wegkamp (2012)[Theorem 2] that there exists a sequence of  $\mathcal{F}_{t-1}$ -measurable copulas  $C_t$ , so that the joint conditional distribution function  $K_t$  of  $\boldsymbol{\varepsilon}_t = (\varepsilon_{1t}, \dots, \varepsilon_{dt})$  given  $\mathcal{F}_{t-1}$  is

$$K_t(\mathbf{x}) = C_t\{\mathbf{F}(\mathbf{x})\}, \quad \mathbf{F}(\mathbf{x}) = (F_1(x_1), \dots, F_d(x_d))^\top, \quad \mathbf{x} = (x_1, \dots, x_d)^\top \in \mathbb{R}^d. \quad (2)$$

In fact,  $\mathbf{U}_t = \mathbf{F}(\boldsymbol{\varepsilon}_t) \sim C_t$ , for every  $t \in \{1, \dots, n\}$ .

Since the generalized errors  $\boldsymbol{\varepsilon}_t$  are not observable,  $\boldsymbol{\theta}$  being unknown, the latter must be estimated by a consistent estimator  $\boldsymbol{\theta}_n$ . One can then compute the pseudo-observations  $\mathbf{e}_{n,t} = (e_{n,1t}, \dots, e_{n,dt})^\top = \mathbf{G}_{\boldsymbol{\theta}_n, t}(\mathbf{X}_t)$ , where  $e_{n,jt} = G_{\boldsymbol{\theta}_n, jt}(X_{jt})$ ,  $j \in \{1, \dots, d\}$  and  $t \in \{1, \dots, n\}$ .

In order to present the main results, and in view of applications, one will consider two cases: time-independent copulas (Case 1) and time-dependent copulas (Case 2).

**Case 1.** In this setting, the distribution of  $\boldsymbol{\varepsilon}_t$  is constant, i.e.,  $K_t \equiv K$ , so  $C_t \equiv C$  for all  $t \in \{1, \dots, n\}$ . It follows that  $\mathbf{U}_1, \dots, \mathbf{U}_n$  are i.i.d. with distribution function  $C$ .

**Case 2.** In this setting, one considers the conditional Rosenblatt transform  $\mathcal{R}_t$  associated with  $C_t$ , i.e.,  $\mathcal{R}_t$  is  $\mathcal{F}_{t-1}$  measurable and  $V_t = \mathcal{R}_t(U_t)$  are i.i.d. with distribution function  $\Pi$  for all  $t \in \{1, \dots, n\}$ , where  $\Pi$  is the independence copula. The Rosenblatt transform is defined more precisely in Section 3.3. Note that this construction also works in the time-independent case.

For implementing the proposed methodologies, we suggest to perform the following steps:

1. For each time series  $(X_{jt})_{t=1}^n$ ,  $j \in \{1, \dots, d\}$ , choose a generalized error model  $G_{\theta,t}$  and estimate  $\theta$  by a consistent estimator  $\hat{\theta}_n$ .
2. Compute the pseudo-observations  $e_{n,jt} = G_{\hat{\theta}_n,jt}(X_{jt})$ ,  $j \in \{1, \dots, d\}$ ,  $t \in \{1, \dots, n\}$ .
3. For each series, one should perform goodness-of-fit tests. For example, one can use the results of Section 2, namely Theorem 1 to test for a specific distribution of each margin, using a parametric bootstrap algorithm (Rémillard, 2011), or Corollary 1 to test for change point in each margin, using statistics  $\mathcal{T}_{n,j}$  or  $\mathcal{S}_{n,j}$ ,  $j \in \{1, \dots, d\}$ .
4. If the univariate models are satisfactory, go to step 5; otherwise, go back to step 1.
5. Use Proposition 1 in Section 2 to test the null hypothesis of a constant copula.
  - If accepted, propose a time-independent copula model and estimate its parameters using the usual pseudo-maximum likelihood (Genest et al., 1995). One can also use classical goodness-of-fit test for copulas (Genest et al., 2009). The validity of these procedures follows from the results of Section 3, i.e., Theorem 2 and Theorem 4. These results extend the ones in Rémillard (2017) obtained in the case of innovations.
  - If rejected, propose a time-dependent copula model (time-dependent parametric model or HMM) and estimate the parameters using Theorem 2 or Theorem 3. Also test for goodness-of-fit using Theorem 4. These results generalize the ones obtained in Nasri et al. (2017) for the serially independent case.

The paper is structured as follows. In Section 2, under minimal conditions, and assuming a time-independent copula model, we study the asymptotic behavior of empirical processes constructed from the pseudo-observations associated with estimated generalized error models. It is shown that the asymptotic behavior of the sequential processes used for testing change point does not depend on the unknown parameter  $\theta$ . Furthermore, the asymptotic distribution of the empirical copula process does not depend on the parameter  $\theta$  either. Then, in Section 3, we consider time-dependent copulas. We start by showing that in two interesting cases, the asymptotic behavior of the maximum pseudo likelihood estimation is independent of the estimated parameters of the generalized errors. As a particular case, one recovers the results of Chen and Fan (2006a) obtained in the case of univariate stochastic volatility models. We also consider goodness-of-fit tests in this context, using conditional Rosenblatt transforms. It is shown that the asymptotic behavior of the associated empirical process does not depend on the unknown parameter  $\theta$ . In Section 4, Monte Carlo experiments are performed to assess the power of the proposed goodness-of-fit test. Finally, an example of application is provided in Section 5.

## 2 Weak convergence of empirical processes in the time-independent case

From now on, convergence of processes means convergence with respect to the Skorohod topology for the space of càdlàg processes, and is denoted by  $\rightsquigarrow$ . The processes studied here are indexed by  $[0, 1] \times [0, 1]^d$ ,  $[0, 1] \times [-\infty, +\infty]^d$ , or product of these spaces. Note that random vectors belong to these spaces, being constant random functions.

In this section, one focuses on Case 1, meaning that  $K_t \equiv K$ , and  $C_t \equiv C$  for all  $t \in \{1, \dots, n\}$ . This is an extension of the results in Rémillard (2017) who considered only stochastic volatility models. The main processes to be studied are the sequential empirical process and copula process constructed from the pseudo-observations  $\mathbf{e}_{n,t}$ . More precisely, define

$$\mathbb{K}_n(s, \mathbf{x}) = \frac{1}{\sqrt{n}} \sum_{t=1}^{\lfloor ns \rfloor} \{\mathbb{I}(\mathbf{e}_{n,t} \leq \mathbf{x}) - K(\mathbf{x})\}, \quad (s, \mathbf{x}) \in [0, 1] \times \bar{\mathbb{R}}^d, \quad (3)$$

where  $\mathbb{I}$  is the indicator function and  $\mathbf{y} \leq \mathbf{x}$  means that the inequality holds componentwise. Further set

$$K_n(\mathbf{x}) = \frac{1}{n} \sum_{t=1}^n \mathbb{I}(\mathbf{e}_{n,t} \leq \mathbf{x}), \quad \mathbf{x} \in \bar{\mathbb{R}}^d,$$

and  $\mathbf{F}_n(\mathbf{x}) = (F_{n1}(1, x_1), \dots, F_{nd}(1, x_d))^\top$ , where

$$F_{nj}(s, x_j) = \frac{1}{n+1} \sum_{t=1}^{\lfloor ns \rfloor} \mathbb{I}(e_{n,jt} \leq x_j), \quad j \in \{1, \dots, d\}, \quad (s, \mathbf{x}) \in [0, 1] \times \bar{\mathbb{R}}^d. \quad (4)$$

For any  $j \in \{1, \dots, d\}$ , and  $(s, y) \in [0, 1] \times \bar{\mathbb{R}}$ , define

$$\mathbb{F}_{nj}(s, y) = \frac{1}{\sqrt{n}} \sum_{t=1}^{\lfloor ns \rfloor} \{\mathbb{I}(e_{n,jt} \leq y) - F_j(y)\} = \sqrt{n} \{F_{nj}(s, y) - sF_j(y)\} + o_P(1).$$

Convergence results for  $\mathbb{K}_n$  are expressed in terms of the (non-observable) auxiliary empirical processes

$$\check{\mathbb{K}}_n(s, \mathbf{x}) = \frac{1}{\sqrt{n}} \sum_{t=1}^{\lfloor ns \rfloor} \{\mathbb{I}(\boldsymbol{\varepsilon}_t \leq \mathbf{x}) - K(\mathbf{x})\}, \quad (s, \mathbf{x}) \in [0, 1] \times \bar{\mathbb{R}}^d,$$

$$\check{\mathbb{C}}_n(s, \mathbf{u}) = \frac{1}{\sqrt{n}} \sum_{t=1}^{\lfloor ns \rfloor} \{\mathbb{I}(\mathbf{U}_t \leq \mathbf{u}) - C(\mathbf{u})\}, \quad (s, \mathbf{u}) \in [0, 1]^{1+d},$$

and  $\check{\mathbb{C}}_{nj}(s, u_j) = \check{\mathbb{C}}_n(s, 1, \dots, 1, u_j, 1, \dots, 1) = \frac{1}{\sqrt{n}} \sum_{t=1}^{\lfloor ns \rfloor} \{\mathbb{I}(U_{jt} \leq u_j) - u_j\}$ ,  $j = 1, \dots, d$ .

According to Bickel and Wichura (1971),  $\check{\mathbb{K}}_n \rightsquigarrow \check{\mathbb{K}}$  where  $\check{\mathbb{K}}$  is a  $K$ -Kiefer process, i.e.,  $\check{\mathbb{K}}$  is a centered continuous Gaussian process with  $\text{Cov} \{\check{\mathbb{K}}(s, \mathbf{x}), \check{\mathbb{K}}(t, \mathbf{y})\} = (s \wedge t) \{K(\mathbf{x} \wedge \mathbf{y}) - K(\mathbf{x})K(\mathbf{y})\}$ ,  $s, t \in [0, 1]$  and  $\mathbf{x}, \mathbf{y} \in \bar{\mathbb{R}}^d$ . Here  $(\mathbf{x} \wedge \mathbf{y})_j = \min(x_j, y_j)$ ,  $j = 1, \dots, d$ . In particular,  $\check{\mathbb{K}}(1, \cdot)$  is a  $K$ -Brownian bridge. Similarly,  $\check{\mathbb{C}}_n \rightsquigarrow \check{\mathbb{C}}$ , where  $\check{\mathbb{C}}$  is a  $C$ -Kiefer process. In particular, for any  $j \in \{1, \dots, d\}$ ,  $\check{\mathbb{C}}^{(j)}(s, u_j) = \check{\mathbb{C}}(s, 1, \dots, 1, u_j, 1, \dots, 1)$  is a classical Kiefer process. Note that for all  $(s, \mathbf{x}) \in [0, 1] \times \bar{\mathbb{R}}^d$ ,  $\check{\mathbb{K}}(s, \mathbf{x}) = \check{\mathbb{C}}\{s, \mathbf{F}(\mathbf{x})\}$ .

The following assumptions are needed in order to prove the convergence results:

(A0) The conditional joint distribution function of  $\boldsymbol{\varepsilon}_t$  given  $\mathcal{F}_{t-1}$  is  $K_t$ , with continuous margins  $F_1, \dots, F_d$  independent of time. Furthermore, assume that  $\mathbf{G}_{\boldsymbol{\theta}, t}$  is continuously differentiable with respect to  $\boldsymbol{\theta} \in \mathcal{O} \subset \mathbb{R}^p$ , with derivative denoted by  $\dot{\mathbf{G}}_{\boldsymbol{\theta}, t}$ , and set  $\mathbf{G}_t = \mathbf{G}_{\boldsymbol{\theta}, t}$ ,  $\dot{\mathbf{G}}_t = \dot{\mathbf{G}}_{\boldsymbol{\theta}, t}$ ,  $\gamma_t(\mathbf{x}) = \dot{\mathbf{G}}_t \circ \mathbf{G}_t^{-1}(\mathbf{x})$ , and  $\dot{\mathbf{G}}_t^\top = \left( \dot{G}_{1t} \mid \dots \mid \dot{G}_{dt} \right)$ .

