SELF-SIMILAR COMMUNICATION MODELS AND VERY HEAVY TAILS

SIDNEY RESNICK AND HOLGER ROOTZÉN

Abstract. Several studies of file sizes either being downloaded or stored in the world wide web have commented that tails can be so heavy that not only are variances infinite, but so are means. Motivated by this fact, we study the infinite node Poisson model under the assumption that transmission times are heavy tailed with infinite mean. The model is unstable but we are able to provide growth rates. Self-similar but non-stationary Gaussian process approximations are provided for the number of active sources, cumulative input, buffer content and time to buffer overflow.

1. Introduction

The identification of self-similarity in various types of teletraffic flow rates has created widespread interest in the possible origins and effects of the self-similarity. Willinger et. al. ([32], [24], [25], [35], [26], [36], [37]) discussed self-similarity of packet counts per unit time in LANS and WANS and a parallel discussion of self-similarity of bytes per unit time in WWW traffic was conducted by Crovella et al ([6, 7, 10, 8]). Crovella, Kim and Park ([9]) conducted a large simulation study to assess the causes and effects of self-similarity in situations that involved slowdown nodes, buffers, varying rates and varying tail parameters. Errammilli and Willinger ([13] used experimental queueing analysis to show why classical models without long range dependence would seriously underestimate delays. Resnick and Samorodnitsky ([28]) constructed an example of a single exponential server fed by a long range dependent input which had queue lengths and waiting times which were heavy tailed. Mathematical studies of the connection between on-off inputs with heavy tailed on-periods appeared in [32] and [18], [19] and [20].

Attempts to explain network self-similarity have largely focussed on heavy tailed transmission times of sources sending data to one or more servers. The common assumption is that transmission times have iid random lengths with common distribution F where F has a Pareto or regularly varying tail. We assume

$$1 - F(x) \sim x^{-\alpha} L(x), \quad x \to \infty, \tag{1.1}$$

and $L(x)$ a slowly varying function, or equivalently

$$\lim_{t\to\infty} \frac{\bar{F}(tx)}{\bar{F}(t)} = x^{-\alpha}, \quad x > 0, \tag{1.2}$$

where $\bar{F} = 1 - F(x)$. The usual assumption on α is that $1 < \alpha < 2$. This means the variance is infinite but the mean is finite. The practical reason for this assumption is the extensive traffic measurements of *on* periods reported in [36] where measured values of α were in the interval $(1, 2)$. The theoretical reason for the assumption is that mathematical analysis has been based on

---

AMS 1991 *subject classifications.* Primary ; secondary 60K250.

Key words and phrases. heavy tails, regular variation, Pareto tails, self-similarity, scaling.

Sidney Resnick's research was partially supported by NSF grant DMS-97-04982 and NSA Grant MDA904-98-1-0041 at Cornell University. Portions of Sidney Resnick's work took place while a guest of the Statistics Department of University of North Carolina, Chapel Hill and grateful acknowledgement is made for hospitality. Collaborative work between S. Resnick and H. Rootzén was supported by the Gothenburg Stochastic Centre.

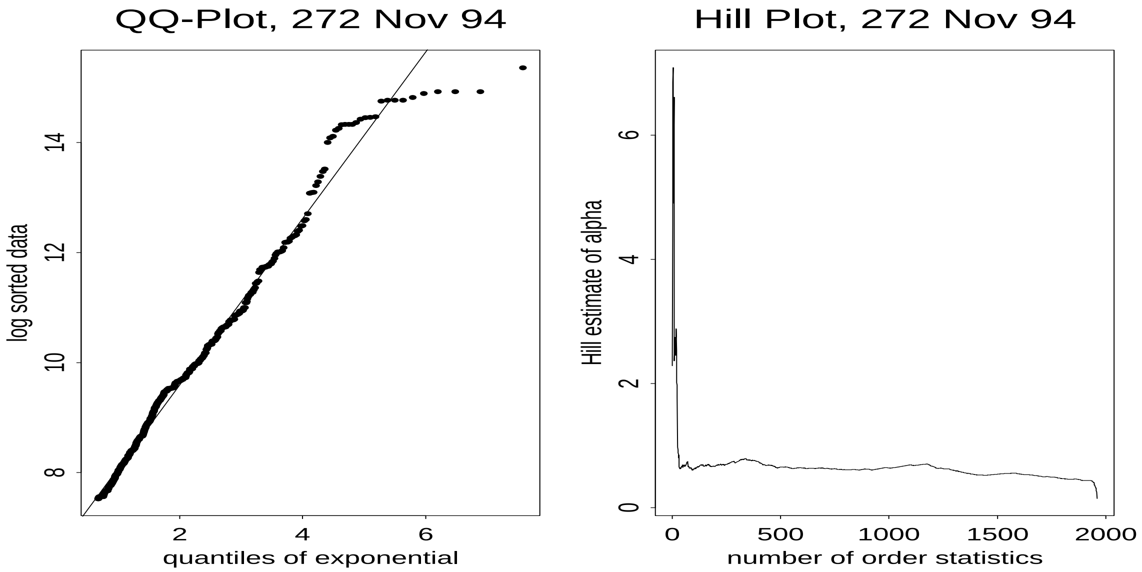

FIGURE 1. QQ and Hill plots of November 1994 file lengths.

renewal theory and without a finite mean, stationary versions of renewal processes do not exist and (uncontrolled) buffer content stochastic processes would not be stable.

Despite the prevalence of this assumption that $1 < \alpha < 2$, it is clear that other assumptions have to be considered. The Boston University study ([6], [7], [11]) suggests self-similarity of web traffic stems from heavy tailed file sizes and reports an overall estimate for a five month measurement period (see [11]) of $\alpha = 1.05$. However, there is considerable month-to-month variation in these estimates and, for instance, the estimate for November 1994 in room 272 places α in the neighborhood of 0.66. Figure 1 gives the QQ and Hill plots ([17, 29, 17, 22, 3]) of the file size data for the month of November in the Boston University study.

Furthermore, studies of sizes of files accessed on various servers by the Calgary study ([1]), report estimates of α from 0.4 to 0.6. So accumulating evidence already exists which suggests values of α outside the range $(1, 2)$ should be considered. Also, as user demands on the web grow and access speeds increase, there may be a drift toward heavier file size distribution tails. However, this is a hypothesis that is currently untested.

This paper focusses on the case $0 < \alpha < 1$. Section 2 reviews a standard infinite node, Poisson based model with heavy tailed transmission times and assumes (1.1) with $\alpha < 1$. A buffer content process is defined and since it will be unstable due to the assumption $\alpha < 1$, we make some comments about the first order content growth and the time to hit high levels. Section 2 develops first order approximations to the number of active nodes, the net input process, the content process and the time to buffer overflow. Section 3 considers a Gaussian approximation to the input process and shows this is self-similar. This approximation is in the spirit of [32]. Sections 4 and 5 give Gaussian process approximations to the content process and time to buffer overflow.

2. An infinite node, Poisson based communication model.

We first review the elements of a communication model used in [19], [20] and [16]. Let $\{\Gamma_k, k \geq 1\}$ be the points of a rate λ homogeneous Poisson process on $\mathbb{R}_+ = [0, \infty)$ so that $\{\Gamma_{k+1} - \Gamma_k, k \geq 1\}$ is a sequence of iid exponentially distributed random variables with parameter λ. We imagine that a communication system has an infinite number of nodes, sometimes called *sources*, and at time Γ_k some node turns on and begins a transmission at unit rate to the server. The length of this transmission is a random variable L_k. We assume $\{L_k, k \geq 1\}$ is iid and independent of $\{\Gamma_k\}$ and

$$P[L_k > x] = \bar{F}(x) = x^{-\alpha}L(x), \quad x \to \infty, \ 0 < \alpha < 1, \tag{2.1}$$

where $L(x)$ is a slowly varying function. We note that

$$M = \sum_{k=1}^{\infty} \epsilon_{(\Gamma_k, L_k)}, \tag{2.2}$$

the counting function on $\mathbb{R}_+ \times (0, \infty]$ corresponding to the points $\{(\Gamma_k, L_k), k \geq 1\}$, is a two dimensional Poisson process on $\mathbb{R}_+ \times (0, \infty]$ with mean measure $\lambda\mathbb{L} \times F$, where $\mathbb{L}$ stands for Lebesgue measure. (Cf. [31].)

