Box-Cox transformation on the estimation of extreme value index (EVI) and high quantiles for heavy-tailed distributions under dependence serials

Mame Birame Diouf, Hadji Deme, Solym M Manou-Abi, Yousri Slaoui

Electronic Journal of Statistics
ISSN: 1935-7524

Box-Cox transformation on the estimation of extreme value index (EVI) and high quantiles for heavy-tailed distributions under dependence serials

Mame Birame Diouf, Elhadji Deme*,

LERSTAD, Département de Mathématiques Appliquées,
UFR SAT, Université Gaston Berger
Saint-Louis, Sénégal
e-mail: mamebiramediouf18@yahoo.com; elhadji.deme@ugb.edu.sn

Solym M. Manou-Abi,

Institut Montpelliérain Alexander Grothendieck, Univ. Montpellier, CNRS, Montpellier, France, Centre Universitaire de Formation et de Recherche, Dembeni, France
e-mail: solym.manou-abi@univ-mayotte.fr

Yousri Slaoui

Laboratoire de Mathématiques et Applications, Université de Poitiers
Poitiers, France
e-mail: Yousri.Slaoui@math.univ-poitiers.fr

Abstract: The Box-Cox transformation is used to make data more suitable for statistical analysis. Its application to extreme value statistics increases the convergence rate of some classical estimators of the tail index and reduces their bias in the context of independent and identically distributed (i.i.d.) random variables. In this paper, we study a bias-reducing estimator of the extreme value index under β-mixing serials dependence using the Box-Cox transformation. Under certain regular conditions, we establish the asymptotic normality of the proposed estimator and derive an asymptotically unbiased estimator of high quantiles. In a simulation study, we highlight the performance of our proposals, which we compare with alternative estimators recently introduced in the literature.

MSC2020 subject classifications: 62E20, 62G30, 62G32.
Keywords and phrases: Estimation, Asymptotic normality, Tail index, Extreme quantiles, Heavy-tailed, Bias reduction, dependent serials, Box-Cox transformations.

Contents

*Corresponding author

1. Introduction and Preliminaries

Quantitative Risk Management (QRM) is a field that aims to build models for understanding financial and environmental risks. In the field of environmental risk management, research related to risk measures such as high quantiles has interested many researchers. Due to the increasing frequency of extreme events and their negative impact on society, the estimation of return levels, which is linked to the estimation of the high quantile of observations, is of great interest. Extreme value theory (EVT) establishes the asymptotic behavior of the largest observations in a sample. It provides methods for extending empirical distribution functions beyond the observed data. This makes it possible to estimate quantities linked to distribution tails, such as high quantiles. The tail distribution function of these observations is characterized by a parameter called the extreme value index (E.V.I.), which indicates the size and frequency of certain extreme phenomena.

Classical extreme value methodology assumes a sample of independent, identically distributed (i.i.d.) random variables $(X_1, ..., X_n)$, with a distribution function $F(x) = \mathbb{P}(X_1 \leq x)$. The main result of extreme value theory is the limiting distribution of the standardized maximum of $n > 1$, a sample of i.i.d. random variables $(X_1, ..., X_n)$:

$$\mathbb{P}\left(a_n^{-1}\left(\max_{1\leq i\leq n} X_i - b_n\right)\right) \to G(x), \quad \text{as } n \to \infty, \tag{1.1}$$

for all points of continuity of G, where $a_n > 0$, $b_n \in \mathbb{R}$ are normalized sequences and G is a non-degenerate limiting distribution function.

Necessary, G is the same type of the following generalized extreme value (GEV) distribution:

$$G_\gamma(x) = \exp\left(-(1+\gamma x)_+^{-1/\gamma}\right),$$

where $y_+ = \max(y, 0)$ and $G_\gamma(x) = \exp(e^{-x})$, for $\gamma = 0$. Here, the real-valued parameter is called the extreme value index γ of F, which in turn is said to belong to the maximum domain of attraction of G_γ, denoted by $F \in D(G_\gamma)$. We refer to [6], for general explanations of extreme value theory.
However, the i.i.d. assumption is often violated in practice, as financial or environmental observations reveal the presence of serial dependence. Let $X_i,\ i \in \mathbb{N}^*$, be a stationary time series defined on a probability space $(\Omega, \mathcal{A}, \mathbb{P})$, with common marginal distribution function (df) $F(x) = \mathbb{P}(X_1 \leq x)$. If $X_i, i \in \mathbb{N}^*$ are weakly dependent, then (1.1) is equivalent to weak convergence of the distribution function of the normalized maximum of n observations to G.

If the $X_i, i \in \mathbb{N}^*$ are weakly dependent, then (1.1) is equivalent to the weak convergence of the distribution function of the standardized maximum of n observations to G. In general, the maximum of a stationary time series is stochastic-ally smaller than the maximum of an independent, identically distributed (i.i.d.) sequence with the same marginal distribution function. Indeed, under certain conditions on the dependency structure, $F \in D(G_\gamma)$ implies,

$$\mathcal{L}\left(a_n^{-1}\left(\max_{1\leq i\leq n} X_i - b_n\right)\right) \to G_\gamma^\theta \quad \text{weakly}, \tag{1.2}$$

for some $\theta \in [0, 1]$, see [19], section 3.7) for more details.

Throughout the article, we assume that the stationary time series $X_i,\ i \in \mathbb{N}$ satisfies the following β-mixing dependence structure condition:

$$\beta(m) := \sup_{p\geq 1} \mathbb{E}\left\{ \sup_{C\in\mathcal{B}_{p+m+1}^{\infty}} |\mathbb{P}(C|\mathcal{B}_1^p) - \mathbb{P}(C)| \right\} \to 0,$$

as $m \to \infty$, where $\mathcal{B}_i^j$ denotes the σ-algebra generated by $(X_i, ..., X_j)$. Without loss of generality, $\beta(m)$ measures the total distance of variation between the unconditional future distribution of the time series and the conditional future distribution given the past of the time series when both are detached by m time points.

Assume that the distribution function F is heavy-tailed (belonging to the the Fréchet maximum domain of attraction), i.e there exist a positive number γ and the tail quantile function $U := (1/1 - F)^{\leftarrow}$ where $^{\leftarrow}$ denotes the left-continuous inverse function, such that

$$\lim_{t\to\infty} \frac{U(tx)}{U(t)} = x^{\gamma},\ \forall\ x > 0. \tag{1.3}$$

According to the relation in (1.3), U is considered to be a regularly varying function at infinity with index $\gamma > 0$ and denoted by $U \in RV_{\gamma}$. The extreme value index γ controls the behavior of the tail distribution function. Its estimation has been mainly studied in the case of i.i.d. random variables, although few papers consider it in the case of time series with serial dependence features. We can mention among others, [4], [7], [8], [9] and [16]. Moreover, in the i.i.d. case, the simplest estimator for $\gamma > 0$ is Hill's estimator [17] defined by

$$\widehat{\gamma}_k^H := \frac{1}{k}\sum_{i=1}^{k} \log X_{n-i+1,n} - \log X_{n-k,n}, \tag{1.4}$$

where $X_{1,n}, \leqslant \cdots \leqslant X_{n,n}$ represents the order statistics and $k = k(n)$ represents an intermediate sequence, i.e. a sequence such that :

$$k \to \infty \quad \text{and} \quad k/n \to 0, \quad \text{as} \quad n \to \infty. \tag{1.5}$$

To prove the asymptotic normality of tail-index estimators such as Hill's, we need a second-order condition that specifies the rate of convergence of the left-hand side in (1.3) towards its limit. This condition can be formulated in various ways, as shown below. We'll use this formulation later.

Second order condition (C_{SO}). Suppose that there exists a positive or negative function A with $\lim_{t\to\infty} A(t) = 0$ and a real number $\rho < 0$ such that :

$$\lim_{t\to\infty} \frac{1}{A(t)}\left(\frac{U(tx)}{U(t)} - x^{\gamma}\right) = x^{\gamma}\frac{x^{\rho}-1}{\rho},\ \forall x > 0. \tag{1.6}$$

The rate of convergence of the function A to 0 is essential, as it illustrates the bias term of the tail index estimators.

In this article, we'll be working with models in the Fréchet attraction domain $\gamma > 0$ which belong to the class of Hall-Welsh models [14], i.e. models with the second-order right-tail expansion :

$$1 - F(x) = \left(\frac{x}{c}\right)^{-1/\gamma}\left\{1 + M_1\left(\frac{x}{c}\right)^{-(-\rho/\gamma)} + o\left(x^{-(-\rho/\gamma+\epsilon)}\right)\right\},\quad x \to \infty; \tag{1.7}$$

with $M_1 \neq 0$, $\rho < 0$ and $c > 0$ is the first-order scaling parameter. For these models, the second-order tail quantile function is as follows:

$$U(t) = ct^{\gamma}\left(1 + D_1 t^{\rho} + o(t^{\rho})\right), \quad t \to \infty, \tag{1.8}$$

with $D_1 = \gamma M_1$. If the second-order condition (C_{SO}) is holds with $A(t)$ equivalent to $\gamma\eta t^{\rho}$, $t \to \infty$, then D_1 equivalent to $A(t)/(\rho t^{\rho})$, $t \to \infty$. Consequently, this implies that $M_1 = \rho^{-1}\eta$. In order to study the behavior of the $1 - F$ tail distribution function, we can estimate the unknown parameters γ, ρ and η with precision $c = 1$.

