Asymptotics of trimmed CUSUM statistics

István Berkes[a], Lajos Horváth[b], Johannes Schauer[c]

[a],[c]*Institute of Statistics, Graz University of Technology, Münzgrabenstraße 11, A-8010 Graz, Austria*
[b]*Department of Mathematics, University of Utah, 155 South 1440 East, Salt Lake City, UT 84112-0090 USA*

Abstract

There is a wide literature on change point tests, but the case of variables with infinite variances is essentially unexplored. In this paper we address this problem by studying the asymptotic behavior of trimmed CUSUM statistics. We show that in a location model with i.i.d. errors in the domain of attraction of a stable law of parameter $0 < \alpha < 2$, the appropriately trimmed CUSUM process converges weakly to a Brownian bridge. Thus after moderate trimming, the classical method for detecting changepoints remains valid also for populations with infinite variance. We note that according to the classical theory, the partial sums of trimmed variables are generally not asymptotically normal and using random centering in the test statistics is crucial in the infinite variance case. We also show that the partial sums of truncated and trimmed random variables have different asymptotic behavior. Finally, we discuss resampling procedures enabling one to determine critical values in case of small and moderate sample sizes.

Keywords: stable distributions, trimming, change point, weak convergence, resampling

MSC 2000: 62G10, 62G20, 60F17

1 Introduction

In this paper we are interested in detecting a possible change in the location of independent observations. We observe $X_1, \ldots, X_n$ and want to test the no change null hypothesis

a) Supported by FWF grant S9603-N23 and OTKA grants K 67961 and K 81928
b) Partially supported by NSF grant DMS 00905400
c) Partially supported by FWF grant S9603-N23

$H_0 \;:\; X_1, X_2, \ldots, X_n$ are independent, identically distributed random variables

against the r changes alternative

$$H_A : X_j = \begin{cases} e_j & 1 \le j \le n_1 \\ e_j + c_1 & n_1 < j \le n_2 \\ e_j + c_2 & n_2 < j \le n_3 \\ \vdots & \\ e_j + c_r & n_r < j \le n. \end{cases}$$

It is assumed that

$$e_1, \ldots, e_n \text{ are independent, identically distributed random variables,} \tag{1.1}$$

and $c_0 = 0$, $c_i \neq c_{i+1}$, $i = 0, \ldots, r-1$, and $1 < n_1 < n_2 < \ldots < n_r < n$ are unknown. In our model the changes are at time n_j, $1 \le j \le r$. Testing H_0 against H_A has been considered by several authors. For surveys we refer to Brodsky and Darkhovsky [7], Chen and Gupta [8] and Csörgő and Horváth [11]. If the observations have finite expected value, the model is referred to as changes in the mean.

Several of the most popular methods are based on the functionals of the CUSUM process (tied down partial sums)

$$M_n(t) = \sum_{j=1}^{\lfloor nt \rfloor} X_j - \frac{\lfloor nt \rfloor}{n} \sum_{j=1}^{n} X_j.$$

If H_0 holds and $0 < \sigma^2 = \operatorname{var} X_1 < \infty$, then

$$\frac{1}{\sqrt{n}} M_n(t) \xrightarrow{\mathcal{D}[0,1]} \sigma B(t), \tag{1.2}$$

where $\{B(t), 0 \le t \le 1\}$ is a Brownian bridge. If $\hat{\sigma}_n$ is a weakly consistent estimator for σ, i.e., $\hat{\sigma}_n \to \sigma$ in probability, then

$$\frac{1}{\hat{\sigma}_n \sqrt{n}} M_n(t) \xrightarrow{\mathcal{D}[0,1]} B(t). \tag{1.3}$$

Functionals of (1.3) can be used to find asymptotically distribution free procedures to test H_0 against H_A. The limit results in (1.2) and (1.3) have been extended into several directions. Due to applications in economics, finance, meteorology, environmental sciences and quality control, several authors studied the properties of $M_n(t)$ and especially (1.3) for dependent observations. For relevant references we refer to Horváth and Steinebach [20]. The case of vector-valued dependent observations is considered in Horváth, Kokoszka and Steinebach [19]. We mention that in case of dependent observations $\sigma^2 = \lim_{n\to\infty} \operatorname{var}\left(n^{-1/2} \sum_{j=1}^{n} X_j\right)$, so the estimation of σ is considerably harder than in the i.i.d. case (cf. Bartlett [3], Grenander and Rosenblatt [13], Parzen [30]). The rate of convergence in (1.3) may be slow, so the asymptotic critical values might be misleading. Hence resampling methods have been advocated in Hušková [21]. With very few exceptions, it has been assumed that at least EX_j^2 is finite. In this paper we are interested in testing H_0 against H_A, when $EX_j^2 = \infty$.

We assume that

$$X_1, X_2, \ldots \text{ belong to the domain of attraction of a stable random variable } \xi_\alpha \text{ with parameter } 0 < \alpha < 2 \tag{1.4}$$

and

$$X_j \text{ is symmetric when } \alpha = 1. \tag{1.5}$$

This means that

$$\left(\sum_{j=1}^{n} X_j - a_n \right) \Big/ b_n \xrightarrow{\mathcal{D}} \xi_\alpha \tag{1.6}$$

for some numerical sequences a_n and b_n. The necessary and sufficient condition for (1.6) is

$$\lim_{t\to\infty} \frac{P\{X_1 > t\}}{L(t)t^{-\alpha}} = p \quad \text{and} \quad \lim_{t\to\infty} \frac{P\{X_1 \le -t\}}{L(t)t^{-\alpha}} = q \tag{1.7}$$

for some numbers $p \ge 0$, $q \ge 0$, $p + q = 1$ and L, a slowly varying function at ∞.

Aue et al. [2] studied the properties of $M_n(t)$ under conditions H_0, (1.4) and (1.5). They used $\max_{1\le j\le n} |X_j|$ as the normalization of $M_n(t)$ and showed that

$$\frac{1}{\gamma_n} M_n(t) \xrightarrow{\mathcal{D}[0,1]} \frac{1}{\mathcal{Z}} B_\alpha(t), \qquad \gamma_n = \max_{1\le j\le n} |X_j|. \tag{1.8}$$

Here $B_\alpha(t) = W_\alpha(t) - tW_\alpha(1)$ is an α-stable bridge, $W_\alpha(t)$ is an α-stable process (cf. also Kasahara and Watanabe [22], Section 9) and $\mathcal{Z}$ is a random norming factor whose joint distribution with $W_\alpha(t)$ is described in [2] explicitly. Nothing is known about the distribution of the functionals of $B_\alpha(t)/\mathcal{Z}$ and therefore it is nearly impossible to determine critical values needed to construct asymptotic test procedures. Hence resampling methods (bootstrap and permutation) have been tried. However, it was proved that the conditional distribution of the resampled $M_n(t)/\gamma_n$, given $X_1, \ldots, X_n$, converges in distribution to a non-degenerate random process depending also on the trajectory $(X_1, X_2, \ldots)$. So resampling cannot be recommended to get asymptotic critical values. This result was obtained by Aue et al. [2] for permutation resampling and Athreya [1], Hall [18] and Berkes et al. [4] for the bootstrap. No efficient procedure has been found to test H_0 against H_A when $EX_j^2 = \infty$.

The reason for the "bad" behavior of the CUSUM statistics described above is the influence of the large elements of the sample. It is known that for i.i.d. random variables $X_1, X_2, \ldots$ in the domain of attraction of a nonnormal stable law, the j-th largest element of $|X_1|, \ldots, |X_n|$ has, for any fixed j, the same order of magnitude as the sum $S_n = X_1 + \ldots + X_n$ as $n \to \infty$. Thus the influence of the large elements in the CUSUM functional does not become negligible as $n \to \infty$ and consequently, the limiting behavior of the CUSUM statistics along different trajectories $(X_1, X_2, \ldots)$ is different, rendering this statistics unpractical for statistical inference. The natural remedy for this trouble is trimming, i.e. removing the $d(n)$ elements with the largest absolute values from the sample, where $d(n)$ is a suitable number with $d(n) \to \infty$, $d(n)/n \to 0$. This type of trimming is usually called modulus trimming in the literature. In another type of trimming, some of the largest and smallest order statistics are removed from the sample, see e.g. Csörgő et al. [10], [12]. Under suitable conditions, trimming leads indeed to a better asymptotic behavior of partial sums, see e.g. Mori [27], [28], [29], Maller [25], [26], Csörgő et al. [9], [10], [12], Griffin and Pruitt [14], [15], Haeusler and Mason [16], [17]. Note, however, that the asymptotic properties of trimmed random variables depend strongly on the type of trimming used. In this paper trimming means modulus trimming as introduced above. Griffin and Pruitt [14] showed that even in the case when the X_j belong to the domain of attraction of a symmetric stable law with parameter $0 < \alpha < 2$, the modulus trimmed partial sums need not be asymptotically normal. Theorem 1.5 reveals the reason of this surprising fact: for nonsymmetric distributions F the center of the sample remains, even after modulus trimming, a nondegenerate random variable and no nonrandom centering can lead to a CLT. In contrast, a suitable random centering will always work and since the CUSUM functional is not affected by centering factors, even in the case of "bad" partial sum behavior, the trimmed CUSUM functional converges to Brownian bridge, resulting in a simple and useful change point test.

