Econometric Theory, **33**, 2017, 292–330.
doi:10.1017/S0266466615000419

GOODNESS-OF-FIT TESTS FOR MULTIVARIATE COPULA-BASED TIME SERIES MODELS

BETINA BERGHAUS AND AXEL BÜCHER
Ruhr-Universität Bochum

In recent years, stationary time series models based on copula functions became increasingly popular in econometrics to model nonlinear temporal and cross-sectional dependencies. Within these models, we consider the problem of testing the goodness-of-fit of the parametric form of the underlying copula. Our approach is based on a dependent multiplier bootstrap and it can be applied to any stationary, strongly mixing time series. The method extends recent i.i.d. results by Kojadinovic et al. (2011) and shares the same computational benefits compared to methods based on a parametric bootstrap. The finite-sample performance of our approach is investigated by Monte Carlo experiments for the case of copula-based Markovian time series models.

1. INTRODUCTION

It is an issue of fundamental importance in econometrics to understand and model both temporal and cross-sectional dependence in time series. Nowadays, due to their capability of describing nonlinear dependencies, copula functions provide a widely accepted tool to approach this problem, see, e.g., Frees and Valdez (1998), Cebrián, Denuit, and Scaillet (2004), McNeil, Frey, and Embrechts (2005), and Denuit, Goderniaux, and Scaillet (2007) for applications in actuarial sciences and insurance, finance or reliability theory. As a common practice, the copula C used to describe the dependence of interest is assumed to belong to a parametric copula family, parametrized by a finite-dimensional parameter $\theta \in \mathcal{P} \subset \mathbb{R}^p$, i.e.,

$$C \in \mathscr{C} = \{C_\theta : \theta \in \mathcal{P}\}.$$

The authors would like to thank three unknown referees and two editors for their constructive comments on an earlier version of this manuscript, which led to a substantial improvement of the paper. The authors are also thankful to Brendan K. Beare for kindly providing us with the gasoline price markup data and Ivan Kojadinovic for helpful discussions regarding this manuscript. Parts of this paper were written when Axel Bücher was a post-doctoral researcher at Université catholique de Louvain, Belgium, a visiting professor at Universität Heidelberg, Germany, and a visiting professor at Technische Universität Dortmund, Germany. This work has been supported in parts by the Collaborative Research Center "Statistical modeling of nonlinear dynamic processes" (SFB 823) of the German Research Foundation (DFG) and by the IAP research network Grant P7/06 of the Belgian government (Belgian Science Policy), which is gratefully acknowledged. Address correspondence to Betina Berghaus, Ruhr-Universität Bochum, Fakultät für Mathematik, Universitätsstr. 150, 44780 Bochum, Germany. e-mail: betina.berghaus@rub.de, axel.buecher@rub.de.

 © Cambridge University Press 2016

https://doi.org/10.1017/S0266466615000419 Published online by Cambridge University Press

In order to achieve maximal robustness with respect to potential marginal misspecification, the marginals are often allowed to follow arbitrary absolutely continuous cumulative distribution functions (c.d.f.s), in which case we speak of a semiparametric copula model.

In most applications hitherto, copulas are used to model the contemporary dependence of a stationary d-dimensional time series. Serial dependence is often either ignored or filtered out in a preliminary step. Relying on i.i.d. assumptions, several estimation techniques for the unknown parameter θ have been proposed, including the method of moments based on inversion of Spearman's rho or Kendall's tau (Berg and Quessy, 2009), minimum distance type estimators (Tsukahara, 2005) or the pseudo-maximum-likelihood estimator (Genest, Ghoudi, and Rivest, 1995). Chen and Fan (2006a) and Rémillard (2010) developed the asymptotics for the pseudo-ML estimator and moment-based estimators, respectively, in case they are based on residuals of marginally fitted time series models. The asymptotics turn out to be the same as in the i.i.d. case which justifies the common practice of filtering out serial dependence features.

In recent years, it also became increasingly popular to model the temporal dependence of a stationary time series by copulas. Building upon models from Darsow, Nguyen, and Olsen (1992), Chen and Fan (2006b) provide asymptotics for the pseudo-ML estimator for univariate stationary Markovian copula models. In these time series models, copulas are used to model the serial dependence at lag one. We refer to Beare (2010), who derived the corresponding mixing properties, for a number of references to applications of these models. Extensions to the multivariate case, which also cover a joint modeling of the contemporary and the serial dependence, are given in Rémillard, Papageorgiou, and Soustra (2012).

All the afore-mentioned estimation methods depend on the underlying assumption that the unknown copula indeed stems from the model $\mathscr{C}$. Not surprisingly, much research has been devoted to test the assumption $H_0 : C \in \mathscr{C}$, see Genest, Rémillard, and Beaudoin (2009) and Fermanian (2013) for recent reviews in the i.i.d. case. As for instance reported in Scaillet (2007), standardization to a pivotal limiting distribution is possible for some of the test statistics, but the approximation is not as accurate as necessary to obtain appropriate test procedures. Therefore, most of the proposed methods, including those that yield to the best finite-sample performance in large-scale simulation studies, are based on the parametric bootstrap. This makes their application quite heavy with respect to computational feasibility. An attractive alternative, based on a multiplier bootstrap technique, has been proposed by Kojadinovic and Yan (2011) and Kojadinovic, Yan, and Holmes (2011). For approximations of the empirical copula process (a building block for many goodness-of-fit tests for copulas), this multiplier bootstrap technique is known to yield superior results compared to the usual resampling bootstrap, as the latter may introduce ties in the bootstrap sample (see, e.g., Bücher and Dette, 2010).

With respect to temporal dependence and Markovian copula models, much less has been done. To the best of our knowledge, the only reference on goodness-of-fit

https://doi.org/10.1017/S0266466615000419 Published online by Cambridge University Press

tests in these models is given by Rémillard et al. (2012). These authors propose a parametric bootstrap procedure to test the goodness-of-fit and mention that, apart from their procedure, *there exist almost no formal goodness-of-fit tests for copulas in a serially dependent context*.

The present paper tries to fill this gap. Our paper can be regarded as an extension of the afore-mentioned, computationally attractive goodness-of-fit tests in Kojadinovic and Yan (2011) and Kojadinovic et al. (2011) to a general serial dependent setting. Our main assumptions are that the observed time series is strongly mixing and that the parametric estimator allows for a linear expansion in terms of a function of the marginal ranks. Among other settings, this covers the case of multivariate Markovian copula models along with the corresponding pseudo-ML estimator (Chen and Fan, 2006b; Rémillard et al., 2012). The method can also be used to test the contemporary dependence of the raw data rather than that of residuals from marginal time series models.

The test statistic is based on a suitable difference between the empirical copula C_n and a parametric estimator C_{θ_n}, details are given in Section 2. We prove weak convergence of the corresponding process $\sqrt{n}(C_n - C_{\theta_n})$. In order to get access to the critical values of the test statistics, we adapt the dependent multiplier bootstrap of Bücher and Kojadinovic (2015) to the present setting, following ideas of Kojadinovic and Yan (2011) and Kojadinovic et al. (2011). As in the iid case mentioned above, the dependent multiplier bootstrap yields better approximations for the empirical copula process than, e.g., a block bootstrap procedure. Among other things, the validation of the procedure is based on a weak convergence result for the (one-dimensional) multiplier empirical process with respect to weighted metrics, which might be of independent interest and which is presented in Appendix 1. The finite-sample performance of our methods are investigated by Monte Carlo experiments, and they are compared with the tests proposed in Rémillard et al. (2012).

The remaining parts of this paper are organized as follows: in Section 2, we present the basic setting (Section 2.1), introduce the test statistics (Section 2.2) and show how critical values can be obtained via the dependent multiplier bootstrap (Section 2.3). The main result of the paper is given in Theorem 2.2. In Section 3, we discuss some implementation issues, while Section 4 presents simulation results. Finally, Section 5 provides an illustrative data application on gasoline price markups, while all technical details and proofs are deferred to a sequence of appendices. An online supplement (Berghaus and Bücher, 2015) contains additional proofs and details.

2. GOODNESS-OF-FIT TESTS

2.1. The general framework

Even though the common copula-based Markovian time series models provide the main motivation of our work, our methods can be easily worked out in a slightly

https://doi.org/10.1017/S0266466615000419 Published online by Cambridge University Press

more general setting at no additional cost. The special Markovian case is treated in Example 2.1.

Let $\boldsymbol{X} = (X_1, \ldots, X_d)'$ be a d-dimensional random vector with joint cumulative distribution function F and continuous marginal c.d.f.s $F_1, \ldots, F_d$. The copula C of F, or, equivalently, the copula of $\boldsymbol{X}$, is defined as the c.d.f. of the random vector $\boldsymbol{U} = (U_1, \ldots, U_d)'$ that arises from marginal application of the probability integral transform, i.e., $U_l = F_l(X_l)$ for $l = 1, \ldots, d$. By construction, the marginal c.d.f.s of C are standard uniform on $[0, 1]$. By Sklar's Theorem, C is the unique function for which we have

$$F(x_1, \ldots, x_d) = C\{F_1(x_1), \ldots, F_d(x_d)\}$$

for all $\boldsymbol{x} = (x_1, \ldots, x_d) \in \mathbb{R}^d$.

Let $(\boldsymbol{X}_i)_{i\in\mathbb{Z}}$ with $\boldsymbol{X}_i = (X_{i1}, \ldots, X_{id})'$ be a stationary sequence of d-variate random vectors with joint c.d.f. F, continuous marginal c.d.f.s $F_1, \ldots, F_d$ and copula C as described above. For $-\infty \le a \le b \le \infty$, let $\mathcal{F}_a^b$ denote the sigma-field generated by those $\boldsymbol{X}_i$ for which $i \in \{a, a+1, \ldots, b\}$ and define, for $k \ge 1$,

$$\alpha_{\boldsymbol{X}}(k) = \sup\left\{|\mathbb{P}(A \cap B) - \mathbb{P}(A)\mathbb{P}(B)| : A \in \mathcal{F}_{-\infty}^{i}, B \in \mathcal{F}_{i+k}^{\infty}, i \in \mathbb{Z}\right\}$$

as the alpha-mixing coefficient of the time series $(\boldsymbol{X}_i)_{i\in\mathbb{Z}}$. Throughout this paper we will assume that $\alpha_{\boldsymbol{X}}(k) \to 0$ for $k \to \infty$ at a certain rate to be specified later on, i.e., we assume that $(\boldsymbol{X}_i)_{i\in\mathbb{Z}}$ is strongly mixing.

Within this general framework, we observe a "sample" $\boldsymbol{X}_1, \ldots, \boldsymbol{X}_n$ and want to do statistical inference on the unknown copula C. In particular, it is the main purpose of this paper to develop a test for the hypothesis

$$H_0 : C \in \mathscr{C} \quad \text{vs.} \quad H_1 : C \notin \mathscr{C}.$$

This shall be done in a semiparametric way treating the marginals as nuisance parameters, i.e., they remain unspecified.

Example 2.1

The general setting described above includes the important special case of a (potentially multivariate) copula-based Markovian time series model (Rémillard et al., 2012), in short *Markovian copula model*. This model is based on the assumption that the d-dimensional stationary time series $(\boldsymbol{Y}_i)_{i\in\mathbb{Z}}$ of interest is Markovian of order 1 with continuous marginal c.d.f.s $F_1, \ldots, F_d$. Obviously, the distribution of the times series is determined through the marginal c.d.f.s and the $2d$-dimensional copula C of the $2d$-dim. time series $\boldsymbol{X}_i = (\boldsymbol{Y}_{i-1}, \boldsymbol{Y}_i), i \in \mathbb{Z}$. It follows from the results in Beare (2010) that $(\boldsymbol{Y}_i)_{i\in\mathbb{Z}}$ is alpha-mixing for many commonly used copula families, and it is easy to see that this property transfers to $(\boldsymbol{X}_i)_{i\in\mathbb{Z}}$.

https://doi.org/10.1017/S0266466615000419 Published online by Cambridge University Press

2.2. The test statistic

The procedure described in the following is a suitable adaptation of two recent publications of Kojadinovic et al. (2011) and Kojadinovic and Yan (2011) which are restricted to the i.i.d. case. As a test statistic, i.e., as a measure of the goodness-of-fit, it is reasonable to calculate a suitable distance between a nonparametric estimator C_n and a parametric estimator C_{θ_n}. It will be a crucial intermediate step to determine the asymptotic distribution of

$$\sqrt{n}(C_n - C_{\theta_n}) = \sqrt{n}(C_n - C) + \sqrt{n}(C - C_{\theta_n}). \tag{2.1}$$

In the following, we will split the discussion into three steps: one for each summand of the latter decomposition and one for assembling both parts.

Step 1: The nonparametric part. For the first summand on the right-hand side of (2.1), the standardized nonparametric part, we can rely on the empirical copula dating back to Deheuvels (1979) and Rüschendorf (1976). For its definition, we first introduce pseudo-observations from the copula C, defined as $\hat{\boldsymbol{U}}_i = (\hat{U}_{i1}, \ldots, \hat{U}_{id})$ with $\hat{U}_{il} = (n+1)^{-1} R_{il}$ and with R_{il} denoting the rank of X_{il} among $X_{1l}, \ldots, X_{nl}$, $l = 1, \ldots, d$. The empirical copula C_n is defined as the empirical distribution function of $\hat{\boldsymbol{U}}_1, \ldots, \hat{\boldsymbol{U}}_n$, i.e.,

$$C_n(\boldsymbol{u}) = \frac{1}{n} \sum_{i=1}^{n} \mathbb{1}(\hat{\boldsymbol{U}}_i \leq \boldsymbol{u}), \qquad \boldsymbol{u} = (u_1, \ldots, u_d) \in [0,1]^d.$$

The associated standardized version, called the empirical copula process,

$$\mathbb{C}_n : [0,1]^d \to \mathbb{R}, \quad \boldsymbol{u} \mapsto \mathbb{C}_n(\boldsymbol{u}) = \sqrt{n}(C_n - C)(\boldsymbol{u}),$$

is well-understood and has been investigated under slightly different conditions in Rüschendorf (1976), Fermanian, Radulović, and Wegkamp (2004), Segers (2012), Bücher and Volgushev (2013), among others. Weak convergence with respect to the supremum-distance holds provided that there are no ties, marginally, that $\alpha_{\boldsymbol{X}}(k) = O(k^{-a})$ for some $a > 1$ and provided a smoothness condition on C originating from Segers (2012) (see Condition I in Appendix 5) is satisfied, see Bücher and Volgushev (2013), Corollary 2.5.

More precisely, we have $\mathbb{C}_n = \sqrt{n}(C_n - C) \rightsquigarrow \mathbb{C}_C$ in $(\ell^\infty[0,1]^d, \|\cdot\|_\infty)$, where $\mathbb{C}_C$ can be represented as

$$\mathbb{C}_C(\boldsymbol{u}) = \mathbb{B}_C(\boldsymbol{u}) - \sum_{l=1}^{d} C^{[l]}(\boldsymbol{u}) \mathbb{B}_C(\boldsymbol{u}^l), \quad \boldsymbol{u} \in [0,1]^d,$$

with $\boldsymbol{u}^l$ denoting the vector with u_l in the l-th component and the other components equal to 1 and with $C^{[l]}$ denoting the lth partial derivative of C.

https://doi.org/10.1017/S0266466615000419 Published online by Cambridge University Press

Moreover, $\mathbb{B}_C$ denotes a centered Gaussian process on $[0,1]^d$ with covariance function

$$\mathrm{Cov}\{\mathbb{B}_{\mathbb{C}}(\boldsymbol{u}), \mathbb{B}_C(\boldsymbol{v})\} = \sum_{i\in\mathbb{Z}} \mathrm{Cov}\{\mathbb{1}(\boldsymbol{U}_0 \le \boldsymbol{u}), \mathbb{1}(\boldsymbol{U}_i \le \boldsymbol{v})\}, \quad \boldsymbol{u}, \boldsymbol{v} \in [0,1]^d,$$

where $\boldsymbol{U}_i = (U_{i1}, \ldots, U_{id})'$ with $U_{il} = F_l(X_{il})$ for $l = 1, \ldots, d$.

Step 2: The parametric part. For the second summand on the right-hand side of (2.1), the parametric part, let us suppose that the hypothesis H_0 is met and let θ_0 denote the true, unknown parameter in Θ. Several semi-parametric estimators for θ_0 have been proposed, which we denote by θ_n in the following. Among the most popular ones are the maximum pseudo-likelihood estimator (Genest et al., 1995) or the moment-based estimator derived from inversion of the sample version of Spearman's rho (see, for instance, Berg and Quessy, 2009). Under suitable smoothness and mixing conditions, the behavior of $\sqrt{n}(C_{\theta_n} - C)$ can be derived from the asymptotics of $\Theta_n = \sqrt{n}(\theta_n - \theta_0)$. For the latter, we will assume that we can write

$$\Theta_n = \sqrt{n}(\theta_n - \theta_0) = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} J_{\theta_0}(\hat{\boldsymbol{U}}_i) + o_p(1),$$

where, for any $\theta \in \mathcal{P}$, $J_\theta : (0,1)^d \to \mathbb{R}^p$ is a (known) score-function satisfying certain smoothness conditions that are specified in Definition V in the appendix. Note that both of the afore-mentioned semi-parametric estimators allow for such a decomposition. We will show later that, under suitable conditions on the model and on the mixing rate, Θ_n converges weakly to some limit which we denote by Θ. Moreover, we will obtain that

$$\sqrt{n}(C_{\theta_n} - C) = \nabla C_{\theta_0} \Theta_n + o_P(1) \rightsquigarrow \nabla C_{\theta_0} \Theta \tag{2.2}$$

in $(\ell^\infty[0,1]^d, \|\cdot\|_\infty)$, where $\nabla C_\theta \in \mathbb{R}^{1\times p}$ denotes the gradient of $\theta \mapsto C_\theta$.

