

THE INFLUENCE OF DEPENDENCE ON DATA NETWORK MODELS

BERNARDO D'AURIA,* *EURANDOM*

SIDNEY I. RESNICK,** *Cornell University*

Abstract

Consider an infinite-source marked Poisson process to model end user inputs to a data network. At Poisson times, connections are initiated. The connection is characterized by a triple (F, L, R) denoting the total quantity of transmitted data in a connection, the length or duration of the connection, and the transmission rate; the three quantities are related by $F = LR$. How critical is the dependence structure of the mark for network characteristics such as burstiness, distribution tails of cumulative input, and long-range dependence properties of traffic measured in consecutive time slots? In a previous publication (D'Auria and Resnick (2006)) we assumed that F and R were independent. Here we assume that L and R are independent. The change in dependence assumptions means that the model properties change dramatically: tails of cumulative input per time slot are dramatically heavier, traffic cannot be approximated by a Gaussian distribution, and the decay of dependence cannot be measured in the traditional way using correlation functions. Different network applications are likely to have different mark dependence structure. We argue that the present independence assumption on L and R is likely to be appropriate for network applications such as streaming media or peer-to-peer networks. Our conclusion is that it is desirable to separate network traffic by application and to model each application with its own appropriate dependence structure.

Keywords: Bursty traffic; M/G/ ∞ input model; infinite-source Poisson model; network modeling; limit distribution; Lévy process; Gaussian limit

2000 Mathematics Subject Classification: Primary 90B15; 60K30

Secondary 90B22; 90B18

1. Introduction

Measurements on data networks exhibit features surprising by the standards of classical queueing and telephone network models. Data studies starting from the 1990s often consisted of analyzing bit counts or packet counts in finely resolved adjacent time windows. The time resolution was at the millisecond or even the microsecond level. Influential studies included Crovella and Bestavros (1996), (1997); Duffy *et al.* (1993); Leland *et al.* (1993); and Willinger *et al.* (1995), (1997).

Data studies have shown that the network traffic typically has important features termed *invariants* or *stylized facts*. These striking features include the following.

- Heavy tails abound (Leland *et al.* (1994); Willinger and Paxson (1998); Willinger *et al.* (1998); Willinger (1998)) for such things as file sizes (Arlitt and Williamson (1996);

Received 4 September 2006; revision received 14 January 2008.

* Current address: Departamento Estadística, Universidad Carlos III de Madrid, Avda. de la Universidad, 30, 28911 Leganés, Madrid, Spain.

** Postal address: School of Operations Research and Information Engineering, Cornell University, Ithaca, NY 14853, USA. Email address: sir1@cornell.edu

Resnick and Rootzén (2000)), transmission rates, transmission durations (Maulik *et al.* (2002); Resnick (2003)), or connection durations.

- The number of bits or packets per time slot exhibits long-range dependence across the time slots (see, for example, Leland *et al.* (1993) and Willinger *et al.* (1995)). There is also a perception of self-similarity as the width of the time slot varies across a range of time scales exceeding a typical round trip time.
- Network traffic is bursty with rare but influential periods of very high transmission rates punctuating typical periods of modest activity. See, for example, Sarvotham *et al.* (2005).

Many models have been proposed to explain empirically observed characteristics in collected network data, although these tended to concentrate on large time scales and cumulative traffic distributional approximations over large time intervals. See, for example, Heath *et al.* (1998); Kaj and Taqqu (2004); Konstantopoulos and Lin (1998); Levy and Taqqu (2000); Maulik and Resnick (2003); Milosch *et al.* (2002); and Taqqu *et al.* (1997). Some of these models attempted to reproduce the physical dynamics behind the measured data while others just tried to match statistical characteristics. Some of the models assumed constant rate user inputs. For these large time scale cumulative input models, it is difficult to find agreement with existing data sets (Guerin *et al.* (2003)).

Since, historically, data collection was over finely resolved time intervals, it is sensible to model cumulative user inputs over adjacent, small time slots. D' Auria and Resnick (2006) began an attempt to explain invariant data features using an infinite-source marked Poisson model in which it was assumed that input traffic consisted of connections initiated at Poisson times and that each connection had an associated random rate and file size which were independent. Our current paper complements the analysis in D' Auria and Resnick (2006) by examining this model with a different assumption on the distributional structure of the sessions. If the dependence structure of the mark is changed so that we assume that the connection duration and the input rate are independent heavy-tailed random variables, how different are the model predictions? The answer is that the model properties change markedly. Tails of cumulative input per time slot become much heavier, cumulative traffic per time slot is no longer Gaussian distributed, and long-range dependence properties can no longer be described in the classical way with correlations.

There are valid statistical reasons for considering the alternative assumption that the duration and rate are independent. See Section 2 and Resnick (2007, Chapter 7). We also suggest in Section 2 that certain Internet applications such as streaming media and peer-to-peer networks may produce data with these dependence properties. One conclusion, explained at greater length in Section 7, is that heterogeneous traffic comprising different types of applications will behave differently to more homogeneous traffic. Applications which require the modeling assumption that the session length and rate are independent will tend to look more bursty, less Gaussian, and have cumulative traffic which is heavier tailed. Applications allowing the assumption that the file size and rate are independent will look more Gaussian, lighter tailed, and less bursty. This suggests that empirical studies of network traffic in the wild should decompose traffic into classes of fairly homogeneous applications and that each class should be studied separately. We can then seek statistical differences in characteristics and compare different applications.

The model we review in Section 2 studies small-scale asymptotic behavior, which means that we measure the amount of data content that arrives in consecutive time slots of length δ . We feel that this approach is more faithful to actual data experiments than the approach of studying large time scaling models that have received the majority attention so far. We study the limit

distribution of content per slot as δ tends to 0 after a centering and scaling. Letting δ tend to 0 allows us to clarify the distributional structure of the stochastic process of cumulative inputs in successive time slots. Our results indicate that, for network traffic satisfying the condition that the connection duration and rate are independent and heavy tailed, cumulative input in successive time slots will be highly dependent and that each time slot has content distributed approximately as a heavy-tailed stable distribution.

Section 2 contains more details on the description of the model as well as a fuller discussion of scenarios where our independence assumption is appropriate. It also presents analysis of the heavy-tail weight for cumulative content per slot and gives a symbol and concept summary list which may be referenced when reading the rest of the paper. In Section 3 we derive the approximating stable distribution of cumulative input per time slot, while in Sections 4 and 5 we describe the dependence structure across time intervals. We provide detailed comparisons in Section 6 between the present model and the one in D'Auria and Resnick (2006), and highlight and review the impact that differing dependence assumptions can have on tail heaviness. In Section 7 we conclude the paper with some final thoughts including the suggestion that it is necessary to statistically characterize different network applications and understand their statistical differences.

2. Model description

The model for data traffic generation is a modification of the $M/G/\infty$ input or infinite-source Poisson model, where the transmission rates are assumed to be random as well. We assume that a homogeneous Poisson process on \mathbb{R} with points $\{\Gamma_k\}$ activates data transmission *connections* or *sessions*. The parameter or rate of the Poisson process is $\lambda = \lambda(\delta)$, and to each transmission activation time Γ_k is associated a mark consisting of a triple (R_k, L_k, F_k) . These three quantities have the following physical interpretations:

- R denotes the rate of the transmission assuming a constant rate across the connection interval;
- L denotes the duration of the transmission or length of the connection;
- F denotes the size of the transmitted file or quantity of transmitted data in a connection.

These three quantities are related by the relation $F = RL$.

We assume that the marks $\{(R_k, L_k, F_k), -\infty < k < \infty\}$ are independent and identically distributed (i.i.d.) and independent of $\{\Gamma_k\}$, and that R_k and L_k are independent for each k . The univariate marginal distributions of the triple are

$$G(x) = \mathbb{P}[F_1 \leq x], \quad F_R(x) = \mathbb{P}[R_1 \leq x], \quad F_L(x) = \mathbb{P}[L_1 \leq x].$$

We suppose that all three distributions are heavy tailed:

$$\bar{G}(x) = x^{-\alpha_F} \ell_F(x), \quad \bar{F}_R(x) = x^{-\alpha_R} \ell_R(x), \quad \bar{F}_L(x) = x^{-\alpha_L} \ell_L(x),$$

where ℓ_F, ℓ_R , and ℓ_L are all slowly varying, and we assume that all three tail parameters satisfy

$$1 < \alpha_F, \alpha_R, \alpha_L < 2.$$

There is empirical and engineering evidence justifying these assumptions and, consequently, we have given priority to models with parameters in this range. See Azzouna *et al.* (2004);

Guerin *et al.* (2003); Heffernan and Resnick (2005); Hernández-Campos *et al.* (2005); Leland *et al.* (1994); Maulik *et al.* (2002); Park and Willinger (2000); Resnick (2003), (2004a); Reidi and Willinger (2000); Sarvotham *et al.* (2005); and Willinger *et al.* (1995).

With these assumptions, the counting function of the points $\{(\Gamma_k, R_k, L_k, F_k)\}$,

$$N = \sum_k \varepsilon_{(\Gamma_k, R_k, L_k, F_k)}$$

on $\mathbb{R} \times [0, \infty)^3$, is a *Poisson random measure* with mean measure

$$\lambda \, ds \, \mathbb{P}[(R_1, L_1, F_1) \in (dr, dl, du)] =: \mu^\#(ds, dr, dl, du).$$

See, for example, Kallenberg (1983); Neveu (1977); and Resnick (2007, p. 122), (1987, Section 3.3.2), (1992). The joint distribution of (R_1, L_1, F_1) is computed using the assumption that R_1 and L_1 are independent and that $F_1 = R_1 L_1$.

For a time window of length δ , we will consider weak limits of the process

$$A(\delta) := \{A(k\delta, (k + 1)\delta), -\infty < k < \infty\} \quad \text{as } \delta \downarrow 0. \tag{1}$$

Here $A(k\delta, (k + 1)\delta]$ represents the total amount of work inputted to the system in the k th time slot $(k\delta, (k + 1)\delta]$. We will define this precisely for $k = 0$, and the definitions for the other values of k will be obvious by analogy.

Distinguish four disjoint regions in $\mathbb{R} \times [0, \infty)^3$ by a decomposition on the arrival time of a session and its duration:

$$\begin{aligned} \{> 0, 1\} &= \{(s, r, l, u) : 0 < s \leq \delta, 0 < s + l \leq \delta\}, \\ \{> 0, 2\} &= \{(s, r, l, u) : 0 < s \leq \delta, s + l > \delta\}, \\ \{< 0, 1\} &= \{(s, r, l, u) : s < 0, 0 < s + l \leq \delta\}, \\ \{< 0, 2\} &= \{(s, r, l, u) : s < 0, s + l > \delta\}. \end{aligned}$$

Region $\{> 0, 1\}$ corresponds to sessions starting and ending in $(0, \delta]$, while region $\{> 0, 2\}$ describes sessions starting in $(0, \delta]$ but ending subsequent to δ . Region $\{< 0, 1\}$ has sessions starting prior to time 0 and ending in $(0, \delta]$, while region $\{< 0, 2\}$ has sessions initiated prior to 0 and ending subsequent to δ . See Figure 1, where the horizontal axis is the initiation time of a session and the vertical axis is session duration.

Corresponding to this decomposition of regions, if we restrict the Poisson random measure to the four regions, we obtain four independent Poisson processes:

$$N(\cdot \cap \{> 0, 1\}), \quad N(\cdot \cap \{> 0, 2\}), \quad N(\cdot \cap \{< 0, 1\}), \quad N(\cdot \cap \{< 0, 2\}), \tag{2}$$

and we use these to express $A(0, \delta) =: A(\delta)$ as the sum of four independent contributions:

$$A(\delta) = A^{>0,1}(\delta) + A^{>0,2}(\delta) + A^{<0,1}(\delta) + A^{<0,2}(\delta), \tag{3}$$

where

$$\begin{aligned} A^{>0,1}(\delta) &= \sum_k R_k L_k \mathbf{1}_{[(\Gamma_k, R_k, L_k, F_k) \in \{>0,1\}]}, \\ A^{>0,2}(\delta) &= \sum_k R_k (\delta - \Gamma_k) \mathbf{1}_{[(\Gamma_k, R_k, L_k, F_k) \in \{>0,2\}]}, \end{aligned}$$

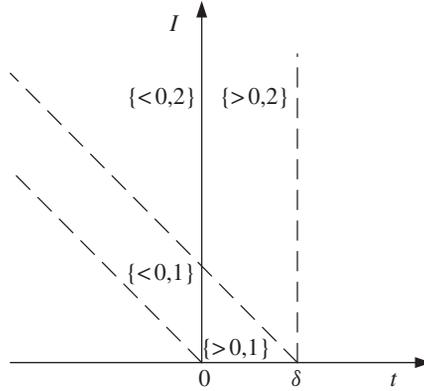


FIGURE 1: Four regions.

$$A^{<0,1}(\delta) = \sum_k R_k(L_k + \Gamma_k) \mathbf{1}_{\{(\Gamma_k, R_k, L_k, F_k) \in \{<0,1\}\}},$$

$$A^{<0,2}(\delta) = \sum_k R_k \delta \mathbf{1}_{\{(\Gamma_k, R_k, L_k, F_k) \in \{<0,2\}\}}.$$

As a further notational device, we will adopt the convention that, for a region \mathcal{R} of the (s, r, l, u) -space, $A^{\mathcal{R}}(t_1, t_2]$ will denote the cumulative work inputted to the system in times $(t_1, t_2]$ from points $(\Gamma_k, R_k, L_k, F_k)$ in region \mathcal{R} .

We can represent the restrictions of N to each of the four regions given in (2) as empirical measures of a Poisson number of i.i.d. points whose joint distributions are the mean measure $\mu^\#$ restricted to that region and normalized to be a probability measure. (See, for instance, Resnick (1992, p. 341)). For example,

$$N(\cdot \cap \{> 0, 1\}) = \sum_{k=1}^{P^{>0,1}(\delta)} \mathcal{E}_{(\Gamma_k^{>0,1}, R_k^{>0,1}, L_k^{>0,1}, F_k^{>0,1})},$$

where $P^{>0,1}(\delta)$ is Poisson with parameter

$$\begin{aligned} \mu^\#(\{> 0, 1\}) &= \int_{\{>0,1\}} \lambda \, ds \, \mathbb{P}[(R_1, L_1, F_1) \in (dr, dl, du)] \\ &= \int_0^\delta \lambda \, ds \, \mathbb{P}[L_1 + s < \delta] \\ &= \int_0^\delta \lambda F_L(\delta - s) \, ds \\ &= \delta \hat{F}_L(\delta) \end{aligned}$$

(where $\hat{F}_L(x) = \int_0^x F_L(y) \, dy$), and $\{(\Gamma_k^{>0,1}, R_k^{>0,1}, L_k^{>0,1}, F_k^{>0,1})\}$ are i.i.d. with joint distribution

$$\frac{\mu^\#(\cdot \cap \{> 0, 1\})}{\mu^\#(\{> 0, 1\})}.$$

In what follows we sometimes use the convention that $P^{\mathcal{R}}(\delta)$ is Poisson distributed with parameter equal to $\mu^\#(\mathcal{R})$, the mean measure of the region \mathcal{R} .

2.1. Specifying the dependence structure for (R, L, F)

We emphasize that one of the points of this paper is that different dependence structures for the triple (R, L, F) induce noticeably different behavior for $A(\delta)$ in (1). In D'Auria and Resnick (2006) we studied the RF case, where the random variables R and F were independent (cf. also Hernández-Campos *et al.* (2005)). In this paper we focus on what we call the RL model, where the random variables R and L are independent (cf. also Maulik *et al.* (2002)).

