LAST PASSAGE TIMES OF MINIMUM CONTRAST ESTIMATORS

ARUP BOSE and SNIGDHANSU CHATTERJEE

Communicated by V. Stefanov

Abstract

We study the last passage time and its asymptotic distribution for minimum contrast estimators defined through the minimization of a convex criterion function based on U-functionals. This includes cases of non-smooth estimators for vector valued parameters. We also derive a Bahadur-type representation and the law of iterated logarithms for such estimators.

2000 *Mathematics subject classification*: primary 62F12; secondary 60F99, 60H99, 62E20.
Keywords and phrases: minimum contrast estimators, last passage time, law of iterated logarithms, Bahadur representation, Oja median, L_1 median.

1. Introduction

Let X be an M valued random variable with distribution F. Let $X_1, \ldots, X_n$, be independent, identically distributed copies of X. Let $q(\theta, Z)$ be a real valued function defined for $\theta \in \mathbb{R}^d$ and $Z \in M^m$ for some m, $1 \leq m < \infty$. We assume that q is measurable and is symmetric in its last m arguments for each θ. Let $Q(\theta) = E_F q(\theta, X_1, \ldots, X_m)$ and θ_0 (unique) be such that $Q(\theta_0) = \inf_\theta E_F q(\theta, X_1, \ldots, X_m)$. Define the sample analogue of $Q(\theta)$ as

$$Q_n(\theta) = \binom{n}{m}^{-1} \sum_{1 \leq i_1 < \cdots < i_m \leq n} q(\theta, X_{i_1}, \ldots, X_{i_m}).$$

The estimator θ_n of θ_0 is the value which satisfies $Q_n(\theta_n) = \inf_\theta Q_n(\theta)$ and is called a minimum contrast estimator. This is the M_m estimate of θ_0 introduced by Huber [9]. For the present paper these estimators are based on minimisation of convex functions of U-functionals. Examples of statistics that come under our set-up include (i) Oja

median (Oja [13]), (ii) univariate location estimators of Maritz *et al.* [11], (iii) univariate Hodges-Lehmann estimators of location, (iv) a univariate robust scale estimator of Bickel and Lehmann [1], (v) a regression coefficient estimator of Theil (see Hollander and Wolfe [8]), (vi) U quantiles (Chowdhury and Serfling [6]), (vii) L_1 median, (viii) geometric quantiles of Chaudhuri [5] and, (ix) Hodges-Lehmann versions of (vii) and (viii).

For $m = 1$, strong consistency and asymptotic normality of θ_n was established by Habermann [7] using the assumption of *convexity* of q in θ. Niemiro [12] utilized this convexity to establish other asymptotic properties of θ_n. Bose [4] extended these results to $m > 1$. In particular, they established a Bahadur type representation of the form

$$\sqrt{n}(\theta_n - \theta_0) = \text{leading term} + R_n,$$

where R_n is of suitable order almost surely under suitable assumptions. If $m = 1$ and q is differentiable then θ_n satisfies $\int q'(\theta, x) F_n(dx) = 0$. For these estimators, which are known as Huber's M-estimators, Stute [16] obtained a similar representation under the restrictions that the score function q' is smooth, bounded and nonincreasing in θ for each $x \in \mathbb{R}$. Here F_n denotes the empirical distribution function and θ_0 is a root of $I(\theta) \equiv \int q'(\theta, x) F_{\theta_0}(dx) = 0$. This representation was used to study asymptotic properties of the *last passage time* of θ_n in Stute [15] for $m = 1$. For a sequence of estimators θ_n, the last passage time is defined as

$$\tau_\epsilon = \sup\{n \geq 1 : |\theta_n - \theta_0| > \epsilon\}.$$

One immediate use of last passage time asymptotics is in computing fixed volume confidence sets for the parameter of interest with prescribed asymptotic coverage probability. Last passage times are also important from the viewpoint of computing asymptotic relative efficiencies of members of a class of estimators and thereby selecting the optimal estimators. They also relate to Pitman efficiency of estimators, see Stute [15] for precise details. They may also be used for computing most other statistics that are relevant in the context of sequential procedures. A truly sequential procedure is often complicated and intractable for non-smooth statistics, so an approximation through last passage time asymptotics can be quite useful.