The first quantity of interest is $N(t)$, the number of active sources at time t. So

$$\begin{aligned} N(t) =& \sum_{k=1}^{\infty} 1_{[\Gamma_k \leq t < \Gamma_k + L_k]} \\ =& M(\{(\gamma, l) \in \mathbb{R}_+ \times (0, \infty] : \gamma \leq t < \gamma + l\}). \end{aligned}$$

The second expression makes it clear that for each t, $N(t)$ is a Poisson random variable with parameter

$$\lambda\mathbb{L} \times F(\{(\gamma, l) \in \mathbb{R}_+ \times (0, \infty] : \gamma \leq t < \gamma + l\}) = \lambda \int_0^t \bar{F}(t - \gamma) d\gamma = \lambda \int_0^t \bar{F}(s) ds =: m(t). \tag{2.3}$$

Because of Karamata's theorem ([5], [31]),

$$m(t) \sim \frac{\lambda}{1 - \alpha} t\bar{F}(t) = \frac{\lambda}{1 - \alpha} t^{1-\alpha} L(t), \quad t \to \infty. \tag{2.4}$$

During a transmission, the transmitting node is sending data to the server at unit rate. In the fluid queueing terminology, the source pours water into the server at unit rate. The total accumulated input in $[0, t]$ is

$$A(t) := \int_0^t N(s) ds.$$

Assume the server works at constant rate r (or assume in fluid queueing terminology that a bucket leaks at rate r). The content of the buffer at time t, $X(t)$, satisfies the storage equation

$$dX(t) = N(t)dt - r1_{[X(t) > 0]}dt, \tag{2.5}$$

or ([15], [2], [27])

$$\begin{aligned} X(t) &= \bigvee_{s=0}^{t} [A(t) - A(s) - r(t - s)] \\ &= \bigvee_{s=0}^{t} \int_s^t (N(s) - r) \, ds, \end{aligned} \tag{2.6}$$

where we have assumed the initial condition $X(0) = 0$.

From these observations, we can rapidly draw some conclusions about first order behavior. Consider the Laplace transform of $N(T)/m(T)$. Since $N(T)$ is Poisson distributed with parameter $m(T) \to \infty$ we have for $\theta > 0$ as $T \to \infty$

$$E(\exp\{-\theta N(T)/m(T)\}) = \exp\{m(T)\big(e^{-\theta/m(T)} - 1\big)\} \to e^{-\theta} = E\big(e^{-\theta 1}\big).$$

This means that as $T \to \infty$,

$$N(T)/m(T) \xrightarrow{P} 1.$$

Therefore, for each fixed $t > 0$, as $T \to \infty$

$$\frac{N(Tt)}{m(T)} = \frac{N(Tt)}{m(Tt)} \frac{m(Tt)}{m(T)} \xrightarrow{P} t^{1-\alpha},$$

by (2.4). It is easy to extend this convergence ([30]) to weak convergence in $D[0, \infty)$:

$$\frac{N(T\cdot)}{m(T)} \Rightarrow (\cdot)^{1-\alpha}. \tag{2.7}$$

Since integration is a continuous functional from $D[0, \infty) \mapsto D[0, \infty)$, we have for any $t > 0$ as $T \to \infty$

$$\int_0^t \frac{N(Tu)}{m(T)} du \Rightarrow \int_0^t u^{1-\alpha} du = \frac{t^{2-\alpha}}{2-\alpha},$$

that is,

$$\int_0^{Tt} \frac{N(s)}{Tm(T)} ds \Rightarrow \frac{t^{2-\alpha}}{2-\alpha},$$

or

$$\frac{A(T\cdot)}{Tm(T)} \Rightarrow \frac{(\cdot)^{2-\alpha}}{2-\alpha}, \tag{2.8}$$

in $D[0, \infty)$.

Since for any $t > 0$,

$$A(t) - rt \leq X(t) \leq A(t),$$

we get, as $T \to \infty$,

$$\frac{X(T)}{Tm(T)} \xrightarrow{P} \frac{1}{2-\alpha},$$

and hence for any $t > 0$, as $T \to \infty$

$$\frac{X(Tt)}{Tm(T)} \xrightarrow{P} \frac{t^{2-\alpha}}{2-\alpha},$$

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

$$\frac{X(T\cdot)}{Tm(T)} \Rightarrow \frac{(\cdot)^{2-\alpha}}{2-\alpha}. \tag{2.9}$$

The map $x(\cdot) \mapsto \vee_{s \leq (\cdot)} x(s)$ is an almost surely continuous map, and therefore $X^{\vee}(t) := \vee_{s \leq t} X(s)$ has the property

$$\frac{X^{\vee}(T\cdot)}{Tm(T)} \Rightarrow \frac{(\cdot)^{2-\alpha}}{2-\alpha} \quad (T \to \infty). \tag{2.10}$$

If we define

$$\begin{aligned} \tau(L) :=& \inf\{t > 0 : X(t) \geq L\} \\ =& \inf\{t > 0 : X^{\vee}(t) \geq L\} = (X^{\vee})^{\leftarrow}(L), \end{aligned}$$

then (2.10) implies that for $L > 0$, as $T \to \infty$,

$$\frac{\tau(TL)}{V^{\leftarrow}(T)} \Rightarrow ((2-\alpha)L)^{1/(2-\alpha)},$$

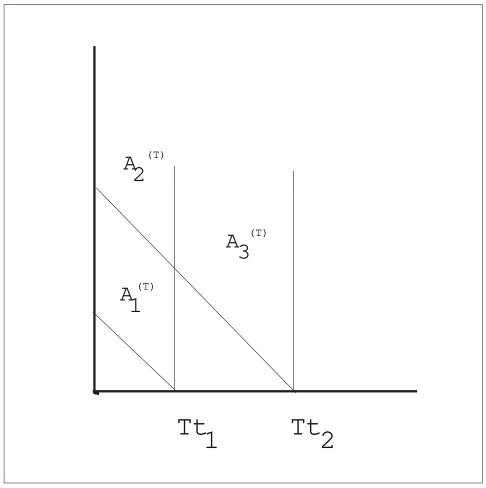

FIGURE 2. Three disjoint regions.

where $V(T) = Tm(T) \sim \lambda T^{2-\alpha} L(T)/(1-\alpha)$ so that the inverse function $V^{\leftarrow}$ is regularly varying with index $1/(2-\alpha)$. So the time necessary for the content to reach a critical high level L is of algebraic order $L^{1/(2-\alpha)}$. This is monotone in α, meaning the smaller the α (that is, the fatter the tail) the quicker a high level is achieved.

3. SELF-SIMILAR GAUSSIAN APPROXIMATIONS.

We begin by considering the family of processes $\{G_T(\cdot), T > 0\}$ defined by

$$G_T(t) = \frac{N(Tt) - m(Tt)}{\sqrt{m(T)}}, \ t \geq 0 \tag{3.1}$$

and showing convergence to a limiting Gaussian process.

First observe that by the central limit theorem for Poisson random variables

$$G_T(1) = \frac{N(T) - m(T)}{\sqrt{m(T)}} \Rightarrow N(0,1)$$

in $\mathbb{R}$, since $m(T) \to \infty$, and hence for any fixed $t \geq 0$

$$G_T(t) = \frac{N(Tt) - m(Tt)}{\sqrt{m(Tt)}} \sqrt{\frac{m(Tt)}{m(T)}} \Rightarrow t^{(2-\alpha)/2} N(0,1) = N(0, t^{2-\alpha}),$$

by (2.4). In fact, $G_T(\cdot)$ is readily seen to converge in the sense of finite dimensional distributions to a limiting Gaussian process. We verify this by illustrating the technique with two time points $t_1 < t_2$.

Write (see Figure 2)

$$\begin{aligned}
A_1^{(T)} =& \{(\gamma, l) \in \mathbb{R}_+ \times (0, \infty] : \gamma \leq Tt_1, Tt_1 < \gamma + l \leq Tt_2\}, \\
A_2^{(T)} =& \{(\gamma, l) \in \mathbb{R}_+ \times (0, \infty] : \gamma \leq Tt_1, Tt_2 < \gamma + l\}, \\
A_3^{(T)} =& \{(\gamma, l) \in \mathbb{R}_+ \times (0, \infty] : Tt_1 < \gamma \leq Tt_2, Tt_2 < \gamma + l\},
\end{aligned}$$

so that $\{A_i^{(T)}, i=1,2,3\}$ are disjoint regions. Therefore, $\{M(A_i^{(T)}), i=1,2,3\}$ are independent Poisson random variables and

$$\begin{aligned} m_1^{(T)} =& E\big(M(A_1^{(T)})\big) = \int_0^{Tt_1} \lambda\left(\bar F(Tt_1-\gamma) - \bar F(Tt_2-\gamma)\right)d\gamma,\\ m_2^{(T)} =& E\big(M(A_2^{(T)})\big) = \int_0^{Tt_1} \lambda \bar F(Tt_2-\gamma)d\gamma,\\ m_3^{(T)} =& E\big(M(A_3^{(T)})\big) = \int_{Tt_1}^{Tt_2} \lambda \bar F(Tt_2-\gamma)d\gamma. \end{aligned}$$