To formulate our results, consider the trimmed CUSUM process

$$T_n(t) = \sum_{j=1}^{\lfloor nt \rfloor} X_j I\{|X_j| \le \eta_{n,d}\} - \frac{\lfloor nt \rfloor}{n} \sum_{j=1}^{n} X_j I\{|X_j| \le \eta_{n,d}\}, \quad 0 \le t \le 1,$$

where $\eta_{n,d}$ is the d^{th} largest value among $|X_1|, \ldots, |X_n|$.

Let

$$F(t) = P\{X_1 \le t\} \quad \text{and} \quad H(t) = P\{|X_1| > t\}.$$

$H^{-1}(t)$ denotes the (generalized) inverse (or quantile) of H. We assume that

$$\lim_{n\to\infty} d(n)/n = 0 \tag{1.9}$$

and

$$\lim_{n\to\infty} d(n)/(\log n)^{7+\varepsilon} = \infty \quad \text{with some } \varepsilon > 0. \tag{1.10}$$

For the sake of simplicity (cf. Mori [27]) we also require that

$$F \text{ is continuous.} \tag{1.11}$$

Let

$$A_n^2 = \frac{\alpha}{2-\alpha}\left(H^{-1}(d/n)\right)^2 d. \tag{1.12}$$

Our first result states the weak convergence of $T_n(t)/A_n$.

Theorem 1.1. *If* H_0, (1.4), (1.5) *and* (1.9)–(1.11) *hold, then*

$$\frac{1}{A_n} T_n(t) \xrightarrow{\mathcal{D}[0,1]} B(t), \tag{1.13}$$

where $\{B(t), 0 \le t \le 1\}$ *is a Brownian bridge.*

Since A_n is unknown, we need to estimate it from the sample. We will use

$$\hat{A}_n^2 = \sum_{j=1}^{n} \left(X_j I\{|X_j| \le \eta_{n,d}\} - \bar{X}_{n,d}\right)^2 \quad \text{and} \quad \hat{\sigma}_n^2 = \frac{1}{n}\hat{A}_n^2,$$

where

$$\bar{X}_{n,d} = \frac{1}{n}\sum_{j=1}^{n} X_j I\{|X_j| \le \eta_{n,d}\}.$$

We note that $\hat{A}_n/A_n \to 1$ almost surely (cf. Lemma 4.7).

Theorem 1.2. *If the conditions of Theorem 1.1 are satisfied, then*

$$\frac{1}{\hat{\sigma}_n\sqrt{n}} T_n(t) \xrightarrow{\mathcal{D}[0,1]} B(t). \tag{1.14}$$

In case of independence and $0 < \sigma^2 = \operatorname{var} X_j < \infty$ we estimate σ^2 by the sample variance. So the comparison of (1.3) and (1.14) reveals that in case of $EX_j^2 = \infty$ we still use the classical CUSUM procedure; only the extremes are removed from the sample. The finite sample properties of tests for H_0 against H_A based on (1.14) are investigated in Section 3.

In case of a given sample, it is difficult to decide if EX_j^2 is finite or infinite. Thus for applications it is important to establish Theorem 1.2 when $EX_j^2 < \infty$.

Theorem 1.3. *If* H_0, (1.9), (1.10) *and* $EX_j^2 < \infty$ *are satisfied, then* (1.14) *holds.*

Putting together Theorems 1.2 and 1.3 we see that the CUSUM based procedures can always be used if the observations with the largest absolute values are removed from the sample.

Now we outline the basic idea of the proofs of Theorems 1.1 and 1.2. It was proved by Kiefer [23] (cf. Shorack and Wellner [33]) that $\eta_{n,d}$ is close to $H^{-1}(d/n)$

and thus it is natural to consider the process obtained from $T_n(t)$ by replacing $\eta_{n,d}$ with $H^{-1}(d/n)$. Let

$$V_n(t) = \sum_{j=1}^{\lfloor nt \rfloor} \left(X_j I\{|X_j| \le H^{-1}(d/n)\} - E(X_j I\{|X_j| \le H^{-1}(d/n)\}) \right)$$

and

$$V_n^*(t) = \sum_{j=1}^{\lfloor nt \rfloor} \left(X_j I\{|X_j| \le \eta_{n,d}\} - E(X_j I\{|X_j| \le \eta_{n,d}\}) \right).$$

Since $V_n(t)$ is a sum of i.i.d. random variables, the classical functional CLT for triangular arrays easily yields

Theorem 1.4. *If the conditions of Theorem 1.1 are satisfied, then*

$$\frac{1}{A_n} V_n(t) \xrightarrow{\mathcal{D}[0,1]} W(t),$$

where $\{W(t), 0 \le t \le 1\}$ is a standard Brownian motion (Wiener process).

In view of the closeness of $\eta_{n,d}$ and $H^{-1}(d/n)$, one would expect that the asymptotic behavior of $V_n(t)/A_n$ and $V_n^*(t)/A_n$ is the same. Surprisingly, this is not the case. Let

$$m(t) = E\left[X_1 I\{|X_1| \le t\} - X_1 I\{|X_1| \le H^{-1}(d/n)\}\right], \quad t \ge 0.$$

Theorem 1.5. *If the conditions of Theorem 1.1 are satisfied, then*

$$\frac{1}{A_n} \max_{1 \le k \le n} \left| \sum_{j=1}^{k} \left[X_j \left(I\{|X_j| \le \eta_{n,d}\} - I\{|X_j| \le H^{-1}(d/n)\} \right) - m(\eta_{n,d}) \right] \right| = o_P(1).$$

By Theorem 1.5, the asymptotic properties of the partial sums of trimmed and truncated variables would be the same if $n|m(\eta_{n,d})| = o_P(A_n)$ were true. However, this is not always the case as the following example shows.

Example 1.1. *Assume that X_1 is concentrated on $(0, +\infty)$ and has a continuous density f which is regularly varying at ∞ with exponent $-(\alpha+1)$ with some $0 < \alpha < 2$. Then*

$$\frac{nm(\eta_{n,d})}{B_n} \xrightarrow{\mathcal{D}} N(0,1),$$

where

$$B_n = \frac{\alpha d^{3/2}}{nH'(H^{-1}(d/n))}.$$

We conjecture that the centering factor $nm(\eta_{n,d})/A_n$ and the partial sum process

$$\sum_{j=1}^{\lfloor nt \rfloor}(X_j I\{|X_j| \le H^{-1}(d/n)\} - E(X_j I\{|X_j| \le H^{-1}(d/n)\})), \quad 0 \le t \le 1,$$

are asymptotically independent under the conditions of Example 1.1. Hence by Theorem 1.5 one would have

$$\frac{1}{A_n}\sum_{j=1}^{\lfloor nt \rfloor}(X_j I\{|X_j| \le \eta_{n,d}\} - c_n) \xrightarrow{\mathcal{D}[0,1]} W(t) + t\left(\frac{2-\alpha}{\alpha}\right)^{1/2}\xi,$$

where $\{W(t), 0 \le t \le 1\}$ and ξ are independent, $W(t)$ is a standard Wiener process, ξ is a standard normal random variable and $c_n = EX_1 I\{|X_1| \le H^{-1}(d/n)\}$.

In view of Theorem 1.5, the normed partial sum processes of $X_j I\{|X_j| \le \eta_{n,d}\} - m(\eta_{n,d})$ and $X_j I\{|X_j| \le H^{-1}(d/n)\}$ have the same asymptotic behavior and thus the same holds for the corresponding CUSUM processes. By Theorem 1.4, the CUSUM process of $X_j I\{|X_j| \le H^{-1}(d/n)\}$ converges weakly to the Brownian bridge and the CUSUM process of $X_j I\{|X_j| \le \eta_{n,d}\} - m(\eta_{n,d})$ clearly remains the same if we drop the term $m(\eta_{n,d})$. Formally,

$$\begin{aligned}\max_{1\le k\le n}&\Bigg|\sum_{j=1}^{k} X_j I\{|X_j| \le \eta_{n,d}\} - \frac{k}{n}\sum_{j=1}^{n} X_j I\{|X_j| \le \eta_{n,d}\}\\ &-\left(\sum_{j=1}^{k} X_j I\{|X_j| \le H^{-1}(d/n)\} - \frac{k}{n}\sum_{j=1}^{n} X_j I\{|X_j| \le H^{-1}(d/n)\}\right)\Bigg| \qquad (1.15)\\ \le 2\max_{1\le k\le n}&\left|\sum_{j=1}^{k}\left[X_j\left(I\{|X_j| \le \eta_{n,d}\} - I\{|X_j| \le H^{-1}(d/n)\}\right) - m(\eta_{n,d})\right]\right|.\end{aligned}$$

Thus, even though the partial sums of trimmed and truncated variables are asymptotically different due to the presence of the random centering $m(\eta_{n,d})$, the asymptotic distributions of the CUSUM processes of the trimmed and truncated variables are the same.

