Article

Goodness-of-Fit Tests for Copulas of Multivariate Time Series

Bruno Rémillard

GERAD, CRM, and Department of Decision Sciences, HEC Montréal, 3000 Chemin de la Côte Sainte-Catherine, Montréal, QC H3T 2A7, Canada; bruno.remillard@hec.ca; Tel.: +1-514-340-6794

Abstract: In this paper, we study the asymptotic behavior of the sequential empirical process and the sequential empirical copula process, both constructed from residuals of multivariate stochastic volatility models. Applications for the detection of structural changes and specification tests of the distribution of innovations are discussed. It is also shown that if the stochastic volatility matrices are diagonal, which is the case if the univariate time series are estimated separately instead of being jointly estimated, then the empirical copula process behaves as if the innovations were observed; a remarkable property. As a by-product, one also obtains the asymptotic behavior of rank-based measures of dependence applied to residuals of these time series models.

Keywords: goodness-of-fit; time series; copulas; GARCH models

1. Introduction

In many financial applications, it is necessary to model both the serial dependence and the dependence between different time series. This can be done instantly by proposing a full parametric model for the multivariate time series, or by modeling the serial dependence for each series and then achieving the interdependence between the series through copulas, which are joint distribution functions of random variables with uniform margins. In any case, one has to deal with the residuals of the model since the innovations are not observable.

Working with residuals complicates the inference procedure since the limiting distribution of statistics and parameters depend in general on unknown parameters. See, e.g., Bai [1] and Ghoudi and Rémillard [2]. In particular, as shown in Bai [3] and Horváth et al. [4], the distribution of the empirical process of GARCH residuals in the univariate case (or their squares) is not trivial. Here, we aim to extend the results of these authors to find the asymptotic behavior of sequential empirical processes constructed from the residuals of multivariate stochastic volatility models. The reason for considering sequential processes is that we want to be able to construct tests for the detection of structural changes in the distribution of the innovations, which is an important problem in practice. Unveiling the punch, one can show that the limiting distribution of the test statistics used for change point analysis will not depend on the estimated parameters of the conditional mean and covariance.

Empirical processes can also be used to test hypotheses about the distribution. In particular, one method involves trying to fit copula models for the dependence between the innovations of several time series. In many applications so far, the serial dependence problem is either ignored, i.e., the data are not "filtered" to remove serial dependence, as in Dobrić and Schmid [5], Dobrić and Schmid [6] and Kole et al. [7], or the data are "filtered" but the potential inferential problems of using these transformed data are not taken into account. For example, Panchenko [8] uses a goodness-of-fit test on "filtered" data (residuals of GARCH models in his case), without proving that his proposed methodology works for residuals. However, he mentioned in passing that working with residuals could destroy the asymptotic properties of his test. A similar situation appears in Breymann et al. [9] where both the problem of working with residuals and the problem of the estimation of the copula parameters are ignored.

It seems that the first paper that rigorously addresses the problems raised by the use of residuals in estimation and goodness-of-fit of copulas is Chen and Fan [10]. Using a multivariate GARCH-like

model with diagonal stochastic volatility matrices, Chen and Fan [10] showed the remarkable result that estimating the copula parameters using the rank-based maximum pseudo-likelihood method [11,12] with the ranks of the residuals instead of the (non-observable) ranks of innovations, leads to the same asymptotic distribution. In particular, the limiting distribution of the estimation of the copula parameters does not depend on the unknown parameters used to estimate the conditional mean and the conditional variance. This property is crucial if one wants to develop goodness-of-fit tests for the copula family of the innovations. In Chen and Fan [10], the authors also proposed ways of selecting or more precisely ranking copulas, based on pseudo-likelihood ratio tests.

Here, under similar assumptions, one can show that the limiting distribution of the empirical copula process does not depend on the conditional mean and conditional variance parameters. As a by-product, the limiting distribution of rank-based dependence measures computed with the residuals are the same as if the dependence measures were computed with the innovations. It is worth noting that, even if Duchesne et al. [13], and Oh and Patton [14] (Section 3.4) were published before this paper, the authors actually build on results proved here.

In what follows, one starts, in Section 2, by describing the model and stating the convergence result for the main sequential empirical process. As a result, tests of structural change and specification tests can be defined for the innovations. Also, as a corollary, one obtains the remarkable result that the limiting copula process does not depend on the estimated parameters of the conditional mean and variance, if the stochastic volatility matrices are diagonal (called Model 2 here and defined in Section 2). This is indeed the case in Chen and Fan [10]. Under this model, specification tests for the copula are proposed in Section 3. These tests are either based on the empirical copula or the empirical Rosenblatt process, whose limiting distribution is studied. Also, since the specification tests rely on estimators of the copula parameters, the asymptotic behavior of rank-based estimators is studied. In particular, one recalls the remarkable result of Chen and Fan [10] concerning maximum pseudo-likelihood estimators. One also obtains the asymptotic behavior of rank-based dependence measures. Finally, a real data example is treated in Section 4, using a data set from Chen and Fan [10]. The main results are proved in a series of Appendices.

2. Weak Convergence of Empirical Processes of Residuals

Consider the stochastic volatility model

$$\mathbf{X}_i = \boldsymbol{\mu}_i(\boldsymbol{\theta}) + \boldsymbol{\sigma}_i(\boldsymbol{\theta})\boldsymbol{\varepsilon}_i,$$

where the innovations $\boldsymbol{\varepsilon}_i = (\varepsilon_{1i}, \ldots, \varepsilon_{di})^\top$ are i.i.d., $E(\varepsilon_{ji}) = 0$, $E(\varepsilon_{ji}^2) = 1$, with continuous distribution function K, and $\boldsymbol{\mu}_i$, $\boldsymbol{\sigma}_i$ are $\mathcal{F}_{i-i}$-measurable and independent of $\boldsymbol{\varepsilon}_i$. Here $\mathcal{F}_{i-1}$ contains information from the past and possibly information from exogenous variables as well. Since the distribution function K is continuous, there exists a unique copula C [15] so that for all $\mathbf{x} = (x_1, \ldots, x_d)^\top \in \mathbb{R}^d$,

$$K(\mathbf{x}) = C\{\mathbf{F}(\mathbf{x})\}, \quad \mathbf{F}(\mathbf{x}) = (F_1(x_1), \ldots, F_d(x_d))^\top, \tag{1}$$

where $F_1, \ldots, F_d$ are the marginal distribution functions of K, i.e., F_j is the distribution function of ε_{ji}, $j \in \{1, \ldots, d\}$. Defining $\mathbf{U}_i = \mathbf{F}(\boldsymbol{\varepsilon}_i)$, $i \in \{1, \ldots, n\}$, one obtains that $\mathbf{U}_1, \ldots, \mathbf{U}_n$ are i.i.d. with distribution function C. However they are not observable since $\mathbf{F}$ is unknown.

Given an estimator $\boldsymbol{\theta}_n$ of $\boldsymbol{\theta}$, compute the residuals $\mathbf{e}_{i,n} = (e_{1i,n}, \ldots, e_{di,n})^\top$, where

$$\mathbf{e}_{i,n} = \boldsymbol{\sigma}_i^{-1}(\boldsymbol{\theta}_n)\{\mathbf{X}_i - \boldsymbol{\mu}_i(\boldsymbol{\theta}_n)\}.$$

The main results of the paper are deduced from the asymptotic behavior of the sequential empirical process

$$\mathbb{K}_n(s,\mathbf{x}) = \frac{1}{\sqrt{n}} \sum_{i=1}^{\lfloor ns \rfloor} \{\mathbf{1}(e_{i,n} \le \mathbf{x}) - K(\mathbf{x})\}, \quad (s,\mathbf{x}) \in [0,1] \times \bar{\mathbb{R}}^d, \tag{2}$$

where $\mathbf{1}$ stands for the indicator function and $\mathbf{y} \le \mathbf{x}$ means that the inequality holds componentwise. Further set

$$K_n(\mathbf{x}) = \frac{1}{n} \sum_{i=1}^{n} \mathbf{1}(e_{i,n} \le \mathbf{x}), \quad \mathbf{x} \in \bar{\mathbb{R}}^d,$$

and $\mathbf{F}_n(\mathbf{x}) = (F_{1n}(1,x_1), \ldots, F_{dn}(1,x_d))^\top$, where

$$F_{jn}(s,x_j) = \frac{1}{n+1} \sum_{i=1}^{\lfloor ns \rfloor} \mathbf{1}(e_{ji,n} \le x_j), \quad j \in \{1,\ldots,d\}, \quad (s,\mathbf{x}) \in [0,1] \times \bar{\mathbb{R}}^d. \tag{3}$$

Finally, for $j \in \{1,\ldots,d\}$, the marginal processes $\mathbb{F}_{j,n}$ are defined, for any $s \in [0,1]$ and any $x_j \in \bar{\mathbb{R}}$ by

$$\mathbb{F}_{j,n}(s,x_j) = \frac{1}{\sqrt{n}} \sum_{i=1}^{\lfloor ns \rfloor} \{\mathbf{1}(e_{ji,n} \le x_j) - F_j(x_j)\} = \sqrt{n}\{F_{j,n}(s,x_j) - sF_j(x_j)\} + o_P(1).$$

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 products of theses spaces. Note that random vectors belong to these spaces because they are constant random functions.

To be able to state the convergence result for $\mathbb{K}_n$, one needs to introduce auxiliary empirical processes. Set

$$\alpha_n(s,\mathbf{x}) = \frac{1}{\sqrt{n}} \sum_{i=1}^{\lfloor ns \rfloor} \{\mathbf{1}(\boldsymbol{\varepsilon}_i \le \mathbf{x}) - K(\mathbf{x})\}, \quad (s,\mathbf{x}) \in [0,1] \times \bar{\mathbb{R}}^d,$$

$$\beta_n(s,\mathbf{u}) = \frac{1}{\sqrt{n}} \sum_{i=1}^{\lfloor ns \rfloor} \{\mathbf{1}(\mathbf{U}_i \le \mathbf{u}) - C(\mathbf{u})\}, \quad (s,\mathbf{u}) \in [0,1]^{1+d},$$

and $\beta_{j,n}(s,u_j) = \beta_n(s,1,\ldots,1,u_j,1,\ldots,1) = \frac{1}{\sqrt{n}} \sum_{i=1}^{\lfloor ns \rfloor} \{\mathbf{1}(U_{ji} \le u_j) - u_j\}$, $j = 1\ldots,d$.

It is well known [16] that $\alpha_n \rightsquigarrow \alpha$ and $\beta_n \rightsquigarrow \beta$ where α is a K-Kiefer process and β is a C-Kiefer process. Recall that α is a K-Kiefer process if it is a continuous centered Gaussian process with $\mathrm{Cov}\{\alpha(s,\mathbf{x}),\alpha(t,\mathbf{y})\} = (s \wedge t)\{K(\mathbf{x} \wedge \mathbf{y}) - K(\mathbf{x})K(\mathbf{y})\}$, $s \in [0,1]$ and $\mathbf{x},\mathbf{y} \in \bar{\mathbb{R}}^d$. Here $(\mathbf{x} \wedge \mathbf{y})_j = \min(x_j, y_j)$, $j = 1\ldots,d$. Note that for all $(s,\mathbf{x}) \in [0,1] \times \bar{\mathbb{R}}^d$, $\alpha(s,\mathbf{x}) = \beta\{s,\mathbf{F}(\mathbf{x})\}$.

The following assumptions are needed in order to prove the convergence of $\mathbb{K}_n$. First, assume that $\boldsymbol{\mu}_i$ and $\boldsymbol{\sigma}_i$ are continuously differentiable with respect to $\boldsymbol{\theta} \in \mathcal{O} \subset \mathbb{R}^p$, and set $\boldsymbol{\gamma}_{0i} = \boldsymbol{\sigma}_i^{-1}\dot{\boldsymbol{\mu}}_i$ and $\boldsymbol{\gamma}_{1ki} = \boldsymbol{\sigma}_i^{-1}\dot{\boldsymbol{\sigma}}_{ki}$, where

$$(\dot{\boldsymbol{\mu}}_i)_{jl} = \partial_{\theta_l}\mu_{ji}, \quad (\dot{\boldsymbol{\sigma}}_{ki})_{jl} = \partial_{\theta_l}\sigma_{jki} = \partial_{\theta_l}(\boldsymbol{\sigma}_i)_{jk}, \quad j,k \in \{1,\ldots,d\}, \quad l \in \{1,\ldots,p\}.$$

Let $\mathbf{d}_{i,n} = \boldsymbol{\varepsilon}_i - \mathbf{e}_{i,n} - (\boldsymbol{\gamma}_{0i}\boldsymbol{\Theta}_n + \sum_{k=1}^d \varepsilon_{ki}\boldsymbol{\gamma}_{1ki}\boldsymbol{\Theta}_n)/\sqrt{n}$, where $\boldsymbol{\Theta}_n = n^{1/2}(\boldsymbol{\theta}_n - \boldsymbol{\theta})$. Note that $\boldsymbol{d}_{i,n}$ is the error term in the first order Taylor expansion of $e_{i,n}$ about $\boldsymbol{\theta}$. These terms are needed in order to be able to measure the difference between $\mathbb{K}_n$ and α_n. Next, assume that for any $j \in \{1,\ldots,d\}$, and any $\mathbf{x} \in \bar{\mathbb{R}}^d$, the following properties hold:

(A1) $\mathbf{\Gamma}_{0,n}(s) = \frac{1}{n}\sum_{i=1}^{\lfloor ns \rfloor} \gamma_{0i} \xrightarrow{Pr} s\mathbf{\Gamma}_0$, $\mathbf{\Gamma}_{1k,n}(s) = \frac{1}{n}\sum_{i=1}^{\lfloor ns \rfloor} \gamma_{1ki} \xrightarrow{Pr} s\mathbf{\Gamma}_{1k}$, uniformly in $s \in [0,1]$, where $\mathbf{\Gamma}_0$ and $\mathbf{\Gamma}_{1k}$ are deterministic, $k = 1, \ldots, d$.

(A2) $\frac{1}{n}\sum_{i=1}^{n} \mathrm{E}\left(\|\gamma_{0i}\|^k\right)$ and $\frac{1}{n}\sum_{i=1}^{n} \mathrm{E}\left(\|\gamma_{1ji}\|^k\right)$ are bounded, for $k = 1, 2$.

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

(A4) $\max_{1\leq i\leq n} \|\gamma_{0i}\|/\sqrt{n} = o_P(1)$ and $\max_{1\leq i\leq n} |\varepsilon_{ji}|\|\gamma_{1ji}\|/\sqrt{n} = o_P(1)$.

