

EMPIRICAL TESTING OF THE INFINITE SOURCE POISSON DATA TRAFFIC MODEL

C.A. GUERIN, H. NYBERG, O. PERRIN, S. RESNICK, H. ROOTZÉN, AND C. STĂRICĂ

ABSTRACT. The infinite source Poisson model is a fluid queue approximation of network data transmission that assumes that sources begin constant rate transmissions of data at Poisson time points for random lengths of time. This model has been a popular one as analysts attempt to provide explanations for observed features in telecommunications data such as self-similarity, long range dependence and heavy tails. We survey some features of this model in cases where transmission length distributions have (a) tails so heavy that means are infinite, (b) heavy tails with finite mean and infinite variance and (c) finite variance. We survey the self-similarity properties of various descriptor processes in this model and then present analyses of four data sets which show that certain features of the model are consistent with the data while others are contradicted. The data sets are 1) the Boston University 1995 study of web sessions, 2) the UC Berkeley home IP HTTP data collected in November 1996, 3) traces collected in end of 1997 at a Customer Service Switch in Munich, and 4) detailed data from a corporate Ericsson WWW server from October 1998.

1. INTRODUCTION

Statistical simulation is of basic importance for the choice of buffer sizes, protocols, network configurations and other aspects of the design of complex telecommunication systems. Simulation, and network analysis in general, must be founded on models which capture important features of the traffic in a realistic and flexible way. Yet the models have to be simple enough to allow for understanding, theoretical analysis, and easy fitting to many kinds of observed and synthetic data. A simple model, here termed “the infinite source Poisson model” and sometimes called the $M/G/\infty$ input model, which has the potential to satisfy these requirements for IP; HTTP, FTP, SMTP and other protocols for file transfers is surveyed and tested in this paper.

Our aim is to explore the statistical properties and limitations of this model, so its potential usefulness can be fully exploited. We examine its fit to a number of traffic measurements in order to understand which aspects of the model agree with reality, where the model is robust to deviations, and in what respects it may require extension and redefinition. A core issue is the relation between the micro-level infinite source Poisson model and limiting aggregated models outlined below.

Data sets similar to the ones analysed in this paper are rapidly being accumulated by the digital networking industry and by academic researchers. A further aim is to survey some statistical methods which we have found helpful, and which may be of use to the engineers and scientists who are coping with these data sets.

The background is the identification of self-similarity in various types of teletraffic flow rates at resolutions above a certain critical threshold. This has created widespread interest in the possible

AMS 1991 subject classifications. Primary 62N99; secondary 60K250.

Key words and phrases. data transmission modelling, internet traffic, heavy tails, regular variation, Pareto tails, self-similarity, scaling.

Research supported by the Gothenburg Stochastic Centre and by the EU TMR network ERB-FMRX-CT96-0095 on “Computational and statistical methods for the analysis of spatial data”. Sidney Resnick’s research was also partially supported by NSF grant DMS-97-04982 and NSA Grant MDA904-98-1-0041 at Cornell University.

origins and effects of the self-similarity. Willinger et. al. ([57, 58, 59, 77, 83, 84, 85]) discussed self-similarity of packet counts per unit time in LANS and WANS and a parallel discussion of self-similarity of bytes per unit time in WWW traffic was conducted by Crovella et al ([15, 16, 19, 17]). Crovella, Kim and Park ([18]) conducted a large simulation study to assess the causes and effects of self-similarity in situations that involved slowdown nodes, buffers, varying rates and varying tail parameters. Erramilli and Willinger ([25]) used experimental queueing analysis to show why classical models without long range dependence would seriously underestimate delays. Resnick and Samorodnitsky ([67]) constructed an example of a single exponential server fed by a long range dependent input which had queue lengths and waiting times which were heavy tailed. Mathematical studies of the connection between on-off inputs with heavy tailed on-periods appeared in [77] and [45, 46, 37, 61, 48, 47]. The infinite source Poisson model was studied in [70] and [61].

Attempts to explain observed self-similarity in network traffic have largely focused on heavy tailed transmission times of sources sending data to one or more servers. The common assumption is that transmission times have iid random lengths with common distribution F . Often F has heavy tails in which case it is assumed F has a Pareto, or more generally, regularly varying tail so that

$$(1.1) \quad \bar{F}(x) := 1 - F(x) \sim x^{-\alpha} L(x), \quad x \rightarrow \infty,$$

where $L(x)$ a slowly varying function, so that

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

where $\bar{F} = 1 - F(x)$. We distinguish three cases which F may satisfy.

- (i) F has such a heavy tail that the mean is infinite and $0 < \alpha < 1$. Such heavy tails appear in the BU study of file sizes for the month of November (see the plots in [70]) and are reported by a Calgary study [7] of file sizes found on various servers.
- (ii) F has a heavy tail with $1 < \alpha < 2$ so that the variance is infinite but the mean is finite. This has been a popular assumption for two reasons. The practical reason is the extensive traffic measurements of *on* periods reported in [84] where measured values of α were almost always in the interval $(1, 2)$. The theoretical reason is that mathematical analysis of models has been based on renewal theory and without a finite mean, stationary versions of renewal processes do not exist and (uncontrolled) buffer content stochastic processes would not be stable. See for example [37].
- (iii) F has relatively thin tails so that the variance is finite. This includes classical models for telecommunication.

Section 2 defines the infinite source Poisson model and defines the basic descriptor processes: number of active sources at time t , cumulative inputted traffic to the server in $[0, t]$, traffic rate, buffer content at t , time for buffer overflow of level $\gamma > 0$. The traffic rate process is the cumulative inputted traffic in small time intervals. This is obtained from the cumulative traffic by differencing. This section also considers known Gaussian and jump process approximations to the basic descriptors. For Brownian motion approximations we could not find a proof in the literature, and hence have provided one.

In Section 3 the statistical methods we have used are presented.

Subsequent sections analyse four data sets in order to see what features of the model are consistent with the data. The data sets are:

- The Boston University data recording http sessions in two labs between November 1994 and February 1995. This data is available at <http://ita.ee.lbl.gov/html/traces.html>. We analysed a 8h 20min part of the trace, with mean traffic rate 30 kbit/s.
- The UC Berkeley data with an 18 day trace collected in Nov. 1996 which contains the home IP HTTP traffic processed by UC Berkeley during this period. It is available at <http://ita.ee.lbl.gov/html/contrib/UCB.home-IP-HTTP.html>. Here we analyzed a three hour peak portion of the data, with mean traffic rate of 341 kbit/s.
- A low-resolution and a high-resolution data set from the two universities in Munich which were kindly supplied to us by Helmuth Gogl. The low-resolution traces were collected around the clock on Wednesday, November 12, 1997 and Wednesday, December 17, 1997, and consist of the total number of cells which passed an ATM link in every two-second interval, in the transmit and receive directions with mean traffic rates 5.6 and 8.8 Mbit/s, respectively. The high-resolution trace records 1690729 inter-arrival times of ATM cells from a measurement that captured all traffic in the sending direction of the link over a period of approximately 137 seconds with a resolution of 1 microsecond. The measurement was done on Tuesday December 23, 1997 starting at 14:48:15 and ending at 14:50:33 with a mean traffic rate of 5 MBit/s.
- The number of bits transmitted and time stamps of starts and completions of HTTP file transfers to and from a corporate www Ericsson server, collected on Thursday and Friday, October 15 and 16, 1998. The trace was very non-stationary, and we restricted the analysis to a 33 minute part of the trace, with a mean traffic rate 273 kbit/s.

For each data set we attempt to assess compatibility of the data with the model. As expected, there is not always a perfect fit to say the least. Depending on the richness of the available data, we wished to sort out the following issues for each data set. Limitations of time and manpower meant we were not always completely successful.

- (1) How do you identify time points which are statistically verified to form a Poisson process? We examined initialisations of sessions and beginnings of data bursts. When humans log in, it is plausible and widely believed to be an internet invariant that the initialization times form a Poisson process but totally implausible that machine generated downloads triggered by an initial http request would follow the Poisson assumption. However, even for events caused by humans, aggregation beyond some scale will have to take place in order for Poisson behaviour to be in force.
- (2) We verify heavy tails for such quantities as file sizes, and file transfer times. Lengths of time depending on human activity such as the length of a login session at a student lab in the BU study is less likely to exhibit heavy tails. We expect to observe that fine scalings appropriate for measuring machine generated activity (milliseconds) are inappropriate for measuring human activity where seconds or even minutes are appropriate and this may explain why file transmission times look heavy tailed but login sessions do not. Further investigations into the nature of distributions where we attempt to fit more than the tail are worthwhile and useful but it should be noted that it is likely that many macroscopic characteristics of a network will only be sensitive to tail behaviour.
- (3) We seek to study the distribution of transmission rates which can be defined as file size divided by the transfer time of the file. A more difficult question concerns dependence between transmission rates and file size and to what extent the dependence on network load is measurable.
- (4) We seek to understand the long range dependence behaviour and the local Hölder behaviour of combined traffic and relate this to the heavy tails found in item (2). An interesting statistical question is whether estimates of the tail parameter α , or equivalently of $\gamma = 1/\alpha$, based on

such techniques as maximum likelihood estimation in generalized Pareto models and QQ-plots ([39, 69, 39, 53, 10]) are consistent with estimates of the Hurst coefficient H and the Hölder exponent, since theoretically, the model guarantees that these two can be expressed as simple functions of α for the basic limit approximation. We estimate H and the Hölder exponent using wavelet and quadratic variation based methods.

Note that for fractional Brownian motion (fBm), H and the Hölder exponent are identical so if in practice the estimates of these two quantities differ significantly, we have reason to doubt that fBm is the appropriate model.

- (5) Many of the descriptors in the infinite source Poisson model have Gaussian or in some cases jump process approximations. We seek to examine the data to see if there are characteristics compatible with these approximations and, where possible, decide which type of approximation is more accurate.
- (6) Much of the data exhibits evident non-stationarities. (For example, loads on the internet are heavy in afternoons, light in early morning.) How do you analyse non-stationary data? One simple technique, because of the abundance of data, is to restrict attention to a subset of the data where behaviour is likely to be more stationary. Although this is our primary technique, a more complete analysis could be performed by trying to view a natural period of time (one day?) as a period and to then deseasonalize the data.

2. BACKGROUND AND BASIC MODELS.

For later convenience, we first collect some basic concepts, beginning with discussions of self-similarity, Hurst and Hölder exponents and long range dependence. We define Lévy stable motion and fractional Brownian motion. We then define the infinite source Poisson model and give basic properties and descriptor quantities, and asymptotic approximations.

The main parameters of the infinite source Poisson model are the connection rate, λ , which is the intensity of starts of file transfers, and the tail parameter α of the transmission times. The three cases $\alpha < 1$, $1 < \alpha < 2$, and $2 < \alpha$ lead to different asymptotic approximations which below are treated in separate sections. The approximations are for large time intervals, $[0, T]$ and for the two last cases also assume that λ is large. A further issue is that in the middle case the nature of the approximation depends on the relation between λ and T .

2.1. Self-similarity, Hurst and Hölder exponents, long-range dependence. A stochastic process $\{X(t); 0 \leq t < \infty\}$ is self-similar if there exists a constant H , the Hurst parameter, such that the finite dimensional distributions of the time changed and rescaled process $\theta^{-H}X(\theta t)$ are the same as for the original process, i.e. in formulas, if

$$\theta^{-H}X(\theta \cdot) \stackrel{fidi}{=} X(\cdot), \text{ for } 0 < \theta.$$

In particular, since a centered Gaussian process is uniquely determined by its covariance function, it is self-similar if and only if its covariance function C satisfies

$$(2.1) \quad \theta^{-2H}C(\theta s, \theta t) = C(s, t) \text{ for } 0 < \theta.$$

A second order process, Gaussian or not, which satisfies (2.1) is called second order self-similar.

A fractional Brownian motion (fBm) B_H is a centered continuous Gaussian process with covariance function

$$C(s, t) = \frac{\sigma^2}{2}(|t|^{2H} + |s|^{2H} - |t - s|^{2H}), \quad 0 < H < 1.$$

Since C satisfies (2.1), B_H is self-similar with Hurst parameter H . By setting $H = 1/2$ an ordinary Brownian motion, with independent increments, is obtained. It can be seen from the form of C that

fBm has stationary (but not independent) increments also for $H \neq 1/2$. It follows that fractional Gaussian noise (fGn), i.e. the difference sequence $Y_k = X((k+1)\Delta) - X(k\Delta)$ is stationary (here Δ is the fixed length of differencing). More details can be found for instance in [75].

As discussed in the introduction, network traces looked at on widely varying time scales above a certain resolution are in the literature often claimed to have similar statistical properties, and all the limiting processes which are discussed below are self-similar. This is completely as should be expected. If a process is obtained as a distributional limit by dilating time linearly and scaling space, then it has to be asymptotically self-similar. See [55, 23].

The semivariogram V of a second order process is defined by

$$V(t, \tau) = \frac{1}{2}E(X(t+\tau) - X(t))^2.$$

It is easy to see that C can be computed from V and vice versa. For a self-similar process, $\theta^{-2H}V(\theta t, \theta \tau) = V(t, \tau)$. If the process has stationary increments, V does not depend on its first variable, $V(t, \tau) = V(\tau)$, and if it also is self-similar, then $V(\tau) = V(1)\tau^{-2H} := c\tau^{-2H}$. If the process is also a centered Gaussian process, this means that it is a fBm. In general, a process is said to have the local Hölder (mean square) index H_o at t if the semivariogram satisfies

$$(2.2) \quad V(t, \tau) = c\tau^{2H_o} + o(\tau^{2H_o}), \text{ as } \tau \rightarrow 0,$$

for each t . For a fBm, $H = H_o$, as is easily seen. The Hölder index of a Gaussian process gives precise information on sample paths (see *e.g.* Adler [5] for extremal distributions, Ibragimov and Rozanov [41] for Hölder smoothness of the sample paths), on the rate of convergence of non-parametric estimates of the covariance function (Istas and Laredo [44]) and on the asymptotic behaviour of wavelet coefficients of X (Istas, [42]). There are other definitions of Hölder indices more suitable to the study of path properties of non-Gaussian processes such as multifractals (cf. [72]). Data network applications are discussed in [26],[29] and [73].

A process X_α is an α -stable Lévy motion if it has stationary independent increments which follow a non-normal stable distribution with index α , $0 < \alpha < 2$. Clearly $\theta^{-1/\alpha}X_\alpha(\theta \cdot)$ also has stationary increments, for any $\theta > 0$. Further, by the characterizing property of the stable distributions, $\theta^{-1/\alpha}X_\alpha(\theta t)$ has the same distribution as $X_\alpha(t)$ for any $t > 0$. It follows that X_α is self-similar with Hurst parameter $H = 1/\alpha$. An iid sequence of α -stable random variables is called α -stable noise. Thus, in particular the sequence $X_\alpha((k+1)\Delta) - X_\alpha(k\Delta); k = 0, 1, \dots$ is α -stable noise for $\Delta > 0$. Extensive information about stable processes may be found in [75].

Since stable variables have infinite variances, the variogram and the local Hölder exponent are not defined for stable Lévy motions.

A (centered) second-order stationary process $X(t)$ displays long-range dependence (LRD) if its covariance function decreases at a polynomial rate at large lags,

$$(2.3) \quad C(\tau) \sim K\tau^{\beta-1}, \quad \tau \rightarrow \infty,$$

where K is a constant and $0 < \beta < 1$. Another way to think about this is that the spectral density $\Gamma(\nu)$ has a polynomial divergence at the origin,

$$\Gamma(\nu) \sim K'|\nu|^{-\beta}, \quad \nu \rightarrow 0,$$

for some other constant K' . The paradigm of such processes is fGn with $1/2 < H < 1$. The LRD parameter then is related to H by $\beta - 1 = 2H - 2$, i.e. $H = (\beta + 1)/2$. By analogy with fractional Gaussian noise, the quantity $(\beta + 1)/2$ is sometimes referred to as the ‘‘Hurst parameter’’ of the process as soon as the behaviour (2.3) is observed, even if the process is not self-similar.

2.2. The infinite source Poisson model. We now review the elements of a data transmission model used in [46], [47], [38], [70], and [61]. Let $\{\Gamma_k, k \geq 1\}$ be the points of a rate λ homogeneous Poisson process on $\mathbb{R}_+ = [0, \infty)$ so that $\{\Gamma_{k+1} - \Gamma_k, k \geq 1\}$ is a sequence of iid exponentially distributed random variables with parameter λ . (In the stationary case the Poisson process instead should be defined on $\mathbb{R} = [-\infty, \infty)$, which leads to some straightforward change of notation.) We imagine that a communication system has *sources* or nodes, and at time Γ_k a *connection* is made and a source begins a transmission at unit rate to or from the server. The duration of this transmission is a random variable L_k with distribution F , usually of the form $\bar{F}(x) = x^{-\alpha}L(x)$. (We find it convenient to use the same letter L to denote a generic slowly varying function, and r.v.'s with the d.f. above.) When F has a finite first moment, it is convenient to set

$$\mu = E(L_1) = \int_0^\infty xF(dx).$$

The input rate could be made to deterministically vary over the transmission time of length L_k as in [52] but we have not pursued this nor the potentially more useful idea of having random input rates where at each transmission initiation a rate is sampled from a random distribution. We sometimes refer to the Poisson rate λ as the *connection intensity*.

We note that

$$(2.4) \quad M = \sum_{k=1}^{\infty} \epsilon_{(\Gamma_k, L_k)},$$

the counting function on $\mathbb{R}_+ \times (0, \infty)$ corresponding to the points $\{(\Gamma_k, L_k), k \geq 1\}$, is a two dimensional Poisson process on $\mathbb{R}_+ \times (0, \infty]$ with mean measure $\lambda \mathbb{L} \times F$, where \mathbb{L} stands for Lebesgue measure. (Cf. [65].)

It is useful to note that this model is stable under aggregation. If the traffic from two independent infinite source Poisson models are superposed, then the result also follows an infinite source Poisson model, with a connection intensity which is equal to the sum of the two intensities, and a transmission length distribution which is a mixture of the two transmission length distributions.

