

A HEAVY TRAFFIC LIMIT THEOREM FOR WORKLOAD PROCESSES WITH HEAVY TAILED SERVICE REQUIREMENTS

SIDNEY RESNICK AND GENNADY SAMORODNITSKY

ABSTRACT. A system with heavy tailed service requirements under heavy load having a single server, has an equilibrium waiting time distribution which is approximated by the Mittag-Leffler distribution. This fact is understood by a direct analysis of the weak convergence of a sequence of negative drift random walks with heavy right tail and the associated all time maxima of these random walks. This approach complements the recent transform view of Boxma and Cohen [4].

1. INTRODUCTION.

Heavy traffic limit theorems were devised to study the behavior of complex networks. For economic reasons, these systems are typically heavily loaded. If one wants to estimate the performance of such a system, a direct simulation may not be efficient, because the system parameters are near the boundary of the parameter set which makes the system stable. This means long excursions are likely. A heavy traffic limit theorem studies a sequence of systems under normalization when system parameters approach the stability boundary.

Original work assumed the component random variables of the model all had finite variance. This research originated with Kingman ([18, 18, 19]) and was nicely surveyed and updated by Whitt ([30, 29]). See also the summaries in [1] and [27]. This early work was built upon by J.M. Harrison, who with coworkers started the subject of diffusion process approximations (see [9]), which is still a subject of active research.

This classical work on heavy traffic approximations has little relevance to recent work in communication networks, which explains self similarity of network traffic by means of on/off models having infinite variance and heavy tailed transmission time distributions. (See [2, 28, 20, 21, 31, 22, 32, 11, 10, 12, 24, 16, 15, 14]). A recent stimulating paper by Boxma and Cohen [4] studies the stationary waiting time for the GI/G/1 queue under the assumptions that the system is under heavy traffic and the service distribution has infinite variance, while the interarrival distribution tail is of smaller order than the service time distribution tail. Assumptions and methodology use Laplace transform techniques. We demonstrate that one can analyze such systems without transforms and this direct attack illuminates and simplifies the various assumptions that must be made to achieve the heavy traffic limit theorem.

The setup is as follows. We assume we have a sequence of Lindley queues (see [1, 27]) indexed by k and the delay process of the k th Lindley queue is given by

$$W_0^{(k)} = 0, \quad W_{n+1}^{(k)} = (W_n^{(k)} + \tau_n^{(k)} - \sigma_{n+1}^{(k)})^+, \quad n \geq 0,$$

Key words and phrases. heavy tails, regular variation, Pareto tails, heavy traffic approximation, random walk, queueing.

The research of Sidney Resnick and Gennady Samorodnitsky was partially supported by NSF Grant DMS-97-04982 and NSA Grant MDA904-98-1-0041 at Cornell University. Portions of this work were completed while Sid Resnick was on sabbatical at the University of North Carolina, Chapel Hill, and the hospitality of the Department of Statistics is gratefully acknowledged.

where $\{\tau_n^{(k)}, n \geq 0\}$ are iid heavy tailed service times with finite mean and $\{\sigma_n^{(k)}, n \geq 1\}$ are iid interarrival times whose tails are of smaller order than the tails of the $\tau_n^{(k)}$'s. For each k , $\{\tau_n^{(k)}, n \geq 0\}$ and $\{\sigma_n^{(k)}, n \geq 1\}$ are independent. For each k , there is an associated random walk and we assume the k th random walk has negative drift so that the stationary waiting time exists; that is, for each fixed k , as $n \rightarrow \infty$ the variable $W_n^{(k)}$ converges in distribution to a nondegenerate and proper limit:

$$W_n^{(k)} \Rightarrow W^{(k)}.$$

If we assume the drift $E(\tau_n^{(k)}) - E(\sigma_{n+1}^{(k)}) \rightarrow 0$ as $k \rightarrow \infty$, we can impose conditions which guarantee that $W^{(k)}$ converges to a heavy traffic limit. The method is to study the sequence of associated random walk processes, show these converge under suitable conditions to a limit stable Lévy motion with negative drift and then prove that this drags along the convergence of the all time maxima. The maximum of the k th random walk has the same distribution as $W^{(k)}$. The maximum of the limit stable Lévy motion with negative drift has a Mittag-Leffler distribution ([8, 7]). The conclusion is that the equilibrium waiting time distribution of a heavily loaded GI/G/1 system whose service time distribution has finite mean, infinite variance and heavy tail is approximated by the Mittag-Leffler distribution.

Section 2 shows how to approximate a sequence of negative drift random walks so that the sequence converges to a limiting zero mean Lévy random motion with drift -1. In Section 3 we show that the distribution of the supremum of the limiting zero mean Lévy random motion with drift -1 approximates the suprema of the sequence of random walks. We specialize these results to the GI/G/1 system in Section 4.

The waiting time distribution of the GI/G/1 queue is only a rough approximation to the content process of a fluid queue and it would be of interest to apply these techniques to studying heavy traffic approximations for such communication models driven by heavy tailed transmission times. We hope to engage this and similar topics in the near future.