Modern network traffic is the superposition of heterogeneous applications. The assumption of independence of the transmission rates from the transmission durations is natural for some applications. Here we describe two possible situations.

The first scenario considers the transmission of streaming flows such as media streams (e.g. video-on-demand). Usually this kind of data is transmitted in real time. This means that the transmission durations approximately coincide with the time length of the data content. For example, if we consider the transmission of a music song, such as by an Internet radio broadcast, the transmission will last as long as the song duration. In some cases, like watching a movie on the Internet from a video-on-demand service, the transmission duration does not exactly coincide with the actual duration of the movie, owing to buffering in the receiver necessitated by the need to prevent bad quality play owing to high jitter in the transmission. Generally, since the required data for the movie greatly exceeds the buffer size, we can neglect the influence of the latter on the duration of the transmission. Therefore, if we consider the rate at which these transmissions take place, they depend on the sampling quality of the media stream and, hence, not on its duration. Returning to the Internet radio example, the user can usually choose a download quality depending on the bandwidth of the Internet access. This choice affects the quality of the content, since it affects the total amount of data that will be transmitted, but it will not alter the duration of the transmission, which as previously noted, will depend on the duration of the content and/or the time the user wants to be connected to listen or watch. In the literature these kinds of transmissions where the rate of the transmission depends on the sampling rate of the content, as well as on the compression scheme, are known as VBR transmissions (Heyman and Lakshman (1996); Park and Willinger (2000)). VBR stands for variable bit rate since usually the compression scheme implies nonconstant transmission rates. By neglecting this feature and assuming for simplicity that the transmission rate stays constant over all the data transmission, we can consider the RL model to be a reasonable model for streaming-flow transmissions.

The second scenario is peer-to-peer (P2P) networks (see Pandurangan *et al.* (2001) and Tanenbaum (1996)). A typical P2P network is composed of a collection of users that are simultaneously online and sharing resources. Usually users dynamically connect to and disconnect from the P2P network so that the size and type of data the network holds change continuously in time. Viewed from the point of view of one particular user, the network contains content that is always changing, and files the user wishes to download are alternatively present and absent. Often it is the case that P2P users open two different communication channels, one for uploading and one for downloading, and for each channel they specify the maximum allowed transmission bandwidth. Usually this specification is done only once at the beginning of the connection and is never modified. Therefore, since the choice of the maximum bandwidth does not depend on the subsequent transferred data, it seems a natural assumption to consider the chosen maximum allowed rates to be independent of the content and duration of the data transmissions. In addition, owing to the large size and high fluctuations of the population comprising the P2P network, it often happens that the upload channel is always fully utilized while the utilization of the download channel is always fluctuating depending on the state

of the network and, therefore, the availability of the desired content. This means that the download rates are fluctuating while the upload rates are constant and equal to the maximum allowed upload rates. Now assuming that the users connect to the P2P network according to a Poisson process, we can associate them with the sources of our infinite-source Poisson model. According to this association, we can consider a source transmission as the total data transfer that one user has transmitted by the upload channel. Therefore, the transmission durations L will be given by the lifetime of a user on the P2P network, while the transmission rates R will be given by the maximum upload bandwidth. In this setting it seems natural to assume that R and L are independent, so that the RL model is appropriate.

Undoubtedly, in practice, it may not be true that R and L are actually independent but rather satisfy some form of asymptotic independence (see, for example, Maulik *et al.* (2002)). However, assuming asymptotic independence rather than full independence would lead to unacceptable complications in the analysis and proofs without changing conclusions and, thus, at this stage, choosing full independence of L and R seems an appropriate modeling assumption.

2.2. The RL model

We assume that the rates of transmissions are independent of the transmission durations. The file sizes are computed by the relation $F = LR$. From Breiman’s theorem (Breiman (1965)), this means that the distribution tail of the random variable F is given by

$$\bar{G}(u) \sim \begin{cases} E(R^{\alpha_L})\bar{F}_L(u) & \text{if } \alpha_R > \alpha_L, \\ E(L^{\alpha_R})\bar{F}_R(u) & \text{if } \alpha_R < \alpha_L. \end{cases}$$

The case in which $\alpha_R = \alpha_L$ is of somewhat less interest in applied probability. This case could be handled by a refinement of Breiman’s theorem which proceeds under the condition that $P[R > x] = o(P[L > x])$ (or vice versa); this result is given in Embrechts and Goldie (1980). A product result in Cline (1983), quoted in Davis and Resnick (1986, p. 542), of a slightly different character describes the case in which $R \stackrel{D}{=} L$ (where ‘ $\stackrel{D}{=}$ ’ denotes equality in distribution), R and L are independent, $P[R > x]$ is regularly varying with index $-\alpha$, and $E(R^\alpha) = \infty$.

By using the property that the random variables R and L are heavy tailed, we derive the tail behavior of the random variables $A^{\mathcal{R}}(\delta)$ with $\mathcal{R} \in \{ \{< 0, 1\}; \{< 0, 2\}; \{> 0, 1\}; \{> 0, 2\} \}$; that is, \mathcal{R} is one of the four regions shown in Figure 1. For a fixed $\delta > 0$, the tails satisfy, as $x \rightarrow \infty$,

$$\frac{P[A^{\mathcal{R}}(\delta) > x]}{\bar{F}_R(x)} \sim \begin{cases} \lambda \int_{s=0}^{\delta} \int_{l=0}^s l^{\alpha_R} F_L(dl) ds, & \mathcal{R} = \{> 0, 1\}, \\ \lambda \int_0^{\delta} s^{\alpha_R} \bar{F}_L(s) ds, & \mathcal{R} = \{> 0, 2\} \text{ or } \{< 0, 1\}, \\ \lambda E(L) \bar{F}_L^{(0)}(\delta) \delta^{\alpha_R}, & \mathcal{R} = \{< 0, 2\}. \end{cases}$$

The tails of all the regions are regularly varying with index $-\alpha_R$. In particular, for fixed $\delta > 0$, as $x \rightarrow \infty$,

$$P[A(0, \delta) > x] \sim (\text{constant})\bar{F}_R(x). \tag{4}$$

We give a sample calculation which explains how to obtain the last relation about the tails of $A^{\mathcal{R}}$ for the case in which $\mathcal{R} = \{> 0, 1\}$. The methodology is similar to the one used in

Section 3. We have

$$\begin{aligned} \frac{P[A^{>0,1}(\delta) > x]}{\bar{F}_R(x)} &\sim \frac{E(P^{>0,1}(\delta)) P[F_0^{>0,1} > x]}{\bar{F}_R(x)} \\ &= \frac{1}{\bar{F}_R(x)} \int_{0 < s < \delta} \int_{l < \delta - s} \int_{rl > x} \lambda \, ds F_R(dr) F_L(dl) \\ &= \lambda \int_{s=0}^{\delta} \int_{l=0}^{\delta-s} \frac{\bar{F}_R(l^{-1}x)}{\bar{F}_R(x)} F_L(dl) \, ds \\ &\rightarrow \lambda \int_{s=0}^{\delta} \int_{l=0}^s l^{\alpha_R} F_L(dl) \, ds \quad \text{as } x \rightarrow \infty. \end{aligned}$$

Since our limiting procedure will shrink the observation window $(0, \delta]$, there is no hope of obtaining a weak limit in (1) unless we increase the arrival rate $\lambda = \lambda(\delta)$ of the sessions. We adopt a heavy-traffic limit theorem philosophy and imagine moving through a family of models indexed by δ as $\delta \downarrow 0$. A convenient and effective choice of λ is

$$\lambda(\delta) = \frac{1}{\delta \bar{F}_R(\delta^{-1})}. \tag{5}$$

Using assumption (5), the behavior of the random variables $A^{(\cdot)}(\delta)$ is as follows.

- $A^{<0,1}(\delta)$, suitably centered, converges weakly to a stable random variable $X_{\alpha_R}^{<0,1}$ with infinite second moment and index $\alpha_R \in (1, 2)$.
- $A^{<0,2}(\delta)$ does not converge weakly without scaling; with centering and scaling it converges to a stable random variable $X_{\alpha_R}^{<0,2}$ with index α_R . We also note that if we suitably decompose region $\{< 0, 2\}$ into two subregions, according to whether the transmission rate is small or large, we can have convergence to a Gaussian random variable in the region where the rate is small, but the required scaling is of smaller order compared with the scaling yielding $X_{\alpha_R}^{<0,2}$.
- $A^{>0,1}(\delta)$ is negligible in the limit under suitable conditions.
- $A^{>0,2}(\delta)$ is equal in distribution to $A^{<0,1}(\delta)$.

2.3. Symbol finder

For convenience and reference, we list some symbols and concepts frequently used.

F_R, F_L, G	The distributions of rate, duration, and file size, respectively.
\bar{F}	For a distribution function $F(x)$, $\bar{F} = 1 - F$ is the distribution tail.
$F^{(0)}$	For a distribution F with finite mean μ , $F^{(0)}(x) = \int_0^x \mu^{-1} \bar{F}(s) \, ds$.
$RV_{-\alpha}$	The class of regularly varying functions with index $-\alpha$.
$l(x)$	A slowly varying function: $\lim_{t \rightarrow \infty} l(tx)/l(t) = 1$ for $x > 0$.
$\varepsilon_x(\cdot)$	The point probability measure putting all mass at the point x .
RL	The model where $R \perp\!\!\!\perp L$; that is, R and L are independent.

RF	The model where $R \perp\!\!\!\perp F$; that is, R and F are independent.
δ	The time slot width.
λ	The Poisson rate of connection arrivals $\lambda = \lambda(\delta)$.
$\mu^\#(ds, dr, dl, du)$	$\lambda ds P[(R_1, L_1, F_1) \in (dr, dl, du)]$.
$P^{\mathcal{R}}(\delta)$	A Poisson random variable with mean $\mu^\#(\mathcal{R})$.
$A^{\mathcal{R}}(I)$	Cumulative work inputted into the system in time interval I by points with characteristics in \mathcal{R} .
$\mu_\delta(dr)$	$F_R(\delta^{-1} dr) / \bar{F}_R(\delta^{-1})$.

3. Limits for the cumulative input $A(\delta)$

Here we analyze the cumulative input in $[0, \delta]$ by analyzing the four pieces separately in the decomposition (3).

3.1. Region $\{> 0, 2\}$

Recall that this is the region contributing input in $(0, \delta]$ from sessions initiated in $(0, \delta]$ but terminating after δ .

3.1.1. *Characteristic function.* For $\theta \in \mathbb{R}$, we compute

$$\begin{aligned}
 & E(\exp(i\theta A^{>0,2}(\delta))) \\
 &= E \exp\left(i\theta \sum_{j=1}^{P^{>0,2}(\delta)} R_j^{>0,2}(\delta - \Gamma_j^{>0,2})\right) \\
 &= \exp(E(P^{>0,2}(\delta)) E(\exp(i\theta R_1^{>0,2}(\delta - \Gamma_1^{>0,2})) - 1)) \\
 &= \exp\left(\int_{0 < s < \delta} \int_{s+l > \delta} \int_{r > 0} (e^{i\theta r(\delta-s)} - 1) \lambda ds F_R(dr) F_L(dl)\right) \\
 &= \exp\left(\int_{0 < s < \delta} \int_{l > s} \int_{r > 0} (e^{i\theta rs} - 1) \lambda ds F_R(dr) F_L(dl)\right) \\
 &= \exp\left(\int_{s=0}^{\delta} \int_{r=0}^{\infty} (e^{i\theta rs} - 1) \bar{F}_L(s) F_R(dr) \lambda ds\right) \\
 &= \exp\left(\int_{r=0}^{\infty} \int_{s=0}^{r\delta} (e^{i\theta s} - 1) \bar{F}_L(r^{-1}s) \lambda r^{-1} ds F_R(dr)\right) \\
 &= \exp\left(\int_{s=0}^{\infty} \int_{r=\delta^{-1}s}^{\infty} (e^{i\theta s} - 1) \bar{F}_L(r^{-1}s) \lambda r^{-1} F_R(dr) ds\right) \\
 &= \exp\left(\lambda \delta \bar{F}_R(\delta^{-1}) \int_{s=0}^{\infty} (e^{i\theta s} - 1) \int_{r=s}^{\infty} \bar{F}_L(\delta r^{-1}s) r^{-1} \frac{F_R(\delta^{-1} dr)}{\bar{F}_R(\delta^{-1})} ds\right),
 \end{aligned}$$

and, finally, using $\lambda = 1/\delta \bar{F}_R(\delta^{-1})$, we obtain

$$E(\exp(i\theta A^{>0,2}(\delta))) = \exp\left(\int_{s=0}^{\infty} (e^{i\theta s} - 1) \int_{r=s}^{\infty} \bar{F}_L(\delta r^{-1}s) r^{-1} \mu_\delta(dr) ds\right).$$

Writing

$$\nu_\delta^{>0,2}(ds) = (\nu_\delta^{>0,2})'(s) ds = \int_{r=s}^\infty \bar{F}_L(\delta r^{-1}s)r^{-1}\mu_\delta(dr) ds,$$

we obtain

$$E(\exp(i\theta A^{>0,2}(\delta))) = \exp\left(\int_{s=0}^\infty (e^{i\theta s} - 1)\nu_\delta^{>0,2}(ds)\right). \tag{6}$$

3.1.2. *Properties of $\nu_\delta^{>0,2}$.*

Proposition 1. *As $\delta \rightarrow 0$,*

$$\nu_\delta^{>0,2} \xrightarrow{v} \nu_0^{>0,2}$$

on $(0, \infty]$; that is, we have vague convergence to the limit measure $\nu_0^{>0,2}$, which is a Lévy measure with density

$$\frac{\alpha_R}{1 + \alpha_R} x^{-\alpha_R-1}.$$

Proof. The proof is similar to Proposition 1 of D’Auria and Resnick (2006). Observe that, for $s \geq 0$,

$$(\nu_\delta^{>0,2})'(s) = \int_{r=s}^\infty r^{-1}\bar{F}_L(\delta r^{-1}s)\mu_\delta(dr) \leq s^{-1}\mu_\delta(s, \infty],$$

and by Potter’s bounds (Bingham *et al.* (1987, p. 25); De Haan (1970); Resnick (1987, p. 23); Seneta (1976)), for some small η , all $s \geq 1$, some $c > 0$, and for all sufficiently small δ , we have an upper bound

$$(\nu_\delta^{>0,2})'(s) \leq cs^{-(\alpha_R-\eta)-1},$$

which is integrable with respect to the Lebesgue measure on any neighborhood of ∞ . Hence, by dominated convergence, for $x > 0$,

$$\begin{aligned} \nu_\delta^{>0,2}(x, \infty] &= \int_x^\infty (\nu_\delta^{>0,2})'(s) ds \\ &\rightarrow \int_x^\infty \int_s^\infty r^{-1}\alpha_R r^{-\alpha_R-1} dr ds \\ &= \nu_0^{>0,2}(x, \infty] \\ &= \frac{\alpha_R}{1 + \alpha_R} \int_x^\infty s^{-\alpha_R-1} ds \\ &= \frac{x^{-\alpha_R}}{1 + \alpha_R}. \end{aligned} \tag{7}$$

To check that $\nu_0^{>0,2}$ is a Lévy measure, note that

$$\int_0^1 s^2 \nu_0^{>0,2}(ds) = \frac{\alpha_R}{1 + \alpha_R} \int_0^1 s^2 s^{-\alpha_R-1} ds < \infty,$$

since $1 < \alpha_R < 2$.

