Non-stationarity and high-order scaling in TCP flow arrivals: a methodological analysis

Steve Uhlig
Computer Science and Engineering Department
Université catholique de Louvain, Belgium
URL: http://www.info.ucl.ac.be/∼suh/
suh@info.ucl.ac.be

ABSTRACT

The last decade has been a very fruitful period in important discoveries in network traffic modeling, uncovering various scaling behaviors. Self-similarity, long-range dependence, multifractal behavior and finally cascades have been studied and convincingly matched to real traffic. The first purpose of this paper is to provide a methodology to go beyond the naive analysis of the second-order wavelet-based estimators of scaling, by performing non-stationarity checks and relying on the information contained in the high-order properties of the wavelet coefficients. Then, we apply this methodology to study the scaling properties of the TCP flow arrivals based on several traffic traces spanning the years from 1993 to early 2002. Our study reveals that the second-order scaling properties of this process describe its dynamics quite well. However, our analysis also provides evidence that high-order scaling in this process appears due to pathological behaviors like rate limitation and non-stationarity.

Categories and Subject Descriptors

C.2.5 [**Computer-communication Networks**]: Local and Wide-Area Networks—*Internet traffic*; G.3 [**Probability and Statistics**]: [Time-series analysis]

Keywords

TCP flow arrivals, non-stationarity, scaling processes, long-range dependence, multiscaling.

1. INTRODUCTION

The last decade has been a very fruitful period in network traffic modeling, uncovering various scaling behaviors. Self-similarity [21], long-range dependence [37], multiscaling and multifractal behavior [29, 12, 14] and finally cascades [12, 15, 36, 14, 34] have been studied and convincingly matched to real traffic. The introduction of these models into the networking world have often brought significant insight about the dynamics of the traffic, but also much misunderstanding concerning their right place within the dynamics of the traffic, their interpretation and practical implications. While all building blocks in terms of the scaling models seem to have been brought to the networking world, there still is a lack of proper understanding of the intrinsic reasons for such a good fit of the traffic to these models and the right place of each model. Two breakthroughs in this respect need to be mentioned though: the ON/OFF sources superposition model [32] that justifies the presence of self-similarity in the traffic, and the multifractal/cascade model [12] that emphasizes the importance of TCP in the traffic dynamics. These two results illustrate quite well that scaling models constitute a natural framework to better capture the invariant properties of network traffic dynamics. In this paper, we focus on the process of the TCP flow arrivals per time interval and analyze its high-order scaling properties as well as its non-stationarity.

This paper is structured as follows. We first discuss the related work in section 2, and provide some background on scaling processes and wavelet analysis. In section 3.1, we describe the traffic traces on which we rely in our analysis of the TCP flow arrivals. We first look at its second-order scaling properties in section 3.2. Then, we tackle its high-order scaling properties in section 3.3. We compare the scaling properties of HTTP and P2P traffic in section 3.4 and study the behavior of internal clients of one of the traces. Finally, section 4 discusses the implications of this paper on the scaling properties of Internet traffic identified in the literature.

2. CONTEXT

This section presents an overview of the current state-of-the-art of scaling in network traffic. Three main roads have been explored concerning the explanation of the presence of scaling in network traffic.

The first concerns distributional properties of the flow activity periods that were shown to be heavy-tailed [28, 7]. The complementary proof in [32] then provided a formal justification for the presence of self-similarity through the superposition of a large number of independent ON/OFF sources with heavy-tailed ON and/or OFF periods. [24] made the connection between distributional properties of the file sizes and the modulating effect of the TCP/IP stack and showed that heavy-tails in the applicative flows were mapped to heavy-tailed activity periods at the IP layer.

The second road was investigated by [12] that studied the short timescales behavior of TCP traffic to show that multiplicative processes described it well, adding multifractals and cascades among the scaling behaviors present in network traffic.

These two roads, heavy-tails and the TCP dynamics, concern traffic volume. The first is related to the size of the files to be carried over the network while the second concerns the way the data of the flows is broken by TCP to be sent as IP packets. The previous two aspects tell that the information to be exchanged as well as the behavior of the network and its protocol can be considered as partly responsible for scaling in the traffic. However, one may wonder whether users of the network have anything to do with scaling in the traffic, allowing us to introduce the third road.

Back in 1995, [28] argued that the connection arrival process of some applications failed to be Poisson. This finding was a surprise due to the belief that users initiating actions independently should lead to a Poisson process. [5] however infirmed the findings of [28] by showing that connection arrivals tend to a non-stationary Poisson process as the rate of the observed link increases. The first study of the scaling properties of TCP flow arrivals was [13] that explained self-similarity in this process by heavy-tails in the number of flows within user sessions. [17] however showed that scaling at the IP level did not depend on the process of the TCP flow arrivals. [10] showed that TCP flow arrivals contained second-order scaling. The aim of [10] however was not to study the timescale dependent properties of the process of the TCP flow arrivals but rather to find a good probability distribution to model the TCP flow arrivals.

To the best of our knowledge, the first paper completely devoted to the scaling properties of the TCP flow arrivals is [35]. [35] showed that this process is by no means simply self-similar, but contains non-stationarity and different scaling components at different timescales. [35] identified correlations between seconds and minutes, second-order scaling at timescales between minutes and hours, and finally a "time of the day" pattern at timescales larger than about half an hour. Because the analysis was limited to second-order scaling properties, [35] could not reveal the exact nature of the components at each timescale. One of the goals of the present paper is to continue [35] with a detailed study of the high-order properties of the process of the TCP flow arrivals, to better identify the scaling nature of this process.

Besides the scaling properties of the traffic, some studies have tried to better understand its non-stationarity. [5] showed that as the rate of the observed link increases, packet and connection arrival processes tend to a non-stationary Poisson process. [19] recently revisited the Poisson assumption in Internet traffic and concluded that current traffic traces are consistent with a non-stationary Poisson process at sub-second timescales while long-range dependence appears at larger timescales. Throughout this paper, we emphasize the importance of non-stationarity to understand the scaling properties of network traffic.

2.1 Scaling Basics

Scaling behavior is a general term meaning that no particular scale controls the process under study. For an introduction to scaling in network traffic see [2]; also see [26] that provides a detailed review of various aspects concerning scaling in network traffic. For a more general discussion of scaling in natural phenomena see [31].

Self-similarity, or more precisely self-affinity, with parameter H ($0 < H < 1$), is defined by

$$\{X(t), t \in \mathbb{R}\} \stackrel{d}{=} \{c^H X(t/c), t \in \mathbb{R}\}, \forall c > 0, \qquad (1)$$

where $\stackrel{d}{=}$ denotes equivalence in finite size distributions. A property of self-similar processes concerns the scaling of their moments:

$$E|X(t)|^q = E|X(1)|^q \ |t|^{qH} \ . \qquad (2)$$

This latter property relates the behavior of moment q (if it exists) as a power law of time, defining high-order moments scaling which is linear in H. A self-similar process will exhibit the same fluctuations independent of the considered scale, hence cannot be stationary due to the dependence of all its moments on time. Recall that for a process to be stationary, its statistical properties must not depend on time.

When one restricts the study of the process to its second-order properties, then comes long-range dependence. Long-range dependence (LRD), or long-memory, is associated with second-order stationary processes. A second-order stationary process displays long-range dependence [4] if its autocovariance function $r(k)$ behaves like a power-law of the time lag k

$$r(k) \sim c_r |k|^{2H-2} \ \ as \ \ |k| \to \infty \qquad (3)$$

for $1/2 < H < 1$. In such a case, the correlations decay so slowly that they sum to infinity. When $H = 1/2$, the process is uncorrelated ($r(k) \sim 0$). When $0 < H < 1/2$, we speak of short-range dependence (SRD), or anti-persistence, due to negative dependence ($r(k) < 0$ for $k \neq 0$) between time lags. The divergence of the correlations when $1/2 < H < 1$ also means that the spectral density behaves like

$$f(\lambda) \sim c_f |\lambda|^{1-2H} \ \ as \ \ |\lambda| \to 0, \qquad (4)$$

hence diverges at the origin ($f(0) = \infty$). But because of stationarity, LRD processes cannot exhibit fluctuations at any arbitrary scale, as do self-similar processes.

An often mistaken relationship between LRD and self-similarity concerns the fact that self-similarity for $H > 0.5$ implies LRD. There are thus two types of long-range dependent behavior: one that is limited to the definitions 3 and 4 for a second-order stationary process but where self-similarity does not hold, and the long-range dependence that is a consequence of self-similarity. This distinction has often been neglected in the literature but has important consequences in practice on the type of process considered. For example, it is not always possible to empirically distinguish between correlations that span more than the length of the studied sample, pure LRD and LRD produced by a self-similar process.

2.2 A Wavelet Primer

This section provides a very short introduction to wavelets. For a complete treatment of wavelet theory and analysis, we refer to [9, 3, 2].

Discrete wavelet signal decomposition consists in analyzing a signal $X(t)$ through a bandpass oscillating function $\psi_{j,k}$ where j represents the timescale and k the time instant. Throughout this paper, small values of j represent the smallest timescales and large values represent the coarsest timescales. By scaling and shifting this function ψ, it is possible to break the signal into its timescale components (at timescale j) and within each timescale along the time axis (at time k):

$$\psi_{j,k}(t) = 2^{-j/2} \psi(2^{-j} t - k) \qquad (5)$$

where the $\psi_{j,k}(t)$ form an orthonormal basis of square integrable functions. Hence any square integrable signal can be approximated by a finite linear combination of the $\psi_{j,k}(t)$.