Step 3: Joining the nonparametric and the parametric part. It will be proved in Theorem 2.2 that the two summands on the right-hand side of (2.1) converge jointly. Hence, we can conclude that

$$\sqrt{n}(C_n - C_{\theta_n}) \rightsquigarrow \mathbb{C}_C - \nabla C_{\theta_0} \Theta$$

in $(\ell^\infty[0,1]^d, \|\cdot\|_\infty)$. The continuous mapping theorem allows to derive the asymptotic distribution of suitable test statistics for H_0. Of particular interest are Kolmogorov-Smirnov-type tests

$$\mathcal{T}_n = \sqrt{n} \sup_{\boldsymbol{u}\in[0,1]^d} |C_n(\boldsymbol{u}) - C_{\theta_n}(\boldsymbol{u})| \tag{2.3}$$

https://doi.org/10.1017/S0266466615000419 Published online by Cambridge University Press

and Cramér-von-Mises tests

$$\mathcal{S}_n = n \int_{[0,1]^d} \{C_n(\boldsymbol{u}) - C_{\theta_n}(\boldsymbol{u})\}^2 \, dC_n(\boldsymbol{u}) = \sum_{i=1}^{n} \{C_n(\hat{\boldsymbol{U}}_i) - C_{\theta n}(\hat{\boldsymbol{U}}_i)\}^2. \tag{2.4}$$

We obtain that $\mathcal{T}_n \rightsquigarrow \mathcal{T}_C$ and $\mathcal{S}_n \rightsquigarrow \mathcal{S}_C$ where

$$\mathcal{T}_C = \sup_{\boldsymbol{u} \in [0,1]^d} |\mathbb{C}_C(\boldsymbol{u}) - \nabla C_{\theta_0}(\boldsymbol{u})\Theta|,$$

$$\mathcal{S}_C = \int_{[0,1]^d} \{\mathbb{C}_C(\boldsymbol{u}) - \nabla C_{\theta_0}(\boldsymbol{u})\Theta\}^2 \, dC(\boldsymbol{u}).$$

We reject the null hypothesis for large values of $\mathcal{T}_n$ or of $\mathcal{S}_n$. To get access to the critical values of the limiting distributions, we propose a bootstrap procedure for the process $\sqrt{n}(C_n - C_{\theta_n})$ in the following subsection.

2.3. Critical values via dependent multiplier bootstrap

In this subsection, we construct a bootstrap approximation for both processes of the decomposition in (2.1).

Step 1: The nonparametric part. The multiplier bootstrap procedure for the empirical copula process is a well-established tool for statistical inference on copulas. It has been proposed in Scaillet (2005) and Rémillard and Scaillet (2009) for i.i.d. data sets, and has been applied to different testing issues in Genest, Nešlehová, and Quessy (2011), Kojadinovic, Segers, and Yan (2011), Bücher, Dette, and Volgushev (2011, 2012), among others. The procedure has been generalized to the alpha-mixing case in Bücher and Ruppert (2013) with several further improvements in Bücher and Kojadinovic (2015).

Let $B \in \mathbb{N}$ be some large integer. For $b \in \{1, \dots, B\}$ and $\boldsymbol{u} \in [0,1]^d$, set

$$\mathbb{C}_n^{(b)}(\boldsymbol{u}) = \mathbb{B}_n^{(b)}(\boldsymbol{u}) - \sum_{l=1}^{d} C_n^{[l]}(\boldsymbol{u}) \mathbb{B}_n^{(b)}(\boldsymbol{u}^l)$$

with

$$\mathbb{B}_n^{(b)}(\boldsymbol{u}) = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} Z_{i,n}^{(b)} \{\mathbb{1}(\hat{\boldsymbol{U}}_i \le \boldsymbol{u}) - C_n(\boldsymbol{u})\}.$$

Here, $(Z_{i,n}^{(b)})_{i \in \mathbb{Z}}$ denote sequences of identically distributed random variables with mean 0 and variance 1, which are independent among $b = 1, \dots, B$ and independent of the original data $(\boldsymbol{X}_i)_{i \in \mathbb{Z}}$. In each sequence, the multipliers can be chosen to be either i.i.d. (provided the original time series is serially independent) or to satisfy some weak dependence condition (see Definition A.1 in Appendix 5). In the latter case, we speak of a *dependent multiplier bootstrap*. A possible way of

https://doi.org/10.1017/S0266466615000419 Published online by Cambridge University Press

simulating these multipliers is specified in Section 3.1. Moreover, $C_n^{[l]}$ denotes an appropriate consistent estimator of $C^{[l]}$, for instance an estimator based on finite-differencing as proposed in Rémillard and Scaillet (2009), i.e.,

$$C_n^{[l]}(\boldsymbol{u}) = \begin{cases} \frac{C_n(\boldsymbol{u}+h_n\boldsymbol{e}_l)-C_n(\boldsymbol{u}-h_n\boldsymbol{e}_l)}{2h_n} & \text{if } u_l \in [h_n, 1-h_n] \\ C_n^{[l]}(u_1,\ldots,u_{l-1},h_n,u_{l+1},\ldots,u_d) & \text{if } u_l \in [0,h_n) \\ C_n^{[l]}(u_1,\ldots,u_{l-1},1-h_n,u_{l+1},\ldots,u_d) & \text{if } u_l \in (1-h_n,1] \end{cases}$$

with a sequence $h_n \downarrow 0$ such that $\inf_n h_n\sqrt{n} > 0$ and with $\boldsymbol{e}_l$ denoting the l-th unit vector in $\mathbb{R}^d$.

Now, under conditions on the mixing rate and on the dependent multipliers, we have

$$\left(\sqrt{n}(C_n - C), \mathbb{C}_n^{(1)}, \ldots, \mathbb{C}_n^{(B)}\right) \rightsquigarrow \left(\mathbb{C}_C, \mathbb{C}_C^{(1)}, \ldots, \mathbb{C}_C^{(B)}\right) \tag{2.5}$$

in $(\ell^\infty[0,1]^d, \|\cdot\|_\infty)^{B+1}$, where $\mathbb{C}_C^{(1)}, \ldots, \mathbb{C}_C^{(B)}$ denote independent copies of the process $\mathbb{C}_C$. This result can be interpreted in the way that, for n large, $\mathbb{C}_n^{(1)}, \ldots, \mathbb{C}_n^{(B)}$ are 'almost' independent copies of $\mathbb{C}_n$ and can hence be used to approximate its distribution. For a proof, see Bücher and Kojadinovic (2015) for the alpha-mixing case or Segers (2012) for the i.i.d. case.

Step 2: The parametric part. For a bootstrap approximation of the parametric component of decomposition (2.1) we will show in the appendix that we can write

$$\Theta_n = \frac{1}{\sqrt{n}}\sum_{i=1}^n J_{\theta_0}(\boldsymbol{U}_i) + \sum_{l=1}^d \left\{ \frac{1}{\sqrt{n}}\sum_{i=1}^n J_{\theta_0}^{[l]}(\boldsymbol{U}_i)(\hat{U}_{i,l} - U_{i,l}) \right\} + o_P(1). \tag{2.6}$$

Observing that $\hat{U}_{i,l} = n^{-1}\sum_{j=1}^n \mathbb{1}(U_{j,l} \le U_{i,l})$ we can write the second (double) sum as

$$\sum_{l=1}^d \left\{ \frac{1}{n^{3/2}}\sum_{i=1}^n\sum_{j=1}^n J_{\theta_0}^{[l]}(\boldsymbol{U}_i)\{\mathbb{1}(U_{j,l} \le U_{i,l}) - U_{i,l}\} \right\}.$$

Introducing the random multipliers $(Z_{i,n}^{(b)})_{i=1,\ldots,n}$ $(b = 1,\ldots,B)$ and replacing each term on the right-hand side of (2.6) by observable estimates we arrive at the following bootstrap approximation for Θ_n,

$$\Theta_n^{(b)} = \frac{1}{\sqrt{n}}\sum_{i=1}^n Z_{i,n}^{(b)} \times \{J_{\theta_n}(\hat{\boldsymbol{U}}_i) + \hat{K}_{i,n,\theta_n}\},$$

where, for any $\theta \in \mathcal{P}$ and $i \in \{1,\ldots,n\}$,

$$\hat{K}_{i,n,\theta} = \sum_{l=1}^d \left\{ \frac{1}{n}\sum_{j=1}^n J_\theta^{[l]}(\hat{\boldsymbol{U}}_j) \times \{\mathbb{1}(\hat{U}_{i,l} \le \hat{U}_{j,l}) - \hat{U}_{j,l}\} \right\}. \tag{2.7}$$

https://doi.org/10.1017/S0266466615000419 Published online by Cambridge University Press

Note that the function J_θ is known for common copula families and common parametric estimators, see also Kojadinovic and Yan (2011) and Kojadinovic et al. (2011). Under H_0 and regularity conditions, it will follow from Theorem 2.2 that

$$\left(\Theta_n, \Theta_n^{(1)}, \ldots, \Theta_n^{(B)}\right) \rightsquigarrow \left(\Theta, \Theta^{(1)}, \ldots, \Theta^{(B)}\right), \tag{2.8}$$

where $\Theta, \Theta^{(1)}, \ldots, \Theta^{(B)}$ denote independent random vectors, each of which has the same distribution as Θ. This result may be of independent interest and can be used to construct confidence sets for the estimation of θ_0 in the alpha-mixing case. In fact, we make use of these confidence bands in the empirical application in Section 5.

Step 3: Joining the nonparametric and the parametric part. As in Section 2.2, the 'nonparametric' convergence in (2.5) and the 'parametric' one in (2.8) actually holds jointly. More precisely, the following theorem summarizes the main results of this paper, including those already mentioned in the previous sections.

THEOREM 2.2. *For $b = 1, \ldots, B$, let $\left(Z_{i,n}^{(b)}\right)_{i\in\mathbb{Z}}$ be independent sequences of multipliers satisfying the conditions in Definition A.1.*

(i) Suppose that Conditions I, II and III in Appendix 5 are met. Then

$$\left(\sqrt{n}(C_n - C), \mathbb{C}_n^{(1)}, \ldots, \mathbb{C}_n^{(B)}\right) \rightsquigarrow \left(\mathbb{C}_C, \mathbb{C}_C^{(1)}, \ldots, \mathbb{C}_C^{(B)}\right) \tag{2.9}$$

in $(\ell^\infty[0,1]^d, \|\cdot\|_\infty)^{B+1}$, where $\mathbb{C}_C^{(1)}, \ldots, \mathbb{C}_C^{(B)}$ denote independent copies of the process $\mathbb{C}_C$.

(ii) If Conditions V and VI hold, if κ from Definition A.1 satisfies $\kappa > 1/\{2(1+\nu)\}$ with ν from Condition V and if $C \in \mathscr{C}$, then

$$\left(\Theta_n, \Theta_n^{(1)}, \ldots, \Theta_n^{(B)}\right) \rightsquigarrow \left(\Theta, \Theta^{(1)}, \ldots, \Theta^{(B)}\right), \tag{2.10}$$

where $\Theta, \Theta^{(1)}, \ldots, \Theta^{(B)}$ denote i.i.d. d-dimensional random vectors that are normally distributed with mean $\mathbf{0}$ and covariance matrix

$$\Sigma = \sum_{i=-\infty}^{\infty} \mathbb{E}[\{J_{\theta_0}(\boldsymbol{U}_0) + K_{0,\theta_0}\}\{J_{\theta_0}(\boldsymbol{U}_i) + K_{i,\theta_0}\}^T],$$

where $K_{i,\theta_0} = \sum_{l=1}^d \int_{[0,1]^d} J_{\theta_0}^{[l]}(\boldsymbol{u})\{\mathbb{1}(U_{i,l} \le u_l) - u_l\} dC_{\theta_0}(\boldsymbol{u})$ and where J_θ is defined in Condition V.

(iii) If Condition IV and all of the conditions in (i) and (ii) are satisfied, then the convergence in (2.9) and (2.10) is jointly in $\{\ell^\infty[0,1]^d\}^{B+1} \times \mathbb{R}^{(B+1)p}$. Moreover,

$$\left(\sqrt{n}(C_n - C_{\theta_n}), \mathbb{C}_n^{(1)} - \nabla C_{\theta_n}\Theta_n^{(1)}, \ldots, \mathbb{C}_n^{(B)} - \nabla C_{\theta_n}\Theta_n^{(B)}\right)$$
$$\rightsquigarrow \left(\mathbb{C}_C - \nabla C_{\theta_0}\Theta, \mathbb{C}_C^{(1)} - \nabla C_{\theta_0}\Theta^{(1)}, \ldots, \mathbb{C}_C^{(B)} - \nabla C_{\theta_0}\Theta^{(B)}\right) \tag{2.11}$$

in $\{\ell^\infty([0,1]^d)\}^{B+1}$.

https://doi.org/10.1017/S0266466615000419 Published online by Cambridge University Press

This result is the basis of the following bootstrap approximations for the Kolomogorov-Smirnov-type statistic $\mathcal{T}_n$ and the Cramér-von-Mises-type statistic $\mathcal{S}_n$ defined in (2.3) and (2.4), respectively, namely

$$\mathcal{T}_n^{(b)} = \sqrt{n} \sup_{\boldsymbol{u} \in [0,1]^d} |\mathbb{C}_n^{(b)}(\boldsymbol{u}) - \nabla C_{\theta_n}(\boldsymbol{u}) \Theta_n^{(b)}|$$

and

$$\begin{aligned}\mathcal{S}_n^{(b)} &= \int_{[0,1]^d} \{\mathbb{C}_n^{(b)}(\boldsymbol{u}) - \nabla C_{\theta_n}(\boldsymbol{u}) \Theta_n^{(b)}\}^2 \, dC_n(\boldsymbol{u}) \\ &= \frac{1}{n} \sum_{i=1}^{n} \{\mathbb{C}_n^{(b)}(\hat{\boldsymbol{U}}_i) - \nabla C_{\theta_n}(\hat{\boldsymbol{U}}_i) \Theta_n^{(b)}\}^2,\end{aligned}$$

for $b = 1, \ldots, B$. For a given level $\alpha \in (0,1)$, let $t_{n,B,1-\alpha}$ and $s_{n,B,1-\alpha}$ denote the empirical $(1-\alpha)$-quantile of the sample $\mathcal{T}_n^{(1)}, \ldots, \mathcal{T}_n^{(B)}$ and $\mathcal{S}_n^{(1)}, \ldots, \mathcal{S}_n^{(B)}$, respectively. The following theorem, whose proof can be found in Section F of the supplement to Bücher and Kojadinovic (2015), shows that the tests which reject H_0 for $\mathcal{T}_n > t_{n,B,1-\alpha}$ or for $\mathcal{S}_n > s_{n,B,1-\alpha}$ asymptotically hold their level, for $B \to \infty$ and $n \to \infty$.

THEOREM 2.3. *Under H_0, if the regularity conditions of Theorem 2.2 are met,*

$$\lim_{B\to\infty} \lim_{n\to\infty} \mathbb{P}(\mathcal{T}_n > t_{n,B,1-\alpha}) = \lim_{B\to\infty} \lim_{n\to\infty} \mathbb{P}(\mathcal{S}_n > s_{n,B,1-\alpha}) = \alpha.$$

An approximate p-value for the test based on $\mathcal{T}_n$ can hence be defined as $\hat{p} = B^{-1} \sum_{b=1}^{B} \mathbb{1}(\mathcal{T}_n^{(b)} > \mathcal{T}_n)$, and similarly for the test based on $\mathcal{S}_n$.

Remark 2.4. For explicitly defined copula families such as, e.g., the Clayton or the Gumbel–Hougaard family, the partial derivative ∇C_θ can easily be calculated. However, in particular for elliptical copulas such as the Gaussian or the t-copula, no explicit formula of the copula function exists. This complicates the calculation of both C_θ and ∇C_θ and hence the application of the developed test statistics. The same computational problem also arises in i.i.d. situations (Kojadinovic and Yan, 2011), and, for the Gaussian and the t-copula, can be resolved by the Plackett formula, see Appendix A in Kojadinovic and Yan (2011), and by applying standard functions from the `R`-package `copula` (Hofert, Kojadinovic, Maechler, and Yan, 2015). The corresponding formulas were also applied in the simulation study in Section 4.

3. IMPLEMENTATION ISSUES

3.1. Practical issues for the choice of the multipliers

The bootstrap version of the test statistics in the previous section depends crucially on a sequence of dependent multipliers satisfying the conditions in Definition A.1 in the appendix. A detailed discussion on how to simulate such a sequence can be found in Bücher and Kojadinovic (2015), Section 5.2, originating

https://doi.org/10.1017/S0266466615000419 Published online by Cambridge University Press

from Bühlmann (1993) and Politis and White (2004). This approach can be easily applied within our context as well, subject to some modifications with respect to the choice of a necessary bandwidth parameter ℓ_n (see Definition A.1). For completeness, the procedure and its modification is described in Appendix E in the supplementary material Berghaus and Bücher (2015).

3.2. Score functions for common estimators

3.2.1. Inversion of Spearman's rho for alpha-mixing observations. One example of an estimator which fulfills the smoothness assumptions in Condition V in the bivariate case is provided by the moment estimator, defined as $\theta_n = \rho^{-1}(\rho_n)$, where ρ_n is the sample version of Spearman's rho and where

$$\rho : \mathcal{P} \to [-1, 1], \quad \rho(\theta) = 12 \int_{[0,1]^2} C_\theta(\boldsymbol{u}) \mathrm{d}\boldsymbol{u} - 3.$$

If the function ρ is bijective and continuously differentiable, this leads to a consistent estimator of the parameter θ_0, with score function given by

$$J_\theta(\boldsymbol{u}) = \frac{1}{\rho'(\theta)} \{12 u_1 u_2 - 3 - \rho(\theta)\}.$$

To see the latter, note that, by the mean value theorem,

$$\sqrt{n}(\theta_n - \theta_0) = \sqrt{n}\{\rho^{-1}(\rho_n) - \rho^{-1}(\rho(\theta_0))\} = A_n + R_n,$$

where $A_n = \sqrt{n} \frac{1}{\rho'(\theta_0)} \{\rho_n - \rho(\theta_0)\}$ and $R_n = \sqrt{n}(\frac{1}{\rho'(\theta_n')} - \frac{1}{\rho'(\theta_0)})\{\rho_n - \rho(\theta_0)\}$ with θ_n' being an intermediate point between θ_0 and θ_n.

We have $R_n = o_P(1)$ by continuity of ρ', consistency of θ_n' and by the fact that $\sqrt{n}(\rho_n - \rho_0) = O_P(1)$ (which follows from the continuous mapping theorem and weak convergence of the empirical copula process for mixing rates $\alpha_{\boldsymbol{X}}(n) = O(n^{-a})$ for some $a > 1$, see Section 2.2).

Moreover, by the definition of ρ_n,

$$\begin{aligned} A_n &= \frac{\sqrt{n}}{\rho'(\theta_0)} \left\{ 1 - \frac{6 \sum_{i=1}^n (\hat{U}_{i1} - \hat{U}_{i2})^2}{n(n-1)^2} - \rho(\theta_0) \right\} \\ &= \frac{1}{\sqrt{n}} \sum_{i=1}^n \left[\frac{1}{\rho'(\theta_0)} \left\{ 12 \hat{U}_{i1} \hat{U}_{i2} - 3 - \rho(\theta_0) \right\} \right] = \frac{1}{\sqrt{n}} \sum_{i=1}^n J_{\theta_0}(\hat{\boldsymbol{U}}_i). \end{aligned}$$

Since J_θ is bounded in $\boldsymbol{u} \in [0,1]^2$, all assumptions in Condition V can be easily shown to hold for most of the most common copula families. For example, for elliptical copulas, we have $\rho(\theta) = \frac{6}{\pi} \arcsin(\frac{\theta}{2})$ whence the corresponding score function is given by

$$J_\theta(\boldsymbol{u}) = \frac{\pi}{3} \sqrt{1 - \frac{\theta^2}{4}} \left\{ 12 u_1 u_2 - 3 - \frac{6}{\pi} \arcsin(\theta/2) \right\}.$$

https://doi.org/10.1017/S0266466615000419 Published online by Cambridge University Press

3.2.2. The pseudo-ML estimator in Markovian copula models. Another estimator satisfying the smoothness assumptions in Condition V is provided by the pseudo-ML estimator in Markovian copula models as described in Example 2.1 above. If M denotes the d-dimensional copula of $\boldsymbol{Y}_i$, then, by stationarity, M is also the copula of $\boldsymbol{Y}_{i-1}$ and for all $\boldsymbol{u} \in [0,1]^d$

$$C(\boldsymbol{u}, \mathbf{1}) = M(\boldsymbol{u}) = C(\mathbf{1}, \boldsymbol{u}).$$

Note that this constraint is satisfied for many commonly used copulas, in particular in lower dimensions. Apparently, it holds for any exchangeable copula such as, e.g., any Archimedean copula or any elliptical copula with a properly chosen covariance matrix. With a view on higher dimensions, parametric families could be conveniently constructed from the class of hierarchical Archimedean copula (Okhrin, Okhrin, and Schmid, 2013) or (simplified) pair copula constructions, in particular from D-vines (Aas, Czado, Frigessi, and Bakken, 2009).