2. APPROXIMATION TO A NEGATIVE DRIFT RANDOM WALK.

In this section we assume that for each $k = 1, 2, \dots$, $\{X_i^{(k)}, i \geq 1\}$ are iid random variables. The k th random walk is

$$S_0^{(k)} = 0, \quad S_n^{(k)} = \sum_{i=1}^n X_i^{(k)}, \quad n \geq 1,$$

so that the k th random walk has steps $X_i^{(k)}$, $i = 1, 2, \dots$

We need to make the following assumptions.

ASSUMPTION 1. There exists a non-negative sequence of integers $d(k) \rightarrow \infty$ such that

$$(2.1) \quad \nu_k(\cdot) := d(k) P[X_1^{(k)} \in \cdot] \xrightarrow{v} \nu(\cdot),$$

vaguely in $[-\infty, \infty] \setminus \{0\}$, where ν is a measure on $[-\infty, \infty] \setminus \{0\}$ satisfying

- (a) ν is a Lévy measure,
- (b) $\nu(-\infty, 0) = 0$,
- (c) $\int_1^\infty x \nu(dx) < \infty$.

ASSUMPTION 2. How much mass is allowed near 0 is controlled by the condition that for any $M > 0$

$$\limsup_{k \rightarrow \infty} d(k) \text{Var} \left(X_1^{(k)} 1_{[|X_1^{(k)}| \leq M]} \right) < \infty$$

and

$$(2.2) \quad \lim_{\epsilon \downarrow 0} \limsup_{k \rightarrow \infty} d(k) \text{Var} \left(X_1^{(k)} 1_{\{|X_1^{(k)}| \leq \epsilon\}} \right) = 0.$$

ASSUMPTION 3. We assume each $X_i^{(k)}$ has a finite mean $\mu^{(k)}$ satisfying

- (a) $0 > \mu^{(k)} \rightarrow 0$ as $k \rightarrow \infty$,
- (b) $\lim_{k \rightarrow \infty} d(k) \mu^{(k)} = -1$.

ASSUMPTION 4. Just assuming ν is a Lévy measure does not provide sufficient control near infinity so we assume further that

$$(2.3) \quad \lim_{M \rightarrow \infty} \limsup_{k \rightarrow \infty} d(k) E \left(|X_i^{(k)}| 1_{\{|X_i^{(k)}| > M\}} \right) = 0.$$

With these assumptions in place, we can state and prove the first result about how a sequence of negative drift random walks can be approximated by a negative drift Lévy process.

Theorem 2.1. *Assume Assumptions 1–4 hold. Define the random element of $D[0, \infty)$*

$$Y^{(k)}(t) = S_{[d(k)t]}^{(k)}, \quad t \geq 0$$

for $k = 1, 2, \dots$. Let $\{\xi^{(\infty)}(t), t \geq 0\}$ be a totally skewed to the right zero mean Lévy process with Lévy measure ν and set $Y^{(\infty)}(t) = \xi^{(\infty)}(t) - t$, $t \geq 0$. Then in $D[0, \infty)$

$$Y^{(k)}(\cdot) \Rightarrow Y^{(\infty)}(\cdot)$$

where ' \Rightarrow ' denotes weak convergence.

Proof. There is, obviously, a $y > 0$ such that the Lévy measure ν has no atoms at $\pm y$. Without changing the argument in order to simplify the notation, we will assume for the duration of the proof that ν has no atoms at ± 1 .

From [26] or [25], (2.1) is equivalent to convergence of a sequence of point measures to a limiting Poisson point measure. We have (2.1) iff

$$(2.4) \quad \mathcal{P}_k := \sum_{i=1}^{\infty} \epsilon_{\left(\frac{i}{d(k)}, X_i^{(k)}\right)} \Rightarrow \sum_m \epsilon_{(t_m, j_m)} =: \mathcal{P}_\infty, \quad k \rightarrow \infty,$$

in $M_p([0, \infty) \times [-\infty, \infty] \setminus \{0\})$, where for a nice set \mathbb{E} , $M_p(\mathbb{E})$ denotes the set of point measures on \mathbb{E} , topologized by the vague metric. (See [26], [17], [23].) Note that the limit random point measure \mathcal{P}_∞ is a Poisson random measure on $[0, \infty) \times [-\infty, \infty] \setminus \{0\}$ with mean measure $L \times \nu$ where L is Lebesgue measure. Note also that because of Assumption 1b, we have $\mathcal{P}_\infty([0, \infty) \times (-\infty, 0)) = 0$ almost surely.

Choosing a $\delta > 0$ to avoid the atoms of ν we have from (2.4) by the continuous mapping theorem that

$$(2.5) \quad \begin{aligned} \sum_{i=1}^{d(k)t} X_i^{(k)} 1_{\{|X_i^{(k)}| > \delta\}} &= \iint_{[s \leq t, |x| > \delta]} x \mathcal{P}_n(ds, dx) \\ &\Rightarrow \iint_{[s \leq t, |x| > \delta]} x \mathcal{P}_\infty(ds, dx) = \sum_{t_m \leq t} j_m 1_{[j_m > \delta]}, \end{aligned}$$

as random elements of $D[0, \infty)$. From (2.1), because of the assumption that ν has no atoms at ± 1 and at $\pm \delta$, we have as $k \rightarrow \infty$

$$(2.6) \quad d(k) E \left(X_1^{(k)} 1_{\{|X_1^{(k)}| \in (\delta, 1]\}} \right) = \int_{[\delta < |x| \leq 1]} x \nu_k(dx) \rightarrow \int_{[\delta < |x| \leq 1]} x \nu(dx)$$

and combining (2.5) and (2.6) yields the $D[0, \infty)$ convergence

$$(2.7) \quad \begin{aligned} X^{(k, \delta)}(t) &= \sum_{i=1}^{[d(k)t]} X_i^{(k)} 1_{[|X_i^{(k)}| > \delta]} - [d(k)t] E(X_1^{(k)} 1_{[|X_1^{(k)}| \in (\delta, 1]}) \\ &\Rightarrow \sum_{t_m \leq t} j_m 1_{[j_m > \delta]} - t \int_{[\delta < |x| \leq 1]} x \nu(dx) =: X^{(\infty, \delta)}(t). \end{aligned}$$

