On tail dependence: a characterization for first-order max-autoregressive processes

Marta Ferreira*

Department of Mathematics,

University of Minho,

Campus de Gualtar,

4710-057 Braga, Portugal

Abstract

In this paper we consider first-order MARMA or $ARMAX$ processes and a modified version of these involving a power transformation, denoted $pARMAX$. We assume Pareto-type tails, the most interesting case for inference within these processes. Some well-known dependence measures of multivariate extreme value theory will be considered in a time series framework, and strong consistency and asymptotic normality of nonparametric estimators will be stated. In calculating these measures, we shall find that $ARMAX$ and $pARMAX$ have opposite behavior in concomitant extremes, covering all types of tail dependence. This characterization will serve modeling purposes.

Keywords: Extreme value theory, Markov chains, max-autoregressive processes, tail dependence.

2000 Mathematics Subject Classification: 60G70, 60J20.

*Electronic address: msferreira@math.uminho.pt

I. INTRODUCTION

Extreme Value Theory (EVT) is an important tool for many applied sciences when faced with modeling high values of certain phenomena. Ocean wave modeling, wind engineering, thermodynamics of earthquakes, risk assessment on financial markets are some examples. The first results were developed considering independent observations but, more recently, models for extreme values have been constructed under the more realistic assumption of temporal dependence. Among these models, stationary Markov chains are very interesting, specially because they may have a somewhat simple treatment in what concerns extremal properties. Linear autoregressive series like ARMA are perhaps the most known and used for applications, even for extreme value modeling. In EVT domains of applications in which heavy tailed distributions are prevalent, taking the maxima or the sum of two heavy tailed random variables (r.v.'s) is basically equivalent for the tail behavior (see [20], Chap. 2). The max-autoregressive moving average processes MARMA [21] are derived from ARMA by replacing summation by the maximum operator. The particular case MAR(1) or *ARMAX*, given by,

$$X_i = c\,X_{i-1} \vee Z_i,$$

with $0 < c < 1$ and $\{Z_i\}$ i.i.d. has been widely studied in literature [3, 11, 12, 18]. Some generalizations of this process have been applied to geophysical and reliability phenomena [4, 8, 21] as well as financial series [24]. Since MARMA finite-dimensional distributions can easily be written explicitly, they are more convenient for analysis.

More recently, some careful attention has been given to the statistical modeling of the tail dependence between consecutive pairs from a stationary first-order Markov chain, since it is important to distinguish asymptotic dependence from asymptotic independence. According to Bortot and Tawn (see [19]), a Markov chain $\{Y_i\}$ is said to be asymptotically tail dependent or independent, whenever $\lambda > 0$ or $\lambda = 0$, respectively, in the limit below:

$$\lim_{y \to y^*} P(Y_2 > y | Y_1 > y) = \lambda\,, \tag{1}$$

where y^* is the right-endpoint of Y_1, i.e., $y^* = \sup\{y : P(Y_1 \leq y) < 1\}$. Observe that coefficient λ above is a formulation of the so-called *tail dependent coefficient* (TDC), introduced in [17], in the context of Markov chains (see Definition 1 below). Within the class of asymptotically dependent variables ($\lambda > 0$), λ increases with increasing degree of dependence at extreme values. For asymptotically tail independent Markov chains ($\lambda = 0$), the dependence between exceedances of y gradually decreases as $y \to y^*$, which leads to an extremal feature increasingly resembling an i.i.d. sequence at high levels. As pointed out in [19], this phenomenon has been noticed in a number of data and theoretical applications. In these cases, the classical procedures of EVT for estimating the probability of an extreme event are not applicable. This problem is overcome by setting the way how $P(Y_2 > y|Y_1 > y)$ converges to zero, as $y \to y^*$, a kind of penultimate dependence which involves the so-called Ledford & Tawn coefficient η [1, 2]. More precisely, Ledford and Tawn's approach is based on assuming that, for a random pair (X, Y), the mapping $t \mapsto P(X > F_X^{-1}(1 - xt), Y > F_Y^{-1}(1 - yt))$, $t \in [0, 1]$, is regularly varying at zero. If (X, Y) is tail dependent, the previous mapping is regularly varying of order 1, and consequently a homogeneity property holds for small t. In case of tail independence, the latter property does not hold. Here, an adjusted homogeneity property can be obtained by assuming regular variation with index $1/\eta$, where $\eta < 1$ is the above mentioned Ledford & Tawn coefficient. For more details see [1, 2, 5]. This is a non-trivial class including many commonly studied processes such as Gaussian Markov chains [17]. Coefficient η is a measure for the strength of dependence within an asymptotic tail independent behavior. More precisely, $\eta = 1$ corresponds to tail dependence whilst $\eta < 1$ means asymptotic tail independence, with $\eta = 1/2$ occurring for the (almost) independent case. Increasing values of η correspond to stronger association. The $ARMAX$ process, which has unit η, is in the group of tail dependent Markov chains [13] and hence is not suitable to model data series expressing the described phenomenon.

In [13] was considered a modified version of MARMA in order to derive an asymptotic tail independent model, the so-called first-order *power max-autoregressive* (in short, $pARMAX$), defined as,

$$X_i = X_{i-1}^c \vee Z_i \ , \ 0 < c < 1, \ i \in \mathbb{Z},$$

with $\{Z_i\}$ i.i.d., for which η is a function of the model parameter c. More precisely, we have $\eta = \max(1/2, c)$ and hence $pARMAX$ is an asymptotic tail independent process, even almost independent in cases $c \leq 1/2$. Hence, it is a suitable model to describe the above mention phenomenon of time series exhibiting asymptotic tail independence. In Figure 1 one can see the similarity between the sample paths of heavy tailed $pARMAX$ and AR(1) processes. The $pARMAX$ process has easily derived extremal properties and also easily explicited finite-dimensional distributions [13, 14]. A generalization of $pARMAX$ has also

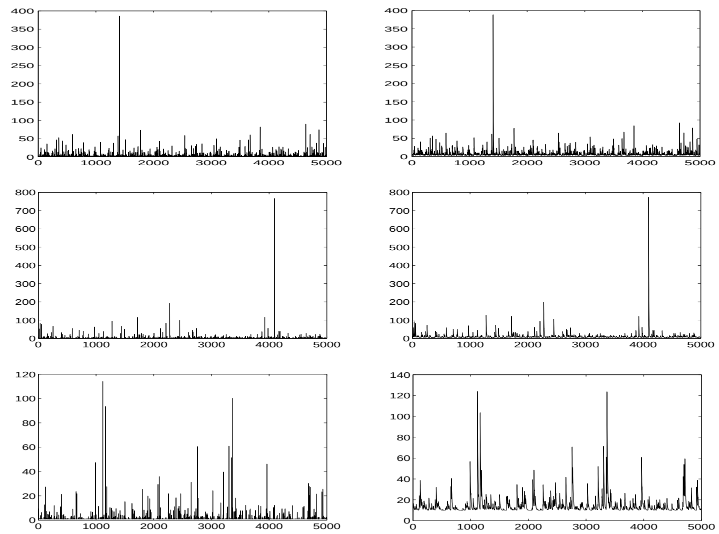

Figure 1: 5000 realizations of $X_i = X_{i-1}^c \ \vee \ Z_i$ (left) and of $X_i = c\ X_{i-1} + Z_i$ (right), for Pareto marginals and (top-to-bottom) $c = 0.7, 0.8, 0.9$, respectively.

been applied in modeling financial data [15]. We remark also the well-known fact that Gumbel-tailed random variables are very common in atmospheric sciences, e.g. temperature maxima and that a Gumbel tailed $ARMAX$ process can be derived by taking the logarithm of a heavy-tailed $pARMAX$.

Another interesting measure for dependence in multivariate extreme values is a generalization of the above mentioned tail dependence coefficient, TDC, to the probability of simultaneous extremes, known as *extremal dependence coefficient (EDC)* [6]. This coefficient accounts for the concomitance of extremal events and can be a useful tool to describe the dependence between extremal data (see, for instance, the studies of [23] and [16]). Both, TDC and EDC, are given in Section II.

In this paper we define all these tail dependence measures in a time series context which allow to assess temporal dependence on extreme values, in a similar manner as the auto-correlation function (ACF) for ARMA models (Section II). As already mentioned above, it is important to distinguish between asymptotic dependence and independence in the tail for inferential purposes. We also state nonparametric estimators and derive consistency and asymptotic normality (Section III). Due to the simplicity in obtaining the finite-dimensional distributions of first-order max-autoregressive series, we derive explicit forms for those measures (Section IV). We will see that $pARMAX$ and $ARMAX$ have opposite performances in the upper tail and cover all the possible tail dependence/independence cases. This characterization can be used for the identification of *ARMAX* and *pARMAX* in max-autoregressive modeling. Finally we will show that the properties of the presented non parametric estimators are valid within these processes. The proofs of this section are in Appendix.