The proofs of the asymptotic results for $\sum_{j=1}^{n} X_j I\{|X_j| \le \eta_{n,d}\}$ in Mori [27], [28], [29], Maller [25], [26] and Griffin and Pruitt [14], [15] are based on classical probability theory. Csörgő et al. [9], [10], [12] and Haeusler and Mason [16], [17] use the weighted approximation of quantile processes to establish the normality of a class of trimmed partial sums. The method of our paper is completely different. We show in Theorem 1.5 that after a suitable random centering, trimmed partial sums can be replaced with truncated ones, reducing the problem to sums of i.i.d. r.v.'s.

2 Resampling methods

Since the convergence in Theorem 1.1 can be slow, critical values in the change point test determined on the basis of the limit distribution may not be appropriate for small sample sizes. To resolve this difficulty, resampling methods can be used to simulate critical values. Let

$$x_j = X_j I\{|X_j| \le \eta_{n,d}\} - \bar{X}_{n,d}, \quad 1 \le j \le n$$

be the trimmed and centered observations. We select m elements from the set $\{x_1, x_2, \dots, x_n\}$ randomly (with or without replacement), resulting in the sample $y_1, \dots, y_m$. If we select with replacement, the procedure is the bootstrap; if we select without replacement and $m = n$, this is the permutation method (cf. Hušková [21]). Now we define the resampled CUSUM process

$$T_{m,n}(t) = \sum_{j=1}^{\lfloor mt \rfloor} y_j - \frac{\lfloor mt \rfloor}{m} \sum_{1 \le j \le m} y_j.$$

We note that conditionally on $X_1, X_2, \dots, X_n$, the mean of y_j is 0 and its variance is $\hat{\sigma}_n^2$.

Theorem 2.1. *Assume that the conditions of Theorem 1.1 are satisfied and draw $m = m(n)$ elements $y_1, \dots, y_m$ from the set $\{x_1, \dots, x_n\}$ with or without replacement, where*

$$m = m(n) \to \infty \quad as \quad n \to \infty \tag{2.1}$$

and $m(n) \le n$ in case of selection without replacement. Then for almost all realizations of $X_1, X_2, \dots$ we have

$$\frac{1}{\hat{\sigma}_n \sqrt{m}} T_{m,n}(t) \xrightarrow{\mathcal{D}[0,1]} B(t),$$

where $\{B(t), 0 \le t \le 1\}$ is a Brownian bridge.

By the results of Aue et al. [2] and Berkes et al. [4], if we sample from the original (untrimmed) observations, then the CUSUM process converges weakly to a non-Gaussian process containing random parameters and thus the resampling procedure is statistically useless.

If we use resampling to determine critical values in the CUSUM test, we need to study the limit also under the the alternative, since in a practical situation we do not know which of H_0 or H_A is valid. As before, assume that the error terms $\{e_j\}$ are in the domain of attraction of a stable law, i.e.

$$\lim_{t \to \infty} \frac{P\{e_1 > t\}}{L(t) t^{-\alpha}} = p \quad \text{and} \quad \lim_{t \to \infty} \frac{P\{e_1 \le -t\}}{L(t) t^{-\alpha}} = q, \tag{2.2}$$

where $p \ge 0$, $q \ge 0$, $p + q = 1$ and L is a slowly varying function at ∞.

Theorem 2.2. *If H_A, (1.1), (1.9)–(1.11), (2.1) and (2.2) hold, then for almost all realizations of $X_1, X_2, \ldots$ we have that*

$$\frac{1}{\hat{\sigma}_n \sqrt{m}} T_{m,n}(t) \xrightarrow{\mathcal{D}[0,1]} B(t),$$

where $\{B(t), 0 \le t \le 1\}$ is a Brownian bridge.

In other words, the limiting distribution of the trimmed CUSUM process is the same under H_0 and H_A and thus the critical values determined by resampling will always work. On the other hand, under H_A the test statistic $\sup_{0<t<1} |T_n(t)|/A_n$ goes to infinity, so using the critical values determined by resampling, we get a consistent test.

We note that Theorems 2.1 and 2.2 remain true if (1.6) is replaced with $EX_j^2 < \infty$. The proofs are similar to that of Theorems 2.1 but much simpler, so no details are given.

3 Simulation study

Consider the model under H_0 with i.i.d. random variables X_j, $j = 1, \ldots, n$, having distribution function

$$F(t) = \begin{cases} q(1-t)^{-1.5} & \text{for } t \le 0 \\ 1 - p(1+t)^{-1.5} & \text{for } t > 0, \end{cases}$$

where $p \ge 0$, $q \ge 0$, $p + q = 1$. We trim the samples using $d(n) = \lfloor n^{0.3} \rfloor$. To simulate the critical values we generate $N = 10^5$ Monte Carlo simulations for each $n \in \{100, 200, 400, 800\}$ according to the model under the no change hypothesis and calculate the values of $\sup_{0<t<1} |T_n(t)|/(\hat{\sigma}_n \sqrt{n})$, where $T_n(t)$ and $\hat{\sigma}_n$ are defined in Section 1. The computation of the empirical quantiles yields the estimated critical values. Table 1 summarizes the results for $p = q = 1/2$ and $1 - \alpha = 0.95$.

$n = 100$	$n = 200$	$n = 400$	$n = 800$	$n = \infty$
1.244	1.272	1.299	1.312	1.358

Table 1: Simulated critical values of $\sup_{0<t<1} |T_n(t)|/(\hat{\sigma}_n \sqrt{n})$ for $1 - \alpha = 0.95$

Figure 1 shows the empirical power of the test of H_0 against H_A based on the statistic $\sup_{0<t<1} |T_n(t)|/(\hat{\sigma}_n \sqrt{n})$ for a single change at time $k = n_1 \in \{n/4, n/2, 3n/4\}$ and each $c_1 \in \{-3, -2.9, \ldots, 2.9, 3\}$ for the same trimming as above ($d(n) = \lfloor n^{0.3} \rfloor$) and a significance level of $1 - \alpha = 0.95$, where the number of repetitions is $N = 10^4$. Note that depending on the sample size we used different simulated quantiles (see

Table 1). The power behaves best for a change-point in the middle of the observation period ($k = n/2$). Due to the differences between the simulated and asymptotic critical values in Table 1, especially for small n the test based on the asymptotic critical values tends to be conservative.

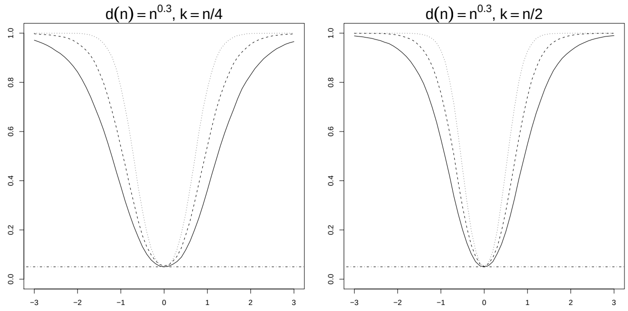

Figure 1: Empirical power curves with $\alpha = 0.05$, $n = 100$ (solid), $n = 200$ (dashed) and $n = 400$ (dotted)

4 Proofs

Throughout this section we assume that H_0 holds. Clearly,

$$H(x) = 1 - F(x) + F(-x), \quad x \geq 0,$$

and by (1.7) we have that

$$H^{-1}(t) = t^{-1/\alpha}K(t), \quad \text{if } t \leq t_0, \tag{4.1}$$

where $K(t)$ is a slowly varying function at 0. We also use

$$d = d(n) \to \infty. \tag{4.2}$$

Lemma 4.1. *If H_0, (1.4), (1.5), (1.9) and (4.2) hold, then*

$$\lim_{n\to\infty} \frac{1}{A_n^2} \operatorname{var} V_n(1) = 1 \tag{4.3}$$

and

$$\lim_{n\to\infty} \sum_{j=1}^{n} E\Big[X_j I\{|X_j| \leq H^{-1}(d/n)\} - E\left[X_j I\{|X_j| \leq H^{-1}(d/n)\}\right]\Big]^4 \times \frac{1}{d(H^{-1}(d/n))^4} = \frac{\alpha}{4-\alpha}. \tag{4.4}$$

Proof. If $1 < \alpha < 2$, then

$$\lim_{n\to\infty} EX_1 I\{|X_1| \le H^{-1}(d/n)\} = EX_1.$$

If $\alpha = 1$, then by the assumed symmetry $EX_1 I\{|X_1| \le H^{-1}(d/n)\} = 0$. In case of $0 < \alpha < 1$ we write

$$\begin{aligned} E|X_1| I\{|X_1| \le H^{-1}(d/n)\} &= \int_{-H^{-1}(d/n)}^{H^{-1}(d/n)} |x| dF(x) \\ &= -\int_0^{H^{-1}(d/n)} x dH(x) \\ &= -xH(x)\big|_{H^{-1}(d/n)} + \int_0^{H^{-1}(d/n)} H(x)dx. \end{aligned}$$