Assuming that the intermediate sequence k is such that $k^{1/2}A(n/k) \to \lambda \in \mathbb{R}$, as $n \to \infty$ and assuming the following regularity conditions on the mixing coefficients β:

Regularity conditions (C_R). There exist $\epsilon > 0$, a bivariate function r and a sequence $\ell = \ell(n)$ such that, as $n \to \infty$,

(a) $\dfrac{\beta(\ell)}{\ell} n + \ell \dfrac{\log^2 k}{\sqrt{k}} \longrightarrow 0$;

(b) $\dfrac{n}{\ell\, k} Cov\left(\sum_{i=1}^{\ell} \mathbb{I}_{\{X_i > F^{\leftarrow}(1-kx/n)\}}, \sum_{i=1}^{\ell_n} \mathbb{I}_{\{X_i > F^{\leftarrow}(1-ky/n)\}}\right) \longrightarrow r(x,y), \quad \forall\, 0 \leqslant x, y \leqslant 1 + \epsilon$;

(c) For some constant C:

$$\frac{n}{\ell\, k}\mathbb{E}\left[\left(\sum_{i=1}^{\ell} \mathbb{I}_{\{F^{\leftarrow}(1-ky/n) < X_i \leqslant F^{\leftarrow}(1-kx/n)\}}\right)^4\right] \leqslant C(y-x), \quad \forall\, 0 \leqslant x < y \leqslant 1+\epsilon \text{ and } n \in \mathbb{N},$$

[8] established the asymptotic normality of $\widehat{\gamma}_k^{(H)}$ as follows

$$\sqrt{k}(\widehat{\gamma}_k^{H} - \gamma) \xrightarrow{d} \mathcal{N}\left(\frac{\lambda}{1-\rho}, \gamma^2 r(1,1)\right), \tag{1.9}$$

where r is the covariance structure in (C_R). But in the i.i.d. context, the asymptotic variance of Hill's estimator $\gamma_k^{(H)}$ is equal to γ^2. In practice, the bias term of $\widehat{\gamma}_{k_n}^{(H)}$ depends on whether ρ is close to zero or not, since under the second-order condition (C_{SO}), the function $|A|$ varies regularly at infinity with the index ρ. This explains all the literature devoted to bias reduction in the i.i.d. context, see, e.g., [2], [11] and [13], etc. However, in the case of β-mixing time series, only the authors [4] and [7] have addressed this problem and proposed bias-reduced estimators for the extreme value index $\gamma > 0$. In addition, they established tasymptotic normality of the proposed estimators under the regularity condition (C_R) and second-order condition (C_{SO}).

These reduced-bias estimators always pose a problem, as they increase the variance of asymptotic distributions. This problem is solved in the i.i.d. context by [21], who used Box Cox transformations and studied a Hill-type estimator with reduced bias and minimal variance. The present paper extends the results obtained by [21] in the case of dependent series.

The remainder of this paper is organized as follows: in section 2 we present the Box-Cox transformation methodology and study a bias-reduced estimator of the extreme value index in the case of mixed β series. In section 3, we derive an unbiased estimator of high quantiles and establish its asymptotic normality. In section 4, we simulate our estimator and compare it with some existing estimators in the literature.

2. Box-Cox Transformation on the EVI estimation

2.1. Box-Cox transformation

Let X be a random variable (r.v.) with d.f. $F = F_X \in D(G_\gamma)$, $\gamma > 0$. The Box-Cox transformation (BC) [21] of X denoted T_X is a function of the parameter $\sigma \in \mathbb{R}$ given as follows:

$$T_X(\sigma) = \begin{cases} \frac{X^\sigma - 1}{\sigma}, & \sigma \neq 0, \\ \log X, & \sigma = 0, \end{cases} \tag{2.1}$$

where $\sigma = 1$ corresponds to a simple change of location, $\sigma = 1/2$ to the square root transformation, and $\sigma = -1$ to the reciprocal transformation.

Remark 2.1. *According to [21], for Hall's class models, the BC transformation of the data increases the rate of convergence of the tail of the distribution to the generalized distribution of extreme values if $\sigma = -\rho/\gamma$, and as a by-product, the bias of the estimation procedure is reduced.*

Consider a transformation BC such that $X^* = X^\sigma + a$ and denote by $(X_1^*, ..., X_n^*)$ the transformed sample, with the d.f. F^* associated with the original sample $(X_1, ..., X_n)$, where $(X_i^* = T_{X_i}(\sigma, a)$, with $T_X(\sigma, a) = X^\sigma + a$. If the original tail quantile function $U := (1/1 - F)^{\leftarrow}$ satisfies the condition (1.3), then the transformed tail quantile function $U^* = (1/1 - F)^{\leftarrow}$ satisfies the condition (1.3). transform $U^* = (1/1 - F^*)^{\leftarrow}$ also satisfies the first-order condition as follows:

$$\lim_{t \to \infty} \frac{U^*(tx)}{U^*(t)} = x^{\gamma\sigma} = x^{\gamma^*} \Longleftrightarrow U^* \in RV_{\gamma^* = \gamma\sigma}. \tag{2.2}$$

Under the validity of Equation (2.2) and for some c^*, $c > 0$, we have:

$$U^*(t) \sim c^* t^{\gamma^*} = (ct^\gamma)^\sigma. \quad t \to \infty.$$

[22] used the theory of extended regular variation (see [6]) to determine the optimal values of σ that maximize the rate of convergence of the second-order condition for values of $\gamma \geq 0$ and when $\gamma > 0$ and $\gamma + \rho > 0$, a BC transformation has no effect on the S_{OC} of U^* unless $\sigma = -\rho/\gamma$; if $\gamma > 0$ and $\gamma + \rho < 0$, the BC transformation can have a negative or positive effect on the S_{OC} of U^* and the positive effect occurs when $\sigma = -\rho/\gamma$. According to [22], if $\gamma > 0$ and $\gamma + \rho = 0$, for any value other than $\sigma = -\rho/\gamma = 1$, there is no improvement in the speed of convergence of S_{OC} . Thus, the value of σ that maximizes the rate of convergence of S_{OC} that will be considered in this work is as follows:

$$\sigma = -\frac{\rho}{\gamma}. \tag{2.3}$$

This choice will improve bias reduction in the estimation procedure in the following sections.

2.2. Box-Cox Hill-type estimator in the β-mixing case

Let $(X_1, X_2, ...)$ be a stationary β-mixing time series with a continuous common marginal distribution function F. Consider a BC transformation such that $X^* = X^\sigma + a$ and denote by $(X_1^*, ..., X_n^*)$ the transformed sample, with d.f. F^*, associated to the original sample $(X_1, X_2, ..., X_n)$, where $X_i^* = T_{X_i}(\sigma, a)$ and $T_{X_i}(\sigma, a) = X_i^\sigma + a$..

Based on the BC transformation $T_{X_i}(\sigma, a) = X_i^\sigma + a$, [22] introduced a Hill-type estimator, called the Hill Box-Cox estimator (Hill BC) defined as follows:

$$\widehat{\gamma}_k^{BC} = \widehat{\gamma}_k^{BC}(\sigma, a) = \frac{1}{k\,\sigma} \sum_{i=1}^{k} \log\left(\frac{X_{n-i+1,n}^\sigma + a}{X_{n-k,n}^\sigma + a} \right), \tag{2.4}$$

where $a > 0$, so that $(X_i^\sigma + a)_{i=1}^n$ is strictly positive, $\sigma = -\rho/\gamma > 0$, as in Equation (2.3). The Hill estimator is a particular case of the Hill Box-Cox estimator which is much more general, one can check that $\widehat{\gamma}_k^{BC}(1, 0) = \widehat{\gamma}_k^H$.

In the following theorem, we begin by deriving the asymptotic behavior of Hill's BC estimator, assuming that the parameters BC σ and a are known.

Theorem 2.1. *In the class of Hall-Welsh models of equation (1.8), let $(X_1, X_2, \dots)$ be a stationary β-mixing time series with a continuous common marginal distribution function F and assume that c_{SO}) and c_R) hold. Let $k = k(n)$ be an intermediate sequence satisfying the condition (1.5), when $n \to \infty$. Then, under a Skorohod construction, there exists a centered Gaussian process $(W(t))_{t\in[0,1]}$ with covariance function $r(.,.)$, such that, when $n \to \infty$, we have the following distributional representation: $(W(t))_{t\in[0,1]}$, we have the following distributional representation:*

$$\widehat{\gamma}_k^{BC}(\sigma, a) \overset{d}{=} \gamma + \frac{\gamma}{\sqrt{k}} \int_0^1 \left(t^{-1} W(t) - W(1) \right) dt + \frac{A(n/k)}{1-\rho} \left\{ 1 - \frac{a}{\eta c^\sigma} \right\} (1 + o_{\mathbb{P}}(1)).$$

More precisely, if $k^{1/2} A(n/k) \to \lambda \in \mathbb{R}$, as $n \to \infty$, we have

$$\sqrt{k} \left(\widehat{\gamma}_k^{BC}(\sigma, a) - \gamma \right) \overset{d}{\to} \mathcal{N}\left(\frac{\lambda}{1-\rho} \left\{ 1 - \frac{a}{\eta c^\sigma} \right\},\ \gamma^2 r(1,1)) \right). \tag{2.5}$$

Theorem 2.1 extends the results of [21] in serial dependence. Next, we note that, if $a = \eta c^\sigma$, then

$$\sqrt{k} \left(\widehat{\gamma}_k^{BC}(\sigma, \eta c^\sigma) - \gamma \right) \overset{d}{\to} N(0, \gamma^2 r(1,1)), \quad \text{as } n \to \infty. \tag{2.6}$$

Clearly, $\widehat{\gamma}_k^{BC}(\sigma, \eta c^\sigma)$ is an asymptotically unbiased estimator for the extreme value index γ in the case of series dependence. From a practical point of view, however, it cannot be obtained directly, as it depends on the unknown parameters σ, η and c.

To solve this problem, we adopt the same approach as in [21], for which we choose the misspecification $c = 1$ and replace σ and η respectively by their semiparametric estimators. The resulting unbiased plug-in estimator is studied in the next section.

2.3. Reduced bias of the Hill Box-Cox estimator

In this section, we introduce a bias-reduced estimator of the extreme value index $\gamma > 0$. As mentioned previously, we first establish consistent estimators for σ and η and derive an estimated version of $\widehat{\gamma}_k^{BC}(\sigma, \eta)$.

Now, it comes from (2.3), that $\sigma = -\rho/\gamma$. Let $\widehat{\rho}_{k_\rho}$ be an external estimator for ρ, consistent in probability, which depends on an intermediate sequence of integers $k_\rho := k_\rho(n)$, greater than k and satisfying:

$$k_\rho \to \infty \quad \text{and} \quad k_\rho/n \to 0 \quad \text{as,} \quad n \to \infty. \tag{2.7}$$

Then, we can estimate σ by

$$\widehat{\sigma}_{k,k_\rho} = -\widehat{\rho}_{k_\rho}/\widehat{\gamma}_k^H, \tag{2.8}$$

where $\widehat{\gamma}_k^H$ is Hill's estimator in (1.4).

