Fractional Brownian motion and data traffic modeling: The other end of the spectrum

Jacques Lévy Véhel and Rudolf Riedi

INRIA - Groupe Fractales, B.P. 105, 78153 Le Chesnay CEDEX, France
e-mail: {Jacques.Levy_Vehel,Rolf.Riedi}@inria.fr
Article as it appeared in "Fractals in Engineering", Eds. J. Lévy Véhel, E. Lutton and C. Tricot, Springer 1997

Summary. We analyze the fractal behavior of the high frequency part of the Fourier spectrum of fBm using multifractal analysis and show that it is not consistent with what is measured on real traffic traces. We propose two extensions of fBm which come closer to actual traffic traces multifractal properties.

Keywords: fractional Brownian motion, Multifractal analysis, TCP

1. Introduction

Fractal analysis of computer traffic has received considerable attention since the seminal work of Leland and al. [11] who provided experimental evidence that some traces of data traffic exhibit long range dependence (LRD). This is a typical fractal feature which is not found with the classical Poisson models. An important issue since then has been to propose "physical" models that lead to such fractal behavior. A popular model [27] is based on the superposition of simple i.i.d ON/OFF sources which ON and/or OFF periods follow a heavy tailed law ($Pr(X > \lambda) \sim c\lambda^{-\alpha}, 1 < \alpha < 2$). When properly normalized, the resulting traffic is a fractional Brownian motion (fBm) of LRD exponent $H = (3 - \alpha)/2$. Several practical implications of LRD traffic have consequently been investigated, e.g. the queuing behavior [15] (see also the two previous papers in this volume). Experimental justifications of the ON/OFF models have also been searched for, one of them being found in the heavy tailed laws of computer file sizes in the context of WWW traffic [6].

In [20], we performed an experimental study of traffic traces to describe fractal features different from LRD and based on Multifractal analysis (MA). Roughly speaking, MA is concerned with "the other end" of the Fourier spectrum, i.e. it looks at the high frequency content of the signal instead of the low frequency one. For some simple fractal models such as fBm, the Hurst exponent H ruling the LRD and, thus, the low frequencies governs also the local singular behavior which is related to the high frequencies. For general processes, however, the local singular behavior depends on time and is best described by the multifractal spectrum of singularity exponents. At this point it is worthwhile noting that studies of the distributions of the sizes of files provide a solid explanation of the LRD behavior of data traffic on large time scales. They do not account, however, for the short time behavior which results from the fragmentation and aggregation in protocol networks and from waiting queues. In some applications such as traffic control, understanding the structure of short term variations, and in particular their maximum amplitude is important. This information is comprised in the multifractal spectrum. Furthermore, MA may enable to evidence differences between traffic types (e.g. incoming/outgoing, LAN/WAN, ftp/WWW, etc...) that remain hidden to a second order analysis such as LRD characterization. Finally, multifractals offer a simple way of building realistic synthetic traces of traffic. MA may thus help to assess the adequacy of previously proposed models by looking at properties different from LRD.

It is our aim in this work to check whether the results of the MA of traffic traces as performed in [20] and [14] is compatible with a modeling based on fBm, and if not, to propose some simple extensions that allow a better account of the multifractal properties. In that view, and after some recalls on MA in Section 2., we compute in Section 3. the multifractal spectrum of fBm, both theoretically and estimated numerically on synthetic traces of fBm.

The comparison of this spectrum with the ones obtained on real traffic traces shows that more refined models are needed. As a first step in this direction, we study in Section 4. generalizations that involve 1) lumpings of fBm-s with different exponents, and 2) multifractional Brownian motion (mBm), which is basically an fBm with continuously varying H [18]. The motivation for considering such models is to account for a possible non stationarity in the data while loosely keeping some of the physical explanation provided by ON/OFF type models. Again, theoretical and numerically estimated spectra are produced and compared to spectra of traces of real traffic.

2. Multifractal analysis

MA deals with the description of the singularity structure of "signals" (which can be measures [3, 4, 12, 13, 19], functions [9] or capacities [26]), both in a local and a global way. The local information is given by the *Hölder exponent* at each point, while the global information is captured through a characterization of the geometrical or statistical distribution of the occurring Hölder exponents, called *multifractal spectrum*. Such an analysis is useful when one deals with very irregular signals (such as traffic traces), and when the singularity structure has practical consequences [24].

There are several ways of measuring the local irregularity, and different global characterizations may be considered. We briefly recall in this section some of the most classical approaches.

2.1 Local irregularity

Among several ways of measuring the local irregularity we will mention two. For simplicity we consider a signal $X(t)$ defined on [0,1] which is nowhere differentiable[1] and define the analysis with respect to the dyadic intervals:

$$I_n^k = [k2^{-n}, (k+1)2^{-n}[, \ k = 0, \dots 2^n - 1, n \in \mathbb{N}$$

Define the *coarse Hölder exponents* through

$$\alpha_n^k = -\frac{1}{n}\log\left|X((k+1)2^{-n}) - X(k2^{-n})\right|$$

where all logarithms are taken to the base 2 and where $\log 0 := -\infty$. For a fixed t in [0,1] let (k_n) be such that $I_n(t) := I_n^{k_n}$ contains t. Then, $k_n 2^{-n} \to t$ as $n \to \infty$. The limiting exponent at t

$$\alpha(t) := \liminf_{n\to\infty} \alpha_n^{k_n}$$

is called the *local* Hölder exponent of X at t. We mention in passing that the exponent considered usually is

$$\tilde{\alpha}(t) := \liminf_{\varepsilon\to 0} \tilde{\alpha}_\varepsilon(t) \quad \text{where} \quad \tilde{\alpha}_\varepsilon(t) = \frac{1}{\log\varepsilon}\log \sup_{s,s'\in B(t,\varepsilon)} |X(s) - X(s')|.$$

Since it is better adapted to numerical computations we will consider mainly α.

Multifractal analysis consists in giving a compact representation of the local singularity structure of a signal as measured through $\alpha(t)$. In that view, two approaches arise naturally: either use a *geometrical* description, or use a *statistical* one. The former leads to the Hausdorff spectrum f_h which we will not address further, while the latter leads to the large deviation spectrum f_g. A third spectrum is also considered at the end of this section, namely the Legendre spectrum, the interest of which is to provide a simple way to compute f_g or f_h when some conditions are met.

