Some Aspects of Modeling Dependence in Copula-based Markov chains

Martial Longla and Magda Peligrad [1]

Department of Mathematical Sciences, University of Cincinnati, PO Box 210025, Cincinnati, Oh 45221-0025, USA.
E-mail addresses: martiala@mail.uc.edu and peligrm@ucmail.uc.edu

Abstract

Dependence coefficients have been widely studied for Markov processes defined by a set of transition probabilities and an initial distribution. This work clarifies some aspects of the theory of dependence structure of Markov chains generated by copulas that are useful in time series econometrics and other applied fields. The main aim of this paper is to clarify the relationship between the notions of geometric ergodicity and geometric ρ-mixing; namely, to point out that for a large number of well known copulas, such as Clayton, Gumbel or Student, these notions are equivalent. Some of the results published in the last years appear to be redundant if one takes into account this fact. We apply this equivalence to show that any mixture of Clayton, Gumbel or Student copulas generate both geometrically ergodic and geometric $\rho-$mixing stationary Markov chains, answering in this way an open question in the literature. We shall also point out that a sufficient condition for $\rho-$mixing, used in the literature, actually implies Doeblin recurrence.

Key words: Markov chains, copula, mixing conditions, reversible processes.

AMS 2000 Subject Classification: Primary 60J20, 60J35, 37A30.

1 Introduction

In recent years copula-based methods became a popular tool for analyzing temporal dependence of time series. A 2-copula is a bivariate distribution function $C(x, y)$ with uniform marginal distributions. Given a stationary Markov chain $(X_n)_{n \in \mathbb{Z}}$ with marginal distribution function $F(x)$, the process is characterized by the bivariate distribution function of (X_1, X_2) denoted by $H(x_1, x_2) = P(X_1 \leq x_1, X_2 \leq x_2)$. Then, by Sklar's theorem (see for instance Nelsen [16]), one can express $H(x_1, x_2)$ in terms of a copula $C(x_1, x_2)$ and $F(x)$ via

$$H(x_1, x_2) = C(F(x_1), F(x_2)) .$$

The copula is uniquely defined on the product of the range of $F(x)$ by itself. So, it is unique if $F(x)$ is continuous, and otherwise can be uniqely constructed by using a bilinear interpolation. Therefore one can specify a stationary Markov process by providing an invariant distribution function and a copula. The copula approach is flexible, since the marginal behavior characterized by $F(x)$ can be separated from the temporal dependence described by $C(x, y)$. In their recent paper, de Vries, C. G. and Zhou, C. [11] point out two examples from economics where this separation is useful. Many interesting patterns of temporal dependence in various applied fields of research can be generated by using certain copula functions. Various procedures for estimating these models have been proposed, ranging from parametric to nonparametric models (see for instance Chen and Fan [7], Chen et al. [8], and the references therein). To establish the asymptotic properties of any of these estimators, one needs to know the temporal dependence properties of the Markov chains, usually described in terms of mixing coefficients. There are a large number of papers in the literature that address this problem. Among them we mention Chen and Fan [7], Gagliardini and Gouriéroux [12], Chen et al. [8], Ibragimov and Lentzas [13], Beare [2].

This work is motivated in fact by the paper by Chen et al. [8]. In their Proposition 2.1, it was shown that Markov processes generated by the Clayton, Gumbel or Student copulas are geometrically ergodic. Their method of proof is based on a sophisticated quantile transformations and construction of small sets for each individual copula. However it is not obvious how to construct small sets to handle for instance the mixture of these copulas. Wei Biao Wu raised the question whether convex combinations of these copulas generate geometrically ergodic Markov chains. We shall positively answer this question. The derivation of this result is based on the theory of the geometric ergodicity of reversible Markov chains developed by Roberts and Rosenthal [18], Roberts and Tweedie [17] and Kontoyiannis and Meyn [14]. This theory stresses the importance of estimating the maximal coefficient of correlation between two consecutive random variables in the Markov chain.

We shall also comment on a class of stationary Markov chains for which Beare [1, Theorem 4.2] showed that is $\rho-$mixing. We shall actually show that this class satisfies a more restrictive condition, namely $\phi-$mixing, and so, the estimators will enjoy richer asymptotic properties. Precisely, we shall show that if the density of the absolutely continuous part of a copula is bounded away from 0 on a set of Lebesgue measure 1, then it generates $\phi-$mixing Markov chains.