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

(A6) $\partial_{x_j} K(\mathbf{x})$ and $x_j \partial_{x_j} K(\mathbf{x})$ are bounded and continuous on $\bar{\mathbb{R}}^d = [-\infty, +\infty]^d$. In addition, $F_1, \ldots, F_d$ have continuous bounded densities $f_1, \ldots, f_d$ respectively.

(A7) For all $k \neq j$, $f_j(x_j)E\{|\varepsilon_{k1}|\mathbf{1}(\boldsymbol{\varepsilon}_1 \leq x)|\varepsilon_{j1} = x_j\}$ and $x_j f_j(x_j)E\{|\varepsilon_{k1}|\mathbf{1}(\boldsymbol{\varepsilon}_1 \leq x)|\varepsilon_{j1} = x_j\}$ are bounded and continuous on $\bar{\mathbb{R}}^d$.

Remark 1. Note that (A1) and (A2) are trivially satisfied if the sequences γ_{0i} and γ_{1ki} are stationary, ergodic and square integrable. Also, if $\frac{1}{n}\sum_{i=1}^{n} \gamma_{0i} \xrightarrow{Pr} \mathbf{\Gamma}_0$, $\frac{1}{n}\sum_{i=1}^{n} \gamma_{1ki} \xrightarrow{Pr} \mathbf{\Gamma}_{1k}$, and $\frac{1}{n}\sum_{i=1}^{n} E(\|\gamma_{0i}\|^2)$, $\frac{1}{n}\sum_{i=1}^{n} E(\|\gamma_{1ki}\|^2)$ converge, then (A1) and (A2) are satisfied.

In order to present the main results, and in view of applications, one needs to consider two models:

- Model 1. In this model, for the sake of identifiability, we assume that the correlation matrix $E\left(\boldsymbol{\varepsilon}_i \boldsymbol{\varepsilon}_i^\top\right)$ is the identity matrix. This means that the conditional covariance matrix of the observations relative to $\mathcal{F}_{i-1}$ is $\sigma_i \sigma_i^\top$. In particular, the conditional correlation is not necessarily constant and it is estimated.
- Model 2. In this model, σ_i is a diagonal matrix for any i, and there is no restriction on the correlation matrix of the innovations. This means $\sigma_i \sigma_i^\top$ is diagonal and for any $j \in \{1, \ldots, d\}$, $\left(\sigma_i \sigma_i^\top\right)_{jj}$ is the conditional variance of the observations X_{ji} relative to $\mathcal{F}_{i-1}$. In particular, this implies that the conditional correlation between the observations is constant, does not depend on $\boldsymbol{\theta}$, and it is implicitly incorporated in the copula C of the innovations. This model appears naturally in practice when the parameters of univariate time series are estimated separately as in Chen and Fan [10]. This model is not the same as the diagonal representation in Engle and Kroner [17], which is included in Model 1.

Remark 2. If in Model 1 the conditional correlations are constant (say $\boldsymbol{R}$), then this model can be transformed into the Model 2 representation. In fact, one can then find a matrix $\boldsymbol{a}$ so that $\boldsymbol{R} = \boldsymbol{a}\boldsymbol{a}^\top$ and $\sigma_i = \tilde{\sigma}_i \boldsymbol{a}$, where $\tilde{\sigma}_i$ diagonal. Then $\tilde{\boldsymbol{\varepsilon}}_i = a\boldsymbol{\varepsilon}_i$ has correlation $\boldsymbol{R}$ and $\boldsymbol{X}_i = \boldsymbol{\mu}_i + \tilde{\sigma}_i \tilde{\boldsymbol{\varepsilon}}_i$.

One can now state the main convergence result. Its proof is given in Appendix A.1.

Theorem 1. *Under Model 1 and assumptions (A1)–(A7), $\mathbb{K}_n \rightsquigarrow \mathbb{K}$, with*

$$\mathbb{K}(s, \mathbf{x}) = \alpha(s, \mathbf{x}) + s\nabla K(\mathbf{x})^\top \mathbf{\Gamma}_0 \Theta + s\sum_{j=1}^{d}\sum_{k=1}^{d} G_{jk}(\mathbf{x})(\mathbf{\Gamma}_{1k}\Theta)_j,$$

where $G_{jk}(\mathbf{x}) = f_j(x_j)E\left\{\varepsilon_{k1}\mathbf{1}(\varepsilon_1 \leq \mathbf{x})|\varepsilon_{j1} = x_j\right\}$. *In particular,* $G_{jj}(\mathbf{x}) = x_j\partial_{x_j}K(\mathbf{x})$. *Furthermore, for all* $j \in \{1,\ldots,d\}$, $\mathbb{F}_{j,n} \rightsquigarrow \mathbb{F}_j$, *where*

$$\mathbb{F}_j(s,x_j) = \beta_j\{s,F_j(x_j)\} + sf_j(x_j)\left\{(\mathbf{\Gamma}_0\Theta)_j + x_j(\mathbf{\Gamma}_{1j}\Theta)_j\right\} + s\sum_{k\neq j} f_j(x_j)E(\varepsilon_{k1}|\varepsilon_{j1} = x_j)(\mathbf{\Gamma}_{1k}\Theta)_j.$$

Under Model 2 and assumptions (A1)–(A6), $\mathbb{K}_n \rightsquigarrow \mathbb{K}$, *where*

$$\mathbb{K}(s,\mathbf{x}) = \alpha(s,\mathbf{x}) + s\nabla K(\mathbf{x})^\top\mathbf{\Gamma}_0\Theta + s\sum_{j=1}^{d} G_{jj}(\mathbf{x})(\mathbf{\Gamma}_{1j}\Theta)_j.$$

Remark 3. One immediate application of Theorem 1 if for goodness-of-fit tests for K. For example, one can be interested in testing the null hypothesis $H_0 : K \in \mathcal{K} = \{K_{\boldsymbol{\phi}}; \boldsymbol{\phi} \in \mathcal{P}\}$, for some parametric family $\mathcal{K}$. Using Theorem 1, such tests could be based on functionals of the empirical process $\sqrt{n}(K_n - K_{\phi_n}) = \mathbb{K}_n(1,\cdot) - \sqrt{n}(K_{\boldsymbol{\phi}_n} - K)$, provided ϕ_n is a "good estimator" of $\boldsymbol{\phi}$. Goodness-of-fit tests could also be based on the so-called Rosenblatt transform of K. See, e.g., Genest and Rémillard [18] and Rémillard [19] for details.

To state the next result, which is crucial for change point problems involving innovations, define, for all $(s,\mathbf{x}) \in [0,1]\times\bar{\mathbb{R}}^d$, the sequential process

$$\mathbb{A}_n(s,\mathbf{x}) = \frac{1}{\sqrt{n}}\sum_{i=1}^{\lfloor ns\rfloor}\{\mathbf{1}(e_{i,n} \leq \mathbf{x}) - K_n(\mathbf{x})\} = \mathbb{K}_n(s,\mathbf{x}) - \frac{\lfloor ns\rfloor}{n}\mathbb{K}_n(1,\mathbf{x}). \tag{4}$$

Note that many test statistics for detecting structural changes in the distribution of innovations are based on functionals of $\mathbb{A}_n$. From the representation of $\mathbb{A}_n$ in terms of $\mathbb{K}_n$ given in (4) and Theorem 1, one obtains a surprising result for the asymptotic behavior of $\mathbb{A}_n$.

Corollary 1. *Under Model 1 and assumptions (A1)–(A7), or under Model 2 and assumptions (A1)–(A6),* $\mathbb{A}_n \rightsquigarrow \mathbb{A}$, *with*

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

In particular $\mathbb{A}$ *is parameter-free, depending only on* K.

Remark 4. 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 innovations. See Rémillard [20] for details.

Empirical Processes Related to the Copula

So far, we have discussed two applications of Theorem 1: specification tests and change point problems for the distribution of the innovations. Next, if one is interested in modeling the dependence between innovations, one has 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 $\mathbf{e}_{i,n}$ with the associated rank vectors

$$\mathbf{U}_{i,n} = (U_{1i,n},\ldots,U_{di,n})^\top, \quad U_{ji,n} = \text{Rank}(e_{ji,n})/(n+1), \quad i \in \{1,\ldots,n\},$$

where $\text{Rank}(e_{ji,n})$ being the rank of $e_{ji,n}$ amongst $e_{j1,n},\ldots,e_{jn,n}$, $j \in \{1,\ldots,d\}$. This can also be written as $\mathbf{U}_{i,n} = \mathbf{F}_n(\mathbf{e}_{i,n})$, $i \in \{1,\ldots,n\}$. Now define the empirical copula

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

together with the sequential copula process

$$\mathbb{C}_n(s,\mathbf{u}) = \frac{1}{\sqrt{n}} \sum_{i=1}^{\lfloor ns \rfloor} \{\mathbf{1}(U_{i,n} \le \mathbf{u}) - C(\mathbf{u})\}, \quad (s,\mathbf{u}) \in [0,1]^{1+d},$$

and set

$$\mathbb{G}_n(s,\mathbf{u}) = \frac{1}{\sqrt{n}} \sum_{i=1}^{\lfloor ns \rfloor} \{\mathbf{1}(U_{i,n} \le \mathbf{u}) - C_n(\mathbf{u})\} = \mathbb{C}_n(s,\mathbf{u}) - \frac{\lfloor ns \rfloor}{n} \mathbb{C}_n(1,\mathbf{u}).$$

In order to work on the space of continuous functions on $[0,1]^d$, one assumes from now on the following additional technical assumption [21] (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, using Genest et al. [22] (Proposition A.1), and the representation of $\mathbb{G}_n$.

Corollary 2. *Under Model 1 and assumptions (A1)–(A7), $\mathbb{C}_n \rightsquigarrow \mathbb{C}$, with*

$$\mathbb{C}(s,\mathbf{u}) = \check{\mathbb{C}}(s,\mathbf{u}) + s \sum_{j \ne k} \tilde{G}_{jk}(\mathbf{u}) (\mathbf{\Gamma}_{1k} \Theta)_j, \tag{5}$$

where $\tilde{G}_{jk}(\mathbf{u}) = f_j \circ F_j^{-1}(u_j) \left[E\left\{\varepsilon_{k1} \mathbf{1}(\mathbf{U}_1 \le \mathbf{u}) | U_{j1} = u_j\right\} - \partial_{u_j} C(\mathbf{u}) E\left\{\varepsilon_{k1} | U_{j1} = u_j\right\}\right]$, $j \in \{1,\dots,d\}$, and

$$\check{\mathbb{C}}(s,\mathbf{u}) = \beta(s,\mathbf{u}) - s \sum_{j=1}^{d} \partial_{u_j} C(\mathbf{u}) \beta_j(1,u_j), \quad (s,\mathbf{u}) \in [0,1]^{1+d}. \tag{6}$$

Moreover, $\mathbb{G}_n \rightsquigarrow \mathbb{G}$, where

$$\mathbb{G}(s,\mathbf{u}) = \beta(s,\mathbf{u}) - s\beta(1,\mathbf{u}), \quad (s,u) \in [0,1]^{1+d}. \tag{7}$$

Furthermore, under Model 2 and assumptions (A1)–(A6), $\mathbb{C}_n \rightsquigarrow \check{\mathbb{C}}$.

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 Rémillard [20].

Remark 5. It is remarkable that under Model 2, $\mathbb{C}_n$ converges to $\check{\mathbb{C}}$ defined by (6), which does not depend on $\mathbf{\Theta}$, even if $\mathbb{K}$ does. This important property will play a major role in the next section, where specification tests for the copula are discussed. Also, recall that $\check{\mathbb{C}}(1,\cdot)$ is the asymptotic limit of the empirical copula process constructed from innovations if they were observable; see, e.g., Gänssler and Stute [23], Fermanian et al. [24], Tsukahara [25]. In fact, setting $\check{\mathbf{U}}_{i,n} = \mathbf{R}_i/(n+1)$, where $\mathbf{R}_1,\dots,\mathbf{R}_n$ are the associated rank vectors of $\mathbf{U}_1,\dots,\mathbf{U}_n$, one easily obtains the result that $\check{\mathbb{C}}$ is the asymptotic limit of $\check{\mathbb{C}}_n(s,\mathbf{u}) = \frac{1}{\sqrt{n}} \sum_{i=1}^{\lfloor ns \rfloor} \left\{\mathbf{1}\left(\check{\mathbf{U}}_{i,n} \le \mathbf{u}\right) - C(\mathbf{u})\right\}$, $(s,\mathbf{u}) \in [0,1]^{1+d}$.

Under Model 1, to obtain a limit $\mathbb{C}$ that does not depend on the estimated parameters, it follows from (5) that the following condition is necessary: $(\mathbf{\Gamma}_{1k}\theta)_j = 0$ for all θ and all $j \ne k$, which is equivalent to the condition: $(\mathbf{\Gamma}_{1k})_{jl} = 0$ for all l and all $j \ne k$. The latter occurs for example if

$$\{\sigma_i(\theta)\}_{jk} = \{\sigma_i(\theta)\}_{kk} (A_i)_{jk}, \text{ with } (A_i)_{jj} = 1 \text{ and } A_i \text{ invertible.}$$

In this case A_i must be known since it is parameter-free. This is true in particular if σ_i is diagonal, in which case A_i is the identity matrix. Setting H_i to be the diagonal matrix with $(H_i)_{jj} = (\sigma_i)_{jj}$,

$j \in \{1,\dots,d\}$, then one can rewrite the model as $X_i = \mu_i + A_i H_i \varepsilon_i$, which is a simple rescaling of Model 2. This justifies the restriction to this family for models in the next section.

Semiparametric Estimation of the Copula

As proposed by Xu [26], another method to estimate the copula C, called IFM (Iterated Function Method), is to find parametric estimators for each margin, and then use the estimator K_n of K with these estimated margins. More precisely, if $\hat{\mathbf{F}}_n$ is the vector of estimated parametric marginal distributions, then the copula C is estimated by $\hat{C}_n$ satisfying $K_n = \hat{C}_n \circ \hat{\mathbf{F}}_n$. Even for Model 2, it is easy to check that the limiting process $\hat{\mathbb{C}}$ of $\hat{\mathbb{C}}_n = n^{1/2}\left(\hat{C}_n - C\right)$ will then depend on all estimated parameters. This is another reason why one should always estimate the copula with the ranks.