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

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

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

$$(2.5) \quad \lambda \mathbb{L} \times F(\{(\gamma, l) \in \mathbb{R}_+ \times (0, \infty] : \gamma \leq t < \gamma + l\}) = \lambda \int_0^t \bar{F}(t - \gamma) d\gamma = \lambda \int_0^t \bar{F}(s) ds =: m(t).$$

During a transmission, the transmitting source is sending data to the server at unit rate. The total *cumulative traffic* in $[0, t]$ is

$$(2.6) \quad A(t) := \int_0^t N(s) ds.$$

The traffic rate process is

$$\{A((k+1)\Delta) - A(k\Delta), k \geq 0\},$$

for some fixed $\Delta > 0$. Assume the server works at constant *output rate* r . The *buffer content* at time t , $X(t)$, satisfies the storage equation

$$(2.7) \quad dX(t) = N(t)dt - r1_{[X(t)>0]}dt,$$

or ([36], [8], [63])

$$(2.8) \quad \begin{aligned} X(t) &= \bigvee_{s=0}^t [A(t) - A(s) - r(t-s)] \\ &= \bigvee_{s=0}^t \int_s^t (N(s) - r) ds, \end{aligned}$$

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

A final quantity which is a useful performance measure but which is hard to make inferences about given the data is *time to buffer overflow* of a level γ which is defined to be

$$(2.9) \quad \begin{aligned} \tau(\gamma) &:= \inf\{t \geq 0 : X(t) \geq \gamma\} \\ &= \inf\{t \geq 0 : \bigvee_{s=0}^t X(s) \geq \gamma\} = \left(\bigvee_{s=0}^{(\cdot)} X(s) \right)^{\leftarrow}(\gamma), \end{aligned}$$

where for a monotone nondecreasing right continuous function f , we write f^{\leftarrow} for the left continuous inverse.

We now describe known behaviour of the basic descriptors for the three cases discussed in the introduction.

2.4. The case $0 < \alpha < 1$. In this section we assume that (1.1), i.e. $\bar{F}(x) = x^{-\alpha}L(x)$, holds with $0 < \alpha < 1$ and survey the asymptotic behavior of the model. The analysis assumes that the connection intensity λ is constant and the time interval considered, $[0, T]$, tends to infinity. However, undoubtedly the same limits apply also if λ and T tend to infinity together. This is in contrast to the next case, $1 < \alpha < 2$, where the form of the limits depend on the precise relation between T and λ .

For the present case the mean transmission time is infinite and the descriptors N, A, X are unstable. From [70] we get the first order behaviour: In the function space $D[0, \infty)$ we have as $T \rightarrow \infty$

$$\begin{aligned} \frac{N(Tt)}{m(T)} &\xrightarrow{P} t^{1-\alpha}, & \frac{A(Tt)}{Tm(T)} &\xrightarrow{P} \frac{t^{2-\alpha}}{2-\alpha}, \\ \frac{X(Tt)}{Tm(T)} &\xrightarrow{P} \frac{t^{2-\alpha}}{2-\alpha}, & \frac{\tau(T\gamma)}{V^{\leftarrow}(T)} &\xrightarrow{P} ((2-\alpha)\gamma)^{1/(2-\alpha)}, \end{aligned}$$

where $V(T) = Tm(T) \sim \lambda T^{2-\alpha}L(T)/(1-\alpha)$ so that the inverse function V^{\leftarrow} is regularly varying with index $1/(2-\alpha)$. So the time necessary for the content to reach a critical high level γ is of algebraic order $\gamma^{1/(2-\alpha)}$.

The second order behaviour which gives Gaussian process approximations to the basic descriptor processes is also provided by [70]. Let $\{G(t), t \geq 0\}$ be a continuous path, Gaussian process with $G(0) = 0$ and with covariance function

$$(2.10) \quad C(s, t) = (s \vee t)^{1-\alpha} - |t-s|^{1-\alpha}, \quad 0 \leq s \leq t.$$

Because of the form of the covariance function, $G(\cdot)$ is self-similar with

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

However, $G(\cdot)$ is not stationary, does not have stationary increments and is not a fractional Brownian motion. The process

$$\hat{G}(t) := \int_0^t G(u) du$$

which has covariance function $\hat{C}(s, t)$ given by

$$\hat{C}(s, t) = \int_{u=0}^s \int_{v=0}^t C(u, v) du dv, \quad 0 \leq s \leq t$$

is also a self-similar, continuous path, Gaussian process satisfying $\hat{G}(0) = 0$ and

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

We have the following Gaussian approximations. If (1.1) holds and the distribution F satisfies a second order condition (given in [70]) then both the number of active sources, $N(t)$, and the cumulative traffic, $A(t)$, are asymptotically normal as $T \rightarrow \infty$,

$$\begin{aligned} \frac{N(Tt) - m(Tt)}{\sqrt{m(T)}} &\Rightarrow G(t), \\ \frac{A(Tt) - \int_0^{Tt} m(s) ds}{T\sqrt{m(T)}} &\Rightarrow \hat{G}(t), \end{aligned}$$

and in fact convergence is in the Skorohod metric on $C[0, \infty)$. The latter result of course at once gives that also the traffic rates $Y_k^{(T)} = A((k+1)T\Delta) - A(kT\Delta)$, for any $\Delta > 0$, are asymptotically normal. Further, if the output rate r is allowed to depend on T in such a way that for fixed $y > 0$, $r = r_T = m(Ty)$, then setting $X_T(t)$ to be the buffer content process at time Tt corresponding to traffic rate $r = r_T$, we have

$$\sqrt{m(T)} \left(\frac{X_T(t)}{Tm(T)} - c_T(t) \right) \Rightarrow \hat{G}(t) - \hat{G}(y), \quad t \geq y,$$

where $c_T(\cdot)$ is given by

$$c_T(t) = \left(\frac{\int_{Ty}^{Tt} m(s) ds}{Tm(T)} \right) - \left(\frac{m(Ty)}{m(T)} (t - y) \right).$$

This says that given an output rate r , define y by $y = m^{\leftarrow}(r)/T$, and then for $t > y$, the buffer content $X_T(t)$ is approximately distributed as

$$(2.11) \quad Tm(T) \left(\frac{\hat{G}(t) - \hat{G}(y)}{\sqrt{m(T)}} + c_T(t) \right).$$

Finally, we have for the first passage process of $X_T(\cdot)$, i.e. for $\tau_T(t) := T^{-1}\tau(t)$, the normalized time to buffer overflow, that

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

where χ is a process obtained from \hat{G} and

$$d_\infty(\gamma) = \frac{(y + \gamma)^{2-\alpha} - y^{2-\alpha}}{2 - \alpha} - y^{1-\alpha}\gamma.$$

2.5. The case $1 < \alpha < 2$. For this case we first derive some basic moment relations and exhibit and discuss the stable noise approximations which are valid for large time scales and small or moderate connection rates (i.e. formally, as $T \rightarrow \infty$ and λ is fixed). Then the fBm limit result of [77] which assumes that first $\lambda \rightarrow \infty$ and then $T \rightarrow \infty$ is briefly mentioned, and finally a theorem which includes these cases as well as intermediate cases where λ and T tend to infinity simultaneously is presented.

We have that $N(T)$, the number of active sources at T , is Poisson distributed with parameter $m(T)$. Since for the case $1 < \alpha < 2$, $m(T) \rightarrow m(\infty) = \lambda\mu < \infty$, it follows that as $T \rightarrow \infty$,

$$N(T) \Rightarrow \text{PO}(\lambda\mu),$$

a Poisson random variable with parameter $\lambda\mu$. Furthermore, for any $s < t$, as $T \rightarrow \infty$, an easy calculation shows

$$\begin{aligned} \text{Corr}(N(Ts), N(Tt)) &= \frac{\text{Cov}(N(Ts), N(Tt))}{\sqrt{m(Ts)m(Tt)}} \\ &= \frac{\int_{T(t-s)}^{Tt} \bar{F}(u) du}{\sqrt{m(Ts)m(Tt)}} \rightarrow 0, \end{aligned}$$

since \bar{F} is integrable on $[0, \infty)$ and $m(v) \rightarrow \lambda\mu$ as $v \rightarrow \infty$. It follows that one cannot expect to get the asymptotic behaviour of $A(\cdot)$ from the asymptotic behaviour of $N(\cdot)$.

Suppose we write the process $A(\cdot)$ defined in (2.6) as

$$\begin{aligned} A(t) &= \sum_{\Gamma_k \leq t} \left\{ \left((\Gamma_k + L_k) \wedge t \right) - \Gamma_k \right\} \\ &= \sum_{\Gamma_k \leq t} (L_k \wedge (t - \Gamma_k)) = \sum_{\Gamma_k \leq t} A_k(t - \Gamma_k), \end{aligned}$$

where $A_k(s) = s \wedge L_k$. This expresses $\{A(t), t \geq 0\}$ as a shot noise process and the results of Klüppelberg and Mikosch ([50, 51] are available. See also [52].

Note that from (2.5) and (2.6), as $T \rightarrow \infty$,

$$\begin{aligned} E(A(T)) &= \lambda \int_0^T E(A_k(s)) ds = \lambda \int_0^T E(L_k \wedge s) ds \\ &= \int_0^T E(N(u)) du = \int_0^T m(u) du \sim \lambda T \mu = \lambda T \int_0^\infty \bar{F}(t) dt \\ \text{Var}(A(T)) &:= \sigma^2(T) = \lambda \int_0^T E(A_k^2(s)) ds = \lambda \int_0^T E(L_k \wedge s)^2 ds \\ &\sim \frac{2\lambda}{(2-\alpha)(3-\alpha)} T^3 \bar{F}(T), \quad T \rightarrow \infty. \end{aligned}$$

Furthermore for $0 < s < t$

$$\begin{aligned} c(s, t) &:= E(A_1(s)A_1(t)) \\ (2.12) \quad &= E(L_1^2 1_{[L_1 \leq s]}) + sE(L_1 1_{[s \leq L_1 \leq t]}) + stP[L_1 > t]. \end{aligned}$$

Now, consider the case λ constant and $T \rightarrow \infty$. Then,

$$(2.13) \quad \frac{c(Ts, Tt)}{T^2 \bar{F}(T)} \rightarrow \frac{s}{\alpha - 1} \left(\frac{\alpha}{2 - \alpha} s^{1-\alpha} - t^{1-\alpha} \right) =: C(s, t).$$

Thus $c(s, t)$ is multivariate regularly varying with index $2 - \alpha$ and limit function $C(s, t)$. From [50, Proposition 2.2, page 129] we get that for $0 < x < y$

$$\hat{c}(x, y) := \text{Cov}(A(x), A(y))$$

is regularly varying with index $3 - \alpha$ and limit function ($x < y, h = y - x$)

$$\hat{C}(x, y) := (\text{const}) (x^{2H} + y^{2H} - h^{2H})$$

where $H = (3 - \alpha)/2$. Thus, $\hat{C}(x, y)$, properly normalised is the correlation function of fractional Brownian motion with Hurst parameter H . As pointed out in [52, page 8–10], this means the process $A(\cdot)$ properly standardised asymptotically is second order self similar. However, (cf [52])

$$(2.14) \quad \hat{G}_T(t) := \frac{A(Tt) - \lambda\mu Tt}{\sigma(T)}$$

does not converge (even in the sense of finite dimensional distributional convergence) to a limiting self similar Gaussian process since ([50, Theorem 3.3, page 132]) as $T \rightarrow \infty$,

$$(2.15) \quad \frac{1}{\sigma^2(T)} \int_{\epsilon\sigma(T)}^{\infty} y P[L_1 \wedge u > y] dy \not\rightarrow 0,$$

for all $\epsilon > 0$. It is not hard to see that in fact the expression on the left converges to

$$\int_0^1 (1 - y) y^{1-\alpha} dy \neq 0.$$

Even though there is no centering and scaling to make $A(T\cdot)$ asymptotically a Gaussian process, it is possible to get stable limits. Define

$$(2.16) \quad b(T) = \left(\frac{1}{1 - F} \right)^{\leftarrow} (T) = T^{1/\alpha} L_1(T),$$

for some slowly varying function L_1 . A variant of [52, Theorem 4] in particular gives, for λ fixed, $T \rightarrow \infty$, that

$$(2.17) \quad \frac{A(Tt) - \lambda\mu Tt}{b(T)} \Rightarrow X_\alpha(t),$$

in the sense of convergence of finite dimensional distributions, where $X_\alpha(\cdot)$ is an α -stable Lévy motion whose marginal distribution is totally skewed to the right. See also [71, 61]. So on large time scales, $A(T\cdot)$ looks like an α -stable Lévy motion.

Let us compare (2.17) and the fact that $\hat{G}_T(t)$ of (2.14) is not asymptotically a Gaussian process, even though by (2.13) the correlation function of $\hat{G}_T(\cdot)$ converges. Note that as $T \rightarrow \infty$

$$\sigma(T) \sim (\text{const}) \sqrt{T^3 \bar{F}(T)} \sim (\text{const}) T^{(3-\alpha)/2} L^{1/2}(T),$$

which may be compared to the asymptotic form of $b(\cdot)$ given in (2.16). Observe that for $\alpha \in (1, 2)$,

$$\frac{1}{\alpha} < \frac{3 - \alpha}{2}.$$

Also from (2.17), we find the first order growth rate of $A(T)$. Observe that the centering in (2.17) is of the form

$$v(T) = T\lambda\mu,$$

so that since $\frac{1}{2} < \frac{1}{\alpha} < 1$,

$$v(T)/b(T) = T\lambda\mu/(T^{1/\alpha} L_1(T)) := T^{1-1/\alpha} L_2(T) \rightarrow \infty,$$

with $L_2(\cdot)$ regularly varying. By (2.17),

$$\frac{A(T)}{v(T)/b(T)} \xrightarrow{P} 1.$$

Consequently, in $C[0, \infty)$,

$$\frac{A(Tt)}{v(T)/b(T)} \Rightarrow t^{1-\frac{1}{\alpha}}.$$

Provided the constant output rate r satisfies $\lambda\mu < r$, the $X(\cdot)$ process of (2.7) and (2.8) has negative drift and is stable. Being regenerative, $X(T)$ will have a limit distribution. However, high levels will still be exceeded by $X(\cdot)$ in algebraic time since, according to ([38]), the expected hitting time of high level γ (see (2.9)) satisfies

$$E\tau(\gamma) \asymp \gamma^{1-k} \left(\frac{1}{1-F(\gamma)} \right)^k$$

where

$$k := \inf\{j \geq 1 : \lambda E(L_1) + j > r\}$$

is the minimum number of sessions running simultaneously which are needed to flip the mean drift from negative to positive. Here we use the notation $f \asymp g$ to mean

$$0 \leq \liminf_{T \rightarrow \infty} f(T)/g(T) \leq \limsup_{T \rightarrow \infty} f(T)/g(T) < \infty.$$

The limit behaviour in (2.17) and the second order behaviour of $\hat{G}_T(\cdot)$ defined in (2.14) assume that the connection rate λ is constant and the time scale T is growing. If T is fixed and $\lambda \rightarrow \infty$, then one may expect $A(\cdot)$ to be asymptotically Gaussian. Thus, on small or moderate time scales, if the input rate is large, the cumulative inputted traffic should be approximately Gaussian. This in fact was found in [77] as well as the result that if after letting $\lambda \rightarrow \infty$ one lets $T \rightarrow \infty$, then $A(\cdot)$ is approximated by a fractional Brownian motion. On the other hand, as seen above, on large time scales, with modest connection rate, the cumulative inputted traffic is approximately a Lévy stable motion. Allowing the connection rates to vary with T produces the following result which give conditions under which either a fBm or Lévy stable motion is an appropriate approximation. The exact statement is quoted from [61].

Proposition 1. *Assume a family of infinite source Poisson models indexed by T where in the T th model, $\lambda = \lambda(T)$ depends on T . Assume the connection length distribution does not depend on T . Recall the definition of $b(\cdot)$ in (2.16). For the T th model, let $A_T(\cdot)$ be the cumulative input and $N_T(t)$ be the number of active sources at time t .*

(i) Assume the Poisson rate $\lambda = \lambda(T)$ depends on T so that one of the following equivalent slow growth conditions is satisfied:

1. $\lim_{T \rightarrow \infty} \frac{b(\lambda T)}{T} = 0.$
2. $\lim_{T \rightarrow \infty} \lambda T \bar{F}(T) = 0.$
3. $\lim_{T \rightarrow \infty} \text{Cov}(N_T(0), N_T(T)) = 0.$

Then the process $(A_T(Tt), t \geq 0)$ describing the cumulative input in $[0, Tt]$, $t \geq 0$, satisfies the limit relation

$$\frac{A_T(T\cdot) - T\lambda\mu(\cdot)}{b(\lambda T)} \xrightarrow{\text{fidi}} X_\alpha(\cdot),$$

where $X_\alpha(\cdot)$ is a Lévy α -stable motion. Here $\xrightarrow{\text{fidi}}$ denotes convergence of the finite dimensional distributions and in fact the result can be strengthened to convergence in the M_1 topology ([71])

(ii) Assume the Poisson rate $\lambda = \lambda(T)$ depends on T so that one of the following equivalent fast growth conditions is satisfied:

1. $\lim_{T \rightarrow \infty} \frac{b(\lambda T)}{T} = \infty$.
2. $\lim_{T \rightarrow \infty} \lambda T \bar{F}(T) = \infty$.
3. $\lim_{T \rightarrow \infty} \text{Cov}(N_T(0), N_T(T)) = \infty$.

Then the process $(A_T(Tt), t \geq 0)$ describing the cumulative input in $[0, Tt]$, $t \geq 0$, when properly normalized as

$$\frac{A_T(T\cdot) - T\lambda\mu(\cdot)}{\sqrt{\lambda T^3 \bar{F}(T)}}$$

converges in $D[0, \infty)$ to a fractional Brownian motion with self similarity parameter $H = (3 - \alpha)/2$.

Heavy traffic approximations for $A(\cdot)$ and $X(\cdot)$ are considered in [54].

2.6. The case $\alpha > 2$. We again assume λ is fixed. However, as for the case $\alpha < 1$, the result undoubtedly holds also if λ is allowed to increase with T . As in Case (ii), $N(T)$ is asymptotically Poisson and as $T \rightarrow \infty$, $N(T\cdot)$ is asymptotically uncorrelated. We cannot expect to get the asymptotic behaviour of $A(\cdot)$ from $N(\cdot)$.

For the case $\alpha > 2$ we show that the cumulative input $A(\cdot)$ can be approximated by a Brownian motion. We set

$$(2.18) \quad A^{(T)}(t) := \frac{A(Tt) - \lambda T t \mu}{\sqrt{\lambda T E(L_1^2)}}$$

and show that, in $C[0, \infty)$, as $T \rightarrow \infty$,

$$(2.19) \quad A^{(T)}(\cdot) \Rightarrow B(\cdot),$$

a standard Brownian motion. Furthermore, this implies a functional central limit theorem for $X(\cdot)$: If $\lambda\mu > r$, so that the system is unstable,

$$(2.20) \quad \frac{X(Tt) - Tt(\lambda\mu - r)}{\sqrt{\lambda T E(L_1^2)}} \Rightarrow B(t)$$

in $C[0, \infty)$. If $\lambda\mu < r$, then the limit is 0.