Our paper is organized as follows. First we give a brief survey of 3-mixing coefficients that are closely related and formulate them in the specific copula terms. In Section 3 we discuss the equivalence between geometric ergodicity and geometric $\rho-$mixing for Markov chains with symmetric copulas. Section 4 treats Doeblin recurrence property. The mathematical arguments are included in Section 5.

Throughout the paper we denote by $I = [0, 1]$, by $\mathcal{R}$ we denote the Borelian sets on R and $\lambda(.)$ denotes the Lebesgue measure. By $||g||_{p,\lambda}$ we de-

note $\left(\int_I |g(x)|^p d\lambda\right)^{1/p}$. For a random variable X defined on a probability space $(\Omega, K, \mathbb{P})$ we denote by $||X||_p = \mathbb{E}(|X|^p)^{1/p}$. The notation a.s. stays for almost sure. By dx, dy, ... we denote the integral with respect to Lebesgue measure on I. For a function $f(x,y)$ we denote by $f_{,1}(x,y)$, $f_{,2}(x,y)$ and $f_{,12}(x,y)$ the partial derivative with respect to x, y, and second mixed derivative respectively. For a set B we denote by B' the complement of B.

2 Three mixing coefficients

In this paper we shall discuss the following 3 mixing coefficients. Let $(\Omega, \mathcal{K}, \mathbb{P})$ be a probability space and let $\mathcal{A}, \mathcal{B}$ be two σ-algebras included in $\mathcal{K}$. Define the absolutely regular coefficient between $\mathcal{A}, \mathcal{B}$ by

$$\beta(\mathcal{A}, \mathcal{B}) = \frac{1}{2} \sup_{\{A_i\},\{B_j\}} \sum_{i=1}^{n} \sum_{j=1}^{m} |\mathbb{P}(A_i \cap B_j) - \mathbb{P}(A_i)\mathbb{P}(B_j)| ,$$

where the supremum is taken over all positive integers n and m, and all finite partitions $\{A_i\}, \{B_j\}$ of Ω with $A_i \in \mathcal{A}$ and $B_j \in \mathcal{B}$.

The maximal coefficient of correlation is defined by

$$\rho(\mathcal{A}, \mathcal{B}) = \sup\{corr(f,g), \ f \in \mathbb{L}_2(\mathcal{A}), \ g \in \mathbb{L}_2(\mathcal{B})\} .$$

where $\mathbb{L}_2(\mathcal{A})$ is the space of random variables that are $\mathcal{A}$ measurable and square integrable.

The uniform mixing coefficient is

$$\phi(\mathcal{A}, \mathcal{B}) = \sup_{B \in \mathcal{B}, A \in \mathcal{A}, \mathbb{P}(A) > 0} |\mathbb{P}(B|A) - P(B)| .$$

For a stationary sequence $(X_n)_{n \in Z}$ let $\mathcal{P} = \sigma(X_k, k \leq 0)$ be the information provided by the past of the process and $\mathcal{F}_n = \sigma(X_k, k \geq n)$ describes the future after n steps. Then define $\beta_n = \beta(\mathcal{P}, \mathcal{F}_n)$, $\rho_n = \rho(\mathcal{P}, \mathcal{F}_n)$, and $\phi_n = \phi(\mathcal{P}, \mathcal{F}_n)$. It is well known that $\beta_n \leq \phi_n$ and $\rho_n \leq 2\sqrt{\phi_n}$. If in addition the sequence is Markov the coefficients simplify and we have $\beta_n = \beta(\sigma(X_0), \sigma(X_n))$, $\rho_n = \rho(\sigma(X_0), \sigma(X_n))$, and $\phi_n = \phi(\sigma(X_0), \sigma(X_n))$. Moreover $\rho_n \leq (\rho_1)^n$ and $(2\phi_n) \leq (2\phi_1)^n$. There are examples of Markov chains such that $\rho_n \to 0$ but $\phi_n \nrightarrow 0$, and also $\rho_n \to 0$ but $\beta_n \nrightarrow 0$ or $\beta_n \to 0$ but $\rho_n \nrightarrow 0$. A convenient reference to all these results are the books by Bradley [3].

