Electronic Journal of Probability

Vol. 13 (2008), Paper no. 32, pages 961–979.

Journal URL
http://www.math.washington.edu/~ejpecp/

Sums of extreme values of subordinated long-range dependent sequences: moving averages with finite variance

Rafał Kulik*

University of Ottawa
and
University of Wrocław

Abstract

In this paper we study the limiting behavior of sums of extreme values of long range dependent sequences defined as functionals of linear processes with finite variance. If the number of extremes in a sum is large enough, we obtain asymptotic normality, however, the scaling factor is relatively bigger than in the i.i.d case, meaning that the maximal terms have relatively smaller contribution to the whole sum. Also, it is possible for a particular choice of a model, that the scaling need not to depend on the tail index of the underlying marginal distribution, as it is well-known to be so in the i.i.d. situation. Furthermore, subordination may change the asymptotic properties of sums of extremes .

Key words: sample quantiles, linear processes, empirical processes, long range dependence, sums of extremes, trimmed sums.

AMS 2000 Subject Classification: Primary 60F05; Secondary: 60G70.

Submitted to EJP on April 16 2007, final version accepted May 16, 2008.

*Department of Mathematics and Statistics, University of Ottawa, K1N 6N5 Ottawa, ON, Canada, email: rkulik@uottawa.ca and Mathematical Institute, Wrocław University, Pl. Grunwaldzki 2/4, 50-384 Wrocław, Poland.
Research supported in part by NSERC Canada Discovery Grants of Miklós Csörgő, Donald Dawson and Barbara Szyszkowicz at Carleton University

1 Introduction

Let $\{\epsilon_i, -\infty < i < \infty\}$ be a centered sequence of i.i.d. random variables. Consider the class of stationary linear processes

$$X_i = \sum_{k=0}^{\infty} c_k \epsilon_{i-k}, \quad i \geq 1. \tag{1}$$

We assume that the sequence c_k, $k \geq 0$, is regularly varying with index $-\beta$, $\beta \in (1/2, 1)$. This means that $c_k \sim k^{-\beta} L_0(k)$ as $k \to \infty$, where L_0 is slowly varying at infinity. We shall refer to all such models as long range dependent (LRD) linear processes. In particular, if the variance of ϵ_1 exists (which is assumed throughout the whole paper), then the covariances $\rho_k := \mathrm{E} X_0 X_k$ decay at the hyperbolic rate, $\rho_k = k^{-(2\beta-1)} L(k)$, where $\lim_{k\to\infty} L(k)/L_0^2(k) = B(2\beta-1, 1-\beta)$ and $B(\cdot,\cdot)$ is the beta-function. Consequently, the covariances are not summable (cf. [11]).

Assume that X_1 has a continuous distribution function F. For $y \in (0,1)$ define $Q(y) = \inf\{x : F(x) \geq y\} = \inf\{x : F(x) = y\}$, the corresponding quantile function, which is assumed to be differentiable. Given the ordered sample $X_{1:n} \leq \cdots \leq X_{n:n}$ of $X_1, \ldots, X_n$, let $F_n(x) = n^{-1} \sum_{i=1}^n 1_{\{X_i \leq x\}}$ be the empirical distribution function and $Q_n(\cdot)$ be the corresponding left-continuous sample quantile function, i.e. $Q_n(y) = X_{k:n}$ for $\frac{k-1}{n} < y \leq \frac{k}{n}$. Define $U_i = F(X_i)$ and $E_n(x) = n^{-1} \sum_{i=1}^n 1_{\{U_i \leq x\}}$, the associated uniform empirical distribution function. Denote by $U_n(\cdot)$ the corresponding uniform sample quantile function.

Assume that $E\epsilon_1^2 < \infty$. Let r be a positive integer and define

$$Y_{n,r} = \sum_{i=1}^{n} \sum_{1 \leq j_1 < \cdots \leq j_r < \infty} \prod_{s=1}^{r} c_{j_s} \epsilon_{i-j_s}, \qquad n \geq 1,$$

so that $Y_{n,0} = n$, and $Y_{n,1} = \sum_{i=1}^n X_i$. If $p < (2\beta-1)^{-1}$, then

$$\sigma_{n,p}^2 := \mathrm{Var}(Y_{n,p}) \sim n^{2-p(2\beta-1)} L_0^{2p}(n). \tag{2}$$

Define now the general empirical, the uniform empirical, the general quantile and the uniform quantile processes respectively as follows:

$$\beta_n(x) = \sigma_{n,1}^{-1} n(F_n(x) - F(x)), \qquad x \in \mathbb{R},$$

$$\alpha_n(y) = \sigma_{n,1}^{-1} n(E_n(y) - y), \qquad y \in (0,1),$$

$$q_n(y) = \sigma_{n,1}^{-1} n(Q(y) - Q_n(y)), \qquad y \in (0,1),$$

$$u_n(y) = \sigma_{n,1}^{-1} n(y - U_n(y)), \qquad y \in (0,1).$$

The aim of this paper is to study the asymptotic behavior of trimmed sums based on the ordered sample $X_{1:n} \leq \cdots \leq X_{n:n}$ coming from the long range dependent sequence defined by (1).

Let $T_n(m,k) = \sum_{j=m+1}^{n-k} X_{j:n}$ and note that (see below for a convention concerning integrals)

$$T_n(m,k) = n \int_{m/n}^{1-k/n} Q_n(y) dy. \tag{3}$$

Ho and Hsing observed in [13] that, under appropriate conditions on F, as $n \to \infty$,

$$\sup_{y\in[y_0,y_1]} \left| q_n(y) + \sigma_{n,1}^{-1} \sum_{i=1}^{n} X_i \right| = o_P(1), \tag{4}$$

where $0 < y_0 < y_1 < 1$. Equation (4) means that, in principle, *the quantile process can be approximated by partial sums, independently of* y. This observation, together with (3), yields the asymptotic normality of the trimmed sums in case of heavy trimming $m = m_n = [\delta_1 n]$, $k = k_n = [\delta_2 n]$, where $0 < \delta_1 < \delta_2 < 1$ and $[\cdot]$ is the integer part (see [13, Corollary 5.2] and [25]). This agrees with the i.i.d. situation (see [23]).

However, the representation (3) requires some additional assumptions on F. In order to avoid them, we may study asymptotics for the trimmed sums via the integrals of the form $\int \alpha_n(y)dQ(y)$. This approach was initiated in two beautiful papers by M. Csörgő, S. Csörgő, Horváth and Mason, [2], [3]. Then, S. Csörgő, Haeusler, Horváth and Mason took this route to provide the full description of the weak asymptotic behavior of the trimmed sums in the i.i.d. case. The list of the papers written by these authors on this particular topic is just about as long as this introduction. Therefore we refer to [7] for an extensive up-to-date discussion and a survey of results.

For the i.i.d. random variables the above mentioned authors approximated the uniform empirical process by an appropriately constructed Brownian bridge $B_0(\cdot)$ and then concluded asymptotic results for the integrals via those for $\int B_0(y)dQ(y)$. In the LRD case, we will use the reduction principle for empirical processes as studied in [11], [13], [15] or [24] (see Lemma 10 below). We can then use an approach that is similar to that of the above mentioned authors to establish asymptotic normality in case of moderate and heavy trimming with the scaling factor $\sigma_{n,1}^{-1}$, which is the same as for the whole partial sum. So, in this context the situation is similar to the i.i.d. case and for details we refer the reader to [16].