3.1.3. *Weak limit for $A^{>0,2}(\delta)$.* Now we use (6) and write

$$\begin{aligned} & E \exp\left(i\theta\left(A^{>0,2}(\delta) - \int_0^1 s v_\delta^{>0,2}(ds)\right)\right) \\ &= \exp\left(\int_1^\infty (e^{i\theta s} - 1) v_\delta^{>0,2}(ds) + \int_0^1 (e^{i\theta s} - 1 - i\theta s) v_\delta^{>0,2}(ds)\right). \end{aligned} \tag{8}$$

Each of the two integrals on the right-hand side of (8) converge when δ tends to 0.

Proposition 2. *As $\delta \rightarrow 0$,*

$$\int_1^\infty (e^{i\theta s} - 1) v_\delta^{>0,2}(ds) \rightarrow \int_1^\infty (e^{i\theta s} - 1) v_0^{>0,2}(ds), \tag{9}$$

$$\int_0^1 (e^{i\theta s} - 1 - i\theta s) v_\delta^{>0,2}(ds) \rightarrow \int_0^1 (e^{i\theta s} - 1 - i\theta s) v_0^{>0,2}(ds). \tag{10}$$

Therefore, as $\delta \rightarrow 0$,

$$A^{>0,2}(\delta) - \int_0^1 s v_\delta^{>0,2}(ds) \Rightarrow X_{\alpha_R}^{>0,2},$$

where the limit random variable is spectrally positive stable with index α_R , Lévy measure $v_0^{>0,2}$ given in (7), and characteristic function given by the right-hand side of (8) with $v_\delta^{>0,2}$ replaced by $v_0^{>0,2}$. Here ‘ \Rightarrow ’ denotes weak convergence or convergence in distribution

Proof. The convergence in (9) follows from standard weak convergence since the integrand is bounded and continuous and

$$\frac{v_\delta^{>0,2}(\cdot)}{v_\delta^{>0,2}(1, \infty]} \Rightarrow \frac{v_0^{>0,2}(\cdot)}{v_0^{>0,2}(1, \infty]}$$

weakly as probability measures on $(1, \infty]$.

For the proof of (10), observe that

$$|e^{i\theta s} - 1 - i\theta s| (v_\delta^{>0,2})'(s) \leq \frac{\theta^2 s^2}{2} s^{-1} \mu_\delta(s, \infty] \leq cs \frac{\bar{F}_R(\delta^{-1}s)}{\bar{F}_R(\delta^{-1})} = c \frac{V(\delta^{-1}s)}{V(\delta^{-1})},$$

where $V(s) = s \bar{F}_R(s)$ is regularly varying with index $-\alpha_R + 1$. Now, as $\delta \rightarrow 0$,

$$|e^{i\theta s} - 1 - i\theta s| (v_\delta^{>0,2})'(s) \rightarrow |e^{i\theta s} - 1 - i\theta s| (v_0^{>0,2})'(s),$$

and

$$\frac{V(\delta^{-1}s)}{V(\delta^{-1})} \rightarrow s^{-\alpha_R+1}.$$

Furthermore, by Karamata’s theorem,

$$\int_0^1 \frac{V(\delta^{-1}s)}{V(\delta^{-1})} ds \rightarrow \int_0^1 s^{-\alpha_R+1} ds = \frac{1}{2 - \alpha_R}.$$

The desired result now follows from Pratt’s lemma (see Pratt (1960) or Resnick (1999, p. 164)), since Pratt’s lemma may be applied to both the real and imaginary parts of

$$(e^{i\theta s} - 1 - i\theta s) (v_\delta^{>0,2})'(s),$$

to obtain convergence to the limit after integrating over $[0, 1]$.

3.2. Region $\{< 0, 1\}$

In this region the contribution to $A(0, \delta]$ is given by the sessions initiated before 0 and terminating in $(0, \delta]$. In D’Auria and Resnick (2006, Proposition 6) it was proved that this contribution is identical in distribution to the one of region $\{> 0, 2\}$. So we have the following result.

Proposition 3. *We have*

$$A^{<0,1}(\delta) \stackrel{D}{=} A^{>0,2}(\delta),$$

and, therefore, as $\delta \rightarrow 0$,

$$A^{<0,1}(\delta) - \int_0^1 s v_\delta^{<0,1}(ds) \Rightarrow X_{\alpha_R}^{<0,1},$$

where $v_\delta^{<0,1} = v_\delta^{>0,2}$ and $X_{\alpha_R}^{<0,1} \stackrel{D}{=} X_{\alpha_R}^{>0,2}$, with the quantities indexed by ‘ $>0,2$ ’ defined as in Proposition 2.

3.3. Region $\{> 0, 1\}$

The region $\{> 0, 1\}$ is a small region decreasing in size with δ and should not make a contribution to the overall traffic. Reasonable conditions which assure this include regular variation of $F_L(x)$ at 0. This includes distributions $F_L(x)$ satisfying

$$F_L(x) \sim x^\beta, \quad x \downarrow 0, \beta > 0,$$

and, hence, distributions F_L with densities $F'_L(x)$ satisfying

$$F'_L(x) \sim cx^{\beta-1}, \quad x \downarrow 0, \beta > 0.$$

Densities which look like gamma densities near 0 are appropriate. In particular, our assumptions allow $\beta = 1$, as is the case for the standard exponential density.

We start analysis of this region by observing that, without loss of generality, we can assume that $F_L(0) = 0$. Indeed if $F_L(0) = a$ with $0 < a < 1$, we can define $F_L^\#(x) = (F_L(x) - a)/(1 - a)$ and decompose the Poisson arrival process into two processes: the first with arrival rate $(1 - a)\lambda$ and the second with arrival rate $a\lambda$. The first arrival process has associated sessions whose lengths are always different from 0, and the second arrival process has sessions whose lengths are always null. The contribution of the second arrival process to cumulative loads is always 0, so the original arrival process will contribute to the process $A(\delta)$ as the first component with the modified arrival rate.

Assuming that $F_L(x)$ is regularly varying at 0 is equivalent to supposing that $W := L^{-1}$ has a distribution tail which is regularly varying at ∞ with index $-\alpha_W < 0$. To roughly estimate the traffic contribution of the region $\{> 0, 1\}$, we compute the mean of $A^{>0,1}(\delta)$. We have

$$\begin{aligned} E(A^{>0,1}(\delta)) &= E(P^{>0,1}(\delta)) E(R^{>0,1} L^{>0,1}) \\ &= \int_{0 < s < \delta} \int_{r > 0} \int_{l \leq \delta - s} r l \lambda \, ds F_R(dr) F_L(dl) \\ &= E(R)\lambda \int_{s=0}^\delta \int_{l=0}^s l F_L(dl) \, ds \end{aligned}$$

$$\begin{aligned}
 &= E(R)\lambda \int_{s=0}^{\delta} E(L \mathbf{1}_{[L \leq s]}) \, ds \\
 &= E(R)\lambda \int_{s=\delta^{-1}}^{\infty} \frac{E(L \mathbf{1}_{[L^{-1} \geq s]}) \, ds}{s^2}.
 \end{aligned}$$

Now suppose that

$$P[W > x] = \bar{F}_W(x) \sim x^{-\alpha_W} L(x) \quad \text{as } x \rightarrow \infty.$$

Then a variant of Karamata's theorem (Resnick (2006, Exercise 2.5)) implies that

$$E(L \mathbf{1}_{[L^{-1} \geq s]}) \sim \frac{\alpha_W}{1 + \alpha_W} s^{-1} \bar{F}_W(s) \quad \text{as } s \rightarrow \infty,$$

and, by Karamata's theorem, as $\delta \rightarrow 0$,

$$E(A^{>0,1}(\delta)) \sim E(R) \frac{\alpha_W}{(1 + \alpha_W)(2 + \alpha_W)} \frac{\bar{F}_W(\delta^{-1})}{\delta^{-1} \bar{F}_R(\delta^{-1})}.$$

Thus, if we assume that F_L is regularly varying at 0 or, equivalently, that \bar{F}_W is regularly varying at ∞ , we find that $E(A^{>0,1}(\delta)) \rightarrow 0$ as $\delta \rightarrow 0$ if and only if

$$\lim_{x \rightarrow \infty} \frac{\bar{F}_W(x)}{x \bar{F}_R(x)} = 0.$$

This implies that $A^{>0,1}(\delta) \xrightarrow{L_1} 0$ and, hence, region $\{> 0, 1\}$ gives a negligible contribution to the cumulative work. A sufficient condition is that

$$1 + \alpha_W > \alpha_R. \tag{11}$$

One reasonable circumstance where (11) holds is the following. Suppose that F_L has a density $F'_L(x)$ which converges to a limit at 0: $F'_L(x) \rightarrow F'_L(0) \in (0, \infty)$ as $x \rightarrow 0$. The standard exponential density satisfies this condition. Then, as $l \rightarrow 0$,

$$F_L(l) \sim l F'_L(0),$$

and, as $x \rightarrow \infty$,

$$\bar{F}_W(x) = P[L^{-1} > x] = P\left[L < \frac{1}{x}\right] \sim F'_L(0) \frac{1}{x},$$

and so $\bar{F}_W(x) \in RV_{-1}$ and $\alpha_W = 1$. So, owing to the condition $1 < \alpha_R < 2$, (11) is satisfied.

For the Boston University data set (Resnick (2007)), we were curious to see what the left-tail behavior of L turned out to be. QQ plots of the empirical quantiles of the log-transformed data plotted against the theoretical quantiles of the exponential distribution are given in Figure 2 and show that a reasonable estimate for α_W is 4.5.

Henceforth, we assume that $A^{>0,1}(\delta) \xrightarrow{L_1} 0$, so that we can neglect the asymptotic contribution to loading from region $\{> 0, 1\}$.

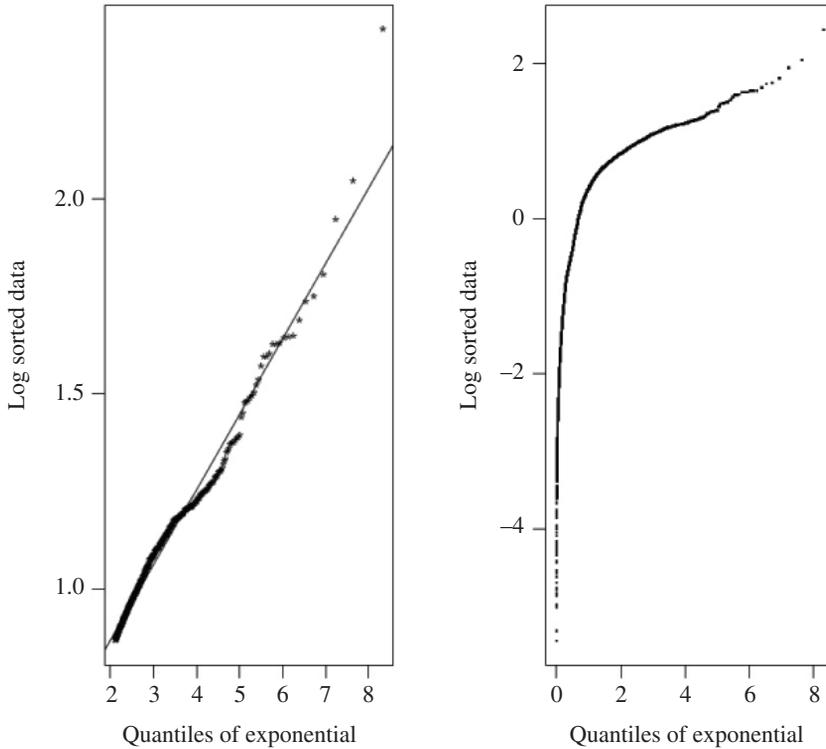


FIGURE 2: QQ plots of the 500 largest order statistics of $1/buL$ (left), the reciprocals of file download durations, and the corresponding plot for the whole data set (right). The slope estimator applied to the left plot gives an estimate for α_W of 4.5.

3.4. The contribution of region $\{< 0, 2\}$

We divide region $\{< 0, 2\}$ into two regions, $\{< 0, 2^-\}$ and $\{< 0, 2^+\}$, defined in the following way:

$$\{< 0, 2^-\} = \left\{ (s, r, l, u) \in \{< 0, 2\} : 0 \leq r < \frac{1}{\delta} \right\}, \quad \{< 0, 2^+\} = \{< 0, 2\} \setminus \{< 0, 2^-\}.$$

The reason for this further splitting is due to the fact that we are looking for a region that asymptotically gives a Gaussian contribution. We show that region $\{< 0, 2^-\}$ is that region. In addition, we show that region $\{< 0, 2^+\}$ dominates in the limit as δ tends to 0, so that in the limit the Gaussian component disappears.

3.4.1. *Characteristic function of $A^{<0, 2^-}(\delta)$.* Since

$$A^{<0, 2^-}(\delta) = \sum_{i=1}^{P^{<0, 2^-}(\delta)} R_i^{<0, 2^-} \delta,$$

the characteristic function of $A^{<0,2^-}(\delta)$ is computed as follows. For $\theta \in \mathbb{R}$,

$$\begin{aligned} \mathbb{E}(\exp(i\theta A^{<0,2^-}(\delta))) &= \exp(\mathbb{E}(P^{<0,2^-}(\delta))) \mathbb{E}(\exp(i\theta R_1^{<0,2^-}(\delta)) - 1) \\ &= \exp\left(\int_{s<0} \int_{0 \leq r < 1/\delta} \int_{l>|s|+\delta} (e^{i\theta r \delta} - 1) \lambda \, ds F_L(dl) F_R(dr)\right) \\ &= \exp\left(\lambda \int_{s>\delta} \int_{l>s} F_L(dl) \, ds \int_{r=0}^{1/\delta} (e^{i\theta r \delta} - 1) F_R(dr)\right) \\ &= \exp\left(\lambda \int_{s>\delta} \bar{F}_L(s) \, ds \int_{r=0}^{1/\delta} (e^{i\theta r \delta} - 1) F_R(dr)\right) \\ &= \exp\left(\lambda \bar{F}_R(\delta^{-1}) \mathbb{E}(L) \bar{F}_L^{(0)}(\delta) \int_{r=0}^1 (e^{i\theta r} - 1) \mu_\delta(dr)\right) \\ &= \exp\left(\lambda \bar{F}_R(\delta^{-1}) \int_0^1 (e^{i\theta r} - 1) \nu_\delta^{<0,2^-}(dr)\right), \end{aligned}$$

where

$$\nu_\delta^{<0,2^-}(dr) = \mathbb{E}(L) \bar{F}_L^{(0)}(\delta) \mu_\delta(dr).$$

Finally, using $\lambda = 1/\delta \bar{F}_R(\delta^{-1})$, we obtain

$$\mathbb{E}(\exp(i\theta A^{<0,2^-}(\delta))) = \exp\left(\delta^{-1} \int_0^1 (e^{i\theta r} - 1) \nu_\delta^{<0,2^-}(dr)\right). \tag{12}$$

3.4.2. *Gaussian limit for $A^{<0,2^-}(\delta)$.* For fixed $\delta > 0$, the quantity

$$m(\delta) := \delta^{-1} \int_0^1 r \nu_\delta^{<0,2^-}(dr) \tag{13}$$

is finite since

$$m(\delta) \leq \mathbb{E}(L) \delta^{-1} \int_0^1 1 \mu_\delta(dr) = \mathbb{E}(L) \delta^{-1} \frac{F_R(\delta^{-1})}{\bar{F}_R(\delta^{-1})} < \infty.$$

Define

$$a(\delta) := \left(\delta^{-1} \int_0^1 r^2 \nu_\delta^{<0,2^-}(dr)\right)^{1/2}. \tag{14}$$