A scaling and shifting operation by a factor (j,k) moves the original central frequency f_0 of the function ψ to frequency $2^{-j}f_0$ while shifts it in time by a factor of $2^j k$. The discrete wavelet transform algorithm performs a dyadic tree decomposition (multiresolution analysis) of the signal

$$X(t) = \sum_k c_X(j_0,k)\phi_{j_0,k} + \sum_{j\leq j_0}^{J}\sum_k d_X(j,k)\psi_{j,k}(t) \quad (6)$$

where the first term represents an approximation of the signal while the second term represents the details, j_0 being the resolution depth, i.e. the coarsest resolution at which the signal is analyzed. The $c_X(j_0,k)$ are called the scaling coefficients and ϕ the scaling function. The $d_X(j,k)$ are called the wavelet coefficients and are of special interest because they are used as a substitute for the increments of the process over the dyadic tree of the time-frequency plane. The orthogonal basis property of ϕ and ψ allows the easy computation of the coefficients by simple inner products with the signal. The nice properties of the wavelet transform make the wavelet coefficients better suited for statistical analysis in comparison to the original increments of the process. In addition, the wavelet transform has a low time complexity in $O(n)$, making it efficient. From now on, we shall exclusively deal with the $d(j,k)$ instead of the process increments $X_\delta(k)$.

Several properties of wavelets make them particularly suited at analyzing scaling processes:

1. scale invariance: the wavelet basis being built through the dilation operator, the analyzing functions family has a built-in scaling property that reproduces scaling present in the data.

2. robustness of the mother wavelet against non-stationarity: the mother wavelet (ψ_0) has a number $N \geq 1$ of vanishing moments $\int t^k\psi_0(t)dt \equiv 0$, $k = 0, ..., N-1$, allowing the wavelet to remove any polynomial trend of order up to $N-1$.

3. almost decorrelation of the wavelet coefficients: under the assumption of $N \geq H + 1/2$, global LRD among the increments of the process can be turned into short-range dependence among the wavelet coefficients [33].

These properties ensure that the wavelet transform captures the scaling properties of the signal (point 1), that bias from LRD in the increments will not bias a statistical analysis of the wavelet coefficients (point 3), and that influence from polynomial trends up to order $N-1$ will be removed by the wavelet (point 2).

3. TCP FLOW ARRIVALS

This section studies the process of the TCP flow arrivals, defined as the number of new TCP flows seen for every time interval. Section 3.1 first describes the traffic traces on which we rely for our analysis. Section 3.2 studies the second-order properties of the TCP flow arrivals, globally over the traces and also in a time-dependent way. Section 3.3 deals with the higher-order properties of the TCP flow arrivals. Finally, section 3.4 studies the behavior of one particular trace to better understand the effect of the application mix on the high-order properties of the TCP flow arrivals. In both sections 3.2 and 3.3, we first present the estimators on which we rely before carrying out the corresponding traffic traces analysis.

Note that because of the different time granularities of the traces, the same octaves on different graphs may represent different timescales. The Auckland trace and the PSC trace use a 10 ms precision, hence octave 1 corresponds to 20 ms. For all other traces, we use a 100 ms time granularity so octave 1 represents 200 ms. Each subsequent octave reduces the amount of data by a factor of two, hence corresponds to a timescale 2 times larger than the previous octave.

3.1 Traffic Traces

3.1.1 LBL-CONN-7

The "lbl-conn-7" trace, available on the Internet Traffic Archive [1] contains thirty days' worth of all TCP connections (606,497 connections) between the Lawrence Berkeley Laboratory (LBL) and the rest of the world. It was collected with *tcpdump* on a Sun Sparcstation using the BPF kernel packet filter between September 16 1993 through October 15 1993. The time precision of the trace is 1 μs but we used in this paper a 100 ms precision aggregation, by not having taken into account in the timestamps the digits corresponding to timescales below 100 ms. Only the TCP connections that ended normally (with both SYN and FIN) were used. The most important applications were e-mail (SMTP) with 34 % of the flows and 3 % of the bytes, FTP with 18 % of the flows and 51 % of the bytes, news (NNTP) with 11 % of the flows and 25 % of the bytes, and finally telnet with 15 % of the flows and 2 % of the bytes. We call this trace the "LBL" trace in the remainder of this paper. For further description of this trace, we refer to [27, 28, 1].

3.1.2 UC Berkeley Home IP Web

The "UC Berkeley Home IP Web" trace is also available from the Internet Traffic Archive [1]. This dataset consists of 18 days' worth of HTTP traffic gathered from the Home IP service offered by UC Berkeley to its students, faculty, and staff. All clients were using modems. This client trace was collected through the use of a packet sniffing machine placed at the head-end of the Home IP modem bank. The tracing program used was a custom module written on top of the Internet Protocol Scanning Engine (IPSE) created by Ian Goldberg. Only traffic destined for port 80 was traced, all non-HTTP protocols and HTTP connections for other ports were excluded from these traces. No TCP flag filtering was performed on this trace, nor were all requests normally terminated (with FIN flag set), so only the fact that the SYN flag was set for the corresponding TCP connections is certain. It contains 9,244,728 flows spanning the period from November 1 1996 through November 19 1996. Although a better precision was available, we used a 100 ms time granularity for the timestamps. Further information about this trace can be found on the ITA website [16]. We refer to this trace as the "UCB" trace in the remainder of this paper.

3.1.3 Auckland IV

The Auckland-IV data set [23] is a continuous 45 days GPS synchronized IP header trace captured with a DAG3 system [22] at the University of Auckland Internet uplink by the WAND research group between February and April 2001. The University of Auckland ITSS department is operating an OC3 ATM link via Clear Communications, which is carrying a variety of services off the main campus. This is a continuous 6 1/2 week GPS-synchronized IP header trace taken with a pair of DAG3 cards at the University of Auckland Internet access link. The timekeeping precision of the DAG timestamping engine, called the DUCK, has been monitored during the entire duration of the trace and was guaranteed to be less than 1 μs to UTC at all times. The tap was installed at an OC3 link

carrying a number of Classical-IP-over-ATM, LANE and POTS services. The trace contains all Classical-IP headers of a pair of redundant VPI/VCI's, which connects the university to their local service provider. Incoming and outgoing directions were captured as different traces. We selected a 7-days sample representative of the whole for the analysis of this paper. There were 11,919,853 incoming flows (12,335,605 outgoing) during the selected 7-days period. In this trace, we solely relied on packets that had the SYN bit set and used a 10 ms precision due to trace size limitations. Most of the bytes and flows of this trace were due to HTTP traffic (see [23] for more details). We refer to this trace as the "Auckland" trace in the remainder of this paper.

3.1.4 Yucom

The "Yucom" trace consists of all flows between modem clients of a small Belgian ISP, Yucom. Yucom is a commercial ISP that provides Internet access to dialup users through regular modem pools. The trace spans less than 5 days between April 17 and April 21 2001. It is a Netflow trace [6], not a packet one. The collecting machine was located within the premises of the ISP and was running *OSU-flow-tools* [30]. Due to the 30 seconds flow timeout policy used by Netflow, we did not rely on the raw flow information as that over-estimates the number of flows. Instead, we considered only the incoming TCP flows that had the SYN flag set. This means that even if we do not have the correct TCP flows information because of the Netflow timeout, we have the same information concerning the new TCP flows (SYN). During the whole period of the trace, there were 59,581,814 incoming TCP flows that had the SYN flag set. Among the 574 GBytes of the incoming TCP flows in the trace, HTTP and HTTPS represent 72 % of the bytes and 77 % of the flows, chat 16.2 % of the bytes and 7.7 % of the flows, and e-mail 5 % of the bytes and 6 % of the flows. We call this trace the "Yucom" trace in the remainder of this paper. We used a 100 ms time granularity for this trace. Note that we only study the incoming flows of this trace in this paper.

3.1.5 PSC

This trace consists of all outgoing flows seen on two monitoring points internal to the Pittsburgh Supercomputing Center, an American commodity and Internet2 network provider to organizations in western Pennsylvania. The flows were collected continuously over 24 hours starting from March 12 2002 16:40 GMT at both monitoring points [35]. These points covered all outbound commodity traffic from non-dormitory hosts at a large university, in addition to all outbound commodity traffic at several smaller organizations. The collection was made using the CoralReef package from CAIDA [20] using typical Netflow settings for flow expiration. During the flow period the network served by one monitoring point experienced considerable congestion, effectively imposing rate limits during peak usage times. We consider the TCP traffic only. Because we had no information about the TCP flags, all terminated flows (FIN or 30 seconds CoralReef timeout) were treated as completed flows. We used a 10 ms time granularity for this trace. There were 39,786,219 TCP flows in the trace, representing 574 GBytes of IP traffic. The important applications we identified were Gnutella with 22 % of the bytes and 12 % of the flows, Kazaa with 20 % of the bytes and 8 % of the flows, HTTP with 10 % of the bytes and 60 % of the flows. We call this trace the "PSC" trace in the remainder of this paper.

3.1.6 Summary

Table 1 summarizes the most important differences between the five traffic traces used in the remainder of this paper.

Table 1: Summary of the five traffic traces.

Trace	Precision	Type	TCP filter	Length
LBL	100 ms	Packet	SYN+FIN	30 d
UCB	100 ms	Packet	HTTP	18 d
Auckland	10 ms	Packet	SYN	7 d
Yucom	100 ms	Netflow	SYN	5 d
PSC	10 ms	CoralReef	All	1 d

3.2 Second-order properties

3.2.1 Second-order scaling

The most widely used wavelet-based estimator of scaling, the *log-scale diagram* (*LD*) [3] consists in plotting $y_j = log_2(\mu_j)$ against j (together with the confidence intervals), where

$$\mu_j = \frac{1}{n_j} \sum_{k=1}^{n_j} |d_X(j,k)|^2, \tag{7}$$

n_j is the number of wavelet coefficients at octave j and $d_X(j,.)$ denotes the wavelet coefficients (of the discrete forward wavelet transform of the time series) at octave j. The *LD* allows the detection of scaling through observation of strict alignment (linear trend) of the confidence intervals of the y_j within some octave range. If, and only if strict alignment is detected, can estimation of scaling parameters be performed. The *LD* is a second-order statistics (sample variance) of the wavelet coefficients, hence it does not capture the higher-order properties of the signal.

Because of the potential bias from non-stationarity, we also rely on a method introduced in [35] that consists in computing the logscale diagram over fixed-length time intervals. In this method, one plots on a three dimensional graph the evolution in time of the *LD*'s computed over fixed size time intervals, hence its name *3D-LD*. On the plot of the *3D-LD*, half of the points of each *3D-LD* overlap with those of the neighboring *LD*'s, to provide a smoother visual aspect of the *3D-LD*.