A possible choice for $\widehat{\rho}_{k_\rho}$ is one of the best-performing consistent estimator among those studied in the i.i.d. case (see, e.g, [10], [12]) and also used in the β-mixing case by [4, 7]. This estimator is defined as follows:

$$\widehat{\rho}^*_{k_\rho} = \frac{6S^{(2)}_{k_\rho} - 4 + \sqrt{3S^{(2)}_{k_\rho} - 2}}{4S^{(2)}_{k_\rho} - 3}, \quad \text{provided} \quad S^{(2)}_{k_\rho} \in \left(\frac{2}{3}, \frac{3}{4}\right), \tag{2.9}$$

where

$$S^{(2)}_{k_\rho} = \frac{3}{4} \frac{\left[M^{(4)}_{k_\rho} - 24\left(M^{(1)}_{k_\rho}\right)^4\right]\left[M^{(2)}_{k_\rho} - 2\left(M^{(1)}_{k_\rho}\right)^2\right]}{\left[M^{(3)}_{k_\rho} - 6\left(M^{(1)}_{k_\rho}\right)^3\right]^2},$$

with

$$M^{(\alpha)}_k = \frac{1}{k}\sum_{i=1}^{k} (\log X_{n-i+1,n} - \log X_{n-k,n})^\alpha, \ \alpha \geq 1.$$

Clearly, Hill's estimator $\widehat{\gamma}_k^H$ corresponds to $M^{(1)}_k$. Under the conditions (C_{SO}) and (C_R) with the additional assumptions $k^{1/2}A(n/k) \to \lambda$ and $k_\rho^{1/2}A(n/k_\rho) \to \infty$, as $n \to \infty$, [7] showed that $\widehat{\rho}^*_{k_\rho}$ is consistent in probability to ρ.

Next, for the estimation of the parameter η, we note that in the class of Hall-Welsh models in equation (1.8), the second-order conditions (C_{SO}) hold with $A(t) = \gamma\eta t^\rho(1 + o(1))$, $t \to \infty$. This implies that $(n/k)^{-\rho}A(n/k)/\gamma \to \eta$, as $n \to \infty$. Therefore, referring to the ¨Lemma 5.1 in the appendix of this document, we have as $n \to \infty$:

$$\frac{(1-\rho)^2\left(M^{(2)}_k - 2\left(M^{(1)}_k\right)^2\right)}{2\gamma\rho A(n/k)} = 1 + o_{\mathbb{P}}(1), \tag{2.10}$$

This implies that

$$\frac{(1-\rho)^2\left(M^{(2)}_k - 2\left(M^{(1)}_k\right)^2\right)}{2\gamma^2\rho\left(\frac{n}{k}\right)^\rho} = \eta(1 + o_{\mathbb{P}}(1)). \tag{2.11}$$

Consequently, the parameter η can be estimated as follows:

$$\widehat{\eta}_{k,k_\rho} = \frac{\left(1-\widehat{\rho}_{k_\rho}\right)^2\left(M^{(2)}_k - 2\left(M^{(1)}_k\right)^2\right)}{2\left(M^{(1)}_k\right)^2\left(\frac{n}{k}\right)^{\widehat{\rho}_{k_\rho}}\widehat{\rho}_{k_\rho}}. \tag{2.12}$$

Finally, we obtain our new unbiased Hill Box-Cox estimator, which is expressed as follows:

$$\widehat{\gamma}^{BC}_{k,k_\rho} := \widehat{\gamma}^{BC}_k(\widehat{\sigma}_{k,k_\rho}, \widehat{\eta}_{k,k_\rho}) = \frac{1}{k\ \widehat{\sigma}_{k,k_\rho}}\sum_{i=1}^{k}\log\left(\frac{X^{\widehat{\sigma}_{k,k_\rho}}_{n-i+1,n} + \widehat{\eta}_{k,k_\rho}}{X^{\widehat{\sigma}_{k,k_\rho}}_{n-k,n} + \widehat{\eta}_{k,k_\rho}}\right). \tag{2.13}$$

In the i.i.d. context, [21] studied the consistency of the estimator $\widehat{\gamma}^{BC}_{k,k_\rho}$, but they did not prove asymptotic normality. Our next result, Theorem 2.2 below, is a more general version, since it solves at the same time the problem of asymptotic normality in the i.i.d. case, which we have established in the β-mixing-context case. Before establishing the main result of the new estimator $\widehat{\gamma}^{BC}_{k,k_\rho}$, we give in the following proposition the consistency of the estimators $\widehat{\sigma}_{k,k_\rho}$ and $\widehat{\eta}_{k,k_\rho}$ with respect to σ and η respectively.

Proposition 2.1. *Assume that the assumptions of Theorem 2.1 hold. Supose that the the intermediate sequence $k := k(n)$ satisfies the condition of (1.5). Further, let $\widehat{\rho}_{k_\rho}$ be an external estimator for ρ, consistent in probability and such that $|\widehat{\rho}_{k_\rho} - \rho| = O_{\mathbb{P}}(n^{-\nu})$, for some $\nu > 0$, where $k_\rho := k_\rho(n)$ is an intermediate sequence of integers, greater than k and satisfying the condition of (2.7). Then, we have:*

$$\widehat{\sigma}_{k,k_\rho} \xrightarrow{\mathbb{P}} \sigma \quad \text{and} \quad \widehat{\eta}_{k,k_\rho} \xrightarrow{\mathbb{P}} \eta, \quad \text{as} \quad n \to \infty.$$

Note that in [4], the assumption $|\widehat{\rho}_{k_\rho} - \rho| = O_{\mathbb{P}}(n^{-\nu})$, for some $\nu > 0$ is required to establish the asymptotic normality of an extreme quantile estimator under the β mixing serials. From the following theorem, we show the asymptotic normality of the unbiased Hill Box-Cox estimator $\widehat{\gamma}^{BC}_{k,k_\rho}$.

Theorem 2.2. *Assume that the assumptions of Theorem 2.1 hold. Supose that the the intermediate sequence $k := k(n)$ satisfies the condition of (1.5). Further, let $\widehat{\rho}_{k_\rho}$ be an external estimator for ρ, consistent in probability and such that $|\widehat{\rho}_{k_\rho} - \rho| = O_{\mathbb{P}}(n^{-\nu})$, for some $\nu > 0$, where $k_\rho := k_\rho(n)$ is an intermediate sequence of integers, greater than k and satisfying the condition of (2.7). Then, we have::*

$$\sqrt{k}\left(\widehat{\gamma}^{BC}_{k,k_\rho} - \gamma\right) \to N(0, \gamma^2 r(1,1)),$$

as $n \to \infty$, where r is the covariance structure given in (C_R).

3. Extreme quantiles estimation

The importance of estimating the extreme value index lies in its usefulness for estimating extreme quantiles, which is of paramount importance in practice. The quantile, at probability level $(1-t)$ $in(0,1)$ with respect to F denoted by $x(t)$, is defined as follows: $x(t) := U(1/t)$. Consequently, the quantile $x(t)$ is estimated by $\widehat{x}(t) := X_{n-[nt],n}$, where $[nt]$ represents the integer part of nt. Let's now consid a positive sequence $t = t(n)$ that tends to 0, such as $n \to \infty$. It is then possible to establish the consistency of the non-parametric estimator $widehatx(t)$, when $t \to 0$ is sufficiently slow. However, in certain areas of life such as the environment, hydrology, finance and reliability, a major requirement is to find values large enough that the chances of exceeding them are very low. This leads to removing the restriction on the rate of convergence of $t = t(n)$ to 0, as $n \to \infty$. Furthermore, the interest is in estimating $x(p)$, an extreme quantile, where p, the tail probability depends on the observed sample size n (i.e. $p := p(n)$) and $p(n)$ is smaller than $1/n$. It is therefore not possible to obtain a non-parametric estimate of such a quantile.

The objective of this section is to address this estimation problem in a β-mixing series framework in order to estimate $x(p) = U(1/p)$, the extreme quantile with $np < 1$. Based on the Box-Cox transformation procedure, we propose, in a bias-reduced method, to estimate $x(p)$, the extreme quantile. The construction of our bias-reduction procedure is based on the second-order condition (C_{SO}), which is stated as follows:

$$\frac{U(tx)}{U(t)} \simeq x^\gamma \exp\left(A(t)\frac{x^\rho - 1}{\rho}\right)$$

Let $tx = 1/p$ and $t = n/k \to \infty$, as $n \to \infty$. We obtain the following approximation:

$$x(p) = U(1/p) \approx U(n/k)\left(\frac{k}{np}\right)^\gamma \exp\left(A(n/k)\frac{\left(\frac{k}{np}\right)^\rho - 1}{\rho}\right), \tag{3.1}$$

where γ, ρ and $A(n/k)$ are unknown. The first part $U(n/k_n)\,(k/(np))^{\gamma}$ in the right side of (3.1) is exactly estimated by the Weissman's estimator $\widehat{x}_k^{(W)}(p)$, [23] and defined as:

$$\widehat{x}_k^{(W)}(p) = X_{n-k,n}\left(\frac{k}{np}\right)^{\widehat{\gamma}_k^{(H)}}, \tag{3.2}$$

where $X_{n-k,n}$ is the empirical estimator of $U(n/k)$ and $\widehat{\gamma}_k^{(H)}$ is the Hill's estimator of γ. Obviously, $\widehat{x}_k^{(W)}(p)$, Weissman's estimator is potentially biased because it depends on Hill's estimator $\widehat{\gamma}_{k_n}^{(H)}$, which presents a similar problem. The expression $\exp(A(n/k)[(k/(np))^{\rho}-1]/\rho)$ can be considered as a correction term, since $A(n/k)$ tends towards 0. The question of estimating $A(n/k)$ and ρ then arises.

