On the central limit theorem for modulus trimmed sums

Alina Bazarova [a], István Berkes [a,*], Lajos Horváth [b]

[a] *Graz University of Technology, Institute of Statistics, Kopernikusgasse 24, 8010 Graz, Austria*
[b] *University of Utah, Department of Mathematics, Salt Lake City, UT 84112-0090, USA*

ARTICLE INFO

ABSTRACT

We prove a functional central limit theorem for modulus trimmed i.i.d. variables in the domain of attraction of a nonnormal stable law. In contrast to the corresponding result under ordinary trimming, our CLT contains a random centering factor which is inevitable in the nonsymmetric case. The proof is based on the weak convergence of a two-parameter process where one of the parameters is time and the second one is the fraction of truncation.

© 2013 Elsevier B.V. All rights reserved.

1. Introduction

Let $X_1, X_2, \ldots$ be independent, identically distributed random variables in the domain of attraction of a stable law G with parameter $0 < \alpha < 2$. That is, assume that the partial sums $S_n = \sum_{k=1}^{n} X_k$ satisfy

$$(S_n - b_n)/a_n \xrightarrow{d} G \tag{1.1}$$

with suitable norming and centering sequences $\{a_n\}$, $\{b_n\}$. The necessary and sufficient condition for (1.1) is that F, the distribution function of X_1, satisfies

$$1 - F(x) + F(-x) = x^{-\alpha} L(x), \quad x > 0 \tag{1.2}$$

and

$$\frac{1 - F(x)}{1 - F(x) + F(-x)} \to p, \qquad \frac{F(-x)}{1 - F(x) + F(-x)} \to q \quad (x \to \infty) \tag{1.3}$$

where L is a function slowly varying at ∞ and $p, q \geq 0$, $p + q = 1$. (See e.g. Feller (1971).) In contrast to the case of finite variances, the contribution of extremal terms in the partial sums S_n is not negligible and dropping a single term can change the asymptotic behavior of the sum. Let $X_{n,1} \leq X_{n,2} \leq \cdots \leq X_{n,n}$ be the order statistics of $(X_1, X_2, \ldots, X_n)$ and put for $d \geq 1$

$$S_n^{(d)} = \sum_{j=d+1}^{n-d} X_{n,j}. \tag{1.4}$$

---

* Corresponding author.
   E-mail addresses: bazarova@tugraz.at (A. Bazarova), berkes@tugraz.at (I. Berkes), horvath@math.utah.edu (L. Horváth).

For fixed d, Le Page et al. (1981) determined the asymptotic distribution of the trimmed sum $S_n^{(d)}$ and Csörgő et al. (1986b) proved that under

$$d_n \to \infty, \quad d_n/n \to 0 \tag{1.5}$$

the trimmed sum $S_n^{(d_n)}$, suitably centered and normalized, is asymptotically normal. These results give a remarkable picture on the partial sum behavior of i.i.d. sequences in the domain of attraction of a non-normal stable law. They show that the contribution of d_n extremal terms under (1.5) already gives the stable limit distribution of the total partial sum S_n and the contribution of the remaining elements will be an asymptotically normal variable with magnitude negligible compared with S_n.

The previous results describe the effects of the extremal elements of an i.i.d. sample on their partial sum. Note, however, that other kinds of trimming lead to different phenomena. For $1 \le d \le n$ let $\eta_{d,n}$ denote the d-th largest of $|X_1|, \ldots, |X_n|$ and let

$$^{(d)}S_n = \sum_{k=1}^{n} X_k I\{|X_k| \le \eta_{d,n}\}. \tag{1.6}$$

If the distribution of X_1 is continuous, then $|X_1|, |X_2|, \ldots$ are different with probability 1, and thus $^{(d)}S_n$ coincides with the usual modulus trimmed sum obtained by discarding from S_n the $d-1$ elements with the largest moduli. Griffin and Pruitt (1987) showed that if X_1 has a symmetric distribution, then $^{(d_n)}S_n$ is asymptotically normal for any $d_n \to \infty, d_n/n \to 0$, but this is generally false in the nonsymmetric case. The purpose of this paper is to describe the asymptotic distribution of $^{(d_n)}S_n$ in the general case. Put

$$H(t) = P(|X| \ge t) \quad \text{and} \quad m(t) = \mathsf{E}XI\{|X| \le t\},$$

and let $H^{-1}(t) = \inf\{x : H(x) \le t\}$ ($0 < t < 1$) denote the generalized inverse of H. Our main result is the following.