When θ_n is the sample mean of independent, identically distributed random variables, asymptotic properties of last passage time estimators have been considered by Robbins and Siegmund [14]. Stute [16] extended these results to M-estimates for $m = 1$. His main assumptions are that the score function is a *non increasing, differentiable function of θ, is bounded uniformly in x, $I(\theta)$ is twice continuously differentiable in a neighbourhood of θ_0 such that $I'(\theta_0) < 0$.*

In this paper we discuss the last passage time asymptotics for the M_m estimators. Thus we extend Stute's results in two directions. First, we consider functionals which

are not necessarily smooth. Second, we allow $m > 1$. Thus we are able to include a significantly wider class of estimators, some examples of which were mentioned earlier.

Our main theorem on last passage times uses a Bahadur representation of M_m estimators. Such a representation is already available in Bose [4]. However, we need a more general statement for the present purpose, and that is given in Theorem 2.1. Following this general Bahadur representation theorem, the law of iterated logarithms is an immediate corollary, so we make a mention of it. The main result on last passage times is given in Theorem 2.2. All proofs are given in Section 3.

2. Results

Let X be an M valued random variable with distribution F. Let $X_1, \ldots, X_n$, be *i.i.d.* F. Let $q(\theta, Z)$ be a real valued function defined for $\theta \in \mathbb{R}^d$ and $Z \in M^m$ for some m, $1 \leq m < \infty$. We assume that q is measurable and is symmetric in its last m arguments for each θ. Let $Q(\theta) = E_F q(\theta, X_1, \ldots, X_m)$ and θ_0 (unique) be such that

$$Q(\theta_0) = \inf_\theta E_F q(\theta, X_1, \ldots, X_m). \tag{2.1}$$

Define the sample analogue of $q(\theta)$ as

$$Q_n(\theta) = \binom{n}{m}^{-1} \sum_{1 \leq i_1 < \cdots < i_m \leq n} q(\theta, X_{i_1}, \ldots, X_{i_m}) \tag{2.2}$$

and the estimator θ_n of θ as the value which satisfies $Q_n(\theta_n) = \inf_\theta Q_n(\theta)$. If no such θ_n exists, take $\theta_n = \infty$. The estimator θ_n can be and is chosen in a measurable way.

Let g be a *subgradient* of q. That is, for all α, β, Z,

$$q(\alpha, Z) + (\beta - \alpha)' g(\alpha, Z) \leq q(\beta, Z). \tag{2.3}$$

Further g is such that it is measurable in Z for each α. We will use the notation $G(\theta) = Eg(\theta, X_1, \ldots X_m)$. The gradient vector and the matrix of second derivatives of Q at θ will be denoted by $\nabla Q(\theta)$ and $\nabla^2 Q(\theta)$ respectively. Let

$$S_n = \sum_{1 \leq i_1 < \cdots < i_m \leq n} g(\theta_0, X_{i_1}, \ldots, X_{i_m}), \quad \text{and} \quad U_n = \binom{n}{m}^{-1} S_n.$$

Define $g_1(\theta_0, X_1) = E[g(\theta_0, X_1, \ldots, X_m)|X_1]$. For a vector a, the notation $|a|$ will denote its Euclidean norm. For a matrix A, the notation A^T will denote its transpose. We now state the conditions that are assumed throughout this paper. Let N be an appropriate neighbourhood of θ_0, $r > 1$ and $\theta \to \theta_0$, $0 \leq s < 1$.

(I) $q(\theta, Z)$ is convex in θ for every Z.
(II) $Q(\theta)$ is finite for all θ.
(III) θ_0 satisfying (2.1) exists and is unique.
(IV) $\mathrm{Var}(g(\theta_0, X_1, \ldots, X_m))$ is positive definite.
(V) $E|g(\theta, X_1, \ldots, X_m)|^r < \infty \; \forall \theta \in N$.
(VI) $H = \nabla^2 Q(\theta_0)$ exists and is positive definite.
(VII) $|\nabla Q(\theta) - \nabla^2 Q(\theta_0)(\theta - \theta_0)| = O(|\theta - \theta_0|^{(3-s)/2})$ as $\theta \to \theta_0$.
(VIII) $E|g(\theta, X_1, \ldots, X_m) - g(\theta_0, X_1, \ldots, X_m)|^2 = O(|\theta - \theta_0|^{(1+s)})$ as $\theta \to \theta_0$.
(IX) $E|g(\theta, X_1, \ldots, X_m)|^r = O(1)$ as $\theta \to \theta_0$.
(X) $V = Eg_1(\theta_0, X_1)g_1^T(\theta_0, X_1)$ is positive definite.