We know from the Itô representation (see [13]) of a Lévy process, that as $\delta \rightarrow 0$, almost surely and locally uniformly

$$X^{(\infty, \delta)}(\cdot) \rightarrow X^{(\infty)}(\cdot),$$

where

$$X^{(\infty)}(t) = \xi^{(\infty)}(t) + t \int_1^\infty x \nu(dx).$$

Define

$$X^{(k)}(t) = \sum_{i=1}^{[d(k)t]} \left(X_i^{(k)} - E(X_i^{(k)} 1_{[|X_i^{(k)}| \leq 1]}) \right)$$

and we claim

$$(2.8) \quad X^{(k)}(\cdot) \Rightarrow X^{(\infty)}(\cdot)$$

in $D[0, \infty)$. Observe that

$$X^{(k)}(t) - X^{(k, \delta)}(t) = \sum_{i=1}^{[d(k)t]} \left(X_i^{(k)} 1_{[|X_i^{(k)}| \leq \delta]} - E(X_i^{(k)} 1_{[|X_i^{(k)}| \leq \delta]}) \right)$$

so to prove the claim (2.8), it suffices to show for any $\eta > 0$ and $K > 0$ that

$$\lim_{\delta \downarrow 0} \limsup_{k \rightarrow \infty} P \left[\sup_{0 \leq t \leq K} |X^{(k)}(t) - X^{(k, \delta)}(t)| > \eta \right] = 0.$$

For simplicity, take $K = 1$ and the previous probability is

$$P \left[\sup_{0 \leq t \leq 1} \left| \sum_{i=1}^{[d(k)t]} \left(X_i^{(k)} 1_{[|X_i^{(k)}| \leq \delta]} - E(X_i^{(k)} 1_{[|X_i^{(k)}| \leq \delta]}) \right) \right| > \eta \right] \leq \eta^{-2} \text{Var} \left(\sum_{i=1}^{[d(k)]} X_i^{(k)} 1_{[|X_i^{(k)}| \leq \delta]} \right),$$

from Kolmogorov's inequality. As $k \rightarrow \infty$ and then $\delta \downarrow 0$, this last expression converges to 0 by (2.2) of Assumption 2.

Now center (2.8) to zero expectations. We have

$$d(k) \left(\mu^{(k)} - E(X_1^{(k)} 1_{[|X_1^{(k)}| \leq 1]}) \right) = d(k) E(X_1^{(k)} 1_{[|X_1^{(k)}| > 1]})$$

and

$$d(k) E(X_1^{(k)} 1_{[|X_1^{(k)}| > 1]}) \rightarrow \int_1^\infty x \nu(dx)$$

since the absolute value of the difference can be bounded by

$$\begin{aligned} &|d(k) E(X_1^{(k)} 1_{[1 < |X_1^{(k)}| \leq M]}) - \int_1^M x \nu(dx)| + d(k) E(X_1^{(k)} 1_{[|X_1^{(k)}| > M]}) + \int_M^\infty x \nu(dx) \\ &= I + II + III, \end{aligned}$$

for an arbitrary M chosen to avoid the atoms of ν . As $k \rightarrow \infty$, $I \rightarrow 0$ by vague convergence (2.1). We can make II as small as desired by (2.3) of Assumption 4 and III is made small by Assumption 1c. We therefore conclude that

$$\sum_{i=1}^{[d(k)t]} X_i^{(k)} - [d(k)t]\mu^{(k)} \Rightarrow X^{(\infty)}(t) - t \int_1^\infty x\nu(dx)$$

in $D[0, \infty)$ and furthermore

$$Y^{(k)}(t) = S_{[d(k)t]}^{(k)} \Rightarrow X^{(\infty)}(t) - t \int_1^\infty x\nu(dx) - t = Y^{(\infty)}(t)$$

in $D[0, \infty)$ where we have used Assumption 3b. This completes the proof. \square

3. APPROXIMATION TO THE SUPREMUM OF A NEGATIVE DRIFT RANDOM WALK.

The supremum of a negative drift random walk is of interest because of its relation to the equilibrium waiting time of GI/G/1 queueing models. In this section we discuss how the approximation of Section 2 to the negative drift random walk, implies an approximation to the supremum. We continue using the notation defined in the previous section.