The most interesting phenomena, however, occur when one deals with the k_n-extreme sums, $\sum_{j=n-k_n+1}^{n} X_{j:n}$. If $F(0) = 0$ and $1 - F(x) = x^{-\alpha}$, $\alpha > 2$, then in the i.i.d situation we have

$$a_n \sum_{j=n-k_n+1}^{n} X_{j:n} - c_n \xrightarrow{\mathrm{d}} Z,$$

where the scaling factor is $a_n = \left(nk_n^{-1}\right)^{1/2-1/\alpha} n^{-1/2}$, c_n is a centering sequence and Z is a standard normal random variable (see [9]). In the LRD case we still obtain asymptotic normality. However, although the Ho and Hsing result (4) does not say anything about the behavior of the quantile process in the neighborhood of 0 and 1, the somewhat imprecise statement that *the quantile process can be approximated by partial sums, independently of* y suggests that

- a required scaling factor would not depend on the tail index α.

Indeed, we will show in Theorem 1 that, under some conditions on k_n, the appropriate scaling in case $1 - F(x) = x^{-\alpha}$ is $(nk_n^{-1})\sigma_{n,1}^{-1}$. Removing the scaling for the whole sums ($n^{-1/2}$ and $\sigma_{n,1}^{-1}$ in the i.i.d. and LRD cases, respectively), we also see that

- the scaling in the LRD situation is greater, meaning that the k_n-extreme sums contribute relatively *less* to the whole sum compared to the i.i.d situation. This also is quite intuitive. Since the dependence is very strong, it is very unlikely that we have few big observations, which is a typical case in the i.i.d. situation. Rather, if we have one big value, we have a lot of them.

One may ask, whether such phenomena are typical for all LRD sequences. Not likely. Define $Y_i = G(X_i)$, $i \ge 1$, with some real-valued measurable function G. In particular, taking $G = F_Y^{-1}F$ we may obtain a LRD sequence with the arbitrary marginal distribution function F_Y. Assume for a while that F, the distribution of X_1, is standard normal and that $q_n(\cdot)$ is the quantile process associated with the sequence $\{Y_i, i \ge 1\}$. Following [6] we observed in [4, Section 2.2] and [5] that $q_n(\cdot)$ is, up to a constant, approximated by $\phi(\Phi^{-1}(y))/f_Y(F_Y^{-1}(y))\sigma_{n,1}^{-1}\sum_{i=1}^n X_i$. Here, f_Y is the density of F_Y and ϕ, Φ are the standard normal density and distribution, respectively. In the non-subordinated case, $Y_i = X_i$, the factor $\phi(\Phi^{-1}(y))/f_Y(F_Y^{-1}(y))$ disappears. Nevertheless, from this discussion it should be clear that the limiting behavior of the extreme sums in the subordinated case $Y_i = G(X_i)$ is different, namely (see Theorem 1)

- the scaling depends on the marginal distributions of both X_i and Y_i.

In particular, if the distribution F of X_1 belongs to the maximal domain of attraction of the Fréchet distribution Φ_α, then though the distribution F_Y of Y_1 belongs to the maximal domain of attraction of the Gumbel distribution, the scaling factor depends on α. This cannot happen in the i.i.d. situation and, intuitively, it means that in the subordinated case *the long range dependent sequence* $\{X_i, i \ge 1\}$ *also contributes information to the asymptotic behavior of extreme sums.*

Moreover, we may have two LRD sequences $\{X_i, i \ge 1\}$, $\{Y_i, i \ge 1\}$, the first one as in (1), the second one defined by $Y_i = G(X_i')$ with a sequence $\{X_i', i \ge 1\}$ defined as in (1), with the same covariance, with the same marginals, but different behavior (i.e., the different scaling) of extremal terms.

It should be pointed out that the above mentioned phenomena for extremal sums of LRD sequences are valid if the number of extremes, k_n, is big enough. In Theorem 1 we have assumed, in particular, that $k_n = [n^\xi]$, $\xi > \beta$. A natural question arises, what happens if $\xi < \beta$. To answer this partially, we assume that $\{\epsilon_i, -\infty < i < \infty\}$ is an i.i.d. sequence of standard normal random variables. We observe that if $k_n = [n^\xi]$, $\xi < 2\beta - 1$, the sums of extremes grow at the same rate as in the corresponding i.i.d. case. However, we are not able to prove the asymptotic normality. We refer to Remark 4 for further discussion.

Of course, it would be desirable to obtain some information about limiting behavior not only of extreme sums, but for sample maxima as well. It should be pointed out that our method is not appropriate. This is still an open problem to derive limiting behavior of maxima in the model (1). A Gaussian case is covered in [17, Chapter 4]. In a different setting, the case of stationary stable processes generated by conservative flow, the problem is treated in [21].

We will use the following convention concerning integrals. If $-\infty < a < b < \infty$ and h, g are left-continuous and right-continuous functions, respectively, then

$$\int_a^b gdh = \int_{[a,b)} gdh \qquad \text{and} \qquad \int_a^b hdg = \int_{(a,b]} hdg,$$

whenever these integrals make sense as Lebesgue-Stjeltjes integrals. The integration by parts formula yields

$$\int_a^b gdh + \int_a^b hdg = h(b)g(b) - h(a)g(a).$$

We shall write $g \in RV_\alpha$ ($g \in SV$) if g is regularly varying at infinity with index α (slowly varying at infinity).

In what follows C will denote a generic constant which may be different at each of its appearances. Also, for any sequences a_n and b_n, we write $a_n \sim b_n$ if $\lim_{n\to\infty} a_n/b_n = 1$. Further, let $\ell(n)$ be a slowly varying function, possibly different at each place it appears. On the other hand, $L(\cdot)$, $L_0(\cdot)$, $L_1(\cdot)$, $L_1^*(\cdot)$, etc., are slowly varying functions of fixed form wherever they appear. Moreover, $g^{(k)}$ denotes the kth order derivative of a function g and Z is a standard normal random variable. For any stationary sequence $\{V_i, i \geq 1\}$, we will denote by V the random variable with the same distribution as V_1.

2 Statement of results

Let F_ϵ be the marginal distribution function of the centered i.i.d. sequence $\{\epsilon_i, -\infty < i < \infty\}$. Also, for a given integer p, the derivatives $F_\epsilon^{(1)}, \ldots, F_\epsilon^{(p+3)}$ of F_ϵ are assumed to be bounded and integrable. Note that these properties are inherited by the distribution function F of X_1 as well (cf. [13] or [24]). Furthermore, assume that $\mathrm{E}\epsilon_1^4 < \infty$. These conditions are needed to establish the reduction principle for the empirical process and will be assumed throughout the paper.

To study sums of k_n largest observations, we shall consider the following forms of F. For the statements below concerning regular variation and domain of attractions we refer to [10, Chapter 3], [12] or [14].

The first assumption is that the distribution F satisfies the following Von-Mises condition:

$$\lim_{x\to\infty} \frac{xf(x)}{1-F(x)} = \alpha > 0. \tag{5}$$

Using notation from [10], the condition (5) will be referred as $X \in MDA(\Phi_\alpha)$, since (5) implies that X belongs to the maximal domain of attraction of the Fréchet distribution with index α. Then

$$Q(1-y) = y^{-1/\alpha} L_1(y^{-1}), \quad \text{as } y \to 0, \tag{6}$$

and the density-quantile function $fQ(y) = f(Q(y))$ satisfies

$$fQ(1-y) = y^{1+1/\alpha} L_2(y^{-1}), \quad \text{as } y \to 0, \tag{7}$$

where $L_2(u) = \alpha(L_1(u))^{-1}$.