Since $m_i^{(T)} \to \infty$ as $T\to\infty$ for $i=1,2,3$, we have by independence and the central limit theorem for Poisson random variables that

$$\left(\frac{M(A_i^{(T)}) - m_i^{(T)}}{\sqrt{m_i^{(T)}}},\ i=1,2,3\right) \Rightarrow (N_i,\ i=1,2,3)$$

where the limit consists of three independent $N(0,1)$ random variables. Now observe that

$$m_1^{(T)} + m_2^{(T)} = m(Tt_1), \quad N(Tt_1) = M(A_1^{(T)}) + M(A_2^{(T)}),$$

and

$$m_2^{(T)} + m_3^{(T)} = m(tT_2), \quad N(Tt_2) = M(A_2^{(T)}) + M(A_3^{(T)}),$$

and that applying Karamata's theorem, as $T\to\infty$,

$$\begin{aligned} \frac{m_1^{(T)}}{m(T)} \to& t_1^{1-\alpha} - [t_2^{1-\alpha} - (t_2-t_1)^{1-\alpha}] =: m_1^{(\infty)},\\ \frac{m_2^{(T)}}{m(T)} \to& t_2^{1-\alpha} - (t_2-t_1)^{1-\alpha} =: m_2^{(\infty)},\\ \frac{m_3^{(T)}}{m(T)} \to& (t_2-t_1)^{1-\alpha} =: m_3^{(\infty)}. \end{aligned}$$

Therefore in $\mathbb{R}^3$

$$\left(\frac{M(A_i^{(T)}) - m_i^{(T)}}{\sqrt{m(T)}},\ i=1,2,3\right) = \left(\frac{M(A_i^{(T)}) - m_i^{(T)}}{\sqrt{m_i^{(T)}}},\ i=1,2,3\right)\sqrt{\frac{m_i^{(T)}}{m(T)}}$$

$$\Rightarrow (\sqrt{m_i^{(\infty)}}N_i,\ i=1,2,3)$$

and so

$$\begin{aligned} \begin{pmatrix} G_T(t_1)\\ G_T(t_2)\end{pmatrix} =& m(T)^{-1/2}\begin{pmatrix} M(A_1^{(T)}) - E\big(M(A_1^{(T)})\big) + M(A_2^{(T)}) - E\big(M(A_2^{(T)})\big)\\ M(A_2^{(T)}) - E\big(M(A_2^{(T)})\big) + M(A_3^{(T)}) - E\big(M(A_3^{(T)})\big)\end{pmatrix}\\ \Rightarrow& \begin{pmatrix} \sqrt{m_1^{(\infty)}}N_1 + \sqrt{m_2^{(\infty)}}N_2\\ \sqrt{m_2^{(\infty)}}N_2 + \sqrt{m_3^{(\infty)}}N_3\end{pmatrix}. \end{aligned}$$

The covariance of the limit random vector is

$$\mathrm{Var}(\sqrt{m_2^{(\infty)}}N_2) = m_2^{(\infty)} = t_2^{1-\alpha} - (t_2-t_1)^{1-\alpha}.$$

Now let

$$C(s,t) = (s \vee t)^{1-\alpha} - |t-s|^{1-\alpha}, \quad 0 \le s \le t, \tag{3.2}$$

and let $\{G(t), t \ge 0\}$ be a zero mean Gaussian process whose covariance function is $C(s,t)$. We have $G(0) = 0$ and the discussion of Section 3.1 using Billingsley's theorem 15.5 ([4]) will show there is a version with continuous paths. Note that $G(\cdot)$ is not stationary, does not have stationary increments and is not a fractional Brownian motion. As we expect from Lamperti's theorem ([23], [12], we do have that $G(\cdot)$ is self-similar, since for any $c > 0$ and $s < t$

$$C(cs, ct) = c^{1-\alpha} C(s,t),$$

so that in the sense of equality of finite dimensional distributions

$$G(c\cdot) \stackrel{d}{=} c^{(1-\alpha)/2} G(\cdot).$$

We now state function space convergence.

Theorem 1. *Assume $0 < \alpha < 1$, and define $G_T(t)$ by (3.1). Then in $D[0,\infty)$*

$$G_T(\cdot) \Rightarrow G(\cdot),$$

where the limit is a zero mean, continuous path, self-similar Gaussian process with covariance function $C(s,t)$ given by (3.2).

Proof. Convergence of the finite dimensional distributions has already been discussed and it remains to verify tightness. This is discussed in Subsection 3.1. □

If $x_n, x \in D[0,\infty)$ and $x(\cdot)$ is continuous and $x_n \to x$ in the Skorohod metric on $D[0,\infty)$, then this convergence is equivalent to local uniform convergence. See [4]. Further, the functional

$$x(\cdot) \mapsto \int_0^{(\cdot)} x(u)du,$$

is continuous with respect to local uniform convergence. From Theorem 1, we thus quickly obtain the following corollary.

Corollary 1. *Assume $0 < \alpha < 1$. In $C[0,\infty)$, as $T \to \infty$,*

$$\hat{G}_T(\cdot) := \int_0^{(\cdot)} G_T(u)du \Rightarrow \int_0^{(\cdot)} G(u)du =: \hat{G}(\cdot),$$

that is,

$$\hat{G}_T(\cdot) := \frac{A(T\cdot) - \int_0^{T(\cdot)} m(s)ds}{T\sqrt{m(T)}} \Rightarrow \hat{G}(\cdot), \tag{3.3}$$

where $\hat{G}$ is a zero mean, continuous path, self-similar Gaussian process satisfying

$$\hat{G}(c\cdot) \stackrel{d}{=} c^{(3-\alpha)/2} \hat{G}(\cdot), \quad c > 0$$

and having covariance function $\hat{C}(s,t)$ given by

$$\hat{C}(s,t) = \int_{u=0}^{s} \int_{v=0}^{t} C(u,v)dudv, \quad 0 \le s \le t.$$

3.1. **Tightness for the Contents Process.** We now verify that the family of processes $\{G_T(\cdot), T > 0\}$ is tight. The proof uses chaining and bracketing arguments as described in [33]. As in Chapter 2.2 of [33], we define

$$\psi(x) = e^x - 1, \quad x \geq 0,$$

and let $\|X\|$ be the Orlicz norm defined by

$$\|X\| = \inf\{c > 0 : \ E\psi(c^{-1}|X|) \leq 1\}.$$

We begin with a simple lemma.

Lemma 1. *Suppose N is a Poisson random variable with mean m. Then for $m \geq 1$, there exists a constant K not depending on $m \geq 1$ such that*

$$\|N - m\| \leq K\sqrt{m}. \tag{3.4}$$

Proof. Write

$$\|N - m\| \leq \|(N-m)_+\| + \|(N-m)_-\|, \tag{3.5}$$

and we claim

$$\|(N-m)_+\| \leq K_1\sqrt{m}, \tag{3.6}$$

with a similar inequality for the other piece. Now, from the definition of the Orlicz norm, keep in mind that (3.6) requires us to show

$$E\exp\{(N-m)_+)/K_1\sqrt{m}\} \leq 2. \tag{3.7}$$

With $p_+ = \sup_{m\geq 1} P[N \leq m] \in (0,1)$, since by the central limit theorem $P[N \leq m] \to 1/2$, as $m \to \infty$, we have have for a constant $c > 0$

$$\begin{aligned} Ee^{c^{-1}(N-m)_+} \leq & Ee^{c^{-1}(N-m)} + P[N \leq m] \\ = & \exp\{m(e^{c^{-1}} - 1 - c^{-1})\} + p_+, \end{aligned}$$

by the standard formula for the moment generating function for the Poisson distribution. So $E\exp\{c^{-1}(N-m)_+\} \leq 2$ if

$$e^{c^{-1}} - 1 - c^{-1} \leq m^{-1}\log(2 - p_+). \tag{3.8}$$

From $m \geq 1$ and $0 < p_+ < 1$ we get $c^{-1} \leq 2$, and then using the inequality $e^x - 1 - x \leq e^x x^2/2$, $x > 0$, we get

$$e^{c^{-1}} - 1 - c^{-1} \leq (c^{-1})^2 e^2/2,$$

and hence (3.8) holds if

$$(c^{-1})^2 \leq m^{-1}\log(2 - p_+)2e^{-2}.$$

The claim (3.6) follows with

$$K_1 = \frac{1}{\sqrt{2e^2\log(1 + P[N > m])}},$$

and since parallel arguments provide a bound for the second piece in (3.5), the lemma is proven. □

To prove tightness of $\{G_T(\cdot)\}$ in $D[0,\infty)$, it is enough to show that $\{G_T(t); 0 \leq t \leq K_0\}$ is tight in $D[0, K_0]$ for any $K_0 > 0$. For simplicity of notation we suppose $K_0 = 1$. By [4], Theorem 15.5, tightness in $D[0,1]$ occurs if for any $\epsilon > 0$, $\eta > 0$, there are T_0 and $\delta > 0$ such that

$$P[\sup_{|s-t|\leq\delta} |G_T(t) - G_T(s)| > \epsilon] < \eta, \quad \text{for } T \geq T_0. \tag{3.9}$$

For convenience we will use δ's which are dyadic rationals of the form 2^{-i} for i a positive integer and define

$$F_j := \{k2^{-j} : 0 \le k \le 2^j\}.$$

Referring to the probability in (3.9), we write for an integer $\nu \ge i$ to be specified to grow with T, that with $\delta = 2^{-i}$,

$$\begin{aligned}\sup_{|s-t|} |G_T(t) - G_T(s)| \le & \max_{s,t \in F_\nu, |s-t| \le \delta} |G_T(t) - G_T(s)| \\ & + 2 \bigvee_{k=0}^{2^\nu - 1} \bigvee_{u-0}^{2^{-\nu}} |G_T(k2^{-\nu} + u) - G_T(k2^{-\nu})|.\end{aligned}$$

So by standard arguments, (3.9) holds if for given $\epsilon, \eta > 0$, there are $T_0, \delta > 0$ such that

$$\max_{s,t \in F_\nu, |s-t| \le \delta} |G_T(t) - G_T(s)| > \epsilon] < \eta, \quad \text{for } T \ge T_0, \tag{3.10}$$

and if

$$\bigvee_{k=0}^{2^\nu - 1} \bigvee_{u=0}^{2^{-\nu}} |G_T(k2^{-\nu} + u) - G_T(k2^{-\nu})| \xrightarrow{P} 0, \tag{3.11}$$

Before proceeding, we state a lemma. We will continue to denote by K generic constants whose specific value is immaterial. The value of K need not be the same with each usage.