Correctly interpreting the *3D-LD* requires some care. The first property of the analyzed trace seen on the *3D-LD* is the evolution in time of the general level of each *LD*. Since the *LD* measures the variance of the wavelet coefficients of the signal at each timescale, the *LD* tells how the energy of the signal is split among the different timescales. Note however that since wavelet coefficients represent the increment of the process from which polynomial trends have been removed, the variances of the wavelet coefficients do not sum to the total energy of the original signal. Changes in the global level of subsequent *LD*'s suggest corresponding changes in the total energy in the signal's increments. A second-order stationary signal, whose mean and variance do not change, should thus have a *3D-LD* that does not change with time. In addition, changes in the *LD*'s with time indicate changes in the relative contribution to the energy of the process at each timescale.

The main information about the variability of the process provided by the *LD* concerns how the variability of the process is decomposed among the different timescales. The slope of the LD over some octave range gives qualitative information about the signal. A slope close to zero points to an uncorrelated process. A positive slope is an indication of potential scaling (LRD or self-similarity). A negative slope means that the process is short-range dependent or self-similar with $H < 1/2$. The interest of the *3D-LD* is to help visualize the changes of the *LD* in time. It has been shown in [35]

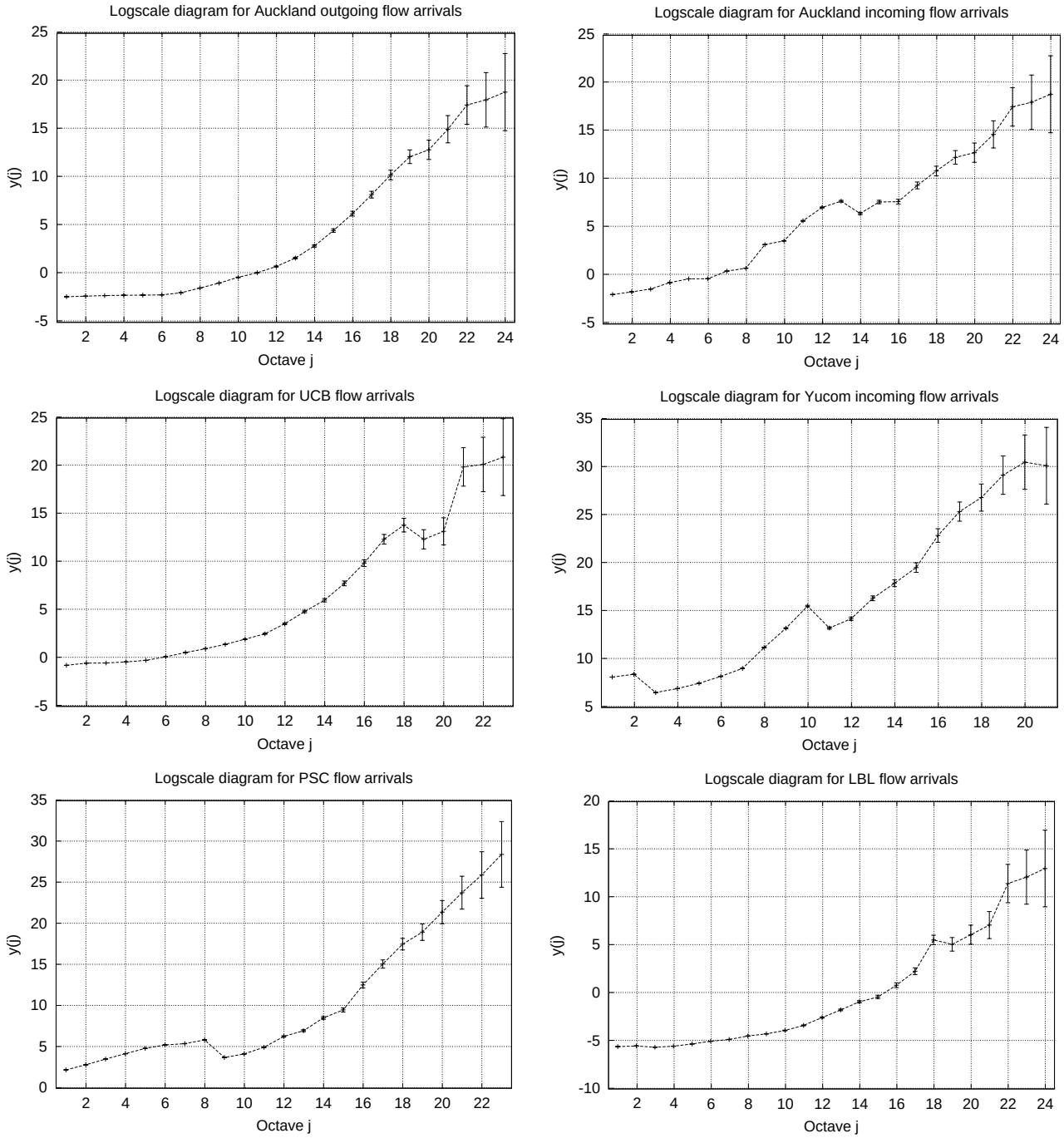

Figure 1: Logscale diagrams.

that whenever several components exist at different timescales and whose behavior change with time, the *LD* alone might be largely biased. If the non-stationarity in the process is important or the different components of the signal at the different timescales interact, then the *LD* might point to a process whose scaling properties are actually different from the real ones. The *3D-LD* is a valuable tool in such situations, to invalidate and eventually correct the information provided by the *LD* over the whole analyzed sequence. Nevertheless, the nature of scaling cannot be automatically inferred simply from the *3D-LD*, it is a graphical method that helps better understand the time-varying nature of the *LD*.

3.2.2 Second-order analysis

In this section, we study the second-order properties of the TCP flow arrivals for the traces presented in section 3.1. We study their *second-order* scaling properties over the total length of the traces, as well as their time-dependent *second-order* scaling properties.

The interest of the logscale diagram is in identifying the presence of strict scaling over a range of timescales in the process taken as a

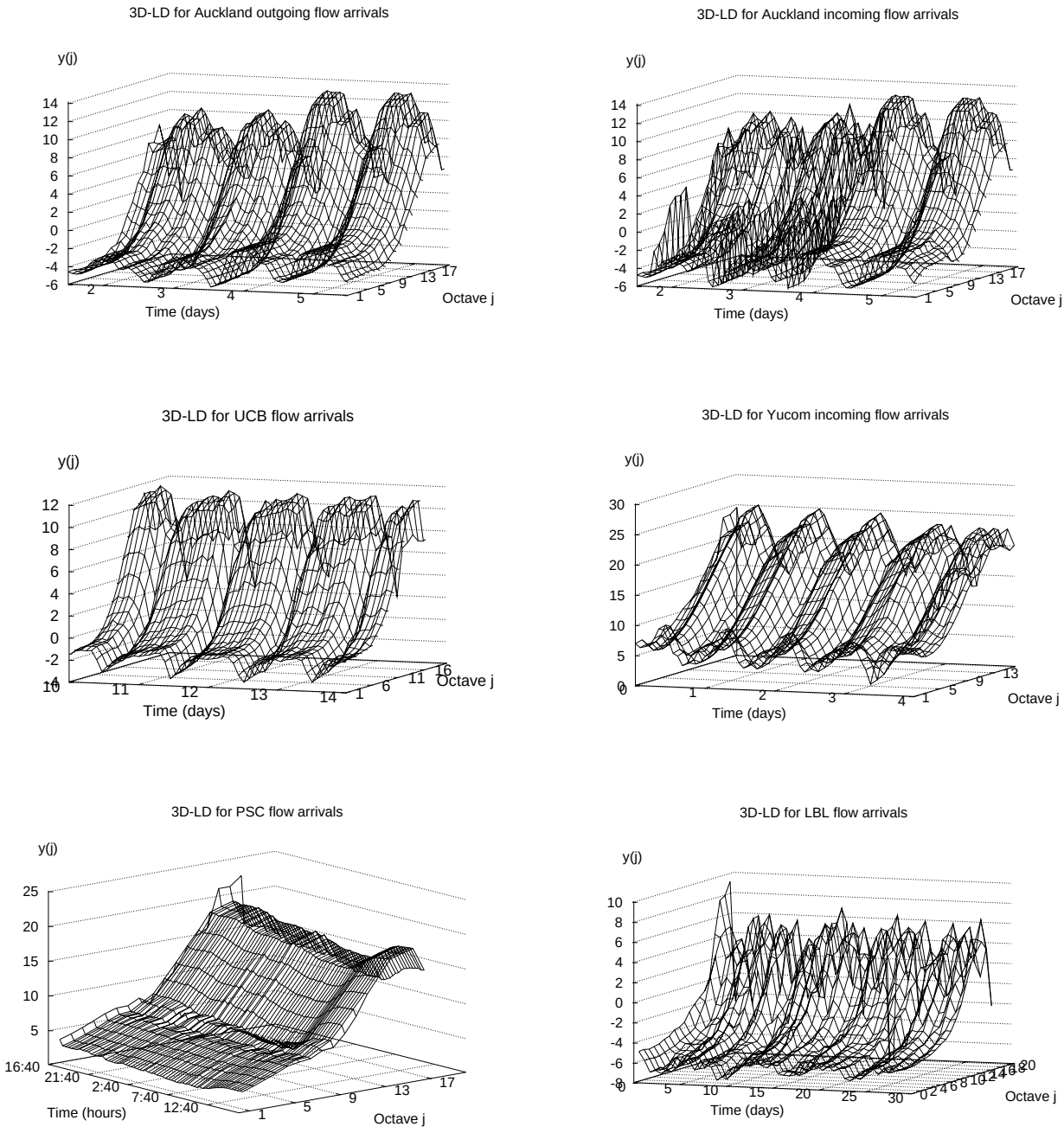

Figure 2: 3D-logscale diagrams.

whole. If a straight line of positive slope can be fitted to a portion of the logscale diagram, this indicates the presence of scaling (be it self-similarity, LRD or simply correlation).