Since we're working here on the class of Hall-Welsh models (1.8), the second-order condition (C_{SO}) hold with $A(t)$ is equivalent to $\gamma\eta t^{\rho}$, as $t\to\infty$. We can therefore estimate $A(n/k)$ as $\widehat{\gamma}_k^H\widehat{\eta}_{k,k_\rho}(n/k)^{\widehat{\rho}_{k_\rho}}$, where $\widehat{\gamma}_k^H$ is Hill's estimator in (1.4), $\widehat{\rho}_{k_\rho}$ is the consistent estimator of ρ defined in (2.9) and $\widehat{\eta}_{k,k_\rho}$ the consistent estimator of η defined in (2.12). Therefore, replacing respectively in the right-hand side of (3.1 $U(n/k)$, γ, ρ and $A(n/k)\sim\gamma\eta(n/k)^{\rho}$ by their estimators $X_{n-k,n}$, $\widehat{\gamma}_{k,k_\rho}^{BC}$, $\widehat{\rho}_{k_\rho}$ and $\widehat{\gamma}_k^H\widehat{\eta}_{k,k_\rho}\,(n/k)^{\widehat{\rho}_{k_\rho}}$, we obtain the following extreme quantile estimator:

$$\widehat{x}_{k,k_\rho}^{BC}(p) := X_{n-k,n}\left(\frac{k}{np}\right)^{\widehat{\gamma}_{k,k_\rho}^{BC}}\exp\left(\widehat{\gamma}_k^H\widehat{\eta}_{k,k_\rho}\left(\frac{n}{k}\right)^{\widehat{\rho}_{k_\rho}}\frac{\left(\frac{k}{np}\right)^{\widehat{\rho}_{k_\rho}}-1}{\widehat{\rho}_{k_\rho}}\right). \tag{3.3}$$

The asymptotic normality of the estimator of the extreme quantile $widehatx_{k,k_\rho}^{BC}$ is established in the following theorem.

Theorem 3.1. *Assume that the assumptions in Theorem 2.1 hold. Let $k=k(n)$ be an intermediate sequence of integers satisfying (1.5) and $\sqrt{k}A(n/k)\to\lambda\in\mathbb{R}$, as $n\to\infty$. Assume in addition that $\sqrt{k}A(n/k)\to\lambda\in\mathbb{R}$, $np/k\to 0$, $\log(np)/\sqrt{k}\to 0$ and $n^{-\alpha}\log p\to 0$ for all $\alpha>0$, $n\to\infty$. Further, let $\widehat{\rho}_{k_\rho}$ be an external estimator for ρ, consistent in probability and such that $|\widehat{\rho}_{k_\rho}-\rho|=O_{\mathbb{P}}(n^{-\nu})$, for some $\nu>0$, where $k_\rho:=k_\rho(n)$ is an intermediate sequence of integers, greater than k and satisfying the condition of (2.7). Then, we have:*

$$\frac{\sqrt{k}}{\log(k/(np))}\left(\frac{\widehat{x}_{k,k_\rho}^{BC}(p)}{x(p)}-1\right)\xrightarrow{d}\mathcal{N}\left(0,\gamma^2 r(1,1)\right),$$

as $n\to\infty$.

Remark 3.1. *Note that the unbiased estimator of the extreme quantiles $\widehat{x}_{k,k_\rho}^{BC}(p)$ is also valid when the observations are i.i.d., for which the asymptotic variance in Theorem 3.1 is equal to γ^2.*

4. Simulation study

In this section, we proceed by generating data from different stationary models, satisfying the (C_R) condition and for which we can estimate the theoretical value of γ and the true extreme quantile $x(p)$, $p=1/1000$. In this spirit, we first consider an i.i.d. sequence of innovations $(\varepsilon_1,...,\varepsilon_n)$ distributed from F_ε given by

$$F_\varepsilon(\varepsilon)=\begin{cases}(1-q)(1-\widetilde{F}(-\varepsilon)) & \text{if}\quad \varepsilon<0,\\ 1-q+q\widetilde{F}(\varepsilon)) & \text{if}\quad \varepsilon>0,\end{cases}$$

where $\widetilde{F}$ is the unit Fréchet distribution function with $\widetilde{F}(\varepsilon)=1-\exp(-1/\epsilon)$, for $\varepsilon>0$, and $q=0.75$. Then F_ε belongs to the Fréchet domain of attraction, with extreme value index $\gamma=1$. Next, we generate

$N = 1000$ stationary time series $(X_1, ..., X_n)$ of size $n = 1000$ based on $(\varepsilon_1, ..., \varepsilon_n)$, by the following models, for $t = 1, ..., n$:

- The stationary AutoRegressive $AR(1)$ model: $X_t = \theta X_{t-1} + \varepsilon_t,\ \theta \in (0, 1)$,

- The stationnary Moving Average $MA(1)$ model: $X_t = \theta \epsilon_{t-1} + \varepsilon_t,\ \theta \in (0, 1)$.

Under the models listed above, we select three time series models as follows:

• **Model 1:** Independence model $X_t = \epsilon_t$, for which the theoretical value of $x(0.001)$ is 749.80.

• **Model 2:** The AR(1) model with $\theta = 0.3$, for whch the theoretical value of $x(0.001)$ is 1072.26.

• **Model 3:** The MA(1) model with $\theta = 0.3$. The theoretical value of $x(0.001)$ is 972.85.

From the models below, we simultaneously apply the original Hill estimator $\widehat{\gamma}_k^H$ as in Eq. 1.4, the reduced-bias estimator of $\gamma > 0$, studied in [7] and denoted by $\widehat{\gamma}_{k,k_\rho}^{dH}$, the reduced-bias estimator of $\gamma > 0$, studied in [4] and denoted by $\widehat{\gamma}_{k,k_\rho}^{Ch}$ and our reduced-bias Hill Box-Cox estimator $\widehat{\gamma}_{k,k_\rho}^{BC}$ as in Eq. (2.13), for $k = 1, ..., m$ with m the number of positive values in each model studied. We also apply the associated extreme quantile estimators, denoted respectively by $\widehat{x}_k^W(p)$, $\widehat{x}_{k,\widehat{\rho}}^{dH}(p)$, $\widehat{x}_{k,k_\rho}^{Ch}(p)$ and $\widehat{x}_{k,k_\rho}^{BC}(p)$.

To calculate the bias-reduced estimators of the tree considered models and the associated extreme quantiles, we use the second-order estimator $\widehat{\rho}_{k_\rho}$ defined in (2.9), where the intermediate sequence k_ρ is selected as follows:

$$k_\rho := \sup\left\{k :\ k \le \min\left(n-1, \frac{2m}{\log\log m})\right) \quad \text{and} \quad \widehat{\rho}_k \quad \text{exists}\right\}.$$

Finally, we compare the performance of the estimators of the tail index in each model. Moreover, we do the same for the estimators of the extreme quantiles. For this reason, we compute the absolute value of the mean bias (ABias) and the root mean square error (RMSE) based on the N samples, defined as follows

$$\text{ABias}(\eta, k) := \left|\frac{1}{N}\sum_{i=1}^{N}\frac{\widehat{\eta}^{(i)}}{\eta} - 1\right| \quad \text{and} \quad \text{RMSE}(\eta, k) := \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(\frac{\widehat{\eta}^{(i)}}{\eta} - 1\right)^2},$$

where η is either γ or $x(p)$, and $\widehat{\eta}^{(i)}$ is the i-th value $(i = 1, ..., N)$ of an estimator of γ or $x(p)$ evaluated at $k = 1, ..., m$.

• In terms of bias in models 1, 2 and 3, our extreme value index $\widehat{\gamma}_k^{BC}$ in (2.13) is much more stable and has a smaller bias than those of the Hill estimator $\widehat{\gamma}_k^H$ in (1.4, the unbiased Hill estimator $\widehat{\gamma}_k^{dH}$ in [7] and the unbiased Hill estimator $\widehat{\gamma}_k^{Ch}$ in [4] for the AR(1) and MA(1) models. In terms of RMSE, our estimator is better and significantly more competitive, as it retains the lowest RMSE values and is very stable compared with the three alternative estimators for the three models.

• Our high-quantile estimator $\widehat{x}_k^{BC}$ has a smaller and more stable bias than $\widehat{x}_k^W$ $\widehat{x}_k^{dH}$ and $\widehat{x}_k^{Ch}$ for smaller values of k and maintains its stability for larger values of k in all the three considered models. In terms of RMSE, our estimator also retains the lowest values and is very stable compared with the three alternative estimators for all three models.