2.2 Large deviation spectrum

The large deviation spectrum f_g measures, loosely speaking, how "fast" the probability of observing a coarse Hölder exponent different from the expected value tends to zero as the resolution tends to ∞. More precisely, f_g is related to the rate function appearing in the large deviation analysis of such quantities. A heuristic explanation is the following: assume we want to assess how much the traffic T_n may vary in a (small) time interval I_n^k of duration 2^{-n}, with respect to n. The strength of the variation is measured by α_n^k, defined as: $T_n \propto |I_n^k|^{\alpha_n^k} = 2^{-n\alpha_n^k}$. If $\alpha_n^k = 1$, the traffic varies smoothly with respect to the scale of measurement ε. The smaller α_n^k is, the more intense the traffic is, while values of α_n^k greater than one correspond to sparse traffic.

It is thus of interest to evaluate the distribution of the α_n^k's when k is picked randomly from $0, \dots, 2^{n-1}$ and n is large. This will allow to characterize the traffic in terms of:

[1] This way we will not have to deal with polynomial trends.

– the highest degree of burstiness (resp. sparseness) occurring in the signal,
– the probability of hitting a given burstiness when measuring the traffic in a "small" time interval picked randomly.

In this view, we set:

$$f_g(\alpha) = \lim_{\varepsilon \to 0} \limsup_{n \to \infty} \frac{\log N_n^\varepsilon(\alpha)}{n}, \tag{2.1}$$

where

$$N_n^\varepsilon(\alpha) = \#\{\alpha_n^k / |\alpha_n^k - \alpha| < \varepsilon\}.$$

The large deviation spectrum f_g describes the distribution of the local singularities, since the number of dyadic intervals of size 2^{-n} with coarse Hölder exponent $\simeq \alpha$ varies roughly as $2^{nf_g(\alpha)}$ for large n. Equivalently:

$$P_n(\alpha_n^k \approx \alpha) \propto 2^{-n(1-f_g(\alpha))}$$

where the probability is related to a random choice of k uniformly in $\{0, \ldots, 2^n - 1\}$, i.e. P_n is the uniform distribution on the set of all dyadic intervals I_n^k of size 2^{-n}.

2.3 Legendre spectrum

It is natural to interpret the spectrum f_g as a rate function in a *large deviation principle* (LDP). The general theory on large deviations provides conditions under which such rate functions may be calculated as the Legendre transform of a limiting moment generating function. As we are about to show this procedure provides a much more robust estimation of f_g than a direct computation via (2.1), provided that the data satisfy the necessary conditions.

From a multifractal point of view it is most convenient to define for $q \in \mathbb{R}$:

$$S_n(q) = \sum_{k=0}^{2^n-1} |X((k+1)2^{-n}) - X(k2^{-n})|^q \tag{2.2}$$

with the convention $0^q := 0$ for all $q \in \mathbb{R}$. Next, define the 'structure function'

$$\tau(q) = \liminf_{n \to \infty} \frac{\log S_n(q)}{-n}$$

and the so-called *Legendre spectrum* of X

$$f_l(\alpha) = \tau^*(\alpha) := \inf_{q \in \mathbb{R}} (\alpha q - \tau(q)).$$

Returning to the large deviation principle, let us follow [19] and define the sequence of random variables $Z_n := \log Y_n^k$ where k is distributed with P_n as before. Consider the corresponding moment generating functions

$$c_n(q) := -\frac{1}{n}\log \mathbb{E}_n[\exp(qZ_n)] = -\frac{1}{n}\log\left(2^{-n}S_n(q)\right).$$

A version of the Gärtner-Ellis theorem on LDP [7] allows then to conclude that if $\lim c_n(q)$ exists (in which case it equals $1+\tau(q)$), and is differentiable, then $c^* = f_g - 1$. In other words, the so-called *weak multifractal formalism*:

$$f_l(\alpha) = f_g(\alpha)$$

holds then. This case is favorable since $\tau(q)$, being computed by averaging and involving a single limit, is easier to estimate than f_g which requires the evaluation of local quantities and of a double limit.

In general, however, these functions may differ, and one has:

Lemma 2.1 ([19, 21]). *f_l is the concave hull of f_g, i.e. $\tau(q) = f_l^*(q)$ and $f_l(\alpha) = f_g^{**}(\alpha)$. Consequently, $\tau(q)$ is concave and $f_g(\alpha) \le f_l(\alpha)$. Furthermore,*

$$\begin{array}{rclr} f_g(\alpha^+) & = & q\alpha^+ - \tau(q) & (q > 0) \\ f_g(\alpha^-) & = & q\alpha^- - \tau(q) & (q < 0) \end{array} \tag{2.3}$$

where $\alpha^+ := \tau'(q+)$ and $\alpha^- := \tau'(q-)$ denote the one-sided derivatives of $\tau(q)$. Finally, due to the special form of P_n f_g cannot assume negative, finite values. Consequently, $\tau(q)$ is finite either in all of $\mathbb{R}$ or exactly in $q \ge 0$.

This shows that f_g contains in general more information than f_l.

3. Multifractal Analysis of fBm

Let $X(t)$ be an fBm of exponent H. As is well known [23] $X(t)$ is a zero mean Gaussian process with stationary increments and is statistically self-similar in the sense of finite-dimensional distributions. It follows that $X(0) = 0$ a.s., and that

$$Y_n^k := X((k+1)2^{-n}) - X(k2^{-n}) \stackrel{d}{=} X(2^{-n}) \stackrel{d}{=} N(0, 2^{-nH}) \tag{3.1}$$

where $\stackrel{d}{=}$ means equality in distribution and where $N(m,\sigma)$ stands for the normal distribution with mean m and variance σ.

The LRD property of fBm manifests itself in the asymptotic behavior of the Fourier spectrum $S(\lambda) \sim |\lambda|^{1-2H}$ as the frequency λ tends to zero. The high frequency content of X is also ruled by H: one can in particular show that with probability one $\tilde{\alpha}(t) = H$ for all t [2]. fBm displays, thus, a very peculiar fractal behavior, where the same parameter H allows to describe both parts of the Fourier spectrum.

