

SMALL AND LARGE SCALE BEHAVIOR OF MOMENTS OF POISSON CLUSTER PROCESSES

NELSON ANTUNES¹, VLADAS PIPIRAS², PATRICE ABRY³ AND DARRYL VEITCH⁴

Abstract. Poisson cluster processes are special point processes that find use in modeling Internet traffic, neural spike trains, computer failure times and other real-life phenomena. The focus of this work is on the various moments and cumulants of Poisson cluster processes, and specifically on their behavior at small and large scales. Under suitable assumptions motivated by the multiscale behavior of Internet traffic, it is shown that all these various quantities satisfy scale free (scaling) relations at both small and large scales. Only some of these relations turn out to carry information about salient model parameters of interest, and consequently can be used in the inference of the scaling behavior of Poisson cluster processes. At large scales, the derived results complement those available in the literature on the distributional convergence of normalized Poisson cluster processes, and also bring forward a more practical interpretation of the so-called slow and fast growth regimes. Finally, the results are applied to a real data trace from Internet traffic.

Mathematics Subject Classification. 60G55, 60G18, 60K30, 60G22.

Received November 24, 2016. Accepted September 27, 2017.

1. INTRODUCTION

A Poisson cluster process (PCP, for short; sometimes also called cluster Poisson process or CPP) consists of points usually defined on the positive half-axis $(0, \infty)$ whose positions are determined by the following construction. Clusters of a finite number of points are assumed to arrive according to a Poisson arrival process with intensity $\lambda > 0$ at times S_j , $j \geq 1$ (with $0 < S_1 < S_2 < \dots$). The clusters are i.i.d. copies with a random but almost surely finite number of points W_j . The focus throughout is on clusters having the following structure: the W_j points are separated in time by i.i.d. sequence of positive interarrival times $A_{j,k}$, $k \geq 1$, and the first point of a cluster is located at the arrival time of the cluster. Such PCPs are also known as the Bartlett–Lewis processes after Bartlett [4] and Lewis [23] (see, for example, Cox and Isham [6], Daley and Vere–Jones [7]).

Keywords and phrases. Poisson cluster process, scaling, moments, cumulants, heavy tails, slow growth regime, fast growth regime, Internet traffic modeling.

¹ Center for Computational and Stochastic Mathematics, University of Lisbon, and University of Algarve, Campus de Gambelas, 8005–139 Faro, Portugal. nantunes@ualg.pt

² Department of Statistics and Operations Research, University of North Carolina, CB 3260, Chapel Hill, NC 27599, USA. pipiras@email.unc.edu

³ Univ Lyon, Ens de Lyon, Univ Claude Bernard, CNRS, Laboratoire de Physique, 69342 Lyon, France. patrice.abry@ens-lyon.fr

⁴ School of Computing and Communications, University of Technology Sydney, P.O. Box 123, Broadway, NSW 2007, Australia. Darryl.Veitch@uts.edu.au

In mathematical terms, if $N(B)$ denotes the number of such PCP points in a set $B \subset (0, \infty)$, then

$$N(B) = \sum_{j=1}^{\infty} \sum_{k=0}^{W_j-1} 1_B \left(S_j + \sum_{m=1}^k A_{j,m} \right), \quad (1.1)$$

where $1_B(x)$ is the indicator function of the set B . The PCP N defined by (1.1) is called *transient* (that is, nonstationary), since the distributions of $N(B)$ and $N(B+T)$ are in general not equal for $T > 0$ and $B \subset (0, \infty)$, where $B+T = \{x+T : x \in B\}$. The *equilibrium* PCP $N_e(B)$ is defined as $N(B+T)$ letting $T \rightarrow \infty$. For the equilibrium process, the distributions of $N_e(B)$ and $N_e(B+h)$ are the same for any $h > 0$ and $B \subset (0, \infty)$. The equilibrium process N_e can be viewed as stationary, and will be the focus throughout this work.

PCPs form an interesting class of point processes which has been studied in theory (*e.g.* in the general context of point process; see Cox and Isham [6], Karr [20], Daley and Vere–Jones [7]) and used successfully in applications (*e.g.* computer failure patterns in Lewis [23], software reliability in Zeephongsekul *et al.* [36], neural spike trains in Grüneis *et al.* [14, 15], physics in Saleh and Teich [31], Lowen and Teich [24], rainfall in Onof *et al.* [28]). The motivating application in this work is the Internet traffic observed on a network link, where points are data packets and clusters are packet flows (essentially packetized document files, web pages, videos or other application contents). The use of PCPs in modeling data packet traffic was popularized by Hohn *et al.* [18], see also Faÿ *et al.* [11], Mikosch and Samorodnitsky [27], Fasen and Samorodnitsky [10], González-Arévalo and Roy [13], Antunes and Pipiras [3]. Models related to PCPs for modeling Internet traffic include the ON/OFF model (*e.g.* Leland *et al.* [22]), the infinite source Poisson arrival process (*e.g.* Mikosch *et al.* [26], Guerin *et al.* [16]), and the renewal point process (*e.g.* Kaj [19], Gaigalas and Kaj [12]).

In this work, we focus on the moments and cumulants of PCPs. On the one hand, moments and cumulants are among the most basic and fundamental quantities of any random object of study and, in fact, have already been studied for PCPs to some extent (see references in Sect. 2 below). We are particularly interested here in their scaling behavior at large (coarse) and small (fine) scales, especially in connection to the use of PCP models motivated by the “self-similar” and multiscale nature of Internet traffic (*e.g.* Abry *et al.* [1], Hohn *et al.* [18]).

More specifically, we will consider the following moments of PCPs: for integer $r \geq 1$,

$$\text{(usual) moments : } m_r(a) = \mathbb{E}N_e(0, a)^r, \quad (1.2)$$

$$\text{factorial moments : } m_{[r]}(a) = \mathbb{E}N_e(0, a)^{[r]}, \quad (1.3)$$

$$\text{central moments : } m_r^0(a) = \mathbb{E}(N_e(0, a) - \mathbb{E}N_e(0, a))^r, \quad (1.4)$$

where $n^{[r]} = n(n-1) \dots (n-r+1)$ for a nonnegative integer n and $a > 0$ will be referred to as “scale.” Central moments are natural to consider in view of some of the large scale limiting results available for centered PCPs (see (5.15) and (5.16) below). Factorial moments are considered because, as will be shown, they may be more informative about PCPs than the usual or central moments.