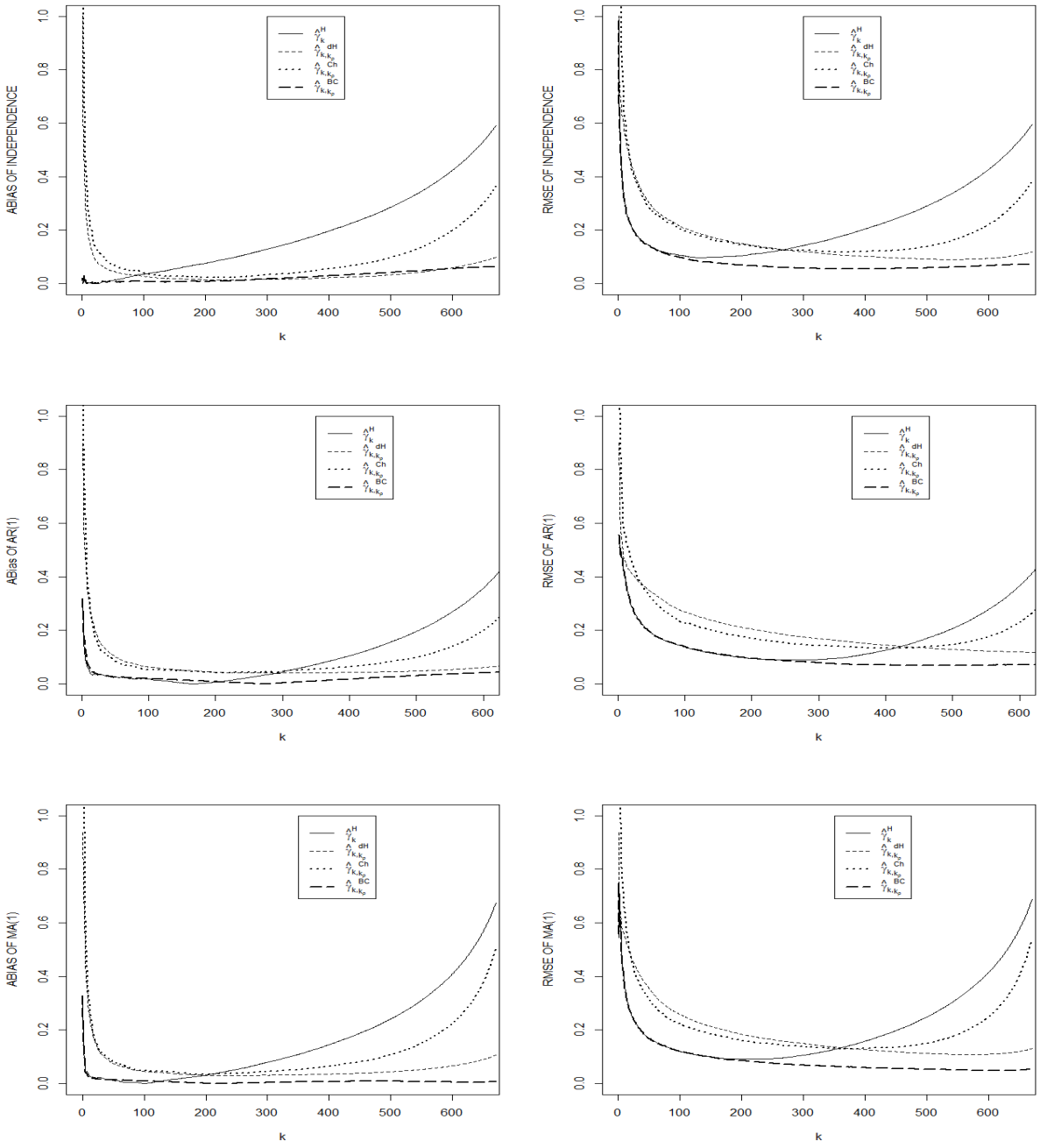

Fig 4.1: Simulation of the tail index: By row, Models 1, 2, 3. By column, ABias (*left*) and RMSE (*right*) as functions of k.

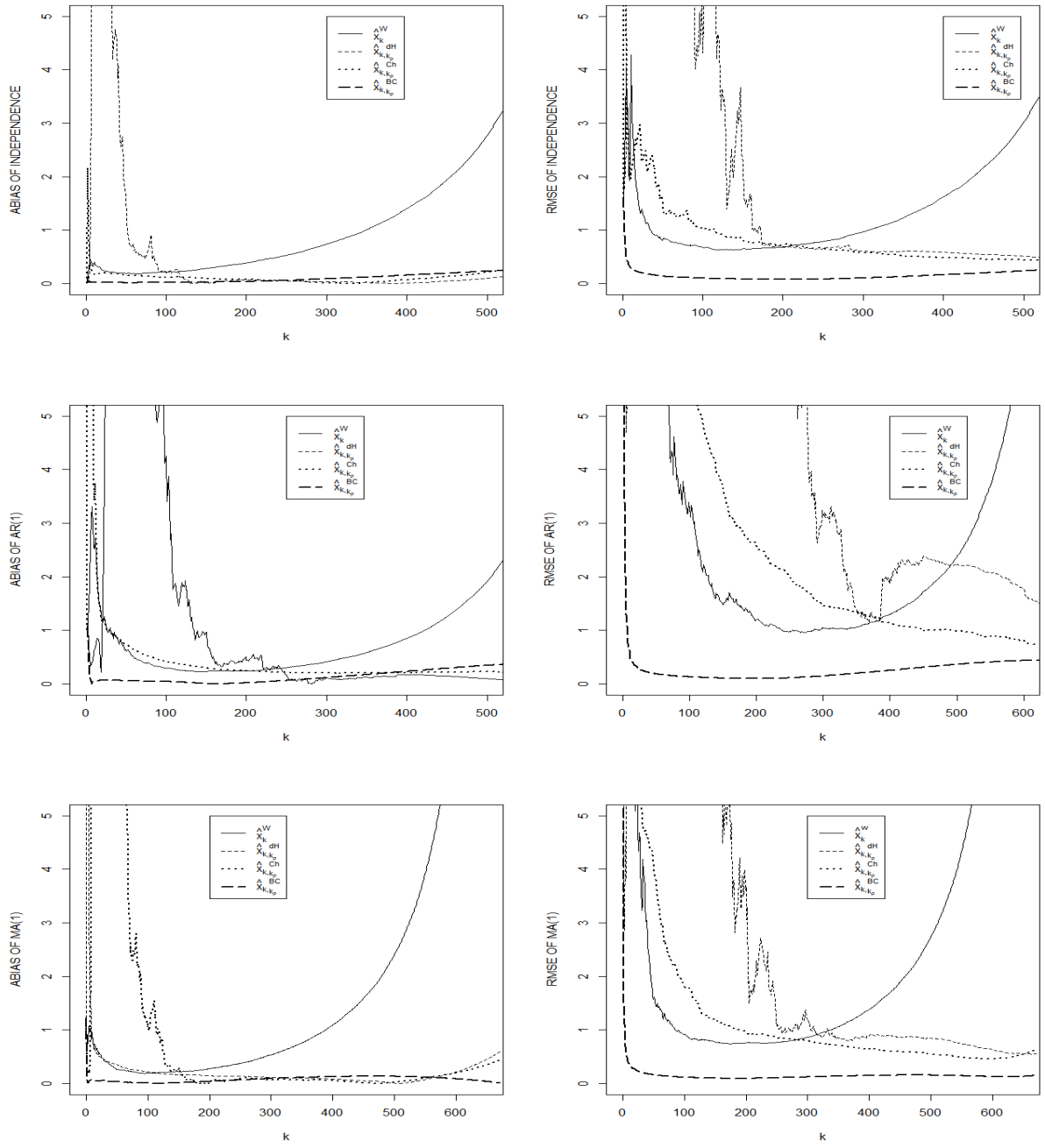

Fig 4.2: Simulation of the extreme quantiles estimators: By row, Models 1, 2, 3. By column, ABias (*left*) and RMSE (*right*) as functions of k.

5. Conclusion

In this paper, we have introduced a new asymptotically unbiased high quantile estimator for β-mixing stationary time series. mixing stationary time series. Comparing the new procedure to the alternative proposed by de Haan et al. (2016), our high-quantile estimator offers, in addition to lower ABias and RMSE in general, greater stability over k, which is an important feature, greater stability over k, an important feature expected in this type of approach to be applicable in practice. practice. In application, the new high-quantile estimator can be proposed to any other stationary model of the type mixing heavy-tailed time series for which high quantiles need to be calculated. This applies to heavy-tailed autoregressive

data encountered in network traffic forecasting for instance and many other applications data in climate change.

Appendix: Proofs of the results

Before establishing the proofs of our Theorems, we need the following preliminary result, which is similar to Corollary A.2 of [7], but assuming $\sqrt{k}A(n/k) = O(1)$ and no third-order condition. In particular, no third-order condition is assumed.

Before establishing the proofs of our Theorems, we need the following preliminary result which is similar to that of Corollary A.2 in [7], but assuming that $\sqrt{k}A(n/k) = O(1)$ and without assuming a third order condition.

Lemma 5.1. *Let $(X_1, X_2, ...)$ be a stationary β-mixing time series with a continuous common marginal distribution function F and assume $(CS0)$ and (C_R). Suppose k is an intermediate sequence satisfying $k \to \infty$, $k/n \to 0$ and $\sqrt{k}A(n/k) = O(1)$, as $n \to \infty$.*

For a given $\epsilon > 0$, under a Skorohod construction, there exists a function $\widetilde{A} \sim A$ and a centered Gaussian process $(W(t))_{t\in[0,1]}$ with covariance structure r defined in (C_R), such that, as $n \to \infty$,

$$\sqrt{k}\left(M_k^{(\alpha)} - \gamma^\alpha\Gamma(\alpha+1)\right) - \alpha\gamma^\alpha P^{(\alpha)} - \sqrt{k}\widetilde{A}\left(\frac{n}{k}\right)\gamma^{\alpha-1}\Gamma(\alpha+1)\frac{1}{\rho}\left(\frac{1}{(1-\rho)^\alpha} - 1\right) \to 0,$$

as $n \to \infty$, where $\Gamma(\alpha+1) = \int_0^1(-\log t)^\alpha dt$ and the random term $P^{(\alpha)} = \int_0^1(-\log t)^{\alpha-1}\left(t^{-1}W(t) - W(1)\right)dt$ is a centered Gaussian process with covariance $c_{\alpha,\alpha'} = Cov\left(P^{(\alpha)}, P^{(\alpha')}\right)$; $\alpha,\ \alpha' > 1$, defined as,

$$c_{\alpha,\alpha'} = \int_0^1\int_0^1(-\log s)^{\alpha-1}(-\log t)^{\alpha'-1}\left(\frac{r(s,t)}{st} - \frac{r(s,1)}{s} - \frac{r(1,t)}{t} + r(1,1)\right)dsdt. \tag{5.1}$$

Proof of Lemma 5.1. It is similar to that of Corollary A.2 in [7] where we restrict ourselves to second-order conditions.