Theorem 1.1. *Let $X_1, X_2, \ldots$ be i.i.d. random variables with distribution function F satisfying* (1.2)*,* (1.3) *and assume that* (1.5) *holds. Then we have*

$$\frac{1}{A_n} \sum_{i=1}^{[nt]} \left(X_i I\{|X_i| \le \eta_{d,n}\} - m(\eta_{d,n})\right) \xrightarrow{\mathcal{D}[0,1]} W(t) \tag{1.7}$$

where

$$A_n^2 = \frac{\alpha}{2-\alpha} d(H^{-1}(d/n))^2 \tag{1.8}$$

and W is the Wiener process.

Theorem 1.1 shows that allowing a random centering factor, the modulus trimmed CLT holds for continuous i.i.d. variables under exactly the same conditions as under ordinary trimming. If F is not continuous, the sample $(X_1, \ldots, X_n)$ may contain equal elements with positive probability; according to the definition in Griffin and Pruitt (1987), 'ties' between elements with equal moduli are broken according to the order in which the variables occur in $(X_1, \ldots, X_n)$. But no matter how we break the ties, it may happen that from a set of sample elements with equal moduli some are discarded and others are not, which is rather unnatural from the statistical point of view, since trimming is mainly used to improve the performance of statistical procedures by removing large elements from the sample. The definition of $^{(d)}S_n$ in (1.6) resolves this difficulty and leads to satisfactory asymptotic results in the general case.

Theorem 1.1 enables one to give, among others, change point tests for heavy tailed processes, while the standard CUSUM test fails under infinite variances. A fairly precise characterization for the modulus trimmed CLT with nonrandom centering and norming factors was given in Berkes and Horváth (2012).

Under additional technical assumptions on the distribution function of X_1 and on the growth speed of d_n, Theorem 1.1 was proved in Berkes et al. (2011) with a fairly complicated argument. The proof of Theorem 1.1 is much simpler and extends to dependent samples as well, as we will show in a subsequent paper. Let

$$\hat{A}_n^2 = \sum_{i=1}^{n} X_i^2 I\{|X_i| \le \eta_{d,n}\} - \frac{1}{n}\left(\sum_{i=1}^{n} X_i I\{|X_i| \le \eta_{d,n}\}\right)^2.$$

Berkes et al. (2011) showed that under the conditions of Theorem 1.1 we have that

$$\hat{A}_n/A_n \xrightarrow{P} 1$$

and therefore Theorem 1.1 yields

$$\frac{1}{\hat{A}_n}\left(\sum_{i=1}^{[nt]} X_i I\{|X_i| \le \eta_{d,n}\} - \frac{[nt]}{n}\sum_{i=1}^{n} X_i I\{|X_i| \le \eta_{d,n}\}\right) \xrightarrow{\mathcal{D}[0,1]} B(t),$$

where $B(t) = W(t) - tW(1)$ denotes a Brownian bridge. Hence standard CUSUM techniques can be used to detect changes in the mean and/or location when in the case of observations without second moments, observations with modulus larger than $\eta_{d,n}$ are excluded from the sample.

Let

$$U_n(t, s) = \sum_{i=1}^{[nt]} \left(X_i I\{|X_i| \leq sH^{-1}(d/n)\} - EX_i I\{|X_i| \leq sH^{-1}(d/n)\} \right) \quad (s \geq 0, t \geq 0).$$

We will deduce (1.7) from the following two-dimensional limit theorem.