Lemma 2. *Define*

$$m_j = \bigvee_{k=1}^{2^j} |G_T(2^{-j}k) - G_T(2^{-j}(k-1))|. \tag{3.12}$$

Then, for T so large that $m(T2^{-j}) \ge 1$, we have for some constant $K = K(m(T2^{-j}))$

$$\|m_j\| \le Kj\sqrt{\frac{m(T2^{-j})}{m(T)}}.$$

Proof. Refer to Figure 2 with $t_1 = (k-1)/2^j$ and $t_2 = k/2^j$. Then

$$\begin{aligned}\|G_T(k2^{-j}) - G_T((k-1)2^{-j})\| \le & \left\| \frac{M(A_1^{(T)}) - E(M(A_1^{(T)}))}{\sqrt{m(T)}} \right\| + \left\| \frac{M(A_3^{(T)}) - E(M(A_3^{(T)}))}{\sqrt{m(T)}} \right\| \\ \le & K \left(\sqrt{\frac{E(M(A_1^{(T)}))}{m(T)}} + \sqrt{\frac{E(M(A_3^{(T)}))}{m(T)}} \right)\end{aligned}$$

by Lemma 1, provided $E(M(A_1^{(T)})) \wedge E(M(A_3^{(T)})) \ge 1$. However both expectations are bounded by $m(T2^{-j})$ (for $A_1^{(T)}$ this requires a moments reflection from the definition) so we end with a bound of the form (remember the K's can change)

$$\|G_T(k2^{-j}) - G_T((k-1)2^{-j})\| \le K\sqrt{\frac{m(T2^{-j})}{m(T)}}.$$

Then from Lemma 2.2.2, page 96 of [33] we have

$$\| \bigvee_{k=1}^{2^j} |G_T(k2^{-j}) - G_T((k-1)2^{-j})| \| \leq K\psi^{\leftarrow}(2^j) \bigvee_{k=1}^{2^j} \|G_T(k2^{-j}) - G_T((k-1)2^{-j})\|$$

and the result follows from the form of ψ. □

We now verify (3.10) and (3.11). First (3.10) which uses a chaining argument. Begin by defining the integer $\nu = \nu(T)$ by

$$2^{-\nu} \geq \frac{m^{\leftarrow}(1)}{T} > 2^{-\nu-1}, \tag{3.13}$$

so that as $T \to \infty$, $\nu \to \infty$. For a constant K', given $\epsilon, \epsilon' < 1-\alpha$ and η, choose i fixed such that

$$2K' \sum_{j=i}^{\infty} j(2^{-j})^{1-\alpha-\epsilon'} < \frac{\epsilon}{\log(1+\eta^{-1})}, \tag{3.14}$$

and assume T is large enough to make $\nu \geq i$. For $t \in [0,1]$, define $t_j := \sup\{w \in F_j : w \leq t\}$. Then $t \in F_\nu$ implies $T_\nu = t$, and t_{j-1} equals t_j or $t_j - 2^{-j}$, and $|t-s| \leq 2^{-i}$ implies $t_i = s_i$ or $t_i = s_i \pm 2^{-i}$. Write

$$G_T(t) = \sum_{j=i+1}^{\nu} \big(G_T(t_j) - G_T(t_{j-1})\big) + G_T(t_i)$$

with a similar expression for s. It follows that if $t, s \in F_\nu$, $|t-s| \leq 2^{-i}$

$$\begin{aligned}|G_T(t) - G_T(s)| \leq& \sum_{j=i+1}^{\nu} |G_T(t_j) - G_T(t_{j-1})| + \sum_{j=i+1}^{\nu} |G_T(s_j) - G_T(s_{j-1})| \\ &+ |G_T(t_i) - G_T(s_i)| \\ \leq& 2\sum_{j=i+1}^{\nu} m_j + m_i \leq 2\sum_{j=i}^{\nu} m_j,\end{aligned}$$

where m_j is defined in (3.12). This bound does not depend on the specific $s, t \in F_\nu$ with $|t-s| \leq 2^{-i}$, so

$$\xi := \bigvee_{t,s \in F_\nu, |t-s| \leq 2^{-i}} |G_T(t) - G_T(s)| \leq 2\sum_{j=i}^{\nu} m_i$$

and therefore

$$\begin{aligned}\|\xi\| =& \| \bigvee_{t,s \in F_\nu, |t-s| \leq 2^{-i}} |G_T(t) - G_T(s)| \| \leq \sum_{j=i}^{\nu} \|m_i\| \\ \leq& 2K \sum_{j=i}^{\nu} j \sqrt{\frac{m(T2^{-j})}{m(T)}},\end{aligned}$$

by Lemma 2, provided $m(T2^{-j}) \geq 1$ for $j = i \ldots, \nu$; but this is guaranteed by the choice of ν in (3.13). The Potter bounds on a regularly varying function (see [5, 14, 31]) guarantee that for given $\epsilon' > 0$, there exists T_0 such that for $T2^{-\nu} \geq T_0$

$$\frac{m(T2^{-\nu})}{m(T)} \leq 2(2^{-\nu})^{1-\alpha-\epsilon'}.$$

Therefore,

$$\|\xi\| \le 2K' \sum_{j=i}^{\nu} j(2^{-j})^{1-\alpha-\epsilon'} \le \epsilon/\log(1+\eta^{-1})$$

from the choice of i in (3.14). Thus,

$$P[\bigvee_{s,t\in F_\nu} |G_T(t)-G_T(s)| > \epsilon] = P[\xi > \epsilon] = P[\psi(\xi/c) > \psi(\epsilon/c)] \le \frac{E\psi(\xi/c)}{\psi(\epsilon/c)}$$

and choosing $c = \inf\{c : E\psi(\xi/c) \le 1\}$ we get the bound

$$\begin{aligned}
&\le \frac{1}{\psi(\epsilon/\|\xi\|)} = \frac{1}{\exp\{\epsilon/\|\xi\|\}-1} \\
&\le \frac{1}{\exp\{\epsilon\log(1+\eta^{-1})/\epsilon\}-1} \\
&= \eta,
\end{aligned}$$

as required for the proof of (3.10).

We now cope with (3.11). Refer to Figure 2 and let $A_1^{(T)}(k2^{-\nu}, (k+u)2^{-\nu}]$ be the $A_1^{(T)}$ region of the figure with $t_1 = k2^{-\nu}$ and $t_2 = (k+u)2^{-\nu}$ with a similar definition of $A_3^{(T)}(k2^{-\nu}, (k+u)2^{-\nu}]$. Then

$$\begin{aligned}
&\sqrt{m(T)} \sup_{0\le u\le 2^{-\nu}} |G_T(k2^{-\nu}+u) - G_T(k2^{-\nu})| \\
&\qquad = \sup_{0\le u\le 2^{-\nu}} \Big| M(A_1^{(T)}(k2^{-\nu},(k+u)2^{-\nu}] - EM\big(A_1^{(T)}(k2^{-\nu},(k+u)2^{-\nu}]\big) \\
&\qquad\qquad + M(A_3^{(T)}(k2^{-\nu},(k+u)2^{-\nu}] - EM\big(A_3^{(T)}(k2^{-\nu},(k+u)2^{-\nu}]\big)\Big| \\
&\qquad \le M(A_1^{(T)}(k2^{-\nu},(k+1)2^{-\nu}] + M\big((k2^{-\nu},(k+1)2^{-\nu}] \times [0,\infty)\big) \\
&\qquad\qquad + m\big(A_1^{(T)}(k2^{-\nu},(k+1)2^{-\nu}]\big) + T2^{-\nu} \\
&\qquad \le \Big| M(A_1^{(T)}(k2^{-\nu},(k+1)2^{-\nu}]) - E\big(M(A_1^{(T)}(k2^{-\nu},(k+1)2^{-\nu}])\big)\Big| \\
&\qquad\qquad + \Big| M\big((k2^{-\nu},(k+1)2^{-\nu}] \times [0,\infty)\big) - E\big(M\big((k2^{-\nu},(k+1)2^{-\nu}] \times [0,\infty)\big)\big)\Big| \\
&\qquad\qquad + 2m(A_1^{(T)}(k2^{-\nu},(k+1)2^{-\nu}]) + 2T2^{-\nu}.
\end{aligned}$$