Recall that for $\alpha \geq 1$,

$$M_k^{(\alpha)} = \frac{1}{k}\sum_{i=1}^{k}(\log X_{n-i+1,n} - \log X_{n-k,n})^\alpha,$$

Let $Q_n(t) = X_{n-[kt],n}$, $0 \leq t \leq 1$, where $[x]$ is the integer part of the value x. Thus $M_k^{(\alpha)}$ can be rewritten as

$$M_k^{(\alpha)} = \int_0^1\left(\log\frac{Q_n(t)}{Q_n(1)}\right)^\alpha dt.$$

Assume that the assumptions of Lemma 5.1 hold. Following the same Skorohod construction as in Proposition 1 in [4], it follows that, for given $\alpha,\ \delta > 0$, there exist a function $\widetilde{A} \sim A$, and a centered Gaussian process $(W(t))_{t\in[0,1]}$ with covariance function r, such that, as $n \to \infty$:

$$\sup_{t\in[0,1]} t^{\frac{1}{2}+\epsilon}\left|\sqrt{k}\left(\log\frac{Q_n(t)}{U(\frac{n}{k})} + \gamma\log t\right) - \gamma t^{-1}W(t) - \sqrt{k}\widetilde{A}(\frac{n}{k})\frac{t^{-\rho}-1}{\rho}\right| \to 0,\ a.s. \tag{5.2}$$

Using the fact that

$$\log\frac{Q_n(t)}{Q_n(1)} = \log\frac{Q_n(t)}{U(\frac{n}{k})} - \log\frac{Q_n(1)}{U(\frac{n}{k})},$$

we get as $n \to \infty$:

$$\sup_{t\in(0,1]} t^{1/2+\delta}\left|\sqrt{k}\left(\log\frac{Q_n(t)}{Q_n(1)} - \gamma(-\log t)\right) - \gamma\left(t^{-1}W(t) - W(1)\right) - \sqrt{k}\tilde{A}(n/k)\frac{t^{-\rho}-1}{\rho}\right| \to 0,\ a.s. \tag{5.3}$$

This means for all $t \in (0,1]$

$$\left(\log \frac{Q_n(t)}{Q_n(1)}\right)^{\alpha} \stackrel{d}{=} \gamma^{\alpha}(-\log t)^{\alpha}\Bigg[1+\frac{(-\log t)^{-1}}{\sqrt{k}}\left(t^{-1}W(t)-W(1)\right) \\ +\gamma^{-1}(-\log t)^{-1}\tilde{A}(n/k)\frac{t^{-\rho}-1}{\rho}+o(k^{-1/2})t^{-1/2-\delta}\Bigg]^{\alpha}.$$

And from the relation $(1+x)^{\alpha}=1+\alpha x+\frac{\alpha(\alpha-1)}{2}x^2+o(x^2), o(x^2)\longrightarrow 0$ when $x\to 0$, we get

$$\sup_{t\in(0,1]} t^{1/2+\delta}\Bigg|\sqrt{k}\left(\log\frac{Q_n(t)}{Q_n(1)}\right)^{\alpha}-\gamma^{\alpha}(-\log t)^{\alpha}-\alpha\gamma^{\alpha}(-\log t)^{\alpha-1}\left(t^{-1}W(t)-W(1)\right) \\ -\sqrt{k}\tilde{A}(n/k)\alpha\gamma^{\alpha-1}(-\log t)^{\alpha-1}\frac{t^{-\rho}-1}{\rho}\Bigg|\longrightarrow 0,\ a.s. \tag{5.4}$$

Without losing generality, some terms tend to 0, as $n\to\infty$. In fact, we have $\sup_{t\in(0,1]} t^{1/2+\epsilon}t^{-1}|W(t)| = O(1)$ a.s. and $\tilde{A}(n/k)\to 0$ as $n\to\infty$.

By using $\delta<1/2$ in (5.4), we can take the integral of $\left(\log\frac{Q_n(t)}{Q_n(1)}\right)^{\alpha}$ on $(0,1]$ and use the fact that $\Gamma(a+1)=a\Gamma(a)$ and $\int_0^1(-\log t)^{a-1}t^{-b}dt=\frac{\Gamma(a)}{(1-b)^a}$ for $b<1$ to obtain the result in the Lemma 5.1. The random term is obtained by taking

$$P^{(\alpha)}=\int_0^1(-\log t)^{\alpha-1}\left(t^{-1}W(t)-W(1)\right)dt.$$

and is normally distributed with mean zero and covariance $c_{\alpha,\alpha'}$. ■

Proof of Theorem 2.1: Clearly, we have $X_{i,n}\stackrel{d}{=}U(Y_{i,n})$, $i=1,...;n$ where each Y_i's follows a standard Pareto distribution, we obtain that $(Y_1,Y_2,...)$ is a stationary β-mixing series satisfying the regularity conditions.

This is a direct consequence of $Y_i=1/(1-F(X_i))$. In what follows, we shall use the following relation for $j>i$, $\frac{Y_{j,n}}{Y_{i,n}}\stackrel{d}{=}Y_{j-i,n-i}$. Thus, remarking that $Q_n(t)=X_{n-[kt],n}\stackrel{d}{=}U(Y_{n-[kt],n})$ and using the expression in (2.4), we have

$$\begin{aligned}\widehat{\gamma}_k^{BC}(\sigma,a) &\stackrel{d}{=} \frac{1}{k\ \sigma}\sum_{i=1}^{k}\log\left(\frac{U^{\sigma}(Y_{n-i+1,n})+a}{U^{\sigma}(Y_{n-k,n})+a}\right)\\ &\stackrel{d}{=} \frac{1}{k\ \sigma}\sum_{i=1}^{k}\log\left(\frac{U(Y_{n-i+1,n})}{U(Y_{n-k,n})}\right)^{\sigma}+\frac{1}{k\ \sigma}\sum_{i=1}^{k}\log\left(\frac{1+a/U^{\sigma}(Y_{n-i+1,n})}{1+a/U^{\sigma}(Y_{n-k,n})}\right).\end{aligned}$$

This implies that

$$\widehat{\gamma}_k^{BC}(\sigma,a)\stackrel{d}{=}\widehat{\gamma}_k^{H}+\frac{1}{k\ \sigma}\sum_{i=1}^{k}\left(\frac{a}{U^{\sigma}(Y_{n-i+1,n}}-\frac{a}{U^{\sigma}(Y_{n-k,n})}\right).$$

Which can be rewritten as

$$\widehat{\gamma}_k^{BC}(\sigma,a)\stackrel{d}{=}\widehat{\gamma}_k^{H}+\frac{1}{k\ \sigma}\sum_{i=1}^{k}\frac{a}{U^{\sigma}(Y_{n-k,n})}\left(-1+\left(\frac{U(Y_{n-k,n})}{U(Y_{n-i+1,n})}\right)^{\sigma}\right). \tag{5.5}$$

From (5.2), we infer that for $n\to\infty$,

$$Q_n(1)\stackrel{d}{=}U\left(\frac{n}{k}\right)\exp\left(\frac{1}{\sqrt{k}}\gamma W(1)+o_{\mathbb{P}}\left(\frac{1}{\sqrt{k}}\right)\right).$$

Observing that $Q_n(t)=X_{n-[kt],n}\stackrel{d}{=}U(Y_{n-[kt],n})$ then

$$U(Y_{n-k,n})\stackrel{d}{=}U\left(\frac{n}{k}\right)\exp\left(\frac{1}{\sqrt{k}}\gamma W(1)+o_{\mathbb{P}}\left(\frac{1}{\sqrt{k}}\right)\right). \tag{5.6}$$

This implies that

$$U^{-\sigma}(Y_{n-k,n}) \overset{d}{=} U^{-\sigma}\left(\frac{n}{k}\right)\exp\left(\frac{-\sigma}{\sqrt{k}}\gamma W(1)+o_{\mathbb{P}}\left(\frac{1}{\sqrt{k}}\right)\right).$$

Moreover, under Hall classes, we have from (1.8), $U(t)\sim ct^{\gamma}$ as $t\to\infty$, this leads to

$$U^{-\sigma}\left(\frac{n}{k}\right)\sim c^{-\sigma}\left(\frac{n}{k}\right)^{-\sigma\gamma}, \quad \text{as} \quad n\to\infty. \tag{5.7}$$

Since $Y_{n-i+1,n}/Y_{n-k,n}\overset{d}{=}Y_{k-i+1,k}$ for $i=1,...,k$, which is also distributed from a standard Pareto distribution. Now, we have $Y_{n-k,n}\to\infty$, as $n\to\infty$, then using the regular varying assumption in (1.3), we get for all large values of n,

$$\frac{U(Y_{n-k,n})}{U(Y_{n-i+1,n})}\overset{d}{=}\left(Y_{k-i+1,k}\right)^{-\gamma}(1+o_{\mathbb{P}}(1)). \tag{5.8}$$

Thus, the combination of (5.5), (5.6), (5.7) and (5.8), ensures that for all n large enough

$$\widehat{\gamma}_k^{BC}(\sigma,a)\overset{d}{=}\widehat{\gamma}_k^{H}+\frac{a}{\sigma c^{\sigma}}\left(\frac{n}{k}\right)^{-\sigma\gamma}\exp\left(\frac{-\sigma}{\sqrt{k}}\gamma W(1)+o_{\mathbb{P}}\left(\frac{1}{\sqrt{k}}\right)\right)\left(-1+\frac{1}{k}\sum_{i=1}^{k}\left(Y_{k-i+1,k}\right)^{-\sigma\gamma}\right).$$

Moreover, we have $-\sigma\gamma=\rho<0$, then by the large law number $\frac{1}{k}\sum_{i=1}^{k}\left(Y_{k-i+1,k}\right)^{-\sigma\gamma}\overset{\mathbb{P}}{\to}\mathbb{E}(Y_1^{-\sigma\gamma})=\frac{1}{1+\sigma\gamma}$, as $n\to\infty$. It comes that,

$$\widehat{\gamma}_k^{BC}(\sigma,a)\overset{d}{=}\widehat{\gamma}_k^{H}-\frac{a}{c^{\sigma}}\left(\frac{n}{k}\right)^{-\sigma\gamma}\frac{\gamma}{1+\sigma\gamma}\exp\left(\frac{-\sigma}{\sqrt{k}}\gamma W(1)+o_{\mathbb{P}}\left(\frac{1}{\sqrt{k}}\right)\right).$$

In the other hand, observing that $A(\frac{n}{k})=\eta\gamma(\frac{n}{k})^{\rho}$, then we have for all n large enough

$$\widehat{\gamma}_k^{BC}(\sigma,a)\overset{d}{=}\widehat{\gamma}_k^{H}-\frac{A(\frac{n}{k})}{1-\rho}\frac{a}{\eta c^{\sigma}}\exp\left(\frac{\rho}{\sqrt{k}}W(1)+o_{\mathbb{P}}\left(\frac{1}{\sqrt{k}}\right)\right).$$

The application of Lemma 5.1, ensures that

$$\widehat{\gamma}_k^{H}=M_k^{(1)}\overset{d}{=}\gamma+\frac{\gamma}{\sqrt{k}}P^{(1)}+\frac{A\left(\frac{n}{k}\right)}{1-\rho}\left(1+o_{\mathbb{P}}(1)\right). \tag{5.9}$$

Finally, since $W(1)$ is bounded, this leads for all n large enough to

$$\widehat{\gamma}_k^{BC}(\sigma,a)\overset{d}{=}\gamma+\frac{\gamma}{\sqrt{k}}P^{(1)}+\frac{A\left(\frac{n}{k}\right)}{1-\rho}\left(1-\frac{a}{\eta c^{\sigma}}\right)+o_{\mathbb{P}}\left(\frac{1}{\sqrt{k}}\right),$$

with $P^{(1)}=\int_0^1\left(t^{-1}W(t)-W(1)\right)dt$ and its variance is $c_{1;1}:=Cov(P^{(1)},P^{(1)})=r(1,1)$. The Theorem 2.1 is then proved. ■

Proof of Proposition 5.4.