Theorem 1.2. *Let $X_1, X_2, \ldots$ be i.i.d. random variables with distribution function F satisfying* (1.2), (1.3) *and assume that* (1.5) *holds. Then*

$$\frac{1}{A_n} U_n(t, s) \longrightarrow W(t, s^{2-\alpha}) \quad \text{weakly in } \mathcal{D}([0, 1] \times [1/2, 3/2]),$$

where A_n is defined by (1.8) *and $\{W(x, y), x \geq 0, y \geq 0\}$ is a two-parameter Wiener process.*

Note that by Kiefer (1972) we have

$$\frac{\eta_{d,n}}{H^{-1}(d/n)} \xrightarrow{P} 1.$$

Since the limit process in Theorem 1.2 has continuous trajectories a.s., Billingsley (1968), p. 144–145 implies that

$$\frac{1}{A_n} U_n(t, \eta_{d,n}/H^{-1}(d/n)) \xrightarrow{\mathcal{D}[0,1]} W(t, 1)$$

which is exactly the functional CLT in (1.7), since $W(t, 1)$ is a Wiener process. Thus Theorem 1.1 is a consequence of Theorem 1.2.

2. Proof of Theorem 1.2

Given a process $Y(s, t)$ defined on a rectangle $H = [a, b] \times [c, d]$, let $Y(H)$ denote the increment of Y over H.

Lemma 2.1. *Let $\{Y_n(t, s), n \geq 1\}$ be processes defined on a rectangle $[a, b] \times [c, d] \subset [0, \infty)^2$ and assume that for some $\gamma > 0$*

$$E|Y_n(B)|^\gamma |Y_n(C)|^\gamma \leq \mu(B)\mu(C), \tag{2.1}$$

where μ denotes area and B and C are rectangles of the form $[t_1, t_2] \times [s_1, s_2]$ having one common edge, but otherwise disjoint. Then the sequence $\{Y_n(t, s), n \geq 1\}$ is tight. If every $Y_n(t, s)$ is piecewise constant in t, i.e. there exists a finite set $H_n \subset [a, b]$ such that $Y_n(t, s)$ is constant on the left closed intervals determined by the elements of $H_n \cup \{a\} \cup \{b\}$, then it suffices to verify (2.1) for rectangles $[t_1, t_2] \times [s_1, s_2]$ where $t_1, t_2 \in H_n$.

This is a special case of a general tightness condition due to Bickel and Wichura (1971, Theorem 3).

As is shown in Csörgő et al. (1986a, Proposition A.3), the conditions of Theorem 1.2 imply $H^{-1}(t) = t^{-1/\alpha}\ell(t)$ $(0 < t < 1)$, where ℓ is slowly varying at 0. Then by (1.8) we have

$$A_n^2 \sim \frac{\alpha}{2 - \alpha} d \, (n/d)^{2/\alpha} \ell^2(d/n) \quad \text{as } n \to \infty \tag{2.2}$$

where $a_n \sim b_n$ means $a_n/b_n \to 1$ as $n \to \infty$.

Lemma 2.2. *If the conditions of Theorem 1.2 are satisfied, then for any $p \geq 2$ and any fixed $0 \leq a < b < \infty$ we have*

$$E|X_1|^p I\{aH^{-1}(d/n) < |X_1| \leq bH^{-1}(d/n)\} \sim \frac{\alpha}{p - \alpha}(b^{p-\alpha} - a^{p-\alpha})\ell^p(d/n)(n/d)^{(p-\alpha)/\alpha} \tag{2.3}$$

as $n \to \infty$. Also, if $b > 0$, then

$$E|X_1|I\{|X_1| \leq bH^{-1}(d/n)\} = \begin{cases} O((n/d)^{(1-\alpha)/\alpha}\ell(d/n)) & \text{if } \alpha < 1, \\ O((n/d)^\varepsilon) & \text{if } \alpha = 1, \\ O(1) & \text{if } \alpha > 1 \end{cases} \tag{2.4}$$

for any $\varepsilon > 0$.

