Available online at www.sciencedirect.com
SCIENCE DIRECT®
ELSEVIER
Computer Networks 48 (2005) 293–313
Computer Networks
www.elsevier.com/locate/comnet

Multifractality in TCP/IP traffic: the case against

Darryl Veitch [a,*], Nicolas Hohn [a], Patrice Abry [b]

[a] ARC Special Research Center for Ultra-Broadband Information Networks (CUBIN), Department of Electrical and Electronic Engineering, The University of Melbourne, Melbourne, Victoria 3010, Australia
[b] CNRS, UMR 5672, Laboratoire de Physique, Ecole Normale Supérieure de Lyon, F-69364 Lyon, France

Available online 18 January 2005

Abstract

The discovery of long-range dependence (a kind of asymptotic fractal scaling) in packet data from LANs and WANs, was followed by further work detailing evidence for multifractal behaviour in TCP/IP traffic in WANs. In terms of networking however, physical mechanisms for such behaviour have never been convincingly demonstrated, leaving open the question of whether multifractal traffic models are of black box type, or alternatively if there is anything 'real' behind them. In this paper we review the evidence for multifractal behaviour of aggregate TCP traffic, and show that in many ways it is weak. Our study includes classic traces and very recent ones. We point out misunderstandings in the literature concerning the scales over which multifractality has been claimed. We explain other pitfalls which have led to the multifractal case being overstated, in particular the possibility of 'pseudo scaling' being confused with true scaling, due to shortcomings in the statistical tools. We argue for an alternative point process model with strong physical meaning. It reproduces the higher order statistics of the data well, despite not being calibrated for them, yet is not multifractal. From its standpoint, the empirical multifractal behaviour is seen as a misinterpretation due to a lack of power in the statistical methodology.

Keywords: Multifractal; Estimation; Internet; Traffic; Cluster

1. Introduction

1.1. Motivation

Teletraffic analysis and practice was transformed by the discovery of scale invariance properties in packet traffic [1]. The presence of large-scale asymptotic scale invariance, or long-range dependence (LRD), is remarkably universal, and has

* Corresponding author.
E-mail addresses: dveitch@unimelb.edu.au (D. Veitch), n.hohn@ee.mu.oz.au (N. Hohn), pabry@ens-lyon.fr (P. Abry).

1389-1286/$ - see front matter
doi:10.1016/j.comnet.2004.11.011

become an indispensable part of traffic modelling, in particular for TCP/IP traffic in the Internet. This role is destined to continue, as the phenomena has a physical underpinning which is both generic and readily understandable in networking terms, namely the heavy tailed nature of file sizes [2], which, through a well-known mechanism [3], results in heavy tailed flows and thereby LRD.

The discovery of evidence for *multifractal* behaviour, a richer form of scaling behaviour associated with non-uniform local variability, raised hopes that another 'traffic invariant' had been found which could lead to a complete, robust model of aggregate wide area network (WAN) traffic over all time scales. There is now a literature which accepts the existence of multifractal traffic, exploring alternative multifractal models [4], traffic generators [5], and related performance studies [6]. More broadly, it has become somewhat accepted that traffic has multifractal characteristics, despite the fact that physical mechanisms, and network meaning, has never been established in the way it has for LRD.

In this paper we review the evidence underlying the adoption of multifractal traffic models. We are motivated primarily by two factors arising from our own work in the modelling of TCP/IP packet traffic: (i) the weakness of the evidence seen when using the available statistical tools in a careful way and (ii) a realisation of the lack of statistical *power* of those same tools, leading to the possibility of erroneous interpretation. The question we wish to answer is whether the original enthusiasm for multifractal models was warranted, or is warranted today, when using the default statistical tools (arguably the best available) in a consistent and thorough way. We conclude that the evidence is not only weak but misleading, and that (in most senses) there has been up to now no compelling reason to conclude that a MF model is indicated, or is particularly natural, to describe traffic.

It is not possible for us here to definitively rule on the deeper question of whether traffic *is* multifractal or not, for three reasons. First, the set of available statistical tools are not powerful enough to clarify all the related issues. Improvements are needed in their performance, the knowledge of their performance under different conditions, and important capabilities such as hypothesis tests are absent. Second, ultimately there is no '*is*', modelling data by mathematical processes with multifractal properties reduces to a philosophical issue of model choice, there may always be *some* sense in which a MF model is correct, or rather, useful and/or appropriate (over some scale range). Finally, traffic is an evolving phenomenon, and so conclusions clearly cannot be final in a temporal sense.

After describing necessary background on multifractals and the statistical tools in the remainder of this introduction, Section 2 provides a succinct overview of the key parts of the literature. In Section 3 we compare and contrast the claims of this prior work, and attempt to clarify the causes of the sometimes contradictory claims, particularly with regard to the scale-range over which evidence of multifractality is found. We then offer our own reexamination of the question for several traces, including two of historical importance. Section 4 discusses drawbacks in the existing statistical procedures and tools, and illustrates circumstances where they can be misleading. Through a non-multifractal point process cluster model we recently proposed in [7], Section 5 completes our discussion by combining issues of physical meaning with estimator limitations to decide against multifractal models. This model is greatly preferable to multifractal alternatives on physical and networking grounds. Although not being fitted for the purpose or designed to do so, it can produce multifractal-like statistical signatures which can be *as convincing* as those for the data, although it is not multifractal. We summarise our findings in Section 6.

1.2. Wavelets, scaling and multifractals

It is not possible here to give a detailed introduction to the field of statistical estimation, or the realm of multifractal processes. We provide a concise practically oriented background sufficient to support our presentation. We follow the wavelet viewpoint, first introduced to traffic analysis in [8], and since become the defacto standard, due to its advantageous statistical and computational properties. We use software we developed ourselves (freely available at [9]) to perform the statistical analysis both at second order and at higher order, and are guided by the methodology outlined in [10]. These procedures are (arguably) the best available, but are nonetheless far from perfect, and in particular the analysis beyond second order is not as well understood as the second order analysis used to study LRD. We describe a theoretical limitation of the techniques we use here below, and discuss other drawbacks in Section 4. Further details of the use of these tools in the networking context can also be found in the review article [11]. A more mathematical introduction to multifractal processes can be found in [12].

Long-range dependence is a form of asymptotic scale invariance in the limit of large scale (low frequency). If $X(t)$ is a continuous time stationary process with power spectral density $\Gamma_X(\nu)$, LRD can be defined as a power law divergence of the spectrum at the origin:

$$\Gamma_X(\nu) \sim c_{\mathrm{f}}|\nu|^{-\alpha}, \quad |\nu| \to 0, \quad \text{with } \alpha \in (0,1). \tag{1}$$

To detect this phenomena using wavelets, first define the discrete wavelet transform coefficients as

$$d_X(j,k) = \int_{-\infty}^{\infty} X(t)\psi_{j,k}(t)\,\mathrm{d}t, \tag{2}$$

where the member $\psi_{j,k}(t) = 2^{-j/2}\psi(2^{-j}t - k)$ of the basis function family is generated from the mother wavelet $\psi(t)$ by dilation by a scale factor $a = 2^j$, and translation by $2^j k$. At fixed octave j, the sequence $\{d_X(j,\cdot)\}$ corresponds to an analysis of $X(t)$ at scale 2^j. It can be shown that the variance of this process satisfies

$$\mathbb{E}|d_X(j,\cdot)|^2 = \int \Gamma_X(\nu)2^j|\Psi(2^j\nu)|^2\,\mathrm{d}\nu, \tag{3}$$

where $\Psi(\nu)$ denotes the Fourier transform of ψ. This equation defines a kind of wavelet energy spectrum, analogous to a Fourier spectrum, but much better suited to the study of fractal processes. In the case of LRD it reads

$$\mathbb{E}|d_X(j,\cdot)|^2 \sim c_{\mathrm{f}}C(\alpha)2^{j\alpha}, \quad j \to +\infty, \tag{4}$$

where $C(\alpha) = \int|\nu|^{-\alpha}|\Psi(\nu)|^2\mathrm{d}\nu$ is close to a constant.

To estimate the wavelet spectrum from data, the time averages

$$S_2(j) = \frac{1}{n_j}\sum_k |d_X(j,k)|^2, \tag{5}$$

where n_j is the number of $d_X(j,k)$ available at octave j (scale $a = 2^j$), perform very well, because of the short range dependence in the wavelet domain. Indeed, the mother wavelet is characterised by an integer $N \geqslant 1$, known as the number of vanishing moments, which satisfies $\int_{\mathbf{R}} t^k\psi(t)\,\mathrm{d}t \equiv 0$ for all $k = 1,2,\ldots,N-1$, and $\int_{\mathbf{R}} t^N\psi(t)\,\mathrm{d}t \neq 0$. It plays a central role in the wavelet based analysis of long memory processes, since the wavelet coefficients $\{d_X(j,k),\ k \in \mathbf{Z}\}$ at a given scale 2^j are short range dependent provided $N > \alpha - 1$. This has been proven in various contexts, see [10] for a review. To enable approximate but analytic analysis of the performance of the scaling exponent estimation procedures, we idealised this whitening of the $d_X(j,k)$ to exact independence. Numerical simulations such as those reported in [8,13] show that this is a useful approximation.

A plot of the logarithm of the estimates $S_2(j)$ against j we call the *Logscale Diagram* (LD):

$$\text{LD}: \quad \log_2 S_2(j) \text{ vs } \log_2 a = j. \tag{6}$$

In these diagrams, straight lines constitute empirical evidence for the presence of scaling. For example, a straight line observed in the range of the largest scales with slope $\alpha \in (0,1)$ (see Fig. 1) both reveals long memory and measures its exponent α.

This wavelet based estimator, beyond its conceptual and practical simplicity, also offers robustness against various types of non-stationarity that may be superimposed onto truly LRD data. This again derives from the possibility of easily performing wavelet decompositions with different vanishing moments and comparing the corresponding estimates, as discussed at length in [8,13]. The estimator can also be used to test for the constancy of the scaling exponent along time. It therefore constitutes a tool enabling one to discriminate between true scaling and certain types of non-stationarities that may conspire to imitate it. These topics are detailed in [14] and illustrated using time series of network traffic.