Theorem 3.1. *Assume Assumptions 1–4 hold. Define*

$$W^{(k)} := \bigvee_{0 \leq t < \infty} Y^{(k)}(t) = \bigvee_{n=0}^{\infty} S_n^{(k)}.$$

Then in \mathbb{R} , we have the convergence in distribution as $k \rightarrow \infty$

$$W^{(k)} \Rightarrow W^{(\infty)} := \bigvee_{t=0}^{\infty} Y^{(\infty)}(t) = \bigvee_{t=0}^{\infty} (\xi^{(\infty)}(t) - t),$$

where recall $\xi^{(\infty)}(\cdot)$ is the zero mean Lévy process of Theorem 2.1.

Proof. We use a method described by Asmussen in [1]. For any $T > 0$ the map $x \mapsto \bigvee_{s=0}^T x(s)$ from $D[0, \infty) \mapsto \mathbb{R}$ is continuous so from Theorem 2.1 we have

$$\bigvee_{s=0}^T Y^{(k)}(s) \Rightarrow \bigvee_{s=0}^T Y^{(\infty)}(s)$$

in \mathbb{R} . The desired result will be proven provided we can show for any $\eta > 0$ that

$$(3.1) \quad \lim_{T \rightarrow \infty} \limsup_{k \rightarrow \infty} P\left[\bigvee_{j \geq d(k)T} S_j^{(k)} > \eta \right] = 0.$$

To prove (3.1), we observe that for any suitably chosen $M > 0$

$$\begin{aligned} P\left[\bigvee_{j \geq d(k)T} S_j^{(k)} > 0 \right] &= P\left[\bigvee_{j \geq d(k)T} \frac{S_j^{(k)}}{j} > 0 \right] \\ &\leq P\left[\bigvee_{j \geq d(k)T} j^{-1} \sum_{i=1}^j X_i^{(k)} 1_{\|X_i^{(k)}\| \leq M} + j^{-1} \sum_{i=1}^j X_i^{(k)} 1_{\|X_i^{(k)}\| > M} > 0 \right] \\ &= P\left[\bigvee_{j \geq d(k)T} j^{-1} \sum_{i=1}^j X_i^{(k)} 1_{\|X_i^{(k)}\| \leq M} - E(X_1^{(k)} 1_{\|X_1^{(k)}\| \leq M}) \right] \end{aligned}$$

$$\begin{aligned}
& + j^{-1} \sum_{i=1}^j X_i^{(k)} 1_{\lfloor |X_i^{(k)}| > M \rfloor} - E(X_1^{(k)} 1_{\lfloor |X_1^{(k)}| > M \rfloor}) > |\mu^{(k)}| \Big] \\
\leq & P\left[\bigvee_{j \geq d(k)T} j^{-1} \sum_{i=1}^j \left(X_i^{(k)} 1_{\lfloor |X_i^{(k)}| \leq M \rfloor} - E(X_i^{(k)} 1_{\lfloor |X_i^{(k)}| \leq M \rfloor}) \right) > |\mu^{(k)}|/2 \right] \\
& + P\left[\bigvee_{j \geq d(k)T} j^{-1} \sum_{i=1}^j \left(X_i^{(k)} 1_{\lfloor |X_i^{(k)}| > M \rfloor} - E(X_i^{(k)} 1_{\lfloor |X_i^{(k)}| > M \rfloor}) \right) > |\mu^{(k)}|/2 \right] \\
= & I + II.
\end{aligned}$$

The centered sample averages are reversed martingales and so we may apply Kolmogorov's inequality. We will use the fact that from Assumption 3b, $d(k) \sim 1/|\mu^{(k)}|$ as $k \rightarrow \infty$. For I we have

$$I \leq (\mu^{(k)}/2)^{-2} \text{Var} \left(\frac{1}{d(k)T} \sum_{i=1}^{\lfloor d(k)T \rfloor} X_i^{(k)} 1_{\lfloor |X_i^{(k)}| \leq M \rfloor} \right) \leq 4 \frac{d(k)T}{d(k)^2 T^2} \frac{1}{(\mu^{(k)})^2} \text{Var}(X_1^{(k)} 1_{\lfloor |X_1^{(k)}| \leq M \rfloor}).$$

Using $(d(k)\mu^{(k)})^2 \rightarrow 1$, as $k \rightarrow \infty$ this is

$$\sim \frac{1}{T} d(k) \text{Var}(X_1^{(k)} 1_{\lfloor |X_1^{(k)}| \leq M \rfloor}).$$

This converges to 0 as $T \rightarrow \infty$ since

$$\limsup_{k \rightarrow \infty} d(k) \text{Var}(X_1^{(k)} 1_{\lfloor |X_1^{(k)}| \leq M \rfloor}) < \infty$$

from (2.2) of Assumption 2. To kill II, we write using the martingale maximal inequality,

$$\begin{aligned}
II & \leq \frac{2}{|\mu^{(k)}|} E \left| \frac{1}{\lfloor d(k)T \rfloor} \sum_{i=1}^{\lfloor d(k)T \rfloor} \left(X_i^{(k)} 1_{\lfloor |X_i^{(k)}| > M \rfloor} - E(X_i^{(k)} 1_{\lfloor |X_i^{(k)}| > M \rfloor}) \right) \right| \\
& \leq \frac{2}{|\mu^{(k)}|} E \left| X_1^{(k)} 1_{\lfloor |X_1^{(k)}| > M \rfloor} - E(X_1^{(k)} 1_{\lfloor |X_1^{(k)}| > M \rfloor}) \right| \\
& \sim 2d(k) E \left| X_1^{(k)} 1_{\lfloor |X_1^{(k)}| > M \rfloor} - E(X_1^{(k)} 1_{\lfloor |X_1^{(k)}| > M \rfloor}) \right| \\
& \leq 4d(k) E(|X_1^{(k)}| 1_{\lfloor |X_1^{(k)}| > M \rfloor}).
\end{aligned}$$