Proof. Assume first $p \geq 2, 0 < a < b < \infty$. Clearly the left hand side of (2.3) equals

$$- \int_{aH^{-1}(d/n)}^{bH^{-1}(d/n)} t^p dH(t) = \int_{H(bH^{-1}(d/n))}^{H(aH^{-1}(d/n))} H^{-1}(u)^p du. \tag{2.5}$$

(Note that H is non-increasing and thus the left hand side of (2.5) is nonnegative.) Since H is regularly varying with exponent $-\alpha$, we have

$$H(aH^{-1}(d/n)) \sim a^{-\alpha}(d/n), \qquad H(bH^{-1}(d/n)) \sim b^{-\alpha}(d/n) \quad \text{as } n \to \infty.$$

Thus using the uniform convergence theorem for regularly varying functions (see e.g. Bingham et al. (1987), Theorem 1.5.2; note that we actually need the analogous result for regular variation at 0), we see that for $n \to \infty$ we have, uniformly for all u in the interval of integration of the second integral in (2.5),

$$H^{-1}(u) = u^{-1/\alpha} \ell(u) \sim u^{-1/\alpha} \ell(d/n).$$

Thus the integral equals

$$(1 + o(1)) \int_{(1+o(1))b^{-\alpha}(d/n)}^{(1+o(1))a^{-\alpha}(d/n)} u^{-p/\alpha} \ell^p(d/n) \, du, \tag{2.6}$$

which yields the right hand side of (2.3) after a simple calculation, since $p \neq \alpha$. If $a = 0$, then the upper limit in the integral on the right hand side of (2.5) and thus also in (2.6) becomes $H(0) = 1$ and by using Theorem 1.5.11 of Bingham et al. (1987) we get the right hand side of (2.3) with $a = 0$.

In the case of (2.4), instead of the integral in (2.6) we get

$$\int_{(1+o(1))b^{-\alpha}(d/n)}^{1} u^{-1/\alpha} \ell(u) \, du. \tag{2.7}$$

By Proposition 1.3.6(i) in Bingham et al. (1987), p. 16 we have $\ell(u) = O(u^{-\varepsilon})$ as $u \to 0$ for any $\varepsilon > 0$ which shows that for $\alpha > 1$ the integral $\int_0^1 u^{-1/\alpha} \ell(u) \, du$ converges and thus the expression (2.7) is $O(1)$. Using the same estimate for $\ell(u)$ for $\alpha = 1$ we get the second bound in (2.4). Finally, for $\alpha < 1$ Theorem 1.5.11 of Bingham et al. (1987) yields the first bound in (2.4), completing the proof of Lemma 2.2.

Proof of Theorem 1.2. Let $\Gamma(t, s)$ denote the limit process in Theorem 1.2 and put

$$Q_n = \frac{1}{A_n} \sum_{m=1}^{M} \sum_{j=1}^{J} \mu_{m,j} U_n([t_{m-1}, t_m] \times [s_{j-1}, s_j])$$

and

$$Z = \sum_{m=1}^{M} \sum_{j=1}^{J} \mu_{m,j} \Gamma([t_{m-1}, t_m] \times [s_{j-1}, s_j])$$

for all $M \geq 1, J \geq 1$, real coefficients $\mu_{m,j}, 1/2 \leq s_1 < s_2 < \cdots < s_J \leq 3/2, 0 < t_1 < \cdots < t_M = 1, t_0 = s_0 = 0$. Clearly, Z is a centered normal r.v. and

$$EZ^2 = \sum_{m=1}^{M} \sum_{j=1}^{J} \mu_{m,j}^2 (s_j^{2-\alpha} - s_{j-1}^{2-\alpha})(t_m - t_{m-1}). \tag{2.8}$$

We claim that

$$Q_n \xrightarrow{d} Z \quad \text{for all considered values of } M, J, \mu_{m,j}, t_m, s_j. \tag{2.9}$$

Since the processes U_n and Γ are equal to 0 on the boundary of the first quadrant, we have

$$U_n(t_m, s_j) = \sum_{m=1}^{M} \sum_{j=1}^{J} U_n([t_{m-1}, t_m] \times [s_{j-1}, s_j])$$

and the same relation holds for Γ. Thus (2.9) implies

$$\frac{1}{A_n} \sum_{m=1}^{M} \sum_{j=1}^{J} \mu_{m,j}^* U_n(t_m, s_j) \xrightarrow{d} \sum_{m=1}^{M} \sum_{j=1}^{J} \mu_{m,j}^* \Gamma(t_m, s_j)$$

for arbitrary real coefficients $\mu_{m,j}^*$ and this, by the Cramér–Wold device, implies the convergence of the finite-dimensional distributions in Theorem 1.2.