By Bingham et al. [6] (p. 26)

$$\lim_{y\to\infty} \frac{\displaystyle\int_0^y H(x)dx}{\dfrac{1}{1-\alpha} y^{1-\alpha} L(y)} = 1$$

and therefore

$$\lim_{n\to\infty} \frac{E|X_1| I\{|X_1| \le H^{-1}(d/n)\}}{\dfrac{\alpha}{1-\alpha} H^{-1}(d/n) d/n} = 1.$$

Similarly,

$$EX_1^2 I\{|X_1| \le H^{-1}(d/n)\} = \int_{-H^{-1}(d/n)}^{H^{-1}(d/n)} x^2 dF(x)$$

$$= -\int_0^{H^{-1}(d/n)} x^2 dH(x) = -x^2 H(x)\big|_{H^{-1}(d/n)} + 2\int_0^{H^{-1}(d/n)} xH(x)dx.$$

Using again Bingham et al. [6] (p. 26), we conclude that

$$\lim_{n\to\infty} \frac{EX_1^2 I\{|X_1| \le H^{-1}(d/n)\}}{\left(H^{-1}(d/n)\right)^2 d/n} = \frac{\alpha}{2-\alpha}.$$

Hence (4.3) is established.

Arguing as above we get

$$\begin{aligned} EX_1^4 I\{|X_1| \le H^{-1}(d/n)\} &= -\int_0^{H^{-1}(d/n)} x^4 dH(x) \\ &= -x^4 H(x)\big|_{H^{-1}(d/n)} + 4\int_0^{H^{-1}(d/n)} x^3 H(x)dx \end{aligned}$$

and therefore

$$\lim_{n\to\infty} \frac{EX_1^4 I\{|X_1| \le H^{-1}(d/n)\}}{(H^{-1}(d/n))^4\, d/n} = \frac{\alpha}{4-\alpha}.$$

Similarly,

$$\lim_{n\to\infty} \frac{E|X_1|^3 I\{|X_1| \le H^{-1}(d/n)\}}{(H^{-1}(d/n))^3\, d/n} = \frac{\alpha}{3-\alpha},$$

completing the proof of (4.4). □

Proof of Theorem 1.4. Clearly, for each n $X_j I\{|X_j| \le H^{-1}(d/n)\}$, $1 \le j \le n$, are independent and identically distributed random variables. By Lemma 4.1 we have that

$$\lim_{n\to\infty} \frac{\sum_{j=1}^{n} E\Big[X_j I\{|X_j| \le H^{-1}(d/n)\} - E\left[X_j I\{|X_j| \le H^{-1}(d/n)\}\right]\Big]^4}{\left(\sum_{j=1}^{n} \mathrm{var}(X_j I\{|X_j| \le H^{-1}(d/n)\})\right)^2} = 0,$$

so the Lyapunov condition is satisfied. Hence the result follows immediately from Skorohod [34]. □

A series of lemmas is needed to establish Theorem 1.5. Let $\eta_{n,1} \ge \eta_{n,2} \ge \ldots \ge \eta_{n,n}$ denote the order statistics of $|X_1|, \ldots, |X_n|$, starting with the largest value.

Lemma 4.2. *If H_0 and (1.11) hold, then*

$$\{H(\eta_{n,k}), 1 \le k \le n\} \stackrel{\mathcal{D}}{=} \{S_k/S_{n+1}, 1 \le k \le n\},$$

where

$$S_k = e_1 + \ldots + e_k, \quad 1 \le k \le n$$

and $e_1, e_2, \ldots, e_{n+1}$ are independent, identically distributed exponential random variables with $Ee_j = 1$.

Proof. The representation in Lemma 4.2 is well-known (cf., for example, Shorack and Wellner [33], p. 335). □

Let $\eta_{n,d}(j)$ denote the d^{th} largest among $|X_1|, \ldots, |X_{j-1}|, |X_{j+1}|, \ldots, |X_n|$.

Lemma 4.3. *If H_0, (1.4), (1.5), (1.9), (1.11) and (2.1) hold, then*

$$\sum_{j=1}^{n} |X_j\,(I\{|X_j| \le \eta_{n,d}\} - I\{|X_j| \le \eta_{n,d}(j)\})| = o_P(A_n).$$

Proof. First we note that $\eta_{n,d}(j) = \eta_{n,d}$ or $\eta_{n,d}(j) = \eta_{n,d+1}$. Hence

$$\frac{H(\eta_{n,d})}{H(\eta_{n,d}(j))} \geq \frac{H(\eta_{n,d})}{H(\eta_{n,d+1})}.$$

By Lemma 4.2 and the law of large numbers we have

$$\frac{H(\eta_{n,d})}{H(\eta_{n,d+1})} \stackrel{\mathcal{D}}{=} \frac{S_d}{S_{d+1}} = \frac{S_d}{S_d + e_{d+1}} = \frac{1}{1 + e_d/S_d} = 1 + O_P(d^{-1}).$$

Furthermore, by the central limit theorem we conclude

$$S_r = r\left(1 + O_P(r^{-1/2})\right)$$

and thus

$$H(\eta_{n,d}) = \frac{d}{n}\left(1 + O_P\left(d^{-1/2}\right)\right).$$

Hence for every $\varepsilon > 0$, there is a constant $C = C(\varepsilon)$ and an event $A = A(\varepsilon)$ such that $P(A) \geq 1 - \varepsilon$ and on A

$$\frac{H(\eta_{n,d})}{H(\eta_{n,d+1})} \geq 1 - \frac{C}{d} \tag{4.5}$$

and

$$H(\eta_{n,d}) \geq \frac{d}{n}\left(1 - \frac{C}{\sqrt{d}}\right). \tag{4.6}$$

We note that $H(|X_j|)$ is uniformly distributed on $[0,1]$ and is independent of $\eta_{n,d}(j)$. So using (4.5) and (4.6) we obtain that

$$\begin{aligned}
&E\Big[|X_j(I\{|X_j| \leq \eta_{n,d}\} - I\{|X_j| \leq \eta_{n,d}(j)\})|I\{A\}\Big] \\
&= E\Big[|X_j|I\{\eta_{n,d}(j) \leq |X_j| \leq \eta_{n,d}\}I\{A\}\Big] \\
&\leq H^{-1}\left(\frac{d}{n}\left(1 - \frac{C}{\sqrt{d}}\right)\right) E\Big[I\left\{H(\eta_{n,d}) \leq H(|X_j|) \leq H(\eta_{n,d}(j))\right\} I\{A\}\Big] \\
&\leq H^{-1}\left(\frac{d}{n}\left(1 - \frac{C}{\sqrt{d}}\right)\right) EI\left\{H(\eta_{n,d}(j))\left(1 - \frac{C}{d}\right) \leq H(|X_j|) \leq H(\eta_{n,d}(j))\right\} \\
&\leq H^{-1}\left(\frac{d}{n}\left(1 - \frac{C}{\sqrt{d}}\right)\right) EH(\eta_{n,d}(j))\frac{C}{d} \leq H^{-1}\left(\frac{d}{n}\left(1 - \frac{C}{\sqrt{d}}\right)\right)\frac{d+1}{n+1}\frac{C}{d},
\end{aligned}$$

since $H(\eta_{n,d}(j)) \leq H(\eta_{n,d+1})$ and by Lemma 4.2 we have $EH(\eta_{n,d+1}) = (d+1)/(n+1)$. The slow variation and monotonicity of H^{-1} yield

$$\lim_{n\to\infty} \frac{H^{-1}\left(\frac{d}{n}\left(1 - \frac{C}{\sqrt{d}}\right)\right)}{H^{-1}(d/n)} = 1,$$

thus we get that

$$\lim_{n\to\infty}\frac{1}{A_n}\sum_{j=1}^{n}E|X_j(I\{|X_j|\le \eta_{n,d}\}-I\{|X_j|\le \eta_{n,d}(j)\})|I\{A\}=0.$$

Since we can choose $\varepsilon>0$ as small as wish, Lemma 4.3 is proven. □

Lemma 4.4. *If the conditions of Lemma 4.3 are satisfied, then*

$$\frac{1}{A_n}\sum_{j=1}^{n}|m(\eta_{n,d})-m(\eta_{n,d}(j))|=o_P(1).$$

Proof. It can be proven along the lines of the proof of Lemma 4.3. □

Let

$$\xi_j=X_j(I\{|X_j|\le \eta_{n,d}(j)\}-I\{|X_j|\le H^{-1}(n/d)\})-m(\eta_{n,d}(j)).$$

Lemma 4.5. *If the conditions of Theorem 1.1 are satisfied, then there is an $a>0$ such that for all $\tau>1/\alpha$ and $0<\varepsilon<1/2$*

$$E\xi_j=0, \tag{4.7}$$
$$E\xi_j^2=E\xi_1^2=\mathcal{O}\left((H^{-1}(d/n))^2(d^{1/2+\varepsilon}/n)+n^{2\tau}\exp(-ad^{2\varepsilon})\right), \tag{4.8}$$
$$E\xi_i\xi_j=E\xi_1\xi_2=\mathcal{O}\left((H^{-1}(d/n))^2(d^{1/2+3\varepsilon}/n^2)+n^{2\tau}\exp(-ad^{2\varepsilon})\right) \tag{4.9}$$

for $1\le j\le n$ and $1\le i<j\le n$ respectively.