The above definition of scaling is second order based. If X were Gaussian then this would be sufficient, but this is far from the case for TCP/IP traffic over small timescales. One can generalise the 2nd order definition and study qth order

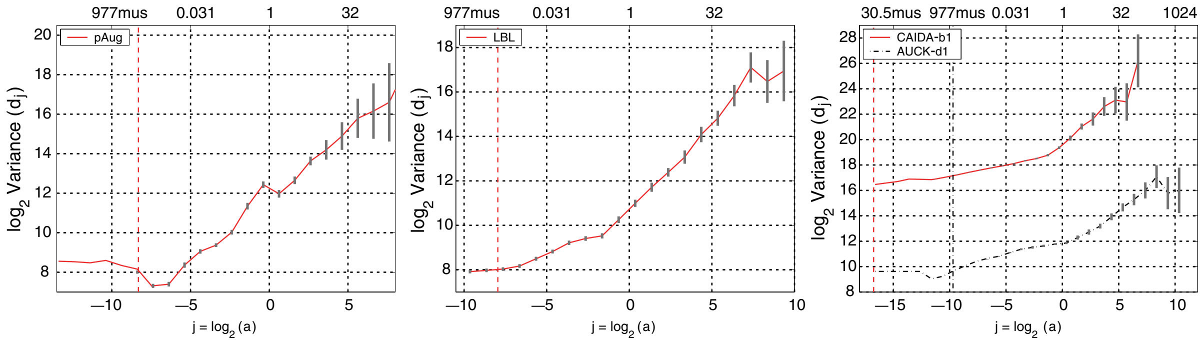

Fig. 1. Wavelet spectra of packet count series $X(t)$. Series are **pAug** (left), **LBL-TCP-3** (middle), and **AUCK-d1** and **CAIDA-b1** (right). For each series LRD is seen. Each of the three WAN traces exhibit biscaling: a second scaling regime to the left of a knee marking the onset of LRD, found respectively at $j^* = \{-6.6, -1, 1, -1\}$. At very small scales the second regime breaks down, eventually tending to a flat point process spectrum at $j \to -\infty$. For calibration, a vertical line is placed at the scale j^{IAT} of the average inter-arrival time.

quantities $\mathbb{E}|d_X(j,\cdot)|^q$, for arbitrary $q \in \mathbf{R}$, by using the estimates

$$S_q(j) = \frac{1}{n_j}\sum_k |d_X(j,k)|^q. \tag{7}$$

Just as the wavelet spectrum serves as a statistically effective summary of second order statistics regardless of whether scaling is at issue or not, a plot of $S_q(j)$ against j, the *qth order Logscale Diagram*:

$$q\text{-LD}: \quad \log_2 S_q(j) \text{ vs } = j \tag{8}$$

is a useful way of examining the raw qth order content of data, independently of any multifractal question, and we use it in this sense below.

For q fixed, a behaviour $\mathbb{E}|d_X(j,\cdot)|^q = c2^{j\alpha_q}$ over some scale range is seen as a straight line in the q-LD, and a measurement of its slope is an estimate of the corresponding q-specific scaling exponent α_q. If, for each q, a straight line is found in the q-LD over the *same* range of scales, then the scaling exponents $\{\alpha_q\}$ are the manifestation of a single underlying scaling phenomenon which we refer to as *multiscaling*. We thereby distinguish it from the conclusion of *multifractality*, which may not necessarily follow, as we will see below. *Multiscaling* will be used synonymously with 'evidence for multifractal behaviour'.

We now explain the connection to multifractals. For many multifractal processes, the collection of exponents $\{\alpha_q\}$ are related to the so called *multifractal spectrum*, which captures the essential details of the multifractal scaling, and can be used to estimate it [12]. We do not attempt to estimate the multifractal spectrum itself, as this would introduce even more estimation difficulties. Instead we adopt the simpler operational approach [10] of testing for linearity of the function $\zeta(q)$. This is because for simple cases such as the exactly self-similar (H-SS) processes with Hurst exponent H, for the corresponding increments processes (which are stationary) $\zeta(q) = qH = \alpha_q + q/2$ is a simple linear function, and one speaks of *monofractality*, whereas for true multifractal processes this is not the case. The same is true for LRD (stationary) processes for which at large scales $\zeta(q) = q(\alpha_2 - 1)/2$. Deviations from linearity can therefore be taken as evidence for the more complex multifractal (MF) behaviour, where a single scaling exponent is insufficient.

In the multifractal case, there is an important theoretical restriction concerning the range of q over which the analysis is to be performed. Exponents $\hat{\zeta}(q)$, estimated according to a multiresolution procedure such as that above, can only be meaningfully measured within a restricted range of q values: $q \in [q_*^-, q_*^+]$, where $q_*^- < 0 < 1 < q_*^+$. This is not a limitation of the estimation procedures themselves. The fact that $\zeta(q) \neq qH$ intrinsically implies that the $\zeta(q)$ are related to the scaling or multifractal properties of the analysed process in a limited range only. This point is studied in detail in [15,16], where theoretical results obtained on Mandelbrot cascades [17] using an aggregation techniques in [18] are extended to a large variety of true multifractal processes such as compound Poisson cascades [19], infinitely divisible random walks [20–23], and fractional Brownian motion in multifractal time [12,24]. The practical consequences are that (q_*^+, q_*^+) must first be estimated (cf. [15,16] for estimation procedures), and second, that the linear behaviour of the $\zeta(q)$ with respect to q should be tested only for $q \in [q_*^-, q_*^+]$.

Following [10], rather than plotting ζ_q against q and looking for linearity, which can be delicate in marginal cases, we plot $h_q \equiv \zeta_q/q$ against q and check for horizontal alignment. This plot:

$$\text{LMD}: \quad h_q \text{ vs } q, \tag{9}$$

we call the *Linear Multiscale Diagram*. Using this approach also has the advantage that the confidence intervals are approximately of the same size, making it easier to assess alignment.

Confidence intervals are a minimal requirement to assess whether slope measurements of lines drawn in any of the LD, q-LD, or LMD defined above are meaningful in any way. Our approach incorporates estimation of confidence intervals. When analysing LRD with our `LDestimate` tool (also available at [9]), confidence intervals are calculated from formulae based on Gaussian approximations. Here we use the `MDestimate` routine and estimate confidence intervals empirically from data. [1] This is because the data at the smallest scales is always non-Gaussian by definition. Indeed, we have routinely observed that at small scales the Gaussian derived intervals are much smaller than those from data, whereas when tested on fractional Gaussian noise, they are almost the same. Even at larger scales when marginals can appear Gaussian, because of the complexity of the data the Gaussian formulae may be misleading over a wide range of scales. Of course assuming non-gaussianity when studying potentially multifractal phenomena is consistent with the well-known fact that Gaussian processes cannot be multifractal.

The confidence intervals follow from an approximate analytic expression for the large n asymptotic form of the variance of $\log_2 S_q(j)$, which relies on the idealised independence property of the wavelet coefficients. A full collection of numerical simulations performed by the authors, as detailed in [10,25], showed that the resulting confidence intervals have acceptable accuracy.

2. A history of MF in traffic

Our aim in this section is to clarify precisely what has been said before, as a backdrop to the analyses that follow. The papers we include in this history of multiscaling in TCP/IP traffic are those which, to the best of our knowledge, really examine the empirical evidence in some detail. Not surprisingly, they are heavily weighted towards the very first discoveries. We do not include papers, for example [26], which are more focused on exploring modelling approaches in a class which is chosen in advance to be of multifractal type. In such cases multiscaling analyses, if they are performed, aim to confirm that plots showing multiscaling can be obtained which are similar to those already seen and interpreted elsewhere, rather than contributing to a debate on the origin and meaning of the evidence itself.

The summaries we give are necessarily brief, and cannot do justice to the full content of the papers. Our main focus is on time indexed traffic processes for which time scales of observation can be meaningfully discussed. We denote these as

- $W(t)$: full byte arrival process (takes values in $\{0, 1\}$).
- $W_\delta(k)$: discretised W, byte count in intervals of width δ.
- $X(t)$: point process of packet arrival times.
- $X_\delta(k)$: discretised X, packet counts in intervals of width δ.

Apart from a desire to report precisely which traffic processes have been studied in the literature, we make the above distinction between

[1] We use the `sigtype=1` option to `MDestimate`, and internally to the `wtspecq_statlog` routine we set `parametric` to 0.

the discretised and continuous time versions (which in practice will also be analysed above some chosen minimum resolution δ) for a technical reason involving the correct initialisation of the algorithm used to calculate the discrete wavelet transform. In essence, discretising before analysing creates a sampling error, whereas a direct analysis (which is possible for point processes like $X(t)$ but not generally), avoids these errors. Further details of this point can be found in [7]. After this section we focus on $X(t)$ and use an exact initialisation procedure to avoid sampling errors (these are generally negligible beyond the smallest three octaves, i.e. for $a > 8\delta$). This allows the important flexibility of choosing δ at will, subject to computational constraints. We use Daubechies wavelets with three vanishing moments

Two early studies asked whether the early Bellcore Ethernet LAN traces, including the celebrated '**pAug**' trace [1,27], were MF or not. Using multiscale diagram methods in the time domain, Ref. [28] studied packet count X_δ and byte count W_δ timeseries with bin sizes of $\delta = 10$ ms. They concluded that each were LRD and monofractal at large scale. The qualifier 'large scale' was needed because one of the traces studied was of incoming WAN traffic for which the LRD regime was seen to begin at a cutoff scale around $10\delta = 100$ ms, below which one saw "*a distinctly different scaling behaviour*". This regime however was not investigated. In [8], a distribution rescaling method in the wavelet domain was used to conclude that the continuous time byte arrival process $W(t)$ in **pAug** was monofractal over timescales above 12 ms.

In [29] we find the first report hailing the discovery of multifractal behaviour in TCP/IP traffic. The traces were collected at LAN gateways. One of these, '**LBL-TCP-3**', is publicly available [27] and is revisited below. Using time domain increment based estimation, MF was concluded for the sequence of packet sizes, of inter-arrival times of packets over all scales, and for both the X_δ and W_δ timeseries. In the case of **LBL-TCP-3** (see [30] for the original discussion of this trace), the base resolution used was $\delta = 150$ ms (corresponding to the largest inter-arrival time). In [31] the same authors discussed a subset of these results, emphasizing the fact that multifractal models are essentially trying to model the 'high frequency' components of data, whereas LRD relates to low frequency. The brief report [32] looked at the number of ATM cells per 1 ms, based on traces two seconds long of outgoing traffic from a university, and found multiscaling using the time domain estimation approach of [29].

In [33] (an early version was published as [34]) it was clearly pointed out for the first time that evidence for scaling behaviour in TCP/IP traffic could be found in two separate regimes, one at 'small scales', and one at 'large scales'. These two regimes were defined empirically by a visually obvious change point in the wavelet spectrum, found to occur at scales of a few 100 ms in the traces examined, which included **LBL-TCP-3** (see Fig. 1). The process observed was X_δ with $\delta = 10$ ms, however a single example was given with $\delta = 1$ ms, where the 'very small scale' behaviour changed again. The conclusion from a wavelet based multiscale analysis was MF behaviour over small scales, and LRD (asymptotic second order self-similarity) over large scales. A protocol driven hierarchal redistribution of data was suggested as a possible mechanism for the MF.