Figure 1 presents the *LD* for the five traffic traces. As it appears on the graphs, there is no evidence of strict scaling except for very limited octave ranges like octaves [7,10] for Yucom. A property of the *LD* concerns the theoretical possibility to distinguish the type of process present within some octave range based on the value of the slope of the logscale diagram [3]. Based on these figures only, there are roughly four different regions on each of the graphs.

Timescales below seconds (octaves in [1,6] for Auckland and PSC, octaves in [1,3] for UCB, Yucom and LBL) show evidence of small correlations (or almost uncorrelated for Auckland outgoing, UCB and LBL). Based on the results of [5], we would expect a non-stationary Poisson process at these timescales, hence no correlation at all and a zero slope of the *LD*.

Timescales between seconds and minutes (octaves in [7,12] for

Auckland and PSC, octaves in [4,9] for UCB, Yucom and LBL), show evidence of dependence ($0 <$ slope < 1) for all traces. The *LD*'s of Auckland (incoming), PSC and Yucom seem to exhibit some peculiar behavior, with negative slopes for very limited segments of their *LD* at these timescales. This behavior will be explained later in this section.

Timescales between minutes and a few hours (octaves in [13,18] for Auckland and PSC, octaves in [10,15] for UCB, Yucom and LBL) show evidence of a self-similar process (slope > 1). The absence of strict alignment however rules out the possibility of strict self-similarity, but rather points to a non-stationary process.

Finally, the largest timescales exhibit very large confidence intervals such that inference is hopeless. These large timescales have however been studied in [35] that identified some periodic "time of the day" pattern for timescales larger than half an hour.

With the knowledge of the flow arrivals process whose second-order properties were shown to be highly non-stationary in [35], we know that the *LD* should not be trusted because it does not show how scaling evolves within the time span of the studied time-series. Henceforth, we now turn to the *3D-LD* of the TCP flow arrivals, which is still a second-order statistic of the wavelet coefficients but which makes explicit non-stationarity in the *LD*.

Figure 2 shows the evolution of the *LD* computed for overlapping blocks of a little less than 3 hours for Auckland, 110 minutes for UCB, 3.5 hours for Yucom, 45 minutes for PSC and 29 hours for LBL. The size of the blocks to compute the *LD* for each trace was chosen large enough to be able to show the largest octaves we were interested in. Note that we do not show the confidence intervals for the *3D-LD* only to prevent visual overloading of the graphs.

The first noticeable pattern visible on the *3D-LD*'s of Figure 2 is the "time of the day" pattern. Recall that the largest octave of the *LD* represents the energy contained in the signal that was not matched by the wavelet transform at the smaller timescales. Hence, non-stationarity in the mean will be visible through a time-varying value of the *LD* for the largest octaves. This non-stationarity already identified in [35] is confirmed here on all traces.

Second, all traces except Auckland incoming and PSC have a similar *3D-LD*. This gives evidence of no fundamental change in the scaling properties of the TCP flow arrivals between 1993 and 2002, nor any particular change of the scaling properties due to changes in the application mix of the traces. Even the PSC trace and Auckland incoming would have a *3D-LD* similar to those of the other traces, were they not biased by some unusual events as will be shown later on.

The other type of non-stationarity present in the traces appears for the incoming flows of Auckland (top right of Figure 2), where days two and three exhibit sharp peaks causing the larger energy of the *LD* for the flows over octaves [8,15] (top right of Figure 1). A single short-term event in the traffic might thus significantly bias the *3D-LD*. Finally, the PSC trace will be shown to be biased by rate limitation enforced for a particular client (see section 3.4). The effect of rate limitation on the *LD* of PSC is to produce a constant positive slope for octaves [1,8].

Although the second-order properties are not sufficient to infer the scaling properties of a process, the *3D-LD* already hints at the possible components of the process at different timescales. For instance, the *3D-LD*'s of Figure 2 strengthen our confidence in the absence of scaling at subsecond timescales, with a zero slope of the *LD* for almost all traces and all time periods. Furthermore, even if we found evidence of correlations between seconds and minutes and of self-similarity between minutes and hours with the *LD*, these timescales do not seem to contain two such clearly distinct components. The time-varying slope of the *LD*'s on Figure 2 for timescales between seconds and hours rather suggest that these timescales contain a mix of correlation, scaling and non-stationarity. The purpose of the next section is however to try to disambiguate the question of which type of process is present at each timescale.

3.3 High-order properties

3.3.1 High-order scaling

The wavelet-based partition function is defined as

$$S(q,j) = \sum_k |2^{-j/2}\ d_X(j,k)|^q \tag{8}$$

where $2^{-j/2}$ is a normalizing factor to remove the L_2 normalization of the wavelet. The partition function captures the local scaling behavior of a signal, in that raising the wavelet coefficients to an exponent magnifies the importance of the largest coefficients that arise due to a local irregularity, while it reduces the importance of small coefficients. Recall that a wavelet is an oscillating function, hence the value of the wavelet coefficient is proportional to the size of the irregularity that is matched by the inner product between the process and the wavelet. The smaller the wavelet coefficients for some scale j, the smaller the value of the partition function for $S(q,j)$ for large q. This permits to study the importance of the local irregularities at some timescale j. A relatively smooth process that has no particularly large local irregularity will have values of the $S(q,j)$ for $q > 2$ larger than $S(2,j)$. We call such a process a *second-order* process in the remainder of this paper. An irregular process on the other hand will tend to have larger values of the $S(q,j)$ for large values of q. We call such a process a *high-order* process in the remainder of this paper. Note that we only study the positive moments because we are only interested in the irregularities of a process.

Multiscaling is a relaxation of strict scaling where the linearity relation among moments does not hold anymore but can vary for each moment q [31, 36]:

- self-similarity : $E|d(j,k)|^q \propto exp(qH\ \ln(2^j))$,
- multiscaling : $E|d(j,k)|^q \propto exp(H(q)\ \ln(2^j))$.

Strict self-similarity implies that all moments behave like a power-law of the timescale, where the power-law is linear with respect to H. Multiscaling on the other hand allows the power-law relationship among moments to depend on the moment q. Self-similarity imposes that a single parameter H controls the behavior of the whole dynamics of the process. Multiscaling on the other hand implies a non-linear behavior of the moments with respect to one another. An example of multiscaling is multifractality [18, 29] where there exists a whole spectrum of local scaling exponents related to the local (ir)regularity of the sample path of the process. Let $\alpha(t)$ denote the local scaling exponent of the process at time t, then a self-similar process has $\alpha(t) = H$ while a multifractal process has a non-constant scaling exponent $\alpha(t) = H(t)$. This exponent

$\alpha(t)$ is related to the local geometrical properties of the function, where $\alpha(t) > 1$ is related to small local variations while $\alpha(t) < 1$ to large local irregularities. Multifractality occurs when $H(q)$ depends non-linearly on q but linearly with respect to j: each moment $S(q, j)$ then tends to diverge relative to one another for the smallest timescales.

Besides power-laws of the moments **in time**, it is also possible to study the power-law relationship among moments (q) at a given timescale j. The partition function is a means to determine the "type of process" that arises at some timescale through the multiscaling paradigm. Multiscaling allows for three possible behaviors related to the linearity of high-order moments relative to low-order ones :

1. sub-linearity of a LRD, correlated or non-scaling process implies that $M(j) < 1$,
2. linearity of a self-similar process implies that $M(j) \sim 1$,
3. super-linearity of a multiscaling process implies that $M(j) > 1$,

with

$$M(j) = \left\langle \left| \frac{log(S(q+1,j)) - log(S(q,j))}{log(S(2,j)) - log(S(1,j))} \right| \right\rangle_q, \qquad (9)$$

where $\langle\rangle_q$ represents the averaging over the values of q. It must be born in mind that the multiscaling paradigm is meaningless unless *second-order* scaling holds. Furthermore, care must be taken when arguing about the "nature" of the process for as with the *LD*, what one sees is often a worst case behavior in the scaling properties of the data. The reason is that summing or averaging the wavelet coefficients leads to a potential bias coming from a few very large wavelet coefficients. For example, the *LD* is largely biased against second- and higher-order non-stationarity. The partition function presents the same defect as the *LD* so that multiscaling in the process could hide self-similarity or LRD if the value of the irregularities of this component are sufficiently large compared to the other components of the process at a given timescale. Hence, care must be taken not to consider too seriously what the $M(j)$ says without carrying a time-dependent analysis in order to check whether non-stationarity in the process does not bias the estimation. Instead of the $M(j)$, we shall rather use $log(M(j))$ that maps the three types of scaling processes, LRD (or positive correlations), self-similarity and multiscaling respectively to the negative, zero and positive values of $log(M(j))$. Another important thing to be noticed is that this scaling among moments cannot be used without considering a certain number of timescales for the qualitative analysis to be meaningful. It is not possible to analyze one single timescale through the $M(j)$ since scaling still has to be present in the *second-order* properties for the *high-order* analysis to make sense. The multiscaling paradigm is useful as a complement to *second-order* analysis, not as a standalone tool.

3.3.2 High-order analysis

The counterpart of the *LD* for the *high-order* properties is the partition function, as described in section 3.3.1. While the *LD* informs one about the energy (variance) of the increments of the process, the partition function gives information about the distribution of large and small wavelet coefficients at each timescale. For this, we plot $log(S(q, j))$ as a function of the scale j, for a range of values of the moments q. If the *low-order* moments ($q \leq 2$) have a larger $log(S(q, j))$ than *high-order* moments ($q > 2$), then in means that the deal with a *second-order* process (according to the terminology defined in section 3.3.1), which is well described by the low order properties of the wavelet coefficients. A *second-order* process must in turn exhibit a relatively smooth sample path at this timescale, without large-valued wavelet coefficients.

Figure 3 presents the logarithm of the partition function $log(S(q, j))$ for all traces. Two important things must be looked at when interpreting the partition function. First, discriminating between *second-order* (smooth) and *high-order* (irregular) processes can be done by determining for which values of q the $log(S(q, j))$ is the largest. A *second-order* process will have larger $log(S(q, j))$ for $q \leq 2$ while a *high-order* process for $q > 2$. Figure 3 shows that Auckland outgoing, UCB and LBL are *second-order* for the timescales above a few seconds. Auckland incoming is *second-order* for timescales larger than 5 minutes only. All other traces are *high-order* processes. This implies that most of the traces contain large wavelet coefficients at each timescale, hence large irregularities.