We now discuss the verification of these claims and continue to apply the results of [50] with $A_k(s) = s \wedge L_k$. We have

$$\begin{aligned} E(A(t)) &= \lambda \int_0^t E(u \wedge L_1) du = \lambda \int_0^t \bar{F}(u) du \sim \lambda\mu, \\ \sigma^2(t) &= E \left(\int_0^t (L_1 \wedge u)^2 du \right) \\ &= \lambda \int_{u=0}^t \left(\int_{v=0}^u \bar{F}(v) 2v dv \right) du \end{aligned}$$

so that,

$$\begin{aligned} \lim_{t \rightarrow \infty} \frac{\sigma^2(t)}{t} &= \lambda \lim_{u \rightarrow \infty} \int_0^u \bar{F}(v) 2v dv \\ &= \lambda \int_0^\infty x^2 F(dx) = \lambda E(L_1^2), \end{aligned}$$

and we conclude

$$(2.21) \quad \sigma^2(t) \sim t\lambda E(L_1^2).$$

Using (2.12) we get, as $T \rightarrow \infty$ that for $s < t$

$$(2.22) \quad E\left(A_1(Ts), A_1(Tt)\right) \rightarrow \int_0^\infty x^2 F(dx) = E(L_1^2),$$

since when $\alpha > 2$ we have $T^2 \bar{F}(Tt) \rightarrow 0$ and

$$Ts \int_{Ts}^{Tt} x F(dx) \leq Ts \int_{Ts}^{Tt} \frac{x^2}{Ts} F(dx) = 1 \cdot \int_{Ts}^{Tt} x^2 F(dx) \rightarrow 0.$$

Therefore, when $\alpha > 2$, the function $E\left(A_1(s)A_2(t)\right)$ is two dimensional regularly varying with index 0 and limit function which is identically zero. Thus ([50, Proposition 2.2]) $E\left(A(s)A(t)\right)$ is regularly varying with index 1 and limit function ($x < y$)

$$\hat{C}(x, y) := \int_0^x 1 du = x,$$

which is the covariance function of Brownian motion. Unlike the case $1 < \alpha < 2$, where convergence in distribution to Brownian motion did not hold, it will hold in the present case $2 < \alpha$ since we may verify that the expression in (2.15) converges to 0. This is relatively easy to verify since

$$\begin{aligned} \frac{1}{t} \int_{\epsilon\sqrt{t}}^\infty y \int_u^t \bar{F}(y) du dy &= \frac{1}{t} \int_{\epsilon\sqrt{t}}^\infty y(t-y) \bar{F}(y) dy \\ &\leq \frac{t}{t} \int_{\epsilon\sqrt{t}}^\infty y \bar{F}(y) dy = \int_{\epsilon\sqrt{t}}^\infty y \bar{F}(y) dy \\ &\rightarrow 0, \quad (t \rightarrow \infty), \end{aligned}$$

since $E(L_1^2) < \infty$. From [50, Theorem 3.3], we get

$$(2.23) \quad A^{(T)}(t) := \frac{A(Tt) - \lambda Tt E(L_1)}{\sqrt{\lambda T E(L_1^2)}} \Rightarrow B(t),$$

in the sense of convergence of finite dimensional distributions where $B(\cdot)$ is a standard Brownian motion.

A version of (2.23) in the J_1 topology also holds according to [50, Proposition 3.4], provided we show as $T \rightarrow \infty$

$$(2.24) \quad \frac{1}{\sqrt{T}} \int_0^T (EA_1(\infty) - EA_1(u)) du \rightarrow 0$$

and

$$(2.25) \quad \frac{1}{\sqrt{T}} \bigvee_{s=0}^T \left| \sum_{i=1}^{M([0,s] \times \mathbb{R}_+)} (A_i(\infty) - A_i(s - \Gamma_i)) \right| \xrightarrow{P} 0.$$

Since $A_1(\infty) = L_1$, (2.24) becomes

$$\frac{1}{\sqrt{T}} \int_0^T E(L_1 - L_1 \wedge u) du = \frac{1}{\sqrt{T}} \int_0^T E\left((L_1 - u) 1_{[L_1 > u]}\right) du$$

which by the Schwartz inequality is bounded by

$$\leq \frac{1}{\sqrt{T}} \int_0^T \sqrt{E(L_1^2) \bar{F}(u)} du = \frac{\sqrt{E(L_1^2)}}{\sqrt{T}} \int_0^T \sqrt{\bar{F}(u)} du.$$

Since \bar{F} is regularly varying, so is $\sqrt{\bar{F}}$ and its index is $-\alpha/2 < -1$. So $\sqrt{\bar{F}(u)}$ is integrable and the ratio converges to 0 as required.

We now show how to verify (2.25) for the case $\alpha > 2$. The left side of (2.25) is

$$\frac{1}{\sqrt{T}} \bigvee_{s=0}^T \left| \sum_{i=1}^{M([0,s] \times \mathbb{R}_+)} \left(L_i - (s - \Gamma_i) \wedge L_i \right) \right|$$

and

$$\begin{aligned} L_i - (s - \Gamma_i) \wedge L_i &= \begin{cases} 0, & \text{if } s - \Gamma_i \geq L_i, \\ L_i - (s - \Gamma_i), & \text{if } s - \Gamma_i < L_i, \end{cases} \\ &= \begin{cases} 0, & \text{if } \Gamma_i + L_i \leq s, \\ \Gamma_i + L_i - s, & \text{if } \Gamma_i + L_i > s. \end{cases} \end{aligned}$$

Therefore (2.25) can be rewritten

$$(2.26) \quad \frac{1}{\sqrt{T}} \bigvee_{s=0}^T \sum_{i=1}^{M([0,s] \times \mathbb{R}_+)} \left(\Gamma_i + L_i - s \right) 1_{[\Gamma_i + L_i > s]} \xrightarrow{P} 0.$$

Since $\Gamma_n \sim n$ almost surely, it is easy to see that (2.26) is implied by

$$(2.27) \quad \frac{1}{\sqrt{n}} \bigvee_{s \in [0, \Gamma_n]} \sum_{i=1}^{M([0,s] \times \mathbb{R}_+)} \left(\Gamma_i + L_i - s \right) 1_{[\Gamma_i + L_i > s]} \xrightarrow{P} 0.$$

The left side of (2.27) is

$$\frac{1}{\sqrt{n}} \bigvee_{j=0}^{n-1} \bigvee_{s \in [\Gamma_j, \Gamma_{j+1}]} \sum_{i=1}^j \left(\Gamma_i + L_i - s \right) 1_{[\Gamma_i + L_i > s]} \leq \frac{1}{\sqrt{n}} \bigvee_{j=0}^{n-1} \sum_{i=1}^j \left(\Gamma_i + L_i - \Gamma_j \right) 1_{[\Gamma_i + L_i > \Gamma_j]}.$$

Therefore, (2.26) holds if

$$(2.28) \quad \sum_{j=0}^n P \left[\sum_{i=1}^j (\Gamma_i + L_i - \Gamma_j) 1_{[\Gamma_i + L_i > \Gamma_j]} > \epsilon \sqrt{n} \right] \rightarrow 0$$

as $n \rightarrow \infty$ for every $\epsilon > 0$. Now the j th probability term in (2.28) is

$$P \left[\sum_{i=1}^j (L_i - (\Gamma_j - \Gamma_i)) 1_{[L_i > \Gamma_j - \Gamma_i]} > \epsilon \sqrt{n} \right] = P \left[\sum_{i=1}^j (L_i - \Gamma_{j-i}) 1_{[L_i > \Gamma_{j-i}]} > \epsilon \sqrt{n} \right]$$

and because $\{\Gamma_n\}$ is independent of $\{L_j\}$, this the same as

$$\begin{aligned} &= P \left[\sum_{i=1}^j (L_i - \Gamma_i) 1_{[L_i > \Gamma_i]} > \epsilon \sqrt{n} \right] \leq P \left[\sum_{i=1}^{\infty} (L_i - \Gamma_i) 1_{[L_i > \Gamma_i]} > \epsilon \sqrt{n} \right] \\ (2.29) \quad &= P \left[\iint_{D^\uparrow} (y - s) M(ds dy) > \epsilon \sqrt{n} \right] \end{aligned}$$

where

$$D^\uparrow = \{(s, y) \in \mathbb{R}_+^2 : y > s\}.$$

Note

$$EM(D^\uparrow) = \int_{s=0}^{\infty} \int_{y>s} \lambda ds F(dy) = \lambda \int_0^{\infty} \bar{F}(s) ds = \lambda \mu$$

and therefore, $M|_{D^\uparrow}$ is a finite point measure. It may be represented as

$$M|_{D^\uparrow} \stackrel{d}{=} \sum_{i=1}^{PO(\lambda\mu)} \epsilon_{(\xi_i, \eta_i)},$$

where $PO(\lambda\mu)$ is a Poisson random variable with parameter $\lambda\mu$, independent of the iid pairs $\{(\xi_i, \eta_i), i \geq 1\}$ which have common distribution

$$\left. \frac{\lambda ds F(dy)}{\lambda\mu} \right|_{D^\uparrow} = \left. \frac{ds F(dy)}{\mu} \right|_{D^\uparrow}.$$

So for $x > 0$,

$$P[\eta_i - \xi_i > x] = \iint_{\substack{s>0 \\ y>s+x}} ds \frac{F(dy)}{\mu} = \int_x^{\infty} \frac{\bar{F}(s)}{\mu} ds.$$

Set $\zeta_i = \eta_i - \xi_i$. The probability in (2.29) is

$$P\left[\sum_{i=1}^{PO(\lambda\mu)} \zeta_i > \epsilon\sqrt{n} \right]$$

and (2.28) is bounded by

$$(2.30) \quad nP\left[\sum_{i=1}^{PO(\lambda\mu)} \zeta_i > \epsilon\sqrt{n} \right].$$

If $\alpha > 2$, then $\text{Var}(\zeta_i) < \infty$ and there exists $\eta > 0$ such that $E(\zeta_i^{2+\eta}) < \infty$. Pick θ such that

$$(2.31) \quad 0 < \theta < \frac{\eta}{2(2+\eta)}.$$

Then

$$\begin{aligned} nP\left[\sum_{i=1}^{PO(\lambda\mu)} \zeta_i > \epsilon\sqrt{n} \right] &= nP\left[\sum_{i=1}^{PO(\lambda\mu)} \zeta_i > \epsilon\sqrt{n}, PO(\lambda\mu) \leq n^\theta \right] + nP\left[\sum_{i=1}^{PO(\lambda\mu)} \zeta_i > \epsilon\sqrt{n}, PO(\lambda\mu) > n^\theta \right] \\ &\leq nP\left[\sum_{i \leq n^\theta} \zeta_i > \epsilon\sqrt{n} \right] + P[PO(\lambda\mu) > n^\theta] \\ &= I + II. \end{aligned}$$

Now by Chebychev's inequality

$$II \leq \frac{nE(e^{PO(\lambda\mu)})}{\exp\{n^\theta\}} = \frac{ne^{\lambda\mu(e-1)}}{\exp\{n^\theta\}} \rightarrow 0,$$

as $n \rightarrow \infty$. For I we have

$$I \leq \frac{nE\left(\sum_{i \leq n^\theta} \zeta_i\right)^{2+\eta}}{\left(\epsilon\sqrt{n}\right)^{2+\eta}} = (\text{const}) \frac{E\left(\sum_{i \leq n^\theta} \zeta_i\right)^{2+\eta}}{n^{\eta/2}}.$$

Let $\|\zeta\|_p$ be the L_p norm of a random variable ζ and the bound in I is of the form

$$\frac{(\text{const})}{n^{\eta/2}} \left(\left\| \sum_{i \leq n^\theta} \zeta_i \right\|_{2+\eta} \right)^{2+\eta}$$

which by Minkowsky's inequality is dominated by

$$\frac{(\text{const})}{n^{\eta/2}} \left(n^\theta \|\zeta_1\|_{2+\eta} \right)^{2+\eta} = (\text{const}') \frac{n^{\theta(2+\eta)}}{n^{\eta/2}} \rightarrow 0$$

as $n \rightarrow \infty$, due to (2.31).

Having shown (2.25) for $\alpha > 2$, we get tightness and hence the functional form of (2.23).

This leads to a functional limit for the content process $X(\cdot)$ of (2.8) and verifies the claims (2.19) and (2.20) as follows. Define

$$\xi_T(t) = A(Tt) - rTt$$

so that from (2.23) we have

$$\frac{\xi_T(t) - Tt(\lambda\mu - r)}{\sqrt{\lambda T E(L_1^2)}} \Rightarrow B(t)$$

in $C[0, \infty)$. If $\lambda\mu > r$, so that the system is unstable since the input overwhelms the traffic rate, then applying [82, Theorem 6.4 (ii)] yields

$$(2.32) \quad \frac{X(Tt) - Tt(\lambda\mu - r)}{\sqrt{\lambda T E(L_1^2)}} \Rightarrow B(t)$$

in $C[0, \infty)$. Note, the condition $\lambda\mu > r$ guarantees $A(Tt) - rTt \xrightarrow{P} \infty$ and seems necessary to get a nontrivial limit in (2.32) due to the denominator becoming infinite. If $\lambda\mu < r$, then the limit in (2.32) is the function which is identically 0 by [82, Theorem 6.4 (iii), page 81].

3. ESTIMATION METHODS

This section gives a brief introduction to the statistical methods we have used. Many are standard, but will still be mentioned briefly and references will be given, in keeping with the aim to give a guide for analysis. However some are of more recent origin. Specifically, we have used rather recently developed quadratic variation and wavelet methods to estimate the Hurst and Hölder exponents. These methods will be presented in more detail.

In some situations, only traces from aggregated traffic are measured but a micro-level model is needed for simulation. Hence, a further topic for the present section is to what extent it is possible to infer parameters of the infinite source Poisson model from measurements of aggregate traffic.

In practice, a frequent goal of network measurement analysis is to construct a simulator for the design of networks. A simulator needs the parameter values for the model as an input. The statistical methods we describe provide such values, which can be used directly, or as a basis for extrapolation to what parameter values may be expected in the future, for new groups of subscribers. In this section, we also make some brief comments about design of simulators.

The methods we discuss throughout assume stationarity of the measured traffic. In reality the characteristics of the traffic, especially of the process of starts of file transfers, vary substantially with the time of day, between different days in the week, and between seasons, and contain strong trends. We circumvent this problem by selecting visually stationary parts of the traces for study. In fact, this may correspond to what is practically interesting: it's the behavior during the (approximately stationary) peak periods which is of primary interest.

3.1. Testing for independence of heavy-tailed variables. The infinite source Poisson model assumes that the transfer times are independent, and similarly the stable limit for aggregated traffic implies that successive traffic rate measurements are independent.

A basic approach to testing for independence is to use the correlation function. However correlations of heavy-tailed data may have a more complex behavior than in the light-tailed case, see e.g. [13, 22, 28, 68, 71] and have to be interpreted with some care. A standard useful way to circumvent this problem is to make an appropriate marginal transformation before computing correlations. In the present situation this amounts to using the logarithms of the data.

A drawback with this approach is that taking logarithms obscures the impact of the very large transfer times which are of major importance. Hence it is desirable to complement with methods which do not use transformations. We employed two methods to check for independence which use the original observations. Both are based on the heavy tailed acf, i.e., on the autocorrelation function computed without subtracting means.

The first, informal, method can detect many forms of nonlinear dependence in heavy tailed observations. It simply is to split the data into parts, say two to five parts, and to compute the heavy tailed acf on each of the parts. If the observations in fact are i.i.d., or come from a linear process, then these acf's should look the same. The second method is to use simulations from the known limiting distribution of the heavy-tailed acf under independence to construct confidence intervals, and to reject independence if the observed acf deviates from these intervals. For more details, see [28] and for a more formal treatment consult [71]. (An alternative, attractive, method is to use permutation tests for, say, the maximum of the heavytailed acf over some number of lags to judge if this is larger than what is caused by randomness alone. However, for very large data sets, such as the present ones, this method is computationally burdensome.)

3.2. Marginal distributions and estimation of means of heavy-tailed variables. The file sizes, file transfer times, file transfer rates ($=(\text{file size})/(\text{file transfer time})$), and traffic rates (cumulative inputted traffic in small time intervals) have similar statistical properties, and are amenable to the same types of analysis.

The precise shapes of the distributions are affected by special, and rapidly changing conditions, such as, for example, automatic hookup to Netscape's homepage. Thus only nonparametric (histogram or kernel density) estimation seem reasonable for the central parts of the distribution. However, at high aggregation levels, these details are less important, and what is needed are good estimates of expected values and tails of the distributions.

For heavy-tailed non-negative iid data X, X_1, \dots, X_n (such as the ones analysed below) with tails $P(X > x) \sim Lx^{-\alpha}$ with $1 < \alpha < 2$, where existing statistical techniques require that we now assume that the slowly varying function L in (1.1) is a constant L , the standard finite variance estimates of variation of the mean are not applicable. It is still possible to estimate the expected value $\mu = E(X)$ by the mean of the observations and an asymptotic confidence 100p% confidence interval for μ may be obtained as

$$(3.1) \quad \bar{X} \pm (\hat{L})^{-1/\hat{\alpha}} n^{-1/\hat{\alpha}-1} c_{\hat{\alpha}}^{1/\alpha} S_{\hat{\alpha},1,1}(1-p),$$

with $c_\alpha = \Gamma(2 - \alpha)|\cos(\pi\alpha/2)/(\alpha - 1)$ (c.f. [75], Theorem 1.8.1). Here $S_{\alpha,1,1}(1 - p)$ is the $1 - p$ -th quantile of the completely (positively) asymmetric stable distribution with tail parameter α , in the parametrization of the cited reference. Estimates of L, α are discussed in the next section.

3.3. Tail estimation. There exists a substantial literature on this topic, see e.g. [24], and development still is rapid. A good source for developments to come is the new journal *Extremes*.

The paradigm is that often, as in the present situation, one cannot hope to infer the behavior of the tail of a distribution from the shape of its central parts. The standard procedure then is to choose an appropriate high level u , and only use those observations which exceed the level for estimation of tail parameters. Since interest usually includes extrapolation outside the range of values already observed, a parametric model for the tail (corresponding to a semiparametric model for the entire distribution) is unavoidable. It is also unavoidable that estimates will be sensitive to the choice of level u .

On theoretical grounds (asymptotic theory, stability under change of level, and “the right amount of flexibility”) the Poisson-generalized Pareto model is preferred. In this model, the exceedances are i.i.d. and follow a generalised Pareto (GP) distribution of the form

$$G_\gamma(x) = 1 - \left(1 + \frac{\gamma x}{\sigma}\right)^{-1/\gamma}, \quad 1 + \frac{\gamma x}{\sigma} > 0, \sigma > 0, \gamma \in \mathbb{R}.$$

This is a heavy tailed distribution for $\gamma > 0$ with $\alpha = 1/\gamma$; for $\gamma = 0$ this is the exponential distribution and for $\gamma < 0$ the distribution has finite upper endpoint. See for example [24]. The times of exceedances occur according to a Poisson process which is independent of the sizes of the exceedances. This model is implied by the stronger assumption that the observations $\{X_n\}$ are iid with a regularly varying tail.

