The Annals of Statistics
1995, Vol. 23, No. 4, 1331–1349

ON THE ASYMPTOTIC BEHAVIOUR OF THE MOVING BLOCK BOOTSTRAP FOR NORMALIZED SUMS OF HEAVY-TAIL RANDOM VARIABLES[1]

BY S. N. LAHIRI

Iowa State University

This paper studies the performance of the moving block bootstrap procedure for normalized sums of dependent random variables. Suppose that $X_1, X_2, \dots$ are stationary ρ-mixing random variables with $\sum \rho(2^i) < \infty$. Let $T_n = (X_1 + \cdots + X_n - b_n)/a_n$, for some suitable constants a_n and b_n, and let $T^*_{m,n}$ denote the moving block bootstrap version of T_n based on a bootstrap sample of size m. Under certain regularity conditions, it is shown that, for X_n's lying in the domain of partial attraction of certain infinitely divisible distributions, the conditional distribution $\hat{H}_{m,n}$ of $T^*_{m,n}$ provides a valid approximation to the distribution of T_n along every weakly convergent subsequence, provided $m = o(n)$ as $n \to \infty$. On the other hand, for the usual choice of the resample size $m = n$, $\hat{H}_{n,n}(x)$ is shown to converge to a nondegenerate random limit as given by Athreya (1987) when T_n has a stable limit of order α, $1 < \alpha < 2$.

1. Introduction. Let $X_1, X_2, \dots$ be a sequence of stationary random variables (r.v.'s) with common (marginal) distribution F. When the X_i's are independent and identically distributed (iid), it is well known that the bootstrap procedure of Efron (1979) (the EB) provides very accurate approximations to the distributions of many commonly used statistics. However, for dependent r.v.'s, the EB fails [cf. Singh (1981), Remark 2.1]. Recently, Künsch (1989) and Liu and Singh (1992) have formulated a moving block bootstrap (MBB) procedure which removes the deficiency of the EB for dependent data. Under suitable conditions, the MBB method is second-order correct for a large class of statistics based on sample means [Lahiri (1991), Götze and Künsch (1993)]. However, the asymptotic properties of the MBB for heavy-tail dependent r.v.'s have remained unknown. This paper investigates the asymptotic behaviour of the MBB when the sample mean (with a suitable normalization) has a nonnormal limit law.

For the sake of completeness, we now briefly describe the MBB procedure. Suppose that the statistic of interest derives from a functional $\tilde{t}(\cdot)$ through

$$\tilde{T}_n = \tilde{t}(F_n), \tag{1.1}$$

Received February 1992; revised December 1994.
[1]Supported in part by NSF Grant DMS-91-07998.
AMS 1991 *subject classifications*. Primary 62E20, 62G05; secondary 60F05.
Key words and phrases. Moving block bootstrap, stable limit, ρ-mixing, stationary, Poisson random measure.

Institute of Mathematical Statistics is collaborating with JSTOR to digitize, preserve, and extend access to *The Annals of Statistics.* www.jstor.org

where F_n denotes the empirical distribution of $(X_1, \dots, X_n)$. Given the data $\mathbf{X}_n = (X_1, \dots, X_n)$ and an integer l, $1 \le l < n$, define the observed blocks $\zeta_1, \dots, \zeta_b$ by

$$(1.2) \qquad \zeta_i = (X_i, \dots, X_{i+l-1}), \qquad i = 1, 2, \dots, b \equiv n - l + 1.$$

Next draw a simple random sample $\zeta_1^*, \dots, \zeta_k^*$, $k \ge 1$, with replacement from $\zeta_n \equiv \{\zeta_i : 1 \le i \le b\}$, and define the MBB version $\tilde{T}_{m,n}^*$ of $\tilde{T}_n$ by

$$(1.3) \qquad \tilde{T}_{m,n}^* = \tilde{t}(F_{m,n}^*),$$

where $m = kl$ and $F_{m,n}^*$ is the empirical distribution of the m components of ζ_i^*, $1 \le i \le k$. Then approximate the unknown distribution of $\tilde{T}_n$ by the conditional distribution of $\tilde{T}_{m,n}^*$ given $\mathbf{X}_n$, letting l and m increase to infinity with n suitably. In practice, one usually draws $k = k_0 \equiv [n/l]$ MBB blocks, where, for any real number x, $[x]$ denotes the largest integer not exceeding x. Although this choice of k is crucial for the second-order correctness of the MBB under higher moment assumptions, it will be shown that as in the iid case this is not necessarily the right choice for sums of heavy-tail dependent r.v.'s.

Note that the EB can also be considered as a moving block resampling scheme with block size $l \equiv 1$ for all $n \ge 1$. In contrast to the (general) MBB method, the performance of the EB for sums of heavy-tail iid r.v.'s is well studied in the literature. Starting with the pioneering work of Athreya (1987a), a number of important papers appeared on this topic over the last few years. See Athreya (1987a, b), Arcones and Giné (1989, 1991), Giné and Zinn (1989, 1990), Knight (1989), Hall (1990), Wu, Carlstein and Cambanis (1990) and references therein.

The main result of this paper (cf. Theorem 2.1) shows that, for F in the domain of *partial* attraction [for the definition, see (2.2)] of a class of infinitely divisible distributions (i.d.d.'s), the MBB works in probability for normalized sums of stationary, ρ-mixing random variables, provided the resample size $m = o(n)$, or equivalently, the number of resampled blocks $k = o(k_0)$, as $n \to \infty$. A direct implication of this result is that the same MBB resampling scheme can be used to approximate the distribution of normalized sums $\{T_n\}$, say, even when $\{T_n\}$ converges to distinct distributions along different subsequences. Since there exist distributions F that lie in the domain of partial attraction of *every* infinitely divisible distribution, using the large-sample distributions to approximate the distribution of T_n can be very unreliable in such cases, particularly in *finite* samples. In comparison, the bootstrap approximation adapts itself to the underlying forms of the distribution of T_n along every convergent subsequence and provides a valid and practically viable way of approximating the distribution of T_n.

It should be pointed out that Arcones and Giné (1989) are the first to prove a similar validity result for the EB for iid r.v.'s when F lies in the domain of partial attraction of an i.d.d. For F lying in the class of distributions considered in this paper, Theorem 2.1 extends their result on the EB for iid r.v.'s to the MBB for dependent r.v.'s.

It is interesting to note that, under the assumed conditions, the EB (which corresponds to blocks of size $l \equiv 1$) also yields a valid approximation for sums of *dependent* heavy-tail r.v.'s whenever $m = o(n)$ as $n \to \infty$. This is somewhat striking in view of the invalidity of the EB under dependence in the finite-variance case [cf. Singh (1981), Remark 2.1]. See Remark 2.2 in Section 2 for further details.

The second main result of the paper concerns the case when the number of resampled blocks k equals k_0, that is, the resample size m satisfies $mn^{-1} \to 1$ as $n \to \infty$. In this case we show that, for F lying in the domain of attraction of a stable law of order α, $1 < \alpha < 2$, the conditional distribution of the bootstrapped sample mean has a random limit as in Athreya (1987a). Thus, for sums of heavy-tail dependent r.v.'s, the MBB fails under the usual choice of resample size. To obtain a valid approximation, one needs to choose a resample size that grows at a slower rate than the size of the original sample.

The paper is organized as follows. Section 2 states the conditions and the main results of this paper. Section 3 gives the proofs of the theorems. Proofs of two technical lemmas used for proving Theorem 2.2 are relegated to Section 4.