The introduction of a multiplicative cascade model was the main aim of [35] (note that such models were already discussed in [29]), however the observation of twin scaling regimes was also confirmed here for different traces. A change point or 'knee' was again found for X_δ with $\delta = 10$ ms at a similar time scale, and it was claimed that WAN traffic is MF at scales below the knee. An attempt was made to locate the source of the multiscaling in the protocol hierarchy, with mixed results.

In [36] new data sets from ISPs were examined and an analysis similar to [35] performed. The same conclusions were reached, but it was noted that the small scale behaviour is complex and can be affected by various factors including network features such as bottlenecks. Most of the paper deals not with real data but with exploring the effects of network parameters on the multiscaling in a detailed TCP simulation over a simple topology.

In [37] (an earlier version appeared as [38]) the term *biscaling* was coined to refer to the presence of twin scaling regimes, again using the empirically clear knee point in a wavelet spectrum to separate the two. Similar results to [33] were found for more recent TCP/IP traces gathered at the Internet access point of the University of Auckland. These are publicly available [39], and are used below. The time series focussed on was X_δ for TCP connection arrivals. The paper's main point was the observation that a class of scaling models known as *infinitely divisible cascades* (IDC), which include multifractals as a special case, could be used to unify the scaling evidence seen over the two regimes, rather than seeing them as entirely separate. For the first time, confidence intervals for the multiscale analysis were calculated and used [9].

In [40] the IDC modelling was explored further for many different time series extracted from Auckland data sets, and the near ubiquity of the biscaling phenomenon highlighted. Again, results were consistent with LRD at large scales, and MF at small. However, it was noted that the quality of the fit and the confidence intervals were such that a monofractal conclusion could almost be drawn at small scales.

The paper [41] does not investigate multifractality in TCP/IP traffic as such, however it contributes to the issue by identifying sources of burstiness in W_δ on scales above $\delta = 500$ ms. A minority of high rate high byte-volume *alpha* flows are identified as the main source of burstiness, whereas the remaining *beta* traffic is LRD but not bursty in amplitude, and is close to Gaussian at high enough aggregation levels. The fact that the *alpha* component can be identified, and carries most of the burstiness, was also verified down to $\delta = 50$ ms [42]. Multifractal spectra were calculated as a means of quantifying the degree of burstiness.

The more recent paper [43] examined traces from high rate Internet backbone links in the Sprint network. The W_δ timeseries was extracted with $\delta = 10$ μs over hour long traces (however only results for $\delta = 1$ ms were presented). The wavelet analysis tools of [9] were used. Scaling ranges 3 octaves wide around 30 ms were used to conclude that the byte level data is monofractal over small scales. The disagreement with the previous literature was explained by observing that in the backbone aggregation levels are much higher, so W_δ is Gaussian, precluding multifractality. The paper goes on to explain the observed second order structure in terms of a classification of flow type.

Finally, Ref. [7] also deals with investigating the network origins of small scale behaviour, for TCP flows from a variety of traces. The focus is mainly on X at second order through the wavelet spectrum (again code from [9] is used), with δ from 5 μs to 5 ms. Using simple point process models which are not scaling at small scales but which empirically show signs of scaling in the wavelet spectrum, the point is made that blind use of the estimation tools can lead to conclusions, such as multifractality, which are not necessarily justified. The main part of the paper then develops the point process cluster model for X which we use below.

3. Reviewing the evidence

As the above history shows, the story of multifractal scaling in TCP/IP traffic contains many variations in terms of traces, time series, time series resolution δ, scale ranges, and finally conclusions. Not all of the latter are compatible.

Of the potential inconsistencies, the most obvious is that not all authors have agreed on the time scale regime to which their conclusions refer. Whereas [29,31] discuss the 'large' scale regime of a few 100 ms and above and conclude in favour of MF as a means to capture the *high* frequency behaviour, Refs. [32,34,33,35,36] are happy to confirm the adequacy of a LRD asymptotically self-similar model over these scales, and instead claim MF behaviour over the 'small' scale regime below a few 100 ms. There is no published clarification of this point, explaining that different authors were in essence not talking about the same thing, and it is still not commonly appreciated today.

The plot comes together yet thickens in the light of [38,37,40], where evidence for multiscaling is admitted over *both* regimes, but given confidence intervals, a monofractal conclusion is handed down over large scales, and a MF conclusion over small—but just barely. Finally, the more recent work of [43,7], whilst again not even considering the possibility of a MF model at large scale, begins to question the MF conclusion even at small scale. Even in [36], it was noted that the MF evidence was not robust in various ways, and that the wavelet spectra reveals a complex combination of effects. In [41,42], multifractality is mentioned in relation to large-scale behaviour, but this time only to describe a small (though important) traffic component.

To attempt to answer further the question of 'what is really going on?', we perform multiscaling analyses for a set of four traces, paying particular attention to the issues mentioned above. In particular, we always compare traces over a common 'real-time' scale, and seek to understand and compare actual values, rather than just looking for slopes. We will in each case analyse the process X. Studies by many groups, including ourselves, show that the scaling properties of X and W are very similar, and these time processes are in many ways also the most important. We briefly consider the time series of inter-arrival times of packets at the end of the section. We employ our wavelet based code [9], using the 'non-Gaussian' option for confidence intervals, as explained above. In this section we will 'trust' the tools, but couch conclusions cautiously in the language of 'multiscaling', and focus on the empirical evidence and the soundness of the analysis methodology. In Section 4 we examine the drawbacks of the wavelet based tools themselves, and in the light of these Section 5 reexamines the conclusions.

The traces, detailed in Table 1, include two traces of historical importance mentioned above, **pAug** and **LBL-TCP-3**. We include a trace, **AUCK-d1**, from the Auckland-IV archive collected at the University of Auckland [39]. This repository of high quality traces (timestamp accuracy 100 ns) has become increasingly popular, and therefore useful. From previous published work (for example [40,41]) we know that traces in this archive are very similar to one another in terms of scaling as well as generally, and that the one we choose here is representative. Finally, we include a recent and high rate trace from an Internet backbone, **CAIDA-b1**, also with high timestamp precision, kindly made available by CAIDA from their MFN network [44].

To give a feeling for the different traces, in addition to the average bitrates, the last three columns in Table 1 give j values of particular significance, which can be used in Fig. 1 to better interpret features of the wavelet spectra.

The octave j^{IAT} gives the scale corresponding to the average inter-arrival time between the points (packet arrival times) of X. To the left of this scale, one is really examining in detail residual correlations of individual inter-arrivals, and packet level effects (such as back to back packets due to some bottleneck in the network) which can be extremely complex and of arguable importance for aggregate traffic modelling.

Octave j^* marks the traditional 'knee' of biscaling, which effectively defines the right boundary of the *small scale* regime (so called by ourselves and most others), and separates it from the *large scale* regime, where LRD can be found. Typically all scales below j^* are considered to be in the small scale regime. We look further, and define a *breakdown scale* regime beginning at octave j^{**}. This is not done in the spirit of an endless and arbitrary classification scheme. Any non-trivial scaling behaviour which may exist must break down when the Poisson-like behaviour of the simple point process limit is reached, that is when individual points

Table 1
The traces used

Traces	Date	Start time	Duration (s)	Link	Network	Rate (Mbps)	j^{IAT}	j^{**}	j^*
pAug	1989-08-29	11:25	3142	10BaseT	LAN Frame	0.138	−8.31	N/A	−6.6
LBL-TCP-3	1994-01-20	14:10	7200	10BaseT	WAN TCP	0.35	−7.96	−7	−1
AUCK-d1	2001-04-02	13:00	10 800	OC3	WAN IP	2.5	−9.75	−7	1
CAIDA-b1	2002-08-14	10:00	600	OC48c	Backbone IP	638	−16.7	−12	−1

Last three columns give three j values ($\log_2$(scale)) of interest—j^{IAT}: scale of isolation of individual packets, j^{**}: breakup of small scale scaling region, j^*: biscaling 'knee' with 'small scales' to its left and LRD to the right over 'large scales'.

are seen as isolated, and may do so earlier. Keeping track of this scale helps compare behaviour across different traces. Comparison with j^{IAT} is useful to build intuition about packet patterns at the finest scales.

Each of j^* and j^{**} are empirically, visually, defined, based on observing the LD plots. Whilst this procedure is far from rigorous and certainly not ideal, it is nonetheless, since we employ confidence intervals in a systematic way, at least as thorough as the approaches found in the literature, which tend to ignore issues of goodness of fit.

No analysis should begin without first considering the primordial issue of stationarity. Although the wavelet tools are particularly forgiving with respect to certain kinds of non-stationarities [8] (including additive polynomial trends), they nonetheless assume stationarity of the key statistical features. Numerous studies, of which [8] is a good example, have verified the stationarity of the **pAug** data. We examined each of the other four data sets ourselves by splitting the data into four equal sized blocks, and comparing (informally) estimates of mean, variance, wavelet spectra, and q-LD plots. In each case the stationarity was convincing at 2nd order, including with respect to scaling features. At other orders the stationarity usually appeared satisfactory but was sometimes doubtful, for example in the q-LD with $q = 0.5$ for **LBL-TCP-3**, but not extremely so.

3.1. Scaling analysis

The wavelet spectra from the traces are shown in Fig. 1. The values on the upper axis are calibrated in (approximate) time. In fact they are nothing other than the scale values $a = 2^j$ corresponding to the j values on the lower x-axis. In each plot straight lines can be drawn at large scale to the right of j^* (see Table 1 for values, which are not marked on the plots to avoid prejudicing the eye), corresponding to LRD. The presence of LRD is neither surprising nor controversial, and we will not discuss it further.

Note that we are not concerned in this paper with the interesting issue of the physical meaning of the coexistence of LRD and multifractal scaling. To treat this question correctly would require a discussion of model processes with such properties, whereas this paper focuses firmly on empirical evidence. We simply ask, in an entirely separate way over two empirically identified and non-overlapping scale regimes, if there is credible evidence for multifractal scaling.

In each of the three WAN traces (middle and right plots) biscaling is evident: in addition to LRD to the right of the knee at j^*, to its left a different straight line can be drawn for over 4 octaves (a minimum practical number). This is not true for **pAug**, which only exhibits a single alignment region corresponding to LRD beginning at $j^* = -6.6$. [2] Again, for these traces these findings are not controversial, and are consistent with the literature which has dealt with scales below the LRD range.

For each trace we calculate the spectrum from very small scales, at least as small as the average packet inter-arrival time j^{IAT}, ensuring an ability to compare against the wide range of sampling rates used in the literature. This broader analysis bandwidth enables us to document clearly for the first time (see however Fig. 8 of [33]), the breakdown of the 'small scale' scaling regime. For these traces, we see this occuring at the j^{**} values marked in Table 1. It also gives us the opportunity to note a leveling off of the spectra at very small scales, corresponding to the Poisson-like point process limit (the ordinate of this level corresponds to the inverse of the arrival intensity, which is just equal to j^{IAT} in value). Note that in other traces the breakdown of the small scale scaling regime can be much more violent than here (see Fig. 9 for a hint as to how this could occur).