An often underestimated issue with wavelet analysis concerns non-stationarity. Wavelet estimators rely on some sufficiently large sample of a time series where the wavelet coefficients on the fixed dyadic grid are assumed to be stationary, at least in the sense of the studied properties of the process. The wavelet coefficients can then be used to compute a meaningful statistic of some property of the process. Wavelet-based statistics however might get biased by a few very large coefficient arising due to a few large irregularities in the process at some timescale. To check for such non-stationarity, we compute the statistics over many subsamples of the time-series and compare them so that one can determine what is a true property of the process and what is just an "outlier", i.e. an abnormal event.

So let us now turn to the time-dependent partition function, to determine whether the *high-order* behavior identified in the partition function arises for limited periods of the process or is an actual property of the process that lasts for its whole duration. Figure 4 shows the distribution of the partition function for limited time intervals for the second and the sixth moments. This figure indicates how frequently the *high-order* properties ($log(S(6, j))$) are more important than ($log(S(2, j))$), by plotting the values of the $log(S(q, j))$ (for $q = 2$ and 6 only) computed over half-overlapping and constant-size time intervals.

The Auckland outgoing plot (top left of figure 4) shows that a single block is *high-order* so that the partition function (top left of figure 3) was not biased by this single event. The Auckland incoming plot (top right of figure 4) on the other hand shows that a few blocks have larger *high-order* moments for octaves [8,14] but this time these events were sufficiently strong to bias the partition function so that one would consider these timescales as containing *high-order* scaling while they are actually *low-order* most of the time. The same bias arises for the PSC trace for octaves [9,14] that are mistaken as *high-order* by two time blocks during which *high-order* moments are larger than the *second-order* one. The other traces do not seem to contain such a pathological behavior.

The raw plot of the partition function provides information about the smoothness of the process at each timescale. To infer the type of scaling that occurs over some timescale range, one must also check the linearity of the moments of the process relative to one another. For this, we compute for small blocks of time the $M(j)$

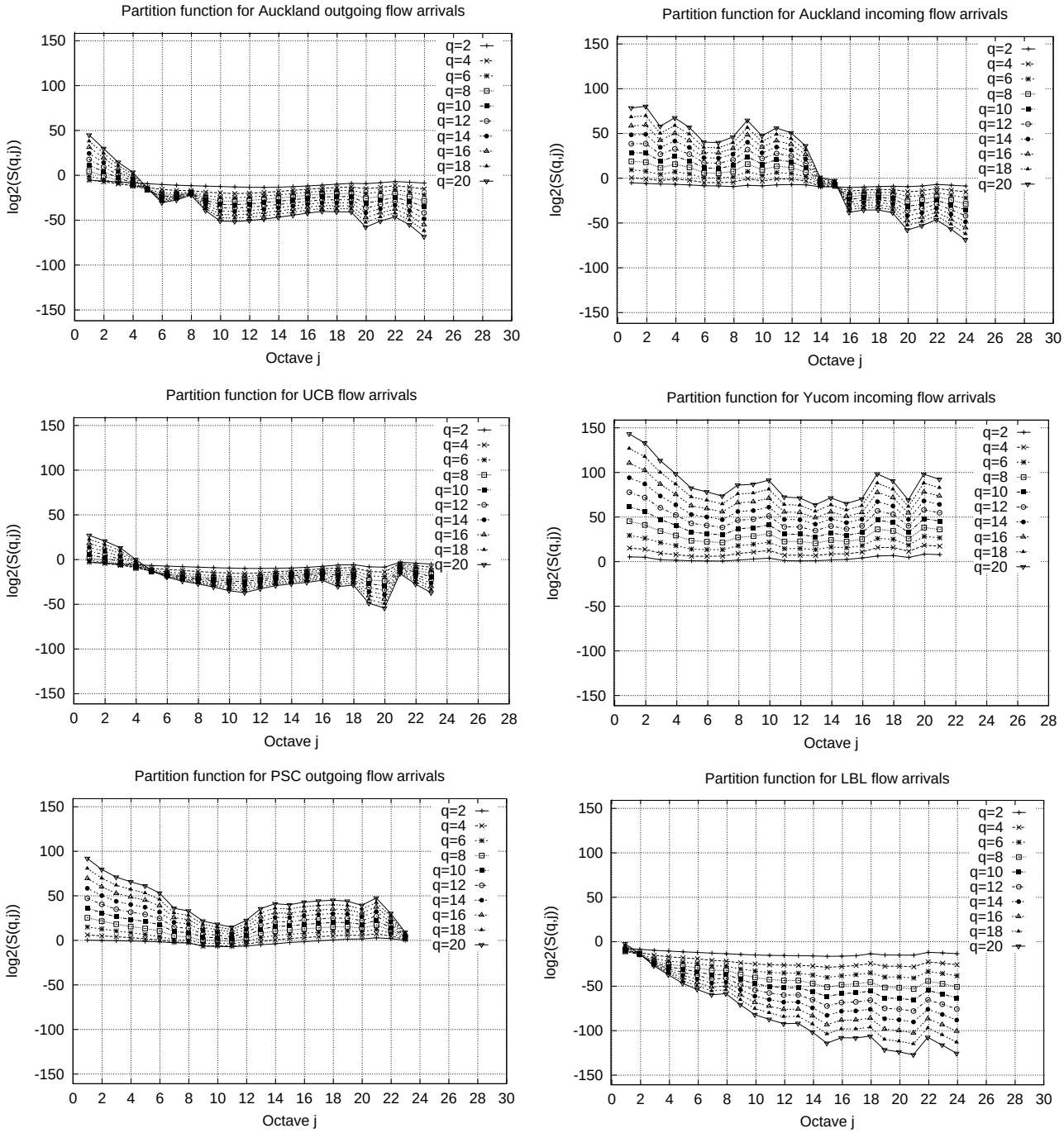

Figure 3: Partition functions.

for each j and each trace, then we plot $log(M(j))$ as a function of j. Figure 5 presents the evolution in time of $log(M(j))$ for each j, for each trace.

Recall that *high-order* scaling cannot be inferred unless *second-order* scaling is also valid. Strict *second-order* scaling together with negative values of $log(M(j))$ are evidence of LRD or positive correlations. Strict *second-order* scaling together with values of $log(M(j))$ around zero are evidence of self-similarity. Finally, strict *second-order* scaling together with positive values of $log(M(j))$ are evidence of multiscaling.

Since no strict *second-order* scaling was identified in the *LD*, one should conclude that there is no evidence of scaling in the TCP flow arrivals except for the "pathological" behavior of the small timescales of the PSC trace. Things are not that simple in practice. The *3D-LD*'s of Figure 2 showed evidence of different components in the TCP flow arrivals, whose scaling properties vary with time and from trace to trace.

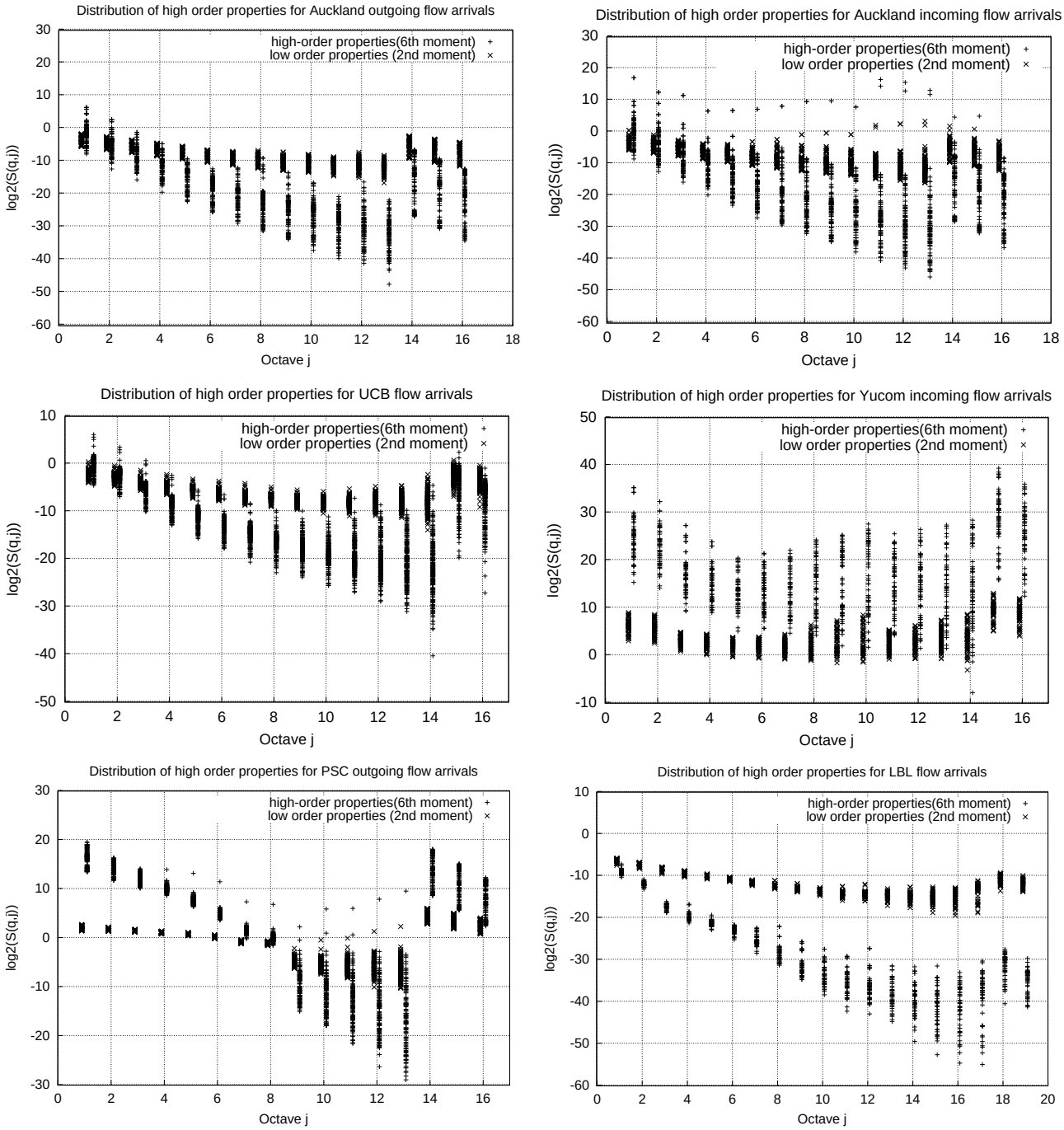

Figure 4: Stationarity check of *high-order* scaling properties.