II. TAIL DEPENDENCE MEASURES

Dependencies between (extreme) financial asset-returns have significantly increased during recent time periods in almost all international markets.

Several empirical surveys such as [16, 23] exhibited that the concept of tail dependence is a useful tool to describe the dependence between extremal data, especially during volatile and bear markets.

The so-called *tail dependence coefficient* (TDC) has become a popular measure in risk management. In the following we will consider dependence in the upper and lower tail.

Definition 1 *The lower tail dependence coefficient (TDC) of two random variables X and Y with marginal d.f.* F_X *and* F_Y*, respectively, is defined as*

$$\lambda^L = \lim_{t\downarrow 0} P(F_Y(Y) \leq t | F_X(X) \leq t) = \lim_{t\downarrow 0} \frac{C(t,t)}{t} \tag{2}$$

where $C(t,t) = P(F_X(X) \leq t, F_Y(Y) \leq t)$ *is the so-called copula function, provided the limit exists. If* $\lambda^L > 0$ *then X and Y are said to be lower tail-dependent.*

Similarly, the upper TDC of X and Y is defined as

$$\lambda^U = \lim_{t\uparrow 1} P(F_Y(Y) > t | F_X(X) > t) \tag{3}$$

provided the limit exists. If $\lambda^U > 0$ *then X and Y are said to be upper tail-dependent.*

Within a tail independence setting (i.e., $\lambda = 0$) we can still graduate a kind of "strength" of the dependence by Ledford and Tawn coefficient η. It has been originally considered for the upper tail but it can be stated for the lower tail as well. We use formulation as in [5], by considering regular variation condition for the bivariate lower tail,

$$\frac{P(X < F_X^{-1}(tx), Y < F_Y^{-1}(ty))}{P(X < F_X^{-1}(t), Y < F_Y^{-1}(t))} \underset{t\downarrow 0}{\longrightarrow} h^L(x,y), \tag{4}$$

where the convergence holds uniformly on $\{(x,y)|\max(x,y) = 1\}$ for some non-degenerate function h^L, homogeneous of order $1/\eta^U$, i.e., $h^L(tx,ty) = t^{1/\eta^L} h^L(x,y)$. It is assumed that the function $t \mapsto P(X > F_X^{-1}(1-t), Y > F_Y^{-1}(1-t))$ is regularly varying at 0 with index $1/\eta^L$ and the limit $\lim_{t\downarrow 0} P(X > F_X^{-1}(1-t), Y > F_Y^{-1}(1-t)) = \lim_{t\downarrow 0} t^{1/\eta^L - 1} L^L(t)$ exists for some slowly varying function L^L. The marginals X and Y are asymptotically dependent if $\eta^L = 1$ and $L^L(t) \not\to 0$ as $t \downarrow 0$ and asymptotically independent otherwise. The cases, $1/\eta^L < 1/2$ and $1/\eta^L > 1/2$, correspond to, respectively, negative and positive association, and if $1/\eta^L = 1/2$ we have (almost) independence.

Analogously, η^U is stated by considering regular variation of the upper bivariate tail,

$$\frac{P(X > F_X^{-1}(1-tx), Y > F_Y^{-1}(1-ty))}{P(X > F_X^{-1}(1-t), Y > F_Y^{-1}(1-t))} \underset{t\downarrow 0}{\longrightarrow} h^U(x,y), \tag{5}$$

with function h^U homogeneous of order $1/\eta^U$.

The lower/upper extremal dependence coefficient (EDC) given below corresponds to a generalization of TDC to the multivariate case [6].

Definition 2 *Let X be a d-dimensional random vector with d.f. F and margins $F_1, \ldots,$ F_d. Furthermore, let $F_{min} := \min(F_1(X_1), ..., F_d(X_d))$ and $F_{max} := \max(F_1(X_1), ..., F_d(X_d))$. The lower extremal dependence coefficient (EDC) of X is defined as*

$$\epsilon^L := \lim_{t\downarrow 0} P(F_{max} \leq t | F_{min} \leq t) = \lim_{t\downarrow 0} \tfrac{P(F_{max}\leq t)}{1-P(F_{min}>t)}, \tag{6}$$

provided the limit exists. If $\epsilon^L > 0$ then the components of X are said to be lower extremal dependent. The upper EDC of X is defined as

$$\epsilon^U := \lim_{t\uparrow 1} P(F_{min} > t | F_{max} > t) = \lim_{t\uparrow 1} \tfrac{P(F_{min}>t)}{1-P(F_{max}\leq t)}, \tag{7}$$

provided the limit exists. If $\epsilon^U > 0$ then the components of X are said to be upper extremal dependent.

The lower extremal dependence coefficient can be interpreted as the probability that the best performer of X is attracted by the worst one provided this one has an extremely bad performance. This interpretation holds vice versa regarding the upper extremal dependence coefficient.

Here we consider these dependence measures in a time series framework, having in mind the assessment of the tail dependence in time, either for consecutive random pairs or separated by a lag m. As already mentioned (see Introduction), it is important to distinguish an asymptotic dependent tail behavior of an asymptotic independent one under penalty of a misspecification. Given a time series $\{X_i\}$, we define the m-lower tail dependence coefficient λ_m^L (respectively, the m-upper tail dependence coefficient λ_m^U) by considering (2) (respectively, (3)) for random pairs (X_1, X_{1+m}). Analogously, we state the m-lower Ledford and Tawn coefficient η_m^L (respectively, the m-upper Ledford and Tawn coefficient η_m^U) by considering (4) (respectively, (5)) also for random pairs (X_1, X_{1+m}).

Similarly, we adapt the EDC coefficient in Definition 2 to a time series $\{X_i\}$. We will say that $\{X_i\}$ is *d-lower (upper) extremal dependent* if (6) (resp. (7)) holds. The d-lower (upper) extremal dependence coefficient corresponds to the probability that the best (worst) performer of X is attracted by the worst (best) one provided this one has a very bad (good) performance, for d consecutive time instants. This can be used, for instance, to assess risk in a single stock market index.

III. NON PARAMETRIC ESTIMATORS

The tail dependent coefficient as well as the extremal dependent coefficient can be defined in terms of the so-called copula function, $C(u_1, ..., u_d) = P(F_1(X_1) \leq u_1, ..., F_d(X_d) \leq u_d)$, of random vector $(X_1, ..., X_d)$, where $F_1, ..., F_d$ are the continuous marginal d.f.'s and $(u_1, ..., u_d) \in [0,1]^d$. The survival copula is given by $\overline{C}(u_1, ..., u_d) = P(F_1(X_1) > u_1, ..., F_d(X_d) > u_d)$.

In [22] it is considered nonparametric estimators of tail copula functions. For simplicity we consider the bivariate case but the following statements also hold for higher dimensions. The bivariate lower and upper tail copula functions are defined as, respectively,

$$\Lambda^L(x,y) = \lim_{t\downarrow 0} \tfrac{1}{t}\, C(tx, ty), \quad \Lambda^U(x,y) = \lim_{t\downarrow 0} \tfrac{1}{t}\, \overline{C}(1-tx, 1-ty).$$

For n i.i.d. replicates of random pair (X_1, X_2), $(X_1^{(1)}, X_2^{(1)}), (X_1^{(2)}, X_2^{(2)}), ..., (X_1^{(n)}, X_2^{(n)})$, the corresponding estimators,

$$\widehat{\Lambda}^L(x,y) = \tfrac{1}{k}\sum_{j=1}^{n} \mathbf{1}_{\{F_{1,n}(X_1^{(j)}) \leq x\frac{k}{n}, F_{2,n}(X_2^{(j)}) \leq y\frac{k}{n}\}}, \; \widehat{\Lambda}^U(x,y) = \tfrac{1}{k}\sum_{j=1}^{n} \mathbf{1}_{\{F_{1,n}(X_1^{(j)}) \geq 1-x\frac{k}{n}, F_{2,n}(X_2^{(j)}) \geq 1-y\frac{k}{n}\}},$$

where $k \equiv k(n) \to \infty$, $k/n \to 0$ as $n \to \infty$, $F_{i,n}$ is the empirical d.f. of X_i $(i = 1, 2)$, are asymptotic normal and strong consistent (Theorem 6 and 7 in [22]). Observe that

$\lambda^L := \Lambda^L(1,1)$ and $\lambda^U := \Lambda^U(1,1)$.