Therefore, from Lemma 1, since $m(T2^{-\nu}) \ge 1$, $T2^{-\nu} \ge 1$,

$$\begin{aligned}
\|\sup_{0\le u\le 2^{-\nu}} |G_T(k2^{-\nu}+u) - G_T(k2^{-\nu}|\| &\le K\sqrt{\frac{m(T2^{-\nu})}{m(T)}} + K\sqrt{\frac{2^{-\nu}}{m(T)}} \\
&\quad + 2\frac{m(T2^{-\nu})}{\sqrt{m(T)}} + \frac{2T2^{-\nu}}{\sqrt{m(T)}} \\
&= A + B + C + D.
\end{aligned}$$

Thus, again by Lemma 2.2.2, page 96 of [33],

$$\|\bigvee_{k=0}^{2^\nu-1} \bigvee_{u=0}^{2^{-\nu}} |G_T(k2^{-\nu}+u) - G_T(k2^{-nu})|\| \le K\psi^{\leftarrow}(2^\nu)(A+B+C+D)$$

where $\psi^{\leftarrow}(2^\nu)$ is bounded by a constant times $\log T$. Further, as $T \to \infty$,

$$\begin{aligned}
A \leq & K \log T \sqrt{\frac{m(2m^{\leftarrow}(1))}{m(T)}} \to 0, \\
B \leq & K \log T \sqrt{\frac{2m^{\leftarrow}(1)}{Tm(T)}} \to 0, \\
C \leq & 2 \log T \frac{m(2m^{\leftarrow}(1)T^{-1})}{\sqrt{m(T)}} \to 0, \\
D \leq & 4 \log T \frac{m^{\leftarrow}(1)}{\sqrt{m(T)}} \to 0,
\end{aligned}$$

and (3.11) and hence tightness is proven.

4. Gaussian approximation for the workload process.

We now investigate a Gaussian approximation to the workload process defined by (2.6). In order to do this we must consider the work rate r as a function of T and so we write $r = r_T$ and set for $T > 0$, $t \geq 0$

$$X_T(t) = \bigvee_{s=0}^{Tt} \Big(A(Tt) - A(s) - r_T(Tt - s) \Big) \tag{4.1}$$

$$= \bigvee_{s=0}^{t} \Big(A(Tt) - A(Ts) - r_T T(t - s) \Big). \tag{4.2}$$

We will continue the practice of putting a hat on a function to indicate the integral and thus

$$\hat{m}(t) = \int_0^t m(s) ds.$$

Another notational convention that we will use is that if $f : \mathbb{R}_+ \mapsto \mathbb{R}_+$, we will write

$$f^{\vee}(t) = \bigvee_{s=0}^{t} f(s)$$

for the supmeasure generated by f evaluated on $[0, t]$.

An appropriate way to let r depend on t is to fix $y > 0$ and set

$$r = r_T = m(Ty), \quad r > 0, \ y > 0. \tag{4.3}$$

Note that from (4.2), the release rate for $X_T(\cdot)$ is $Tr_T = Tm(Ty)$ and since (2.8) holds, this form of the release rule is needed to provide some balance to the input.

From (3.3) and (4.3) we can modify (4.2) to get

$$X_T(t) = T\sqrt{m(T)}\hat{G}_T(t) + \hat{m}(Tt) - Tm(Ty)t + \bigvee_{s=0}^{t} [Tm(Ty)s - A(Ts)]$$

and after dividing through by $Tm(T)$ and rearranging terms we get

$$\frac{X_T(t)}{Tm(T)} - \left(\frac{\hat{m}(Tt)}{Tm(T)} - \frac{m(Ty)}{m(T)} t \right) = \frac{\hat{G}_T(t)}{\sqrt{m(T)}} + \bigvee_{s=0}^{t} \xi_T(s) \tag{4.4}$$

where

$$\xi_T(s) := \frac{m(Ty)}{m(T)}s - \frac{A(Ts)}{Tm(T)}. \tag{4.5}$$

Note that $\xi_T(\cdot)$ is continuous, $\xi_T(0) = 0$, and the derivative is

$$\frac{d}{ds}\xi_T(s) =: \xi_T'(s) = \frac{m(Ty)}{m(T)} - \frac{N(Ts)}{m(T)}.$$

Furthermore, from the definition (4.5) and (3.3) we have

$$\xi_T(s) = \frac{m(Ty)}{m(T)}s - \left(\frac{\hat{G}_T(s)}{\sqrt{m(T)}} + \frac{\hat{m}(Ts)}{Tm(T)}\right) \tag{4.6}$$

so that

$$\sqrt{m(T)}\big(\xi_T(s) - b_T(s)\big) = -\hat{G}_T(s) \Rightarrow -\hat{G}(s) \tag{4.7}$$

in $D[0,\infty)$, where we have written

$$b_T(s) = \frac{m(Ty)}{m(T)}s - \frac{\hat{m}(Ts)}{Tm(T)}.$$

To evaluate the supremum in (4.4), the idea is the following. Note that

$$\frac{d}{ds}b_T(s) =: b_T'(s) = \frac{m(Ty)}{m(T)} - \frac{m(Ts)}{m(T)}$$

is strictly positive for $s < y$ and strictly negative for $s > y$ and thus $b_T(\cdot)$ has a unique maximum at y. The process $\xi_T(\cdot)$ is (locally) uniformly close to $b_T(\cdot)$ and hence the place where $\xi_T(\cdot)$ achieves its maximum should be close to y and $\xi_T^\vee(t)$ should be of the order of $b_T(y)$ and $\sqrt{m(T)}(\xi_T^\vee(t) - b_T(y))$ should be near $-\hat{G}(y)$ from (4.7).

We now examine these ideas more carefully.

Set

$$\begin{aligned} s_T = s_T(y) &:= \inf\{s \geq 0 : N(Ts) \geq m(Ty)\} \\ &= \inf\{s \geq 0 : N^\vee(Ts) \geq m(Ty)\} = \frac{(N^\vee)^\leftarrow(m(Ty))}{T} \end{aligned} \tag{4.8}$$

where $(N^\vee)^\leftarrow$ is the left continuous inverse of the monotone function $N^\vee$. Observe that on $(0, s_T)$, $\xi_T(\cdot)$ has a nonnegative derivative and hence is nondecreasing. Therefore, for $t < s_T$, the maximum in (4.4) is achieved at t. From (4.2) it is immediate that $X(Tt) = 0$ for $t \leq s_T$. So when analysing the asymptotic behavior of $X_T(\cdot)$, we will concentrate on the region where $t > s_T$.

It turns out that s_T is approximately equal to y. To see this, observe that from (2.7), we get by applying the maximum functional that

$$\frac{N^\vee(Ts)}{m(T)} \xrightarrow{P} s^{1-\alpha}$$

in $D[0,\infty)$ and since the limit is continuous and increasing we may invert to get

$$\frac{(N^\vee)^\leftarrow(sm(T))}{T} \xrightarrow{P} s^{1/(1-\alpha)}.$$

Since the limit is continuous, the convergence is in the topology of local uniform convergence and hence we may replace s with

$$m(Ty)/m(T) \to y^{1-\alpha}$$

to get

$$s_T(y) = \frac{(N^\vee)^{\leftarrow}(m(Ty))}{T} \to y \tag{4.9}$$

in $D[0,\infty)$.

The process $N(\cdot)$ has only finitely many jumps of size ± 1 in any finite interval and therefore $\xi_T'(\cdot)$ is piecewise constant with jumps of size $\pm\frac{1}{m(T)}$. Therefore $\xi_T(\cdot)$ is piecewise linear. Define a local maximum to be any point $s_T^{(0)}$ such that for some neighborhood Λ of $s_T^{(0)}$

$$\xi_T(s_T^{(0)}) \geq \bigvee_{v\in\Lambda} \xi_T(v),$$

and for some $v \in \Lambda$, $\xi_T(s_T^{(0)}) > \xi_T(v)$. The smallest local maximum is at s_T. At any local maximum $s_T^{(0)}$, there must be a turning point; that is,

$$\xi_T'(s_T^{(0)}+) \leq 0, \quad \xi_T'(s_T^{(0)}-) \geq 0$$

and because $\xi_T'(\cdot)$ changes by jumps of $\pm\frac{1}{m(T)}$,

$$0 \leq \xi_T'(s_T^{(0)}) = \frac{N(Ts_T^{(0)})}{m(T)} - \frac{m(Ty)}{m(T)} \leq \frac{1}{m(T)}$$

or

$$0 \leq N(Ts_T^{(0)}) - m(Ty) \leq 1, \tag{4.10}$$

where we have assumed $N(\cdot)$ is right continuous.

Fix a large $t' > y$ and let $M_T(t')$ be the set of local maxima in $[0,t']$. Since $s_T \xrightarrow{P} y$ from (4.9), we have for $t' > y$ that $P[M_T(t') \neq \emptyset] \to 1$, as $T \to \infty$. We begin with an analysis of the asymptotics of any $s_T^{(0)} \in M_T(t')$ for $t' > y$. We state the results as a proposition.