The intensity of the Poisson process is simply estimated by the number of exceedances divided by the total number of observations and the parameters of the GP distribution may be estimated by the method of maximum likelihood. S+ programs for this estimation have been made available by A. McNeill, <http://www.math.ethz.ch/mcneil/>. Another relevant software package, *Xtremes*, is a PC package for analyzing and graphing extreme values and comes on a disk with the book [64]. Straightforward asymptotic confidence intervals based on the information matrix do not perform well for a small or moderate number of exceedances, but instead profile likelihood intervals do, see [76].

In addition to maximum likelihood analysis, we have used semiparametric graphical methods for variables with a Pareto-like tail, such as qq-plots of quantiles of a standard exponential distribution against the logarithms of the ordered exceedances and estimating α by the slope of an ordinary regression line. This method of QQ-plotting does not use the detailed form of the GP distribution and is asymptotically less efficient than other similar estimators, such as the Hill estimator. However, it can be used with any standard statistical software, and has the advantage of being directly linked to a readily interpreted graphical display. It also gives a useful impression of the size of deviations from a Pareto model, since a Pareto tail corresponds to a straight line. The method is discussed in detail in [53].

Finally a practical consideration, similar to the choice of bandwidth in density estimation, and a subject of much current research, is the choice of the level u , or, equivalently, the number of order statistics used for the tail estimation. In this paper we have used informal, graphical methods as discussed e.g. in [69], which look at changes in the estimators as the level u is changed.

3.4. Starts of file transfers. That the file requests appear as a Poisson process is equivalent to the interarrival times being i.i.d. and exponentially distributed. We have investigated the Poisson

process assumptions by making qq-plots of the interarrival times against exponential quantiles, and by looking at the correlation functions of the interarrival times.

3.5. Stable Lévy processes. At low aggregation levels, and for $1 < \alpha < 2$ traffic rates in the infinite source Poisson model are asymptotically stable and independent. For maximum likelihood estimation of the parameters and model checking we have followed the prescriptions in [62] and used the very convenient accompanying software which can be downloaded from <http://www.cas.american.edu/jpnolan/stable.html>.

We have used the most common parametrization of stable distributions, which corresponds to the characteristic function

$$E \exp(itX) = \begin{cases} \exp\left(-\sigma^\alpha |t|^\alpha \left(1 - i\beta \tan\left(\frac{\pi\alpha}{2}\right) \text{sign}(t)\right) + i\delta t\right) & \alpha \neq 1 \\ \exp\left(-\sigma |t| \left(1 + i\beta \frac{2}{\pi} \text{sign}(t) \ln |t|\right) + i\delta t\right) & \alpha = 1 \end{cases}.$$

3.6. The Hölder index. For a Gaussian process $\{X(t), t \geq 0\}$, a short survey of estimators for the local Hölder index H_o is given in Lang ([56]). Hall and Wood ([34]) study the so-called box counting estimation method based directly on the capacity of fractal dimension ([40]) and show that the estimator has a large asymptotic bias. Hall *et al.* ([35]) and Feuerverger *et al.* studied estimators based on level crossings. Constantine and Hall ([14]), cf. also Gladyshev ([32]) discussed estimators based on simple empirical quadratic variation. At scale $1/n$ and for a process with stationary increments on $[0, 1]$, this is

$$V_n = \frac{1}{2} \sum_{k=1}^n (X(k/n) - X((k-1)/n))^2,$$

which suitably normalized converges with probability 1,

$$\lim_{n \rightarrow \infty} n^{2H_o-1} V_n = c,$$

where c is given by (2.2), i.e. by $V(\tau) = c\tau^{2H_o} + o(\tau^{2H_o})$.

However, these quadratic variation estimators are not scale invariant and all of the estimators mentioned so far converge slowly. Istas and Lang ([43]) and Kent and Wood ([49]) propose an improved method, where the “simple” difference $X(k/n) - X((k-1)/n)$ is replaced by general discrete differencing of X . More precisely, they introduce the weights $\mathbf{a} = (a_0, a_1, \dots, a_p)$ of a discrete difference operator of degree $p \geq 1$, and consider the empirical quadratic variations associated with it,

$$U_{\mathbf{a},n} = \sum_{j=1}^{n-p+1} \left(\sum_{i=0}^p a_i X((i+j-1)/n) \right)^2.$$

In this paper we use an estimator of H_o which is based on two such “general” quadratic variations. Specifically, from the sequence $\mathbf{a} = (1, -5, 10, -10, 5, -1)$ we define a sequence \mathbf{b} with “double time mesh”, i.e. by putting $b_{2i} = a_i$ and $b_{2i+1} = 0$ for $0 \leq i \leq 5$ and estimate H_o by

$$\hat{H}_o = \frac{1}{2} \ln_2 \left(\frac{U_{\mathbf{b},n}}{U_{\mathbf{a},n}} \right).$$

This estimator is discussed in Istas and Lang ([43]) p. 432). This sequence \mathbf{a} corresponds to the binomial coefficients in the expansion of $(u-v)^5$, $(u, v) \in \mathbb{R}^2$. An advantage of the estimator is that it filters out polynomial trends up to order 4, and hence is rather robust to smooth non-stationarities of the mean.

3.7. The Hurst exponent. In this section we describe a recently developed wavelet estimation methodology. The method works under a variety of assumptions. If the process under investigation is stationary with finite variance, the method yields an estimate of the parameter β introduced in (2.3), with $\beta > 0$ corresponding to LRD. When applied to the increments of a selfsimilar process, the wavelet method yields an estimate of $2H - 1$, as shown below. Now, for a selfsimilar process with finite second moments and stationary increments the increments process has $\beta = 2H - 1$. Hence also for this case LRD corresponds to $\beta > 0$.

However, if the selfsimilar process has infinite second moments, $2H - 1$ can be greater than zero even if there is no LRD. For example, stable Lévy processes have Hurst parameter $H = 1/\alpha$, so that $2H - 1 > 0$ (since $0 < \alpha < 2$) even though the process has independent increments and hence is as far from LRD as possible. To summarize, for processes with finite variance, if the wavelet method gives an estimate clearly greater than zero, then there is LRD. However, for processes with heavy tails an estimate of which is greater than zero does not necessarily mean that there is LRD.

A naive estimator for β may be obtained from the empirical covariance function. For strongly correlated data this is known to be unreliable. Several alternative estimation methods have been developed: the aggregated variance, Whittle, R/S, absolute value, periodogram, variance of residuals, aggregated Whittle, local Whittle and wavelet methods. The first three methods are treated in [11]. For discussion and comparisons of all of the methods we refer to [6],[3] and references therein. Here we use the wavelet method because it provides an appealing compromise between low computational cost and good statistical performance. It is also more flexible than maximum-likelihood-based estimators such as Whittle's estimators since it does not require an exact parametric model for the spectral density. In addition, it is based on identification of scaling in a log-log diagram, which makes it possible to judge the range of scales on which the model fits. A final advantage is that it is robust to smooth non-stationarities. In many cases we have also computed, as a sanity check, estimates using traditional techniques such as the R/S statistic and associated plot and the variance-time plot, and in all cases agreement with the wavelet method was good.

We here outline the main ideas and refer to [4],[1], [80],[3] and [81] for an exhaustive presentation. A *wavelet* ψ is a smooth function which is well-localised both in position and frequency and which satisfies the admissibility condition $0 = \int \psi = \mathcal{F}(\psi)(0)$, where $\mathcal{F}(\psi)$ is the Fourier transform of ψ . Usually, it is also required that the wavelet has some vanishing moments, $\int t^n \psi(t) = 0$, $n = 0, 1 \dots N$. By “well-localised” is meant that the function has compact support or at least is rapidly decaying. To a reference wavelet ψ , usually called the “mother wavelet”, is associated a two-parameter family of functions ψ_{ba} obtained by translation and dilation,

$$\psi_{ba}(t) := \frac{1}{\sqrt{a}} \psi \left(\frac{t-b}{a} \right), b \in \mathbb{R}, a > 0.$$

If the mother wavelet ψ is localized around some central position t_0 , then it is clear from the definition that ψ_{ba} is shifted to position $at_0 + b$. Similarly, if the Fourier transform $\mathcal{F}(\psi)$ is centred around some central frequency ν_0 , then $\mathcal{F}(\psi_{ba})$ is centered around $a^{-1}\nu_0$. The admissibility condition $\hat{\psi}(0) = 0$ ensures that the function stays localised in frequency after dilation. The wavelet transform $W_\psi X$ of a process X is a function on the position-frequency plane $\mathbb{R} \times \mathbb{R}^+$,

$$(3.2) \quad W_\psi X(b, a) := \int \frac{1}{\sqrt{a}} \psi \left(\frac{t-b}{a} \right) X(t) dt = \int \sqrt{a} \psi(s) X(as + b) ds.$$

Roughly speaking, the filter $W_\psi X(b, a)$ retains the part of the process which contributes to the frequency a at the position b . No information is lost, in the sense that the original signal can be retrieved from (3.2) by a reconstruction formula.

For random processes, the wavelet transform captures the stationarity and scaling properties. If the process $X(\cdot)$ is stationary, or has stationary increments then the process $W_\psi X(\cdot, a)$ is again stationary as is clear from the second formula in (3.2). If a process Y is H -self-similar with stationary increments and finite second moments, then it follows readily from the definition that

$$E|W_\psi Y(b, a)|^2 = a^{2H+1} E|W_\psi Y(b/a, 1)|^2$$

If we define $X(t) = Y(t + \Delta) - Y(t)$ for $\Delta > 0$, then by change of variables, and using a first order Taylor expansion and self-similarity in the third step,

$$\begin{aligned} (3.3) \quad W_\psi X(b, a) &= \int \frac{1}{\sqrt{a}} \psi\left(\frac{t-b}{a}\right) \{Y(t+\Delta) - Y(t)\} dt \\ &= \int \sqrt{a} \left\{ \psi\left(s - \frac{b+\Delta}{a}\right) - \psi\left(s - \frac{b}{a}\right) \right\} Y(as) ds \\ &\sim \int \sqrt{a} \left\{ -\frac{\Delta}{a} \psi'(s) \right\} a^H Y(s) ds \\ &= -a^{H-1/2} \Delta \int \psi'(s) Y(s) ds, \quad a \rightarrow \infty, \end{aligned}$$

under suitable regularity conditions. Thus, if second moments exist

$$(3.4) \quad C(a) = E|W_\psi X(b, a)|^2 \sim K a^{2H-1}, \quad a \rightarrow \infty,$$

where $K = \Delta^2 E|\int \psi'(s) Y(s) ds|^2$. Thus, in particular this holds for fGn.

Now assume $X(\cdot)$ is any stationary process with the LRD property (2.3). The large scale behaviour of the wavelet coefficients is

$$(3.5) \quad C(a) = E|W_\psi X(b, a)|^2 \sim K_\beta a^\beta, \quad a \rightarrow +\infty,$$

where $K_\beta = K \int d\nu |\nu|^{-\beta} |\mathcal{F}(\psi)(\nu)|^2$, as shown by a straightforward computation. This of course agrees with (3.4) for fGn, since $H = (\beta + 1)/2$.

Even if the original process $X(\cdot)$ is long range dependent, the corresponding wavelet transform is short-range dependent as a function of b . The idea of the wavelet estimator of the LRD parameter for LRD stationary processes is to take advantage of this decorrelation and to compute an estimator $\hat{\beta}$ using (3.5). In practice, this is done using so-called multiresolution analysis, which provides a fast algorithm to compute the wavelet coefficients on a dyadic grid in the position-time plane, i.e. the coefficients $d(j, k) = W_\psi X(2^j k, 2^j)$, $j, k \in \mathbb{Z}$. The information contained in these coefficients is sufficient to reconstruct the process (see [21]). By (3.5)

$$(3.6) \quad C(2^j) = E|d(j, k)|^2 \sim K_\beta 2^{j\beta}, \quad j \rightarrow +\infty,$$

and the decrease of correlation between the wavelet coefficients $d(j, k)$ is controlled by the number of vanishing moments, N , in the following way ([79],[27]),

$$(3.7) \quad |E d(j, k) d(j, k')| = O(|2^{-j}(k - k')|^{+\beta-1-2N}), \quad |2^{-j}(k - k')| \rightarrow \infty.$$

Now, $C(2^j)$ may be estimated by the sum of the coefficients at fixed scale,

$$\hat{C}(2^j) = \frac{1}{n_j} \sum_k |d_{j,k}|^2,$$

where n_j is the number of available coefficients at scale 2^j . The parameter β in (3.5) or equivalently $H = (\beta + 1)/2$ is then simply estimated from a linear regression in the log-log diagram of $\widehat{C}(2^j)$.

However, taking the logarithm introduces a bias ($E \log \neq \log E$). Under the simplifying assumptions that the process is Gaussian and the wavelet coefficients are perfectly decorrelated (instead of (3.7)), $\hat{C}(2^j)$ is a sum of chi-squared independent variables, and the bias can be explicitly computed and removed.

Since the variance of the wavelet coefficients increases with the scale, the quality of the estimator is improved by performing a weighted linear regression, which gives more weight to small scales. Altogether, the estimator of β is defined as, see [81],

$$\hat{\beta} = \frac{\sum y_j(jS - S_j)/\sigma_j^2}{SS_{jj} - S_j^2},$$

where

$$S = \sum 1/\sigma_j^2, \quad S_j = \sum j/\sigma_j^2, \quad S_{jj} = \sum j^2/\sigma_j^2,$$

with

$$\begin{aligned} g(j) &= \frac{\Gamma'(n_j/2)}{\Gamma(n_j/2) \log 2} - \log_2(n_j/2) \simeq -\frac{1}{n_j \log 2}, \\ y_j &= \log(\hat{C}(2^j)) - g(j), \\ \sigma_j^2 &= \text{Var}(y_i) = \frac{\zeta(2, n_j/2)}{\log^2 2} \simeq \frac{2}{n_j \log^2 2}, \end{aligned}$$

Γ being the Gamma function and ζ a generalized Riemann Zeta function. Here, the sums run over some selected range of scales $[j_{\min}, j_{\max}]$, which is chosen a priori. The smallest scale j_{\min} should be large enough for the asymptotic regime to be reached, while j_{\max} is limited by the lack of coefficients at the coarsest scale. Using the relation $H = (\beta + 1)/2$ the estimator of β at once gives an estimator for H .

In [81] it is shown that, if decorrelation actually were perfect, then the asymptotic variance of the estimator for β would achieve the Cramer-Rao lower bound. Numerical simulations exhibited in [81] show agreement with this approximate result. Moreover, the estimator is empirically shown to be robust with respect to some deviations from the Gaussian assumption. An approximate confidence interval for H is (see [4]):

$$\hat{H}_u - \sigma_{\hat{H}_u} z_\epsilon \leq \beta \leq \hat{H}_u + \sigma_{\hat{H}_u} z_\epsilon,$$

where z_ϵ is the $1 - \epsilon$ quantile of the normal distribution and

$$\sigma_{\hat{H}_u}^2 = \frac{2}{n_{j_{\min}} \log^2 2} \frac{1 - 2^{-J}}{1 - 2^{-J+1}(J^2 + 4) + 2^{-2J}},$$

$J = j_{\max} - j_{\min}$ is the number of scales which is used in the regression. This of course also gives a confidence interval for H .

We also performed a small simulation study. On 500 simulated paths of length 4096 of a fGn with $H = 0.8$ we computed the estimator of H using $j_1 = 1, j_2 = 10$ and the Daubechies1 wavelet. The empirical means and standard deviations of the estimator were 0.797, and 0.038, where the latter should be compared with the value 0.012 obtained for the theoretical standard deviation.

An appealing feature is that the wavelet transform performs a “smooth differentiation” of the signal, with the degree of differentiation equal to the number of vanishing moments. Thus, as already mentioned, means and smooth trends are removed, and non-stationary processes which have stationary increments of order N produce stationary wavelet coefficients ([12], see also [33] for the fractional increments). If such nonstationary processes exhibit a scaling of type (3.5), then the corresponding parameter can be estimated by the same procedure as before.

Suppose now X is self-similar with Hurst parameter H , but not necessarily long range dependent. Using (3.3) above, it may be seen that then, also in cases with infinite variance, the wavelet method gives an estimator of H , when taking logarithms makes moments finite. In particular, for a stable Lévy motion, the method estimates $H = 1/\alpha$. This is discussed in detail in [2]. Thus, as already pointed out, an estimated value of β which is clearly different from 0.5 does not necessarily indicate LRD — it could also be caused by heavy tails. This comment also applies to, for example, the R/S statistic. See [78].

3.8. Inferring the parameters of the infinite source Poisson model from aggregated data. Assume that a trace of cumulative traffic has been observed during a time interval $[0, T]$ in a situation which is well described by the infinite node Poisson model with $1 < \alpha < 2$ and at a high aggregation level (i.e. with $\lambda T \bar{F}(T)$ large). Further, recall the notation from Section 2.5 and suppose the regularly varying factor in (1.1) is (at least approximately, and over the range of interest) equal to a constant, L .

By Proposition 1, the observed cumulative traffic then is approximately distributed as an expectation term $\mu(T) = \lambda\mu T$ added to a centered fBm with $H = (3 - \alpha)/2$ and variance $\sigma^2(T) = \lambda T^{3-\alpha} L(4-\alpha)(2-\alpha)^{-1}(3-\alpha)^{-1}$. From the observations it is possible to find estimates $\hat{\mu}(T), \hat{H}, \hat{\sigma}^2(T)$ of $\mu(T), H, \sigma^2(T)$ as described above. This immediately leads to the estimate

$$\hat{\alpha} = 3 - 2\hat{H}.$$

Knowledge of the traffic rates alone does not contain enough information to make it possible to untangle λ, μ and L . We hence assume that we have more detailed information or experience from other data sets which allows a reasonable guess $\tilde{\mu}$ of the value of μ . The remaining parameters may then be estimated as

$$\hat{\lambda} = \hat{\mu}(T)/(\tilde{\mu}T)$$

and

$$\hat{L} = \hat{\sigma}^2(T) \hat{\lambda}^{-1} T^{-(3-\hat{\alpha})} (4 - \hat{\alpha})^{-1} (2 - \hat{\alpha}) (3 - \hat{\alpha}).$$

If we instead assume that the data are at a low aggregation level, then the trace of the cumulative traffic is approximately distributed as the same mean term $\mu(T) = \lambda\mu T$ as above, plus $(\lambda t/L)^{1/\alpha} X_\alpha := \gamma(t) X_\alpha$. From the observations we obtain estimates $\hat{\mu}(T), \hat{\alpha}, \hat{\gamma}(T)$, of $\mu(T), \alpha, \gamma(T)$. Proceeding in the same way as before this leads to the estimators

$$\hat{\lambda} = \hat{\mu}(T)/(\tilde{\mu}T)$$

and

$$\hat{L} = \hat{\gamma}(T)^{-\alpha} \hat{\lambda}^{-1} T^{-1}.$$

3.9. Simulation methods. As discussed in the introduction, an important use of traffic models is to produce simulated traces, and such traces are also useful for testing estimation methods. For each of the three main models, stable noise, fBn, and the infinite source Poisson noise, simulated traces and the wavelet regression estimator for estimating the Hurst parameter are shown, in Figures 3.1–3.3.