We continue this section by computing the large deviation and the Legendre spectrum of fBm, elaborating especially on the complications induced by the randomness. Indeed, measured along paths, f_g and f_l are random functions and hard to compute. As a first step towards this goal, we take expectations before going to the limits (Subsection 3.1). The spectra F_g and F_l obtained in this way are deterministic and represent the rate functions of a LDP in a higher probability space. Using estimates between f_l and F_l we are then able to treat the pathwise spectra in Subsection 3.2. It will not surprise that all spectra are essentially determined by H.

3.1 Deterministic spectra

For the calculation of the multifractal spectra one would like to exploit this self-similarity in distribution. Therefore, we take expectation in (2.2) and set

$$\overline{S}_n(q) := \mathbb{E}_\omega S_n(q,\omega) = \mathbb{E}_\omega \sum_{k=0}^{2^n-1} |Y_n^k|^q = 2^n \mathbb{E}_\omega |Y_n^1|^q, \tag{3.2}$$

where we used the stationarity of increments. Letting

$$c_q := \mathbb{E}_\omega |Y_1^1|^q = \mathbb{E}_\omega |X(1)|^q = \frac{2}{\sqrt{2\pi}} \int_0^\infty y^q e^{-y^2/2} dy,$$

which is finite exactly for $q > -1$, we find $\mathbb{E}_\omega |Y_n^1|^q = 2^{-nqH} c_q$. Thus, one finds the scaling law

$$T(q) := \lim_{n\to\infty} \frac{\log \overline{S}_n(q)}{-n} = \begin{cases} -\infty & \text{for } q \le -1 \\ qH - 1 & \text{for } q > -1, \end{cases}$$

which is — up to the constant 1 — the structure function of an LDP on the sequence of probability spaces $(\Omega \times \{1,\ldots,2^n\}, P_\omega \times P_n)$. Note in particular the built-in averaging on the dyadic intervals through S_n. The Legendre transform of T is

$$F_l(\alpha) = \inf_q (q\alpha - T(q)) = \inf_{q>-1} (q\alpha - T(q)) = \begin{cases} -\infty & \text{for } \alpha < H \\ 1 + H - \alpha & \text{for } \alpha \ge H. \end{cases}$$

We will address F_l as the 'deterministic Legendre spectrum'. To define the corresponding 'deterministic spectrum of large deviations' set

$$\pi_n^\varepsilon(\alpha) := P_\omega \times P_n[\alpha_n^k \in (\alpha - \varepsilon, \alpha + \varepsilon)]$$

and define

$$F_g(\alpha) = 1 + \lim_{\varepsilon\to 0} \limsup_{n\to\infty} \frac{\log \pi_n^\varepsilon(\alpha)}{n}.$$

Note, first of all, that F_g may assume non-trivial negative values, whence the last conclusion of Lemma 2.1 is not valid. The other claims, though, continue to hold for this more general 'random procedure' of choosing α_n^k's.

We argue that $F_g(\alpha)$ should be a good estimator of $f_g(\alpha,\omega)$, the usual spectrum measured along paths. Indeed, equality holds a.s. at least for $H = 1/2$ (Lemma 3.2). In general, the increments of fBm with constant lag form a stationary process which is ergodic under time-shift [8]. This implies that $N_n^\varepsilon(\alpha)$ divided by the total number 2^n of dyadic intervals considered should be well approximated by $\pi_n^\varepsilon(\alpha)$. Also,

$$\mathbb{E}_\omega[N_n^\varepsilon(\alpha)/2^n] = \frac{1}{2^n}\mathbb{E}_\omega \sum_{k=0}^{2^n-1} \mathbf{1}_{(\alpha-\varepsilon,\alpha+\varepsilon)}(\alpha_n^k) = \pi_n^\varepsilon(\alpha),$$

which closes the circle started with (3.2). For a computation of F_g note that

$$\pi_n^\varepsilon(\alpha) = \frac{2^{nH+1}}{\sqrt{2\pi}} \int_{2^{-(\alpha+\varepsilon)n}}^{2^{-(\alpha-\varepsilon)n}} \exp\left(-\frac{x^2}{2^{-2nH+1}}\right) dx.$$

A crude estimate of the monotonous integrand gives [25]

$$F_g(\alpha) = \begin{cases} -\infty & \text{for } \alpha < H \\ 1 + H - \alpha & \text{for } \alpha \geq H. \end{cases} \tag{3.3}$$

Note that F_g assumes negative values, which is not possible for f_g. Consequently, F_g may be expected to be a good estimator of f_g only where it is positive. Negative $F_g(\alpha)$ values correspond to probabilities of observing a coarse Hölder exponent α which decay faster than the number of 'samples' α_n^k considered. Oversampling the process, i.e. analyzing several independent realizations will, thus, result in larger observed α_n^k. In fact, in $\exp(-n\ln(2)F_g(\alpha))$ independent traces one has a fair chance to see at least one increment $|Y_k^n|$ with coarse Hölder exponent α.

Let us note that

$$F_l = F_g$$

which is a weak form of the so-called multifractal formalism.

3.2 Path-wise spectra of fBm

Let us return now to the spectra calculated along a path of fBm. Though $2^{nqH-n}S_n$ is in general not equally distributed to

$$\frac{1}{2^n}\sum_{k=0}^{2^n-1} |Y_1^k|^q. \tag{3.4}$$

it is at least equal in expectation. As we will show now in a general argument the study of $\overline{S}$ leads to the right normalization term. For any q with finite $T(q)$ and $\varepsilon > 0$

$$\begin{aligned} \mathbb{E}\limsup_{n\to\infty} 2^{n(T(q)-\varepsilon)}S_n(q,\omega) &\leq \mathbb{E}\sum_{n\in\mathbb{N}} 2^{n(T(q)-\varepsilon)}S_n(q,\omega) \\ &\leq \sum_{n\in\mathbb{N}} 2^{n(T(q)-\varepsilon)}\mathbb{E}S_n(q,\omega) \leq \frac{b(q,\varepsilon)}{1-2^{-\varepsilon/2}} < \infty, \end{aligned}$$

by definition of T. Here b is a suitable constant depending on q and ε only. This allows to conclude that almost surely $\limsup_{n\to\infty} 2^{n(T(q)-\varepsilon)} S_n(q,\omega) < \infty$. Whence, $\tau(q) \geq T(q) - \varepsilon$, which is trivial if $T(q) = -\infty$. It is clear that this estimate holds with probability one for all $\varepsilon = 1/m$ ($m \in \mathbb{N}$) and some countable, dense set of q values with $T(q) < \infty$ simultaneously. Using that $\tau(q)$ is always continuous on open sets due to Lemma 2.1 we find the following result valid for any process:

Lemma 3.1. *With probability one*

$$\tau(q,\omega) \geq T(q) \qquad \text{for all } q \text{ with } T(q) < \infty.$$

As an immediate consequence we obtain with Lemma 2.1 that with probability one

$$f_g(\cdot,\omega) \leq f_l(\cdot,\omega) \leq F_l \tag{3.5}$$

As a matter of fact, we have equality at least for $H = 1/2$, the ordinary Brownian motion due to

Lemma 3.2 ([25]). *For any process with independent and stationary increments and any α*

$$f_g(\alpha,\omega) = F_g(\alpha) \qquad a.s.$$

Due to the linearity of T the bound provided by lemma 3.1 is all we need. Indeed, for each n, $S_n(0,\omega)$ counts the number of non-vanishing increments which is 2^n almost surely. Thus, $\tau(0) = -1$ a.s. and since $\tau(q)$ must be concave the obtained bound is necessarily sharp.

In conclusion, with probability one:

$$\tau(q,\omega) = qH - 1 \qquad \text{for } q > -1, \tag{3.6}$$

and

$$f_g(\alpha,\omega) \begin{cases} = f_l(\alpha,\omega) & = -\infty & \text{for } \alpha < H \\ = f_l(\alpha,\omega) & = 1 & \text{for } \alpha = H \\ \leq f_l(\alpha,\omega) & \leq 1 + H - \alpha & \text{for } \alpha \geq H. \end{cases}$$

Finally, the implicit restrictions on $\tau(q)$ imply bounds for $q \leq -1$: with probability one

$$qH + q \leq \tau(q,\omega) \leq qH - 1 \qquad \text{for } q \leq -1. \tag{3.7}$$

The upper bound follows from concavity combined with (3.6), while the lower bound reflects the fact that $f_l(\alpha,\omega)$ is either positive or $-\infty$ (Lemma 2.1). This together with $f_l(\alpha,\omega) \leq 1 + H - \alpha$ implies that the slopes of τ have to be smaller than $1 + H$.

Let us end the section with a remark on the scaling behavior for $q \leq -1$. Following the intuitive approximation (3.4) one may argue with Aaronson [10, p 15] that there is no power γ which makes $2^{n\gamma} S_n$ converge a.s. to a positive, finite number, due to the infinite expectation of $|Y_1^1|^q$. For $H = 1/2$ we have independence which allows to apply the theory of infinitely divisible laws.

With $|Y_1^1|^q$ lying in the domain of attraction of an α-stable law with $\alpha = -1/q$ this suggests that a renormalization of the type $2^{n(qH+q)} S_n(q,\omega) - A_n$ might provide convergence in distribution to a stable law. In this result, we recognize the limiting value $Hq + q$ for $\tau(q)$. However, the weak sense of 'convergence in distribution' and the divergence of $A_n \simeq 2^{n(qH+q-(q+1)/q)}$ hint to a poor scaling behavior for $q \leq -1$. Indeed, in Subsection 3.3 we provide numerical evidence for unreliable estimates starting at $q \approx -1/2$ which corresponds to infinite variance of $|Y_1^1|^q$.

3.3 Experimental results

We show in this section estimated large deviation and Legendre spectra computed on synthetic fBm traces. The traces consist of 2^{14} points and were generated using the Choleski decomposition method with $H = 0.2, 0.6, 0.8$. The estimation procedures for both f_g and f_l are the ones described in [20]. We first compare log-log plots of the partition sum $S_n(q)$ for an fBm and a real traffic trace on Fig. 3.1. To illustrate more clearly whether the assumption of linearity of these log-log plots holds, Fig. 3.2 displays their increments $\log S_n(q) - \log S_{n+1}(q)$. If constant over a considerable range of n they provide a reliable estimate of $\tau(q)$.

From these figures it becomes apparent that fBm displays a good scaling behavior for q larger $\simeq -1/2$ while $\tau(q)$ does not seem to converge for lower values of q, a result which is consistent with (3.7). In contrast, the computation on real traces reveals a good scaling even for $q = -2$ and sufficient level of aggregation. This is a first point where the fBm model departs significantly from the data.[2]

Finally, Fig. 3.3 shows the estimated spectra. Both f_g and f_l almost perfectly match the theoretical expressions obtained in the previous sections, supporting in particular the fact that the positive part of F_l is indeed a good approximation of f_g. Again, the shapes differ clearly from the typical ones obtained on real traces [20].

3.4 Discussion

The preceding results show that neither spectrum of fBm, f_g nor f_l, displays the shape estimated on real traffic traces as presented in [20] and [14]. This calls for fractal traffic models which are more refined than fBm. The fact that an estimation of H alone is not sensitive to certain features observed with a multifractal analysis [20] further supports this point of view.

---

[2] One may argue that the bad scaling for negative q comes from the fact that fBm is a centered process. But adding a constant to fBm in order to obtain a positive trace worsens the situation by resulting in the trivial multifractal behavior $\tau(q) = q - 1$. Furthermore, the cut-off value for the divergence of τ_n is not 0 but lies between $-1/2$ and -1.