The second type of assumption is that F belongs to the maximal domain of attraction of the double exponential Gumbel distribution, written as $X \in MDA(\Lambda)$. Then the corresponding Von-Mises condition implies

$$\lim_{y\to 0} \frac{fQ(1-y)\int_{1-y}^{1}(1-u)/fQ(u)du}{y^2} = 1. \tag{8}$$

Thus, with $L_3(y^{-1}) = \left(y^{-1}\int_{1-y}^{1}(1-u)/fQ(u)du\right)^{-1}$ one has

$$fQ(1-y) = yL_3(y^{-1}),$$

and L_3 is slowly varying at infinity.

We note in passing that the conditions on f can be expressed (in certain cases) in terms of those for f_ϵ (see Remark 9).

To study the effect of subordination, we will consider the corresponding assumptions on F_Y and $Q_Y = F_Y^{-1}$, referred to later as $Y \in MDA(\Phi_{\alpha_0})$ and $Y \in MDA(\Lambda)$, respectively:

$$Q_Y(1-y) = y^{-1/\alpha_0}L_1^*(y^{-1}) \text{ and } f_YQ_Y(1-y) = y^{1+1/\alpha_0}L_2^*(y^{-1}), \quad \text{as } y \to 0, \tag{9}$$

with $L_2^*(u) = \alpha_0(L_1^*(u))^{-1}$, and

$$f_YQ_Y(1-y) = yL_3^*(y^{-1}),$$

where L_3^* is defined in the corresponding way as L_3.

Recall that $Q_n(y) = \inf\{x : F_n(x) \geq y\} = X_{k:n}$ if $\frac{k-1}{n} < y \leq \frac{k}{n}$. Let $T_n(m,k) = \sum_{j=m+1}^{n-k} Y_{j:n}$ and

$$\mu_n(m,k) = n\int_{m/n}^{1-k/n} Q_Y(y)dy.$$

The main result of this paper is the following theorem.

Theorem 1. *Let $G(x) = Q_Y(F(x))$. Let $k_n = [n^\xi]$, where $\xi \in (0,1)$ is such that*

$$\xi > \begin{cases} \frac{\beta+1/\alpha}{1+1/\alpha-1/\alpha_0}, & \text{if } X \in MDA(\Phi_\alpha),\ Y \in MDA(\Phi_{\alpha_0}), \quad (*) \\ \frac{\beta+1/\alpha}{1+1/\alpha}, & \text{if } X \in MDA(\Phi_\alpha),\ Y \in MDA(\Lambda), \quad (**) \\ \frac{\beta}{1-1/\alpha_0}, & \text{if } X \in MDA(\Lambda),\ Y \in MDA(\Phi_{\alpha_0}), \quad (***) \\ \beta, & \text{if } X \in MDA(\Lambda),\ Y \in MDA(\Lambda), \quad (****). \end{cases}$$

Assume that $\mathrm{E}Y < \infty$. *Let p be the smallest positive integer such that $(p+1)(2\beta-1) > 1$ and assume that for $r = 1, \ldots, p$,*

$$\int_{1/2}^{1} F^{(r)}(Q(y))dQ_Y(y) = \int_{1/2}^{1} \frac{F^{(r)}(Q(y))}{f_YQ_Y(y)}dy < \infty. \tag{10}$$

Let

$$A_n = \begin{cases} \left(\frac{n}{k_n}\right)^{1+1/\alpha-1/\alpha_0} L_{21}\left(\frac{n}{k_n}\right), & \text{if } X \in MDA(\Phi_\alpha), Y \in MDA(\Phi_{\alpha_0}), \\ \left(\frac{n}{k_n}\right)^{1+1/\alpha} L_{22}\left(\frac{n}{k_n}\right), & \text{if } X \in MDA(\Phi_\alpha), Y \in MDA(\Lambda), \\ \left(\frac{n}{k_n}\right)^{1-1/\alpha_0} L_{23}\left(\frac{n}{k_n}\right), & \text{if } X \in MDA(\Lambda), Y \in MDA(\Phi_{\alpha_0}), \\ \left(\frac{n}{k_n}\right) L_{24}\left(\frac{n}{k_n}\right), & \text{if } X \in MDA(\Lambda), Y \in MDA(\Lambda). \end{cases}$$

where $L_{21}, L_{22}, L_{23}, L_{24}$ *are slowly varying functions to be specified later on. Then*

$$A_n \sigma_{n,1}^{-1}\left(\sum_{j=n-k_n+1}^{n} Y_{j:n} - n\int_{1-k_n/n}^{1} Q_Y(y)dy\right) \xrightarrow{\text{d}} Z.$$

The corresponding cases concerning assumptions on X and Y will be referred as Case 1, Case 2, Case 3 and Case 4.

In the non-subordinated case we have the following result.

Corollary 2. *Under the conditions of Theorem* 1*, if either* $X \in MDA(\Phi_\alpha)$ *or* $X \in MDA(\Lambda)$*, then*

$$\left(\frac{n}{k_n}\right)\sigma_{n,1}^{-1}\left(\sum_{j=n-k_n+1}^{n} X_{j:n} - n\int_{1-k_n/n}^{1} Q(y)dy\right) \xrightarrow{\text{d}} Z.$$

In the subordinated case we have chosen to work with $G = Q_Y F$ to illustrate phenomena rather then deal with technicalities. One could work with general functions G, but then one would need to assume that G has the power rank 1 (see [13] for the definition). Otherwise the scaling $\sigma_{n,1}^{-1}$ is not correct. To see that $G(\cdot) = Q_Y F(\cdot)$ has the power rank 1, note that for $G_\infty(x) := \int_{-\infty}^{\infty} G(x+t)dF(t)$ we have

$$\frac{d}{dx}G_\infty(x) = \int_{-\infty}^{\infty} \frac{f(x+t)}{f_Y Q_Y F(x+t)} dF(t).$$

Substituting $x = 0$ and changing variables $y = F(t)$ we obtain

$$\frac{d}{dx}G_\infty(x)|_{x=0} = \int_0^1 \frac{fQ(y)}{f_Y Q_Y(y)} dy \neq 0.$$

Furthermore, we must assume that the distribution of $Y = G(X)$ belongs to the appropriate domain of attraction. For example, if $X \in MDA(\Phi_\alpha)$ and $Y_i = X_i^\rho$, ρ being a positive integer, then $Y \in MDA(\Phi_{\alpha/\rho})$, provided that the map $x \to x^\rho$ is increasing on $\mathbb{R}$. Otherwise, if for example $\rho = 2$, one needs to impose conditions not only on the right tail of X, but on the left one as well.

Nevertheless, to illustrate flexibility for the choice of G, let $G(x) = \log(x^+)^\alpha$, $\alpha > 0$. If $X \in MDA(\Phi_\alpha)$, then $Y = G(X)$ belongs to $MDA(\Lambda)$. Further, since $\mathrm{E}X = 0$, the quantile function $Q(u)$ of X must be positive for $u > u_0$ with some $u_0 \in (0,1)$. Since the map $x \to \log(x^+)^\alpha$ is increasing, $Q_Y(u) = Q_{\alpha\log(X^+)}(u) = \alpha \log Q(u)$ for $u > u_0$. Consequently, from Theorem 1 we obtain the following corollary.

Corollary 3. *If (**) holds and $X \in MDA(\Phi_\alpha)$, then*

$$A_n \sigma_{n,1}^{-1} \left(\sum_{j=n-k_n+1}^{n} \log(X_{j:n}^{+})^{\alpha} - n \int_{1-k_n/n}^{1} \log Q(y)dy \right) \xrightarrow{\mathrm{d}} Z,$$

where $A_n = \left(\frac{n}{k_n}\right)^{1+1/\alpha} L_{22}\left(\frac{n}{k_n}\right)$.