Under assumptions, we have from (5.9), $\widehat{\gamma}_k^{H}\overset{\mathbb{P}}{\to}\gamma$, as $n\to\infty$. According to [22], for a given intermediate sequences $k_\rho:=k_\rho(n)$ satisfy the conditions in (2.7) and such that $\sqrt{k_\rho}A(n/k_\rho)\to\infty$, as $n\to\infty$, then the estimator $\widehat{\rho}_{k_\rho}$ defined in (2.9) is consistent in probability to the second order parameter ρ as $n\to\infty$. Since $\widehat{\sigma}_{k,k_\rho}=-\widehat{\rho}_{k,k_\rho}/\widehat{\gamma}_k^{H}$, this implies that $\widehat{\sigma}_{k,k_\rho}\overset{\mathbb{P}}{\to}\sigma=-\rho/\gamma$.

In other hand, for Hall-Welsh class of models in Equation (1.8), we have for all n large enough, $A(n/k)\approx\gamma\eta\left(\frac{n}{k}\right)^{\rho}$, and further from Lemma 5.1:

$$M_k^{(2)}=2\gamma^2+\frac{2\gamma^2}{\sqrt{k}}P^{(2)}+2\gamma^2\eta\left(\frac{n}{k}\right)^{\rho}\frac{1}{\rho}\left(\frac{1}{(1-\rho)^2}-1\right)(1+o_{\mathbb{P}}(1)),$$

and

$$\left(M_k^{(1)}\right)^2=\gamma^2\left(1+\frac{P^{(1)}}{\sqrt{k}}+\frac{\eta\left(\frac{n}{k}\right)^{\rho}}{1-\rho}(1+o_{\mathbb{P}}(1))\right)^2.$$

Using Taylor's expansion, we get as $n \to \infty$,

$$\left(M_k^{(1)}\right)^2 = \gamma^2 \left(1 + \frac{2P^{(1)}}{\sqrt{k}} + \frac{2\eta \left(\frac{n}{k}\right)^\rho}{1-\rho}(1 + o_{\mathbb{P}}(1))\right).$$

Hence

$$M_k^{(2)} - 2(M_k^{(1)})^2 \overset{d}{=} 2\gamma^2 k^{-1/2}\left(P^{(2)} - 2P^{(1)}\right) + 2\eta\gamma^2 \left(\frac{n}{k}\right)^\rho \frac{\rho}{(1-\rho)^2}(1 + o_{\mathbb{P}}(1)). \tag{5.10}$$

Therefore

$$\frac{(1-\rho)^2\left(M_k^{(2)} - 2\left(M_k^{(1)}\right)^2\right)}{2\gamma^2\left(\frac{n}{k}\right)^\rho \rho} \xrightarrow{\mathbb{P}} \eta, \text{ as } n \to \infty.$$

And this leads to the estimation of η as

$$\widehat{\eta}_{k,k_\rho} = \frac{\left(1-\widehat{\rho}_{k_\rho}\right)^2\left(M_k^{(2)} - 2\left(M_k^{(1)}\right)^2\right)}{2\left(M_k^{(1)}\right)^2\left(\frac{n}{k}\right)^{\widehat{\rho}_{k_\rho}}\widehat{\rho}_{k_\rho}}.$$

In addition, if $k_\rho := k_\rho(n)$ satisfies the condition (2.7) and $\sqrt{k_\rho}A(n/k_\rho) \to \infty$, as $n \to \infty$, we have from [7], $\widehat{\rho}_{k_\rho} \xrightarrow{\mathbb{P}} \rho$, which implies that $\widehat{\eta}_{k,k_\rho} \xrightarrow{\mathbb{P}} \eta$, as $n \to \infty$. ■

Proof of Theorem 2.2:

We follow the same approach as in the proof of Theorem 2.1. Firstly, as in (5.5), we have

$$\widehat{\gamma}_{k,k_\rho}^{BC} \overset{d}{=} \widehat{\gamma}_k^H + \frac{1}{k\widehat{\sigma}_{k,k_\rho}} \sum_{i=1}^{k} \frac{\widehat{\eta}_{k,k_\rho}}{(U(Y_{n-k,n}))^{\widehat{\sigma}_{k,k_\rho}}}\left(-1 + \left(\frac{U(Y_{n-k,n})}{U(Y_{n-i+1,n})}\right)^{\widehat{\sigma}_{k,k_\rho}}\right)$$

The application of Corollary A.2 in [7], ensures that

$$\widehat{\gamma}_k^H = \gamma + \frac{\gamma}{\sqrt{k}}P_1^{(1)} + \frac{1}{1-\rho}A\left(\frac{n}{k}\right) + o_{\mathbb{P}}\left(\frac{1}{\sqrt{k}}\right)$$

then we have

$$\widehat{\gamma}_{k,k_\rho}^{BC} \overset{d}{=} \gamma + \frac{\gamma}{\sqrt{k}}P_1^{(1)} + \frac{1}{1-\rho}A\left(\frac{n}{k}\right) + \frac{\widehat{\eta}_{k,k_\rho}}{U^{\widehat{\sigma}_{k,k_\rho}}(Y_{n-k,n})}\frac{1}{k\widehat{\sigma}_{k,k_\rho}}\sum_{i=1}^{k}\left(-1 + \left(\frac{U(Y_{n-k,n})}{U(Y_{n-i+1,n})}\right)^{\widehat{\sigma}_{k,k_\rho}}\right) + o_{\mathbb{P}}\left(\frac{1}{\sqrt{k}}\right)$$

$$\left(\frac{U(Y_{n-k,n})}{U(Y_{n-i+1,n})}\right)^{\widehat{\sigma}_{k,k_\rho}} = \left(Y_{k-i+1,k}\right)^{-\widehat{\sigma}_{k,k_\rho}\gamma} + o_{\mathbb{P}}(1)$$

$$\widehat{\gamma}_{k,k_\rho}^{BC} \overset{d}{=} \gamma + \frac{\gamma}{\sqrt{k}}P_1^{(1)} + \frac{1}{1-\rho}A\left(\frac{n}{k}\right) + \frac{\widehat{\eta}_{k,k_\rho}}{U^{\widehat{\sigma}_{k,k_\rho}}(Y_{n-k,n})}\frac{-\gamma}{1+\widehat{\sigma}_{k,k_\rho}\gamma} + o_{\mathbb{P}}\left(\frac{1}{\sqrt{k}}\right)$$

It follows from (5.6)

$$U^{\widehat{\sigma}_{k,k_\rho}}(Y_{n-k,n}) = U^{\widehat{\sigma}_{k,k_\rho}}\left(\frac{n}{k}\right)\exp\left(\frac{\widehat{\sigma}_{k,k_\rho}}{\sqrt{k}}\gamma W(1) + o_{\mathbb{P}}\left(\frac{1}{\sqrt{k}}\right)\right).$$

Moreover, we have 1.8 under Hall classes ensures that $U(t) \sim ct^\gamma$ then

$$U^{\widehat{\sigma}_{k,k_\rho}}\left(\frac{n}{k}\right) \sim \left(\frac{n}{k}\right)^{\widehat{\sigma}_{k,k_\rho}\gamma}.$$

Hence

$$\widehat{\eta}_{k,k_\rho}U^{-\widehat{\sigma}_{k,k_\rho}}(Y_{n-k,n}) = \widehat{\eta}_{k,k_\rho}\left(\frac{n}{k}\right)^{-\widehat{\sigma}_{k,k_\rho}\gamma}\exp\left(\frac{-\widehat{\sigma}_{k,k_\rho}}{\sqrt{k}}\gamma W(1) + o_{\mathbb{P}}\left(\frac{1}{\sqrt{k}}\right)\right).$$