In terms of conditional probabilities, denoted by $P^n(x,B) = \mathbb{P}(X_n \in B|X_0 = x)$, and marginal distribution function $F(x)$, which generates the invariant measure $\pi(A) = P(X_0 \in A)$, by using the equivalent definitions of the mixing coefficients (see [3, Theorem 3.32 and Lemma 4.3]) we have

$$\beta_n = \int_R \sup_{B \in \mathcal{R}} |P^n(x,B) - \pi(B)| dF,$$

3

$$\rho_n = \sup\{\left(\int_R \left(\int_R g(y)P^n(x,dy)\right)^2 dF\right)^{1/2}, \ \int_R g^2(y)dF(y) = 1, \ \mathbb{E}g = 0\},$$

and

$$\phi_n = \sup_{B \in \mathcal{B}} ess\sup_{x \in R} |P^n(x,B) - \pi(B)|.$$

Remark 1 *We should mention that, all these mixing coefficients for stationary Markov chains are invariant under strictly increasing and continuous transformations of the variables. Then, without restricting the generality, we can replace in their computations X_n by $U_n = F(X_n)$, where $F(x)$, the distribution function of X_0, is assumed continuous and strictly increasing. Since U_0 and U_n are both uniformly distributed on $[0,1]$ these coefficients are characterized only by copulas.*

We shall express next the mixing coefficients of a Markov chain in the specific terms of copula characteristics. One of the most important notions that facilitates the link is the fold product of copulas, defined by Relation (2.10) in Darsow et al. [10] as follows:

Definition 2 *Let $C_1(x,y)$ and $C_2(x,y)$ be two copulas. Their fold product is*

$$A(x,y) = C_1 * C_2(x,y) = \int_I \frac{\partial C_1}{\partial y}(x,t)\frac{\partial C_2}{\partial x}(t,y)dt .$$

This operation is associative, distributive over convex combinations of copulas and the set of copulas is closed under it. For more details about the product of copulas, see Darsow et al. [10] and also Nelsen [16], where it is also proven that copulas are almost everywhere differentiable. Furthermore, for all $n \geq 1$ and $y \in [0,1]$ the transition probabilities of the stationary Markov chain, $(U_i)_{i \in \mathbb{Z}}$, with uniform marginal distributions and copula $C(x,y)$ is given by

$$\mathbb{P}(U_n \leq y|U_0 = x) = C_{,1}^n(x,y) \text{ a.s. }, \tag{1}$$

where $C^n(x,y)$ is the n-th fold product of $C(x,y) = C^1(x,y)$ with itself. Then, we can construct a set Ω of Lesbegue measure 1, such that for all $x \in \Omega$ we have $\mathbb{P}(U_n \leq y|U_0 = x) = C_{,1}^n(x,y)$ for all y rational, and we deduce that for any x in Ω and any Borelian A

$$P^n(x,A) = \mathbb{P}(U_n \in A|U_0 = x) = C_{,1}^n(x,A) , \tag{2}$$

where by $C_{,1}^n(x,A)$ we denote the measure induced by $C_{,1}^n(x,y) = C_{,1}^n(x,[0,y])$.

By using these notations we have the following reformulation of the mixing coefficients for $(U_n)_{n \in \mathbb{Z}}$, a stationary Markov chain with uniform marginal distributions, in terms of copula $C^n(x,y)$ associated to variables (U_0, U_n):

$$\beta_n = \int_0^1 \sup_{B \in \mathcal{R} \cap I} |C_{,1}^n(x,B) - \lambda(B)|dx,$$

$$\rho_n = \sup\{ \left(\int_0^1 \left(\int_0^1 g(y) C_{,1}^n(x, dy) \right)^2 dx \right)^{1/2} , \ ||g||_{2,\lambda} = 1, \ \mathbb{E}g = 0\}$$

and

$$\phi_n = \sup_{B \in \mathcal{B}} ess \sup_{x \in I} |C_{,1}^n(x, B) - \lambda(B)| \ .$$

If in addition the copula $C^n(x, y)$ is absolutely continuous with respect to λ^2, and denoting its density by $c_n(x, y)$ then, these coefficients become

$$\beta_n = \int_0^1 \sup_{B \in \mathcal{R} \cap I} | \int_B (c_n(x, y) - 1) dy | dx \ .$$

$$\rho_n = \sup\{ \int_0^1 \int_0^1 c_n(x, y) f(x) g(y) dx dy : ||g||_{2,\lambda} = ||f||_{2,\lambda} = 1, \ \mathbb{E}f = \mathbb{E}g = 0\} \ ,$$

$$\phi_n = \sup_{B \subset \mathcal{R} \cap I} ess \sup_{x \in I} | \int_B (c_n(x, y) - 1) dy | \ . \tag{3}$$

