EXTREMAL BEHAVIOR OF HEAVY-TAILED ON-PERIODS IN A SUPERPOSITION OF ON/OFF PROCESSES

ALWIN STEGEMAN,* *University of Groningen*

Abstract

Empirical studies of data traffic in high-speed networks suggest that network traffic exhibits self-similarity and long-range dependence. Cumulative network traffic has been modeled using the so-called ON/OFF model. It was shown that cumulative network traffic can be approximated by either fractional Brownian motion or stable Lévy motion, depending on how many sources are active in the model. In this paper we consider exceedances of a high threshold by the sequence of lengths of ON-periods. If the cumulative network traffic converges to stable Lévy motion, the number of exceedances converges to a Poisson limit. The same holds in the fractional Brownian motion case, provided a very high threshold is used. Finally, we show that the number of exceedances obeys the central limit theorem.

Keywords: Exceedances; extreme value theory; point process; point process of exceedances; Poisson random measure; Poisson process; martingale; stopping time; heavy tails; regular variation; Pareto tails; ON/OFF process; ON/OFF model

AMS 2000 Subject Classification: Primary 60G55
Secondary 60F05; 60F17; 60G70; 60K25

1. Introduction

Recent measurements of high-speed network traffic have shown three characteristic properties: heavy tails, self-similarity and long-range dependence (LRD). This implies that traditional traffic models, based on classical queueing theory with exponential interarrival times, are not appropriate for describing high-speed network traffic (see for example [5] and [12]). Empirical evidence of the presence of self-similarity and LRD in traffic measurements can be found in [1], [2], [10] and [18]. A common explanation for the observed self-similarity and LRD of network traffic is heavy-tailed transmission times. In [1] and [2] evidence is found of heavy tails in file lengths, causing heavy-tailed transmission times.

In this paper we will use the framework of the celebrated ON/OFF model, introduced by Willinger *et al.* [18]. It is used to give a 'physical explanation' for the observed self-similarity and LRD, using the assumption of heavy-tailed transmission lengths. In the ON/OFF model, traffic is generated by M independent ON/OFF sources such as workstations in a big computer lab or hosts in the Internet. An ON/OFF source transmits data at unit rate onto the network if it is ON and remains silent if it is OFF. Every ON/OFF source generates an individual ON/OFF process consisting of independent alternating ON- and OFF-periods. The lengths of the ON-periods are identically distributed and so are the lengths of OFF-periods. For a single ON/OFF source, we introduce the following notation. Let $(X_i)_{i \geq 1}$ and $(Y_i)_{i \geq 1}$ be independent i.i.d.

---

sequences of nonnegative random variables, having continuous distribution functions F_{on} and F_{off}, respectively. The random variables X_i and Y_i are the lengths of ON- and OFF-periods, respectively. We assume that, as $x \to \infty$,

$$\bar{F}_{on}(x) = 1 - F_{on}(x) = x^{-\alpha} L(x) \quad \text{and} \quad \bar{F}_{off}(x) = 1 - F_{off}(x) = o(\bar{F}_{on}(x)), \qquad (1.1)$$

where $\alpha > 1$ and L is slowly varying at infinity. Hence, F_{on} is regularly varying and F_{off} has a lighter tail than F_{on}. Denote

$$\mathrm{E}\,X = \mu_{on}, \qquad \mathrm{E}\,Y = \mu_{off}, \quad \text{and} \quad \mu = \mu_{on} + \mu_{off}.$$

Let

$$(S_n, \ n \geq 0) = \left(D, D + \sum_{i=1}^{n} (X_i + Y_i), \ n \geq 1 \right) \qquad (1.2)$$

be a stationary renewal sequence, where the non-negative delay random variable D is independent of the X- and Y-sequences and has distribution given by

$$\mathrm{P}(D > x) = \frac{1}{\mu} \int_x^{\infty} \mathrm{P}(X + Y > s)\, ds.$$

Define the renewal counting process

$$\xi_t = \sum_{n=0}^{\infty} 1_{[0,t]}(S_n), \qquad t \geq 0, \qquad (1.3)$$

and set $\mu_t = \mathrm{E}\,\xi_t$. We denote the ON/OFF process generated by the source by W_t, where

$$W_t = \begin{cases} 1 & \text{if } t \text{ is in an ON-period,} \\ 0 & \text{if } t \text{ is in an OFF-period.} \end{cases}$$

Heath *et al.* [7] give an explicit construction of D in terms of ON- and OFF-periods. Using this construction, W_t is also defined for $0 \leq t < D$. Heath *et al.* [7] show that the ON/OFF process W is strictly stationary with mean $\mathrm{E}\,W_t = \mu_{on}/\mu$. The main result of [7] yields an asymptotic relation for the autocovariance function $\gamma(k)$ of W if $1 < \alpha < 2$: as $k \to \infty$,

$$\gamma(k) \sim \text{const.}\ k^{-(\alpha-1)}, \qquad k = 0, 1, 2, \ldots. \qquad (1.4)$$

Hence, if $1 < \alpha < 2$, the autocorrelations are not absolutely summable. In this sense, the ON/OFF process W exhibits LRD.

Next, consider a network of M i.i.d. ON/OFF sources. We adapt the notation introduced so far to the mth source by adding a superscript (m), e.g. $W_t^{(m)}$ is the ON/OFF process generated by source m. The total traffic in the network at time t is defined by

$$W_M(t) = \sum_{m=1}^{M} W_t^{(m)}, \qquad t \geq 0.$$

We call W_M the *workload process*. Notice that, since we assume unit rate transmissions, $W_M(t)$ equals the number of active sources at time t. Since the sources are i.i.d., (1.4) yields that the workload process exhibits LRD. The total traffic until time t is then given by

$$W_M^*(t) = \int_0^t \left(\sum_{m=1}^{M} W_u^{(m)} \right) du, \qquad t \geq 0.$$

We call W_M^* the *cumulative workload process*.

For large T, we think of $(W_M^*(Tt),\ t \geq 0)$ as the process on large time scales. Self-similar approximations of $W_M^*(T\cdot)$, for large M and T, have been studied in [11], [17] and [18]. Recall that a self-similar process is invariant in distribution under rescaling both in time and space (see [16, Section 7.1]). In [17] and [18], sequential limits in M and T are taken. Depending on the order of the limits, either fractional Brownian motion or stable Lévy motion is obtained as an approximation. In [11], the case $M = M_T$ is studied, where M_T has an infinite limit in T. Here, it depends on the growth rate of M_T whether fractional Brownian motion or stable Lévy motion is obtained as the limiting process.

Recall that fractional Brownian motion $B_H(\cdot)$ is a mean-zero Gaussian process, which is self-similar with parameter $H \in (0, 1)$ and has LRD in its increment sequence for $H \in (\frac{1}{2}, 1)$. Stable Lévy motion $\Lambda_\alpha(\cdot)$ is self-similar with parameter $1/\alpha$ and has independent increments; $\Lambda_\alpha(t)$ has an α-stable distribution, where $\alpha \in (0, 2]$. For $\alpha < 2$, $\Lambda_\alpha(t)$ has infinite variance. For more properties of fractional Brownian motion and stable Lévy motion, we refer the reader to the monograph [16].

Let $M = M_T$ be a nondecreasing integer-valued function such that

$$M_T \to \infty \quad \text{as } T \to \infty.$$

Denote the quantile function of F_{on} by

$$b(x) = (1/\bar{F}_{\mathrm{on}}(x))^{\leftarrow}, \qquad x > 0. \tag{1.5}$$

For a nondecreasing function U, we write $U^{\leftarrow}$ for the left-continuous inverse of U, i.e.

$$U^{\leftarrow}(y) = \inf\{s : U(s) \geq y\}.$$

Notice that $b(x) = x^{1/\alpha}\widetilde{L}(x)$, for an appropriate slowly varying function $\widetilde{L}$. We introduce the following conditions on M, using the function b defined in (1.5).

Slow growth condition.

$$\lim_{T\to\infty} \frac{b(MT)}{T} = 0 \quad \Longleftrightarrow \quad \lim_{T\to\infty} MT\bar{F}_{\mathrm{on}}(T) = 0.$$

Fast growth condition.

$$\lim_{T\to\infty} \frac{b(MT)}{T} = \infty \quad \Longleftrightarrow \quad \lim_{T\to\infty} MT\bar{F}_{\mathrm{on}}(T) = \infty.$$

The equivalence of the above conditions was shown in [11, Lemma 1].
Define the centered and normalized cumulative workload

$$V_T(t) = \frac{W_{M_T}^*(Tt) - \mathrm{E}\, W_{M_T}^*(Tt)}{d_T},$$

where d_T is strictly increasing in T. The main result of [11] is the following.

Theorem 1.1. *Let* $1 < \alpha < 2$, $B_H(t)$ *be fractional Brownian motion with* $H = \frac{1}{2}(3 - \alpha)$ *and* $\Lambda_\alpha(t)$ *be a totally skewed to the right* α-*stable Lévy motion. Set*

$$\sigma_0 = \frac{\mu_{\mathrm{off}}}{\mu^{1+1/\alpha}} \left(\frac{1-\alpha}{\Gamma(2-\alpha)\cos(\pi\alpha/2)} \right)^{-1/\alpha}, \qquad \sigma^2 = \frac{2\mu_{\mathrm{off}}\Gamma(2-\alpha)}{(\alpha-1)\mu^3\Gamma(4-\alpha)}$$

and

$$(d_T, V) = \begin{cases} (b(M_T T), \sigma_0 \Lambda_\alpha) & \text{if } M_T \text{ satisfies the slow growth condition,} \\ ([T^{3-\alpha} L(T) M_T]^{1/2}, \sigma B_H) & \text{if } M_T \text{ satisfies the fast growth condition.} \end{cases}$$

Then,

$$V_T(t) \xrightarrow{\text{fidi}} V(t) \quad \text{as } T \to \infty, \tag{1.6}$$

where $\xrightarrow{\text{fidi}}$ *stands for convergence of the finite-dimensional distributions.*

In the case of fractional Brownian motion as limit process, the convergence in (1.6) can be strengthened to a functional central limit theorem. We also mention that, for $\alpha > 2$, V_T satisfies the functional central limit theorem with Brownian motion as limit.

The theorem above shows that the growth rate of M is responsible for whether the limiting process is fractional Brownian motion or stable Lévy motion. Since $\alpha < 2$, we have $H \in (\frac{1}{2}, 1)$. In this case, fractional Brownian motion exhibits self-similarity and the corresponding increment sequence enjoys LRD, thus giving one possible explanation of the empirically observed phenomena in teletraffic. Stable Lévy motion is also self-similar, but has independent increments. In contrast to the fractional Brownian motion case, LRD is lost in the limit. Also, fractional Brownian motion is a Gaussian process, while the marginal distributions of stable Lévy motion have heavy tails.

The difference in the dependence structures of the limiting processes can be explained as follows. The cumulative workload $W_M^*(Tt)$ can be decomposed as

$$W_M^*(Tt) = \sum_{m=1}^{M} B_{Tt}^{(m)} 1_{[0, S_0^{(m)})}(Tt) + \sum_{m=1}^{M} \sum_{i=1}^{\xi_{Tt}^{(m)}} X_i^{(m)} - \sum_{m=1}^{M} R_{Tt}^{(m)} 1_{[S_0^{(m)}, \infty)}(Tt), \tag{1.7}$$

where B_{Tt} and R_{Tt} are the contributions of the zeroth renewal interval and the interval $[Tt, S_{\xi_{Tt}})$, respectively. The latter term needs to be subtracted, since the contribution of $[S_{\xi_{Tt}-1}, S_{\xi_{Tt}})$ is contained in the second term.