From (2.3) of Assumption 4, if we choose M sufficiently large, then $\limsup_{k \rightarrow \infty} II$ can be made as small as desired and this completes the proof. \square

4. HEAVY TRAFFIC FOR QUEUES WITH HEAVY TAILED SERVICES.

In this section we suppose the steps of the k th random walk are of the form

$$X_i^{(k)} = \frac{\tau_i^{(k)} - \sigma_{i+1}^{(k)}}{a(k)}, \quad i \geq 1,$$

where $\{\tau_i^{(k)}, i \geq 1\}$ is a non-negative iid sequence (of service lengths) with common distribution $B^{(k)}(x)$ and $\{\sigma_i^{(k)}, i \geq 1\}$ is an independent sequence of non-negative iid interarrival times with

common distribution $A^{(k)}(x)$ and $a(k)$ is a suitable scaling sequence to be determined. Then the step mean is

$$\mu^{(k)} = \frac{E(\tau_1^{(k)}) - E(\sigma_1^{(k)})}{a(k)}$$

and the traffic intensity is

$$\rho^{(k)} = E(\tau_1^{(k)})/E(\sigma_1^{(k)}).$$

We seek reasonable assumptions on $A^{(k)}(x)$ and $B^{(k)}(x)$ so as to be able to apply Theorems 2.1 and 3.1.

We assume the following conditions.

CONDITION (A). Suppose there exists a distribution function F concentrating on $[0, \infty)$ such that $\bar{F} := 1 - F \in RV_\alpha$; that is, \bar{F} is regularly varying with index $-\alpha$. It follows that the quantile function

$$(4.1) \quad b(t) = \left(\frac{1}{1 - F} \right)^{\leftarrow} (t)$$

is regularly varying with index $1/\alpha$. Suppose further that $B^{(k)}(x)$ satisfies

$$(4.2) \quad \lim_{x \rightarrow \infty} \frac{\bar{B}^{(k)}(x)}{\bar{F}(x)} = 1,$$

uniformly in $k = 1, 2, \dots$. This means that given $\delta > 0$, there exists $x_0 = x_0(\delta)$ independent of k such that for $x > x_0$ and all k we have

$$(4.3) \quad 1 - \delta < \frac{\bar{B}^{(k)}(x)}{\bar{F}(x)} \leq 1 + \delta.$$

CONDITION (B). The tails of the distribution of $\sigma^{(k)}$ are always smaller than the tail of F . A convenient way we ensure this is by supposing that there exists $\eta > \alpha$ such that

$$(4.4) \quad c^\vee := \sup_{k \geq 1} E(\sigma^{(k)})^\eta < \infty.$$

CONDITION (C). Assume

$$(4.5) \quad 0 > m(k) := E(\tau_1^{(k)}) - E(\sigma_1^{(k)}) \rightarrow 0,$$

as $k \rightarrow \infty$ and set

$$(4.6) \quad X_i^{(k)} := \frac{\tau_i^{(k)} - \sigma_{i+1}^{(k)}}{b(d(k))}, \quad i \geq 1,$$

where the specification of $d(k)$ is given next.

Definition of $d(k)$. The function $d(k)$ must satisfy Assumption 3 of Section 2. So we need

$$\frac{d(k)m(k)}{b(d(k))} \rightarrow -1,$$

as $k \rightarrow \infty$. The function

$$(4.7) \quad H(t) := \frac{t}{b(t)} \in RV_{1-\frac{1}{\alpha}},$$

and has an asymptotic inverse

$$H^{\leftarrow} \in RV_{\alpha/(\alpha-1)}.$$

The sequence $d(k)$ must satisfy

$$H(d(k)) \sim \frac{1}{|m(k)|}.$$

Therefore, we choose the sequence $\{d(k)\}$ to be any sequence satisfying

$$(4.8) \quad d(k) \sim H^{\leftarrow} \left(\frac{1}{|m(k)|} \right),$$

where H is specified in (4.7).

Theorem 4.1. *Assume Assumptions (A)–(C) hold. Then with $\{X_i^{(k)}, i \geq 1\}$ defined by (4.6) and $\{d(k)\}$ satisfying (4.8) we have in $D[0, \infty)$*

$$(4.9) \quad Y^{(k)}(t) := \sum_{i=1}^{[d(k)t]} X_i^{(k)} = \frac{1}{b(d(k))} \sum_{i=1}^{[d(k)t]} \left(\tau_i^{(k)} - \sigma_{i+1}^{(k)} \right) \Rightarrow Y^{(\infty)}(t),$$

where the limit is a zero mean, α -stable Lévy motion modified to have drift -1.