The LBL trace (bottom right of Figure 5) is the easiest to interpret since all values of the $log(M(j))$ are negative, showing evidence for no high-order scaling. The Yucom trace (middle right of Figure 5) exhibits the opposite behavior, with only positive values of its $log(M(j))$, pointing to some form of high-order scaling. Since no *second-order* scaling could be identified in the Yucom trace, we do not know what type of process this trace may contain particularly with the absence of structure in its *high-order* properties (no straight line in the partition function nor linear moment scaling in the $M(j)$). The PSC trace (bottom left of Figure 5) exhibits a mixed behavior of its $log(M(j))$, with positive values for the subsecond timescales while negative values for the large timescales. The intermediate region of the $M(j)$ for PSC is due to the crossing of the second and high-order moments at timescales around a few seconds, that explains the large positive values of the $log(M(j))$ caused by a very small denominator of the $M(j)$. The same behavior of positive values of the $log(M(j))$ caused by a very small denominator of the $M(j)$ arises at the smallest timescales for the Auckland and UCB traces. We conjecture that this behavior is probably not more than some kind of "phase transition" occurring

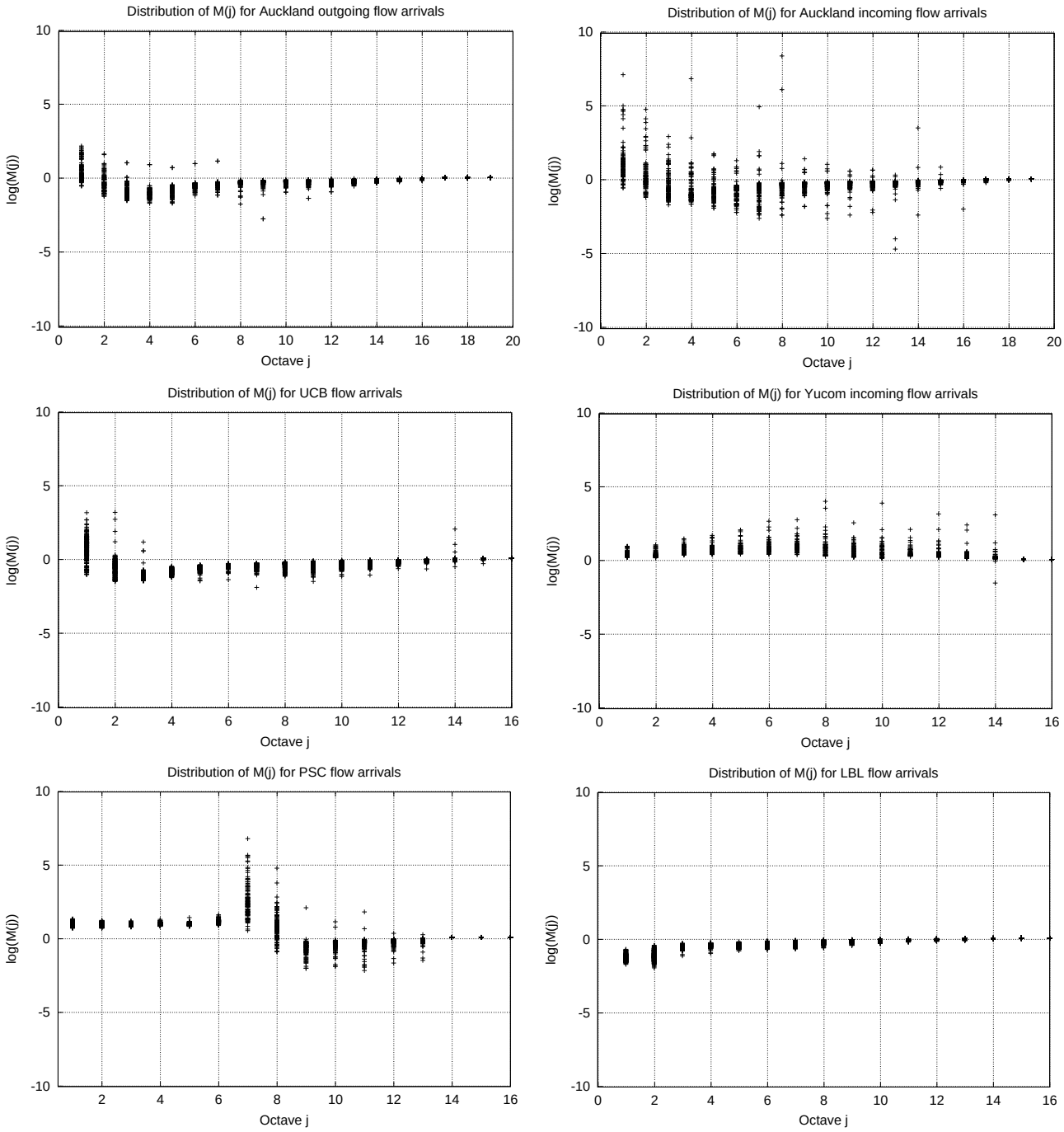

Figure 5: Stationarity check of M(j) estimator.

in the statistical properties of the TCP flow arrivals, maybe due to a too small number of TCP flow arrivals at the timescales that are mistaken by the wavelet analysis as large changes in the process. We do not believe that this behavior has anything to do with high-order scaling.

As most of the $log(M(j))$'s of the traces are negative, this suggests that except for the small timescales of PSC, there is no *high-order* scaling in the TCP flow arrivals. The *second-order* scaling that arises in the TCP flow arrivals at timescales larger than seconds is hence more likely to be connected to correlations or LRD. The smoothness of the TCP flow arrivals at timescales larger than seconds, as shown by the stationarity check of the partition function (Figure 4), seems to rule out the presence of self-similarity in this process.

3.4 Applications behavior

In this section we study one particular trace, the PSC trace. The reasons to focus on this trace are numerous: first we have the information about the TCP flow arrivals of the internal clients of PSC.

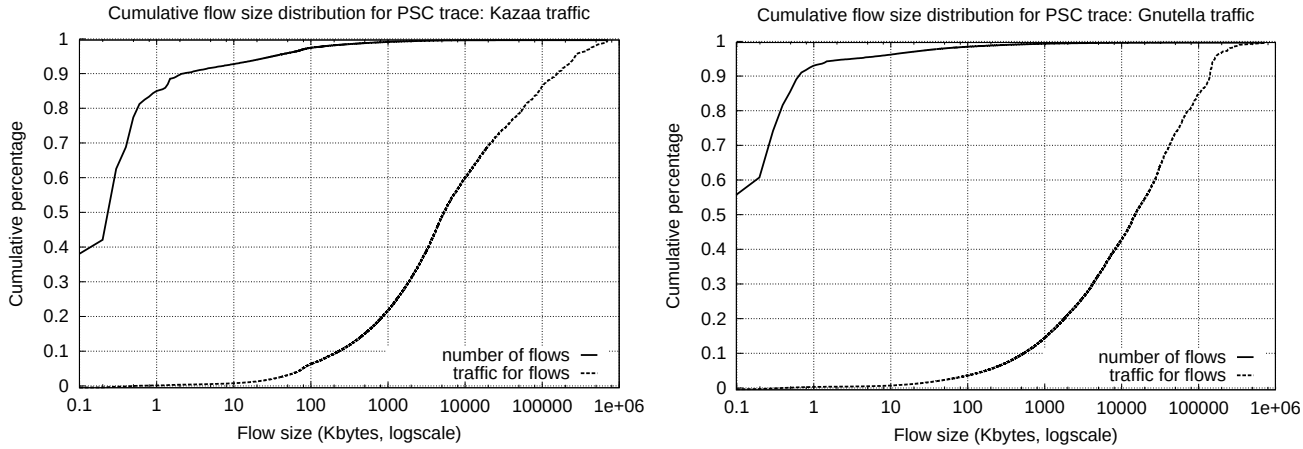

Figure 6: Flow size distributions for Kazaa (left) and Gnutella (right).

Table 2: Comparison of flows mix for the two largest PSC clients.

Client	HTTP	P2P	Total
1	9,767,651	2,265,086	12,032,737
2	2,038,776	2,769,587	4,808,363

Second, we know that the largest client undergoes rate limitation. We can thus compare the impact of rate limitation on the TCP flow arrivals. Finally, we know that a significant percentage of the traffic is P2P so that we can also compare the difference in scaling behavior between P2P and HTTP flows.

Table 2 compares the number of flows for the two largest clients, for HTTP and P2P (only Kazaa and Gnutella, ports 1214 and 6346). The number of P2P flows is quite large, representing 18 % of the TCP flows for the largest client and 57 % for the second largest client. It is well known that there are two types of TCP flows in P2P traffic: those that are used for "control" purposes and those that carry the real traffic exchanged between P2P hosts. P2P traffic has a particular TCP flow size distribution due to two different types of requests, as illustrated by Figure 6. Figure 6 shows that there are many very small flows whose size is smaller than 1 KBytes (85 % for Kazaa, 93 % for Gnutella), while these flows do not represent a significant traffic volume (0.7 % for Kazaa, 0.8 % for Gnutella). These small P2P flows correspond to arrivals that did not induce a complete transfer (control messages or aborted transfers). Most of the traffic volume for P2P arises from very large flow sizes. P2P flows of size larger than 1 MBytes represent more than 93 % of the Kazaa traffic and more than 95 % of the Gnutella traffic. Due to space limitations, we do not show the scaling analysis of the small and large P2P flows, but there is no important difference between the scaling properties of the arrivals of these flows.