If M satisfies the slow growth condition, the first and third term in (1.7) vanish in the limit. After centering and normalizing, the second term in (1.7) consists of an asymptotically negligible part (called $A_{22} + A_{23}$ in Section 5.4 of [11]) and the random sum

$$\frac{1}{b(MT)} \sum_{m=1}^{M} \sum_{i=1}^{\xi_{Tt}^{(m)}} \left[X_i^{(m)} - \frac{\mu_{\text{on}}}{\mu} (X_i^{(m)} + Y_i^{(m)}) \right] = \frac{1}{b(MT)} \sum_{m=1}^{M} \sum_{i=1}^{\xi_{Tt}^{(m)}} \left[\frac{\mu_{\text{off}}}{\mu} X_i^{(m)} - \frac{\mu_{\text{on}}}{\mu} Y_i^{(m)} \right].$$

Next, in each of the random sums, the renewal counting process $\xi_{Tt}^{(m)}$ is replaced by its mean μ_{Tt}. This is possible because we can use an Anscombe-type condition, which is satisfied by virtue of the slow growth condition. Notice that, since the renewal process (S_n) is stationary, we have $\mu_{Tt} = Tt/\mu$. In this way, the centered and normalized cumulative workload behaves asymptotically like a sum of $M[Tt/\mu]$ mean-zero random variables with a heavy right tail. The normalization $b(MT) = (MT)^{1/\alpha} \tilde{L}(MT)$ is precisely the one needed for convergence to a totally skewed to the right α-stable distribution. The possibility of replacing $\xi_{Tt}^{(m)}$ by μ_{Tt} also explains why the limiting process has independent increments.

If M satisfies the fast growth condition, all three terms in (1.7) contribute to the limit. Also, it is not possible to replace $\xi_{Tt}^{(m)}$ by μ_{Tt}. The normalization $[T^{3-\alpha} L(T) M]^{1/2}$ is asymptotic

to the standard deviation of $W_M^*(T)$. In this way, the LRD present in the workload process remains intact as T is large and a Gaussian limit is obtained.

In this paper, we will use the framework of the ON/OFF model to study the number of exceedances by the sequence of ON-periods $(X_i^{(m)})$. As in [11], we study the case $M = M_T$ and distinguish between fast and slow growth of M. We use a threshold x_T which has an infinite limit in T. The number of exceedances is counted up to time T, which means that we only consider the completed ON-periods

$$X_1^{(m)}, \ldots, X_{\xi_T^{(m)}-1}^{(m)} \quad \text{and} \quad \min\left(T - S_{\xi_T^{(m)}-1}^{(m)}, X_{\xi_T^{(m)}}^{(m)}\right), \quad \text{for } m = 1, \ldots, M.$$

The total number of exceedances up to time T is given by

$$A_T = \sum_{m=1}^{M} \left[\sum_{i=1}^{\xi_T^{(m)}-1} 1_{[x_T,\infty)}(X_i^{(m)}) + 1_{[x_T,\infty)}\left(\min\left(T - S_{\xi_T^{(m)}-1}^{(m)}, X_{\xi_T^{(m)}}^{(m)}\right)\right) \right].$$

We are interested in the limit distribution of A_T as $T \to \infty$. It is clear that the ON-periods $X_i^{(m)}$ and the counting process $\xi_T^{(m)}$ are heavily dependent. From the definition of $\xi_T^{(m)}$ in (1.3) it follows that

$$X_1^{(m)}, \ldots, X_{\xi_T^{(m)}-1}^{(m)}, \min\left(T - S_{\xi_T^{(m)}-1}^{(m)}, X_{\xi_T^{(m)}}^{(m)}\right) \leq T, \qquad m = 1, \ldots, M.$$

This implies that the threshold x_T must be less than T to obtain a nondegenerate limit for A_T.

If M satisfies the slow growth condition, we obtain a Poisson limit for A_T, for certain thresholds x_T. The intuition behind the proof is as follows. Define

$$A_T^{\text{lo}} = \sum_{m=1}^{M} \sum_{i=1}^{\xi_T^{(m)}-1} 1_{[x_T,\infty)}(X_i^{(m)}) \quad \text{and} \quad A_T^{\text{up}} = \sum_{m=1}^{M} \sum_{i=1}^{\xi_T^{(m)}} 1_{[x_T,\infty)}(X_i^{(m)}). \tag{1.8}$$

Observe that, for all ω,

$$A_T^{\text{lo}} \leq A_T \leq A_T^{\text{up}}. \tag{1.9}$$

As in the proof of Theorem 1.1, with Lévy motion in the limit, it is possible to replace all $\xi_T^{(m)}$ by μ_T. For large T, the distributions of A_T^{lo} and A_T^{up} are approximately binomial with $M[T/\mu]$ trials and success probability $\bar{F}_{\text{on}}(x_T)$. The Poisson approximation to the binomial distribution then guarantees that $A_T^{\text{lo,up}}$ converges in distribution to $\text{Poi}(1/\mu)$ (and by (1.9) also that A_T converges in distribution to $\text{Poi}(1/\mu)$) if

$$\mathbb{E} A_T^{\text{lo,up}} \approx M \frac{T}{\mu} \bar{F}_{\text{on}}(x_T) \to \frac{1}{\mu} \quad \text{as } T \to \infty.$$

This implies that $x_T \sim b(MT)$. Moreover, since the slow growth condition is satisfied, $x_T = o(T)$.

For $1 < \alpha < 2$, there is a clear connection with Theorem 1.1. Let x_n be such that $n\bar{F}_{\text{on}}(x_n) \sim 1$, that is, $x_n \sim b(n)$. Since $\bar{F}_{\text{on}}$ is regularly varying with tail parameter α, as $n \to \infty$,

$$x_n^{-1} \sum_{i=1}^{n} (X_i - \mu_{\text{on}}) \xrightarrow{\text{D}} S_\alpha, \tag{1.10}$$

where S_α is a totally skewed to the right α-stable distribution. Also, as $n \to \infty$,

$$x_n^{-1} \max(X_1, \ldots, X_n) \xrightarrow{D} \Phi_\alpha, \tag{1.11}$$

where $P(\Phi_\alpha \leq y) = \exp\{-y^{-\alpha}\}$ is the Fréchet distribution. Finally, due to the Poisson limit theorem, the definition of x_n guarantees that, as $n \to \infty$,

$$\sum_{i=1}^{n} 1_{[x_n, \infty)}(X_i) \xrightarrow{D} \mathrm{Poi}(1). \tag{1.12}$$

For $\alpha \geq 2$, (1.11) and (1.12) still hold, but (1.10) translates into the central limit theorem where the normalization is $(n \, \mathrm{var}(X))^{1/2}$ (if $\alpha = 2$ and $\mathrm{E}\, X^2 = \infty$ a different normalization has to be used).

If M satisfies the fast growth condition the situation is more complicated. In order to apply the Poisson limit theorem, we must have $\mathrm{E}\, A_T \sim$ const. Since $\xi_T^{(m)}$ cannot be replaced by μ_T, there is no straightforward method of calculating $\mathrm{E}\, A_T$. However, notice that $\xi_T^{(m)}$ is a stopping time with respect to the filtration

$$\mathcal{F}_n^{(m)} = \sigma(D^{(m)}, X_1^{(m)}, Y_1^{(m)}, \ldots, X_n^{(m)}, Y_n^{(m)}), \qquad n \geq 1.$$

Therefore, we also include the ON-periods $X_{\xi_T^{(m)}}^{(m)}$, which allows us to use Wald's identity for the expectation of random sums (see [15, Section 1.8.1] or [6, Theorem 5.3 in Chapter I]). We consider the following number of exceedances:

$$\widehat{A}_T = \sum_{m=1}^{M} \sum_{i=1}^{\xi_T^{(m)}} 1_{[x_T, T)}(X_i^{(m)}).$$

We need the restriction $X_i^{(m)} < T$ to obtain a nondegenerate limit for $\widehat{A}_T$. Using Wald's identity, we have

$$\begin{aligned}
\mathrm{E}\, \widehat{A}_T &= M \frac{T}{\mu}[\bar{F}_{\mathrm{on}}(x_T) - \bar{F}_{\mathrm{on}}(T)] \\
&= M \frac{T}{\mu} \bar{F}_{\mathrm{on}}(T)\left[\frac{\bar{F}_{\mathrm{on}}(x_T)}{\bar{F}_{\mathrm{on}}(T)} - 1\right].
\end{aligned} \tag{1.13}$$

Since the fast growth condition holds, $MT\bar{F}_{\mathrm{on}}(T) \to \infty$. To have $\mathrm{E}\, \widehat{A}_T \sim$ const., we must choose x_T such that $\bar{F}_{\mathrm{on}}(x_T) \sim \bar{F}_{\mathrm{on}}(T)$. Using the monotonicity of $\bar{F}_{\mathrm{on}}$ and a property of regularly varying functions (see Proposition B.1), we see that $x_T \sim T$ must hold. Moreover, since $x_T < T$, we can write $x_T = T - a_T$, where a_T is a positive sequence satisfying $a_T = o(T)$. We obtain a Poisson limit for $\widehat{A}_T$ by balancing M and a_T such that $\mathrm{E}\, \widehat{A}_T \sim$ const.

It appears that, if M satisfies the fast growth condition, the space where exceedances can occur must be chosen very small in order to obtain a nondegenerate limit. Since the ON-periods cannot be larger than T, there will be more and more of them near T as M increases. The faster M grows, the smaller the region $[T - a_T, T)$ has to be to ensure a nondegenerate limiting number of exceedances. A large M must be compensated by a small a_T. Thus, for M very fast, Poisson convergence of $\widehat{A}_T$ is due to the number of ON-periods with lengths which are practically indistinguishable from T.

This paper is organized as follows. In Section 2, we consider the case when M satisfies the slow growth condition. We use the theory of point processes to show that the number of exceedances of the threshold $x_T \sim b(MT) = o(T)$ converges to a Poisson random measure. We apply this result to show that the number of exceedances up to time Tt converges weakly to a homogeneous Poisson process in $(\mathbb{D}[0, \infty), J_1)$ as $T \to \infty$.

In Section 3, we consider the case when M satisfies the fast growth condition, using point processes as in Section 2. We obtain convergence to a Poisson random measure for a threshold $x_T = T - a_T$ with $a_T = o(T)$ under a balancing condition on M and a_T.

In Section 4, we show that

$$\sum_{m=1}^{M} \sum_{i=1}^{\xi_T^{(m)}} 1_{[x_T, T]}(X_i^{(m)})$$

obeys the central limit theorem under the condition that $MT[\bar{F}_{\text{on}}(x_T) - \bar{F}_{\text{on}}(T)] \to \infty$. This ensures an increasing number of exceedances. Here, M does not have to satisfy a particular growth condition.

2. Slow growth of M

Here we consider the case where M is either fixed or $M = M_T \to \infty$ is a nondecreasing integer-valued function satisfying the slow growth condition

$$\lim_{T \to \infty} \frac{b(MT)}{T} = 0 \quad \Longleftrightarrow \quad \lim_{T \to \infty} MT\bar{F}_{\text{on}}(T) = 0.$$

We derive the limit distribution of

$$A_T = \sum_{m=1}^{M} \left[\sum_{i=1}^{\xi_T^{(m)}-1} 1_{[x_T, \infty)}(X_i^{(m)}) + 1_{[x_T, \infty)}\left(\min\left(T - S_{\xi_T^{(m)}-1}^{(m)}, X_{\xi_T^{(m)}}^{(m)} \right) \right) \right],$$

which is the total number of exceedances, up to time T, by completed and running ON-periods of all M sources. The threshold x_T is such that, as $T \to \infty$,

$$MT\bar{F}_{\text{on}}(x_T) \sim 1. \tag{2.1}$$

Notice that $x_T \sim b(MT)$, where b is defined in (1.5). Since the slow growth condition holds, we have $x_T = o(T)$.

In Section 2.1, we use the theory of point processes and obtain convergence to a Poisson random measure. The convergence is illustrated by means of simulations. Section 2.2 considers the number of exceedances up to time Tt and shows weak convergence to the homogeneous Poisson process.