To test for the hypothesis $C \in \{C_\theta : \theta \in \mathcal{P}\}$, where C_θ has density c_θ and where m_θ denotes the density corresponding to the marginal copula $M_\theta = C_\theta(\mathbf{1}, \cdot) = C_\theta(\cdot, \mathbf{1})$ for any $\theta \in \mathcal{P}$, one may use the maximum-pseudo-likelihood estimator for the parameter θ defined as

$$\hat{\theta}_n = \arg\max_{\theta \in \mathcal{P}} \sum_{i=2}^{n} \log\left\{ \frac{c_\theta(\hat{\boldsymbol{U}}_i, \hat{\boldsymbol{U}}_{i-1})}{m_\theta(\hat{\boldsymbol{U}}_{i-1})} \right\}.$$

From the results in Rémillard et al. (2012) it follows that the corresponding score function $J_\theta : (0,1)^{2d} \to \mathbb{R}^p$ is given by

$$J_\theta(\boldsymbol{u}) = \left\{ \mathbb{E}_{C_\theta}\left[\nabla^2 \frac{c_\theta(\boldsymbol{U}_1, \boldsymbol{U}_0)}{m_\theta(\boldsymbol{U}_0)} \right] \right\}^{-1} \nabla\left[\log\left\{ \frac{c_\theta(\boldsymbol{u})}{m_\theta(u_{d+1}, \ldots, u_{2d})} \right\} \right].$$

For example, for the bivariate Gaussian copula with correlation parameter θ, the score function is given by

$$J_\theta(u_1, u_2) = \frac{\theta(1-\theta^2) - \theta\{\Phi^{-1}(u_1)^2 + \Phi^{-1}(u_2)^2\} + (1+\theta^2)\Phi^{-1}(u_1)\Phi^{-1}(u_2)}{1+\theta^2}, \tag{3.1}$$

where $\Phi : \mathbb{R} \to [0,1]$ denotes the c.d.f. of the standard normal distribution (corresponding higher dimensional score functions in the Gaussian case can be calculated based on Appendix A in Kojadinovic and Yan, 2011). Obviously, the latter function is unbounded on $(0,1)^2$, which is typical for most of the commonly used copula families and which is why the assumptions in Condition V are designed to allow for unbounded functions. In fact, it can be shown that the assumptions in Condition V hold for most of the commonly used copulas (see Example A.4 for details on the Gaussian case). For a discussion of very similar assumptions in the case of a one-dimensional Markovian time series, we refer the reader to Chen and Fan (2006b).

https://doi.org/10.1017/S0266466615000419 Published online by Cambridge University Press

4. NUMERICAL RESULTS

In this section, we study the finite-sample performance of the goodness-of-fit test by means of Monte Carlo simulations. We restrict ourselves to the investigation of the one- and two-dimensional Markovian copula model of order one described in Example 2.1, which also allows for a comparison with the tests in Rémillard et al. (2012).

The experimental design for the one-dimensional case is as follows: we consider four common bivariate copula models (the Gaussian, the t_4, the Clayton and the Gumbel–Hougaard copula), six different levels of dependence as measured through Kendall's rank correlation ($\tau = 0.05, 0.1, 0.15, 0.2, 0.25, 0.3$) and five different sample sizes ($n = 50, 100, 200, 400, 800$). For higher values of τ, analogue simulations did not lead to satisfying results, both in terms of the approximation of the nominal level and the power properties of the test. This can be explained by the fact that even moderately dependent copulas induce a rather high level of serial dependence, which results in larger error terms within the proofs and hence a less accurate approximation of the asymptotic distributions for the sample sizes under consideration. The effect of the choice of the strength of dependence on the serial dependence is illustrated in Figure 1, where empirical autocorrelations are plotted as a function of Kendalls tau, for various lags and copula families.

In the two-dimensional case, we consider the same experimental design. The values of Kendall's tau now correspond to the dependence within the bivariate marginal copulas M. Within the elliptical models (Gaussian and t_4), we choose all off-diagonal correlations equal to obtain an exchangeable copula. We needed to omit the Gumbel–Hougaard copula since the simulation of the corresponding time series, based on Algorithm 2 in Rémillard et al. (2012), was not feasible

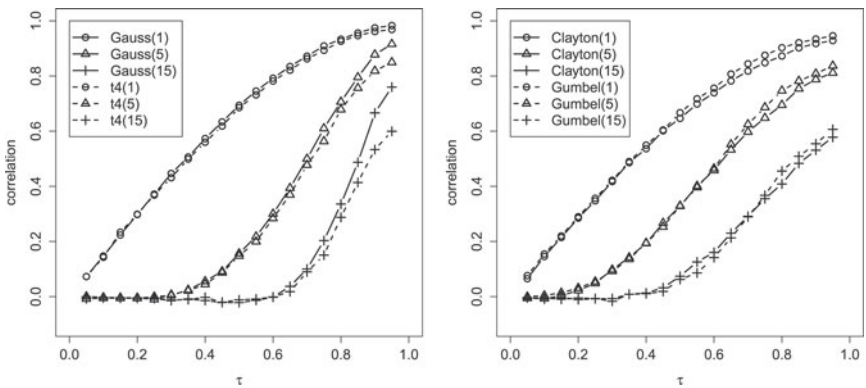


FIGURE 1. Estimated autocorrelations at lag 1 (○), 5 (△) and 15 (+) as a function of Kendall's τ for the Markovian copula model based on the Gaussian copula (left panel, solid lines), the t_4-copula (left panel, dashed lines), the Clayton copula (right panel, solid lines) and the Gumbel copula (right panel, dashed lines).

https://doi.org/10.1017/S0266466615000419 Published online by Cambridge University Press

from a computational perspective: the algorithm is based on the inverse Rosenblatt transform of a certain conditional copula, which, for the Gumbel–Hougaard family, is nonexplicit. The computational cost associated with a numerical evaluation is increasing with the dimension, and becomes unfeasible for even small dimension such as $d = 4$.

For the parametric estimator, we choose the pseudo-ML estimator from Chen and Fan (2006b) as described in Section 3.2.2. The multipliers are generated as described in Section 3.1 (with ξ_i standard normally distributed) and the bandwidth parameter is chosen according to the description in Section E.2 in the supplementary material Berghaus and Bücher (2015).

The corresponding results can be found in Tables 1–3 where we state the empirical rejection probabilities for a nominal level of 5% based on $N = 1,000$ simulation runs, each of which is based on $B = 200$ bootstrap replications. For the sake of brevity, we do not present results for $\tau \in \{0.05, 0.15, 0.25\}$ and for the two-dimensional time series based on the t_4-copula, since the tests

TABLE 1. Simulated rejection probabilities for the goodness-of-fit test in the one-dimensional Markovian copula model based on the Gaussian and the t_4-copula

		True copula: Gaussian					
H_0	n	$\tau = .1$		$\tau = .2$		$\tau = .3$	
Gaussian	50	10.7	(2.4)	5.9	(3.1)	3.1	(3.5)
	100	5.2	(3.4)	3.1	(4.5)	1.3	(5.6)
	200	4.2	(4.7)	3.3	(5.6)	2.2	(5.6)
	400	4.6	(4.5)	3.8	(4.4)	2.4	(4.2)
	800	4.8	(3.5)	2.6	(5.5)	4.4	(4.5)
t_4	50	6.7	(2.2)	4.5	(2.4)	3.9	(2.9)
	100	2.9	(3.4)	4.2	(4.5)	1.6	(5.3)
	200	2.9	(6.1)	3.4	(7.4)	3.1	(6.4)
	400	8.7	(10.5)	9.1	(11.5)	8.5	(11.0)
	800	15.7	(35.0)	29.5	(32.5)	28.7	(29.0)
Clayton	50	28.3	(4.1)	33.2	(5.1)	38.5	(9.1)
	100	23.3	(5.4)	35.3	(11.0)	50.7	(18.9)
	200	25.9	(9.9)	56.2	(21.3)	79.1	(38.3)
	400	35.1	(15.0)	80.0	(42.0)	97.2	(69.5)
	800	56.0	(34.0)	95.9	(82.0)	100.0	(97.5)
Gumbel	50	6.3	(-)	5.9	(-)	6.8	(-)
	100	6.6	(-)	8.7	(-)	7.5	(-)
	200	12.3	(-)	16.9	(-)	21.8	(-)
	400	18.5	(-)	38.3	(-)	49.3	(-)
	800	34.4	(-)	71.5	(-)	87.2	(-)

Table continues on overleaf

https://doi.org/10.1017/S0266466615000419 Published online by Cambridge University Press

TABLE 1. *continued*

		True copula: t_4					
H_0	n	$\tau = .1$		$\tau = .2$		$\tau = .3$	
Gaussian	50	15.6	(4.7)	12.3	(7.5)	7.3	(10.4)
	100	10.6	(7.0)	8.8	(10.2)	6.0	(12.4)
	200	11.5	(12.1)	8.4	(16.8)	7.9	(21.3)
	400	14.3	(20.0)	12.6	(26.0)	8.8	(35.0)
	800	28.4	(44.5)	27.2	(57.0)	23.9	(68.5)
t_4	50	5.6	(3.2)	3.9	(4.8)	2.9	(5.8)
	100	2.1	(3.0)	1.8	(4.4)	0.6	(5.1)
	200	2.4	(4.4)	1.7	(4.6)	1.0	(4.5)
	400	1.6	(4.5)	2.0	(5.0)	1.4	(3.0)
	800	2.2	(9.5)	1.3	(5.0)	1.9	(4.0)
Clayton	50	21.0	(5.5)	30.0	(10.7)	34.2	(15.6)
	100	18.7	(9.5)	25.3	(19.1)	40.9	(27.3)
	200	20.8	(16.2)	42.0	(35.1)	68.3	(49.6)
	400	27.4	(33.5)	64.6	(67.5)	89.1	(82.0)
	800	46.3	(71.5)	92.0	(92.0)	99.9	(99.5)
Gumbel	50	7.6	(-)	8.1	(-)	8.0	(-)
	100	7.2	(-)	8.2	(-)	9.4	(-)
	200	7.9	(-)	12.0	(-)	19.2	(-)
	400	14.3	(-)	23.9	(-)	34.8	(-)
	800	27.2	(-)	49.5	(-)	69.7	(-)

behave similarly, but are slightly more conservative as for the Gaussian copula. The values in brackets refer to the simulated rejection probabilities of the parametric bootstrap-based goodness-of-fit test in Rémillard et al. (2012). Due to computational limitations, we need to omit the two-dimensional case and testing for the Gumbel–Hougaard family. As mentioned above, the simulation of a Markovian time series based on this family is computationally extensive and makes a parametric bootstrap procedure based on many simulated repetitions unfeasible. Note that this problem does not occur for the multiplier bootstrap. The main findings can be summarized as follows:

- The approximation of the nominal level is accurate and gets better with increasing sample size. For the t_4 and the Gaussian copula, the test seems to be globally conservative, while it is slightly too liberal for the Clayton copula.
- We observe decent power properties for Gaussian or t_4 vs. Clayton or Gumbel and vice versa. Also, Gumbel and Clatyon can be well-distinguished. On the other hand, a distinction between the t_4 and the Gaussian copula seems to be more difficult. The latter behavior has also been observed by Kojadinovic et al. (2011) in the i.i.d. context.

https://doi.org/10.1017/S0266466615000419 Published online by Cambridge University Press

- The power of the test is increasing in the strength of the dependence as measured through Kendall's tau. This can be explained by the fact the independence copula is included in all four models under consideration, whence the difference between the models tends to be very small for small values of τ.
- If the dimension increases, the test becomes more conservative whereas the power increases at the same time. This is in accordance with the findings in Kojadinovic and Yan (2012).
- Compared to the test proposed by Rémillard et al. (2012), our test seems to be slightly more conservative in most (but not all, as e.g., for the Clayton copula) situations, with particular advantages for the former test for strongly dependent data. In terms of power, there are situations which are in favor of our test (e.g., H_0 : Gaussian vs. H_1: Clayton) and some which are in favor of the parametric bootstrap procedure (e.g., $H_0 : t_4$ vs. H_1: Gaussian). In general, none of the tests is clearly superior to the others. In terms of runtime, however, the new test clearly outperforms the test by Rémillard et al. (2012),

TABLE 2. Simulated rejection probabilities for the goodness-of-fit test in the one-dimensional Markovian copula model based on the Clayton and the Gumbel–Hougaard copula

		True copula: Clayton					
H_0	n	$\tau = .1$		$\tau = .2$		$\tau = .3$	
Gaussian	50	11.5	(2.7)	9.3	(3.4)	6.7	(7.0)
	100	5.9	(5.2)	9.1	(9.1)	12.6	(16.9)
	200	8.8	(9.1)	23.4	(21.0)	34.2	(40.1)
	400	16.0	(18.0)	50.7	(44.5)	72.8	(73.0)
	800	34.2	(26.5)	85.3	(79.5)	98.3	(97.5)
t_4	50	5.7	(1.6)	4.0	(2.3)	3.2	(4.3)
	100	4.4	(3.6)	4.3	(6.8)	3.4	(10.1)
	200	5.3	(9.1)	11.1	(18.2)	15.0	(30.1)
	400	14.2	(22.0)	30.0	(47.5)	41.0	(64.5)
	800	35.5	(51.0)	64.0	(85.0)	73.0	(97.5)
Clayton	50	17.4	(3.7)	13.0	(2.7)	13.8	(5.2)
	100	7.6	(4.0)	8.2	(5.4)	9.3	(6.6)
	200	5.0	(5.1)	6.6	(5.9)	9.3	(6.1)
	400	5.7	(3.5)	5.7	(5.0)	7.8	(7.5)
	800	5.0	(5.0)	6.5	(3.5)	8.1	(6.0)
Gumbel	50	10.2	(-)	17.6	(-)	17.5	(-)
	100	13.2	(-)	31.6	(-)	40.0	(-)
	200	30.4	(-)	64.2	(-)	83.9	(-)
	400	56.5	(-)	94.0	(-)	99.1	(-)
	800	84.3	(-)	100.0	(-)	100.0	(-)

Table continues on overleaf

https://doi.org/10.1017/S0266466615000419 Published online by Cambridge University Press

TABLE 2. *continued*

		True copula: Gumbel–Hougaard					
H_0	n	$\tau = .1$		$\tau = .2$		$\tau = .3$	
Gaussian	50	13.4	(5.4)	10.7	(6.0)	4.6	(6.1)
	100	7.4	(7.2)	6.1	(8.9)	3.3	(9.6)
	200	10.3	(8.5)	11.6	(13.2)	9.2	(15.9)
	400	13.7	(11.0)	19.3	(20.0)	23.6	(24.0)
	800	22.0	(24.0)	46.1	(31.5)	59.9	(46.0)
t_4	50	6.5	(4.1)	5.6	(4.8)	5.0	(6.3)
	100	3.9	(6.7)	4.9	(9.2)	3.7	(9.9)
	200	4.7	(9.4)	6.5	(15.2)	5.8	(18.7)
	400	10.2	(21.0)	13.4	(26.0)	13.0	(35.5)
	800	19.1	(43.0)	32.7	(50.5)	33.8	(66.0)
Clayton	50	30.0	(8.2)	36.0	(12.8)	35.5	(18.0)
	100	25.0	(14.7)	45.0	(26.3)	52.7	(38.4)
	200	38.4	(24.6)	65.2	(52.3)	84.7	(71.7)
	400	55.8	(44.5)	90.9	(87.0)	97.9	(93.5)
	800	81.7	(71.0)	99.4	(98.5)	100.0	(100.0)
Gumbel	50	5.1	(-)	6.1	(-)	4.2	(-)
	100	5.6	(-)	4.6	(-)	4.8	(-)
	200	6.2	(-)	5.5	(-)	4.6	(-)
	400	4.8	(-)	5.5	(-)	3.8	(-)
	800	6.2	(-)	5.5	(-)	5.7	(-)

the computational cost being up to 70 times higher for the latter. More details on a runtime comparison can be found in Appendix F in the supplementary material on this paper.

5. DATA APPLICATION

In this section, we illustrate the use of our goodness-of-fit test by applying it to a time series of 267 weekly gasoline price markups in Windsor, Ontario, from August 20, 1989, to September 25, 1994. As described in Beare and Seo (2014), these markups were obtained by dividing the average retail price across several gasoline stations in Windsor by the wholesale price of unbranded gasoline at the terminal in Toronto, Ontario. A plot of the time series can be found in Figure 2. The data set was investigated by several authors, see, e.g., Eckert (2002) and McCausland (2007). Recently, Beare and Seo (2014) used Markovian copula models to investigate time reversibility of the time series of price markups. Economic and financial time series frequently exhibit time irreversibility and these authors found indeed overwhelming evidence for time irreversibility of the markup data.

https://doi.org/10.1017/S0266466615000419 Published online by Cambridge University Press

TABLE 3. Simulated rejection probabilities for the goodness-of-fit test in the two-dimensional Markovian copula model based on the Gaussian and the Clayton copula

		True copula: Gaussian			True copula: Clayton		
H_0	n	$\tau = .1$	$\tau = .2$	$\tau = .3$	$\tau = .1$	$\tau = .2$	$\tau = .3$
Gaussian	50	1.8	0.2	0.1	1.3	0.4	0.5
	100	1.7	0.0	0.0	1.1	0.9	1.0
	200	0.9	0.5	0.5	4.2	6.8	5.2
	400	0.6	0.8	0.5	17.5	39.3	41.5
	800	1.6	1.0	2.0	58.8	82.8	86.7
t_4	50	1.9	0.7	1.1	0.9	0.4	0.1
	100	1.0	1.2	1.6	0.4	0.1	0.2
	200	2.9	8.7	14.2	2.6	1.8	0.4
	400	12.4	37.7	51.4	9.1	7.9	6.1
	800	51.2	80.8	87.9	32.5	24.7	20.3
Clayton	50	17.6	26.3	34.9	5.7	4.2	4.6
	100	23.6	46.8	68.6	4.0	3.9	4.8
	200	41.2	82.3	95.3	4.2	4.0	6.9
	400	68.9	98.7	99.9	4.0	5.0	8.3
	800	92.7	100.0	100.0	4.9	5.8	7.9
Gumbel	50	14.2	13.6	7.8	10.0	8.7	6.0
	100	20.9	24.1	16.5	15.5	15.9	12.3
	200	35.6	52.4	51.7	25.7	43.3	39.7
	400	60.2	89.8	92.7	51.1	87.2	87.6
	800	90.8	100.0	100.0	90.5	99.9	100.0

Loosely speaking, time reversibility means that the probabilistic behavior of a time series $(X_i)_{i\in\mathbb{Z}}$ does not alter by reversing the direction of time. Formally, it is defined by the requirement that

$$(X_{i_1}, \ldots, X_{i_n}) \stackrel{d}{=} (X_{i_n}, \ldots, X_{i_1}) \text{ for all } i_1 < \cdots < i_n.$$

For a Markovian copula model, it is equivalent to exchangeability of the copula of (X_{i-1}, X_i), i.e., to the fact that $C(u, v) = C(v, u)$ for all $u, v \in [0, 1]$, see, e.g., Beare and Seo (2014). Therefore, the results on the markup data in the last-named paper suggest that our goodness-of-fit tests should reject any exchangeable copula model for C.