Proof. It follows from the independence of X_j and $\eta_{n,d}(j)$ that

$$E\xi_j=E(E(\xi_j|\eta_{n,d}(j)))=E(m(\eta_{n,d}(j))-m(\eta_{n,d}(j)))=0,$$

so (4.7) is proven.

The first relation in (4.8) is clear. For the second part we note that

$$E\xi_1^2\le 2EX_1^2\left(I\{|X_1|\le \eta_{n,d}(1)\}-I\{|X_1|\le H^{-1}(d/n)\}\right)^2+2Em^2(\eta_{n,d}(1))$$

and

$$\begin{aligned}
&EX_1^2\left(I\{|X_1|\le \eta_{n,d}(1)\}-I\{|X_1|\le H^{-1}(d/n)\}\right)^2\\
&\qquad\le EX_1^2I\{\eta_{n,d}(1)\le |X_1|\le H^{-1}(d/n)\}+EX_1^2I\{H^{-1}(d/n)\le |X_1|\le \eta_{n,d}(1)\}\\
&\qquad\le \left(H^{-1}(d/n)\right)^2P\{\eta_{n,d}(1)\le |X_1|\le H^{-1}(d/n)\}\\
&\qquad\quad+E((\eta_{n,d}(1))^2I\{H(\eta_{n,d}(1))\le H(|X_1|)\le d/n\}).
\end{aligned}$$

There are constants c_1 and c_2 such that

$$P\{|S_d - d| \geq x\sqrt{d}\} \leq \exp(-c_1 x^2), \quad \text{if } 0 \leq x \leq c_2 d. \tag{4.10}$$

Let $0 < \varepsilon < 1/2$. Using Lemma 4.2 and (4.10), there is a constant c_3 such that

$$P(A) \geq 1 - c_3 \exp(-c_1 d^{2\varepsilon}), \tag{4.11}$$

where

$$A = \left\{ \omega : \frac{d}{n}\left(1 - \frac{1}{d^{1/2-\varepsilon}}\right) \leq H(\eta_{n,d}(1)) \leq \frac{d}{n}\left(1 + \frac{1}{d^{1/2-\varepsilon}}\right)\right\}.$$

Let A^c denote the complement of A. By (4.11) we have

$$\begin{aligned}
&(H^{-1}(d/n))^2 P\{\eta_{n,d}(1) \leq |X_1| \leq H^{-1}(d/n)\} \\
&= (H^{-1}(d/n))^2 \left(P(A^c) + P\{\eta_{n,d}(1) \leq |X_1| \leq H^{-1}(d/n), A\}\right) \\
&\leq (H^{-1}(d/n))^2 \left(c_3 \exp(-c_1 d^{2\varepsilon}) + P\left\{\frac{d}{n} \leq H(|X_1|) \leq \frac{d}{n}\left(1 + \frac{1}{d^{1/2-\varepsilon}}\right)\right\}\right) \\
&= \mathcal{O}\left((H^{-1}(d/n))^2 \left(\exp(-c_1 d^{2\varepsilon}) + \frac{d^{1/2+\varepsilon}}{n}\right)\right).
\end{aligned}$$

Similarly, by the independence of $|X_1|$ and $\eta_{n,d}(1)$ we have

$$\begin{aligned}
&E((\eta_{n,d}(1))^2 I\{H(\eta_{n,d}(1)) \leq H(|X_1|) \leq d/n\}) \\
&\leq E(\eta^2_{n,d}(1) I\{A^c\}) \\
&\qquad + E\left(\left(H^{-1}(d/n(1 - d^{\varepsilon-1/2}))\right)^2 I\{d/n(1 - d^{\varepsilon-1/2}) \leq H(|X_1|) \leq d/n\}\right) \\
&= E(\eta^2_{n,d}(1) I\{A^c\}) + \left(H^{-1}(d/n(1 - d^{\varepsilon-1/2}))\right)^2 \frac{d}{n} d^{\varepsilon-1/2}.
\end{aligned}$$

Since $H^{-1}(t)$ is a regularly varying function at 0 with index $-1/\alpha$, for any $\tau > 1/\alpha$ there is a constant c_4 such that

$$H^{-1}(t) \leq c_4 t^{-\tau}, \quad 0 < t \leq 1. \tag{4.12}$$

By the Cauchy-Schwarz inequality we have

$$E\eta^2_{n,d}(1) I\{A^c\} \leq (E\eta^4_{n,d}(1))^{1/2} (P(A^c))^{1/2} \leq (E\eta^4_{n,d}(1))^{1/2} c_3^{1/2} \exp\left(-\frac{c_1}{2} d^{2\varepsilon}\right).$$

Next we use (4.12) and Lemma 4.2 to conclude

$$\begin{aligned}
E\eta^4_{n,d}(1) \leq E\eta^4_{n,d} &\leq c_4^4 E\left(\frac{S_d}{S_{n+1}}\right)^{-4\tau} = c_4^4 E\left(1 + \frac{S_{n+1} - S_d}{S_d}\right)^{4\tau} \\
&\leq c_5 \left(1 + E(S_{n+1} - S_d)^{4\tau} E\frac{1}{S_d^{4\tau}}\right) \leq c_6 n^{4\tau}
\end{aligned} \tag{4.13}$$

since S_d has a Gamma distribution with parameter d and therefore $ES_d^{-4\tau} < \infty$ if $d \geq d_0(\tau)$. Thus we have that

$$\begin{aligned} EX_1^2(I\{|X_1| \leq \eta_{n,d}(1)\} - I\{|X_1| \leq H^{-1}(d/n)\})^2 \\ = \mathcal{O}\left((H^{-1}(d/n))^2 \left(d^{\varepsilon+1/2}/n\right) + n^{2\tau} \exp\left(-\frac{c_1}{2} d^{2\varepsilon}\right)\right). \end{aligned}$$

Similar arguments give

$$Em^2(\eta_{n,d}(1)) = \mathcal{O}\left((H^{-1}(d/n))^2 \left(d^{\varepsilon+1/2}/n\right) + n^{2\tau} \exp\left(-\frac{c_1}{2} d^{2\varepsilon}\right)\right).$$

The proof of (4.8) is now complete.

The first relation of (4.9) is trivial. To prove the second part we introduce $\eta_{n,d}(1,2)$, the d^{th} largest among $|X_3|, |X_4|, \ldots, |X_n|$. Set

$$\xi_{1,2} = X_1(I\{|X_1| \leq \eta_{n,d}(1,2)\} - I\{|X_1| \leq H^{-1}(d/n)\}) - m(\eta_{n,d}(1,2))$$

and

$$\xi_{2,1} = X_2(I\{|X_2| \leq \eta_{n,d}(1,2)\} - I\{|X_2| \leq H^{-1}(d/n)\}) - m(\eta_{n,d}(1,2)).$$

Using the independence of $|X_1|$, $|X_2|$ and $\eta_{n,d}(1,2)$ we get

$$E\xi_{1,2}\xi_{2,1} = 0. \tag{4.14}$$

Next we observe

$$\begin{aligned} &\xi_1\xi_2 = \\ &= X_1(I\{|X_1| \leq \eta_{n,d}(1)\} - I\{|X_1| \leq \eta_{n,d}(1,2)\}\xi_2) - (m(\eta_{n,d}(1)) - m(\eta_{n,d}(1,2)))\xi_2 \\ &+ X_2(I\{|X_2| \leq \eta_{n,d}(2)\} - I\{|X_2| \leq \eta_{n,d}(1,2)\})\xi_{1,2} - (m(\eta_{n,d}(2)) - m(\eta_{n,d}(1,2)))\xi_{1,2} \\ &+ \xi_{1,2}\xi_{2,1}. \end{aligned}$$

So by (4.14) we have

$$\begin{aligned} E\xi_1\xi_2 &= E(X_1 I\{\eta_{n,d}(1,2) < |X_1| \leq \eta_{n,d}(1)\}\xi_2) + E((m(\eta_{n,d}(1,2)) - m(\eta_{n,d}(1)))\xi_2) \\ &\quad + E(X_2 I\{\eta_{n,d}(1,2) < |X_2| \leq \eta_{n,d}(2)\}\xi_{1,2}) + E((m(\eta_{n,d}(1,2)) - m(\eta_{n,d}(2)))\xi_{1,2}) \\ &\quad = a_{n,1} + \ldots + a_{n,4}. \end{aligned}$$