Proposition 1. *Suppose F satisfies the regular variation condition* (1.1) *with $0 < \alpha < 1$. Fix $t' > y$. If $M_T(t') \neq \emptyset$, let $s_T^{(0)}$ be any local maximum in $M_T(t')$. If $M_T(t') = \emptyset$, set $s_T^{(0)} = t'$.*

(i) **First order behavior:** *As $T \to \infty$,*

$$s_T^{(0)} \xrightarrow{P} y.$$

In fact

$$|s_T^{(0)} - y| \leq \epsilon_1(T,t',y) \xrightarrow{P} 0, \quad T \to \infty, \tag{4.11}$$

where $\epsilon_1(T,t',y)$ is a bound depending only on T,t',y.

(ii) **Second order behavior:** *Assume in addition to* (1.1) *that the following second order condition holds. For some function $\psi : \mathbb{R}_+ \mapsto \mathbb{R}$, we have*

$$\lim_{T\to\infty} \sqrt{m(T)}\left(\frac{m(Ts)}{m(T)} - s^{1-\alpha}\right) = \psi(s), \quad s \geq 0. \tag{4.12}$$

Then as $T \to \infty$,

$$G_T^*(y) := \sqrt{m(T)}\big(s_T^{(0)} - y\big) \Rightarrow -\frac{1}{1-\alpha} y^\alpha G(y) =: G^*(y). \tag{4.13}$$

In fact,

$$\left| \sqrt{m(T)}\big(s_T^{(0)} - y\big) + y^{\alpha} \frac{G_T(y)}{1-\alpha} \right| \le \epsilon_2(T,t',y) \xrightarrow{P} 0 \tag{4.14}$$

as $T \to \infty$, where $\epsilon_2(T,t',y)$ is a bound depending only on T, t', y.

Proof. (i) From (4.10), on $[M(t') \neq \emptyset]$

$$0 \le \sqrt{m(T)} G_T(s_T^{(0)}) + m(Ts_T^{(0)}) - m(Ty) \le 1,$$

so that

$$0 \le \frac{G_T(s_T^{(0)})}{\sqrt{m(T)}} + \frac{m(Ts_T^{(0)})}{m(T)} - \frac{m(Ty)}{m(T)} \le \frac{1}{m(T)}. \tag{4.15}$$

Set

$$v_T(u) = \frac{m(Tu)}{m(T)}, \tag{4.16}$$

so that $v_T(u) \to u^{1-\alpha}$, locally uniformly in $u \ge 0$ as $T \to \infty$ since the index of regular variation is positive ([5, 14]. Also, define

$$B_1 = \frac{G_T(s_T^{(0)})}{\sqrt{m(T)}} + v_T(s_T^{(0)}) - v_T(y),$$

so that

$$0 \le B_1 1_{[M(t') \neq \emptyset]} \le \frac{1}{m(T)}. \tag{4.17}$$

On $[M(t') \neq \emptyset]$

$$v_T(s_T^{(0)}) = B_1 + v_T(y) - \frac{G_T(s_T^{(0)})}{\sqrt{m(T)}}$$

and

$$\begin{aligned} |s_T^{(0)} - y| =& \left| v_T^{\leftarrow}\big(B_1 + v_T(y) - \frac{G_T(s_T^{(0)})}{\sqrt{m(T)}}\big) - y \right| \\ \le& \sup\{|v_T^{\leftarrow}(s) - y| : v_T(y) - \frac{|G_T|^{\vee}(t')}{\sqrt{m(T)}} \le s \le \frac{1}{m(T)} + v_T(y) + \frac{|G_T|^{\vee}(t')}{\sqrt{m(T)}}\} \\ =& : B_2. \end{aligned}$$

On Ω

$$|s_T^{(0)} - y| \le B_2 + (t' - y)1_{[s_T \ge t']}.$$

The right side is a bound which is independent of the particular local maximum and which converges in probability to 0. This follows from the uniform convergence of $v_T(u)$ in (4.16) and $G_T \Rightarrow G$ which implies $|G_T|^{\vee}(t') \Rightarrow |G|^{\vee}(t')$ in $\mathbb{R}$.

(ii) Rewrite (4.12) as

$$v_T(s) = \frac{\psi_T(s)}{\sqrt{m(T)}} + s^{1-\alpha}, \tag{4.18}$$

where $\psi_T \to \psi$ locally uniformly as $T \to \infty$. From (4.15), on $[M(t') \neq \emptyset]$,

$$\sqrt{m(T)}\big(v_T(s_T^{(0)}) - v_T(y)\big) + G_T(y) = \sqrt{m(T)} B_1 + G_T(y) - G_T(s_T^{(0)})$$

or using (4.18), on $[M(t') \neq \emptyset]$

$$\begin{aligned}\sqrt{m(T)}\Big((s_T^{(0)})^{1-\alpha} - y^{1-\alpha}\Big) &+ G_T(y) \\ &= \sqrt{m(T)}B_1 + G_T(y) - G_T(s_T^{(0)}) - \big(\psi_T(s_T^{(0)}) - \psi_T(y)\big) \\ &=: B_3. \end{aligned} \tag{4.19}$$

Note, as in (i),

$$|B_3| 1_{[M(t') \neq \emptyset]} \leq \epsilon_2(T, t', y) \xrightarrow{P} 0.$$

Furthermore,

$$\sqrt{m(T)}\Big((s_T^{(0)})^{1-\alpha} - y^{1-\alpha}\Big) = B_3 - G_T(y)$$

and by the mean value theorem the left side is

$$\sqrt{m(T)}\Big(s_T^{(0)} - y\Big)(1-\alpha)(y^*)^{-\alpha}$$

where y^* is between $s_T^{(0)}$ and y, so

$$\begin{aligned}\sqrt{m(T)}\Big(s_T^{(0)} - y\Big) &+ y^\alpha \frac{G_T(y)}{1-\alpha} \\ &= \big(B_3 + G_T(y)\big(y^\alpha - (y^*)^\alpha\big)\big)/(1-\alpha) \\ &=: B_4\end{aligned}$$

on $[M(t') \neq \emptyset]$. Therefore, on Ω

$$\sqrt{m(T)}\Big(s_T^{(0)} - y\Big) + y^\alpha \frac{G_T(y)}{1-\alpha} = B_4 1_{[M(t')\neq\emptyset]} + \left(\sqrt{m(T)}(t'-y) + y^\alpha \frac{G_T(y)}{1-\alpha}\right) 1_{[M(t')=\emptyset]} =: B_5$$

and

$$|B_5| \leq \epsilon_3(T, t', y) \xrightarrow{P} 0.$$

□

Remark 1. *Remarks on the second order condition* (4.12). The condition (4.12) is equivalent to

$$\lim_{T\to\infty} \frac{s^{\alpha-1}T^{\alpha-1}m(Ts) - T^{\alpha-1}m(T)}{T^{\alpha-1}\sqrt{m(T)}} = s^{\alpha-1}\psi(s).$$

This places it in a standard framework of extended regular variation theory. See [5], [14]. Since the denominator $T^{\alpha-1}\sqrt{m(T)}$ is regularly varying with index

$$\alpha - 1 + \frac{1-\alpha}{2} = -\left(\frac{1-\alpha}{2}\right),$$

there is a second order index

$$\rho := -\left(\frac{1-\alpha}{2}\right) < 0,$$

and $l \in \mathbb{R}_+$ such that

$$\lim_{T\to\infty} T^{\alpha-1}m(T) = l$$

exists and for a constant c

$$l - T^{\alpha-1}m(T) \sim cT^{\alpha-1}\sqrt{m(T)}$$

which is regularly varying with index $-(1-\alpha)/2$. This means that for some other constant k the limit ψ is of the form

$$\psi(s) = ks^{1-\alpha}(s^{\rho} - 1).$$

□

Remark 2. If we consider $s_T^{(0)}$ as a function of y, the results of Proposition 1 could be formulated in the space $D[0,\infty)$.

We are now ready to evaluate the limit law for $X_T(\cdot)$. Fix $t' \geq t \geq y$. The process $\xi_T(\cdot)$ is continuous on $[0,t]$ and hence assumes its maximum somewhere in $[0,t]$. With probability approaching 1, the maximum cannot be at 0, since $\xi_T(0) = 0$ and

$$\begin{aligned} P[\xi_T^{\vee}(t) = 0] =& P[\xi_T(s) \leq 0, \forall 0 \leq s \leq t] \\ \leq& P[\xi_T(y) \leq 0] = P[\frac{A(Ty)}{Tm(T)} \geq \frac{m(Ty)}{m(T)}y]. \end{aligned}$$

Since regular variation implies

$$\frac{m(Ty)}{m(T)}y \to y^{1-\alpha}y = y^{2-\alpha},$$

and from (2.8)

$$\frac{A(Ty)}{Tm(T)} \xrightarrow{P} \frac{y^{2-\alpha}}{2-\alpha},$$

and $2-\alpha > 1$, we have $P[\xi_T^{\vee}(t) = 0] \to 0$, as $T \to \infty$.