Since stable noise simply consists of i.i.d. stable variables, simulation is straightforward. We used the built in simulator in the program package Splus. The parameter values were chosen as the estimated values for the UCB 10s traffic rate trace, cf. Section 4.3.2 below. The maximum likelihood estimates which were obtained from the simulated trace were $\hat{\alpha} = 1.49 \pm .09, \hat{\beta} = 1.00 \pm .00, \hat{\sigma} = 90,000 \pm 600, \hat{\delta} = 446,000 \pm 1,100$, in reasonable agreement with the true parameters (given in the caption to Figure 3.1). It may be noted that the wavelet estimate of the Hurst parameter (Figure

3.1) is close to $1/\alpha = 0.66$ and hence is well away from $1/2$ although the variables are independent, rather than long-range dependent, c.f. the discussion at the end of Section 2.1.

The best available method to simulate fBn is imbedding in a circulant process. The method is described in [86], and we used software which is made available by G. Chan at <http://www.maths.unsw.EDU.AU/grace/>.

The infinite source Poisson model was simply simulated by building up traces from i.i.d. exponential starts of file transfers, and i.i.d. transfer times. For the present purpose of model evaluation and illustration, we only used the simplest possible transfer time distributions, i.e. Pareto distributed variables with parameters roughly adjusted to match the measured traces. For use of simulation to aid in design, a more sophisticated choice would be desirable, at the very least means should be adjusted to have correct values (c.f. Section 3.8). For the purposes of Section 4, we call the resulting trace $\text{simM}/G/\infty$.

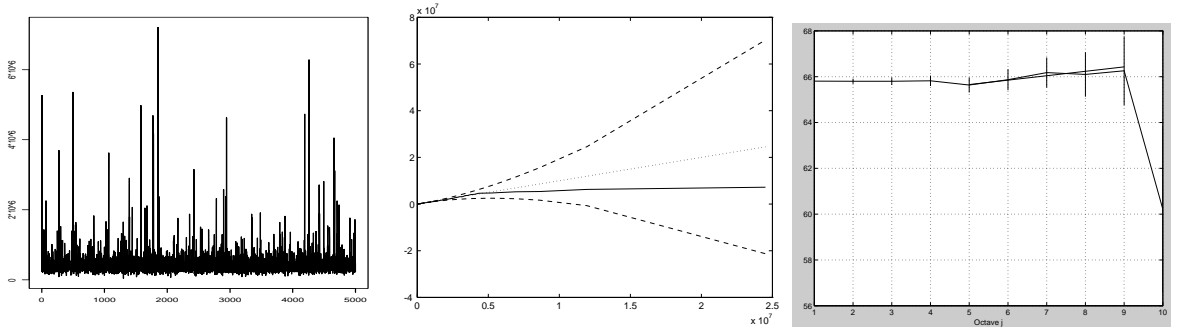


FIGURE 3.1. Synthetic trace from stable noise ($\alpha = 1.52, \beta = 1, \sigma = 86,000, \delta = 438,000$) used as the traffic rate: *left*) the traffic rate, *middle*) qq-plot against fitted stable distribution, and *right*) wavelet regression estimation of the Hurst parameter, $\hat{H} = .60 \pm .06$.

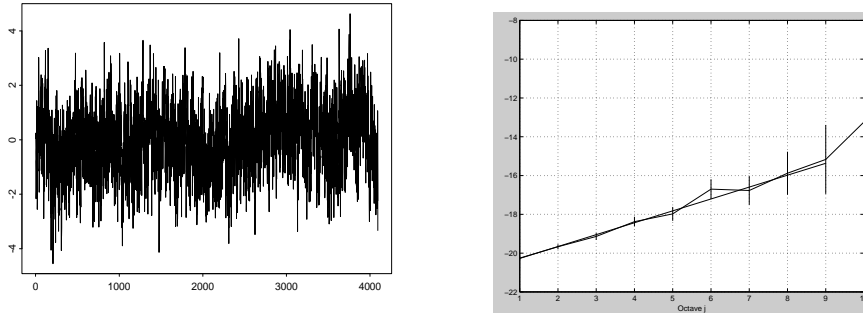


FIGURE 3.2. Synthetic trace from fBn, $H = .8$, used as the traffic rate: *left*) the traffic rate, and *right*) wavelet regression estimation of Hurst parameter, $\hat{H} = .81 \pm .01$.

4. DATA ANALYSIS

4.1. Overview. This section contains the statistical analysis of the traces which were briefly introduced in Section 1. The plan is as follows: Tables 1 and 2 summarize the results. The traces

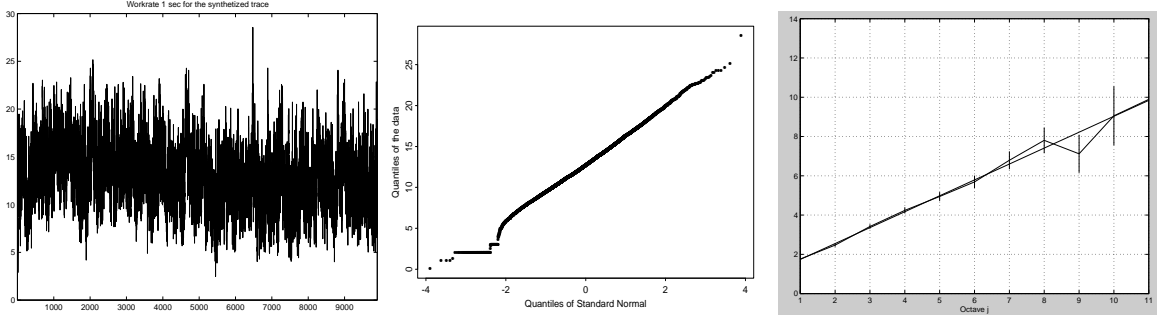


FIGURE 3.3. $\text{SimM/G}/\infty$, Synthetic trace from the infinite source Poisson model, Poisson intensity 8, $\alpha = 1.2$ used as the traffic rate: *left*) traffic rate (resolution 1 second), *middle*) qqplot against normal distribution, and *right*) wavelet regression estimation of Hurst parameter,

Data set	$\gamma (=1/\alpha)$	\hat{H}	H^*	\hat{H}_o	$\lambda T \bar{F}(T)$	Gauss	stable	dep
simM/G/ ∞	$-.13 \pm .03$	$.90 \pm .01$.90	.88	8	good	bad	str/long
BUburst 10s	$-.36 \pm .13$	$.89 \pm .02$	$.67 \pm .01$.73	.09	med	good	str/long
BUburst 1s	$.17 \pm .03$	$.81 \pm .01$	$.67 \pm .01$.87	.07	med	bad	nonstat
UCB 10s	$.05 \pm .18$	$.58 \pm .03$	$.62 \pm .01$.65	16	bad	good	indep
UCB syn 10s	$-.60 \pm .14$	$.95 \pm .07$	$.62 \pm .01$	1.36	16	good	bad	str/long
Munich lo TX	$.09 \pm .04$	$.89 \pm .01$	-	.86	-	bad	bad	str/long
Munich lo RX	$-.03 \pm .05$	$.97 \pm .01$	-	.85	-	bad	bad	nonstat
Munich hi .1s	$.17 \pm .12$	$1.02 \pm .03$	-	.66	-	bad	bad	med/int
Munich hi .01s	$.10 \pm .03$	$1.03 \pm .04$	-	.56	-	bad	bad	sm/long
Ericsson	$-.47 \pm .09$	$.88 \pm .02$	$.86 \pm .08$	1.21	.5	bad	bad	nonstat
Eri syn 1s	$-.31 \pm .12$	$1.48 \pm .02$	$.86 \pm .08$	1.51	.5	medium	medium	str/long

TABLE 1. Traffic rates: summary of statistical analysis. Whenever possible, numerical results are given as point estimate \pm “standard deviation”. Abbreviations are explained in Section 4.1.

are then discussed one by one in separate subsections. This includes a more detailed description of the data, some further discussion of special issues for the separate traces, and conclusions trace by trace.

Table 1 contains results for the traffic rate measurements. The left most column contains the shape parameter γ (related to the tail index α by $\alpha = 1/\gamma$) estimated by maximum likelihood using a generalized Pareto model (Subsection 3.3). Standard deviations are calculated assuming independence. In all cases the top 5% of the observations were used, and the fit to a Generalized Pareto distribution, as judged by QQ-plots was good (except Eri syn 1s). It might be worth recalling that a value of γ close to zero means light tails, that the variance is finite for $\gamma < 1/2$, and that the important case with finite mean and infinite variance ($1 < \alpha < 2$) corresponds to $\gamma \in [1/2, 1)$.

After this comes the Hurst exponent H with standard deviation, estimated by the wavelet method with the Daubechies wavelet and the number of vanishing moments and j_{\min}, j_{\max} chosen to give good fit to the regression on scales (Subsection 3.6). We also calculated the Hurst exponent from the tails of the file sizes as $H^* = (3 - \alpha)/2$ with $\alpha = 1/\gamma$ taken from the file size column in Table

Data set	intens	exp	dep	time	dep	file size	dep	rate	dep
BUburst	.14	no	nonstat	$.60 \pm .02$	tiny/sh	$.69 \pm .13$	tiny/sh	$1.01 \pm .14$	tiny/sh
UCB	6.72	yes	tiny/long	$.57 \pm .02$	tiny/int	$.52 \pm .02$	tiny/sh	$.79 \pm .04$	tiny/sh
Ericsson	1.39	no	sm/sh	$.78 \pm .16$	tiny/sh	$1.15 \pm .18$	sm/sh	-	sm/sh

TABLE 2. Starts of transfers and file sizes: summary of statistical analysis. The columns “time”, “file size” and “rate”, show the shape parameter γ ($=1/\alpha$) for the transfer times, file sizes, and transfer rates. Numerical results are given as point estimate \pm “standard deviation”. Abbreviations explained in Section 4.1.

2 and with standard deviations obtained by the delta method. The following column shows the Hölder parameter H_0 , estimated by the quadratic variation method (Subsection 3.7). The next column shows an estimate of $\lambda T \bar{F}(T)$ obtained by using λ from Table 2, with T as 1 second for “BUburst 1s” and T as 10 seconds for “BUburst 10s” and “UCB 10s”, etc, and $\bar{F}(T)$ estimated by $\#(\text{observations} > T)/\#\text{observations}$ (except for simM/G/ ∞ where parameters were known).

The next two columns contain subjective judgements of the fit of the marginal distributions to a Gaussian and a stable distribution. The family of stable distributions is quite rich, and even in the cases which are labelled “bad” in the “stable” column of the table, the visual discrepancies in the QQ-plot between histograms and fitted densities were small. In all cases the tail estimates of α in column 2 were larger than those in the mle fitted stable distribution. It may be noted that the Pareto behaviour of stable distributions in some cases is only apparent far out in the tail. We have not investigated this further.

The last column classifies the dependence in the traffic rate measurements. The entries are based on the estimated correlation function of the log traffic rates, using the standard 95% asymptotic confidence limits. Two scales are used. The first one is “str” if of the first 200 correlations at least 50 are $> .1$, “med” if between 20 and 50 are $> .1$, “sm” if less than 20 are $> .1$, and “tiny” means that all correlations are $< .1$. For the second scale “long” means that the area where most correlations are significantly different from 0 extend more than 200 lags, “int” that this area lasts for between 50 and 200 lags, and “sh” that it is shorter than 50 lags. It should be noted that for different data sets, “lags” may correspond to rather different amounts of real time. The entry “indep” means the traffic rates were judged to be compatible with independence.

Finally, an entry “-” means that the estimate was not available.

Table 2 only contains the data sets which contained information about individual file transfers. In the first column, the intensity λ of starts of transfers is estimated as the total number of transfers divided by total time in seconds, and in the next the fit of the interarrival times to an exponential distribution as checked by qq-plots is given. In the remaining columns, the shape parameter γ is estimated and dependence is checked in the same way as for Table 4.1. The column marked “time” refers to the download time of a file and “filesize” refers to the size of the downloaded file. In the column marked rate, we computed the transmission rate as

$$(\text{file size})/(\text{time to download the file})$$

and then estimated the shape parameter γ . The parameter γ for the rate in the Ericsson trace could not be reasonably estimated; see the data description below.

4.2. The BU traces.

4.2.1. *Data description.* This data contains four components and are described and analyzed in [17] and [20]. Below we describe the BU and the derived BUburst traces, which were obtained from the posted data. Both contain times of requests, file sizes, transfer rates and transfer times, before and after “burstification” for a period of approximately 28,000 seconds (8 hours and 20 min) with the most intense traffic in the February 1995 part of the trace. The construction of the traffic rate data sets *BUburst 1s* and *BUburst 10s* is explained below. The calculation of traffic rates assumes that the transmission rate did not change during the transmission of a file, and hence the constructed traffic rates deviate from the real ones.

The traffic was generated from two rooms of users during the period October 1994 to February 1995. The statistical characteristics of data vary considerably from one month to the next. We have only used data from the room containing 32 work-stations used by undergraduates, with all the cache-file requests removed. This trace is the most complete among the public domain data sets we are aware of. It was recorded at a time when Mosaic was the most common browser. Unlike Netscape, the source code of Mosaic is publicly available and could be altered for measurements purposes. A followup study by a BU team is reported in [9].

The data consists of the record of all the individual sessions generated by the different users. A session is a succession of URL requests (http, ftp, gopher, . . .) made by one user from logging in until logout. Every request corresponds to a line in the session file which contains the following information: machine number, starting time of the request (in micro-seconds since January 1, 1970), URL of the requested document, size of the document (in bytes) and transfer time. The cache-files, that is files already stored on the user disk, are marked with a zero transmission time and file size.

The *BUburst* data resulted from the need to distinguish between file requests made by humans and machine generated requests. Typically a human initiated request, e.g. for a web page, made triggers a cascade of file transfers, and hence very small intervals between file transfers usually are machine generated. If we think of users as being sources in the infinite source Poisson model, we must correct for this cascading. Hence we lumped together requests which arrived less than .5 seconds apart into a single “request” which we will refer to as a “burst”. The selection of the “threshold” .5 seconds was based on a close look at the data. The size of the burst is then the sum of the sizes of the files lumped together, and the duration is from the beginning of the first file transmission in the burst to the termination of the last file transmission in it. *BUburst 1s* and *10s* are obtained from the *BUburst* data by calculating traffic rates at 1s and 10s resolutions respectively. A typical session before and after creating the bursts is presented below. The 100682 initial requests resulted in 56516 bursts, out of which the period we studied contains 4161.

BU				BUburst		
Begin download	URL	File size	Transfer time	Begin download	File size	transfer time
43817.159177	"http://cs-www.bu.edu/"	2069	2.994023	43817.159177	4591	4.903682
43820.586374	"http://cs-www.bu.edu/lib/pics/bu-logo.gif"	1805	0.551812			
43821.538385	"http://cs-www.bu.edu/lib/pics/bu-label.gif"	717	0.524474			
43835.427403	"http://cs-www.bu.edu/courses/Home.html"	3382	0.510081	43835.427403	3382	0.948917
43836.36075	"http://cs-www.bu.edu/lib/pics/bu-logo.gif"	0	0.0			
43836.37632	"http://cs-www.bu.edu/lib/pics/bu-label.gif"	0	0.0			
43842.134286	"http://cs-www.bu.edu/students/grads/tahir/CS111/"	1065	0.404507	43842.134286	1065	0.404507
43846.456855	"http://cs-www.bu.edu/students/grads/tahir/CS111/hw6.ps"	32246	0.498012	43846.456855	32246	0.498012

4.2.2. *The number of logged-on work-stations.* Due to the diurnal cycle (strongly visible because of the small number of users), the traffic is highly non-stationary, see Figure 4.1 which shows the number of active sessions. The periods with 0 logged-on stations typically occur during nights and week-ends.

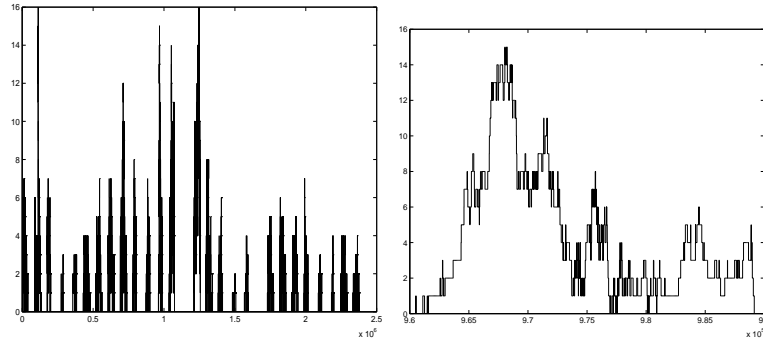


FIGURE 4.1. **BUburst**: *left*) Number of active sessions versus time (in seconds), January 1995. *right*) Detail of the left plot showing one day.

4.2.3. *Discussion.* The interarrival times for the BU and BUburst traces over the entire month were clearly not exponentially distributed. For BUburst there was some indication of exponential-like behavior for small interarrival times, but the distribution was heavy tailed as revealed by a QQ-plot (see Section 3.3) yielding the estimate $\hat{\alpha} = 1.54$, cf Figure 4.2. In view of the low aggregation level this is not surprising.

For the BUburst trace the transmission times, file sizes, and transfer rates had heavy tails with $\hat{\gamma} > .5$. The tails of the transmission rates dominated the tails of the transmission times and file sizes, in contrast to our expectations.

The BUburst 1s and BUburst 10s traffic rate data sets have rather light-tailed marginal distributions. However, the distributions still look much more stable than normal. Since the traffic rates are quite dependent (or non-stationary), they cannot be well modelled by a stable Lévy noise.

The Hurst and Hölder estimates were rather similar, while the Hurst estimate derived from the fBm model and the $\hat{\gamma}$ for the transfer times was smaller.

We selected a short period with high traffic. Realistic models should include the variation in the number of logged on workstations.

An extensive statistical analysis of these data has been carried out by the authors of the trace ([15]). In particular they present similar estimates for tails and traffic rates and explain the discrepancy with the theory through the low traffic level.