2.1 Remarks

Remark 4. To see what happens if the number of extremes is small, let us assume that $\{\epsilon_i, -\infty < i < \infty\}$ is an i.i.d. sequence of standard normal random variables, $\sum_{k=0}^{\infty} c_k^2 = 1$ and $\sup_{k \geq 1} |\rho_k| < 1$. Let $G(x) = Q_Y(\Phi(x))$ and $k_n = [n^{\xi}]$, where $\xi \in (0,1)$ is such that $\xi < 2\beta - 1$. Let

$$B_n = \begin{cases} \left(\frac{n}{k_n}\right)^{1/2-1/\alpha_0} \left(L_1^{*}\left(\frac{n}{k_n}\right)\right)^{-1} c_{\alpha_0}^{-1}, & \text{if } Y \in MDA(\Phi_{\alpha_0}), \\ \left(\frac{n}{k_n}\right)^{1/2}, & \text{if } Y \in MDA(\Lambda), \end{cases}$$

where $c_{\alpha_0}^2 = \frac{2(1/\alpha_0)^2}{(1-1/\alpha_0)(1-2/\alpha_0)}$. Then

$$B_n n^{-1/2} \left(\sum_{j=n-k_n+1}^{n} Y_{j:n} - n \int_{1-k_n/n}^{1} Q_Y(y)dy \right) = O_P(1).$$

The meaning of this is that for small ξ extremal sums grow at the same rate as in the corresponding i.i.d. situation. It follows from the Normal Comparison Lemma, see e.g. [17, p. 81].

This is not quite unexpected. In view of Theorem 4.3.3 in [17], asymptotic distribution of properly normalized maxima of LRD Gaussian sequences (with a covariance ρ_k decaying faster than $(\log k)^{-1}$) is the same as for the corresponding i.i.d. sequences (i.e., Gaussian sequences with the same marginals as X_i). In particular, large values of the sequence $\{X_i, i \geq 1\}$ do not cluster. We conjecture that $O_P(1)$ above can be replaced with an asymptotic normality. On the other hand, however, it is not clear if the similar statement will be valid if we assume that $\epsilon \in MDA(\Phi_\alpha)$ (which implies that $X \in MDA(\Phi_\alpha)$, see Remark 9 below). It is well known that if

$$\sum_{k=0}^{\infty} |c_k|^{\min(\alpha,1)} < \infty, \tag{11}$$

then large values cluster and the asymptotic distribution of $\max(X_1, \ldots, X_n)$ is different from the corresponding i.i.d. sequence (see [10] for more details). Thus, clustering of extremes should influence the asymptotic behavior of sums of extremes even in the short range dependent case (11).

Remark 5. Wu in his paper [24] considered a weighted approximation of empirical processes. In principle, using a weighted version of Lemma 10 below, one could expect to have weaker constraints on ξ in Theorem 1. However, this is not the case and with this method we cannot go beyond $\xi > \beta$. See Remark 14 below for more details.

Remark 6. From the beginning we assumed that $\mathrm{E}\epsilon_1^4 < \infty$, thus, in Cases 1 and 2 we have the requirement $\alpha \geq 4$ and this is the only constraint on this parameter. Condition $\mathrm{E}Y < \infty$ requires $\alpha_0 > 1$ in case of $Y \in MDA(\Phi_{\alpha_0})$. In view of (*), to be able to choose $\xi < 1$ we need to have $\alpha_0 > (1-\beta)^{-1} > 2$. The same restriction appears in Case 3.

Remark 7. The conditions $D_r := \int_{1/2}^1 F^{(r)}(Q(y))/f_Y Q_Y(y)dy < \infty$ are not restrictive at all, since they are fulfilled for most distributions with a regularly varying density-quantile function $fQ(1-y)$, for those we refer to [20]. Consider for example Case 1, and assume that the density f is non-increasing on some interval $[x_0, \infty)$. Then $F^{(r)}$ is regularly varying at infinity with index $r+\alpha$. Thus, for some $x_1 > x_0$

$$\int_{1/2\vee x_1}^1 F^{(r)}(Q(y))/f_Y Q_Y(y)dy = \int_{1/2\vee x_1}^1 (1-y)^{r/\alpha - 1/\alpha_0}\ell(y)dy < \infty$$

for all $r \geq 1$ provided $\alpha_0 > 1$. If, additionally, we impose the following *Csörgő-Révész-type conditions* (cf. [1, Theorem 3.2.1]):

(CsR1): f_Y exists on (a,b), where $-\infty \leq a < b \leq \infty$, $a = \sup\{x : F(x) = 0\}$, $b = \inf\{x : F(x) = 1\}$,
(CsR2): $f_Y(x) > 0$ for $x \in (a,b)$,

then in view of (CsR2) and the assumed boundness of derivatives $F^{(r)}(\cdot)$, the integral D_r is finite.

Remark 8. In the proof of Theorem 1 we have to work with both $Q(\cdot)$ and $fQ(\cdot)$. Therefore, we assumed the Von-Mises condition (5) since it implies both (6) and (7). If one assumes only (6), then (5) and, consequently, (7) hold, provided a monotonicity of f is assumed. Moreover, the von-Mises condition is natural, since the existence of the density f is explicitly assumed.

Remark 9. In some applications one knows the properties of f_ϵ, rather than of f.

Assume that $F_\epsilon \in MDA(\Phi_\alpha)$. Then also $F \in MDA(\Phi_\alpha)$ since

$$\lim_{x\to\infty} \frac{P(X_1 > x)}{P(|\epsilon| > x)} = const. \in (0, \infty).$$

For $\alpha > 2$ the above result is valid as long as $\sum_{j=0}^\infty c_j^2 < \infty$, in particular, in case of long range dependence (see [19] for details).

If ϵ_1 is normally distributed, then X too, thus in this special case both F_ϵ and F belong to $MDA(\Lambda)$.

Furthermore, as for the condition $\int_0^1 F^{(r)}(Q(y))dQ(y) < \infty$. Once again, if $F_\epsilon \in MDA(\Phi_\alpha)$ then the latter condition is fulfilled for both F_ϵ and F in view of the discussion in the previous remark.

3 Proofs

3.1 Consequences of the reduction principle

Let p be a positive integer and let

$$\begin{aligned} S_{n,p}(x) &= \sum_{i=1}^{n}(1_{\{X_i\le x\}} - F(x)) + \sum_{r=1}^{p}(-1)^{r-1}F^{(r)}(x)Y_{n,r} \\ =: & \sum_{i=1}^{n}(1_{\{X_i\le x\}} - F(x)) + V_{n,p}(x), \end{aligned}$$

where $F^{(r)}$ is the rth order derivative of F. Setting $U_i = F(X_i)$ and $x = Q(y)$ in the definition of $S_n(\cdot)$, we arrive at its uniform version,

$$\begin{aligned} \tilde{S}_{n,p}(y) &= \sum_{i=1}^{n}(1_{\{U_i\le y\}} - y) + \sum_{r=1}^{p}(-1)^{r-1}F^{(r)}(Q(y))Y_{n,r} \\ =: & \sum_{i=1}^{n}(1_{\{U_i\le y\}} - y) + \tilde{V}_{n,p}(y). \end{aligned}$$

Denote

$$d_{n,p} = \begin{cases} n^{-(1-\beta)}L_0^{-1}(n)(\log n)^{5/2}(\log\log n)^{3/4}, & (p+1)(2\beta-1)\ge 1 \\ n^{-p(\beta-\frac{1}{2})}L_0^{p}(n)(\log n)^{1/2}(\log\log n)^{3/4}, & (p+1)(2\beta-1)< 1 \end{cases}.$$

We shall need the following lemma, referred to as the reduction principle.