In view of the definition of $U_n(t, s)$ in Section 1 we have

$$U_n([t_{m-1}, t_m] \times [s_{j-1}, s_j]) = \sum_{i=[nt_{m-1}]+1}^{[nt_m]} (v_{i,j} - \mathsf{E}v_{i,j})$$

where

$$v_{i,j} = X_i I\{s_{j-1} H^{-1}(d/n) < |X_i| \le s_j H^{-1}(d/n)\}.$$

Thus relation (2.9) can be written equivalently as

$$\frac{1}{A_n} \sum_{k=1}^{n} (z_{k,n} - \mathsf{E}z_{k,n}) \xrightarrow{d} N(0, EZ^2) \quad \text{as } n \to \infty, \tag{2.10}$$

where

$$z_{k,n} = \sum_{j=1}^{J} \mu_{m,j} X_k I\{s_{j-1} H^{-1}(d/n) < |X_k| \le s_j H^{-1}(d/n)\}, \quad [nt_{m-1}] + 1 \le k \le [nt_m].$$

Since the terms in the last sum are random variables with disjoint support, we get from Lemma 2.2

$$\mathsf{E}z_{k,n}^2 = (1 + o_n(1)) \frac{\alpha}{2 - \alpha} (n/d)^{(2-\alpha)/\alpha} \ell^2(d/n) \sum_{j=1}^{J} \mu_{m,j}^2 (s_j^{2-\alpha} - s_{j-1}^{2-\alpha}), \quad [nt_{m-1}] + 1 \le k \le [nt_m]$$

and similarly

$$\mathsf{E}z_{k,n}^4 = (1 + o_n(1)) \frac{\alpha}{4 - \alpha} (n/d)^{(4-\alpha)/\alpha} \ell^4(d/n) \sum_{j=1}^{J} \mu_{m,j}^4 (s_j^{4-\alpha} - s_{j-1}^{4-\alpha}), \quad [nt_{m-1}] + 1 \le k \le [nt_m].$$

Thus using $d = d_n \to \infty$ we get by a simple calculation

$$\lim_{n \to \infty} \frac{\displaystyle\sum_{k=1}^{n} \mathsf{E}z_{k,n}^4}{\left(\displaystyle\sum_{k=1}^{n} \mathsf{E}z_{k,n}^2\right)^2} = 0. \tag{2.11}$$

On the other hand, the previous asymptotics for $\mathsf{E}z_{k,n}^2$ and the statement of Lemma 2.2 for $p = 1$ imply

$$\mathsf{E}^2 |z_{k,n}| = o_n(1) \mathsf{E}z_{k,n}^2, \quad 1 \le k \le n$$

and thus by Minkowski's inequality

$$\mathsf{E}|z_{k,n} - \mathsf{E}z_{k,n}|^2 = (1 + o_n(1)) \mathsf{E}z_{k,n}^2, \qquad \mathsf{E}|z_{k,n} - \mathsf{E}z_{k,n}|^4 = (1 + o_n(1)) \mathsf{E}z_{k,n}^4. \tag{2.12}$$

Thus (2.11) remains valid if we replace $z_{k,n}$ with $z_{k,n} - \mathsf{E}z_{k,n}$. Further by (2.2) and (2.8)

$$\sum_{k=1}^{n} \mathsf{E}z_{k,n}^2 = (1 + o_n(1)) \frac{\alpha}{2 - \alpha} n(n/d)^{(2-\alpha)/\alpha} \ell^2(d/n) \sum_{m=1}^{M} \sum_{j=1}^{J} \mu_{m,j}^2 (s_j^{2-\alpha} - s_{j-1}^{2-\alpha})(t_m - t_{m-1})$$

$$= (1 + o_n(1)) A_n^2 \mathsf{E}Z^2.$$

The last relation, together with (2.11), (2.12) and Lyapunov's CLT for triangular arrays, implies (2.10).