The quantities most convenient to work with in the context of PCPs are not any of the moments above but rather factorial cumulants. Moreover, the (usual) cumulants are often considered in practice, in addition to the (usual) moments. We will thus also consider: for integer $r \geq 1$,

$$\text{(usual) cumulants : } \kappa_r(a) = \left. \frac{\partial^r \log M_a(t)}{\partial t^r} \right|_{t=0}, \quad (1.5)$$

$$\text{factorial cumulants : } \kappa_{[r]}(a) = \left. \frac{\partial^r \log P_a(z)}{\partial z^r} \right|_{z=1}, \quad (1.6)$$

where $P_a(z)$ is the probability generating function of the equilibrium PCP on the interval $(0, a)$ (see Sect. 2 for definition) and $M_a(t)$ is the moment generating function of the equilibrium PCP on the interval $(0, a)$. In fact, the results of interest will be derived first for factorial cumulants $\kappa_{[r]}(a)$ and then used to obtain analogous results for the remaining quantities (1.2)–(1.5).

TABLE 1. The asymptotic behavior of the various moments and cumulants ($r \geq 2$).

Asymptotic behavior	Small scales ($a \rightarrow 0^+$)	Large scales ($a \rightarrow \infty$)
cumulants $\kappa_r(a)$	if $a \ll a_{\kappa,r}$: $\sim c_{\kappa,r} \lambda a$	if $a_{\kappa,r} \ll a$: (*) $\sim C_{\kappa,r} \lambda a^{r-\alpha+1}$
factorial cumulants $\kappa_{[r]}(a)$	if $a \ll a_{\kappa,[r]}$: $\sim c_{\kappa,[r]} \lambda a^{1+(r-1)\theta}$	if $a_{\kappa,[r]} \ll a$: $\sim C_{\kappa,[r]} \lambda a^{r-\alpha+1}$
moments $m_r(a)$	if $a \ll a_{m,r}$: $\sim c_{m,r} \lambda a$	if $a_{m,r} \ll a$: $\sim C_{m,r} \lambda^r a^r$
factorial moments $m_{[r]}(a)$	if $a \ll a_{m,[r]}$: $\sim c_{m,[r]} \lambda a^{1+(r-1)\theta}$ ($\theta < 1$) $\sim c_{m,[r]} \lambda^r a^r$ ($\theta > 1$)	if $a_{m,[r]} \ll a$: $\sim C_{m,[r]} \lambda^r a^r$
central moments $m_r^0(a)$	if $a \ll a_{m,r,1}^0$: $\sim c_{m,r}^0 \lambda a$	if $a_{m,r,1}^0 \ll a \ll a_{m,r,2}^0$: (fast growth*) $\sim C_{m,r,1}^0 \lambda^{r/2} a^{(3-\alpha)r/2}$, for even r $\sim C_{m,r,2}^0 \lambda^{(r-1)/2} a^{(3-\alpha)(r-1)/2+1}$, for odd r if $a_{m,r,2}^0 \ll a$: (slow growth*) $\sim C_{m,r,3}^0 \lambda a^{r-\alpha+1}$

Small and large scale behaviors of the quantities (1.2)–(1.4) and (1.5)–(1.6) refer, respectively, to $a \rightarrow 0^+$ and $a \rightarrow \infty$. As indicated above, in connection to Internet traffic and especially its self-similar (multiscale) nature, it has become common and useful to examine various quantities as analysis scale changes. In our study, we shall make the following assumptions motivated by the applications to Internet traffic. At large scales ($a \rightarrow \infty$), we shall assume, in particular, that the cluster size distribution of $W = W_j$ is heavy-tailed with exponent $\alpha \in (1, 2)$ in the sense that

$$\mathbb{P}(W > w) \sim C_W w^{-\alpha}, \quad \text{as } w \rightarrow \infty, \tag{1.7}$$

where \sim denotes the asymptotic equivalence and $C_W > 0$ is a constant. This is a common assumption in the Internet traffic models, based on empirical findings (e.g. Abry *et al.* [2]). At small scales ($a \rightarrow 0^+$), we shall assume that the cumulative distribution F of interarrival times $A = A_{j,m}$ has a density f satisfying: for $\theta > 0$ and $C_f > 0$,

$$f(t) \sim C_f t^{\theta-1}, \quad \text{as } t \rightarrow 0^+. \tag{1.8}$$

In the applications to Internet traffic, F is often taken as a gamma distribution, that satisfies (1.8). For the Internet traffic data considered in this work, the parameter θ of the gamma distribution satisfying (1.8) is estimated to be smaller than 1.

Under the assumptions (1.7) and (1.8), the asymptotic behavior of the various cumulants and moments at small and large scales is summarized in Table 1 for $r \geq 2$. In each case, the asymptotic behavior is expressed in terms of λ and a , and a constant which depend only on r , and possibly the distributions of W (at large scales) and A (at small scales), with the exception of the factorial moments at small scales when $\theta = 1$. The latter case is not included in the table but is treated in our analysis below. We also specify the scales for which the asymptotic results are expected to hold in practice: e.g. $a \ll a_{\kappa,r}$ and $a_{\kappa,r} \ll a$ for cumulants where $a_{\kappa,r}$ is the transition scale between small and large scales obtained by equating the two asymptotic behaviors at small and large scales, and solving for a . (The notation \ll stands for the heuristic “much smaller.”) The exact forms of all the constants and transition scales will be given below.

One reason that we single out λ in the asymptotic results, in addition to a , is that the asymptotic behavior of the central moments is more delicate at large scales, depending on the magnitude of λ compared to a . This fact is well-known in the studies of large scale behavior of the distributional properties of PCPs, where one

distinguishes between the so-called fast growth regime corresponding to large $\lambda/a^{\alpha-1}$ and the slow growth regime corresponding to small $\lambda/a^{\alpha-1}$ (see Sect. 5 for more information). The different asymptotic behaviors of the central moments at large scales in Table 1 correspond to these two regimes. We should stress that all the other stated results at large scales are valid in both slow and fast growth regimes. We also note that the results indicated by (*) in Table 1, have been verified up to the order $r = 10$ only but otherwise conjectured to hold for all r .

Several interesting conclusions can be drawn from Table 1. For example, on the moments side and at large scales, note that the exponent α in (1.7) is captured by the central moments only, while at small scales, the exponent θ in (1.8) is captured by the factorial moments and in the case $\theta < 1$ only. The factorial cumulants, on the other hand, have the exponents α and θ at large and small scales, respectively. One natural interest in the obtained scaling relations is that they could be used for robust estimation procedures while inferring the scaling behavior. Our results for the central moments at large scales provide a more practical interpretation of the slow and fast growth regimes. Indeed, as argued below, these regimes reflect naturally the changing nature of the central moments as a increases for fixed λ ; whereas in the earlier literature, λ was associated with the length of the time window where the PCP was observed.