We now present a multiscale analysis for the three WAN traces. There is no disagreement in the literature over the monofractal character of **pAug** (at scales above j^*, for X or W), and as below j^* we simply observe a rapid convergence to the point process limit, we do not investigate that trace any further.

[2] The energy spike near $j = -1.5$ may surprise some readers, who are familiar with the Bellcore traces and their invariably 'perfect' scaling. This same spike was also seen in [8] for X. Much 'straighter' spectra are found for almost any other time series, in particular W_δ which is the most commonly studied.

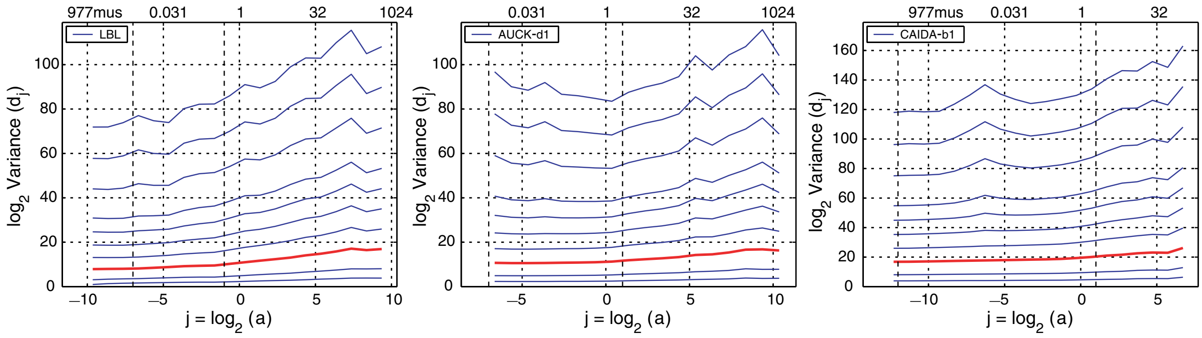

Fig. 2. Set of q-LDs for **LBL-TCP-3** (left), **AUCK-d1** (middle), and **CAIDA-b1** (right). From the bottom up, orders are $q = \{0.5, 1, 2, 3, 4, 5, 6, 8, 10, 12\}$. It is difficult to see from this representation if biscaling identified in the spectrum at $q = 2$ (thicker line) where the cutoff scales j^{**} and j^{*} (vertical lines) were identified, extends to multiscaling over either small or large scales.

For each of the three WAN traces in Fig. 2 the q-LDs for orders $q = \{0.5, 1, 2, 3, 4, 5, 6, 8, 10, 12\}$ are plotted together. This representation is commonly used (for example [29,33]) as it is very compact, and allows a direct comparison between different orders. However it is not practical to include confidence intervals in such a plot, and furthermore the compression of scale makes it very difficult to make clear judgements. For example the detail in the $q = 2$ curves (the thicker line third from the bottom), is virtually impossible to read—compare the same curves plotted in Fig. 1!

We regroup in Table 2 the boundaries of the scaling zones both at small or fine scales (FS), and at large or coarse scales (CS), as determined from the wavelet spectrum. Even after marking these boundaries by vertical lines in the respective plots in Fig. 2, it is not very clear if the conclusion of biscaling made at $q = 2$ is contradicted or not at other orders, that is if multiscaling is present both in the fine scale regime and at coarse scales (recall that *biscaling* is merely a convenient term to refer, for a given q, to the observation of twin scaling regimes).

Table 2
Scales defining twin scaling (biscaling) regimes as measured from the LD (spectrum), and used for estimation in the q-LDs, over fine (FS) and coarse (CS) scales

Trace	Biscaling regimes from spectrum	
	FS	CS
LBL-TCP-3	$[j^{**}, j^{*}] = [-7, -1]$	$[j^{*}, j^{\max}] = [-1, 9]$
AUCK-d1	$[j^{**}, j^{*}] = [-7, 1]$	$[j^{*}, j^{\max}] = [1, 6]$
CAIDA-b1	$[j^{**}, j^{*}] = [-12, -1]$	$[j^{*}, j^{\max}] = [-1, 7]$

To investigate the underlying evidence for multiscaling more carefully, it is necessary to look individually at each qth order Logscale Diagram, and to carefully check that a range of scales exists where scaling is present, separately over the fine and coarse scales. We have done this for each trace, and show the results for **AUCK-d1** in Fig. 3. As detailed in the caption, multiscaling is in fact found over both FS and CS. The evidence for it seems quite convincing.

Having established a case for multiscaling, we proceed to estimate the slopes over the FS and CS scales following Table 2. We must first however determine the range of q values to use. Using tools proposed in [15], an analysis that will not be detailed here indicates that, for the **LBL-TCP-3**, **CAIDA-b1** and **AUCK-d1** time series, the cutoff parameter q_*^+ lies in the range $15 \leqslant q_*^+ \leqslant 20$ for each scale regime, suggesting the range $q \in [0, 15]$. This is mainly good news for the literature on traffic modelling, which has not taken into account the need to estimate below q_*^+ prior to now, and which has tended to use values no higher than 20. It is nonetheless possible that the issue of q_*^+ has biased some published results. In practical terms however, a more pressing problem is that high order moments are notoriously difficult to estimate accurately, and will tend to have poor robustness even to minor non-stationarities. A *possible* example is the spike in Fig. 3 near $j = -3$, which grows steadily with increasing q. Although apparently

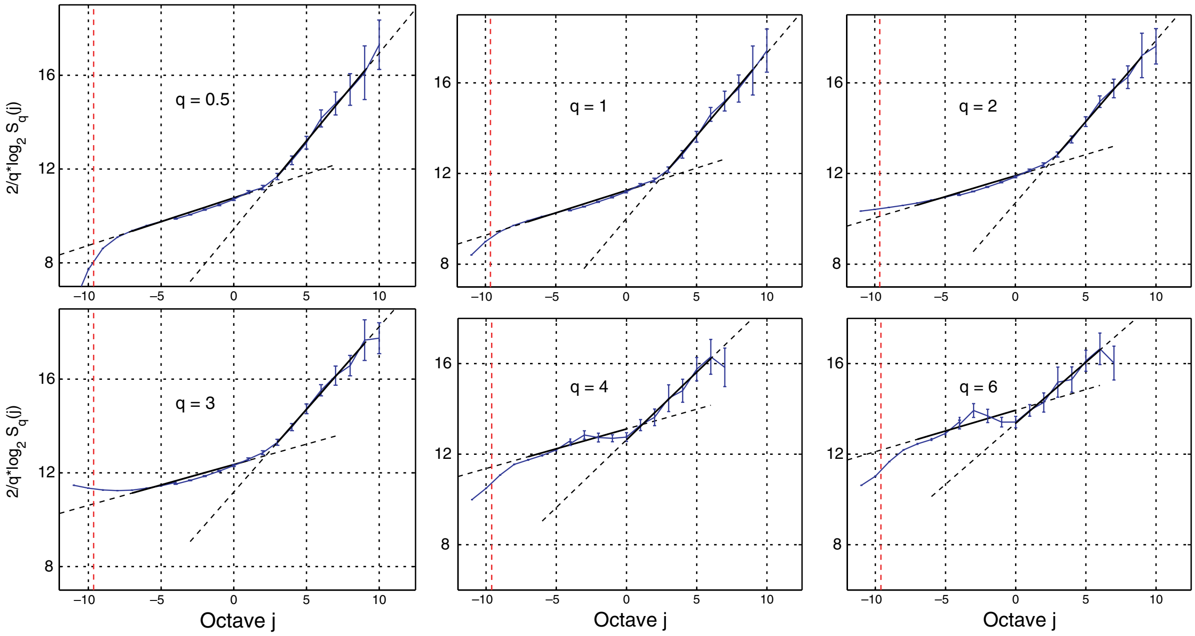

Fig. 3. qth order logscale diagrams for **AUCK-d1**. Plots of $2/q\log_2(S_q(j))$ vs j for $q = [0.5, 1, 2, 3, 4, 6]$ show evidence of twin scaling regimes: at coarse scales (CS), $[j^{**}, j^*] = [1, 6]$, and fine scales (FS), $[j^{**}, j^*] = [-7, 1]$. The data therefore exhibits multiscaling in each scale range. The solid black lines show the fit (and dotted extensions for visualisation). The vertical solid lines show confidence intervals based on data which become quasi-constant at small scale. The vertical dashed line is $j = j^{\mathrm{IAT}}$.

indicating a deviation from the multiscaling model at FS, it could also be due to a non-stationarity whose effect on the estimators becomes more pronounced as order increases. (Complementary analyses not detailed here show that the feature is associated with the repartition of packets within long flows.) As it is not possible to precisely determine the value of q for which estimation, using existing tools, becomes unreliable, we make the conservative choice of $q \in [0, 6]$.

From the estimated slopes, we obtain finally the LMD plots of Fig. 4. From these plots alone, taking into account the confidence intervals therein, we would be obliged to conclude the following. Over the CS regime, for each trace, given the confidence intervals a hypothesis of monofractality

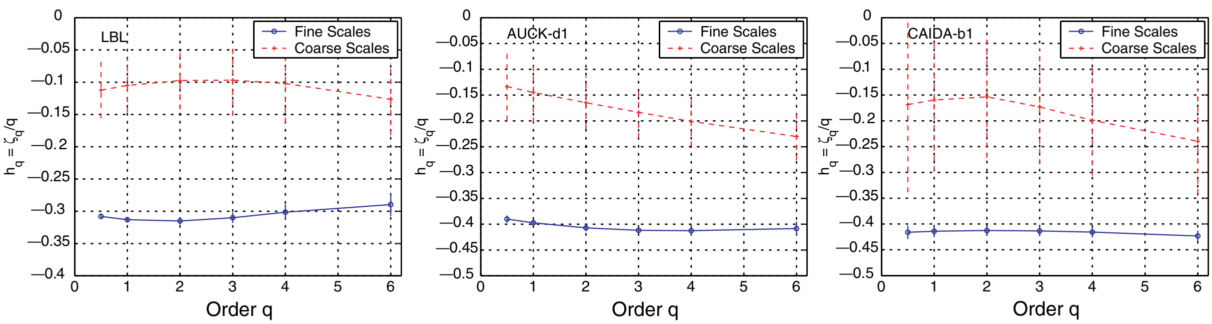

Fig. 4. LMDs for **LBL-TCP-3** (left), **AUCK-d1** (middle), and **CAIDA-b1** (right) over fine scales and coarse scales.

cannot be rejected. Over the FS regime, a careful look at the confidence intervals shows that monofractality cannot be rejected for **CAIDA-b1**, whereas for either of **LBL-TCP-3** or **AUCK-d1** there is only a small deviation from it. It is however marginal, and certainly not convincing evidence for a non-trivial multiscaling consistent with multifractality.