2. Main results. As defined in Section 1, let $X_1, X_2, \ldots$ be a sequence of stationary r.v.'s, defined on a common probability space $(\Omega, \mathscr{F}, P)$, and let F denote the marginal distribution of X_1 under P. Write $S_n = X_1 + \cdots + X_n$, $n \geq 1$. For easy reference later on, define the strong-mixing coefficient $\alpha(\cdot)$ and the ρ-mixing coefficient $\rho(\cdot)$ of $X_1, X_2, \ldots$ as

$$\begin{aligned}\alpha(n) &= \sup\Big\{|P(A \cap B) - P(A)P(B)|: \\ &\qquad\qquad A \in \mathscr{F}_1^i,\ B \in \mathscr{F}_{i+n}^{\infty},\ i \geq 1\Big\}, \\ \rho(n) &= \sup\Big\{|Efg|\big((Ef^2)(Eg^2)\big)^{-1/2}: \\ &\qquad\qquad f \in L_2'(\mathscr{F}_1^i),\ g \in L_2'(\mathscr{F}_{i+n}^{\infty}),\ i \geq 1\Big\},\end{aligned} \tag{2.1}$$

where $\mathscr{F}_i^j$ is the σ-field generated by $X_i, \ldots, X_j$, $1 \leq i \leq j \leq \infty$, and

$$L_2'(\mathscr{F}_i^j) = \left\{ f\colon \Omega \to \mathbb{R} \,\middle|\, \int f^2\, dP < \infty,\ \int f dP = 0 \text{ and } f \text{ is } \mathscr{F}_i^j\text{-measurable} \right\}.$$

We say that the marginal distribution F of the stationary sequence $\{X_n\}_{n \geq 1}$ belongs to the *domain of partial attraction* of an i.d.d. F_0 if there exist a subsequence $\{N_n\}_{n \geq 1}$ and constants $b_{N_n} \in \mathbb{R}$, $a_{N_n} > 0$ such that

$$\left(\tilde{X}_1 + \cdots + \tilde{X}_{N_n} - b_{N_n}\right)/a_{N_n} \to_d W, \tag{2.2}$$

where $\{\tilde{X}_n\}_{n \geq 1}$ is the "associated" sequence of iid r.v.'s with common distribution F, W is a r.v. with distribution F_0 and $\to_d$ denotes convergence in distribution.

In this paper we will consider i.d.d.'s F_0 with characteristic functions (ch.f.'s) of the form

$$\text{(2.3)} \quad \phi_0(t) = \exp\left(\int x^{-2}(\exp(itx) - 1 - it\tau_c(x))M(dx)\right), \qquad t \in \mathbb{R},$$

where M is a canonical measure on $\mathbb{R}$ [i.e., $M(I) < \infty$ for all bounded intervals I, and $M^+(x) \equiv \int_{(x,\infty)} y^{-2}M(dy) < \infty$, $M^-(-x) \equiv \int_{(-\infty,-x]} y^{-2}M(dy) < \infty$, for all $x > 0$], with $M(\{0\}) = 0$, and $\tau_c(x) \equiv xI(|x| \le c)$, $x \in \mathbb{R}$, $c > 0$. Here and throughout the paper $I(A)$ denotes the indicator of a set A. Note that the restriction on M in (2.3) excludes the normal distribution.

Let $C(M)$ denote the set of continuity points of M. Then a necessary and sufficient condition for the convergence in (2.2) of the sums of iid r.v.'s to W having ch.f. ϕ_0 of (2.3) and $c \in C(M)$ is that as $n \to \infty$, for all $x \in C(M)$,

$$\text{(2.4)} \quad \begin{aligned} N_n\big(1 - F(xa_{N_n})\big) &\to M^+(x) \quad \text{if } x > 0,\\ N_nF(xa_{N_n}) &\to M^-(x) \quad \text{if } x < 0,\\ N_n a_{N_n}^{-2} EX_1^2 I\big(|X_1| \le ca_{N_n}\big) &\to M([-c,c]), \end{aligned}$$

and

$$a_{N_n}^{-1} b_{N_n} - N_n E\tau_c(X_1/a_{N_n})) \to 0.$$

In this case, one can centre the X_i's at $E\tau_c(X_1/a_{N_n})$ and get the convergence of $\sum_{1 \le i \le N_n}[a_{N_n}^{-1}X_i - E\tau_c(X_1/a_{N_n})]$ to W. In general, it is not possible further to replace $E\tau_c(X_1/a_{N_n})$ by EX_1/a_{N_n}. However, if one additionally assumes that

$$\text{(2.5)} \qquad \lim_{\lambda\to\infty} \limsup_{n\to\infty} N_n a_{N_n}^{-1} E|X_1| I\big(|X_1| > \lambda a_{N_n}\big) = 0,$$

then it can be shown that, under some regularity conditions,

$$(X_1 + \cdots + X_{N_n} - N_n\mu)/a_{N_n} \to_d W_1,$$

where $\mu = EX_1$ and W_1 has ch.f. (2.3) with $c = +\infty$. From the statistical point of view, this seems to be the most important case where one can make inference on the population mean μ.

In recent years, a number of papers have appeared in the literature dealing with the weak convergence of normalized sums of weakly dependent r.v.'s to i.d.d.'s. See Davis (1983), Samur (1984), Jakubowski and Kobus (1989), Denker and Jakubowski (1989) and references therein. For proving such results, in addition to the usual assumptions on the tails of F, one needs to impose different regularity conditions depending on the form of the canonical measure of the limiting i.d.d. For simplicity, we restrict attention to i.i.d.'s which can appear as weak limits of normalized sums of X_i's [in (2.2)], satisfying (2.5), and for which $M(\{0\}) = 0$. One can treat the other cases similarly.

For proving the results of this paper, we will assume the following dependence structure: the X_i's are ρ-mixing with

$$\sum_{n=1}^{\infty} \rho(2^n) < \infty, \tag{2.6}$$

and

$$\Psi^* \equiv \limsup_{x\to\infty} \sup_{n\geq 1} \frac{P(X_1 > x, X_{n+1} > x)}{(P(X_1 > x))^2} < \infty. \tag{2.7}$$

Condition (2.6) is quite common for proving central limit theorems for ρ-mixing r.v.'s [cf. Ibragimov (1975) and Peligrad (1982)]. The quantity Ψ^* in (2.7) is closely related to the well-known Ψ-mixing coefficient, defined by

$$\Psi(n) = \sup\left\{ \frac{|P(A\cap B) - P(A)P(B)|}{P(A)P(B)} : A \in \mathscr{F}_1^i,\, B \in \mathscr{F}_{i+n}^{\infty},\, i \geq 1 \right\}.$$

Both (2.6) and (2.7) appear in Jakubowski and Kobus (1989) and Denker and Jakubowski (1989) in the context of proving the weak convergence of S_n to a nonnormal stable law.

Next, following the description of the MBB in Section 1, define the bootstrap version of the normalized sample mean

$$T_n = (S_n - n\mu)/a_n \tag{2.8}$$

as

$$T^*_{m,n} = (S^*_{m,n} - E_* S^*_{m,n})/a_m, \tag{2.9}$$

where $S^*_{m,n} = \zeta_1^* + \cdots + \zeta_k^*$, $m = kl$ and ζ_i^* is the sum of the l components of the resampled block $\boldsymbol{\zeta}_i^*$, $1 \leq i \leq k$. Here and in the following, E_* denotes the conditional expectation under the MBB resampling scheme, given $\mathbf{X}_n$. Also let ρ_0 denote a metric, metricizing the topology of weak convergence on the set of all probability measures on $\mathbb{R}$. Unless otherwise specified, limits in all order symbols are taken as $n \to \infty$. Furthermore, for notational simplicity, whenever feasible we drop n and other subsequential indices from the subscripts.

Now we are ready to state the main result.

THEOREM 2.1. *Suppose that* (2.2) *holds with* W *having ch.f.* ϕ_0 *of* (2.3). *Let* $k \to \infty$ *such that* $mn^{-1} = o(1)$, *and let* (2.5), (2.6) *and* (2.7) *hold. If the subsequence* $\{m_{N_n}: n \geq 1\}$ *is contained in* $\{N_n: n \geq 1\}$ *and* $k_{N_n}^{-1/2}(m_{N_n} N_n^{-1})(a_{N_n}/a_{m_{N_n}}) = o(1)$, *then*

$$\rho_0\left(\hat{H}_{m_{N_n}, N_n}, H_{N_n}\right) \to 0 \quad \textit{in probability as } n \to \infty,$$

where $\hat{H}_{j,n}(x) = E_* I(T^*_{j,n} \leq x)$ *and* $H_n(x) = P(T_n \leq x)$, $x \in \mathbb{R}$.

Thus, Theorem 2.1 shows that if F belongs to the domain of partial attraction of the i.d.d. in (2.3), then under suitable conditions the MBB can capture all subsequential weak limits of the normalized sums T_n, provided the resample size m grows slower than n through the given subsequence. In particular, when T_n converges in distribution to a stable distribution of order $1 < \alpha < 2$, then (2.2), (2.3) and (2.5) hold and the MBB works for any choice of k, l and m satisfying $k^{-1} = o(1)$ and $m = o(n)$.

REMARK 2.1. Under the conditions of Theorem 2.1, $H_{N_n}(x)$ converges weakly to an i.d.d. F_0 having ch.f. $\phi_0(t)$ of (2.3) with $c = +\infty$ (cf. Lemma 3.5 below). Hence, $\hat{H}_{m_{N_n}, N_n}$ also converges weakly to F_0 in probability [i.e., $\rho_0(\hat{H}_{m_{N_n}, N_n}, F_0) = o_p(1)$].