Let  $\mathbf{d}_{n,t} = \mathbf{e}_{n,t} - \boldsymbol{\varepsilon}_t - \gamma_t(\boldsymbol{\varepsilon}_t)\boldsymbol{\Theta}_n/\sqrt{n}$ , where  $\boldsymbol{\Theta}_n = n^{1/2}(\boldsymbol{\theta}_n - \boldsymbol{\theta})$ .

Next, assume that for any  $j \in \{1, \dots, d\}$ , and any  $\mathbf{x} \in \bar{\mathbb{R}}^d$ , the following properties hold:

(A1)  $\boldsymbol{\Gamma}_n(s, \mathbf{x}) = \frac{1}{n} \sum_{t=1}^{\lfloor ns \rfloor} \gamma_t(\mathbf{x}) \xrightarrow{Pr} s\boldsymbol{\Gamma}(\mathbf{x})$ , uniformly in  $s \in [0, 1]$ , where  $\boldsymbol{\Gamma}(\mathbf{x})$  is deterministic.

(A2)  $\frac{1}{n} \sum_{t=1}^n \mathbb{E}(\|\gamma_t(\boldsymbol{\varepsilon}_t)\|^k)$  is bounded, for  $k = 1, 2$ .

(A3) There exists a sequence of positive terms  $r_t > 0$  so that  $\sum_{t \geq 1} r_t < \infty$  and such that the sequence  $\max_{1 \leq t \leq n} \|\mathbf{d}_{n,t}\|/r_t$  is tight.

(A4)  $\max_{1 \leq t \leq n} \|\gamma_t(\boldsymbol{\varepsilon}_{jt})\|/\sqrt{n} = o_P(1)$ .

(A5)  $(\check{\mathbb{K}}_n, \Theta_n) \rightsquigarrow (\check{\mathbb{K}}, \Theta)$  in  $\mathcal{D}([-\infty, \infty]^d) \times \mathbb{R}^p$ .

(A6)  $\partial_{x_j} K(\mathbf{x})$  is bounded and continuous on  $\bar{\mathbb{R}}^d = [-\infty, +\infty]^d$ . In addition, for any  $j \in \{1, \dots, d\}$ ,  $F_j$  has a continuous bounded density  $f_j$ , and for any  $\delta > 0$ ,

$$\mathcal{L}_{n,j}(\delta) = \frac{1}{n} \sum_{t=1}^n \sup_{|y_j - x_j| < \delta} \|\gamma_{jt}(y_j) - \gamma_{jt}(x_j)\| \partial_{x_j} K(\mathbf{x}) = O_P(1).$$

One can now state the main convergence result. Its proof is given in [A.1](#).

**Theorem 1** *Under Case 1 and assumptions (A0)–(A6),  $(\mathbb{K}_n, \check{\mathbb{K}}_n, \Theta_n) \rightsquigarrow (\mathbb{K}, \check{\mathbb{K}}, \Theta)$ , with*

$$\mathbb{K}(s, \mathbf{x}) = \check{\mathbb{K}}(s, \mathbf{x}) - s \nabla K(\mathbf{x})^\top \Gamma(\mathbf{x}) \Theta, \quad (s, \mathbf{x}) \in [0, 1] \times \bar{\mathbb{R}}^d.$$

Furthermore, for all  $j \in \{1, \dots, d\}$ ,  $\mathbb{F}_{j,n} \rightsquigarrow \mathbb{F}_j$ , where

$$\mathbb{F}_j(s, x_j) = \check{\mathbb{C}}^{(j)}\{s, F_j(x_j)\} - s f_j(x_j) \Gamma_j(x_j)^\top \Theta, \quad (s, \mathbf{x}) \in [0, 1] \times \bar{\mathbb{R}}.$$

**Remark 1** *The univariate case studied in Bai (2003) corresponds to uniform margins here, so one obtains the same results in this case, namely that his empirical processes  $\hat{V}_{nj}(u_j) = \mathbb{F}_{nj}(1, u_j)$  converge in law to  $\mathbb{F}_j(1, u_j) = \check{\mathbb{C}}^{(j)}(1, u_j) - \Gamma_j(u_j)^\top \Theta$ ,  $u_j \in [0, 1]$ , for any  $j \in \{1, \dots, d\}$ . In addition, since the margins are all uniform,  $K$  is a copula and  $K_n$  corresponds to the so-called IFM estimator of the copula, as proposed in Xu (1996). Then Theorem 1 yields that the limiting empirical copula process in this case depends on the estimated parameter error  $\Theta$  of the marginal distributions. One will see later that taking the ranks solves this annoying problem.*

When dealing with time series, it is often important to check that the distribution of the series stay the same over a given time period. To test that there is no change point, define the sequential processes

$$\mathbb{A}_n(s, \mathbf{x}) = \frac{1}{\sqrt{n}} \sum_{t=1}^{\lfloor ns \rfloor} \{\mathbb{I}(e_{n,t} \leq \mathbf{x}) - K_n(\mathbf{x})\} = \mathbb{K}_n(s, \mathbf{x}) - \frac{\lfloor ns \rfloor}{n} \mathbb{K}_n(1, \mathbf{x}), \quad (s, \mathbf{x}) \in [0, 1] \times \bar{\mathbb{R}}^d, \quad (5)$$

$$\mathbb{B}_{nj}(s, y) = \frac{1}{\sqrt{n}} \sum_{t=1}^{\lfloor ns \rfloor} \{\mathbb{I}(e_{n,jt} \leq \mathbf{x}) - F_{nj}(y)\} = \mathbb{F}_{nj}(s, y) - \frac{\lfloor ns \rfloor}{n} \mathbb{F}_{nj}(1, y), \quad (s, y) \in [0, 1] \times \bar{\mathbb{R}}, \quad (6)$$

$j \in \{1, \dots, d\}$ . Many test statistics for detecting structural changes in the distribution of the generalized errors  $\varepsilon_t$  are functionals of  $\mathbb{A}_n$  or  $\mathbb{B}_{n1}, \dots, \mathbb{B}_{nd}$ . From (5), (6), and Theorem 1, one obtains the following interesting result, already obtained in Rémillard (2017) in the case of innovations.

**Corollary 1** *Under Case 1 and assumptions (A0)–(A6),  $\mathbb{A}_n \rightsquigarrow \mathbb{A}$ , and  $\mathbb{B}_{nj} \rightsquigarrow \mathbb{B}_j$ , with*

$$\mathbb{A}(s, \mathbf{x}) = \check{\mathbb{K}}(s, \mathbf{x}) - s \check{\mathbb{K}}(1, \mathbf{x}), \quad (s, \mathbf{x}) \in [0, 1] \times \bar{\mathbb{R}}^d,$$

$$\mathbb{B}_j(s, y) = \mathcal{K}_j\{s, F_j(y)\}, \quad (s, y) \in [0, 1] \times \bar{\mathbb{R}},$$

where  $\mathcal{K}_j(s, v) = \check{\mathbb{C}}^{(j)}(s, v) - s \check{\mathbb{C}}^{(j)}(1, v)$ ,  $j \in \{1, \dots, d\}$ ,  $v \in [0, 1]$ . In particular  $\mathbb{A}$ ,  $\mathbb{B}_1, \dots, \mathbb{B}_d$  do not depend on the estimated error  $\Theta$ . In addition,  $\text{Cov}\{\mathbb{A}(s, \mathbf{x}), \mathbb{A}(t, \mathbf{y})\} = (s \wedge t - st) \{K(\mathbf{x} \wedge \mathbf{y}) - K(\mathbf{x})K(\mathbf{y})\}$ ,  $s, t \in [0, 1]$  and  $\mathbf{x}, \mathbf{y} \in \bar{\mathbb{R}}^d$ , and  $\text{Cov}\{\mathcal{K}_j(s, v), \mathcal{K}_j(t, u)\} = (s \wedge t - st)(u \wedge v - uv)$ ,  $s, t, u, v \in [0, 1]$ .

**Remark 2** *Although the distribution of  $\mathbb{A}$  depends on the unknown distribution function  $K$ , it is still possible to bootstrap  $\mathbb{A}$ , i.e., to generate asymptotically independent copies of  $\mathbb{A}$ , making it possible to detect structural changes in the distribution of the generalized errors. See Nasri and Rémillard (2018) for details. However, in the one-dimensional case, tests statistics like*

$$\mathcal{T}_{nj} = \sup_{(s,y) \in [0,1] \times \bar{\mathbb{R}}} |\mathbb{B}_{nj}(s, y)|, \quad \mathcal{S}_{nj} = \sup_{s \in [0,1]} \int_{\bar{\mathbb{R}}} \mathbb{B}_{nj}^2(s, y) ds dF_{nj}(y), \quad j \in \{1, \dots, d\}, \quad (7)$$



converge in law to the distribution-free random variables

$$\mathcal{T}_j = \sup_{(s,v) \in [0,1]^2} |\mathcal{K}_j(s,v)|, \quad \mathcal{S}_j = \sup_{s \in [0,1]} \int_{\mathbb{R}} \mathcal{K}_j^2(s,v) ds dv,$$

whose law are already tabulated (Carlstein, 1988, Ghoudi and Rémillard, 2015).

**Remark 3** As a particular case, assume the same setting as in Bai (2003), i.e., the margins are all uniform. This corresponds to an extension of the so-called IFM method (Xu, 1996), who only considered the time-independent case. Then the empirical copula and the sequential empirical process are respectively  $K_n(\mathbf{1}, \mathbf{u})$  and  $K_n(s, \mathbf{u})$ ,  $(s, \mathbf{u}) \in [0, 1]^{1+d}$ , and the limiting process depends on the estimated parameter error  $\Theta$  of the marginal distributions. However, as we will see next taking the ranks makes this nuisance vector disappear.

## 2.1 Empirical processes related to the copula

Since one of the main results of this paper is to model the dependence between time series, we have to deal with the (unique) copula  $C$  associated with  $K$ . Since the copula is independent of the margins, one way to estimate it is to remove their effect by replacing  $\mathbf{e}_{n,t}$  with the associated rank vectors

$$\mathbf{U}_{n,t} = (U_{n,1t}, \dots, U_{n,dt})^\top, \quad U_{n,jt} = \text{Rank}(e_{n,jt})/(n+1), \quad t \in \{1, \dots, n\},$$

where  $\text{Rank}(e_{n,jt})$  being the rank of  $e_{n,jt}$  amongst  $e_{n,j1}, \dots, e_{n,jn}$ ,  $j \in \{1, \dots, d\}$ . This can also be written as  $\mathbf{U}_{n,t} = \mathbf{F}_n(\mathbf{e}_{n,t})$ ,  $t \in \{1, \dots, n\}$ .

Next, define the empirical copula

$$C_n(\mathbf{u}) = \frac{1}{n} \sum_{t=1}^n \mathbb{I}(\mathbf{U}_{n,t} \leq \mathbf{u}), \quad \mathbf{u} \in [0, 1]^d,$$

together with the sequential copula process

$$\mathbb{C}_n(s, \mathbf{u}) = \frac{1}{\sqrt{n}} \sum_{t=1}^{\lfloor ns \rfloor} \{\mathbb{I}(\mathbf{U}_{n,t} \leq \mathbf{u}) - C(\mathbf{u})\}, \quad (s, \mathbf{u}) \in [0, 1]^{1+d},$$

and set

$$\mathbb{G}_n(s, \mathbf{u}) = \frac{1}{\sqrt{n}} \sum_{t=1}^{\lfloor ns \rfloor} \{\mathbb{I}(\mathbf{U}_{n,t} \leq \mathbf{u}) - C_n(\mathbf{u})\} = \mathbb{C}_n(s, \mathbf{u}) - \frac{\lfloor ns \rfloor}{n} C_n(\mathbf{1}, \mathbf{u}).$$

In order to work on the space of continuous functions on  $[0, 1]^d$ , one assumes from now on the following additional technical assumption (Segers, 2012, Condition 2.1) on the partial derivatives of  $C$ .