Next we prove tightness in Theorem 1.2. Consider two pairs of sets $B_{11} = [t_1, t] \times [s_1, s]$, $B_{12} = [t_1, t] \times [s, s_2]$ and $B_{11} = [t_1, t] \times [s_1, s]$, $B_{21} = [t, t_2] \times [s_1, s]$, where $t_1 < t < t_2$, $s_1 < s < s_2$. In view of Lemma 2.1, it suffices to show that

$$\mathsf{E} \left| \frac{1}{A_n} U_n(B_{11}) \right|^2 \left| \frac{1}{A_n} U_n(B_{ij}) \right|^2 \le C\mu(B_{11})\mu(B_{ij}), \tag{2.13}$$

holds for each $ij \in \{12, 21\}$ with some constant $C > 0$. Moreover, since $U_n(t, s)$ is constant on intervals $k/n \le t < (k+1)/n$, by the last statement of Lemma 2.1 we may assume that nt, nt_1 and nt_2 are all integers. Using the independence of the X_j's,

relation (2.2), Lemma 2.2 and the fact that the function $x^{2-\alpha}$ has a bounded derivative on $[1/2, 3/2]$, we get

$$
\begin{aligned}
\mathsf{E}\left|\frac{1}{A_n}U_n(B_{11})\right|^2\left|\frac{1}{A_n}U_n(B_{21})\right|^2 &= \mathsf{E}\left(\frac{1}{A_n}\sum_{i=nt_1+1}^{nt}\left(X_iI\{s_1H^{-1}(d/n) < |X_i| \le sH^{-1}(d/n)\} - m_i\right)\right)^2 \\
&\quad \times \left(\frac{1}{A_n}\sum_{i=nt+1}^{nt_2}\left(X_iI\{s_1H^{-1}(d/n) < |X_i| \le sH^{-1}(d/n)\} - m_i\right)\right)^2 \\
&= \mathsf{E}\left(\frac{1}{A_n}\sum_{i=nt_1+1}^{nt}\left(X_iI\{s_1H^{-1}(d/n) < |X_i| \le sH^{-1}(d/n)\} - m_i\right)\right)^2 \\
&\quad \times \mathsf{E}\left(\frac{1}{A_n}\sum_{i=nt+1}^{nt_2}\left(X_iI\{s_1H^{-1}(d/n) < |X_i| \le sH^{-1}(d/n)\} - m_i\right)\right)^2 \\
&\le \frac{1}{A_n^4}\left(\sum_{i=nt_1+1}^{nt}\mathsf{E}X_i^2I\{s_1H^{-1}(d/n) < |X_i| \le sH^{-1}(d/n)\}\right) \\
&\quad \times \left(\sum_{i=nt+1}^{nt_2}\mathsf{E}X_i^2I\{s_1H^{-1}(d/n) < |X_i| \le sH^{-1}(d/n)\}\right) \\
&\le C_1(t - t_1)(t_2 - t)(s^{2-\alpha} - s_1^{2-\alpha})^2 \le C_2(t - t_1)(t_2 - t)(s - s_1)^2 \\
&= C_2\mu(B_{11})\mu(B_{21}),
\end{aligned}
\tag{2.14}
$$

where

$$
m_i = m_i(s_1, s) = \mathsf{E}X_iI\{s_1H^{-1}(d/n) < |X_i| \le sH^{-1}(d/n)\}
$$