2.1. Convergence to a Poisson random measure

In this section, we use the theory of point processes to derive the limit distribution of the number of exceedances of the sequences $(X_i^{(m)})$. Recall the definition of the stationary renewal sequence $(S_n^{(m)})$ from (1.2). For $T > 0$ we define

$$N_T^{\text{lo}} = \sum_{m=1}^{M} \sum_{i=1}^{\infty} \varepsilon_{S_i^{(m)}/T, X_i^{(m)}/x_T} \quad \text{and} \quad N_T^{\text{up}} = \sum_{m=1}^{M} \sum_{i=1}^{\infty} \varepsilon_{S_{i-1}^{(m)}/T, X_i^{(m)}/x_T}, \tag{2.2}$$

which are point processes on the state space $E := [0, \infty) \times (0, \infty]$. For $(u, v) \in E$ and $C \times D \subset E$, the Dirac measure ε is defined by

$$\varepsilon_{(u,v)}(C \times D) = \begin{cases} 1 & \text{if } (u, v) \in C \times D, \\ 0 & \text{if } (u, v) \notin C \times D. \end{cases}$$

For fixed T, the point process N_T^{lo} resembles the point process of exceedances (see [4, Example 5.1.3]). The difference is that N_T^{lo} also takes into account the 'real' times, in the form of the renewal sequences $(S_n^{(m)})$, at which the exceedances occur. An additional complication is that the $X_i^{(m)}$ and $S_n^{(m)}$ are heavily dependent. Let A_T^{lo} and A_T^{up} be defined by (1.8). Notice that

$$N_T^{lo}([0, 1) \times [1, \infty)) = \sum_{m=1}^{M} \sum_{i=1}^{\xi_T^{(m)}-1} 1_{[x_T, \infty)}(X_i^{(m)}) = A_T^{lo},$$

$$N_T^{up}([0, 1) \times [1, \infty)) = \sum_{m=1}^{M} \sum_{i=1}^{\xi_T^{(m)}} 1_{[x_T, \infty)}(X_i^{(m)}) = A_T^{up}.$$

Let $M_p(E)$ denote the space of all point measures defined on E, equipped with the vague topology (see [14, Section 3.4]). Our goal is to show that both N_T^{lo} and N_T^{up} converge in distribution to a Poisson random measure (PRM) in $M_p(E)$. Recall that for a PRM N on E with mean measure ν, $N(A)$ has a Poisson distribution with intensity $\nu(A)$ for any Borel set $A \subset E$, and $N(A_1)$ and $N(A_2)$ are independent if A_1 and A_2 are disjoint.

Let $\mathbb{L}$ denote Lebesgue measure. Our main result is the following.

Theorem 2.1. *Suppose that M satisfies the slow growth condition and x_T satisfies (2.1). Then, as $T \to \infty$,*

$$N_T^{lo} \xrightarrow{D} N \quad \text{and} \quad N_T^{up} \xrightarrow{D} N \quad \text{in } M_p(E),$$

where N is PRM with mean measure $\mu^{-1}\mathbb{L} \times \nu$ and, for $0 < a < b$,

$$\nu[a, b) = \int_a^b \alpha t^{-(\alpha+1)} \, dt.$$

Notice that, since $\mu^{-1}(\mathbb{L} \times \nu)([0, 1) \times [1, \infty)) = \mu^{-1}$, Theorem 2.1 implies that, as $T \to \infty$,

$$A_T^{lo} \xrightarrow{D} \text{Poi}(\mu^{-1}) \quad \text{and} \quad A_T^{up} \xrightarrow{D} \text{Poi}(\mu^{-1}).$$

By (1.9), $A_T \xrightarrow{D} \text{Poi}(\mu^{-1})$ also.

Proof of Theorem 2.1. We start with proving convergence of N_T^{lo}. We use [9] (see Theorem A.1 below) and consider the class $\mathcal{A}$ of sets

$$A = \bigcup_{j=1}^{k} [a_j, b_j) \times (c_j, d_j],$$

for $k \geq 1$, $[a_j, b_j) \times (c_j, d_j] \subset E$, $j = 1, \ldots, k$. We may assume that the sets $[a_j, b_j) \times (c_j, d_j]$ are mutually disjoint. We first show that

$$P(N_T^{lo}(A) = 0) \to P(N(A) = 0).$$

Notice that

$$N_T^{\text{lo}}(A) = \sum_{j=1}^{k} \sum_{m=1}^{M} \sum_{i=\xi_{Ta_j}^{(m)}}^{\xi_{Tb_j}^{(m)}-1} 1_{(c_j x_T, d_j x_T]}(X_i^{(m)}).$$

Let $\epsilon_T \to 0$ satisfy (C.1) below. Define the event

$$B_T = \{|\xi_{Tt}^{(m)} - \mu_{Tt}| \le \epsilon_T \mu_{Tt}, t = a_j, b_j, j = 1, \ldots, k, m = 1, \ldots, M\}.$$

From Lemma C.1 it follows that $P(B_T^c) \to 0$. Define

$$N_{\mu_T^{\pm}}^{\text{lo}}(A) = \sum_{j=1}^{k} \sum_{m=1}^{M} \sum_{i=\lfloor (1\mp\epsilon_T)\mu_{Ta_j}\rfloor}^{\lfloor (1\pm\epsilon_T)\mu_{Tb_j}\rfloor-1} 1_{(c_j x_T, d_j x_T]}(X_i^{(m)}),$$

where $\lfloor\cdot\rfloor$ denotes the integer part function. We have the inequalities

$$P(N_T^{\text{lo}}(A) = 0, B_T) \le P\Big(N_{\mu_T^{-}}^{\text{lo}}(A) = 0\Big),$$

and

$$P(N_T^{\text{lo}}(A) = 0, B_T) \ge P\Big(N_{\mu_T^{+}}^{\text{lo}}(A) = 0, B_T\Big) \ge P\Big(N_{\mu_T^{+}}^{\text{lo}}(A) = 0\Big) - P(B_T^c).$$

We will show that

$$P\Big(N_{\mu_T^{+}}^{\text{lo}}(A) = 0\Big) \to P(N(A) = 0).$$

The proof for μ_T^{-} is analogous. First, we assume that the sets $[a_j, b_j)$ are mutually disjoint. This assumption is relaxed later on. Moreover, assume that

$$0 \le a_1 < b_1 \le a_2 < b_2 \le \cdots \le a_k < b_k < \infty.$$

Then, for large T, $\lfloor (1+\epsilon_T)\mu_{Tb_j}\rfloor - 1 < \lfloor (1-\epsilon_T)\mu_{Ta_{j+1}}\rfloor$, for $j = 1, \ldots, k-1$. Using (2.1), we have

$$P\Big(N_{\mu_T^{+}}^{\text{lo}}(A) = 0\Big)$$

$$= \prod_{j=1}^{k}[P(X_i/x_T \notin (c_j, d_j], i = \lfloor (1-\epsilon_T)\mu_{Ta_j}\rfloor, \ldots, \lfloor (1+\epsilon_T)\mu_{Tb_j}\rfloor - 1)]^M$$

$$\sim \prod_{j=1}^{k}[1 - \bar{F}_{\text{on}}(c_j x_T) + \bar{F}_{\text{on}}(d_j x_T)]^{MT(b_j-a_j)\mu^{-1}}$$

$$\sim \prod_{j=1}^{k}\Big[1 - \frac{c_j^{-\alpha}}{MT} + \frac{d_j^{-\alpha}}{MT}\Big]^{MT(b_j-a_j)\mu^{-1}}$$

$$\sim \prod_{j=1}^{k}\Big[1 - \frac{(b_j - a_j)(c_j^{-\alpha} - d_j^{-\alpha})}{\mu MT}\Big]^{MT}$$

$$\sim \exp\Big\{-\mu^{-1}\sum_{j=1}^{k}(b_j - a_j)(c_j^{-\alpha} - d_j^{-\alpha})\Big\}$$

$$= P(N(A) = 0).$$

The final step consists of dropping the assumption that the $[a_j, b_j)$ are mutually disjoint. We only consider the case $k = 2$. The general case is completely analogous. Let

$$A = \bigcup_{j=1}^{2} [a_j, b_j) \times (c_j, d_j],$$

where $a_1 < a_2 \leq b_1 < b_2$ and, for example, $c_1 < d_1 \leq c_2 < d_2$. Then

$$P\left(N_{\mu_T^{-}}^{\text{lo}}(A) = 0\right)$$

$$= \Big[P(X_i/x_T \notin (c_1, d_1], \; i = \lfloor (1 - \epsilon_T)\mu_{Ta_1} \rfloor, \ldots, \lfloor (1 - \epsilon_T)\mu_{Ta_2} \rfloor - 1,$$

$$X_i/x_T \notin (c_1, d_1] \cup (c_2, d_2], \; i = \lfloor (1 - \epsilon_T)\mu_{Ta_2} \rfloor, \ldots, \lfloor (1 + \epsilon_T)\mu_{Tb_1} \rfloor - 1,$$

$$X_i/x_T \notin (c_2, d_2], \; i = \lfloor (1 + \epsilon_T)\mu_{Tb_1} \rfloor, \ldots, \lfloor (1 + \epsilon_T)\mu_{Tb_2} \rfloor - 1) \Big]^{M}$$

$$\sim \Big[(1 - \bar{F}_{\text{on}}(c_1 x_T) + \bar{F}_{\text{on}}(d_1 x_T))^{T(a_2 - a_1)\mu^{-1}}$$

$$\cdot (1 - \bar{F}_{\text{on}}(c_1 x_T) + \bar{F}_{\text{on}}(d_1 x_T))(1 - \bar{F}_{\text{on}}(c_2 x_T) + \bar{F}_{\text{on}}(d_2 x_T))^{T(b_1 - a_2)\mu^{-1}}$$

$$\cdot (1 - \bar{F}_{\text{on}}(c_2 x_T) + \bar{F}_{\text{on}}(d_2 x_T))^{T(b_2 - b_1)\mu^{-1}} \Big]^{M}$$

$$= \prod_{j=1}^{2} [1 - \bar{F}_{\text{on}}(c_j x_T) + \bar{F}_{\text{on}}(d_j x_T)]^{MT(b_j - a_j)\mu^{-1}},$$

which converges to $P(N(A) = 0)$ as before.

It remains to show that, for $I = [a, b) \times (c, d] \subset E$, we have

$$\limsup_{T \to \infty} P(N_T^{\text{lo}}(I) > 1) \leq P(N(I) > 1).$$

Since $I \in \mathcal{A}$, by the first part of the proof,

$$\limsup_{T \to \infty} P(N_T^{\text{lo}}(I) > 1) \leq 1 - \liminf_{T \to \infty} P(N_T^{\text{lo}}(I) = 0) - \liminf_{T \to \infty} P(N_T^{\text{lo}}(I) = 1)$$

$$= 1 - P(N(I) = 0) - \liminf_{T \to \infty} P(N_T^{\text{lo}}(I) = 1).$$

Thus it suffices to show that

$$\liminf_{T \to \infty} P(N_T^{\text{lo}}(I) = 1) \geq P(N(I) = 1).$$

As before, let $\epsilon_T \to 0$ satisfy (C.1) and define

$$B_T = \{|\xi_{Tt}^{(m)} - \mu_{Tt}| \leq \epsilon_T \mu_{Tt}, \; t = a, b, m = 1, \ldots, M\}.$$

Write (slightly abusing notation)

$$N_{\mu_T^{-}}^{\text{lo}}(I) = \sum_{m=1}^{M} \sum_{i=\lfloor (1+\epsilon_T)\mu_{Ta} \rfloor}^{\lfloor (1-\epsilon_T)\mu_{Tb} \rfloor - 1} 1_{(cx_T, dx_T]}(X_i^{(m)}),$$

$$N_{\epsilon_T}^{a} = \sum_{m=1}^{M} \sum_{i=\lfloor (1-\epsilon_T)\mu_{Ta} \rfloor}^{\lfloor (1+\epsilon_T)\mu_{Ta} \rfloor - 1} 1_{(cx_T, dx_T]}(X_i^{(m)})$$

and

$$N^b_{\epsilon_T} = \sum_{m=1}^{M} \sum_{i=\lfloor (1-\epsilon_T)\mu_{Tb}\rfloor}^{\lfloor (1+\epsilon_T)\mu_{Tb}\rfloor - 1} 1_{(cx_T, dx_T]}(X^{(m)}_i).$$

Since $N^{lo}_{\mu_T^-}(I)$, $N^a_{\epsilon_T}$ and $N^b_{\epsilon_T}$ are mutually independent, we have

$$P(N^{lo}_T(I) = 1) \geq P(N^{lo}_T(I) = 1, B_T)$$

$$\geq P\left(N^{lo}_{\mu_T^-}(I) = 1, N^a_{\epsilon_T} = 0, N^b_{\epsilon_T} = 0, B_T\right)$$

$$\geq P\left(N^{lo}_{\mu_T^-}(I) = 1, N^a_{\epsilon_T} = 0, N^b_{\epsilon_T} = 0\right) - P(B^c_T)$$

$$= P\left(N^{lo}_{\mu_T^-}(I) = 1\right)P(N^a_{\epsilon_T} = 0)P(N^b_{\epsilon_T} = 0) - P(B^c_T).$$