Furthermore, the sequence of stationary waiting times indexed by k converges in distribution as well

$$(4.10) \quad W^{(k)} = \bigvee_{n=0}^{\infty} \frac{1}{b(d(k))} \sum_{i=1}^n \left(\tau_i^{(k)} - \sigma_{i+1}^{(k)} \right) \Rightarrow W^{(\infty)} = \bigvee_{t=0}^{\infty} Y^{(\infty)}(t).$$

Remark 4.1. It is important to note that the distribution of the maximum $W^{(\infty)}$ of a negative drift α -stable Lévy motion has been computed by Furrer (see [8], [7]) using work of Zolotarev [33]. The limit distribution is a Mittag-Leffler distribution. Thus, the theorem states that a queueing system with heavy tailed service requirements under heavy load, has an equilibrium waiting time distribution which is approximated by the Mittag-Leffler distribution. Specifically, for every $t > 0$,

$$(4.11) \quad P(W^{(k)} \leq t) \rightarrow P(W^{(\infty)} \leq t) = 1 - \sum_{n=0}^{\infty} \frac{(-a)^n}{\Gamma(1 + n(\alpha - 1))} t^{n(\alpha-1)},$$

where $a = (\alpha - 1)/\Gamma(2 - \alpha)$, and for every $\lambda \geq 0$,

$$(4.12) \quad Ee^{-\lambda W^{(k)}} \rightarrow Ee^{-\lambda W^{(\infty)}} = \frac{a}{a + \lambda^{\alpha-1}}.$$

See e.g. (3.20) of Furrer [7].

Proof. Both assertions in the statement of the theorem will be proven if we verify that Conditions (A)–(C) and the definition of $d(k)$ given in (4.8) imply Assumptions 1-4. We begin by showing Assumption 1 is valid with

$$\nu(dx) = \alpha x^{-\alpha-1} dx 1_{(0, \infty)}(x).$$

On the one hand we have for $x > 0$

$$d(k)P[X_1^{(k)} > x] \leq d(k)P[\tau_1^{(k)} > b(d(k))x] \rightarrow x^{-\alpha}$$

and on the other, for any $\delta > 0$

$$\begin{aligned} d(k)P[X_1^{(k)} > x] &\geq d(k)P[\tau_1^{(k)} - \sigma_2^{(k)} > b(d(k))x, \sigma_2^{(k)} \leq \delta b(d(k))] \\ &\geq d(k)P[\tau_1^{(k)} > b(d(k))(x + \delta), \sigma_2^{(k)} \leq \delta b(d(k))] \\ &= d(k)P[\tau_1^{(k)} > b(d(k))(x + \delta)] - d(k)P[\tau_1^{(k)} > b(d(k))(x + \delta), \sigma_2^{(k)} > \delta b(d(k))] \\ &\rightarrow (x + \delta)^{-\alpha} - 0, \end{aligned}$$

where the last 0 results from

$$(4.13) \quad \begin{aligned} \lim_{k \rightarrow \infty} d(k)P[\sigma_2^{(k)} > \delta b(d(k))] &\leq \lim_{k \rightarrow \infty} \frac{d(k)}{b(d(k))^\eta} E(\sigma_2^{(k)})^\eta \\ &\leq \lim_{t \rightarrow \infty} c^\vee \frac{t}{b(t)^\eta} = 0, \end{aligned}$$

since $\eta/\alpha > 1$. Thus, since $\delta > 0$ is arbitrary, we conclude for $x > 0$ that $d(k)P[X_1^{(k)} > x] \rightarrow \nu(x, \infty]$.

For $x < 0$, note

$$\begin{aligned} d(k)P[X_1^{(k)} < x] &= d(k)P[\sigma_2^{(k)} - \tau_1^{(k)} > b(d(k))|x] \\ &\leq d(k)P[\sigma_2^{(k)} > b(d(k))x] \rightarrow 0, \end{aligned}$$

due to (4.13). This verifies that Assumption 1 holds.

To check Assumption 2 is valid, observe that we have after integration by parts

$$\begin{aligned} d(k)E\left((X_1^{(k)})^2 1_{\{|X_1^{(k)}| \leq \epsilon\}}\right) &\leq 2d(k) \int_0^\epsilon xP[|X_1^{(k)}| > x] dx \\ &\leq 2d(k) \int_0^\epsilon xP[\tau_1^{(k)} > b(d(k))x] dx + 2d(k) \int_0^\epsilon xP[\sigma_1^{(k)} > b(d(k))x] dx \\ &=: I + II. \end{aligned}$$

For I we have with x_0 and δ as in Condition (A) (see (4.3))

$$\begin{aligned} I &= 2d(k) \int_0^{x_0/b(d(k))} xP[\tau_1^{(k)} > b(d(k))x] dx + 2d(k) \int_{x_0/b(d(k))}^\epsilon xP[\tau_1^{(k)} > b(d(k))x] dx \\ &\leq d(k) \left(\frac{x_0}{b(d(k))}\right)^2 + 2(1 + \delta)d(k) \int_{x_0/b(d(k))}^\epsilon x\bar{F}(b(d(k))x) dx \\ &\leq \frac{d(k)}{b(d(k))^2} (\text{const}) + 2(1 + \delta) \int_0^\epsilon x d(k) \bar{F}(b(d(k))x) dx \\ &=: Ia + Ib. \end{aligned}$$

Since $t/b^2(t) \rightarrow 0$, we have $Ia \rightarrow 0$ with $k \rightarrow \infty$. By Karamata's theorem ([3, 5, 6, 26]) we have $Ib \rightarrow (1 + \delta)\epsilon^{2-\alpha}$ as $k \rightarrow \infty$ which goes to 0 as $\epsilon \rightarrow 0$. We therefore conclude that

$$\lim_{\epsilon \rightarrow 0} \limsup_{k \rightarrow \infty} I = 0.$$

For II note

$$\begin{aligned} II &\leq 2d(k) \int_0^\epsilon x E(\sigma_2^{(k)})^\eta x^{-\eta} dx / b(d(k))^\eta \\ &\leq (\text{const}) \frac{d(k)}{b(d(k))^\eta} \epsilon^{2-\alpha}. \end{aligned}$$

This goes to 0 as $k \rightarrow \infty$ since $\eta/\alpha > 1$. This completes the verification that Assumption 2 holds.