and C_1, C_2 are positive constants. On the other hand,

$$
\begin{aligned}
\mathsf{E}\left|\frac{1}{A_n}U_n(B_{11})\right|^2\left|\frac{1}{A_n}U_n(B_{12})\right|^2 &= \frac{1}{A_n^4}\mathsf{E}\left(\sum_{i=nt_1+1}^{nt}\left(X_iI\{s_1H^{-1}(d/n) < |X_i| \le sH^{-1}(d/n)\} - m_i^{(s_1,s)}\right)\right)^2 \\
&\quad \times \left(\sum_{i=nt_1+1}^{nt}\left(X_iI\{sH^{-1}(d/n) < |X_i| \le s_2H^{-1}(d/n)\} - m_i^{(s,s_2)}\right)\right)^2 \\
&= \frac{1}{A_n^4}\mathsf{E}\left(\sum_{i=nt_1+1}^{nt}(X_i^{(s_1,s)} - m_i^{(s_1,s)})\right)^2\left(\sum_{i=nt_1+1}^{nt}(X_i^{(s,s_2)} - m_i^{(s,s_2)})\right)^2
\end{aligned}
\tag{2.15}
$$

where we put

$$
X_i^{(u,v)} = X_iI\{uH^{-1}(d/n) < |X_i| \le vH^{-1}(d/n)\}, \quad m_i^{(u,v)} = \mathsf{E}X_i^{(u,v)}.
$$

Expanding the product expectation in (2.15), we get the sum of all expressions

$$
\mathsf{E}(X_i^{(s_1,s)} - m_i^{(s_1,s)})(X_j^{(s_1,s)} - m_j^{(s_1,s)})(X_k^{(s,s_2)} - m_k^{(s,s_2)})(X_\ell^{(s,s_2)} - m_\ell^{(s,s_2)}),
\tag{2.16}
$$

where $nt_1 + 1 \le i, j, k, \ell \le nt$. By the independence of the X_v's, the product expectation in (2.16) equals 0 if one of the i, j, k, ℓ differs from the other three. Thus it suffices to estimate the contribution of the terms where i, j, k, ℓ are pairwise equal, or all are equal. Assume first that $i = j, k = \ell$ and $i \ne k$; the other cases $i = k, j = \ell, i \ne j$ and $i = \ell, j = k, i \ne j$ can be handled similarly as the case $i = j = k = \ell$ below. Then X_i and X_k are independent, and thus using Lemma 2.2, the product expectation (2.16) becomes

$$
\begin{aligned}
\mathsf{E}\left[(X_i^{(s_1,s)} - m_i^{(s_1,s)})^2(X_k^{(s,s_2)} - m_k^{(s,s_2)})^2\right] &= \mathsf{E}(X_i^{(s_1,s)} - m_i^{(s_1,s)})^2\mathsf{E}(X_k^{(s,s_2)} - m_k^{(s,s_2)})^2 \\
&\le \mathsf{E}(X_i^{(s_1,s)})^2\mathsf{E}(X_k^{(s,s_2)})^2 \sim \frac{\alpha^2}{(2-\alpha)^2}(s^{2-\alpha} - s_1^{2-\alpha})(s_2^{2-\alpha} - s^{2-\alpha})\ell^4(d/n)(n/d)^{(4-2\alpha)/\alpha} \\
&\le C_3(s - s_1)(s_2 - s)\ell^4(d/n)(n/d)^{(4-2\alpha)/\alpha}.
\end{aligned}
\tag{2.17}
$$

The number of such pairs (i, k) is at most $(nt - nt_1)^2$ and thus dividing by A_n^4 and using (2.2) we get that the contribution of such terms (2.16) is not greater than

$$
C_4(t - t_1)^2(s - s_1)(s_2 - s) = C_4\mu(B_{11})\mu(B_{12}).
$$

Consider now the case $i = j = k = \ell$. In this case (2.16) becomes, expanding and introducing new letters to lighten the notations,

$$
\begin{aligned}
\mathsf{E}\left[(X_i^{(s_1,s)} - m_i^{(s_1,s)})^2 (X_i^{(s,s_2)} - m_i^{(s,s_2)})^2\right] &= \mathsf{E}(\xi - m^{(1)})^2 (\eta - m^{(2)})^2 \\
&= \mathsf{E}\xi^2\eta^2 - 2m^{(2)}\mathsf{E}\xi^2\eta + (m^{(2)})^2\mathsf{E}\xi^2 - 2m^{(1)}\mathsf{E}\xi\eta^2 + 4m^{(1)}m^{(2)}\mathsf{E}\xi\eta \\
&\quad - 2m^{(1)}(m^{(2)})^2\mathsf{E}\xi + (m^{(1)})^2\mathsf{E}\eta^2 - 2(m^{(1)})^2 m^{(2)}\mathsf{E}\eta + (m^{(1)})^2(m^{(2)})^2,
\end{aligned}
\tag{2.18}
$$

where

$$
\xi = X_i^{(s_1,s)}, \qquad \eta = X_i^{(s,s_2)}, \qquad m^{(1)} = \mathsf{E}\xi, \qquad m^{(2)} = \mathsf{E}\eta.
$$