From Figure 4.2 one sees that the left tail (near 0) of the inter-arrival times looks like an exponential or Weibull tail while the right tail looks Pareto. The autocorrelations of the inter-arrival times seem significant but rather small. Figures 4.3 and 4.4 show the marginal distributions of the traffic rates and the Hurst parameter estimations. For heavy tailed data it is difficult to plot an informative histogram so we show the histogram of the log of the data.

Finally, this data set is rather old on the timescale of the internet evolution, and in particular the parameter values obtained from it may be rather different from current ones.

4.3. The UCB traces.

4.3.1. *Data description.* The UCB data is an 18 day trace of the home IP HTTP traffic processed by UC Berkeley during November 1996. It consists of all the internet connections through the server of the university established by individual dial-up users. The traces together with a detailed description of them is available at <http://ita.ee.lbl.gov/html/contrib/UCB.home-IP-HTTP.html>.

The data content is similar to the BU traces (initiation time of a file transfer, file size, transfer times of a request, IP address of client). We have not attempted to burstify the UCB data, as

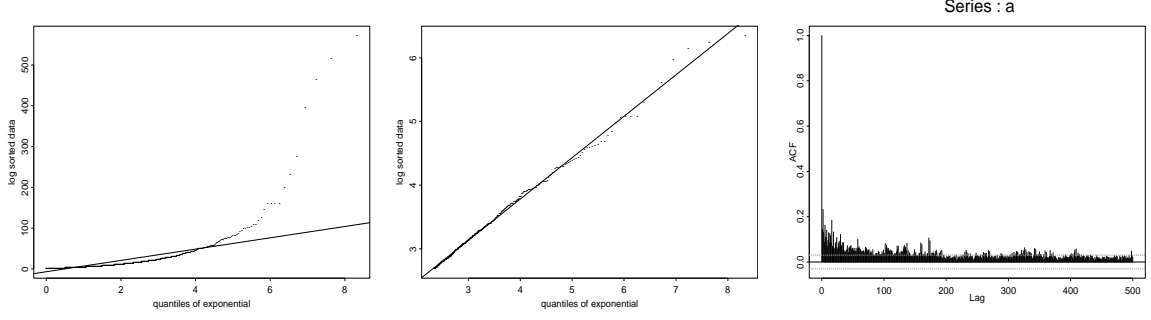


FIGURE 4.2. **BUburst**. Inter-arrival times of bursts: *left*) qq-plot against exponential distribution, *middle*) Pareto fit for the 400 largest values (last 10 %), ($\alpha = 1.54$), *right*) autocorrelation function.

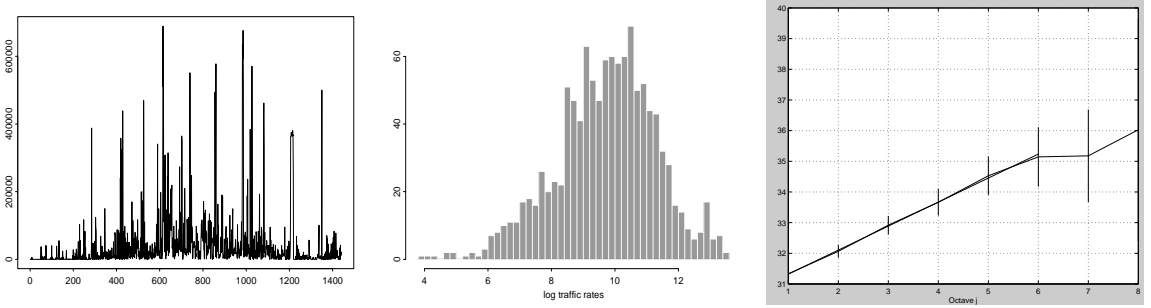


FIGURE 4.3. **BUburst 10s**. Traffic rates: *left*) trace, *middle*) histogram of log traffic rates; *right*) wavelet regression estimation of the Hurst parameter.

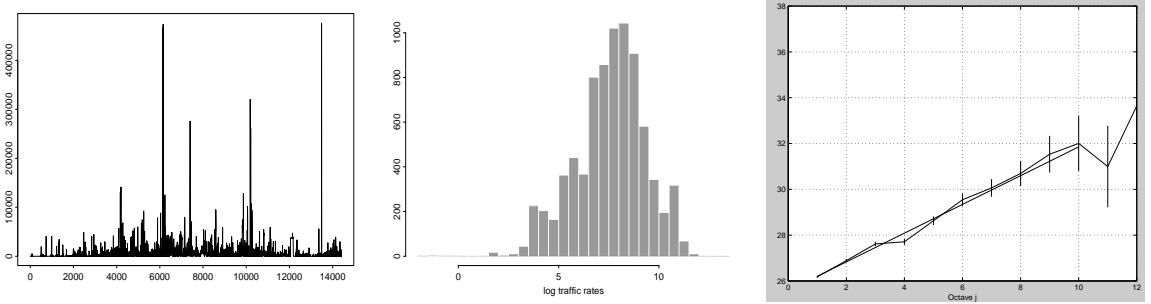


FIGURE 4.4. **BUburst 1s**. Traffic rates: *left*) trace, *middle*) histogram of log traffic rates; *right*) wavelet regression estimation of the Hurst parameter.

we did with the BU data. However, a much larger number of users are included and the traffic rate is higher. Due to the non-stationarity and the diurnal cycle, we chose to restrict the analysis to several hours of peak traffic on a weekday, i.e. the period 5-8 p.m. on Thursday November 7. This part of the trace consists of about 80,000 requests. We would like to emphasize the need to carefully select the period for analysis, since some network outages occurred during the recording of the trace. E.g. one outage can be observed on November 6, from 5 pm to 8 pm.

As for the BU data, the actual traffic rates were not available and we again constructed approximate traffic rates by assuming constant (but different from transmission to transmission) transmission times.

4.3.2. Discussion. An exponential distribution fits the inter-arrival times fairly well. While rather many of the estimated correlations for the log interarrival times are significantly different from zero, they are quite small (and perhaps caused by a small nonstationarity) and independent interarrival times seems a reasonable approximation (Figure 4.6).

The file size and transmission time distributions were close to the borderline between finite variance and infinite variance while the transmission rate was considerably more heavy tailed. The distribution of log transfer rate is clearly bi-modal (Figure 4.5, *left*), perhaps due to different modem speeds. A natural consequence is that the distribution of log transfer times is also bi-modal (Figure 4.5, *right*). There seems to be no hope of finding simple parametric forms which accommodate these distributions.

The UCB 10s traffic rates show a quite good fit to the stable Lévy noise model, actually the qq-plot and density estimates were quite similar to those for the simulated stable noise. The ML estimates of the marginal parameters were $\alpha = 1.52 \pm .02$, $\beta = 1.00 \pm .00$, $\sigma = 86000 \pm 1200$, $\delta = 438000 \pm 2300$. The α estimate may be compared to the α -value $1/.57 = 1.75$ of the tail of the transfer times, and was quite different from the α of $1/.05=20$ obtained for the traffic rate.

The stable Lévy motion model fitted somewhat less well for traffic rates computed using shorter time interval. This aspect of the data agrees with the analysis in Section 2.5, Proposition 1. However, the seemingly large value of $\lambda T \bar{F}(T)$ in Table 1, perhaps indicates the difficulty of interpreting the conditions $\lambda T \bar{F}(T) \rightarrow 0$ or ∞ in Proposition 1. The effects of varying block size on the marginal distribution of the traffic rates is illustrated in Figures 4.9. The distributional shape is clearly different for different sizes of the time intervals and hence indicates a lack of distributional self-similarity. Figure 4.10 shows the wavelet regression estimation of the Hurst parameters, which were close to the independent increments Brownian motion value 0.5.

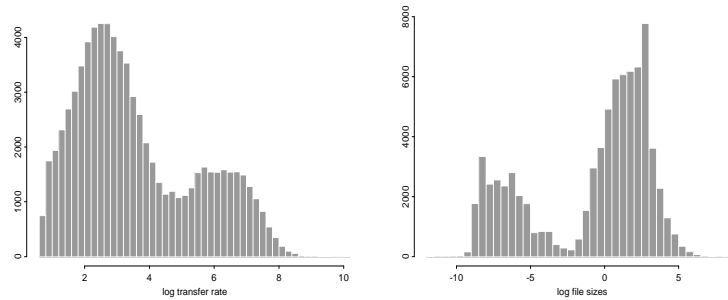


FIGURE 4.5. **UCB** Histograms of *left*) log transfer rates and *right*) log file sizes.

4.4. The Munich traces.

4.4.1. Data description. The *Munich lo* data set contains measurements of cell rates for both the sending (TX) and receiving (RX) directions of an ATM link. The link, a Customer Service Switch (GDC APEX 200) with a line speed of 155 Mbit/s, connects the Munich University network with the German Scientific Broadband Network. The data, kindly provided to us by Helmuth Gogl, was collected around the clock on Wednesday, November 12, 1997 (TX) and Wednesday, December 17,

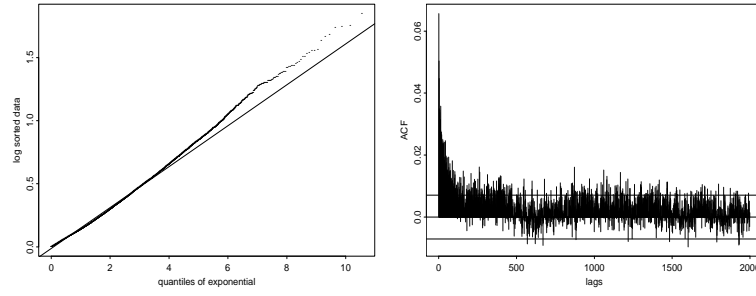


FIGURE 4.6. **UCB** inter-arrival times: *left*) qqplot against exponential distribution, *right*) autocorrelation function of interarrival times.

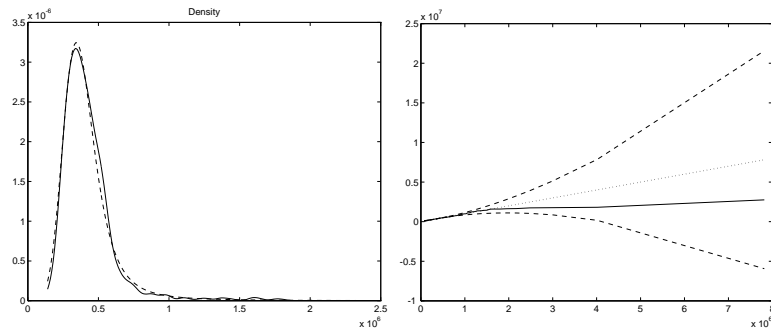


FIGURE 4.7. **UCB 10s** traffic rate: *left*) qqplot against estimated stable distribution, and *right*) density estimate and density of estimated stable distribution.

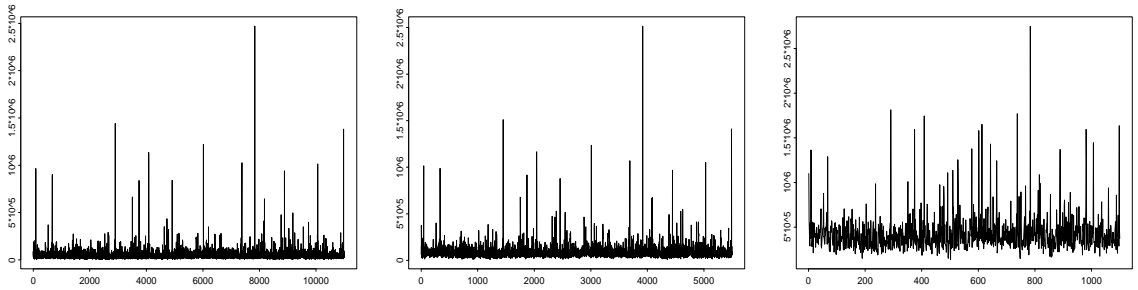


FIGURE 4.8. **UCB** traffic rates at different aggregation levels: *left*) 1 seconds, *middle*) 2 seconds, and *right*) 10 seconds.

1997 (RX) with a temporal resolution of 2 seconds, *i.e.* the total number of cells that passed the ATM link every 2 seconds was recorded. The maximum bandwidth which was available was about 20 Mbit/s.

The traffic recorded was pure IP (mainly HTTP, FTP, and NNTP) data traffic, without any audio/video components. A shorter sample covering the period between 10 a.m. until 1 p.m. was selected for analysis. The shorter period was chosen to obtain a roughly stationary data set. However, for the RX direction, the influence of the lunch hour is still clearly visible.

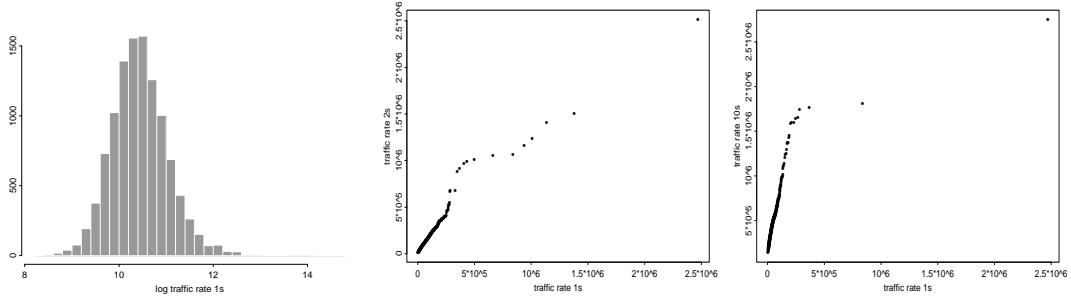


FIGURE 4.9. **UCB**: traffic rate over time blocks of varying lengths: *left*) histogram of log traffic rate for 1 second blocks ($\hat{\gamma} = .43 \pm .06$) *middle*) qqplot of log traffic rate for 1 second blocks against 2 second blocks ($\hat{\gamma} = .40 \pm .08$ for the 2 second blocks), and *right*) qqplot of log traffic rate for 1 second blocks against 10 second blocks ($\hat{\gamma} = .05 \pm .17$ for the 10 second blocks).

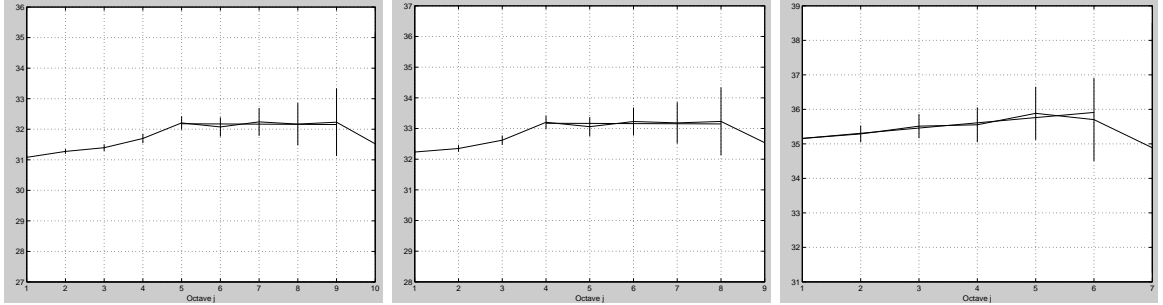


FIGURE 4.10. **UCB** wavelet regression estimation of the Hurst exponent for traffic rates over different time intervals: *left*) 1 second, $\hat{H} = 0.50 \pm .04$; *middle*) 2 seconds, $\hat{H} = 0.50 \pm .01$; *right*) 10 seconds, $\hat{H} = 0.58 \pm .03$.

The *Munich hi* data set was recorded at the same ATM link, but with a much higher time resolution. It contains 1690729 inter-arrival times of ATM cells from all traffic in the TX direction over a period of approximatively 137 seconds with a resolution of 1 micro-second. The measurement was done on Tuesday, December 23, 1997 starting at 14:48:15 and ending at 14:50:33. A mean utilisation of 5.062 MBit/s (11938 cells/s) was recorded, which corresponds to about 25% utilisation. A detailed description of the data is given in [30]

4.4.2. Discussion. All four Munich traces come from the same server, and similar traffic situations. The RX cells had typically passed through and been reassembled into frames in more routers than the TX traces. The temporal resolution varied by a factor 200, from the high aggregation level of 2 seconds down to a low one of .01 second. The different kinds of traffic, HTTP, FTP, and NNTP, had rather different characteristics, as exhibited in figures in [30].

The traffic rate traces have similar light tails with γ close to zero. The distributions are neither normal nor stable, with heavier tails than for normals but not as heavy as for a stable distribution; cf. Table 1 and Figures 4.11 – 4.14. In fact, lognormal QQ-plots showed rather good fit. The marginal distributions had different shapes for the different traces.

There is a clear and strong long range dependence, which persists over a wide range of scales as seen by autocorrelation functions and wavelet regression plots and also by the fact that for all four

traces the estimates of the Hurst exponent are close to the same value, .95. In addition, the Munich lo RX trace was clearly nonstationary. However, based on theoretical considerations and numerical experiments, this is not expected to influence the Hurst and Hölder estimates significantly.

The Hölder exponents throughout are smaller than the Hurst ones, and decrease from values of around .85 corresponding to strong local dependence for the high aggregation levels down to .56 which is close to independent increments for the .01 second trace.

In the fBm model, the Hurst and Hölder exponents are the same. Here, this is clearly not the case. One explanation could be that although the tails of the marginal distributions are lighter than Pareto, they still are rather heavy, which may inflate the Hurst exponents (recall that the i.i.d. stable variables considered in this paper have Hurst exponents $> 1/2$). However, what is not explained by the infinite source Poisson model is that the data seem more independent on a finer time resolution. A possible explanation of this discrepancy is the slow variation of λ due to a varying number of users.

An extensive analysis of these data at the cell level, and after collection into “bursts”, is made in [30, 31]. Both on the cell and the burst level, estimated Hurst exponents were smaller than our estimates – since this analysis concerned short time distances it is in agreement with the finding that the Hölder exponents were smaller than the Hurst ones. An expected result was that the lengths of very short interarrival times between cells ($< 10\mu s$) were not exponential, but had a discrete distribution determined by the system clock frequency.

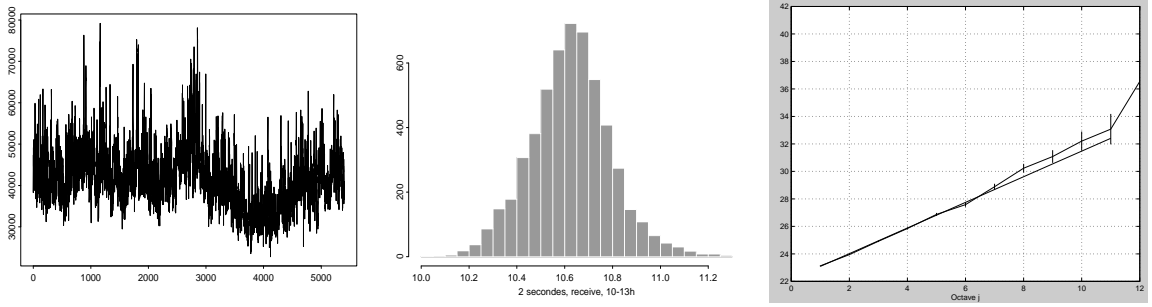


FIGURE 4.11. **Munich lo, RX:** *left*) trace, *middle*) histogram of log traffic rate, and *right*) wavelet regression estimation of the Hurst exponent.