Lemma 10 ([24])**.** *Let p be a positive integer. Then, as $n\to\infty$,*

$$\mathrm{E}\sup_{x\in\mathbf{R}}\left|\sum_{i=1}^{n}(1_{\{X_i\le x\}} - F(x)) + \sum_{r=1}^{p}(-1)^{r-1}F^{(r)}(x)Y_{n,r}\right|^2 = O(\Xi_n + n(\log n)^2),$$

where

$$\Xi_n = \begin{cases} O(n), & (p+1)(2\beta-1)>1 \\ O(n^{2-(p+1)(2\beta-1)}L_0^{2(p+1)}(n)), & (p+1)(2\beta-1)<1 \end{cases}.$$

Using Lemma 10 we obtain (cf. [4])

$$\begin{aligned} &\sigma_{n,p}^{-1}\sup_{x\in\mathbf{R}}|S_n(x)| \\ &= \begin{cases} O_{a.s}(n^{-(\frac{1}{2}-p(\beta-\frac{1}{2}))}L_0^{-p}(n)(\log n)^{5/2}(\log\log n)^{3/4}), & (p+1)(2\beta-1)>1 \\ O_{a.s}(n^{-(\beta-\frac{1}{2})}L_0(n)(\log n)^{1/2}(\log\log n)^{3/4}), & (p+1)(2\beta-1)<1 \end{cases}. \end{aligned}$$

Since (see (2))

$$\frac{\sigma_{n,p}}{\sigma_{n,1}} \sim n^{-(\beta-\frac{1}{2})(p-1)}L_0^{p-1}(n)$$

we obtain

$$\sup_{x\in\mathbf{R}} |\beta_n(x) + \sigma_{n,1}^{-1} V_{n,p}(x)| =$$

$$= \frac{\sigma_{n,p}}{\sigma_{n,1}} \sup_{x\in\mathbf{R}} \left| \sigma_{n,p}^{-1} \sum_{i=1}^{n} (1_{\{X_i \le x\}} - F(x)) + \sigma_{n,p}^{-1} V_{n,p}(x) \right| = o_{a.s}(d_{n,p}).$$

Consequently, via $\{\alpha_n(y), y \in (0,1)\} = \{\beta_n(Q(y)), y \in (0,1)\}$,

$$\sup_{y\in(0,1)} |\alpha_n(y) + \sigma_{n,1}^{-1} \tilde{V}_{n,p}(y)| = O_{a.s}(d_{n,p}). \tag{12}$$

We have for any $a_n \to 0$ and by (10),

$$\begin{aligned} A_n \sigma_{n,1}^{-1} \int_{1-a_n/n}^{1-1/n} \tilde{V}_{n,p}(y) dQ_Y(y) &= A_n \sigma_{n,1}^{-1} \int_{1-a_n/n}^{1-1/n} \frac{\tilde{V}_{n,p}(y)}{f_Y Q_Y(y)} dy \\ &= -\left(A_n \int_{1-a_n/n}^{1-1/n} \frac{fQ(y)}{f_Y Q_Y(y)} dy \right) \left[\left(\sigma_{n,1}^{-1} \sum_{i=1}^{n} X_i \right) + o_P(\sigma_{n,1}^{-1}) \right]. \end{aligned} \tag{13}$$

Let

$$L_{11}(u) = L_2^*(u)/L_2(u), \qquad L_{21}(u) = (1/\alpha - 1/\alpha_0 + 1) L_{11}(u),$$
$$L_{12}(u) = L_3^*(u)/L_2(u), \qquad L_{22}(u) = (1/\alpha + 1) L_{12}(u),$$
$$L_{13}(u) = L_2^*(u)/L_3(u), \qquad L_{23}(u) = (-1/\alpha + 1) L_{13}(u),$$
$$L_{14}(u) = L_3^*(u)/L_3(u), \qquad L_{24}(u) = L_{14}(u).$$

Lemma 11. *Let p be a positive integer. Assume that* (10) *holds for $r = 1, \ldots, p$. Then*

$$A_n \sigma_{n,1}^{-1} \int_{1-k_n/n}^{1-1/n} \tilde{V}_{n,p}(y) dQ_Y(y) \xrightarrow{\mathrm{d}} Z.$$

Proof. In view of (13), we need only to study the asymptotic behavior, as $n \to \infty$, of $A_n \int_{1-k_n/n}^{1-1/n} \frac{fQ(y)}{f_Y Q_Y(y)} dy =: A_n K_n$ and to show that $A_n K_n \sim 1$.

We have by Karamata's Theorem:

In Case 1,

$$\begin{aligned} K_n &= \int_{1-k_n/n}^{1-1/n} (1-y)^{1/\alpha - 1/\alpha_0} \left(L_{11}((1-y)^{-1}) \right)^{-1} dy \\ &\sim (1/\alpha - 1/\alpha_0 + 1)^{-1} \left(\frac{k_n}{n} \right)^{1+1/\alpha - 1/\alpha_0} \left(L_{11}\left(\frac{n}{k_n} \right) \right)^{-1} \\ &\sim \left(\frac{k_n}{n} \right)^{1+1/\alpha - 1/\alpha_0} \left(L_{21}\left(\frac{n}{k_n} \right) \right)^{-1}. \end{aligned}$$

In Case 2,

$$\begin{aligned} K_n &= \int_{1-k_n/n}^{1-1/n} (1-y)^{1/\alpha} \left(L_{12}((1-y)^{-1})\right)^{-1} dy \\ &\sim \quad (1/\alpha+1)^{-1} \left(\frac{k_n}{n}\right)^{1+1/\alpha} \left(L_{12}\left(\frac{n}{k_n}\right)\right)^{-1} \sim \left(\frac{k_n}{n}\right)^{1+1/\alpha} \left(L_{22}\left(\frac{n}{k_n}\right)\right)^{-1}. \end{aligned}$$

In Case 3,

$$\begin{aligned} K_n &= \int_{1-k_n/n}^{1-1/n} (1-y)^{-1/\alpha_0} \left(L_{13}((1-y)^{-1})\right)^{-1} dy \\ &\sim \quad \frac{1}{-1/\alpha_0+1} \left(\frac{k_n}{n}\right)^{1-1/\alpha_0} \left(L_{13}\left(\frac{n}{k_n}\right)\right)^{-1} \sim \left(\frac{k_n}{n}\right)^{1-1/\alpha_0} \left(L_{23}\left(\frac{n}{k_n}\right)\right)^{-1}. \end{aligned}$$

In Case 4,

$$\begin{aligned} K_n &= \int_{1-k_n/n}^{1-1/n} \left(L_{14}((1-y)^{-1})\right)^{-1} dy \\ &\sim \quad \left(\frac{k_n}{n}\right) \left(L_{14}\left(\frac{n}{k_n}\right)\right)^{-1} \sim \left(\frac{k_n}{n}\right) \left(L_{14}\left(\frac{n}{k_n}\right)\right)^{-1}. \end{aligned}$$

Thus, in either case, $A_n K_n \sim 1$.