We shall prove that this approach still holds in a dependent framework, more precisely, considering consecutive random pairs, $(X_1, X_{1+m}), (X_2, X_{2+m}), \dots$, for a given stationary sequence $\{X_i\}$. Therefore, we consider the lag-m lower and upper tail copula functions, respectively,

$$\Lambda_m^L(x,y) = \lim_{t\downarrow 0} \tfrac{1}{t} C_m(tx,ty), \quad \Lambda_m^U(x,y) = \lim_{t\downarrow 0} \tfrac{1}{t} \overline{C}_m(1-tx,1-ty), \tag{8}$$

where,

$$C_m(u,u) = P(F(X_1) < u, F(X_{1+m}) < u) \text{ and } \overline{C}_m(u,u) = P(F(X_1) > u, F(X_{1+m}) > u), \tag{9}$$

with corresponding estimators,

$$\widehat{\Lambda}_m^L(x,y) \sim \tfrac{n}{k}\widehat{C}_m\big(x\tfrac{k}{n}, y\tfrac{k}{n}\big), \ \widehat{\Lambda}_m^U(x,y) \sim \tfrac{n}{k}\widehat{\overline{C}}_m\big(1-x\tfrac{k}{n}, 1-y\tfrac{k}{n}\big), \tag{10}$$

where,

$$\widehat{C}_m(u,v) = \tfrac{1}{n-m}\sum_{j=1}^{n-m} \mathbf{1}_{\{F_n(X_j)\le u, F_n(X_{j+m})\le v\}}, \ \widehat{\overline{C}}_m(u,v) = \tfrac{1}{k}\sum_{j=1}^{n-m} \mathbf{1}_{\{F_n(X_j)\ge 1-u, F_n(X_{j+m})\ge 1-v\}}.$$

Analogously, $\lambda_m^L := \Lambda_m^L(1,1)$ and $\lambda_m^U := \Lambda_m^U(1,1)$. In this case we must drop independence which does not affect strong consistency (see Theorem 7 in [22]). The following results are stated for the upper tail copula, but they are similar to the lower one.

Next result states asymptotic normality of estimators in (10), for β-mixing sequences under suitable conditions on β-coefficients and satisfying some regularity conditions on the joint tail, similar to [7], in order to ensure that a limiting covariance function exists. Recall that $\{X_i\}$ is β-mixing (or absolutely regular) if

$$\beta(l) := \sup_{p\in\mathbb{N}} E\Big(\sup_{A\in\mathcal{B}_{p+l+1}^{\infty}} |P(A|\mathcal{B}_1^p) - P(A)|\Big) \to 0 \tag{11}$$

as $l \to \infty$, where $\mathcal{B}_1^p$ and $\mathcal{B}_{p+l+1}^{\infty}$ denote the σ-fields generated by $(X_i)_{1\le i\le p}$ and $(X_i)_{i\ge p+l+1}$, respectively.

Theorem 1 *Let $\{X_i\}$ be a stationary sequence, with continuous marginal d.f. F, satisfying β-mixing condition (11) such that,* $\lim_{n\to\infty} \frac{\beta(l_n)}{l_n} n + l_n k_n^{-1/2} \log^2(k_n) = 0$, *where* $l_n \to \infty$,

$l_n = o(n/k_n)$, $k_n \to \infty$ and $k_n/n \to 0$, as $n \to \infty$. Consider that, there exist $\epsilon > 0$ and functions c_p, $p \in \mathbb{N}$, such that

$$\begin{aligned}\lim_{n\to\infty} \frac{n}{k_n} P\Big(X_1 > F^{-1}\Big(1 - \frac{k_n}{n}x_r\Big), X_{1+m} > F^{-1}\Big(1 - \frac{k_n}{n}y_r\Big),\\ X_{1+p} > F^{-1}\Big(1 - \frac{k_n}{n}x_s\Big), X_{1+p+m} > F^{-1}\Big(1 - \frac{k_n}{n}y_s\Big)\Big) \to c_p(x_r, y_r, x_s, y_s),\end{aligned} \tag{12}$$

$\forall p \in \mathbb{N}$, some fixed positive integer m, $0 < x_r, x_s, y_r, y_s \leq 1 + \epsilon$, and

$$\frac{n}{k_n} P\big(X_1 \in I_n(x,y), X_{1+p} \in I_n(x,y)\big) \leq (y - x)\Big(\tilde{\rho}(p) + D_1 \frac{k_n}{n}\Big), \tag{13}$$

$\forall p \in \mathbb{N}$, where $I_n(x,y) = \big]F^{-1}(1 - yk_n/n), F^{-1}(1 - xk_n/n)\big]$, $0 < x, y \leq 1 + \epsilon$, constant $D_1 \geq 0$, with sequence $\tilde{\rho}(p)$, $p \in \mathbb{N}$, satisfying $\sum_{p=1}^{\infty} \tilde{\rho}(p) < \infty$. If the tail copula $\Lambda_m^U \not\equiv 0$ exist, possesses continuous partial derivatives and the second order condition

$$\lim_{w\to\infty} \frac{\Lambda_m^U(x,y) - w\overline{C}_m(1 - x/w, 1 - y/w)}{A(w)} = g(x,y) < \infty \tag{14}$$

holds locally uniformly for $(x,y) \in [0,\infty]^2 \backslash \{(\infty,\infty)\}$ and some non constant function g, where function $A : \mathbb{R}_+ \to \mathbb{R}_+$ is such that $A(w) \to 0$ as $w \to \infty$, then for $\sqrt{k}A(n/k) \to 0$ as $n \to \infty$,

$$\sqrt{k}(\widehat{\Lambda}_m^U(x,y) - \Lambda_m^U(x,y)) \to \mathbb{G}(x,y)$$

weakly, where $\mathbb{G}(x,y)$ is a centered Gaussian random field, with covariance function $\Lambda_m^U(\min(x_r,x_s), \min(y_r,y_s)) + \sum_{p=1}^{\infty}(c_p(x_r,y_r,x_s,y_s) + c_p(x_s,y_s,x_r,y_r))$.

Proof. The proof is straightforward from Theorem 6 in [22]. We only need to look at the covariance structure, since now we have dependent random pairs $(X_1, X_{1+m}), (X_2, X_{2+m}), \ldots$. This is done along the same steps of Proposition 2.1 in [7]. Observe that, by (13),

$$\begin{aligned}&\lim_{n\to\infty} \frac{n}{k_n} P\Big(X_1 > F^{-1}\Big(1 - \frac{k_n}{n}(1+\epsilon)\Big), X_{1+m} > F^{-1}\Big(1 - \frac{k_n}{n}(1+\epsilon)\Big),\\ &\qquad X_{1+p} > F^{-1}\Big(1 - \frac{k_n}{n}(1+\epsilon)\Big), X_{1+p+m} > F^{-1}\Big(1 - \frac{k_n}{n}(1+\epsilon)\Big)\Big) \\ &\leq (1+\epsilon)\big(\widetilde{\rho}(p) + D_1 \tfrac{k_n}{n}\big),\end{aligned} \tag{15}$$

and, as $l_n k_n/n \to 0$, $\lim_{n\to\infty} \sum_{p=1}^{l_n} (\widetilde{\rho}(p) + D_1 k_n/n) = \sum_{p=1}^{\infty} \widetilde{\rho}(p) < \infty$. By (12) and applying

Pratt's lemma in [10],

$$\lim_{n\to\infty}\frac{n}{k_n}\sum_{p=1}^{l_n}cov\Big(\mathbf{1}_{\left\{X_1>F^{-1}(1-\frac{k_n}{n}x_r),X_{1+m}>F^{-1}(1-\frac{k_n}{n}y_r)\right\}},$$

$$\mathbf{1}_{\left\{X_{1+p}>F^{-1}(1-\frac{k_n}{n}x_s),X_{1+p+m}>F^{-1}(1-\frac{k_n}{n}y_s)\right\}}\Big) \tag{16}$$

$$=\sum_{p=1}^{\infty}c_p(x_r,y_r,x_s,y_s).$$

Therefore, we have that,

$$\frac{n}{l_nk_n}\sum_{1\leq i<j\leq l_n}cov\Big(\mathbf{1}_{\left\{X_i>F^{-1}(1-\frac{k_n}{n}x_r),X_{i+m}>F^{-1}(1-\frac{k_n}{n}y_r)\right\}},\mathbf{1}_{\left\{X_j>F^{-1}(1-\frac{k_n}{n}x_s),X_{j+m}>F^{-1}(1-\frac{k_n}{n}y_s)\right\}}\Big)$$