Note that, as δ tends to 0, we have $a(\delta) \rightarrow \infty$ since, by Fatou's lemma,

$$\begin{aligned} \lim_{\delta \rightarrow 0} \delta a^2(\delta) &= \lim_{\delta \rightarrow 0} \mathbb{E}(L) \bar{F}_L^{(0)}(\delta) \int_0^1 r^2 \mu_\delta(dr) \\ &\geq \mathbb{E}(L) \int_0^1 r^2 \alpha_R r^{-\alpha_R-1} \, dr \\ &= \mathbb{E}(L) \frac{\alpha_R}{2 - \alpha_R}. \end{aligned} \tag{15}$$

Now we use (12) and write

$$\begin{aligned} & \mathbb{E} \exp\left(\frac{i\theta(A^{<0,2^-}(\delta) - m(\delta))}{a(\delta)}\right) \\ &= \exp\left(\int_0^1 \delta^{-1}(\exp(ia^{-1}(\delta)\theta r) - 1)v_\delta^{<0,2^-}(dr) - i\frac{\theta}{a(\delta)}\int_0^1 \delta^{-1}r v_\delta^{<0,2^-}(dr)\right) \\ &= \exp\left(\int_0^1 \delta^{-1}\left(\exp(ia^{-1}(\delta)\theta r) - 1 - i\frac{\theta}{a(\delta)}r\right)v_\delta^{<0,2^-}(dr)\right), \end{aligned}$$

and the exponent in the last expression converges to $-\theta^2/2$ since

$$\begin{aligned} & \left| \int_0^1 \delta^{-1}\left(\exp(ia^{-1}(\delta)\theta r) - 1 - i\frac{\theta}{a(\delta)}r\right)v_\delta^{<0,2^-}(dr) + \frac{\theta^2}{2} \right| \\ &= \left| \int_0^1 \delta^{-1}\left(\exp(ia^{-1}(\delta)\theta r) - 1 - i\frac{\theta}{a(\delta)}r - \frac{1}{2}\left(\frac{i\theta r}{a(\delta)}\right)^2\right)v_\delta^{<0,2^-}(dr) \right| \\ &\leq \frac{\delta^{-1}}{a^3(\delta)}\int_0^1 \frac{1}{3!}|\theta|^3 r^3 v_\delta^{<0,2^-}(dr) \end{aligned}$$

and

$$\frac{\delta^{-1}}{a^3(\delta)}\int_0^1 r^3 v_\delta^{<0,2^-}(dr) \leq \frac{\delta^{-1}}{a^3(\delta)}\int_0^1 r^2 v_\delta^{<0,2^-}(dr) = \frac{1}{a(\delta)} \rightarrow 0.$$

We summarize the previous result by the following proposition.

Proposition 4. *With $m(\delta)$ given by (13) and $a(\delta)$ given by (14), we have*

$$\frac{A^{<0,2^-}(\delta) - m(\delta)}{a(\delta)} \Rightarrow N^{<0,2^-} \sim N(0, 1) \quad \text{as } \delta \rightarrow 0.$$

3.4.3. *Characteristic function of $A^{<0,2^+}(\delta)$.* We have

$$A^{<0,2^+}(\delta) = \sum_{i=1}^{P^{<0,2^+}(\delta)} R_i^{<0,2^+} \delta.$$

Define

$$v_\delta^{<0,2^+}(dr) = \mathbb{E}(L)\bar{F}_L^{(0)}(\delta)\mu_{\delta/b(\delta)}(dr),$$

where

$$\mu_{\delta/b(\delta)}(dr) := \frac{F_R(\delta^{-1}b(\delta) dr)}{\bar{F}_R(\delta^{-1}(\delta))},$$

with $b(\delta)$ satisfying the relation

$$\gamma(\delta) := \frac{\bar{F}_R(\delta^{-1}b(\delta))}{\delta \bar{F}_R(\delta^{-1})} \rightarrow 1. \tag{16}$$

In Section 3.4.4, below, we study the function $b(\delta)$ and show that

$$b(\delta) = \left(\frac{1}{\delta}\right)^{1/\alpha_R} \ell\left(\frac{1}{\delta}\right) \rightarrow \infty \quad \text{as } \delta \downarrow 0 \tag{17}$$

for a function ℓ which is slowly varying at ∞ .

Now define

$$n(\delta) := b(\delta)\gamma(\delta) \int_{1/b(\delta)}^1 r v_\delta^{<0,2^+}(dr).$$

The characteristic function of $(A^{<0,2^+}(\delta) - n(\delta))/b(\delta)$ is computed as follows. For $\theta \in \mathbb{R}$,

$$\begin{aligned} & \mathbb{E}\left(\exp\left(i\theta \frac{A^{<0,2^+}(\delta) - n(\delta)}{b(\delta)}\right)\right) \\ &= \exp\left(\mathbb{E}(P^{<0,2^+}(\delta)) \mathbb{E}\left(\exp\left(i\frac{\theta}{b(\delta)} R_1^{<0,2^+} \delta\right) - 1\right) - i\theta \frac{n(\delta)}{b(\delta)}\right) \\ &= \exp\left(\int_{s<0} \int_{r \geq 1/\delta} \int_{|l|>|s|+\delta} (e^{i\theta r \delta/b(\delta)} - 1) \lambda ds F_L(dl) F_R(dr) - i\theta \frac{n(\delta)}{b(\delta)}\right) \\ &= \exp\left(\lambda \int_{s>\delta} \int_{l>s} F_L(dl) ds \int_{r=1/\delta}^\infty (e^{i\theta r \delta/b(\delta)} - 1) F_R(dr) - i\theta \frac{n(\delta)}{b(\delta)}\right) \\ &= \exp\left(\lambda \int_{s>\delta} \bar{F}_L(s) ds \int_{r=1/\delta}^\infty (e^{i\theta r \delta/b(\delta)} - 1) F_R(dr) - i\theta \frac{n(\delta)}{b(\delta)}\right) \\ &= \exp\left(\lambda \bar{F}_R(\delta^{-1}b(\delta)) \mathbb{E}(L) \bar{F}_L^{(0)}(\delta) \int_{r=1/b(\delta)}^\infty (e^{i\theta r} - 1) \mu_{\delta/b(\delta)}(dr) - i\theta \frac{n(\delta)}{b(\delta)}\right) \\ &= \exp\left(\lambda \bar{F}_R(\delta^{-1}b(\delta)) \int_{r=1/b(\delta)}^\infty (e^{i\theta r} - 1) v_\delta^{<0,2^+}(dr) - i\theta \frac{n(\delta)}{b(\delta)}\right), \end{aligned}$$

and, using $\lambda = 1/(\delta \bar{F}_R(\delta^{-1}))$, we have

$$\begin{aligned} & \mathbb{E}\left(\exp\left(i\theta \frac{A^{<0,2^+}(\delta) - n(\delta)}{b(\delta)}\right)\right) \\ &= \exp\left(\gamma(\delta) \left(\int_{1/b(\delta)}^1 (e^{i\theta r} - 1 - i\theta r) v_\delta^{<0,2^+}(dr) + \int_1^\infty (e^{i\theta r} - 1) v_\delta^{<0,2^+}(dr)\right)\right) \\ &\rightarrow \exp\left(\int_0^1 (e^{i\theta r} - 1 - i\theta r) \nu(dr) + \int_1^\infty (e^{i\theta r} - 1) \nu(dr)\right), \end{aligned}$$

where $\nu := \nu_0^{<0,2^+}$ is a Lévy measure with density $\mathbb{E}(L)\alpha_R x^{-\alpha_R-1}$.

We summarize the previous result by the following proposition.

Proposition 5. *As $\delta \rightarrow 0$,*

$$\frac{A^{<0,2^+}(\delta) - n(\delta)}{b(\delta)} \Rightarrow X_{\alpha_R}^{<0,2^+},$$

where the limit random variable is stable with Lévy measure ν whose density is $\mathbb{E}(L)\alpha_R x^{-\alpha_R-1}$.

3.4.4. *On the function $b(\delta)$.* The definition of $b(\cdot)$ in (16) is related to the concept of *conjugate inverses* of regularly varying functions. See Bingham *et al.* (1987, p. 48). The relationship in (16) can be rephrased as follows. Define

$$t = \delta^{-1}, \quad V(t) = \frac{1}{1 - F_R(t)}, \quad h\left(\frac{1}{\delta}\right) = b(\delta),$$

and require that

$$\frac{tV(t)}{V(th(t))} \rightarrow 1 \quad \text{as } t \rightarrow \infty$$

or

$$\frac{V(th(t))}{tV(t)} \rightarrow 1 \quad \text{as } t \rightarrow \infty.$$

Now define

$$c(t) = th(t).$$

We need

$$V \circ c(t) \sim tV(t),$$

so an obvious solution is

$$c(t) \sim V^{\leftarrow}(tV(t)).$$

Therefore,

$$h(t) = \frac{c(t)}{t} = \frac{V^{\leftarrow}(tV(t))}{t} \quad \text{as } t \rightarrow \infty.$$

Observe that

1. $tV(t) \in RV_{\alpha_R+1}$;
2. $V^{\leftarrow} \in RV_{1/\alpha_R}$ and, therefore, $V^{\leftarrow}(tV(t)) \in RV_{(1+\alpha_R)/\alpha_R} = RV_{1+1/\alpha_R}$;
3. $h(t) = V^{\leftarrow}(tV(t))/t \in RV_{1/\alpha_R+1-1} = RV_{1/\alpha_R}$.

Therefore,

$$b(\delta) = h\left(\frac{1}{\delta}\right) = \left(\frac{1}{\delta}\right)^{1/\alpha_R} \ell\left(\frac{1}{\delta}\right)$$

for a function ℓ which is slowly varying at ∞ .

The connection to conjugate pairs of slowly varying functions (Bingham *et al.* (1987, Section 1.5.7)) is as follows. Two slowly varying functions (ℓ, ℓ_*) are conjugate pairs if, as $x \rightarrow \infty$,

$$\ell(x)\ell_*(x\ell(x)) \rightarrow 1 \quad \text{and} \quad \ell_*(x)\ell(x\ell_*(x)) \rightarrow 1. \tag{18}$$

Given the functions V and V^{\leftarrow} , we may write

$$V(t) \sim t^{\alpha_R} \ell(t^{\alpha_R}) \quad \text{and} \quad V^{\leftarrow}(t) \sim t^{1/\alpha_R} \ell_*^{1/\alpha_R}(t)$$

for a conjugate pair (ℓ, ℓ_*) . This representation is possible (Bingham *et al.* (1987, Proposition 1.5.15)) because

$$V \circ V^{\leftarrow}(t) \sim t \quad \text{and} \quad V^{\leftarrow} \circ V(t) \sim t$$

re-expresses (18). This allows the following expression for h in terms of (ℓ, ℓ_*) :

$$\begin{aligned} h(t) &= \frac{1}{t} V^{\leftarrow}(tV(t)) \\ &\sim \frac{1}{t} (tV(t))^{1/\alpha_R} \ell_*^{1/\alpha_R}(tV(t)) \\ &= t^{1/\alpha_R} [\ell^{1/\alpha_R}(t^{\alpha_R}) \ell_*^{1/\alpha_R}(t^{1+\alpha_R} \ell(t^{\alpha_R}))]. \end{aligned}$$

The expression in square brackets is slowly varying and expressed in terms of (ℓ, ℓ_*) .

3.5. Discussion and summary

We summarize the contributions of the four regions to cumulative traffic in $(0, \delta)$.

1. For region $\{> 0, 2\}$, we have, as $\delta \rightarrow 0$,

$$X^{>0,2}(\delta) := A^{>0,2}(\delta) - \int_0^1 s\nu_\delta^{>0,2}(ds) \Rightarrow X_{\alpha_R}^{>0,2},$$

a spectrally positive, stable random variable with index α_R and Lévy measure $\nu_0^{>0,2}$ given in (7).

2. For region $\{> 0, 1\}$, we have, under suitable conditions on F_L , $A^{>0,1}(\delta) \xrightarrow{L_1} 0$. The contribution of this region is negligible in the limit.

3. For region $\{< 0, 2^-\}$, we have

$$X^{<0,2^-}(\delta) := \frac{A^{<0,2^-}(\delta) - m(\delta)}{a(\delta)} \Rightarrow N^{<0,2^-} \sim N(0, 1).$$

4. For region $\{< 0, 2^+\}$, we have

$$X^{<0,2^+}(\delta) := \frac{A^{<0,2^+}(\delta) - n(\delta)}{b(\delta)} \Rightarrow X_{\alpha_R}^{<0,2^+} =: X_{\alpha_R}^{<0,2},$$

a spectrally positive, stable random variable with index α_R and Lévy measure ν with density $E(L)\alpha_R x^{-\alpha_R-1}$ on $(0, \infty]$.

5. For region $\{< 0, 1\}$, we have

$$A^{<0,1}(\delta) \stackrel{D}{=} A^{>0,2}(\delta),$$

so

$$X^{<0,1}(\delta) := A^{<0,1}(\delta) - \int_0^1 s\nu_\delta^{>0,2}(ds) \Rightarrow X_{\alpha_R}^{<0,1} \stackrel{D}{=} X_{\alpha_R}^{>0,2}.$$

Thus, we may write

$$\begin{aligned} A(\delta) &= X^{>0,2}(\delta) + \int_0^1 s\nu_\delta^{>0,2}(ds) + A^{>0,1}(\delta) + a(\delta)X^{<0,2^-}(\delta) \\ &\quad + m(\delta) + b(\delta)X^{<0,2^+}(\delta) + n(\delta) + X^{<0,1}(\delta) + \int_0^1 s\nu_\delta^{>0,2}(ds). \end{aligned}$$

We conclude that

$$\begin{aligned} A(\delta) - m(\delta) - n(\delta) - 2 \int_0^1 s\nu_\delta^{>0,2}(ds) \\ = X^{>0,2}(\delta) + A^{>0,1}(\delta) + a(\delta)X^{<0,2^-}(\delta) + b(\delta)X^{<0,2^+}(\delta), \end{aligned}$$

where the summands on the right are independent and

$$X^{<0,1}(\delta) \stackrel{D}{=} X^{>0,2}(\delta) \Rightarrow X_{\alpha_R}^{>0,2} \quad (\text{spectrally positive, stable random variable with index } \alpha_R),$$

$$A^{>0,1}(\delta) \xrightarrow{L_1} 0 \quad (\text{negligible}),$$

$$X^{<0,2^+}(\delta) \Rightarrow X_{\alpha_R}^{<0,2^+} \quad (\text{spectrally positive, stable random variable with index } \alpha_R),$$

$$X^{<0,2^-}(\delta) \Rightarrow N^{<0,2^-} \quad (\text{normal}).$$

Also,

$$\frac{A(\delta) - d(\delta)}{b(\delta)} \Rightarrow X_{\alpha_R}^{<0,2},$$

which is stable with index α_R , where

$$d(\delta) := m(\delta) + n(\delta) + 2 \int_0^1 s v_\delta^{>0,2}(ds). \tag{19}$$

Using (15) and (17), we see that

$$\lim_{\delta \rightarrow 0} \frac{b(\delta)}{a(\delta)} \leq (\text{constant}) \lim_{\delta \rightarrow 0} \frac{\delta^{-1/\alpha_R} l(1/\delta)}{1/\delta^{1/2}} = (\text{constant}) \lim_{\delta \rightarrow 0} \delta^{1/2-1/\alpha_R} \left(\frac{1}{\delta}\right) = 0,$$

since $\frac{1}{2} < 1/\alpha_R < 1$. So our final traffic representation for cumulative traffic on a small interval for the RL model is

$$A(0, \delta] - d(\delta) \stackrel{D}{=} (X_{\alpha_R}^{>0,2} + o_p(1)) + (X_{\alpha_R}^{<0,1} + o_p(1)) + (X_{\alpha_R}^{<0,2+} + o_p(1))b(\delta) + (N^{<0,2-}(0, 1) + o_p(1))o(b(\delta)),$$

where $o_p(1)$ is a factor converging in probability to 0.