**Condition 1** For each  $j \in \{1, \dots, d\}$ , the  $j$ -th first-order partial derivative  $\partial_{u_j} C$  exists and is continuous on  $\{u \in [0, 1]^d; 0 < u_j < 1\}$ .

The next result follows directly from Theorem 1, Condition 1, using the results in Bücher and Volgushev (2013), Bücher and Kojadinovic (2016), and the representation of  $\mathbb{G}_n$ .

**Corollary 2** Under Case 1, Condition 1, and assumptions (A0)–(A6),  $\mathbb{C}_n \rightsquigarrow \mathbb{C}$ , with

$$\mathbb{C}(s, \mathbf{u}) = \check{\mathbb{C}}(s, \mathbf{u}) - s \sum_{j=1}^d \partial_{u_j} C(\mathbf{u}) \check{\mathbb{C}}^{(j)}(1, u_j), \quad (s, \mathbf{u}) \in [0, 1]^{1+d}. \quad (8)$$

Moreover,  $\mathbb{G}_n \rightsquigarrow \mathbb{G}$ , where  $\mathbb{G}(s, \mathbf{u}) = \check{\mathbb{C}}(s, \mathbf{u}) - s \check{\mathbb{C}}(1, \mathbf{u})$ ,  $(s, \mathbf{u}) \in [0, 1]^{1+d}$ .

Note that  $\mathbb{C}$  is the limiting sequential copula process of the generalized errors  $\varepsilon_t$ , even though the latter are not observed. This result generalizes the one in Rémillard (2017) which was applied to innovations of stochastic volatility models. Furthermore, taking the ranks eliminates the dependency on the estimated parameters of the margin, in contrast with the IFM method (see Remark 1)

An immediate application of Corollary 2 is that tests for detecting structural change in the copula of the innovations can be based on the process  $\mathbb{G}_n$  and that the limiting process  $\mathbb{G}$  is parameter free, depending only on the unknown copula  $C$ . However, as it was also true for  $\mathbb{A}$ , it is easy to simulate asymptotically independent copies of  $\mathbb{G}$ . See Nasri and Rémillard (2018). Therefore change point tests can be based on  $\mathbb{G}_n$ . For example, one could define

$$\mathcal{T}_n = \max_{1 \leq k \leq n} \max_{1 \leq t \leq n} |\mathbb{G}_n(k/n, \mathbf{U}_{n,t})| \quad (9)$$

and reject the null hypothesis for large values of  $\mathcal{T}_n$ . The limiting distribution of  $\mathcal{T}_n$  is given next.

**Proposition 1** *Under Case 1 and assumptions (A0)–(A6),  $\mathcal{T}_n \rightsquigarrow \mathcal{T}$ , where  $\mathcal{T} = \sup_{s \in [0,1], \mathbf{u} \in [0,1]^d} |\mathbb{G}(s, \mathbf{u})|$ .*

**Remark 4** *Note that  $\max_{1 \leq k \leq n} \max_{1 \leq t \leq n} |\mathbb{A}_n(k/n, \mathbf{e}_{n,t})| \rightsquigarrow \sup_{s \in [0,1], \mathbf{x} \in \mathbb{R}^d} |\mathbb{A}(s, \mathbf{x})| = \mathcal{T}$ , since for all  $(s, \mathbf{x}) \in [0,1] \times \mathbb{R}^d$ ,  $\check{\mathbb{K}}(s, \mathbf{x}) = \check{C}\{s, \mathbf{F}(\mathbf{x})\}$ .*

### 3 Weak convergence of empirical processes in the time-dependent case

For the time-dependent case (Case 2), one considers two settings: the first one is a time-dependent parametric family setting, while the second one is a regime-switching model for the copula. In both cases, under reasonable assumptions, the asymptotic error of the copula parameter does not depend of the error term  $\Theta$ . The results proved here are somewhat an extension of those of Nasri et al. (2017), where independent observations were considered.

#### 3.1 Parametric time-dependent setting

Here, we consider the same parametric model as in Nasri et al. (2017), namely  $c_{\psi,t} = c_{\phi(\psi,t)}$ , with  $\phi(\psi, t) = \mathbf{h}_t \psi$ , for a given matrix function  $\mathbf{h}_t$  and  $\psi$  belongs to the parameter space denoted  $\mathcal{P}$ . It follows that the log-likelihood of  $\mathbf{U}_1, \dots, \mathbf{U}_n$  at  $\mathbf{u}_1, \dots, \mathbf{u}_n$ , is

$$\sum_{t=1}^n \log c_{\psi,t}(\mathbf{u}_1, \dots, \mathbf{u}_n) = \sum_{t=1}^n \log c_{\phi(\psi,t)}(\mathbf{u}_t).$$

As a result, the pseudo-maximum likelihood estimator  $\psi_n$  of the real parameter  $\psi_0$ , is defined by

$$\psi_n = \arg \max_{\psi \in \mathcal{P}} \sum_{t=1}^n \log c_{\phi(\psi,t)}(\mathbf{u}_{n,t}).$$

Each copula  $C_\phi$  is assumed to admit a density  $c_\phi$  satisfying the following assumptions:

- (B1) For every  $\phi \in \mathcal{P}$ , the density  $c_\phi$  of  $C_\phi$  admits first and second order derivatives with respect to all components of  $\phi$ . The gradient (column) vector with respect to  $\phi$  is denoted  $\dot{c}_\phi$ , and the Hessian matrix is represented by  $\ddot{c}_\phi$ .
- (B2) For arbitrary  $\mathbf{u} \in (0,1)^d$  and every  $\phi_0 \in \mathcal{P}$ , the mappings  $\phi \mapsto \dot{c}_\phi(\mathbf{u})/c_\phi(\mathbf{u})$  and  $\phi \mapsto \ddot{c}_\phi(\mathbf{u})/c_\phi(\mathbf{u})$  are continuous at  $\phi_0$ .
- (B3) For every  $\phi_0 \in \mathcal{P}$ , there exist a neighborhood  $\mathcal{N}$  of  $\phi_0$  and  $C_{\phi_0}$ -integrable functions  $h_1, h_2 : \mathbb{R}^d \rightarrow \mathbb{R}$  such that for every  $\mathbf{u} \in (0,1)^d$ ,

$$\sup_{\phi \in \mathcal{N}} \left\| \frac{\dot{c}_\phi(\mathbf{u})}{c_\phi(\mathbf{u})} \right\| \leq h_1^{1/2}(\mathbf{u}) \quad \text{and} \quad \sup_{\phi \in \mathcal{N}} \left\| \frac{\ddot{c}_\phi(\mathbf{u})}{c_\phi(\mathbf{u})} \right\| \leq h_2(\mathbf{u}).$$

(B4) As in Genest and Rémillard (2008), assume, for identifiability purposes, that for every  $\delta > 0$ ,

$$\inf \left\{ \sup_{\mathbf{u} \in [0,1]^d} |C_\phi(\mathbf{u}) - C_{\phi_0}(\mathbf{u})| : \phi \in \mathcal{P} \text{ and } |\phi - \phi_0| > \delta \right\} > 0.$$

Furthermore, the mapping  $\phi \mapsto C_\phi$  is assumed to be Fréchet differentiable with derivative  $\dot{C}$ , i.e., for all  $\phi_0 \in \mathcal{P}$ ,

$$\lim_{\mathbf{h} \rightarrow 0} \sup_{\mathbf{u} \in (0,1)^d} \frac{|C_{\phi_0+\mathbf{h}}(\mathbf{u}) - C_{\phi_0}(\mathbf{u}) - \dot{C}(\mathbf{u})\mathbf{h}|}{r(\mathbf{u})\|\mathbf{h}\|} = 0, \quad (10)$$

for some function  $r$  such that  $\inf_{\mathbf{u} \in (0,1)^d} r(\mathbf{u}) > 0$  and  $E\{r(\mathbf{U}_1)\} < \infty$ .

For simplicity, set  $c_t = c_{\phi(\psi_0, t)}$ , and define

$$\mathcal{W}_n = n^{-1/2} \sum_{t=1}^n \frac{\mathbf{h}_t^\top \dot{c}_t(\mathbf{U}_t)}{c_t(\mathbf{U}_t)}, \quad (11)$$

$$\mathcal{I}_n = n^{-1} \sum_{t=1}^n \frac{\mathbf{h}_t^\top \dot{c}_t(\mathbf{U}_t) \dot{c}_t(\mathbf{U}_t)^\top \mathbf{h}_t}{c_t^2(\mathbf{U}_t)}. \quad (12)$$

For any  $j \in \{1, \dots, d\}$ , and any  $t \in \{1, \dots, n\}$ , set  $\zeta_{jt} = \mathbf{h}_t^\top \left\{ \frac{\partial_{u_j} \dot{c}_t(\mathbf{U}_t)}{c_t(\mathbf{U}_t)} - \frac{\dot{c}_t(\mathbf{U}_t) \partial_{u_j} c_t(\mathbf{U}_t)}{c_t^2(\mathbf{U}_t)} \right\}$ , and define  $\mathbf{J}_n = (\mathbf{J}_{n,1}, \dots, \mathbf{J}_{n,d})^\top$ , where  $\mathbf{J}_{n,j}(u_j) = n^{-1} \sum_{t=1}^n E\{\zeta_{jt} | U_{jt} = u_j\}$ . Further set

$$\mathbf{L}_n = n^{-1} \sum_{j=1}^d \sum_{t=1}^n \zeta_{jt} f_j(\varepsilon_{jt}) \{\gamma_{jt}(\varepsilon_{jt}) - \Gamma_j(\varepsilon_{jt})\}.$$

**Theorem 2** Suppose that Assumptions (A0)–(A6) and (B1)–(B4) hold. Suppose also that Lindeberg's condition is satisfied for the sequence  $\frac{\mathbf{h}_t^\top \dot{c}_t(\mathbf{U}_t)}{c_t(\mathbf{U}_t)}$ ,  $t \in \{1, \dots, n\}$ . Assume that  $(\mathbf{J}_n, \mathcal{I}_n)$  converges in probability to  $(\mathbf{J}, \mathcal{I})$ , with  $\mathcal{I}$  invertible, while  $\mathbf{L}_n$ ,  $n^{-1} \sum_{t=1}^n \frac{\mathbf{h}_t^\top \ddot{c}_t(\mathbf{U}_t) \mathbf{h}_t}{c_t(\mathbf{U}_t)}$ , and  $n^{-2} \sum_{t=1}^n \sum_{j=1}^d E\{\zeta_{jt}^2\}$  converge in probability to 0. Then, as  $n \rightarrow \infty$ ,  $(\check{\mathcal{C}}_n^{(1)}, \dots, \check{\mathcal{C}}_n^{(d)}, \mathcal{W}_n, \Psi_n)$  converges in law to  $(\check{\mathcal{C}}^{(1)}, \dots, \check{\mathcal{C}}^{(d)}, \mathcal{W}, \Psi)$ , where  $\mathcal{W} \sim N(0, \mathcal{I})$  and

$$\Psi = \mathcal{I}^{-1} \left\{ \mathcal{W} + \sum_{j=1}^d \int_0^1 \mathbf{J}_j(u_j) \check{\mathcal{C}}^{(j)}(1, u_j) du_j \right\}.$$

Moreover,  $(\check{\mathcal{C}}^{(1)}, \dots, \check{\mathcal{C}}^{(d)})$  and  $\mathcal{W}$  are independent Gaussian processes, and  $E(\Psi \mathcal{W}^\top) = I$ .

The proof of the theorem is given in A.2. Note that the time-independent case corresponds to  $\mathbf{h}_t \equiv 1$ , so Theorem 2 extends the results of Chen and Fan (2006a) obtained in the case of innovations. When  $\mathbf{h}_t \equiv 1$ , the conditions of Theorem 2 are much easier to verify.

