Statistical Papers (2024) 65:1041–1063
https://doi.org/10.1007/s00362-023-01418-z

REGULAR ARTICLE

A novel copula-based approach for parametric estimation of univariate time series through its covariance decay

Guilherme Pumi[1] · Taiane S. Prass[1] · Sílvia R. C. Lopes[2]

Received: 6 July 2021 / Revised: 11 November 2022 / Accepted: 25 February 2023 /
Published online: 29 March 2023

Abstract
In this note we develop a new technique for parameter estimation of univariate time series by means of a parametric copula approach. The proposed methodology is based on a relationship between a process' covariance decay and parametric bivariate copulas associated to lagged variables. This relationship provides a way for estimating parameters that are identifiable through the process' covariance decay, such as in long range dependent processes. We provide a rigorous asymptotic theory for the proposed estimator. We also present a Monte Carlo simulation study to asses the finite sample performance of the proposed estimator.

Keywords Copulas · Covariance decay · Time series · Parametric estimation

Mathematics Subject Classification Primary 62M10 · 62F12; Secondary 62E20

Although recognized as an important issue in many applications, the copula literature involving long range dependence is very sparse and mostly focused on empirical investigation (Mendes and Kolev 2008; Härdle and Mungo 2008). The work of Beran (2016) is an exception and formally studies the problem of long range dependence in the estimation of extreme value copulas. In the weakly dependent case, an account on general results can be found in Bücher and Volgushev (2013) and references therein.

✉ Guilherme Pumi
guilherme.pumi@ufrgs.br

Taiane S. Prass
taiane.prass@ufrgs.br

Sílvia R. C. Lopes
silvia.lopes@ufrgs.br

[1] Programa de Pós-Graduação em Estatística-Universidade Federal do Rio Grande do Sul, Porto Alegre, Brazil

[2] Programa de Pós-Graduação em Matemática-Universidade Federal do Rio Grande do Sul, Porto Alegre, Brazil

The works mentioned are, however, related to multivariate time series. The study of dependence on univariate time series by means of copula has received a fair amount of attention in the case of weakly dependent processes. Darsow et al. (1992) provide the grounds of the modern use of copulas in stationary Markov processes. Lagerås (2010) discusses some non-standard behavior presented by some copula-based Markov processes. Some methods for constructing short memory time series based on conditional copulas are discussed in chapters 4 and 8 of Joe (1997). Chen and Fan (2006) and Chen et al. (2009) discusses semiparametric estimation in copula-based one dimensional stationary Markovian process. Other interesting properties of copula-based Markov chain such as geometrical ergodicity, ρ-mixing, β-mixing, among others, are discussed in Chen et al. (2009), Beare (2010) and Beare (2012) (see also references therein). Ibragimov (2009) studies higher order Markov processes in terms of copulas and conditions under which a copula-based Markov process of some given order exhibit the so-called m-dependence, r-independence and conditional symmetry. Although copulas are mainly applied to model nonlinear dependence, the literature on covariance decay is remarkably sparse, especially in the context of long range dependent time series. The recent literature on the subject either considers only the case of copula-based Markov processes or focus on non-standard definition of long-range dependence. For instance, Ibragimov and Lentzas (2017) attempts to understand long range dependence in terms of copulas, but only non-standard definitions in terms of copula-based dependence measures are discussed. Their approach, does not encompass the classical definition of long range dependence.

In this note we aim to shed some light on this problem by exploring a connection between the covariance decay in an univariate time series and arbitrary parametric bivariate copulas associated to lagged variables. More specifically, let $\{X_t\}_{t=0}^{\infty}$ be a univariate time series of interest. In its simplest form, the idea behind our approach is as follows. Suppose that associated to the pair (X_0, X_t) is a copula C_{θ_t} from a family of parametric family, say $\{C_\theta\}_{\theta\in\Theta}$. Assume that $\text{Cov}(X_0, X_t) \sim R(t) \to 0$, say, as t increases, for a continuous function R. So in one hand, to the covariance "eyes", we have the behavior $R(t)$ as t increases, while on the other hand, to the copula's family point of view, we have the behavior θ_t in Θ. The question we would like to answer is, can we infer from one regarding the other and vice-versa? Under some easily verifiable conditions, the answer is yes, we can. This is achieved by studying a connection between the covariance decay in an univariate time series and the parametric bivariate copulas associated to lagged variables. Our approach is based on parameterizing the copulas related to pairs (X_{n_0}, X_{n_0+h}) for a fixed n_0, assuming they come from a given parametric family and it can be shown to be free from the so-called compatibility problem, which hinders most attempts to solve this problem. We show that under suitable simple conditions, the parameterization on the copula family will ultimately determine the covariance decay of the pair (X_{n_0}, X_{n_0+h}), as h increases.

The uncovered connection allows us to define a copula-based estimator of some parameter of interested, identifiable through its covariance decay. An immediate application is, of course, estimation of the dependence parameter in long range dependent time series, but our approach covers any type of covariance decay or even convergence to values other than 0. To the best of our knowledge, the proposed estimator is the first copula-based one designed to estimate long range dependence in the context of univariate time series in the literature. The main strengths of our approach are that, being copula-based, it naturally accommodates non-Gaussian time series and, by design, applies directly to time series with missing data. We provide a rigorous asymptotic theory for the proposed estimator, including conditions for its consistency and for a central limit theorem to hold. We also present a complete Monte Carlo study on the propose estimator, including a comparison with other commonly applied estimators in the literature. The results show that not only the estimator is very competitive in all proposed scenarios, but also outperforms the competitors in several instances, especially for moderate to large samples.

1 Preliminaries

In this section we recall a few concepts and results we shall need in what follows. An n-dimensional copula is a distribution function defined in the n-dimensional hypercube I^n, where $I := [0, 1]$, and whose marginals are uniformly distributed. More details on the theory of copulas can be found in the monographs by Nelsen (2013) and Joe (1997).

Copulas have been successfully applied and widely spread in several areas in the last decade. In finances, copulas have been applied in major topics such as asset pricing, risk management and credit risk analysis among many others (see the books McNeil et al. 2010; Cherubini et al. 2004, for details). In econometrics, copulas have been widely employed in constructing multidimensional extensions of complex models (see Lee and Long 2009, and references therein). In statistics, copulas have been applied in all sort of problems, such as development of dependence measures, modeling, testing, just to cite a few. The main result in the theory is the so-called Sklar's theorem (Nelsen 2013) which elucidates the role copulas play as a tool for statistical analysis and modeling.

Another result we shall need is the so-called copula version of the Hoeffding's lemma, which states that for X and Y, two continuous random variables with marginal distributions F and G, respectively, and copula C,

$$\mathrm{Cov}(X,Y) = \iint_{I^2} \frac{C(u,v) - uv}{F'\big(F^{(-1)}(u)\big)G'\big(G^{(-1)}(v)\big)}\,dudv. \tag{1}$$

In this work, $\mathbb{N}$ denotes the set of natural numbers, defined as $\mathbb{N} := \{0, 1, 2, \ldots\}$ for convenience, while $\mathbb{N}^* := \mathbb{N} \setminus \{0\}$. For a given set $A \subseteq \mathbb{R}$, $\overline{A}$ denotes the closure of A and A' denotes the set of all accumulation points. For a vector $\boldsymbol{x} \in \mathbb{R}^k$, $\boldsymbol{x}'$ denotes the transpose of $\boldsymbol{x}$. The measure space behind the notion of measurable sets and functions is always assumed (without further mention) to be $\big(\mathbb{R}^n, \mathscr{B}(\mathbb{R}^n), \mathfrak{m}\big)$ (or some appropriate restriction of it), where $\mathscr{B}(\mathbb{R}^n)$ denotes the Borel σ-field in $\mathbb{R}^n$ and $\mathfrak{m}$ is the Lebesgue measure in $\mathbb{R}^n$.

2 Relationship between copulas and decay of covariance

Suppose $\{C_\theta\}_{\theta\in\Theta}$ is a family of parametric copulas, for $\Theta \subseteq \mathbb{R}$ with non-empty interior and that the independence copula Π, defined as $\Pi(u, v) = uv$, is a member of the family, say $C_a = \Pi$ with $a \in \text{int}(\Theta)$. Assume for now that no other point in Θ yields a null covariance. Let $\{\theta_n\}_{n\in\mathbb{N}^*}$ be a sequence in Θ converging to a and let $\{X_n\}_{n\in\mathbb{N}}$ be a sequence of continuous random variables with finite second moment for which the copula associated with (X_0, X_n) is C_{θ_n}, for all $n \in \mathbb{N}^*$. For simplicity let us assume for the moment that $\{X_n\}_{n\in\mathbb{N}}$ is identically distributed and that $\text{Cov}(X_0, X_n) = \gamma_n \to 0$, as n tends to infinity. With this set up, the question we want to answer is the following: can we make a connection between how fast/slow γ_n decays to zero and how fast/slow the sequence θ_n approaches to a in Θ? In order words, can we related the covariance decay of X_n and the way C_{θ_n} approaches Π for large n? Under some mild conditions, an answer is given by Theorem 2.1 below.

The precise (and more general) mathematical formulation is the following: let $\{C_{\boldsymbol{\theta}}\}_{\boldsymbol{\theta}\in\Theta}$, $\Theta \subseteq \mathbb{R}^k$, be a family of copulas for which $C_{\boldsymbol{\theta}}$ is twice continuously differentiable with respect to $\boldsymbol{\theta}$ on an open neighborhood $U \subseteq \Theta$ of a point $\boldsymbol{a} = (a_1, \ldots, a_k)' \in \text{int}(\Theta)$. Recall that the *differential* of $C_{\boldsymbol{\theta}}$ with respect to $\boldsymbol{\theta}$ at $\boldsymbol{a} \in \mathbb{R}^k$ is the linear functional $d_{\boldsymbol{\theta}} C_{\boldsymbol{a}}(u, v) : \mathbb{R}^k \to \mathbb{R}$ whose value at a point $\boldsymbol{b} = (b_1, \ldots, b_k)' \in \mathbb{R}^k$ is

$$d_{\boldsymbol{\theta}} C_{\boldsymbol{a}}(u, v) \cdot \boldsymbol{b} = \sum_{i=1}^{k} \frac{\partial}{\partial \theta_i} C_{\boldsymbol{\theta}}(u, v) b_i \bigg|_{\boldsymbol{\theta}=\boldsymbol{a}}.$$

The *second differential* of $C_{\boldsymbol{\theta}}$ with respect to $\boldsymbol{\theta}$ at $\boldsymbol{a} \in \mathbb{R}^k$ applied to $\boldsymbol{b} = (b_1, \ldots, b_k)' \in \mathbb{R}^k$ is given by

$$d^2_{\boldsymbol{\theta}} C_{\boldsymbol{a}}(u, v) \cdot \boldsymbol{b}^2 = \sum_{i,j=1}^{k} \frac{\partial^2}{\partial \theta_i \partial \theta_j} C_{\boldsymbol{\theta}}(u, v) b_i b_j \bigg|_{\boldsymbol{\theta}=\boldsymbol{a}}.$$

Let $\{C_{\boldsymbol{\theta}}\}_{\boldsymbol{\theta}\in\Theta}$, for $\Theta \subseteq \mathbb{R}^{k+s}$ with non-empty interior, $k \in \mathbb{N}^*$ and $s \in \mathbb{N}$, be a family of parametric copulas. The following assumptions will be needed.