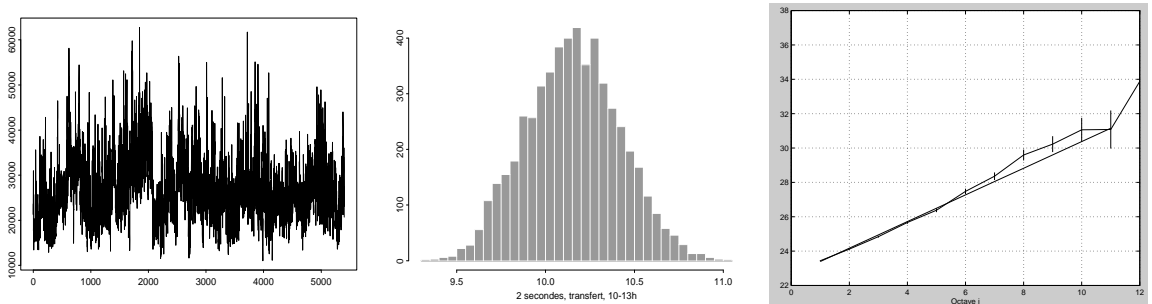


FIGURE 4.12. **Munich lo, TX:** *left*) trace, *middle*) histogram of log traffic rate, and *right*) wavelet regression estimation of the Hurst exponent.

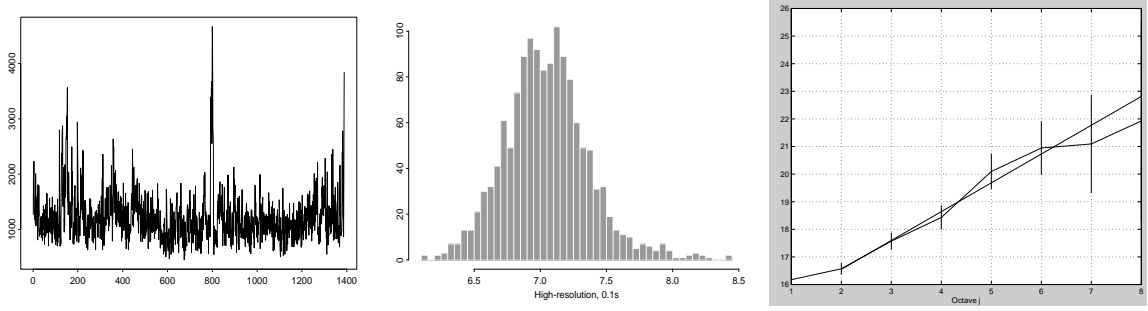


FIGURE 4.13. **Munich hi .1 s:** *left*) trace, *middle*) histogram of log traffic rate, and *right*) wavelet regression estimation of the Hurst exponent.

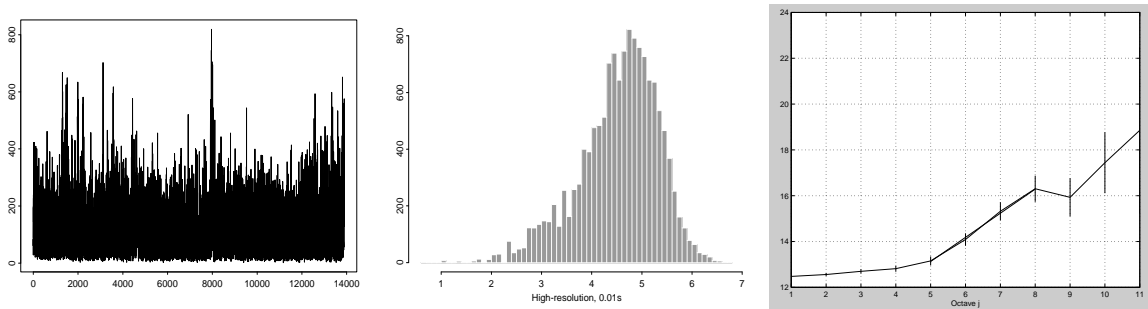


FIGURE 4.14. **Munich hi .01s:** *left*) trace, *middle*) histogram of log traffic rate, and *right*) wavelet regression estimation of the Hurst exponent.

4.5. The Ericsson trace.

4.5.1. Data description. The *Ericsson* trace consist of time stamps of starts and completions of the TCP connections that correspond to HTTP file transfers to and from a corporate WWW server which holds home pages and information primarily directed to about 2000 company users. In addition, people at other Ericsson companies around the world have access to the server. The recording was started Thursday Oct 15, 1998, at 15:20 and was ended Friday Oct 16 at 15:49. The information extracted from the data gives the times of connection starts, connection durations, number of bytes transferred (from server to user as well as the opposite direction), and client identification for each connection. Approximately 2% of the connections resulted in missing data for server-to-client transfers. The reason was that the connection was abnormally terminated and the number of bytes transferred could not be obtained from the data saved in the trace. The data set is quite nonstationary, and hence a more stationary subset covering 2000 seconds was chosen for analysis. As for the BU and UCB traces approximate traffic rates were constructed from the data described above.

For unexplained reasons the measurement system erroneously added 30 milliseconds to about 18% of the transfer times.

4.5.2. Discussion. There were both stationarity and measurement problems with this data set, but it was also the most recent one. It had the most heavytailed transfer times and file sizes. However, the transfer rates were lighttailed – this may have been particularly influenced by the measurement errors, and we have hence not included the estimated value in Table 2. The interarrival times were

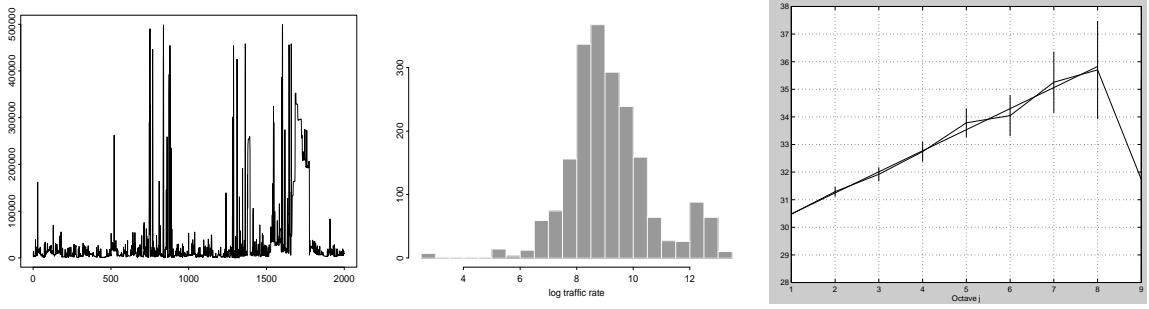


FIGURE 4.15. **Ericsson traffic rates:** *left*) trace, *middle*) histogram of log traffic rate, and *right*) wavelet regression estimation of Hurst exponent.

close to independent, but their distribution had heavier than exponential tails, as expected for this relatively low aggregation level.

The two Hurst estimates were similar, while the Hölder estimate was in the differentiable region.

The traffic rates were lighttailed and the wavelet regression showed good agreement with the fBm model, but the marginal distributions were clearly nonnormal.

4.6. The synthetic UCB and Ericsson traces.

4.6.1. Data description. The *UCB syn 10s* and *Eri syn 1s* traces were constructed from the same segments of the UCB and Ericsson data sets which were used in the analysis of work rates in Sections 4.5 and 4.3. The new traces were constructed as follows: File transfers started and ended as they did in the real trace, but the transfers were changed to make transfer rates equal to one. From the resulting “data sets” we then calculated 10 and 1 second traffic rates, respectively to obtain the “UCB syn 10s” and “Eri syn 1s” traces.

The infinite source Poisson model assumes that transfer rates are constant, and we hence thought it interesting to compare a situation where this indeed was the case with the real traces where the transfer rates had a heavy tailed distribution.

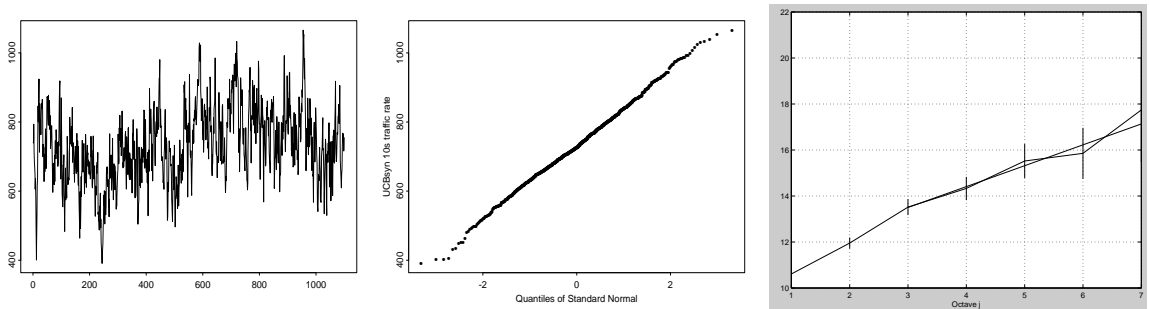


FIGURE 4.16. **UCB syn 10s traffic rates:** *left*) trace, *middle*) qqplot vs normal variable, *right*) wavelet regression estimation of the Hurst parameter.

4.6.2. Discussion. The UCB syn 10s traffic rate conformed with a fBm in except that the Hölder estimate was in the differentiable region, $H_o \geq 1$. For the Eri syn 1s the marginal distribution were close to normal except for a few observations in the tails. The Hurst and Hölder estimates were

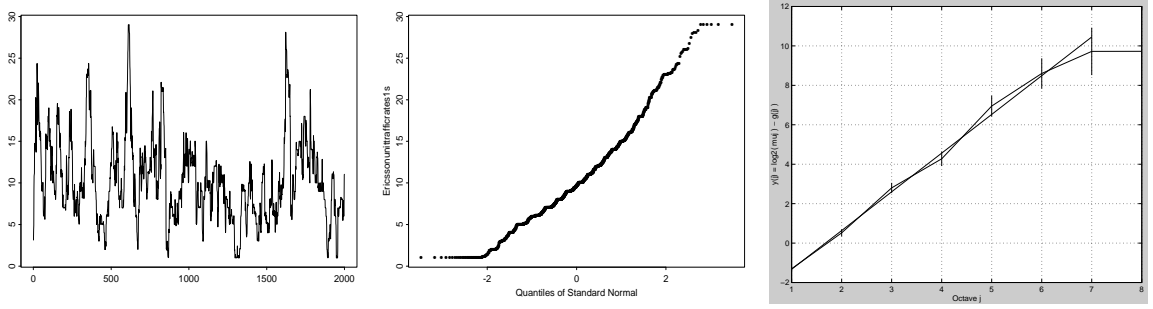


FIGURE 4.17. **Eri syn 1s** traffic rates *left*) trace, *middle*) histogram, *right*) regression estimation of the Hurst parameter.

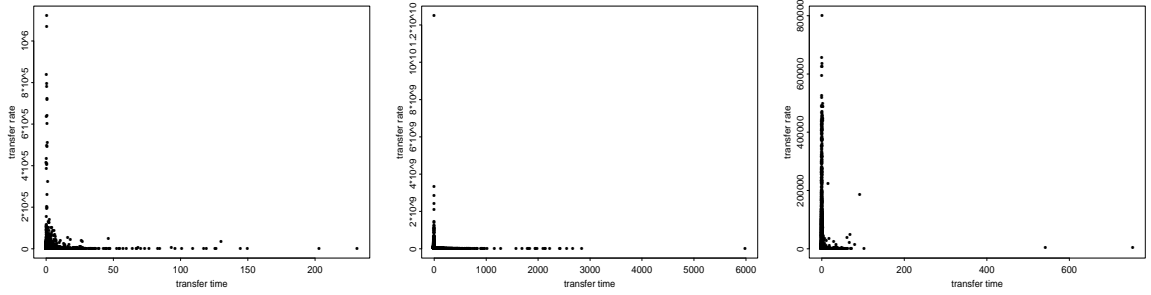


FIGURE 4.18. Plot of transfer times against transfer rates for *left*) the BUburst trace, *middle*) the UCB trace, and *right*) and the Ericsson trace.

both close to 1.5. We believe this is a consequence of non-stationarities in the data, which clearly show up in Figures 4.15 and 4.17. Both synthetic traces were very different from the real ones.

As shown in Figure 4.18, large values of transfer times and transfer rates were fairly independent; the scatter plot of transfer times vs transfer rates hugs the two axes which shows that either the two quantities are independent or if dependent, their joint distribution is multivariate regularly varying satisfying a condition of asymptotic independence [65, page 290ff].

4.7. Summary and conclusions. We have provided a summary in Section 2 of the description and properties of a fairly flexible model which we have called the *infinite source Poisson model*. This model predicts that on large time scales (i.e. as $T \rightarrow \infty$), traffic will have certain properties and we catalogued these properties according to whether the regular variation index α of the connection length distribution F satisfied (i) $\alpha < 1$, (ii) $1 < \alpha < 2$ or (iii) $\alpha > 2$. We also surveyed in Section 3 statistical methods for estimating model parameters in order that we could diagnose whether our model provided a good fit to four data sets.

Our assessment of the statistical methods selected is that they were convenient to apply and aided understanding of the data. However, we have downplayed difficulties in estimating tail parameters such as α or γ . Such estimates are sensitive to either the choice of threshold or choice of number of upper order statistics as has been amply documented; see, for example, [28, 69, 66]. Similar choices have to be made in the estimation of Hurst parameters. More development is needed to provide reliable confidence intervals for estimates, especially of long range dependence parameters and Hölder exponents. Better understanding of the estimation of Hölder exponents would be useful

as well as a clearer understanding of the relationship between treatments using the second order definition (2.2) and the pathwise treatments in, for example, [72].

Global statistical properties such as heavy tails and long range dependence were amply in evidence in our data as expected and as predicted by the model. Transfer times, file sizes and transfer rates were consistently heavy tailed, usually with $1 < \alpha < 2$. (See Table 2.) Traffic rates frequently displayed evidence of long-range dependence as shown in Table 1. However, the estimated marginal distributions of traffic rates changed markedly with aggregation level for all of the traces (see, for example, Figure 4.9) and thus we conclude that none of the traffic rates was fully distributionally self-similar. Either the model does not adequately fit the data or considered time scales are too small for the asymptotic behavior discussed in Section 2.2 to hold.

The scaling behavior of the actual traffic rate traces as summarized by the wavelet estimator of the Hurst parameter was compatible both with the fBn model and with the stable noise model (see the column for \hat{H} in Table 1). The fBn and stable noise model are the asymptotic limits given in Section 2.1 when $1 < \alpha < 2$. However, all measured marginal distributions were far from normal, and hence the fBn model does not fit the data. The stable distributions are much more flexible and generally fitted the data better. In one case, the UCB 10s traffic rates, the stable noise model provided a good description of the trace (Figure 4.7). However, generally speaking, estimates of α given in Table 1 (remember $\alpha = 1/\gamma$) using the maximum likelihood estimation in the generalized Pareto distribution model produced lighter tails than using Nolan's maximum likelihood estimators ([62]) to fit stable distributions directly to the data, so there is doubt that the stable model adequately fits the tails of the data.

An overall impression is that the infinite source Poisson model struggles to adequately describe our data. The assumption of constant transfer rates in the model is at the center of the problem. This is clearly shown by the UCB syn and Eri syn traces where the real transfer rates were changed to a fixed rate ($=1$). In particular, the UCB trace conformed fairly well with the limiting fBm model in the properties studied by us (Table 1 and Section 4.6). This was in complete contrast to the actual measured traces which had widely varying transfer rates and were quite far from being a fBm (Sections 4.3 and 4.5). Examining file transfer rates in, for example, the BU data shows that such rates should more realistically be modelled as random with a heavy tailed distribution. Furthermore, file transfer rates seem to be only asymptotically independent of file transfer times (Figure 4.18). However, implementations of TCP which include features designed to make bandwidth sharing more equitable may make these effects less pronounced.

Another difficulty in the model is identifying Poisson time points from the data. This is impossible with certain data sets such as the Munich data and difficult with the BU data. A common paradigm is that activities initiated by humans is well modelled by Poisson processes. This was in agreement with our analyses of the UCB data (Table 2 and Figure 4.6). However this simple paradigm has to be informed by the non-stationarity of most collected data and by the fact that a Poisson process model is only expected (and observed) when activities initiated by many humans are aggregated. Furthermore, many types of web based activity are initiated by machine and cannot be expected to follow the Poisson process except perhaps at extremely high aggregation levels. A possible refinement of the infinite source Poisson model is a Poisson cluster model where activity triggered by humans is modelled by Poisson cluster initiations and machine triggered actions are the cluster points associated to the Poisson cluster starts. Another alternative is the Markov modulated Poisson process or Markovian arrival process.

Here are some additional final comments.

- All the traces had very clear diurnal and weekly variations and trends. An informal technique which is widely used is to select a subset of the data for analysis which is visually stationary.

An often used heuristic rule is to not consider data over time intervals greater than 4 hours. In situations where data is often copious, this waste of data may not be serious but thought needs to be given to models which incorporate the non-stationarity explicitly. If data subset selection is used, the choice of subset should be determined by stationary situations of particular interest; for example the desire to model times of peak load.

- It is important to know which time scale is of interest. We were able to detect behavioral differences over time scales in the range 0.01s–10s. Also, at very fine resolutions of the order of microseconds, protocols and clock frequencies are very influential as was seen in the Munich high resolution data. Pareto behavior of tails of transfer times is not apparent until times exceed tens of seconds and therefore correlation behavior of the cumulative input will not match those of the limiting models for lags of smaller order. Hence, one should not expect limiting approximations to be applicable at resolutions finer than tens of seconds.
- Without added refinements, the infinite source Poisson model is clearly not capable of describing behavior on very fine time scales. Fine time scale behavior is presumably affected by the passage through many protocol layers and control mechanisms such as TCP and thought to modelling the effect of such controls is urgently needed. For some results in this direction using the concept of multifractals see [72, 73]. Investigating such a refinement of the model is a long term goal.
- Models should be quite different for high and low numbers of users or active nodes. For relatively low numbers of active nodes, say up to several hundred users, rather detailed models are needed. These should incorporate at least the varying number of active users, the activity levels of the users, the specific kinds of tasks of interest, and machine generated bursts of transfers of several files caused by one user request. In addition it is likely that models and approximations should be rather different at high and low utilization levels; that is when traffic rates constitute a high percentage of the maximum rate allowed by a link.