The reason that Assumption 3 holds is amply clear from Conditions (3) and (4) and the accompanying discussion so we turn to verifying why Assumption 4 holds. Referring to the form of Assumption 4 in (2.3) and using the decomposition

$$(4.14) \quad 1_{\{|X_1^{(k)}| \leq \epsilon\}} = 1_{\{|X_1^{(k)}| \leq \epsilon, \tau_1^{(k)} > \sigma_2^{(k)}\}} + 1_{\{|X_1^{(k)}| \leq \epsilon, \tau_1^{(k)} < \sigma_2^{(k)}\}}$$

we see that

$$\begin{aligned} d(k)E\left(|X_i^{(k)}|1_{[|X_i^{(k)}|>M]}\right) &\leq d(k)E\left(|\frac{\tau_1^{(k)}}{b(d(k))}|1_{[\tau_1^{(k)}>b(d(k))M]}\right) + d(k)E\left(|\frac{\sigma_2^{(k)}}{b(d(k))}|1_{[\sigma_2^{(k)}>b(d(k))M]}\right) \\ &= d(k)\int_M^\infty P\left[\frac{\tau_1^{(k)}}{b(d(k))} > x\right]dx + d(k)MP\left[\frac{\tau_1^{(k)}}{b(d(k))} > M\right] \\ &\quad + d(k)\int_M^\infty P\left[\frac{\sigma_2^{(k)}}{b(d(k))} > x\right]dx + d(k)MP\left[\frac{\sigma_2^{(k)}}{b(d(k))} > M\right] \end{aligned}$$

and again using the definition of x_0 and δ from (4.3) of Condition (A) we have the bound

$$\begin{aligned} &\leq (1+\delta)\int_M^\infty d(k)\bar{F}(b(d(k))x)dx + (1+\delta)d(k)M\bar{F}(b(d(k))M) \\ &\quad + c^\vee \frac{d(k)}{b(d(k))^\eta}\int_M^\infty x^{-\eta}dx + o(1) \end{aligned}$$

using (4.13), and as $k \rightarrow \infty$, this is asymptotic to

$$\sim (1+\delta)\int_M^\infty x^{-\alpha}dx + (1+\delta)MM^{-\alpha} = O(M^{-\alpha+1})$$

which converges to 0 as $M \rightarrow \infty$. This verifies that Assumption 4 holds and completes the proof of Theorem 4.1. \square

A simple circumstance where Conditions (A)–(D) hold is the following. Suppose $\{\tau_i, i \geq 1\}$ are iid non-negative random variables with common distribution F where \bar{F} satisfies the regular variation assumptions of Condition (A). Similarly suppose $\{\sigma_i, i \geq 1\}$ are iid non-negative random variables with common distribution $A(x)$. Both τ_1 and σ_2 are assumed to have finite means with $E(\tau_1) < E(\sigma_2)$. Define

$$X_i^{(k)} = \frac{\tau_i - \theta_k \sigma_{i+1}}{b(d(k))}$$

and assume

$$\theta_k = \frac{E(\tau_1)}{E(\sigma_2)}(1 + \epsilon_k),$$

where $\epsilon_k > 0$, and $\epsilon_k \rightarrow 0$ as $k \rightarrow \infty$. It is straightforward to verify that Conditions (A)–(D) hold.

REFERENCES

- [1] S. Asmussen. *Applied Probability and Queues*. Wiley, Chichester, West Sussex, UK, 1987.
- [2] J. Beran, R. Sherman, W. Willinger, and M.S. Taqqu. Long range dependence in variable-bit-rate video. *IEEE Transactions on Communications*, 43:1566–1579, = 1995.
- [3] N.H. Bingham, C.M. Goldie, and J.L. Teugels. *Regular Variation*. Cambridge University Press, 1987.
- [4] O. Boxma and J.W. Cohen. Heavy traffic analysis for the GI/G/1 queue with heavy = tailed distributions. PNA-R9710, ISSN 1386-3711. Available at dbs.cwi.nl/cwwwi/owa/cwwwi.print_reports?ID=3D449&type=3D1., 1997.
- [5] L. de Haan. *On Regular Variation and its Application to the Weak = Convergence of Sample Extremes*. Mathematisch Centrum Amsterdam, 1970.
- [6] W. Feller. *An Introduction to Probability Theory and its = Applications*, volume 2. Wiley, New York, 2nd edition, 1971.
- [7] H. Furrer. *Risk Theory and Heavy-Tailed Lévy Processes*. PhD thesis, ETH, Zurich, 1997. Diss. ETH No. 12408.
- [8] H. Furrer. Risk processes perturbed by α -stable Lévy motion. *Scand. Actuar. J.*, 1998(1):59–74, 1998.
- [9] J. Michael Harrison. *Brownian Motion and Stochastic Flow Systems*. Wiley, New York, 1985.