3. Specification Tests for the Copula

We have seen in Corollary 2 that, under Model 2, the sequential copula process $\mathbb{C}_n$ converges to a limit $\check{\mathbb{C}}$ not depending on the parameters of conditional mean and covariance. In fact, as already said in Section 2 this is the model considered by Chen and Fan [10], where

$$X_{ji} = \mu_{ji}(\boldsymbol{\theta}) + h_{ji}(\boldsymbol{\theta})^{1/2}\varepsilon_{ji}, \quad i \geq 1, \quad j \in \{1,\dots,d\},$$

the innovations $\boldsymbol{\varepsilon}_i = (\varepsilon_{1i},\dots,\varepsilon_{di})^\top$ are independent, and ε_{ji} has mean zero and variance 1 for any $j \in \{1,\dots,d\}$. It amounts to fit the d univariate stochastic volatility models separately, which is often the case in applications. Then, $\boldsymbol{\theta}$ is either estimated by maximum likelihood, which requires assuming parametric families for the marginal distributions F_j of ε_{ji}, or by quasi maximum likelihood. The dependence between the innovations components $\epsilon_{1i},\dots,\epsilon_{di}$ is then modeled by the copula C of $\boldsymbol{\varepsilon}_i$.

The aim of this section is to study tests of goodness-of-fit for parametric copula families, i.e., we want to test the null hypothesis

$$H_0 : C \in \mathcal{C} = \{C_{\boldsymbol{\phi}}; \boldsymbol{\phi} \in \mathcal{P}\},$$

for some parametric family of copula $\mathcal{C}$. Typical families are of the elliptical type (Gaussian and Student copulas) and Archimedean copulas (Clayton, Frank, Gumbel). See, e.g., Joe [27] and Cherubini et al. [28] for general references. As proposed in Dias and Embrechts [29], Chen and Fan [10] and Patton [30], one could also consider mixtures of copulas, while for high-dimensional data, parametric vine models would be useful; see, e.g., Aas et al. [31], Kurowicka and Joe [32], and references therein.

Under H_0, each copula $C_{\boldsymbol{\phi}}$ is assumed to admit a density $c_{\boldsymbol{\phi}}$ satisfying the following assumptions:

(B1) For every $\boldsymbol{\phi} \in \mathcal{P}$, the density $c_{\boldsymbol{\phi}}$ of $C_{\boldsymbol{\phi}}$ admits first and second order derivatives with respect to all components of $\boldsymbol{\phi}$. The gradient (column) vector with respect to $\boldsymbol{\phi}$ is denoted $\dot{c}_{\boldsymbol{\phi}}$, and the Hessian matrix is represented by $\ddot{c}_{\boldsymbol{\phi}}$.

(B2) For arbitrary $\mathbf{u} \in (0,1)^d$ and every $\boldsymbol{\phi}_0 \in \mathcal{P}$, the mappings $\boldsymbol{\phi} \mapsto \dot{c}_{\boldsymbol{\phi}}(\mathbf{u})/c_{\boldsymbol{\phi}}(\mathbf{u})$ and $\boldsymbol{\phi} \mapsto \ddot{c}_{\boldsymbol{\phi}}(\mathbf{u})/c_{\boldsymbol{\phi}}(\mathbf{u})$ are continuous at $\boldsymbol{\phi}_0$.

(B3) For every $\boldsymbol{\phi}_0 \in \mathcal{P}$, there exist a neighborhood $\mathcal{N}$ of $\boldsymbol{\phi}_0$ and $C_{\boldsymbol{\phi}_0}$-integrable functions $h_1, h_2 : \mathbb{R}^d \to \mathbb{R}$ such that for every $\mathbf{u} \in (0,1)^d$,

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

In order to test the null hypothesis that $C \in \mathcal{C}$, i.e., $C = C_{\boldsymbol{\phi}}$ for some $\boldsymbol{\phi} \in \mathcal{P}$, it is natural to consider functionals of the empirical process $\mathbb{P}_n = \sqrt{n}(C_n - C_{\boldsymbol{\phi}_n})$, where $\boldsymbol{\phi}_n$ is an estimator of $\boldsymbol{\phi}$. Specification tests based on $\mathbb{P}_n$ are described in Section 3.1 together with a bootstrapping method to estimate P-values. In Section 3.2, one describes processes constructed from the Rosenblatt transform; their

asymptotic behavior is studied, tests statistics are proposed together with a bootstrapping method. All these tests use the estimation $\boldsymbol{\phi}_n$, hence the importance of finding the asymptotic behavior of $\boldsymbol{\Phi}_n = \sqrt{n}(\boldsymbol{\phi}_n - \boldsymbol{\phi})$. In Section 3.3, we consider the most common estimation methods based on ranks, i.e., $\boldsymbol{\phi}_n = \mathcal{T}_n(\mathbf{U}_{1,n}, \ldots, \mathbf{U}_{n,n})$, for some deterministic function $\mathcal{T}_n$. Finally, estimation methods based on common dependence measures are described in Section 3.4, and arguments in favor of the Rosenblatt transform vs. the copula are given in Section 3.5.

3.1. Test Statistics Based on the Empirical Copula

For testing H_0, one could use the Cramér-von Mises type statistic based on

$$S_n = \int_{[0,1]^d} \mathbb{P}_n^2(\mathbf{u}) dC_n(\mathbf{u}) = \sum_{i=1}^{n} \left\{ C_n(\mathbf{U}_{i,n}) - C_{\boldsymbol{\phi}_n}(\mathbf{U}_{i,n}) \right\}^2. \tag{8}$$

According to Genest et al. [33], S_n is one of the best statistics constructed from $\mathbb{P}_n$ for an omnibus test, and is much more powerful and easier to compute than the Kolmogorov-Smirnov type statistic $\|\mathbb{P}_n\| = \sup_{\mathbf{u}\in[0,1]^d} |\mathbb{P}_n(\mathbf{u})|$. Of course, if the parametric family under H_1 is specified, then one can find better tests statistics than S_n. See, e.g., Berg and Quessy [34]. As in Genest and Rémillard [18], assume, for identifiability purposes, that for every $\delta > 0$,

$$\inf \left\{ \sup_{\mathbf{u}\in[0,1]^d} |C_{\boldsymbol{\phi}}(\mathbf{u}) - C_{\boldsymbol{\phi}_0}(\mathbf{u})| : \boldsymbol{\phi} \in \mathcal{P} \text{ and } |\boldsymbol{\phi} - \boldsymbol{\phi}_0| > \delta \right\} > 0.$$

Furthermore, the mapping $\boldsymbol{\phi} \mapsto C_{\boldsymbol{\phi}}$ is assumed to be Fréchet differentiable with derivative $\dot{C}$, i.e., for all $\boldsymbol{\phi}_0 \in \mathcal{P}$,

$$\lim_{h\to 0} \sup_{\mathbf{u}\in(0,1)^d} \frac{|C_{\boldsymbol{\phi}_0+h}(\mathbf{u}) - C_{\boldsymbol{\phi}_0}(\mathbf{u}) - \dot{C}(\mathbf{u})h|}{r(\mathbf{u})\|h\|} = 0, \tag{9}$$

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 example, in the Gaussian copula case, one can take $r(\mathbf{u}) \equiv 1$.

Remark 6. Although S_n was the best test statistic amongst those considered in Genest et al. [33], it might not remain the case here due to the extra level of variability induced by using residuals. It would be interesting to reproduce the study in Genest et al. [33] by using residuals of GARCH models to check if the same hierarchy of tests is obtained.

Before stating the main result of the section, one needs to extend the notion of regularity of $\boldsymbol{\phi}_n$ as defined in Genest and Rémillard [18]. To this end, define

$$\mathbb{W}_n = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \frac{\dot{c}(\mathbf{U}_i)}{c(\mathbf{U}_i)}. \tag{10}$$

One says that $\boldsymbol{\phi}_n$ is regular for $\boldsymbol{\phi}$ if $(\alpha_n, \mathbb{W}_n, \boldsymbol{\Phi}_n) \rightsquigarrow (\alpha, \mathbb{W}, \boldsymbol{\Phi})$ where the latter is centered Gaussian with $E\left(\boldsymbol{\Phi}\mathbb{W}^\top\right) = I$, and $\boldsymbol{\Phi}$ does not depend on $\boldsymbol{\theta}$ or $\boldsymbol{\Theta}$. It is an immediate consequence of the delta method that the property of being regular is preserved by homeomorphisms. The basic result for testing goodness-of-fit using $\mathbb{P}_n$ is stated next and its proof is given in the Appendix B.

Proposition 1. *Under Model 2 and assumptions (A1)–(A6), if $\boldsymbol{\phi}_n$ is regular for $\boldsymbol{\phi}$, then $\mathbb{P}_n \rightsquigarrow \mathbb{P}$, and $S_n \rightsquigarrow S = \int_{[0,1]^d} \mathbb{P}^2(\mathbf{u}) dC(\mathbf{u})$, where $\mathbb{P}(\mathbf{u}) = \check{\mathbb{C}}(1, \mathbf{u}) - \dot{C}(\mathbf{u})^\top \boldsymbol{\Phi}$, $\mathbf{u} \in [0,1]^d$. In fact, if ψ is a continuous function on the space $C([0,1])$, then $T_n = \psi(\mathbb{P}_n) \rightsquigarrow T = \psi(\mathbb{P})$. Moreover, the parametric bootstrap algorithm described next or the two-level parametric bootstrap proposed in Genest et al. [33] can be used to estimate P-values of S_n or T_n.*

Parametric bootstrap for S_n

The following procedure leads to an approximate P-value for the test based on S_n. The changes required for any other function of $\mathbb{P}_n$ are obvious. It can be used only if there is an explicit expression for $C_{\boldsymbol{\phi}}$. Otherwise, the 2-level parametric bootstrap must be used [18].

Algorithm 1: Parametric bootstrap for the empirical copula process.

For some large integer N, do the following steps:

1.- Compute C_n and estimate $\boldsymbol{\phi}$ with $\boldsymbol{\phi}_n = \mathcal{T}_n\left(\mathbf{U}_{1,n}, \ldots, \mathbf{U}_{n,n}\right)$.

2.- Compute the value of S_n, as defined by (8).

3.- Repeat the following steps for every $k \in \{1, \ldots, N\}$:

(a) Generate a random sample $\mathbf{Y}_{1,n}^{(k)}, \ldots, \mathbf{Y}_{n,n}^{(k)}$ from distribution $C_{\boldsymbol{\phi}_n}$ and compute the pseudo-observations $\mathbf{U}_{i,n}^{(k)} = \mathbf{R}_{i,n}^{(k)} / (n+1)$, where $\mathbf{R}_{1,n}^{(k)}, \ldots, \mathbf{R}_{n,n}^{(k)}$ are the associated rank vectors of $\mathbf{Y}_{1,n}^{(k)}, \ldots, \mathbf{Y}_{n,n}^{(k)}$.

(b) Set

$$C_n^{(k)}(\mathbf{u}) = \frac{1}{n} \sum_{i=1}^{n} \mathbf{1}\left(\mathbf{U}_{i,n}^{(k)} \leq \mathbf{u}\right), \quad \mathbf{u} \in [0,1]^d$$

and estimate $\boldsymbol{\phi}$ by $\boldsymbol{\phi}_n^{(k)} = \mathcal{T}_n\left(\mathbf{U}_{1,n}^{(k)}, \ldots, \mathbf{U}_{n,n}^{(k)}\right)$.

(c) Compute

$$S_n^{(k)} = \sum_{i=1}^{n} \left\{ C_n^{(k)}\left(\mathbf{U}_{i,n}^{(k)}\right) - C_{\boldsymbol{\phi}_n^{(k)}}\left(\mathbf{U}_{i,n}^{(k)}\right) \right\}^2.$$

An approximate P-value for the test is then given by $\sum_{k=1}^{N} \mathbf{1}\left(S_n^{(k)} > S_n\right) / N$.

An important feature of Algorithm 1 is that one generates observations from the copula $C_{\boldsymbol{\phi}_n}$, instead of having to generate the whole process $\mathbf{X}_i$ and re-estimate $\boldsymbol{\theta}$ and $\boldsymbol{\phi}$ each time. This is possible only because $\mathbb{P}$ does not depend on $\boldsymbol{\Theta}$ or $\boldsymbol{\theta}$. However, for Model 1 where $\mathbb{P}$ depends on $\boldsymbol{\Theta}$ and possibly on margin parameters, fitting the copula requires as much work as fitting K since one needs to generate the whole process $\mathbf{X}_i$ each time!

Remark 7. Some authors, e.g., Kole et al. [7], proposed tests statistics of the Anderson-Darling type, dividing $\mathbb{P}_n(\mathbf{u})$ by $\sqrt{C_{\boldsymbol{\phi}_n}(\mathbf{u})\{1 - C_{\boldsymbol{\phi}_n}(\mathbf{u})\}}$, and then integrating or taking the supremum. As argued in Genest et al. [33] and Ghoudi and Rémillard [35], these tests should be avoided. First, the denominator only makes sense in the univariate case when parameters are not estimated. In the present context, the limiting distribution of such weighted processes has not been proven and in fact, Ghoudi and Rémillard [35] gave an example where the limiting variance of the weighted process is infinite.

3.2. Tests Statistics Based on the Rosenblatt Transform

Instead of using the empirical copula process, one might also use goodness-of-fit tests constructed from the Rosenblatt's transform [36]. Based on recent results of Genest et al. [33], such tests were among the most powerful omnibus tests.