**Remark 5** It makes sense to assume that  $\mathbf{L}_n$  converges in probability to 0 since  $\gamma_t$  is  $\mathcal{F}_{t-1}$ -measurable with mean very close to  $\Gamma$  and both are independent of  $\varepsilon_t$ . In particular, this is true if  $\gamma_t$  is stationary and square integrable, as in the case of innovations.

**Remark 6** Note that it is quite easy to consider nonlinear parametric models as well, i.e.,  $\phi_t = \phi(\mathbf{h}_t, \psi)$ , for a known function  $\phi$ . The only changes would be to replace  $\mathbf{h}_t$  with the gradient of  $\phi$  whenever required.

## 3.2 HMM setting

Here,  $\tau_t$  is a finite Markov chain on  $\{1, \dots, \ell\}$  with transition matrix  $Q$ , and given  $\tau_1 = j_1, \dots, \tau_n = j_n$ , the conditional distribution of  $\mathbf{U}_1, \dots, \mathbf{U}_n$  at  $\mathbf{u}_1, \dots, \mathbf{u}_n$  is

$$\prod_{t=1}^n c_\alpha^{(j_t)}(\mathbf{u}_t),$$

where for any  $j \in \{1, \dots, \ell\}$ ,  $c_{\alpha}^{(j)}(\cdot)$  is a copula family with parameter  $\alpha \in \mathcal{A}$ . To simplify the notations, let  $\psi = (\alpha, Q)$  be the vector of parameters. It follows that for any  $t \geq 2$ , the conditional density  $c_{\psi,t}$  of  $\mathbf{U}_t$  at  $\mathbf{u}_t$ , given  $\mathbf{U}_{t-1} = \mathbf{u}_{t-1}, \dots, \mathbf{U}_1 = \mathbf{u}_1$ , is expressed as a mixture viz.

$$c_{\psi,t}(\mathbf{u}_t) = \sum_{i=1}^l c_{\alpha}^{(i)}(\mathbf{u}_t) \left\{ \sum_{j=1}^l \eta_{\psi,t-1}(j) Q_{ji} \right\} = \sum_{i=1}^l c_{\alpha}^{(i)}(\mathbf{u}_t) W_{\Psi,t-1}(i), \quad (13)$$

where  $W_{\Psi,t-1}(i) = \sum_{j=1}^l \eta_{\Psi,t-1}(j) Q_{ji}$ ,  $i \in \{1, \dots, l\}$ , and

$$\eta_{\Psi,t}(i) = \frac{c_{\alpha}^{(i)}(\mathbf{u}_t)}{Z_{t|t-1}} \sum_{j=1}^l \eta_{\Psi,t-1}(j) Q_{ji}, \quad t \geq 1. \quad (14)$$

The pseudo-maximum likelihood estimator  $\psi_n$  of the unknown parameter  $\psi_0$  is defined by

$$\psi_n = \arg \max_{\psi \in \mathcal{P}} \sum_{t=2}^n \log c_{\psi,t}(\mathbf{u}_{n,t}).$$

Assume now that for any  $i \in \{1, \dots, \ell\}$ , the copula  $C_{\alpha}^{(i)}$  satisfies Assumptions (B1)–(B4). For simplicity, set  $c_t = c_{\phi(\psi_0,t)}$ , and define

$$\mathcal{W}_n = n^{-1/2} \sum_{t=1}^n \frac{\dot{c}_t(\mathbf{U}_t)}{c_t(\mathbf{U}_t)}, \quad (15)$$

$$\mathcal{I}_n = n^{-1} \sum_{t=1}^n \frac{\dot{c}_t(\mathbf{U}_t) \dot{c}_t(\mathbf{U}_t)^{\top}}{c_t^2(\mathbf{U}_t)}. \quad (16)$$

For any  $j \in \{1, \dots, d\}$ , and any  $t \in \{1, \dots, n\}$ , set  $\zeta_{jt} = \left\{ \frac{\partial_{u_j} \dot{c}_t(\mathbf{U}_t)}{c_t(\mathbf{U}_t)} - \frac{\dot{c}_t(\mathbf{U}_t) \partial_{u_j} c_t(\mathbf{U}_t)}{c_t^2(\mathbf{U}_t)} \right\}$ , and define  $\mathbf{J}_n = (\mathbf{J}_{n,1}, \dots, \mathbf{J}_{n,d})^{\top}$ , where  $\mathbf{J}_{n,j}(u_j) = n^{-1} \sum_{t=1}^n E(\zeta_{jt} | U_{jt} = u_j)$ . Further set

$$\mathbf{L}_n = n^{-1} \sum_{j=1}^d \sum_{t=1}^n \zeta_{jt} f_j(\varepsilon_{jt}) \{ \gamma_{jt}(\varepsilon_{jt}) - \Gamma_j(\varepsilon_{jt}) \}.$$

**Theorem 3** *Suppose that Assumptions (A0)–(A6) and (B1)–(B4) hold. Suppose also that  $(\mathbf{J}_n, \mathcal{I}_n)$  converges in probability to  $(\mathbf{J}, \mathcal{I})$ , with  $\mathcal{I}$  invertible, while  $\mathbf{L}_n$  converges in probability to 0. Then, as  $n \rightarrow \infty$ ,  $(\check{\mathcal{C}}_n^{(1)}, \dots, \check{\mathcal{C}}_n^{(d)}, \mathcal{W}_n, \Psi_n)$  converges in law to  $(\check{\mathcal{C}}^{(1)}, \dots, \check{\mathcal{C}}^{(d)}, \mathcal{W}, \Psi)$ , where  $\mathcal{W} \sim N(0, \mathcal{I})$  and*

$$\Psi = \mathcal{I}^{-1} \left\{ \mathcal{W} + \sum_{j=1}^d \int_0^1 \mathbf{J}_j(u_j) \check{\mathcal{C}}^{(j)}(1, u_j) du_j \right\}.$$

Moreover,  $(\check{\mathcal{C}}^{(1)}, \dots, \check{\mathcal{C}}^{(d)})$  and  $\mathcal{W}$  are independent Gaussian processes, and  $E(\Psi \mathcal{W}^{\top}) = \mathbf{I}$ .

The proof of the results are given in [A.2.1](#)

### 3.3 Tests statistics based on the Rosenblatt transform

Instead of using the empirical copula process which does not make sense for time-varying copulas, we will use goodness-of-fit tests constructed from the Rosenblatt transform (Rosenblatt, 1952). Recall that the Rosenblatt transform of a  $d$ -dimensional copula  $C$  is the mapping  $\mathcal{R}$  from  $(0, 1)^d \rightarrow (0, 1)^d$  so that  $\mathbf{u} = (u_1, \dots, u_d) \mapsto \mathcal{R}(\mathbf{u}) = (v_1, \dots, v_d)$  with  $v_1 = u_1$  and

$$v_k = \frac{\partial^{k-1} C(u_1, \dots, u_k, 1, \dots, 1)}{\partial u_1 \cdots \partial u_{k-1}} \bigg/ \frac{\partial^{k-1} C(u_1, \dots, u_{k-1}, 1, \dots, 1)}{\partial u_1 \cdots \partial u_{k-1}},$$

$k \in \{2, \dots, d\}$ . Rosenblatt transforms for Archimedean copulas and meta-elliptic copulas are quite easy to compute for any dimension; see, e.g., Rémillard et al. (2012). Based on Genest et al. (2009), goodness-of-fit tests based on the Rosenblatt transform were among the most powerful omnibus tests.

Going back to our setting, let  $C_t$  be the copula of  $\mathbf{U}_t$ , with associated Rosenblatt transform  $\mathcal{R}_t$ ,  $t \in \{1, \dots, n\}$ . Then  $\mathbf{V}_t = \mathcal{R}_t(\mathbf{U}_t) \sim \Pi$  if and only if  $\mathbf{U}_t \sim C_t$ ,  $t \in \{1, \dots, n\}$ . It follows that the null hypothesis  $\mathcal{H}_0 : C_t \in \{C_{\phi(t)}; \phi(t) \in \mathcal{P}\}$ ,  $t \in \{1, \dots, n\}$ , can be restated as follows:

$$\mathcal{H}_0 : \mathcal{R}_t \in \{\mathcal{R}_{\phi(t)}; \phi(t) \in \mathcal{P}\}, \quad t \in \{1, \dots, n\}.$$

One can build tests for  $\mathcal{H}_0$  by comparing the empirical distribution function of  $\mathbf{V}_{n,t} = \mathcal{R}_{\phi_n(t)}(\mathbf{U}_{n,t})$ ,  $t \in \{1, \dots, n\}$ , with  $\Pi$ , since under  $\mathcal{H}_0$ ,  $\mathbf{V}_{n,t}$  has approximately distribution  $\Pi$ . More precisely, set

$$\mathbb{D}_n(\mathbf{u}) = \frac{1}{\sqrt{n}} \sum_{t=1}^n \{\mathbb{I}(\mathbf{V}_{n,t} \leq \mathbf{u}) - \Pi(\mathbf{u})\}, \quad \mathbf{u} \in [0, 1]^d, \quad (17)$$

and define

$$S_n^{(B)} = \int_{[0,1]^d} \mathbb{D}_n^2(\mathbf{u}) d\mathbf{u} = \frac{n}{3^d} - \frac{1}{2^{d-1}} \sum_{t=1}^n \prod_{k=1}^d (1 - V_{n,kt}^2) + \frac{1}{n} \sum_{t=1}^n \sum_{j=1}^n \prod_{k=1}^d (1 - V_{n,kt} \vee V_{n,kj}), \quad (18)$$

where  $a \vee b = \max(a, b)$ . Further set  $\check{\mathbb{D}}_n(\mathbf{u}) = \frac{1}{\sqrt{n}} \sum_{t=1}^n \{\mathbb{I}(\mathbf{V}_t \leq \mathbf{u}) - \Pi(\mathbf{u})\}$ ,  $\mathbf{u} \in [0, 1]^d$ . It is easy to check that  $(\check{\mathbb{D}}_n, \mathcal{W}_n) \rightsquigarrow (\check{\mathbb{D}}, \mathcal{W})$ , where the joint law is Gaussian, and  $\check{\mathbb{D}}$  is a  $\Pi$ -Brownian bridge.

As in the case of copula processes studied in the previous section, in order to prove the next result, one must assume the following assumption:

(R1)  $\mathcal{R}_{\Psi,t}$  is Fréchet differentiable, i.e.,

$$\lim_{h \rightarrow 0} \sup_{\mathbf{u} \in (0,1)^d} \frac{|\mathcal{R}_{\psi_0+h,t}(\mathbf{u}) - \mathcal{R}_{\psi_0,t}(\mathbf{u}) - \dot{\mathcal{R}}_t(\mathbf{u})h|}{r(\mathbf{u})\|h\|} = 0, \quad (19)$$

for some function  $r$  such that  $\inf_{\mathbf{u} \in (0,1)^d} r(\mathbf{u}) > 0$  and  $E\{r(\mathbf{U}_1)\} < \infty$ .