Often, the finiteness condition (II) is satisfied for only a subset of $\mathbb{R}^d$. All our results remain valid if θ_0 is an interior point of this subset. Assumptions (I)–(IX) are same as that of Niemiro (1992) and Bose (1998). Assumption (X) in particular implies that the U-statistic U_n is non-degenerate. The matrix $V = Eg_1(\theta_0, X_1)g_1^T(\theta_0, X_1)$ being positive definite, it can be written as $V = \Sigma\Sigma^T$, where Σ can be chosen to be symmetric, and further, positive definite. We will retain the notation Σ for such a choice. As remarked in the introduction, our main result Theorem 2.2 requires a Bahadur representation of minimum contrast estimators. The general representation result involve four sequences of non-negative reals that satisfy certain interrelationships. We consider the following: δ_n, η_n, v_n and ϵ_n are four nonnegative sequences satisfying

(a) $\eta_n = o(n^{1/2})$, $\{\eta_n\}$ bounded away from 0,
(b) $\eta_n/\delta_n \geq 1$ for all $n \geq 1$,
(c) $n^\delta \geq v_n \geq Cn^{-(1+s)/4}(\eta_n)^{(3+s)/2}$ for a constant $C \geq 1$ and some $\delta \geq 1/2$,
(d) $\epsilon_n = 2K_0 v_n(\log n)^{1/2} + 4K_1\delta_n$, where K_0, K_1 are constants.

THEOREM 2.1. *Assume the conditions* (I)–(IX) *with* $r > 2$ *and some* $s \in [0, 1)$. *Then*

$$n^{1/2}(\theta_n - \theta_0) = -n^{1/2}H^{-1}U_n + R_n$$

with

$$P(|R_n| > K\epsilon_n) = O\left(\left(\frac{\eta_n}{\delta_n}\right)^d n^{1-r/2}v_n^{-r}(\log n)^{r/2}\right),$$

where $1/K = \inf_{|e|=1} e'He/2$.

The corresponding representations available in Bose [4] or Niemiro [12] can be worked from Theorem 2.1. A law of iterated logarithm follows immediately from Theorem 2.1. We record this result in the following corollary.

COROLLARY 2.1. *Suppose assumptions* (I)–(X) *hold and let* $r > 4+d(1+s)/(3+s)$. *Then for any* $c \in \mathbb{R}^d$, $\|c\| = 1$, *we have*

$$P\left[\limsup_n \frac{n^{1/2}c^T\Sigma^{-1}H(\theta_n - \theta_0)}{\sqrt{2m^2(\log\log n)}} = 1\right] = 1.$$

We now state our main result.

THEOREM 2.2. *Suppose assumptions* (I)–(X) *hold with* $r > 2$. *Also let* W_0 *be a standard Brownian motion on* $C_d[0,1]$ *and* $G(y) = P[\sup_{0\le t\le 1}|\Sigma W_0(t)| < y^{1/2}/m]$. *Then*

(i) $\lim_{\epsilon\to 0} P[\epsilon^2\tau_\epsilon < y] = G(y)$ *and*
(ii) $\lim_{\epsilon\to 0} E(\epsilon^2\tau_\epsilon) = \int_0^\infty y\, dG(y)$.

The above theorem requires higher moment conditions than what is needed in establishing similar results for *i.i.d.* real random variables or Huber's M-estimators. This is a trade off in that we have dropped the smoothness conditions and boundedness conditions on the criterion function.

3. Proofs

For all the proofs, assume without loss of generality that $\theta_0 = 0$ and $Q(\theta_0) = 0$. For a real quantity x, $[x]$ will denote its integral part.

In order to prove the Bahadur representation theorem, we need a lemma that we state here. This is essentially Lemma 2 of Bose [3]. We omit the proof of this lemma.