Recall that the Rosenblatt's mapping of a d-dimensional copula C is the mapping $\boldsymbol{\mathcal{R}}$ from $(0,1)^d \to (0,1)^d$ so that $\mathbf{u} = (u_1, \ldots, u_d) \mapsto \boldsymbol{\mathcal{R}}(\mathbf{u}) = (v_1, \ldots, v_d)$ with $v_1 = u_1$ and

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

$k \in \{2, \ldots, d\}$. Rosenblatt's transforms for Archimedean copulas and meta-elliptic copulas are quite easy to compute for any dimension; see, e.g., Rémillard et al. [37]. The usefulness of Rosenblatt's transform lies in the following properties [38]: suppose that $\mathbf{V} \sim \Pi$, where Π is the independence copula, which is equivalent to saying that $\mathbf{V}$ is uniformly distributed over $(0,1)^d$. Recall that the independence copula Π is given by

$$\Pi(u_1, \ldots, u_d) = \prod_{j=1}^{d} u_j, \quad u_1, \ldots, u_d \in [0,1].$$

Then $\boldsymbol{\mathcal{R}}(\mathbf{U}) \sim \Pi$ if and only if $\mathbf{U} \sim C$. In addition, $\boldsymbol{\mathcal{R}}^{-1}(\mathbf{V}) \sim C$. Since $\mathbf{U} = \boldsymbol{\mathcal{R}}^{-1}(\mathbf{V})$ can be computed in a recursive way, this is particularly useful for simulation purposes. It follows that the null hypothesis $H_0 : C \in \{C_{\boldsymbol{\phi}};\ \boldsymbol{\phi} \in \mathcal{P}\}$ can be stated in terms of Rosenblatt's transforms viz.

$$H_0 : \boldsymbol{\mathcal{R}} \in \{\boldsymbol{\mathcal{R}}_{\boldsymbol{\phi}};\ \boldsymbol{\phi} \in \mathcal{P}\}.$$

Using an idea of Breymann et al. [9], extending previous ideas of Durbin [39] and Diebold et al. [40,41], one can build tests for H_0 by comparing the empirical distribution function of $\mathbf{E}_{i,n} = \boldsymbol{\mathcal{R}}_{\boldsymbol{\phi}_n}(\mathbf{U}_{i,n})$, $i \in \{1, \ldots, n\}$, with Π, since under H_0, $\mathbf{E}_{i,n}$ has approximately distribution Π. More precisely, set

$$\mathbb{D}_n(\mathbf{u}) = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \{\mathbf{1}(\mathbf{E}_{i,n} \leq \mathbf{u}) - \Pi(\mathbf{u})\}, \quad \mathbf{u} \in [0,1]^d, \tag{11}$$

and define

$$\begin{aligned} S_n^{(B)} &= \int_{[0,1]^d} \mathbb{D}_n^2(\mathbf{u}) d\mathbf{u} \\ &= \frac{n}{3^d} - \frac{1}{2^{d-1}} \sum_{i=1}^{n} \prod_{k=1}^{d} \left(1 - E_{ki,n}^2\right) + \frac{1}{n} \sum_{i=1}^{n} \sum_{j=1}^{n} \prod_{k=1}^{d} \left(1 - E_{ki,n} \vee E_{kj,n}\right), \end{aligned} \tag{12}$$

where $a \vee b = \max(a, b)$.

To say what is a regular estimators in this setting, one needs to define

$$\mathbb{B}_n(\mathbf{u}) = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \{\mathbf{1}(\mathbf{E}_i \leq \mathbf{u}) - \Pi(\mathbf{u})\}, \quad \mathbf{u} \in [0,1]^d,$$

in terms of $\mathbf{E}_i = \boldsymbol{\mathcal{R}}_{\boldsymbol{\phi}}(\mathbf{U}_i) \sim \Pi$, $i \in \{1, \ldots, n\}$. It is easy to check that check that $(\mathbb{B}_n, \mathbb{W}_n) \rightsquigarrow (\mathbb{B}, \mathbb{W})$, where the joint law is Gaussian, and $\mathbb{B}$ is a Π-Brownian bridge. One then says that $\boldsymbol{\phi}_n$ is regular for $\boldsymbol{\phi}$ if $(\mathbb{B}_n, \mathbb{W}_n, \boldsymbol{\Phi}_n) \rightsquigarrow (\mathbb{B}, \mathbb{W}, \boldsymbol{\Phi})$ where the latter is centered Gaussian with $E\left(\boldsymbol{\Phi}\mathbb{W}^\top\right) = I$, and $\boldsymbol{\Phi}$ does not depend on $\boldsymbol{\theta}$ or $\boldsymbol{\Theta}$.

As in the case of copula processes studied in the previous section, in order to prove the next result, one must assume that $\boldsymbol{\mathcal{R}}_{\boldsymbol{\phi}}$ is Fréchet differentiable, i.e.,

$$\lim_{h \to 0} \sup_{\mathbf{u} \in (0,1)^d} \frac{\|\boldsymbol{\mathcal{R}}_{\boldsymbol{\phi}_0 + h}(\mathbf{u}) - \boldsymbol{\mathcal{R}}_{\boldsymbol{\phi}_0}(\mathbf{u}) - \dot{\boldsymbol{\mathcal{R}}}(\mathbf{u}) h\|}{r(\mathbf{u}) \|h\|} = 0, \tag{13}$$

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 example, in the Gaussian copula case, one can take $r(\mathbf{u}) = 1 + \sum_{j=1}^{d-1} \{\mathcal{N}^{-1}(u_j)\}^2$. One also has to assume that $\boldsymbol{\mathcal{R}}$ is continuously differentiable with respect to $\mathbf{u} \in (0,1)$. One can now state the main result of the section: giving the convergence of the empirical Rosenblatt process.

Recall that $\check{\mathbf{U}}_{i,n} = \mathbf{R}_i/(n+1)$, where $\mathbf{R}_1, \ldots, \mathbf{R}_n$ are the associated rank vectors of $\mathbf{U}_1, \ldots, \mathbf{U}_n$, and let $\check{\mathbf{E}}_{i,n} = \boldsymbol{\mathcal{R}}_{\check{\boldsymbol{\phi}}_n}(\check{\mathbf{U}}_i)$, where $\check{\boldsymbol{\phi}}_n$ is the estimator of $\boldsymbol{\phi}$ calculated with $\check{\mathbf{U}}_{i,n} = \mathbf{R}_i/(n+1)$, $i \in \{1, \ldots, n\}$. Further set

$$\check{\mathbb{D}}_n(\mathbf{u}) = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \{\mathbf{1}(\check{\mathbf{E}}_{i,n} \leq \mathbf{u}) - \Pi(\mathbf{u})\}, \quad \mathbf{u} \in [0,1]^d. \tag{14}$$

Theorem 2. *Under Model 2 and assumptions (A1)–(A6), if $(\boldsymbol{\phi}_n)$ is regular for $\boldsymbol{\phi}$, then $\mathbb{D}_n - \check{\mathbb{D}}_n \rightsquigarrow 0$ and $\check{\mathbb{D}}_n \rightsquigarrow \check{\mathbb{D}}$, with $\check{\mathbb{D}}$ given by*

$$\check{\mathbb{D}}(\mathbf{u}) = \mathbb{B}(\mathbf{u}) - \kappa(\mathbf{u}) - \boldsymbol{\Phi}^\top \boldsymbol{\varrho}(\mathbf{u}),$$

where $\mathbb{B}$ is a Π-Brownian bridge, $E\{\mathbb{B}(\mathbf{u})\mathbb{W}\} = \boldsymbol{\varrho}(\mathbf{u})$, $E\{\kappa(\mathbf{u})\mathbb{W}\} = 0$, and

$$\kappa(\mathbf{u}) = \sum_{j=1}^{d} \sum_{k=1}^{j} E\left\{\mathbf{1}(\tilde{\mathbf{E}} \leq u)\beta_j(1, \tilde{U}_k)\partial_{u_k}\boldsymbol{\mathcal{R}}^{(j)}(\tilde{\mathbf{U}})|\tilde{E}_j = u_j\right\},$$

where $\tilde{\mathbf{U}} \sim C = C_{\boldsymbol{\phi}}$ and $\tilde{\mathbf{E}} = \boldsymbol{\mathcal{R}}(\tilde{\mathbf{U}})$, with $\tilde{\mathbf{U}}$ independent of all other observations.

The proof of Theorem 2 is given in Appendix A.2. Note that as in the case of the copula process, the limiting process $\check{\mathbb{D}}$ does not depend on $\boldsymbol{\theta}$, as if $\boldsymbol{\theta}$ were known. The following result is then a direct application of the continuous mapping theorem.

Proposition 2. *Under the assumptions of Theorem 2, $S_n^{(B)} \rightsquigarrow S^{(B)} = \int_{[0,1]^d} \check{\mathbb{D}}^2(\mathbf{u})d\mathbf{u}$. In fact, if ψ is a continuous function on the space $C([0,1])$, then the statistic $T_n = \psi(\mathbb{D}_n)$ converges in law to $T = \psi(\check{\mathbb{D}})$. Moreover, the parametric bootstrap algorithm described next in Section 3.2 can be used to estimate P-values of $S_n^{(B)}$ or T_n.*

A Parametric Bootstrap for $S_n^{(B)}$

The following algorithm is described in terms of statistic $S_n^{(B)}$ but can be applied easily to any statistic of the form $T_n = \psi(\mathbb{D}_n)$.

Algorithm 2: Parametric bootstrap for the empirical Rosenblatt process.

For some large integer N, do the following steps:

1. Estimate $\boldsymbol{\phi}$ by $\boldsymbol{\phi}_n = \mathcal{T}_n(\mathbf{U}_{1,n}, \ldots, \mathbf{U}_{n,n})$, compute $\mathbb{D}_n$ and $S_n^{(B)}$ according to formulas (11) and (12).
2. For some large integer N, repeat the following steps for every $k \in \{1, \ldots, N\}$:
   (a) Generate a random sample $\mathbf{Y}_{1,n}^{(k)}, \ldots, \mathbf{Y}_{n,n}^{(k)}$ from distribution $C_{\boldsymbol{\phi}_n}$ and compute the pseudo-observations $\mathbf{U}_{i,n}^{(k)} = \mathbf{R}_{i,n}^{(k)}/(n+1)$, where $\mathbf{R}_{1,n}^{(k)}, \ldots, \mathbf{R}_{n,n}^{(k)}$ are the associated rank vectors of $\mathbf{Y}_{1,n}^{(k)}, \ldots, \mathbf{Y}_{n,n}^{(k)}$.
   (b) Estimate $\boldsymbol{\phi}$ by $\boldsymbol{\phi}_n^{(k)} = \mathcal{T}_n\left(\mathbf{U}_{1,n}^{(k)}, \ldots, \mathbf{U}_{n,n}^{(k)}\right)$, and compute and compute $\mathbf{E}_{i,n}^{(k)} = \boldsymbol{\mathcal{R}}_{\boldsymbol{\phi}_n^{(k)}}\left(\mathbf{U}_{i,n}^{(k)}\right)$, $i \in \{1, \ldots, n\}$.
   (c) Let

$$\mathbb{D}_n^{(k)}(\mathbf{u}) = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \left\{\mathbf{1}\left(\mathbf{E}_{i,n}^{(k)} \leq \mathbf{u}\right) - \Pi(\mathbf{u})\right\}, \quad \mathbf{u} \in [0,1]^d$$

   and compute

$$S_{n,k}^{(B)} = \int_{[0,1]^d} \left\{\mathbb{D}_n^{(k)}(\mathbf{u})\right\}^2 d\mathbf{u}.$$

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

3.3. Estimation of Copula Parameters

In what follows, it is shown that many estimation methods produce regular estimators, as needed in Propositions 1–2 and Theorem 2.

3.3.1. Maximum Pseudo-Likelihood Estimators

In Chen and Fan [10], it is shown that under smoothness conditions on the densities $c_{\boldsymbol{\phi}}$ (conditions D, C, and N in their article), the maximum pseudo-likelihood estimator

$$\boldsymbol{\phi}_n = \arg\max_{\boldsymbol{\phi}\in\mathcal{P}} \left\{ \sum_{i=1}^{n} \log c_{\boldsymbol{\phi}}(\mathbf{U}_{i,n}) \right\}$$

is asymptotically Gaussian with covariance matrix depending only on $c_{\boldsymbol{\phi}}$. Therefore, the asymptotic behavior does not depend on the estimation of the parameter $\boldsymbol{\theta}$ required for the evaluation of the residuals! In fact, it has the same representation as the estimator studied by Genest et al. [11] in the serially independent case, i.e., if $\boldsymbol{\theta}$ were known. More precisely, one has

$$\boldsymbol{\Phi}_n = J^{-1}(\mathbb{W}_n - \mathbb{Z}_n) + o_P(1), \tag{15}$$

where $\mathbb{W}_n$ is defined by (10), $\mathbb{Z}_n = \frac{1}{\sqrt{n}} \sum_{j=1}^{d} \sum_{i=1}^{n} Q_j(U_{ji})$, with

$$Q_j(u_j) = \int_{(0,1)^d} \frac{\dot{c}(\mathbf{v})\partial_{v_j} c(\mathbf{v})}{c(\mathbf{v})} \{\mathbf{1}(u_j \le v_j) - v_j\} d\mathbf{v},$$

$j \in \{1, \ldots, d\}$, and where J is the Fisher's information matrix $\int_{(0,1)^d} \frac{\dot{c}(\mathbf{u})\dot{c}(\mathbf{u})^\top}{c(\mathbf{u})} d\mathbf{u}$. Note that $\begin{pmatrix} \mathbb{W}_n \\ \mathbb{Z}_n \end{pmatrix}$ converges in law to $\begin{pmatrix} \mathbb{W} \\ \mathbb{Z} \end{pmatrix} \sim N(0, \Sigma)$, with $\Sigma = \begin{pmatrix} J & 0 \\ 0 & \mathcal{J} \end{pmatrix}$. In particular, $\mathbb{W} \sim N(0, J)$ is independent of $\mathbb{Z} \sim N(0, \mathcal{J})$. It follows that $\boldsymbol{\Phi}_n$ converges in law to $\boldsymbol{\Phi} \sim N\left(0, J^{-1} + J^{-1}\mathcal{J}J^{-1}\right)$. Therefore $\boldsymbol{\phi}_n$ is a regular estimator for $\boldsymbol{\phi}$ since $(\boldsymbol{\Phi}_n, \mathbb{W}_n) \rightsquigarrow (\boldsymbol{\Phi}, \mathbb{W})$ which is centered Gaussian, and $E\left(\boldsymbol{\Phi}\mathbb{W}^\top\right) = I$. Note also that $(\mathbb{B}_n, \mathbb{W}_n, \boldsymbol{\Phi}_n) \rightsquigarrow (\mathbb{B}, \mathbb{W}, \boldsymbol{\Phi})$ where the latter is centered Gaussian, so the assumptions of Theorem 2 are also met.

Remark 8. For Model 1, it is easy to check that under the same conditions, the limiting distribution of $\boldsymbol{\Phi}_n$ depends on $(\Gamma_{1j}\boldsymbol{\Theta})_k$ for all $j \neq k$.

3.3.2. Two-Stage Estimators

In addition to maximum pseudo-likelihood estimators, one may consider a two-stage estimator. That is, suppose that $\boldsymbol{\phi} = \begin{pmatrix} \boldsymbol{\phi}_1 \\ \boldsymbol{\phi}_2 \end{pmatrix}$, and that $\boldsymbol{\phi}_1$ is estimated first by $\boldsymbol{\phi}_{1,n}$, and then $\boldsymbol{\phi}_2$ is estimated, using a pseudo-likelihood with $\boldsymbol{\phi}_{1,n}$ instead of $\boldsymbol{\phi}_1$. Two-stage estimation is often used for elliptical copulas which depend on a correlation matrix $\mathbf{r}$ and possibly other parameters. It is known that $\mathbf{r}$ can be expressed in terms of functions of Kendall's tau, playing the role of $\boldsymbol{\phi}_1$, while the remaining parameters are defined as $\boldsymbol{\phi}_2$. In fact, $\tau_{jk} = \tau(U_{ji}, U_{ki}) = \frac{2}{\pi}\arcsin(r_{jk})$ [42]. For example, in the Student copula case, $\boldsymbol{\phi}_2$ would be the degrees of freedom. Since many estimators could be functions of dependence measures, and the regularity of estimators is preserved by homeomorphisms, one should check that the latter are regular. This is done next in Section 3.4.