The above conclusions of weak, or no evidence in favour of a multifractal scaling, are summarised in Table 3. They differ from the multifractal conclusions found in the literature at small or large scales, as detailed in Section 2. The principle reason for these disagreements is the use of confidence intervals. By using confidence intervals which reflect the true variations of non-Gaussian data, the LMD plots suddenly paint a far less convincing picture than if one uses very optimistic confidence intervals, or none at all. Another reason is the individual treatment that we give for different q values. Using plots such as those in Fig. 3 makes manifest the many deviations from perfect alignment that exist in the data, and avoids the overly optimistic impression one can obtain from representations like that of Fig. 2.

Table 3
Formal conclusion on scaling at fine (FS) and coarse (CS) scales, using confidence intervals estimated from data

Trace	Biscaling?	MultiScaling?	
		FS	CS
LBL-TCP-3	Yes	Yes: ≈MonoF	Yes: MonoF
AUCK-d1	Yes	Yes: ≈MonoF	Yes: MonoF
CAIDA-b1	Yes	Yes: MonoF	Yes: MonoF

We add two caveats to the summary above. First, the conclusions do in fact agree with those of [43]. However, we disagree with the methodology used. At fine scales a regime only three octaves wide was chosen in [43], which was at some distance from the knee j^*, and no results were given over the scales above and below this range. It is also not clear whether the data based 'non-Gaussian' confidence intervals were used or not. For these reasons their conclusions may not be robust. We return to the question of gaussianity in Section 5.2.

Second, our results do not in fact contradict those of [41], as the *alpha* traffic is only one traffic component. Whilst the overall traffic may not support a multifractal model at large scale, a specific component may. To resolve this, it must be analysed separately. However, we believe that more work is needed to clarify whether the burstiness of *alpha* traffic is best seen as non-stationaries at the time scales considered. We note however that [41] does not claim that *alpha* traffic is multifractal.

We emphasize that the conclusions of Table 3 are *formal*, in the sense of a conscientious but straightforward application of standard wavelet tools and a heuristic approach to goodness of fit based on confidence intervals. In the next section we consider new issues which can potentially revise these conclusions.

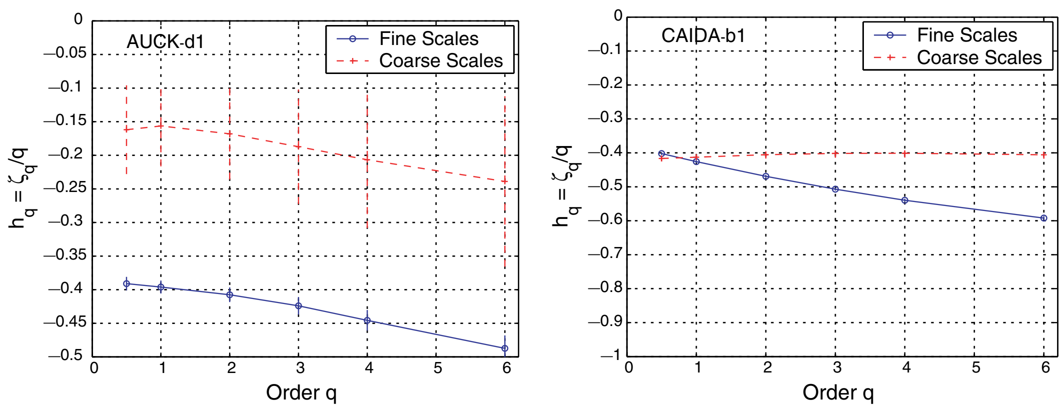

Fig. 5. LMDs for **AUCK-d1** (left), and **CAIDA-b1** (right) over FS and CS, for the inter-arrival time series. A formal conclusion of multifractality applies at FS.

We quickly include some analyses for a different time series, the discrete series of packet inter-arrival times. The q-LD plots (not shown) reveal clear high quality multiscaling, and the corresponding LMD plots in Fig. 5 show monofractal scaling at large scales, but pronounced *multifractal scaling* at small scales. Similar results were found for many other traces. The inter-arrival time process is closely linked to X, in fact they each determine the other. We therefore conjecture that the multifractal behaviour of these two, if present, must be closely linked. It is easier however to deal with X, as time scales can be directly interpreted as 'real'-time. Our final conclusions for the inter-arrival time series therefore follow those for X as laid out below.

4. Statistical limitations

The conclusions made in Section 3 are based on the formal use of an estimator. In the current section we discuss the limitations of that estimator and the associated methodology. It is useful to frame the discussion in terms of the language of hypothesis testing, where the null hypothesis is that there exists a true multifractal scaling. This is particularly instructive when we discuss how the estimator can be fooled in Section 4.2. Our discussion remains qualitative however. No well defined statistical test exists for the detection of multifractal scaling, and it is beyond the scope of this paper to develop such.

4.1. Estimator performance

For second order analysis under Gaussian hypotheses, analytic (approximate) expressions are known [13] for the confidence intervals, and also bias correction factors which account for the logarithmic based analysis (since $\mathbb{E}(\log) \neq \log(\mathbb{E})$). Neither of these are available at qth order and/or in the non-Gaussian context. The estimates at each (q,j) may be biased, as therefore may be the ζ_q, and furthermore it is not known how this bias may vary as a function of signal type. The quality of the variance estimates which underlie the confidence intervals is likewise not precisely known. These limitations of bias and confidence interval determination increase the chance of making an 'error of type I', that is, of incorrectly concluding that multifractal scaling is not present, when in fact it is. Should MF scaling be correctly detected, its characteristics, namely the estimates ζ_q and the multifractal spectrum one could estimate from them, may nonetheless be distorted.

Another drawback is that there is no formal procedure for the choice of cutoff scales defining the scaling regime, nor a well defined goodness of fit statistic which is appropriate to multifractal data and the q-order analysis. Instead these choice are made informally, based on observed alignment 'within the confidence intervals'.

We should at this point mention that other, particularly non-wavelet based, estimators typically have even more drawbacks, and that this area of statistics needs further development. In particular earlier work dealing with multifractal scaling in traffic suffers from the same limitations.

4.2. Analysis power and pseudo-slopes

While the presence of some estimation bias will surprise nobody, the possibility of measuring multiscaling behaviour in a process which does not possess true multifractal scaling is more insidious, and has not been widely discussed. In the hypothesis test language it corresponds to an 'error of type II'. As we show below, signals exist which can give signatures which are difficult, or impossible to distinguish from true multifractal scaling. This weakness can be thought of as a lack of *power* of the estimation procedure, that is, high probability of an error of type II. We give two mechanisms by which this can occur.

The first mechanism arises from discontinuities. The thick grey line in Fig. 6(a) is a single isolated discontinuity, whose LD appears as the grey line in Fig. 6(b). This simple signal possesses a clear scaling behaviour, which reflects the natural invariance of a discontinuity under dilation. The black square wave of Fig. 6(a) shows that we may split the energy of the isolated jump into a number of smaller discontinuities, and still obtain a wavelet spectrum showing scaling, and the

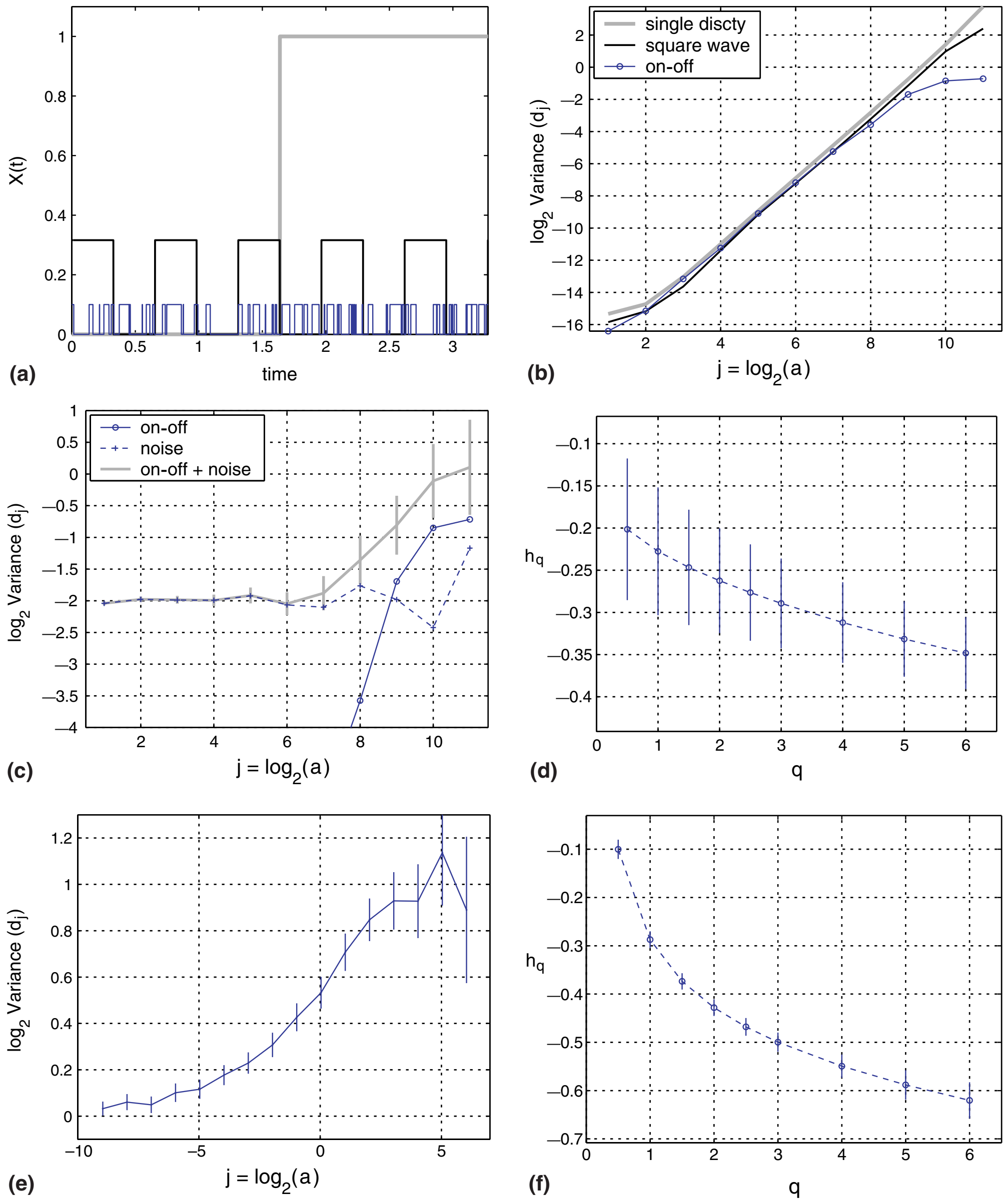

Fig. 6. Lack of power of LD and LMD: two mechanisms. [1] Distortion of scaling due to isolated discontinuities: (a) three signals with discontinuities, (b) corresponding LDs show clear scaling with $\alpha = 2$, (c) adding noise leads to apparent slope with $\alpha < 1$, (d) slopes look convincing over $q \in [1, 6]$, corresponding LMD indicates weak multifractal scaling. [2] Pseudo-scaling due to transition between two levels: (e) transition zone of a Gamma Renewal process exhibits multiscaling, (f) corresponding LMD over $q \in [-4, 0]$ suggests multifractal scaling.

remaining curve in Fig. 6(a) and (b), which is of an *on–off* process with exponentially distributed (and mutually independent) on and off periods, shows that this remains true even after further splitting combined with randomization. Discontinuities can represent the arrival of new, large traffic sources such as alpha flows, or fundamentally non-stationary features. On–off processes can be used to model different traffic sub-processes, such as the arrival of packets within a flow.