- [10] D. Heath, S. Resnick, and G. Samorodnitsky. Patterns of buffer overflow in a class of queues with long = memory in the input stream. *Ann. Applied Probability*, 7:1021–1057, 1997. Available as TR1169.ps.Z at <http://www.orie.cornell.edu/trlist/trlist.html>.
- [11] D. Heath, S. Resnick, and G. Samorodnitsky. Heavy tails and long range dependence in on/off processes and associated fluid models. *Math. of OR*, 23:145–165, 1998.
- [12] D. Heath, S. Resnick, and G. Samorodnitsky. How system performance is affected by the interplay of = averages in a fluid queue with long range dependence induced by heavy tails. Available as TR1204.ps.Z at <http://www.orie.cornell.edu/trlist/trlist.html>; To appear: *Annals = Appl. Probab.*, 1998.
- [13] K. Itô. *Stochastic Processes*, volume 16. Aarhus University Lecture Notes Series, 1969.
- [14] P.R. Jelenković and A.A. Lazar. Subexponential asymptotics of a markov–modulated G/G/1 = queue. Technical report, Columbia University, 1995.
- [15] P.R. Jelenković and A.A. Lazar. Asymptotic results for multiplexing subexponential on–off = sources. Technical report, Columbia University, 1997.
- [16] Predrag R. Jelenković and Aurel A. Lazar. A network multiplexer with multiple time scale and = subexponential arrivals. In *Stochastic Networks (New York, 1995)*, pages 215–235. Springer, New York, 1996.
- [17] O. Kallenberg. *Random Measures*. Akademie-Verlag, Berlin, third edition, 1983.
- [18] J. F. C. Kingman. The single server queue in heavy traffic. *Proc. Cambridge Philos. Soc.*, 57:902–904, 1961.
- [19] J. F. C. Kingman. The heavy traffic approximation in the theory of queues. = (With discussion). In *Proc. Sympos. Congestion Theory (Chapel Hill, N.C., = 1964)*, pages 137–169. Univ. North Carolina Press, Chapel Hill, N.C., 1965.
- [20] W.E. Leland, M.S. Taqqu, W. Willinger, and D.V. Wilson. On the self-similar nature of Ethernet traffic. *ACM/SIGCOMM Computer Communications Review*, 1993. Proceedings of the ACM/SIGCOMM93. To appear.
- [21] W.E. Leland, M.S. Taqqu, W. Willinger, and D.V. Wilson. On the self-similar nature of Ethernet traffic (extended = version). *IEEE/ACM Transactions on Networking*, 2:1–15, 1994.
- [22] Will E. Leland, Murad S. Taqqu, Walter Willinger, and Daniel V. Wilson. Statistical analysis of high time-resolution ethernet Lan = traffic measurements. In *PCmpScSt25*, pages 146–155, 1993.
- [23] J. Neveu. *Processus ponctuels*, pages 249–445. Lecture Notes in = Math., Vol. 598. Springer-Verlag, Berlin, 1977.
- [24] S. Resnick and G. Samorodnitsky. Activity periods of an infinite server queue and performance = of certain heavy tailed fluid queues. Available as TR1201.ps.Z at <http://www.orie.cornell.edu/trlist/trlist.html>, 1998.
- [25] S.I. Resnick. Point processes, regular variation and weak convergence. *Adv. Appl. Prob.*, 18:66–138, 1986.
- [26] S.I. Resnick. *Extreme Values, Regular Variation and Point Processes*. Springer-Verlag, New York, 1987.
- [27] S.I. Resnick. *Adventures in Stochastic Processes*. Birkhäuser, Boston, 1992.
- [28] M. Taqqu, W. Willinger, and R. Sherman. Proof of a fundamental result in self-similar traffic = modeling. *Computer Communications Review*, 27, 1997.
- [29] Ward Whitt. The continuity of queues. *Advances in Appl. Probability*, 6:175–183, 1974.
- [30] Ward Whitt. Heavy traffic limit theorems for queues: a survey. pages 307–350. *Lecture Notes in Econom. and Math. Systems*, = Vol. 98, 1974.
- [31] W. Willinger, M.S. Taqqu, M. Leland, and D. Wilson. Self–similarity in high–speed packet traffic: analysis and modelling of ethernet traffic measurements. *Statistical Science*, 10:67–85, 1995.
- [32] W. Willinger, M.S. Taqqu, M. Leland, and D. Wilson. Self–similarity through high variability: statistical = analysis of ethernet lan traffic at the source level. *Computer Communications Review*, 25:100–113, 1995. Proceedings of the ACM/SIGCOMM’95, Cambridge, MA.
- [33] V. M. Zolotarev. The first passage time of a level and the behaviour at = infinity of a class of processes with independent increments. *Theory Probab. Appl.*, 9:653–661, 1964.

SIDNEY RESNICK AND GENNADY SAMORODNITSKY, SCHOOL OF OPERATIONS RESEARCH AND INDUSTRIAL ENGINEERING, CORNELL UNIVERSITY, ITHACA, NY 14853

E-mail address: sid@orie.cornell.edu gennady@orie.cornell.edu