Now, decompose also $\mathbb{W}_n$ and $\mathbb{Z}_n$ accordingly. Suppose that $\boldsymbol{\phi}_{1,n}$ is an estimator of $\boldsymbol{\phi}_1$ that is regular in the sense that $\boldsymbol{\Phi}_{1,n} = \sqrt{n}(\boldsymbol{\phi}_{1,n} - \boldsymbol{\phi}_1) \rightsquigarrow \boldsymbol{\Phi}_1 \sim N(0, \Sigma_1)$ and $E(\boldsymbol{\Phi}_1 \mathbb{W}_1^\top) = I$, $E(\boldsymbol{\Phi}_1 \mathbb{W}_2^\top) = 0$. Next, define $\boldsymbol{\phi}_{2,n}$ as the pseudo-likelihood estimator of the reduced log-likelihood viz.

$$\boldsymbol{\phi}_{2,n} = \arg \max_{\boldsymbol{\phi}_2 \in \mathcal{O}_2} \left[\sum_{i=1}^{n} \log c_{\boldsymbol{\phi}_{1,n}, \boldsymbol{\phi}_2}(\mathbf{U}_{i,n}) \right].$$

It is then easy to check that $\mathbb{W}_{2,n} - \mathbb{Z}_{2,n} = J_{21}\boldsymbol{\Phi}_{1,n} + J_{22}\boldsymbol{\Phi}_{2,n} + o_P(1)$, so $\boldsymbol{\Phi}_n \rightsquigarrow \boldsymbol{\Phi} = \begin{pmatrix} \boldsymbol{\Phi}_1 \\ \boldsymbol{\Phi}_2 \end{pmatrix}$, with $\boldsymbol{\Phi}_2 = J_{22}^{-1}(\mathbb{W}_2 - \mathbb{Z}_2 - J_{21}\boldsymbol{\Phi}_1)$. As a result, $E(\boldsymbol{\Phi}_2 \mathbb{W}_1^\top) = 0$ and $E(\boldsymbol{\Phi}_2 \mathbb{W}_2^\top) = I$. This proves that $\boldsymbol{\phi}_n$ is a regular estimator of $\boldsymbol{\phi}$ since $(\boldsymbol{\Phi}_n, \mathbb{W}_n) \rightsquigarrow (\boldsymbol{\Phi}, \mathbb{W})$ which is a centered Gaussian vector with $E(\boldsymbol{\Phi}\mathbb{W}^\top) = I$. Also if $\boldsymbol{\Phi}_1$ does not depend on $\boldsymbol{\theta}$, then $\boldsymbol{\Phi}$ does not either. This is the case if $\boldsymbol{\phi}_{1,n}$ is a function of C_n.

3.4. Estimators Based on Measures of Dependence

In this section, we investigate the asymptotic behavior of four well-known rank-based dependence measures constructed from the residuals: Kendall's tau, Spearman's rho, van der Waerden and Blomqvist's coefficients. The main result is that these measures behave asymptotically like the ones computed from innovations, extending the results of Chen and Fan [10]. This property is remarkable and justifies many results in the literature, where the dependence measures were computed on the residuals. The proofs depend on the asymptotic behavior of the empirical copula process and they are given in Appendix B. Another important property is that these estimators are all regular, which in the present context is equivalent to the following property: If ρ_n and ρ are respectively the empirical and theoretical dependence measures, then $\mathcal{K}_n = \sqrt{n}(\rho_n - \rho) \rightsquigarrow \mathcal{K}$, with $E(\mathcal{K}\mathbb{W}) = \partial_{\boldsymbol{\phi}}\rho$. Moreover, in all cases, $(\mathbb{B}_n, \mathbb{W}_n, \mathcal{K}_n) \rightsquigarrow (\mathbb{B}, \mathbb{W}, \mathcal{K})$ which is centered Gaussian, since $\mathcal{K}_n = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \psi(\mathbf{U}_i) + o_P(1)$, with $E\{\psi(\mathbf{U}_i)\} = 0$ and $E\{\psi^2(\mathbf{U}_i)\} < \infty$.

3.4.1. Kendall's Tau

The empirical Kendall's coefficient for the pairs $(e_{ji,n}, e_{ki,n})$, $i \in \{1, \ldots, n\}$, denoted $\tau_{jk,n}$, is

$$\tau_{jk,n} = \frac{2}{n(n-1)} \left(\text{number of concordant pairs} - \text{number of concordant pairs} \right),$$

where the pairs $(e_{ji,n}, e_{ki,n})$ and $(e_{jl,n}, e_{kl,n})$ are concordant if $(e_{ji,n} - e_{jl,n})(e_{ki,n} - e_{kl,n}) > 0$, $i \neq l$. Otherwise, they are discordant. Its theoretical counterpart is

$$\tau_{jk} = 4 \int_0^1 \int_0^1 C^{(j,k)}(u_j, u_k) dC^{(j,k)}(u_j, u_k) - 1,$$

with values in $[-1, 1]$ and with value 0 under independence.

Proposition 3. *Under Model 2 and assumptions (A1)–(A6), for all* $1 \leq j < k \leq d$,

$$\sqrt{n}(\tau_{jk,n} - \tau_{jk}) = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \left\{ 8C^{(j,k)}(U_{ji}, U_{ki}) - 8U_{ji} - 8U_{ki} + 6 - 2\tau_{jk} \right\} + o_P(1).$$

converge to centered Gaussian variables $\mathcal{Z}_{jk}^K$, *with*

$$E\left(\mathcal{Z}_{jk}^K \mathbb{W}\right) = 8 \int_0^1 \int_0^1 \dot{C}^{(j,k)}(u_j, u_k) dC^{(j,k)}(u_j, u_k) = \partial_{\boldsymbol{\phi}} \left(\tau_{jk}\right).$$

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 16 March 2017 doi:10.20944/preprints201703.0117.v1

Peer-reviewed version available at Econometrics 2017, 5, 13; doi:10.3390/econometrics5010013

3.4.2. Spearman's Rho

Spearman's empirical coefficient $\rho^S_{jk,n}$ is the correlation coefficient of the pairs $(U_{ji,n}, U_{ki,n})$, $i \in \{1,\dots,n\}$, while its theoretical counterpart ρ^S_{jk} is $\text{Cor}(U_{ji}, U_{ki}) = 12\text{Cov}(U_{ji}, U_{ki}) = 12\int_0^1\int_0^1 \left\{C^{(j,k}(u_j,u_k) - u_j u_k\right\} du_j du_k$. It has values in $[-1,1]$ and has value 0 under independence.

Proposition 4. *Under Model 2 and assumptions (A1)–(A6),*

$$\sqrt{n}\left(\rho^S_{jk,n} - \rho^S_{jk}\right) = \frac{12}{\sqrt{n}}\sum_{i=1}^n \Big\{12(U_{ji}-1/2)(U_{ki}-1/2) - \rho^S_{jk} + 6(U_{ji}-1/2)^2 + 6(U_{ki}-1/2)^2 - 1\Big\} + o_P(1).$$

converge to centered Gaussian variables $\mathcal{Z}^S_{jk}$ with

$$E\left(\mathcal{Z}^S_{jk}\mathbb{W}\right) = 12\int_0^1\int_0^1 \dot{C}^{(j,k)}(u_j,u_k)du_j du_k = \partial_{\boldsymbol{\phi}}\left(\rho^S_{jk}\right).$$

3.4.3. Van der Waerden's Coefficient

Let $\mathcal{N}$ and $\mathcal{N}^{-1}$ be respectively the distribution function and the quantile function of the standard Gaussian distribution. Then the van der Waerden's empirical coefficient $\rho^W_{jk,n}$ is the correlation coefficient of the pairs $(Z_{ji,n}, Z_{ki,n})$, $i \in \{1,\dots,n\}$, where $Z_{ji,n} = \mathcal{N}^{-1}(U_{ji,n})$. Its theoretical counterpart ρ^W_{jk} is defined by

$$\text{Cor}(Z_{ji}, Z_{ki}) = E(Z_{ji}Z_{ki}) = \int_0^1\int_0^1 \left\{C^{(j,k}(u_j,u_k) - u_j u_k\right\} d\mathcal{N}^{-1}(u_j)d\mathcal{N}^{-1}(u_k),$$

with $Z_{ji} = \mathcal{N}^{-1}(U_{ji})$. It has values in $[-1,1]$ and has value 0 under independence.

Proposition 5. *Under Model 2 and assumptions (A1)–(A6),*

$$\sqrt{n}\left(\rho^W_{jk,n} - \rho^W_{jk}\right) = \frac{1}{\sqrt{n}}\sum_{i=1}^n \left\{Z_{ji}Z_{ki} - \rho^W_{jk} - \kappa_{jk}(Z_{ji}) - \kappa_{kj}(Z_{ki})\right\} + o_P(1)$$

converge to $\mathcal{Z}^W_{jk}$ centered Gaussian variables, with

$$E\left(\mathcal{Z}^W_{jk}\mathbb{W}\right) = 12\int_0^1\int_0^1 \dot{C}^{(j,k)}(u_j,u_k)d\mathcal{N}^{-1}(u_j)d\mathcal{N}^{-1}(u_k) = \partial_{\boldsymbol{\phi}}\left(\rho^W_{jk}\right),$$

where

$$\kappa_{jk}(z_j) = \int_{\mathbb{R}}\{\mathbf{1}(z_j \le x) - \mathcal{N}(x)E(Z_{k1}|Z_{j1} = x)\}dx$$

and

$$\kappa_{kj}(z_k) = \int_{\mathbb{R}}\{\mathbf{1}(z_k \le y) - \mathcal{N}(y)\}E(Z_{j1}|Z_{k1} = y)dy.$$

3.4.4. Blomqvist's Coefficient

Blomqvist's empirical coefficient $\rho^B_{jk,n}$ is defined as

$$\rho^B_{jk,n} = \frac{4}{n}\sum_{i=1}^n \mathbf{1}(U_{ji,n} \le 1/2, U_{ki,n} \le 1/2) - 1.$$

Its theoretical counterpart $\rho^B_{jk} = 4P(U_{ji} \le 1/2, U_{ki} \le 1/2) - 1$, has values in $[-1,1]$ with value zero under independence.

Proposition 6. *Under Model 2 and assumptions (A1)–(A6),*

$$\begin{aligned}\sqrt{n}\left(\rho^B_{jk,n}-\rho^B_{jk}\right) &= \frac{4}{\sqrt{n}}\sum_{i=1}^{n}\Big[\mathbf{1}(U_{ji}\le 1/2, U_{ki}\le 1/2)-C^{(j,k)}(1/2,1/2)\\ &\quad -\{\mathbf{1}(U_{ji}\le 1/2)-1/2\}\partial_{u_j}C^{(j,k)}(1/2,1/2)\\ &\quad -\{\mathbf{1}(U_{ki}\le 1/2)-1/2\}\partial_{u_k}C^{(j,k)}(1/2,1/2)\Big]+o_P(1)\end{aligned}$$

converge to centered Gaussian variables $\mathcal{Z}^B_{jk}$ with

$$E\left(\mathcal{Z}^B_{jk}\mathbb{W}\right)=4\dot{C}^{(j,k)}(1/2,1/2)=\partial_{\boldsymbol{\phi}}\left(\rho^B_{jk}\right).$$

3.5. Copula vs. Rosenblatt Transform

In general, specification tests based on $\mathbb{P}_n$ are quite powerful [33]; however, if there is no explicit formula for $C_{\boldsymbol{\phi}}$ as in the multivariate Gaussian or Student cases, then one has to rely on simulation methods that are time-consuming [18]. This is why one should use tests based on the Rosenblatt transform. In general, these tests are easier to perform since the Rosenblatt transform is simple to compute, even for multivariate distributions. Moreover, they are also among the most powerful omnibus tests according to Genest et al. [33]. In fact, according to the aforementioned article, the best test statistic is $S_n^{(B)}$. This is the one used here for the application presented next.

4. Example of Application

In this example, we want to model the dependence between the innovations of two time series. In order to be able to make comparisons with Chen and Fan [10], we take the Deutsche Mark/US and Japanese Yen/US exchanges rates, from 28 April 1988 to 31 December 1998. AR(3)-GARCH(1,1) and AR(1)-GARCH(1,1) models were fitted on the 2684 log-returns.

For such a large sample size, one must be sure that there is no structural change point. To this end, univariate change point tests were performed first on the standardized residuals and the null hypothesis was not rejected each time. Then, the copula change point test was performed, leading once again to the non rejection of the null hypothesis, since the p-value was estimated to be 33%, using $N=100$ replications. See Rémillard [20] and [38] (Algorithm 8.E.2) for details.

Since there is no significant structural change in the distribution of the innovations, one can now try to model the dependence between the innovations.

First, the usual standard copula models (Gaussian, Student, Clayton, Frank, Gumbel) were checked for goodness-of-fit, using statistic $S_n^{(B)}$ constructed from the Rosenblatt process. In each case, the null hypothesis was rejected since the P-value was estimated to be 0, using N=100 replications. With this number of replications, one may conclude that the true P-value is less than 5%, rejecting the null hypothesis. Note that, according to Chen and Fan [10], the proposed model was a Student copula. Here, this hypothesis is rejected. Having rejected the standard copula models, one can try to fit a mixture of Gaussian copulas. In this case, the copula is given by

$$\pi_1\Phi_2\left\{\Phi^{-1}(u),\Phi^{-1}(v);\rho_1\right\}+\pi_2\Phi_2\left\{\Phi^{-1}(u),\Phi^{-1}(v);\rho_2\right\},\quad u,v\in(0,1),$$

where $\Phi_2(x,y;\rho)$ is the bivariate distribution function of two standard Gaussian random variables with correlation $\rho\in(-1,1)$, Φ is the distribution function of the univariate standard Gaussian, $\pi_2=1-\pi_1\in(0,1)$, $\rho_1,\rho_2\in(-1,1)$, with $\rho_1\neq\rho_2$. Similar models were proposed by Dias and Embrechts [29], Chen and Fan [10] and Patton [30].

For the mixture of two Gaussian copulas, the null hypothesis is not rejected since the P-value is 84% corresponding to $S_n^{(B)} = 0.0183$, computed with $N = 100$ replications. The parameters of the two Gaussian copulas are $\hat{\rho}_1 = 0.8205$, $\hat{\rho}_2 = 0.3749$, and $\hat{\pi}_1 = 0.4017$, $\hat{\pi}_2 = 0.5983$.