By Lemma C.1, $P(B^c_T) \to 0$. Using (2.1), we observe that as $T \to \infty$

$$P(N^a_{\epsilon_T} = 0) = [1 - \bar{F}_{on}(cx_T) + \bar{F}_{on}(dx_T)]^{M(2\epsilon_T\mu_{Ta}-1)}$$

$$\sim \left[1 - \frac{c^{-\alpha}}{MT} + \frac{d^{-\alpha}}{MT}\right]^{M(2\epsilon_T Ta\mu^{-1}-1)}$$

$$\sim \left[1 - \frac{2a(c^{-\alpha} - d^{-\alpha})}{\mu MT}\right]^{\epsilon_T MT} \to 1.$$

Analogously, $P(N^b_{\epsilon_T} = 0) \to 1$. Set

$$r_T = M(\lfloor (1 - \epsilon_T)\mu_{Tb}\rfloor - 2 - \lfloor (1 + \epsilon_T)\mu_{Ta}\rfloor).$$

Finally, by (2.1) and the first part of the proof,

$$P\left(N^{lo}_{\mu_T^-}(I) = 1\right) = r_T[\bar{F}_{on}(cx_T) - \bar{F}_{on}(dx_T)][1 - \bar{F}_{on}(cx_T) + \bar{F}_{on}(dx_T)]^{r_T-1}$$

$$\sim \frac{MT(b-a)}{\mu}\frac{(c^{-\alpha} - d^{-\alpha})}{MT}P(N(I) = 0)$$

$$= \frac{(b-a)(c^{-\alpha} - d^{-\alpha})}{\mu}\exp\{\mu^{-1}(b-a)(c^{-\alpha} - d^{-\alpha})\}$$

$$= P(N(I) = 1).$$

This completes the proof for N^{lo}_T. Since the proof is based upon the fact that the counting process ξ_T can be replaced by its mean μ_T, the same techniques can be used to show convergence of N^{up}_T.

Remark 2.1. To prove convergence of N^{up}_T we can also apply Theorem 4.7 of [8] (see also [14, Proposition 3.22]). This involves showing that $E N^{up}_T(A) \to E N(A)$, which follows from Wald's identity and (2.1). An alternative way to prove Theorem 2.1 is by starting with convergence of either N^{lo}_T or N^{up}_T and then showing that the distance between N^{lo}_T and N^{up}_T in the vague topology on $M_p(E)$ (see [14, Section 3.4]) becomes arbitrarily small as $T \to \infty$. By virtue of Theorem 4.2 of [8], this means showing that for each continuous function $f : E \to \mathbb{R}$ with compact support, we have

$$P(|N^{lo}_T(f) - N^{up}_T(f)| > \epsilon) \to 0,$$

for every $\epsilon > 0$.

In Figure 1 the exceedances of simulated series of ON-periods are depicted for $M_T = \log(T)$.

$T = 100, M = 5, x = 2.6$

$T = 1000, M = 7, x = 25$

$T = 5000, M = 9, x = 112$
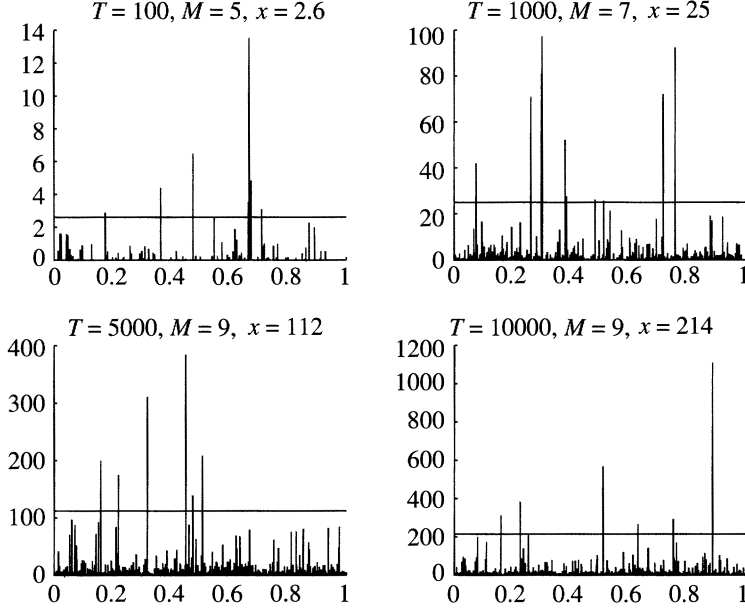
$T = 10000, M = 9, x = 214$

FIGURE 1: Illustration of the exceedances of simulated series of ON-periods in the slow growth situation. Both F_{on} and F_{off} are Pareto with tail parameters $\alpha = 1.2$ and 1.8 and means $\mu_{on} = 3$ and $\mu_{off} = 5$, respectively. Here $M_T = \log(T)$ and x_T is chosen such that $M T \bar{F}_{on}(x_T)/\mu \sim 10$. It is clear from the plots that the number of exceedances stabilizes as T becomes large. Moreover, the exceedances do not appear in clusters, which is what might be expected in the case of a Poisson limit.

2.2. Convergence to a homogeneous Poisson process

In this section, we will use the convergence of N_T^{lo} and N_T^{up} to a PRM to show that, as $T \to \infty$, the number of exceedances up to time Tt by the double array $X_i^{(m)}$ converges to a homogeneous Poisson process. Again we assume that there are M i.i.d. sources, where M is either fixed or increasing as $T \to \infty$. In the latter case we also suppose that M satisfies the slow growth condition. The total number of exceedances up to time Tt is given by

$$A_T(t) = \sum_{m=1}^{M} \left[\sum_{i=1}^{\xi_{Tt}^{(m)}-1} 1_{[x_T,\infty)}(X_i^{(m)}) + 1_{[x_T,\infty)}\left(\min\left(T - S_{\xi_{Tt}^{(m)}-1}^{(m)}, X_{\xi_{Tt}^{(m)}}^{(m)} \right) \right) \right].$$

We assume that the threshold x_T satisfies (2.1). Notice that the processes A_T contain more information on the times when exceedances occur, while N_T^{lo} and N_T^{up} hold more information on their sizes.

Let $\mathbb{D}[0, \infty)$ denote the space of functions on $[0, \infty)$ which are right-continuous and have left limits. We assume that $\mathbb{D}[0, \infty)$ is equipped with the J_1-topology (see [13, Chapter VI]). We have the following result.

Proposition 2.1. Let $(Z(t), t \geq 0)$ be a homogeneous Poisson process with intensity μ^{-1}. If M satisfies the slow growth condition and x_T satisfies (2.1), then as $T \to \infty$

$$(A_T(t), t \geq 0) \xrightarrow{D} (Z(t), t \geq 0),$$

where $\xrightarrow{D}$ denotes weak convergence in $(\mathbb{D}[0, \infty), J_1)$.

Proof. From Theorem 2.1 it follows that the families of point processes $N_T^{lo}(\cdot \times [1, \infty))$ and $N_T^{up}(\cdot \times [1, \infty))$ converge weakly in $M_p([0, \infty))$ to $N(\cdot \times [1, \infty))$. Define for $t \geq 0$

$$A_T^{lo}(t) = \sum_{m=1}^{M} \sum_{i=1}^{\xi_{Tt}^{(m)}-1} 1_{[x_T, \infty)}(X_i^{(m)}) \quad \text{and} \quad A_T^{up}(t) = \sum_{m=1}^{M} \sum_{i=1}^{\xi_{Tt}^{(m)}} 1_{[x_T, \infty)}(X_i^{(m)}).$$

Since
$$A_T^{lo}(t) = N_T^{lo}([0, t) \times [1, \infty)) \quad \text{and} \quad A_T^{up}(t) = N_T^{up}([0, t) \times [1, \infty)),$$

A_T^{lo} and A_T^{up} are the corresponding families of cumulative processes. Notice that, for all $t \geq 0$ and all ω,

$$A_T^{lo}(t) \leq A_T(t) \leq A_T^{up}(t).$$

The result now follows from these bounds, Lemma 9.1.X of [3], and the fact that $N([0, t) \times [1, \infty))$ is a homogeneous Poisson process with intensity μ^{-1}.

3. Fast growth of M

In this section, we derive the limit distribution of the number of exceedances if M satisfies the fast growth condition

$$\lim_{T \to \infty} \frac{b(MT)}{T} = \infty \quad \Longleftrightarrow \quad \lim_{T \to \infty} MT\bar{F}_{on}(T) = \infty,$$

where the quantile function b is defined in (1.5). Recall from Section 1 that we consider the following number of exceedances:

$$\widehat{A}_T = \sum_{m=1}^{M} \sum_{i=1}^{\xi_T^{(m)}} 1_{[T-a_T, T)}(X_i^{(m)}),$$

where a_T is a positive sequence satisfying $a_T = o(T)$. We have

$$E\,\widehat{A}_T = M\frac{T}{\mu}\bar{F}_{on}(T)\left[\frac{\bar{F}_{on}(T - a_T)}{\bar{F}_{on}(T)} - 1\right].$$

For any M satisfying the fast growth condition we choose a_T such that $E\,\widehat{A}_T \sim$ const. Proposition B.1 gives that $\bar{F}_{on}(T - a_T)/\bar{F}_{on}(T) = 1 + o(1)$, but in order to have $E\,\widehat{A}_T \sim$ const. we need to know the rate at which the $o(1)$ term converges to 0. This requires us to impose a second order regular variation condition on $\bar{F}_{on}$. In practice, however, such a condition often cannot be verified. Therefore, and since we are only interested in a qualitative characterization concerning the limit of $\widehat{A}_T$, we will assume from now on that for some x_0 and some constant $c > 0$,

$$\bar{F}_{on}(x) = cx^{-\alpha} \quad \text{for } x \geq x_0. \tag{3.1}$$

As before, $\alpha > 1$ and $\bar{F}_{off}(x) = o(\bar{F}_{on}(x))$. Then, for T large enough,

$$E\,\widehat{A}_T = M\frac{T}{\mu}\bar{F}_{on}(T)\left[\left(1 - \frac{a_T}{T}\right)^{-\alpha} - 1\right].$$

Using a first order Taylor expansion, we obtain

$$
\begin{aligned}
\mathrm{E}\,\widehat{A}_T &= M\frac{T}{\mu}\bar{F}_{\mathrm{on}}(T)\Big[\alpha\frac{a_T}{T} + o\Big(\frac{a_T}{T}\Big)\Big] \\
&= \frac{\alpha}{\mu}Ma_T\,\bar{F}_{\mathrm{on}}(T) + o(Ma_T\,\bar{F}_{\mathrm{on}}(T)).
\end{aligned}
$$

Therefore, we choose a_T such that

$$
a_T = o(T) \quad \text{and} \quad Ma_T\,\bar{F}_{\mathrm{on}}(T) \sim 1. \tag{3.2}
$$

In this way,

$$
\lim_{T\to\infty}\mathrm{E}\,\widehat{A}_T = \frac{\alpha}{\mu}.
$$

We will derive the limit distribution of $\widehat{A}_T$ as $T \to \infty$. In Section 3.1, we will use weak convergence of point processes to a PRM to show that $\widehat{A}_T \xrightarrow{\mathrm{D}} \mathrm{Poi}(\alpha\mu^{-1})$. This convergence is illustrated by means of simulations. In Section 3.2, we consider convergence to a homogeneous Poisson process.

3.1. Convergence to a Poisson random measure

Define the point processes

$$
\widehat{N}_T = \sum_{m=1}^{M}\widehat{N}_T^{(m)},
$$

where

$$
\widehat{N}_T^{(m)} = \sum_{i=1}^{\infty}\varepsilon_{S_{i-1}^{(m)}/T,(T-X_i^{(m)})/a_T}.
$$

Let $E := [0,\infty) \times (0,1]$ be the state space of $\widehat{N}_T$. Notice that

$$
\widehat{N}_T([0,1)\times(0,1]) = \sum_{m=1}^{M}\sum_{i=1}^{\xi_T^{(m)}}1_{[T-a_T,T)}(X_i^{(m)}) = \widehat{A}_T.
$$

In the following theorem we show that $\widehat{N}_T$ weakly converges to a PRM.