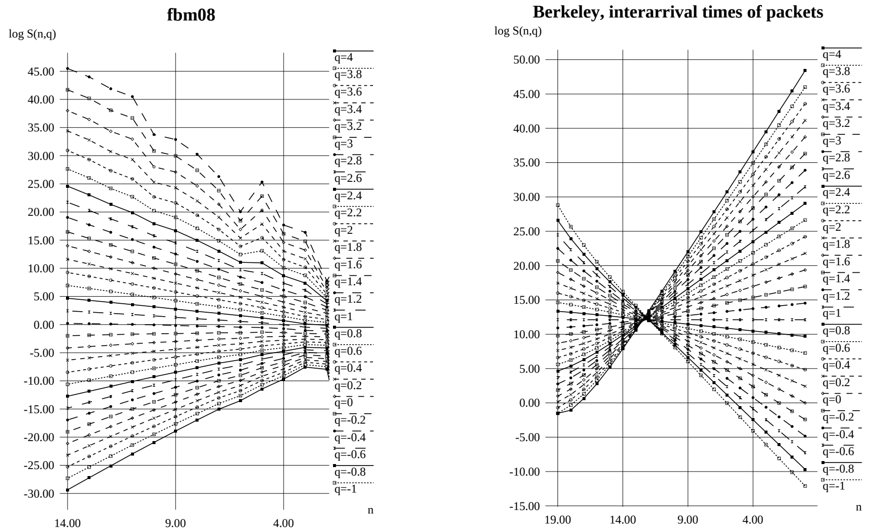

Fig. 3.1. Log-log plots of the partition sum $S_n(q)$ for fractional Brownian motion with $H = 0.8$ (left) and for the interarrival times of packets as observed at the gateway of Berkeley [20] (right).

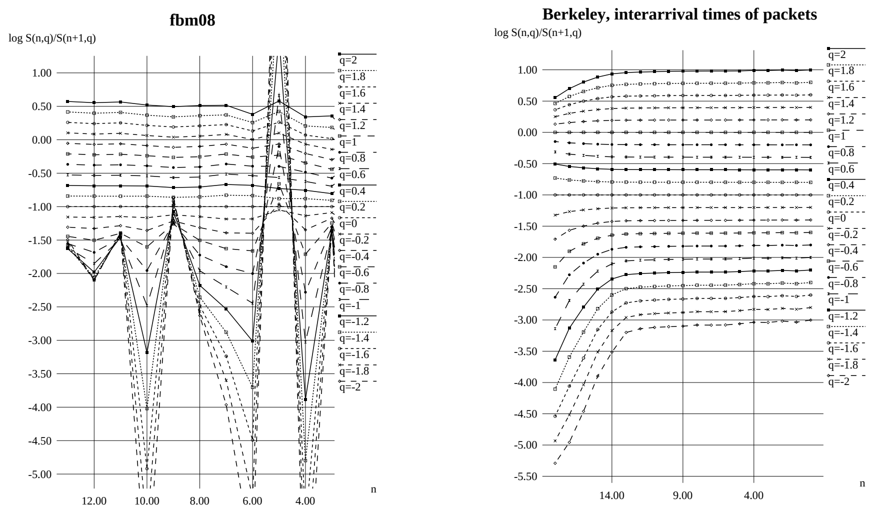

Fig. 3.2. Increments of the log-log plots of the partition sum $S_n(q)$ for fractional Brownian motion with $H = 0.8$ (left) and for the interarrival times of packets as observed at the gateway of Berkeley [20] (right).

The most simple multifractal processes bearing some analogies with real traffic traces are *random multiplicative cascades*. For a detailed introduction see [20]. The striking resemblance of such cascades with actual traces of real data traffic makes them a natural candidate for data traffic modeling.

If random cascades are to be invoked for justifying the multifractal behavior of real traffic, it is important to find "physical" reasons why such a multiplicative structure should occur. Loose equivalents of the multiplicative structure can be found in the fact that Telecommunication protocols are organized in a hierarchy and that these protocols work by splitting wholes into pieces. For instance, in the case of file transfer, we find at the highest level the decision to send a file. At the next level, the TCP control comes into play, and time sharing with other processes results in a random variability. The maximal window size and the speed regulation (e.g. slow start) still emphasizes this effect. Further levels are the maximal IP packet size, and finally, if the network is an ATM one, the breaking of IP packets into cells, which depends on the rate of the ATM virtual connection ([16]).

However, such explanations remain somewhat vague and much work is needed before one can come up with a physical model for multiplicative cascades as neat as the ON/OFF source model for fBm. Here, we propose a different approach which exhibits multifractal properties similar to real traffic traces, while retaining some of the physical explanation of the ON/OFF model.

The basic idea is simply to relax the hypothesis of stationarity. Although some experimental studies [1] conclude that the LRD parameter seems to remain constant over large time intervals, we argue that the exponent ruling the heavy tail of each individual ON/OFF source may undergo some changes in time, due to important intra-day variations (lunch time, busy hours, night-time). In addition, while H, as a parameter of LRD, should not depend much on the protocol used, it is conceivable that it varies with the content of the communication. Finally, it is likely that even minor intra-day variations can cause modifications in the local singular behavior, leading to a Hölder exponent that would depend on time.

4. Extensions of fBm

4.1 Lumpings

The most simple way of obtaining a non stationary process is to assume that, depending on the time of the day, we are observing different fBm processes with various exponents H. Formally, if $X_H(t)$ denotes an fBm with exponent H, the traffic model X is given by:

$$X(t) = X_{H_k}(t) \qquad \text{for } t \in [t_{k-1}, t_k[$$

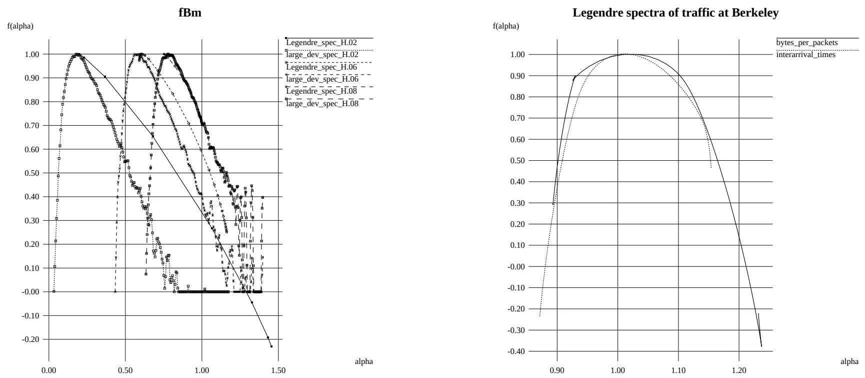

Fig. 3.3. Estimated large deviation spectrum and Legendre spectrum for fractional Brownian motion with $H = 0.2$, 0.6 and 0.8 (left) and for the interarrival times of packets as observed at the gateway of Berkeley [20] (right).

for $k = 1, \ldots, m$, where m is a fixed number and $t_0 := 0 < t_1 < t_2 \ldots < t_{m-1} < t_m := 1$. For simplicity let us assume that $t_k 2^n$ is integer for all k from some large n on. Since m is fixed and finite, and due to the built-in averaging over dyadic intervals in their definitions, $\overline{S}_n(q)$ and $\pi_n^\varepsilon(\alpha)$ will be composed of m terms corresponding to the various $X_{H_k}(t)$; e.g.

$$\pi_n^\varepsilon(\alpha) = \sum_{k=1}^{m} \frac{t_k - t_{k-1}}{2^n} \pi_n^\varepsilon[X_{H_k}](\alpha)$$

where we used stationarity in the intervals $[t_{k-1}, t_k[$. It is clear that largest scaling exponent will govern the asymptotic behavior of the overall sum. In short, $T(q)$ will be the minimum of the corresponding rate functions of the $X_{H_k}(t)$, while F_g will be the corresponding maximum.