$$=\tfrac{1}{l_n}\sum_{i=1}^{l_n}\tfrac{n}{k_n}\sum_{j=i+1}^{i+l_n-1}cov\Big(\mathbf{1}_{\left\{X_i>F^{-1}(1-\frac{k_n}{n}x_r),X_{i+m}>F^{-1}(1-\frac{k_n}{n}y_r)\right\}},\mathbf{1}_{\left\{X_j>F^{-1}(1-\frac{k_n}{n}x_s),X_{j+m}>F^{-1}(1-\frac{k_n}{n}y_s)\right\}}\Big)$$

$$-\tfrac{1}{l_n}\sum_{i=2}^{l_n}\tfrac{n}{k_n}\sum_{j=l_n+1}^{i+l_n-1}cov\Big(\mathbf{1}_{\left\{X_i>F^{-1}(1-\frac{k_n}{n}x_r),X_{i+m}>F^{-1}(1-\frac{k_n}{n}y_r)\right\}},\mathbf{1}_{\left\{X_j>F^{-1}(1-\frac{k_n}{n}x_s),X_{j+m}>F^{-1}(1-\frac{k_n}{n}y_s)\right\}}\Big)$$

$$\to\sum_{p=1}^{\infty}c_p(x_r,y_r,x_s,y_s),\ \text{as } n\to\infty,$$

since, the second term, by (15) and taking $i-1=p$, can be bounded by $\sum_{p=1}^{l_n-1}p(\widetilde{\rho}(p)+D_1k_n/n)(1+\epsilon)/l_n$, which converges to 0, as $n\to\infty$. $\square$

Corollary 1 *Under the prerequisites of Theorem 1,*

$$\sqrt{k}(\widehat{\lambda}_m^U-\lambda_m^U)\to N(0,\sigma_U^2)$$

weakly, where

$$\sigma_U^2=a+\big(\tfrac{\partial}{\partial x}\Lambda_m^U(1,1)\big)^2+\big(\tfrac{\partial}{\partial y}\Lambda_m^U(1,1)\big)^2+2a\Big(\big(\tfrac{\partial}{\partial x}\Lambda_m^U(1,1)-1\big)\big(\tfrac{\partial}{\partial y}\Lambda_m^U(1,1)-1\big)-1\Big), \tag{17}$$

with $a=\lambda_m^U+\sum_{p=1}^{\infty}2c_p(1,1,1,1)$.

Remark 1 *If the common marginal d.f. F is known, then Corollary 1 holds with $\sigma_U^2=a$. For details see [22] (Theorem 5, Corollary 1 and 2).*

We will see that both processes considered in this paper, $ARMAX$ and $pARMAX$, satisfy the above conditions, but other processes well know in the literature as linear ARMA and ARCH(1) also satisfy (c.f. [7]).

Theorem 1 is similar for lower and for higher dimensional empirical copulas (Theorem 10 in [22]). Strong consistency also holds for upper and lower multidimensional copulas (Theorem 11 in [22]). Hence these results can be applied to the lower and upper EDC,

$$\epsilon_L = \lim_{t\downarrow 0} \frac{C(t,\dots,t)}{1-\overline{C}(t,\dots,t)} \quad \text{and} \quad \epsilon_d^U = \lim_{t\uparrow 1} \frac{\overline{C}(t,\dots,t)}{1-C(t,\dots,t)} \,. \tag{18}$$

With respect to lag-m Ledford & Tawn coefficient, based on, respectively, (4) and (5), the upper η_m^U can be estimated as the tail index of

$$T_i = \min\left(\frac{1}{1-F(X_i)}, \frac{1}{1-F(X_{i+m})}\right), \tag{19}$$

and the lower, η_m^L, as the tail index of $\min\left(\frac{1}{F(X_i)}, \frac{1}{F(X_{i+m})}\right)$.

In the sequel we shall treat one of the coefficients since it is similar to the other one. We consider the upper coefficient η_m^U.

Observe that sequence $\{T_i\}$ in (19) is dependent and properties of the usual tail index estimators, e.g. Hill, maximum likelihood, Pickands, moments and probability weighted moments estimators, are stated for independent ones. Yet it can be proved consistency and asymptotical normality of the latters, using Drees approach in [7]. In this work, Drees has shown that the tail empirical quantile function (q.f.) of a β-mixing stationary sequence, under some convenient conditions on the joint tail similar to (12) and (13), converges weakly to a centered Gaussian process, and from this, he states the consistency and asymptotic normality of the above mentioned estimators, which can be represented as a functional of the tail empirical q.f.. In the sequel we shall refer to these as, the Drees class of tail index estimators.

If the marginal d.f. is unknown, the result still holds for sequence

$$T_i^{(n)} := \min\left(\frac{n+1}{n+1-nF_n(X_i)}, \frac{n+1}{n+1-nF_n(X_{i+m})}\right), \, i = 1, \dots, n, \tag{20}$$

as we shall see on the next theorem.

Theorem 2 *Consider sequences $\{T_i\}$ and $\left\{T_i^{(n)}\right\}_i$ defined in (19) and (20), respectively, where $\{X_i\}$ is stationary. If $\{T_i\}$ is β-mixing under the same prerequisites of Theorem 1,* *satisfies the regularity condition on the joint tail of [7], i.e., there exist* $\epsilon > 0$ *and functions* r_p, $p \in \mathbb{N}$, *such that*

$$\lim_{n\to\infty} \frac{n}{k_n} P\Big(T_1 > F_T^{-1}\Big(1 - \frac{k_n}{n}x\Big), T_{1+p} > F_T^{-1}\Big(1 - \frac{k_n}{n}y\Big)\Big) \to r_p(x,y), \tag{21}$$

$\forall p \in \mathbb{N}$, *and the uniform bound condition (13), for the extreme interval* $I_n(x,y) = \big]F_T^{-1}(1 - yk_n/n), F_T^{-1}(1 - xk_n/n)\big]$, *the Drees class of tail index estimators based on sequence* $\big\{T_i^{(n)}\big\}_i$ *and on sequence* $\{T_i\}$, *are consistent and asymptotically normal.*

Proof. In what concerns sequence $\{T_i\}$, the result follows from Theorem 2.2 in [7]. Only remains to show that

$$P\big(T_i^{(n)} > x, T_{i+p}^{(n)} > y\big) \xrightarrow{d} P(T_i > x, T_{i+p} > y). \tag{22}$$

Observe that,

$$P\big(T_i^{(n)} > x\big) \sim P\Big(X_i > F_n^{-1}\big(1 - \tfrac{1}{x}\big), X_{i+m} > F_n^{-1}\big(1 - \tfrac{1}{x}\big)\Big), \text{ as } n \to \infty. \tag{23}$$

Note also that, according to Theorem 1 in [9], the following convergence always hold, for any y in the support of X_i,

$$F_n(y) = \frac{1}{n}\sum_{j=1}^{n} 1_{\{X_j \le y\}} \xrightarrow{P} F(y)\,, \tag{24}$$

since $\sum_{j=1}^{n} 1_{\{X_j \le y\}}$ is a sum of dependent Bernoulli trials, denoted by $Binomial(n,p,\varrho)$ in Klotz (1973), with $p = P(X_j \le y) = F(y)$ and $\varrho = P(X_j \le y | X_{j-1} \le y)$. Based on the theoretical result, usually attributed to Slutsky, in which, any given random elements, Y_n, Y and Y_n', such that $Y_n \xrightarrow{d} Y$ and $Y_n' \xrightarrow{P} a$, then, $(Y_n, Y_n') \xrightarrow{d} (Y, a)$, we have that, for each i, and by the continuous mapping theorem,

$$\begin{aligned} &P\Big(X_i > F_n^{-1}\big(1 - \tfrac{1}{x}\big), X_{i+m} > F_n^{-1}\big(1 - \tfrac{1}{x}\big)\Big) \\ \underset{n\to\infty}{\longrightarrow}\ &P\Big(X_i > F^{-1}\big(1 - \tfrac{1}{x}\big), X_{i+m} > F^{-1}\big(1 - \tfrac{1}{x}\big)\Big) \end{aligned} \tag{25}$$

Therefore, given (23), we conclude that, $T_i^{(n)} \xrightarrow{d} T_i$. Similarly, (22) holds for all p. □

IV. *ARMAX* AND *PARMAX*: TAIL DEPENDENCE CHARACTERIZATION

Consider $\{Z_i\}$ a sequence of i.i.d. copies of a r.v., Z, having real positive support and marginal d.f. F_Z. A sequence $\{X_i\}$ is said to be a *pARMAX* process if,

$$X_i = X_{i-1}^c \vee Z_i \ , \ 0 < c < 1, \ \ i = 0, \pm 1, \pm 2, ... \tag{26}$$

and is said to be an *ARMAX* process if,

$$X_i = cX_{i-1} \vee Z_i \ , \ 0 < c < 1, \ \ i = 0, \pm 1, \pm 2, ... \tag{27}$$

with X_i independent of Z_j, for all integer $i < j$. For the sake of stationarity in the *pARMAX* case, the innovations $\{Z_i\}$ have support in $[1, \infty[$.