Theorem 3.1. *Assume that $\bar{F}_{\mathrm{on}}$ satisfies (3.1). If M satisfies the fast growth condition and a_T satisfies (3.2), then as $T \to \infty$*

$$
\widehat{N}_T \xrightarrow{\mathrm{D}} \widehat{N} \quad \text{in } M_p(E),
$$

where $\widehat{N}$ is PRM with mean measure $\alpha\mu^{-1}\mathbb{L}\times\mathbb{L}$.

Since $\alpha\mu^{-1}(\mathbb{L}\times\mathbb{L})([0,1)\times(0,1]) = \alpha\mu^{-1}$, Theorem 3.1 implies that, as $T \to \infty$,

$$
\widehat{A}_T = \sum_{m=1}^{M}\sum_{i=1}^{\xi_T^{(m)}}1_{[T-a_T,T)}(X_i^{(m)}) \xrightarrow{\mathrm{D}} \mathrm{Poi}(\alpha\mu^{-1}).
$$

The proof of Theorem 3.1 is presented below. We need the following lemma.

Lemma 3.1. *Assume that $\bar{F}_{\text{on}}$ satisfies (3.1). Suppose that M satisfies the fast growth condition and a_T satisfies (3.2). Let $0 \leq a < b \leq 1$ and define*

$$B_i = \{(T - X_i)/a_T \in (a, b]\}, \qquad i \geq 1.$$

As $T \to \infty$,

$$M \, \mathrm{E}\left(\sum_{i=2}^{\xi_T} \sum_{l=1}^{i-1} 1_{B_i \cap B_l}\right) \to 0.$$

Proof. Using a first order Taylor expansion, we have for T large enough

$$
\begin{aligned}
\mathrm{P}(B_1) &= \bar{F}_{\text{on}}(T - ba_T) - \bar{F}_{\text{on}}(T - aa_T) \\
&= \bar{F}_{\text{on}}(T)\left(\frac{\bar{F}_{\text{on}}(T - ba_T)}{\bar{F}_{\text{on}}(T)} - \frac{\bar{F}_{\text{on}}(T - aa_T)}{\bar{F}_{\text{on}}(T)}\right) \\
&= \bar{F}_{\text{on}}(T)\left(\left(1 - b\frac{a_T}{T}\right)^{-\alpha} - \left(1 - a\frac{a_T}{T}\right)^{-\alpha}\right) \\
&= \bar{F}_{\text{on}}(T)\left(\alpha(b - a)\frac{a_T}{T} + o\left(\frac{a_T}{T}\right)\right).
\end{aligned}
$$

Hence, by (3.1) and (3.2),

$$MT \, \mathrm{P}(B_1) \sim \alpha(b - a)Ma_T \, \bar{F}_{\text{on}}(T) \sim \alpha(b - a). \tag{3.3}$$

We have

$$
\begin{aligned}
2\sum_{i=2}^{\xi_T} \sum_{l=1}^{i-1} 1_{B_i \cap B_l} &= \left(\sum_{i=1}^{\xi_T} 1_{B_i}\right)^2 - \sum_{i=1}^{\xi_T} 1_{B_i} \\
&= \left(\sum_{i=1}^{\xi_T} (1_{B_i} - \mathrm{P}(B_1))\right)^2 + 2\mathrm{P}(B_1)\xi_T \sum_{i=1}^{\xi_T} 1_{B_i} - [\mathrm{P}(B_1)]^2 \xi_T^2 - \sum_{i=1}^{\xi_T} 1_{B_i} \\
&=: I_T + 2II_T - III_T - IV_T.
\end{aligned}
$$

From [6, Theorem 5.1 in Chapter II], we have that for $r > 0$

$$\mathrm{E}\,\xi_T^r \sim \frac{T^r}{\mu^r} \quad \text{as } T \to \infty. \tag{3.4}$$

Using this fact and (3.3), we obtain

$$M \, \mathrm{E}\, III_T \sim \text{const. } M^{-1} \to 0.$$

Theorem 5.3 of [6, Chapter I], gives

$$\mathrm{E}\left(\sum_{i=1}^{\xi_T} (1_{B_i} - \mathrm{P}(B_1))\right)^2 = \text{var}(1_{B_1}) \, \mathrm{E}\,\xi_T. \tag{3.5}$$

Using this, we find that

$$
\begin{aligned}
M(\mathrm{E}\, I_T - \mathrm{E}\, IV_T) &= M \, \text{var}(1_{B_1}) \, \mathrm{E}\,\xi_T - M \, \mathrm{P}(B_1) \, \mathrm{E}\,\xi_T \\
&= M(\mathrm{P}(B_1) - [\mathrm{P}(B_1)]^2) \, \mathrm{E}\,\xi_T - M \, \mathrm{P}(B_1) \, \mathrm{E}\,\xi_T \\
&= -M[\mathrm{P}(B_1)]^2 \, \mathrm{E}\,\xi_T,
\end{aligned}
$$

which is $o(1)$ by (3.3) and (3.4). For II_T we have

$$M \, \mathrm{E} \, II_T = M \, \mathrm{P}(B_1) \, \mathrm{E} \left(\xi_T \sum_{i=1}^{\xi_T} (1_{B_i} - \mathrm{P}(B_1)) \right) + M[\mathrm{P}(B_1)]^2 \, \mathrm{E} \, \xi_T^2.$$

The second term on the right-hand side is $o(1)$ by (3.3) and (3.4). Using the Cauchy–Schwarz inequality and (3.5), we obtain for the first term the upper bound

$$M \, \mathrm{P}(B_1) (\mathrm{E} \, \xi_T^2)^{1/2} (\mathrm{P}(B_1)[1 - \mathrm{P}(B_1)] \, \mathrm{E} \, \xi_T)^{1/2},$$

which is $o(1)$ by (3.3) and (3.4). This completes the proof.

Proof of Theorem 3.1. We consider sets of the form

$$A = \bigcup_{j=1}^{k} [r_j, s_j) \times (a_j, b_j],$$

for $k \geq 1$, $[r_j, s_j) \times (a_j, b_j] \subset E$, $j = 1, \ldots, k$. We may assume that the sets $[r_j, s_j) \times (a_j, b_j]$ are mutually disjoint. Notice that $\widehat{N}_T$ are the row sums of a triangular array of point processes. According to Theorem A.2 below, it suffices to show that

$$M \, \mathrm{P}(\widehat{N}_T^{(1)}(A) \geq 1) \to \mathrm{E} \, \widehat{N}(A) \quad \text{and} \quad M \, \mathrm{P}(\widehat{N}_T^{(1)}(A) \geq 2) \to 0.$$

Since

$$\mathrm{E} \, \widehat{N}_T(A) = \sum_{n=1}^{\infty} M \, \mathrm{P}(\widehat{N}_T^{(1)}(A) \geq n),$$

it also suffices to show that

$$\mathrm{E} \, \widehat{N}_T(A) \to \mathrm{E} \, \widehat{N}(A) \quad \text{and} \quad M \, \mathrm{P}(\widehat{N}_T^{(1)}(A) \geq 1) \to \mathrm{E} \, \widehat{N}(A).$$

Define

$$B_{ij} = \{(T - X_i)/a_T \in (a_j, b_j]\}, \qquad i \geq 1, j = 1, \ldots, k.$$

Notice that

$$\widehat{N}_T(A) = \sum_{m=1}^{M} \sum_{i=\xi_{Tr_j}^{(m)}+1}^{\xi_{Ts_j}^{(m)}} 1_{[T-b_j a_T, T-a_j a_T)}(X_i^{(m)}).$$

Using the fact that ξ_T is a stopping time, Wald's identity and (3.3), we obtain

$$\mathrm{E} \, \widehat{N}_T(A) = \mu^{-1} \sum_{j=1}^{k} (s_j - r_j) M T \, \mathrm{P}(B_{1j})$$

$$\sim \alpha \mu^{-1} \sum_{j=1}^{k} (s_j - r_j)(b_j - a_j) = \mathrm{E} \, \widehat{N}(A).$$

It remains to show that $M \, \mathrm{P}(\widehat{N}_T^{(1)}(A) \geq 1) \to \mathrm{E} \, \widehat{N}(A)$. By Markov's inequality we have

$$\limsup_{T \to \infty} M \, \mathrm{P}(\widehat{N}_T^{(1)}(A) \geq 1) \leq \mathrm{E} \, \widehat{N}(A).$$

For nonnegative vectors $\boldsymbol{n}_r = (n_r^{(1)}, \ldots, n_r^{(k)})$ and $\boldsymbol{n}_s = (n_s^{(1)}, \ldots, n_s^{(k)})$, let

$$A_{\boldsymbol{n}_r}^{\boldsymbol{n}_s} = \{\xi_{Tr_j} = n_r^{(j)}, \xi_{Ts_j} = n_s^{(j)}, \ j = 1, \ldots, k\}.$$

Notice that for $\boldsymbol{n}_r \leq \boldsymbol{n}_s$ (which means that $n_r^{(j)} \leq n_s^{(j)}$ for $j = 1, \ldots, k$), the $A_{\boldsymbol{n}_r}^{\boldsymbol{n}_s}$ constitute a partition of Ω. We write

$$P(\widehat{N}_T^{(1)}(A) \geq 1) = P\left(\sum_{j=1}^{k} \sum_{i=\xi_{Tr_j}+1}^{\xi_{Ts_j}} 1_{B_{ij}} \geq 1\right) = P\left(\bigcup_{j=1}^{k} \bigcup_{i=\xi_{Tr_j}+1}^{\xi_{Ts_j}} B_{ij}\right)$$

$$= \sum_{\boldsymbol{n}_r \leq \boldsymbol{n}_s} P\left(\bigcup_{j=1}^{k} \bigcup_{i=n_r^{(j)}+1}^{n_s^{(j)}} (B_{ij} \cap A_{\boldsymbol{n}_r}^{\boldsymbol{n}_s})\right).$$

For fixed $\boldsymbol{n}_r$ and $\boldsymbol{n}_s$, we apply the inclusion–exclusion formula on the two unions as a whole, which yields the lower bound

$$\sum_{j=1}^{k} \sum_{i=n_r^{(j)}+1}^{n_s^{(j)}} P(B_{ij} \cap A_{\boldsymbol{n}_r}^{\boldsymbol{n}_s}) - \sum_{j=2}^{k} \sum_{n=1}^{j-1} \sum_{i=n_r^{(j)}+1}^{n_s^{(j)}} \sum_{l=n_r^{(n)}+1}^{n_s^{(n)}} P(B_{ij} \cap B_{ln} \cap A_{\boldsymbol{n}_r}^{\boldsymbol{n}_s})$$

$$- \sum_{j=1}^{k} \sum_{i=n_r^{(j)}+2}^{n_s^{(j)}} \sum_{l=n_r^{(j)}+1}^{i-1} P(B_{ij} \cap B_{lj} \cap A_{\boldsymbol{n}_r}^{\boldsymbol{n}_s}).$$

Using $P(B) = E(1_B)$ and the linearity of the expectation, this equals

$$\sum_{j=1}^{k} E\left(\sum_{i=n_r^{(j)}+1}^{n_s^{(j)}} 1_{B_{ij} \cap A_{\boldsymbol{n}_r}^{\boldsymbol{n}_s}}\right) - \sum_{j=2}^{k} \sum_{n=1}^{j-1} E\left(\sum_{i=n_r^{(j)}+1}^{n_s^{(j)}} \sum_{l=n_r^{(n)}+1}^{n_s^{(n)}} 1_{B_{ij} \cap B_{ln} \cap A_{\boldsymbol{n}_r}^{\boldsymbol{n}_s}}\right)$$

$$- \sum_{j=1}^{k} E\left(\sum_{i=n_r^{(j)}+2}^{n_s^{(j)}} \sum_{l=n_r^{(j)}+1}^{i-1} 1_{B_{ij} \cap B_{lj} \cap A_{\boldsymbol{n}_r}^{\boldsymbol{n}_s}}\right).$$