LEMMA 3.1 (Bose [3]). *Let* $\{h_n\}$ *be a sequence of (symmetric) kernels of order* m *and let* $\{X_{ni}, 1 \le i \le n\}$ *be i.i.d. real valued random variables for each* n. *Let* $U_n(h_n) = \binom{n}{m}^{-1}\sum_{1\le i_1<\cdots<i_m\le n} U_n(h_n(X_{ni_1},\ldots,X_{ni_m}))$. *Further, suppose that for some* $\delta > 0$, *and some* $v_n \le n^\delta$, $E\,U_n(h_n(X_n,\ldots,X_{nm})) = 0$, $E|h_n(X_{ni},\ldots,X_{nm})|^2 \le v_n^2$ *and* $E|h_n(X_{ni},\ldots,X_{nm})|^r \le b < \infty$ *for some* $r > 2$. *Then for all large* K,

$$P(n^{1/2}|U_n(h_n)| > Kv_n(\log n)^{1/2}) \le Dn^{1-r/2}v_n^{-r}(\log n)^{r/2}.$$

PROOF OF THEOREM 2.1. Let S denote the set of all m element subsets of $\{1,\ldots,n\}$. For any $s = \{i_1,\ldots,i_m\} \in S$, let Y_s denote the random vector $(X_{i_1},\ldots,X_{i_m})$. Define $G_n(\alpha) = \binom{n}{m}^{-1}\sum_{s\in S} g(\alpha, Y_s)$, $X_{ns} = g(\alpha/n^{1/2}, Y_s) - g(0, Y_s)$. Note that $E(X_{ns}) = G(\alpha/n^{1/2})$ and $\binom{n}{m}^{-1}\sum_{s\in S} X_{ns} = \left[G_n(\alpha/n^{1/2}) - \binom{n}{m}^{-1}S_n\right]$. By (VIII),

$$E|X_{ns}|^2 = O((n^{-1/2}\eta_n)^{1+s}) \quad \text{uniformly for } |\alpha| \le M\eta_n, \tag{3.1}$$

where $n^{-1/2}\eta_n \to 0 \text{as} n \to \infty$ from (a). Now note that (3.1) in conjunction with the upper bound in (c) and assumption (V) ensure that the conditions of Lemma 3.1 are satisfied. Therefore for large K_0

$$(3.2) \quad \sup_{|\alpha|\le M\eta_n} P\left[\left|n^{1/2}G_n\left(\frac{\alpha}{n^{1/2}}\right) - n^{1/2}\binom{n}{m}^{-1} S_n - n^{1/2}G\left(\frac{\alpha}{n^{1/2}}\right)\right| > K_0 v_n (\log n)^{1/2}\right] = O\left(n^{1-r/2}v_n^{-r}(\log n)^{r/2}\right).$$

Moreover, from (VII) we have $\sup_{|\alpha|\le M\eta_n} |n^{1/2}G(\alpha/n^{1/2}) - H\alpha| = O(n^{-(1+s)/4}\eta_n^{(3+s)/2})$. Hence, since $n^{-(1+s)/4}\eta_n^{(3+s)/2} \le v_n(\log n)^{1/2}$, the term $n^{1/2}G(\alpha/n^{1/2})$ can be replaced in (3.2) by $H\alpha$.

Now consider a finite δ_n triangulation of the ball $\{|\alpha| \le M\eta_n + 1\}$ consisting of $O((\eta_n/\delta_n)^d)$ points. From (3.2) it follows that

$$\left|n^{1/2}G_n\left(\frac{\alpha}{n^{1/2}}\right) - n^{1/2}\binom{n}{m}^{-1} S_n - H\alpha\right| \le K_0 v_n \log n^{1/2}$$

holds simultaneously for all α belonging to the triangulation with probability equal to $1 - O((\eta_n/\delta_n)^d n^{1-r/2}v_n^{-r}(\log n)^{r/2}$. Now apply Lemma 6 of Niemiro [12] to extend this inequality to all points of the ball. Let $\epsilon_n = 2K_0 v_n(\log n)^{1/2} + 4K_1\delta_n$, where K_0 is as earlier, and K_1 is a constant from (VII). Then

$$(3.3) \quad P\left[\sup_{|\alpha|\le M\eta_n}\left|n^{1/2}G_n\left(\frac{\alpha}{n^{1/2}}\right) - n^{1/2}U_n - H\alpha\right| \ge \epsilon_n\right] = O\left(\left(\frac{\eta_n}{\delta_n}\right)^d n^{1-r/2}v_n^{-r}(\log n)^{r/2}\right).$$