In the sequel, we will always consider that both have Pareto-type marginal d.f. F (i.e. heavy tail), with left-end-point x_* and right-end-point $x^* = +\infty$, such that,

$$1 - F(x) = x^{-1/\gamma} L_F(x), \tag{28}$$

where L_F is a slow varying function at $+\infty$ and γ (the tail index) is positive. Formulation (28) means also that $1 - F$ is a regularly varying function at infinity.

In the following, we derive the tail dependence measures, defined in Section II, for *ARMAX* and *pARMAX*. This will give us a characterization tool for first-order max-autoregressive models, in a similar manner as the *auto-correlation function* (ACF) for linear models ARMA. The proofs can be seen in Appendix.

Proposition 1 *Let m be a positive integer. The* pARMAX *process is lower tail dependent if the left-end-point $x_* = 1$, with $\lambda_m^L = c^m$, and lower tail independent otherwise. The* ARMAX *process is lower tail dependent if the left-end-point $x_* = 0$, with $\lambda_m^L = 1$, and lower tail independent otherwise.*

Proposition 2 *Let m be a positive integer. The* pARMAX *process is upper tail independent whereas the* ARMAX *is upper tail dependent with $\lambda_m^U = c^{m/\gamma}$, for all positive integer m.*

In what concerns the Ledford & Tawn coefficient η, according to Proposition 2, for both processes we will have $\eta^L < 1$, except if the left-end-point x_* is 1, respectively 0, which will lead to unit η^L. This particular measure do not distinguish *pARMAX* from *ARMAX* as expressed in the next result. In view of Proposition 2, *pARMAX* must have $\eta^U < 1$, whilst for *ARMAX* it must be unit. Proposition 4 is already stated in [13, 14].

Proposition 3 *For both* pARMAX *and* ARMAX *processes we have* $\eta^L_m = 1/2$*, unless* $x_* = 1$ *in* pARMAX *and* $x_* = 0$ *in* ARMAX *leading to* $\eta^L_m = 1$*, for all positive integer* m*.*

Proposition 4 *For the* pARMAX *process we have* $\eta_m = \max(1/2, c^m)$*, whereas for* ARMAX *we have* $\eta_m = 1$*, for all positive integer* m*.*

Note that, in the upper tail case, η_m in *pARMAX* presents also a power decay as the ACF of an AR(1) with a cut-off as in the MA process.

The next two results concern the EDC.

Proposition 5 *Let* d *be a positive integer. The* pARMAX *process is* d*-lower extremal dependent if the left-end-point* $x_* = 1$*, with* $\epsilon^L_d = c^{d-1}/(d-(d-1)c)$*, and* d*-lower extremal independent otherwise. The* pARMAX *process is* d*-lower extremal dependent if the left-end-point* $x_* = 0$*, with* $\epsilon^L_d = 1$*, and* d*-lower extremal independent otherwise.*

Proposition 6 *Let* d *be a positive integer. The* pARMAX *process is* d*-upper extremal independent whereas the* ARMAX *is* d*-upper extremal dependent, with* $\epsilon^U_d = \frac{c^{(d-1)/\gamma}}{d-(d-1)c^{1/\gamma}}$*.*

Note as a curiosity, whenever a non null and non unit coefficient is obtained, the tail dependence presents a power decay as the ACF of an AR(1) process.

Proposition 7 *Both processes,* ARMAX *and* pARMAX*, satisfy the conditions of Theorem 1 and 2.*

Appendix: PROOFS

The stationarity equation of *pARMAX* in (26) is given by

$$F(x) = F(x^{1/c})F_Z(x)\,. \tag{A.1}$$

whilst for *ARMAX* in (27) it is given by

$$F(x) = F(x/c)F_Z(x)\,, \tag{A.2}$$

Using the latest, we derive the respective *m-step* transition probability functions (t.p.f.) from x to $]-\infty, y]$: for *pARMAX* process we have,

$$Q^m(x,]-\infty, y]) := P\big(X_{n+m} \le y | X_n = x\big) = \tfrac{F(y)}{F(y^{1/c^m})}\mathbf{1}_{\{x \le y^{1/c^m}\}}. \tag{A.3}$$

and for $ARMAX$ process it is given by,

$$Q^m(x,]-\infty,y]) := P\big(X_{n+m} \le y | X_n = x\big) = \tfrac{F(y)}{F(y/c^m)} \mathbf{1}_{\{x \le y/c^m\}}, \tag{A.4}$$

In the following we present the fundamental relations within each process for the proofs below.

By (28), we have,

$$F^{-1}(1-t) = t^{-\gamma} L_{F^{-1}}(t), \tag{A.5}$$

with function $L_{F^{-1}}$ slowly varying. Since,

$$F(F^{-1}(1-t)) \sim F\big(t^{-\gamma} L_{F^{-1}}(t)\big) = 1 - t\big[L_{F^{-1}}(t)\big]^{-1/\gamma} L_F\big(t^{-\gamma} L_{F^{-1}}(t)\big), \tag{A.6}$$

we have the following relation between L_F and $L_{F^{-1}}$:

$$\big[L_{F^{-1}}(t)\big]^{-1/\gamma} L_F\big(t^{-\gamma} L_{F^{-1}}(t)\big) \sim 1,\ t \downarrow 0. \tag{A.7}$$

- for $pARMAX$ recursion in (26), we have

$$\begin{aligned} P(X_i \le F^{-1}(t), X_j \le F^{-1}(t)) &= \textstyle\int_{x_*}^{F^{-1}(t)} Q^{j-i}(x,]-\infty,F^{-1}(t)])F(dx) \\ &= \frac{F(F^{-1}(t))^2}{F\big(F^{-1}(t)^{1/c^{j-i}}\big)} \end{aligned} \tag{A.8}$$

  where in the last step we have used the t.p.f. given in (A.3). Moreover, for the multivariate case,

$$\begin{aligned} &P(X_{i_1} \le F^{-1}(t), .., X_{i_k} \le F^{-1}(t)) \\ =&\int_{x_*}^{F^{-1}(t)} \cdots \int_{x_*}^{F^{-1}(t)} Q^{i_k - i_{k-1}}\big(x_{i_{k-1}},]-\infty,F^{-1}(t)]\big) \prod_{j=2}^{k-1} Q^{i_{k-j}-i_{k-j+1}}(x_{i_{k-j}}, dx_{i_{k-j+1}}) F(dx_{i_1}) \\ =& \frac{F(F^{-1}(t))^k}{\prod_{j=2}^k F\big(F^{-1}(t)^{1/c^{i_j - i_{j-1}}}\big)} \end{aligned} \tag{A.9}$$

  Observe now that,

$$F(F^{-1}(t)^{1/c^j}) = F\big((1-t)^{-\gamma/c^j}(L_{F^{-1}}(1-t))^{1/c^j}\big) = 1-(1-t)^{1/c^j}\mathcal{L}_j(1-t), \tag{A.10}$$

  where

$$\mathcal{L}_j(1-t) = (L_{F^{-1}}(1-t))^{-1/(\gamma c^j)} L_F\big((1-t)^{-\gamma/c^j}(L_{F^{-1}}(1-t))^{-1/c^j}\big) \tag{A.11}$$

By (A.5) and (A.6),

$$\mathcal{L}_j(1-t) \sim \big(1 - F\big(x_*^{1/c^j}\big)\big), \text{ as } t \downarrow 0, \tag{A.12}$$

and by (A.7),

$$\mathcal{L}_j(1-t) \text{ is slow varying, as } t \uparrow 1. \tag{A.13}$$

- for $ARMAX$ recursion in (27), in a similar manner we derive,

$$P(X_i \le F^{-1}(t), X_j \le F^{-1}(t)) = \frac{F(F^{-1}(t))^2}{F\big(F^{-1}(t)/c^{j-i}\big)} \tag{A.14}$$

and for the more general multivariate case,

$$P(X_{i_1} \le F^{-1}(t), .., X_{i_k} \le F^{-1}(t)) = \frac{F(F^{-1}(t))^k}{\prod_{j=2}^k F\big(F^{-1}(t)/c^{i_j - i_{j-1}}\big)} \tag{A.15}$$

where

$$F(F^{-1}(t)/c^j) = F\big((1-t)^{-\gamma}/c^j L_{F^{-1}}(1-t)\big) = 1 - (1-t)c^{j/\gamma}\mathfrak{L}_j(1-t). \tag{A.16}$$

and

$$\mathfrak{L}_j(1-t) = (L_{F^{-1}}(1-t))^{-1/\gamma} L_F\big((1-t)^{-\gamma}(L_{F^{-1}}(1-t))/c^j\big) \tag{A.17}$$