□

Lemma 12. *For any $k_n \to \infty$, $k_n = o(n)$*

$$\frac{U_{n-k_n:n}}{1-k_n/n} \xrightarrow{p} 1.$$

Proof. In view of (12) one obtains

$$\sup_{y\in(0,1)} |u_n(y)| = \sup_{y\in(0,1)} |\alpha_n(y)| = O_P(1).$$

Consequently,

$$\begin{aligned} \sup_{y\in(0,1)} |y - U_n(y)| &= \sup_{y\in(0,1)} \sigma_{n,1} n^{-1} |u_n(y)| = \sup_{y\in(0,1)} \sigma_{n,1} n^{-1} |\alpha_n(y)| \\ &= O_P(\sigma_{n,1} n^{-1}). \end{aligned}$$

Thus, the result follows by noting that $U_n(1-k_n/n) = U_{n-k_n:n}$.

□

An easy consequence of (12) is the following result.

Lemma 13. *For any $k_n \to 0$,*

$$\sup_{y\in(1-k_n/n,1)} |\alpha_n(y)| = O_{a.s.}(d_{n,p}) + O_P(f(Q(1-k_n/n))).$$

3.2 Proof of Theorem 1

To obtain the limiting behavior of sums of extremes, we shall use the following decomposition: Since $E_n(\cdot)$ has no jumps after $U_{n:n}$ and $Y_j = Q_Y F(X_j) = Q_Y(U_j)$, we have

$$
\begin{aligned}
& A_n \sigma_{n,1}^{-1} \left(\sum_{j=n-k_n+1}^{n} Y_{j:n} - n \int_{1-k_n/n}^{1} Q_Y(y)dy \right) \\
= \; & A_n \sigma_{n,1}^{-1} \left(\sum_{j=n-k_n+1}^{n} Q_Y(U_{j:n}) - n \int_{1-k_n/n}^{1} Q_Y(y)dy \right) \\
= \; & A_n \sigma_{n,1}^{-1} \left(n \int_{U_{n-k_n:n}}^{U_{n:n}} Q_Y(y)dE_n(y) - n \int_{1-k_n/n}^{1} Q_Y(y)dy \right) \\
= \; & A_n \sigma_{n,1}^{-1} \left(n \int_{U_{n-k_n:n}}^{1} Q_Y(y)dE_n(y) - n \int_{1-k_n/n}^{1} Q_Y(y)dy \right) \\
= \; & A_n \sigma_{n,1}^{-1} n \left\{ \int_{1-\frac{k_n}{n}}^{1-\frac{1}{n}} (y - E_n(y))dQ_Y(y) \right. \\
& \left. + \int_{1-\frac{1}{n}}^{1} (y - E_n(y))dQ_Y(y) + \int_{U_{n-k_n:n}}^{1-k_n/n} (1 - \frac{k_n}{n} - E_n(y))dQ_Y(y) \right\} \\
= \; & -A_n \int_{1-\frac{k_n}{n}}^{1-\frac{1}{n}} \alpha_n(y)dQ_Y(y) - A_n \int_{1-1/n}^{1} \alpha_n(y)dQ_Y(y) \\
& + A_n \sigma_{n,1}^{-1} n \int_{U_{n-k_n:n}}^{1-k_n/n} (1 - \frac{k_n}{n} - E_n(y))dQ_Y(y) =: I_1 + I_2 + I_3.
\end{aligned}
$$

We will show that I_1 yields the asymptotic normality. Further, we will show that the latter two integrals are asymptotically negligible.

Each term will be treated in a separate section. Let p be the smallest integer such that $(p+1)(2\beta - 1) > 1$, so that $d_{n,p} = n^{-(1-\beta)}\ell(n)$.

3.2.1 First term

Let $\psi_\mu(y) = (y(1-y))^\mu$, $y \in [0,1]$, $\mu > 0$.

For $k_n = [n^\xi]$ and arbitrary small $\delta > 0$ one has by (12),

$$
\begin{aligned}
& A_n \sup_{y \in (0,1)} \left| \alpha_n(y) + \sigma_{n,1}^{-1} \tilde{V}_{n,p}(y) \right| = O_{a.s}\left(A_n d_{n,p}\right) \\
= \; & \begin{cases}
n^{-(\xi+\xi/\alpha-\xi/\alpha_0-1/\alpha+1/\alpha_0-\beta-\delta)}, & \text{if } X \in MDA(\Phi_\alpha), Y \in MDA(\Phi_{\alpha_0}), \\
n^{-(\xi+\xi/\alpha-1/\alpha-\beta-\delta)}, & \text{if } X \in MDA(\Phi_\alpha), Y \in MDA(\Lambda), \\
n^{-(\xi-\xi/\alpha_0+1/\alpha_0-\beta-\delta)}, & \text{if } X \in MDA(\Lambda), Y \in MDA(\Phi_{\alpha_0}), \\
n^{-(\xi-\beta-\delta)}, & \text{if } X \in MDA(\Lambda), Y \in MDA(\Lambda).
\end{cases}
\end{aligned}
$$

Let

$$J_n = A_n \left| \int_{1-\frac{k_n}{n}}^{1-\frac{1}{n}} \frac{\left|\alpha_n(y) + \sigma_{n,1}^{-1}\tilde{V}_{n,p}(y)\right|}{\psi_\mu(y)} \psi_\mu(y) dQ_Y(y) \right|.$$

Case 1: Since condition (*) on ξ holds,

$$1/\alpha_0 < \xi + \xi(1/\alpha - 1/\alpha_0) - 1/\alpha + 1/\alpha_0 - \beta.$$

Set $\mu = (\alpha_0 - \delta)^{-1}$ with $\delta > 0$ so small that

$$\mu < \xi + \xi(1/\alpha - 1/\alpha_0) - 1/\alpha + 1/\alpha_0 - \beta - \delta.$$

Then, we have $\mathrm{E}(Y^+)^{1/\mu+\delta/2} < \infty$. The latter condition is sufficient for the finiteness of $\int_{x_1}^1 \psi_\mu(y) dQ_Y(y)$, where $x_1 = \inf\{y : Q_Y(y) \geq 0\}$, (see [22, Remark 2.4]). Thus,

$$J_n = o_{a.s}(A_n d_{n,p} n^\mu) \int_{x_1}^1 \psi_\mu(y) dQ_Y(y) = o_{a.s}(1) O(1).$$

Since in Case 3, (***) holds, a similar approach yields that in this case $J_n = o_{a.s}(1)$.

Case 2: If $Y \in MDA(\Lambda)$ then $\mathrm{E}(Y^+)^{\alpha_0} < \infty$ for all $\alpha > 0$ (see [10, Corollary 3.3.32]). Thus, in view of (**), choose arbitrary small $\delta > 0$ and α_0 so big that $\mathrm{E}(Y^+)^{\alpha_0} < \infty$ and

$$\frac{1}{\alpha_0 - \delta} < \xi + \xi/\alpha - 1/\alpha - \beta - \delta.$$

Set $\mu = (\alpha_0 - \delta)^{-1}$ and continue as in the Case 1. A similar reasoning applies to Case 4, provided $\xi > \beta$. Thus, in either case

$$A_n \left| \int_{1-\frac{k_n}{n}}^{1-\frac{1}{n}} \left(\alpha_n(y) + \sigma_{n,1}^{-1} \tilde{V}_{n,p}(y) \right) dQ_Y(y) \right| = o_{a.s}(1).$$

Now, the asymptotic normality of I_1 follows from Lemma 11.