- **C0** There exists $\boldsymbol{a} \in \Theta'$ such that $\lim_{\boldsymbol{\theta}\to\boldsymbol{a}} C_{\boldsymbol{\theta}}(u, v) = uv$, for all $u, v \in I$.
- **C1** There exists a set $D \subseteq \Theta$ with non-empty interior such that $\boldsymbol{a} \in D'$ and $C_{\boldsymbol{\theta}}$ is twice continuously differentiable with respect to $\{\theta_1, \ldots, \theta_k\}$ in D.

Assumption **C0** is a very mild and mathematically convenient assumption. It can be replaced by the following assumption: there exists a point $\boldsymbol{a} \in \Theta$ such that if $X \sim F$, $Y \sim G$ and the copula of (X, Y) is $C_{\boldsymbol{a}}$, then $\text{Cov}(X, Y) = 0$. The limit in assumption **C0** is to be understood as the coordinatewise adequate lateral limits in case $\boldsymbol{a} \notin \text{int}(\Theta)$. We can also allow for s coordinates to remain fixed, that is, $\boldsymbol{\theta} = (\theta_1, \ldots, \theta_k, \theta^0_{k+1}, \ldots, \theta^0_{k+s}) \longrightarrow (a_1, \ldots, a_k, \theta^0_{k+1}, \ldots, \theta^0_{k+s}) = \boldsymbol{a}$. Assumption **C1** is a mild regularity condition necessary to apply a second order Taylor expansion in the proof of Theorem 2.1. It is easily verifiable for the majority of commonly applied copula families.

Theorem 2.1 *Let $\{C_{\boldsymbol{\theta}}\}_{\boldsymbol{\theta}\in\Theta}$, for $\Theta \subseteq \mathbb{R}^{k+s}$ with non-empty interior, $k \in \mathbb{N}^*$ and $s \in \mathbb{N}$, be a family of parametric copulas satisfying assumptions **C0** and **C1**. Let $\{F_n\}_{n\in\mathbb{N}}$ be* *a sequence of absolutely continuous distribution functions and define the sequences*

$$K_1^{(i)}(n) = \iint_{I^2} \frac{1}{l_0(u)l_n(v)} \lim_{\boldsymbol{\theta}\to\boldsymbol{a}} \frac{\partial C_{\boldsymbol{\theta}}(u,v)}{\partial\theta_i}\, dudv, \quad i = 1,\dots,k,$$
$$K_2^{(i,j)}(n) = \iint_{I^2} \frac{1}{l_0(u)l_n(v)} \lim_{\boldsymbol{\theta}\to\boldsymbol{a}} \frac{\partial^2 C_{\boldsymbol{\theta}}(u,v)}{\partial\theta_i\partial\theta_j}\, dudv, \quad i,j = 1,\dots,k,$$

where $l_m(x) := F_m'\big(F_m^{(-1)}(x)\big)$. *Let* $\{\boldsymbol{\theta}_n\}_{n\in\mathbb{N}^*}$ *be a sequence in D converging to* $\boldsymbol{a}$, *and let* $\{X_n\}_{n\in\mathbb{N}}$ *be a sequence of random variables such that* $X_n \sim F_n$, *and the copula associated with* (X_0, X_n) *is* $C_{\boldsymbol{\theta}_n}$. *Given a measurable function* $R : \mathbb{R} \to \mathbb{R}$ *satisfying* $\lim_{n\to\infty} R(n) = 0$, *suppose that*

$$\sum_{i=1}^{k} K_1^{(i)}(n)(\theta_n^{(i)} - a_i) \sim R(n) \quad \text{and} \quad \sum_{i,j=1}^{k} K_2^{(i,j)}(n)(\theta_n^{(i)} - a_i)(\theta_n^{(j)} - a_j) = o\big(R(n)\big). \tag{2}$$

Then, $\operatorname{Cov}(X_0, X_n) \sim R(n)$ *as n goes to infinity.*

Proofs of all mathematical results are deferred to the Appendix.

Remark 2.1 If $\Theta \subseteq \mathbb{R}$ and $\{X_n\}_{n\in\mathbb{N}}$ is stationary, then the converse of Theorem 2.1 holds, since, in this case, $l_0(x) = l_n(x)$ for all x and $K_1^{(i)}(n)$ and $K_2^{(i,j)}(n)$ do not depend on n, for all i, j, so that the result follows trivially.

We emphasize that our approach does not rely on a full probability model for the time series of interest, but only on the copulas related to $\{(X_0, X_h)\}_{h=1}^{\infty}$. No requirement is made on the copulas related to any other pair of random variable. In this way, we are only assuming minimal knowledge of the process' dependence structure, avoiding tackling the hard compatibility problem that arises when more control on the time series dependence structure is needed. The compatibility-free nature of our approach follows from the so-called pair-copula construction (Bedford and Cooke 2001, 2002), since the require structure can be used as a starting point for the construction of a pair-copula.

3 Definition of the estimator

To take advantage of the relationship presented in Proposition 2.1, in the sequel, we shall work in the following mathematical framework.

3.1 Framework A

Let $\{C_\theta\}_{\theta\in\Theta}$ be a family of parametric copulas, for $\Theta \subseteq \mathbb{R}$ with non-empty interior. Assume that there exists a point $a \in \Theta'$ such that $\lim_{\theta\to a} C_\theta = \Pi$, where the limit is to be understood as the adequate lateral limit if $a \notin \operatorname{int}(\Theta)$. Also assume that there exist a set $D \subseteq \Theta$ with non-empty interior such that $a \in D'$ and C_θ, seen as a function of the parameter θ, is of class C^2 in D. Let $\{\theta_n\}_{n\in\mathbb{N}^*}$ be a sequence in D such that $\lim_{n\to\infty} \theta_n = a$. Let $\{X_n\}_{n\in\mathbb{N}}$ be a process for which X_n is identically distributed with common absolutely continuous distribution F, for all $n \in \mathbb{N}$, satisfying $\mathrm{Cov}(X_0, X_n) \sim R(n, \boldsymbol{\eta})$, where $R(n, \boldsymbol{\eta})$ is a given continuous function such that $R(n, \boldsymbol{\eta}) \to 0$, as n goes to infinity and $\boldsymbol{\eta} \in S \subseteq \mathbb{R}^p$ is some (identifiable) parameter of interest. Also assume that $\theta_n - a \sim L(n, \boldsymbol{\eta})$, where $L(n, \boldsymbol{\eta})$ is a given continuous function satisfying $L(n, \boldsymbol{\eta}) \to 0$, as n goes to infinity.

Remark 3.1 In the context of **Framework A**, the functions $K_1^{(i)}(n) = K_1$ and $K_2^{i,j}(n) = K_2$ in (2) are constants (provided they exist). Furthermore, $R(h, \boldsymbol{\eta}) \sim K_1 L(h, \boldsymbol{\eta})$ and, if in addition, the process is weakly stationary, then $\mathrm{Cov}(X_t, X_{t+h}) \sim K_1 L(h, \boldsymbol{\eta})$ for all (fixed) $t \geq 0$.

Suppose we observe a realization (time series) $x_1, \ldots, x_n$ from a weakly stationary process $\{X_n\}_{n\in\mathbb{N}}$, under **Framework A**. To estimate the parameter $\boldsymbol{\eta}$, the theory developed in the last sections suggests the following multistage estimator.

1. Chosen a parametric family of copulas, $\{C_\theta\}_{\theta\in\Theta}$, we start by obtaining estimates $\hat{F}_n$, $\hat{F}_n^{-1}$ and $\hat{F}_n'$ of the underlying unknown distribution F, the quantile function F^{-1} and the density function F', respectively.
2. With $\hat{F}_n^{-1}$ and $\hat{F}_n'$ at hand, we can obtain $\hat{K}_1$ and $\hat{K}_2$, which must be finite and $\hat{K}_1 \neq 0$. We then form a new time series by setting $y_i := \hat{F}_n(x_i)$, for $i = 1, \ldots, n$. Notice that y_i will lie on the unit interval.
3. Let $s \geq 1$ and $m \geq 0$ be two integers satisfying $1 < s < m < n$. We shall call s the *starting lag of estimation* and m the *maximum desired lag*. Next, we form a bivariate time series $\{\boldsymbol{u}_k^{(s)}\}_{k=1}^{n-s}$ by setting $\boldsymbol{u}_i^{(s)} := (y_i, y_{i+s})$, $i = 1, \ldots, n-s$. By Sklar's theorem, $\{\boldsymbol{u}_k^{(s)}\}_{k=1}^{n-s}$ can be regarded as a correlated sample from C_{θ_s}. From these pseudo observations, θ_s can be estimated by using any reasonable method. Let $\hat{\theta}_s(n)$ be the estimated θ_s. Notice that $\hat{K}_1(\hat{\theta}_s(n) - a)$ is an estimate of $K_1 L(s, \boldsymbol{\eta}) \sim R(s, \boldsymbol{\eta})$.
4. Proceeding analogously for each $\ell \in \{s+1, \ldots, m\}$, we form the sequence $\{\boldsymbol{u}_k^{(\ell)}\}_{k=1}^{n-\ell}$ by setting $\boldsymbol{u}_i^{(\ell)} := (y_i, y_{i+\ell})$, $i = 1, \ldots, n-\ell$, from which we obtain the estimate $\hat{\theta}_\ell(n)$. For each ℓ, $\hat{K}_1(\hat{\theta}_\ell(n) - a)$ is an estimate of $K_1 L(\ell, \boldsymbol{\eta}) \sim R(\ell, \boldsymbol{\eta})$.
5. Let $\mathscr{D} : \mathbb{R}^{m-s+1} \times \mathbb{R}^{m-s+1} \to [0, \infty)$, be a given function measuring the distance between two vectors in $\mathbb{R}^{m-s+1}$. Let $\widehat{\boldsymbol{L}}_{s,m}(n) := \hat{K}_1(\hat{\theta}_s(n) - a, \ldots, \hat{\theta}_m(n) - a)'$ and $\boldsymbol{R}_{s,m}(\boldsymbol{\eta}) := \big(R(s, \boldsymbol{\eta}), \ldots, R(m, \boldsymbol{\eta})\big)'$. The estimator $\hat{\boldsymbol{\eta}}_{s,m}(n)$ of $\boldsymbol{\eta}$ is then defined as

$$\hat{\boldsymbol{\eta}}_{s,m}(n) := \underset{\boldsymbol{\eta}\in S}{\mathrm{argmin}}\big\{\mathscr{D}\big(\widehat{\boldsymbol{L}}_{s,m}(n), \boldsymbol{R}_{s,m}(\boldsymbol{\eta})\big)\big\}. \tag{3}$$

One of the main perks of being copula-based is that the proposed estimator can naturally accommodate for non-Gaussian time series as well as multimodality, bounds, and many other marginal behavior. It also easily handles missing data. To do that, we only need to apply an estimator $\hat{F}_n$ that is capable of handling missing data in step 1 (as, for instance, the empirical distribution) and then perform step 4 considering only lagged pairs available in the pseudo-sample to estimate the copula parameter. Of course, as most copula-based approach, we pay a price for this flexibility by being forced to estimate the marginal behavior. Choosing the copula family that is best suited for a given application is also a problem shared with other copula-based approaches.