It is easy to see that the coefficients for $(U_i)_{i \in \mathbb{Z}}$ are robust in the following sense:

Remark 3 *The mixing coefficients of a Markov chain generated by a given copula $C(x, y)$ and marginal distribution uniform on $[0, 1]$, are larger than or equal to those of a Markov chain generated by the same copula and another marginal distribution $F(x)$. As a matter of fact they are the supremum over all $F(x)$, of the mixing coefficients of stationary Markov chains generated by the copula.*

To see this we consider the generalized inverse,

$$F^{-1}(u) = \inf\{x, u \le F(x)\} \ .$$

Notice that $x \ge F^{-1}(u)$ if and only if $F(x) \ge u$. Given the stationary Markov chain $(U_i)_{i \in \mathbb{Z}}$ generated by copula $C(x, y)$ and uniform on $[0, 1]$ marginal distribution, then, the stationary Markov chain $(F^{-1}(U_i))_{i \in \mathbb{Z}}$ has the marginal distribution function $F(x)$ and the same copula. It remains to notice that $\sigma(F^{-1}(U_i)) \subset \sigma(U_i)$.

We shall make the following convention:

Convention: Given a copula $C(x, y)$ we shall refer to the Markov chain it generates, without specifying its marginal distribution, if this distribution is uniform on $[0, 1]$.

3 Geometric ergodicity

An important notion for the Markov chains is the notion of absolute regularity. A stationary sequence is said to be absolutely regular if $\beta_n \to 0$ as $n \to \infty$. It is well known (see for instance Corollary 21.7 in Bradley vol. 2 [3]) that a strictly stationary Markov chain is absolutely regular (i.e. $\beta_n \to 0$) if and only if it is irreducible, (i.e. Harris recurrent) and aperiodic. A Markov chain is irreducible if there exists a set B, such that $\pi(B) = 1$ and the following holds: for all $x \in B$ and every set $A \in \mathcal{R}$ such that $\pi(A) > 0$, there is a positive integer $n = n(x, A)$ for which $P^n(x, A) > 0$. An irreducible stationary Markov chain is aperiodic if and only if there is A with $\pi(A) > 0$ and a positive integer n such that $P^n(x, A) > 0$ and $P^{n+1}(x, A) > 0$ for all $x \in A$ (see Chan and Tong [6, Theorem 3.3.1]).

By using these definitions along with measure theoretical arguments we shall prove the following general result, where we impose a less restrictive condition than Assumption 1 in Chen and Fan [7].

Proposition 4 *If the absolutely continuous part of a copula has a strictly positive density on a set of measure 1, then it generates an absolutely regular Markov chain.*

It is well known that any convex combination of copulas is still a copula. We shall comment next on the absolute regularity of such a mixture of copulas and point out that it will inherit this property from one of the copulas in the combination. We present this fact as a lemma that is needed for our proofs.

Lemma 5 *Let $(C_k(x, y); 1 \le k \le n)$ be n copulas such that $C_1(x, y)$ generates an absolutely regular Markov chain. Any stationary Markov chain generated by a convex combination, $\sum_{k=1}^{n} a_k C_k(x, y)$ with $\sum_{k=1}^{n} a_k = 1$, $0 \le a_k \le 1$, $a_1 \neq 0$, is absolutely regular.*

3.1 Speed of convergence

The speed of convergence to 0 of the mixing coefficients is a very important question for establishing limit theorems for estimators and their speed of convergence.

We shall say that a sequence is geometric $\beta-$mixing (or geometric absolutely regular) if there is $0 < \gamma < 1$ such that $\beta_n < \gamma^n$.

We say that the sequence is geometric $\rho-$mixing if there is $0 < \delta < 1$ such that $\rho_n \le \delta^n$. For a stationary Markov chain, because $\rho_n \le \rho_1^n$, we have that $\rho_1 < 1$ implies $\rho_n \le \delta^n$ with $\delta = \rho_1$.