The RL model predicts that cumulative traffic over a small time interval is approximated by a stable random variable. We separated region $\{< 0, 2\}$ into two parts in order to find a region capable of giving a normal limit. In actual measurements (Sarvotham *et al.* (2005)) a component termed the β -traffic is observed which seems well approximated by a Gaussian distribution. At first, this observation seems to discourage the use of the RL model since the alternate RF model does give a normal approximation. However, as previously mentioned in Section 2, the RL model seems appropriate for specific Internet applications. This suggests that, by looking at these specific kinds of data flows, we could expect to find at high aggregation levels behavior different from Gaussian.

4. Dependence structure: asymptotic distributions

We now analyze the weak limits of the stochastic process

$$A(\delta) := \{A(k\delta, (k + 1)\delta], -\infty < k < \infty\},$$

defined in (1). We will see that the \mathbb{R}^∞ family indexed by δ converges to a limiting stable sequence

$$X_\infty = \{X_\infty(k), -\infty < k < \infty\}$$

with

$$P[X_\infty(0) = X_\infty(k)] = 1.$$

The price paid for letting δ tend to 0 is thus a limit sequence with degenerate dependence structure. The consequence of sampling at too high a frequency (using economic terminology) is perfect dependence.

Before starting the analysis, we state Lemma 1, below, which will considerably simplify subsequent computations. Its proof follows by the same computations carried out for Proposition 5.

Lemma 1. Let $\mathcal{R} = \mathcal{R}(\delta)$ be a subset of $\{< 0, 2\}$ of the form $\mathcal{R}(\delta) = \{(s, r, l, u) \in \{< 0, 2\}: (s, l) \in \mathcal{B}^{\mathcal{R}}(\delta)\}$, where $\mathcal{B}^{\mathcal{R}}(\delta)$ is a Borel subset of $\mathbb{R} \times \mathbb{R}_+$, and define

$$A^{\mathcal{R}}(\delta) = \sum_{i=1}^{P^{\mathcal{R}}(\delta)} R_i^{\mathcal{R}} \delta.$$

Then, as $\delta \rightarrow 0$,

$$\frac{1}{b(\delta)} \left(A^{\mathcal{R}}(\delta) - \frac{|\mathcal{B}^{\mathcal{R}}(\delta)|}{E(L)\bar{F}_L^{(0)}(\delta)}(n(\delta) + m(\delta)) \right) \Rightarrow \frac{|\mathcal{B}^{\mathcal{R}}(0)|}{E(L)} X,$$

where $|\mathcal{B}^{\mathcal{R}}(\delta)|$ is the measure of the set $\mathcal{B}^{\mathcal{R}}(\delta)$ under the measure $ds \times F_L(dl)$, $|\mathcal{B}^{\mathcal{R}}(0)| = \lim_{\delta \rightarrow 0} |\mathcal{B}^{\mathcal{R}}(\delta)| \geq 0$ is assumed to exist, and the limit random variable is stable with Lévy measure $\nu|\mathcal{B}^{\mathcal{R}}(0)|/E(L)$ whose density is $\alpha_{R^X}^{-\alpha_{R^X}-1}|\mathcal{B}^{\mathcal{R}}(0)|/E(L)$.

Corollary 1. Lemma 1 implies that if $|\mathcal{B}^{\mathcal{R}}(\delta)| \rightarrow 0$ as $\delta \rightarrow 0$ then

$$A^{\mathcal{R}}(\delta) - \frac{|\mathcal{B}^{\mathcal{R}}(\delta)|}{E(L)\bar{F}_L^{(0)}(\delta)}(n(\delta) + m(\delta)) = o_p(b(\delta)).$$

4.1. Convergence of finite-dimensional distributions of $\{A(i\delta, (i + 1)\delta], i \geq 1\}$

In this section we prove the following result.

Proposition 6. For any nonnegative integer k , as δ tends to 0 we have, in \mathbb{R}^{k+1} ,

$$\frac{1}{b(\delta)} \begin{pmatrix} A(0, \delta] - d(\delta) \\ A(\delta, 2\delta] - d(\delta) \\ \vdots \\ A(k\delta, (k + 1)\delta] - d(\delta) \end{pmatrix} \Rightarrow \begin{pmatrix} X_{\infty}(0) \\ X_{\infty}(1) \\ \vdots \\ X_{\infty}(k) \end{pmatrix},$$

where

$$b(\delta) = \left(\frac{1}{\delta}\right)^{1/\alpha_{R^X}} \ell\left(\frac{1}{\delta}\right) \rightarrow \infty \text{ as } \delta \downarrow 0,$$

$$d(\delta) = 2 \int_0^1 v \int_{r=v}^{\infty} \bar{F}_L(\delta r^{-1}v)r^{-1}\mu_{\delta}(dr) dv + b(\delta)\gamma(\delta) E(L)\bar{F}_L^{(0)}(\delta) \int_0^1 r\mu_{\delta b^{-1}(\delta)}(dr), \tag{20}$$

and each $X_{\infty}(i)$ for $0 \leq i \leq k$ are stable with Lévy measure ν whose density is $E(L)\alpha_{R^X}^{-\alpha_{R^X}-1}$. In addition, $P[X_{\infty}(i) = X_{\infty}(j)] = 1$ for $0 \leq i, j \leq k$.

Remark 1. The function $d(\delta)$ is the same function defined in (19). In (20) we have used the fact that

$$m(\delta) + n(\delta) = \delta^{-1} \int_0^{b(\delta)} r\nu_{\delta}^{<0, 2^-}(dr) = b(\delta)\gamma(\delta) \int_0^1 r\nu_{\delta}^{<0, 2^+}(dr),$$

which is straightforward to check.

Proof of Proposition 6. Along with the regions $\{< 0, 1\}$, $\{< 0, 2\}$, $\{> 0, 1\}$, and $\{> 0, 2\}$ used to analyze the convergence in distribution of $A(0, \delta]$, we need the analogously defined regions $\{< k\delta, 1\}$, $\{< k\delta, 2\}$, $\{> k\delta, 1\}$, and $\{> k\delta, 2\}$, where, for example,

$$\begin{aligned} \{< k\delta, 2\} &= \{(s, r, l, u) : s < k\delta, s + l > (k + 1)\delta\}, \\ \{> k\delta, 2\} &= \{(s, r, l, u) : k\delta < s < (k + 1)\delta, s + l > (k + 1)\delta\}. \end{aligned}$$

See Figure 3.

Additionally, for analyzing the dependence between $A(0, \delta]$ and $A(k\delta, (k + 1)\delta]$, we will need the regions \mathcal{R}_{11} , \mathcal{R}_{12} , \mathcal{R}_{21} , and \mathcal{R}_{22} which contain points $(\Gamma_k, R_k, L_k, F_k)$ contributing to both $A(0, \delta]$ as well as $A(k\delta, (k + 1)\delta]$. (See the left graphic in Figure 4.) In particular, points in $\mathcal{R}_{22} = \{< 0, 2\} \cap \{< k\delta, 2\}$ contribute

$$A^{\mathcal{R}_{22}} = \sum_{k: (\Gamma_k, R_k, L_k, F_k) \in \mathcal{R}_{22}} R_k \delta$$

to both $A(0, \delta]$ and $A(k\delta, (k + 1)\delta]$, which shows that a high degree of dependence is expected.

By applying Lemma 1 we have

$$\frac{1}{b(\delta)} \left(A^{\mathcal{R}_{22}} - \frac{|\mathcal{B}^{\mathcal{R}_{22}}(\delta)|}{E(L)\bar{F}_L^{(0)}(\delta)(n(\delta) + m(\delta))} \right) \Rightarrow X_{\alpha_R}^{\mathcal{R}_{22}} \quad \text{as } \delta \rightarrow 0$$

with $X_{\alpha_R}^{\mathcal{R}_{22}}$ stable with Lévy measure ν .

As for the other regions (see the right graphic in Figure 4), set

$$\mathcal{R}^{<0, (\delta, (k+1)\delta]} = \{(s, r, l, u) : s < 0, \delta < |s| + l \leq (k + 1)\delta\},$$

and write

$$\begin{aligned} A(0, \delta] &= A^{>0,1}(0, \delta] + A^{>0,2}(0, \delta] + A^{<0,1}(0, \delta] + A^{<0,2}(0, \delta] \\ &= A^{>0,1}(0, \delta] + A^{>0,2}(0, \delta] + A^{<0,1}(0, \delta] + A^{\mathcal{R}^{<0, (\delta, (k+1)\delta]}}(0, \delta] + A^{\mathcal{R}_{22}}(0, \delta]. \end{aligned}$$

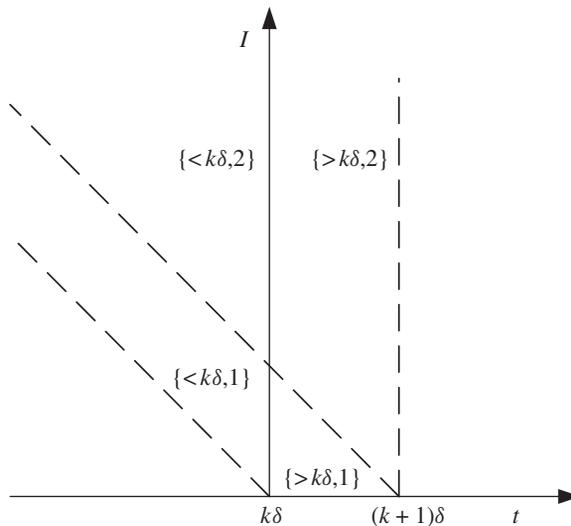


FIGURE 3: Four regions for analyzing contributions in the k th slot.

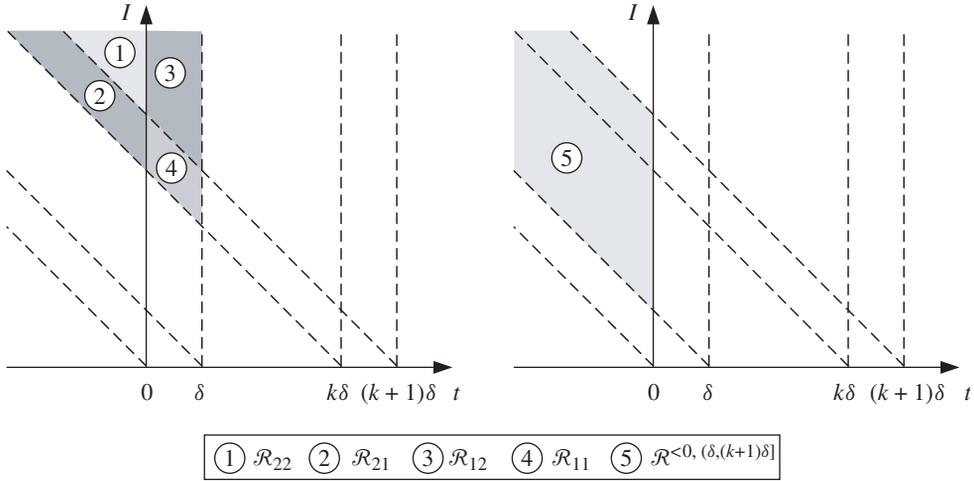


FIGURE 4: Regions for dependence analysis.

To analyze region $\mathcal{R}^{<0, (\delta, (k+1)\delta]}$, observe that we have

$$|\mathcal{B}^{\mathcal{R}^{<0, (\delta, (k+1)\delta]}}(\delta)| = E(L)(\bar{F}_L^{(0)}(\delta) - \bar{F}_L^{(0)}((k+1)\delta)) \rightarrow 0.$$

Hence, by applying Corollary 1 we have

$$A^{\mathcal{R}^{<0, (\delta, (k+1)\delta]}}(0, \delta] - \left(1 - \frac{|\mathcal{B}^{\mathcal{R}^{22}}|}{E(L)\bar{F}_L^{(0)}(\delta)}\right)(m(\delta) + n(\delta)) = o_p(b(\delta)).$$

Therefore, we conclude that

$$A(0, \delta] - d(\delta) = A^{\mathcal{R}^{22}}(0, \delta] - d(\delta) + o_p(b(\delta)).$$

Likewise, we consider $A(i\delta, (i+1)\delta]$ for $1 \leq i \leq k$. We set

$$\begin{aligned} \mathcal{R}^{<0, ((i+1)\delta, (k+1)\delta]} &= \{(s, r, l, u) : s < 0, (i+1)\delta < s+l < (k+1)\delta\}, \\ \mathcal{R}^{(0, i\delta], ((i+1)\delta, \infty]} &= \{(s, r, l, u) : 0 < s \leq i\delta, s+l > (i+1)\delta\} \end{aligned}$$

(see Figure 5), and write

$$\begin{aligned} A(i\delta, (i+1)\delta] &= A^{\mathcal{R}^{>i\delta, 1}}(i\delta, (i+1)\delta] + A^{\mathcal{R}^{>i\delta, 2}}(i\delta, (i+1)\delta] + A^{\mathcal{R}^{<i\delta, 1}}(i\delta, (i+1)\delta] \\ &\quad + (A^{\mathcal{R}^{<0, ((i+1)\delta, (k+1)\delta]}}(i\delta, (i+1)\delta] + A^{\mathcal{R}^{22}}(i\delta, (i+1)\delta] \\ &\quad + A^{\mathcal{R}^{(0, i\delta], ((i+1)\delta, \infty]}}(i\delta, (i+1)\delta]). \end{aligned}$$

Again, by Corollary 1 we have

$$\begin{aligned} A^{\mathcal{R}^{<0, ((i+1)\delta, (k+1)\delta]}}(i\delta, (i+1)\delta] - (m(\delta) + n(\delta)) \frac{|\mathcal{B}^{\mathcal{R}^{<0, ((i+1)\delta, (k+1)\delta]}}(\delta)|}{E(L)\bar{F}_L^{(0)}(\delta)} &= o_p(b(\delta)), \\ A^{\mathcal{R}^{(0, i\delta], ((i+1)\delta, \infty]}}(i\delta, (i+1)\delta] - (m(\delta) + n(\delta)) \frac{|\mathcal{B}^{\mathcal{R}^{(0, i\delta], ((i+1)\delta, \infty]}}(\delta)|}{E(L)\bar{F}_L^{(0)}(\delta)} &= o_p(b(\delta)). \end{aligned}$$

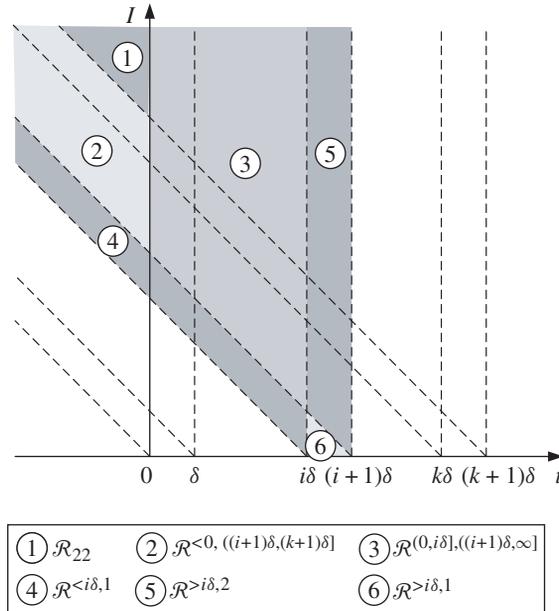


FIGURE 5: Regions for dependence analysis.