REMARK 2.2. Note that Theorem 2.1 admits the choice $l \equiv 1$ for all $n \geq 1$. Thus Theorem 2.1 shows that, under the above condition, the EB (which corresponds to the MBB with blocks of size $l \equiv 1$) provides a valid approximation to the distribution of the normalized sum T_n under weak dependence in the heavy-tail case. This may seem somewhat surprising since the EB is known to fail drastically under dependence when X_1 has enough finite moments and T_n is asymptotically normal. A simple justification for this phenomenon comes from the observation that, under the assumed regularity conditions, the limit distribution of T_n depends *only* on the characteristics of the marginal distribution F of X_1 as opposed to depending on the *joint* distribution of X_i's in the finite-variance case. As a result, the EB, which resamples from an estimator of the marginal distribution of the X_i's, can approximate the limit distribution of T_n in our case, but fails in the finite-variance case.

REMARK 2.3. When the constants $\{a_{N_n}\}$ are unknown, Theorem 2.1 may not be very useful for constructing confidence intervals (CI's) for μ. If, however, there exist estimators $\hat{a}_n \equiv \hat{a}_n(\mathbf{X})$ satisfying $\hat{a}_N/a_N \to 1$ in probability, then it is possible to apply a "hybrid" MBB to approximate the distribution H_{1N} of the "studentized" statistic $T_{1N} = (S_N - N\mu)/\hat{a}_N$. Let $\hat{H}_{1N}$ denote the conditional distribution of $T^*_{1m_N} = a_{m_N} T^*_{m_N, N}/\hat{a}_{m_N}$ given $\mathbf{X}_n$. Then, under the conditions of Theorem 2.1, $\rho_0(\hat{H}_{1N}, H_{1N}) \to 0$ in probability and, hence, one can use the quantiles of $\hat{H}_{1N}$ to construct bootstrap CI's for μ.

REMARK 2.4. An alternative data-based approximation for the distributions of normalized sample means can be obtained using the subsampling method [cf. Politis and Romano (1994) and Hall and Jing (1994)], which corresponds to the MBB with $k = 1$ for all n. Validity of the approximation in the present problem essentially follows from a very general result (namely Theorem 3.1) of Politis and Romano (1994).

Next we briefly consider the asymptotic behaviour of the MBB when the number of resampled blocks k equals $k_0 \equiv [n/l]$, so that $mn^{-1} \to 1$ as $n \to \infty$. As one might expect, the MBB does not provide a valid approximation to $H_n(\cdot)$ in this case; under suitable conditions the bootstrap distribution of

$T^*_{m,n}$ converges to a random distribution. Here we establish this fact when F belongs to the domain of attraction of a stable distribution F_α of order α, $1 < \alpha < 2$. More specifically, assume that there exist constants $p \geq 0$, $q \geq 0$, $p + q = 1$, such that, for all $x > 0$,

$$(2.10)\quad P(X_1 > x) = px^{-\alpha}L(x) \quad \text{and} \quad P(X_1 < -x) = qx^{-\alpha}L(x),$$

where $1 < \alpha < 2$, and $L(x)$ is a function varying slowly at ∞, that is,

$$(2.11)\qquad \lim_{x \to \infty} \frac{L(tx)}{L(x)} = 1 \quad \text{for all } t > 0.$$

Let $\{a_n\}_{n \geq 1}$ be a sequence of constants such that

$$(2.12)\qquad nL(a_n)/a_n^\alpha \to 1 \quad \text{as } n \to \infty.$$

Then T_n [defined by (2.8) with a_n as in (2.12)] converges in distribution to a stable law of order α with ch.f. $\phi_\alpha(t)$, where

$$(2.13)\qquad \log \phi_\alpha(t) = \int (\exp(itx) - 1 - itx)\, d\lambda_\alpha(x), \qquad t \in \mathbb{R},$$

and, for any Borel subset A of $\mathbb{R}$,

$$(2.14)\qquad \lambda_\alpha(A) = \alpha\left[\int_{A \cap (0,\infty)} px^{-1-\alpha}\, dx + \int_{A \cap (-\infty, 0)} q|x|^{-1-\alpha}\, dx\right].$$

However, the (random) distribution function $\hat{H}_{m,n}(x)$ of $T^*_{m,n}$ converges in distribution to a random limit distribution $\hat{H}$ (say) as in Athreya (1987a, b). To describe it, let $\mathscr{B}(\mathbb{R})$ denote the Borel σ-field on $\mathbb{R}$. Also, let $N(\cdot)$ be a Poisson random measure on $(\mathbb{R}, \mathscr{B}(\mathbb{R}))$ with mean measure $\lambda_\alpha(\cdot)$, that is:

(a) $\{N(A) : A \in \mathscr{B}(\mathbb{R})\}$ is a collection of r.v.'s defined on some probability space $(\tilde{\Omega}, \tilde{\mathscr{F}}, \tilde{P})$, such that, for each $\tilde{\omega} \in \tilde{\Omega}$, $N(\cdot)(\tilde{\omega})$ is a measure on $(\mathbb{R}, \mathscr{B}(\mathbb{R}))$,

and

(b) for every disjoint collection of sets $A_1, \dots, A_j \in \mathscr{B}(\mathbb{R})$, $N(A_1), \dots, N(A_j)$ are independent Poisson r.v.'s with respective means $\lambda_\alpha(A_1), \dots, \lambda_\alpha(A_j)$, where $\lambda_\alpha(\cdot)$ is as defined in (2.14), and $j \geq 1$.

Then, the ch.f. $\hat{\phi}$ of the random limit $\hat{H}$ is given by

$$(2.15)\qquad \log \hat{\phi}(t) = \int (\exp(itx) - 1 - itx) N(dx), \qquad t \in \mathbb{R}.$$

THEOREM 2.2. *Suppose that* (2.6), (2.7) *and* (2.10) *hold and that* T_n *and* $T^*_{m,n}$ *are defined by* (2.8) *and* (2.9) *with* a_n *from* (2.12). *Let* $l^{-1} + n^{-1/2}l = o(1)$ *as* $n \to \infty$, *and* $m = lk_0 \equiv l[n/l]$, $n \geq 1$. *If, in addition,* $n\alpha(l)/l = O(1)$, *then, for every* $x_1, \dots, x_r \in \mathbb{R}$, $r \geq 1$,

$$\left(\hat{H}_{m,n}(x_1), \dots, \hat{H}_{m,n}(x_r)\right) \to_d \left(\hat{H}(x_1), \dots, \hat{H}(x_r)\right) \quad \text{as } n \to \infty,$$

where $\alpha(\cdot)$ *is as in* (2.1).

Thus, for the usual choice of the resample size, the MBB fails to provide a valid approximation to the distribution of the normalized sum T_n.

3. Proofs. Let $S_n(x, y) = \sum_{i=1}^n X_i I(x \le |X_i| < y)$, $S_n(x, \infty) = S_n(x)$, $0 \le x \le y \le \infty$, $n \ge 1$. Also, let $S_{ij} = X_i + \cdots + X_j$, $1 \le i \le j < \infty$. For any two sequences $\{r_n\}_{n\ge 1}$ and $\{t_n\}_{n \ge 1}$ of positive real numbers, write $r_n \ll t_n$ if $r_n = o(t_n)$, as $n \to \infty$. Thus, "$1 \ll t_n$" is equivalent to "$t_n \to \infty$ as $n \to \infty$." Let D denote a generic positive constant, not depending on n.

LEMMA 3.1. *Assume that* (2.4), (2.5), (2.6) *and* (2.7) *hold. Let* $l_n = o(n)$, $1 \ll k_n$ *be such that* $m_n \equiv k_n l_n \le n$ *for all* $n \ge 1$ *and* $\{m_{N_n}\colon n \ge 1\} \subseteq \{N_n\colon n \ge 1\}$. *Then, the following hold*:

(a) *for all* $y > 0$ *and* $\Delta = \pm 1$ *with* $\Delta \cdot y \in C(M)$,

$$\max_{1 \le r \le l_N} k_N \big| P\big(\Delta \cdot (S_r - r\mu) > a_{m_N} y\big) - rP\big(\Delta \cdot (X_1 - \mu) > a_{m_N} y\big)\big| = o(1);$$

(b) $$\lim_{\delta \downarrow 0} \lim_{n \to \infty} k_N a_{m_N}^{-2} E(S_{l_N} - l_N \mu)^2 I\big(|S_{l_N} - l_N\mu| < \delta a_{m_N}\big) = 0;$$

(c) $$\lim_{\lambda \to \infty} \lim_{n\to\infty} k_N a_{m_N}^{-1} E|S_{l_N} - l_N\mu| I\big(|S_{l_N} - l_N \mu| > \lambda a_{m_N}\big) = 0.$$

PROOF. Without loss of generality, assume that $\mu = 0$. Since in parts (a)–(c), l_n, k_n and m_n occur only along the subsequence $\{N\}$, for further notational simplicity, set $l_N = l$, $k_N = k$ and $m_N = m$.