3.2.2 Second term

We have

$$\begin{aligned} & A_n \int_{1-1/n}^1 \alpha_n(y) dQ_Y(y) \\ = \quad & -A_n \sigma_{n,1}^{-1} n \int_{1-1/n}^1 (1 - E_n(y)) dQ_Y(y) + A_n \sigma_{n,1}^{-1} n \int_{1-1/n}^1 (1-y) dQ_Y(y) \\ := \quad & J_1 + J_2. \end{aligned}$$

Since $\mathrm{E}J_1 = J_2$, it suffices to show that $J_2 = o(1)$.

Case 1: We have by Karamata's Theorem

$$J_2 = A_n \sigma_{n,1}^{-1} n \int_{1-1/n}^{1} \frac{(1-y)}{(1-y)^{1+1/\alpha_0} L_2^*(y^{-1})} dy$$
$$\sim \left(\frac{n}{k_n}\right)^{1+1/\alpha-1/\alpha_0} n^{\beta-3/2} n (\frac{1}{n})^{1-1/\alpha_0} \ell(n)\ell(n/k_n)$$

which converges to 0 using the assumption (*).

Likewise, in Case 3,

$$J_2 \sim \left(\frac{n}{k_n}\right)^{1-1/\alpha_0} n^{\beta-3/2} n (\frac{1}{n})^{1-1/\alpha_0} \ell(n)\ell(n/k_n)$$

which converges to 0 using the assumption (***).

Case 2: We have,

$$J_2 = A_n \sigma_{n,1}^{-1} n \int_{1-1/n}^{1} \frac{1-y}{f_Y Q_Y(y)} dy \sim A_n \sigma_{n,1}^{-1} \ell(n)\ell(n/k_n)$$

which converges to 0, using the assumption (**). The same argument applies to Case 4. Therefore, in either case, $I_2 = o_P(1)$.

3.2.3 Third term

To prove that $I_3 = o_P(1)$, let y be in the interval with the endpoints $U_{n-k_n:n}$ and $1 - k_n/n$. Then

$$\left|1 - E_n(y) - \frac{k_n}{n}\right| \leq |E_n(1-k_n/n) - (1-k_n/n)|.$$

Case 1: By Lemma 12 and $Y \in MDA(\Phi_{\alpha_0})$, we have

$$Q_Y(1-k_n/n)/Q_Y(U_{n-k_n:n}) \xrightarrow{p} 1. \tag{14}$$

Hence, by condition (*),

$$\left(\frac{n}{k_n}\right)^{1+1/\alpha-1/\alpha_0} \ell(n/k_n) Q_Y(1-k_n/n) d_{n,p}$$
$$= n^{1+1/\alpha} \ell(n)\ell(n/k_n) n^{-\xi(1+1/\alpha)} d_{n,p} \to 0. \tag{15}$$

Also, by (7) and (9),

$$A_n Q_Y(1-k_n/n) f Q(1-k_n/n) \sim C L_{21} \left(\frac{n}{k_n}\right) \frac{L_1^*(n/k_n)}{L_1(n/k_n)} \sim C \tag{16}$$

Thus, by (14), (15), (16) and Lemma 13

$$\begin{aligned} I_3 &\leq A_n Q_Y(1-k_n/n)|\alpha_n(1-k_n/n)|\frac{|Q_Y(1-k_n/n)-Q_Y(U_{n-k_n:n})|}{Q_Y(1-k_n/n)} \\ &= A_n Q_Y(1-k_n/n)\alpha_n(1-k_n/n)o_p(1) \\ &= o_p\left(A_n Q_Y(1-k_n/n) fQ(1-k_n/n)\right) + o_p\left(A_n Q(1-k_n/n) d_{n,p}\right) = o_P(1). \end{aligned}$$

Case 3: By Lemma 12 and $Y \in MDA(\Phi_{\alpha_0})$ we have (14). Since $\xi > \beta > \beta/(1-1/\alpha_0)$,

$$\left(\frac{n}{k_n}\right)^{1-1/\alpha_0} \ell(n/k_n) Q_Y(1-k_n/n) d_{n,p} = n^{\beta-\xi} \to 0. \tag{17}$$

Also, by (8) and (9),

$$\begin{aligned} &\left(\frac{n}{k_n}\right)^{1-1/\alpha_0} L_{23}\left(\frac{n}{k_n}\right) Q_Y(1-k_n/n) fQ(1-k_n/n) \\ &\qquad \sim \ C L_{23}\left(\frac{n}{k_n}\right)\frac{L_3(n/k_n)}{L_2^*(n/k_n)} \sim C. \end{aligned} \tag{18}$$

Thus, by (17), (18), we conclude as above that $I_3 = o_P(1)$.

Cases 2 and 4:

$$T_n(\lambda) = A_n|\alpha_n(1-k_n/n)|\left|Q_Y(r_n^+(\lambda)) - Q_Y(r_n^-(\lambda))\right|,$$

where $r_n^+(\lambda) = 1 - \frac{k_n}{\lambda n}$, $r_n^-(\lambda) = 1 - \lambda\frac{k_n}{n}$ and $1 < \lambda < \infty$ is arbitrary. Applying an argument as in the proof of Theorem 1 in [8], we have

$$\liminf_{n\to\infty} P(|I_3| < |T_n(\lambda)|) \geq \liminf_{n\to\infty} P(r_n^-(\lambda) \leq U_{n-k_n:n} \leq r_n^+(\lambda)).$$

In view of Lemma 12, the lower bound is 1. Thus, $\lim_{n\to\infty} P(|I_3| < |T_n(\lambda)|) = 1$. Further, by Lemma 4 in [18],

$$\lim_{n\to\infty}(Q_Y(r_n^+(\lambda)) - Q_Y(r_n^-(\lambda)))L_3^*(n/k_n) = -\log\lambda.$$

Thus, for large n,

$$\begin{aligned} T_n(\lambda) &= A_n|\alpha_n(1-k_n/n)|(L_3^*(n/k_n))^{-1}|Q_Y(r_n^+(\lambda)) - Q_Y(r_n^-(\lambda))|L_3^*(n/k_n) \\ &\leq C_1\frac{A_n}{L_3^*(n/k_n)} fQ(1-k_n/n)(\log\lambda) + C_2\frac{A_n}{L_3^*(n/k_n)} d_{n,p}\log\lambda \end{aligned}$$

almost surely with some constants C_1, C_2. The second term, for arbitrary λ, converges to 0 by the choice of ξ. Also,

$$A_n\frac{fQ(1-k_n/n)}{L_3^*(n/k_n)} \leq \begin{cases} \left(\frac{n}{k_n}\right)^{1+1/\alpha} L_{22}\left(\frac{n}{k_n}\right)\frac{\left(\frac{k_n}{n}\right)^{1+1/\alpha} L_2\left(\frac{n}{k_n}\right)}{L_3^*\left(\frac{n}{k_n}\right)}, & \text{in Case 2,} \\ \left(\frac{n}{k_n}\right) L_{24}\left(\frac{n}{k_n}\right)\frac{\left(\frac{k_n}{n}\right) L_3\left(\frac{n}{k_n}\right)}{L_3^*\left(\frac{n}{k_n}\right)}, & \text{in Case 4.} \end{cases}$$

In either case, the above expressions are asymptotically equal to 1. Thus, we have for sufficiently large n, $T_n(\lambda) \le C_1 \log\lambda$ almost surely. Thus, $\lim_{n\to\infty} P(|T_n(\lambda)| \le C_1\log\lambda) = 1$. Consequently,

$$
\begin{aligned}
&\lim_{n\to\infty} P(|I_3| > C_1\log\lambda) = \\
&\le \lim_{n\to\infty} P(|I_3| > C_1\log\lambda, |T_n(\lambda)| \le C_1\log\lambda) + \lim_{n\to\infty} P(|T_n(\lambda)| > C_1\log\lambda) \\
&\le \lim_{n\to\infty} P(|I_3| > |T_n(\lambda)|) + 0 = 0
\end{aligned}
$$

and thus $I_3 = o_P(1)$ by taking $\lambda \to 1$.