In Step 4, the choices of s and m depend highly on the nature of the parameter $\boldsymbol{\eta}$. As a rule of thumb, s should be at least equal to the same dimension of $\boldsymbol{\eta}$ (1 in most applications) and m should be a small fraction of n. For $\boldsymbol{u}, \boldsymbol{v} \in \mathbb{R}^k$, $k > 1$, say $\boldsymbol{u} = (u_1, \ldots, u_k)'$ and $\boldsymbol{v} = (v_1, \ldots, v_k)'$, usual choices for the function $\mathscr{D}$ in Step 5 are $\mathscr{D}(\boldsymbol{u}, \boldsymbol{v}) = \sum_{i=1}^k |u_i - v_i|$, $\mathscr{D}(\boldsymbol{u}, \boldsymbol{v}) = \sum_{i=1}^k (u_i - v_i)^2$ and $\mathscr{D}(\boldsymbol{u}, \boldsymbol{v}) = \max_{1 \le i \le k} \{|u_i - v_i|\}$. However, the choice of $\mathscr{D}$ is mainly of theoretical importance (for large sample theory). It has negligible impact in applications.

Examples of processes that present a covariance decay estimable by the proposed approach are the classical ARMA processes, as well as general long range dependent process, for which the covariance decay is of the form $\mathrm{Cov}(X_t, X_{t+h}) \sim v(h)h^{-\beta}$, as h goes to infinity, for some $\beta \in (0, 1)$, where v is a slowing varying function. For the classical ARFIMA(p, d, q) processes, $v(h) = \Gamma(1 - d)/\Gamma(d)$, while for Fractional Gaussian noise with Hurst parameter H, $v(h) = H(2H - 1)$.

4 Large sample theory

In this section we shall study large sample properties of the proposed estimator. We start by proving its consistency. **Framework A** is assumed throughout this section. We also need the following assumptions:

A0 $\hat{F}_n$, $\hat{F}_n'$ and $\hat{F}_n^{-1}$ are consistent estimators of F, F' and F^{-1}, in the sense that $\hat{F}_n(x) \xrightarrow{\mathbb{P}} F(x)$, $\hat{F}_n'(x) \xrightarrow{\mathbb{P}} F'(x)$, for all $x \in \mathbb{R}$, and $\hat{F}_n^{-1}(u) \xrightarrow{\mathbb{P}} F^{-1}(u)$, for all $u \in I$, and such that $\hat{K}_1 \xrightarrow{\mathbb{P}} K_1$, as n tends to infinity.

A1 The estimator of the copula parameter at lag k, $\hat{\theta}_k(n)$, satisfies $\hat{\theta}_k(n) \xrightarrow{\mathbb{P}} \theta_k^0$, as $n \to \infty$, for all $s \le k \le m$, where θ_k^0 denotes the true copula parameter at lag k.

A2 The space $(\mathbb{R}^{m-s+1}, \mathscr{D})$ is a metric space and $\mathscr{D}$ is equivalent to the usual metric in $\mathbb{R}^{m-s+1}$.

The consistency requirements in Assumption **A0** are very mild ones. In particular, general sufficient conditions for the consistency of $\hat{K}_1$ are provided in Lemma 4.1 below. Assumption **A1** is a “high level” one in the sense that we require that the copula estimator applied is consistent for the particular scenario applied. This is mathematically convenient by keeping the list of assumptions simple, without having to rely on a particular estimator (or class of estimators) for the theory to hold. Of course, other assumptions might be needed to assure that **A1** holds in a case-by-case fashion. See Remark 4.1.

Lemma 4.1 *Let **Framework A** hold and suppose that $\hat{F}_n$, $\hat{F}_n^{-1}$ and $\hat{F}_n'$ are consistent estimators of F, F^{-1} and F'. Consider the conditions:*

$$\text{(a)} \iint_{I^2} \left| \frac{1}{\hat{F}_n'\big(\hat{F}_n^{(-1)}(u)\big)\hat{F}_n'\big(\hat{F}_n^{(-1)}(v)\big)} - \frac{1}{F'\big(F^{(-1)}(u)\big)F'\big(F^{(-1)}(v)\big)} \right| du dv \stackrel{\mathbb{P}}{\longrightarrow} 0;$$

$$\text{(b)} \sup_{u,v \in I} \left\{ \left| \frac{1}{\hat{F}_n'\big(\hat{F}_n^{(-1)}(u)\big)\hat{F}_n'\big(\hat{F}_n^{(-1)}(v)\big)} - \frac{1}{F'\big(F^{(-1)}(u)\big)F'\big(F^{(-1)}(v)\big)} \right| \right\} \stackrel{\mathbb{P}}{\longrightarrow} 0.$$

If either conditions (a) or (b) hold, then $\hat{K}_1 \stackrel{\mathbb{P}}{\longrightarrow} K_1$, as $n \to \infty$.

Theorem 4.1 *Let **Framework A** hold and assume that the process $\{X_n\}_{n\in\mathbb{N}}$ is strongly stationary. Under assumptions **A0–A2**, $\hat{\boldsymbol{\eta}}_{s,m}(n) \stackrel{\mathbb{P}}{\longrightarrow} \boldsymbol{\eta}_0$, as n tends to infinity.*

In order to prove a central limit theorem for the proposed estimator, we shall need a different set of assumptions.

A3 There exist a positive integer k_0 such that, as a function of $\boldsymbol{\eta}$, $L(k,\cdot) : \mathbb{R}^p \to \mathbb{R}$ is twice differentiable in a neighborhood $\Omega_0 \subseteq \mathbb{R}^p$ of $\boldsymbol{\eta}_0$ and $\boldsymbol{a}_k\boldsymbol{a}_k'$ is positive definite, where, $\boldsymbol{a}_k = \frac{\partial L(k,\boldsymbol{\eta})}{\partial \boldsymbol{\eta}}$, for all $k > k_0$ and $\boldsymbol{\eta} \in \Omega_0$.

A4. There exists a positive integer k_1, a neighborhood $\Omega_1 \subseteq \mathbb{R}^p$ of $\boldsymbol{\eta}_0$ and a sequence $b_n \to \infty$ such that the copula parameter estimator at lag k, $\hat{\theta}_k(n)$, satisfies,

$$b_n\big(\hat{\theta}_k(n) - a - L(k,\boldsymbol{\eta})\big) \stackrel{d}{\longrightarrow} Z_k, \quad \forall\, k \geq k_1, \quad \boldsymbol{\eta} \in \Omega_1,$$

with $\mathbb{E}(Z_k^2) < \infty$. Furthermore, we assume that the random variables $\{\hat{\theta}_k(n)\}_{k,n}$ and $\{Z_k\}_k$ are defined in the same probability space for all $k \geq k_1$ and n.

Assumptions **A3** and **A4** are necessary to guarantee that the limit distribution is well defined. **A3** is a minimal condition and it is actually hard to come up with examples of time series of practical importance for which **A3** does not hold in the context of **Framework A**. Assumption **A4** also guarantees the existence of the limit distribution. Often, Z_k in **A4** will be normally distributed, but, even in this case the limiting distribution for the proposed estimator may be non-standard as the sequence $\{\widehat{\theta}_k(n)\}_k$ may not be independent nor jointly normally distributed. Furthermore, observe that the estimation of the underlying distribution F is incidently imbedded in **A4**.

The rate of convergence in assumption **A4** defines the convergence rate for the proposed estimator. Also, observe that **A1** is implied by **A4**. Finally, the limiting distribution depends heavily on the metric $\mathscr{D}$ applied. To prove a CLT for the proposed estimator, we shall consider the Euclidean distance in $\mathbb{R}^{m-s+1}$.

Theorem 4.2 *Let **Framework A** hold and assume that the process $\{X_n\}_{n\in\mathbb{N}}$ is strongly stationary. Also suppose that assumptions **A3** and **A4** hold. Then, for $\mathscr{D}(\boldsymbol{u},\boldsymbol{v}) = \sum_{i=1}^m (v_i - u_i)^2$,*

$$b_n\big(\widehat{\boldsymbol{\eta}}_{s,m}(n) - \boldsymbol{\eta}_0\big) \stackrel{d}{\longrightarrow} \sum_{k=0}^{m-s} \boldsymbol{\tau}_{s+k} Z_{s+k}, \tag{4}$$

as n tends to infinity, for all s and m, where $\boldsymbol{\tau}_{s+k} = \left[\sum_{j=0}^{m-s} \boldsymbol{a}_{s+j}\boldsymbol{a}'_{s+j}\right]^{-1} \boldsymbol{a}_{s+k}$.

Remark 4.1 Assumptions **A1** and **A4** are high level ones. Under weak dependence, such results are readily available in the literature [see Bücher and Volgushev (2013) and Beran (2016), and references therein]. Consistency and the central limit theorem of copula estimators under long range dependence, on the other hand, are difficult to obtain, and the literature on the subject is sparse. For instance, Bücher and Volgushev (2013) study the problem of weak convergence of the empirical copula under general assumptions which can be extended to the case of long range dependence using the results presented in Marinucci (2005). As an application of their findings, the authors present asymptotic results related to Spearman's ρ, showing that, under long range dependence, the limiting distribution is non-Gaussian and the convergence rate is slower than $\sqrt{n}$ (see the discussion about condition 2.1, Remark 2.6(a) and example 2.7 in Bücher and Volgushev 2013).