In each of the three cases above, the LD does not reveal the isolated nature of the discontinuities generating the scaling, which could therefore be erroneously taken to be evidence of a fractal process. In these examples of 'pure discontinuity' however, the slope in the LDs is a characteristic $\alpha = 2$, which, being beyond $\alpha = 1$ and therefore incompatible with stationarity, is a hint that something is awry. We illustrate in Fig. 6(c) how the $\alpha = 2$ signature can be distorted. We consider the same on–off sample path to which we add white Gaussian noise. [3] The noise increases the energy at all scales, hiding the signature entirely at small scales and lowering the slope at medium scales. On the right, the characteristic short-range dependence of the on–off process begins to have an influence on the larger scales, again reducing the slope. The result is a slope around $\alpha = 0.5$ when measured over $j \in [6, 11]$ in Fig. 6(c), and a corresponding multiscaling behaviour with a LMD shown in Fig. 6(d). The evidence for multifractality in this example seems 'clear but weak' over a relatively narrow range of scales, precisely the situation we find ourselves in for real traffic over small scales.

The second mechanism arises from a transition giving the appearance of alignment. We examine a renewal process with (over dispersed) Gamma distributed inter-arrival times. One can show that the LD of this point process is a monotonically increasing function, which transits between a constant value in the small scale limit to another constant in the limit of large scales. Fig. 6(e) shows the LD of a Gamma renewal process centered over scales in the transition region. As noted also in [7], the transition gives the appearance of alignment in this region, an artifact of what in reality is a smooth cross over between two levels. In fact one observes multiscaling over the range $[-4, 0]$, and would be led to the formal conclusion of multifractality from the corresponding LMD of Fig. 6(f).

Each of the above mechanisms give rise to apparent multiscaling, which does not have true multifractal scaling underlying it. We refer to this phenomenon as *pseudo-scaling*.

5. Adding physical meaning

The main aim of this section is to review the conclusions of Section 3 in the light of the possibilities for error revealed in the previous section. To do so requires the introduction of a new criteria. This will be physical meaning, which we explore through the Poisson cluster process [7], recently introduced as a powerful model of the packet arrival process $X(t)$. A second aim then naturally arises, to use the physical insights which this model brings as a tool to help distinguish real from apparent traffic evolution, and to learn how a fair comparison between different traces can be made.

5.1. A case for multifractality?

Poisson (Barlett Lewis) cluster processes (PCP) are a class of stationary point processes defined as follows [46]: seeds are positioned according to a Poisson process with rate λ_F, and each seed marks the starting point of a cluster which takes the form of a finite renewal process with rate λ_A. The number of points in a cluster is a random variable P. The mapping from the cluster model to a traffic model is very intuitive. The seeds model the arrival of TCP connections as a Poisson process, and the cluster of points associated with a seed correspond to the packets belonging to that connection, or *flow*. The variable P thus denotes the number of packets per connection, and is taken to be heavy tailed (infinite variance) to generate the LRD found in X. Within a connection, packets are spread out according to a renewal process. We choose a Gamma inter-arrival variable, thereby allowing the in-flow burstiness to be controlled by the Gamma shape parameter. Flows are independent and identically distributed.

The significance of the PCP model lies in the fact that its underlying assumptions are not arbitrary, in the sense of black box modelling, but descend directly from real data in a very definite way.

[3] A detailed examination of the effect in LDs of mixing discontinuities with fractional Gaussian noise can be found in [45].

In [7], a series of *semi-experiments* were performed on a number of traces, including **AUCK-d1**, whereby selected aspects of the real data were replaced by neutral model substitutes. The most important experiment involved replacing the true flow arrival process with a simple Poisson process of the same rate, whilst keeping the internal packet structures within each flow intact. Using LDs to compare $X(t)$ before and after, *essentially no change was observed*—a very strong indication that flows can be treated as independent in so far as modelling $X(t)$ is concerned (the analysis was restricted to links of low utilisation, such as those we study here). Other key features of the PCP model were also strongly motivated by physical observations of this semi-experimental type. Therefore, the parameters in the model are meaningful not only in the sense of possessing clear interpretations, but also in the much stronger sense of having verifiable physical meaning, with values which are also meaningful.

Following the procedures outlined in [7], we fitted the PCP model to the **AUCK-d1** dataset, and repeated the multiscaling analysis of Section 3. Fig. 7 shows the q-LD plots, where we notice that, although the PCP model involves no matching of 'other' order moments of any kind, it reproduces the q-LD curves of Fig. 3 remarkably well. The match is not perfect, however the q-LD plots satisfy the multiscaling criteria to a similar level of quality as in the real data. Accordingly, using the same scaling regimes as for **AUCK-d1**, we calculated the LMD, and placed it next to the LMD of **AUCK-d1** in Fig. 8 to favour a direct comparison. The agreement is excellent. Not only are the general shapes comparable, but so are the values, and the formal conclusions regarding multifractality are close to those of Table 3. The difference is that the 'marginally monofractal' conclusion over FS from the data becomes 'weakly multifractal' in the case of the model, mainly since confidence intervals are smaller.

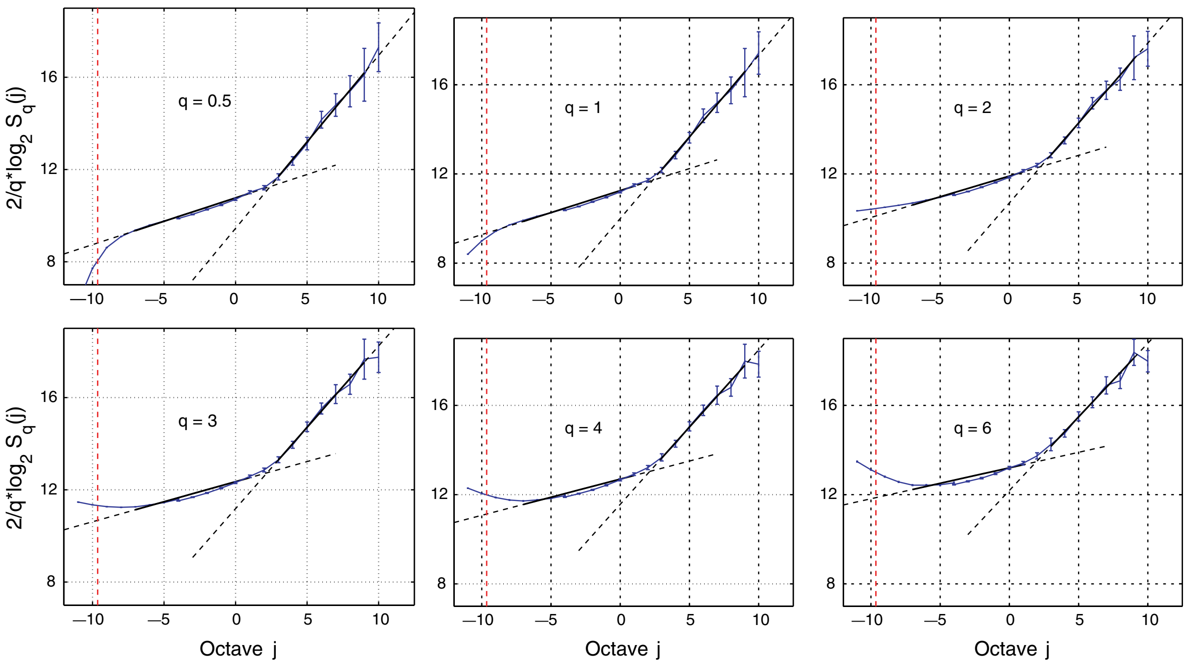

Fig. 7. qth order logscale diagrams for the PCP model fitted to **AUCK-d1**. The details of the plots are exactly as for Fig. 3. The curves show remarkable agreement with the corresponding ones for the data, given that the model was not calibrated to match higher moments. Multiscaling of comparable quality is found over the same scaling regimes (the barely visible small jumps at $j = -4$ are due to a separate calculation for the very small scales, performed over only a subset of the data for computational reasons).

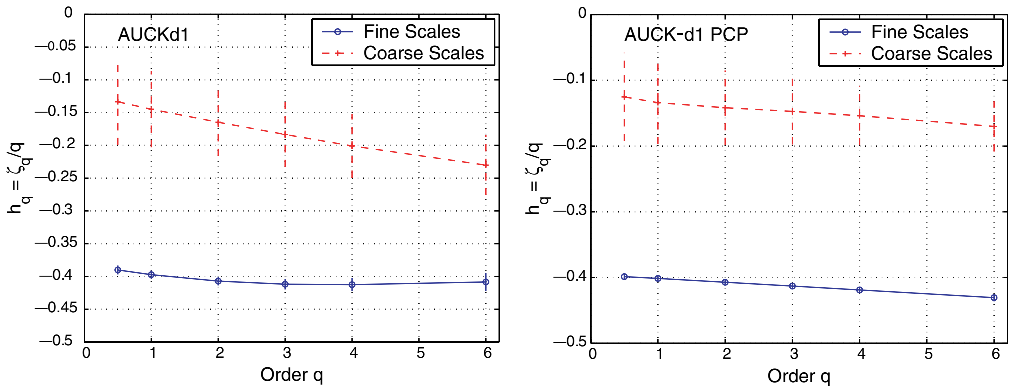

Fig. 8. Multiscaling comparison between **AUCK-d1** and the fitted PCP model. The left hand plot is a copy from Section 3. In both the fine and coarse scale regimes the multiscaling signatures are very similar, and similar formal conclusions follow: monofractal at CS, 'marginally monofractal' or 'weakly multifractal' at FS. However, the PCP is *not* multifractal.