Our results on the large scale asymptotic behavior of the various moments are closest in the spirit to those of Dombry and Kaj [9] who considered moment measures in the parallel context of renewal point processes. But it should be noted that our approach and proofs are different, and some of the issues considered here are not addressed in Dombry and Kaj [9]. Further comparison with the work of Dombry and Kaj [9] will be provided (see Rem. 5.3 below).

Our results at small scales are somewhat connected to the so-called multifractal analysis which similarly focuses on the scaling behavior of the usual moments and cumulants of various quantities at small scales (*e.g.* Wendt *et al.* [33]). We shall pursue these connections in greater detail elsewhere. But we would like to note here that the behavior of PCPs and related models in connection to Internet traffic and multifractals, was explored in Hohn *et al.* [18], Veitch *et al.* [32], Ribeiro *et al.* [30], Krishna *et al.* [21]. The use of factorial moments in the multifractal (intermittency) analysis of point process data, instead of the usual moments, can be found in Carruthers *et al.* [5], de Wolf *et al.* [8], in connection to high-energy multiparticle collisions.

In summary, the structure of the paper is as follows. In Section 2, we provide the known formulae for the various cumulants and moments of PCPs. The behavior of the moments of PCP at large scales is studied in Section 3. Section 4 concerns the behavior of the moments at small scales. The large scale behavior in the slow and fast regimes mentioned above is discussed in Section 5. The transition between the different scaling behaviors is discussed in Section 6. The application to Internet traffic is given in Section 7, where we find the various empirical moments of the Internet traffic data set to be described quite well by the derived formulae for the moments of PCP and their asymptotic relations. Finally, in Appendix A, we provide the formulae relating the first 7 central moments and factorial cumulants, which are used in Section 5, and in Appendix B, we derive the formulae for the factorial cumulants of PCP, adapting the approach of Westcott [34].

2. MOMENTS AND CUMULANTS OF POISSON CLUSTER PROCESSES

The definition of moments and cumulants and their relations for general point processes are given in Daley and Vere-Jones [7] (Sect. 5.2). We gather here these formulae for the Poisson cluster processes (PCPs). We also introduce some notation used throughout this work. The focus is on the PCP N given by (1.1), and the corresponding equilibrium PCP N_e discussed following (1.1).

The interarrival times $A_{j,m}$ in (1.1) between the points in a cluster are independent and identically distributed as a random variable A having distribution function

$$F(t) = \mathbb{P}(A \leq t), \quad t > 0. \quad (2.1)$$

Its k th convolution will be denoted F_k , $k \geq 1$, with $F_1 = F$. The numbers of points W_j in clusters are independent and identically distributed as a random variable W with probability mass function

$$p_W(w) = \mathbb{P}(W = w), \quad w \geq 1. \tag{2.2}$$

Its tail probability will be denoted

$$R_w = \mathbb{P}(W \geq w), \quad w \geq 1. \tag{2.3}$$

As in (1.1), the starting points S_j of the clusters are the arrival times of a Poisson process with intensity $\lambda > 0$.

Let

$$P_t(z) = \mathbb{E}z^{N_e(0,t)} \tag{2.4}$$

be the probability generating function of the equilibrium PCP on the interval $(0, t)$. *Factorial cumulants* of $N_e(0, t)$ are defined as

$$\kappa_{[r]}(t) = \left. \frac{\partial^r \log P_t(z)}{\partial z^r} \right|_{z=1}, \quad r \geq 1. \tag{2.5}$$

The factorial cumulants of the equilibrium PCP N_e can also be obtained as

$$\kappa_{[1]}(t) = \lambda \mathbb{E}Wt, \quad \kappa_{[2]}(t) = 2\lambda \sum_{k=1}^{\infty} \int_0^t F_k(u)du \sum_{j=1}^{\infty} R_{j+k} \tag{2.6}$$

and, for $r > 2$,

$$\kappa_{[r]}(t) = (r-1)r\lambda \sum_{k=r-1}^{\infty} (k-1)(k-2)\dots(k-r+2) \int_0^t F_k(u)du \sum_{j=1}^{\infty} R_{j+k}. \tag{2.7}$$

The formulas (2.6)–(2.7) appear in Westcott [34] when the first points of the clusters are excluded. When the first points are included, the formulae are derived in Appendix B below. The *usual cumulants* of $N_e(0, t)$, on the other hand, are denoted $\kappa_r(t)$ and defined as in (1.5).

Remark 2.1. In computing the factorial cumulants through (2.7) for underlying choices of F , we truncate the sum (2.7) at large k . Evaluating the integrals $\int_0^t F_k(u)du$ is computationally more expensive, especially as k increases. One possibility we explored in computing these integrals for large k is to use normal approximations. That is, by central limit theorem, F_k is approximately normal with mean $k\mu$ and variance $k\sigma^2$, where μ and σ^2 are the mean and variance of F , respectively. Supposing this normal distribution for F_k , one can show that

$$\begin{aligned} \int_0^t F_k(u)du &= \frac{1}{2\sqrt{2\pi k}\sigma} e^{-(t^2+k^2\mu^2)/(2k\sigma^2)} \sum_{i=0}^{\infty} \sum_{j=0}^{\lfloor i/2 \rfloor} t^{2+i} \frac{(\mu/\sigma^2)^{i-2j}}{(i-2j)!(2k\sigma^2)^j} \\ &\times \left(\frac{\Gamma(i/2-j+1/2)}{\Gamma(i/2+3/2)} - \frac{\Gamma(i/2-j+1)}{\Gamma(i/2+2)} \right), \end{aligned} \tag{2.8}$$

where $\lfloor \cdot \rfloor$ is the floor function. (The proof of (2.8) uses simple algebraic manipulations and is omitted). The series on the right-hand side of (2.8) converges quickly and makes the evaluation of $\int_0^t F_k(u)du$ computationally convenient. We found this approximation to work well in practice, but also not to make significant difference for our purposes.