We run our procedure based on $B = 1,000$ bootstrap replications for the four copulas considered in the simulation study in Section 4. The corresponding p-values are $p = 0.002$ for the Gaussian copula, $p = 0.471$ for the t_4-copula, $p = 0.016$ for the Clayton copula, and $p = 0.011$ for the Gumbel–Hougaard

https://doi.org/10.1017/S0266466615000419 Published online by Cambridge University Press

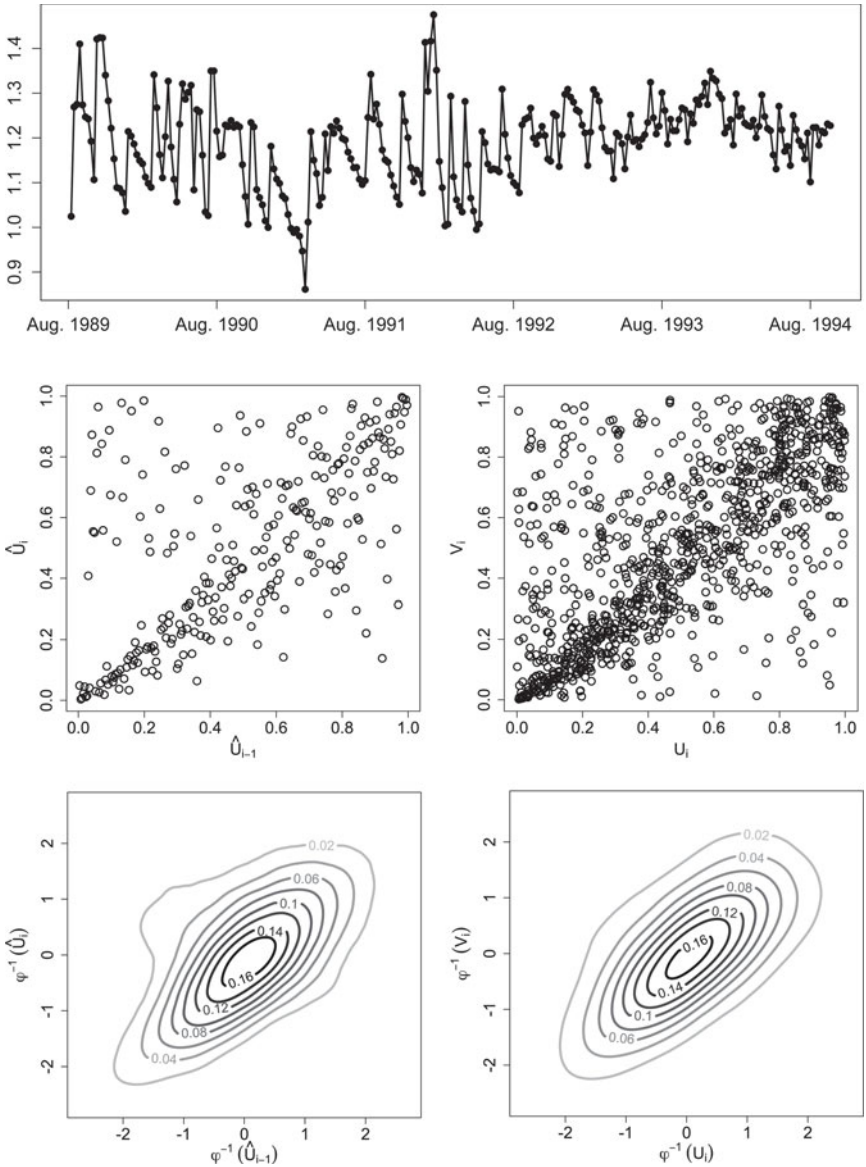


FIGURE 2. Top: the time series of gasoline markups. Middle: pseudo observations $(\hat{U}_{i-1}, \hat{U}_i)$ for the markups (left) and for a simulated sample (U_i, V_i) of size $1,000$ from the fitted nonexchangeable extension of the Clayton copula (right). Bottom: level plots of kernel density estimates on a normal scale for the margins, for the markups (left) and for a simulated sample of size $4,000$ (right).

https://doi.org/10.1017/S0266466615000419 Published online by Cambridge University Press

copula. Hence, there is indeed rather strong evidence against three out of four of these models.

Nonexchangeable copula models can for instance be derived by Khoudraji's device, see Khoudraji (1995). Based on two arbitrary copulas C_1 and C_2, one can construct a new copula through

$$C(u,v) = C_1(u^\alpha, v^\beta)C_2(u^{1-\alpha}, v^{1-\beta}), \quad (u,v) \in [0,1]^2,$$

where $\alpha, \beta \in [0,1]$, see also Liebscher (2008). We focus on the special case where C_1 is either from the Gumbel–Hougaard or the Clayton family and where C_2 is equal to the independence copula (see also Beare and Seo, 2014, for the Gumbel–Hougaard choice). For both of the resulting 3-parameter copula families, we did not find any evidence against H_0, with p-values of $p = 0.561$ and $p = 0.941$ for the Gumbel and Clayton extensions, respectively. For the extension of the Clayton copula, we obtained parameter estimates $\alpha = 0.868$, $\beta = 0.726$ and $\theta = 7.362$ (the parameter of the Clayton copula), with corresponding 95%-multiplier bootstrap confidence bands $(0.77, 0.97)$, $(0.63, 0.83)$ and $(3.49, 11.25)$, respectively.

In the left panels of Figure 2, we show a plot of the pseudo-observations $(\hat{U}_{i-1}, \hat{U}_i)$ based upon the observations (X_{i-1}, X_i), along with a contour plot of a kernel density estimate being based on the sample $(\Phi^{-1}(\hat{U}_{i-1}), \Phi^{-1}(\hat{U}_i))$ with the standard normal cdf Φ (i.e., margins are transformed to a normal scale). These plots can be compared with the right panels in Figure 2, where the same plots are based on simulated samples from the extension of the Clayton copula based on the fitted parameters. Indeed, we observe a very good approximation of the data by the model, with the contour plots only differing for the lower level curves as a consequence of small sample deviations. Overall, our results are consistent with the findings in Beare and Seo (2014).

Supplementary Material

A detailed description of a procedure to simulate a dependent multiplier sequence, a section illustrating the runtime performance of the proposed tests and a section containing the proofs of Lemma C.4, C.6 and C.7 are provided online at Cambridge Journals Online in supplementary material to this article (Berghaus and Bücher, 2015). Readers may refer to the supplementary material associated with this article, available at Cambridge Journals Online (journals.cambridge.org/ect).

REFERENCES

Aas, K., C. Czado, A. Frigessi, & H. Bakken (2009) Pair-copula constructions of multiple dependence. *Insurance, Mathematics and Economics* 44(2), 182–198.

Beare, B.K. (2010) Copulas and temporal dependence. *Econometrica* 78(1), 395–410.

Beare, B.K. & J. Seo (2014) Time irreversible copula-based markov models. *Econometric Theory* 30, 923–960.

Berg, D. & J.-F. Quessy (2009) Local power analyses of goodness-of-fit tests for copulas. *Scandinavian Journal of Statistics* 36(3), 389–412.

Berghaus, B. & A. Bücher (2015) Online Supplement to "Goodness-of-fit tests for multivariate copula-based time series models".

https://doi.org/10.1017/S0266466615000419 Published online by Cambridge University Press

Billingsley, P. (1999) *Convergence of Probability Measures*, 2nd ed. Wiley Series in Probability and Statistics: Probability and Statistics. Wiley.

Bücher, A. & H. Dette (2010) A note on bootstrap approximations for the empirical copula process. *Statistics and Probability Letters* 80, 1925–1932.

Bücher, A., H. Dette, & S. Volgushev (2011) New estimators of the Pickands dependence function and a test for extreme-value dependence. *Annals of Statistics* 39(4), 1963–2006.

Bücher, A., H. Dette, & S. Volgushev (2012) A test for Archimedeanity in bivariate copula models. *Journal of Multivariate Analysis* 110, 121–132.

Bücher, A. & I. Kojadinovic (2015) A dependent multiplier bootstrap for the sequential empirical copula process under strong mixing. *Bernoulli*, in press.

Bücher, A. & M. Ruppert (2013) Consistent testing for a constant copula under strong mixing based on the tapered block multiplier technique. *Journal of Multivariate Analysis* 116, 208–229.

Bücher, A. & S. Volgushev (2013) Empirical and sequential empirical copula processes under serial dependence. *Journal of Multivariate Analysis* 119, 61–70.

Bühlmann, P.L. (1993) The blockwise bootstrap in time series and empirical processes. Thesis (Dr.Sc.Math)–Eidgenoessische Technische Hochschule Zürich (Switzerland), ProQuest LLC.

Cebrián, A.C., M. Denuit, & O. Scaillet (2004) Testing for concordance ordering. *Astin Bulletin* 34(1), 151–173.

Chen, X. & Y. Fan (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. & Y. Fan (2006b) Estimation of copula-based semiparametric time series models. *Journal of Econometrics* 130(2), 307–335.

Darsow, W.F., B. Nguyen, & E.T. Olsen (1992) Copulas and Markov processes. *Illinois Journal of Mathematics* 36(4), 600–642.

Deheuvels, P. (1979) La fonction de dépendance empirique et ses propriétés. Un test non paramétrique d'indépendance. *Academie Royale de Belgique. Bulletin de la Classe des Sciences (5)* 65(6), 274–292.

Dehling, H. & W. Philipp (2002) Empirical process techniques for dependent data. In *Empirical Process Techniques for Dependent Data*, pp. 3–113. Birkhäuser Boston.

Denuit, M., A.-C. Goderniaux, & O. Scaillet (2007) A Kolmogorov-Smirnov-type test for shortfall dominance against parametric alternatives. *Technometrics* 49(1), 88–98.

Eckert, A. (2002) Retail price cycles and response asymmetry. *Canadian Journal of Economics/Revue canadienne d'conomique* 35(1), 52–77.

Fermanian, J.-D. (2013) An overview of the goodness-of-fit test problem for copulas. In P. Jaworski, F. Durante, & W.K. Hrdle (eds.), *Copulae in Mathematical and Quantitative Finance*. Lecture Notes in Statistics, pp. 61–89. Springer.

Fermanian, J.-D., D. Radulović, & M. Wegkamp (2004) Weak convergence of empirical copula processes. *Bernoulli* 10(5), 847–860.

Frees, E.W. & E.A. Valdez (1998) Understanding relationships using copulas. *North American Actuarial Journal* 2(1), 1–25.

Genest, C., K. Ghoudi, & L.-P. Rivest (1995) A semiparametric estimation procedure of dependence parameters in multivariate families of distributions. *Biometrika* 82(3), 543–552.

Genest, C., J. Nešlehová, & J.-F. Quessy (2011) Tests of symmetry for bivariate copulas. *The Annals of the Institute of Statistical Mathematics* 64, 811–834.

Genest, C., B. Rémillard, & D. Beaudoin (2009) Goodness-of-fit tests for copulas: A review and a power study. *Insurance, Mathematics and Economics* 44(2), 199–213.

Hofert, M., I. Kojadinovic, M. Maechler, & J. Yan (2015) *Copula: Multivariate Dependence with Copulas*. R package version 0.999-13.

Khoudraji, A. (1995) *Contributions à l'étude des copules et à la modélisation de valeurs extrêmes bivariées*. Ph.D thesis, Université Laval.

Kojadinovic, I., J. Segers, & J. Yan (2011) Large-sample tests of extreme-value dependence for multivariate copulas. *Canadian Journal of Statistics* 39(4), 703–720.

https://doi.org/10.1017/S0266466615000419 Published online by Cambridge University Press

Kojadinovic, I. & J. Yan (2011) A goodness-of-fit test for multivariate multiparameter copulas based on multiplier central limit theorems. *Statistics and Computing* 21(1), 17–30.
Kojadinovic, I. & J. Yan (2012) A non-parametric test of exchangeability for extreme-value and left-tail decreasing bivariate copulas. *Scandinavian Journal of Statistics* 39(3), 480–496.
Kojadinovic, I., J. Yan, & M. Holmes (2011) Fast large-sample goodness-of-fit tests for copulas. *Statistica Sinic* 21(2), 841–871.
Kosorok, M.R. (2008) *Introduction to Empirical Processes and Semiparametric Inference*. Springer Series in Statistics. Springer.
Liebscher, E. (2008) Construction of asymmetric multivariate copulas. *Journal of Multivariate Analysis* 99(10), 2234–2250.
McCausland, W.J. (2007) Time reversibility of stationary regular finite-state markov chains. *Journal of Econometrics* 136(1), 303–318.
McNeil, A.J., R. Frey, & P. Embrechts (2005) *Quantitative Risk Management*. Princeton Series in Finance. Princeton University Press. Concepts, techniques and tools.
Móricz, F. (1982) A general moment inequality for the maximum of partial sums of single series. *Acta Scientiarum Mathematicarum (Szeged)* 44(1–2), 67–75.
Okhrin, O., Y. Okhrin, & W. Schmid (2013) On the structure and estimation of hierarchical archimedean copulas. *Journal of Econometrics* 173(2), 189–204.
Politis, D.N. & H. White (2004) Automatic block-length selection for the dependent bootstrap. *Econometric Reviews* 23(1), 53–70.
Radulovic, D., M. Wegkamp, & Y. Zhao (2015) Weak convergence of empirical copula processes indexed by functions, *arXiv:1410.4150*.
Rémillard, B. (2010) Goodness-of-fit tests for copulas of multivariate time series. Technical report, HEC Montréal.
Rémillard, B., N. Papageorgiou, & F. Soustra (2012) Copula-based semiparametric models for multivariate time series. *Journal of Multivariate Analysis* 110, 30–42.
Rémillard, B. & O. Scaillet (2009) Testing for equality between two copulas. *Journal of Multivariate Analysis* 100(3), 377–386.
Rio, E. (2000) *Théorie asymptotique des processus aléatoires faiblement dépendants*. Mathématiques & Applications (Berlin) [Mathematics & Applications], vol. 31. Springer-Verlag.
Rüschendorf, L. (1976) Asymptotic distributions of multivariate rank order statistics. *Annals of Statistics* 4, 912–923.
Ruymgaart, F.H. (1973) *Asymptotic Ttheory of Rank Tests for Independence*. Mathematical Centre Tracts, vol. 43. Mathematisch Centrum.
Scaillet, O. (2005) A Kolmogorov-Smirnov type test for positive quadrant dependence. *Canadian Journal of Statistics* 33(3), 415–427.
Scaillet, O. (2007) Kernel-based goodness-of-fit tests for copulas with fixed smoothing parameters. *Journal of Multivariate Analysis* 98(3), 533–543.
Segers, J. (2012) Asymptotics of empirical copula processes under non-restrictive smoothness assumptions. *Bernoulli* 18(3), 764–782.
Shao, Q.-M. & H. Yu (1996) Weak convergence for weighted empirical processes of dependent sequences. *Annals of Probability* 24(4), 2098–2127.
Tsukahara, H. (2005) Semiparametric estimation in copula models. *Canadian Journal of Statistics* 33(3), 357–375.

APPENDIX A. Technical assumptions

In this section, we will introduce the technical conditions needed for the proof of Theorem 2.2. Moreover, we will briefly comment on each of the conditions.

First of all, weak convergence of the empirical copula process in Section 2.2 requires a smoothness condition on C, see Segers (2012).

https://doi.org/10.1017/S0266466615000419 Published online by Cambridge University Press

Condition I. *For $l = 1, \dots, d$, the partial derivative $C^{[l]}(\boldsymbol{u}) = \frac{\partial}{\partial u_l} C(\boldsymbol{u})$ exists and is continuous for any $\boldsymbol{u} = (u_1, \dots, u_d) \in [0,1]^d$ with $u_l \in (0,1)$.*

For completeness, we define $C^{[l]}(\boldsymbol{u})$ as zero for any $\boldsymbol{u}$ such that $u_l \in \{0,1\}$, see also Segers (2012) and Bücher and Volgushev (2013).

Throughout the proofs, we need to be able to easily go back and forth between normalized ranks and empirical quantile functions. For that purpose, ties must not occur. In the case of serial independence, it is sufficient to assume that the marginal distributions are continuous. In the case of serial dependence, continuity of the marginal distributions is not sufficient to guarantee the absence of ties which brings us to the following condition, see also Bücher and Kojadinovic (2015).

Condition II. *For any $l \in \{1, \dots, d\}$, there are no ties in the marginal time series $X_{1l}, \dots, X_{nl}$ with probability one.*

Since we allow for alpha-mixing observations, the multipliers $Z_{i,n}^{(b)}$ have to be chosen serially dependent in an appropriate way. We follow Bücher and Kojadinovic (2015), relying on previous results from Bühlmann (1993) and Bücher and Ruppert (2013).