Proof of (a). Without loss of generality, let $\Delta = +1$. Fix $y > 0$. Then for any δ and δ_1 satisfying $0 < \delta_1 < \delta + \delta_1 < y$, (2.7) implies

$$\begin{aligned} P(S_r > a_m y) &\le P(S_r(a_m(y - \delta_1)) > 0) + P(S_r(0, \delta a_m) > a_m \delta_1/2) \\ &\quad + P\big(S_r(\delta a_m, (y - \delta_1)a_m) > a_m(y - 2^{-1}\delta_1)\big) \\ &\le P\left(\bigcup_{i=1}^r \{X_i > a_m(y - \delta_1)\}\right) + P(|S_r(0, \delta a_m)| > a_m \delta_1/2) \\ &\quad + P\big(|X_i| > \delta a_m, |X_j| > \delta a_m \text{ for some } 1 \le i \ne j \le r\big) \\ &\le rP(X_1 > a_m(y - \delta_1)) + P(|S_r(0, \delta a_m)| > a_m \delta_1/2) \\ &\quad + Dl^2 P(|X_1| > \delta a_m)^2. \end{aligned} \tag{3.1}$$

Next, using the inequality $P(A) \ge P(B) - P(A' \cap B)$, for any $\delta_1 > 0$ one gets

$$\begin{aligned} P(S_r > a_m y) \ge{}& P\big(S_l(a_m(y + \delta_1)) > 0\big) \\ & - P\big(S_r < a_m y, S_r(a_m(y + \delta_1)) > 0\big). \end{aligned} \tag{3.2}$$

Also, for any $0 < \delta < \delta_1 + y$, as in the derivation of (3.1),

$$\begin{aligned}
&P\big(S_r < a_m y, S_r(a_m(y+\delta_1)) > 0\big)\\
&\quad \le P\big(S_r(0, \delta a_m) < -\delta_1 a_m/2\big)\\
&\qquad + \sum_{i=1}^{r} P\big(S_r(\delta a_m) < a_n(y + 2^{-1}\delta_1), X_i > a_m(y+\delta_1)\big)\\
&\quad \le P\big(|S_r(0, \delta a_m)| > \delta_1 a_m/2\big) + Dl^2 P\big(|X_1| > \delta a_m\big)^2.
\end{aligned} \tag{3.3}$$

By Bonferonni's inequality and Lemma 3.4 of Peligrad (1982),

$$\begin{aligned}
&P\big(S_r(a_m(y+\delta_1)) > 0\big)\\
&\quad \ge rP\big(X_1 > a_m(y+\delta_1)\big) - Dl^2 P\big(X_1 > a_m(y+\delta_1)\big)^2,
\end{aligned} \tag{3.4}$$

and

$$\begin{aligned}
&\max\big\{P\big(|S_r(0, \delta a_m)| > \delta_1 a_m/2\big) \colon 1 \le r \le l\big\}\\
&\quad \le D(\delta_1 a_m)^{-2}\Big[lEX_1^2 I(|X_1| < \delta a_m) + \big(lEX_1 I(|X_1| < \delta a_m)\big)^2\Big].
\end{aligned} \tag{3.5}$$

Since $\{m_N\} \subseteq \{N\}$, by (2.4) and (2.5), for any $\delta > 0$ with $M(\{-\delta, \delta\}) = 0$,

$$\begin{aligned}
&\lim_{n\to\infty} a_{m_N}^{-2} klEX_1^2 I(|X_1| < \delta a_m)\\
&\quad = \lim_{n\to\infty} a_{m_N}^{-2} m_N EX_1^2 I\big(|X_1| < \delta a_{m_N}\big) = M((-\delta, \delta)),
\end{aligned} \tag{3.6}$$

$$\begin{aligned}
&\lim_{n\to\infty} k a_m^{-2} l^2 \big(EX_1 I(|X_1| < \delta a_m)\big)^2\\
&\quad \le \lim_{n\to\infty} 2k_N^{-1}\big[\lambda m_N P\big(|X_1| > \delta a_{m_N}\big)\big]^2\\
&\qquad + \lim_{n\to\infty} 2k_N^{-1}\big[m_N a_{m_N}^{-1} E|X_1| I\big(|X_1| > \lambda a_{m_N}\big)\big]^2\\
&\quad = 0 \quad \text{if one chooses } \lambda > 0 \text{ large.}
\end{aligned} \tag{3.7}$$

By (3.1)–(3.7), for any $0 < \delta_1 < \delta + \delta_1 < y$ with $y \pm \delta_1$, $\pm\delta \in C(M)$,

$$\begin{aligned}
&\lim_{n\to\infty} \max_{1\le r\le l} k|P(S_r > a_m y) - rP(X_1 > a_m y)|\\
&\quad \le D\lim_{n\to\infty}\Big[m|P(X_1 > (y-\delta_1)a_m) - P(X_1 > ya_m)|\\
&\qquad\qquad + m|P(X_1 > (y+\delta_1)a_m) - P(X_1 > ya_m)|\\
&\qquad\qquad + kl^2 P\big(|X_1| > \delta a_m\big)^2\Big]\\
&\qquad + D\delta_1^{-2} M((-\delta, \delta))\\
&\quad \le D\Big[\big(M^+(y-\delta_1) - M^+(y+\delta_1)\big)\\
&\qquad + \lim_{n\to\infty}\big(k_N^{-1}(M^+(\delta) + M^-(-\delta))^2\big)\Big]\\
&\qquad + D\delta_1^{-2} M((-\delta, \delta)).
\end{aligned} \tag{3.8}$$

Letting $\delta \to 0+$ and then $\delta_1 \to 0+$ [such that $\pm\delta, y \pm \delta_1 \in C(M)$], the result follows.

Proof of (b). By (3.5), for any $\delta_1, \delta > 0$,

$$\begin{aligned}
&ka_m^{-2}E(S_l)^2 I(|S_l| < \delta a_m)\\
&\quad = ka_m^{-2}\big[E(S_l)^2 I(|S_l| < \delta a_m, |S_l(\delta_1 a_m)| > 0)\\
&\qquad\qquad + E(S_l)^2 I(|S_l| < \delta a_m, |S_l(\delta_1 a_m)| = 0)\big]\\
&\quad \le ka_m^{-2}\big[(\delta a_m)^2 lP(|X_1| \ge \delta_1 a_m) + E(S_l(0, \delta_1 a_m))^2\big]\\
&\quad \le D\delta^2 mP(|X_1| \ge \delta_1 a_m)\\
&\qquad + Dka_m^{-2}\big[lEX_1^2 I(|X_1| < \delta_1 a_m) + (lEX_1 I(|X_1| < \delta_1 a_m))^2\big].
\end{aligned}$$

Hence, using (3.6) and (3.7), as in the proof of part (a), for $\pm\delta, \pm\delta_1 \in C(M)$,

$$\begin{aligned}
&\lim_{\delta\to 0+}\ \limsup_{n\to\infty} k_N a_{m_N}^{-2} E(S_{l_N} - l_N\mu)^2 I(|S_{l_N} - l_N\mu| < \delta a_{m_N})\\
&\quad \le \lim_{\delta\to 0+}\big[D\delta^2(M^+(\delta_1) + M^-(-\delta_1)) + DM((-\delta_1, \delta_1))\big]\\
&\quad = DM((-\delta_1, \delta_1)),
\end{aligned}$$

which tends to zero as $\delta_1 \to 0+$. This proves part (b).