The *factorial moments* are defined as

$$m_{[r]}(t) = \mathbb{E}N_e(0, t)^{[r]}, \quad r \geq 1, \tag{2.9}$$

where $n^{[r]} = n(n-1)\dots(n-r+1)$. In view of (1.5)–(1.6), they are related to the factorial cumulants in the same way that the usual moments and cumulants relate, that is,

$$m_{[1]}(t) = \kappa_{[1]}(t), \quad m_{[2]}(t) = \kappa_{[2]}(t) + \kappa_{[1]}(t)^2, \quad m_{[3]}(t) = \kappa_{[3]}(t) + 3\kappa_{[2]}(t)\kappa_{[1]}(t) + \kappa_{[1]}(t)^3 \tag{2.10}$$

and, in general,

$$m_{[r]}(t) = \sum_{k=0}^{r-1} \binom{r-1}{k} \kappa_{[r-k]}(t)m_{[k]}(t) \tag{2.11}$$

and also

$$m_{[r]}(t) = \sum_{k=1}^r B_{r,k}(\kappa_{[1]}(t), \kappa_{[2]}(t), \dots, \kappa_{[r-k+1]}(t)), \tag{2.12}$$

where $B_{r,k}$ are the Bell polynomials given by

$$B_{r,k}(x_1, x_2, \dots, x_{r-k+1}) = \sum_{(n_1, n_2, \dots, n_{r-k+1}) \in S_{r,k}} \frac{r!}{n_1!n_2!\dots n_{r-k+1}!} \left(\frac{x_1}{1!}\right)^{n_1} \left(\frac{x_2}{2!}\right)^{n_2} \dots \left(\frac{x_{r-k+1}}{(r-k+1)!}\right)^{n_{r-k+1}} \tag{2.13}$$

with $S_{r,k}$ consisting of all $(n_1, n_2, \dots, n_{r-k+1}) \in (\mathbb{N} \cup \{0\})^{r-k+1}$ such that $n_1 + n_2 + \dots + n_{r-k+1} = k$ and $n_1 + 2n_2 + \dots + (r-k+1)n_{r-k+1} = r$. In the case of the usual moments and cumulants, the formula (2.12) appears, for example, in Peccati and Taqqu [29], Proposition 3.3.1 (see the relation (3.3.26) of the proposition and also the definition of the complete Bell polynomials in (2.4.13)).

The usual moments

$$m_r(t) = \mathbb{E}N_e(0, t)^r, \quad r \geq 1, \tag{2.14}$$

are related to the factorial moments through the relation

$$m_r(t) = \sum_{j=1}^r \Delta_{j,r} m_{[j]}(t), \tag{2.15}$$

where $\Delta_{j,r}$ are the Stirling numbers of the second kind (e.g. Daley and Vere–Jones [7], pp. 114–115). In our analysis, we shall be working with factorial cumulants through the formulas (2.6)–(2.7), and then relate them to factorial and usual moments by using the relations above.

We shall also present results for *central moments*

$$m_r^0(t) = \mathbb{E}(N_e(0, t) - \mathbb{E}N_e(0, t))^r, \quad r \geq 1, \tag{2.16}$$

which are related to the usual moments through

$$m_r^0(t) = \sum_{j=0}^r \binom{r}{j} (-1)^{r-j} m_j(t) m_1(t)^{r-j}. \tag{2.17}$$

Finally, we use the following recursion formula relating cumulants and central moments (e.g. Willink [35]):

$$\kappa_r(t) = m_r^0(t) - \sum_{j=1}^{r-2} m_j^0(t) \kappa_{r-j}(t), \quad r \geq 2, \tag{2.18}$$

and $\kappa_1(t) = \lambda \mathbb{E}Wt$.

3. MOMENT AND CUMULANT BEHAVIOR AT LARGE SCALES

In this section, we study the asymptotic behavior of the cumulants $\kappa_{[r]}(a)$, $\kappa_r(a)$ and the various moments $m_{[r]}(a)$, $m_r(a)$ and $m_r^0(a)$ at large scales, that is, as $a \rightarrow \infty$. We assume that the distribution of the number of points in a cluster is heavy-tailed in the following sense.

Assumption 3.1. The distribution of W is heavy-tailed, that is,

$$P(W > w) \sim C_W w^{-\alpha}, \quad \text{as } w \rightarrow \infty, \tag{3.1}$$

where $1 < \alpha < 2$ and $C_W > 0$.

The assumption $\alpha \in (1, 2)$ can be relaxed to $\alpha > 1$ but at the expense of more involved formulae. The range $\alpha \in (1, 2)$ is motivated by typical estimated values of α in applications to Internet traffic, and corresponds to W having finite mean but infinite variance.

Proposition 3.2. *Suppose that the distribution of the number of points in a cluster of PCP satisfies Assumption 3.1 above. Suppose also that $\mathbb{E}A < \infty$. The factorial cumulants $\kappa_{[r]}(a)$, $r \geq 1$, of PCP then satisfy:*

$$\kappa_{[1]}(a) = \lambda \mathbb{E}W a, \quad \kappa_{[r]}(a) \sim C_{\kappa, [r]} \lambda a^{r-\alpha+1}, \quad r \geq 2, \quad \text{as } a \rightarrow \infty, \tag{3.2}$$

where

$$C_{\kappa, [r]} = \frac{r(r-1)C_W}{(\alpha-1)(r-\alpha)(r+1-\alpha)(\mathbb{E}A)^{r-\alpha}}. \tag{3.3}$$

Proof. We denote by \tilde{S}_k the sum of k i.i.d. interarrival times A , associated with the distribution function F_k , $k \geq 1$. We also let $\tilde{N}(u) = \sum_{k=1}^{\infty} 1_{\{\tilde{S}_k \leq u\}}$, $u \geq 0$, be the respective infinite renewal process. The first relation in (3.2) is just the first relation in (2.6). For $r = 2$, the second relation in (2.6) and Assumption 3.1 yield

$$\kappa_{[2]}(a) = \kappa_{[2],1}(a) + o(a^{1+\epsilon} \vee \kappa_{[2],1}(a)), \tag{3.4}$$

where $\epsilon > 0$ is arbitrarily small but fixed, $a \vee b = \max\{a, b\}$ and

$$\begin{aligned} \kappa_{[2],1}(a) &= 2\lambda C_W \sum_{k=1}^{\infty} \int_0^a F_k(u) du \sum_{j=1}^{\infty} (j+k)^{-\alpha} \\ &= 2\lambda C_W \sum_{k=1}^{\infty} \int_0^a F_k(u) du k^{1-\alpha} \sum_{j=1}^{\infty} \left(1 + \frac{j}{k}\right)^{-\alpha} \frac{1}{k}. \end{aligned}$$

Indeed, note that

$$\kappa_{[2]}(a) - \kappa_{[2],1}(a) = 2\lambda C_W \sum_{k=1}^{\infty} \int_0^a F_k(u) du \sum_{j=1}^{\infty} (j+k)^{-\alpha} \left(\frac{\mathbb{P}(W \geq j+k)}{C_W(j+k)^{-\alpha}} - 1 \right).$$

The latter sum over $\sum_{k=K+1}^{\infty}$, when divided by $\kappa_{[2],1}(a)$, can be made arbitrarily small for large enough K , since $\mathbb{P}(W \geq j+k) \sim C_W(j+k)^{-\alpha}$ as $k \rightarrow \infty$. For the sum over $\sum_{k=1}^K$, on the other hand, note that it is bounded up to a multiplicative constant by a since $\int_0^a F_k(u) du \leq \int_0^a 1 du = a$. Hence, when divided by $a^{1+\epsilon}$, the sum converges to zero as $a \rightarrow \infty$. Thus, the relation (3.4) indeed holds.