Since the $A_{\boldsymbol{n}_r}^{\boldsymbol{n}_s}$ are a partition of Ω, we have shown that

$$M P(\widehat{N}_T^{(1)}(A) \geq 1) \geq M \sum_{j=1}^{k} E\left(\sum_{i=\xi_{Tr_j}+1}^{\xi_{Ts_j}} 1_{B_{ij}}\right) - M \sum_{j=2}^{k} \sum_{n=1}^{j-1} E\left(\sum_{i=\xi_{Tr_j}+1}^{\xi_{Ts_j}} \sum_{l=\xi_{Tr_n}+1}^{\xi_{Ts_n}} 1_{B_{ij} \cap B_{ln}}\right)$$

$$- M \sum_{j=1}^{k} E\left(\sum_{i=\xi_{Tr_j}+2}^{\xi_{Ts_j}} \sum_{l=\xi_{Tr_j}+1}^{i-1} 1_{B_{ij} \cap B_{lj}}\right)$$

$$= I_T - II_T - III_T.$$

Notice that $I_T = E \widehat{N}_T(A)$, which converges to $E \widehat{N}(A)$. It remains to show that II_T and III_T converge to 0. Since, for fixed j,

$$\sum_{i=\xi_{Tr_j}+2}^{\xi_{Ts_j}} \sum_{l=\xi_{Tr_j}+1}^{i-1} 1_{B_{ij} \cap B_{lj}} \leq \sum_{i=2}^{\xi_{Ts_j}} \sum_{l=1}^{i-1} 1_{B_{ij} \cap B_{lj}},$$

the convergence $III_T \to 0$ follows from Lemma 3.1. Next we deal with II_T. Fix j and n. We have $n < j$. Assume that the sets $[r_j, s_j)$ are ordered such that

$$n < j \quad \implies \quad r_n \leq r_j.$$

Denote

$$D_T^{(j,n)} = \sum_{i=\xi_{Tr_j}+1}^{\xi_{Ts_j}} \sum_{l=\xi_{Tr_n}+1}^{\xi_{Ts_n}} 1_{B_{ij} \cap B_{ln}},$$

and define the following disjoint partition of Ω:

$$C_1(j, n) = \{\xi_{Ts_n} \leq \xi_{Tr_j}\},$$

$$C_2(j, n) = \{\xi_{Ts_j} \geq \xi_{Ts_n} > \xi_{Tr_j}\},$$

$$C_3(j, n) = \{\xi_{Ts_n} > \xi_{Ts_j}\}.$$

We will show that

$$D_T^{(j,n)} \leq 3 \sum_{i=2}^{\xi_{Ts_j}} \sum_{l=1}^{i-1} 1_{\widetilde{B}_i \cap \widetilde{B}_l} + 2 \sum_{l=2}^{\xi_{Ts_n}} \sum_{i=1}^{l-1} 1_{\widetilde{B}_i \cap \widetilde{B}_l}, \tag{3.6}$$

where

$$\widetilde{B}_i = \{(T - X_i)/a_T \in (\min(a_j, a_n), \max(b_j, b_n)]\}, \qquad i \geq 1.$$

Then $II_T \to 0$ follows from Lemma 3.1. It is not difficult to see that

$$D_T^{(j,n)} 1_{C_1(j,n)} \leq \sum_{i=\xi_{Tr_j}+1}^{\xi_{Ts_j}} \sum_{l=\xi_{Tr_n}+1}^{\xi_{Ts_n}} 1_{\widetilde{B}_i \cap \widetilde{B}_l} \leq \sum_{i=2}^{\xi_{Ts_j}} \sum_{l=1}^{i-1} 1_{\widetilde{B}_i \cap \widetilde{B}_l}. \tag{3.7}$$

Next consider $C_2(j, n)$. From the definition of ξ_T in (1.3) it follows that if $C_2(j, n) \neq \varnothing$, then $s_n > r_j$. The ordering of the sets $[r_j, s_j)$ now implies that $[r_j, s_j)$ and $[r_n, s_n)$ are not disjoint. Since $[r_j, s_j) \times (a_j, b_j]$ are mutually disjoint, it follows that $(a_j, b_j] \cap (a_n, b_n] = \varnothing$. Hence, $B_{ij} \cap B_{ln} = \varnothing$ if $i = l$. The reasoning above implies that

$$D_T^{(j,n)} 1_{C_2(j,n)} \leq \sum_{i=\xi_{Tr_n}+1}^{\xi_{Ts_j}} \sum_{\substack{l=\xi_{Tr_n}+1 \\ l \neq i}}^{\xi_{Ts_j}} 1_{B_{ij} \cap B_{ln}} \leq \sum_{i=\xi_{Tr_n}+1}^{\xi_{Ts_j}} \sum_{\substack{l=\xi_{Tr_n}+1 \\ l \neq i}}^{\xi_{Ts_j}} 1_{\widetilde{B}_i \cap \widetilde{B}_l} \leq 2 \sum_{i=2}^{\xi_{Ts_j}} \sum_{l=1}^{i-1} 1_{\widetilde{B}_i \cap \widetilde{B}_l}. \tag{3.8}$$

Since $C_3(j, n) \neq \varnothing$ also implies that $[r_j, s_j)$ and $[r_n, s_n)$ are not disjoint, we have analogously

$$D_T^{(j,n)} 1_{C_3(j,n)} \leq \sum_{i=\xi_{Tr_n}+1}^{\xi_{Ts_n}} \sum_{\substack{l=\xi_{Tr_n}+1 \\ l \neq i}}^{\xi_{Ts_n}} 1_{B_{ij} \cap B_{ln}} \leq 2 \sum_{i=2}^{\xi_{Ts_n}} \sum_{l=1}^{i-1} 1_{\widetilde{B}_i \cap \widetilde{B}_l}. \tag{3.9}$$

The proof now follows by observing that (3.7)–(3.9) together imply (3.6).

In Figure 2 the exceedances of simulated series of ON-periods are depicted for $M_T = T$.
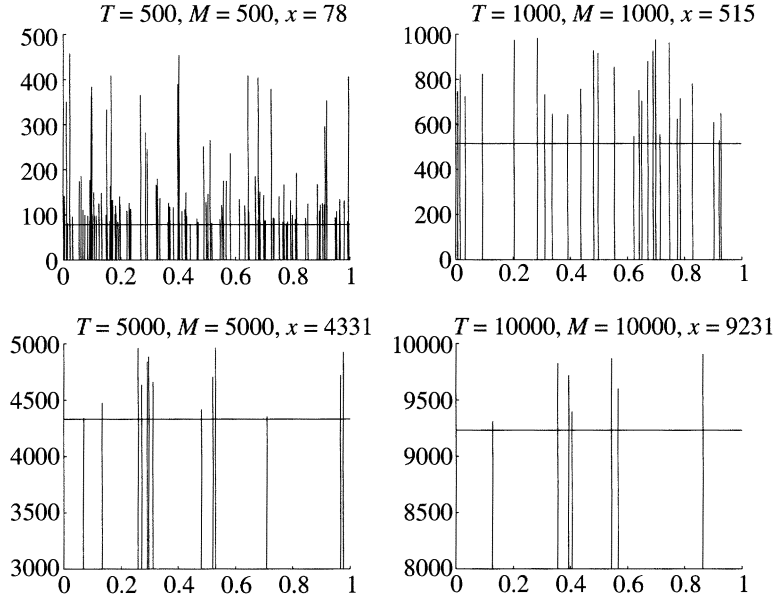
FIGURE 2: Illustration of the exceedances of simulated series of ON-periods in the fast growth situation. Both F_{on} and F_{off} are Pareto with tail parameters $\alpha = 1.2$ and 1.8 and means $\mu_{\mathrm{on}} = 3$ and $\mu_{\mathrm{off}} = 5$, respectively. Here $M_T = T$ and x_T is chosen such that $MT[\bar{F}_{\mathrm{on}}(x_T) - \bar{F}_{\mathrm{on}}(T)]/\mu \sim 10$. To make the plots more clear, only the ON-period lengths that lie in $[x_T, T]$ are depicted. As in the slow growth situation the number of exceedances stabilizes and there are no clusters, which is consistent with a Poisson limit.

Remark 3.1. In Figure 2, almost all exceedances in the plots with $T = 5000$ and $T = 10000$ are due to the ON-periods $X^{(m)}_{\xi^{(m)}_T}$. This can be seen as follows. Let $n \in \{1, \ldots, \xi^{(m)}_T - 1\}$. By definition, $S^{(m)}_{\xi_T - 1} \le T$ and hence the maximum length of an ON-period $X^{(m)}_n$ is $T - S^{(m)}_{n-1}$. For an ON-period $X^{(m)}_n$ exceeding x_T we must have $T - S^{(m)}_{n-1} \ge x_T$, which is equivalent to $S^{(m)}_{n-1} \le T - x_T$. In the plots we normalized the time scale to the interval $[0, 1]$. Hence, for any ON-period $X^{(m)}_i$ exceeding x_T and starting at a time later than $(T - x_T)/T$, we have $i = \xi^{(m)}_T$. For the plots with $T = 5000$ and $T = 10000$ we have $(T - x_T)/T = 0.134$ and 0.077, respectively. So for $T = 10000$ all 7 exceedances are due to the ON-periods $X^{(m)}_{\xi^{(m)}_T}$. For $T = 5000$ this is true for at least 11 of the 13 exceedances. This raises the question as to whether the same holds for $\widehat{A}_T$, i.e. as $T \to \infty$, does

$$\sum_{m=1}^{M} 1_{[T - a_T, T)}(X^{(m)}_{\xi^{(m)}_T}) \xrightarrow{\mathrm{D}} \mathrm{Poi}(\alpha \mu^{-1})?$$

We were, however, not able to obtain a conclusive answer.

3.2. Convergence to a homogeneous Poisson process

The following result is analogous to Proposition 2.1. Since the proof is similar, it is omitted. Define for $t \in [0, 1]$

$$\widehat{A}_T(t) = \sum_{m=1}^{M} \sum_{i=1}^{\xi^{(m)}_{Tt}} 1_{[T - a_T, T)}(X^{(m)}_i).$$

Proposition 3.1. *Assume that $\bar{F}_{\mathrm{on}}$ satisfies (3.1). Let $(Z(t),\ t \geq 0)$ be a homogeneous Poisson process with intensity $\alpha\mu^{-1}$. If M satisfies the fast growth condition and a_T satisfies (3.2), then as $T \to \infty$*

$$(\widehat{A}_T(t),\ t \geq 0) \xrightarrow{\mathrm{D}} (Z(t),\ t \geq 0),$$

where $\xrightarrow{\mathrm{D}}$ denotes weak convergence in $(\mathbb{D}[0, 1], J_1)$.

Remark 3.2. From the steps in the proof of Theorem 3.1 it is clear that similar results hold if we impose a second order regular variation condition on $\bar{F}_{\mathrm{on}}$, instead of assuming (3.1). Proposition 3.1 will also follow in this case. Such a proof would, however, be more technical.

4. Gaussian approximations

In this section, we show that the number of exceedances

$$\sum_{m=1}^{M} \sum_{i=1}^{\xi_T^{(m)}} 1_{[x_T, T]}(X_i^{(m)})$$

satisfies the central limit theorem for certain thresholds $x_T \to \infty$. The number of sources $M = M_T \to \infty$ is a nondecreasing integer-valued function. Together, M and x_T have to satisfy

$$x_T \leq T \quad \text{and} \quad MT(\bar{F}_{\mathrm{on}}(x_T) - \bar{F}_{\mathrm{on}}(T)) \to \infty. \tag{4.1}$$

The latter is needed to have an increasing amount of exceedances (see (1.13)). Define

$$Z_T = \sum_{m=1}^{M} Z_T^{(m)}$$

$$= \sum_{m=1}^{M} [MT(\bar{F}_{\mathrm{on}}(x_T) - \bar{F}_{\mathrm{on}}(T))]^{-1/2} \sum_{i=1}^{\xi_T^{(m)}} (1_{[x_T, T]}(X_i^{(m)}) - (\bar{F}_{\mathrm{on}}(x_T) - \bar{F}_{\mathrm{on}}(T))).$$

Our goal is to show that Z_T has a Gaussian limit as $T \to \infty$. In doing so, we will use the central limit theorem. Since each $\xi_T^{(m)}$ is a stopping time, we have $\mathrm{E}\,Z_T = 0$. Using (3.5) and the fact that

$$\mathrm{var}(1_{[x_T, T]}(X_1)) = (\bar{F}_{\mathrm{on}}(x_T) - \bar{F}_{\mathrm{on}}(T))(1 - (\bar{F}_{\mathrm{on}}(x_T) - \bar{F}_{\mathrm{on}}(T))) \sim \bar{F}_{\mathrm{on}}(x_T) - \bar{F}_{\mathrm{on}}(T),$$

we obtain

$$\mathrm{var}(Z_T) = \frac{M\,\mathrm{E}(\xi_T)\,\mathrm{var}(1_{[x_T, T]}(X_1))}{MT(\bar{F}_{\mathrm{on}}(x_T) - \bar{F}_{\mathrm{on}}(T))} \sim \mu^{-1}.$$

Notice that, if we replaced each $\xi_T^{(m)}$ by μ_T, Z_T would actually be a centered and normalized sum of M i.i.d. binomial random variables with $[T/\mu]$ trials and success probability $\bar{F}_{\mathrm{on}}(x_T) - \bar{F}_{\mathrm{on}}(T)$. Contrary to Sections 1–3, we do not need the assumption that $\bar{F}_{\mathrm{on}}$ is regularly varying. In fact, it suffices to have

$$\mu = \int_0^{\infty} (\bar{F}_{\mathrm{on}}(s) + \bar{F}_{\mathrm{off}}(s))\,\mathrm{d}s < \infty.$$

In this way, the existence of a stationary renewal sequence (S_n), defined by (1.2), is guaranteed. In Section 4.1, we will show a central limit theorem for Z_T, using the fact that $\xi_T^{(m)}$ are stopping times.