Thus, F_g assumes a zig-zag shape, consisting of m lines of slope -1 which start in $(H_k, 1)$ and continue to $(H_{k+1}, 1 + H_k - H_{k+1})$ (with the convention $H_{m+1} = \infty$). Letting $\underline{H} = \min_{1 \le i \le m} H_i, \overline{H} = \max_{1 \le i \le m} H_i$, we have

$$F_l(\alpha) = \begin{cases} -\infty & \text{for } \alpha < \underline{H} \\ 1 & \text{for } \underline{H} \le \alpha \le \overline{H} \\ 1 + \overline{H} - \alpha & \text{for } \alpha > \overline{H}. \end{cases}$$

In Section 3. we argued that the spectra numerically measured on a path of an fBm with Hurst parameter H_k should be well approximated by (3.3). In any case, (3.5) provides an upper bound. This is indeed confirmed by our experiments (see Fig. 4.2). In particular, f_g is not concave and differs from f_l, comprising more information than the latter.

4.2 Cantor superpositions of fBm-s

With a whole range of observable Hölder exponents the previous simple model shows indeed a non trivial multifractal behavior, although different from the one observed on real traffic. Its peculiar shape stems from the fact that the different values of the self-similarity parameter H occur in entire intervals (which are of dimension 1). In order to obtain the approximate $\cap$ shape corresponding to actual traces, one needs to build a superposition of two fBm processes with exponents $H_1 < H_2$ such that the 'size' of the support of X_{H_1} is 'essentially smaller' than the one of X_{H_2}.

A somewhat artificial way to achieve this is to set:

$$X(t) = Z_{H_1}(t) + X_{H_2}(t)$$

where $X_{H_2}(t)$ is an fBm of exponent H_2 and $Z_{H_1}(t)$ is a 'Cantor Brownian motion' obtained as follows: let $X_{H_1}(t)$ be an fBm of exponent $H_1 < H_2$ and independent of $X_{H_2}(t)$. Let C be a Cantor-like set of dimension $D_1 < 1$ constructed with dyadic intervals. Set $Z_{H_1}(t) = X_{H_1}(t)$ for $t \in C$ and interpolate Z_{H_1} linearly elsewhere. Since C is a closed set without isolated points, Z_{H_1} is well defined and continuous.

The computation of the deterministic Legendre spectrum follows the lines of Subsections 3.1 and 4.1, observing that $\overline{S}_n$ can again be split into two parts, one coming from dyadic intervals used in the construction of C and one from the remaining intervals, i.e. those contained in a gap (the components of the closure of the complement of C). Bearing this in mind, a lengthy but straightforward calculation gives [25]

$$F_l(\alpha) = \begin{cases} -\infty & \text{for } \alpha < H_1 \\ (1-D_1)/(H_2-H_1)(\alpha - H_1) + D_1 & \text{for } H_1 \le \alpha \le H_2 \\ 1 + H_2 - \alpha & \text{for } \alpha > H_2. \end{cases} \tag{4.1}$$

Adding more than two fBm with well chosen pairs (H_i, D_i) would allow to design spectra that would more or less look like the ones obtained on real traffic. A loose "physical" interpretation of this model could be that different processes corresponding to various kinds of applications and/or traffic types and/or users (machines) coexist on the network. Some, more regular ones, being active most of the time (the ones with the largest values of H and D), and the others appearing intermittently and dominating the traffic during these periods (low values of H and D). Such an interpretation is also supported by other types of analysis [5]. Of course, this model is very unstable because any shift in the origin of the analysis would affect the obtained spectrum. Although it would be possible to refine the construction, we prefer to turn to another generalization which possesses nicer analytical properties.

4.3 Multifractional Brownian motion

While the above extensions remain quite simple, they are not fully satisfactory from a theoretical point of view. A step towards a better model is *multifractional Brownian motion* (mBm), a process introduced in [18]. Roughly, mBm is an fBm where the parameter H is a function of time. The precise definition is as follows:

Let $H : (0, \infty) \to (0, 1)$ be a Hölder function with exponent $\beta > 0$. We will write H_t instead of $H(t)$ for the ease of notation. Multifractional Brownian motion (mBm) is the Gaussian process $Y_t = X_{H_t}(t)$ defined on $(0, \infty)$ by:

$$Y_t = \frac{1}{\Gamma(H_t + \frac{1}{2})} \int_{-\infty}^{t} [(t-s)_+^{H_t - 1/2} - (-s)_+^{H_t - 1/2}] dW(s),$$

where $(x)_+$ equals x if positive and 0 otherwise.

The main properties of mBm are the following:

1. $(Y_t)_{t>0}$ is an a. s. continuous process.
2. Under the technical assumption $H_t < \beta$ for all t, the graph of mBm has the 'intuitively correct dimension', i.e. with probability one,
$$\dim_H\{(t, Y_t) : t \in [a, b]\} = 2 - \min\{H_t, t \in [a, b]\}.$$
3. With probability one, the pointwise Hölder exponent of Y_t at t is H_t for all t provided $H_t < \beta$ for all t.