It is easy to see that

$$\eta_{n,d+2} \leq \eta_{n,d}(1,2) \leq \eta_{n,d}(1) \leq \eta_{n,d} \quad \text{and} \quad \eta_{n,d+2} \leq \eta_{n,d}(1,2) \leq \eta_{n,d}(2) \leq \eta_{n,d}.$$

Hence

$$\frac{H(\eta_{n,d}(1))}{H(\eta_{n,d}(1,2))} \geq \frac{H(\eta_{n,d})}{H(\eta_{n,d+2})} \stackrel{\mathcal{D}}{=} \frac{S_d}{S_{d+2}} = 1 - \frac{e_{d+1} + e_{d+2}}{S_{d+2}}$$

according to Lemma 4.2. Using (4.10) we get for any $0 < \varepsilon < 1/2$

$$P\{|S_{d+2} - (d+2)| \geq d^{2\varepsilon}\sqrt{d+2}\} \leq \exp(-c_1 d^{2\varepsilon}).$$

The random variables e_{d+1} and e_{d+2} are exponentially distributed with parameter 1 and therefore

$$P\{e_{d+1} \geq d^{2\varepsilon}\} = P\{e_{d+2} \geq d^{2\varepsilon}\} \leq \exp(-d^{2\varepsilon}).$$

Thus we obtain for any $0 < \varepsilon < 1/2$

$$P\left\{\frac{H(\eta_{n,d}(1))}{H(\eta_{n,d}(1,2))} \geq 1 - \frac{c_7 d^{2\varepsilon}}{d}\right\} \geq 1 - c_8 \exp(-c_9 d^{2\varepsilon})$$

and similar arguments yield

$$P\left\{\frac{H(\eta_{n,d}(2))}{H(\eta_{n,d}(1,2))} \geq 1 - \frac{c_7 d^{2\varepsilon}}{d}\right\} \geq 1 - c_8 \exp(-c_9 d^{2\varepsilon})$$

and

$$P\left\{\frac{d}{n}\left(1 - \frac{1}{d^{1/2-\varepsilon}}\right) \leq H(\eta_{n,d}) \leq \frac{d}{n}\left(1 + \frac{1}{d^{1/2-\varepsilon}}\right)\right\} \geq 1 - c_8 \exp(-c_9 d^{2\varepsilon})$$

with some constants c_7, c_8 and c_9. Now we define the event A as the set on which

$$\frac{H(\eta_{n,d}(1))}{H(\eta_{n,d}(1,2))} \geq 1 - \frac{c_7}{d^{1-2\varepsilon}}, \quad \frac{H(\eta_{n,d}(2))}{H(\eta_{n,d}(1,2))} \geq 1 - \frac{c_7}{d^{1-2\varepsilon}}$$

and

$$\frac{d}{n}\left(1 - \frac{1}{d^{1/2-\varepsilon}}\right) \leq H(\eta_{n,d}) \leq \frac{d}{n}\left(1 + \frac{1}{d^{1/2-\varepsilon}}\right)$$

hold. Clearly,

$$P(A^c) \leq 3c_8 \exp(-c_9 d^{2\varepsilon}).$$

Using the definition of ξ_2 we get that

$$\begin{aligned}
a_{n,1} &\leq E\big(|X_1| I\{\eta_{n,d}(1,2) \leq |X_1| \leq \eta_{n,d}(1)\} \\
&\qquad \times |X_2|\, |I\{|X_2| \leq \eta_{n,d}(2)\} - I\{|X_2| \leq H^{-1}(n/d)\}|\big) \\
&\quad + E|X_1| I\{\eta_{n,d}(1,2) \leq |X_1| \leq \eta_{n,d}(1)\}|m(\eta_{n,d}(2))| \\
&\leq E|X_1||X_2| I\{\eta_{n,d}(1,2) \leq |X_1| \leq \eta_{n,d}(1)\} I\{H^{-1}(d/n) \leq |X_2| \leq \eta_{n,d}(2)\} \\
&\quad + E|X_1||X_2| I\{\eta_{n,d}(1,2) \leq |X_1| \leq \eta_{n,d}(1)\} I\{\eta_{n,d}(2) \leq |X_2| \leq H^{-1}(d/n)\} \\
&\quad + E|X_1| I\{\eta_{n,d}(1,2) \leq |X_1| \leq \eta_{n,d}(1)\}|m(\eta_{n,d}(2))| \\
&= a_{n,1,1} + a_{n,1,2} + a_{n,1,3}.
\end{aligned}$$

Using the definition of A we obtain that

$$a_{n,1,1} \leq E|X_1 X_2| I\{\eta_{n,d}(1,2) \leq |X_1| \leq \eta_{n,d}(1)\}$$

$$
\begin{aligned}
&\quad \times I\{H^{-1}(d/n) \le |X_2| \le \eta_{n,d}(2)\} I\{A\} \\
&+ E|X_1X_2| I\{\eta_{n,d}(1,2) \le |X_1| \le \eta_{n,d}(1)\} \\
&\quad \times I\{H^{-1}(d/n) \le |X_2| \le \eta_{n,d}(2)\} I\{A^c\} \\
\le{}& E\Bigg(|X_1X_2| I\left\{ H(\eta_{n,d}(1,2)) \left(1 - \frac{c_7}{d^{1-2\varepsilon}}\right) \le H(|X_1|) \le H(\eta_{n,d}(1,2)) \right\} \\
&\times I\{A\} I\{H^{-1}(d/n) \le |X_2| \le \eta_{n,d}(2)\} \Bigg) + E(\eta_{n,d}^2 I\{A^c\}) \\
\le{}& \left(H^{-1} \left(\frac{d}{n} \left(1 - \frac{c_{10}}{d^{1/2-\varepsilon}} \right) \right) \right)^2 \\
&\quad \times E\Bigg(I\left\{ H(\eta_{n,d}(1,2)) \left(1 - \frac{c_7}{d^{1-2\varepsilon}}\right) \le H(|X_1|) \le H(\eta_{n,d}(1,2)) \right\} \\
&\quad \times I\left\{ \frac{d}{n} \left(1 - \frac{1}{d^{1/2-\varepsilon}} \right) \le H(|X_2|) \le \frac{d}{n} \right\} \Bigg) + E(\eta_{n,d}^2 I\{A^c\}).
\end{aligned}
$$

Using again the independence of $|X_1|$, $|X_2|$ and $\eta_{n,d}(1,2)$ we conclude that

$$
\begin{aligned}
&E\left(I\left\{ H(\eta_{n,d}(1,2)) \left(1 - \frac{c_7}{d^{1-2\varepsilon}}\right) \le H(|X_1|) \le H(\eta_{n,d}(1,2)) \right\} \right. \\
&\qquad \left. \times I\left\{ \frac{d}{n} \left(1 - \frac{1}{d^{1/2-\varepsilon}} \right) \le H(|X_2|) \le \frac{d}{n} \right\} \right) \\
&\quad = EH(\eta_{n,d}(1,2)) \frac{c_7}{d^{1-2\varepsilon}} \frac{d}{n} \frac{1}{d^{1/2-\varepsilon}} \le \frac{d}{n-1} \frac{c_7}{n} \frac{1}{d^{1/2-3\varepsilon}}.
\end{aligned}
$$

The Cauchy-Schwarz inequality yields

$$
E(\eta_{n,d}^2 I\{A^c\}) \le \left(E\eta_{n,d}^4 \right)^{1/2} (P(A^c))^{1/2} = \mathcal{O}\left(n^{2\tau} \exp\left(-\frac{c_9}{2} d^{2\varepsilon} \right) \right)
$$

for all $\tau > 1/\alpha$ on account of (4.13). Thus we conclude

$$
a_{n,1,1} = \mathcal{O}\left((H^{-1}(d/n))^2 \left(d^{1/2+3\varepsilon}/n^2 \right) + n^{2\tau} \exp\left(-\frac{c_9}{2} d^{2\varepsilon} \right) \right).
$$

Similar but somewhat simpler arguments imply

$$
a_{n,1,2} + a_{n,1,3} = \mathcal{O}\left((H^{-1}(d/n))^2 \left(d^{1/2+3\varepsilon}/n^2 \right) + n^{2\tau} \exp\left(-\frac{c_9}{2} d^{2\varepsilon} \right) \right),
$$

resulting in

$$
a_{n,1} = \mathcal{O}\left((H^{-1}(d/n))^2 \left(d^{1/2+3\varepsilon}/n^2 \right) + n^{2\tau} \exp\left(-\frac{c_9}{2} d^{2\varepsilon} \right) \right). \tag{4.15}
$$

Following the proof of (4.15), the same rates can be obtained for $a_{n,2}$ and $a_{n,3}$. □