In this section we are going to use an equivalent definition for $\rho-$mixing coefficients in terms of the operator associated to the Markov chain. As before, denote the marginal distribution by $\pi(A) = \mathbb{P}(X_0 \in A)$ and assume there is a regular conditional distribution for X_1 given X_0 denoted by $P(x, A) = \mathbb{P}(X_1 \in A | X_0 = x)$. In addition P denotes the Markov operator acting via $(Pf)(x) = \int_S f(s)P(x, ds)$. Next let $\mathbb{L}_2^0(\pi)$ be the set of measurable functions such that

$\int f^2 d\pi < \infty$ and $\int f d\pi = 0$. With these notations, the coefficient ρ_1 is simply the norm operator of $P : \mathbb{L}_2^0(\pi) \to \mathbb{L}_2^0(\pi)$,

$$\rho_1 = ||P||_{\mathbb{L}_2^0(\pi)} = \sup_{g \in \mathbb{L}_2^0(\pi)} \frac{||P(g)||_2}{||g||_2} \ . \tag{4}$$

Still in this Markov setting, geometric $\beta-$mixing is equivalent to the notion of geometric ergodicity that means there exists a measurable function $A(x)$ such that for some $0 < \gamma < 1$ and for all $n \geq 1$

$$||P^n(x, .) - \pi(.)||_{\text{tot var}} \leq A(x)\gamma^n \text{ a.s.}$$

A convenient reference to these results is Theorem 21.19 in Bradley [3], or Meyn and Tweedie [15].

We say that the stationary Markov chain is reversible if (X_0, X_1) and (X_1, X_0) are identically distributed. Equivalently P is self-adjoint. In the context of reversible irreducible and aperiodic Markov chains $1 - \rho_1$ equals the so called spectral gap, and if $\rho_1 < 1$ we say that the operator P has a spectral gap in $\mathbb{L}_2$. For a convenient reference to spectral theory we mention the book by Conway [9]. See also the remarks above and after Theorem 2.1 in Roberts and Rosenthal [18] and Lemma 2.2 in Kontoyannis and Meyn [14].

Based partially on results of Roberts and Rosenthal [18], Roberts and Tweedie [17], Kontoyannis and Meyn [14], in their Proposition 1.2, state that any irreducible and aperiodic reversible Markov chain is geometrically ergodic if and only if has a spectral gap in $\mathbb{L}_2(\pi)$. In view of previous comments we formulate their result in the following language which is familiar to researchers in applied areas:

Theorem 6 *Any irreducible and aperiodic reversible Markov chain is geometrically ergodic if and only if $\rho_1 < 1$.*

This important result is the key for obtaining the following statement.

Theorem 7 *Let $(C_k(x, y); 1 \leq k \leq n)$ be n symmetric copulas that generate geometrically ergodic Markov chains. Any stationary Markov chain generated by a convex combination of these copulas is geometrically ergodic and geometric $\rho-$mixing.*

These results have rich implications. We shall give two corollaries that are useful in applications. Combining Proposition 4 and Theorem 6 leads to:

Corollary 8 *A symmetric copula with the density of its absolutely continuous part strictly positive on a set of Lebesgue measure 1 generates a geometrically ergodic stationary Markov chain if and only if $\rho_1 < 1$.*

By combining now Lemma 5 with Theorem 6 one obtains:

Corollary 9 *Assume $(C_k(x, y); 1 \leq k \leq n)$ are n symmetric copulas and $C_1(x, y)$ has the density of its absolute continuous part strictly positive on a set of Lebesgue measure 1. Assume each one generates a $\rho-$mixing Markov chain. Then, any convex combination, $\sum_{k=1}^{n} a_k C_k(x, y)$ with $\sum_{k=1}^{n} a_k = 1, 0 \leq a_i \leq 1$, $a_1 \neq 0$ generates a geometrically ergodic Markov chain.*

Based on these results we can give the following examples:

3.2 Examples

1. The Student t-copula, Clayton and Gumbel copulas generate geometric ρ-mixing Markov chains. It was shown by Chen et al. [8] that these copulas generate geometrically ergodic stationary Markov chains, and then, an application of Corollary 8 proves our statement. It should be noticed that Beare [1], also states that the t copula generates geometric ρ-mixing, but his proof has a gap. It is based on a theorem that does not apply to the t-copula, since its density is not bounded away from 0. He also made a numerical study that confirms our statement that Clayton and Gumbel copulas generate geometric ρ-mixing Markov chains.

Student t-copula is given by

$$C_{\rho,\nu}(u, v) = t_{\rho,\nu}(t_\nu^{-1}(u), t_\nu^{-1}(v)), \quad |\rho| < 1, \quad \nu \in (1, \infty) ,$$

where $t_{\rho,\nu}(.,.)$ is the distribution function of the bivariate Student-t distribution with mean zero, the correlation matrix having off-diagonal element ρ, and ν degrees of freedom, and $t_\nu(.)$ is the distribution function of a univariate Student-t distribution with mean zero, and ν degrees of freedom.