If $t > y$, with probability approaching 1 as $T \to \infty$, the maximum in $[0,t]$ cannot be assumed at t since then $\xi_T(t) \geq \xi_T(y)$. However,

$$P[\xi_T(t) \geq \xi_T(y)] = P[\frac{m(Ty)}{m(T)}(t-y) \geq \frac{A(Tt) - A(Ty)}{Tm(T)}].$$

Since (2.8) implies

$$\frac{A(Tt) - A(Ty)}{Tm(T)} \xrightarrow{P} \frac{t^{2-\alpha} - y^{2-\alpha}}{2-\alpha} > y^{1-\alpha}(t-y) = \lim_{T\to\infty} \frac{m(Ty)}{m(T)}(t-y),$$

we have $P[\xi_T(t) \geq \xi_T(y)] \to 0$, as $T \to \infty$.

So apart from events whose probability approach 0 as $T \to \infty$, either the maximum is assumed at an internal point of $[0,t]$ and is hence a local maximum, or if $t = y$, the maximum might be assumed at y or be assumed at an internal point. Let $s_T^{(0)}$ be the point where the maximum is assumed on $[0,t]$. If $s_T^{(0)}$ is internal to $[0,t]$,then it is a local maximum and

$$\begin{aligned} \sqrt{m(T)}\big(\xi_T^{\vee}(t) - b_T(y)\big) =& \sqrt{m(T)}\big(\xi_T(s_T^{(0)}) - b_T(y)\big) \\ =& \sqrt{m(T)}\big(\xi_T(s_T^{(0)}) - b_T(s_T^{(0)})\big) + \sqrt{m(T)}\big(b_T(s_T^{(0)}) - b_T(y)\big) \\ =& -\hat{G}_T(s_T^{(0)}) + \sqrt{m(T)}\big(b_T(s_T^{(0)}) - b_T(y)\big). \end{aligned}$$

while if $t = y$ and $s_T^{(0)} = y$, then

$$\sqrt{m(T)}\Big(\xi_T^{\vee}(t) - b_T(y)\Big) = \sqrt{m(T)}\Big(\xi_T(y) - b_T(y)\Big) = -\hat{G}_T(y).$$

In either case,

$$\begin{aligned}\Big|\sqrt{m(T)}\Big(\xi_T^{\vee}(t)-b_T(y)\Big)+\hat{G}_T(s_T^{(0)})\Big|&\\ \leq&\sqrt{m(T)}|b_T(s_T^{(0)})-b_T(y)|\\ =&\Big|v_T(y)\sqrt{m(T)}(s_T^{(0)}-y)-v_T(y^{\#})\sqrt{m(T)}(s_T^{(0)}-y)\Big|,\end{aligned}$$

where, by the mean value theorem, $y^{\#}$ is between y and $s_T^{(0)}$. Note the second absolute value term converges in probability to 0. Summarizing, we observe that by (4.13)

$$\Big|\sqrt{m(T)}\big(\xi_T^{\vee}(t)-b_T(y)\big)+\hat{G}_T(s_T^{(0)})\Big|\leq\epsilon_4(T,t',y)\xrightarrow{P}0. \tag{4.20}$$

From (4.4), we get

$$\begin{aligned}\frac{X_T(t)}{Tm(T)}-\left(\frac{\hat{m}(Tt)}{Tm(T)}-\frac{m(Ty)}{m(T)}t\right)=&\xi_T^{\vee}(t)+\frac{\hat{G}_T(t)}{\sqrt{m(T)}}\\ =&\frac{\hat{G}_T(t)-\hat{G}_T(s_T^{(0)})}{\sqrt{m(T)}}+\frac{o_p(1)}{\sqrt{m(T)}}+b_T(y)\end{aligned}$$

and therefore, setting

$$c_T(t):=\frac{\hat{m}(Tt)}{Tm(T)}-\frac{m(Ty)}{m(T)}t+\left(\frac{m(Ty)y}{m(T)}-\frac{\hat{m}(Ty)}{Tm(T)}\right) \tag{4.21}$$

we conclude that in $D[y,\infty)$

$$\begin{aligned}\sqrt{m(T)}\left(\frac{X_T(t)}{Tm(T)}-c_T(t)\right)=&\hat{G}_T(t)-\hat{G}_T(s_T^{(0)})+o_p(1)\\ \Rightarrow&\hat{G}(t)-\hat{G}(y).\end{aligned}$$

As a consequence of the second order assumption (4.12), the centering by $c_T(t)$ can be replaced by a function independent of T at the cost of a translation in the limit. To see this, observe that

$$\begin{aligned}c_T(t)&=\left(\frac{\hat{m}(Tt)-\hat{m}(Ty)}{Tm(T)}\right)-\left(\frac{m(Ty)}{m(T)}(t-y)\right)\\ &\rightarrow\frac{t^{2-\alpha}-y^{2-\alpha}}{2-\alpha}-y^{1-\alpha}(t-y)=:c_\infty(t)\end{aligned} \tag{4.22}$$

and

$$\begin{aligned}\sqrt{m(T)}\big(c_T(t)-c_\infty(t)\big)=&\sqrt{m(T)}\left(\frac{\hat{m}(Tt)-\hat{m}(Ty)}{Tm(T)}-\frac{t^{2-\alpha}-y^{2-\alpha}}{2-\alpha}-\left(\frac{m(Ty)}{m(t)}-y^{1-\alpha}\right)(t-y)\right)\\ =&\sqrt{m(T)}\left(\int_y^t\frac{m(Tu)}{m(T)}du-\int_y^t u^{1-\alpha}du\right)-\sqrt{m(T)}\left(\frac{m(Ty)}{m(t)}-y^{1-\alpha}\right)(t-y)\\ =&\int_y^t\sqrt{m(T)}\Big[\frac{m(Tu)}{m(T)}-u^{1-\alpha}\Big]du-\sqrt{m(T)}\left(\frac{m(Ty)}{m(t)}-y^{1-\alpha}\right)(t-y)\\ \rightarrow&\int_y^t\psi(u)du-\psi(y)(t-y).\end{aligned}$$

Therefore

$$\begin{aligned}&\sqrt{m(T)}\left(\frac{X_T(t)}{Tm(T)}-c_\infty(t)\right)\\&=\sqrt{m(T)}\left(\frac{X_T(t)}{Tm(T)}-c_T(t)\right)+\sqrt{m(T)}\,(c_T(t)-c_\infty(t))\\&\Rightarrow\hat{G}(t)-\hat{G}(y)+\int_y^t\psi(u)du-\psi(y)(t-y).\end{aligned}$$

We summarize by stating the following theorem.

Theorem 2. *Suppose $\{X_T(t), t\geq 0\}$ is the contents process of (4.1) and $1-F$ satisfies the first order condition (1.1) and second order condition (4.12). Suppose the work rate r depends on T so that (4.3) holds for a fixed $y>0$. Then for $t\leq s_T(y)$, $X_T(t)=0$ and for $t<y$ we have $X_T(t)\Rightarrow 0$ in $D[0,y)$. Furthermore in $D[y,\infty)$*

$$\sqrt{m(T)}\left(\frac{X_T(\cdot)}{Tm(T)}-c_T(\cdot)\right)\Rightarrow\hat{G}(\cdot)-\hat{G}(y)\tag{4.23}$$

where $c_T(\cdot)$ is given in (4.21) and

$$\sqrt{m(T)}\left(\frac{X_T(t)}{Tm(T)}-c_\infty(t)\right)\Rightarrow\hat{G}(t)-\hat{G}(y)+\int_y^t\psi(u)du-\psi(y)(t-y)\tag{4.24}$$

where $c_\infty(\cdot)$ is defined by (4.22) and ψ is from the second order condition (4.12).

So given a work rate r, define y by $y=m^{\leftarrow}(r)/T$, and then for $t>y$, the content $X_T(t)$ is approximately distributed as

$$Tm(T)\left(\frac{\hat{G}(t)-\hat{G}(y)}{\sqrt{m(T)}}+c_T(t)\right).$$

5. Gaussian approximation for buffer overflow times.

We now apply the results of the previous section to obtain Gaussian approximations to the buffer overflow probabilities. We begin with the following lemma which is a minor variant of Lemma 3.3.3, page 45 and Lemma 3.1.4, page 39 of [34].