5. Conclusions

The asymptotic behavior of empirical processes constructed from residuals of stochastic volatility models was studied. The results show that one can easily perform tests of change point on the full distribution, the margins or the copula, as if the parameters of the conditional mean and volatility were known. It was also shown that, under Model 2, when the stochastic volatility matrices are diagonal, the empirical copula process and the associated Rosenblatt process also behave as if the parameters were known. This remarkable property makes it possible to construct consistent tests of specification for the copula of innovations using the residuals, as if they were the innovations. Then one can apply all the methodologies recently developed for goodness-of-fit of copulas in a serially independent context.

Acknowledgments: Partial funding in support of this work was provided by the Natural Sciences and Engineering Research Council of Canada and the Fonds de Recherche Nature et Technologie du Québec. The author would also like to thank the referees for helpful comments and suggestions.

Conflicts of Interest: Please disclose conflicts of interest, or add 'The authors declare no conflicts of interest'. .

Appendix A. Proofs of the Main Results

Appendix A.1. Proof of Theorem 1

For any $A \subset S_d = \{1, \ldots, d\}$, set

$$\mu_{A,n}(s,\mathbf{x}) = \frac{1}{\sqrt{n}} \sum_{i=1}^{\lfloor ns \rfloor} \prod_{j \in A} \{\mathbf{1}(e_{ji,n} \le x_j) - \mathbf{1}(\varepsilon_{ji} \le x_j)\} \prod_{k \in A^c} \mathbf{1}(\varepsilon_{ki} \le x_k),$$

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

$$\begin{aligned}
\mathbb{K}_n(s,\mathbf{x}) &= \frac{1}{\sqrt{n}} \sum_{i=1}^{\lfloor ns \rfloor} \left[\sum_{A \subset S_d} \prod_{j \in A} \{\mathbf{1}(e_{ji,n} \le x_j) - \mathbf{1}(\varepsilon_{ji} \le x_j)\} \prod_{j \in A^c} \mathbf{1}(\varepsilon_{ji} \le x_j) - K(\mathbf{x}) \right] \\
&= \frac{1}{\sqrt{n}} \sum_{i=1}^{\lfloor ns \rfloor} \left\{ \prod_{j=1}^{d} \mathbf{1}(\varepsilon_{ji} \le x_j) - K(\mathbf{x}) \right\} \\
&\quad + \frac{1}{\sqrt{n}} \sum_{i=1}^{\lfloor ns \rfloor} \sum_{j=1}^{d} \{\mathbf{1}(e_{ji,n} \le x_j) - \mathbf{1}(\varepsilon_{ji} \le x_j)\} \prod_{k \ne j} \mathbf{1}(\varepsilon_{ki} \le x_k) \\
&\quad + \frac{1}{\sqrt{n}} \sum_{i=1}^{\lfloor ns \rfloor} \sum_{|A|>1} \prod_{j \in A} \{\mathbf{1}(e_{ji,n} \le x_j) - \mathbf{1}(\varepsilon_{ji} \le x_j)\} \prod_{k \in A^c} \mathbf{1}(\varepsilon_{ki} \le x_k) \\
&= \alpha_n(s,\mathbf{x}) + \sum_{j=1}^{d} \mu_{j,n}(s,\mathbf{x}) + \sum_{|A|>1} \mu_{A,n}(s,\mathbf{x}).
\end{aligned}$$

To prove the theorem, it suffices to show that for any $1 \le j \le d$, uniformly in $(s,\mathbf{x})$, $\mu_{j,n}(s,\mathbf{x})$ converges in probability to $s\partial_{x_j} K(\mathbf{x})(\mathbf{\Gamma}_0\Theta)_j + s\sum_{k=1}^{d} G_{jk}(\mathbf{x})(\mathbf{\Gamma}_{1k}\Theta)_j = s\mathbf{\Theta}^\top \mathbf{L}_j(\mathbf{x})$, 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 (A2), (A3) and (A5), one can find $M > 0$ such that if n is large enough, then $P(B_{M,n}) > 1-\delta$, where

$$B_{M,n} = \{\|\Theta_n\| \leq M\} \cap_{i=1}^n \{\|d_{i,n}\| \leq Mr_i\} \cap \left\{\frac{1}{n}\sum_{i=1}^n \|\gamma_{0i}\| \leq M\right\} \cap_{j=1}^d \left\{\frac{1}{n}\sum_{i=1}^n \|\gamma_{1ji}\| \leq M\right\}.$$

Let $\lambda \in (0,1/2)$ be given. Further set $C_{\lambda,n} = \{\max_{1\leq i\leq n}(\|\gamma_{0i}\| + \max_{j\in S_d}|\varepsilon_{ji}|\|\gamma_{1ji}\|)/\sqrt{n} \leq \lambda\}$. By (A4), $P(C_\lambda) \geq 1-\delta$ if n is large enough.

Next, for $\boldsymbol{\kappa} = (\kappa_1, \boldsymbol{\kappa}_2, \kappa_3) \in \mathbb{R}\times\mathbb{R}^p\times\mathbb{R}$, set

$$\eta_{i,n}(\boldsymbol{\kappa}) = \kappa_1 r_i + \{(\gamma_{0i}\boldsymbol{\kappa}_2)_1 + \sum_{k=1}^d \varepsilon_{ki}(\gamma_{1ki}\boldsymbol{\kappa}_2)_1 + \kappa_3\|\gamma_{0i}\| + \kappa_3\sum_{k=1}^d |\varepsilon_{ki}|\|\gamma_{1ki}\|\}/\sqrt{n},$$

and define

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

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

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

The main problem working with residuals is the fact 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 λ 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 $B_{M,n}\cap C_{\lambda,n}\cap\{\|\Theta_n - \boldsymbol{\zeta}\| < \lambda\}$, one has $\eta_{i,n}(-M,\boldsymbol{\zeta},-\lambda) \leq \varepsilon_{1i} - e_{1i,n} \leq \eta_{i,n}(M,\boldsymbol{\zeta},\lambda)$, and $|\varepsilon_{1i}-e_{1i,n}| \leq a_{i,n}$, with $a_{i,n} = \eta_{i,n}(M,0,c)$, where $c = M+1$. Hence $\mathbf{1}\{\varepsilon_{1i}\leq x_1 - \eta_{i,n}(M,-\boldsymbol{\zeta},\lambda)\} \leq \mathbf{1}(e_{1i,n}\leq x_1) \leq \mathbf{1}\{\varepsilon_{1i}\leq x_1+\eta_{i,n}(M,\boldsymbol{\zeta},\lambda)\}$, $|\mathbf{1}(e_{1i,n}\leq x_1) - \mathbf{1}(\varepsilon_{1i}\leq x_1)| \leq \mathbf{1}(x_1 - a_{i,n} < \varepsilon_{1i} \leq x_1 + a_{i,n})$, and

$$\begin{aligned}|\mathbf{1}(e_{2i,n}\leq x_2) - \mathbf{1}(\varepsilon_{2i}\leq x_2)| &\leq \mathbf{1}(x_2 - Mr_i - c\lambda \leq \varepsilon_{2i} \leq x_2 + Mr_i + c\lambda)\\ &\leq \mathbf{1}(x_2 - 2c\lambda < \varepsilon_{2i} \leq x_2 + 2c\lambda),\end{aligned}$$

if $i \geq i_0$, for some i_0, since $r_i \to 0$. As a result, for any $A \supset \{1,2\}$,

$$|\mu_{A,n}(s,\mathbf{x})| \leq \frac{i_1}{\sqrt{n}} + \frac{1}{\sqrt{n}}\sum_{i=1}^n \mathbf{1}(x_1 - a_{i,n} < \varepsilon_{1i} \leq x_1 + a_{i,n})\mathbf{1}(x_2 - 2c\lambda < \varepsilon_{2i} \leq x_2 + 2c\lambda).$$

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

$$\begin{aligned}|\mu_{A,n}(s,\mathbf{x})| \leq\ & \frac{i_1}{\sqrt{n}} + \tilde{\mu}_{12,n}(x_1,x_2+2c\lambda;M,0,c\lambda) - \tilde{\mu}_{12,n}(x_1,x_2-2c\lambda;M,0,c\lambda)\\ & -\tilde{\mu}_{12,n}(x_1,x_2+2c\lambda;-M,0,-c\lambda) + \tilde{\mu}_{12,n}(x_1,x_2-2c\lambda;-M,0,-c\lambda).\end{aligned}$$

The proof will be completed if one can prove the following properties: if κ_3 small enough, then on $B_{M,n}$, and uniformly in $(s,\mathbf{x})$, both $\tilde{\mathbf{L}}_{1,n}(s,\mathbf{x};\boldsymbol{\kappa}) - s\boldsymbol{\kappa}_2^\top\mathbf{L}_1(\mathbf{x})$ and $\check{\mu}_{1,n}(s,\mathbf{x};\boldsymbol{\kappa}) = \tilde{\mu}_{1,n}(s,\mathbf{x};\boldsymbol{\kappa}) - \tilde{\mathbf{L}}_{1,n}(s,\mathbf{x})$ 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 λ can be chosen as small as needed, one may conclude that as $n\to\infty$, $\mu_{j,n}(s,\mathbf{x})$ 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. For the rest of the proof, to simplify notations, assume that $d = 2$.

To prove the first statement, $P\left(\varepsilon_{1i} \leq x_1 + \eta_{i,n}(\boldsymbol{\kappa}), \varepsilon_{2i} \leq z_1, \ldots, \varepsilon_{di} \leq z_{d-1} \middle| \mathcal{F}_{i-1}\right)$ is given by

$$P\left\{\varepsilon_{1i} > 0, \varepsilon_{1i} \leq \frac{x_1 + \kappa_1 r_i + (\gamma_{0i}\boldsymbol{\kappa}_2)_1/\sqrt{n} + \varepsilon_{2i}(\gamma_{12i}\boldsymbol{\kappa}_2)_1/\sqrt{n} + \kappa_3|\varepsilon_{2i}|\|\gamma_{12i}\|/\sqrt{n}}{1 - (\gamma_{11i}\boldsymbol{\kappa}_2)_1/\sqrt{n} - \kappa_3\|\gamma_{11i}\|/\sqrt{n}}, \varepsilon_{2i} \leq x_2 \middle| \mathcal{F}_{i-1}\right\}$$
$$+ P\left\{\varepsilon_{1i} \leq 0, \varepsilon_{1i} \leq \frac{x_1 + \kappa_1 r_i + (\gamma_{0i}\boldsymbol{\kappa}_2)_1/\sqrt{n} + \varepsilon_{2i}(\gamma_{12i}\boldsymbol{\kappa}_2)_1/\sqrt{n} + \kappa_3|\varepsilon_{2i}|\|\gamma_{12i}\|/\sqrt{n}}{1 - (\gamma_{11i}\boldsymbol{\kappa}_2)_1/\sqrt{n} + \kappa_3\|\gamma_{11i}\|/\sqrt{n}}, \varepsilon_{2i} \leq x_2 \middle| \mathcal{F}_{i-1}\right\}$$

It follows from (A1), (A2), (A6) and (A7) that on $B_{M,n}$, there is a constant c_0 so that

$$\sup_{s\in[0,1]} \sup_{\mathbf{x}\in\bar{\mathbb{R}}^d} \left|\tilde{\mathbf{L}}_{1,n}(s,\mathbf{x};\boldsymbol{\kappa}) - s\boldsymbol{\kappa}_2^\top \mathbf{L}_1(\mathbf{x})\right| \leq c_0 \left\{\frac{|\kappa_1|}{\sqrt{n}}\sum_{i=1}^n r_i + \frac{|\kappa_3|}{n}\sum_{i=1}^n \{\|\gamma_{0i}\| + \|\gamma_{11i}\| + \|\gamma_{12i}\|\}\right\},$$

which can be made arbitrarily small with large probability by choosing κ_3 small enough. Under Model 2, $(\gamma_{12i}\boldsymbol{\kappa}_2)_1 = 0$ for any $\boldsymbol{\kappa}_2$, so (A7) is not necessary.

It only remains to show that the partial sum of martingale differences $\check{\mu}_{1,n}(s,\mathbf{x};\boldsymbol{\kappa})$ can be made arbitrarily small by choosing κ_3 small enough. The proof is similar to the proof of Lemmas 7.1–7.2 in Ghoudi and Rémillard [2]. Suppose $1/2 < \nu < 1$ and set $N_n = \lfloor n^\nu \rfloor$. Then, set $y_k = F_1^{-1}(k/N_n)$, $1 \leq k < N_n$. Further, set $y_0 = -\infty$ and $y_{N_n} = +\infty$. Now, if $y_k \leq x_1 < y_{k+1}$, and $z = (x_2, \ldots, x_d)$. First, note that one can cover $\bar{\mathbb{R}}^d$ by a finite number $N_n \times J$ of intervals of the form $[a,b) = [y_k, y_{k+1}) \times [u_l, v_l)$, for which $0 \leq K(y_{k+1}, z) - K(y_k, z) \leq F_1(y_{k+1}) - F_1(y_k) = 1/N_n$.