Under Hall classes we have

$$\widehat{A}\left(\frac{n}{k}\right) = \widehat{\eta}_{k,k_\rho}\left(\frac{n}{k}\right)^{\widehat{\rho}_{k_\rho}}\widehat{\gamma}^H \quad \text{then} \quad \left(\frac{n}{k}\right)^{\widehat{\rho}_{k_\rho}} = \frac{\widehat{A}\left(\frac{n}{k}\right)}{\widehat{\eta}_{k,k_\rho}\widehat{\gamma}^H}$$

$$\widehat{\eta}_{k,k_\rho}\left(\frac{n}{k}\right)^{-\widehat{\sigma}_{k,k_\rho}\gamma}=\widehat{\eta}_{k,k_\rho}\left[\left(\frac{n}{k}\right)^{\widehat{\rho}_{k_\rho}}\right]^{\frac{\gamma}{\widehat{\gamma}^H}}=\widehat{\eta}_{k,k_\rho}\left[\frac{\widehat{A}\left(\frac{n}{k}\right)}{\widehat{\eta}_{k,k_\rho}\widehat{\gamma}^H}\right]^{\frac{\gamma}{\widehat{\gamma}^H}}$$

$$\widehat{\gamma}_k^{BC}(\widehat{\sigma}_{k,k_\rho},\widehat{\eta}_{k,k_\rho})\overset{d}{=}\gamma+\frac{\gamma}{\sqrt{k}}P_1^{(1)}+\frac{1}{1-\rho}A\left(\frac{n}{k}\right)+\widehat{\eta}_{k,k_\rho}\left[\frac{\widehat{A}\left(\frac{n}{k}\right)}{\widehat{\eta}_{k,k_\rho}\widehat{\gamma}^H}\right]^{\frac{\gamma}{\widehat{\gamma}^H}}\frac{-\gamma}{1+\widehat{\sigma}_{k,k_\rho}\gamma}\exp\left(-\frac{\widehat{\sigma}_{k,k_\rho}}{\sqrt{k}}\gamma W(1)+o_{\mathbb{P}}\left(\frac{1}{\sqrt{k}}\right)\right)$$
$$+o_{\mathbb{P}}\left(\frac{1}{\sqrt{k}}\right)$$

since $\widehat{\sigma}_{k,k_\rho}\overset{\mathbb{P}}{\to}\sigma$, $\frac{\widehat{A}\left(\frac{n}{k}\right)}{A\left(\frac{n}{k}\right)}\overset{\mathbb{P}}{\to}1$, $\widehat{\eta}_{k,k_\rho}\overset{\mathbb{P}}{\to}\eta$, $\widehat{\gamma}^H\overset{\mathbb{P}}{\to}\gamma$ and $\frac{\gamma}{\widehat{\gamma}^H}\overset{\mathbb{P}}{\to}1$ then

$$\sqrt{k}\left(\widehat{\gamma}_k^{BC}(\widehat{\sigma}_{k,k_\rho},\widehat{\eta}_{k,k_\rho})-\gamma\right)\overset{d}{=}\gamma P_1^{(1)}+o_{\mathbb{P}}(1)$$

then we obtain the result. ■

Proof of theorem 3.1:

First of all, we consider the equation 3.3

$$\widehat{x}_{k,k_\rho}^{BC}(p):=X_{n-k,n}\left(\frac{k}{np}\right)^{\widehat{x}_{k,k_\rho}^{BC}(p)}\exp\left(\widehat{\gamma}_k^H\widehat{\eta}_{k,k_\rho}\left(\frac{n}{k}\right)^{\widehat{\rho}_{k_\rho}}\frac{\left(\frac{k}{np}\right)^{\widehat{\rho}_{k_\rho}}-1}{\widehat{\rho}_{k_\rho}}\right).$$

$$\begin{aligned}\frac{\sqrt{k}}{\log(\frac{k}{np})}\log\frac{\widehat{x}_{k,k_\rho}^{BC}(p)}{x(p)}&=\frac{\sqrt{k}}{\log(\frac{k}{np})}\left[\log X_{n-k,n}+\widehat{\gamma}_{k,k_\rho}^{BC}\log\left(\frac{k}{np}\right)+\widehat{\eta}_{k,k_\rho}\widehat{\gamma}_k^H\left(\frac{n}{k}\right)^{\widehat{\rho}_{k_\rho}}\frac{\left(\frac{k}{np}\right)^{\widehat{\rho}_{k_\rho}}-1}{\widehat{\rho}_{k_\rho}}-\log x(p)\right]\\
&=\sqrt{k}(\widehat{\gamma}_{k,k_\rho}^{BC}-\gamma)+\frac{\sqrt{k}}{\log(\frac{k}{np})}\log\frac{Q_n(1)}{U\left(\frac{n}{k}\right)}-\frac{\sqrt{k}}{\log(\frac{k}{np})}\left[\log\frac{U\left(\frac{1}{p}\right)}{U\left(\frac{n}{k}\right)}-\gamma\log(\frac{k}{np})\right]\\
&+\frac{\sqrt{k}}{\log(\frac{k}{np})}\widehat{\eta}_{k,k_\rho}\widehat{\gamma}_k^H\left(\frac{n}{k}\right)^{\widehat{\rho}_{k_\rho}}\frac{\left(\frac{k}{np}\right)^{\widehat{\rho}_{k_\rho}}-1}{\widehat{\rho}_{k_\rho}}\\
&=\sqrt{k}(\widehat{\gamma}_{k,k_\rho}^{BC}-\gamma)+\frac{\sqrt{k}}{\log(\frac{k}{np})}\log\frac{Q_n(1)}{U\left(\frac{n}{k}\right)}-\frac{\sqrt{k}}{\log(\frac{k}{np})}\overline{A}\left(\frac{n}{k}\right)\frac{\left(\frac{k}{np}\right)^{\rho}-1}{\rho}\\
&-\frac{\sqrt{k}}{\log(k/(np))}\overline{A}\left(\frac{n}{k}\right)\left[\frac{\log U\left(\frac{1}{p}\right)-\log U\left(\frac{n}{k}\right)-\gamma\log\frac{k}{np}}{\overline{A}\left(\frac{n}{k}\right)}-\frac{\left(\frac{k}{np}\right)^{\rho}-1}{\rho}\right]\\
&+\frac{\sqrt{k}}{\log(\frac{k}{np})}\widehat{\eta}_{k,k_\rho}\widehat{\gamma}_k^H\left(\frac{n}{k}\right)^{\widehat{\rho}_{k_\rho}}\frac{\left(\frac{k}{np}\right)^{\widehat{\rho}_{k_\rho}}-1}{\widehat{\rho}_{k_\rho}}\\
&=I_1+I_2-I_3-I_4+I_5\end{aligned}$$

$$I_1\xrightarrow{d}N\left(0,\gamma^2r(1,1)\right).$$

It was proved in [4] that I_2, I_3 et $I_4\to0$

Now, if we consider $I_5=\widehat{\eta}_{k,k_\rho}\widehat{\gamma}_k^H\left(\frac{n}{k}\right)^{\widehat{\rho}}\frac{\left(\frac{k}{np}\right)^{\widehat{\rho}_{k_\rho}}-1}{\widehat{\rho}_{k_\rho}}$

we have $\widehat{\eta}_{k,k_\rho}=\eta(1+o_{\mathbb{P}}(1))$, $\widehat{\gamma}_k^H=\gamma(1+o_{\mathbb{P}}(1))$, $\frac{\left(\frac{k}{np}\right)^{\widehat{\rho}_{k_\rho}}-1}{\widehat{\rho}_{k_\rho}}=\frac{-1}{\rho}(1+o_{\mathbb{P}}(1))$ and $\left(\frac{n}{k}\right)^{\widehat{\rho}_{k_\rho}}\to0$ then

$$I_5\to0$$

which conclude the proof ■

References

[1] Bali, T. G. (2003). The generalized extreme value distribution, *Econom. Lett.* **79** 423–427.

[2] Beirlant, J., Goegebeur, Y., Segers, J. and Teugels, J. (2004). *Statist. Extremes.*, Wiley.

[3] Caeiro, F., Gomes, M. I. and Pestana, D. (2005). Direct reduction of bias of the classical Hill estimator, *Revstat* **3** 111–136.

[4] Chavez-Demoulin, V. and Guillou, A. (2018). Extreme quantile estimation for β-mixing time series and applications. *Insurance Math. Econom.* **83** 59-74.

[5] de Haan, L. (1984), Slow Variation and Characterization of Domains of Attraction, Springer Netherlands, Dordrecht.

[6] de Haan, L. and Ferreira, A. (2006), Extreme Value Theory: An Introduction, Springer, New York.

[7] de Haan, L. (2016), Mercadier, C. and Zhou, C. (2016). Adapting extreme value statistics to financial time series: dealing with bias and serial dependence, *Finance Stoch.* **20** 321–354.

[8] Drees, H. (2000). Weighted approximations of tail processes for β-mixing random variables. *Ann. Appl.Probab.* **10** 1274–1301.

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

[10] Deme, E. H., Gardes, L. and Girard, S. (2013). On the estimation of the second order parameter for heavy-tailed distributions. *Statist. J.* **3** 277-299.

[11] Feuerverger, A. and Hall, P. (1999). Estimating a tail exponent by modeling departure from a pareto discutribution. *Ann. Stat.* **27** 760-781.

[12] Gomes, M., de Haan, L. and Peng, L. (2002), Semi-parametric estimation of the second order parameter in statistics of extremes, *Extremes.* **5** 387–414.

[13] Gomes, M. I. and Martins, M. and Neves, M. (2007). Improving second order reduced bias extreme value index estimator. *REVSTAT.* **5** 177-207.

[14] Hall, P. and Welsh, A. (1985), Adaptive estimates of parameters of regular variation, *Ann. Stat.* **13**, 331–341.

[15] Hall, P. (1982), On some simple estimates of an exponent of regular variation, *J. R. Stat. Soc. Ser. B Stat. Methodol.*, **44** 37–42.

[16] Hsing, T. (1991), On tail index estimation using dependent data, *Ann. Stat.* **19** 1547–1569.

[17] Hill, B. (1975). A simple general approach to inference about the tail of a distribution. *Ann. Stat.* **3** 1163–1174.

[18] Wadsworth, J. L., Tawn, J. A. and Jonathan, P. (2010), Accounting for choice of measurement scale in extreme value modeling, *Ann. Appl. Stat.* **4** 1558–1578.

[19] Leadbetter, M. R., Lindgren, G. and Rootzen, H. (1983). Extremes and Related Properties of Random Sequences and Processes. Berlin: Springer-Verlag.

[20] Reiss, R. D., Thomas, M.(2007). Analysis of Extreme Values with Applications to Insurance, Finance, Hydrology and Other Fields. 3rd ed. Birkhauser.

[21] Lígia Henriques-Rodrigues and M. I. Gomes (2022). Box-Cox Transformations and Bias Reduction in Extreme Value Theory. *Comput. Math. Methods Med.* **2022**, Article ID 3854763, 15 pages, https://doi.org/10.1155/2022/3854763.

[22] Teugels, J. L. and Vanroelen, G. (2004). Box-Cox transformations and heavy-tailed distributions, *Chinese J. Appl. Probab. Statist.* **41** 213–227.

[23] Weissman, I. (1978). Estimation of parameters and large quantiles based on the k-largest observations.

J. Amer. Statist. Assoc. **73** 812-815.