2. Any convex combination of Clayton, Gumbel and t-copulas generates a geometrically ergodic stationary Markov chain (and thus, geometric ρ-mixing). This is due to the fact that all these copulas are symmetric in their variables and we apply then Theorem 7. This statement positively answers the question posed by Wei Biao Wu on this topic. The Clayton and Gumbel copulas are respectively

$$C_\theta(u, v) = (u^{-\theta} + v^{-\theta} - 1)^{-1/\theta}, \quad \theta \in (0, \infty) ,$$

$$C_\beta(u, v) = \exp(-[(-\ln u)^\beta + (-\ln v)^\beta]^{1/\beta}), \quad \beta \in [1, \infty) .$$

3. All Archimedian copulas that were shown to be geometrically ergodic by Beare [2] and their convex combinations also generate geometric ρ-mixing by Theorem 7.

4 Doeblin recurrence

Beare , in [1, Theorem 4.2], based on arguments related to results in Breiman and Friedman [4] and Bryc [5], showed that if the density of the absolutely

continuous part of a copula is bounded away from 0 a.s., then $\rho_1 < 1$. Actually we shall prove that more can be said under this condition, namely this condition implies ϕ-mixing and therefore geometric ergodicity for the generated Markov chain.

Theorem 10 *Assume the density of the absolutely continuous part of the copula $C(x, y)$ exists and is bounded away from 0 on a set of Lebesgue measure 1, (that is $c(x, y) \geq c > 0$ a.s.). Then the stationary Markov chain generated by the copula is $\phi-$mixing. This is equivalent to saying there are constants D and $0 < r < 1$ such that for every $n \geq 1$, and $B \in \mathcal{R} \cap I$*

$$|P(U_n \in B|U_0 = x) - \lambda(B)| \leq Dr^{-n} \ \text{a.s.}$$

Remark 11 *This result also implies that the sequence is geometrically ergodic since $\beta_n \leq \phi_n \leq Dr^{-n}$.*

Example Marshall-Olkin copula is given by formula

$$C_{\alpha,\beta}(u, v) = \min{(uv^{1-\alpha}, vu^{1-\beta})}, \quad 0 \leq \alpha, \beta \leq 1 \ ,$$

is geometric $\phi-$mixing for $0 \leq \alpha, \beta < 1$.

5 Proofs

Proof of Proposition 4

Because for almost all x we know that $C_{,1}(x, y)$ exists and is increasing in y, we have that $C_{,12}(x, y)$ exists a.s. It follows that for all y there is a set Ω_y with $\lambda(\Omega_y) = 1$ such that for all $x \in \Omega_y$

$$\mathbb{P}(U_1 \leq y|U_0 = x) = C_{,1}(x, y) = \int_0^y C_{,12}(x, v)dv + S_1(x, y) \ ,$$

where $C_{,12}(x, v)$ is the density of the absolute continuous part of the copula and $S_1(x, y) = C_{,1}(x, y) - \int_0^y C_{,12}(x, v)dv$ is the singular part of $C_{,1}(x, y)$. Since, by Lebesgue Theorem, $\int_0^y C_{,12}(x, v)dv \leq C_{,1}(x, y) - C_{,1}(x, 0)$, we have $S_1(x, y) \geq 0$. In the same way we argued the relation (2), we find a set Ω of measure 1 such that for all $x \in \Omega$ and all Borelians A,

$$\mathbb{P}(U_1 \in A|U_0 = x) = C_{,1}(x, A) = \int_A C_{,12}(x, v)dv + S_1(x, A) \geq \int_A C_{,12}(x, v)dv > 0 \ ,$$

(5)

and irreducibility follows.