In situations with superposition of a large number of users (thousands or more) the assumptions of the infinite source Poisson model gave a good description of user behavior during stationary periods. In fact the asymptotic fractional Brownian motion described well the simplified version of the traffic obtained by enforcing the assumption of constant transfer rates, for the UCB data, as discussed in Section 4.6 above.

- For our data, fBm was inappropriate for modelling the real traffic rates with varying transfer rates. When the transfer rates were artificially set to be constant, in particular for the UCB syn trace, there was good fit. It is clear that also at coarse resolutions actual network traffic is strongly influenced by flow control mechanisms such as TCP, server behavior, congestion, caching strategies and other factors. To obtain useful and realistic models, these factors cannot be ignored. Since much of the network traffic passes through a large and varying number of routers and switches, simplistic modelling of the behavior of the queue in one router is unlikely to achieve acceptable levels of realism. To find realistic and useful models for highly loaded systems serving large number of users which are subject to flow control, protocol modifications and congestion is an urgent area for further modelling research. The urgency is emphasized by the rapidly changing nature of the internet.

Acknowledgement: We are very grateful to Helmut Gogl, for providing us with the Munich lo and Munich hi data sets, and to P. Abry, for letting us use his wavelet estimation program, and to both for illuminating discussions. We also want to thank M. Greiner, Å. Arvidsson, P. Abry, W. Willinger, Steve Gribble, and J. Istas for useful conversations and help with details and references.

REFERENCES

- [1] P. Abry. *Ondelettes et Turbulences - Multirésolutions, Algorithmes de Décompositions, Invariance d'Echelle et Signaux de Pression*. Diderot, Editeur des Sciences et des Arts, Paris, 1997.
- [2] P. Abry, L. Delbecke and P. Flandrin. Waveletbased estimator for the self-similarity parameter of α -stable processes. Preprint 1999
- [3] P. Abry, P. Flandrin, M. S. Taqqu, and D. Veitch. Wavelets for the analysis, estimation and synthesis of scaling data. 1998.
- [4] P. Abry and D. Veitch. Wavelet analysis of long-range-dependent traffic. *IEEE Trans. Inform. Theory*, **44**, 2–15, 1998.
- [5] R. J. Adler. *An introduction to continuity, extrema, and related topics for general Gaussian processes*. I.M.S. Lecture Notes-Monograph Series, **12**, 1990.
- [6] R.J. Adler, R. Feldman, and M.S. Taqqu, *A practical guide to heavy tails: statistical techniques and applications* Birkhauser, Boston, 1998.
- [7] M. Arlitt and C.L. Williamson. Web server workload characterization: The search for invariants (extended version). In *Proceedings of the ACM Sigmetrics Conference, Philadelphia, Pa*, 1996. Available from {mfal26,carey}@cs.usask.ca.
- [8] S. Asmussen. *Applied Probability and Queues*. Wiley, Chichester, West Sussex, UK, 1987.
- [9] P. Barford, A. Bestavros, A. Bradley, and M. E. Crovella. Changes in Web Client Access Patterns: Characteristics and Caching Implications. *World Wide Web, Special Issue on Characterization and Performance Evaluation*, **2**, 15–28, 1999.
- [10] J. Beirlant, P. Vynckier, and J. Teugels. Tail index estimation, Pareto quantile plots, and regression diagnostics. *J. Amer. Statist. Assoc.*, **91**, 1659–1667, 1996.
- [11] J. Beran. *Statistics for long-memory processes* Chapman & Hall, New York, 1994.
- [12] S. Cambanis and C. Houdré. On the continuous wavelet transform of second-order random processes. *IEEE Trans. Inform. Theory*, **41**, 628–642, 1995.
- [13] J. Cohen, S. Resnick, and G. Samorodnitsky. Sample correlations of infinite variance time series models: and empirical and theoretical study. *Journal of Applied Mathematics and Stochastic Analysis*, **1**, 255–282, 1998.
- [14] A. G. Constantine and P. Hall. Characterizing surface smoothness via estimation of effective fractal dimension. *J. R. Statist. Soc. B*, **56**, 97–113, 1994.
- [15] M. Crovella and A. Bestavros. Explaining world wide web traffic self-similarity. Preprint available as TR-95-015 from {crovella,best}@cs.bu.edu, 1995.
- [16] M. Crovella and A. Bestavros. Self-similarity in world wide web traffic: evidence and possible causes. In *Proceedings of the 1996 ACM SIGMETRICS International Conference on Measurement and Modelling of Computer Systems*, 1996.
- [17] M. Crovella, A. Bestavros, and M. Taqqu. Heavy-tailed distributions in the world wide web. In *A practical guide to heavy tails: statistical techniques and applications* Birkhauser, Boston, 1998.
- [18] M. Crovella, G. Kim, and K. Park. On the relationship between file sizes, transport protocols, and self-similar network traffic. In *Proceedings of the Fourth International Conference on Network Protocols (ICNP'96)*, pages 171–180, 1996.
- [19] M. Crovella, K. Park, and G. Kim. On the effect of traffic self-similarity on network performance. In *Proceedings of the SPIE International Conference on Performance and Control of Network Systems*, 1997.
- [20] C.R. Cunha, M. Crovella and A. Bestavros. Characteristics of www client-based traces. Preprint BU-CS-95-010, Boston University, 1995.
- [21] I. Daubechies. *Ten lectures on wavelets* SIAM, Philadelphia, PA, 1992.
- [22] R. Davis and S.I. Resnick. Limit theory for bilinear processes with heavy tailed noise *Ann. Applied Probability*, **6**, 1191–1210, 1996.
- [23] R. Durrett and S. Resnick. Weak convergence with random indices. *J. Stoch. Proc. Appl.*, **5**, 213–220, 1977.
- [24] P. Embrechts, C. Klüppelberg, and T. Mikosh. *Modelling extremal events for insurance and finance*. Springer, Heidelberg, 1997.
- [25] A. Erramilli, O. Narayan, and W. Willinger. Experimental queueing analysis with long-range dependent packet traffic. *IEEE/ACM Transactions on Network Computing*, **4**, 209–223, 1996.
- [26] A. Feldmann, A. C. Gilbert, and W. Willinger. Data networks as cascades: Investigating the multifractal nature of Internet WAN traffic. *Proc. of the ACM/SIGCOMM'98*, Vancouver, B.C., 25–38, 1998.
- [27] P. Flandrin. Wavelets analysis and synthesis of fractional Brownian motion. *IEEE Trans. Inform. Th.*, **38**, 910–917, 1992.

- [28] P. Feigin and S. Resnick. Pitfalls of fitting autoregressive models for heavy-tailed time series. *Extremes* **1**, 391–422, 1999.
- [29] A.C. Gilbert, W. Willinger, and A. Feldmann. Scaling analysis of conservative cascades, with applications to network traffic. *IEEE Transactions on Information Theory*, **45**, 3, 971–991, 1999.
- [30] H. Gogl. Data Acquisition and Data Analysis in ATM based Networks. Report 1998, Technical University of Munich, available at http://www.jessen.informatik.tu-muenchen.de/forschung/leistungsanalyse/ATM_Project/atm_index.html.
- [31] H. Gogl. Traffic statistics and application profiles in ATM networks. Report 1998, Technical University of Munich, available at http://www.jessen.informatik.tu-muenchen.de/forschung/leistungsanalyse/ATM_Project/atm_index.html.
- [32] E.G. Gladyshev. A new limit theorem for processes with Gaussian increments. *Theory Probab. Appl.*, **6**, 52–61, 1961.
- [33] C.A. Guérin. Wavelet analysis and covariance structure of some classes of non-stationary processes. Technical report, Department of Mathematics, Chalmers University, Göteborg, 1998. To appear: *J. Fourier Anal. Appl.*
- [34] P. Hall and A. Wood. On the performance of box-counting estimators of fractal dimension. *Biometrika*, **80**, 246–252, 1993.
- [35] P. Hall, A. Wood and A. Feuerverger. Estimation of fractal index and fractal dimension of a Gaussian process by counting the number of level crossings. *J. of Time Ser. Anal.*, **15**, n° 6, 587–606, 1994.
- [36] J. Michael Harrison. *Brownian Motion and Stochastic Flow Systems*. Wiley, New York, 1985.
- [37] D. Heath, S. Resnick, and G. Samorodnitsky. Heavy tails and long range dependence in on/off processes and associated fluid models. *Math. Oper. Res.*, **23**, 145–165, 1998.
- [38] D. Heath, S. Resnick, and G. Samorodnitsky. How system performance is affected by the interplay of averages in a fluid queue with long range dependence induced by heavy tails. *Annals of Applied Probability*, **9**, 352–375, 1999.
- [39] B.M. Hill. A simple general approach to inference about the tail of a distribution. *Ann. Statist.*, **3**, 1163–1174, 1975.
- [40] F. Hunt. Error analysis and convergence of capacity dimension algorithms. *SIAM J. Appl. Math.*, **50**, 307–321.
- [41] I. A. Ibragimov and Y. A. Rozanov. *Gaussian random processes*, Springer Verlag, 1978.
- [42] J. Istas. Wavelets coefficients of a Gaussian process and applications. *Ann. Inst. Henri Poincaré*, **28**, 537–556, 1992.
- [43] J. Istas and G. Lang. Quadratic variations and estimation of the local Hölder index of a Gaussian process. *Ann. Inst. Henri Poincaré*, **33**, 407–436, 1997.
- [44] J. Istas and C. Laredo. Estimation de la fonction de covariance d'un processus gaussien stationnaire par méthodes d'échelles. *C. R. Acad. Sci., Tome 316, Série I*, 495–498, 1993.
- [45] P.R. Jelenković and A.A. Lazar. Subexponential asymptotics of a Markov-modulated G/G/1 queue. Technical report, Columbia University, 1995.
- [46] P.R. Jelenković and A.A. Lazar. Multiplexing on/off sources with subexponential on-periods. Technical Report CTR 457-96-23, Columbia University, 1996.
- [47] P.R. Jelenković and A.A. Lazar. Asymptotic results for multiplexing subexponential on-off sources. Technical report, Columbia University, 1997.
- [48] I. Kaj. Convergence of scaled renewal processes to fractional Brownian motion. Technical report, available at <http://www.math.uu.se/~ikaj/papers.html>, 1999.
- [49] J. T. Kent and A. T. A Wood. Estimating the fractal dimension of a locally self-similar Gaussian process by using increments. *J. R. Statist. Soc. B*, **59**, n° 3, 679–699, 1997.
- [50] C. Klüppelberg and T. Mikosch. Explosive Poisson shot noise processes with applications to risk reserves. *Bernoulli*, **1**, 125–148, 1995.
- [51] C. Klüppelberg and T. Mikosch. Modelling delay in claim settlement. *Scand. Actuar. Journal*, 154–168, 1995.
- [52] T. Konstantopoulos and Si-Jian Lin. Macroscopic models for long-range dependent network traffic. *Queueing Systems Theory Appl.*, **28**, 215–243, 1998.
- [53] M. Kratz and S. Resnick. The qq-estimator and heavy tails. *Stochastic Models*, 12:699–724, 1996.
- [54] T. Kurtz. Limit theorems for workload input models. In S. Zachary F.P. Kelly and I. Ziedins, editors, *Stochastic Networks: Theory and Applications*, Oxford, 1996. Clarendon Press.
- [55] J.W. Lamperti. Semi-stable stochastic processes. *Trans. Amer. Math. Soc.*, 62–78, 1962
- [56] G. Lang. Estimation de la régularité et de la longue dépendance de processus gaussiens. Thèse de l'université Toulouse III, 1994.

- [57] W.E. Leland, M.S. Taqqu, W. Willinger, and D.V. Wilson. On the self-similar nature of Ethernet traffic. *ACM/SIGCOMM Computer Communications Review*, 1993. Proceedings of the ACM/SIGCOMM93. To appear.
- [58] W.E. Leland, M.S. Taqqu, W. Willinger, and D.V. Wilson. On the self-similar nature of Ethernet traffic (extended version). *IEEE/ACM Transactions on Networking*, **2**, 1–15, 1994.
- [59] Will E. Leland, Murad S. Taqqu, Walter Willinger, and Daniel V. Wilson. Statistical analysis of high time-resolution ethernet Lan traffic measurements. In *PCmpScSt25*, 146–155, 1993.
- [60] B. Mandelbrot and J. Van Ness. Fractional Brownian motions, fractional Brownian noises and applications. *SIAM Review*, **10**, 422–437, 1968.
- [61] T. Mikosch, S. Resnick, H. Rootzén, and A. Stegeman. Is network traffic approximated by stable Lévy motion or fractional Brownian motion? Preprint 1999, available as #1999:32 at <http://www.md.chalmers.se/Stat/Research/Preprints>.
- [62] J.P.Nolan Maximum likelihood estimation and diagnostics for stable processes. Available at <http://www.cas.american.edu/jpnolan/stable.html>.
- [63] N.U. Prabhu. *Stochastic Storage Processes: Queues, Insurance risk, Dams and Data Communication*. Springer, New York, 1998.
- [64] R.D. Reiss and M. Thomas. *Statistical analysis of extreme values*. Birghäuser, Basel, 1997.
- [65] S.I. Resnick. *Extreme Values, Regular Variation and Point Processes*. Springer, New York, 1987.
- [66] S.I. Resnick. *Heavy tail modelling and teletraffic data*. Ann. Statist., **25**, 1805–1869, 1997.
- [67] S. Resnick and G. Samorodnitsky. Performance decay in a single server exponential queueing model with long range dependence. *Operations Research*, **45**, 235–243, 1997.
- [68] S. Resnick and G. Samorodnitsky and F. Xue. How misleading can sample acf's of stable ma's be? (Very!). Technical report 1209; available at www.orie.cornell.edu/trlist/trlist.html; to appear: Annals of Applied Probability, 1998
- [69] S. Resnick and C. Stărică. Smoothing the Hill estimator. *Advances in Applied Probability*, **29**, 271–293, 1997.
- [70] S.I. Resnick and H. Rootzén. Self-similar communication models and very heavy tails. *Ann. Appl. Probab.* to appear. Available as TR1228.ps.Z at www.orie.cornell.edu/trlist/trlist.html, 1998.
- [71] S. Resnick and E. van den Berg. A test for nonlinearity of time series with infinite variance. Technical report 1222, available at www.orie.cornell.edu/trlist/trlist.html; To appear: Extremes, 1998.
- [72] R.H. Riedi. Multifractal Processes. Technical report, ECE Dept., Rice University, TR 99-06, 1999.
- [73] R. H. Riedi and W. Willinger. Toward an improved understanding of network traffic dynamics. In: Self-similar Network Traffic and Performance Evaluation; to appear with Wiley, June 1999.
- [74] G. Samorodnitsky and M.S. Taqqu. $(1/\alpha)$ -self-similar processes with stationary increments. *J. Multivariate Analysis*, **35**, 308–313, 1990.
- [75] G. Samorodnitsky and M. Taqqu. *Stable non-Gaussian random processes* Chapman and Hall, New York, 1994.
- [76] N. Tajvidi. Confidence Intervals and Accuracy Estimation for Heavy tailed Generalised Pareto Distribution. Part of PhD thesis, 1996, Chalmers, available at <http://www.mai.liu.se/nataj/>
- [77] M. Taqqu, W. Willinger, and R. Sherman. Proof of a fundamental result in self-similar traffic modelling. *Computer Communications Review*, **27**, 1997.
- [78] B.B. Mandelbrot and M.S. Taqqu. Robust R/S analysis of long-run serial correlation. *Proceedings of the 42nd Session of the International Statistical Institute, Manila*, Vol.48, Book 2, pp. 69–104, Bulletin of the ISI, 1979.
- [79] A.H. Tewfik and M. Kim. Correlation structure of the discrete wavelet coefficients of fractional Brownian motion. *IEEE Trans. Inform. Th.*, **38**, 904–909, 1992.
- [80] D. Veitch and P. Abry. Estimation conjointe en ondelettes des paramètres du phénomène de dépendance longue. In *Proc. 16ième Colloque GRETSI, Grenoble, France*, 1451–1454, 1997.
- [81] D. Veitch and P. Abry. A wavelet based joint estimator for the parameters of lrd. *IEEE Trans. Info. Th.*, 1998. special issue "Multiscale Statistical Signal Analysis and its Application".
- [82] Ward Whitt. Some useful functions for functional limit theorems. *Math. Oper. Res.*, **5**, 67–85, 1980.
- [83] W. Willinger, M.S. Taqqu, M. Leland, and D. Wilson. Self-similarity in high-speed packet traffic: analysis and modelling of ethernet traffic measurements. *Statistical Science*, **10**, 67–85, 1995.
- [84] W. Willinger, M.S. Taqqu, M. Leland, and D. Wilson. Self-similarity through high variability: statistical analysis of ethernet lan traffic at the source level. *Computer Communications Review*, **25**, 100–113, 1995. Proceedings of the ACM/SIGCOMM'95, Cambridge, MA.

- [85] W. Willinger, M.S. Taqqu, M. Leland, and D. Wilson. Self-similarity through high variability: statistical analysis of ethernet lan traffic at the source level (extended version). *IEEE/ACM Transactions on Networking*, 1996. To appear.
- [86] A.T.A. Wood, and G. Chan. Simulation of stationary Gaussian Processes in $[0, 1]^d$. *Journal of Computational and Graphical Statistics*, **3**, 409-432, 1994.

CHARLES-ANTOINE GUERIN, LABORATOIRE D'OPTIQUE ELECTROMAGNETIQUE, CASE 162, FACULTE DES SCIENCES DE SAINT-JEROME, F-13397 MARSEILLE CEDEX 20, FRANCE

E-mail address: `caguerin@loe.u-3mrs.fr`

HENRIK NYBERG, ERICSSON RADIO SYSTEMS AB, ERICSSON RESEARCH, ACCESS NETWORKS, KISTA, SWEDEN

E-mail address: `henrik.nyberg@era-t.ericsson.se`

OLIVIER PERRIN, GREMAQ - UNIVERSITY TOULOUSE I, MANUFACTURE DES TABACS, 21, ALLE DE BRIENNE, 31000 TOULOUSE, FRANCE

E-mail address: `perrin@cict.fr`

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

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

HOLGER ROOTZÉN, DEPARTMENT OF MATHEMATICS, CHALMERS UNIVERSITY OF TECHNOLOGY, S-412 96 GOTHENBURG, SWEDEN

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

CATALIN STĂRICĂ, DEPARTMENT OF MATHEMATICS, CHALMERS UNIVERSITY OF TECHNOLOGY, S-412 96 GOTHENBURG, SWEDEN

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