DEFINITION A.1. *A sequence $(Z_{i,n})_{i \in \mathbb{Z}}$ is called a dependent multiplier sequence if the following three conditions are met:*

(M1) *$(Z_{i,n})_{i \in \mathbb{Z}}$ is independent of the process $(\boldsymbol{X}_i)_{i \in \mathbb{N}}$.*

(M2) *The sequence $(Z_{i,n})_{i \in \mathbb{Z}}$ in ℓ_n-dependent, i.e., $Z_{i,n}^{(b)}$ is independent of $Z_{i+h,n}^{(b)}$ for any $h > \ell_n$ and any $i \in \mathbb{Z}$, where $\ell_n \in \mathbb{N}$ tends to infinity and $\ell_n = O(n^{1/2-\kappa})$ for some $\kappa \in (0, 1/2)$, for $n \to \infty$.*

(M3) *The sequence $(Z_{i,n})$ is strictly stationary with $\mathbb{E}[Z_{i,n} Z_{i+h,n}] = \varphi(h/\ell_n)$, where φ is a symmetric function which is strictly positive on $(-1,1)$, zero on $\mathbb{R} \setminus (-1,1)$ and continuous in 0. Furthermore, $\sup_{n \in \mathbb{N}} \mathbb{E}[|Z_{i,n}|^p] < \infty$ for all $p \geq 1$.*

For the proof of (2.5), Bücher and Kojadinovic (2015) imposed the following condition on the mixing rate of $(\boldsymbol{U}_i)_{i \in \mathbb{Z}}$, where $\boldsymbol{U}_i = (\boldsymbol{U}_{i1}, \dots, \boldsymbol{U}_{id})'$ with $U_{il} = F_l(X_{il})$ for $l \in \{1, \dots, d\}$.

Condition III. *$(\boldsymbol{U}_i)_{i \in \mathbb{Z}}$ is a stationary, strongly mixing process satisfying $\alpha_{\boldsymbol{U}}(n) = O(n^{-a})$ for some $a > 3 + 3d/2$.*

The next condition, ensuring uniform continuity of the partial derivative of the copula with respect to θ, is needed for the proof of (2.2).

Condition IV. *As $\varepsilon \downarrow 0$,*

$$\sup_{\|\theta' - \theta\| < \varepsilon} \sup_{\boldsymbol{u} \in [0,1]^d} \|\nabla C_{\theta'}(\boldsymbol{u}) - \nabla C_{\theta}(\boldsymbol{u})\| \to 0.$$

Regarding the parametric estimator θ_n, we opt for general conditions that include the pseudo-maximum likelihood estimator. Following Tsukahara (2005), we begin by defining q-functions and u-shaped functions.

DEFINITION A.2 (q-functions). *Let $\mathcal{Q}$ be the class of all continuous functions $q : [0,1] \to [0,1]$ which are positive on $(0,1)$, symmetric around $1/2$, increasing on $(0, 1/2]$ and for which $\int_{[0,1]} q(u)^{-2} du < \infty$. The functions in $\mathcal{Q}$ are called q-functions.*

https://doi.org/10.1017/S0266466615000419 Published online by Cambridge University Press

For instance, for any $0 < \delta < 1/2$, the function $q(u) = \{u(1-u)\}^{1/2-\delta}$, is an element of $\mathcal{Q}$.

DEFINITION A.3 (u-shaped functions). *A function $r : (0,1) \to (0,\infty)$, which is symmetric around $1/2$, is called a u-shaped function if it is decreasing on $(0, 1/2]$.*

A typical example of a u-shaped function is $r(u) = \{u(1-u)\}^{-\tau}$ for $\tau \geq 0$.

Condition V. *The estimator θ_n allows for the decomposition*

$$\Theta_n := \sqrt{n}(\theta_n - \theta_0) = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} J_{\theta_0}(\hat{\boldsymbol{U}}_i) + o_p(1),$$

where, for $\theta \in \mathcal{P}$, $J_\theta : (0,1)^d \to \mathbb{R}^p$ denotes a score-function with the following properties:

(i) *J_θ is standardized, i.e.,*

$$\int_{[0,1]^d} J_\theta(\boldsymbol{u}) dC_\theta(\boldsymbol{u}) = 0.$$

(ii) *For any $\theta \in \mathcal{P}$ and for any $l = 1, \ldots, d$ the derivatives $J_\theta^{[l]} = \frac{\partial}{\partial u_l} J_\theta \in \mathbb{R}^p$, $\nabla J_\theta \in \mathbb{R}^{p \times p}$ and $\nabla(J_\theta^{[l]}) \in \mathbb{R}^{p \times p}$ exist, are continuous on $(0,1)^d$ and there is a constant k_θ such that for all $\theta' \in \mathcal{P}$ with $\|\theta - \theta'\| \leq k_\theta$*

$$\|J_{\theta'}(\boldsymbol{u})\| \leq \prod_{l=1}^{d} r_l(u_l) \quad \text{and} \quad \|\nabla J_{\theta'}(\boldsymbol{u})\|_{op} \leq \prod_{l=1}^{d} r_l(u_l),$$

where, for $A \in \mathbb{R}^p \times \mathbb{R}^p$, $\|A\|_{op} = \sup_{\|x\|=1} \|Ax\|$ and where $r_1, \ldots, r_d$ are u-shaped functions with

$$\int_{[0,1]^d} \prod_{l=1}^{d} r_l(u_l)^{2+\nu} dC_{\theta_0}(\boldsymbol{u}) < \infty, \tag{A.1}$$

$$\int_{[0,1]^d} \prod_{l=1}^{d} \sup_{\frac{|v-u_l|}{q_l(u_l)} \leq \delta} r_l(v)^2 dC_{\theta_0}(\boldsymbol{u}) < \infty \tag{A.2}$$

for some ν and $\delta > 0$. Additionally, for any $l = 1, \ldots, d$, there exists a u-shaped function $\tilde{r}_l$ with

$$\|J_{\theta_0}^{[l]}(\boldsymbol{u})\| \leq \tilde{r}_l(u_l) \prod_{l' \neq l} r_{l'}(u_{l'}) \quad \text{and} \quad \|\nabla(J_{\theta'}^{[l]})(\boldsymbol{u})\|_{op} \leq \tilde{r}_l(u_l) \prod_{l' \neq l} r_{l'}(u_{l'})$$

such that for some q-function q_l with $\int_0^1 q_l(u)^{-(2+\nu)} du < \infty$

$$\int_{[0,1]^d} q_l(u_l) \sup_{\frac{|v-u_l|}{q_l(u_l)} \leq \delta} \tilde{r}_l(v) \prod_{l' \neq l} \sup_{\frac{|v-u_{l'}|}{q_{l'}(u_{l'})} \leq \delta} r_{l'}(v) dC_{\theta_0}(\boldsymbol{u}) < \infty. \tag{A.3}$$

A discussion of estimators satisfying these assumptions is given in Section 3.2. The following example shows that it is in fact necessary to allow for nonconstant functions q and r.

https://doi.org/10.1017/S0266466615000419 Published online by Cambridge University Press

Example A.4
The score function for the pseudo maximum likelihood estimator for the one-dimensional Markovian copula model based on a Gaussian copula is given in (3.1). It is easy to show that this function satisfies (A.1) and (A.2). To see this, note that $|\Phi^{-1}(u)| \leq \text{const} \times \{u(1-u)\}^{-\varepsilon}$ for any $\varepsilon > 0$. As we will show next, Condition (A.3) is satisfied with nonconstant function q only. Denoting the standard normal density function by ϕ, the first first-order partial derivative of the score function is given by

$$J_\theta^{[1]}(u_1, u_2) = \frac{(1+\theta^2)\Phi^{-1}(u_2) - 2\theta\Phi^{-1}(u_1)}{(1+\theta^2)\phi(\Phi^{-1}(u_1))}.$$

It can easily be seen that $\int_{(0,1)^2} |J_{\theta_0}^{[1]}(u_1, u_2)| \mathrm{d}C_{\theta_0}(u_1, u_2) = \infty$. However, from $[\phi\{\Phi^{-1}(u)\}]^{-1} \leq \{u(1-u)\}^{-1}$, we obtain that $|J_\theta^{[1]}(u_1, u_2)| \leq K_\theta \times \{u_1(1-u_1)\}^{-1-\varepsilon}\{u_2(1-u_2)\}^{-\varepsilon}$ for any $\varepsilon > 0$ (see also Chen and Fan, 2006b), such that Condition (A.3) is satisfied with $q(u) = \{u(1-u)\}^{\delta}$ for any $\delta \in (\varepsilon, 1/2)$.

As discussed above, the components of the score function J_θ are in general not bounded. Therefore, in order to derive the asymptotic behavior of Θ_n by a suitable application of the central limit theorem for alpha-mixing sequences, we need a stronger condition on the mixing coefficients of $(\boldsymbol{U}_i)_{i\in\mathbb{Z}}$ which depends on the moments of the score function.

Condition VI. *$(\boldsymbol{U}_i)_{i\in\mathbb{Z}}$ is a stationary, strongly mixing process satisfying $\alpha_{\boldsymbol{U}}(n) = O(n^{-a})$ for some $a > \frac{2(2+\nu)(1+\nu)}{\nu^2} = 2 + \frac{6}{\nu} + \frac{4}{\nu^2}$, where ν is defined in Condition V.*

APPENDIX B. Proofs of the main results

Let us first recall and introduce some notation. For $\boldsymbol{u} \in [0,1]^d$, let

$$\tilde{\mathbb{B}}_n(\boldsymbol{u}) = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \{\mathbb{1}(\boldsymbol{U}_i \leq \boldsymbol{u}) - C_{\theta_0}(\boldsymbol{u})\}, \tag{B.1}$$

$$\tilde{\mathbb{B}}_n^{(b)}(\boldsymbol{u}) = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} Z_{i,n}^{(b)} \{\mathbb{1}(\boldsymbol{U}_i \leq \boldsymbol{u}) - C_{\theta_0}(\boldsymbol{u})\}, \tag{B.2}$$

and, recalling the definition of K_{i,θ_0} in Theorem 2.2, set

$$\tilde{\Theta}_n = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \{J_{\theta_0}(\boldsymbol{U}_i) + K_{i,\theta_0}\}, \qquad \tilde{\Theta}_n^{(b)} = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} Z_{i,n}^{(b)} \{J_{\theta_0}(\boldsymbol{U}_i) + K_{i,\theta_0}\}. \tag{B.3}$$

Proof of Theorem 2.2. As the other two items can be shown by similar, but simpler proofs, we only prove (iii). Based on the central result in Lemma C.6, we begin by showing that

$$\begin{aligned} \mathbb{H}_n &= \left(\mathbb{C}_n, \Theta_n, \mathbb{C}_n^{(1)}, \Theta_n^{(1)}, \ldots, \mathbb{C}_n^{(B)}, \Theta_n^{(B)}\right) \\ &\rightsquigarrow \mathbb{H}_C = \left(\mathbb{C}_C, \Theta, \mathbb{C}_C^{(1)}, \Theta^{(1)}, \ldots, \mathbb{C}_C^{(B)}, \Theta^{(B)}\right) \end{aligned} \tag{B.4}$$

https://doi.org/10.1017/S0266466615000419 Published online by Cambridge University Press

in $\left\{\ell^\infty([0,1]^d)\times\mathbb{R}^p\right\}^{B+1}$. For the parametric components of $\mathbb{H}_n$ in (B.4), note that $\Theta_n = \tilde{\Theta}_n + o_p(1)$ for $n\to\infty$ as a consequence of Lemma C.1 and C.2. Similarly, $\Theta_n^{(b)} = \tilde{\Theta}_n^{(b)} + o_p(1)$ for any $b = 1,\dots,B$ from Lemma C.3, C.4 and C.5.

For the nonparametric components of $\mathbb{H}_n$, we have

$$\sup_{\boldsymbol{u}\in[0,1]^d}\left|\mathbb{C}_n(\boldsymbol{u}) - \tilde{\mathbb{B}}_n(\boldsymbol{u}) + \sum_{l=1}^d C_{\theta_0}^{[l]}(\boldsymbol{u})\tilde{\mathbb{B}}_n(\boldsymbol{u}^l)\right| = o_p(1)$$

and

$$\sup_{\boldsymbol{u}\in[0,1]^d}\left|\mathbb{C}_n^{(b)}(\boldsymbol{u}) - \tilde{\mathbb{B}}_n^{(b)}(\boldsymbol{u}) + \sum_{l=1}^d C_{\theta_0}^{[l]}(\boldsymbol{u})\tilde{\mathbb{B}}_n^{(b)}(\boldsymbol{u}^l)\right| = o_p(1),$$

by simple adaptations of the proof of Proposition 3.1 in Segers (2012) and the proof of Proposition 4.2 in Bücher and Kojadinovic (2015), respectively. As a consequence,

$$\mathbb{H}_n(\boldsymbol{u}) = \Big(\tilde{\mathbb{B}}_n(\boldsymbol{u}) - \sum_{l=1}^d C_{\theta_0}^{[l]}(\boldsymbol{u})\tilde{\mathbb{B}}_n(\boldsymbol{u}^l), \tilde{\Theta}_n, \tilde{\mathbb{B}}_n^{(1)}(\boldsymbol{u}) - \sum_{l=1}^d C_{\theta_0}^{[l]}(\boldsymbol{u})\tilde{\mathbb{B}}_n^{(1)}(\boldsymbol{u}^l), \tilde{\Theta}_n^{(1)},\dots,$$
$$\tilde{\mathbb{B}}_n^{(B)}(\boldsymbol{u}) - \sum_{l=1}^d C_{\theta_0}^{[l]}(\boldsymbol{u})\tilde{\mathbb{B}}_n^{(B)}(\boldsymbol{u}^l), \tilde{\Theta}_n^{(B)}\Big) + o_P(1),$$

where the remainder is uniform in $\boldsymbol{u}\in[0,1]^d$. An application of the continuous mapping Theorem and Lemma C.6 proves (B.4).

It remains to prove assertion (2.11) in Theorem 2.2. Applying Lemma C.7 and the continuous mapping Theorem to (B.4) yields that the process

$$\Big(\sqrt{n}(C_n - C_{\theta_n}), \mathbb{B}_n^{(1)}(\boldsymbol{u}) - \sum_{l=1}^d C_{\theta_0}^{[l]}(\boldsymbol{u})\mathbb{B}_n^{(1)}(\boldsymbol{u}^l) - \nabla C_{\theta_0}(\boldsymbol{u})\Theta_n^{(1)},\dots$$
$$\dots, \mathbb{B}_n^{(B)}(\boldsymbol{u}) \sum_{l=1}^d C_{\theta_0}^{[l]}(\boldsymbol{u})\mathbb{B}_n^{(B)}(\boldsymbol{u}^l) - \nabla C_{\theta_0}(\boldsymbol{u})\Theta_n^{(B)}\Big)$$

converges in $\{\ell^\infty([0,1]^d)\}^{B+1}$ to the limiting process in (2.11). Hence, it remains to be shown that replacing $C_{\theta_0}^{[l]}$ and ∇C_{θ_0} by $C_n^{[l]}$ and ∇C_{θ_n}, respectively, in the last display does not change the limiting distribution. To see this, fix l and bound

$$\sup_{\boldsymbol{u}\in[0,1]^d}\left|C_{\theta_0}^{[l]}(\boldsymbol{u})\mathbb{B}_n^{(b)}(\boldsymbol{u}^l) + \nabla C_{\theta_0}(\boldsymbol{u})\Theta_n^{(b)} - C_n^{[l]}(\boldsymbol{u})\mathbb{B}_n^{(b)}(\boldsymbol{u}^l) - \nabla C_{\theta_n}(\boldsymbol{u})\Theta_n^{(b)}\right|$$
$$\le \max\left\{\sup_{u_l\in[\delta,1-\delta]}|\mathbb{B}_n^{(b)}(\boldsymbol{u}^l)||C_n^{[l]}(\boldsymbol{u}) - C_{\theta_0}^{[l]}(\boldsymbol{u})|,\right.$$
$$\left.\sup_{u_l\notin[\delta,1-\delta]}|\mathbb{B}_n^{(b)}(\boldsymbol{u}^l)||C_n^{[l]}(\boldsymbol{u}) - C_{\theta_0}^{[l]}(\boldsymbol{u})|\right\}$$
$$+ \sup_{\boldsymbol{u}\in[0,1]^d}|\nabla C_{\theta_n}(\boldsymbol{u}) - \nabla C_{\theta_0}(\boldsymbol{u})||\Theta_n^{(b)}|.$$

https://doi.org/10.1017/S0266466615000419 Published online by Cambridge University Press

Due to Condition III and consistency of θ_n as a consequence of Lemma C.1, C.2 and C.6, ∇C_{θ_n} provides a uniformly consistent estimator of ∇C_{θ_0}. Therefore, the last summand on the right-hand side of the last display converges to zero in probability as $n \to \infty$. The first argument in the maximum in that display converges to zero in probability since the estimator $C_n^{[l]}$ is uniformly consistent on the set $\{\boldsymbol{u} \in [0,1]^d : u_l \in [\delta, 1-\delta]\}$ for any $0 < \delta < 1/2$ and since $\mathbb{B}_n^{(b)} = O_P(1)$. For the estimation of the second argument in the maximum note that, since $\|C_n^{[l]}\|_\infty \leq 2$ and $\|C_\theta^{[l]}\|_\infty \leq 1$, we have

$$\mathbb{P}\left(\sup_{u_l \notin [\delta, 1-\delta]} |\mathbb{B}_n^{(b)}(\boldsymbol{u}^l)||C_n^{[l]}(\boldsymbol{u}) - C_{\theta_0}^{[l]}(\boldsymbol{u})| > \eta \right) \leq \mathbb{P}\left(\sup_{u_l \notin [\delta, 1-\delta]} |\mathbb{B}_n^{(b)}(\boldsymbol{u}^l)| > \eta/3 \right)$$

which can be made arbitrarily small by decreasing δ. This proves the Theorem. ■

APPENDIX C. Auxiliary results

Let us first introduce some additional notation: for $\boldsymbol{u} \in [0,1]^d$, let

$$G_n(\boldsymbol{u}) = \frac{1}{n} \sum_{i=1}^n \mathbb{1}(\boldsymbol{U}_i \leq \boldsymbol{u}), \qquad G_n^\star(\boldsymbol{u}) = \frac{1}{n+1} \sum_{i=1}^n \mathbb{1}(\boldsymbol{U}_i \leq \boldsymbol{u})$$

and $G_{nl}(u_l) = G_n(1, \ldots, 1, u_l, 1, \ldots, 1)$ with u_l at the lth coordinate, for $l = 1, \ldots, d$, and similar for $G_{nl}^\star$. By Theorem 2.2 in Shao and Yu (1996) and a simple adaptation of Lemma 2.3.1 in Ruymgaart (1973) to the alpha-mixing case we can find, for any $\gamma > 0$, some constant $D > 0$ such that $\mathbb{P}(\Omega_{n,D}) \geq 1 - \gamma$ for any $n \in \mathbb{N}$, where

$$\Omega_{n,D} = \left\{ \omega : \sup_{u \in [U_{1:n,l}, U_{n:n,l}]} \left| \frac{\sqrt{n}\{G_{nl}^\star(u) - u\}}{q_l(u)} \right| \vee \sup_{u \in (0,1)} \left| \frac{\sqrt{n}\{G_{nl}(u) - u\}}{q_l(u)} \right| \leq D \right.$$
$$\left. \text{for all } l = 1, \ldots, d \right\} \tag{C.1}$$

and where $U_{i:n,l}$ denotes the ith order statistic of $U_{1,l}, \ldots, U_{n,l}$. Note that, on the event $\Omega_{n,D}$, we have $\max_{i=1}^n |\hat{U}_{i,l} - U_{i,l}|/q_l(U_{i,l}) \leq D/\sqrt{n}$ for all $l = 1, \ldots, d$.