Figure 7 presents the *LD* (top), *3D-LD* (middle) and M(j) estimator (bottom) for the two largest clients of PSC. The difference between the *second-order* properties of the two clients is obvious: the largest client has the mark of its rate limitation. The *LD* of the largest client exhibits a larger variance of its wavelet coefficients for subsecond timescales while the second client has a more "common" *LD*. The *3D-LD* confirms the information provided by the *LD*, in that rate limitation affects the TCP flow arrivals all the time. The $log(M(j))$ confirms that the *high-order* scaling is linear in q for octaves [1,5]. This behavior corresponds to multifractality since the moments diverge linearly from one another at the smaller timescales, hence the almost constant $log(M(j))$. [11] showed that the small timescales of congested TCP traffic was multifractal. Here we show that such models as multiscaling and multifractals are even applicable for TCP flow arrivals, not only for traffic volume at the packet-level.

To conclude this section, the two main applications present in the PSC trace do not exhibit the same scaling properties. Figure 8 provides the *3D-LD*'s of the two largest clients, by splitting their flows into HTTP (port 80) and P2P (ports 1214 and 6346). As evidenced by Figure 8, only the HTTP flows of the largest client of PSC (top left of Figure 8) exhibit multifractal scaling at timescales below seconds. P2P flows on the other hand do not seem to be sensitive to rate limitation.

4. EVALUATION

In this section, we discuss the implications of the findings of this paper on the scaling properties of Internet traffic identified in the literature.

[7, 24] have shown the influence of the distributional properties of web traffic and file sizes on the statistical self-similarity of the traffic. This self-similarity was a second-order one, not strict self-similarity as used in this paper, and it involved a single value H for all timescales. Other papers [25, 12, 15, 11] dealing with scaling in Internet traffic have convincingly showed the influence of the TCP protocol and its dynamics on the traffic scaling properties.

In this paper, we showed that no strict scaling could be identified in the TCP flow arrivals. However, we showed evidence of four different components in this process of the TCP flow arrivals. At the subsecond timescales, an uncorrelated process takes place. Between seconds and minutes, we identified a *second-order* non-stationary process that could be caused by correlations. Then, at timescales between minutes and hours, we found evidence for another *second-order* process. The fourth component of the TCP flow arrivals at timescales beyond hours confirmed [35] that found a "time of the day" pattern. Relying on *high-order* scaling and non-stationarity checks revealed that TCP flow arrivals were everything but stationary and that this process has relatively smooth sample path proper-

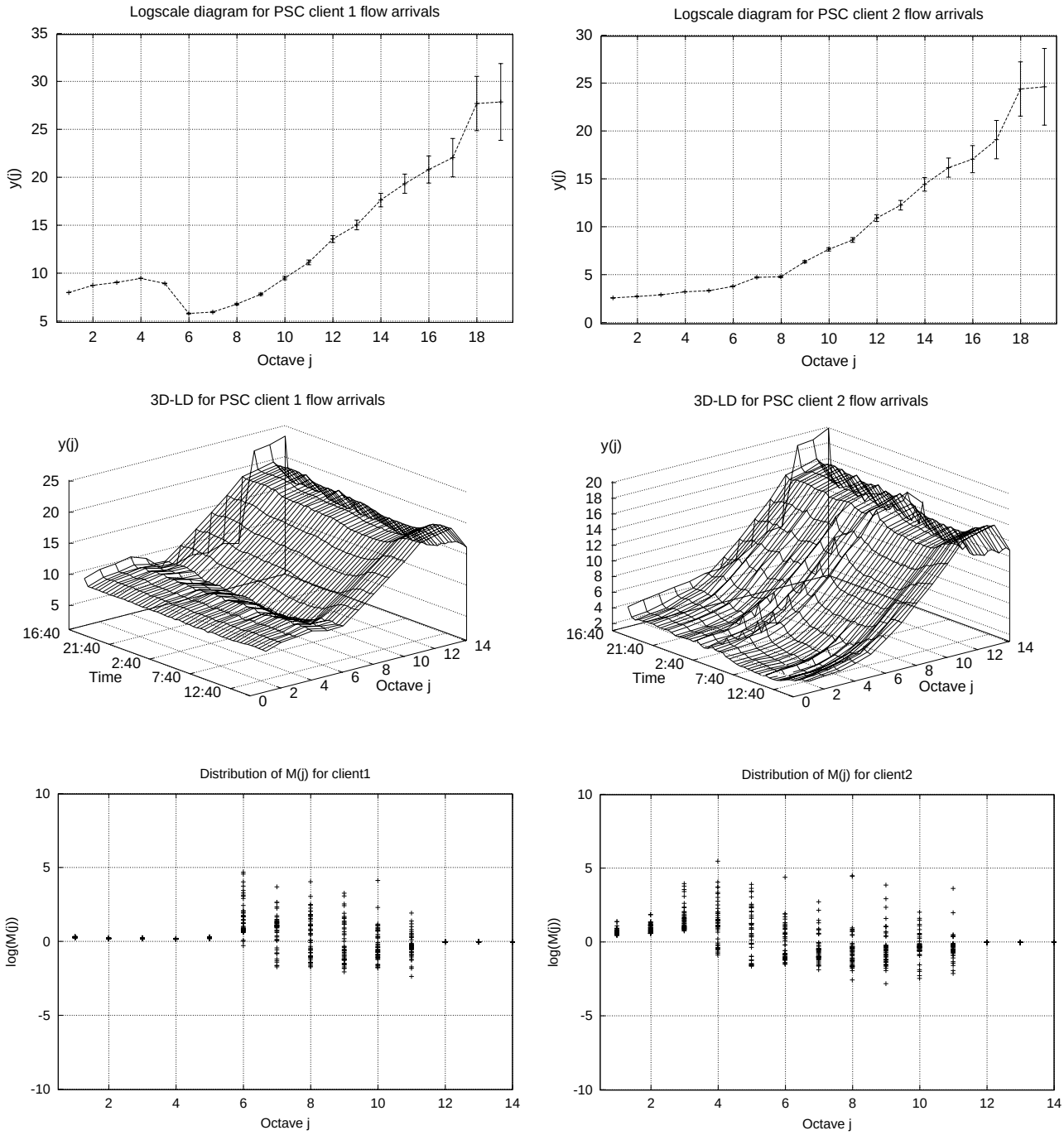

Figure 7: *LD, 3D-LD* and *M(j)* estimator for largest two PSC clients.

ties.

Our study of the non-stationarity of the scaling properties of the TCP flow arrivals also provided a methodological contribution to the study of scaling in network traffic. We applied this methodology on the TCP flow arrivals and demonstrated its ability to remove the intrinsic bias of non-stationarity when studying the scaling properties of processes with wavelet-based estimators. We showed in this paper that inferring scaling properties cannot be performed without properly caring about issues like non-stationarity or the presence of several components in the process under study. The assumption of stationarity is too often made in the literature without carefully checking for it. This is not to say that previous works in the literature should be questioned. The presence of scaling in network traffic is not questionable. However, studying the scaling properties of some process requires a thorough methodology, mainly because of the very nature of scaling processes whose sample path properties tend to bias many scaling estimators [8].

In this paper, we confirmed the presence of *second-order* scal-

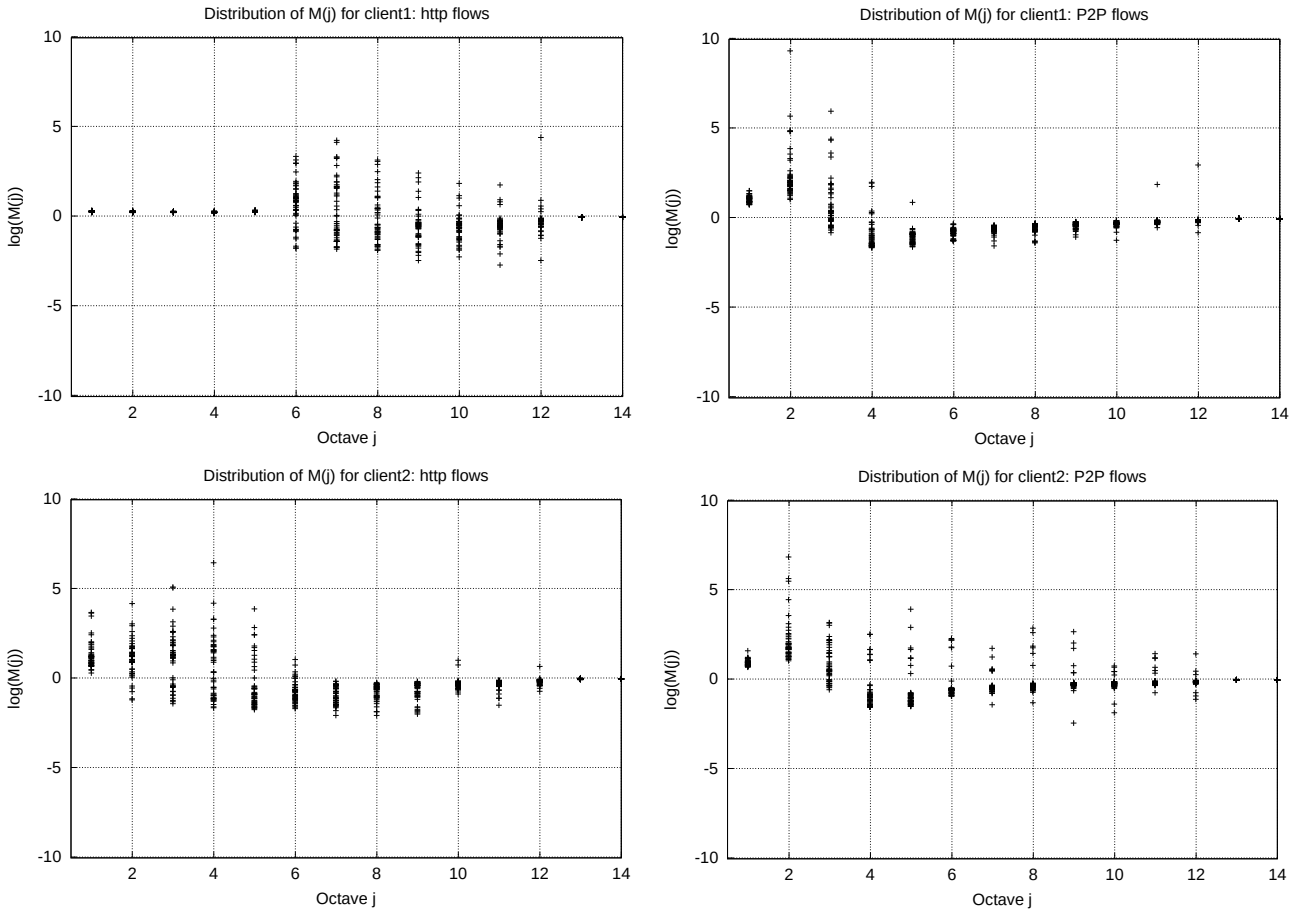

Figure 8: *3L-LD*'s for HTTP and P2P flows of largest two PSC clients.