One can now state the main result of the section, giving the convergence of the empirical Rosenblatt process  $\mathbb{D}_n$ . But first, recall that  $\mathring{\mathbf{U}}_{n,t} = \mathbf{R}_t/(n+1)$ , where  $\mathbf{R}_1, \dots, \mathbf{R}_n$  are the associated rank vectors of  $\mathbf{U}_1, \dots, \mathbf{U}_n$ , and let  $\mathring{\mathbf{V}}_{n,t} = \mathcal{R}_{\mathring{\phi}_n(t)}(\mathring{\mathbf{U}}_t)$ , where  $\mathring{\phi}_n(t)$  is the estimator of  $\phi(t)$  calculated with  $\mathring{\mathbf{U}}_{n,t} = \mathbf{R}_t/(n+1)$ ,  $t \in \{1, \dots, n\}$ . Further set

$$\mathring{\mathbb{D}}_n(\mathbf{u}) = \frac{1}{\sqrt{n}} \sum_{t=1}^n \left\{ \mathbb{I}(\mathring{\mathbf{V}}_{n,t} \leq \mathbf{u}) - \Pi(\mathbf{u}) \right\}, \quad \mathbf{u} \in [0, 1]^d. \quad (20)$$

Next, let  $\mathcal{M}$  be the class of functions  $\mathbf{m} = (m_1, \dots, m_d)$ , where for any  $j \in \{1, \dots, d\}$ ,  $m_j$  is continuous on  $[0, 1]$  with  $m_j(0) = m_j(1) = 0$ . For any  $\mathbf{m} \in \mathcal{M}$  and  $\mathbf{u} \in [0, 1]^d$ , set

$$\begin{aligned} \boldsymbol{\kappa}_n(\mathbf{m}, \mathbf{u}) &= \frac{1}{n} \sum_{t=1}^n \sum_{j=1}^d \sum_{k=1}^j E[\mathbb{I}(\mathbf{V}_t \leq \mathbf{u}) \partial_{u_k} \mathcal{R}_{jt}(\mathbf{U}_t) m_k(\varepsilon_{kt}) | V_{jt} = u_j], \\ \boldsymbol{\varrho}_n(\mathbf{u}) &= \frac{1}{n} \sum_{t=1}^n \sum_{j=1}^d \mathbf{h}_t^\top E[\mathbb{I}(\mathbf{V}_t \leq \mathbf{u}) \dot{\mathcal{R}}_{jt}(\mathbf{U}_t) | V_{jt} = u_j], \\ \boldsymbol{\delta}_n(\mathbf{u}) &= \frac{1}{n} \sum_{t=1}^n \sum_{j=1}^d E[\mathbb{I}(\mathbf{V}_t \leq \mathbf{u}) \mathcal{L}_{n,jt} | V_{jt} = u_j]. \end{aligned}$$

**Theorem 4** Suppose that the assumptions (A0)–(A6), (B1)–(B4), and (R1) are met. Suppose also that  $\tilde{\gamma}_t = \mathbf{h}_t^\top \tilde{\mathbf{R}}_t \circ \mathbf{R}_t^{-1}$  satisfies Assumptions (A0)–(A6), with  $\gamma$  replaced by  $\tilde{\gamma}$  and  $\varepsilon_t$  replaced by  $\mathbf{U}_t$ ,  $t \in \{1, \dots, n\}$ . Further assume that uniformly in any  $(\mathbf{m}, \mathbf{u}) \in \mathcal{M} \times [0, 1]^d$ ,  $\kappa_n(\mathbf{m}, \mathbf{u})$  converges in probability to  $\kappa(\mathbf{m}, \mathbf{u})$ ,  $\mathbf{g}_n(\mathbf{u})$  converges in probability to  $\mathbf{g}(\mathbf{u})$ , and  $\delta_n(\mathbf{u})$  converges in probability to 0. Under Case 2,  $\mathbb{D}_n - \overset{\circ}{\mathbb{D}}_n \rightsquigarrow 0$  and  $\overset{\circ}{\mathbb{D}}_n \rightsquigarrow \overset{\circ}{\mathbb{D}}$ , with  $\overset{\circ}{\mathbb{D}}$  given by

$$\overset{\circ}{\mathbb{D}}(\mathbf{u}) = \check{\mathbb{D}}(\mathbf{u}) - \check{\kappa}(\mathbf{u}) - \check{\Phi}^\top \mathbf{g}(\mathbf{u}),$$

where  $\check{\mathbb{D}}$  is a  $\Pi$ -Brownian bridge,  $\check{\kappa}(\mathbf{u}) = \kappa \{ \check{C}^{(1)}(1, \cdot), \dots, \check{C}^{(d)}(1, \cdot), \mathbf{u} \}$ ,  $\mathbf{u} \in [0, 1]^d$ ,  $E\{\check{\mathbb{D}}(\mathbf{u})\mathcal{W}\} = \mathbf{g}(\mathbf{u})$ , and  $E\{\check{\kappa}(\mathbf{u})\mathcal{W}\} = 0$ . In particular  $E\left(\overset{\circ}{\mathbb{D}}\mathcal{W}\right) \equiv 0$ .

The proof of Theorem 4 is given in A.3. The following result is then a direct application of the continuous mapping theorem.

**Proposition 2** Under the assumptions of Theorem 4,  $S_n^{(B)} \rightsquigarrow S^{(B)} = \int_{[0,1]^d} \left\{ \overset{\circ}{\mathbb{D}}(\mathbf{u}) \right\}^2 d\mathbf{u}$ . In fact, if  $\mathcal{H}$  is a continuous function on the space  $C([0, 1])$ , then the statistic  $T_n = \mathcal{H}(\mathbb{D}_n)$  converges in law to  $T = \mathcal{H}\left(\overset{\circ}{\mathbb{D}}\right)$ .

### 3.4 Parametric bootstrap

In order to test for goodness-of-fit using statistics based on  $\mathbb{D}_n$ , one needs a way to compute  $P$ -values. To do so, we propose the parametric bootstrap algorithm described next. It can be used to estimate  $P$ -values of  $S_n^{(B)}$  or any continuous functional of  $\mathbb{D}_n$ . Before stating the main result of the section, one needs to extend the notion of regularity of  $\phi_n$  as defined in Genest and Rémillard (2008) and extended in Nasri et al. (2017).

$\psi_n$  is regular for  $\psi$  if  $(\check{\mathbb{K}}_n, \mathcal{W}_n, \check{\Phi}_n) \rightsquigarrow (\check{\mathbb{K}}, \mathcal{W}, \check{\Psi})$  where the latter is centered Gaussian with  $E(\check{\Psi}\mathcal{W}^\top) = I$ , and  $\check{\Psi}$  does not depend on  $\theta$  or  $\Theta$ . As shown in Genest and Rémillard (2008) and Nasri et al. (2017), regularity of estimators is a sufficient condition for the parametric bootstrap to work. The following algorithm is described in terms of the statistic  $S_n^{(B)}$  but can be applied easily to any statistic of the form  $T_n = \psi(\mathbb{D}_n)$  as in Proposition 2.

**Algorithm 1 (Parametric bootstrap for the empirical Rosenblatt process)** For some large integer  $N$ , do the following steps:

1. Estimate  $\psi$  by  $\psi_n = \mathcal{T}_n(\mathbf{U}_{n,1}, \dots, \mathbf{U}_{n,n})$ , compute  $\mathbb{D}_n$  and  $S_n^{(B)}$  according to (17) and (18).
2. For some large integer  $N$ , repeat the following steps for every  $k \in \{1, \dots, N\}$ :
  - (a) Generate a random sample  $\mathbf{Y}_{n,1}^{(k)}, \dots, \mathbf{Y}_{n,n}^{(k)}$  from distribution  $C_{\psi_{n,t}}$ ,  $t \in \{1, \dots, n\}$ , and compute the pseudo-observations  $\mathbf{U}_{n,t}^{(k)} = \mathbf{R}_{n,t}^{(k)} / (n+1)$ , where  $\mathbf{R}_{n,1}^{(k)}, \dots, \mathbf{R}_{n,n}^{(k)}$  are the associated rank vectors of  $\mathbf{Y}_{n,1}^{(k)}, \dots, \mathbf{Y}_{n,n}^{(k)}$ .
  - (b) Estimate  $\psi$  by  $\psi_n^{(k)} = \mathcal{T}_n(\mathbf{U}_{n,1}^{(k)}, \dots, \mathbf{U}_{n,n}^{(k)})$ , define  $\mathbf{V}_{n,t}^{(k)} = \mathcal{R}_{\psi_n^{(k)}, t}(\mathbf{U}_{n,t}^{(k)})$ ,  $i \in \{1, \dots, n\}$ , and compute  $\mathbb{D}_n^{(k)}(\mathbf{u}) = \frac{1}{\sqrt{n}} \sum_{t=1}^n \left\{ \mathbb{I}(\mathbf{V}_{n,t}^{(k)} \leq \mathbf{u}) - \Pi(\mathbf{u}) \right\}$ ,  $\mathbf{u} \in [0, 1]^d$ .
  - (c) Calculate  $S_{n,k}^{(B)} = \int_{[0,1]^d} \left\{ \mathbb{D}_n^{(k)}(\mathbf{u}) \right\}^2 d\mathbf{u}$ .

An approximate  $P$ -value for the test is then given by  $\sum_{k=1}^N \mathbb{I}(S_{n,k}^{(B)} > S_n^{(B)}) / N$ .

Using Theorem 2 or Theorem 3, together with Theorem 4, one may conclude that the pseudo-maximum estimator  $\psi_n$  is a regular estimator, and so that Algorithm 1 works.

## 4 Simulation study

In this section we consider two Monte Carlo experiments for assessing the level and power of the proposed goodness-of-fit tests based on the Cramér-von Mises type statistics defined in Section 3.3. We generated random samples of size  $n \in \{250, 500\}$  from four bivariate copula families: Clayton, Gumbel, Gaussian and Student (with  $\nu = 5$ ). In each experiment, we considered the “linear” case, i.e.,  $\tau_t = 1 / \left(1 + e^{-H_t^{(1)}\theta^{(1)}}\right)$ , where  $H_t^{(1)}\theta^{(1)} = .4055t/n$ , so that  $\tau_1 = .5004$  and  $\tau_n = 0.6$ . The same four families were used under the null hypothesis. The dynamic model for  $X_1$  is a GARCH(1,1) model with parameters  $\mu = .05$ ,  $\omega = 0.1$ ,  $\alpha = .05$ ,  $\beta = .94$ , and Gaussian innovations. For  $X_2$ , we assumed a Gaussian HMM with two regimes, with means  $-0.1, 0.8$ , standard deviations  $0.5, 0.1$ , and transition matrix  $Q = \begin{pmatrix} 0.7 & 0.3 \\ 0.05 & 0.95 \end{pmatrix}$ . For each experiment, we performed 1000 replications and in each replication, 100 bootstrap samples ( $N = 100$ ) were used to compute the  $P$ -value of the test statistic. The results of the Monte Carlo experiments are displayed in Table 1. One can see that the levels are not significantly different from the target level of 5%. Also the power of the goodness-of-fit test is satisfactory, even with  $n = 250$ , except when simulating from the Student copula; in this case the null hypothesis of a Gaussian copula is not rejected often. However, In general, as expected, the power increases with the sample size.

**Table 1: Percentage of rejection of the Cramér-von Mises statistic in the linear case with  $n \in \{250, 500\}$  for a target level of 5%, when the margins are GARCH(1,1) with Gaussian innovations and Gaussian HMM with two regimes. The levels of the tests are displayed in bold.**

$H_0$	$n = 250$				$n = 500$			
	Copula family under $H_1$				Copula family under $H_1$			
	Clayton	Gumbel	Gaussian	Student	Clayton	Gumbel	Gaussian	Student
Clayton	<b>6.1</b>	100.0	99.5	91.6	<b>5.0</b>	100	100	100
Gumbel	99.0	<b>4.9</b>	19.2	22.2	100	<b>3.8</b>	35.5	52.7
Gaussian	79.9	31.0	<b>5.3</b>	5.3	99.6	69.0	<b>5.7</b>	5.9
Student	77.5	45.2	24.5	<b>4.6</b>	99.1	80.0	45.8	<b>4.4</b>

**Remark 7** When estimating the parameter  $\nu$  of the Student copula, we restricted its value to be lower than 25, since the Student copula converges to the Gaussian copula when  $\nu$  tends to infinity. Furthermore, we did not consider the Frank copula family in these simulations because there is no explicit formula expressing its parameter in terms of Kendall’s tau.

## 5 Example of application

In this example, we want to model the dependence between the innovations of two time series considered in Adams et al. (2017). They used the Dynamic Conditional Correlation (DCC) GARCH(1,1) model of Engle (2002) on the daily log-returns of the S&P 500 index and Nasdaq index, from 1990 to 2014 ( $n = 6300$  observations). Their model is a bit restrictive since they implicitly assume that the GARCH and ARCH parameters are equal. As can be seen from Figure 1, Gaussian HMM could probably be a better fit than the DCC GARCH(1,1) for both time series since there are periods of high volatility around 2000 and 2008 which cannot be captured easily by univariate GARCH models.