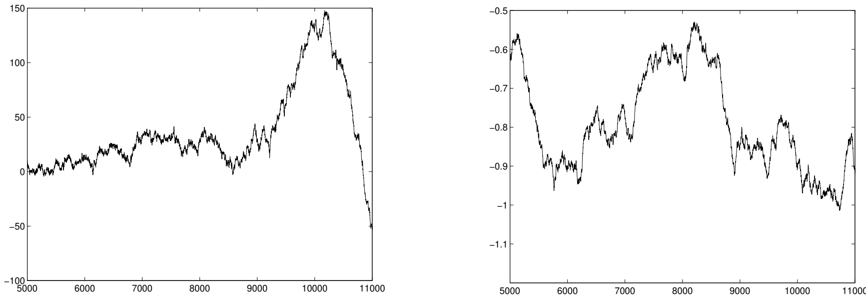

Fig. 4.1. Normalized mBm with $H(t) = t$ for $t \in [0.4, 0.8]$ (left) and fractional Brownian motion with $H = 0.6$ (right).

Multifractional Brownian motion allows, thus, to take into account a possible non stationarity of the exponent H as long as its evolution is smooth ($H(t)$ is a Hölder function). Let us now move to the computation of the deterministic Legendre spectrum of mBm. We present a heuristic computation to make the basic idea clear. A rigorous argument is contained in [25].

Recall the definition (3.2) of $\overline{S}_n$ in the random case. For n large, using property 3 above, we see that $Y_n^k = X((k+1)2^{-n}) - X(k2^{-n})$ is approximately equal to an $N(0, 2^{-nH(k2^{-n})})$ random variable. As earlier

$$\mathbb{E}(Y_n^k)^q \approx \begin{cases} \infty & \text{for } q \leq -1 \\ c_q 2^{-nqH(k2^{-n})} & \text{for } -1 < q. \end{cases}$$

In the special case $H(t) = t$ for $t \in (0,1)$ we obtain

$$\overline{S}_n(q) \simeq c_q \frac{1 - 2^{-nq}}{1 - 2^{-nq2^{-n}}} = c_q \frac{1 - 2^{-nq}}{nq2^{-n}\xi_n} \qquad (q \neq 0),$$

where $\xi_n \in [1/2, 2]$ for large n. With $\overline{S}_n(0) = c_q 2^n$ this leads to

$$T(q) = \begin{cases} -\infty & \text{for } q \leq -1 \\ q - 1 & \text{for } -1 < q < 0 \\ -1 & \text{for } 0 \leq q. \end{cases}$$

In general, for any non constant C^1 function $H\colon \mathbb{R}^{+*} \to [a,b] \subset]0,1[$ one

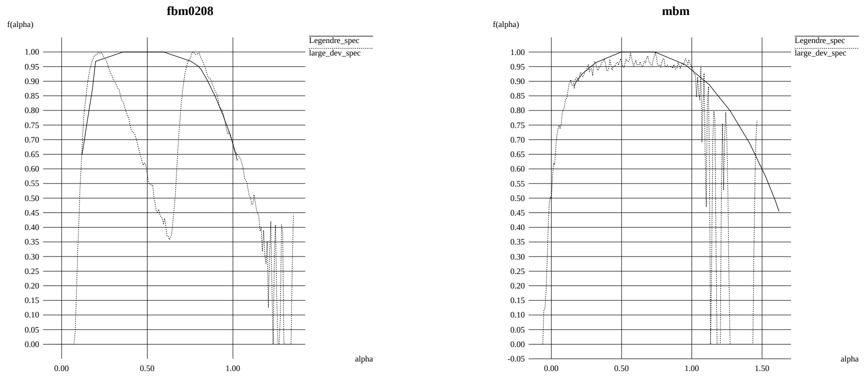

Fig. 4.2. Estimated large deviation spectrum and Legendre spectrum for the lumping of fBm $H = 0.2$ and $H = 0.8$ (left) as well as for the same mBm as before (right).

gets:

$$F_l(\alpha) = \begin{cases} -\infty & \text{for } \alpha < a \\ 1 & \text{for } a \leq \alpha \leq b \\ 1 + b - \alpha & \text{for } b < \alpha. \end{cases}$$

One finds, thus, the same spectrum F_l as the one of the lumping of two fBm with exponents corresponding to the minimum and maximum of the function $H(t)$.

Instead of $H(t) = t$, let us now consider a Hölder continuous function H which assumes the value H_1 on a Cantor set C as before and the value

$H_2 > H_1$ in points of distance larger than 2^{-N} from C. Any numerical computation of the spectra restricted to resolutions coarser than 2^{-N} would yield the same spectrum as in (4.1).

4.4 Experimental results

Fig. 4.1 displays a sample path of a normalized mBm (for all t, $X(t)$ is divided by the instantaneous variance) for which $H(t) = t$ for $t \in [0.4, 0.8]$, and, for comparison, a trace of an fBm with $H = 0.6$. Fig. 4.2 displays the estimated spectra on a lumping of fBm-s and on an mBm. As in Subsection 3.3, a good match with the theoretical expression for F_l is found. As for f_g, the estimated result support the heuristic argument presented in Subsection 4.1. Comparing with Fig. 3.3, one sees that these spectra indeed bear some analogies with the ones obtained on real traces (approximate $\cap$ shape, asymmetry), and could explain the findings reported in [20] and [14].

Finally, Fig. 4.3 displays estimates of the parameter H, based on the simple relation: $H = 1 + \tau(1, \omega)$ a.s. (3.6). For 'pure' fBm-s, the obtained value is quite good. For a lumping, the lowest H is estimated as predicted by theory. Lastly, no meaningful value is obtained in the case of mBm, which is again consistent with the non stationarity of H in this situation.

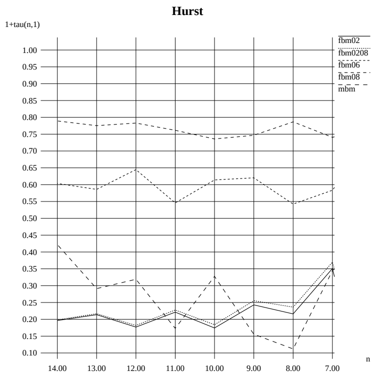

Fig. 4.3. Estimation of the Hurst parameter H through $1+\tau(1)$ for fractional Brownian motion with $H = 0.2$, 0.6 and 0.8 as well as for mBm and a lumping of fBm-s. Displayed is an estimation of $\tau(1)$ by $\tau(n, 1) := \log(S_n(1) - \log(S_{n+1}(1))$, i.e. as the local slope of the log-log plot of $S_n(1)$, where n is as in (2.2). The scaling was excellent at high resolutions, i.e. $n = 14, \ldots, 7$.