Lemma 4.6. *If the conditions of Theorem 1.1 are satisfied, then*

$$\frac{1}{A_n}\max_{1\le k\le n}\left|\sum_{j=1}^{k}\xi_j\right| = o_P(1).$$

Proof. It is easy to see that for any $1 \le \ell_1 \le \ell_2 \le n$ we have

$$\begin{aligned} E\left(\sum_{j=\ell_1}^{\ell_2}\xi_j\right)^2 &= (\ell_2-\ell_1+1)E\xi_1^2 + (\ell_2-\ell_1)(\ell_2-\ell_1+1)E\xi_1\xi_2 \\ &\le (\ell_2-\ell_1+1)(E\xi_1^2 + nE\xi_1\xi_2). \end{aligned}$$

Lemma 4.5 and (1.12) yield

$$E\xi_1^2 \le c_1\frac{A_n^2}{n}\left[d^{-1/2+\varepsilon} + n^{2\tau+1}\exp(-ad^{2\varepsilon})\right]$$

and

$$E\xi_1\xi_2 \le c_2\frac{A_n^2}{n^2}\left[d^{-1/2+3\varepsilon} + n^{2\tau+2}\exp(-ad^{2\varepsilon})\right]$$

for all $0 < \varepsilon < 1/6$. Hence we conclude

$$E\left(\sum_{j=\ell_1}^{\ell_2}\xi_j\right)^2 \le c_3(\ell_2-\ell_1+1)\frac{A_n^2}{n}\left[d^{-1/2+3\varepsilon} + n^{2\tau+2}\exp(-ad^{2\varepsilon})\right].$$

So using an inequality of Menshov (cf. Billingsley [5], p. 102) we get that

$$\begin{aligned} &E\left(\max_{1\le k\le n}\left|\sum_{j=1}^{k}\xi_j\right|\right)^2 \le c_4(\log n)^2A_n^2\left[d^{-1/2+3\varepsilon} + n^{2\tau+2}\exp(-ad^{2\varepsilon})\right] \\ &\le c_4A_n^2\left[(\log n)^2d^{-2/7} + \exp((2\tau+2)\log n + 2\log\log n - ad^{2\varepsilon})\right] \\ &= A_n^2o(1) \quad \text{as} \quad n\to\infty, \end{aligned}$$

where $\varepsilon = 1/14$ and $d = (\log n)^\gamma$ with any $\gamma > 7$, resulting in

$$\frac{1}{A_n^2}E\left(\max_{1\le k\le n}\left|\sum_{j=1}^{k}\xi_j\right|\right)^2 = o(1).$$

Now Markov's inequality completes the proof of Lemma 4.6. □

Proof of Theorem 1.5. It follows immediately from Lemmas 4.3, 4.4 and 4.6. □

Proof of Theorem 1.1. According to (1.15), Theorems 1.4 and 1.5 imply Theorem 1.1. □

Lemma 4.7. *If the conditions of Theorem 1.1 are satisfied, then*

$$\frac{\hat{A}_n}{A_n} \longrightarrow 1 \quad a.s.$$

Proof. It is an immediate consequence of Haeusler and Mason [16]. □

Proof of Theorem 1.2. By Slutsky's lemma, Lemma 4.7 and Theorem 1.1 imply the result. □

Proof of Example 1.1. Since $H'(x) = -f(x)$, our assumptions imply that $H'(x)$ is also regularly varying at ∞. By elementary results on regular variation (see e.g. Bingham et al. [6]), it follows that

$$H(x) = 1 - F(x) = \int_x^\infty f(t)dt \sim \frac{1}{\alpha} x f(x) \quad \text{as } x \to \infty.$$

Hence H^{-1} is regularly varying at 0, therefore the function $(H^{-1}(t))' = 1/H'(H^{-1}(t))$ is also regularly varying at 0. Also,

$$m'(x) = \frac{d}{dx}\int_0^x tf(t)dt = xf(x) \sim \alpha H(x) \quad \text{as } x \to \infty$$

and therefore $m'(H^{-1}(t)) \sim t\alpha$. Using Lemma 4.2, the mean value theorem gives

$$\begin{aligned}\frac{nm(\eta_{n,d})}{B_n} \stackrel{\mathcal{D}}{=} \frac{nm(H^{-1}(S_d/S_{n+1}))}{B_n} &= \frac{n(\ell(S_d/S_{n+1}) - \ell(d/n))}{B_n} \\ &= \frac{n}{B_n}\ell'(\xi_n)\left(\frac{S_d}{S_{n+1}} - \frac{d}{n}\right),\end{aligned}$$

where ξ_n is between S_d/S_{n+1} and d/n and $\ell(t) = m(H^{-1}(t))$. It follows from the central limit theorem for central order statistics that

$$\frac{n}{d^{1/2}}\left(\frac{S_d}{S_{n+1}} - \frac{d}{n}\right) \xrightarrow{\mathcal{D}} N(0,1). \tag{4.16}$$

The regular variation of ℓ' and (4.16) yield

$$\ell'(\xi_n)/\ell'(d/n) \to 1 \quad \text{in probability.}$$

The result now follows from (4.16) by observing that

$$\frac{n}{B_n}\ell'(d/n) \sim \frac{n}{d^{1/2}}.$$

□

The proof of Theorem 1.3 is based on analogues of Theorems 1.4, 1.5 and Lemmas 4.3–4.7 when $EX_j^2 < \infty$.

Lemma 4.8. *If the conditions of Theorem 1.3 are satisfied, then*

$$\frac{1}{\sqrt{n}}\sum_{j=1}^{\lfloor nt\rfloor}(X_jI\{|X_j|\le H^{-1}(d/n)\}-E[X_1I\{|X_1|\le H^{-1}(d/n)\}])\xrightarrow{\mathcal{D}[0,1]}\sigma W(t),$$

where $\sigma^2=\operatorname{var}X_1$.

Proof. By $EX_1^2<\infty$ we have

$$E\bigg[X_1I\{|X_1|\le H^{-1}(d/n)\}-E[X_1I\{|X_1|\le H^{-1}(d/n)\}]-(X_1-EX_1)\bigg]^2\longrightarrow 0$$

as $n\to\infty$. So using Lévy's inequality [24, p. 248] we get

$$\frac{1}{\sqrt{n}}\max_{1\le k\le n}\bigg|\sum_{j=1}^{k}(X_jI\{|X_j|\le H^{-1}(d/n)\}$$
$$-E[X_1I\{|X_1|\le H^{-1}(d/n)\}]-(X_j-EX_1))\bigg|=o_P(1).$$

Now Donsker's theorem (cf. Billingsley [5, p. 137]) implies the result. □

Lemma 4.9. *If the conditions of Theorem 1.3 are satisfied, then*

$$\frac{1}{\sqrt{n}}\sum_{j=1}^{n}|X_j(I\{|X_j|\le\eta_{n,d}\}-I\{|X_j|\le\eta_{n,d}(j)\})|=o_P(1)$$

and

$$\frac{1}{\sqrt{n}}\sum_{j=1}^{n}|m(\eta_{n,d})-m(\eta_{n,d}(j))|=o_P(1).$$

Proof. We adapt the proof of Lemma 4.3. We recall that A is an event satisfying (4.5), (4.6) and $P(A)\ge 1-\varepsilon$, where $\varepsilon>0$ is an arbitrary small positive number. We also showed that

$$E(|X_j(I\{|X_j|\le\eta_{n,d}\}-I\{|X_j|\le\eta_{n,d}(j)\})|I\{A\})$$
$$\le H^{-1}\left(\frac{d}{n}\left(1-\frac{C}{\sqrt{d}}\right)\right)\frac{d+1}{n+1}\frac{C}{d}$$

with some constant C. Assumption $EX_1^2<\infty$ yields

$$\limsup_{x\to 0}x^{1/2}H^{-1}(x)<\infty,$$

and therefore

$$\lim_{n\to\infty}\sqrt{n}H^{-1}\left(\frac{d}{n}\left(1-\frac{C}{\sqrt{d}}\right)\right)\frac{d+1}{n+1}\frac{C}{d}=0$$

for all $C > 0$. Thus we have for all $\varepsilon > 0$

$$\lim_{n\to\infty} \frac{1}{\sqrt{n}} \sum_{j=1}^{n} E|X_j(I\{|X_j| \le \eta_{n,d}\} - I\{|X_j| \le \eta_{n,d}(j)\})|I\{A\} = 0.$$

Since we can choose $\varepsilon > 0$ as small as we wish, the first result is proven. The second part of the lemma can be established similarly. □

Lemma 4.10. *If the conditions of Theorem 1.3 are satisfied, then for all* $0 < \varepsilon < 1/2$

$$\begin{aligned} E\xi_j &= 0, \quad 1 \le j \le n, \\ E\xi_j^2 = E\xi_1^2 &= \mathcal{O}\big((H^{-1}(d/n))^2 d^{1/2+\varepsilon}/n + n\exp(-ad^{2\varepsilon})\big), 1 \le j \le n, \\ E\xi_i\xi_j = E\xi_1\xi_2 &= \mathcal{O}\big((H^{-1}(d/n))^2 d^{1/2+3\varepsilon}/n^2 + n\exp(-ad^{2\varepsilon})\big), 1 \le i \ne j \le n. \end{aligned}$$