Set $U_{i,n}(\mathbf{x}) = [\mathbf{1}\{\varepsilon_{i1} \leq x_1 + \eta_{i,n}(\boldsymbol{\kappa})\} - \mathbf{1}(\varepsilon_{1i} \leq x_1)]\,\mathbf{1}(\varepsilon_{2i} \leq x_2)$ and set $V_{i,n}(\mathbf{x}) = E\{U_{i,n}(\mathbf{x})|\mathcal{F}_{i-1}\}$. One cannot work directly with $U_{i,n} - V_{i,n}$. Better bounds are obtained by decomposing $U_{i,n}$ and $V_{i,n}$ as follows: set

$$U_{i,n}^+(\mathbf{x}) = \left[\mathbf{1}\{\varepsilon_{i1} \leq x_1 + \eta_{i,n}^+(\boldsymbol{\kappa})\} - \mathbf{1}(\varepsilon_{1i} \leq x_1)\right]\mathbf{1}(\varepsilon_{2i} \leq x_2),$$

and

$$U_{i,n}^-(\mathbf{x}) = \left[\mathbf{1}(\varepsilon_{1i} \leq x_1) - \mathbf{1}\{\varepsilon_{i1} \leq x_1 - \eta_{i,n}^-(\boldsymbol{\kappa})\}\right]\mathbf{1}(\varepsilon_{2i} \leq x_2).$$

Similarly, set $V_{i,n}^\pm(\mathbf{x}) = E\{U_{i,n}^\pm(\mathbf{x})|\mathcal{F}_{i-1}\}$ and define $\check{\mu}_{1,n}^\pm(s,\mathbf{x};\boldsymbol{\kappa}) = \frac{1}{\sqrt{n}}\sum_{i=1}^{\lfloor ns\rfloor}\{U_{i,n}^\pm(\mathbf{x}) - V_{i,n}^\pm(\mathbf{x})\}$. Then $U_{i,n} - V_{i,n} = U_{i,n}^+ - V_{i,n}^+ - \{U_{i,n}^- - V_{i,n}^-\}$, so $\check{\mu}_{1,n} = \check{\mu}_{1,n}^+ - \check{\mu}_{1,n}^-$. To complete the proof, it is enough to show that $\check{\mu}_{1,n}^\pm$ can be made arbitrarily small. Only the proof for the $+$ part is given, the other one being similar. Now, for $\mathbf{x} \in [y_k, y_{k+1}) \times [u_l, v_l)$, observe that

$$U_{i,n}^+(y_k, u_l) - \mathbf{1}(y_k < \varepsilon_{1i} \leq y_{k+1}) \leq U_{i,n}^+(\mathbf{x}) \leq U_{i,n}^+(y_{k+1}, v_l) + \mathbf{1}(y_k < \varepsilon_{1i} \leq y_{k+1}).$$

Taking expectations over the last inequality and summing over i yield the following bound:

$$\sup_{s\in[0,1]} \sup_{\mathbf{x}\in[y_k,y_{k+1})\times[u_l,v_l)} \left|\check{\mu}_{1,n}^\pm(s,\mathbf{x};\boldsymbol{\kappa})\right|$$
$$\leq \sup_{s\in[0,1]} \max\left\{|\check{\mu}_{1,n}^\pm(s,y_{k+1},v_l;\boldsymbol{\kappa})|, |\check{\mu}_{1,n}^\pm(s,y_k,u_l;\boldsymbol{\kappa})|\right\} + 2\frac{\sqrt{n}}{N_n}$$
$$+ \sup_{s\in[0,1]} |\beta_{1,n}(s,y_{k+1}) - \beta_{1,n}(s,y_k)| + \frac{1}{\sqrt{n}}\sum_{i=1}^n \left\{V_{i,n}^+(y_{k+1},v_l) - V_{i,n}^+(y_k,u_l)\right\}.$$

Next $\xi_{i,n} = U_{i,n}^+ - V_{i,n}^+$ is a martingale difference such that $|\xi_{i,n}| \leq 2$ and $E(\xi_{i,n}^2|\mathcal{F}_{i-1}) = V_{i,n}^+(1 - V_{i,n}^+) \leq V_{i,n}^+$. As a result, from the maximum inequality for martingales,

$$P\left\{\sup_{s\in[0,1]} \max_{1\leq k\leq N_n} \max_{1\leq l\leq J} \left|\check{\mu}_{1,n}^\pm(s,y_k,u_l;\boldsymbol{\kappa})\right| > \lambda_0\right\} \leq (N_n \times j)\lambda_0^{-4} \sup_{\mathbf{x}\in\bar{\mathbb{R}}^d} E\left[\left\{\check{\mu}_{1,n}^\pm(1,\mathbf{x};\boldsymbol{\kappa})\right\}^4\right]$$

which is bounded by $c(N_n \times j)\lambda_0^{-4} \sup_{\mathbf{x}\in\mathbb{R}^d}\left[\frac{16}{n} + \frac{1}{n^2}E\left\{\left(\sum_{i=1}^n V_{i,n}^+(\mathbf{x})\right)^2\right\}\right]$, for some universal constant c.

Using (A3), (A6) and (A7), the latter is $O(N_n/n)$, proving that $\sup_{s\in[0,1]}\max_{1\le k\le N_n}\max_{1\le l\le J}\left|\check{\mu}_{1,n}^{\pm}(y_k,u_l;\boldsymbol{\kappa})\right|$ converges in probability to zero.

Similarly, $\sup_{s\in[0,1]}\max_{1\le k\le N_n}\max_{1\le l\le J}\left|\check{\mu}_{1,n}^{\pm}(y_k,v_l;\boldsymbol{\kappa})\right|$ also converges in probability to zero. Next,

$$\sup_{s\in[0,1]}\max_{1\le k\le N_n}|\beta_{1,n}(s,y_{k+1})-\beta_{1,n}(s,y_k)| = \sup_{s\in[0,1]}\max_{1\le k\le N_n}\left|\tilde{\beta}_{1,n}(s,(k+1)/N_n)-\tilde{\beta}_{1,n}(s,k/N_n)\right|$$

converges in probability to zero, where $\tilde{\beta}_{1,n}$ is the empirical Kiefer process constructed from uniform variables. Finally, set $\tilde{f}_1(\mathbf{x}) = \partial_{x_1}K(\mathbf{x})$ and $g_1(\mathbf{x}) = x_1\tilde{f}_1(\mathbf{x})$. From (A1), (A2), (A6) and (A7), one may conclude that for some constants $c_1,\ldots,c_4$ depending on $\|\tilde{f}_1\|$ and $\|g_1\|$,

$$\begin{aligned}\frac{1}{\sqrt{n}}\sum_{i=1}^n\left\{V_{i,n}^+(y_{k+1},v_l)-V_{i,n}^+(y_k,u_l)\right\}&\\ \le c_1|\boldsymbol{\kappa}_1|\frac{\sum_{i=1}^n r_i}{n} + c_2\max_{0\le k<N_n}\max_{1\le l\le J}|f_1(y_{k+1},v_l)-f_1(y_k,u_l)|&\\ + c_3\max_{0\le k<N_n}\max_{1\le l\le J}|g_1(y_{k+1},v_l)-g_1(y_k,u_l)| + c_4|\kappa_3|.&\end{aligned}$$

This can be made as small as necessary, provided n is large, κ_3 is small and the mesh of the covering is small enough. Hence the result. □

Appendix A.2. Proof of Theorem 2

First, note that $\mathbb{B}_n(\mathbf{u}) = \frac{1}{\sqrt{n}}\sum_{i=1}^n[\mathbf{1}\{\boldsymbol{\mathcal{R}}(\mathbf{U}_i)\le\mathbf{u}\}-\Pi(\mathbf{u})] \rightsquigarrow \mathbb{B}$, where $\mathbb{B}$ is a Π-Brownian bridge. Next, set $\check{H}_n(\mathbf{x}) = \boldsymbol{\mathcal{R}}_{\check{\boldsymbol{\phi}}_n}\{\check{\mathbf{F}}_n(\mathbf{x})\}$ and $H(\mathbf{x}) = \boldsymbol{\mathcal{R}}\{\mathbf{F}(\mathbf{x})\}$, where $\check{\mathbf{F}}_n = (\check{F}_{1,n},\ldots,\check{F}_{d,n})$, with

$$\check{F}_{j,n}(x_j) = \frac{1}{n+1}\sum_{i=1}^n\mathbf{1}(\varepsilon_{ji}\le u_j),\quad j\in\{1,\ldots,d\}.$$

Note that $\check{\mathbf{U}}_{i,n} = \check{\mathbf{F}}_n(\boldsymbol{\varepsilon}_i)$, $i\in\{1,\ldots,n\}$. By hypothesis, $\boldsymbol{\phi}_n = \mathcal{T}_n(\mathbf{U}_{1,n},\ldots,\mathbf{U}_{n,n})$. Further set $\check{\boldsymbol{\phi}}_n = \mathcal{T}_n\{\check{\mathbf{U}}_{1,n},\ldots,\check{\mathbf{U}}_{n,n}\}$. Then $\check{\mathbf{V}}_{i,n} = \check{H}_n(\boldsymbol{\varepsilon}_i)$ and $\mathbf{V}_i = H(\boldsymbol{\varepsilon}_i)$, for all $i\in\{1,\ldots,n\}$.

Now $\sqrt{n}(\check{\boldsymbol{\phi}}_n-\boldsymbol{\phi}) \rightsquigarrow \boldsymbol{\Phi}$, using the results in Sections 3.3–3.4, $\check{\mathbb{H}}_n = \sqrt{n}(\check{H}_n - H) \rightsquigarrow \check{\mathbb{H}}$, and $\|\check{\mathbb{H}}_n\|_r = \sup_{\mathbf{x}}|\check{\mathbb{H}}_n(\mathbf{x})|/r\circ\mathbf{F}(\mathbf{x})$ is tight, where for $j\in\{1,\ldots,d\}$,

$$\check{\mathbb{H}}^{(j)}(\mathbf{x}) = \boldsymbol{\Phi}^\top\dot{\boldsymbol{\mathcal{R}}}^{(j)}\{\mathbf{F}(\mathbf{x})\} + \sum_{k=1}^j\partial_{u_k}\boldsymbol{\mathcal{R}}^{(j)}\{\mathbf{F}(\mathbf{x})\}\beta_k\{1,F_k(u_k)\}.$$

It then follows from the results in Ghoudi and Rémillard [2] that $\check{\mathbb{D}}_n \rightsquigarrow \check{\mathbb{D}}$, where

$$\check{\mathbb{D}}(\mathbf{u}) = \mathbb{B}(\mathbf{u}) - \kappa(\mathbf{u}) - \boldsymbol{\Phi}^\top\left\{\sum_{j=1}^d\boldsymbol{\varrho}_j(\mathbf{u})\right\},\qquad u\in[0,1]^d,$$

with

$$\boldsymbol{\varrho}_j(\mathbf{u}) = E\left\{\dot{\boldsymbol{\mathcal{R}}}^{(j)}(\tilde{\mathbf{U}})\mathbf{1}(\tilde{\mathbf{E}}\le\mathbf{u})|\tilde{E}_j=u_j\right\},\qquad j\in\{1,\ldots,d\}. \tag{16}$$

From computations in Rémillard [19] (Lemma 1), one gets $\boldsymbol{\varrho} = \sum_{j=1}^{d} \boldsymbol{\varrho}_j$, so $\check{\mathbb{D}} = \mathbb{B} - \boldsymbol{\kappa} - \boldsymbol{\Phi}^\top \boldsymbol{\varrho}$. Hence $E\{\mathbb{B}(\mathbf{u})\mathbb{W}\} = \boldsymbol{\varrho}(\mathbf{u})$, as claimed. Next, since we already know that for any $j \in \{1, \ldots, d\}$, $E\{\beta_j(1, u_j)\mathbb{W}\} = 0$, it follows that $E\{\boldsymbol{\kappa}(\mathbf{u})\mathbb{W}\} = 0$, for all $\mathbf{u} \in [0,1]^d$. As a result, for any $\mathbf{u} \in [0,1]^d$,

$$E\left\{\check{\mathbb{D}}(\mathbf{u})\mathbb{W}\right\} = \boldsymbol{\varrho}(\mathbf{u}) - E\left(\boldsymbol{\Phi}^\top \mathbb{W}\right)\boldsymbol{\varrho}(\mathbf{u}) = 0,$$

since any $\boldsymbol{\phi}_n$ in Sections 3.3–3.4 is a regular estimator of $\boldsymbol{\phi}$, implying that $E\left(\boldsymbol{\Phi}\mathbb{W}^\top\right) = I$. It then follows from Genest and Rémillard [18] that the parametric bootstrap work for $\check{\mathbb{D}}_n$. To complete the proof, it only remains to show that $\mathbb{D}_n - \check{\mathbb{D}}_n \rightsquigarrow 0$. To this end, note that $\mathbf{V}_{i,n} = H_n(\mathbf{e}_{i,n})$, where $H_n = \boldsymbol{\mathcal{R}}_{\boldsymbol{\phi}_n} \circ \mathbf{F}_n$, so if $\mathbb{H}_n = \sqrt{n}(H_n - H)$, then $\mathbb{H}_n \rightsquigarrow \mathbb{H}$, where, for all $j \in \{1, \ldots, d\}$,

$$\begin{aligned}
\mathbb{H}^{(j)}(\mathbf{x}) &= \boldsymbol{\Phi}^\top \dot{\boldsymbol{\mathcal{R}}}^{(j)}\{\mathbf{F}(\mathbf{x})\} + \sum_{k=1}^{j} \partial_{u_k} \boldsymbol{\mathcal{R}}^{(j)}\{\mathbf{F}(\mathbf{x})\}\mathbb{F}_k\{1, F_k(u_k)\} \\
&= \check{\mathbb{H}}^{(j)} + \sum_{k=1}^{j} \partial_{u_k} \boldsymbol{\mathcal{R}}^{(j)}\{\mathbf{F}(\mathbf{x})\} f_k(x_k)\left\{(\boldsymbol{\Gamma}_0\Theta)_k + x_k(\boldsymbol{\Gamma}_{1k}\Theta)_k\right\}.
\end{aligned}$$

Next, for $i \in \{1, \ldots, n\}$,

$$\begin{aligned}
V_{ji} - V_{ji,n} &= -\frac{\mathbb{H}_n^{(j)}(\mathbf{e}_{i,n})}{\sqrt{n}} + H^{(j)}(\boldsymbol{\varepsilon}_i) - H^{(j)}(\mathbf{e}_{i,n}) \\
&= -\frac{\mathbb{H}_n^{(j)}(\mathbf{e}_{i,n})}{\sqrt{n}} + \sum_{k=1}^{j} \partial_{u_k} \boldsymbol{\mathcal{R}}^{(j)}(\mathbf{U}_i)\left\{d_{ki,n} + (\gamma_{0i}\Theta_n + \sum_{l=1}^{d} \varepsilon_{li}\gamma_{1li}\Theta_n)_k/\sqrt{n}\right\}.
\end{aligned}$$

It then follows from the proof of Theorem 1, the tightness of $\mathbb{H}$ and Ghoudi and Rémillard [2] (Lemma 5.1) that $\mathbb{D}_n - \check{\mathbb{D}}_n \rightsquigarrow 0$. □

Appendix B. Other Proofs

Before starting the proofs, the following lemma is quite useful in some proofs.