Similarly,

$$\kappa_{[2],1}(a) = \kappa_{[2],2}(a) + o(a^{1+\epsilon} \vee \kappa_{[2],2}(a)), \tag{3.5}$$

where, by using $\int_0^\infty (1 + u)^{-\alpha} du = (\alpha - 1)^{-1}$,

$$\begin{aligned} \kappa_{[2],2}(a) &= \frac{2\lambda C_W}{\alpha - 1} \sum_{k=1}^\infty \int_0^a F_k(u) du k^{1-\alpha} \\ &= \frac{2\lambda C_W}{\alpha - 1} \int_0^a \sum_{k=1}^\infty \mathbb{P}(\tilde{S}_k \leq u) k^{1-\alpha} du \\ &= \frac{2\lambda C_W}{\alpha - 1} \int_0^a \sum_{k=1}^\infty \mathbb{P}(\tilde{N}(u) \geq k) k^{1-\alpha} du \\ &= \frac{2\lambda C_W}{\alpha - 1} \int_0^a \sum_{j=1}^\infty \mathbb{P}(\tilde{N}(u) = j) \sum_{k=1}^j k^{1-\alpha} du \end{aligned}$$

and

$$\kappa_{[2],2}(a) = \kappa_{[2],3}(a) + o(a^{1+\epsilon} \vee \kappa_{[2],3}(a)), \tag{3.6}$$

where

$$\begin{aligned} \kappa_{[2],3}(a) &= \frac{2\lambda C_W}{(\alpha - 1)(2 - \alpha)} \int_0^a \sum_{j=1}^\infty \mathbb{P}(\tilde{N}(u) = j) j^{2-\alpha} du \\ &= \frac{2\lambda C_W}{(\alpha - 1)(2 - \alpha)} \int_0^a \mathbb{E}(\tilde{N}(u)^{2-\alpha}) du. \end{aligned}$$

By Theorem 5.1, (ii), in Gut [17], Chapter 2, $\mathbb{E}((\tilde{N}(u)/u)^{2-\alpha}) \rightarrow (1/\mathbb{E}A)^{2-\alpha}$ as $u \rightarrow \infty$ and hence

$$\kappa_{[2],3}(a) \sim \frac{2\lambda C_W}{(\alpha - 1)(2 - \alpha)(3 - \alpha)(\mathbb{E}A)^{2-\alpha}} a^{3-\alpha}. \tag{3.7}$$

The relation (3.2) with $r = 2$ now follows from (3.4)–(3.7) as long as $\epsilon > 0$ is such that $1 + \epsilon < 3 - \alpha$. The relation (3.2) with $r > 2$ can be proved similarly by arguing that

$$\kappa_{[r]}(a) \sim \frac{r(r - 1)\lambda C_W}{(\alpha - 1)(r - \alpha)} \int_0^a \mathbb{E}(\tilde{N}(u)^{r-\alpha}) du$$

and again using the same result of Gut [17]. □

The next two results describe the asymptotic behavior of the moments and cumulants at large scales.

Corollary 3.3. *Under the assumptions of Proposition 3.2, the factorial moments $m_{[r]}(a)$, the moments $m_r(a)$ and the central moments $m_r^0(a)$, $r \geq 2$, of PCP satisfy:*

$$m_{[r]}(a) \sim C_{m,[r]} \lambda^r a^r, \tag{3.8}$$

$$m_r(a) \sim C_{m,r} \lambda^r a^r, \tag{3.9}$$

$$m_r^0(a) \sim C_{m,r}^0 \lambda a^{r-\alpha+1}, \quad \text{as } a \rightarrow \infty, \tag{3.10}$$

where

$$C_{m,[r]} = C_{m,r} = (\mathbb{E}W)^r, \quad C_{m,r}^0 = C_{\kappa,[r]} \tag{3.11}$$

with $C_{\kappa,[r]}$ given in (3.3). (When $r = 1$, $m_{[1]}(a) = m_1(a) = \kappa_{[1]}(a) = \lambda \mathbb{E}W a$ and $m_1^0(a) = 0$.)

Proof. The relation (3.8) can be shown recursively by using (3.2) and (2.10)–(2.11). Indeed, when $r = 2$, it follows from the second relation in (2.10) and (3.2). Supposing it holds for $2, \dots, r - 1$, it also holds for r since the term $\kappa_{[r-k]}(a)m_{[k]}(a)$ in (2.11) behaves as $(\lambda \mathbb{E}W a)(C_{m, [r-1]} \lambda^{r-1} a^{r-1}) = C_{m, [r]} \lambda^r a^r$ when $k = r - 1$, and is of the smaller order $a^{r-k-\alpha+1} \cdot a^k = a^{r-\alpha+1}$ when $k < r - 1$. The relation (3.9) follows immediately from (2.15) and (3.8), and the fact that $\Delta_{r,r} = 1$.