Proof of (c). Fix $\lambda_1 > 0$ such that $\pm\lambda_1 \in C(M)$. Then, by Lemma 3.4 of Peligrad (1982) and the proper convergence of $Nx^2\, dF(xa_N)$ to $M(dx)$ [cf. Feller (1966), page 527],

$$\begin{aligned}
&\limsup_{n\to\infty} ka_m^{-1}E(|S_l| I(|S_l| > \lambda a_m))\\
&\quad \le \limsup_{n\to\infty} 2ka_m^{-1}\{E|S_l(0, \lambda_1 a_m) - ES_l(0, \lambda_1 a_m)| I(|S_l(0, \lambda_1 a_m)\\
&\qquad\qquad - ES_l(0, \lambda_1 a_m)| > 2^{-1}a_m) + 2E|S_l(\lambda_1 a_m)|\}\\
&\quad \le \limsup_{n\to\infty}\big[D\lambda^{-1}ka_m^{-2}\operatorname{Var}(S_l(0, \lambda_1 a_m)) + Dka_m^{-1}lE|X_1| I(|X_1| > \lambda_1 a_m)\big]\\
&\quad \le D\lambda^{-1}M([-\lambda_1, \lambda_1]) + D\lim_{n\to\infty} m_N a_{m_N}^{-1} E|X_1| I(|X_1| > \lambda_1 a_{m_N}).
\end{aligned}$$

Hence, the result follows by (2.5). □

LEMMA 3.2. *Under the conditions of Lemma* 3.1, *for all* $x > 0$ *with* $\Delta \cdot x \in C(M)$,

$$\operatorname{Var}\left(k_N b_N^{-1}\sum_{j=1}^{b_N}\big(I(\Delta\cdot\zeta_j > xa_m) - P(\Delta\cdot\zeta_1 > xa_m)\big)\right) = O(m_N N^{-1}),$$

where $\Delta = \pm 1$, $\zeta_j = X_j + \cdots + X_{j+l_N-1} - l_N\mu$, $j \ge 1$, *and* $b_N = N - l_N + 1$.

PROOF. As in the proof of Lemma 3.1, set $l_N = l$, $k_N = k$, $m_N = m$, $b_N = b$ and $\mu = 0$. Without loss of generality, take $\Delta = 1$ and write $Y_j \equiv I(\zeta_j > xa_m)$

$-P(\zeta_1 > xa_m)$, $j \geq 1$. With $l_0 = l + 1$, define the block sums of the Y_j's as $\hat{Y}_r = \sum_{(r-1)l_0 < j \leq rl_0} Y_j$, $r \geq 1$. Also, let $K = [b/2l_0]$. Then $\sum_{j=1}^{b} Y_j = \sum_{r=1}^{K} \hat{Y}_{2r} + \sum_{r=1}^{K} \hat{Y}_{2r-1} + \hat{R}_N$, where $\hat{R}_N$ is defined by subtraction. Note that $\{\hat{Y}_{2r}\}_{r\geq 1}$ is a mean-zero stationary sequence with ρ-mixing coefficient $\hat{\rho}(r) \leq \rho(rl_0 - l) \leq \rho(r)$ for all $r \geq 1$. Hence, by Lemma 3.4 of Peligrad (1982),

$$E\left(\sum_{r=1}^{K} \hat{Y}_{2r}\right)^2 \leq 8000\left[\prod_{r=1}^{\infty}(1 + \rho(2^r))\right] KE\left(\hat{Y}_2\right)^2.$$

Hence, using similar arguments for $\sum_{r=1}^{K} \hat{Y}_{2r-1}$, by Lemma 3.1 one gets

$$\begin{aligned}
\operatorname{Var}&\left(kb^{-1}\sum_{j=1}^{b} Y_j\right)\\
&\leq 4(kb^{-1})^2\left[E\left(\sum_{r=1}^{K}\hat{Y}_{2r}\right)^2 + E\left(\sum_{r=1}^{K}\hat{Y}_{2r-1}\right)^2 + E\left(\hat{R}_N\right)^2\right]\\
&\leq Dk^2N^{-1}l^{-1}\left[E\left(\hat{Y}_2\right)^2 + E\left(\hat{Y}_1\right)^2 + E\left(\hat{R}_n\right)^2\right]\\
&\leq Dk^2N^{-1}l^{-1}\left[l^2EY_1^2\right]\\
&= O\left(m_N N^{-1}\right).
\end{aligned}$$

This completes the proof of Lemma 3.2. □

For Lemmas 3.3 and 3.4 and the proof of Theorem 2.2, define

$$\begin{aligned}
l_1 &= \min\left\{\left[l\left(\rho([l^{1/2}])\right)^{-1/2}\right], n^{3/4}\right\},\\
l_1' &= l + \left[(l_1 l)^{1/2}\right], \qquad l_2 = l_1' + l_1, \qquad m_1 = [b/l_2],\\
\Gamma_i(J) &= \sum_{il_2 < j \leq il_2 + l_1} k_0 b^{-1} I\left((\zeta_j - l\mu)a_n^{-1} \in J\right), \qquad 0 \leq i < m_1,\\
\Delta_i(J) &= \sum_{il_2 + l_1 < j \leq (i+1)l_2} k_0 b^{-1} I\left((\zeta_j - l\mu)a_n^{-1} \in J\right), \qquad 0 \leq i < m_1, \qquad (3.9)\\
\Delta_{m_1}(J) &= \sum_{m_1 l_2 \leq j \leq b} k_0 b^{-1} I\left((\zeta_j - l\mu)a_n^{-1} \in J\right), \quad J \subseteq \mathbb{R}\setminus\{0\},\\
N_{1n} &= \sum_{0 \leq i < m_1} \Gamma_i \quad \text{and} \quad N_{2n} = \sum_{0 \leq i \leq m_1} \Delta_i.
\end{aligned}$$

LEMMA 3.3. *Suppose that the conditions of Theorem 2.2 hold. Then, for any two intervals $J_i \subseteq \mathbb{R}\setminus\{0\}$, $i = 1, 2$, with* $\sup\{x\colon x \in J_1\} < \inf\{x\colon x \in J_2\}$,

$$m_1 E\Gamma_1(J_1)\Gamma_1(J_2) = o(1).$$

See Section 4 for the proof.

LEMMA 3.4. *Assume that the conditions of Theorem* 2.2 *hold, and let* $J \subseteq \mathbb{R} \setminus \{0\}$ *be an interval. Then we have the following*:

(a) $m_1 E\Gamma_1(J)^2 I(|\Gamma_1(J) - 1| > \delta) = o(1)$ *for every* $\delta > 0$;
(b) $m_1 E\Gamma_1(J)^2 = \lambda_\alpha(J) + o(1)$.

See Section 4 for the proof.

LEMMA 3.5. *Assume that* (2.4)–(2.7) *hold. Then*

$$a_N^{-1}(S_N - N\mu) \to_d W_1 \quad \textit{as } n \to \infty,$$

where W_1 *has ch.f.*

$$\phi(t) = \exp\left(\int (\exp(itx) - 1 - itx) x^{-2} M(dx)\right), \qquad t \in \mathbb{R}. \tag{3.10}$$

PROOF. We only outline a proof here. Let $l_3 \equiv l_{3N}$ and $l_4 \equiv l_{4N}$ be positive integers such that $l_4 l_3^{-1} + l_3 N^{-1} = o(1)$ and $\rho(l_4)(Nl_3^{-1}) = o(1)$. Set $\mu = 0$ and write $l_5 = l_3 + l_4$, $K = [N/l_5]$, $V_i = S_{(i-1)l_5 + l_3} - S_{(i-1)l_5}$, $i \geq 1$ (where $S_0 \equiv 0$), $S_{1N} = \sum_{1 \leq i \leq K} V_i$, $S_{2N} = S_N - S_{1N}$. Then, using (2.4), Lemma 3.1 (with $m_N = N$), a truncation at some $c > 0$ with $\pm c \in C(M)$ and Lemma 3.4 of Peligrad (1982), one can show that $S_{2N}/a_N \to 0$ in probability and that

$$\left| E\exp\left(\frac{itS_{1N}}{a_N}\right) - \left(E\exp\left(\frac{itV_1}{a_n}\right)\right)^K \right| \leq DK\rho(l_4) = o(1),$$

$$\lim_{\delta \downarrow 0} \limsup_{n \to \infty} Ka_N^{-2} EV_1^2 I(|V_1| \leq \delta a_N) = 0 = \lim_{\delta \downarrow 0} M([-\delta, \delta]),$$

$$\lim_{n \to \infty} KP(\Delta \cdot V_1 > xa_N) = M^\Delta(\Delta \cdot x) \quad \text{for all } \Delta \cdot x \in C(M),\ x > 0,$$

$$\lim_{\lambda \to \infty} \limsup_{n \to \infty} Ka_N^{-1} E|V_1| I(|V_1| > \lambda a_N) = 0,$$

and by Fatou's Lemma, integration by parts, and (2.5),

$$\begin{aligned}
&\lim_{\lambda \to \infty} \int |x|^{-1} I(|x| > \lambda) M(dx) \\
&\quad \leq \lim_{\lambda \to \infty} 2\int_{\lambda/2}^{\infty} [M^+(y) + M^-(-y)]\, dy \\
&\quad \leq \lim_{\lambda \to \infty} \limsup_{n \to \infty} 2\int_{\lambda/2}^{\infty} NP(|X_1| > a_N y)\, dy \\
&\quad \leq \lim_{\lambda \to \infty} \limsup_{n \to \infty} 2Na_N^{-1} E|X_1| I\left(|X_1| > \frac{\lambda a_N}{2}\right) = 0.
\end{aligned} \tag{3.11}$$