For sake of comparison with Adams et al. (2017), we first tried to model both series by GARCH(1,1) models with Gaussian innovations. We then performed a goodness-of-fit test for Gaussian innovations (Ghoubi and Rémillard, 2013). In both cases, as predicted, the null hypothesis of normality was rejected. Next, we also tried to fit GARCH(1,1), GARCH(2,1) and GARCH (2,2) with Gaussian, weighted-symmetric Gaussian, GED, Student and weighted-symmetric Student with no success. Finally, we fitted Gaussian HMM models for both series, choosing the number of regimes with goodness-of-fit tests, as proposed in Rémillard et al. (2017). It then follows that for the Nasdaq returns, a Gaussian HMM with 6 regimes is needed ( $P$ -value of 45%), while a Gaussian HMM with 5 regimes on the S&P 500 returns has a  $P$ -value of 6.1%. The  $P$ -values were

computed with a parametric bootstrap of 100 samples. The estimated parameters are given in Tables 2-3. We also performed change points tests using statistic  $\mathcal{T}_{n,1}$  and  $\mathcal{T}_{n,2}$  described in Remark 2. For both series, the null hypothesis of no change point was not rejected (19% and 9.5%  $P$ -values respectively). Hence the generalized error models are both Gaussian HMM with 6 regimes and 5 regimes respectively, and both  $F_1, F_2$  are uniform margins.

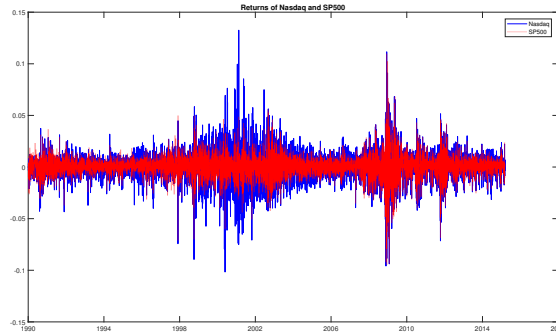


Figure 1: Daily log-returns of Nasdaq and S&P 500.

Table 2: Estimated parameters  $\mu_j, \sigma_j$  for regime  $j \in \{1, \dots, 6\}$ , and transition matrix  $Q$  for the daily log-returns of the Nasdaq.

Parameter	Regime					
$\mu_j$	0.0024	-0.0039	-0.0020	0.0083	-0.0033	0.0087
$\sigma_j$	0.0046	0.0076	0.0311	0.0119	0.0147	0.0068
Q	0.6246	0.3090	0.0000	0.0000	0.0000	0.0664
	0.1825	0.5700	0.0000	0.0000	0.0368	0.2107
	0.0000	0.0000	0.9830	0.0139	0.0031	0.0000
	0.0043	0.0002	0.0000	0.4336	0.4935	0.0684
	0.0000	0.0000	0.0084	0.2264	0.7647	0.0004
	0.7137	0.0714	0.0000	0.0001	0.0007	0.2141

Table 3: Estimated parameters  $\mu_j, \sigma_j$  for regime  $j \in \{1, \dots, 5\}$ , and transition matrix  $Q$  for the daily log-returns of the S&P500.

Parameter	Regime				
$\mu_j$	0.0008	0.0058	-0.0003	-0.0012	-0.0029
$\sigma_j$	0.0042	0.0067	0.0115	0.0260	0.0066
Q	0.7336	0.1111	0.0060	0.0000	0.1493
	0.6644	0.1524	0.0412	0.0000	0.1420
	0.0002	0.0357	0.9232	0.0064	0.0344
	0.0000	0.0010	0.0258	0.9715	0.0017
	0.1481	0.3488	0.2020	0.0000	0.3011

Next, we tested the null hypothesis of a constant copula, using Proposition 1 with the pseudo-observations. We obtained a  $P$ -value of 0%, based on 1000 bootstrap samples. As a result, the null hypothesis is rejected. When then separated the sample in two parts, according to the estimation of the change point (September 1st, 2000). With these two new series, we performed change points tests and found two other change points, one on May 10, 1995, and the other on March 3rd, 2009. So finally we have 4 periods where one cannot reject the null hypothesis of a constant copula. The length of these series are respectively 1354, 1343, 2138, and 1465, and the scatter plots of the corresponding pseudo-observations are displayed in Figure 2.

We can then use the results of Theorem 2 with  $\mathbf{h}_t \equiv 1$  to estimate the parameters. Theorem 4 can then be used for choosing the appropriate copula family.



To choose an appropriate copula model for each of the four periods, we first tested the simplest model that the copula is the independence copula; it was rejected each time. Then, using the statistic  $S_n^{(B)}$ , we performed goodness-of-fit tests for the following families: Clayton, Frank, Gumbel, Student and Gaussian, a mixture of the Clayton and its survival, a mixture of two Frank copulas (equivalent to a mixture of a Frank copula with its survival), and a mixture of the Gumbel and its survival. The best models, according to the  $P$ -value and the number of parameters are displayed in Table 4. While Adams et al. (2017) found four change points using a DCC-GARCH(1,1) model, we find three change points, and we fitted different copula models for the dependence in each period, not just a Gaussian copula. From Table 4, we can note that the dependence as measured by Kendall's tau is stronger in the last two periods which started in 2000. This can also be observed from the scatter plots in Figure 2.

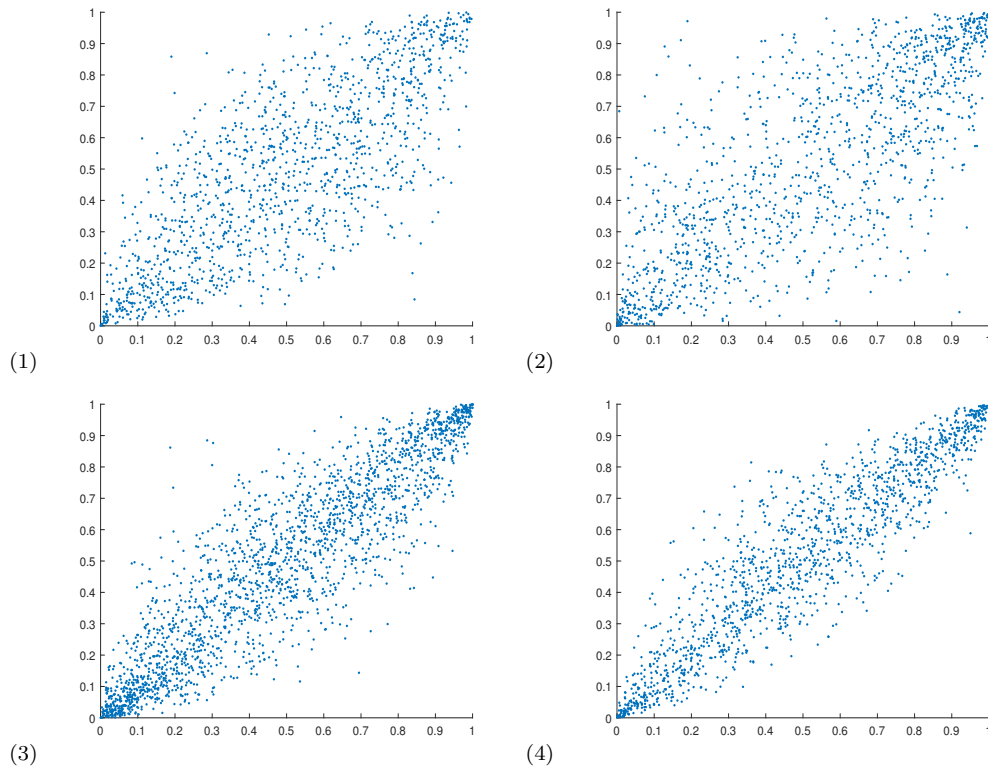


Figure 2: Scatter plots of the pseudo-observations for each period.

Table 4: Estimated parameters expressed as Kendall's tau for each model. The 95% confidence intervals are computed using the quantiles of the  $B = 1000$  bootstrap samples. For the Clayton mixture,  $p$  represents the probability of choosing the Clayton copula, so  $1 - p$  is the probability of choosing the Clayton survival copula.

Period	Model	Estimated parameters and 95% CI	$P$ -value (%)
1	Gaussian	$\tau = .587 \in [.564, .608]$	9.7%
2	Student	$\tau = .578 \in [.553, .600]$ , $\nu = 9.33 \in [5.91, 24.29]$	13.6%
3	Student	$\tau = .731 \pm .011$ , $[\cdot720, \cdot742]$ $\nu = 21.83 \in [11.34, 59.99]$	8.6%
4	Clayton mixture	$\tau_1 = .728 \in [.702, .756]$ , $\tau_2 = .7166 \in [.691, .749]$ , $p = .577 \in [.5, .653]$	54%

## 6 Conclusion

In this paper, we proposed a way to model multivariate dynamic time series by introducing copulas between the so-called generalized errors for each univariate dynamic model. Two cases were considered: time-independent and time-dependent copulas. In both cases, we proved that the limiting behavior of the empirical

processes of interest, especially those used for the change point tests and the goodness-of fit tests does not depend on the estimation errors of the generalized error models. This property also holds true for the asymptotic distribution of the estimation errors of the copula parameters. As an example of application, we analyzed the same dataset as Adams et al. (2017) and we found three change points, instead of four, fitting a constant copula model on each of the four periods. The difference between the number of change points might be due to the misidentification of the generalized error models.

## A Proofs of the main results

### A.1 Proof of Theorem 1

The proof uses the same techniques as in Ghoudi and Rémillard (2004) and Rémillard (2017), so we only indicate the main ideas. For any  $A \subset S_d = \{1, \dots, d\}$ , set

$$\mu_{A,n}(s, \mathbf{x}) = \frac{1}{\sqrt{n}} \sum_{t=1}^{\lfloor ns \rfloor} \prod_{j \in A} \{\mathbb{I}(e_{n,jt} \leq x_j) - \mathbb{I}(\varepsilon_{jt} \leq x_j)\} \prod_{k \in A^c} \mathbb{I}(\varepsilon_{kt} \leq x_k),$$

with  $\mu_{j,n} = \mu_{\{j\},n}$  for any  $j \in S_d$ . Using the multinomial formula, one has

$$\mathbb{K}_n(s, \mathbf{x}) = \check{\mathbb{K}}_n(s, \mathbf{x}) + \sum_{j=1}^d \mu_{j,n}(s, \mathbf{x}) + \sum_{|A|>1} \mu_{A,n}(s, \mathbf{x}), \quad (s, \mathbf{x}) \in [0, 1] \times [-\infty, +\infty]^d.$$

To prove the theorem, it suffices to show that for any  $1 \leq j \leq d$ , uniformly in  $(s, \mathbf{x})$ ,  $\mu_{j,n}(s, \mathbf{x})$  converges in probability to  $s \partial_{x_j} K(\mathbf{x}) \mathbf{\Gamma}_j(x_j)^\top \boldsymbol{\Theta} = s \boldsymbol{\Theta}^\top \mathbf{L}_j(\mathbf{x})$ , and that for any  $|A| > 1$ ,  $\mu_{A,n}(s, \mathbf{x})$  converges in probability to zero. These proofs will be done for  $j = 1$  and  $A \supset \{1, 2\}$ , the other cases being similar.