ing (although not strict scaling) in the TCP flow arrivals, confirming [10, 35]. Furthermore, we showed that *high-order* scaling was also necessary to properly describe its dynamics. We showed evidence for multifractal-like scaling in the HTTP flows of the largest client of PSC undergoing rate limitation. This paper therefore points out the importance of the process of the TCP flow arrivals as another potential cause for the scaling identified in Internet traffic. While flow size distribution in web traffic and network mechanisms are undoubtedly involved in the scaling nature of network traffic, flow arrivals could also be involved, even if masked by other types of scaling. This paper hence asks for more investigations about the relationships between scaling in network traffic and the process of the flow arrivals. We showed in this paper that our understanding of the process of the flow arrivals is still extremely limited. Much interesting information could be gathered by thoroughly studying this process, especially the relationships between the network conditions and the interactive behavior of the users. The *high-order* moments properties provide largely unexploited information about the fine-grained dynamics of network traffic.

5. CONCLUSION

In this paper, we have studied the process of the TCP flow arrivals through the wavelet lens and analyzed in details its scaling properties. We studied the non-stationarity of the TCP flow arrivals and provided a methodology to understand the type of scaling present in this process as well as remove the bias from non-stationarity. We showed that such a methodology is necessary to guarantee a sound inference of the scaling properties of such a complex process as the TCP flow arrivals.

We first described the different scaling processes and the wavelet tools aimed at analyzing them. Then, we studied the *second-order* scaling properties of the TCP flow arrivals and showed its non-stationarity. We also showed that no strict scaling was present over the whole range of timescales of the traces. Our analysis pointed to the nature of the process at each timescale, with an uncorrelated process at the subsecond timescales, correlation between seconds to minutes, LRD or self-similarity between minutes and hours, and finally the "time of the day" pattern for timescales larger than hours. Knowing that *second-order* properties are not sufficient to infer the scaling nature of a process, we then analyzed the *high-order* properties of the TCP flow arrivals, to discover that *high-order* scaling was present in the HTTP flows of one of the traces, caused by rate limitation. The P2P flows of the same trace were however not affected by this rate limitation. We showed that this rate limitation produced a stationary multifractal-like scaling at the subsecond timescales.

Acknowledgements

The author acknowledges all the people who contributed to the collection of the traffic traces used, and particularly among them (in no particular order): B. Piret from BT Ignite Belgium for the Yucom trace, C. Rapier from the Pittsburgh Supercomputing Center for the

PSC trace, V. Paxson from ACIRI for the LBL trace, S. Gribble from UCB for the UCB HOME IP trace and finally the NLANR MOAT and the WAND research groups for the AUCKLAND IV data set, as well as the members of the DAG project [22] that made possible the collection of this trace.

The author also acknowledges the anonymous referee for his insightful comments.

6. REFERENCES

[1] Internet Traffic Archive. `http://www.acm.org/sigcomm/ITA/`.

[2] P. Abry, R. Baraniuk, P. Flandrin, R. Riedi, and D. Veitch. The Multiscale Nature of Network Traffic: Discovery, Analysis, and Modelling. *IEEE Signal Processing Magazine*, May 2002.

[3] P. Abry, P. Flandrin, M. Taqqu, and D. Veitch. Wavelets for the analysis, estimation, and synthesis of scaling data. *In [26]*.

[4] J. Beran. *Statistics for Long-Memory Processes*. Monographs on Statistics and Applied Probability, Chapman & Hall, 1994.

[5] J. Cao, W. Cleveland, D. Lin, and D. Sun. On the nonstationarity of Internet traffic. In *Proc. of* ACM SIGMETRICS 2001, pages 102–112, June 2001.

[6] Cisco. NetFlow services and applications. White paper, available from `http://www.cisco.com/warp/public/732/netflow`, 1999.

[7] M. Crovella and A. Bestavros. Self-similarity in world wide web traffic evidence and possible causes. In *Proc. of* ACM SIGMETRICS'96, pages 160–169, May 1996.

[8] T. Dang and S. Molnar. On the effects of non-stationarity in long-range dependence tests. Periodica Polytechnica *Ser. El.*, 43(4):227–250, 1999.

[9] I. Daubechies. *Ten Lectures on Wavelets*. Number 61 in CMBS-NSF Series in Applied Mathematics, SIAM, Philadelphia, 1992.

[10] A. Feldmann. Characteristics of TCP connection arrivals. *In [26]*.

[11] A. Feldmann, A. Gilbert, P. Huang, and W. Willinger. Dynamics of IP traffic: A study of the role of variability and the impact of control. In *Proc. of ACM SIGCOMM'99*, pages 301–313, 1999.

[12] A. Feldmann, A. Gilbert, and W. Willinger. Data networks as cascades: Investigating the multifractal nature of internet WAN traffic. In *Proc. of ACM SIGCOMM'98*, pages 42–55, 1998.

[13] A. Feldmann, A. Gilbert, W. Willinger, and T. Kurtz. The changing nature of network traffic: scaling phenomena. *ACM Computer Communication Review*, 28(2):5–29, 1998.

[14] A. Gilbert. Multiscale analysis and data networks. *Appl. Comp. Harmon. Anal.*, 10(3):185–202, 2001.

[15] A. Gilbert, W. Willinger, and A. Feldmann. Scaling analysis of conservative cascades, with applications to network traffic. *IEEE Trans. on Information Theory*, 45(3):971–992, 1999.

[16] S. D. Gribble. UC Berkeley Home IP HTTP Traces. Available at `http://www.acm.org/sigcomm/ITA/`, July 1997.

[17] N. Hohn, D. Veitch, and P. Abry. Does fractal scaling at the IP level depend on TCP flow arrival processes? In *Proc. of ACM SIGCOMM Internet Measurement Workshop*, 2002.

[18] S. Jaffard. Multifractal formalism for functions Part I : Results valid for all functions. *SIAM J. Math. Anal.*, 28(4):944–970, July 1997.

[19] T. Karagianis, M. Molle, M. Faloutsos, and A. Broido. A nonstationary Poisson view of Internet traffic. In *Proc. of IEEE INFOCOM*, March 2004.

[20] K. Keys, D. Moore, R. Roga, E. Lagache, M. Tesch, and K. Claffy. The architecture of Coralreef: an Internet traffic monitoring software suite. In *Proc. of* PAM2001, April 2001.

[21] W. Leland, M. Taqqu, W. Willinger, and D. Wilson. On the Self-similar Nature of Ethernet Traffic (Extended Version). IEEE/ACM *Transactions on Networking*, 1994.

[22] U. of Waikato. The DAG project. `http://dag.cs.waikato.ac.nz/`.

[23] J. Org, M. Ian, and G. Brownlee. The auckland data set: an access link observed. In *Proc. of the 14th ITC Specialist Seminar*, April 2001.

[24] K. Park, G. Kim, and M. Crovella. On the relationship between file sizes, transport protocols, and self-similar network traffic. In ICNP'96, 1996.

[25] K. Park, G. Kim, and M. Crovella. On the effect and control of self-similar network traffic. In *Proc. SPIE International Conference on Performance and Control of Network Systems*, pages 296–310, 1997.

[26] K. Park and W. Willinger. *Self-Similar Network Traffic and Performance Evaluation*. Wiley-Interscience, 2000.

[27] V. Paxson. Empirically-derived analytic models of wide-area tcp connections. IEEE/ACM *Transactions on Networking*, 2(4):316–336, 1994.

[28] V. Paxson and S. Floyd. Wide-area traffic: the failure of Poisson modeling. IEEE/ACM *Transactions on Networking*, 3(3):226–244, 1995.

[29] R. H. Riedi. An improved multifractal formalism and self-similar measures. *J. Math. Anal. Appl.*, 189:462–490, 1995.

[30] S. Romig, M. Fullmer, and R. Luman. The osu flow-tools package and cisco netflow logs. In *Proc. of* USENIX LISA, December 1995.

[31] D. Sornette. *Critical Phenomena in Natural Sciences*. Springer-Verlag, 2000.

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

[33] A. Tewfik and A. Kim. Correlation structure of the discrete wavelet coefficients of fractional brownian motion. IEEE *Trans. Inform. Theory*, 38:904–909, 1992.

[34] S. Uhlig. Conservative cascades: an Invariant of Internet traffic. In *Proc. of IEEE ISSPIT2003*, December 2003.

[35] S. Uhlig, O. Bonaventure, and C. Rapier. 3D-LD : a graphical wavelet-based method for analyzing scaling processes. In *Proc. of the 15th ITC Specialist Seminar*, July 2002.

[36] D. Veitch, P. Abry, P. Flandrin, and P. Chainais. Infinitely divisible cascade analysis of network traffic data. In *Proc. of ICASSP, Istanbul, Turkey*, June 2000.

[37] W. Willinger, V. Paxson, R. H. Riedi, and M. S. Taqqu. Long-range dependence and data network traffic. *In "Long range Dependence : Theory and Applications", edited by Doukhan, Oppenheim and Taqqu*, 2001.