Clearly ξ and η have disjoint support and thus $\xi\eta = 0$, showing that the first, second, fourth and fifth term of the last sum in (2.18) are equal to 0. Thus the sum equals

$$
\begin{aligned}
&(m^{(2)})^2\mathsf{E}\xi^2 - 2m^{(1)}(m^{(2)})^2\mathsf{E}\xi + (m^{(1)})^2\mathsf{E}\eta^2 - 2(m^{(1)})^2 m^{(2)}\mathsf{E}\eta + (m^{(1)})^2(m^{(2)})^2 \\
&= (m^{(2)})^2\mathsf{E}\xi^2 - 2(m^{(1)})^2(m^{(2)})^2 + (m^{(1)})^2\mathsf{E}\eta^2 - 2(m^{(1)})^2(m^{(2)})^2 + (m^{(1)})^2(m^{(2)})^2.
\end{aligned}
$$

By the Cauchy–Schwarz inequality we have $(m^{(1)})^2 \leq \mathsf{E}\xi^2$, $(m^{(2)})^2 \leq \mathsf{E}\eta^2$ and thus the absolute value of the last sum is at most $7\mathsf{E}(\eta^2)E(\xi^2)$, which, apart from the coefficient, is exactly the third expression in (2.17), leading to the same estimate as there. The number of choices for i in (2.18) is $nt - nt_1 \leq (nt - nt_1)^2$, so for the contribution of all terms in (2.18) we get the same estimate as for (2.17), i.e. $C_5 \mu(B_{11})\mu(B_{12})$. Thus we proved (2.13) for $B_{ij} = B_{12}$ and the proof of Theorem 1.2 is completed.

Acknowledgments

References

Berkes, I., Horváth, L., 2012. The central limit theorem for sums of trimmed variables with heavy tails. Stochastic Process. Appl. 122, 449–465.

Berkes, I., Horváth, L., Schauer, J., 2011. Asymptotics of trimmed CUSUM statistics. Bernoulli 17, 1344–1367.

Bickel, P.J., Wichura, M.J., 1971. Convergence criteria for multiparameter stochastic processes and some applications. Ann. Math. Statist. 42, 1656–1670.

Billingsley, P., 1968. Convergence of Probability Measures. Wiley, New York.

Bingham, N.H., Goldie, C.M., Teugels, J.L., 1987. Regular Variation. Cambridge University Press.

Csörgő, S., Csörgő, M., Horváth, L., Mason, D., 1986a. Normal and stable convergence of integral functions of the empirical distribution function. Ann. Probab. 14, 86–118.

Csörgő, S., Horváth, L., Mason, D., 1986b. What portion of the sample makes a partial sum asymptotically stable or normal? Z. Wahrscheinlichkeitstheor. Verwandte Geb. 72, 1–16.

Feller, W., 1971. An Introduction to Probability Theory and its Applications, Vol. II, second ed. Wiley, New York.

Griffin, P.S., Pruitt, W.E., 1987. The central limit problem for trimmed sums. Math. Proc. Cambridge Philos. Soc. 102, 329–349.

Kiefer, J., 1972. Iterated logarithm analogues for sample quantiles when $p_n \to 0$. In: Proceedings of the Sixth Berkeley Symposium on Mathematical Statistics and Probability, Vol. I. University of California Press, pp. 227–244.

Le Page, R., Woodrofe, M., Zinn, J., 1981. Convergence to a stable distribution via order statistics. Ann. Probab. 9, 624–632.