The relation (3.10) is slightly more difficult to deal with. We shall use the relation (2.17) to express $m_r^0(a)$ in terms of the moments $m_j(a)$, $j = 1, \dots, r$, and the relations (2.15) and (2.12) to express $m_j(a)$ in terms of the factorial cumulants $\kappa_{[1]}(a), \dots, \kappa_{[j]}(a)$. Changing the indices to avoid confusion, note that (2.15) and (2.12) yield

$$\begin{aligned} m_j(a) &= \sum_{p=1}^j \Delta_{j,p} m_{[p]}(a) = \sum_{p=1}^j \Delta_{j,p} \sum_{q=1}^p B_{p,q}(\kappa_{[1]}(a), \kappa_{[2]}(a), \dots, \kappa_{[p-q+1]}(a)) \\ &= \sum_{p=1}^j \sum_{q=1}^p \sum_{(n_1, \dots, n_{p-q+1}) \in S_{p,q}} \Delta_{j,p} \frac{p!}{n_1! \dots n_{p-q+1}!} \left(\frac{\kappa_{[1]}(a)}{1!} \right)^{n_1} \dots \left(\frac{\kappa_{[p-q+1]}(a)}{(p-q+1)!} \right)^{n_{p-q+1}} \\ &=: \sum_{p=1}^j \sum_{q=1}^p \sum_{(n_1, \dots, n_{p-q+1}) \in S_{p,q}} T_{p,q}(n_1, \dots, n_{p-q+1}), \end{aligned} \tag{3.12}$$

where integers $n_1, n_2, \dots, n_{p-q+1} \geq 0$ are such that $n_1 + n_2 + \dots + n_{p-q+1} = q$ and $n_1 + 2n_2 + \dots + (p - q + 1)n_{p-q+1} = p$. By using these two relations for $n_1, n_2, \dots, n_{p-q+1}$ and (3.2), note that the order of the term $T_{p,q}(n_1, \dots, n_{p-q+1})$ in (3.12) is

$$a^{n_1} a^{n_2(2-\alpha+1)} \dots a^{n_{p-q+1}(p-q+1-\alpha+1)} = a^{p-(q-n_1)(\alpha-1)}. \tag{3.13}$$

This order is largest when

$$p = j, \quad q = j, \quad n_1 = j, \quad n_2 = \dots = n_{p-q+1} = 0, \tag{3.14}$$

which corresponds to

$$T_{j,j}(j, 0, \dots, 0) = \kappa_{[1]}(a)^j.$$

But, when substituted into (2.17), this term yields

$$\sum_{j=0}^r \binom{r}{j} (-1)^{r-j} T_{j,j}(j, 0, \dots, 0) m_1(a)^{r-j} = \kappa_{[1]}(a)^r \sum_{j=0}^r \binom{r}{j} (-1)^{r-j} = 0,$$

and hence the case (3.14) can be eliminated from the sum in (3.12). The next largest order in (3.13) occurs when

$$q - n_1 = 1 \quad (n_1 = q - 1), \quad p = j. \tag{3.15}$$

The rest of the integers $n_2, \dots, n_{p-q+1} \geq 0$ then satisfy $n_2 + \dots + n_{p-q+1} = 1$ and $2n_2 + \dots + (p - q + 1)n_{p-q+1} = p - q + 1$ which is possible only when $n_{p-q+1} = 1, n_2 = \dots = n_{p-q} = 0$. The corresponding terms in (3.12) are then

$$\sum_{q=1}^{j-1} T_{j,q}(q - 1, 0, \dots, 0, 1) = \sum_{q=1}^{j-1} \frac{j!}{(q-1)!(j-q+1)!} \kappa_{[1]}(a)^{q-1} \kappa_{[j-q+1]}(a).$$

When substituted into (2.17), this yields

$$\begin{aligned} & \sum_{j=2}^r \binom{r}{j} (-1)^{r-j} \sum_{q=1}^{j-1} T_{j,q}(q-1, 0, \dots, 0, 1) \kappa_{[1]}(a)^{r-j} \\ &= \sum_{j=2}^r \binom{r}{j} (-1)^{r-j} \sum_{q=1}^{j-1} \frac{j!}{(q-1)!(j-q+1)!} \kappa_{[1]}(a)^{q-1+r-j} \kappa_{[j-q+1]}(a) \\ &= \sum_{j=2}^r \binom{r}{j} (-1)^{r-j} \sum_{\ell=2}^j \frac{j!}{(j-\ell)!\ell!} \kappa_{[1]}(a)^{r-\ell} \kappa_{[\ell]}(a) \\ &= \sum_{\ell=2}^r \kappa_{[1]}(a)^{r-\ell} \kappa_{[\ell]}(a) \sum_{j=\ell}^r (-1)^{r-j} \binom{r}{j} \binom{j}{\ell} = \kappa_{[r]}(a), \end{aligned}$$

since, for $\ell < r$,

$$\sum_{j=\ell}^r (-1)^{r-j} \binom{r}{j} \binom{j}{\ell} = \binom{r}{\ell} \sum_{k=0}^{r-\ell} \binom{r-\ell}{k} (-1)^{r-\ell-k} = 0.$$

This yields (3.10) in view of (3.2). □

Corollary 3.4. *Under the assumptions of Proposition 3.2, the cumulants $\kappa_r(a)$, $r \geq 2$, of PCP satisfy:*

$$\kappa_r(a) \sim C_{\kappa,r} \lambda a^{r-\alpha+1}, \quad \text{as } a \rightarrow \infty, \tag{3.16}$$

where

$$C_{\kappa,r} = C_{\kappa,[r]} \tag{3.17}$$

with $C_{\kappa,[r]}$ given in (3.3). (When $r = 1$, $\kappa_1(a) = \lambda \mathbb{E}W a$.)

Proof. To show (3.16), we shall use the relation (2.18) between cumulants and central moments and the asymptotic behavior (3.10) of the central moments at large scales. The relation (3.16) is trivial for the second and third cumulants since $\kappa_2(a) = m_2^0(a)$ and $\kappa_3(a) = m_3^0(a)$ (which follow from (2.18)). By induction, if (3.16) holds for $2, \dots, r-1$, then it also holds for r since in (2.18) the term $m_j^0(a) \kappa_{r-j}(a)$ is of the order $a^{j-\alpha+1} \cdot a^{r-j-\alpha+1} = a^{r-2\alpha+2}$ and the term $m_r^0(a)$ has the order $a^{r-\alpha+1}$. □

At large scales, the scaling behaviors of (factorial) cumulants and central moments include the tail parameter α in the exponent of a , and in that sense they are more informative about the scaling behavior of PCPs than moments and factorial moments whose behavior does not involve α in the exponents.

4. MOMENT AND CUMULANT BEHAVIOR AT SMALL SCALES

We are interested here in the asymptotic behavior of the cumulants $\kappa_{[r]}(a)$, $\kappa_r(a)$ and the various moments $m_{[r]}(a)$, $m_r(a)$ and $m_r^0(a)$ at small scales, that is, as $a \rightarrow 0^+$. We focus on the following class of distributions of the interarrival times between points in clusters.