To prove aperiodicity, by Theorem 3.2 in Darsow et al. [10], we know that

$$C^2(x, y) = \mathbb{P}(U_0 \leq x, U_2 \leq y) = \int_I C_{,2}(x, t)C_{,1}(t, y)dt \ .$$

9

By Fatou lemma we obtain,

$$C^2_{,12}(x,y) \geq \int_I C_{,21}(x,t)C_{,12}(t,y)dt \ .$$

Then, by Proposition 3.5 in Šremr [19], (see also Lemma 1 of Walczak [20]), we have $C_{,21}(x,y) = C_{,12}(x,y)$ a.s. and by our assumption they are strictly positive a.s. Furthermore, by Fubini Theorem, for almost all x, $\lambda\{(t : C_{,21}(x,t) > 0)'\} = 0$. Then we easily find a set of Lebesque measure 1 such that, on that set, we have $C^2_{,12}(x,y) > 0$. By repeating the arguments above we find a set Ω' of measure 1 such that for all $x \in \Omega'$ and all Borelians A

$$\mathbb{P}(U_2 \in A|U_0 = x) > 0 \ , \tag{6}$$

The aperiodicity follows from (5) and (6), by taking $A = \Omega \cap \Omega'$. $\diamond$

Proof of Lemma 5

For simplicity, we shall argue the conclusion of the lemma only for two copulas. Define $C(x,y) = aC_1(x,y) + (1-a)C_2(x,y)$, with $0 < a < 1$. Their n-steps transition kernels are $\frac{\partial C_1^n}{\partial x}(x,A)$ a.s. and $\frac{\partial C_2^n}{\partial x}(x,A)$ a.s., as mentioned in relation (2). The n-steps transition kernel of the Markov chain generated by $C(x,y)$ is

$$P^n(x,A) = \frac{\partial}{\partial x}C^n(x,A) = \frac{\partial}{\partial x}(aC_1 + (1-a)C_2)^n(x,A) \ ,$$

for $x \in B$ with $\lambda(B) = 1$ and all $A \in \mathcal{R} \cap I$. Due to distributivity and associativity of the fold product from Definition 2, we easily obtain

$$P^n(x,A) \geq a^n \frac{\partial C_1^n}{\partial x}(x,A) + (1-a)^n \frac{\partial C_2^n}{\partial x}(x,A) \geq a^n \frac{\partial C_1^n}{\partial x}(x,A).$$

for all $n \geq 1$, $x \in B$ with $\lambda(B) = 1$ and all $A \in \mathcal{R} \cap I$. Therefore the conclusion of this lemma follows by the definitions of irreducibility and aperiodicity given at the beginning of Section 3. $\diamond$

Proof of Theorem 7

The convex combination generates an absolutely regular Markov chain by Lemma 5. Because this combination is still a symmetric copula, it generates a stationary and reversible Markov chain. By Theorem 6, in order to proof that it is geometrically ergodic, we have to show that its first $\rho-$mixing coefficient is strictly less than 1. We shall argue that this holds and for simplicity we shall consider the case $n = 2$. Denote by ρ'_1, ρ''_1 and ρ_1 the corresponding first $\rho-$mixing coefficients for the stationary Markov chains generated by $C_1(x,y)$, $C_2(x,y)$ and by $C(x,y) = aC_1(x,y) + (1-a)C_2(x,y)$ with $0 \leq a \leq 1$,

respectively. According to Theorem 6, we have $\rho_1' < 1$ and $\rho_1'' < 1$. Then, by definition (4) we easily derive that

$$\rho_1 \le a\rho_1' + (1-a)\rho_1'' < 1$$

and the result follows. $\Diamond$

Proof of Theorem 10

The proof is based on Doeblin theory. We mention first that Doeblin's condition, in the basic form (see Bradley, vol. 2 page 330 [3]), is implied by

Condition 12 *There exists $A \subset I$ with $\lambda(A) = 1$ and $\varepsilon \in (0,1)$ such that for all x in A and all $B \in \mathcal{R} \cap I$, the relation $\lambda(B) \le \varepsilon$ implies $C_{,1}(x, B) \le 1 - \varepsilon$.*

This condition implies that $\varphi_1 < 1 - \varepsilon$. Here is a short argument in terms of copula. Since $C_{,1}(x, B) - \lambda(B) = \lambda(B') - C_{,1}(x, B')$, we notice we do not need the absolute value in the definition of ϕ_1. By Condition 12,

$$\sup_B |C_{,1}(x, B) - \lambda(B)| = \sup_B (C_{,1}(x, B) - \lambda(B)) \le$$
$$\max\{ \sup_{B, \lambda(B) \le \varepsilon} (C_{,1}(x, B) - \lambda(B)), \sup_{B, \lambda(B) > \varepsilon} C_{,1}(x, B) - \lambda(B)) \}$$
$$\le \max(\sup_{B, \lambda(B) \le \varepsilon} C_{,1}(x, B), \sup_{B, \lambda(B') \le 1-\varepsilon} \lambda(B')) \le 1 - \varepsilon \quad \text{a.s.}$$

This gives
$$\varphi_1 = ess \sup_x \sup_B |C_{,1}(x, B) - \lambda(B)| \le 1 - \varepsilon \ .$$

On the other hand, by Proposition 4, we already know that the process is absolutely regular and thus is ergodic and aperiodic. Then, according to Doeblin theorem (see Comment 6 in Bradley, vol. 2, page 331 [3]) we have only to verify Condition 12.