4.1. A central limit theorem

The following result shows that Z_T converges weakly to a normal distribution as $T \to \infty$.

Theorem 4.1. *Suppose that $\mu < \infty$. If M and x_T satisfy (4.1), then as $T \to \infty$*

$$\sum_{m=1}^{M} Z_T^{(m)} \overset{D}{\to} N(0, \mu^{-1}).$$

Notice that there is no distinction between fast and slow growth of M.

The proof of Theorem 4.1 is given below. For ease of presentation, we introduce some notation. Set

$$\widetilde{X}_i = 1_{[x_T, T]}(X_i), \qquad \tilde{\mu} = \mathrm{E}\,\widetilde{X}, \qquad \tilde{\sigma}^2 = \mathrm{var}(\widetilde{X}),$$
$$\tilde{\mu}_3 = \mathrm{E}(\widetilde{X} - \tilde{\mu})^3, \qquad \tilde{\mu}_4 = \mathrm{E}(\widetilde{X} - \tilde{\mu})^4.$$

For $a \le b$, write

$$\bar{F}_{\mathrm{on}}(a, b) = \bar{F}_{\mathrm{on}}(a) - \bar{F}_{\mathrm{on}}(b).$$

Notice that

$$\tilde{\sigma}^2 \sim \tilde{\mu}_3 \sim \tilde{\mu}_4 \sim \bar{F}_{\mathrm{on}}(x_T, T). \tag{4.2}$$

For $n \ge 1$, let

$$\widetilde{S}_0 = 0, \qquad \widetilde{S}_n = \widetilde{X}_1 + \cdots + \widetilde{X}_n, \quad \text{and} \quad \mathcal{F}_n = \sigma(D, X_1, Y_1, \ldots, X_n, Y_n).$$

Recall that ξ_T is a stopping time with respect to the filtration $(\mathcal{F}_n)$.

Proof of Theorem 4.1. We will derive Lyapunov's condition (with $\delta = 2$) for the central limit theorem, i.e. as $T \to \infty$

$$\frac{M\,\mathrm{E}(Z_T^{(1)})^4}{[M\,\mathrm{E}(Z_T^{(1)})^2]^2} \sim \mu^2 M\,\mathrm{E}(Z_T^{(1)})^4 \to 0.$$

We have

$$M\,\mathrm{E}(Z_T^{(1)})^4 = \frac{\mathrm{E}(\widetilde{S}_{\xi_T} - \xi_T\,\tilde{\mu})^4}{MT^2[\bar{F}_{\mathrm{on}}(x_T, T)]^2}. \tag{4.3}$$

The approach is to calculate $\mathrm{E}(\widetilde{S}_{\xi_T} - \xi_T\,\tilde{\mu})^4$ by first constructing a martingale and then using the optional stopping theorem. It can be seen that

$$\mathrm{E}[(\widetilde{S}_n - n\tilde{\mu})^4 \mid \mathcal{F}_{n-1}]$$
$$= (\widetilde{S}_{n-1} - (n-1)\tilde{\mu})^4 + 6\tilde{\sigma}^2(\widetilde{S}_{n-1} - (n-1)\tilde{\mu})^2 + 4\tilde{\mu}_3(\widetilde{S}_{n-1} - (n-1)\tilde{\mu}) + \tilde{\mu}_4.$$

Define

$$A_N = \sum_{n=1}^{N} [(\widetilde{S}_n - n\tilde{\mu})^4 - \mathrm{E}[(\widetilde{S}_n - n\tilde{\mu})^4 \mid \mathcal{F}_{n-1}]].$$

By definition, A_N is a martingale with respect to the filtration $(\mathcal{F}_N)$ and from the optional stopping theorem (see [6, Theorem A2.4]) it follows that $\mathrm{E}\,A_{\xi_T} = 0$. This implies that

$$\mathrm{E}(\widetilde{S}_{\xi_T} - \xi_T\tilde{\mu})^4$$
$$= 6\tilde{\sigma}^2\,\mathrm{E}\left[\sum_{n=1}^{\xi_T}(\widetilde{S}_{n-1} - (n-1)\tilde{\mu})^2\right] + 4\tilde{\mu}_3\,\mathrm{E}\left[\sum_{n=1}^{\xi_T}(\widetilde{S}_{n-1} - (n-1)\tilde{\mu})\right] + \tilde{\mu}_4\,\mathrm{E}\,\xi_T.$$

$$(4.4)$$

From (D.1) below, it follows that

$$\mathrm{E}\left[\sum_{n=1}^{\xi_T}(\widetilde{S}_{n-1} - (n-1)\tilde{\mu})\right] = \mathrm{E}\left[\sum_{i=1}^{\xi_T-1}(\widetilde{X}_i - \tilde{\mu})(\xi_T - i)\right] = \mathrm{E}\left[\sum_{i=1}^{\xi_T-1}(\widetilde{X}_i - \tilde{\mu})i\right]$$
$$= \mathrm{E}\left[\sum_{i=1}^{\xi_T}(\widetilde{X}_i - \tilde{\mu})i\right] - \mathrm{E}[\xi_T(\widetilde{X}_{\xi_T} - \tilde{\mu})] = -\mathrm{E}[\xi_T(\widetilde{X}_{\xi_T} - \tilde{\mu})].$$

The final step follows from the fact that $\sum_{i=1}^{n}(\widetilde{X}_i - \tilde{\mu})i$ is a martingale with respect to $(\mathcal{F}_n)$ and the optional stopping theorem.

Continuing the analysis, we write

$$\mathrm{E}\left[\sum_{n=1}^{\xi_T}(\widetilde{S}_{n-1} - (n-1)\tilde{\mu})^2\right]$$
$$= \mathrm{E}\left[\sum_{n=2}^{\xi_T}\sum_{i=1}^{n-1}(\widetilde{X}_i - \tilde{\mu})^2\right] + 2\,\mathrm{E}\left[\sum_{n=3}^{\xi_T}\sum_{i=2}^{n-1}\sum_{j=1}^{i-1}(\widetilde{X}_i - \tilde{\mu})(\widetilde{X}_j - \tilde{\mu})\right]. \quad (4.5)$$

By (D.1), the first term equals

$$\mathrm{E}\left[\sum_{i=1}^{\xi_T-1}(\widetilde{X}_i - \tilde{\mu})^2(\xi_T - i)\right] = \mathrm{E}\left[\sum_{i=1}^{\xi_T}(\widetilde{X}_i - \tilde{\mu})^2 i\right] - \mathrm{E}[\xi_T(\widetilde{X}_{\xi_T} - \tilde{\mu})^2]$$
$$= \frac{\tilde{\sigma}^2}{2}\,\mathrm{E}\,\xi_T^2 + \frac{\tilde{\sigma}^2}{2}\,\mathrm{E}\,\xi_T - \mathrm{E}[\xi_T(\widetilde{X}_{\xi_T} - \tilde{\mu})^2]. \quad (4.6)$$

In the last step we used the fact that $\sum_{i=1}^{n}(\widetilde{X}_i - \tilde{\mu})^2 i - \frac{1}{2}\tilde{\sigma}^2 n(n+1)$ is a martingale with respect to $(\mathcal{F}_n)$ and the optional stopping theorem.

By (D.2), the second term in (4.5) equals

$$2\,\mathrm{E}\left[\sum_{i=2}^{\xi_T-1}(\widetilde{X}_i - \tilde{\mu})(\xi_T - i)\sum_{j=1}^{i-1}(\widetilde{X}_j - \tilde{\mu})\right]$$
$$= 2\,\mathrm{E}\left[\sum_{i=2}^{\xi_T}(\widetilde{X}_i - \tilde{\mu})\sum_{j=1}^{i-1}(\widetilde{X}_j - \tilde{\mu})j\right] - 2\,\mathrm{E}\left[(\widetilde{X}_{\xi_T} - \tilde{\mu})\sum_{j=1}^{\xi_T-1}(\widetilde{X}_j - \tilde{\mu})j\right]$$
$$= -2\,\mathrm{E}\left[(\widetilde{X}_{\xi_T} - \tilde{\mu})\sum_{j=1}^{\xi_T}(\widetilde{X}_j - \tilde{\mu})j\right] + 2\,\mathrm{E}[\xi_T(\widetilde{X}_{\xi_T} - \tilde{\mu})^2]. \quad (4.7)$$

In the final step we used the fact that $\sum_{i=1}^{n}(\widetilde{X}_i - \tilde{\mu})\sum_{j=1}^{i-1}(\widetilde{X}_j - \tilde{\mu})j$ is a martingale with respect to $(\mathcal{F}_n)$ and the optional stopping theorem.

Combining (4.4)–(4.7) yields

$$
\mathrm{E}(\widetilde{S}_{\xi_T} - \xi_T\tilde{\mu})^4 = 3\tilde{\sigma}^4\,\mathrm{E}\,\xi_T^2 + (3\tilde{\sigma}^4 + \tilde{\mu}_4)\,\mathrm{E}\,\xi_T - 4\tilde{\mu}_3\,\mathrm{E}[\xi_T(\widetilde{X}_{\xi_T} - \tilde{\mu})]
$$

$$
+ 6\tilde{\sigma}^2\,\mathrm{E}[\xi_T(\widetilde{X}_{\xi_T} - \tilde{\mu})^2] - 12\tilde{\sigma}^2\,\mathrm{E}\left[(\widetilde{X}_{\xi_T} - \tilde{\mu})\sum_{j=1}^{\xi_T}(\widetilde{X}_j - \tilde{\mu})j\right].
$$

We will now show that

$$
\mathrm{E}(\widetilde{S}_{\xi_T} - \xi_T\tilde{\mu})^4 = o(MT^2[\bar{F}_{\mathrm{on}}(x_T, T)]^2).
$$

The proof of the Lyapunov condition then follows from (4.3).