By (A.5) and (A.6),

$$\mathfrak{L}_j(1-t) \sim \big(1 - F\big(x_*/c^j\big)\big), \text{ as } t \downarrow 0, \tag{A.18}$$

and by (A.7),

$$\mathfrak{L}_j(1-t) \sim 1, \text{ as } t \uparrow 1. \tag{A.19}$$

Proof of Proposition 1

Consider $\{X_i\}$ a $pARMAX$ process as defined in (26). Using (A.8), we have,

$$P(F(X_{1+m}) \le t | F(X_1) \le t) = \frac{P(X_{1+m} \le F^{-1}(t), X_1 \le F^{-1}(t))}{P(X_1 \le F^{-1}(t))} = \frac{F(F^{-1}(t))}{F(F^{-1}(t)^{1/c^m})} \frac{F(F^{-1}(t))}{t}$$

Hence, from relations (A.10)-(A.12), the lower TDC defined in (2), is given by,

$$\lambda_m^L = \lim_{t\downarrow 0} P(F(X_{1+m}) \leq t | F(X_1) \leq t) \sim \frac{t}{1-(1-t)^{1/c^m}\left(1-F\left(x_*^{1/c^m}\right)\right)}. \qquad \text{(A.20)}$$

Observe that, if $x_* = 1$, we have $\left(1 - F\left(x_*^{1/c^m}\right)\right) = 1$ and hence we obtain $\lambda_m^L = c^m$. Otherwise the limit is null.

Using the same reasoning for an $ARMAX$ process $\{X_i\}$ as defined in (27), we obtain,

$$\lambda_m^L = \lim_{t\downarrow 0} P(F(X_{1+m}) \leq t | F(X_1) \leq t) \sim \frac{t}{1-(1-t)(1-F(x_*/c^m))}, \qquad \text{(A.21)}$$

which is null unless $x_* = 0$, being unitary in this case. $\square$

Proof of Proposition 2

In order to derive the upper TDC of a $pARMAX$ process $\{X_i\}$ satisfying (26), observe that,

$$\begin{aligned} P(F(X_{1+m}) > t | F(X_1) > t) &= \frac{P(X_{1+m} > F^{-1}(t), X_1 > F^{-1}(t))}{P(X_1 > F^{-1}(t))} = \frac{\int_{F^{-1}(t)}^{\infty} \left\{1 - Q(x,]-\infty, F^{-1}(t)])\right\} F(dt)}{1-t} \\ &= \frac{1 - F(F^{-1}(t)) - \frac{F(F^{-1}(t))}{F(F^{-1}(t)^{1/c^m})}(F(F^{-1}(t)^{1/c^m}) - F(F^{-1}(t)))}{1-t} \end{aligned} \qquad \text{(A.22)}$$

Hence, applying relations (A.10)-(A.12), the upper TDC defined in (3) is given by,

$$\begin{aligned} \lambda_m^U &= \lim_{t\uparrow 1} P(F(X_{1+m}) > t | F(X_1) > t) = \lim_{t\uparrow 1} \frac{1 - 2t + \frac{t^2}{1-(1-t)^{1/c^m}\mathcal{L}_m(1-t)}}{1-t} \\ &\sim 1 - \frac{t}{1-(1-t)^{1/c^m}\mathcal{L}_m(1-t)} + \frac{t(1-t)^{1/c^m - 1}\mathcal{L}_m(1-t)}{1-(1-t)^{1/c^m}\mathcal{L}_m(1-t)} \sim 0, \text{ as } t \uparrow 1. \end{aligned} \qquad \text{(A.23)}$$

Using the same reasoning for a process $\{X_i\}$ satisfying $ARMAX$ recursion (27), applying now relations (A.16)-(A.18), we derive,

$$\begin{aligned} \lambda_m^U &= \lim_{t\uparrow 1} P(F(X_{1+m}) > t | F(X_1) > t) \\ &\sim 1 - \frac{t}{1-(1-t)c^{m/\gamma}\mathfrak{L}_m(1-t)} + \frac{tc^{m/\gamma}\mathfrak{L}_m(1-t)}{1-(1-t)\mathfrak{L}_m(1-t)c^{m/\gamma}} \sim c^{m/\gamma}, \text{ as } t \uparrow 1. \ \square \end{aligned} \qquad \text{(A.24)}$$

Proof of Proposition 3

According to (4), we must prove that the limiting function $h^L(x,y)$, given by

$$h^L(x,y) = \lim_{t\downarrow 0} \frac{P\big(F(X_1) < tx, F(X_{1+m}) < ty\big)}{P\big(F(X_1) < t, F(X_{1+m}) < t\big)}, \tag{A.25}$$

with $x, y \geq 0$, is homogeneous of order 1/2 (i.e. $h(sx, sy) = s^2 h(x,y)$).

Observe that the same procedure of (A.22) and (A.23) leads to

$$\begin{aligned} P\big(F(X_1) < tx, F(X_{1+m}) < ty)\big) &= \int_{x_*}^{F^{-1}(xt)} Q^m(u,]-\infty, F^{-1}(xt)])F(du) \\ = \int_{x_*}^{F^{-1}(xt)} & Q^m(u,]-\infty, F^{-1}(xt)])F(du)\mathbf{1}_{\{F^{-1}(xt)\leq F^{-1}(yt)^{1/c^m}\}} \\ + \int_{x_*}^{F^{-1}(yt)^{1/c^m}} & Q^m(u,]-\infty, F^{-1}(xt)])F(du)\mathbf{1}_{\{F^{-1}(xt)> F^{-1}(yt)^{1/c^m}\}} \\ = \tfrac{F(F^{-1}(xt))F(F^{-1}(yt))}{F\left(F^{-1}(yt)^{1/c^m}\right)} & \mathbf{1}_{\{F^{-1}(xt)\leq F^{-1}(yt)^{1/c^m}\}} + F(F^{-1}(yt))\mathbf{1}_{\{F^{-1}(xt)> F^{-1}(yt)^{1/c^m}\}}. \end{aligned} \tag{A.26}$$

Relation $F^{-1}(xt) \leq F^{-1}(yt)^{1/c}$ always holds for small enough t, since $F(F^{-1}(xt)) \leq F(F^{-1}(yt)^{1/c^m}) \Leftrightarrow xt \leq 1-(1-(yt)^{1/c}\mathcal{L}_m(1-yt)$ by (A.6) and $\mathcal{L}_m(1-yt) \to 1-F(x_*^{1/c^m})$, as $t \downarrow 0$. Exception made for $x_* = 1$ since $1-F(x_*^{1/c^m}) = 1$.

Considering first $x_* \neq 1$, the function $h^L(x,y)$ in (A.25) (for the denominator just take $x = y = 1$), is derived:

$$h^L(x,y) = \lim_{t\downarrow 0} \frac{P\big(F(X_1) < tx, F(X_{1+m}) < ty\big)}{P\big(F(X_1) < t, F(X_{1+m}) < t\big)} = \lim_{t\downarrow 0} \frac{\frac{xyt^2}{1-(1-yt)^{1/c^m}}\mathcal{L}_m(1-yt)}{\frac{t^2}{1-(1-t)^{1/c^m}}\mathcal{L}_m(1-t)} = xy. \tag{A.27}$$

which is homogeneous of order 2.

If $x_* = 1$, and given the last line of development (A.26), we have now,

$$\lim_{t\downarrow 0} \frac{P\big(F(X_1) < tx, F(X_{1+m}) < ty\big)}{P\big(F(X_1) < t, F(X_{1+m}) < t\big)} = \lim_{t\downarrow 0} \frac{\frac{xyt^2}{1-(1-yt)^{1/c^m}} + yt}{\frac{t^2}{1-(1-t)^{1/c^m}}} = x + y/c. \tag{A.28}$$

In this case, $h^L(x,y) = x + y/c$ is homogeneous of order 1.