LEMMA C.1. *Suppose that $(\boldsymbol{U}_i)_{i \in \mathbb{Z}}$ satisfies $\alpha_{\boldsymbol{U}}(n) = O(n^{-a})$ for some $a > 1 + \sqrt{2}$ (which is a consequence of Condition III). Moreover, suppose that Condition V holds. Then*

$$L_{n1} = \Theta_n - \frac{1}{\sqrt{n}} \sum_{i=1}^n \{J_{\theta_0}(\boldsymbol{U}_i) + K_{i,n,\theta_0}\} = o_P(1),$$

where, for $\theta \in \mathcal{P}$ and $i \in \{1, \ldots, n\}$,

$$K_{i,n,\theta} = \frac{1}{n} \sum_{j=1}^n \sum_{l=1}^d J_\theta^{[l]}(\boldsymbol{U}_j) \times \{\mathbb{1}(U_{i,l} \leq U_{j,l}) - U_{j,l}\}.$$

Proof. Recall the definition of $\Omega_{n,D}$ in (C.1). Since we can make $\mathbb{P}(\Omega \setminus \Omega_{n,D})$ arbitrary small by increasing D, it is sufficient to show that $L_{n1} \mathbb{1}_{\Omega_{n,D}} = o_P(1)$, for any $D > 0$. For the ease of notation, we will suppress this indicator in the following and we will assume without loss of generality, that the inequality in (C.1) holds for any $\omega \in \Omega$.

https://doi.org/10.1017/S0266466615000419 Published online by Cambridge University Press

Now, by the mean value theorem

$$J_{\theta_0}(\hat{\boldsymbol{U}}_i) = J_{\theta_0}(\boldsymbol{U}_i) + \sum_{l=1}^{d} J_{\theta_0}^{[l]}(\boldsymbol{U}_i)(\hat{U}_{i,l} - U_{i,l}) + R_{ni},$$

where the remainder term is given by

$$R_{ni} = \sum_{l=1}^{d} \{J_{\theta_0}^{[l]}(\bar{\boldsymbol{U}}_{i,n}) - J_{\theta_0}^{[l]}(\boldsymbol{U}_i)\}(\hat{U}_{i,l} - U_{i,l})$$

and where $\bar{\boldsymbol{U}}_{i,n}$ denotes an intermediate point lying between $\boldsymbol{U}_i$ and $\hat{\boldsymbol{U}}_i$. Let us first show that $\frac{1}{\sqrt{n}} \sum_{i=1}^{n} R_{ni} = o_P(1)$. We have

$$\left\| \frac{1}{\sqrt{n}} \sum_{i=1}^{n} R_{ni} \right\| \leq D \sum_{l=1}^{d} \left\{ \frac{1}{n} \sum_{i=1}^{n} S_{nil} \right\}$$

where $S_{nil} = \| J_{\theta_0}^{[l]}(\bar{\boldsymbol{U}}_{i,n}) - J_{\theta_0}^{[l]}(\boldsymbol{U}_i) \| q_l(U_{i,l})$. It remains to be shown that each of the d sums inside the curly brackets converges to 0 in probability. For $\eta > 0$, set $M_\eta = [\eta, 1-\eta]^d$, and write

$$\frac{1}{n} \sum_{i=1}^{n} S_{nil} = \frac{1}{n} \sum_{i=1}^{n} S_{nil} \mathbb{1}(\boldsymbol{U}_i \in M_\eta) + \frac{1}{n} \sum_{i=1}^{n} S_{nil} \mathbb{1}(\boldsymbol{U}_i \in M_\eta^C). \tag{C.2}$$

Let us first consider the second sum on the right. By Condition V, for sufficiently large n such that $D/\sqrt{n} < \delta$ with δ as in the condition, we can bound its expected value by

$$2 \int_{M_\eta^C} q_l(u_l) \sup_{\frac{|v-u_l|}{q_l(u_l)} \leq \delta} \tilde{r}_l(v) \prod_{l' \neq l} \sup_{\frac{|v-u_{l'}|}{q_{l'}(u_{l'})} \leq \delta} r_{l'}(v) \, \mathrm{d}C_{\theta_0}(\boldsymbol{u}) < \infty.$$

The latter can be made arbitrary small by decreasing η, whence it remains to consider the first sum on the right-hand side of (C.2) for fixed $\eta > 0$. Obviously, the sum can be bounded by

$$\max\{\| J_{\theta_0}^{[l]}(\bar{\boldsymbol{U}}_{i,n}) - J_{\theta_0}^{[l]}(\boldsymbol{U}_i) \| : i = 1, \ldots, n, \boldsymbol{U}_i \in M_\eta\}$$

This expression converges to 0 by uniform continuity of $J_{\theta_0}^{[l]}$ on $M_{\eta/2}$ and by the fact that $\max_{i=1}^{n} \| \bar{\boldsymbol{U}}_{i,n} - \boldsymbol{U}_i \| \leq D/\sqrt{n}$ as a consequence of assuming $\Omega = \Omega_{n,D}$. To conclude, $\| \frac{1}{\sqrt{n}} \sum_{i=1}^{n} R_{ni} \| = o_P(1)$.

To finalize the proof of Lemma C.1, it remains to be shown that

$$\frac{1}{\sqrt{n}} \sum_{i=1}^{n} \sum_{l=1}^{d} J_{\theta_0}^{[l]}(\boldsymbol{U}_i)(\hat{U}_{i,l} - U_{i,l}) - \frac{1}{\sqrt{n}} \sum_{i=1}^{n} K_{i,n,\theta_0}$$

https://doi.org/10.1017/S0266466615000419 Published online by Cambridge University Press

converges to 0 in probability. By exchanging the indices in the double sum resulting from the sum over the K_{i,n,θ_0}, the last display can be written as

$$\frac{1}{n}\sum_{i=1}^{n}\sum_{l=1}^{d} J_{\theta_0}^{[l]}(\boldsymbol{U}_i) \times \sqrt{n}[\{G_{nl}^{\star}(U_{i,l}) - U_{i,l}\} - \{G_{nl}(U_{i,l}) - U_{i,l}\}].$$

We can bound the norm of the last display by $\frac{2D}{n}\sum_{i=1}^{n}\sum_{l=1}^{d} J_{\theta_0}^{[l]}(\boldsymbol{U}_i)q(U_{i,l})$, which is integrable by Condition V. Then, since $\sup_{u\in[0,1]}|\sqrt{n}\{G_{nl}^{\star}(u) - G_{nl}(u)| \le n^{-1/2}$, the assertion follows from dominated convergence. ■

LEMMA C.2. *Suppose that $(\boldsymbol{U}_i)_{i\in\mathbb{Z}}$ satisfies $\alpha_{\boldsymbol{U}}(n) = O(n^{-a})$ for some $a > 1+\sqrt{2}$ (which is a consequence of Condition III). Moreover, suppose that Condition V holds. Then*

$$\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\left(K_{i,n,\theta_0} - K_{i,\theta_0}\right) \xrightarrow{P} 0,$$

where K_{i,n,θ_0} and K_{i,θ_0} are defined in Lemma C.1 and Theorem 2.2, respectively.

Proof. As in the proof of Lemma C.1, we may assume without loss of generality that $\Omega = \Omega_{n,D}$ for some $D > 0$. Note that we can write

$$\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\left(K_{i,n,\theta_0} - K_{i,\theta_0}\right) = \sum_{l=1}^{d} I_{n,l}([0,1]^d),$$

where, for some Borel-measurable subset $A \subset [0,1]^d$,

$$I_{n,l}(A) = \int_A J_{\theta_0}^{[l]}(\boldsymbol{u})\sqrt{n}\{G_{nl}(u_l) - u_l\}\,\mathrm{d}\{G_n(\boldsymbol{u}) - C_{\theta_0}(\boldsymbol{u})\}.$$

It suffices to show that $I_{n,l}([0,1]^d) = o_P(1)$ for fixed $l \in \{1,\dots,d\}$. For $\eta \in (0,1/2)$, set $M_\eta = [\eta, 1-\eta]^d$, decompose $I_{n,l}([0,1]^d) = I_{n,l}(M_\eta) + I_{n,l}(M_\eta^C)$ and treat both summands separately.

For $I_{n,l}(M_\eta^C)$, we proceed similar to the proof of Lemma C.1: since we can work on the event $\Omega_{n,D}$, we can bound $\|I_{n,l}(M_\eta^C)\|$ by

$$\frac{D}{n}\sum_{i=1}^{n}\|J_{\theta_0}^{[l]}(\boldsymbol{U}_i)\|q_l(U_{i,l})\mathbb{1}(\boldsymbol{U}_i \in M_\eta^C) + D\int_{M_\eta^C}\|J_{\theta_0}^{[l]}(\boldsymbol{u})\|q_l(u_l)\,\mathrm{d}C_{\theta_0}(\boldsymbol{u}),$$

and the expectation of this expression can be made arbitrary small by decreasing η.

Thus, it remains to be shown that $I_{n,l}(M_\eta) = o_P(1)$ for each fixed $\eta > 0$. This is a consequence of Lemma C.8, the continuous mapping theorem, weak convergence of $\boldsymbol{u} \mapsto J_{\theta_0}^{[l]}(\boldsymbol{u})\sqrt{n}\{G_{nl}(u_l) - u_l\}$ as an element of $\ell^\infty(M_\eta)$ and the fact that $\sup_{\boldsymbol{u}\in[0,1]^d}|G_n(\boldsymbol{u}) - C_{\theta_0}(\boldsymbol{u})| = o_P(1)$ as a consequence of, for instance, Theorem 7.3 in Rio (2000). ■

https://doi.org/10.1017/S0266466615000419 Published online by Cambridge University Press

LEMMA C.3. *Suppose that* $(\boldsymbol{U}_i)_{i\in\mathbb{Z}}$ *satisfies* $\alpha_{\boldsymbol{U}}(n) = O(n^{-a})$ *for some* $a > 6$ *(which is a consequence of Condition III). Moreover, suppose that Condition V holds. Then, for any* $b = 1, \ldots, B$,

$$L_{n1}^{(b)} = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} Z_{i,n}^{(b)} \left[\{ J_{\theta_n}(\hat{\boldsymbol{U}}_i) + \hat{K}_{i,n,\theta_n} \} - \{ J_{\theta_0}(\hat{\boldsymbol{U}}_i) + \hat{K}_{i,n,\theta_0} \} \right] = o_P(1),$$

where $\hat{K}_{i,n,\theta}$ *is defined in (2.7).*

Proof. As in the proof of Lemma C.1, we may assume without loss of generality that $\Omega = \Omega_{n,D}$ for some $D > 0$. By definition, we have

$$\hat{K}_{i,n,\theta_n} - \hat{K}_{i,n,\theta_0} = \sum_{l=1}^{d} \frac{1}{n} \sum_{j=1}^{n} \{ J_{\theta_n}^{[l]}(\hat{\boldsymbol{U}}_j) - J_{\theta_0}^{[l]}(\hat{\boldsymbol{U}}_j) \} \times \{ \mathbb{1}(\hat{U}_{i,l} \le \hat{U}_{j,l}) - \hat{U}_{j,l} \}.$$

Therefore, by the mean value theorem, there exist some $\theta'_{i,n}$ and $\theta^{l}_{j,n}$, $i, j = 1, \ldots, n$, $l = 1, \ldots, d$, all lying between θ_n and θ_0, such that

$$J_{\theta_n}(\hat{\boldsymbol{U}}_i) - J_{\theta_0}(\hat{\boldsymbol{U}}_i) = \nabla J_{\theta'_{i,n}}(\hat{\boldsymbol{U}}_i)(\theta_n - \theta_0)$$

$$\hat{K}_{i,n,\theta_n} - \hat{K}_{i,n,\theta_0} = \sum_{l=1}^{d} \frac{1}{n} \sum_{j=1}^{n} \nabla J_{\theta^{l}_{j,n}}^{[l]}(\hat{\boldsymbol{U}}_j)(\theta_n - \theta_0) \times \{ \mathbb{1}(U_{i,l} \le U_{j,l}) - \hat{U}_{j,l} \},$$

where we used the fact that $\mathbb{1}(U_{i,l} \le U_{j,l}) = \mathbb{1}(\hat{U}_{i,l} \le \hat{U}_{j,l})$. Then, $L_{n1}^{(b)} = A_n + \sum_{l=1}^{d} B_{n,l}$, where

$$A_n = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} Z_{i,n}^{(b)} \nabla J_{\theta'_{i,n}}(\hat{\boldsymbol{U}}_i)(\theta_n - \theta_0),$$

$$B_{n,l} = \frac{1}{n} \sum_{j=1}^{n} \nabla J_{\theta^{l}_{j,n}}^{[l]}(\hat{\boldsymbol{U}}_j)(\theta_n - \theta_0) \times \frac{1}{\sqrt{n}} \sum_{i=1}^{n} Z_{i,n}^{(b)} \left\{ \mathbb{1}(U_{i,l} \le U_{j,l}) - \hat{U}_{j,l} \right\}.$$

Let us first treat A_n, which we decompose as $A_n = A_{n,1} \times \Theta_n + A_{n,2}$, where

$$A_{n,1} = \frac{1}{n} \sum_{i=1}^{n} Z_{i,n}^{(b)} \nabla J_{\theta'_{i,n}}(\hat{\boldsymbol{U}}_i) \mathbb{1} \left\{ \| \nabla J_{\theta'_{i,n}}(\hat{\boldsymbol{U}}_i) \|_{op} \le \prod_{l=1}^{d} r_l(\hat{U}_{i,l}) \right\}$$

$$A_{n,2} = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} Z_{i,n}^{(b)} \nabla J_{\theta'_{i,n}}(\hat{\boldsymbol{U}}_i)(\theta_n - \theta_0) \mathbb{1} \left\{ \| \nabla J_{\theta'_{i,n}}(\hat{\boldsymbol{U}}_i) \|_{op} > \prod_{l=1}^{d} r_l(\hat{U}_{i,l}) \right\},$$

with $r_1, \ldots, r_d$ being defined in Condition V. We begin with the estimation of the first term in this decomposition. Let $A_{n,1}^{(s,s')}$ denote the (s, s')-entry of the $p \times p$-matrix $A_{n,1}$.

https://doi.org/10.1017/S0266466615000419 Published online by Cambridge University Press

Obviously, its mean is equal to 0 and its variance can be estimated by

$$\begin{aligned}
&\mathbb{E}\left[\left\{A_{n,1}^{(s,s')}\right\}^2\right] \\
&\quad\le \frac{1}{n}\sum_{i,j=1}^{n}\left|\mathbb{E}\left[Z_{i,n}^{(b)}Z_{j,n}^{(b)}\right]\right|\left\{\mathbb{E}[\prod_{l=1}^{d} r_l(\hat{U}_{i,l})^2]\right\}^{1/2}\left\{\mathbb{E}\left[\prod_{l=1}^{d} r_l\left(\hat{U}_{j,l}\right)^2\right]\right\}^{1/2} \\
&\quad\le \frac{1}{n^2}\sum_{i,j=1}^{n}\left|\mathbb{E}\left[Z_{i,n}^{(b)}Z_{j,n}^{(b)}\right]\right|\int_{[0,1]^d}\prod_{l=1}^{d}\sup_{\frac{|v-u_l|}{q_l(u_l)}\le Bn^{-1/2}} r_l(v)^2\,\mathrm{d}C_{\theta_0}(\boldsymbol{u}) \\
&\quad= O\left(n^{-2}\right)O(n\ell_n) = o(1),
\end{aligned}$$

where we used the Cauchy-Schwarz inequality, (A.2) and the fact that

$$\begin{aligned}
\frac{1}{n}\sum_{i,j=1}^{n}\left|\mathbb{E}\left[Z_{i,n}^{(b)}Z_{j,n}^{(b)}\right]\right| &= \frac{1}{n}\sum_{i,j=1}^{n}\varphi\{(i-j)/\ell_n\} \\
&= \sum_{k=-\ell_n}^{\ell_n}(1-|k|/n)\varphi(k/\ell_n) \le \{2\ell_n+1\}\varphi(0). \qquad \textbf{(C.3)}
\end{aligned}$$

Thus, $A_{n,1}\times\sqrt{n}(\theta_n-\theta_0) = o_P(1)$. For the estimation of $A_{n,2}$, let $\varepsilon>0$ be arbitrary. By Condition V, we can find $n_0\in\mathbb{N}$ such that, for any $n\ge n_0$,

$$\begin{aligned}
1-\varepsilon < \mathbb{P}\Big(\|\theta_n-\theta_0\|<k_\theta\Big) &\le \mathbb{P}\Big(\cap_{i=1}^{n}\{\|\theta'_{i,n}-\theta_0\|<k_\theta\}\Big) \\
&\le \mathbb{P}\big(\cap_{i=1}^{n}\{\|\nabla J_{\theta'_{i,n}}(\hat{\boldsymbol{U}}_i)\|_{op}\le\prod_{l=1}^{d} r_l(\hat{U}_{i,l})\}\big).
\end{aligned}$$

This implies that, for any $\eta>0$ and for all $n\ge n_0$,

$$\mathbb{P}(|A_{n,2}|>\eta)\le\mathbb{P}\big(\cup_{i=1}^{n}\{\|\nabla J_{\theta'_{i,n}}(\hat{\boldsymbol{U}}_i)\|_{op}>\prod_{l=1}^{d} r_l(\hat{U}_{i,l})\}\big)<\varepsilon,$$

that is $A_{n,2}=o_p(1)$ and hence $A_n=o_P(1)$.