Therefore, keeping in mind that

$$A^{\mathcal{R}22}(k\delta, (k + 1)\delta] = A^{\mathcal{R}22}(i\delta, (i + 1)\delta] = A^{\mathcal{R}22}(0, \delta]$$

and that

$$|\mathcal{B}^{\mathcal{R}^{(0, i\delta], (i+1)\delta, \infty)}}(\delta)| + |\mathcal{B}^{\mathcal{R}^{<0, ((i+1)\delta, (k+1)\delta]}}(\delta)| + |\mathcal{B}^{\mathcal{R}22}(\delta)| = E(L)\bar{F}_L^{(0)}(\delta),$$

we have

$$A(i\delta, (i + 1)\delta] - d(\delta) = A^{\mathcal{R}22}(0, \delta] - d(\delta) + o_p(b(\delta)).$$

Thus, we have

$$\begin{pmatrix} A(0, \delta] - d(\delta) \\ A(\delta, 2\delta] - d(\delta) \\ \vdots \\ A(k\delta, (k + 1)\delta] - d(\delta) \end{pmatrix} = \begin{pmatrix} A^{\mathcal{R}22}(0, \delta] - d(\delta) \\ A^{\mathcal{R}22}(0, \delta] - d(\delta) \\ \vdots \\ A^{\mathcal{R}22}(0, \delta] - d(\delta) \end{pmatrix} + \begin{pmatrix} o_p(b(\delta)) \\ o_p(b(\delta)) \\ \vdots \\ o_p(b(\delta)) \end{pmatrix},$$

and the conclusion of Proposition 6 follows.

4.2. Asymptotic distribution over intervals a fixed distance apart

The previous section discusses dependence over successive slots of length δ . The asymptotic statement in Proposition 6 leads to a degenerate limit because $\delta \downarrow 0$ shrinks the distance between $A(0, \delta]$ and $A(k\delta, (k + 1)\delta]$. Here we investigate $(A(0, \delta], A(t, t + \delta])$ for $t > \delta$ with t fixed and find that, as $\delta \downarrow 0$, this vector is asymptotically stable with a nondegenerate dependence structure.

Proposition 7. *Suppose that $t > 0$ is fixed. As $\delta \downarrow 0$,*

$$\frac{1}{b(\delta)} \begin{pmatrix} A(0, \delta] - d(\delta) \\ A(t, t + \delta] - d(\delta) \end{pmatrix} \Rightarrow \begin{pmatrix} X_1 + X_t \\ X_2 + X_t \end{pmatrix}, \tag{21}$$

where $d(\delta)$ is given by (20) in Proposition 6, $X_1, X_2,$ and X_t are independent, stable random variables with respective Lévy measures

$$\begin{aligned} \nu_1(dx) &= \nu_2(dx) = E(L)F_L^{(0)}(t)\alpha_R x^{-\alpha_R-1} dx, \\ \nu_t(dx) &= E(L)\bar{F}_L^{(0)}(t)\alpha_R x^{-\alpha_R-1} dx. \end{aligned}$$

Proof. As in the proof of Proposition 6, we decompose

$$\begin{aligned} A(0, \delta] &= A^{<0, (\delta, t+\delta]}(\delta) + A^{<0, (t+\delta, \infty]}(\delta) + o_p(b(\delta)), \\ A(t, t + \delta] &= A^{(0, t], (t+\delta, \infty]}(t, t + \delta] + A^{<0, (t+\delta, \infty]}(t, t + \delta] + o_p(b(\delta)), \end{aligned}$$

and keep in mind that

$$A^{<0, (t+\delta, \infty]}(\delta) = A^{<0, (t+\delta, \infty]}(t, t + \delta]$$

and

$$A^{<0, (\delta, t+\delta]}(\delta) \stackrel{D}{=} A^{(0, t], (t+\delta, \infty]}(t, t + \delta].$$

Then we apply Lemma 1 by noting that

$$|\mathcal{B}^{<0, (\delta, t+\delta]}(0)| = E(L)F_L^{(0)}(t)$$

and

$$|\mathcal{B}^{<0, (t+\delta, \infty]}(0)| = E(L)\bar{F}_L^{(0)}(t).$$

5. Dependence structure: long-range dependence analysis

Correlations are not defined for either the pair $(A(0, \delta], A(k\delta, (k + 1)\delta])$ or for the limits $(X_1 + X_t, X_2 + X_t)$ in Proposition 7. This precludes a conventional discussion of the long-range dependence of variables lagged by k or t . In such circumstances the following alternative methods for defining and identifying long-range dependence have been proposed and used.

- General sample moment based methods proposed and discussed in Heyde and Yang (1997) using the sample Allen variance.
- Methods geared to analyzing the rate of decay of dependence in heavy-tailed processes such as the *extremal dependence measure* (Hernández-Campos *et al.* (2005); Resnick (2004b)).
- Methods specific to stable processes such as *covariation* or *codimension* (Samorodnitsky and Taqqu (1994)).
- Methods based on phase change ideas as promoted in Samorodnitsky (2002).

We will consider the first two methods here. The methods in the second two bullets often provide decay rates which look similar to long-range dependence discussions in finite-variance contexts using correlations. We emphasize that in non-Gaussian contexts, such as considered here, correlation or its surrogates contain only limited information about dependence.

5.1. Dependence structure: long-range dependence measured by sample Allen variance

One way to consider the dependence structure of the process $A(\delta)$ defined in (1) is by examining whether long-range dependence is present as measured by the sample Allen variance (see Heyde and Yang (1997)). A stationary, centered process $\{X_k, k \in \mathbb{Z}\}$ (which is not necessarily assumed to have finite-marginal second moments) is long-range dependent by sample Allen variance if

$$\frac{(\sum_{j=1}^m X_j)^2}{\sum_{j=1}^m X_k^2} \xrightarrow{P} \infty \quad \text{as } m \rightarrow \infty, \tag{22}$$

where ‘ \xrightarrow{P} ’ denotes convergence in probability.

Applying this to our stationary process $A(\delta)$, we find, for fixed δ , that the ratio in (22) reduces to

$$\frac{(\sum_{j=0}^{m-1} \tilde{A}(j\delta, (j+1)\delta))^2}{\sum_{j=0}^{m-1} (\tilde{A}(j\delta, (j+1)\delta))^2}, \tag{23}$$

where the tilde on a quantity like $\tilde{A}(j\delta, (j+1)\delta)$ indicates that the variable has been centered.

The numerator of (23) is equal to $(\tilde{A}(0, m\delta))^2$, that is the centered contribution of the sources in the interval $(0, m\delta]$. A modification of Theorem 5.1 of Maulik *et al.* (2002) to the stationary case shows that

$$\frac{\tilde{A}(0, m\delta]}{b_F(m)} \Rightarrow X_{\alpha_F} \quad \text{as } m \rightarrow \infty$$

for fixed δ , where X_{α_F} is a stable random variable with index α_F and as usual

$$b_F(m) := \inf \left\{ z : P[F > z] < \frac{1}{m} \right\} = \bar{G}^{\leftarrow} \left(\frac{1}{m} \right).$$

Note that $\alpha_F = \alpha_R \wedge \alpha_F$ and

$$(\tilde{A}(0, m\delta])^2 = O_p(b_F^2(m)),$$

where $b_F(m)$ is regularly varying with index $1/\alpha_F$. Therefore, (23) holds if

$$\frac{\sum_{j=0}^{m-1} (\tilde{A}(j\delta, (j+1)\delta))^2}{b_F^2(m)} \xrightarrow{P} 0.$$

As pointed out in Heyde and Yang (1997), the stability theorem of Loève (1978, Theorem E, p. 53) is relevant. Equation (4) implies that $E(A(0, \delta]^p) < \infty$ for $p < \alpha_R$, and from Loève (1978),

$$\sum_{j=0}^{m-1} \frac{(\tilde{A}(j\delta, (j+1)\delta))^2}{m^{2/p}} \rightarrow 0 \quad \text{as } m \rightarrow \infty. \tag{24}$$

When $\alpha_F = \alpha_L < \alpha_R$, we obtain (24), since b_F is regularly varying with index $1/\alpha_F$. We have not been able to resolve this when $\alpha_F = \alpha_R < \alpha_L$.

5.2. Dependence structure: long-range dependence measured by the extremal dependence measure

The extremal dependence measure (EDM) is defined for two nonnegative random variables whose joint distribution is multivariate regularly varying on the cone $[0, \infty)^2$ with limit measure ν and angular probability measure S . (This terminology is reviewed in Resnick (2007, Chapter 6).)

For a random vector $\mathbf{Z} = (Z^{(1)}, Z^{(2)})$, this means that there must exist a normalizing sequence $b_n \rightarrow \infty$ and a Radon measure ν on $[0, \infty]^2 \setminus \{\mathbf{0}\}$ such that

$$n \mathbb{P} \left[\frac{\mathbf{Z}}{b_n} \in \cdot \right] \xrightarrow{\nu} \nu(\cdot), \tag{25}$$

where ‘ $\xrightarrow{\nu}$ ’ denotes vague convergence on $[0, \infty]^2 \setminus \{\mathbf{0}\}$. The limit measure ν has the property that the polar coordinate transformation converts ν into a product measure,

$$\nu\{\mathbf{x} \in [0, \infty]^2 \setminus \{\mathbf{0}\} : r(\mathbf{x}) > r_0, \theta(\mathbf{x}) \in \cdot\} = cr_0^{-\alpha} S(\cdot)$$

for some $c > 0$, some index $\alpha > 0$, and S a probability measure on $[0, \pi/2]$. Here $(r(\mathbf{x}), \theta(\mathbf{x}))$ are the usual polar coordinates of \mathbf{x} , and S is called the angular measure.

5.2.1. *Definition and a useful property.* The EDM of the two random variables $(Z^{(1)}, Z^{(2)})$ is

$$\text{EDM}(Z^{(1)}, Z^{(2)}) := 1 - \frac{\int_0^{\pi/2} (\theta - \pi/4)^2 S(d\theta)}{(\pi/4)^2} = \frac{1}{(\pi/4)^2} \int_0^{\pi/2} \theta \left(\frac{\pi}{2} - \theta \right) S(d\theta). \tag{26}$$

The EDM has some of the desirable properties possessed by correlation (Resnick (2004b)). The EDM is 0 if $(Z^{(1)}, Z^{(2)})$ are independent or asymptotically independent, and the EDM is 1 if $(Z^{(1)}, Z^{(2)})$ are comonotone, $\mathbb{P}[Z^{(1)} = Z^{(2)}] = 1$, or if asymptotic full dependence holds. Another useful property which we need is summarized next.

Proposition 8. *Suppose that \mathbf{Z}_1 and \mathbf{Z}_2 are two independent random vectors of dimension 2 satisfying (25) with the same scaling sequence $\{b_n\}$. Then, with obvious notation regarding subscripting, we have the EDM of $\mathbf{Z}_1 + \mathbf{Z}_2$ satisfies*

$$\begin{aligned} \text{EDM}(Z_1^{(1)} + Z_2^{(1)}, Z_1^{(2)} + Z_2^{(2)}) &= \left(\frac{\nu_1\{\mathbf{x} : \|\mathbf{x}\| > 1\}}{\nu_1\{\mathbf{x} : \|\mathbf{x}\| > 1\} + \nu_2\{\mathbf{x} : \|\mathbf{x}\| > 1\}} \right) \text{EDM}(Z_1^{(1)}, Z_1^{(2)}) \\ &+ \left(\frac{\nu_2\{\mathbf{x} : \|\mathbf{x}\| > 1\}}{\nu_1\{\mathbf{x} : \|\mathbf{x}\| > 1\} + \nu_2\{\mathbf{x} : \|\mathbf{x}\| > 1\}} \right) \text{EDM}(Z_2^{(1)}, Z_2^{(2)}). \end{aligned} \tag{27}$$

Proof. The limit measure for $\mathbf{Z}_1 + \mathbf{Z}_2$ is $\nu_1 + \nu_2$ (Resnick (2007, Proposition 7.4, Section 7.3)) and, therefore, the angular measure of $\nu_1 + \nu_2$ can be written as

$$\frac{(\nu_1 + \nu_2)(\{\mathbf{x} : r(\mathbf{x}) > 1, \theta(\mathbf{x}) \in \cdot\})}{(\nu_1 + \nu_2)(\{\mathbf{x} : r(\mathbf{x}) > 1\})}.$$

The rest follows from algebra.

Remark 2. One practical implication of Proposition 8 is that if we can decompose a bivariate vector into an independent sum then any summand which possesses independent or asymptotically independent components will not contribute to the EDM and can be neglected.

5.2.2. *The EDM of $(X_1 + X_t, X_2 + X_t)$.* To see how an EDM calculation works, consider the stable vector

$$(X_1 + X_t, X_2 + X_t) = (X_1, X_2) + (X_t, X_t).$$

Its two-dimensional Lévy measure is $\nu_1 + \nu_2$, where ν_1 is the Lévy measure of (X_1, X_2) and ν_2 is the Lévy measure of (X_t, X_t) . Since (X_1, X_2) has independent components, its two-dimensional Lévy measure concentrates on the axes and will contribute 0 to the overall EDM of the sum. The measure ν_2 concentrates on the diagonal $\{(x, x) : x > 0\}$ and puts mass $E(L)\bar{F}_L^{(0)}(t)\nu_{\alpha_R}(dx)$ along the diagonal. The angular measure S corresponding to $\nu_1 + \nu_2$ has the form

$$S(d\theta) = c'\varepsilon_0(d\theta) + c'\varepsilon_{\pi/2}(d\theta) + c_t\varepsilon_{\pi/4}(d\theta),$$

where to make S a probability measure we require that

$$2c' + c_t = 1.$$

Proposition 8 implies that the constant c_t is given by

$$c_t := \nu_2\{\mathbf{x} : \|\mathbf{x}\| \geq 1\} = (\text{constant})\bar{F}_L^{(0)}(t).$$

Therefore, the EDM of $(X_1 + X_t, X_2 + X_t)$, denoted $\text{EDM}(t)$, is

$$\begin{aligned} \text{EDM}(t) &= 1 - \frac{\int_0^{\pi/2} (\theta - \pi/4)^2 S(d\theta)}{(\pi/4)^2} \\ &= 1 - \frac{2c'(\pi/4)^2}{(\pi/4)^2} \\ &= 1 - 2c' \\ &= c_t \\ &= (\text{constant})\bar{F}_L^{(0)}(t). \end{aligned}$$

Thus, $\text{EDM}(t)$ decays as $ct^{-(\alpha_L-1)}\ell_L(t)$ for some $c > 0$.

5.2.3. *The EDM of $(A(0, \delta], A(k\delta, (k + 1)\delta])$ as a function of k .* In this section we fix δ and compute an asymptotic form for the EDM of the bivariate random vector $(A(0, \delta], A(k\delta, (k + 1)\delta])$ and show a power law decay as k tends to ∞ .