Let $\varepsilon = c/(1+c)$. Let $A \in \mathcal{R} \cap I$ with $\lambda(A) \le \varepsilon$ or equivalently $\lambda(A') > 1 - \varepsilon$. Then, by the definition of ε, for all x in a set of measure 1,

$$1 - C_{,1}(x, A) = C_{,1}(x, A') \ge \int_{A'} c(x, y) dy \ge c\lambda(A') \ge c(1 - \varepsilon) = c/(1 + c) = \varepsilon \ .$$

So, for almost all x
$$C_{,1}(x, A) \le 1 - \varepsilon \ .$$

The conclusion of Doeblin's theorem is that the Markov chain is $\phi-$mixing. (see Bradley, vol. 2 page 331, Comments 4 and 5 and 6 [3]). $\Diamond$

References

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

[2] B. K. Beare, Archimedean copulas and temporal dependence, University of California at San Diego, Economics Working Paper Series 1549539, (2010).

[3] R.C. Bradley, *Introduction to strong mixing conditions.* Vol 1, 2, 3. Kendrick Press, Heber City, (2007).

[4] L. Breiman, J.H. Friedman, Estimating optimal transformations for multiple regression and correlation, *J. Amer. Statist. Assoc.*, **80**, (1985), 580-598.

[5] W. Bryc, Conditional moment representations for dependent random variables, *Electron. J. Probab.* ,**1**, (1996), 1–14.

[6] K. Chan, H. Tong, *Chaos: A Statistical Perspective.* Springer-Verlag, New York, (2001).

[7] X. Chen, Y. Fan, Estimation of copula-based semiparametric time series models, *J. Econometrics*, **130**, (2006), 307–335.

[8] X. Chen, W.B. Wu, Y. Yi, Efficient estimation of copula-based semiparametric Markov models, *Ann. Statist.*, **37**, (2009), 4214-4253.

[9] J.B. Conway, *A course in functional analysis*, 2nd ed., Springer-Verlag, New York (1990).

[10] W.F. Darsow, B. Nguyen, E.T. Olsen, Copulas and Markov processes, *Illinois J. Math.*, **36**, (1992), 600-642.

[11] C.G. de Vries, C. Zhou, Discussion of "Copulas: Tales and facts", by Thomas Mikosch, *Extremes*, **9**, (2006), 23-25.

[12] P. Gagliardini, C. Gouriéroux, Duration time-series models with proportional hazard, *Journal of Time Series Analysis*, **29**, (2008), 74-124.

[13] R. Ibragimov, G. Lentzas, Copulas and long memory, Harvard Institute of Economic Research Discussion Paper No. 2160, (2009).

[14] I. Kontoyannis, S.P. Meyn, Geometric ergodicity and spectral gap of non-reversible real valued Markov chains, *Probab. Theory and related Fields* (2011)(papers to appear in subsequent numbers).

[15] S.P. Meyn, R.L. Tweedie, *Markov Chains and Stochastic Stability,* 2nd ed., Cambridge University Press, London, (2009).

[16] R.B. Nelsen, *An introduction to copulas.* 2nd ed., Springer-Verlag, New York, (2006).

[17] G.O. Roberts, R.L. Tweedie, Geometric L_2 and L_1 convergence are equivalent for reversible Markov chains, *J. Appl.Probab.* **38**A, (2001), 37–41.

[18] G.O. Roberts, J.S. Rosenthal, Geometric ergodicity and hybrid Markov chains, *Electron. Commun. Probab.* **2**, (1997), 13-25.

[19] J. Šremr, Absolutely continuous functions of two variables in the sense of Carathéodory, *Electron. J. Differential Equations,* **154**, (2010),1-11.

[20] S. Walczak, On the differentiability of absolutely continuous functions of several variables, remarks on the Rademacher theorem, *Bull. Polish Acad. Sci. Math.,* **36**, (1988), 513–520.