5 Numerical results

In this section we present a Monte Carlo simulation study to assess the finite sample behavior of the proposed estimator. As stated in Theorems 4.1 and 4.2, the asymptotic properties of the proposed estimator $\hat{\boldsymbol{\eta}}_{s,m}(n)$ is directly connected to the estimation of the copula's parameter. However, the literature regarding the finite sample performance of the latter in the context of correlated samples is very scarce, especially under long-range dependence. Given its importance, we also evaluate the finite sample performance of three well-known estimators for the copulas' parameters. Another long-standing issue when working with copulas is deciding which parametric family to use. To address this problem and evaluate the influence of misspecification in the proposed estimator, we consider four parametric copula families in the simulation. Besides these points, we also explore the influence of the step size, s and m, the metric choice $\mathscr{D}$, the estimation of K_1, and asymptotic related results. Here we present our main findings. Complete results can be found in the supplementary material accompanying the paper.

5.1 Data generating process

We simulate 1000 replicas of a Gaussian ARFIMA$(0, d, 0)$ for $d \in \{0.1, 0.2, 0.3, 0.4\}$ using the traditional MA(∞) representation

$$X_t = \sum_{k=0}^{\infty} c_k \varepsilon_{t-k}, \quad \varepsilon_t \sim N(0,1), \quad c_0 = 1, \; c_k = \frac{\Gamma(k+d)}{\Gamma(k+1)\Gamma(d)}, \; \text{for } k \geq 1,$$

truncated at lag 100,000 in all cases. Observe that in this case X_t is stationary and ergodic with $X_t \sim N(0, \sum_{k=0}^{\infty} c_k^2)$, and the copula related to (X_t, X_{t+h}) is the

Gaussian copula with parameter

$$\rho_h = \frac{\Gamma(d+h)\Gamma(1-d)}{\Gamma(h-d+1)\Gamma(d)} \sim \frac{\Gamma(1-d)}{\Gamma(d)} h^{2d-1}, \text{ as } h \to \infty, \quad \text{for all } t \in \mathbb{Z}. \tag{5}$$

Sample size was set to $n = 2{,}000$ in all cases. More details can be found in Brockwell and Davis (1991) and Palma (2007).

5.2 Parameter estimation

Given a sample $x_1, \ldots, x_n$ generated as described above, to estimate d we apply the procedure outlined in Sect. 3. The estimation was performed by considering the entire time series and subsamples of sizes 1000 and 500. The settings for each step are outlined in the sequel.

In step 1, we consider four parametric families of copulas: the Ali–Mikhail–Haq (AMH), Farlie–Gumbel–Morgenstern (FGM), Frank and Gaussian. For detailed information regarding these families, see Nelsen (2013). The Gaussian copula corresponds to the true underlying family, so there is no misspecification. The others were chosen due to their frequent use in applications and also because they all have closed formulas for $\partial C_\theta(u, v)/\partial\theta$, making the calculation of K_1 easier.

Remark 5.1 Upon applying (1) considering standard Gaussian marginals, it can be shown that for the AMH copula, the correlation varies approximately in the range $(-0.26, 0.5)$ while for the FGM copula, in the range $(-0.32, 0.32)$. These narrow ranges for the correlation hinder the use of AMH and FGM in applications, especially under strong long range dependence.

In order to estimate the marginal distributions F, the quantile function F^{-1} and the density function F', we proceed as follows:

- F is estimated upon applying the rescaled empirical distribution function, namely, $\hat{F}_n(x) := (n+1)^{-1} \sum_{i=1}^{n} I(X_i \le x)$. In the context of this simulation study, by the generalized Glivenko-Cantelli theorem for stationary and ergodic sequences (Stute and Schumann 1980), the empirical distribution is a strong consistent estimator of the underlying distribution, and so is $\hat{F}_n$.
- F^{-1} is estimated by considering convex combinations of consecutive order statistics, namely, $\hat{F}_n^{-1}(p) := (1-\alpha)x_{(\lfloor\tau\rfloor)} + \alpha x_{(\lceil\tau\rceil)}$, with $\tau = 1 + (n-1)p$ and $\alpha = \tau - \lfloor\tau\rfloor$. This interpolation technique is the default in the R function `quantile` and ensures that $\hat{F}_n^{-1}$ is a continuous function of p and also a consistent estimator of F^{-1}, as an application of Giraitis and Surgailis (1999).
- F' is estimated using a kernel density approach. More specifically, first $y_i = \hat{f}(x_i^*)$ is estimated using $T = 512$ (default for the R function `density`) equally spaced points x_i^*, $1 \le i \le T$, in the interval $[x_{(1)} - 3b, x_{(n)} + 3b]$, where b is the bandwidth for the Gaussian kernel density estimator, chosen by applying the Silverman's rule of thumb (the default procedure in `density`). A cubic spline interpolation (the default method for the R function `spline`) is then applied to the pairs $\{(x_i^*, y_i)\}_{i=1}^{T}$ to obtain $\hat{F}_n'(x)$ for all $x \in [x_{(1)} - 3b, x_{(n)} + 3b]$. The consistency of $\hat{F}_n'(x)$ follows

from the consistency of the kernel density estimator for infinite-order moving average processes (Hall and Hart 1990).

Since these estimators can be applied in any context, it is interesting to compare their performance to the respective true Gaussian counterparts, with mean and variance replaced by the sample estimators.

In step 2, if $\hat{K}_1$ cannot be calculated analytically, we apply numerical integration using the Gauss-Kronrod algorithm as implemented in the function `integral2` in R package `pracma` (Borchers 2021). In steps 3 and 4 we apply 11 different combinations of the starting lag of estimation s and the maximum desired lag m, given by $\big\{(s,m) \in \{1,3,6\} \times \{6,12,24,50\} : s < m\big\}$. The goal is to investigate if there exist any evidence that a specific combination of s and m provides better results in practice.

From the pseudo observations, we estimate the copula parameter by using three different methods implemented in R package `copula` (Hofert et al. 2020), namely, the maximum pseudo-likelihood estimator (`mpl`) and the estimators based on the inversion of Kendall's τ (`itau`) and Spearman's ρ (`irho`). Notice that `itau` and `irho` are rank-based methods, hence, using the empirical or Gaussian distribution to obtain the pseudo observations will lead to the exact same estimates for θ_h. To investigate how the dependence between the pseudo observations affects the estimation of the copulas' parameters, we only used pseudo observations that are far apart by a fixed lag (thinning), which we call step. When step $= 1$, all pseudo observations are used for estimation purposes, while when step $= 10$ only pseudo observations that are (exactly) 10 lags apart are used. Of course, as there are more pseudo observations for step $= 1$ compared to step $= 10$, the estimates based on the latter will present higher variance.

Upon denoting by θ_h the parameter of the copula associated to (X_t, X_{t+h}), let $\hat{\theta}_h$ be the estimate of θ_h based on the pseudo observations. In step 5, we apply the Minkowski distance as metric, namely

$$\mathscr{D}(\boldsymbol{x},\boldsymbol{y}) = \left(\sum_{i=1}^{n} |x_i - y_i|^r\right)^{\frac{1}{r}}, \quad \text{for } \boldsymbol{x},\boldsymbol{y} \in \mathbb{R}^n, \quad \text{and } r > 0.$$

Upon observing that for all copula families considered we have $a = 0$, d is estimated as

$$\hat{d} := \underset{|d|<0.5}{\operatorname{argmin}} \left\{ \sum_{h=s}^{m} \left| \hat{K}_1 \hat{\theta}_h - \frac{\Gamma(1-d)}{\Gamma(d)} h^{2d-1} \right|^r \right\}, \quad r > 0. \tag{6}$$

The metric choice should only be relevant for derivations of asymptotic results, such as Theorem 4.2, and should not significantly affect pointwise estimation. In order to verify if that is really the case, we consider $r = 2$ (Euclidian distance) and also $r = 1/2$. The latter might be advantageous in scenarios where the objective function is globally too close to zero due to the value of d or the choices of s and m. Optimization (6) is performed using a combination of golden section search and successive parabolic interpolation, as implemented in R function `optimize` (R Core Team 2020).

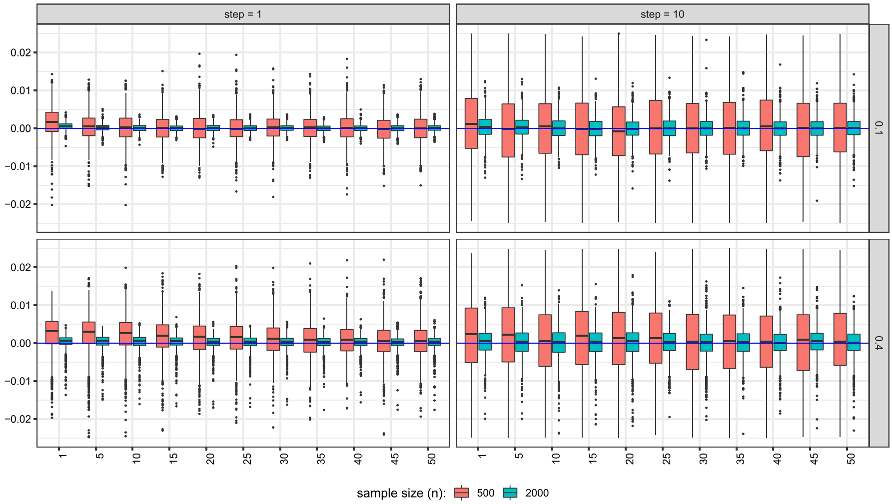

Fig. 1 Box-plots of the differences $\hat{\theta}_h^E - \hat{\theta}_h^G$ for $d \in \{0.1, 0.4\}$ (row panels), $h \in \{1, 5, 10, \ldots, 45, 50\}$ (all panels), $n \in \{500, 2000\}$ (all panels) and step sizes 1 and 10 (column panels)

5.3 Simulation results

Estimation of the copulas' parameters

In the sequel we discuss results obtained by considering the Gaussian copula, which corresponds to the scenario were there is no misspecification in the copula's family. Results for AMH, FGM and Frank copula are discussed in the supplementary material.

Figure 1 shows the box-plots of the differences $\hat{\theta}_h^E - \hat{\theta}_h^G$ based on 1000 replications, where $\hat{\theta}_h^E$ and $\hat{\theta}_h^G$ denote the estimated values $\hat{\theta}_h$ obtained from the `mpl` method with the Gaussian copula, considering pseudo observations obtained with the empirical and the Gaussian distribution, respectively. To save space this figure only reports the results for $d \in \{0.1, 0.4\}$ (row panels), $h \in \{1, 5, 10, \ldots, 45, 50\}$ (all panels) and $n \in \{500, 2000\}$ (all panels). In all cases the results are presented for step sizes 1 and 10 (column panels). Also, for better visualization, the y-axis is restricted to $(-0.025, 0.025)$.