By virtue of (3.4), (4.1) and (4.2) we obtain

$$
3\tilde{\sigma}^4\,\mathrm{E}\,\xi_T^2 + (3\tilde{\sigma}^4 + \tilde{\mu}_4)\,\mathrm{E}\,\xi_T = o(MT^2[\bar{F}_{\mathrm{on}}(x_T, T)]^2).
$$

Furthermore,

$$
|\tilde{\mu}_3\,\mathrm{E}[\xi_T(\widetilde{X}_{\xi_T} - \tilde{\mu})]| \le |\tilde{\mu}_3|\,\mathrm{E}\,\xi_T \sim \mu^{-1}T\,\bar{F}_{\mathrm{on}}(x_T, T) = o(MT^2[\bar{F}_{\mathrm{on}}(x_T, T)]^2),
$$
$$
\tilde{\sigma}^2\,\mathrm{E}[\xi_T(\widetilde{X}_{\xi_T} - \tilde{\mu})^2] \le \tilde{\sigma}^2\,\mathrm{E}\,\xi_T \sim \mu^{-1}T\,\bar{F}_{\mathrm{on}}(x_T, T) = o(MT^2[\bar{F}_{\mathrm{on}}(x_T, T)]^2).
$$

Finally,

$$
\tilde{\sigma}^2\left|\mathrm{E}\left[(\widetilde{X}_{\xi_T} - \tilde{\mu})\sum_{j=1}^{\xi_T}(\widetilde{X}_j - \tilde{\mu})j\right]\right|
$$

$$
\le \tilde{\sigma}^2\,\mathrm{E}\left[\sum_{j=1}^{\xi_T}|\widetilde{X}_j - \tilde{\mu}|j\right]
$$

$$
= \tilde{\sigma}^2\,\mathrm{E}\left[\sum_{j=1}^{\xi_T}(|\widetilde{X}_j - \tilde{\mu}|j - \mathrm{E}|\tilde{X} - \tilde{\mu}|j)\right] + \tilde{\sigma}^2\,\mathrm{E}\left[\sum_{j=1}^{\xi_T}j\right]\mathrm{E}|\tilde{X} - \tilde{\mu}|
$$

$$
\le 2\tilde{\sigma}^2\,\mathrm{E}\,\xi_T^2\,\bar{F}_{\mathrm{on}}(x_T, T) \sim 2\mu^{-2}T^2[\bar{F}_{\mathrm{on}}(x_T, T)]^2
$$

$$
= o(MT^2[\bar{F}_{\mathrm{on}}(x_T, T)]^2).
$$

In the third step, we again used the fact that ξ_T is a stopping time. This completes the proof.

Appendix A. Convergence to a simple point process

Let $M_p(E)$ denote the space of all point measures defined on the state space $E := [0, \infty) \times (0, \infty]$. The space $M_p(E)$ is equipped with the vague topology (see [14, Section 3.4]). A point process N on E is called *simple* if

$$
\mathrm{P}(N(\{x\}) \le 1 \text{ for all } x \in E) = 1,
$$

that is, if there are no points coinciding, with probability 1. Notice that a Poisson random measure N with mean measure ν is simple if ν is atomless.

Let $\mathcal{I}$ be the class of rectangles $[a, b) \times (c, d] \subset E$ and let $\mathcal{A}$ be the collection of sets

$$A = \bigcup_{j=1}^{k} I_j, \qquad I_j \in \mathcal{I}, \ j = 1, \ldots, k, \ k \geq 1.$$

Observe that we may assume that the sets $[a_j, b_j) \times (c_j, d_j]$ are mutually disjoint. If two such sets intersect, they are the union of at most three mutually disjoint sets of the form $[a, b) \times (c, d]$. Notice that $\mathcal{A}$ is closed under finite unions and intersections.

It can be seen that for any compact $K \subset E$ and open $G \subset E$ with $K \subset G$, there exists an $A \in \mathcal{A}$ such that $K \subset A \subset G$. In [9] a class $\mathcal{A}$ with this property is called a *separating class*. Evidently, finite unions of elements in $\mathcal{I}$ constitute a separating class. Any such class $\mathcal{I}$ is called a *preseparating class*.

The following result is due to Kallenberg [9] and is an improved version of Theorem 4.7 of [8].

Theorem A.1. *Suppose that N and N_T, $T > 0$, are point processes on E and N is simple. If, for all $A \in \mathcal{A}$ and $I \in \mathcal{I}$,*

$$\lim_{T \to \infty} P(N_T(A) = 0) = P(N(A) = 0) \quad and \quad \limsup_{T \to \infty} P(N_T(I) > 1) \leq P(N(I) > 1),$$

then

$$N_T \xrightarrow{D} N \quad in \ M_p(E).$$

Next we state a result on the convergence of a triangular array of point processes to a PRM. Let M_T be a nondecreasing integer-valued function such that $M_T \to \infty$ as $T \to \infty$. Let

$$(N_T^{(m)}, \ m = 1, \ldots, M_T, T > 0)$$

be a triangular array of point processes on E, such that for each T the processes $(N_T^{(m)}, \ m = 1, \ldots, M_T)$ are mutually independent. The array is *uniformly asymptotically negligible* if

$$\lim_{T \to \infty} \sup_{m=1,\ldots,M_T} P(N_T^{(m)}(A) > 0) = 0,$$

for all sets $A \in \mathcal{A}$. Define the row sums

$$N_T = \sum_{m=1}^{M_T} N_T^{(m)}, \qquad T > 0.$$

The following result can be found in [3, Theorem 9.2.V].

Theorem A.2. *Let N be a simple PRM with mean measure ν. If the triangular array $(N_T^{(m)})$ is uniformly asymptotically negligible, then*

$$N_T \xrightarrow{D} N \ in \ M_p(E)$$

if and only if, for all sets $A \in \mathcal{A}$,

$$\lim_{T \to \infty} \sum_{m=1}^{M_T} P(N_T^{(m)}(A) \geq 1) = \nu(A) \quad and \quad \lim_{T \to \infty} \sum_{m=1}^{M_T} P(N_T^{(m)}(A) \geq 2) = 0.$$

Appendix B. A property of regularly varying functions

Let $U(x)$ be a regularly varying function with index $\rho \in \mathbb{R}$, i.e. for $x > 0$

$$\lim_{t \to \infty} \frac{U(tx)}{U(t)} = x^\rho.$$

The following result can be found in [14, Proposition 0.8 (iii)].

Proposition B.1. *Let (a_n) and (a'_n) be positive sequences, converging to infinity, such that $a_n \sim a'_n c$, as $n \to \infty$, with $0 < c < \infty$. Then*

$$\lim_{n \to \infty} \frac{U(a_n)}{U(a'_n)} = c^\rho.$$

Appendix C. Replacement of the counting process ξ_T by its mean

For $1 < \alpha < 2$, the following result was proved in [11, Lemma 4]. However, a close inspection of the proof shows that the result holds for any $\alpha > 1$.

Lemma C.1. *Suppose that M satisfies the slow growth condition. Let $\epsilon_T \to 0$ such that*

$$b(MT) = o(\epsilon_T T) \quad and \quad \frac{1}{\log(T)} = o(\epsilon_T). \tag{C.1}$$

Then, for any $t \geq 0$,

$$M \, \mathrm{P}(|\xi_{Tt} - \mu_{Tt}| > \epsilon_T \mu_{Tt}) \to 0,$$

as $T \to \infty$.

Appendix D. An identity in law for stopped random sums

Here we use the notation introduced at the beginning of Section 4.1.

Lemma D.1. *Let $f : \mathbb{R} \to \mathbb{R}$ be a measurable function. Then*

$$\sum_{i=1}^{\xi_T - 1} f(\widetilde{X}_i)(\xi_T - i) \overset{\mathrm{D}}{=} \sum_{i=1}^{\xi_T - 1} f(\widetilde{X}_i)i, \tag{D.1}$$

and

$$\sum_{i=1}^{\xi_T - 1} f(\widetilde{X}_i)(\xi_T - i) \sum_{j=1}^{i-1} f(\widetilde{X}_j) \overset{\mathrm{D}}{=} \sum_{i=1}^{\xi_T - 1} f(\widetilde{X}_i) \sum_{j=1}^{i-1} f(\widetilde{X}_j)j. \tag{D.2}$$

Proof. First we prove (D.1). We have

$$p_T = \mathrm{P}\left(\sum_{i=1}^{\xi_T - 1} f(\widetilde{X}_i)(\xi_T - i) > x\right) = \sum_{n=0}^{\infty} \mathrm{P}\left(\sum_{i=1}^{\xi_T - 1} f(\widetilde{X}_i)(\xi_T - i) > x, \; \xi_T = n\right)$$

$$= \sum_{n=0}^{\infty} \mathrm{P}\left(\sum_{i=1}^{n-1} f(\widetilde{X}_i)(n - i) > x, \; S_{n-1} \leq T, \; S_n > T\right)$$

$$= \sum_{n=0}^{\infty} \mathrm{P}\left(\sum_{i=1}^{n-1} f(\widetilde{X}_{n-i})i > x, \; S_{n-1} \leq T, \; S_n > T\right).$$

Notice that the probabilities above are invariant under a permutation on $X_1, \ldots, X_{n-1}$. We apply the permutation $X_i \rightarrow X_{n-i}$, $i = 1, \ldots, n - 1$. This yields

$$p_T = \sum_{n=0}^{\infty} \mathrm{P}\left(\sum_{i=1}^{n-1} f(\widetilde{X}_i)i > x, \, D + \sum_{i=1}^{n-1}(X_{n-i} + Y_i) \leq T, \right.$$

$$\left. D + \sum_{i=1}^{n-1}(X_{n-i} + Y_i) + (X_n + Y_n) > T \right)$$

$$= \sum_{n=0}^{\infty} \mathrm{P}\left(\sum_{i=1}^{\xi_T - 1} f(\widetilde{X}_i)i > x, \, \xi_T = n \right) = \mathrm{P}\left(\sum_{i=1}^{\xi_T - 1} f(\widetilde{X}_i)i > x \right).$$

This completes the proof of (D.1). The proof of (D.2) is along the same lines and therefore omitted.

Acknowledgements

References

[1]   CROVELLA, M. E. AND BESTAVROS, A. (1996). Self-similarity in world wide web traffic: evidence and possible causes. *Perf. Evaluation Rev.* **24**, 160–169.

[2]   CROVELLA, M. E., TAQQU, M. S. AND BESTAVROS, A. (1996). Heavy-tailed probability distributions in the world wide web. In *A Practical Guide to Heavy Tails: Statistical Techniques and Applications*, eds R. Adler, R. Feldman and M. S. Taqqu, Birkhäuser, Boston, pp. 3–25.

[3]   DALEY, D. J. AND VERE-JONES, D. (1988). *An Introduction to the Theory of Point Processes.* Springer, New York.

[4]   EMBRECHTS, P., KLÜPPELBERG, C. AND MIKOSCH T. (1997). *Modelling Extremal Events.* Springer, Berlin.

[5]   FOWLER, H. J. AND LELAND, W. E. (1991). Local area network traffic characteristics with implications for broadband network congestion management. *IEEE J. Sel. Areas Commun.* **9**, 1139–1149.

[6]   GUT, A. (1988). *Stopped Random Walks: Limit Theorems and Applications.* Springer, New York.

[7]   HEATH, D., RESNICK, S. I. AND SAMORODNITSKY, G. (1998). Heavy tails and long range dependence in ON/OFF processes and associated fluid models. *Math. Operat. Res.* **23**, 145–165.

[8]   KALLENBERG, O. (1986). *Random Measures,* 4th edn. Akademie-Verlag, Berlin.

[9]   KALLENBERG, O. (1996). Improved criteria for distributional convergence of point processes. *Stoch. Process. Appl.* **64**, 93–102.

[10] LELAND, W. E., WILLINGER, W., TAQQU, M. S. AND WILSON, D. V. (1993). On the self-similar nature of Ethernet traffic. *Comput. Commun. Rev.* **23**, 183–193.

[11] MIKOSCH, T., RESNICK, S. I., ROOTZÉN, H. AND STEGEMAN, A. W. (2002). Is network traffic approximated by stable Lévy motion or fractional Brownian motion? To appear in *Ann. Appl. Prob.*

[12] PAXTON, V. AND FLOYD, S. (1995). Wide-area traffic: the failure of Poisson modeling. *IEEE/ACM Trans. Networking* **3**, 226–244.

[13] POLLARD, D. (1984). *Convergence of Stochastic Processes.* Springer, Berlin.

[14] RESNICK, S. I. (1987). *Extreme Values, Regular Variation and Point Processes.* Springer, New York.

[15] RESNICK, S. I. (1992). *Adventures in Stochastic Processes.* Birkhäuser, Boston.

[16] SAMORODNITSKY, G. AND TAQQU, M. S. (1994). *Stable Non-Gaussian Random Processes.* Chapman and Hall, New York.

[17] TAQQU, M. S., WILLINGER, W. AND SHERMAN, R. (1997). Proof of a fundamental result in self-similar traffic modeling. *Comput. Commun. Rev.* **27**, 5–23.

[18] WILLINGER, W., TAQQU, M. S., SHERMAN, R. AND WILSON, D. V. (1995). Self-similarity through high-variability: statistical analysis of Ethernet LAN traffic at the source level (extended version). *Comput. Commun. Rev.* **25**, 100–113.