It remains to be shown that $B_{n,l}=o_P(1)$ for all $l=1,\dots,d$. For this purpose, we bound $B_{n,l}$ as follows

$$\begin{aligned}
&\|B_{n,l}\| \\
&\quad= \left\|\frac{1}{n}\sum_{j=1}^{n}\nabla J_{\theta_{j,l}^{l}}^{[l]}(\hat{\boldsymbol{U}}_j)(\theta_n-\theta_0)\frac{1}{\sqrt{n}}\sum_{i=1}^{n}Z_{i,n}^{(b)}\{I(U_{i,l}\le U_{j,l})-U_{j,l}+U_{j,l}-\hat{U}_{j,l}\}\right\| \\
&\quad\le \|\theta_n-\theta_0\|\times\left\{\frac{1}{n}\sum_{j=1}^{n}\|\nabla J_{\theta_{j,n}^{l}}^{[l]}(\hat{\boldsymbol{U}}_j)\|_{op}\,q_l(U_{j,l})\right\} \\
&\qquad\times\left\{\sup_{u\in[0,1]}\left|\frac{1}{\sqrt{n}}\sum_{i=1}^{n}Z_{i,n}^{(b)}\frac{\mathbb{1}(U_{i,l}\le u)-u}{q_l(u)}\right|\right.
\end{aligned}$$

https://doi.org/10.1017/S0266466615000419 Published online by Cambridge University Press

$$+ \sup_{u\in[U_{1:n,l},U_{n:n,l}]} \sqrt{n}\left|\frac{G^{\star}_{nl}(u)-u}{q_l(u)}\right| \times \left|\frac{1}{n}\sum_{i=1}^{n} Z^{(b)}_{i,n}\right|\Bigg\}.$$

Since $\|\theta_n-\theta_0\|$ converges to zero in probability, it suffices to show that both factors in curly brackets are stochastically bounded. For the second factor, this follows from Lemma D.1, Theorem 2.2 in Shao and Yu (1996), a simple adaptation of Lemma 2.3.1 in Ruymgaart (1973) to the alpha-mixing case and from (C.3). Regarding the first factor, we decompose

$$\frac{1}{n}\sum_{j=1}^{n}\left\|\nabla J^{[l]}_{\theta^l_{j,n}}\left(\hat{\boldsymbol{U}}_j\right)\right\|_{op} q_l(U_{j,l}) = \frac{1}{n}\sum_{j=1}^{n}\left\|\nabla J^{[l]}_{\theta^l_{j,n}}(\hat{\boldsymbol{U}}_j)\right\|_{op} q_l(U_{j,l}) \times \mathbb{1}(\|\theta_n-\theta_0\|\le k_{\theta_0})$$
$$+ \frac{1}{n}\sum_{j=1}^{n}\left\|\nabla J^{[l]}_{\theta^l_{j,n}}(\hat{\boldsymbol{U}}_j)\right\|_{op} q_l(U_{j,l}) \times \mathbb{1}(\|\theta_n-\theta_0\|> k_{\theta_0}).$$

Clearly, the second sum on the right is $o_P(1)$ by consistency of θ_n. By Condition V (in particular (A.3)), Markov's inequality and $|\hat{U}_{j,l}-U_{j,l}|/q_l(U_{j,l}) \le D/\sqrt{n}$ for all l as a consequence of the fact that we may assume $\Omega=\Omega_{n,D}$, we have

$$\frac{1}{n}\sum_{j=1}^{n}\left\|\nabla J^{[l]}_{\theta^l_{j,n}}(\hat{\boldsymbol{U}}_j)\right\|_{op} q_l(U_{j,l}) \times \mathbb{1}(\|\theta_n-\theta_0\|\le k_{\theta_0})$$
$$\le \frac{1}{n}\sum_{j=1}^{n}\tilde{r}_l(\hat{U}_{j,l})q_l(U_{j,l})\prod_{l'\neq l} r_{l'}(\hat{U}_{j,l'}) = O_P(1).$$

To conclude, $B_{n,l}=o_P(1)$ which proves the lemma. ■

LEMMA C.4. *Suppose that $(\boldsymbol{U}_i)_{i\in\mathbb{Z}}$ satisfies $\alpha_{\boldsymbol{U}}(n)=O(n^{-a})$ for some $a>6$ (which is a consequence of Condition III). Moreover, suppose that Condition V holds. Then, for any $b=1,\dots,B$,*

$$L^{(b)}_{n2} = \frac{1}{\sqrt{n}}\sum_{i=1}^{n} Z^{(b)}_{i,n}\left[\{J_{\theta_0}(\hat{\boldsymbol{U}}_i)+\hat{K}_{i,n,\theta_0}\}-\{J_{\theta_0}(\boldsymbol{U}_i)+K_{i,n,\theta_0}\}\right] = o_P(1).$$

Proof. The proof is similar to the proof of Lemma C.1 and is carried out explicitly in Section G in the supplementary material Berghaus and Bücher (2015). ■

LEMMA C.5. *Suppose that $(\boldsymbol{U}_i)_{i\in\mathbb{Z}}$ satisfies $\alpha_{\boldsymbol{U}}(n)=O(n^{-a})$ for some $a>6$ (which is a consequence of Condition III). Moreover, suppose that Condition V holds. Then, for any $b=1,\dots,B$,*

$$\frac{1}{\sqrt{n}}\sum_{i=1}^{n} Z^{(b)}_{i,n}\left\{K_{i,n,\theta_0}-K_{i,\theta_0}\right\} \xrightarrow{P} 0.$$

Proof. For $l\in\{1,\dots,d\}$ and $A\subset[0,1]^d$, set

$$\tilde{I}_{n,l}(A) = \int_A J^{[l]}_{\theta_0}(\boldsymbol{u})\frac{1}{\sqrt{n}}\sum_{i=1}^{n} Z^{(b)}_{i,n}\{\mathbb{1}(U_{i,l}\le u_l)-u_l\}\mathrm{d}\{G_n(\boldsymbol{u})-C_{\theta_0}(\boldsymbol{u})\}.$$

https://doi.org/10.1017/S0266466615000419 Published online by Cambridge University Press

As in the proof of Lemma C.2, it suffices to show that $\tilde{I}_{n,l}([0,1]^d) = o_P(1)$ for any fixed $l = 1, \dots, d$. This follows along the same lines as in the proof of Lemma C.2 with the weak convergence of $\sqrt{n}\{G_{nl}(u_l) - u_l\}$ and Theorem 2.2 in Shao and Yu (1996) replaced by the weak convergence of the process $\frac{1}{\sqrt{n}} \sum_{i=1}^n Z_{i,n}^{(b)} \{I(U_{i,l} \le u_l) - u_l\}$ and Lemma D.1, respectively. ■

LEMMA C.6. *If Conditions III, V and VI hold and if* $\kappa > \frac{1}{2(1+\nu)}$ *with* κ *and* ν *as defined in Definition A.1 and Condition V, respectively, then*

$$\left(\tilde{\mathbb{B}}_n, \tilde{\Theta}_n, \tilde{\mathbb{B}}_n^{(1)}, \tilde{\Theta}_n^{(1)}, \dots, \tilde{\mathbb{B}}_n^{(B)}, \tilde{\Theta}_n^{(B)}\right) \rightsquigarrow \left(\mathbb{B}_C, \Theta, \mathbb{B}_C^{(1)}, \Theta^{(1)}, \dots, \mathbb{B}_C^{(B)}, \Theta^{(B)}\right).$$

in $\left\{\ell^\infty([0,1]^d) \times \mathbb{R}^p\right\}^{B+1}$*, where the expressions on the left are defined in (B.1), (B.2) and (B.3).*

Proof. Using Condition III, tightness of the vector of processes follows from marginal tightness of $\tilde{\mathbb{B}}_n$ and $\tilde{\mathbb{B}}_n^{(b)}$, see Theorem 3.1 in Bücher and Kojadinovic (2015). Regarding weak convergence of the finite dimensional distributions, the proof is similar to the proof of Lemma A.1 in Bücher and Kojadinovic (2015). It is explicitly carried out in the supplementary material Berghaus and Bücher (2015). ■

LEMMA C.7. *Under Conditions IV, V and VI, we have*

$$\sqrt{n}\{C_{\theta_n}(\boldsymbol{u}) - C_{\theta_0}(\boldsymbol{u})\} = \nabla C_{\theta_0}(\boldsymbol{u})\Theta_n + R_n(\boldsymbol{u}),$$

where $\Theta_n = \sqrt{n}(\theta_n - \theta_0)$ *and* $\sup_{\boldsymbol{u} \in [0,1]^d} \|R_n(\boldsymbol{u})\| = o_p(1)$.

Proof. The proof is given in Section G in the supplementary material Berghaus and Bücher (2015). ■

LEMMA C.8. *Let* A *be a continuous function on* $[\boldsymbol{a}, \boldsymbol{b}]$ *and* B_n *be a sequence of functions on* $[\boldsymbol{a}, \boldsymbol{b}]$ *with* $\int_{[\boldsymbol{a},\boldsymbol{b}]} |dB_n| \le V$ *(where* $\int_{[\boldsymbol{a},\boldsymbol{b}]} |dB_n|$ *refers to the Vitali-variation of* B_n*, see Radulovic, Wegkamp, and Zhao, 2015 for a precise definition), uniformly in* n*, for some* $V > 0$*. If* $\sup_{\boldsymbol{x} \in [\boldsymbol{a},\boldsymbol{b}]} |B_n(\boldsymbol{x})| \to 0$ *and if there is a sequence of functions* A_n *such that* $\sup_{\boldsymbol{x} \in [\boldsymbol{a},\boldsymbol{b}]} |A_n(\boldsymbol{x}) - A(\boldsymbol{x})| \to 0$ *as* $n \to \infty$*, then we have*

$$\Phi(A_n, B_n) \to 0$$

as $n \to \infty$*, where* $\Phi(A, B) = \int_{[\boldsymbol{a},\boldsymbol{b}]} A(\boldsymbol{x}) dB(\boldsymbol{x})$.

Remark C.9. Note that, by integration by parts, it is possible to obtain $\Phi(B_n, A_n) \to 0$ as $n \to \infty$ as well. However, this requires additional assumptions on the function A_n and B_n, namely that both functions are of bounded variation in the sense of Hardy-Krause. We refer to Radulovic et al. (2015) for a precise definition and a proper integration by parts formula.

Proof. For $m \in \mathbb{N}$, $j = 1, \dots, m$ and $l = 1, \dots d$, let $y_{j,l,m} = a_l + (b_l - a_l)\, j/m$. Define a piecewise constant approximation of A through

$$\tilde{A}_m(\boldsymbol{x}) = \sum_{k_1,\dots,k_d=1}^{m} \mathbb{1}(\boldsymbol{x} \in (\boldsymbol{y}_{k_1-1,\dots,k_d-1}, \boldsymbol{y}_{k_1,\dots,k_d}]) A(\boldsymbol{y}_{k_1,\dots k_d}),$$

https://doi.org/10.1017/S0266466615000419 Published online by Cambridge University Press

where $\boldsymbol{y}_{k_1,\ldots,k_d} = (y_{k_1,1,m},\ldots,y_{k_d,d,m})$. Note, that $r(m) = \sup_{\boldsymbol{x}\in[\boldsymbol{a},\boldsymbol{b}]}|A(\boldsymbol{x}) - \tilde{A}_m(\boldsymbol{x})| \to 0$ as $m \to \infty$ by uniform continuity of A on $[\boldsymbol{a},\boldsymbol{b}]$.

Now, write $\Phi(A_n, B_n) = I_{n1} + I_{n2} + I_{n3}$, where

$$I_{n1} = \int_{[\boldsymbol{a},\boldsymbol{b}]} \{A_n(\boldsymbol{x}) - A(\boldsymbol{x})\}\mathrm{d}B_n(\boldsymbol{x}),$$

$$I_{n2} = \int_{[\boldsymbol{a},\boldsymbol{b}]} \{A(\boldsymbol{x}) - \tilde{A}_m(\boldsymbol{x})\}\mathrm{d}B_n(\boldsymbol{x}), \quad I_{n3} = \int_{[\boldsymbol{a},\boldsymbol{b}]} \tilde{A}_m(\boldsymbol{x})\mathrm{d}B_n(\boldsymbol{x}).$$

We begin by bounding I_{n1}. We have

$$|I_{n1}| \le \int_{[\boldsymbol{a},\boldsymbol{b}]} |A_n(\boldsymbol{x}) - A(\boldsymbol{x})||\mathrm{d}B_n(\boldsymbol{x})| \le \sup_{\boldsymbol{x}\in[\boldsymbol{a},\boldsymbol{b}]} |A_n(\boldsymbol{x}) - A(\boldsymbol{x})| \times V$$

which converges to zero as $n \to \infty$ by uniform convergence of A_n to A. Similarly, $|I_{n2}|$ can be bounded by $r(m) \times V$, which can be made smaller than any given $\varepsilon > 0$ by increasing m. Finally, for any fixed m, $|I_{n3}|$ can be bounded by $m^d \times \sup_{\boldsymbol{x}\in[\boldsymbol{a},\boldsymbol{b}]}|A(\boldsymbol{x})| \times \sup_{\boldsymbol{x}\in[\boldsymbol{a},\boldsymbol{b}]}|B_n(\boldsymbol{x})|$. The latter expression converges to 0 for $n \to \infty$ by uniform convergence of B_n. Assembling everything, for any $\varepsilon > 0$, $\limsup_{n\to\infty}|\Phi(A_n, B_n)| \le \varepsilon$, which proves the Lemma. ■

APPENDIX D. Weighted Convergence of a dependent multiplier bootstrap empirical process

LEMMA D.1. *Let $(U_i)_{i\in\mathbb{Z}}$ be a stationary, strongly mixing sequence of uniformly distributed random variables with $\alpha_U(k) = O(k^{-a})$, where $a > 6$, and suppose that $Z_{i,n}^{(b)}$ satisfies the conditions in Definition A.1. For $u \in [0,1]$, set*

$$\bar{\mathbb{A}}_n^{(b)}(u) = \frac{1}{\sqrt{n}}\sum_{i=1}^{n} Z_{i,n}^{(b)}\{\mathbb{1}(U_i \le u) - u\}$$

and let q be a q-function with $q(u) \ge K\{u(1-u)\}^{\gamma}$ for some constant $K > 0$ and some $0 \le \gamma < (a-3)/(2a)$. Then, in $\ell^\infty([0,1])$,

$$\bar{\mathbb{A}}_n^{(b)}(u)/q(u) \rightsquigarrow \mathbb{A}(u)/q(u)$$

where $\mathbb{A}$ denotes the weak limit of the process $u \mapsto n^{-1/2}\sum_{i=1}^{n}\{\mathbb{1}(U_i \le u) - u\}$.

Proof. The proof is similar to the proof of Theorem 2.1 in Shao and Yu (1996). First, note that $\bar{\mathbb{A}}_n^{(b)} \rightsquigarrow \mathbb{A}$ by Theorem 3.1 in Bücher and Kojadinovic (2015). Hence, by Theorem 3.2 in Billingsley (1999) it suffices to show that, for any $\varepsilon > 0$,

$$\lim_{\eta\to 0}\limsup_{n\to\infty}\mathbb{P}\left(\sup_{0<u\le\eta}\left|\bar{\mathbb{A}}_n^{(b)}(u)/q(u)\right| \ge \varepsilon\right) = 0 \quad \text{and}$$

$$\lim_{\eta\to 0}\mathbb{P}\left(\sup_{0<u\le\eta}|\mathbb{A}(u)/q(u)| \ge \varepsilon\right) = 0, \tag{D.4}$$

https://doi.org/10.1017/S0266466615000419 Published online by Cambridge University Press

and that

$$\lim_{\eta\to 0}\limsup_{n\to\infty}\mathbb{P}\left(\sup_{1-\eta\le u<1}\left|\bar{\mathbb{A}}_n^{(b)}(u)/q(u)\right|\ge\varepsilon\right)=0 \quad \text{and}$$

$$\lim_{\eta\to 0}\mathbb{P}\left(\sup_{1-\eta\le u<1}|\mathbb{A}(u)/q(u)|\ge\varepsilon\right)=0. \tag{D.5}$$

Since the proofs for (D.4) and (D.5) are very similar, we only treat (D.4).

We begin with a proof of the left-hand side of (D.4). Let $K\ge 1$ be a constant and let us first assume that, for any $n\ge 1$ and $i\in\{1,\dots,n\}$, we have $Z_{i,n}^{(b)}\ge -K$. The general case will be treated afterwards. We begin by showing the following three inequalities, which we will need in the course of this proof. For any $p>2a/(a-1)$ and for any $u,v\in[0,1]$

$$\mathbb{E}\left[|\bar{\mathbb{A}}_n^{(b)}(u)-\bar{\mathbb{A}}_n^{(b)}(v)|^2\right]\le \text{const}\times|u-v|^{\frac{2}{p}}. \tag{D.6}$$

For any $p'\in\left(\frac{2a}{a-3},4\right)$ and for any $u,v\in[0,1]$

$$\mathbb{E}\left[|\bar{\mathbb{A}}_n^{(b)}(u)-\bar{\mathbb{A}}_n^{(b)}(v)|^4\right]\le \text{const}\times\left(|u-v|^{4/p'}+n^{-1}|u-v|^{2/p'}\right). \tag{D.7}$$

For any $0\le u\le v\le u+h\le 1$,

$$\begin{aligned}\left|\bar{\mathbb{A}}_n^{(b)}(u)-\bar{\mathbb{A}}_n^{(b)}(v)\right|&\le\left|\bar{\mathbb{A}}_n^{(b)}(u+h)-\bar{\mathbb{A}}_n^{(b)}(u)\right|\\&\quad+K|\mathbb{A}_n(u+h)-\mathbb{A}_n(u)|+\sqrt{n}h\left\{2K+n^{-1}\textstyle\sum_{i=1}^n Z_{i,n}^{(b)}\right\},\end{aligned} \tag{D.8}$$

where $\mathbb{A}_n(u)=n^{-1/2}\sum_{i=1}^n\{\mathbb{1}(U_i\le u)-u\}$.

The assertion in (D.7) follows directly from Lemma B.2 in Bücher and Kojadinovic (2015), whereas (D.8) can be shown by the same arguments as in the proof of (31) in Bücher and Kojadinovic (2015).

Hence, it remains to show (D.6). We have, for any $u\le v$,

$$\begin{aligned}\mathbb{E}\left[\left|\bar{\mathbb{A}}_n^{(b)}(u)-\bar{\mathbb{A}}_n^{(b)}(v)\right|^2\right]&=\mathbb{E}\left[\left|\frac{1}{\sqrt{n}}\sum_{i=1}^n Z_{i,n}^{(b)}\{\mathbb{1}(u<U_i\le v)-(v-u)\}\right|^2\right]\\&=\left|\frac{1}{n}\sum_{i,j=1}^n\mathbb{E}\left[Z_{i,n}^{(b)}Z_{j,n}^{(b)}\right]\times\mathbb{E}\left[\{\mathbb{1}(u<U_i\le v)-(v-u)\}\{\mathbb{1}(u<U_j\le v)-(v-u)\}\right]\right|.\end{aligned}$$

The fact that $|\mathbb{E}[Z_{i,n}^{(b)}Z_{j,n}^{(b)}]|$ is bounded by 1 together with an application of Lemma 3.11 in Dehling and Philipp (2002) with $r=s=p$ and $t=p/(p-2)$ allows to bound the right-hand side of the last display by

$$\begin{aligned}&\frac{10}{n}\sum_{i,j=1}^n\alpha(|i-j|)^{\frac{p-2}{p}}\mathbb{E}\left[\mathbb{1}(u<U_i\le v)\right]^{\frac{2}{p}}\\&=|v-u|^{\frac{2}{p}}\frac{10}{n}\sum_{i,j=1}^n\alpha(|i-j|)^{\frac{p-2}{p}}\end{aligned}$$

https://doi.org/10.1017/S0266466615000419 Published online by Cambridge University Press

$$= |v-u|^{\frac{2}{p}} \frac{10}{n} \sum_{i=0}^{n} (n-|i|)\alpha(i)^{\frac{p-2}{p}} \le |v-u|^{\frac{2}{p}} K_1,$$

where $K_1 < \infty$ since $p > 2a/(a-1)$.

In the following, we choose $p = p' \in (\frac{2a}{a-3}, 4 \wedge \gamma^{-1})$.