Although the close agreement between the data and the model is very satisfying, the point we wish to make here is something quite different. The PCP model is **not** multifractal, and yet, it reproduced a non-trivial multiscaling behaviour (at least to the same extent as the data). It therefore provides another, rather powerful, example of pseudo-scaling. In fact, it is shown in [7] that the PCP model, at small scales, is closely related to the Gamma renewal example shown in Section 4, and the pseudo-scaling arises from the transition effect (in [7] this was shown for the LD, here we extend this through other orders to multiscaling).

From the above, we see that pseudo-scaling is not only a possible explanation for the weak evidence found in Section 3, it actually *is* responsible for the empirical scaling within a model of that data which has a strong physical foundation. If we accept this model as preferable on physical grounds, which we do here, then we are led to conclude that the evidence for the multiscaling itself (whether it be monofractal or not) is misleading. It is to be interpreted in fact as a pernicious example of the pseudo-scaling described in Section 4.

Table 4 summarises our final conclusion regarding scaling from the point of view just discussed. The verdicts at fine scales have changed character, from monofractal or weak multifractal multiscaling, to no true scaling at all, effectively due to a lack of power on the part of the statistical procedure. The conclusions at coarse scales are unaffected.

Before leaving this section, we reiterate that we have not demonstrated in a formal sense that the data is not multifractal, nor have we formally demonstrated a lack of power on the part of the estimator. Instead, our main point is that, given the limitations of the current tools, and the weak nature of the empirical evidence, there is no case for concluding *for* multifractality, particularly when there are conceptually simpler models which can account for the empirical observations equally well, as far as the statistical tools are concerned.

5.2. Traffic evolution, insights and misconceptions

In this subsection we use the cluster model to guide a discussion on how different traces can be compared. This is important when one wishes to address questions such as whether evidence for scaling is changing or not over time.

Table 4
Physically interpreted conclusions on scaling at fine (FS) and coarse (CS) scales, using confidence intervals estimated from data, and insights from the Poisson cluster model and underlying semi-experiments

Trace	Biscaling?	MultiScaling?	
		FS	CS
LBL-TCP-3	Yes	No: pseudo-scaling	Yes: MonoF
AUCK-d1	Yes	No: pseudo-scaling	Yes: MonoF
CAIDA-b1	Yes	No: pseudo-scaling	Yes: MonoF

It has been recently argued in [43] that Internet backbone traffic can no longer be multifractal since its rate is now so high that it has Gaussian marginals even at 'small' scales of the order of 1 ms. While it is true that backbone traffic marginals may seem close to Gaussian at such scales, they can never be exactly Gaussian and therefore conclusions cannot be so readily drawn. To see this consider the following argument. The independence of flows inherent in the cluster model implies the important fact that the parameter λ_F, the arrival rate of flows, plays the role of the 'amount' of traffic. Doubling λ_F simply yields 'twice as much of the same thing', and corresponds, in terms of the wavelet spectrum, to a simple addition of 1. Thus, varying λ_F simply shifts the wavelet spectrum vertically: it does not change the underlying structure of traffic dependencies (in particular it leaves correlations invariant), but it of course does changes the nature of the marginals of X_δ, making them closer to or further from Gaussian.

Similarly, backbone traffic has been said to tend to a Poisson process with increasing traffic rate [47]. While it is true that the distribution of inter-arrival times tends to exponential as λ_F increases, the inter-arrivals remain correlated both in data and within the cluster model. In the latter this can be seen by the fact that the wavelet spectrum is simply shifted. This contradicts the necessary assumption of independent inter-arrival times of a Poisson process. Indeed, what is at work here is nothing other than the smoothing effect of multiplexing gain, which speaks of the relative size of traffic variability, but not of its nature. One must be careful of the subtle fact that when examining inter-arrival times as traffic rates increase, one is in fact shifting the focus of observation to smaller and smaller scales.

The very high rates of backbone traces make it difficult to compare today's traffic with what it was a decade ago. Although much has changed, biscaling is still present, and firm conclusions about traffic evolution cannot be made unless it is clear what constitutes a fair comparison. A key element of this is the meaning of terms such as 'large' or 'small' scales. A priori, these could be defined in terms of link bandwidth, flow rate, distance from Gaussianity in some sense, or via phenomenological criteria such as the scale of an observed 'knee'.

Whilst, in the light of the above discussion, one can compare the wavelet spectra of the CAIDA and LBL traces by adjusting for λ_F, there are other issues. A comparison should also be 'fair' in terms of marginals. However one cannot simply zoom in on a trace until the marginals look the same for the two, for at least two reasons. First, any fair comparison should be made in analogous regimes: comparing the LRD behaviour of LBL with what happens before the onset of LRD in CAIDA does not make any sense. Another important reason is that one cannot safely zoom in indefinitely. Fig. 9 shows the LD and an averaged periodogram of the very small scale regime of the **CAIDA-b1** trace (the two can be linked by reinterpreting Eq. (3) as a spectral estimator, and setting $\nu = 1/a$). The Fourier analysis reveals periodicities in the packet arrival process at scales $j \leqslant 10$ due to physical network effects, such as back to back packets on upstream bottleneck links, which translate to shaped, roughly periodic traffic on the observed link (the wavelet analysis averages these out and leads to a roughly flat spectrum consistent with a Poisson process). It would be unfair to compare a scale where the physical layer is very strong, such as

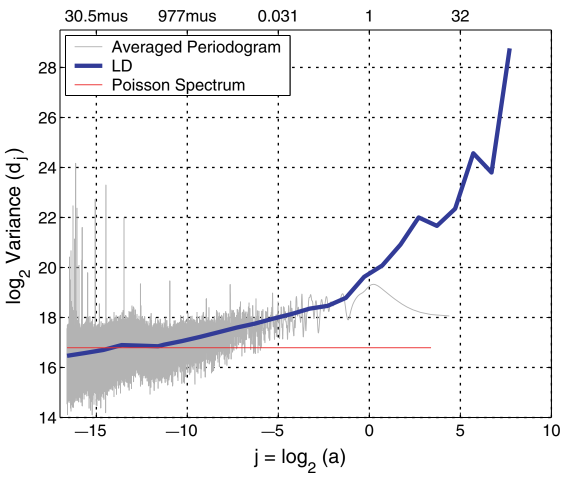

Fig. 9. Periodicities at small scale. LD and averaged (15 windows each with 50% overlap) periodogram of **CAIDA-b1**. The Fourier analysis reveals periodicities in the packet arrival process due to the physical network effects.

around $j = -15$ for **CAIDA-b1**, with one where it is not.

One can make further use of the PCP model for meaningful comparisons. For instance a random flow based thinning of a high rate trace could be performed to equalise λ_F whilst preserving the auto-correlation structure and, in principle, generating a new marginal which could be fairly compared to that of LBL.

Moving beyond traffic volume and marginals, fairness with respect to scaling regime can be accessed through comparing the positions of the knee scale j^* in the light of the formulae given in [7]. In this way, the PCP model can be used not only to clarify the issue of multifractality in traffic but also to track traffic evolution in a way that does not ignore the undeniable presence of empirical multiscaling.

6. Conclusion

This paper investigated in depth the case against multifractal behaviour in TCP/IP traffic, beginning with a careful literature review showing that inconsistencies and confusion exist in the prevailing views, justifying a new look at the evidence and its interpretation.

Using wavelet based estimation methods, we provided a thorough analysis of the process of packet arrivals for four representative traces. In terms of scaling regimes, we defined two, one at 'large' and one at 'small' scales, and explored for the first time the breakdown of the latter at very small scales. In terms of scaling beyond second order we clearly separated out the empirical concept of 'multiscaling' from the much stronger conclusion of multifractality or monofractality. Our formal conclusions, in terms of the careful use of the available analysis tools, were that multiscaling exists at both large and small scales, with monofractal and long-range dependent behaviour at large scale, in agreement with most, but not all, of the literature. In contradiction with most of the literature, over small scales we concluded monofractal (or very marginal multifractal) behaviour. The difference is due mainly to a more rigorous use of confidence intervals.

We then examined the wavelet tools themselves, detailed their weaknesses, and showed how they could be fooled. Most importantly, data which is not multifractal can nonetheless display multiscaling features indistinguishable from genuine monofractal or multifractal scaling, in effect revealing a lack of power, in the sense of informal hypothesis testing, of the statistical methodology. We emphasized this point by showing that a Poisson cluster model, with very strong data based physical justification, could successfully model the higher order statistics of the data, and even its multiscaling wavelet signatures, even though it is not multifractal. This led to a revised conclusion at small scale, that the observed multiscaling does not have a physical basis but is best regarded in effect as a misinterpretation, a *pseudo-scaling*. We emphasized that these conclusions, although persuasive, cannot be entirely definitive, as they are linked to the philosophical issue of model choice, as well as limited by the available statistical machinery. Finally the cluster model was used to clarify misconceptions regarding the evolution of traffic and its relation to gaussianity and the nature of fair comparisons across scale.

Acknowledgments

We thank CAIDA for kindly making their MFN data available. This work was supported by the Australian Research Council. This work was partially supported by the French MENRT grant ACI Jeune Chercheur 2329, 1999.

References

[1] Will E. Leland, Murad S. Taqqu, Walter Willinger, Daniel V. Wilson, On the self-similar nature of Ethernet traffic (extended version), IEEE/ACM Transactions on Networking 2 (1) (1994) 1–15.
[2] Mark E. Crovella, Azer Bestavros, Self-similarity in world wide web traffic: evidence and possible causes, IEEE/ACM Transactions on Networking 6 (5) (1997) 835–846.
[3] K. Park, W. Willinger (Eds.), Self-Similar Network Traffic and Performance Evaluation, Wiley, New York, 2000.
[4] R.H. Riedi, M.S. Crouse, V.J. Ribeiro, R.G. Baraniuk, A multifractal wavelet model with application to network traffic, IEEE Transactions on Information Theory, special

issue "Multiscale Statistical Signal Analysis and its Applications" 45 (3) (1999) 992–1018.

[5] Kant, On aggregate traffic generation with multifractal properties, in: Global Telecommunications Conference, 1999, GLOBECOM'99, 1999, pp. 1179–1183.

[6] A. Erramilli, O. Narayan, A. Neidhardt, I. Saniee, Performance impacts of multi-scaling in wide area TCP/IP traffic, in: Proceedings of IEEE Infocom'2000, Tel Aviv, Israel, March 2000.

[7] Nicolas Hohn, Darryl Veitch, Patrice Abry, Cluster processes, a natural language for network traffic, IEEE Transactions on Signal Processing, special issue "Signal Processing in Networking" 51 (8) (2003) 2229–2244.

[8] P. Abry, D. Veitch, Wavelet analysis of long-range dependent traffic, IEEE Transactions on Information Theory 44 (1) (1998) 2–15.

[9] D. Veitch, P. Abry, Matlab code for the wavelet based analysis of scaling processes. Available from <http://www.cubinlab.ee.mu.oz.au/~darryl/>.