Conclusion

LRD analysis and associated fBm models allow to describe "robust" features of the traffic, i.e. features that are not sensitive to local variations in time or space. Multifractal analysis, on the other hand, is well fitted to the investigation of the fine structure of the traffic, that reflects some particular aspects of the network at a given location and hour of the day. Such a description is important because it allows to obtain more precise models and to characterize different types of traffic. A fine scale analysis is generally difficult for various reasons, and it seems that multifractal analysis may be a good tool for this task.

In as much as one is interested in the high frequency content of the data, we have shown that fBm is not a good model for computer traffic: The multifractal spectra of fBm, which we computed theoretically and estimated on synthetic traces, do not coincide with the ones of real traffic data. We have proposed to refine this model by taking into account a possible non stationarity in the singularity exponent.

Two situations have been studied. In the first one, several consecutive regimes are considered by lumping together plain fBm processes. The second model allows for a continuous variation of H. An adapted choice of the various parameters associated to these models allows to obtain spectra that look like real ones, at least when estimated on a finite number of resolutions. This result along with the fact that some 'physical' mechanism can be invoked which lead to such processes makes them possible candidates for modeling. A further argument is that a non stationary model is probably better suited to the high intra-day variability of the data.

One important discrepancy with real traces remains, namely that a non multifractal behavior is predicted for $q \leq -1$ in both models, the lumping of fBm-s and the mBm, while traces of real data traffic exhibit good scaling even for q lower than -4 [20]. In this respect, a radically different modeling based on multinomial cascades, which still lacks precise physical interpretation, would be more adapted. We should add, however, that such multinomial processes exhibit an everywhere discontinuous H function.

Despite of its drawbacks, it would be worthwhile to pursue the mBm type modeling and to study, for instance, the associated queueing behavior.

Acknowledgement. The authors gratefully acknowledge the financial support of CNET, France, under contract # 95 8B 069 and express their thanks to Fabrice Clerot, Ilkka Norros, Murad Taqqu and Marie-Claude Viano for helpful comments.

References

1. P. Abry and D. Veitch. Wavelet analysis of lange range dependent traffic. *preprint*, 1996.
2. R. Adler. *The Geometry of Random Fields.* John Wiley & Sons, New York, 1981.
3. M. Arbeiter and N. Patzschke. Self-similar random multifractals. *Math. Nachr.*, 181:5–42, 1996.
4. G. Brown, G. Michon, and J. Peyriere. On the multifractal analysis of measures. *J. Stat. Phys.*, 66:775–790, 1992.
5. F. Clerot. private communication, 1997.

6. M. Crovella and A. Bestavros. Self-similarity in world wide web traffic. evidence and possible causes. In *Proceedings of SIGMETRICS '96*, May 1996.
7. Richard Ellis. Large deviations for a general class of random vectors. *Ann. Prob.*, 12:1–12, 1984.
8. K. Itô. On the ergodicity of a certain stationary process. *Proc. Imp. Acad. Tokyo*, 20:54–55, 1944.
9. S. Jaffard. Multifractal formalism for functions. *CRAS*, 317:745–750, 1993.
10. U. Krengel. *Ergodic Theorems.* Walter de Gruyter, Berlin, 1985.
11. W. Leland, M. Taqqu, W. Willinger, and D. Wilson. On the self-similar nature of ethernet traffic (extended version). *IEEE/ACM Transactions on Networking*, pages 1–15, 1994.
12. B. B. Mandelbrot. Intermittent turbulence in self similar cascades : divergence of high moments and dimension of the carrier. *J. Fluid. Mech.*, 62:331, 1974.
13. B. B Mandelbrot and R. H. Riedi. Inverse measures, the inversion formula and discontinuous multifractals. *Adv. Appl. Math.*, 18:50–58, 1997.
14. P. Mannersalo and I. Norros. Multifractal analysis of real ATM traffic: a first look. *COST257TD*, 1997.
15. I. Norros. A storage model with self-similar input. *Queueing Systems*, 16:387–396, 1994.
16. I. Norros. private communication, 1997.
17. V. Paxson. Emperically-derived analytic models of wide-area TCP connections. *IEEE/ACM Transactions on Networking*, 2(4):316–326, August 1994.
18. R. Peltier and J. Lévy Véhel. Multifractional Brownian motion: Definition and preliminary results. *INRIA Research Report, No 2645*, 1995.
19. R. H. Riedi. An improved multifractal formalism and self-similar measures. *J. Math. Anal. Appl.*, 189:462–490, 1995.
20. R. Riedi and J. Lévy Véhel. TCP traffic is multifractal: a numerical study. *submitted IEEE/ACM Transactions on Networking*, 1997. (also INRIA Research Report No. 3129,1997 http://www-syntim.inria.fr/fractales/).
21. Riedi, R.H. (1997): Seven definitions of multifractal scaling exponents. *In preparation.*
22. J. Roberts, U. Mocci, and J. Virtamo (eds.). Braodband network teletraffic. In *Lecture Notes in Computer Science, No 1155.* Springer, 1996.
23. G. Samorodnitsky and M. Taqqu. *Stable non-Gaussian random processes.* Chapman and Hall, New York ISBN 0-412-05171-0, 1994.
24. J. Lévy Véhel. Fractal approaches in signal processing. *Fractal Geometry and Analysis, The Mandelbrot Festschrift, Curacao 1995, C.J.G. Evertsz, H.-O. Peitgen, R.F. Voss (eds.), World Scientific*, 1996.
25. J. Lévy Véhel and R. Riedi. Fractional Brownian motion and data traffic modeling: The other end of the spectrum. *Technical Report No .., INRIA Rocquencourt, France*, 1997. http://www-syntim.inria.fr/fractales/ Extended version to be published as 'Large deviation spectra of some multifractal processes'.
26. J. Lévy Véhel and R. Vojak. Multifractal analysis of Choquet capacities: Preliminary results. *Adv. Appl. Math.*, to appear 1997.
27. W. Willinger, M. Taqqu, R. Sherman, and D. Wilson. Self-similarity through high-variability: Statistical analysis of ethernet lan traffic at the source level. *IEEE/ACM Transactions on Networking (Extended Version)*, 1995 (to appear).