Lemma 3. *Suppose $f(\cdot)\in C[0,\infty)$, $x_n(\cdot)\in D[0,\infty)$ and g satisfies $g(0)=0$, $g(\infty)=\infty$, g is differentiable and $g'(x)>0$ on $[0,\infty)$. Then if $\{c_n\}$ is a sequence satisfying $c_n\to\infty$*

$$c_n(x_n-g)\to f,$$

locally uniformly on $[0,\infty)$, implies

$$c_n(x_n^{\leftarrow}(\gamma)-g^{\leftarrow}(\gamma))\to\left(-f\circ g(\gamma)\right)\left(\frac{1}{g'(\gamma)}\right),$$

for each $\gamma>0$ where

$$x_n^{\leftarrow}(\gamma)=\inf\{s\geq 0: x_n(s)\geq\gamma\}.$$

We seek to apply this Lemma to Theorem 2. Define

$$d_\infty(s)=c_\infty(y+s),\quad D_T(s)=X_T(y+s),\quad s\geq 0.\tag{5.1}$$

From Theorem 2 we get in $D[0,\infty)$

$$\sqrt{m(T)}\left(\frac{D_T(s)}{Tm(T)} - d_\infty(s)\right) \Rightarrow H(s),$$

where

$$H(s) := \hat{G}(y+s) - \hat{G}(y) + \int_y^{y+s} \psi(u)du - \psi(y)s.$$

Lemma 3 yields the consequence that for $\gamma > 0$

$$\sqrt{m(T)}\left(\left(\frac{D_T(\cdot)}{Tm(T)}\right)^{\leftarrow}(\gamma) - d_\infty^{\leftarrow}(\gamma)\right) \Rightarrow -H(d_\infty(\gamma))\left(\frac{1}{c'_\infty(\gamma)}\right) := L(\gamma). \tag{5.2}$$

Observe

$$\begin{aligned}\left(\frac{D_T(\cdot)}{Tm(T)}\right)^{\leftarrow}(\gamma) &= \inf\{s \geq 0 : X_T(y+s) \geq Tm(T)\gamma\} \\ &= (\inf\{v \geq 0 : X_T(v) \geq Tm(T)\gamma\} - y)\, 1_{[\vee_{u=0}^{y} X_T(u) \leq Tm(T)\gamma]}.\end{aligned}$$

However, it is unlikely that in $[0,y]$ the process $X_T(\cdot)$ can exceed $Tm(T)\gamma$ since

$$P[\bigvee_{u=0}^{y} X_T(u) \geq Tm(T)\gamma] \leq P[s_T < y, \bigvee u = s_T{}^y X_T(u) \geq Tm(T)\gamma]$$

(since $X_T(u) = 0$ for $u < s_T$)

$$\leq P[y-\epsilon \leq s_T < y, \bigvee_{u=s_T}^{y} X_T(u) \geq Tm(T)\gamma] + o_p(1)$$

(since $s_T \xrightarrow{P} y$)

$$\begin{aligned}&\leq P[y-\epsilon \leq s_T < y, \bigvee_{u=y-\epsilon}^{y} X_T(u) \geq Tm(T)\gamma] \\ &\leq P[A(Ty) - A(T(y-\epsilon)) \geq Tm(T)\gamma] \\ &\to 0,\end{aligned}$$

as $T \to \infty$ for sufficiently small $\epsilon > 0$ by (2.8).

We summarize this discussion.

Theorem 3. *Suppose the assumptions of Theorem 2 hold and set*

$$\tau_T(v) = \inf\{s \geq 0 : X_T(s) \geq v\}$$

for the first time buffer content exceeds v. Then as $T \to \infty$, we have for each $\gamma > 0$

$$\sqrt{m(T)}\,(\tau_T(Tm(T)\gamma) - (y + d_\infty^{\leftarrow}(\gamma)) \Rightarrow L(\gamma),$$

where L is defined in (5.2) and $d_\infty(\gamma)$ is given in (5.1) and (4.22).

6. Concluding Remarks.

It remains to be seen how useful and accurate these Gaussian approximations will be. We are currently examining telecommunication packet flow rate data, some of which do and some of which do not look Gaussian. As remarked in [32] (see also [21]) one has a choice of whether to try to approximate by Gaussian processes or by jump processes. We will be investigating the fit of Gaussian processes to data and also seeking alternate approximations.

References

[1] M. Arlitt and C.L. Williamson. Web server workload characterization: The search for invariants (extended version). In *Proceedings of the ACM Sigmetrics Conference, Philadelphia, Pa*, 1996. Available from {mfa126,carey}@cs.usask.ca.

[2] S. Asmussen. *Applied Probability and Queues*. Wiley, Chichester, West Sussex, UK, 1987.

[3] J. Beirlant, P. Vynckier, and J. Teugels. Tail index estimation, Pareto quantile plots, and regression diagnostics. *J. Amer. Statist. Assoc.*, 91(436):1659–1667, 1996.

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

[5] N.H. Bingham, C.M. Goldie, and J.L. Teugels. *Regular Variation*. Cambridge University Press, 1987.

[6] M. Crovella and A. Bestavros. Explaining world wide web traffic self–similarity. Preprint available as TR-95-015 from {crovella,best}cs.bu.edu, 1995.

[7] M. Crovella and A. Bestavros. Self-similarity in world wide web traffic: evidence and possible causes. In *Proceedings of the 1996 ACM SIGMETRICS International Conference on Measurement and Modeling of Computer Systems*, 1996.

[8] M. Crovella, A. Bestavros, and M. Taqqu. Heavy-tailed probability distributions in the world wide web. Preprint, 1996.

[9] M. Crovella, G. Kim, and K. Park. On the relationship between file sizes, transport protocols, and self-similar network traffic. In *Proceedings of the Fourth International Conference on Network Protocols (ICNP'96)*, pages 171–180, 1996.

[10] M. Crovella, K. Park, and G. Kim. On the effect of traffic self-similarity on network performance. In *Proceedings of the SPIE International Conference on Performance and Control of Network Systems*, 1997.

[11] Cunha, A. Besavros, and M. Crovella. Characteristics of www client–based traces. Preprint available as BU-CS-95-010 from {crovella,best}cs.bu.edu, 1995.

[12] R. Durrett and S. Resnick. Weak convergence with random indices. *Journal of Stochastic Processes and Their Applications*, 5:213–220, 1977.

[13] A. Erramilli, O. Narayan, and W. Willinger. Experimental queueing analysis with long–range dependent packet traffic. *IEEE/ACM Transactions on Network Computing*, 4:209–223, 1996.

[14] J. L. Geluk and L. de Haan. *Regular Variation, Extensions and Tauberian Theorems*, volume 40 of *CWI Tract*. Stichting Mathematisch Centrum Centrum voor Wiskunde en Informatica, Amsterdam, 1987.

[15] J. Michael Harrison. *Brownian Motion and Stochastic Flow Systems*. Wiley, New York, 1985.

[16] 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.

[17] B.M. Hill. A simple general approach to inference about the tail of a distribution. *Ann. Statist.*, 3:1163–1174, 1975.

[18] P.R. Jelenković and A.A. Lazar. Subexponential asymptotics of a markov–modulated G/G/1 queue. Technical report, Columbia University, 1995.

[19] P.R. Jelenković and A.A. Lazar. Multiplexing on/off sources with subexponential on–periods. Technical Report CTR 457-96-23, Columbia University, 1996.

[20] P.R. Jelenković and A.A. Lazar. Asymptotic results for multiplexing subexponential on–off sources. Technical report, Columbia University, 1997.

[21] Takis Konstantopoulos and Si-Jian Lin. Macroscopic models for long-range dependent network traffic. *Queueing Systems Theory Appl.*, 28(1-3):215–243, 1998.

[22] M. Kratz and S. Resnick. The qq–estimator and heavy tails. *Stochastic Models*, 12:699–724, 1996.

[23] J.W. Lamperti. Semi-stable stochastic processes. *Transaction of the American Mathematical Society*, 1962:62–78, 1962.

[24] 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.

[25] 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.

[26] 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.

[27] N.U. Prabhu. *Stochastic Storage Processes: Queues, Insurance risk, Dams and Data Communication.* Springer, New York, 1998.

[28] S. Resnick and G. Samorodnitsky. Performance decay in a single server exponential queueing model with long range dependence. *Operations Research*, 45:235–243, 1997.

[29] S. Resnick and C. Stărică. Smoothing the hill estimator. *Advances in Applied Probability*, 29:271–293, 1997.

[30] S.I. Resnick. Point processes, regular variation and weak convergence. *Adv. Appl. Prob.*, 18:66–138, 1986.

[31] S.I. Resnick. *Extreme Values, Regular Variation and Point Processes.* Springer-Verlag, New York, 1987.

[32] M. Taqqu, W. Willinger, and R. Sherman. Proof of a fundamental result in self-similar traffic modeling. *Computer Communications Review*, 27, 1997.

[33] Aad W. van der Vaart and Jon A. Wellner. *Weak Convergence and Empirical Processes with Applications to Statistics.* Springer-Verlag, New York, 1996.

[34] W. Vervaat. *Success Epochs in Bernoulli Trials (with applications in number theory).* Mathematisch Centrum, Amsterdam, 1972. Mathematical Centre Tracts, No. 42.

[35] 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.

[36] 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.

[37] 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 (extended version). *IEEE/ACM Transactions on Networking*, 1996. To appear.

Sidney Resnick, School of Operations Research and Industrial Engineering, Cornell University, Ithaca, NY 14853

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

Holger Rootzén, Mathematical Institute, Chalmers University, Göteborg, Sweden

E-mail address: `rootzen@math.chalmers.se`