The results presented in Fig. 1 indicate that, for small values of h, the values of $\hat{\theta}_h^E$ tend to be slightly higher than $\hat{\theta}_h^G$, for any sample size. The variability of the differences $\hat{\theta}_h^E - \hat{\theta}_h^G$ does not appear to have any relation to h, being slightly higher for $d = 0.4$, significantly increasing with the step size and decreasing with n. Analogous results were found for $d \in \{0.2, 0.3\}$ and $n = 1000$ (see the supplementary material).

Figure 2 shows the box-plots of the estimated values $\hat{\theta}_h$ obtained from methods `mpl`, `itau` and `irho` (all panels), considering the Gaussian copula and pseudo observations obtained with the empirical distribution and step sizes 1 and 10 (column panels). To save space, this figure only reports the results for $d \in \{0.1, 0.4\}$ (row panels), $h \in \{1, 5, 10, \ldots, 45, 50\}$ (all panels), $n \in \{500, 2000\}$ (row panels). The true values of $\theta_h = \rho_h$, given by (5) are also reported (blue dots).

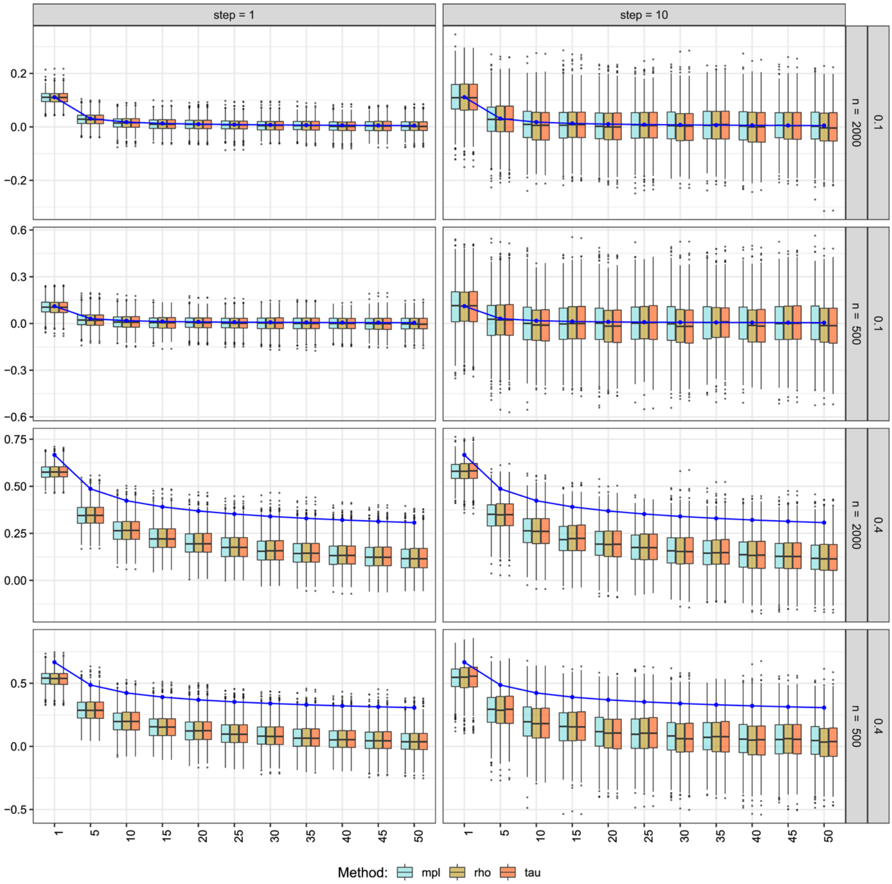

Fig. 2 Box-plots of the estimated values $\hat{\theta}_h$ obtained from methods `mpl`, `itau` and `irho` (all panels) with the Gaussian copula, considering pseudo observations obtained using the empirical distribution and step sizes 1 and 10 (column panels), for $d \in \{0.1, 0.4\}$ (row panels), $h \in \{1, 5, 10, \ldots, 45, 50\}$ (all panels), $n \in \{500, 2000\}$ (row panels). The connected blue lines represent the true value of θ_h

The results presented in Fig. 2 indicate that the three estimation methods perform similarly. The estimation bias increases with d and decreases with h and n. Overall, for large values of d, the estimation of the copula parameter present considerable bias, which remains fairly constant as the lag increases. The estimation pattern for $\hat{\theta}_h$ resembles the overall expected theoretical decay (blue line), with an almost constant bias. This behavior is likely to be a consequence of the well-known ill behavior of the covariance under strong long range dependence, which is reflected in the estimation of the copula parameter and, ultimately, in the estimation of d using the proposed approach. The variability in the estimates quickly increases from lag 1 to 10 and increases at a much slower rate for $h > 10$. The step size does not appear to have any influence on the bias, only on the variability of the estimates. More simulation results are provided in the supplementary material.

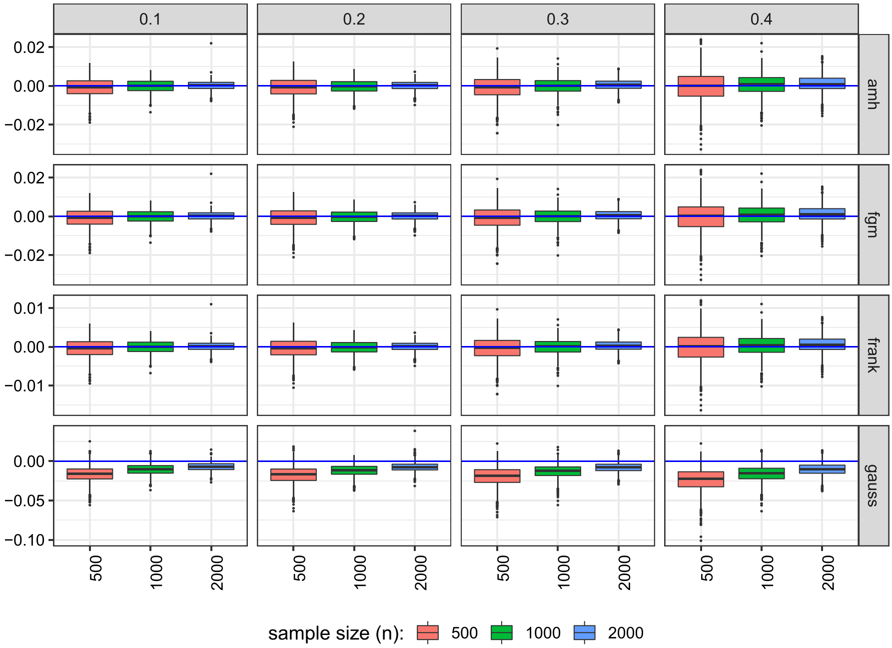

Fig. 3 Box-plots of the difference $\hat{K}_1^E - \hat{K}_1^G$ for $d \in \{0.1, 0.2, 0.3, 0.4\}$ (column panels), sample sizes $n \in \{500, 1000, 2000\}$ (all panels), considering the AMH, FGM, Frank and Gaussian copula families (row panels)

Estimation of K_1

Let $\hat{K}_1^E$ denote the estimator of K_1 obtained by using the estimators of F^{-1} and F' presented in Sect. 5.2 and $\hat{K}_1^G$ the estimator of K_1 based on the Gaussian quantile and density functions with mean and variance replaced by their sample estimators.

Figure 3 presents the box-plot of the difference $\hat{K}_1^E - \hat{K}_1^G$ for $d \in \{0.1, 0.2, 0.3, 0.4\}$ (column panels), sample sizes $n \in \{500, 1000, 2000\}$ (all panels), considering the AMH, FGM, Frank and Gaussian copula families. From Fig. 3 we observe that for the AMH, FGM, and Frank copulas, the box plots are fairly symmetric, centered at 0, and showing small variability. This indicates that using the empirical or Gaussian marginals makes little difference in the estimation of K_1. For the correctly specified Gaussian copula with Gaussian marginals, K_1 coincides with the variance. In most cases, we observe that $\hat{K}_1^G > \hat{K}_1^E$. However, the difference is very small and does not significantly affect the estimation of $\hat{d}$ (Figure 9 in the supplementary material).

Estimation of d and the influence of $\mathscr{D}$

Figure 4 shows the box-plots of the estimated values $\hat{d}$ considering $\hat{\theta}_h$ obtained from the `mpl` method, with the Gaussian copula and pseudo observations obtained with the empirical distribution, with step sizes 1 and 10 (column panels). To save space, we only report the results for $d \in \{0.1, 0.4\}$ and $n \in \{500, 2000\}$ (row panels).

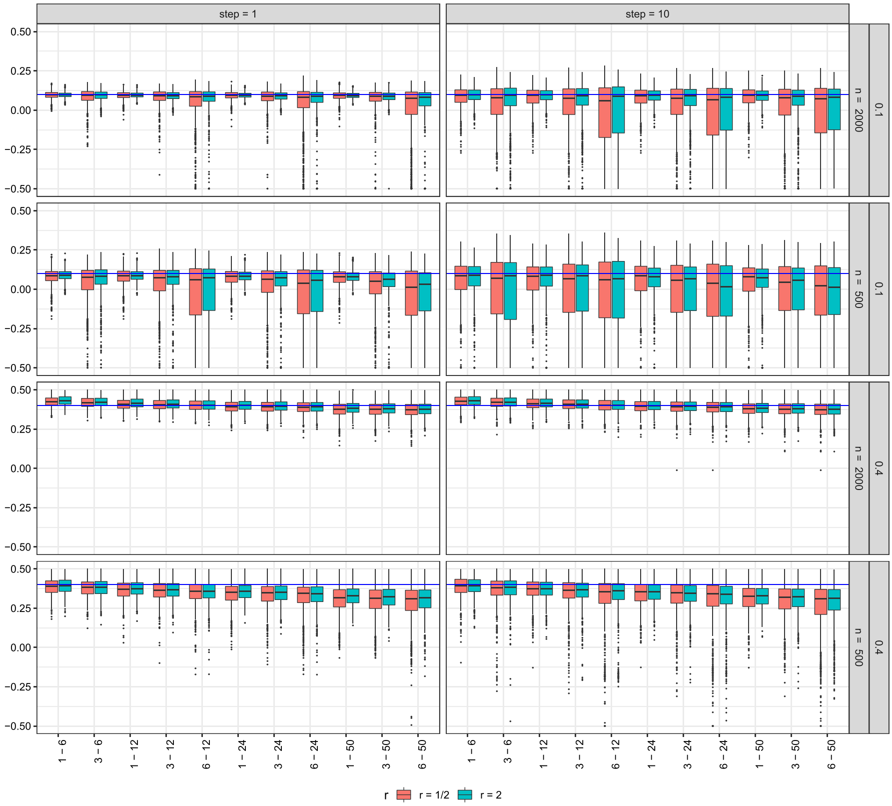

Fig. 4 Box-plots of the estimated values $\hat{d}$ considering $\hat{\theta}_h$ obtained from the `mpl` method, with Gaussian copula, pseudo observations obtained with the empirical distribution and step sizes 1 and 10 (column panels), for $d \in \{0.1, 0.4\}$ (row panels), $n \in \{500, 2000\}$ (row panels), $r \in \{1/2, 2\}$ (all panels) and each combination of s and m in $\{(s, m) \in \{1, 3, 6\} \times \{6, 12, 24, 50\} : s < m\}$ (labeled $s - m$ in all panels). In all panels, the horizontal blue line represents the true value of d

For all panels $r \in \{1/2, 2\}$, $\{(s, m) \in \{1, 3, 6\} \times \{6, 12, 24, 50\} : s < m\}$ (labeled $s - m$) and the horizontal line represents the true value of d. For methods `itau` and `irho` (presented in the supplementary material) the results are analogous.