[10] P. Abry, P. Flandrin, M.S. Taqqu, D. Veitch, Wavelets for the analysis, estimation, and synthesis of scaling data, in: K. Park, W. Willinger (Eds.), Self-Similar Network Traffic and Performance Evaluation, Wiley, New York, 2000, pp. 39–88.

[11] P. Abry, R. Baraniuk, P. Flandrin, R. Riedi, D. Veitch, The multiscale nature of network traffic: discovery, analysis, and modelling, IEEE Signal Processing Magazine, special issue "Analysis and Modeling of High-Speed Data Network Traffic" 19 (3) (2002) 28–46.

[12] R.H. Riedi, Multifractal processes, in: P. Doukhan, G. Oppenheim, M.S. Taqqu (Eds.), Theory and Applications of Long-Range Dependence, Birkhäuser, Basel, 2002, pp. 625–715.

[13] Darryl Veitch, Patrice Abry, A wavelet based joint estimator of the parameters of long-range dependence, IEEE Transactions on Information Theory, special issue "Multiscale Statistical Signal Analysis and its Applications" 45 (3) (1999) 878–897.

[14] Darryl Veitch, Patrice Abry, A statistical test for the time constancy of scaling exponents, IEEE Transactions on Signal Processing 49 (10) (2001) 2325–2334.

[15] B. Lashermes, P. Abry, P. Chainais, Scaling exponents estimation for multiscaling processes, in: Int. Conf. on Acoustics and Signal Proc., Montreal, Canada, 2004.

[16] B. Lashermes, P. Abry, P. Chainais, New insights into the estimation of scaling exponents, International Journal of Wavelets, Multiresolution and Information Processing 2 (4) (2004) 497–523.

[17] B.B. Mandelbrot, Intermittent turbulence in self similar cascades: divergence of high moments and dimension of the carrier, Journal of Fluid Mechanics 62 (1974) 331.

[18] M. Ossiander, E.C. Waymire, Statistical estimation for multiplicative cascades, The Annals of Statistics 28 (6) (2000) 1533–1560.

[19] J. Barral, B. Mandelbrot, Multiplicative products of cylindrical pulses, Probability Theory and Related Fields 124 (2002) 409–430.

[20] E. Bacry, J. Delour, J.F. Muzy, Multifractal random walk, Physical Review E 64 (2001) 026103.

[21] E. Bacry, J.F. Muzy, Log-infinitely divisible multifractal processes, Communications in Mathematical Physics 236 (2003) 449–475.

[22] P. Chainais, R. Riedi, P. Abry, On non scale invariant infinitely divisible cascades, in: Actes du GRETSI, 2003.

[23] P. Chainais, R. Riedi, P. Abry, On non scale invariant infinitely divisible cascades, 2003, submitted for publication.

[24] B.B. Mandelbrot, A multifractal walk down wall street, Scientific American 280 (2) (1999) 70–73.

[25] P. Chainais, Cascades log-infiniment divisibles et analyse multirésolution. Application à l'étude des intermittences en turbulence, Ph.D. thesis, ENS Lyon, 2001.

[26] N. Desaulniers-Soucy, A. Iuoras, Traffic modeling with universal multifractals, in: Proceedings of Globecom'99, 1999.

[27] Lawrence Berkeley National Laboratory (LBL), Internet traffic archive, `http://ita.ee.lbl.gov/index.html`.

[28] M.S. Taqqu, V. Teverovsky, W. Willinger, Is network traffic self-similar or multifractal? Fractals (5) (1997) 63–73.

[29] R.H. Riedi, J. Lévy Véhel, Tcp traffic is multifractal: a numerical study, Tech. Rep. 3129, INRIA Rocquencourt, March 1997.

[30] V. Paxson, S. Floyd, Wide-area traffic: the failure of Poisson modeling, IEEE/ACM Transactions on Networking 3 (3) (1994) 226–244.

[31] J. Lévy Véhel, R.H. Riedi, in: J. Lévy Véhel, E. Lutton, C. Tricot (Eds.), Fractals in Engineering'97, Springer, Berlin, 1997, Chapter: Fractional Brownian motion and data traffic modeling: the other end of the spectrum.

[32] P. Mannersalo, I. Norros, Multifractal analysis of real ATM traffic: a first look, COST257TD(97)19 version 1, VTT Information Technology, 21 January 1997.

[33] A. Feldmann, A.C. Gilbert, W. Willinger, T.G. Kurtz, The changing nature of network traffic: scaling phenomena, Computer Communication Review (April) (1998).

[34] A. Feldmann, A.C. Gilbert, W. Willinger, T.G. Kurtz, Looking behind and beyond self-similarity: on scaling phenomenon in measured WAN traffic, in: Proceedings of the 35th Allerton Conference on Communication, Control and Computing, October 1997.

[35] Anja Feldmann, Anna Gilbert, Walter Willinger, Data networks as cascades: investigating the multifractal nature of Internet WAN traffic, in: ACM/Sigcomm'98, Vancouver, Canada, 1998.

[36] A. Feldmann, A.C. Gilbert, P. Huang, W. Willinger, Dynamics of IP traffic: a study of the role of variability and the impact of control, in: Proceedings of the ACM SIGCOMM 99, Boston, MA, 1999, pp. 301–313.

[37] D. Veitch, P. Abry, P. Flandrin, P. Chainais, Infinitely divisible cascade analysis of network traffic data, in: Proceedings of ICASSP 2000, Istanbul, Turkey, June 2000.

[38] Darryl Veitch, Patrick Flandrin, Patrice Abry, Infinitely divisible cascade analysis of network traffic data, in: Institute for Mathematics and its Applications (IMA) 'Hot Topics' Workshop on Scaling Phenomena, Minneapolis, October 1999.
[39] Waikato Applied Network Dynamics. Available from <http://wand.cs.waikato.ac.nz/wand/wits/>.
[40] S. Roux, D. Veitch, P. Abry, L. Huang, P. Flandrin, J. Micheel, Statistical scaling analysis of TCP/IP data, in: ICASSP 2001, special session, Network Inference and Traffic Modeling, Salt Lake City, Utah, 7–11 May 2001.
[41] S. Sarvotham, R. Riedi, R. Baraniuk, Connection-level analysis and modeling of network traffic, in: Proceedings of the ACM SIGCOMM Internet Measurement Workshop, 2001.
[42] S. Sarvotham, R. Riedi, R. Baraniuk, Connection-level analysis and modeling of network traffic, Tech. Rep., ECE Dept., Rice University, Houston, 30 June 2001.
[43] Z.-L. Zhang, V. Ribeiro, S. Moon, C. Diot, Small-time scaling behaviors of internet backbone traffic: an empirical study, in: Proceedings of IEEE Infocom, San Francisco, March 2003, to appear.
[44] Cooperative Association for Internet Data Analysis. Available from <http://www.caida.org/analysis/workload/oc48/>.
[45] Matthew Roughan, Darryl Veitch, Measuring long-range dependence under changing traffic conditions, in: Annual Joint Conference of the IEEE Computer and Communications Societies (Infocom'99), Manhattan, NY, March 1999, IEEE Computer Society Press, Los Alamitos, CA, 1999, pp. 1513–1521.
[46] D.J. Daley, D. Vere-Jones, An Introduction to the Theory of Point Processes, Springer, Berlin, 1988.
[47] Jin Cao, William S. Cleveland, Dong Lin, Don X. Sun, in: C. Holmes, D. Denison, M. Hansen, B. Yu, B. Mallick (Eds.), Nonlinear Estimation and Classification, Springer, New York, 2002, Chapter: Internet traffic tends toward poisson and independent as the load increases.

Darryl Veitch was born in Melbourne, Australia where he completed a Bachelor of Science Honours degree at Monash University in 1985. His doctorate is in mathematics from the University of Cambridge, England, graduating in 1990. In 1991 he joined the research laboratories of Telstra in Melbourne where he became interested in long-range dependence of tele-traffic. He spent 1 year at the CNET of France Telecom in Paris, and then held visiting positions at the KTH in Stockholm, INRIA Sophia-Antipolis in France, and Bellcore in New Jersey, before returning to acadmia at RMIT, Melbourne. In 2000 he then joined the Electrical and Electronic Engineering department at the University of Melbourne where he directed the EMULab, an Ericsson funded networking research group. He is now a principal research fellow in the ARC Special Research Centre for Ultra-Broadband Information Networks (CUBIN) in the department, and a project leader in NICTA (National ICT Australia). His research interests include the statistical and dynamic nature of Internet traffic, parameter estimation problems and queueing theory for traffic processes, the theory and practice of active measurement, and traffic inversion and sampling. He is a prominent member of the Internet measurement community, and is the IMC 2005 co-chair. He is a senior member of the IEEE.

Nicolas Hohn received the Ingénieur degree in electrical engineering in 1999 from Ecole Nationale Supérieure d'Electronique et de Radio-électricité de Grenoble, Institut National Polytechnique de Grenoble, France. In 2000 he received a M.Sc. degree in biophysics from the University of Melbourne, Australia, while working for the Bionic Ear Institute on signal processing techniques in auditory neurons. In 2004 he completed a Ph.D. in the Department of Electrical Engineering at the University of Melbourne as a member of the ARC Special Research Centre for Ultra-Broadband Information Networks (CUBIN). His thesis was on the statistical nature of Internet traffic, traffic sampling and inversion techniques, as well as modelling of Internet routers. He received the best student paper award at the 2003 ACM Internet Measurement Conference. He is currently working as a researcher in a hedge fund in Sydney, Australia.

Patrice Abry was born in Bourg-en-Bresse, France in 1966.

He received the degree of Professeur-Agrégé de Sciences Physiques, in 1989 at Ecole Normale Supérieure de Cachan and completed a PhD in Physics and Signal Processing, at Ecole Normale Supérieure de Lyon and Université Claude-Bernard Lyon I, in 1994.

Since October 1995, he is a permanent CNRS researcher, at the laboratoire de Physique of Ecole Normale Supérieure de Lyon.

He received the AFCET-MESR-CNRS prize for best PhD in Signal Processing for the years 1993–1994 and is the author of the book "Ondelettes et Turbulences—Multirésolution, algorithmes de décompositions, invariance d'échelle et signaux de pression", published in October 1997, by Diderot, éditeur des Sciences et des Arts, Paris, France. He also is the coeditor of the book in French entitled "Lois d'échelle, Fractales et Ondelettes", Hermès, Paris, France, 2002.

His current research interests include wavelet-based analysis and modelling of scaling phenomena and related topics (self-similarity, stable processes, multi-fractal, $1/f$ processes, long-range dependence, local regularity of processes, infinitely divisible cascades, departures from exact scale invariance...). Hydrodynamic turbulence and the analysis and modelling of computer network teletraffic are the main applications under current investigation. He is also involved in the study of baroreflex sensitivity with a French medical group at University Claude Bernard Lyon I. He recently started a wavelet based detection/analysis of Acoustic Gravity Waves in Ionosphere.