We first decompose the components of $(A(0, \delta], A(k\delta, (k + 1)\delta])$ in the following terms that refer to the various regions depicted in Figure 6:

$$\begin{aligned} A(0, \delta] &= A_0^{\mathcal{R}_{I_0}} + A_0^{\mathcal{R}_{11}} + A_0^{\mathcal{R}_{12}} + A_0^{\mathcal{R}_{21}} + A_0^{\mathcal{R}_{22}}, \\ A(k\delta, (k + 1)\delta] &= A_k^{\mathcal{R}_{I_k}} + A_k^{\mathcal{R}_{11}} + A_k^{\mathcal{R}_{12}} + A_k^{\mathcal{R}_{21}} + A_k^{\mathcal{R}_{22}}, \end{aligned}$$

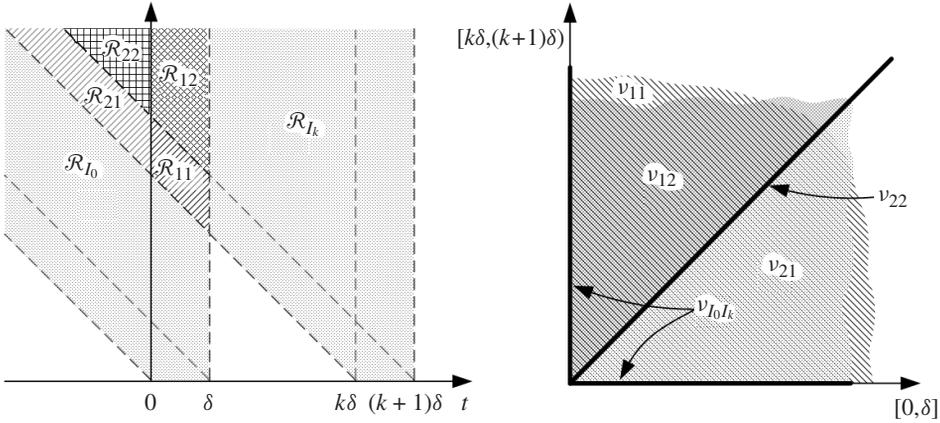


FIGURE 6: Regions and Lévy measures.

where

$$\begin{aligned}
 A_0^{\mathcal{R}_{11}} &= \sum_{i=1}^{p^{\mathcal{R}_{11}}} R_i^{\mathcal{R}_{11}} (\delta - \Gamma_i^{\mathcal{R}_{11}}), & A_k^{\mathcal{R}_{11}} &= \sum_{i=1}^{p^{\mathcal{R}_{11}}} R_i^{\mathcal{R}_{11}} (L_i^{\mathcal{R}_{11}} + \Gamma_i^{\mathcal{R}_{11}} - k\delta), \\
 A_0^{\mathcal{R}_{12}} &= \sum_{i=1}^{p^{\mathcal{R}_{12}}} R_i^{\mathcal{R}_{12}} (\delta - \Gamma_i^{\mathcal{R}_{12}}), & A_k^{\mathcal{R}_{12}} &= \sum_{i=1}^{p^{\mathcal{R}_{12}}} R_i^{\mathcal{R}_{12}} \delta, \\
 A_0^{\mathcal{R}_{21}} &= \sum_{i=1}^{p^{\mathcal{R}_{21}}} R_i^{\mathcal{R}_{21}} \delta, & A_k^{\mathcal{R}_{21}} &= \sum_{i=1}^{p^{\mathcal{R}_{21}}} R_i^{\mathcal{R}_{21}} (L_i^{\mathcal{R}_{21}} + \Gamma_i^{\mathcal{R}_{21}} - k\delta), \\
 A_0^{\mathcal{R}_{22}} &= \sum_{i=1}^{p^{\mathcal{R}_{22}}} R_i^{\mathcal{R}_{22}} \delta, & A_k^{\mathcal{R}_{22}} &= \sum_{i=1}^{p^{\mathcal{R}_{22}}} R_i^{\mathcal{R}_{22}} \delta.
 \end{aligned}$$

This allows us to decompose the bivariate random vector in the following way:

$$\begin{aligned}
 (A(0, \delta], A(k\delta, (k + 1)\delta]) &= (A_0^{\mathcal{R}_{I_0}}, A_k^{\mathcal{R}_{I_k}}) + (A_0^{\mathcal{R}_{11}}, A_k^{\mathcal{R}_{11}}) + (A_0^{\mathcal{R}_{12}}, A_k^{\mathcal{R}_{12}}) \\
 &\quad + (A_0^{\mathcal{R}_{21}}, A_k^{\mathcal{R}_{21}}) + (A_0^{\mathcal{R}_{22}}, A_k^{\mathcal{R}_{22}}),
 \end{aligned}$$

where all the random vectors on the right-hand side are compound Poisson and independent of each other. Furthermore, using the same argument as the one used to prove equality in distribution of the random variables $A^{>0,2}(\delta)$ and $A^{<0,1}(\delta)$, we have $A_0^{\mathcal{R}_{I_0}} \stackrel{D}{=} A_k^{\mathcal{R}_{I_k}}$ and $(A_0^{\mathcal{R}_{12}}, A_k^{\mathcal{R}_{12}}) \stackrel{D}{=} (A_k^{\mathcal{R}_{21}}, A_0^{\mathcal{R}_{21}})$.

We claim the following.

1. The dominant contribution to the EDM is from $(A_0^{\mathcal{R}_{22}}, A_k^{\mathcal{R}_{22}})$, which, as a function of k , decays as a constant times $\bar{F}_L^{(0)}(k)$.
2. The contribution of $(A_0^{\mathcal{R}_{I_0}}, A_k^{\mathcal{R}_{I_k}})$ to the EDM can be neglected because the random vector consists of independent components.
3. The contribution of $(A_0^{\mathcal{R}_{21}}, A_k^{\mathcal{R}_{21}})$ to the EDM is of lower order when k tends to ∞ . This is also true for $(A_0^{\mathcal{R}_{12}}, A_k^{\mathcal{R}_{12}})$ because of the distributional identities.

5.2.4. *The contribution of $(A_0^{\mathcal{R}22}, A_k^{\mathcal{R}22})$ to the EDM.* This is a relatively easy calculation since

$$A_0^{\mathcal{R}22} = A_k^{\mathcal{R}22} = \sum_{i=1}^{\mathcal{P}^{\mathcal{R}22}} R_1^{\mathcal{R}22} \delta,$$

and, therefore, using (27), the contribution to the EDM from $(A_0^{\mathcal{R}22}, A_k^{\mathcal{R}22})$ is proportional to $\nu_{22}\{\mathbf{x} : \|\mathbf{x}\| \geq 1\}$, where ν_{22} is the limit measure of $(A_0^{\mathcal{R}22}, A_k^{\mathcal{R}22})$. For some constant $c > 0$ (not necessarily the same with each use) and b_n satisfying $n\bar{F}_R(b_n) \rightarrow 1$, we have

$$\begin{aligned} \nu_{22}\{\mathbf{x} : \|\mathbf{x}\| \geq 1\} &= \lim_{n \rightarrow \infty} n \mathbb{P}\left[\frac{\|(A_0^{\mathcal{R}22}, A_k^{\mathcal{R}22})\|}{b_n} \geq 1\right] \\ &= \lim_{n \rightarrow \infty} cn \mathbb{P}\left[\sum_{i=1}^{\mathcal{P}^{\mathcal{R}22}} R_i^{\mathcal{R}22} \delta > b_n\right] \\ &= \lim_{n \rightarrow \infty} c \mathbb{E}(\mathcal{P}^{\mathcal{R}22}) n \mathbb{P}[R_1^{\mathcal{R}22} \delta > b_n] \\ &= \lim_{n \rightarrow \infty} \frac{c \mathbb{E}(\mathcal{P}^{\mathcal{R}22}) \mathbb{P}[R_1^{\mathcal{R}22} \delta > b_n]}{\bar{F}_R(b_n)} \\ &= c \int_{r\delta > b_n} \int_{s < 0} \int_{s+l > (k+1)\delta} \lambda \, ds \frac{F_R(dr)}{\bar{F}_R(b_n)} F_L(dl) \\ &= c\lambda \frac{\bar{F}_R(x/\delta)}{\bar{F}_R(b_n)} \int_{s < 0} \int_{l > (k+1)\delta - s} F_L(dl) \, ds \\ &= c\bar{F}_L^{(0)}((k+1)\delta). \end{aligned}$$

Our conclusion is that, as k tends to ∞ , the contribution to the composite EDM will be proportional to $\bar{F}_L^{(0)}(k)$.

5.2.5. *The contribution of $(A_0^{\mathcal{R}21}, A_k^{\mathcal{R}21})$ to the EDM is of lower order when k tends to ∞ .* We start with a marginal calculation analogous to the one just completed. As $x \rightarrow \infty$,

$$\begin{aligned} \mathbb{P}[A_0^{\mathcal{R}21} > x] &\sim \int \int \int_{\substack{s < 0, r\delta > x \\ k\delta < s+l \leq (k+1)\delta}} \lambda \, ds F_R(dr) F_L(dl) \\ &= \lambda \bar{F}_R\left(\frac{x}{\delta}\right) \int_{s=-\infty}^0 \int_{l=k\delta-s}^{(k+1)\delta-s} F_L(dl) \lambda \, ds, \end{aligned}$$

and changing variables again, $-s \mapsto s$, yields

$$\mathbb{P}[A_0^{\mathcal{R}21} > x] = \mathbb{E}(L) \bar{F}_R\left(\frac{x}{\delta}\right) \lambda (\bar{F}_L^{(0)}(k\delta) - \bar{F}_L^{(0)}((k+1)\delta)).$$

As k tends to ∞ , because of the difference, this is asymptotic to a constant times $\bar{F}_L(k)$ which is of lower order than $\bar{F}_L^{(0)}$.

For the EDM analysis of $(A_0^{\mathcal{R}21}, A_k^{\mathcal{R}21})$, note that

$$A_0^{\mathcal{R}21} \geq A_k^{\mathcal{R}21}. \tag{28}$$

Referring back to (25), let b_n be the appropriate scaling constant (which could be the quantile function of F_R). Then, from definition (26) and the ordering (28), we have

$$\begin{aligned} \text{EDM}(A_0^{\mathcal{R}21}, A_k^{\mathcal{R}21}) &= \int_0^{\pi/2} \theta \left(\frac{\pi}{2} - \theta \right) \lim_{n \rightarrow \infty} n \mathbb{P} \left[\sqrt{(A_0^{\mathcal{R}21})^2 + (A_k^{\mathcal{R}21})^2} > b_n, \Theta \in d\theta \right] \\ &\leq (\text{constant})n \mathbb{P}[(A_0^{\mathcal{R}21})^2 + (A_k^{\mathcal{R}21})^2 > b_n^2] \\ &\leq n \mathbb{P}[\sqrt{2}A_0^{\mathcal{R}21} > b_n] \\ &\sim (\text{constant})(\bar{F}_L^{(0)}(k\delta) - \bar{F}_L^{(0)}((k+1)\delta)). \end{aligned}$$

Therefore, as $k \rightarrow \infty$,

$$\text{EDM}(A_0^{\mathcal{R}21}, A_k^{\mathcal{R}21}) = o(\bar{F}_L^{(0)}(k)).$$

5.2.6. *Conclusion.* We have the following conclusion. The EDM satisfies, as $k \rightarrow \infty$,

$$\text{EDM}(A(0, \delta], A(k\delta, (k+1)\delta]) \sim (\text{constant})\bar{F}_L^{(0)}(k) = (\text{constant})k^{-(\alpha_L-1)}\ell_L(k).$$

So α_R controls the heaviness of tails of cumulative input and α_L controls the dependence decay for fixed δ as k tends to ∞ .

6. Comparison of the RL and RF models

In this section we compare the RL and RF models, in order to emphasize their differences. As noted in Section 2, these two models make different assumptions about the joint distribution of the mark vector (R, L, F) even though in both cases $F = LR$. The RL model assumes that the components R and L are independent while the RF model chooses R and F to be independent.

6.1. Asymptotic distribution of $A(0, \delta)$

For each of the two models, Table 1 summarizes the results for the limit distribution of the random variable $A(\delta)$ and its components $A^{\mathcal{R}}(\delta)$ relative to the regions $\mathcal{R} \in \{ \{> 0, 2\}; \{> 0, 1\}; \{< 0, 2\}; \{< 0, 1\} \}$. For reading Table 1, recall that

$$\mathbb{P}[F \leq x] = G(x), \quad \mathbb{P}[L \leq x] = F_L(x), \quad \mathbb{P}[R \leq x] = F_R(x),$$

and that the tails of the three distributions are regularly varying with parameters α_F, α_L , and α_R , respectively, all assumed to be strictly between 1 and 2. Furthermore, $\mu_\delta(dr) := F_R(\delta^{-1} dr) / \bar{F}_R(\delta^{-1})$ and

$$\bar{G}_0(x) := \frac{1}{E(F)} \int_x^\infty \bar{G}(u) du, \quad \bar{F}_L^{(0)}(x) := \frac{1}{E(L)} \int_x^\infty \bar{F}_L(u) du.$$

The following additional notation is used in Table 1.

$$\begin{aligned} v_{\text{RL}}^\delta(ds) &:= \left(\int_s^\infty \bar{F}_L(\delta r^{-1}s)r^{-1} \mu_\delta(dr) \right) ds, & v_{\text{RF}}^\delta(ds) &:= \left(\bar{G}(s) \int_{r=s}^\infty r^{-1} \mu_\delta(dr) \right) ds, \\ n_{\text{RL}}(\delta) &:= \delta^{-1} \int_\delta^\infty \bar{F}_L(l) dl \int_0^{b_{\text{RL}}(\delta)} r \mu_\delta(dr), & n_{\text{RF}}(\delta) &:= E(F) \int_0^1 \bar{G}_0(r) \mu_\delta(dr), \\ d_{\text{RL}}(\delta) &:= n_{\text{RL}}(\delta) + 2 \int_0^1 s v_{\text{RL}}^\delta(ds), & d_{\text{RF}}(\delta) &:= n_{\text{RF}}(\delta) + 2 \int_0^1 s v_{\text{RF}}^\delta(ds), \\ b_{\text{RL}}(\delta) &:= \left(\frac{1}{\delta} \right)^{1/\alpha_R} \ell \left(\frac{1}{\delta} \right), & b_{\text{RF}}(\delta) &:= \left(E(F) \int_0^1 r \bar{G}_0(r) \mu_\delta(dr) \right)^{1/2}. \end{aligned}$$

TABLE 1: Comparison of the RL and RF models: weak limit of $A(\delta)$.

\mathcal{R}	RL model	RF model
$\{> 0, 1\}$	Negligible (under conditions)	$A^{>0,1}(\delta) \Rightarrow X_{\text{RF}}^{>0,1}$ Infinitely divisible $P[X_{\text{RF}}^{>0,1} > x] \in RV_{-(\alpha_R + \alpha_F)}$
$\{< 0, 2\}$	$\frac{A^{<0,2}(\delta) - n_{\text{RL}}(\delta)}{b_{\text{RL}}(\delta)} \Rightarrow X_{\text{RL}}^{<0,2}$ Stable(α_R)	$\frac{A^{<0,2}(\delta) - n_{\text{RF}}(\delta)}{b_{\text{RF}}(\delta)} \Rightarrow N_{\text{RF}}^{<0,2} \sim N(0, 1)$
$\{> 0, 2\}$	$A^{>0,2}(\delta) - \int_0^1 s v_{\text{RL}}^\delta(ds) \Rightarrow X_{\text{RL}}^{>0,2}$ Stable(α_R)	$A^{>0,2}(\delta) - \int_0^1 s v_{\text{RF}}^\delta(ds) \Rightarrow X_{\text{RF}}^{>0,2}$ Infinitely divisible $P[X_{\text{RF}}^{>0,2} > x] \in RV_{-(\alpha_R + \alpha_F)}$
$\{< 0, 1\}$	$A^{<0,1}(\delta) \stackrel{D}{=} A^{>0,2}(\delta)$	$A^{<0,1}(\delta) \stackrel{D}{=} A^{>0,2}(\delta)$
All	$\frac{A(\delta) - d_{\text{RL}}(\delta)}{b_{\text{RL}}(\delta)} \Rightarrow X_{\text{RL}}^{<0,2}$ Stable(α_R)	$\frac{A(\delta) - d_{\text{RF}}(\delta)}{b_{\text{RF}}(\delta)} \Rightarrow N_{\text{RF}}^{<0,2} \sim N(0, 1)$

Table 1 shows that in both models $A(0, \delta]$ converges weakly after centering and scaling. In the RF model the weak limit is a normal random variable while in the RL model the weak limit is a heavy-tailed stable random variable with index α_R .