Comparing the results for steps 1 and 10, we observe that step size 10 produces estimates with higher bias and variance than step size 1 (see detailed graphs in the supplementary material). This may be related to the variability observed in the estimates of θ_h. The variability clearly decreases with d, for any pair (s, m) and any r. This is expected, since θ_h is close to zero for small values of d (including 0.1), causing the objective function to become flat in the vicinities of d. The Euclidean metric ($r = 2$) presents smaller variability, but pointwise estimation is similar for both metrics. Choosing $r = 2$, $s = 1$, $m = 24$ and step size 1 seem to produce estimates with the smallest bias and variability. Using methods `itau` and `irho` yield analogous results (see the Supplementary Material).

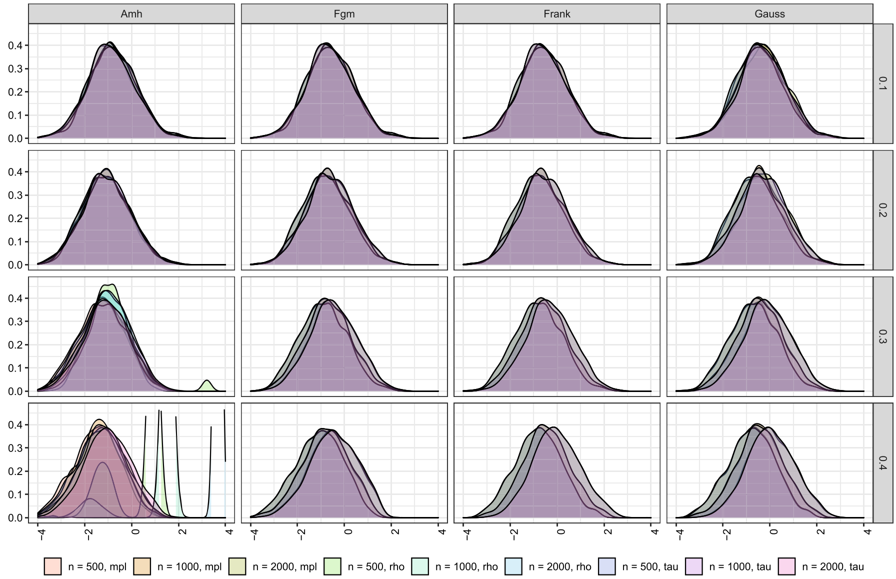

Fig. 5 Kernel density estimator of the standardized values of $\hat{d}$ considering $\hat{\theta}_h$ obtained from the `mpl`, `itau` and `irho` methods (all panels), using copulas AMH, FGM, Frank and Gaussian (column panels), pseudo observations obtained with the empirical distribution and step size 1, for $d \in \{0.1, 0.2, 0.3, 0.4\}$ (row panels), $n \in \{500, 1000, 2000\}$ (all panels), $r = 2$, $s = 1$ and $m = 24$

We also assess the asymptotic normality of the proposed estimator using 3 different copula estimator and under various scenarios. Figure 5 presents the kernel density estimates of standardized values of $\hat{d}$ obtained from `mpl`, `itau` and `irho` methods (all panels), using copulas AMH, FGM, Frank and Gaussian (column panels), considering pseudo observations obtained using the empirical distribution and step size 1, for $d \in \{0.1, 0.2, 0.3, 0.4\}$ (row panels), $n \in \{500, 1000, 2000\}$ (all panels), $r = 2$, $s = 1$ and $m = 24$.

Although with some visible bias, the plots show that applying the FGM and Frank copulas yield overall good results, despite the misspecified scenario and the copula estimator applied. Not surprisingly, the best results were obtained for the correctly specified Gaussian copula. For strong long range dependence ($d \in \{0.3, 0.4\}$) and considering the `irho` method for the AMH copula, we observe obvious departures from normality for all sample sizes considered. As the true value of d increases, the impact of the sample size n on the densities also increases in all cases. Overall, the simulation results suggest that the proposed methodology is fairly robust to copula misspecification and to the copula estimator applied.

Comparison with other estimators

We now compare the proposed estimator with some of the most commonly used ones in the literature. We consider five estimators for d, namely, the rescaled range estimator (R/S) proposed by Hurst (1951), the GPH estimator proposed by Geweke and Porter-Hudak (1983), the regression method based on the detrended fluctuation analysis (DFA) proposed by Peng et al. (1994), the local Whittle estimator (local.W) of Robinson (1995) and the Exact local Whittle estimator (ELW) of Shimotsu and Phillips (2005). The GPH, local.W and ELW were estimated using `R` package `LongMemoryTS` (Leschinski 2019). The bandwidth required in these estimators was set at $1+\sqrt{n}$ for all three. Estimators R/S and DFA were implemented by the authors. The detrended variance necessary to the estimator DFA was calculated using the `R` package `DCCA` (Prass and Pumi 2020). In calculating the detrended variance, we apply non-overlapping windows of sizes $\{50, 51, \dots, 100\}$ (Prass and Pumi 2021).

We apply the proposed estimator using $r = 2$, $s = 1$, $m = 24$ and step size 1, with copulas AMH, Frank, FGM and Gaussian, considering pseudo-observations obtained using the empirical distribution, for $n \in \{500, 1000, 2000\}$. The DGP is the same as before. For brevity, we only present the results considering the `mpl` estimator. The results using the `irho` and `itau` methods, presented in the supplementary material, are very similar.

The results are presented in Table 1. Regarding point estimation, for the proposed estimator the Gaussian copula is the best performer, as expected, followed closely by the Frank and FGM copulas. For $n = 500$ and $d = 0.1$ the proposed estimator performs uniformly better than the competitors regardless the copula. For other values of d, the proposed estimator is very competitive, always in the top three. The worst performer among all seems to be the local Whittle estimator, which presents a considerable bias.

As n increases, the results for all estimator improve, as expected, but especially so for the proposed estimator. For instance, for $n = 1000$, the proposed model present the smallest bias for all values of d, except for $d = 0.4$, for which the smallest bias is achieved by the GPH estimator. For $n = 2000$, the proposed estimator with the Gaussian copula performs uniformly better than the competitors with the Frank copula usually in the top 2 often followed by the FGM in the top 4. Finally, in terms of variability, the proposed estimator is almost uniformly better than all other competitors, regardless the copula. In terms of variability, the only competitor on par with the proposed one is the R/S. Curiously the proposed estimator using the AMH copula is the best overall performer in this regard.

Computational aspects

The simulation was performed on a PC equipped with 8GB of RAM and a Intel Core i7-8700 processor (3.20GHz, 6 cores, 12 threads), running linux Ubuntu 20.04. Simulations were performed using version 4.1.3 of R (R Core Team 2020), using the package `doParallel` (Microsoft Corporation and Weston 2020) for parallel execution, considering 4 cores (one for each value of d).

The estimation of the copula's parameters is the most demanding task. Each replication consists of estimating θ_h, for $h \in \{1, \dots, 50\}$, for a given scenario. The `itau` and `irho` methods are much faster than `mpl`. In the slowest case, running 1000 replications takes about 1 min, for `itau`, and 4 min, for `irho`. For the `mpl` method, the same task may take anywhere between 4 and 83 min, depending on d, step size and the marginal applied. In contrast, once the copula parameters are obtained for the 1000 replications of any given scenario, it only takes about 1.6 s on average to estimate d.

Table 1 Comparison between the proposed estimator for copulas AMH, Frank, FGM and Gaussian and the traditional R/S, GPH, DFA, local.W and ELW estimators, for the `mpl` method. Presented are the estimated value and standard deviation (in parenthesis)

d	DFA	GPH	local.W	ELW	R/S	Gaussian	Frank	FGM	AMH
	$n = 500$								
0.1	0.070	0.051	−0.037	0.046	0.157	0.085	0.079	0.079	0.072
	(0.179)	(0.163)	(0.138)	(0.112)	(0.056)	(0.037)	(0.036)	(0.036)	(0.034)
0.2	0.164	0.163	0.064	0.128	0.222	0.175	0.167	0.166	0.153
	(0.200)	(0.164)	(0.139)	(0.151)	(0.058)	(0.045)	(0.044)	(0.044)	(0.042)
0.3	0.257	0.277	0.167	0.242	0.283	0.265	0.259	0.254	0.235
	(0.219)	(0.165)	(0.139)	(0.173)	(0.060)	(0.052)	(0.053)	(0.052)	(0.050)
0.4	0.350	0.392	0.270	0.366	0.340	0.354	0.352	0.338	0.316
	(0.239)	(0.163)	(0.137)	(0.168)	(0.059)	(0.058)	(0.060)	(0.057)	(0.054)
	$n = 1{,}000$								
0.1	0.074	0.034	0.007	0.041	0.144	0.092	0.085	0.085	0.080
	(0.111)	(0.138)	(0.114)	(0.107)	(0.042)	(0.026)	(0.026)	(0.026)	(0.025)
0.2	0.169	0.152	0.108	0.122	0.213	0.187	0.179	0.177	0.165
	(0.122)	(0.138)	(0.115)	(0.140)	(0.044)	(0.033)	(0.033)	(0.033)	(0.032)
0.3	0.264	0.271	0.211	0.245	0.280	0.283	0.277	0.271	0.253
	(0.131)	(0.138)	(0.115)	(0.156)	(0.045)	(0.040)	(0.041)	(0.040)	(0.038)
0.4	0.360	0.392	0.312	0.373	0.340	0.380	0.378	0.363	0.341
	(0.140)	(0.139)	(0.112)	(0.148)	(0.045)	(0.046)	(0.048)	(0.045)	(0.042)
	$n = 2{,}000$								
0.1	0.077	0.040	0.034	0.038	0.133	0.095	0.089	0.089	0.084
	(0.073)	(0.111)	(0.091)	(0.083)	(0.037)	(0.019)	(0.018)	(0.018)	(0.018)
0.2	0.173	0.158	0.135	0.123	0.208	0.194	0.186	0.185	0.173
	(0.079)	(0.112)	(0.091)	(0.119)	(0.040)	(0.025)	(0.025)	(0.025)	(0.024)
0.3	0.271	0.278	0.238	0.256	0.280	0.297	0.290	0.285	0.267
	(0.085)	(0.112)	(0.091)	(0.139)	(0.041)	(0.031)	(0.032)	(0.031)	(0.029)
0.4	0.368	0.398	0.340	0.388	0.345	0.401	0.399	0.383	0.360
	(0.090)	(0.113)	(0.089)	(0.117)	(0.041)	(0.037)	(0.039)	(0.035)	(0.032)

6 Conclusion

In this work we investigate how long range dependence can be understood from the perspective of copulas. We uncovered a relationship between the covariance decay in univariate time series and the parametric copulas associated to lagged variables. Inspired by this relationship, a copula-based estimator for parameters identifiable through the covariance structure in univariate time series was proposed, excelling in the context of long range dependence. Being copula-based, the proposed methodology is very flexible, naturally accommodating non-Gaussian time series, missing data, as well as a wide range of marginal behavior. To the best of our knowledge, in this context, the proposed estimator is the first copula-based one in the literature.