Suppose that the event

$$(3.4) \quad \sup_{|\alpha|\le M\eta_n}\left|n^{1/2}G_n\left(\frac{\alpha}{n^{1/2}}\right) - n^{1/2}\binom{n}{m}^{-1} S_n - H\alpha\right| \le \epsilon_n$$

holds for some n. Consider radial directional derivatives of the convex function $nQ_n(\alpha/n^{1/2}) - nQ_n(0)$ on the sphere $|\alpha - n^{1/2}H^{-1}U_n| = K\epsilon_n$ setting the value $1/K = \inf_{|e|=1} e'He/2$. Under (3.4), for $|e| = 1$ we have

$$\left|n^{1/2}G_n\left(\frac{-n^{1/2}H^{-1}U_n + K\epsilon_n e}{n^{1/2}}\right) - n^{1/2}U_n - H(-n^{1/2}H^{-1}U_n + K\epsilon_n e)\right| \le \epsilon_n,$$

so that

$$\left|e'n^{1/2}G_n\left(\frac{-n^{1/2}H^{-1}U_n + K\epsilon_n e}{n^{1/2}}\right) - e'HeK\epsilon_n\right| \le \epsilon_n.$$

and hence we have

$$e'n^{1/2}G_n\left(\frac{-n^{1/2}H^{-1}U_n + K\epsilon_n e}{n^{1/2}}\right) \geq e'HeK\epsilon_n - \epsilon_n > 0.$$

Thus these radial directional derivatives are positive whenever (3.4) holds. Therefore we have the result

$$P(|R_n| > K\epsilon_n) = O\left(\left(\frac{\eta_n}{\delta_n}\right)^d n^{1-r/2}v_n^{-r}(\log n)^{r/2}\right). \qquad \square$$

REMARK. The above proof utilised an idea of Niemiro [12] where he used particular choices of δ_n, η_n, v_n and ϵ_n sequences. Our result provides a more precise rate of the error term R_n in the Bahadur representation of M- estimates.

PROOF OF COROLLARY 2.1. Using Theorem 2.1 one has

$$n^{1/2}(\theta_n - \theta_0) = -n^{1/2}H^{-1}U_n + R_n.$$

Thus

$$\frac{n^{1/2}c^T\Sigma^{-1}H(\theta_n - \theta_0)}{(2m^2(\log\log n))^{1/2}} = -\frac{n^{1/2}c^T\Sigma^{-1}U_n}{(2m^2(\log\log n))^{1/2}} + \frac{c^T\Sigma^{-1}HR_n}{(2m^2(\log\log n))^{1/2}}.$$

Now take $\eta_n = O((\log\log n)^{1/(3+s)}(\log n)^{-1/(3+s)}n^{(1+s)/2(3+s)})$, $\delta_n = O((\log\log n)^{1/2})$ and $v_n = O(n^{-(1+s)/4}\eta_n^{(3+s)/2})$ with the constants appropriately adjusted, to get $\epsilon_n = O((\log\log n)^{1/2})$. Now observe that for the sequences thus fixed and given value of r, we have from Theorem 2.1 that $P(|R_n| > K\epsilon_n) = O(n^{-\gamma})$ with $\gamma > 1$. Hence $\sum_n P(|R_n| > K\epsilon_n) < \infty$, and $(2m^2(\log\log n))^{-1/2}c^T\Sigma^{-1}HR_n \to 0$ almost surely. Now use the law of iterated logarithms for U-statistics to get the result. $\square$

PROOF OF THEOREM 2.2. (i) Recall that we have assumed that $\theta_0 = 0$ and $Q(\theta_0) = 0$. For $\epsilon > 0, y > 0$, put $r = 2x^2$ and $j = y^s\epsilon^{-r/s}$. Also recall the Bahadur representation $n^{1/2}(\theta_n) = -n^{1/2}H^{-1}U_n + R_n$. Then we have $P[\epsilon^2\tau_\epsilon \geq y] = P[\tau_\epsilon \geq j^{1/s}] = P[j^{s/r}\sup_{n\geq j^{1/s}}|\theta_n| > y^{1/2}]$. Now for fixed $\delta > 0$,

$$P\left[j^{s/r}\sup_{n\geq j^{1/s}}|\theta_n| > y^{1/2}\right]$$
$$\leq P\left[j^{s/r}\sup_{n\geq j^{1/s}}|-H^{-1}U_n| > y^{1/2} - \delta\right] + \sum_{n\geq j^{1/s}} P\left[j^{s/r}n^{-1/2}|R_n| > \delta\right].$$

As $\epsilon \to 0$ we have $j \to \infty$, and hence the sum $\sum_{n\geq j^{1/s}}$ actually represents a tail sum. Since for our choice of r we have $\sum_{n\geq 1} P[j^{s/r}n^{-1/2}|R_n| > \delta] < \infty$, so that