Lemma 3.5 now follows by standard arguments. Details are omitted. □

PROOF OF THEOREM 2.1. By Lemma 3.5, it is enough to show that

$$\rho_0\left(\hat{H}_{m_N, N}, H\right) = o_P(1), \tag{3.12}$$

where $\int \exp(itx)\,dH(x) = \phi(t)$, $t \in \mathbb{R}$ [cf. (3.10)]. Let $\hat{\phi}_n(t) = E_* \exp(it(\zeta_1^* - E_* S_1^*))$, $t \in \mathbb{R}$. Since a sequence of r.v.'s converges in probability if and only if, given a subsequence, there is a further subsequence through which it converges almost surely, (3.12) would hold if, for any subsequence $\{N''\}$ of $\{N\}$, there is a subsequence $\{N'\}$ of $\{N''\}$ such that, for every sample point $w \in A$ for some $A \in \mathscr{F}$ with $P(A) = 1$,

$$K_{N'}\left(E_* \frac{(\zeta_1^* - l_{N'}\mu)}{a_{m_{N'}}}\right)^2 = o(1), \tag{3.13}$$
$$K_{N'}\int f_t(x)\,dG_{m_{N'}}(x) = \int f_t(x)x^{-2}M(dx) + o(1) \quad \text{for all } t \in \mathbb{R},$$

where $f_t(x) = \exp(itx) - 1 - itx$, $x \in \mathbb{R}$, and $G_m \equiv G_{m,n}$ denotes the conditional distribution of $(\zeta_1^* - l\mu)/a_m$ given $\mathbf{X}_n$.

Let $\mu = 0$. Then, by a diagonalization argument involving suitable countable dense subsets of x, δ and λ's, it is enough to show that

$$k_N E_*\left(\frac{\zeta_1^*}{a_{m_N}}\right)^2 = o_P(1), \tag{3.14}$$

$$k_N P_*(\Delta\zeta_1^* > xa_{m_N}) = M^{\Delta}(\Delta x) + o_P(1), \tag{3.15}$$

$$k_N E_*\left(\frac{\zeta_1^*}{a_{m_N}}\right)^2 I(|\zeta_1^*| < \delta a_{m_N}) = M((-\delta,\delta)) + o_P(1), \tag{3.16}$$

$$k_N E_*\left|\frac{\zeta_1^*}{a_{m_N}}\right| I(|\zeta_1^*| > \lambda a_{m_N}) = \int_{|x|>\lambda} |x|^{-1}M(dx) + o_P(1), \tag{3.17}$$

for all $x > 0$, $\delta > 0$, $\lambda > 0$, $\Delta = \pm 1$ such that Δx, $\pm\delta$ and $\pm\lambda \in C(M)$.

For $\varepsilon > 0$ by (2.4), Lemma 3.4 of Peligrad (1982) and Lemma 3.5,

$$k_N^{1/2} a_{m_N}^{-1} b_N^{-1} l_N |S_N| = O_P\left(k_N^{-1/2}(M_N N^{-1})\left(\frac{a_N}{a_{m_N}}\right)\right) = o_P(1)$$

and

$$\begin{aligned} &P\left(k_N^{1/2} a_{m_N}^{-1} b_N^{-1}\left|\sum_{1\le j\le l_N}(j - l_N)(X_j + X_{N-j+1})\right| > \varepsilon\right)\\ &\quad\le 2l_N P(|X_1| > ca_{m_N})\\ &\qquad + 2\varepsilon^{-2}k_N a_{m_N}^{-2} b_N^{-2} E\left(\sum_{1\le j< l_N}(l_N - j)X_j I(|X_j| \le ca_{m_N})\right)^2\\ &\quad\le Dk_N^{-1} + D\varepsilon^{-2}(N^{-2}l_N^2 + c^{-2}N^{-3}l_N^3) = o(1) \quad \text{where } c \in C(M),\ c > 0. \end{aligned}$$

This proves (3.14). Also, (3.15) follows from (2.4) and Lemma 3.2. Next, let $Y_{1i} = a_{m_n}^{-2}\zeta_i^2 I(|\zeta_i| < \delta a_{m_N})$ and $Y_i = Y_{1i} - EY_{1i}$, $i \ge 1$. Then, by (2.4), Lemma 3.1 and the blocking argument in the proof of Lemma 3.2,

$$\begin{aligned} k_N^2 \operatorname{Var}\left(E_* \left(\frac{\zeta_1^*}{a_{m_N}} \right)^2 I(|\zeta_1^*| < \delta a_{m_N}) \right) &= k_N^2 b_N^{-2} E\Big(\sum_{1 \le i \le b_N} Y_i \Big)^2 \\ &\le D k_N^2 b_N^{-2} (b_N l_N) EY_1^2 \\ &\le D M_N b_N^{-1} (2\delta)(k_N EY_{11}) \\ &= O\left(\frac{M_N}{N} \right), \end{aligned}$$

which implies (3.16). Hence, it remains to prove (3.17). Let $\{\lambda_n\}$ be a sequence such that $\lambda_n^{-1} = o(1)$ and $\lambda_N^2 \le (N/m_N)$ for all N. Define $Y_{2i} = a_{m_N}^{-1}|\zeta_i| I(|\zeta_i| \ge \lambda_N)$, $Y_{3i} = a_{m_N}^{-1}|\zeta_i| I(\lambda < |\zeta_i/a_{m_N}| < \lambda_N)$, $i \ge 1$. Then, by (2.4), (3.11) and Lemma 3.1, the expected value of the l.h.s. of (3.17) is equal to $k_N(EY_{21} + EY_{31})$, which tends to $\int |x|^{-1} I(|x| > \lambda) M(dx)$. Hence, (3.17) follows by noting that $k_N EY_{21} = o(1)$ so that, for every $\varepsilon > 0$, for sufficiently large n,

$$\begin{aligned} &P\left(k_N b_N^{-1} \Big| \sum_{1 \le i \le b_N} (Y_{2i} + Y_{3i} - E(Y_{21} + Y_{31})) \Big| > 3\varepsilon \right) \\ &\quad \le P\Big(k_N b_N^{-1} \sum_{1 \le i \le b_N} Y_{2i} > \varepsilon \Big) + P\left(k_N b_N^{-1} \Big| \sum_{1 \le i \le b_N} (Y_{3i} - EY_{3i}) \Big| > \varepsilon \right) \\ &\quad \le \varepsilon^{-1}(k_N EY_{21}) + \varepsilon^{-2} k_N^2 b_N^{-2} \operatorname{Var}\Big(\sum_{1 \le i \le b_N} Y_{3i} \Big) \\ &\quad \le \varepsilon^{-2}(M_N N^{-1} \lambda_N)\left\{ k_N a_{m_N}^{-1} E|\zeta_1| I(|\zeta_1| > \lambda a_{m_N}) \right\} + o(1) = o(1), \end{aligned}$$

by the blocking argument of Lemma 3.2. This completes the proof of Theorem 2.1. □

PROOF OF THEOREM 2.2. By Theorem A of Athreya (1987a), Lemma 3.1 and (3.14), it is enough to show that, for any disjoint collection $J_1, \dots, J_r$ of closed intervals in $\mathbb{R} \setminus \{0\}$ and for any $c_1, \dots, c_r$, $r \ge 1$,

$$k_0 \sum_{j=1}^{r} c_j G_n(J_j) \to_d \sum_{j=1}^{r} c_j N(J_j), \tag{3.18}$$

where $G_n(x) = E_* I(|\zeta_1^*| \le x a_n)$, $x \in \mathbb{R}$. Note that $k_0 G_n(\cdot) = N_{1n}(\cdot) + N_{2n}(\cdot)$, where the N_{in}'s are as in (3.9). For any $J = [\beta, \gamma] \subseteq \mathbb{R} \setminus \{0\}$,

$$EN_{2n}(J) \le m_1 l_1' k_0 b^{-1} P(|\zeta_1| > Da_n) + l_2 k_0 b^{-1} P(|\zeta_1| > Da_n) = o(1).$$