We derive a rigorous asymptotic theory related to the proposed estimator. Under mild assumptions, we show its consistency and a central limit theorem. Its finite sample performance is investigated in great detail through a Monte Carlo simulation study. Among our findings, our simulation suggests that the proposed methodology is robust against misspecification of the copula family and the choice of the copula parameter estimator. We found that using the correctly specified marginals or using sample estimators has little affect on the estimator's performance. We also compared the proposed estimator with some of the most commonly applied ones in literature. We found that the proposed estimator is very competitive even in small samples, outperforming the competitor ones in most scenarios studied.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s00362-023-01418-z.

Acknowledgements T.S. Prass gratefully acknowledges the support of FAPERGS (ARD 01/2017, Processo 17/2551-0000826-0). S.R.C. Lopes' research was partially supported by CNPq - Brazil (303453/2018-4). The constructive comments of two reviewers are gratefully acknowledged.

Appendix 1: Mathematical proofs

Proof of Theorem 2.1: We present the proof for the case where $\boldsymbol{a} \notin \text{int}(\Theta)$. The other cases are dealt analogously. Let $\{\boldsymbol{\alpha}_m\}_{m\in\mathbb{N}^*}$ be an arbitrary sequence of parameters in D such that $\boldsymbol{\alpha}_m \to \boldsymbol{a}$ (assuming the adequate lateral limit when necessary, allowing for s coordinates to remain fixed). Applying a second order Taylor expansion in $\boldsymbol{\theta}_n$ around $\boldsymbol{a}$, apart from an $o(\|\boldsymbol{\theta}_n - \boldsymbol{a}\|^2)$ remainder, we have

$$
\begin{aligned}
C_{\boldsymbol{\theta}_n}(u,v) &= \lim_{m\to\infty} C_{\boldsymbol{\alpha}_m}(u,v) + \lim_{m\to\infty} d_{\boldsymbol{\theta}} C_{\boldsymbol{\alpha}_m}(u,v)(\boldsymbol{\theta}_n - \boldsymbol{a}) + \frac{1}{2}(\boldsymbol{\theta}_n - \boldsymbol{a})' \lim_{m\to\infty} d^2_{\boldsymbol{\theta}} C_{\boldsymbol{\alpha}_m}(u,v)(\boldsymbol{\theta}_n - \boldsymbol{a}) \\
&= uv + \sum_{i=1}^{k} \lim_{m\to\infty} \left[\frac{\partial C_{\boldsymbol{\theta}}(u,v)}{\partial\theta_i} \bigg|_{\boldsymbol{\theta}=\boldsymbol{\alpha}_m} \right] (\theta_n^{(i)} - a_i) + \\
&\quad + \frac{1}{2} \sum_{i,j=1}^{k} \lim_{m\to\infty} \left[\frac{\partial^2 C_{\boldsymbol{\theta}}(u,v)}{\partial\theta_i \partial\theta_j} \bigg|_{\boldsymbol{\theta}=\boldsymbol{\alpha}_m} \right] (\theta_n^{(i)} - a_i)(\theta_n^{(j)} - a_j).
\end{aligned} \tag{7}
$$

Let $\{X_n\}_{n\in\mathbb{N}}$ and $\{F_n\}_{n\in\mathbb{N}}$ be as in the enunciate. Hoeffding's lemma combined with (7) yields

$$
\begin{aligned}
\text{Cov}(X_0, X_n) &= \sum_{i=1}^{k} \left[\iint_{I^2} \frac{1}{l_0(u) l_n(v)} \lim_{m\to\infty} \frac{\partial C_{\boldsymbol{\theta}}(u,v)}{\partial\theta_i} \bigg|_{\boldsymbol{\theta}=\boldsymbol{\alpha}_m} du dv \right] (\theta_n^{(i)} - a_i) + \\
&+ \frac{1}{2} \sum_{i,j=1}^{k} \left[\iint_{I^2} \frac{1}{l_0(u) l_n(v)} \lim_{m\to\infty} \frac{\partial^2 C_{\boldsymbol{\theta}}(u,v)}{\partial\theta_i \partial\theta_j} \bigg|_{\boldsymbol{\theta}=\boldsymbol{\alpha}_m} du dv \right] \times (\theta_n^{(i)} - a_i)(\theta_n^{(j)} - a_j)
\end{aligned}
$$

$$= \sum_{i=1}^{k} K_1^{(i)}(n)(\theta_n^{(i)} - a_i) + \frac{1}{2}\sum_{i,j=1}^{k} K_2^{(i,j)}(n)(\theta_n^{(i)} - a_i)(\theta_n^{(j)} - a_j) \sim R(n) + o\big(R(n)\big) \sim R(n),$$

by the hypothesis on $K_1^{(i)}$ and $K_2^{(i,j)}$. ■

Proof of Lemma 4.1: Under **Framework A**, there exists $M > 0$ such that $\left|\lim_{\theta\to a} \frac{\partial C_\theta(u,v)}{\partial\theta}\right| \leq M$ for all $u, v \in I$, since it is a continuous function defined on a compact set. Now, since

$$\begin{aligned}
&|\hat{K}_1 - K_1| \\
&\leq \iint_{I^2} \left|\frac{1}{\hat{F}_n'\big(\hat{F}_n^{(-1)}(u)\big)\hat{F}_n'\big(\hat{F}_n^{(-1)}(v)\big)} - \frac{1}{F'\big(F^{(-1)}(u)\big)F'\big(F^{(-1)}(v)\big)}\right|\left|\lim_{\theta\to a}\frac{\partial C_\theta(u,v)}{\partial\theta}\right| du dv, \\
&\leq M\iint_{I^2} \left|\frac{1}{\hat{F}_n'\big(\hat{F}_n^{(-1)}(u)\big)\hat{F}_n'\big(\hat{F}_n^{(-1)}(v)\big)} - \frac{1}{F'\big(F^{(-1)}(u)\big)F'\big(F^{(-1)}(v)\big)}\right| du dv, \qquad (8)
\end{aligned}$$

given $\varepsilon > 0$, it follows from (8) that

$$\begin{aligned}
&\mathbb{P}\big(|\hat{K}_1 - K_1| > \varepsilon\big) \leq \\
&\leq \mathbb{P}\left(\iint_{I^2}\left|\frac{1}{\hat{F}_n'\big(\hat{F}_n^{(-1)}(u)\big)\hat{F}_n'\big(\hat{F}_n^{(-1)}(v)\big)} - \frac{1}{F'\big(F^{(-1)}(u)\big)F'\big(F^{(-1)}(v)\big)}\right| du dv > \frac{\varepsilon}{M}\right) \to 0,
\end{aligned}$$

as $n \to \infty$, by condition (a). Hence $\hat{K}_1 \xrightarrow{\mathbb{P}} K_1$ as desired. To complete the proof, observe that (b) $\Rightarrow$ (a) trivially. ■

Proof of Theorem 4.1: Under **A0**, $\hat{K}_1 \xrightarrow{\mathbb{P}} K_1$, while under **A1**, $\hat{\theta}_k(n) - a \xrightarrow{\mathbb{P}} \theta_k^0 - a \in \mathbb{R}$, for all $s \leq k \leq m$, as n tends to infinity, so that $\widehat{\boldsymbol{L}}_{s,m}(n) \xrightarrow{\mathbb{P}} \boldsymbol{R}_{s,m}(\boldsymbol{\eta}_0)$. Now, by assumption **A2** $\mathscr{D}$ is equivalent to the usual metric in $\mathbb{R}^{m-s+1}$ and since $(\mathbb{R}^{m-s+1}, \mathscr{D})$ is a complete metric space, it follows that,

$$\mathscr{D}\big(\widehat{\boldsymbol{L}}_{s,m}(n), \boldsymbol{R}_{s,m}(\boldsymbol{\eta}_0)\big) \xrightarrow{\mathbb{P}} 0, \qquad (9)$$

as n tends to infinity. Let $\hat{\boldsymbol{\eta}}_{s,m}(n)$ be as in (3) and notice that, for sufficiently large n,

$$\begin{aligned}
\mathscr{D}\big(\boldsymbol{R}_{s,m}(\boldsymbol{\eta}_0), \boldsymbol{R}_{s,m}(\hat{\boldsymbol{\eta}}_{s,m}(n))\big) &< \mathscr{D}\big(\boldsymbol{R}_{s,m}(\boldsymbol{\eta}_0), \widehat{\boldsymbol{L}}_{s,m}(n)\big) + \mathscr{D}\big(\widehat{\boldsymbol{L}}_{s,m}(n), \boldsymbol{R}_{s,m}(\hat{\boldsymbol{\eta}}_{s,m}(n))\big) \\
&< \varepsilon + \mathscr{D}\big(\widehat{\boldsymbol{L}}_{s,m}(n), \boldsymbol{R}_{s,m}(\hat{\boldsymbol{\eta}}_{s,m}(n))\big), \qquad (10)
\end{aligned}$$