Let $A_n \subset \Omega$ denote the event that $\{n^{-1} \sum_{i=1}^{n} Z_{i,n}^{(b)} \le K\}$. By Markov's inequality, we have $\mathbb{P}(\Omega \setminus A_n) = o(1)$.

Now, consider the probability on the left-hand side of (D.4). We have

$$\mathbb{P}\left(\sup_{0<u\le\eta} \left| \bar{\mathbb{A}}_n^{(b)}(u)/q(u) \right| \ge \varepsilon \right) = \mathbb{P}\left(\sup_{0<u\le\eta} \left| \bar{\mathbb{A}}_n^{(b)}(u)/q(u) \right| \ge \varepsilon, A_n \right) + o(1),$$

for $n \to \infty$. The probability on the right-hand side can be bounded by

$$\begin{aligned}
&\textstyle\sum_{j=1}^{\infty} \mathbb{P}\left(\sup_{\eta 2^{-j} < u \le \eta 2^{-j+1}} \left| \bar{\mathbb{A}}_n^{(b)}(u)/q(u) \right| \ge \varepsilon, A_n \right) \\
&\quad \le \textstyle\sum_{j=1}^{\infty} \mathbb{P}\left(\sup_{\eta 2^{-j} < u \le \eta 2^{-j+1}} \left| \bar{\mathbb{A}}_n^{(b)}(u) \right| \ge \varepsilon q(\eta 2^{-j}), A_n \right) \\
&\quad \le \textstyle\sum_{j=1}^{\infty} \mathbb{P}\left(\sup_{0 < u \le \eta 2^{-j+1}} \left| \bar{\mathbb{A}}_n^{(b)}(u) \right| \ge \varepsilon q(\eta 2^{-j}), A_n \right).
\end{aligned}$$

We will now split the sum according to $j \in G_n$ or $j \in H_n$, where

$$\begin{aligned}
G_n &= \{ j \in \mathbb{N} : n^{1/2} 3K \eta 2^{-j+1} \le \varepsilon_j/2 \}, \\
H_n &= \{ j \in \mathbb{N} : n^{1/2} 3K \eta 2^{-j+1} > \varepsilon_j/2 \},
\end{aligned}$$

and where $\varepsilon_j = \varepsilon q(\eta 2^{-j})$. First, consider those summands with $j \in G_n$. In this case, (D.8) yields

$$\begin{aligned}
&\mathbb{P}\left(\sup_{0<u\le\eta 2^{-j+1}} |\bar{\mathbb{A}}_n^{(b)}(u)| \ge \varepsilon_j, A_n \right) \\
&\quad \le \mathbb{P}\left(\left| \bar{\mathbb{A}}_n^{(b)}\left(\eta 2^{-j+1}\right) \right| + \left| \mathbb{A}_n\left(\eta 2^{-j+1}\right) \right| + \sqrt{n}\eta 2^{-j+1} \left\{ 2K + \frac{1}{n} \sum_{i=1}^{n} Z_{i,n}^{(b)} \right\} \ge \varepsilon_j, A_n \right) \\
&\quad \le \mathbb{P}\left(\left| \bar{\mathbb{A}}_n^{(b)}\left(\eta 2^{-j+1}\right) \right| + \left| \mathbb{A}_n\left(\eta 2^{-j+1}\right) \right| \ge \varepsilon_j/2 \right) \\
&\quad \le \mathbb{P}\left(\left| \bar{\mathbb{A}}_n^{(b)}\left(\eta 2^{-j+1}\right) \right| \ge \varepsilon_j/4 \right) + \mathbb{P}\left(\left| \mathbb{A}_n(\eta 2^{-j+1}) \right| \ge \varepsilon_j/4 \right).
\end{aligned}$$

The first probability on the right can be bounded with the help of (D.6) by

$$\begin{aligned}
(\varepsilon_j/4)^{-2} \mathbb{E}\left[\left| \bar{\mathbb{A}}_n^{(b)}\left(\eta 2^{-j+1}\right) \right|^2 \right] &\le \text{const} \times (\varepsilon_j/4)^{-2} \times \left(\eta 2^{-j+1}\right)^{\frac{2}{p}} \\
&\le \text{const} \times \left\{ \varepsilon q\left(\eta 2^{-j}\right) \right\}^{-2} \times \left(\eta 2^{-j}\right)^{\frac{2}{p}} \\
&\le \text{const} \times \left(\eta 2^{-j}\right)^{2\left(\frac{1}{p}-\gamma\right)} \\
&= \text{const} \times \eta^{2\beta} \left(2^{2\beta}\right)^{-j},
\end{aligned}$$

https://doi.org/10.1017/S0266466615000419 Published online by Cambridge University Press

where $\beta = 1/p - \gamma > 0$, by the choice of p. Hence, for any $n \in \mathbb{N}$,

$$\sum_{j \in G_n} \mathbb{P}\left(\left|\bar{\mathbb{A}}_n^{(b)}\left(\eta 2^{-j+1}\right)\right| \geq \varepsilon_j/4\right) \leq \text{const} \times \eta^{2\beta} \sum_{j \in G_n} \left(2^{2\beta}\right)^{-j} \leq \text{const} \times \eta^{2\beta}$$

which converges to 0 for $\eta \to 0$. For the estimation of $\mathbb{P}(|\mathbb{A}_n(\eta 2^{-j+1})| \geq \varepsilon_j/4)$ notice that the conditions of Theorem 2.2 in Shao and Yu (1996) are met (since $\gamma < (a-3)/(2a)$ and since $\alpha(n) = O(n^{-a})$ for some $a > 6$). Therefore, also the conditions in Theorem 2.1 in the latter reference are met and we can estimate the probability exactly as in the proof of this Theorem.

Now, let $j \in H_n$ and define

$$\delta = \delta_{n,j} = \frac{1}{4} \frac{\varepsilon_j}{3Kn^{1/2}} = \frac{\varepsilon}{4} \frac{q\left(\eta 2^{-j}\right)}{3Kn^{1/2}}$$

It follows form (D.8) that

$$\begin{aligned}
&\mathbb{P}\left(\sup_{0<u\leq\eta 2^{-j+1}} \left|\bar{\mathbb{A}}_n^{(b)}(u)\right| \geq \varepsilon_j, A_n\right) \\
&\quad \leq \mathbb{P}\left(\max_{1\leq i\leq\lfloor\delta^{-1}\eta 2^{-j+1}\rfloor} \left|\bar{\mathbb{A}}_n^{(b)}(i\delta)\right| \geq \varepsilon_j/2\right) \\
&\qquad + \mathbb{P}\left(\max_{0\leq i\leq\lfloor\delta^{-1}\eta 2^{-j+1}\rfloor} \sup_{i\delta<u\leq(i+1)\delta} \left|\bar{\mathbb{A}}_n^{(b)}(u) - \bar{\mathbb{A}}_n^{(b)}(i\delta)\right| \geq \varepsilon_j/2, A_n\right)
\end{aligned}$$

The second probability on the right can be further estimated by

$$\begin{aligned}
&\mathbb{P}\Bigg(\max_{0\leq i\leq\lfloor\delta^{-1}\eta 2^{-j+1}\rfloor} |\bar{\mathbb{A}}_n^{(b)}(i\delta) - \bar{\mathbb{A}}_n^{(b)}((i+1)\delta)| + |\mathbb{A}_n(i\delta) - \mathbb{A}_n((i+1)\delta)| \\
&\qquad + \sqrt{n}\delta\{2K + \frac{1}{n}\sum_{i=1}^{n} Z_{i,n}^{(b)}\} \geq \varepsilon_j/2, A_n\Bigg) \\
&\quad \leq \mathbb{P}\left(\max_{0\leq i\leq\lfloor\delta^{-1}\eta 2^{-j+1}\rfloor} \left|\bar{\mathbb{A}}_n^{(b)}(i\delta) - \bar{\mathbb{A}}_n^{(b)}((i+1)\delta)\right| + |\mathbb{A}_n(i\delta) - \mathbb{A}_n((i+1)\delta)| \geq \varepsilon_j/4\right) \\
&\quad \leq 2\,\mathbb{P}\left(\max_{1\leq i\leq\lfloor\delta^{-1}\eta 2^{-j+2}\rfloor} \left|\bar{\mathbb{A}}_n^{(b)}(i\delta)\right| \geq \varepsilon_j/16\right) + 2\,\mathbb{P}\left(\max_{1\leq i\leq\lfloor\delta^{-1}\eta 2^{-j+2}\rfloor} |\mathbb{A}_n(i\delta)| \geq \varepsilon_j/16\right).
\end{aligned}$$

By (D.7) (note that $p' = p$) we have, for any $0 \leq i < k \leq \lfloor\delta^{-1}\eta 2^{-j+2}\rfloor$, that

$$\begin{aligned}
\mathbb{E}\left[\left|\bar{\mathbb{A}}_n^{(b)}(i\delta) - \bar{\mathbb{A}}_n^{(b)}(k\delta)\right|^4\right] &\leq \text{const} \times \left\{(|k-i|\delta)^{4/p} + n^{-1}(|k-i|\delta)^{2/p}\right\} \\
&\leq \text{const} \times \left\{(|k-i|\delta)^{4/p} + n^{-1}|k-i|\delta^{2/p}\right\}.
\end{aligned}$$

https://doi.org/10.1017/S0266466615000419 Published online by Cambridge University Press

Then, the main result in Móricz (1982) (with $\gamma = 4$, $f(k,m) = \text{const} \times m\delta$ and $\varphi(t,m) = \{(m\delta)^{4/p-1} + n^{-1}\delta^{2/p-1}\}^{1/4}$), which can be applied since $\bar{\mathbb{A}}_n^{(b)}(i\delta) = \sum_{k=1}^{i} \bar{\mathbb{A}}_n^{(b)}(k\delta) - \bar{\mathbb{A}}_n^{(b)}((k-1)\delta)$, implies

$$
\begin{aligned}
&\mathbb{E}\left[\max_{0\le i\le \lfloor \delta^{-1}\eta 2^{-j+2}\rfloor} \left|\bar{\mathbb{A}}_n^{(b)}(i\delta)\right|\right] \\
&\quad\le \text{const}\times \eta 2^{-j+2} \times \left\{\sum_{k=0}^{\lfloor \log(\lfloor \delta^{-1}\eta 2^{-j+2}\rfloor)\rfloor -1} \left\{n^{-1}\delta^{2/p-1} + \left(\eta 2^{-j-k+1}\right)^{4/p-1}\right\}^{1/4}\right\}^4 \\
&\quad\le \text{const}\times \eta 2^{-j+2} \times \left\{\left(\sum_{k=0}^{\lfloor \log(\delta^{-1}\eta 2^{-j+2})\rfloor} \left(n^{-1}\delta^{2/p-1}\right)^{1/4}\right)^4\right. \\
&\qquad\qquad \left. + \left(\sum_{k=0}^{\lfloor \log(\delta^{-1}\eta 2^{-j+2})\rfloor} \left(\eta 2^{-j+1-k}\right)^{1/p-1/4}\right)^4\right\} \\
&\quad\le \text{const}\times \eta 2^{-j+2} \times \left\{n^{-1}\delta^{2/q-1}\log^4(\delta^{-1}\eta 2^{-j+2})\right. \\
&\qquad\qquad \left. + (\eta 2^{-j+1})^{4/p-1}\left(\sum_{k=0}^{\lfloor \log(\delta^{-1}\eta 2^{-j+2})\rfloor} 2^{-k(1/p-1/4)}\right)^4\right\} \\
&\quad\le \text{const}\times \left\{n^{-1}\delta^{2/p-1}(\eta 2^{-j})\log^4(\eta 2^{-j+2}\delta^{-1}) + (\eta 2^{-j})^{4/p}\right\}
\end{aligned}
$$

Now, we can estimate by Markov's inequality and the fact, that $\log(x)^4 \le \text{const}\times x^{1+2/p}$ for all $x \ge 1$ and that, for $j \in H_n$, $\eta 2^{-j+2}\delta^{-1} > 4$,

$$
\begin{aligned}
&\mathbb{P}\left(\max_{0\le i\le \lfloor \delta^{-1}\eta 2^{-j+2}\rfloor} \left|\bar{\mathbb{A}}_n^{(b)}(i\delta)\right| \ge \varepsilon_j/16\right) \\
&\quad\le \text{const}\times \varepsilon_j^{-4}\left\{n^{-1}\delta^{2/p-1}\eta 2^{-j}\log^4\left(\eta 2^{-j+2}\delta^{-1}\right) + \left(\eta 2^{-j}\right)^{4/p}\right\} \\
&\quad\le \text{const}\times \varepsilon_j^{-4}\left\{n^{-1}\delta^{2/p-1}\eta 2^{-j}\left(\eta 2^{-j+2}\right)^{1+2/p}\delta^{-1-2/p} + \left(\eta 2^{-j}\right)^{4/p}\right\} \\
&\quad\le \text{const}\times \varepsilon_j^{-4}\left\{\varepsilon_j^{-2}\left(\eta 2^{-j}\right)^{2+2/p} + \left(\eta 2^{-j}\right)^{4/p}\right\} \\
&\quad= \text{const}\times \varepsilon^{-4}\left\{\varepsilon^{-2}q\left(\eta 2^{-j}\right)^{-6}\left(\eta 2^{-j}\right)^{2+2/p} + q\left(\eta 2^{-j}\right)^{-4}\left(\eta 2^{-j}\right)^{4/p}\right\} \\
&\quad\le \text{const}\times \left\{\left(\eta 2^{-j}\right)^{-6\gamma+2+2/p} + (\eta 2^{-j})^{4/p-4\gamma}\right\}
\end{aligned}
$$

Since $\gamma < 1/p$, the right-hand side is bounded by $\text{const}\times \eta^{\beta}\times (2^{\gamma})^{-j}$, where $\beta = 1/p - \gamma > 0$.

https://doi.org/10.1017/S0266466615000419 Published online by Cambridge University Press

Thus, for any $n \in \mathbb{N}$,

$$\sum_{j \in H_n} \mathbb{P}\left(\max_{0 \le i \le \lfloor \delta^{-1} \eta 2^{-j+2} \rfloor} |\bar{\mathbb{A}}_n^{(b)}(i\delta)| \ge \varepsilon_j / 16 \right) \le \text{const} \times \eta^\beta \sum_{j \in H_n} \left(2^\beta\right)^{-j} \le \text{const} \times \eta^\beta$$

which, as to be shown, converges to 0 for $\eta \to 0$.

For the estimation of $\mathbb{P}(\max_{1 \le i \le \lfloor \delta^{-1} \eta 2^{-j+2} \rfloor} |\mathbb{A}_n(i\delta)| \ge \varepsilon_j / 16)$ we can again proceed exactly as in the proof of Theorem 2.1 in Shao and Yu (1996).

Now, consider the general case where the multipliers are not bounded from below. Let $Z_{i,n}^{(b),+} = \max\{0, Z_{i,n}^{(b)}\}$, $Z_{i,n}^{(b),-} = \max\{0, -Z_{i,n}^{(b)}\}$, $K^+ = \mathbb{E}[Z_{i,n}^{(b),+}]$ and $K^- = \mathbb{E}[Z_{i,n}^{(b),-}]$ and note that $K^+ = K^-$. Furthermore, set $M_{i,n}^{(b),+} = Z_{i,n}^{(b),+} - K^+$ and $M_{i,n}^{(b),-} = Z_{i,n}^{(b),-} - K^-$, such that $Z_{i,n}^{(b)} = Z_{i,n}^{(b),+} - Z_{i,n}^{(b),-}$. Then we can write $\mathbb{A}_n^{(b)}(u) = \mathbb{A}_n^{(b),+}(u) - \mathbb{A}_n^{(b),-}(u)$, where $\bar{\mathbb{A}}_n^{(b),\pm}(u) = \frac{1}{\sqrt{n}} \sum_{i=1}^n M_{i,n}^{(b),\pm} \{\mathbb{1}(U_i \le u) - u\}$, and it is sufficient to show that (D.4) and (D.5) hold with $\mathbb{A}_n^{(b)}$ replaced by $\mathbb{A}_n^{(b),+}$ and $\mathbb{A}_n^{(b),-}$. This, however, follows from what has been shown so far, observing that both $(M_{i,n}^{(b),+})_{i \in \mathbb{Z}}$ and $(M_{i,n}^{(b),-})_{i \in \mathbb{Z}}$ satisfy the conditions in Definition A.1 and are bounded below by $-K^* = -K^-$.

Finally, let us treat $\mathbb{P}(\sup_{0 < u \le \eta} |\mathbb{A}(u)/q(u)| \ge \varepsilon)$. Since $\mathbb{A}$ is sub-Gaussian with respect to the semimetric $d(u, v) = |u - v|^{1/p}$, we can apply Corollary 8.5 in Kosorok (2008) to get the bound

$$\mathbb{E}\left[\sup_{0 \le u \le \eta 2^{-j+1}} |\mathbb{A}(u)| \right] \le \mathbb{E}|\mathbb{A}(0)| + \text{const} \int_0^{\eta 2^{-j+1}} \sqrt{\log(2x^{-p})} \mathrm{d}x,$$

for any $j \in \mathbb{N}_0$. Since $\mathbb{A}(0) = 0$, almost surely, and since $\log(x) \le \text{const} \times x^{(1-\gamma)/p}$ for all $x \ge 1$, we can bound the right-hand side by

$$\text{const} \times \int_0^{\eta 2^{-j+1}} x^{(\gamma-1)/2} \mathrm{d}x \le \text{const} \times (\eta 2^{-j})^{(\gamma+1)/2}.$$

Therefore, by Markov's inequality,

$$\begin{aligned}
\mathbb{P}\left(\sup_{0 < u \le \eta} |\mathbb{A}(u)/q(u)| \ge \varepsilon \right) &\le \sum_{j=0}^\infty \mathbb{P}\left(\sup_{\eta 2^{-j} < u \le \eta 2^{-j+1}} |\mathbb{A}(u)/q(u)| \ge \varepsilon \right) \\
&\le \sum_{j=0}^\infty \mathbb{P}\left(\sup_{0 \le u \le \eta 2^{-j+1}} |\mathbb{A}(u)| \ge \varepsilon q\left(\eta 2^{-j}\right) \right) \\
&\le \sum_{j=0}^\infty \mathbb{E}\left[\sup_{0 \le u \le \eta 2^{-j+1}} |\mathbb{A}(u)| \right] q\left(\eta 2^{-j}\right)^{-1} \varepsilon^{-1} \\
&\le \text{const} \times \sum_{j=0}^\infty \left(\eta 2^{-j}\right)^{(\gamma+1)/2} q\left(\eta 2^{-j}\right)^{-1} \\
&\le \text{const} \times \eta^{(1-\gamma)/2} \sum_{j=0}^\infty 2^{-j(1-\gamma)/2},
\end{aligned}$$

which converges to 0 for $\eta \to 0$ since $(1 - \gamma)/2 > 0$. ■

https://doi.org/10.1017/S0266466615000419 Published online by Cambridge University Press