For both models, the main component in $A(0, \delta]$ comes from region $\{< 0, 2\}$. We need to understand the difference in treatment of this region by the two models. Sources that contribute in this region are ones whose durations are relatively long, since they start from the past and continue past δ and their contributions are given by $R\delta$ in each case. In the RF model R and F are independent and it is therefore unlikely they are both large. The relationship $L = F/R$ means that a long L may be associated with a small R and large F . To have a contribution of a long L limits the values of R in such a way that the central limit theorem holds. In the RL model, on the other hand, the independence of R and L makes it unlikely that R and L are both large, but a large value of R makes cumulative input in $(0, \delta]$ asymptotically stable with index α_R , while a large value of L induces dependence measured by decay governed by α_L .

For both models, we have always assumed that

$$\lambda(\delta) = \frac{1}{\delta \bar{F}_R(\delta^{-1})}.$$

If we compare the normalizing functions $b_{\text{RF}}(\delta)$ and $b_{\text{RL}}(\delta)$, we have, as $\delta \rightarrow 0$,

$$b_{\text{RF}}(\delta) = o(b_{\text{RL}}(\delta)).$$

This means that if we construct a combined model that mixes the RF and the RL models, we would find that its limit behavior coincides with that of the RL model. In order to obtain, in the limit, a linear combination of a normal distribution and an infinite divisible distribution, the Poisson intensity functions $\{\lambda(\delta)\}$ would have to grow at different speeds.

TABLE 2: Comparison of the RL and RF models: dependence structure.

	RL model	RF model
$t \rightarrow \infty$	EDM($X_{RL}^{<0,2}(0), X_{RL}^{<0,2}(t)$) $\sim ct^{-(\alpha_L-1)}\ell_L(t)$	cov($X_{RF}^{<0,2}(0), X_{RF}^{<0,2}(t)$) $\sim ct^{-(\alpha_F-1)}\ell_F(t)$
$k \rightarrow \infty$, fixed δ	EDM($A(0, \delta], A(k\delta, (k+1)\delta]$) $\sim ck^{-(\alpha_L-1)}\ell_L(k)$	cov($A(0, \delta], A(k\delta, (k+1)\delta]$) $\sim ck^{-(\alpha_F-1)}\ell_F(k)$

6.2. Dependence structure

If we look at the dependence structure, we have, in the limit, the effect of high-frequency sampling for both models. For any nonnegative integer k and $h \in \{RL, RF\}$, we have, as $\delta \rightarrow 0$,

$$\frac{1}{b_h(\delta)} \begin{pmatrix} A(0, \delta] - d_h(\delta) \\ A(\delta, 2\delta] - d_h(\delta) \\ \vdots \\ A(k\delta, (k+1)\delta] - d_h(\delta) \end{pmatrix} \Rightarrow \begin{pmatrix} X_h^{<0,2}(0) \\ X_h^{<0,2}(1) \\ \vdots \\ X_h^{<0,2}(k) \end{pmatrix}$$

in \mathbb{R}^{k+1} , with $P[X_h^{<0,2}(i) = X_h^{<0,2}(j)] = 1$ for $0 \leq i, j \leq k$. If $h = RF$, the limit is Gaussian. If $h = RL$, the limit is stable.

The time slots $(i\delta, (i+1)\delta]$, $i = 0, \dots, k$, are at a distance from each other which converges to 0 as δ tends to 0. For time slots $(0, \delta]$ and $(t, t + \delta]$ at a minimum distance t apart, we have, for fixed $t > 0$ and $h \in \{RF, RL\}$, as $\delta \downarrow 0$,

$$\frac{1}{b_h(\delta)} \begin{pmatrix} A(0, \delta] - d_h(\delta) \\ A(t, t + \delta] - d_h(\delta) \end{pmatrix} \Rightarrow \begin{pmatrix} X_h^{<0,2}(0) \\ X_h^{<0,2}(t) \end{pmatrix}. \tag{29}$$

When $h = RL$, the limit is the dependent stable vector given in Proposition 7; see (21). When $h = RF$, the limit is a dependent Gaussian pair (D’Auria and Resnick (2006)).

Assessing the decay of dependence requires different techniques for the two models. For the RF model, tails are relatively thin allowing traditional correlation techniques to be used to claim long-range dependence. For the heavy-tailed RL model, correlations do not exist and an alternative technique based on the EDM is used to show the slow decay of dependence at power-law rate. These results are summarized in Table 2, which uses the notation for the limit in (29).

7. Final thoughts

Heterogeneous traffic comprising different types of applications may behave differently to more homogeneous traffic. In particular, it looks sensible to decompose traffic into classes of fairly homogeneous applications and to study each separately, seeking statistical differences in their characteristics. Our results suggest that the tail behavior, dependence structure, and approximating distributions may depend on the statistical characteristics of each application component of the network traffic. This suggests the detailed statistical studies are required to understand how statistical characteristics of the traffic vary by application.

Acknowledgements

We thank an anonymous referee for pointing out the interesting reference of Heyde and Yang (1997), and suggesting we examine the dependence structure of our process from the point of view of Allen variance.

Sidney Resnick's research was partially supported by NSA grant MSPF-05G-0492. Much of this research took place while Sidney Resnick was a wandering academic on sabbatical, and grateful acknowledgement for support and hospitality go to Eurandom, The Netherlands; Columbia University, Department of Statistics, Department of Industrial Engineering and Operations Research, Graduate School of Business; University of North Carolina, Chapel Hill, Department of Statistics and Operations Research; and SAMSI, Research Triangle Park.

References

- ARLITT, M. AND WILLIAMSON, C. L. (1996). Web server workload characterization: the search for invariants (extended version). In *Proc. ACM SIGMETRICS Conf.* (Philadelphia, PA), ACM, New York, pp. 126–137.
- BEN AZZOUNA, N., CLÉROT, F., FRICKER, C. AND GUILLEMIN, F. (2004). A flow-based approach to modeling ADSL traffic on an IP backbone link. *Ann. Telecommun.*, **59**, 1260–1299.
- BINGHAM, N. H., GOLDIE, C. M., AND TEUGELS, J. L. (1987). *Regular Variation*. Cambridge University Press.
- BREIMAN, L. (1965). On some limit theorems similar to the arc-sin law. *Theory Prob. Appl.* **10**, 323–331.
- CLINE, D. B. H. (1983). Estimation and linear prediction for regression, autoregression and ARMA with infinite variance data. Doctoral Thesis, Colorado State University.
- CROVELLA, M. AND BESTAVROS, A. (1996). Self-similarity in world wide web traffic: evidence and possible causes. In *Proc. ACM SIGMETRICS Conf.* (Philadelphia, PA), ACM, New York, pp. 160–169.
- CROVELLA, M. AND BESTAVROS, A. (1997). Self-similarity in world wide web traffic: evidence and possible causes. *IEEE/ACM Trans. Networking* **5**, 835–846.
- D'AURIA, B. AND RESNICK, S. I. (2006). Data network models of burstiness. *Adv. Appl. Prob.* **38**, 373–404.
- DAVIS, R. A., AND RESNICK, S. I. (1986). Limit theory for the sample covariance and correlation functions of moving averages. *Ann. Statist.* **14**, 533–558.
- DE HAAN, L. (1970). *On Regular Variation and Its Application to the Weak Convergence of Sample Extremes*. Mathematisch Centrum, Amsterdam.
- DUFFY, D. E., MCINTOSH, A. A., ROSENSTEIN, M. AND WILLINGER, W. (1993). Analyzing telecommunications traffic data from working common channel signaling subnetworks. In *Computing Science and Statistics Interface* (Proc. 25th Symp. Interface; San Diego, CA), eds M. E. Tarter and M. D. Lock, Interface, Fairfax Station, VA, pp. 156–165.
- EMBRECHTS, P. AND GOLDIE, C. M. (1980). On closure and factorization properties of subexponential and related distributions. *J. Austral. Math. Soc. Ser. A* **29**, 243–256.
- GUERIN, C. A. *et al.* (2003). Empirical testing of the infinite source Poisson data traffic model. *Stoch. Models*, **19**, 151–200.
- 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.
- HEFFERNAN, J. E. AND RESNICK, S. I. (2005). Hidden regular variation and the rank transform. *Adv. Appl. Prob.* **37**, 393–414.
- HERNÁNDEZ-CAMPOS, F. *et al.* (2005). Extremal dependence: internet traffic applications. *Stoch. Models* **21**, 1–35.
- HEYDE, C. C. AND YANG, Y. (1997). On defining long-range dependence. *J. Appl. Prob.* **34**, 939–944.
- HEYMAN, D. AND LAKSHMAN, T. V. (1996). What are the implications of long-range dependence for VBR-video traffic engineering? *IEEE/ACM Trans. Networking* **4**, 301–317.
- KAJ, I. AND TAQQU, M. S. (2008). Convergence to fractional Brownian motion and to the Telecom process: the integral representation approach. To appear in *Brazilian Prob. School, 10th Anniversary Volume*, eds M.E. Vares and V. Sidoravicius.
- KALLENBERG, O. (1983). *Random Measures*, 3rd edn. Akademie-Verlag, Berlin.
- KONSTANTOPOULOS, T. AND LIN, S. J. (1998). Macroscopic models for long-range dependent network traffic. *Queueing Systems* **28**, 215–243.
- LELAND, W. E., TAQQU, M. S., WILLINGER, W. AND WILSON, D.V. (1993). Statistical analysis of high time-resolution ethernet LAN traffic measurements. In *Proc. 25th Symp. Interface Statist. Comput. Sci.*, Interface Foundation of North America, Virginia, pp. 146–155.
- LELAND, W. E., TAQQU, M. S., WILLINGER, W. AND WILSON, D.V. (1994). On the self-similar nature of ethernet traffic (extended version). *IEEE/ACM Trans. Networking* **2**, 1–15.

- LEVY, J. AND TAQQU, M. (2000). Renewal reward processes with heavy-tailed interrenewal times and heavy-tailed rewards. *Bernoulli* **6**, 23–44.
- LOÈVE, M. (1978). *Probability Theory* (Graduate Texts Math. **46**), Vol. II, 4th edn. Springer, New York.
- MAULIK, K. AND RESNICK, S. I. (2003). The self-similar and multifractal nature of a network traffic model. *Stoch. Models* **19**, 549–577.
- MAULIK, K., RESNICK, S. I. AND ROOTZÉN, H. (2003). Asymptotic independence and a network traffic model. *J. Appl. Prob.* **39**, 671–699.
- MIKOSCH, T., RESNICK, S. I., ROOTZÉN, H. AND STEGEMAN, A. W. (2002). Is network traffic approximated by stable Lévy motion or fractional Brownian motion? *Ann. Appl. Prob.* **12**, 23–68.
- NEVEU, J. (1977). Processus ponctuels. In *École d'Été de Probabilités de Saint-Flour, VI—1976* (Lecture Notes Math. **598**), Springer, Berlin, pp. 249–445.
- PANDURANGAN, G., RAGHAVAN, P. AND UPFAL, E. (2001). Building low-diameter P2P networks. In *Proc. 42nd IEEE Symp. Foundations Comput. Sci.* (Las Vegas, NV), IEEE, Los Alamitos, CA, pp. 492–499.
- PARK, K. AND WILLINGER, W. (2000). Self-similar network traffic: an overview. In *Self-Similar Network Traffic and Performance Evaluation*, eds K. Park and W. Willinger, John Wiley, New York, pp. 1–38.
- PRATT, J. W. (1960). On interchanging limits and integrals. *Ann. Math. Statist.* **31**, 74–77.
- RESNICK, S. I. (1987). *Extreme Values, Regular Variation and Point Processes*. Springer, New York.
- RESNICK, S. I. (1992). *Adventures in Stochastic Processes*. Birkhäuser, Boston, MA.
- RESNICK, S. I. (1999). *A Probability Path*. Birkhäuser, Boston, MA.
- RESNICK, S. I. (2003). Modeling data networks. In *SemStat: Seminaire Europeen de Statistique, Extreme Values in Finance, Telecommunications, and the Environment*, eds B. Finkenstadt and H. Rootzén, Chapman & Hall, London, pp. 287–372.
- RESNICK, S. I. (2004a). On the foundations of multivariate heavy tail analysis. In *Stochastic Methods and Their Applications* (J. Appl. Prob. Spec. Vol. **41A**), eds J. Gani and E. Seneta, Applied Probability Trust, Sheffield, pp. 191–212.
- RESNICK, S. I. (2004b). The extremal dependence measure and asymptotic independence. *Stoch. Models* **20**, 205–227.
- RESNICK, S. I. (2007). *Heavy Tail Phenomena: Probabilistic and Statistical Modeling*. Springer, New York.
- RESNICK, S. I. AND ROOTZÉN, H. (2000). Self-similar communication models and very heavy tails. *Ann. Appl. Prob.* **10**, 753–778.
- RIEDI, R. H. AND WILLINGER, W. (2000). Toward an improved understanding of network traffic dynamics. In *Self-Similar Network Traffic and Performance Evaluation*, John Wiley, New York, pp. 507–530.
- SAMORODNITSKY, G. (2002). Long range dependence, heavy tails and rare events. Lecture Note MPS-LN 2002-12, Centre for Mathematical Physics and Stochastics, Department of Mathematical Sciences, University of Aarhus. Available at <http://www.maphysto.dk/cgi-bin/gp.cgi?publ=412>.
- SAMORODNITSKY, G. AND TAQQU, M. (1994). *Stable Non-Gaussian Random Processes. Stochastic Models with Infinite Variance*. Chapman & Hall, New York.
- SARVOTHAM, S., RIEDI, R. AND BARANIUK, R. (2005). Network and user driven on-off source model for network traffic. *Comput. Networks* **48**, 335–350.
- SENETA, E. (1976). *Regularly Varying Functions* (Lecture Notes Math. **508**), Springer, New York.
- TANENBAUM, A. (1996). *Computer Networks*, 3rd edn. Prentice Hall PTR, Upper Saddle River, NJ.
- 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.
- WILLINGER, W. (1998). Data network traffic: heavy tails are here to stay. Presentation at Extremes – Risk and Safety, Nordic School of Public Health, Gothenberg, Sweden, August 1998.
- WILLINGER, W. AND PAXSON, V. (1998). Where mathematics meets the Internet. *Notices Amer. Math. Soc.* **45**, 961–970.
- WILLINGER, W., PAXSON, V. AND TAQQU, M. S. (1998). Self-similarity and heavy tails: structural modeling of network traffic. In *A Practical Guide to Heavy Tails. Statistical Techniques and Applications*, eds R. J. Adler *et al.* Birkhäuser, Boston, MA, pp. 27–53.
- WILLINGER, W., TAQQU, M. S., LELAND, M. AND WILSON, D. (1995). Self-similarity in high-speed packet traffic: analysis and modelling of ethernet traffic measurements. *Statist. Sci.* **10**, 67–85.
- WILLINGER, W., TAQQU, M. S., LELAND, M. AND WILSON, D. (1997). Self-similarity through high variability: statistical analysis of ethernet LAN traffic at the source level (extended version). *IEEE/ACM Trans. Networking* **5**, 71–96.