Let  $\delta \in (0, 1)$  be given. From (A0), (A2), (A3) and (A5), one can find  $M > 0$  such that if  $n$  is large enough, then  $P(B_{M,n}) > 1 - \delta$ , where

$$B_{M,n} = \{\|\boldsymbol{\Theta}_n\| \leq M\} \cap \left\{ \max_{1 \leq t \leq n} \frac{\|d_{n,t}\|}{r_t} \leq M \right\} \cap \left\{ \frac{1}{n} \sum_{t=1}^n \|\gamma_t(\boldsymbol{\varepsilon}_t)\| \leq M \right\}.$$

Let  $\lambda > 0$  be given. Further set  $C_{\lambda,n} = \{\max_{1 \leq t \leq n} (\|\gamma_t(\boldsymbol{\varepsilon}_t)\|/\sqrt{n} \leq \lambda)\}$ . By (A4),  $P(C_\lambda) \geq 1 - \delta$  if  $n$  is large enough. Next, for  $\boldsymbol{\kappa} = (\kappa_1, \kappa_2, \kappa_3) \in \mathbb{R} \times \mathbb{R}^p \times \mathbb{R}$ , define

$$\tilde{\mu}_{n,1}(s, \mathbf{x}; \boldsymbol{\kappa}) = \frac{1}{\sqrt{n}} \sum_{t=1}^{\lfloor ns \rfloor} [\mathbb{I}\{\varepsilon_{1t} \leq x_1 + \eta_{n,t}(\boldsymbol{\kappa})\} - \mathbb{I}(\varepsilon_{1t} \leq x_1)] \prod_{k=2}^d \mathbb{I}(\varepsilon_{kt} \leq x_k),$$

where  $\eta_{n,t}(\boldsymbol{\kappa}) = \kappa_1 r_t + \{\gamma_{1t}(\varepsilon_{1t})^\top \boldsymbol{\kappa}_2 + \kappa_3 \|\gamma_{1t}(\varepsilon_{1t})\|\} / \sqrt{n}$ .

Further set  $\tilde{\mu}_{n,12}(x_1, x_2; \boldsymbol{\kappa}) = \tilde{\mu}_{n,1}(1, x_1, x_2, \infty, \dots, \infty; \boldsymbol{\kappa})$ , and define

$$\tilde{\mathbb{L}}_{n,1}(s, \mathbf{x}; \boldsymbol{\kappa}) = \frac{1}{\sqrt{n}} \sum_{t=1}^{\lfloor ns \rfloor} [P(\varepsilon_{1t} \leq x_1 + \eta_{n,t}(\boldsymbol{\kappa}), \varepsilon_{2t} \leq x_2, \dots, \varepsilon_{dt} \leq x_d | \mathcal{F}_{t-1}) - K(\mathbf{x})].$$

The main problem working with pseudo-observations is that they depend on  $\boldsymbol{\theta}_n$ , making them dependent. However, because the closed ball of radius  $M$  in  $\mathbb{R}^p$  is compact, it can be covered by finitely many balls of radius  $\lambda$  centered at  $\boldsymbol{\zeta} \in \mathcal{C}$ , for some finite subset  $\mathcal{C}$  of  $\mathbb{R}^p$ , so one can replace the random vector  $\boldsymbol{\Theta}_n$  by any of these centers. So let  $\boldsymbol{\zeta} \in \mathcal{C}$  be given. On  $B_{M,n} \cap C_{\lambda,n} \cap \{\|\boldsymbol{\Theta}_n - \boldsymbol{\zeta}\| < \lambda\}$ , one has  $-\eta_{n,t}(M, \boldsymbol{\zeta}, \lambda) \leq \varepsilon_{1t} - e_{n,1t} \leq \eta_{n,t}(M, -\boldsymbol{\zeta}, \lambda)$ , and  $|\varepsilon_{1t} - e_{n,1t}| \leq a_{n,t}$ , with  $a_{n,t} = M \eta_{n,t}(1, 0, 1)$ . Hence  $\mathbb{I}\{\varepsilon_{1t} \leq x_1 - \eta_{n,t}(M, \boldsymbol{\zeta}, \lambda)\} \leq \mathbb{I}(e_{n,1t} \leq x_1) \leq \mathbb{I}\{\varepsilon_{1t} \leq x_1 + \eta_{n,t}(M, -\boldsymbol{\zeta}, \lambda)\}$ , and

$$\begin{aligned} |\mathbb{I}(e_{n,2t} \leq x_2) - \mathbb{I}(\varepsilon_{2t} \leq x_2)| &\leq \mathbb{I}(x_2 - Mr_t - M\lambda \leq \varepsilon_{2t} \leq x_2 + Mr_t + M\lambda) \\ &\leq \mathbb{I}(x_2 - 2M\lambda < \varepsilon_{2t} \leq x_2 + 2M\lambda), \end{aligned}$$

if  $t \geq t_0$ , for some  $t_0 > 0$ , since  $r_t \rightarrow 0$  as  $t \rightarrow \infty$ . As a result, for some  $t_1 > 0$  and for any  $A \supset \{1, 2\}$ ,

$$|\mu_{A,n}(s, \mathbf{x})| \leq \frac{t_1}{\sqrt{n}} + \frac{1}{\sqrt{n}} \sum_{t=1}^n \mathbb{I}(x_1 - a_{n,t} < \varepsilon_{1i} \leq x_1 + a_{n,t}) \mathbb{I}(x_2 - 2M\lambda < \varepsilon_{2i} \leq x_2 + 2M\lambda).$$

Then  $\tilde{\mu}_{n,1}(s, \mathbf{x}; -M, -\zeta, -\lambda) \leq \mu_{n,1}(s, \mathbf{x}) \leq \tilde{\mu}_{n,1}(s, \mathbf{x}; M, -\zeta, \lambda)$ , and

$$\begin{aligned} |\mu_{A,n}(s, \mathbf{x})| &\leq \frac{t_1}{\sqrt{n}} + \tilde{\mu}_{12,n}(x_1, x_2 + 2M\lambda; M, 0, M\lambda) - \tilde{\mu}_{12,n}(x_1, x_2 - 2M\lambda; M, 0, M\lambda) \\ &\quad - \tilde{\mu}_{12,n}(x_1, x_2 + 2M\lambda; -M, 0, -M\lambda) + \tilde{\mu}_{12,n}(x_1, x_2 - 2M\lambda; -M, 0, -M\lambda). \end{aligned}$$

The proof will be completed if one can prove the following properties: if  $\kappa_3$  small enough, then on  $B_{M,n}$ , and uniformly in  $(s, \mathbf{x})$ , both  $\tilde{\mathbf{L}}_{n,1}(s, \mathbf{x}; \boldsymbol{\kappa}) - s\boldsymbol{\kappa}_2^\top \mathbf{L}_1(\mathbf{x})$  and  $\tilde{\mu}_{n,1}(s, \mathbf{x}; \boldsymbol{\kappa}) - \tilde{\mu}_{n,1}(s, \mathbf{x}; \boldsymbol{\kappa}) - \tilde{\mathbf{L}}_{n,1}(s, \mathbf{x}; \boldsymbol{\kappa})$  can be made arbitrarily small with probability close to 1. For if these two statements are proved, then because of the previous inequalities, and because  $\lambda$  can be chosen as small as needed, one may conclude that as  $n \rightarrow \infty$ ,  $\mu_{j,n}(s, \mathbf{x})$  converges in probability to  $-\boldsymbol{\Theta}^\top \mathbf{L}_j(\mathbf{x})$ , and that for any  $A \supset \{1, 2\}$ ,  $\mu_{A,n}(s, \mathbf{x})$  converges in probability to zero. To prove the first statement, note that  $P(\varepsilon_{1t} \leq x_1 + \eta_{n,t}(\boldsymbol{\kappa}), \varepsilon_{2t} \leq x_2, \dots, \varepsilon_{dt} \leq x_d | \mathcal{F}_{t-1})$  is given by

$$\int \partial_{x_1} K(z_1, x_2, \dots, x_d) \mathbb{I}\{z_1 \leq \kappa_1 r_t + \kappa_2^\top \boldsymbol{\gamma}_{1t}(z_1) / \sqrt{n} + \kappa_2 \|\boldsymbol{\gamma}_{1t}(z_1)\| / \sqrt{n}\} dz_1.$$

It follows from (A0), (A1), (A2), and (A6) that on  $B_{M,n}$ , there is a constant  $c_0$  so that

$$\sup_{s \in [0,1]} \sup_{\mathbf{x} \in \mathbb{R}^d} \left| \tilde{\mathbf{L}}_{n,1}(s, \mathbf{x}; \boldsymbol{\kappa}) - s\boldsymbol{\kappa}_2^\top \mathbf{L}_1(\mathbf{x}) \right| \leq c_0 \left\{ \frac{|\kappa_1|}{\sqrt{n}} \sum_{t=1}^n r_t + |\kappa_3| \mathcal{L}_{n,1}(\delta) \right\},$$

which can be made arbitrarily small with large probability by choosing  $\kappa_3$  small enough. It only remains to show that the partial sum of martingale differences  $\tilde{\mu}_{n,1}(s, \mathbf{x}; \boldsymbol{\kappa})$  can be made arbitrarily small by choosing  $\kappa_3$  small enough. This is rather technical and can be easily done as in the proof of Lemmas 7.1-7.2 in Ghoudi and Rémillard (2004) or the proof of Theorem 1 in Rémillard (2017), so it is omitted.  $\square$

## A.2 Proof of Theorem 2

It follows from the proof of Theorem 1 that for any  $j \in \{1, \dots, d\}$ ,  $\mathbb{F}_{n,j}(1, x_j) = \check{\mathbf{C}}_n^{(j)} \{1, F_j(x_j)\} - f_j(x_j) \boldsymbol{\Theta}_n^\top \boldsymbol{\Gamma}_j(x_j) + o_P(1)$ , so

$$\begin{aligned} \sqrt{n}(U_{n,jt} - U_{jt}) &= \sqrt{n} \{F_{n,j}(1, e_{n,jt}) - F_j(\varepsilon_{jt})\} = \mathbb{F}_{n,j}(1, \varepsilon_{jt}) + f_j(\varepsilon_{jt}) \boldsymbol{\Theta}_n^\top \boldsymbol{\gamma}_{jt}(\varepsilon_{jt}) + o_P(1) \\ &= \check{\mathbf{C}}_n^{(j)}(1, U_{jt}) - f_j(\varepsilon_{jt}) \boldsymbol{\Theta}_n^\top \boldsymbol{\Gamma}_j(\varepsilon_{jt}) + f_j(\varepsilon_{jt}) \boldsymbol{\Theta}_n^\top \boldsymbol{\gamma}_{jt}(\varepsilon_{jt}) + o_P(1). \end{aligned}$$

Next, note that for any  $j \in \{1, \dots, d\}$ ,  $n^{-1} \sum_{t=1}^n \zeta_{jt} \check{\mathbf{C}}_n^{(j)}(1, U_{jt}) = \int_0^1 \mathbf{J}_j(u_j) \check{\mathbf{C}}_n^{(j)}(1, u_j) du_j + o_P(1)$ , using the assumptions of Theorem 2. Under the smoothness assumptions on  $c_\phi$ , and  $\phi(t, \boldsymbol{\psi})$ , one gets

$$\begin{aligned} 0 &= n^{-1/2} \sum_{t=1}^n \frac{\mathbf{h}_t^\top \dot{c}_{\phi(t, \boldsymbol{\psi}_n)}(\mathbf{U}_{n,t})}{c_{\phi(t, \boldsymbol{\psi}_n)}(\mathbf{U}_{n,t})} = \mathcal{W}_n + n^{-1/2} \sum_{j=1}^d \sum_{t=1}^n \zeta_{jt} (U_{n,jt} - U_{jt}) - \mathcal{I}_n \boldsymbol{\Psi}_n + o_P(1) \\ &= \mathcal{W}_n - \mathcal{I} \boldsymbol{\Psi}_n + n^{-1} \sum_{j=1}^d \sum_{t=1}^n \zeta_{jt} \check{\mathbf{C}}_n^{(j)}(1, U_{jt}) + \mathbf{L}_n^\top \boldsymbol{\Theta}_n + o_P(1) \\ &= \mathcal{W}_n + \sum_{j=1}^d \int_0^1 \mathbf{J}_j(u_j) \check{\mathbf{C}}_n^{(j)}(1, u_j) du_j - \mathcal{I} \boldsymbol{\Psi}_n + o_P(1). \end{aligned}$$