Assumption 4.1. Suppose that the cumulative distribution F of interarrival times between points in clusters has a density f satisfying: for $\theta > 0$ and $C_f > 0$,

$$f(t) \sim C_f t^{\theta-1}, \quad \text{as } t \rightarrow 0^+. \tag{4.1}$$

An example is the gamma distribution with parameters $\theta > 0$ and $\beta > 0$ having density

$$f(t) = \frac{\beta(\beta t)^{\theta-1}e^{-\beta t}}{\Gamma(\theta)}, \quad t > 0, \tag{4.2}$$

where $\Gamma(\cdot)$ denotes the usual gamma function. The gamma distribution will be used in the application to Internet traffic in Section 7 below. Note that for the gamma distribution, $C_f = \beta^\theta/\Gamma(\theta)$ in (4.1).

The next result provides the small scale behavior of the factorial cumulants. The behavior of the moments and cumulants will follow from this result, as stated in the subsequent corollaries.

Proposition 4.2. *Suppose that the distribution of the interarrival times of PCP satisfies Assumption 4.1 above. The factorial cumulants $\kappa_{[r]}(a)$, $r \geq 2$, of PCP then satisfy:*

$$\kappa_{[r]}(a) \sim c_{\kappa,[r]} \lambda a^{1+(r-1)\theta}, \quad \text{as } a \rightarrow 0^+, \tag{4.3}$$

where

$$c_{\kappa,[r]} = \frac{r!C_{F,r-1}\bar{R}_r}{(r-1)\theta+1} \tag{4.4}$$

with $\bar{R}_r = \sum_{w=r}^\infty R_w = \sum_{w=r}^\infty \mathbb{P}(W \geq w)$ and

$$C_{F,r-1} = C_{F,r-2}C_f B((r-2)\theta+1, \theta) = \frac{C_f^{r-1}\Gamma(\theta)^{r-1}}{\Gamma((r-1)\theta+1)}, \quad C_{F,1} = \frac{C_f}{\theta} \tag{4.5}$$

for the beta function $B(\cdot, \cdot)$. (When $r = 1$, $\kappa_{[1]}(a) = \lambda \mathbb{E}W a$).

Proof. We shall use the formulas (2.6)–(2.7) for the factorial cumulants $\kappa_{[r]}(a)$, which involve the integrals $\int_0^a F_k(u)du$, $k \geq 1$. When $k = 1$, we have from (4.1) that $F_1(u) \sim C_f u^\theta/\theta =: C_{F,1}u^\theta$, as $u \rightarrow 0^+$. In fact, for any $k \geq 1$,

$$F_k(u) \sim C_{F,k}u^{k\theta}, \quad \text{as } u \rightarrow 0^+, \tag{4.6}$$

where $C_{F,k} = C_{F,k-1}C_f B((k-1)\theta+1, \theta)$ with the beta function $B(\cdot, \cdot)$. Indeed, supposing by induction that (4.6) hold for k , note that

$$\begin{aligned} F_{k+1}(u) &= \int_0^u F_k(y)f(u-y)dy \sim C_{F,k}C_f \int_0^u y^{k\theta}(u-y)^{\theta-1}dy \\ &= C_{F,k}C_f \int_0^1 z^{k\theta}(1-z)^{\theta-1}dz u^{(k+1)\theta} = C_{F,k}C_f B(k\theta+1, \theta) u^{(k+1)\theta} = C_{F,k+1}u^{(k+1)\theta}. \end{aligned}$$

The relation (4.6) now implies that

$$\int_0^a F_k(u)du \sim \frac{C_{F,k}}{k\theta+1} a^{1+k\theta}, \quad \text{as } a \rightarrow 0^+. \tag{4.7}$$

In view of (4.7), the leading term for $\kappa_{[r]}(a)$ in (2.6)–(2.7) is of the desired order $a^{1+(r-1)\theta}$ and with the specified constant $c_{\kappa,[r]}\lambda$. To show that the sum of the remaining terms is negligible, one can use the argument above to conclude that, for any $\epsilon > 0$, $f(a) \leq C a^{\theta-\epsilon-1}$, $a \in (0, a_0)$, and hence

$$\int_0^a F_k(u)du \leq \frac{C'_{F,k}}{k(\theta-\epsilon)+1} a^{1+k(\theta-\epsilon)}, \quad a \in (0, a_0), \tag{4.8}$$

where $C'_{F,k}$ has the same structure as $C_{F,k}$ but with θ replaced by $\theta - \epsilon$. The remaining terms in (2.6)–(2.7) (that is, without the leading term $a^{1+(r-1)\theta}$) are thus bounded by a function of the order $a^{1+r(\theta-\epsilon)}$, which is negligible compared to $a^{1+(r-1)\theta}$ for small enough ϵ .

The last equality in the first relation of (4.5) follows from using the recursion relation $C_{F,r-1} = C_{F,r-2}C_f B((r-2)\theta+1, \theta)$ along with the definition of the beta function $B(a, b) = \Gamma(a)\Gamma(b)/\Gamma(a+b)$. \square

Corollary 4.3. *Suppose the distribution of the interarrival times of PCP satisfies Assumption 4.1 above. The moments $m_r(a)$, factorial moments $m_{[r]}(a)$ and central moments $m_r^0(a)$, $r \geq 2$, of PCP then satisfy:*

$$m_{[r]}(a) \sim \begin{cases} c_{m,[r]}(\theta)\lambda a^{1+(r-1)\theta}, & 0 < \theta < 1, \\ (\sum_{k=1}^r B_{r,k}(c_{\kappa,[1]}, c_{\kappa,[2]}, \dots, c_{\kappa,[r-k+1]})\lambda^k) a^r, & \theta = 1, \\ c_{m,[r]}(\theta)\lambda^r a^r, & \theta > 1, \end{cases} \tag{4.9}$$

$$m_r(a) \sim c_{m,r}\lambda a, \tag{4.10}$$

$$m_r^0(a) \sim c_{m,r}^0\lambda a, \text{ as } a \rightarrow 0^+, \tag{4.11}$$

where

$$c_{m,[r]}(\theta) = \begin{cases} c_{\kappa,[r]}, & 0 < \theta < 1, \\ (\mathbb{E}W)^r, & \theta > 1, \end{cases} \quad c_{m,r} = \mathbb{E}W, \quad c_{m,r}^0 = \mathbb{E}W \tag{4.12}$$

with $B_{r,k}$ the Bell polynomials and $c_{\kappa,[r]}$ appearing in (4.4). (When $r = 1$, $m_{[1]}(a) = m_1(a) = \kappa_{[1]}(a) = \lambda \mathbb{E}W a$ and $m_1^0(a) = 0$.)