hence $\lim_{n\to\infty} \mathscr{D}\big(\boldsymbol{R}_{s,m}(\boldsymbol{\eta}_0), \boldsymbol{R}_{s,m}(\hat{\boldsymbol{\eta}}_{s,m}(n))\big) \leq \lim_{n\to\infty} \mathscr{D}\big(\widehat{\boldsymbol{L}}_{s,m}(n), \boldsymbol{R}_{s,m}(\hat{\boldsymbol{\eta}}_{s,m}(n))\big)$. By the definition of $\hat{\boldsymbol{\eta}}_{s,m}(n)$, given $\delta > 0$, $\mathscr{D}\big(\widehat{\boldsymbol{L}}_{s,m}(n), \boldsymbol{R}_{s,m}(\hat{\boldsymbol{\eta}}_{s,m}(n))\big) \leq \mathscr{D}\big(\widehat{\boldsymbol{L}}_{s,m}(n), \boldsymbol{R}_{s,m}(\boldsymbol{\eta})\big)$, for all $\boldsymbol{\eta} \in \overline{B_\delta(\hat{\boldsymbol{\eta}}_{s,m}(n))}$, the closed ball in $\mathbb{R}^p$ with radius δ centered at $\hat{\boldsymbol{\eta}}_{s,m}(n)$. Now, by (9), it follows that for sufficiently large n, $\boldsymbol{\eta}_0 \in \overline{B_\delta(\hat{\boldsymbol{\eta}}_{s,m}(n))}$, so that

$$\mathscr{D}\big(\widehat{\boldsymbol{L}}_{s,m}(n), \boldsymbol{R}_{s,m}(\hat{\boldsymbol{\eta}}_{s,m}(n))\big) \leq \mathscr{D}\big(\widehat{\boldsymbol{L}}_{s,m}(n), \boldsymbol{R}_{s,m}(\boldsymbol{\eta}_0)\big) \xrightarrow{\mathbb{P}} 0.$$

Now, since $\mathscr{D}$ is a metric in $\mathbb{R}^{m-s+1}$, by the continuity of R and by the identifiability of $\boldsymbol{\eta}_0$, it follows that $\mathbb{P}\big(\|\hat{\boldsymbol{\eta}}_{s,m}(n) - \boldsymbol{\eta}_0\| < \varepsilon\big) \longrightarrow 1$, as n tends to infinity. ∎

Proof of Theorem 4.2: Without loss of generality, we shall assume that $a = 0$. Let $s_0 = \max\{k_0, k_1\}$ and $\Omega = \Omega_0 \cap \Omega_1$ in **A3** and **A4** and let $s > s_0$. Under the hypothesis, upon defining

$$S_{s,m}(\boldsymbol{\eta}; n) = \sum_{k=0}^{m-s} \big(\widehat{\theta}_{s+k}(n) - L(s+k, \boldsymbol{\eta})\big)^2,$$

as n tends to infinity, with probability 1

$$\mathbf{0} = \left.\frac{\partial S_{s,m}(\boldsymbol{\eta}; n)}{\partial \boldsymbol{\eta}}\right|_{\widehat{\boldsymbol{\eta}}} = \left.\frac{\partial S_{s,m}(\boldsymbol{\eta}; n)}{\partial \boldsymbol{\eta}}\right|_{\boldsymbol{\eta}_0} + \left(\left.\frac{\partial^2 S_{s,m}(\boldsymbol{\eta}; n)}{\partial \boldsymbol{\eta} \partial \boldsymbol{\eta}'}\right|_{\overline{\boldsymbol{\eta}}}\right)(\widehat{\boldsymbol{\eta}} - \boldsymbol{\eta}_0),$$

for some $\overline{\boldsymbol{\eta}} \in \Omega$ such that $\|\overline{\boldsymbol{\eta}} - \boldsymbol{\eta}_0\| \leq \|\widehat{\boldsymbol{\eta}} - \boldsymbol{\eta}_0\|$. In order to prove the result, it suffices to show that

$$b_n \left(\left.\frac{\partial S_{s,m}(\boldsymbol{\eta}; n)}{\partial \boldsymbol{\eta}}\right|_{\boldsymbol{\eta}_0}\right) \xrightarrow{d} -2 \sum_{k=0}^{m-s} \boldsymbol{a}_{s+k} Z_{s+k} \quad \text{and} \quad \left.\frac{\partial^2 S_{s,m}(\boldsymbol{\eta}; n)}{\partial \boldsymbol{\eta} \partial \boldsymbol{\eta}'}\right|_{\overline{\boldsymbol{\eta}}} \xrightarrow{\mathbb{P}} 2 \sum_{k=0}^{m-s} \boldsymbol{a}_{s+k} \boldsymbol{a}'_{s+k} \tag{11}$$

and to observe that the right hand side in the second relation in (11) is positive definite by **A3**. On one hand, by **A3**, we can write

$$\frac{\partial S_{s,m}(\boldsymbol{\eta}; n)}{\partial \boldsymbol{\eta}} = -2 \sum_{k=0}^{m-s} \boldsymbol{a}_{s+k} \big(\widehat{\theta}_{s+k}(n) - L(s+k, \boldsymbol{\eta})\big),$$

so that the first equation in (11) follows from **A4** by multiplying both sides by b_n and taking the limit as n tends to infinity. On the other hand, by **A3** and **A4**, for $\boldsymbol{\eta} \in \Omega$,

$$\begin{aligned} \frac{\partial^2 S_{s,m}(\boldsymbol{\eta}; n)}{\partial \boldsymbol{\eta} \partial \boldsymbol{\eta}'} &= -2 \sum_{k=0}^{m-s} \left\{ \frac{\partial^2 L(s+k, \boldsymbol{\eta})}{\partial \boldsymbol{\eta} \partial \boldsymbol{\eta}'} \big(\widehat{\theta}_{s+k}(n) - L(s+k, \boldsymbol{\eta})\big) - \boldsymbol{a}_{s+k} \boldsymbol{a}'_{s+k} \right\} \\ &= 2 \sum_{k=0}^{m-s} \boldsymbol{a}_{s+k} \boldsymbol{a}'_{s+k} + o_P(1), \end{aligned}$$

and the proof is complete. ∎

References

Beare BK (2010) Copulas and temporal dependence. Econometrica 78(1):395–410
Beare BK (2012) Archimedean copulas and temporal dependence. Economet Theor 28(6):1165–1185

Bedford T, Cooke RM (2001) Probability density decomposition for conditionally dependent random variables modeled by vines. Ann Math Artif Intell 32(1):245–268
Bedford T, Cooke RM (2002) Vines—a new graphical model for dependent random variables. Ann Stat 30(4):1031–1068
Beran J (2016) On the effect of long-range dependence on extreme value copula estimation with fixed marginals. Commun Stat 45(19):5590–5618
Borchers HW (2021) pracma: practical numerical math functions. R package version 2.3.3. https://CRAN.R-project.org/package=pracma
Brockwell PJ, Davis RA (1991) Time series: theory and methods, 2nd edn. Springer, New York
Bücher A, Volgushev S (2013) Empirical and sequential empirical copula processes under serial dependence. J Multivar Anal 119:61–70
Chen X, Fan Y (2006) Estimation of copula-based semiparametric time series models. J Econ 130(2):307–335
Chen X, Wu WB, Yi Y (2009) Efficient estimation of copula-based semiparametric Markov models. Ann Stat 37(6B):4214–4253
Cherubini U, Luciano E, Vecchiato W (2004) Copula methods in finance. Wiley, Hoboken
Darsow WF, Nguyen B, Olsen ET (1992) Copulas and Markov processes. Ill J Math 36(4):600–642
Geweke J, Porter-Hudak S (1983) The estimation and application of long memory time series models. J Time Ser Anal 4(4):221–238
Giraitis L, Surgailis D (1999) Central limit theorem for the empirical process of a linear sequence with long memory. J Stat Plan Inference 80:81–93
Hall P, Hart JD (1990) Convergence rates in density estimation for data from infinite-order moving average processes. Probab Theory Relat Fields 87:253–274
Härdle WK, Mungo J (2008) Value-at-risk and expected shortfall when there is long range dependence. SFB 649 Discussion Paper, 2008–006
Hofert M, Kojadinovic I, Maechler M, Yan J (2020) copula: multivariate dependence with copulas. R package version 1.0-1. https://CRAN.R-project.org/package=copula
Hurst HE (1951) Long-term storage capacity of reservoirs. Trans Am Soc Civ Eng 116(1):770–799
Ibragimov R (2009) Copula-based characterizations for higher order Markov processes. Economet Theor 25(3):819–846
Ibragimov R, Lentzas G (2017) Copulas and long memory. Probab Surv 14:289–327
Joe H (1997) Multivariate models and multivariate dependence concepts. Chapman & Hall/CRC Monographs on Statistics & Applied Probability. Taylor & Francis, London
Lagerås AN (2010) Copulas for Markovian dependence. Bernoulli 16(2):331–342
Lee T-H, Long X (2009) Copula-based multivariate Garch model with uncorrelated dependent errors. J Econ 150(2):207–218
Leschinski C (2019) LongMemoryTS: Long Memory Time Series. R package version 0.1.0. https://CRAN.R-project.org/package=LongMemoryTS
Marinucci D (2005) The empirical process for bivariate sequences with long memory. Stat Infer Stoch Process 8:205–223
McNeil A, Frey R, Embrechts P (2010) Quantitative risk management: concepts, techniques, and tools: concepts, techniques, and tools. Princeton Series in Finance. Princeton University Press
Mendes BV, Kolev N (2008) How long memory in volatility affects true dependence structure. Int Rev Fin Anal 17(5):1070–1086
Microsoft Corporation, Weston S (2020) doParallel: Foreach Parallel Adaptor for the 'parallel' Package. R package version 1.0.16. https://CRAN.R-project.org/package=doParallel
Nelsen R (2013) An introduction to copulas. Lecture notes in statistics, 2nd edn. Springer, New York
Palma W (2007) Long-memory time series: theory and methods. Wiley series in probability and statistics. Wiley, Hoboken
Peng C-K, Buldyrev SV, Havlin S, Simons M, Stanley HE, Goldberger AL (1994) Mosaic organization of DNA nucleotides. Phys Rev E 49:1685–1689
Prass TS, Pumi G (2020) DCCA: detrended fluctuation and detrended cross-correlation analysis. R package version 0.1.1. https://CRAN.R-project.org/package=DCCA
Prass TS, Pumi G (2021) On the behavior of the DFA and DCCA in trend-stationary processes. J Multivar Anal 182:104703
R Core Team (2020) R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/

Robinson PM (1995) Gaussian semiparametric estimation of long range dependence. Ann Stat 23(5):1630–1661

Shimotsu K, Phillips PCB (2005) Exact local Whittle estimation of fractional integration. Ann Stat 33(4):1890–1933

Stute W, Schumann G (1980) A general Glivenko–Cantelli theorem for stationary sequences of random observations. Scand J Stat 7(2):102–104