Then, as  $n \rightarrow \infty$ ,  $(\check{\mathbf{C}}_n, \mathcal{W}_n, \boldsymbol{\Psi}_n)$  converges in law to  $(\check{\mathbf{C}}, \mathcal{W}, \boldsymbol{\Psi})$ , the limit of  $\boldsymbol{\Psi}_n$  can be deduced. Note that for any  $j \in \{1, \dots, d\}$  and any  $u_j \in [0, 1]$ ,  $E\{\check{\mathbf{C}}^{(j)}(1, u_j) \mathcal{W}\} = 0$ , so  $\{\check{\mathbf{C}}(1, \cdot), \dots, \check{\mathbf{C}}(1, \cdot)\}$  and  $\mathcal{W}$  are independent. As a result,  $E(\boldsymbol{\Psi} \mathcal{W}^\top) = \mathbb{I}$ .  $\square$

### A.2.1 Proof of Theorem 3

In this case,  $\mathbf{h}_t \equiv 1$ , most terms depending on  $t$  are stationary, and the processes are ergodic since  $Q$  is ergodic. The proof is similar to the previous one, so it is omitted.  $\square$

### A.3 Proof of Theorem 4

Recall that  $\check{\mathbb{D}}_n(\mathbf{u}) = \frac{1}{\sqrt{n}} \sum_{t=1}^n [\mathbb{I}\{\mathbf{V}_t \leq \mathbf{u}\} - \Pi(\mathbf{u})]$ ,  $\mathbf{u} \in [0, 1]^d$ , where  $\mathbf{V}_t = \mathcal{R}_t(\mathbf{U}_t)$ ,  $t \in \{1, \dots, n\}$ . Then  $\check{\mathbb{D}}_n \rightsquigarrow \check{\mathbb{D}}$ , where  $\check{\mathbb{D}}$  is a  $\Pi$ -Brownian bridge.

Next, for  $t \in \{1, \dots, n\}$ , and any  $j \in \{1, \dots, d\}$ ,

$$\begin{aligned} \sqrt{n}(V_{n,jt} - V_{jt}) &= \sqrt{n} \{ \mathcal{R}_{\phi_{n(t),j}}(\mathbf{U}_{n,t}) - \mathcal{R}_{jt}(\mathbf{U}_{n,t}) + \mathcal{R}_{jt}(\mathbf{U}_{n,t}) - \mathcal{R}_{jt}(\mathbf{U}_t) \} \\ &= \dot{\mathcal{R}}_{jt}(\mathbf{U}_t)^\top \mathbf{h}_t \boldsymbol{\Psi}_n + \sqrt{n} \sum_{k=1}^d \partial_{u_k} \mathcal{R}_{jt}(\mathbf{U}_t) [F_{n,k} \{1, G_{\boldsymbol{\theta}_{n,kt}}(X_{kt})\} - F_k \{G_{kt}(X_{kt})\}] + o_P(1) \\ &= \dot{\mathcal{R}}_{jt}(\mathbf{U}_t)^\top \mathbf{h}_t \boldsymbol{\Psi}_n + \sum_{k=1}^d \partial_{u_k} \mathcal{R}_{jt}(\mathbf{U}_t) [\mathbb{F}_{n,k}(1, \varepsilon_{kt}) + f_k(\varepsilon_{kt}) \boldsymbol{\gamma}_{kt}(\varepsilon_{kt})^\top \boldsymbol{\Theta}_n] + o_P(1) \\ &= \dot{\mathcal{R}}_{jt}(\mathbf{U}_t)^\top \mathbf{h}_t \boldsymbol{\Psi}_n + \sum_{k=1}^d \partial_{u_k} \mathcal{R}_{jt}(\mathbf{U}_t) \check{\mathbb{C}}_n^{(k)}(1, U_{kt}) + \mathcal{L}_{n,jt}^\top \boldsymbol{\Theta}_n + o_P(1), \end{aligned}$$

where  $\mathcal{L}_{n,jt} = \sum_{k=1}^d \partial_{u_k} \mathcal{R}_{jt}(\mathbf{U}_t) f_k(\varepsilon_{kt}) \{ \boldsymbol{\gamma}_{kt}(\varepsilon_{kt}) - \boldsymbol{\Gamma}_k(\varepsilon_{kt}) \}$ . One can then replicated the proof of Theorem 1 to obtain the convergence of the process. Hence

$$\mathbb{D}_n(\mathbf{u}) = \check{\mathbb{D}}_n(\mathbf{u}) - \kappa_n \left\{ \check{\mathbb{C}}_n^{(1)}(1, \cdot), \dots, \check{\mathbb{C}}_n^{(d)}(1, \cdot), \mathbf{u} \right\} - \boldsymbol{\Psi}_n^\top \boldsymbol{\varrho}_n(\mathbf{u}) - \boldsymbol{\Theta}_n^\top \boldsymbol{\delta}_n(\mathbf{u}) + o_P(1).$$

It then follows that  $\mathbb{D}_n \rightsquigarrow \mathring{\mathbb{D}} = \check{\mathbb{D}} - \kappa - \boldsymbol{\Psi}^\top \boldsymbol{\varrho}$ . Similarly,  $\mathring{\mathbb{D}}_n \rightsquigarrow \mathring{\mathbb{D}}$ , since this case corresponds to a known  $\boldsymbol{\theta}$ , and the estimation of  $\boldsymbol{\psi}$  in this case converges to the same limit. Next, from Theorem 2 or Theorem 3, one may conclude that  $\boldsymbol{\psi}_n$  is a regular estimator of  $\boldsymbol{\psi}$ , implying that  $E(\boldsymbol{\Phi} \mathcal{W}^\top) = I$ . Finally, it follows from Rémillard (2011)[Lemma 1] that  $E\{\check{\mathbb{D}}_n(\mathbf{u}) \mathcal{W}_n\} = \boldsymbol{\varrho}_n(\mathbf{u})$  for any  $\mathbf{u} \in [0, 1]^d$ . As a result,  $E\{\mathring{\mathbb{D}} \mathcal{W}\} = \boldsymbol{\varrho}$ .  $\square$

## References

- Adams, Z., Füßs, R., and Glück, T. (2017). Are correlations constant? Empirical and theoretical results on popular correlation models in finance. *Journal of Banking & Finance*, 84:9–24.
- Bai, J. (2003). Testing parametric conditional distributions of dynamic models. *The Review of Economics and Statistics*, 85(3):531–549.
- Bickel, P. J. and Wichura, M. J. (1971). Convergence criteria for multiparameter stochastic processes and some applications. *Ann. Math. Statist.*, 42:1656–1670.
- Bücher, A. and Kojadinovic, I. (2016). A dependent multiplier bootstrap for the sequential empirical copula process under strong mixing. *Bernoulli*, 22(2):927–968.
- Bücher, A. and Volgushev, S. (2013). Empirical and sequential empirical copula processes under serial dependence. *Journal of Multivariate Analysis*, 119:61–70.
- Carlstein, E. (1988). Nonparametric change-point estimation. *Ann. Statist.*, 16:188–197.
- Chen, X. and Fan, Y. (2006a). Estimation and model selection of semiparametric copula-based multivariate dynamic models under copula misspecification. *Journal of Econometrics*, 135(1–2):125–154.
- Chen, X. and Fan, Y. (2006b). Estimation of copula-based semiparametric time series models. *Journal of Econometrics*, 130:307–335.

- Du, Z. (2016). Nonparametric bootstrap tests for independence of generalized errors. *The Econometrics Journal*, 19(1):55–83.
- Embrechts, P., McNeil, A. J., and Straumann, D. (2002). Correlation and dependence in risk management: properties and pitfalls. In *Risk Management: Value at Risk and Beyond* (Cambridge, 1998), pages 176–223. Cambridge Univ. Press, Cambridge.
- Engle, R. (2002). Dynamic conditional correlation. *Journal of Business & Economic Statistics*, 20(3):339–350.
- Fermanian, J.-D. and Wegkamp, M. H. (2012). Time-dependent copulas. *Journal of Multivariate Analysis*, 110:19–29. Special Issue on Copula Modeling and Dependence.
- Genest, C., Ghoudi, K., and Rivest, L.-P. (1995). A semiparametric estimation procedure of dependence parameters in multivariate families of distributions. *Biometrika*, 82:543–552.
- Genest, C. and Rémillard, B. (2008). Validity of the parametric bootstrap for goodness-of-fit testing in semiparametric models. *Ann. Inst. H. Poincaré Sect. B*, 44:1096–1127.
- Genest, C., Rémillard, B., and Beaudoin, D. (2009). Omnibus goodness-of-fit tests for copulas: A review and a power study. *Insurance Math. Econom.*, 44:199–213.
- Ghoudi, K. and Rémillard, B. (2004). Empirical processes based on pseudo-observations. II. The multivariate case. In *Asymptotic Methods in Stochastics*, volume 44 of *Fields Inst. Commun.*, pages 381–406. Amer. Math. Soc., Providence, RI.
- Ghoudi, K. and Rémillard, B. (2013). Comparison of specification tests for GARCH models. *Computational Statistics & Data Analysis*, 76:291–300.
- Ghoudi, K. and Rémillard, B. (2015). Diagnostic tests for innovations of ARMA models using empirical processes of residuals. In Dawson, D., Kulik, R., Haye, M. O., Szyszkowicz, B., and Zhao, Y., editors, *Asymptotic Laws and Methods in Stochastics*, number 76 in *Fields Institute Communications*, pages 239–282. Springer, New York.
- Jondeau, E. and Rockinger, M. (2006). The Copula-GARCH model of conditional dependencies: An international stock market application. *Journal of International Money and Finance*, 25(5):827–853.
- Nasri, B., Rémillard, B., and Bouezmarni, T. (2017). Semi-parametric copula-based models under non-stationarity. Technical Report G-2017-102, GERAD.
- Nasri, B. R. and Rémillard, B. (2018). Non-parametric change point problems using multipliers. Technical report, SSRN Working Paper Series No. 2043632.
- Patton, A. J. (2004). On the out-of-sample importance of skewness and asymmetric dependence for asset allocation. *Journal of Financial Econometrics*, 2(1):130–168.
- Patton, A. J. (2006). Modelling asymmetric exchange rate dependence. *International Economic Review*, 47(2):527–556.
- Patton, A. J. (2009). Copula-based models for financial time series. In Mikosch, T., Kreiß, J.-P., Davis, R. A., and Andersen, T. G., editors, *Handbook of Financial Time Series*, pages 767–785. Springer Berlin Heidelberg.
- Patton, A. J. (2012). A review of copula models for economic time series. *J. Multivariate Anal.*, 110:4–18.
- Rémillard, B. (2011). Validity of the parametric bootstrap for goodness-of-fit testing in dynamic models. Technical report, SSRN Working Paper Series No. 1966476.
- Rémillard, B. (2017). Goodness-of-fit tests for copulas of multivariate time series. *Econometrics*, 5(1):13.
- Rémillard, B., Hocquard, A., Lamarre, H., and Papageorgiou, N. A. (2017). Option pricing and hedging for discrete time regime-switching model. *Modern Economy*, 8:1005–1032.
- Rémillard, B., Papageorgiou, N., and Soutra, F. (2012). Copula-based semiparametric models for multivariate time series. *J. Multivariate Anal.*, 110:30–42.
- Rosenblatt, M. (1952). Remarks on a multivariate transformation. *Ann. Math. Stat.*, 23:470–472.

- Segers, J. (2012). Asymptotics of empirical copula processes under non-restrictive smoothness assumptions. *Bernoulli*, 18(3):764–782.
- van den Goorbergh, R. W. J., Genest, C., and Werker, B. J. M. (2005). Bivariate option pricing using dynamic copula models. *Insurance: Mathematics and Economics*, 37:101–114.
- Xu, J. J. (1996). *Statistical Modelling and Inference for Multivariate and Longitudinal Discrete Response Data*. PhD thesis, University of British Columbia.