Hence, (3.18) holds if $\sum_{j=1}^{r} c_j N_{1n}(J_j) \to_d \sum_{j=1}^{r} c_j N(J_j)$. By the α-mixing condition on the X_i's,

$$\left| E\exp\left(it\sum_{j=1}^{r} c_j N_{1n}(J_j)\right) - \left(E\exp(itU_1)\right)^{m_1}\right|$$
$$\leq Dm_1\alpha(l_1' - l) \leq D\left(\frac{n\alpha(l)}{l}\right)(l_1^{-1}l) = o(1),$$

where $\{U_i: i \geq 1\}$ are iid and $U_1 =_d \sum_{j=1}^{r} c_j\Gamma_1(J_j)$. Since $N(J_1), \dots, N(J_r)$ are independent Poisson r.v.'s, it is enough to show that, for $\delta > 0$ small,

$$m_1 EU_1 = \sum_{j=1}^{r} c_j\lambda_j + o(1), \tag{3.19}$$

$$m_1\int_{|x-c_j|<\delta} x^2\, d\tilde{G}_n(x) = \lambda_j c_j^2 + o(1), \qquad 1 \leq j \leq r, \tag{3.20}$$

and

$$m_1\int_{A_\delta} x^2\, d\tilde{G}_n(x) = o(1), \tag{3.21}$$

where $\lambda_j = \lambda_\alpha(J_j)$ and $A_\delta = \{x: |x - c_j| \geq \delta \text{ for all } 1 \leq j \leq r\}$.

Equation (3.19) is implied by (2.10) and Lemma 3.1. Note that, by Lemma 3.3 and Lemma 3.4, for all $0 < \delta < 1$,

$$\begin{aligned}
&m_1\int_{A_\delta} x^2\, d\tilde{G}_n(x)\\
&\quad = m_1 E\left(\sum c_j\Gamma_1(J_j)\right)^2 I\left(\left|\sum c_j\Gamma_1(J_j) - c_{j_1}\right| > \delta \text{ for all } j_1\right)\\
&\quad \leq Dm_1\left[\sum_{i\neq j} E\Gamma_1(J_i)\Gamma_1(J_j) + \sum_{j=1}^{r} E\Gamma_1(J_j)^2 I\left(\left|\sum_{i=1}^{r} c_i\Gamma_1(J_i) - c_j\right| > \delta\right)\right]\\
&\quad \leq o(1) + \sum_{j=1}^{r} m_1 E\Gamma_1(J_j)^2 I\left(|c_j|\cdot|\Gamma_1(J_j) - 1| > \frac{\delta}{2}\right)\\
&\qquad + \sum_{j=1}^{r} m_1 E\Gamma_1(J_j)^2 I\left(\sum_{i\neq j}|c_i\Gamma_1(J_i)| + |c_j(\Gamma_1(J_j) - 1)| > \delta,\right.\\
&\qquad\qquad\qquad \left. |c_j(\Gamma_1(J_j) - 1)| < \frac{\delta}{2}\right)\\
&\quad \leq o(1) + D\delta^{-1}\sum_{j=1}^{r} m_1 E\Gamma_1(J_j)\left(\sum_{i\neq j}|c_i\Gamma_1(J_i)|\right) = o(1),
\end{aligned}$$

proving (3.20). One can prove (3.21) similarly. Hence, Theorem 2.2 follows. □

4. Proofs of Lemmas 3.3 and 3.4. Without loss of generality, let $\mu = 0$.

PROOF OF LEMMA 3.3. Without loss of generality, let $J_i = (\beta_i, \delta_i)$, $i = 1, 2$, where $0 < \beta_1 < \gamma_1 < \beta_2 < \delta_2 \le \infty$. Note that $J_1 \cap J_2 = \varnothing$ implies

$$\begin{aligned} &m_1 E\Gamma_1(J_1)\Gamma_1(J_2) \\ &\quad \le Dnl^{-2} \sum_{1<i\le l} P(\zeta_1 \in J_1 a_n, \zeta_i \in J_2 a_n) \\ &\qquad + Dnl^{-2} \sum_{l<i\le l_1} P(|\zeta_1| > \beta_1 a_n)(\rho(l-i) + P(|\zeta_1| > \beta_1 a_n)) \\ &\quad \equiv R_{1n} + R_{2n} \quad \text{(say)}. \end{aligned} \tag{4.1}$$

Note that, for any sequence $\{l'\}$, $1 \ll l' \ll n$, by Lemma 3.1(a) and (2.12),

$$\max_{1\le r\le l'} P(|S_r| > Da_n) = O(n^{-1}l'). \tag{4.2}$$

For any r.v.'s W_1, W_2 and W_3 and real numbers $0 < a < b_1 < c < d$ (with $a_0 = a/2$),

$$\begin{aligned} &P(a \le W_1 + W_2 \le b_1, c \le W_2 + W_3 \le d) \\ &\quad \le P(W_1 \ge a_0, W_2 \le b_1 - W_1, c \le b_1 - W_1 + W_3) \\ &\qquad + P(W_1 < a_0, a - a_0 < W_2 \le b_1 - W_1, c \le b_1 - W_1 + W_3) \\ &\quad \le P(2W_1 \ge a, W_3 \ge c + a_0 - b_1) \\ &\qquad + P(2W_1 \le -(c - b_1), 2W_2 \ge a) \\ &\qquad + P(a \le 2W_2, (c - b_1) \le 2W_3). \end{aligned} \tag{4.3}$$

Let $l_6 = [l^{1/2}]$, $l_7 = [l_6^{1/2}]$ and $A_i = \{|S_i - S_{i-l_7}| < \delta_0 a_n\}$, $i > l_7$, where $8\delta_0 = \min\{\beta_1, \beta_2 - \gamma_1\}$. Then, by (4.2) [see Lahiri (1993)],

$$\begin{aligned} R_{1n} &\le Dnl^{-2} \sum_{l_6<i<l-l_6} P(\{\zeta_1 \in J_1 a_n, \zeta_i \in J_2 a_n\} \cap A_i \cap A_{i-l_7}) + o(1) \\ &\equiv R_{11n} + o(1) \quad \text{(say)}. \end{aligned}$$

Using (4.1) and (4.3) (with $W_{1i} = S_{i-l_7}$, $W_{2i} = S_{l-l_7} - S_{i-1}$ and $W_{3i} = \zeta_i - S_{l-1}$), one gets

$$\begin{aligned} R_{11n} &\le Dnl^{-2} \sum_{l_6<i<l-l_6} [P(W_{1i} > \delta_0 a_n, W_{3i} > \delta_0 a_n) \\ &\qquad + P(W_{1i} < -\delta_0 a_n, W_{2i} > \delta_0 a_n) \\ &\qquad + P(W_{2i} > \delta_0 a_n, W_{3i} > \delta_0 a_n)] \\ &\le Dnl^{-2} l\left[(n^{-1}l)^2 + \rho(l_7)(n^{-1}l)\right] \\ &\le D\left[n^{-1}l + \rho(l_7)\right] = o(1). \end{aligned}$$

By similar arguments,

$$
\begin{aligned}
R_{2n} &\le Dnl^{-2}P(|\zeta_1| > \beta_1 a_n)\cdot\left[l_6 + l_1\,\rho(l_6) + n^{-1}l_1 l\right] \\
&= O\big(l^{-1}l_6 + l^{-1}l_1\,\rho(l_6) + n^{-1}l_1\big).
\end{aligned}
\tag{4.4}
$$

This proves the lemma. □

PROOF OF LEMMA 3.4. Using Lemmas 3.1 and 3.3, one can show [see Lahiri (1993)] that it is enough to prove Lemma 3.4 only for $J = [x, \infty)$, $x > 0$.