□

Remark 14. Wu [24] proved a stronger version of Lemma 10 above:

$$
\begin{aligned}
\mathrm{E}\sup_{x\in\mathbf{R}}(1+|x|)^\gamma \left|\sum_{i=1}^n (1_{\{X_i\le x\}} - F(x)) + \sum_{r=1}^p (-1)^{r-1}F^{(r)}(x)Y_{n,r}\right|^2 \\
= O(\Xi_n + n(\log n)^2),
\end{aligned}
$$

where $\gamma > 0$ is such that $\mathrm{E}|\epsilon_1|^{4+\gamma} < \infty$. Applying it to the uniform random variables $U_i = F(X_i)$,

$$
\begin{aligned}
\mathrm{E}\sup_{y\in(0,1)}(1+|Q(y)|)^\gamma \left|\sum_{i=1}^n (1_{\{U_i\le y\}} - y) + \sum_{r=1}^p (-1)^{r-1}F^{(r)}(Q(y))Y_{n,r}\right|^2 \\
= O(\Xi_n + n(\log n)^2).
\end{aligned}
$$

Now, let's look at Case 4. The constraint $\xi > \beta$ comes from the estimation of the third term: we need to control $A_n\alpha_n(1-k_n/n)$ and thus, in view of Lemma 13, we have to estimate $A_n d_{n,p}$. This converges to 0 if $\xi > \beta$.

Using the weighted version of Lemma 10, we would have obtained in Lemma 13:

$$
\sup_{y\in(1-k_n/n,1)} (1+Q(y))^{\gamma/2}|\alpha_n(y)| = O_{a.s.}(d_{n,p}) + O_P(fQ(1-k_n/n)).
$$

However, in Case 4, $Q(\cdot)$ is slowly varying at 1. Consequently, the approach via weighted approximation does not improve constraints on ξ. (A slight improvement can be achieved in Cases 1 and 2, where $Q(\cdot)$ is regularly varying at 1).

Acknowledgement.

This work was done partially during my stay at Carleton University. I am thankful to Professors Barbara Szyszkowicz and Miklós Csörgő for the support and helpful remarks. Also, the Referee's comments are greatly appreciated.

References

[1] Csörgő, M. (1983). *Quantile Processes with Statistical Applications.* CBMS-NSF Regional Conference Series in Applied Mathematics. MR0745130

[2] Csörgő, M., Csörgő, S., Horváth, L., Mason, D. M. (1986). Weighted empirical and quantile processes. *Ann. Probab.* **14**, 31–85. MR0815960

[3] Csörgő, M., Csörgő, S., Horváth, L., Mason, D. M. (1986). Normal and stable convergence of integral functions of the empirical distribution function. *Ann. Probab.* **14**, 86–118. MR0815961

[4] Csörgő, M. and Kulik, R. (2008). Reduction principles for quantile and Bahadur-Kiefer processes of long-range dependent linear sequences. *Probab. Theory Relat. Fields.* DOI: 10.1007/s00440-007-0107-9.

[5] Csörgő, M. and Kulik, R. (2008). Weak convergence of Vervaat and Vervaat Error processes of long-range dependent Gaussian sequences. *J. Theoret. Probab.* DOI: 10.1007/s10959-007-0124-8.

[6] Csörgő, M., Szyszkowicz, B. and Wang, L. (2006). Strong Invariance Principles for Sequential Bahadur-Kiefer and Vervaat Error Processes of Long-Range Dependence Sequences. *Ann. Statist.* **34**, 1013–1044. MR2283402

[7] Csörgő, S., Haeusler, E. and Mason, D. M. (1991). The quantile-transform–empirical-process approach to limit theorems for sums of order statistics. *Sums, trimmed sums and extremes*, 215–267, Progr. Probab., **23**, Birkhäuser, Boston, MA. MR1117272

[8] Csörgő, S., Horváth, L., Mason, D. M. (1986). What portion of the sample makes a partial sum asymptotically stable or normal? *Probab. Theory Relat. Fields* **72**, 1–16. MR0835156

[9] Csörgő, S. and Mason, D. M. (1986). The asymptotic distribution of sums of extreme values from a regularly varying distribution. *Ann. Probab.* **11**, 974–983. MR0841597

[10] Embrechts, P., Kluppelberg, C. and Mikosch, T. (1997). *Modelling extremal events. For insurance and finance.* Applications of Mathematics **33**. Springer-Verlag, Berlin. MR1458613

[11] Giraitis, L. and Surgailis, D. (2002). The reduction principle for the empirical process of a long memory linear process. *Empirical process techniques for dependent data*, 241–255, Birkhäuser Boston, Boston, MA. MR1958784

[12] de Haan, L. (1975). On Regular Variation and Its Applications to the Weak Convergence of Sample Extremes. Math. Centre Tracts 32, Amsterdam.

[13] Ho, H.-C. and Hsing, T. (1996). On the asymptotic expansion of the empirical process of long-memory moving averages. *Ann. Statist.* **24**, 992–1024. MR1401834

[14] Horváth, L. (1987). On the tail behaviour of quantile processes. *Stochastic Process. Appl.* **25**, 57–72. MR0904264

[15] Koul, H.L. and Surgailis, D. (2002). Asymptotic expansion of the empirical process of long memory moving averages. *Empirical process techniques for dependent data*, 213–239, Birkhäuser Boston, Boston, MA. MR1958783

[16] Kulik, R. and Ould Haye, M. (2008). Trimmed sums of long-range dependent moving averages. To appear in *Statist. Probab. Letters.* DOI: 10.1016/j.spl.2008.02.033.

[17] Leadbetter, M.R., Lindgren, G. and Rootzén, H. (1983). *Extremes and Related Properties of Random Sequences and Processes*. Springer-Verlag, New York. MR0691492

[18] Lo, G. S. (1989). A note on the asymptotic normality of sums of extreme values. *J. Statist. Plann. Inference* **22**, 127–136. MR0996806

[19] Mikosch, T. and Samorodnitsky, G. (2000). The supremum of a negative drift random walk with dependent heavy-tailed steps. *Ann. Appl. Probab.* **10**, 1025–1064. MR1789987

[20] Parzen, E. (1979). Nonparametric statistical data modeling. With comments by John W. Tukey, Roy E. Welsch, William F. Eddy, D. V. Lindley, Michael E. Tarter and Edwin L. Crow, and a rejoinder by the author. *J. Amer. Statist. Assoc.* **74**, 105–131. MR0529528

[21] Samorodnitsky, G. (2004). Extreme value theory, ergodic theory and the boundary between short memory and long memory for stationary stable processes. *Ann. Probab.* **32**, 1438–1468. MR2060304

[22] Shao, Q.-M. and Yu, H. (1996). Weak convergence for weighted empirical processes of dependent sequences. *Ann. Probab.* **24**, 2098–2127. MR1415243

[23] Stigler, S. M. (1973). The asymptotic distribution of the trimmed mean. *Ann. Statist.* **1**, 472–477. MR0359134

[24] Wu, W.B. (2003). Empirical processes of long-memory sequences. Bernoulli **9**, 809–831. MR2047687

[25] Wu, W.B. (2005). On the Bahadur representation of sample quantiles for dependent sequences. *Ann. Statist.* **4**, 1934–1963. MR2166566