$\sum_{n\geq j^{1/x}} P[j^{x/r}n^{-1/2}|R_n| > \delta] \to 0$, and hence

$$\limsup_{\epsilon\to 0} P[\epsilon^2\tau_\epsilon \geq y] \leq \lim_{j\to\infty} P\left[j^{x/r} \sup_{n\geq j^{1/x}} |-H^{-1}U_n| > y^{1/2} - \delta\right]. \tag{3.5}$$

Similarly,

$$P\left[j^{x/r} \sup_{n\geq j^{1/x}} |\theta_n| > y^{1/2}\right]$$
$$\geq P\left[j^{x/r} \sup_{n\geq j^{1/x}} |-H^{-1}U_n| > y^{1/2} + \delta\right] - \sum_{n\geq j^{1/x}} P\left[j^{x/r}n^{-1/2}|R_n| > \delta\right],$$

and hence

$$\liminf_{\epsilon\to 0} P[\epsilon^2\tau_\epsilon \geq y] \geq \lim_{j\to\infty} P\left[j^{x/r} \sup_{n\geq j^{1/x}} |-H^{-1}U_n| > y^{1/2} + \delta\right]. \tag{3.6}$$

For any $c > 1$ and fixed j, define $\lambda_{\min}(H^{-2})$ to be the minimum eigenvalue of H^{-2}, we have

$$\begin{aligned} &P\left[j^{x/r} \sup_{n\geq cj^{1/x}} |-H^{-1}U_n| > y^{1/2} - \delta\right] \\ &\leq P\left[j^{x/r} \sup_{n\geq cj^{1/x}} |U_n|\lambda_{\min}^{1/2}(H^{-2}) > y^{1/2} - \delta\right] \\ &\leq \sum_{n\geq cj^{1/x}} P\left[|U_n| > (y^{1/2} - \delta)j^{-x/r}\lambda_{\min}^{-1/2}(H^{-2})\right]. \end{aligned} \tag{3.7}$$

The above sum in (3.7) is finite for every fixed c, and therefore it can be made arbitrarily small by choosing c large enough. Choose a large c such that $cj^{1/x}$ is an integer. It remains to deal with $P[j^{x/r} \sup_{j^{1/x}\leq n\leq cj^{1/x}} |-H^{-1}U_n| > y^{1/2} - \delta]$. Define

$$U_{cj^{1/x}}(t) = \begin{cases} 0, & \text{if} \quad 0 \leq t \leq (m-1)/cj^{1/x}; \\ \dfrac{l\Sigma^{-1}H^{-1}U_l}{m(cj^{1/x})^{1/2}}, & \text{if} \quad t = l/cj^{1/x};\ l = m, m+1, \ldots, cj^{1/x}; \end{cases}$$

and linearly in between. Then $U_{cj^{1/x}}(t) \Rightarrow W_0$ on $C_d[0,\infty]$ as $j \to \infty$ for fixed c (see, for example, Lee [10, pages 136–137] for a statement and proof of the univariate version of this theorem). The d-dimensional version is a direct generalisation. Using this convergence, it is easy to see that

$$\begin{aligned} j^{1/2x} \sup_{j^{1/x}\leq n\leq cj^{1/x}} |H^{-1}U_n| &= \sup_{c^{-1}\leq t\leq 1} |\Sigma U_{cj^{1/x}}(t)| \frac{m(cj^{1/x})^{1/2}j^{1/2x}}{[cj^{1/x}t]} \\ &\to mc^{-1/2} \sup_{c^{-1}\leq t\leq 1} \frac{|\Sigma W_0(t)|}{t} \quad \text{as } j \to \infty. \end{aligned}$$

Now $mc^{-1/2}\sup_{c^{-1}\le t\le 1}|\Sigma W_0(t)|/t$ has the same law as $m\sup_{1\le t\le c}c^{-1/2}|\Sigma W_0(t)|$ by a time-reversal of Brownian motion, and this in turn has the same law as $m\sup_{c^{-1}\le t\le 1}|\Sigma W_0(t)|$ by a time scaling of Brownian motion. Therefore, as $j\to\infty$,

$$P\left[j^{x/r}\sup_{j^{r/x}\le n\le j^{1/x}}|-H^{-1}U_n|>y^{1/2}-\delta\right]\to P\left[\sup_{c^{-1}\le t\le 1}|\Sigma W_0(t)|\ge\frac{y^{1/2}-\delta}{m}\right].$$