Lemma B.1. *Suppose that C and D and distribution functions on $[0,1]^2$, so that C is a copula and D has mean $1/2$ for each marginal distribution. Then*

$$\int D(u,v)dC(u,v) = \int C(u,v)dD(u,v).$$

Proof. Suppose $(U,V) \sim C$, $(\tilde{U},\tilde{V}) \sim D$, and (U,V) is independent of $(\tilde{U},\tilde{V})$. Then, since C is a copula, $P(U < \tilde{U}) = E(\tilde{U}) = 1/2$, $P(V < \tilde{V}) = E(\tilde{U}) = 1/2$, by hypothesis. As a result,

$$\begin{aligned}
\int D(u,v)dC(u,v) &= E\{D(U,V)\} = E\left\{\mathbf{1}(\tilde{U} \le U, \tilde{V} \le V)\right\} \\
&= 1 - P(U < \tilde{U}) - P(V < \tilde{V}) + E\left\{\mathbf{1}(U < \tilde{U}, V < \tilde{V})\right\} \\
&= 1 - E(\tilde{U}) - E(\tilde{V}) + E\{C(\tilde{U},\tilde{V})\} \\
&= 1 - \frac{1}{2} - \frac{1}{2} + \int C(u,v)dD(u,v) = \int C(u,v)dD(u,v).
\end{aligned}$$

□

Appendix B.1. Proof of Proposition 1

The convergence of $\sqrt{n}(C_n - C)$ follows from Corollary 2 and the joint convergence of $(\alpha_n, \mathbf{\Phi}_n)$ follows from the representation of α_n and the estimators of Sections 3.3–3.4. Using the smoothness of $c_{\boldsymbol{\phi}}$, it follows that $\dot{C} = \partial_{\boldsymbol{\phi}} C_{\boldsymbol{\phi}}$ is continuous and under H_0,

$$\mathbb{P}_n = \sqrt{n}(C_n - C_{\boldsymbol{\phi}_n}) = \mathbb{C}_n(1, \cdot) - \sqrt{n}(C_{\boldsymbol{\phi}_n} - C_{\boldsymbol{\phi}}) = \mathbb{C}_n - \dot{C}^\top \mathbf{\Phi}_n.$$

As a result, $\mathbb{A}_n \rightsquigarrow \mathbb{A} = \mathbb{C} - \dot{C}^\top \mathbf{\Phi} = \check{\mathbb{C}} - \dot{C}^\top \mathbf{\Phi}$. Following Genest and Rémillard [18], the parametric bootstrap approach will work since $E\left(\mathbf{\Phi}\mathbb{W}^\top\right) = I$, as shown in Sections 3.3–3.4. □

Appendix B.2. Proof of Propositions 3–6

To prove Proposition 3, note that

$$\begin{aligned}\int C_n^{(j,k)} dC_n^{(j,k)} &= \int \left\{C_n^{(j,k)} - C^{(j,k)}\right\} dC_n^{(j,k)} + \int C^{(j,k)} dC_n^{(j,k)} \\ &= \int \left\{C_n^{(j,k)} - C^{(j,k)}\right\} dC_n^{(j,k)} + \int C_n^{(j,k)} dC^{(j,k)},\end{aligned}$$

using Lemma B.1, since $C_n^{(j,k)}$ and $C^{(j,k)}$ satisfy the assumptions. Then

$$\begin{aligned}\mathcal{Z}_{jk,n}^K &= \sqrt{n}(\tau_{jk,n} - \tau_{jk}) = 4\sqrt{n}\left\{\int C_n^{(j,k)} dC_n^{(j,k)} - \int C^{(j,k)} dC^{(j,k)}\right\} + o_P(1) \\ &= 4\int \mathbb{C}_n^{(j,k)} dC_n^{(j,k)} + 4\int \mathbb{C}_n^{(j,k)} dC^{(j,k)} + o_P(1) \\ &= 8\int \mathbb{C}_n^{(j,k)} dC_n^{(j,k)} + o_P(1).\end{aligned}$$

Similarly, $\sqrt{n}(\check{\tau}_{jk,n} - \tau_{jk}) = 8\int \check{\mathbb{C}}_n^{(j,k)} dC_n^{(j,k)} + o_P(1)$. By Corollary 2, $\mathbb{C}_n = \check{\mathbb{C}}_n + o_P(1) \rightsquigarrow \check{\mathbb{C}}$, proving that $\mathcal{Z}_{jk,n}^K$ converges to $8\int \mathbb{C}^{(j,k)} dC_n^{(j,k)}$.

Next, it is easy to check that $\hat{\mathbb{C}}_n(\mathbf{u}) = \frac{1}{\sqrt{n}} \sum_{i=1}^n \left\{\mathbf{1}(\mathbf{U}_i \leq \mathbf{u}) - C(\mathbf{u}) - \sum_{j=1}^d \mathbf{1}(U_{ji} \leq u_j)\partial_{u_j} C(\mathbf{u})\right\}$ converges to $\check{\mathbb{C}}$. Hence, for any $1 \leq j < k \leq d$,

$$8\int \hat{\mathbb{C}}_n dC = \frac{1}{\sqrt{n}} \sum_{i=1}^n \left\{8C^{(j,k)}(U_{ji}, U_{ki}) - 4U_{ji} - 4U_{ki} + 2 - 2\tau_{jk}\right\}$$

converges to $\mathcal{Z}_{jk}^K$.

To compute the covariance between $\mathcal{Z}_{jk}^K$ and $\mathbb{W}$, note that $E\left\{\check{\mathbb{C}}(\mathbf{u})\mathbb{W}^\top\right\} = \dot{C}(\mathbf{u})$, and the latter is 0 if $u_j = 1$ for at least $d-1$ indices. As a result,

$$E\left(\mathcal{Z}_{jk}^K \mathbb{W}\right) = 8\int \dot{C}^{(j,k)} dC^{(j,k)} = \partial_{\boldsymbol{\phi}}\left\{\tau_{jk}\right\},$$

using integration by parts, since $\tau_{jk} = 4\int C^{(j,k)} dC^{(j,k)} - 1$. The proof of Propositions 4–6 is similar. It is sufficient to note that for the three estimators, one has

$$\begin{aligned}\sqrt{n}(\rho_{jk,n} - \rho_{jk}) &= \sqrt{n}\left[\int \{L(u_j) - \bar{L}\}\{L(u_k) - \bar{L}\} dC_n^{(j,k)}(u_j, u_k)\right. \\ &\qquad \left. - \int \{L(u_j) - \bar{L}\}\{L(u_k) - \bar{L}\} dC^{(j,k)}(u_j, u_k)\right] + o_P(1) \\ &= \int \mathbb{C}_n^{(j,k)} \{J(x), J(y)\} dxdy + o_P(1),\end{aligned}$$

for an appropriate distribution function J with left-continuous inverse L. For example, $J = \mathcal{N}$ for van der Waerden, J is the distribution of the uniform over $[0, \sqrt{12}]$ for Spearman's rho while J is the distribution function of the discrete random variable taking values 0 and 2 with $p = 1/2$ for Blomqvist's coefficient.

According to Genest and Rémillard [43] and Corollary 2, the latter converges to

$$\int \mathbb{C}(j,k)\{J(x), J(y)\}dxdy. = \int \check{\mathbb{C}}(j,k)\{J(x), J(y)\}dxdy.$$

The representations come from the convergence of $\hat{\mathbb{C}}_n$ to $\check{\mathbb{C}}$. The proof of the covariance with $\mathbb{W}$ can be dealt with in a similar way to the one involving Kendall's tau. □

References

1. Bai, J. Weak convergence of the sequential empirical processes of residuals in ARMA models. *Ann. Stat.* **1994**, *22*, 2051–2061.
2. Ghoudi, K.; Rémillard, B. Empirical processes based on pseudo-observations. II. The multivariate case. In *Asymptotic Methods in Stochastics*; Fields Institute Communications; American Mathematical Society: Providence, RI, USA, 2004; Volume 44, pp. 381–406.
3. Bai, J. Testing Parametric Conditional Distributions of Dynamic Models. *Rev. Econ. Stat.* **2003**, *85*, 531–549.
4. Horváth, L.; Kokoszka, P.; Teyssière, G. Empirical process of the squared residuals of an ARCH sequence. *Ann. Stat.* **2001**, *29*, 445–469.
5. Dobrić, J.; Schmid, F. Testing Goodness of Fit for Parametric families of copulas: Application to financial data. *Commun. Stat. Simul. Comput.* **2005**, *34*, 1053–1068.
6. Dobrić, J.; Schmid, F. A goodness of fit test for copulas based on Rosenblatt's transformation. *Comput. Stat. Data Anal.* **2007**, *51*, 4633–4642.
7. Kole, E.; Koedijk, K.; Verbeek, M. Selecting copulas for risk management. *J. Bank. Financ.* **2007**, *31*, 2405–2423.
8. Panchenko, V. Goodness-of-fit test for copulas. *Physic A* **2005**, *355*, 176–182.
9. Breymann, W.; Dias, A.; Embrechts, P. Dependence structures for multivariate high-frequency data in finance. *Quant. Financ.* **2003**, *3*, 1–14.
10. Chen, X.; Fan, Y. Estimation and model selection of semiparametric copula-based multivariate dynamic models under copula misspecification. *J. Econom.* **2006**, *135*, 125–154.
11. Genest, C.; Ghoudi, K.; Rivest, L.P. A semiparametric estimation procedure of dependence parameters in multivariate families of distributions. *Biometrika* **1995**, *82*, 543–552.
12. Shih, J.H.; Louis, T.A. Inferences on the association parameter in copula models for bivariate survival data. *Biometrics* **1995**, *51*, 1384–1399.
13. Duchesne, P.; Ghoudi, K.; Rémillard, B. On testing for independence between the innovations of several time series. *Can. J. Stat.* **2012**, *40*, 447–479.
14. Oh, D.H.; Patton, A.J. Simulated method of moments estimation for copula-based multivariate models. *J. Am. Stat. Assoc.* **2013**, *108*, 689–700.
15. Sklar, M. Fonctions de répartition à N dimensions et leurs marges. *Publ. Inst. Stat. Univ. Paris* **1959**, *8*, 229–231.
16. Bickel, P.J.; Wichura, M.J. Convergence criteria for multiparameter stochastic processes and some applications. *Ann. Math. Stat.* **1971**, *42*, 1656–1670.
17. Engle, R.F.; Kroner, K.F. Multivariate Simultaneous Generalized ARCH. *Econom. Theory* **1995**, *11*, 122–150.
18. Genest, C.; Rémillard, B. Validity of the parametric bootstrap for goodness-of-fit testing in semiparametric models. *Ann. Inst. H. Poincaré Probab. Statist.* **2008**, *44*, 1096–1127.
19. Rémillard, B. Validity of the Parametric Bootstrap for Goodness-of-Fit Testing in Dynamic Models. SSRN Working Paper Series No. 1966476, 2011. Available online: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=1966476 .
20. Rémillard, B. Non-Parametric Change Point Problems Using Multipliers. SSRN Working Paper Series No. 2043632, 2012. Available online: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2043632 .
21. Segers, J. Asymptotics of empirical copula processes under non-restrictive smoothness assumptions. *Bernoulli* **2012**, *18*, 764–782.

22. Genest, C.; Ghoudi, K.; Rémillard, B. Rank-based extensions of the Brock Dechert Scheinkman test for serial dependence. *J. Am. Stat. Assoc.* **2007**, *102*, 1363–1376.
23. Gänssler, P.; Stute, W. *Seminar on Empirical Processes*; DMV Seminar; Birkhäuser Verlag: Basel, Switzerland, 1987; Volume 9.
24. Fermanian, J.D.; Radulović, D.; Wegkamp, M.H. Weak convergence of empirical copula processes. *Bernoulli* **2004**, *10*, 847–860.
25. Tsukahara, H. Semiparametric estimation in copula models. *Can. J. Stat.* **2005**, *33*, 357–375.
26. Xu, J.J. Statistical Modelling and Inference for Multivariate and Longitudinal Discrete Response Data. Ph.D. Thesis, University of British Columbia, Vancouver, BC, Canada, 1996.
27. Joe, H. *Multivariate Models and Dependence Concepts*; Monographs on Statistics and Applied Probability; Chapman & Hall: London, UK, 1997; Volume 73.
28. Cherubini, U.; Mulinacci, S.; Gobbi, F.; Romagnoli, S. *Dynamic Copula Methods in Finance*; Wiley Finance: New York, NY, USA, 2011.
29. Dias, A.; Embrechts, P. *Dynamic Copula Models for Multivariate High-Frequency Data in Finance*; Technical Report; ETH Zürich: Zurich, Switzerland, 2004.
30. Patton, A.J. Modelling Asymmetric Exchange Rate Dependence. *Int. Econ. Rev.* **2006**, *47*, 527–556.
31. Aas, K.; Czado, C.; Frigessi, A.; Bakken, H. Pair-copula constructions of multiple dependence. *Insur. Math. Econ.* **2009**, *44*, 182–198.
32. Kurowicka, D.; Joe, H. (Eds.) *Dependence Modeling*; Vine Copula Handbook; World Scientific Publishing Co. Pte. Ltd.: Hackensack, NJ, USA, 2011.
33. Genest, C.; Rémillard, B.; Beaudoin, D. Omnibus goodness-of-fit tests for copulas: A review and a power study. *Insur. Math. Econ.* **2009**, *44*, 199–213.
34. Berg, D.; Quessy, J.F. Local power analyses of goodness-of-fit tests for copulas. *Scand. J. Stat.* **2009**, *36*, 389–412.
35. Ghoudi, K.; Rémillard, B. Diagnostic tests for innovations of ARMA models using empirical processes of residuals. In *Asymptotic Laws and Methods in Stochastics*; Number 76 in Fields Institute Communications; Dawson, D., Kulik, R., Haye, M.O., Szyszkowicz, B., Zhao, Y., Eds.; Springer: New York, NY, USA, 2015; pp. 239–282.
36. Rosenblatt, M. Remarks on a multivariate transformation. *Ann. Math. Stat.* **1952**, *23*, 470–472.
37. Rémillard, B.; Papageorgiou, N.; Soustra, F. Copula-based semiparametric models for multivariate time series. *J. Multivar. Anal.* **2012**, *110*, 30–42.
38. Rémillard, B. *Statistical Methods for Financial Engineering*; Chapman and Hall/CRC Financial Mathematics Series; Taylor & Francis: Boca Raton, FL, USA, 2013.
39. Durbin, J. Weak convergence of the sample distribution function when parameters are estimated. *Ann. Stat.* **1973**, *1*, 279–290.
40. Diebold, F.X.; Gunther, T.A.; Tay, A.S. Evaluating Density Forecasts with Applications to Financial Risk Management. *Int. Econ. Rev.* **1998**, *39*, 863–883.
41. Diebold, F.X.; Hahn, J.; Tay, A.S. Multivariate Density Forecast Evaluation and Calibration In Financial Risk Management: High-Frequency Returns on Foreign Exchange. *Rev. Econ. Stat.* **1999**, *81*, 661–673.
42. Fang, H.B.; Fang, K.T.; Kotz, S. The meta-elliptical distributions with given marginals. *J. Multivar. Anal.* **2002**, *82*, 1–16.
43. Genest, C.; Rémillard, B. Tests of independence or randomness based on the empirical copula process. *Test* **2004**, *13*, 335–369.