Proof. The proof of Lemma 4.5 can be repeated; only (4.12) should be replaced with

$$H^{-1}(t) \le Ct^{-1/2}, \quad 0 < t \le 1. \tag{4.17}$$

□

Lemma 4.11. *If the conditions of Theorem 1.3 are satisfied, then*

$$\frac{1}{\sqrt{n}} \max_{1\le k\le n} \left| \sum_{j=1}^{k} \xi_j \right| = o_P(1).$$

Proof. Following the proof of Lemma 4.6 we get

$$E\left(\max_{1\le k\le n} \left| \sum_{j=1}^{k} \xi_j \right| \right)^2 \le c_1 n(\log n)^2 \left[d^{-1/2+3\varepsilon} + n^3 \exp(-ad^{2\varepsilon}) \right] = no(1) \tag{4.18}$$

as $n \to \infty$. Markov's inequality completes the proof of Lemma 4.11. □

Lemma 4.12. *If the conditions of Theorem 1.3 are satisfied, then*

$$\frac{1}{\sqrt{n}} \max_{1\le k\le n} \left| \sum_{j=1}^{k} [X_j(I\{|X_j| \le \eta_{n,d}\} - I\{|X_j| \le H^{-1}(d/n)\}) - m(\eta_{n,d})] \right| = o_P(1).$$

Proof. It follows immediately from Lemmas 4.9 and 4.11. □

Proof of Theorem 1.3. By Lemmas 4.8 and 4.12 we have that

$$\frac{T_n(t)}{\sigma\sqrt{n}} \xrightarrow{\mathcal{D}[0,1]} B(t).$$

It is easy to see that

$$\frac{\hat{A}_n^2}{n} \xrightarrow{P} \sigma^2,$$

which completes the proof of Theorem 1.3. □

Proof of Theorem 2.1. We show that

$$\frac{\max_{1\le j\le n}|x_j|}{\sqrt{\sum_{j=1}^n x_j^2}} \longrightarrow 0 \quad \text{a.s.} \tag{4.19}$$

By Lemma 4.7 it is enough to prove that

$$\frac{\max_{1\le j\le n}|x_j|}{A_n} \longrightarrow 0 \quad \text{a.s.}$$

It follows from the definition of x_j that

$$\max_{1\le j\le n}|x_j| \le \eta_{d,n} + |\bar{X}_{n,d}| \le 2\eta_{d,n}.$$

Using Kiefer [23] (cf. Shorack and Wellner [33]) we get

$$\frac{\eta_{d,n}}{A_n} \longrightarrow 0 \quad \text{a.s.}$$

Since (4.19) holds for almost all realizations of $X_1, X_2, \ldots$, Rosén [32] implies Theorem 2.1 when we sample without replacement and Prohorov [31] when we sample with replacement (bootstrap). □

Proof of Theorem 2.2. It can be established along the lines of the proof of Theorem 2.1. □

References

[1] K. Athreya. Bootstrap of the mean in the infinite variance case. *Ann. Statist.*, 15:724–731, 1987.

[2] A. Aue, I. Berkes, and L. Horváth. Selection from a stable box. *Bernoulli*, 14:125–139, 2008.

[3] M. S. Bartlett. Periodogram analysis and continuous spectra. *Biometrika*, 37:1–16, 1950.

[4] I. Berkes, L. Horváth, and J. Schauer. Non-central limit theorems for random selections. *Probab. Theory Related Fields*, 147:449–479, 2010.

[5] P. Billingsley. *Convergence of Probability Measures.* Wiley, New York, 1968.

[6] N. H. Bingham, C. M. Goldie, and J. L. Teugels. *Regular Variation*, volume 27 of *Encyclopedia of Mathematics and its Applications.* Cambridge University Press, Cambridge, 1989.

[7] B. E. Brodsky and B. S. Darkhovsky. *Non-parametric Statistical Diagnosis*, volume 509 of *Mathematics and its Applications*. Kluwer Academic Publishers, Dordrecht, 2000.

[8] J. Chen and A. K. Gupta. *Parametric Statistical Change Point Analysis*. Birkhäuser Boston Inc., Boston, MA, 2000.

[9] S. Csörgő, E. Haeusler and D.M. Mason. A probabilistic approach to the asymptotic distribution of sums of independent, identically distributed random variables. *Adv. Appl. Math.* **9**, 259–333, 1988.

[10] S. Csörgő, E. Haeusler and D.M. Mason. The asymptotic distribution of trimmed sums. *Ann. Probab.* **16**, 672–699, 1988.

[11] M. Csörgő and L. Horváth. *Limit Theorems in Change-Point Analysis*. Wiley, Chichester, 1997.

[12] S. Csörgő, L. Horváth, and D. M. Mason. What portion of the sample makes a partial sum asymptotically stable or normal? *Probab. Theory Related Fields*, 72:1–16, 1986.

[13] U. Grenander and M Rosenblatt. *Statistical Analysis of Stationary Time Series*. John Wiley & Sons, New York, 1957.

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

[15] P. S. Griffin and W. E. Pruitt. Asymptotic normality and subsequential limits of trimmed sums. *Ann. Probab.* **17**, 1186–1219, 1989.

[16] E. Haeusler and D. M. Mason. Laws of the iterated logarithm for sums of the middle portion of the sample. *Math. Proc. Cambridge Philos. Soc.* **101**, 301–312, 1987.

[17] E. Haeusler and D. M. Mason. A law of the iterated logarithm for modulus trimming. In: *Limit theorems in probability and statistics (Pécs, 1989)*, pp. 241–254. Coll. Math. Soc. János Bolyai 57, North-Holland, Amsterdam, 1990.

[18] P. Hall. Asymptotic properties of the bootstrap for heavy-tailed distributions. *Ann. Probab.*, 18:1342–1360, 1990.

[19] L. Horváth, P. Kokoszka, and J. Steinebach. Testing for changes in multivariate dependent observations with an application to temperature changes. *J. Multivariate Anal.*, 68:96–119, 1999.

[20] L. Horváth and J. Steinebach. Testing for changes in the mean or variance of a stochastic process under weak invariance. *J. Statist. Plann. Inference*, 91:365–376, 2000. Prague Workshop on Perspectives in Modern Statistical Inference: Parametrics, Semi-parametrics, Non-parametrics (1998).

[21] M. Hušková. Permutation principle and bootstrap in change point analysis. In *Asymptotic Methods in Stochastics*, volume 44 of *Fields Inst. Commun.*, pages 273–291. Amer. Math. Soc., Providence, RI, 2004.

[22] Y. Kasahara and S. Watanabe. Limit theorems for point processes and their functionals. *J. Math. Soc. Japan*, 38:543–574, 1986.

[23] J. Kiefer. Iterated logarithm analogues for sample quantiles when $p_n \downarrow 0$. In *Proceedings of the Sixth Berkeley Symposium on Mathematical Statistics and Probability (Univ. California, Berkeley, Calif., 1970/1971), Vol. I: Theory of statistics*, pages 227–244, Berkeley, Calif., 1972. Univ. California Press.

[24] M. Loève. *Probability Theory.* 2nd ed. The University Series in Higher Mathematics. D. Van Nostrand Co., Inc., Princeton, N. J.-Toronto-New York-London, 1960.

[25] R. A. Maller. Asymptotic normality of lightly trimmed means - a converse. *Math. Proc. Cambridge Philos. Soc.*, 92:535–545, 1982.

[26] R. A. Maller. Relative stability of trimmed sums. *Z. Wahrsch. Verw. Gebiete*, 66:61–80, 1984.

[27] T. Mori. The strong law of large numbers when extreme terms are excluded from sums. *Z. Wahrsch. Verw. Gebiete*, 36:189–194, 1976.

[28] T. Mori. Stability for sums of i.i.d. random variables when extreme terms are excluded. *Z. Wahrsch. Verw. Gebiete*, 40:159–167, 1977.

[29] T. Mori. On the limit distributions of lightly trimmed sums. *Math. Proc. Cambridge Philos. Soc.*, 96:507–516, 1984.

[30] E. Parzen. On consistent estimates of the spectrum of a stationary time series. *Ann. Math. Statist.*, 28:329–348, 1957.

[31] Y. V. Prohorov. Convergence of random processes and limit theorems in probability theory. *Teor. Veroyatnost. i Primenen.*, 1:177–238, 1956.

[32] B. Rosén. Limit theorems for sampling from finite populations. *Ark. Mat.*, 5:383–424, 1965.

[33] G. Shorack and J. Wellner. *Empirical Processes with Applications to Statistics.* Wiley, New York, 1986.

[34] A. V. Skorohod. Limit theorems for stochastic processes. *Theory Probab. Appl.*, 1:261–290, 1956.