The procedure is similar for $ARMAX$. □

Proof of Proposition 5

Consider EDC Definition 2. We start by noting that,

$$P(F_{max} \leq t) \ = \ P(X_1 \leq F^{-1}(t), ..., X_d \leq F^{-1}(t)) \tag{A.29}$$

and also,

$$P(F_{min} > t) \;=\; 1 - \sum_{1<i<d} P(X_i \leq F^{-1}(t)) + \sum_{1<i<j<d} P(X_i \leq F^{-1}(t), X_j \leq F^{-1}(t)) + ... + \tag{A.30}$$

$$+(-1)^d P(X_1 \leq F^{-1}(t), ..., X_d \leq F^{-1}(t)) \tag{A.31}$$

An expression for (A.29) is immediate from (A.9)-(A.12), leading to

$$P(F_{max} \leq t) = t\left(\frac{F(F^{-1}(t))}{F(F^{-1}(t)^{1/c})}\right)^{d-1} = t\left(\frac{t}{1-(1-t)^{1/c}\left(1-F\left(x_*^{1/c}\right)\right)}\right)^{d-1} \tag{A.32}$$

The computation of probability in (A.30) is not so straightforward. For simplicity we illustrate below the calculations for $d = 3$. Observe that, by (A.9),

$$\begin{aligned} P(F_{min} > t) \;=\; & 1 - 3F(F^{-1}(t)) + 2\frac{F(F^{-1}(t))}{F\left(F^{-1}(t)^{1/c}\right)}F(F^{-1}(t)) \\ & + \frac{F(F^{-1}(t))}{F\left(F^{-1}(t)^{1/c^2}\right)}F(F^{-1}(t)) - \left(\frac{F(F^{-1}(t))}{F\left(F^{-1}(t)^{1/c}\right)}\right)^2 F(F^{-1}(t)) \end{aligned} \tag{A.33}$$

and hence, using (A.10)-(A.12), the lower EDC defined in (6), is given by,

$$\begin{aligned} \epsilon_d^L = \lim_{t\downarrow 0} t\left(\frac{t}{1-(1-t)^{1/c}\left(1-F\left(x_*^{1/c}\right)\right)}\right)^2 \Big/ \Bigg[3t - 2t\frac{t}{1-(1-t)^{1/c}\left(1-F\left(x_*^{1/c}\right)\right)} \\ -t\frac{t}{1-(1-t)^{1/c^2}\left(1-F\left(x_*^{1/c^2}\right)\right)} + t\left(\frac{t}{1-(1-t)^{1/c}\left(1-F\left(x_*^{1/c}\right)\right)}\right)^2\Bigg]. \end{aligned} \tag{A.34}$$

If $x_* = 1$, then $\left(1 - F\left(x_*^{1/c}\right)\right) = 1$ and so the above limit becomes

$$\frac{t\left(\frac{t}{1-(1-t)^{1/c}}\right)^2}{3t-2t\frac{t}{1-(1-t)^{1/c}} - t\frac{t}{1-(1-t)^{1/c^2}} + t\left(\frac{t}{1-(1-t)^{1/c}}\right)^2} \sim \frac{c^2}{3-2c}, \text{ as } t \downarrow 0, \tag{A.35}$$

otherwise we have $\left(1 - F\left(x_*^{1/c}\right)\right) \neq 1$ and the limit is null.

Extending the result to any given d, after some more calculations, one can derive the following:

$$\epsilon_d^L = \lim_{t\downarrow 0}\Big[t\Big(\tfrac{t}{1-(1-t)^{1/c}\left(1-F\left(x_*^{1/c}\right)\right)}\Big)^{d-1}\Big]\Big/\Big[dt-(d-1)t\tfrac{t}{1-(1-t)^{1/c}\left(1-F\left(x_*^{1/c}\right)\right)}$$

$$-(d-2)t\tfrac{t}{1-(1-t)^{1/c^2}\left(1-F\left(x_*^{1/c^2}\right)\right)}-\ldots-t\tfrac{t}{1-(1-t)^{1/c^{d-1}}\left(1-F\left(x_*^{1/c^{d-1}}\right)\right)}+\ldots$$

$$+(-1)^{d-1}t\Big(\tfrac{t}{1-(1-t)^{1/c}\left(1-F\left(x_*^{1/c}\right)\right)}\Big)^{d-1}\Big]$$

$$=\tfrac{c^{d-1}}{d-(d-1)c},$$

Hence, for $pARMAX$, if $x_* = 1$ we obtain $\epsilon_d^L = \frac{c^{d-1}}{d-(d-1)c}$, otherwise $\epsilon_d^L = 0$.

For the $ARMAX$ process, and now applying (A.15)-(A.18), expression in (A.29) is given by

$$P(F_{max} \le t) = t\left(\tfrac{F(F^{-1}(t))}{F(F^{-1}(t)/c)}\right)^{d-1} = t\left(\tfrac{t}{1-(1-t)(1-F(x_*/c))}\right)^{d-1} \tag{A.36}$$

and following the same procedure as above, we derive,

$$\begin{aligned}\epsilon_d^L = \lim\nolimits_{t\downarrow 0} t\big(\tfrac{t}{(1-(1-t)(1-F(x_*/c)))}\big)^{d-1}\Big]\Big/\Big[dt-(d-1)t\tfrac{t}{1-(1-t)(1-F(x_*/c))} \\ -(d-2)t\tfrac{t}{1-(1-t)(1-F(x_*/c^2))}-\ldots-t\tfrac{t}{1-(1-t)(1-F(x_*/c^{(d-1)}))}+\ldots \\ +(-1)^{d-1}t\Big(\tfrac{t}{1-(1-t)(1-F(x_*/c))}\Big)^{d-1}\Big].\end{aligned} \tag{A.37}$$

This limit is null unless $x_* = 0$, leading to $1 - F(x_*/c^j) = 1$, $\forall j \ge 1$, and hence to an unit result. □

Proof of Proposition 6

Just observe that the d-upper EDC can be obtained by arithmetically inverting the lower ϵ_d^L and taking the complementary of the numerator and of the denominator, now considering that $t \uparrow 1$. □

Proof of Proposition 7 The β-mixing property has already been stated for $ARMAX$ (see, for instance, [11]), and is straightforward for $pARMAX$ (by taking logarithms of $ARMAX$). The second order condition holds for both $ARMAX$ and $pARMAX$, by considering $A(w) = w^{-1}$.

Conditions (12) and (21) hold for both processes by Propositions 3.3 and 3.6 in [13]. Condition (13) is straightforward for $ARMAX$ by Proposition 3.4 in the latter reference. We need to show that (13) in Theorem 1 and in Theorem 2 also holds in the $pARMAX$ case. Consider $I_n(x,y) =]a_{n,y}, a_{n,x}]$, with $a_{n,y} = F^{-1}\big(1-\frac{k_n}{n}y\big)$ and $a_{n,x} = F^{-1}\big(1-\frac{k_n}{n}x\big)$. We have that,

$$\begin{aligned}&\tfrac{n}{k_n}P(X_1 \in I_n(x,y), X_{1+m} \in I_n(x,y)) \leq \tfrac{n}{k_n}P(X_1 \in I_n(x,y), X_{1+m} > a_{n,y}) \\ = \;& \tfrac{n}{k_n}\Big[F(a_{n,x}) - F(a_{n,y}) - \int_{a_{n,y}}^{a_{n,x}} Q^m(w,]-\infty, a_{n,y}])dF(w)\Big]\end{aligned} \tag{A.38}$$

and applying (A.3) then,

$$\tfrac{n}{k_n}P(X_1 \in I_n(x,y), X_{1+m} \in I_n(x,y)) \leq \tfrac{n}{k_n}\Big[F(a_{n,x}) - F(a_{n,y})\Big](1-F(a_{n,y})), \tag{A.39}$$

and hence we can state,

$$\tfrac{n}{k_n}P(X_1 \in I_n(x,y), X_{1+m} \in I_n(x,y)) \leq (y-x)\big(\tfrac{k_n}{n}(1+\epsilon)\big) \tag{A.40}$$

for some $\epsilon > 0$. In what concerns Theorem 2, we must take interval $I_n(x,y) \equiv I_n^{(T)}(x,y) =]F_T^{-1}\big(1-\frac{k_n}{n}y\big), F_T^{-1}\big(1-\frac{k_n}{n}x\big)]$, where F_T is the d.f. of T_i for each $i \in \mathbb{N}$. Moreover, we will see that the slightly modified but similar condition

$$\tfrac{n}{k_n}P\big(X_1 \in I_n^{(T)}(x,y), X_{1+p} \in I_n^{(T)}(x,y)\big) \leq (y-x)\Big(\tilde{\rho}(p) + D_1\Big(\tfrac{k_n}{n}\Big)^{\alpha}\Big), \tag{A.41}$$

holds for some $0 < \alpha < 1$ with $l_n = o((n/k_n)^{\alpha})$. Observe that, by Proposition 4, F_T has regularly varying tail with index $\eta_m = \max(1/2, c^m)$ and, for the sake of simplicity, it is assumed that, for some real constants, $d > 0$ and $d^* > 0$,

$$F_T^{-1}(1-t) \sim d^* t^{-\eta_m} \quad \text{and} \quad F^{-1}(1-t) \sim d t^{-\gamma}\,,\ \text{ as } t \downarrow 0. \tag{A.42}$$