Then as $c\to\infty$,

$$P\left[\sup_{c^{-1}\le t\le 1}|\Sigma W_0(t)|\ge\frac{y^{1/2}-\delta}{m}\right]\to P\left[\sup_{0\le t\le 1}|\Sigma W_0(t)|\ge\frac{y^{1/2}-\delta}{m}\right] \tag{3.8}$$

and as $\delta\to 0$, this converges to $P[\sup_{0\le t\le 1}|\Sigma W_0(t)|\ge y^{1/2}/m]$ where W_0 is standard Brownian motion on $C_d[0,1]$. This proves (i). □

PROOF OF THEOREM 2.2. (ii) It suffices to prove that for some $\epsilon_0>0$, the family $\{\epsilon^2\tau_\epsilon : 0<\epsilon\le\epsilon_0\}$ is uniformly integrable. Recall that $P[\epsilon^2\tau_\epsilon\ge y]=P[j^{x/r}\sup_{n\ge j^{1/x}}|\theta_n|>y^{1/2}]$, where $j=y^x/\epsilon^{r/x}$. By application of Lemma 3.2 of Bickel and Yahav [2], we only need to prove

$$\sum_{k\ge 1}\sup_j P\left[j^{x/r}\sup_{n\ge j^{1/x}}|\theta_n|>k^{1/2}\right]<\infty, \tag{3.9}$$

where the first supremum extends over all j such that $j\ge k^x/\epsilon_0^{r/x}$. This is proved along lines similar to the proof of part (i) of this result. □

Acknowledgement

We thank the referee for his helpful comments.

References

[1] P. J. Bickel and E. L. Lehmann, 'Descriptive statistics for nonparametric models. IV. Spread', in: *Contributions to statistics* (ed. J. Jurečková) (Academia, Prague, 1979) pp. 33–40.

[2] P. J. Bickel and J. A. Yahav, 'Asymptotically optimal Bayes and minimax procedures in sequential estimation', *Ann. Math. Statist.* **39** (1968), 442–456.

[3] A. Bose, 'Bahadur representation and other asymptotic properties of M estimates based on U functionals', Tech. Rep. no. 19/94, (Stat.-Math. Unit, ISI, Calcutta, 1994).

[4] ———, 'Bahadur representation of M_m estimates', *Ann. Statist.* **26** (1998), 771–777.

[5] P. Chaudhuri, 'On a geometric notion of quantiles for multivariate data', *J. Amer. Statist. Assoc.* **91** (1996), 862–872.

[6] J. Choudhury and R. J. Serfling, 'Generalized order statistics, Bahadur representations, and sequential nonparametric fixed-width confidence intervals', *J. Statist. Plann. Inference* **19** (1988), 269–282.

[7] S. J. Habermann, 'Concavity and estimation', *Ann. Statist.* **17** (1989), 1631–1661.

[8] M. Hollander and D. A. Wolfe, *Nonparametrical statistical methods* (John Wiley and Sons, New York, 1973).

[9] P. J. Huber, 'Robust estimation of a location parameter', *Ann. Math. Statist.* **35** (1964), 73–101.

[10] A. J. Lee, *U-Statistics* (Marcel Dekker, New York, 1990).

[11] J. S. Maritz, M. Wu and R. G. Staudte, 'A location estimator based on U statistic', *Ann. Statist.* **5** (1977), 779–786.

[12] W. Niemiro, 'Asymptotics for M-estimators defined by convex minimization', *Ann. Statist.* **20** (1992), 1514–1533.

[13] H. Oja, 'Descriptive statistics for multivariate distribution', *Statist. Probab. Letters* **1** (1983), 327–333.

[14] H. Robbins and D. Siegmund, 'Boundary crossing probabilities for the Wiener process and sample sums', *Ann. Math. Statist.* **41** (1970), 1410–1429.

[15] W. Stute, 'Last passage times of M-estimators', *Scandinavian J. Statist.* **10** (1983), 301–305.

[16] ———, 'On a class of stopping times for M-estimators', *J. Multivariate Anal.* **14** (1984), 83–93.

Theoretical Statistics and Mathematics Unit
Indian Statistical Institute
203 B. T. Road
Calcutta 700 035
India
e-mail: abose@isical.ac.in

Department of Mathematics
The University of Manchester
Oxford Road
Manchester M13 9PL
UK
e-mail: s.chatterjee@maths.man.ac.uk