Proof of (a). Let $Z_i = \sum_{(i-1)l < j \le il} k_0 b^{-1} I(\zeta_j \in Ja_n)$; let $\tilde{S}_1$ (respectively, $\tilde{S}_2$) denote the sum of Z_i's over all even i (odd i), $1 \le i \le 2l_8$; and let $R_{3n} = \Gamma_1(J) - \sum_{1 \le i \le 2l_8} Z_i$, where $l_8 = [l_1/2l]$. Clearly,

$$
m_1 E R_{3n}^2 \le D m_1 l^2 \big(k_0 b^{-1}\big)^2 E\big(I(\zeta_1 \in Ja_n)\big)^2 \le D l_1^{-1} l,
$$

and

$$
\begin{aligned}
&m_1 \sum_{2 \le i < 2l_8} (2l_8 - i) E Z_1 Z_i \\
&\quad \le D n l_1^{-1}\left[l_8 E Z_1 Z_2 + l_8^2 (E Z_1)^2 + l_8^2 E Z_1\,\rho(l)\right] = o(1),
\end{aligned}
\tag{4.5}
$$

by (4.2). Note that $|Z_i| = O(1)$ a.s. Hence by (4.5), for large n,

$$
\begin{aligned}
&m_1 E \tilde{S}_1^2 I(|\Gamma_1(J) - 1| > \delta) \\
&\quad \le D n l_1^{-1} \sum_{1 \le i \le l_8} E Z_{2i}^2 \left[I\left(|Z_{2i-1} + Z_{2i} + Z_{2i+1} - 1| > \frac{\delta}{2}\right) \right. \\
&\qquad\qquad \left. + I\left(|\Gamma_1(J) - Z_{2i-1} - Z_{2i} - Z_{2i+1}| > \frac{\delta}{2}\right)\right] + o(1) \\
&\quad \le D n l^{-2} \delta^{-2} \sum_{i=1}^{l} E I(\zeta_{i+l} \in Ja_n)(Z_1 + Z_2 + Z_3 - 1)^2 + o(1) \\
&\quad \le D n l^{-1} \delta^{-2} \left[E I(\zeta_l \in Ja_n)\left(l^{-1} \sum_{j=1}^{2l} I(\zeta_j \in Ja_n) - 1\right)^2 \right] \\
&\qquad + D n l^{-2} \delta^{-2} \sum_{l < j < 2l} P(\zeta_1 \in Ja_n, \zeta_j \in Ja_n) + o(1) \\
&\quad = D n l^{-1} \delta^{-2} E I(\zeta_l \in Ja_n)\left(l^{-1} \sum_{j=1}^{2l} I(\zeta_j \in Ja_n) - 1\right)^2 + o(1),
\end{aligned}
\tag{4.6}
$$

where the last step follows from an argument similar to (4.4).

Next, a lengthy but straightforward calculation shows that

$$
\begin{aligned}
&nl^{-1}EI(\zeta_2 \in Ja_n)\left(l^{-1}\sum_{j=1}^{2l} I(\zeta_j \in Ja_n)\right)^2 \\
&\quad = 6nl^{-3}\sum\sum_{1\le i<j\le l} P(\zeta_1 \in Ja_n, \zeta_i \in Ja_n, \zeta_j \in Ja_n) + o(1). \\
&\quad = \left[6l^{-3}\sum_{1\le i<j\le l}(l-j)\right]\left[nP(X_1 > xa_n)\right] + o(1) \\
&\quad = px^{-\alpha} + o(1).
\end{aligned} \tag{4.7}
$$

Also, by similar arguments,

$$
2nl^{-2}EI(\zeta_l \in Ja_n)\left(\sum_{j=1}^{2l} I(\zeta_j \in Ja_n)\right) = 2px^{-\alpha} + o(1). \tag{4.8}
$$

Now, part (a) can be proved using (4.6), (4.7) and (4.8). See Lahiri (1993).

Proof of (b). As in the proof of Lemma 3.3 [cf. (4.4)],

$$
\begin{aligned}
&m_1 E\Gamma_1([x,\infty))^2 \\
&\quad = m_1(k_0 b^{-1})^2\left[l_1 P(\zeta_1 > xa_n) + 2\sum_{i=2}^{l}(l_1 - i)P(\zeta_1 > xa_n, \zeta_{1+i} > xa_n)\right] \\
&\qquad + o(1) \\
&\quad = 2nl^{-2}\sum_{i=1}^{l} P(\zeta_1 > xa_n, \zeta_{i+1} > xa_n) + o(1).
\end{aligned}
$$

The rest is as in (4.8). See Lahiri (1993) for further details. □

Acknowledgments. The author would like to thank an Associate Editor and an anonymous referee for a number of constructive suggestions that led to substantial improvement of an earlier draft of the paper.

REFERENCES

ARCONES, M. A. and GINÉ, E. (1989). The bootstrap of the mean with arbitrary sample size. *Ann. Inst. H. Poincaré* **25** 457–481.

ARCONES, M. A. and GINÉ, E. (1991). Additions and correction to "The bootstrap of the mean with arbitrary bootstrap sample size." *Ann. Inst. H. Poincaré* **27** 583–595.

ATHREYA, K. B. (1987a). Bootstrap of the mean in the infinite variance case. *Ann. Statist.* **15** 724–731.

ATHREYA, K. B. (1987b). Bootstrap of the mean in the infinite variance case. In *Proceedings of the 1st World Congress of the Bernoulli Society* (Y. Prohorov and V. V. Sazonov, eds.) **2** 95–98. VNU Science Press, Amsterdam.

DAVIS, R. A. (1983). Stable limits for partial sums of dependent random variables. *Ann. Probab.* **11** 262–269.

DENKER, M. and JAKUBOWSKI, A. (1989). Stable limit distributions for strongly mixing sequences. *Statist. Probab. Lett.* **8** 477–483.

EFRON, B. (1979). Bootstrap methods: Another look at the jackknife. *Ann. Statist.* **7** 1–26.

FELLER, W. (1966). *An Introduction to Probability Theory and Its Applications* **2**. Wiley, New York.

GINÉ, E. and ZINN, J. (1989). Necessary conditions for the bootstrap of the mean. *Ann. Statist.* **17** 684–691.

GINÉ, E. and ZINN, J. (1990). Bootstrapping general empirical measures. *Ann. Probab.* **18** 851–869.

GÖTZE, F. and KÜNSCH, H. R. (1993). Blockwise bootstrap for dependent observation: higher order approximations for studentized statistics. Technical report, Univ. Bielefeld, Germany.

HALL, P. (1990). Asymptotic properties of the bootstrap for heavy-tailed distributions. *Ann. Probab.* **18** 1342–1360.

HALL, P. and JING, B.-Y. (1994). On the sample re-use methods. Research Report SR8-94, Centre for Mathematics and its Applications, Australian National Univ.

IBRAGIMOV, I. A. (1975). A note on the central limit theorem for dependent random variables. *Theory Probab. Appl.* **20** 135–140.

JAKUBOWSKI, A. and KOBUS, M. (1989). α-stable limit theorems for sums of dependent random vectors. *J. Multivariate Anal.* **29** 219–251.

KNIGHT, K. (1989). On the bootstrap of the sample mean in the finite variance case. *Ann. Statist.* **17** 1168–1175.

KÜNSCH, H. R. (1989). The jackknife and the bootstrap for general stationary observations. *Ann. Statist.* **17** 1217–1241.

LAHIRI, S. N. (1991). Second order optimality of stationary bootstrap. *Statist. Probab. Lett.* **11** 335–341.

LAHIRI, S. N. (1993). On the asymptotic behaviour of the moving block bootstrap for normalized sums of heavy-tail random variables: II. Extensions. Preprint No. 93-19, Dept. Statistics, Iowa State Univ, Ames.

LIU, R. and SINGH, K. (1992). Moving blocks jackknife and bootstrap capture weak dependence. In *Exploring the Limits of Bootstrap* (R. Lepage and L. Billard, eds.) 225–248. Wiley, New York.

PELIGRAD, M. (1982). Invariance principles for mixing sequences of random variables. *Ann. Probab.* **10** 968–981.

POLITIS, D. and ROMANO, J. (1994). Large sample confidence regions based on subsamples under minimal assumptions. *Ann. Statist.* **22** 2031–2050.

SAMUR, J. D. (1984). Convergence of sums of mixing triangular arrays for random vectors with stationary laws. *Ann. Probab.* **12** 390–426.

SINGH, K. (1981). On the asymptotic accuracy of Efron's bootstrap. *Ann. Statist.* **9** 1187–1195.

WU, W., CARLSTEIN, E. and CAMBANIS, S. (1990). Bootstrapping the sample mean for data with infinite variance. Preprint 296, Dept. Statistics, Univ. North Carolina, Chapel Hill.

DEPARTMENT OF STATISTICS
IOWA STATE UNIVERSITY
AMES, IOWA 50011