Denoting, $F^{-1}\big(1-1/F_T^{-1}\big(1-\frac{k_n}{n}x\big)\big) = a_{n,x}$, we have,

$$F\big(a_{n,x}^{1/c^j}\big) \sim 1 - \big(\tfrac{k_n}{n}x\big)^{\eta m} A^{1/c^j} d^{1/\gamma},\ \ \forall j \geq 0, \tag{A.43}$$

where $A = d^{-1/\gamma}(d^*)^{-1}$. Now observe that,

$$\begin{aligned}&\tfrac{n}{k_n}P\Big(X_1 \in I_n^{(T)}(x,y), X_{1+m} \in I_n^{(T)}(x,y)\Big) \\ &\leq \tfrac{n}{k_n}\Big[P\Big(a_{n,y} < X_1 \leq a_{n,x}, X_1 > a_{n,y}^{1/c^m}\Big) + P\Big(a_{n,y} < X_1 \leq a_{n,x}, \bigvee_{j=0}^{m-1}(Z_{m-j+1}^{c^j}) > a_{n,y}\Big)\Big].\end{aligned}$$

By the independence assumptions, the last expression becomes,

$$\frac{n}{k_n}\Big\{\Big[F(a_{n,x})-F\big(a_{n,y}^{1/c^m}\big)\Big]+\Big[F(a_{n,x})-F(a_{n,y})\Big]\cdot\Big[1-\prod_{j=0}^{m-1}F_Z\big(a_{n,y}^{1/c^j}\big)\Big]\Big\}$$

$$\leq \frac{n}{k_n}\Big\{\Big[F(a_{n,x}^{1/c^m})-F(a_{n,y}^{1/c^m})\Big]+\Big[F(a_{n,x})-F(a_{n,y})\Big]\Big[1-F(a_{n,y})\Big]\Big\}$$

Considering (A.43) and conditions in (A.41), then,

$$\Big[F(a_{n,x}^{1/c^m})-F(a_{n,y}^{1/c^m})\Big]+\Big[F(a_{n,x})-F(a_{n,y})\Big]\Big[1-F(a_{n,y})\Big]$$

$$=\Big[\big(\tfrac{k_n}{n}\big)^{\eta_m/c^m}d^{1/\gamma}A^{1/c^m}\big(y^{\eta_m/c^m}-x^{\eta_m/c^m}\big)\Big]+\Big[\big(\tfrac{k_n}{n}\big)^{\eta_m}d^{1/\gamma}A(y^{\eta_m}-x^{\eta_m})\Big]\big(\tfrac{k_n}{n}y\big)^{\eta_m}d^{1/\gamma}A$$

$$\leq (y-x)\tfrac{k_n}{n}\Big(d^{1/\gamma}A^{1/c^m}\Big[\tfrac{k_n}{n}(1+\epsilon)\Big]^{\eta_m/c^m-1}\tfrac{1+\epsilon}{\delta}+\big(\tfrac{k_n}{n}\big)^{2\eta_m-1}\tfrac{(1+\epsilon)^{2\eta_m}}{\delta}d^{2/\gamma}A^2\Big)$$

$$\leq (y-x)\tfrac{k_n}{n}\Big(\widetilde{\rho}(m)+\big(\tfrac{k_n}{n}\big)^{\alpha}D_1\Big).$$

where $\delta = y-x$, $\widetilde{\rho}(m) = d^{1/\gamma}A^{1/c^m}\Big[\frac{k_n}{n}(1+\epsilon)\Big]^{\eta_m/c^m-1}\frac{1+\epsilon}{\delta}$, $D_1 = \frac{(1+\epsilon)^{2\eta_m}}{\delta}d^{2/\gamma}A^2 > 0$ and $\alpha = 2\eta_m - 1$. Note that, $\sum_{m=0}^{\infty}\widetilde{\rho}(m) < \infty$ since, from some order n,

$$\lim_{m\to\infty}\widetilde{\rho}(m+1)/\widetilde{\rho}(m) \sim \Big\{A\Big[\tfrac{k_n}{n}(1+\epsilon)\Big]^{\eta_m}\Big\}^{\frac{1}{c^m}(1/c-1)} < 1. \ \square$$

---

[1] A. Ledford and J.A. Tawn, (1996). J.A. Statistics for near independence in multivariate extreme values, *Biometrika* **83**, 169-187.

[2] A. Ledford and J.A. Tawn, (1997). Modelling Dependence within joint tail regions," *J. Royal Statist. Soc.Ser. B* **59**, 475-499.

[3] A.V. Lebedev, (2008), Statistical analysis of first-order MARMA Processes, *Math. Notes* **83**, 506-511.

[4] D. Daley and J. Haslet, (1982). A thermal energy storage with controlled input, *Adv. in Appl. Probab.* **14** (2), 257-271.

[5] G. Draisma, H. Drees, (2004). A. Ferreira and L. de Haan, Bivariate tail estimation: dependence in asymptotic independence, *Bernoulli* **10**, 251-280.

[6] G. Frahm, "On the Extremal Dependence Coefficient of Multivariate Distributions," Statist. Probab. Lett. **76**, 1470-1481 (2006).

[7] H. Drees, (2003). Extreme quantile estimation for dependent data with applications to finance, *Bernoulli* **9**, 617-657.

[8] I. Helland and T. Nilsen, (1976). On a general random exchange model, *J. Appl. Probab.* **13** (4), 781-790.

[9] J. Klotz, (1973). Statistical inference in Bernoulli trials with dependence, *Ann. Statist.* **1**, 373-379.

[10] J.W. Pratt, (1960). On interchanging limits and integrals. *Ann. Math. Statist.* **31**, 74-77.

[11] L. Canto e Castro, (1992). *Sobre a Teoria Assintótica de Extremos*, Ph.D. Thesis, FCUL.

[12] M.A. Ancona-Navarrete and J.A. Tawn, (2000). A comparison of methods for estimating the extremal index, *Extremes* **3**, 5-38.

[13] M. Ferreira and L. Canto e Castro, (2008a). Tail and dependence behavior of levels that persist for a fixed period of time, *Extremes* **11**, 113-133.

[14] M. Ferreira and L. Canto e Castro, (2008b). *Asymptotic and pre-asymptotic tail behavior of a power max-autoregressive model*, Tech. Rep. Center Stat. Applic. (Univ. Lisbon, 2008); `http://www.ceaul.fc.ul.pt/notas.html?ano=2008`.

[15] M. Ferreira and L. Canto e Castro, (2008b). *Modeling rare events through a pRARMAX process*, Tech. Rep. Center Stat. Applic. (Univ. Lisbon, 2009); `http://www.ceaul.fc.ul.pt/notas.html?ano=2009`.

[16] M. Junker and A. May, (2005). Measurement of aggregate risk with copulas, *Econom. J.* **8**, 428-454.

[17] M. Sibuya, (1960). Bivariate extreme statistics, *Ann. Inst. Statist. Math.* **11**, 195-210.

[18] M. T. Alpuim, (1989). An extremal Markovian sequence, *J. Appl. Probab.* **26** (2), 219-232.

[19] P. Bortot and J.A. Tawn, (1998). Models for the extremes of Markov Chains, *Biometrika* **85**, 851-867.

[20] P. Embrechts, C. Klüppelberg, and T. Mikosch,(1997). *Modelling Extremal Events for Insurance and Finance*, in *Applications of Mathematics*, Springer-Verlag, Berlin, Vol. 33.

[21] R. A. Davis and S. I. Resnick, (1989). Basic properties and prediction of max-ARMA processes, *Adv. in Appl. Probab.* **21** (4), 781-803.

[22] R. Schmidt and U. Stadtmüller, (2006). Nonparametric estimation of tail dependence," *Scand. J. Statist.* **33**, 307-335.

[23] T. Ané and C. Kharoubi, (2003). Dependence structure and risk measure,. *J. Bus. Econom.*

Statist. **76**, 411-438.

[24] Z. Zhang and R. L. Smith, (2001). *Modelling Financial Time Series Data as Moving Maxima Processes*, Tech. Rep. Dept. Stat. (Univ. North Carolina, Chapel Hill, NC); `http://www.stat.unc.edu/faculty/rs/ papers/RLS_Papers.html.`