Proof. To show (4.9), we suppose first that $\theta \neq 1$ and argue by induction. The relation (4.9) with $r = 2$ holds in view of (2.10), (2.6) and (4.3), and taking into account the fact that $0 < \theta < 1$ or $\theta > 1$. Supposing it holds for $1, \dots, r - 1$, it also holds for r by using the recursion formula (2.11) relating factorial moments and factorial cumulants and (4.3). Indeed, for $0 < \theta < 1$, the term $\kappa_{[r-k]}(a)m_{[k]}(a)$ in the sum (2.11) behaves as $c_{\kappa,[r]}\lambda a^{1+(r-1)\theta}$ when $k = 0$, and is of the smaller order $a^{1+(r-k-1)\theta} a^{1+(k-1)\theta} = a^{2+(r-1)\theta-\theta}$ when $k \geq 1$. Similarly, for $\theta > 1$, the term $\kappa_{[r-k]}(a)m_{[k]}(a)$ in the sum (2.11) behaves as $(\lambda \mathbb{E}W a)(c_{m,[r-1]}(\theta)\lambda^{r-1}a^{r-1}) = c_{m,[r]}(\theta)\lambda^r a^r$ when $k = r - 1$, and is of the smaller order $a^{1+(r-k-1)\theta} a^k = a^{r\theta-(\theta-1)(k+1)}$ when $k < r - 1$.

When $\theta = 1$, it is more convenient to use the direct relation (2.12) between factorial moments and factorial cumulants. By using (4.3) and (2.13), the Bell polynomials in (2.12) behave as

$$B_{r,k}(\kappa_{[1]}(a), \kappa_{[1]}(a), \dots, \kappa_{[r-k+1]}(a)) \sim \sum_{(n_1, n_2, \dots, n_{r-k+1}) \in S_{r,k}} \frac{r!}{n_1! n_2! \dots n_{r-k+1}!} \left(\frac{c_{\kappa,[1]}\lambda a}{1!} \right)^{n_1} \left(\frac{c_{\kappa,[r]}\lambda a^2}{2!} \right)^{n_2} \dots \left(\frac{c_{\kappa,[r-k+1]}\lambda a^{r-k+1}}{(r-k+1)!} \right)^{n_{r-k+1}} \tag{4.13}$$

and since $\lambda^{n_1+n_2+\dots+n_{r-k+1}} = \lambda^k$ and $a^{n_1+2n_2+\dots+(r-k+1)n_{r-k+1}} = a^r$, the relation (4.9) follows for $\theta = 1$.

The relation (4.10) follows from (2.15), (4.9) and the fact that $m_{[1]}(a) = \kappa_{[1]}(a) = \lambda \mathbb{E}W a$, since the term $m_{[1]}(a)$ dominates in the sum in (2.15). Similarly, the relation (4.11) follows from (2.17) and (4.10) since the term $m_r(a)$ when $j = r$ is dominant in (2.17). \square

Corollary 4.4. *Suppose the distribution of the interarrival times of PCP satisfies Assumption 4.1 above. The cumulants $\kappa_r(t)$, $r \geq 2$, of PCP then satisfy:*

$$\kappa_r(a) \sim c_{\kappa,r}\lambda a, \text{ as } a \rightarrow 0^+, \tag{4.14}$$

where

$$c_{\kappa,r} = \mathbb{E}W. \tag{4.15}$$

(When $r = 1$, $\kappa_1(a) = \lambda \mathbb{E}W a$.)

Proof. The relation (4.14) can be shown by using (2.18) and (4.11). From (2.18) we have $\kappa_2(a) = m_2^0(a)$ and $\kappa_3(a) = m_3^0(a)$ and the result follows immediately by (4.11). By induction, if (4.14) holds for $2, 3, \dots, r - 1$, then it also holds for r since in (2.18) the term $m_j^0(a)\kappa_{r-j}(a)$ in the sum is of the order $a \cdot a = a^2$ and the term $m_r^0(a)$ has the order a . \square

The asymptotic results (4.9)–(4.11) show that using regular moments and central moments of PCP at small scales does not reveal the underlying interarrival distribution, since the dominating behavior is governed by a for all the moments; in contrast, the behavior of the factorial moments (when $\theta < 1$) is more informative. A similar conclusion can be drawn for cumulants and factorial cumulants.

Example 4.5. Proposition 4.2 describes the asymptotic behavior of the factorial cumulants of PCP under Assumption 4.1. A more explicit, non-asymptotic expression of the factorial cumulants can be obtained in the special case of the gamma distribution with parameters $\theta > 0$ and $\beta > 0$ (4.2) used in practice (Abry *et al.* [2]), yielding a result of independent interest. Indeed, observe that in this case,

$$F_k(x) = \frac{\gamma(k\theta, \beta x)}{\Gamma(k\theta)},$$

where $\gamma(s, x) = \int_0^x a^{s-1} e^{-s} ds$ is the lower incomplete gamma function. By using integration by parts, note that

$$\int_0^a \gamma(k\theta, \beta u) du = \frac{1}{\beta} \int_0^{\beta a} \gamma(k\theta, u) du = \frac{1}{\beta} (\beta a \gamma(k\theta, \beta a) - \gamma(k\theta + 1, \beta a))$$

and hence

$$\begin{aligned} \int_0^a F_k(u) du &= \frac{1}{\Gamma(k\theta)\beta} (\beta a \gamma(k\theta, \beta a) - \gamma(k\theta + 1, \beta a)) \\ &= \frac{1}{\Gamma(k\theta)\beta} \left(\beta a (\beta a)^{k\theta} \Gamma(k\theta) e^{-\beta a} \sum_{j=0}^{\infty} \frac{(\beta a)^j}{\Gamma(k\theta + j + 1)} - (\beta a)^{k\theta+1} \Gamma(k\theta + 1) e^{-\beta a} \sum_{j=0}^{\infty} \frac{(\beta a)^j}{\Gamma(k\theta + 1 + j + 1)} \right) \\ &= \frac{e^{-\beta a}}{\beta} \sum_{j=0}^{\infty} (\beta a)^{k\theta+j+1} \left(\frac{1}{\Gamma(k\theta + j + 1)} - \frac{k\theta}{\Gamma(k\theta + 1 + j + 1)} \right) \\ &= \frac{e^{-\beta a}}{\beta} \sum_{j=0}^{\infty} (\beta a)^{k\theta+j+1} \frac{j + 1}{\Gamma(k\theta + 1 + j + 1)} = \frac{e^{-\beta a}}{\beta} \sum_{j=1}^{\infty} (\beta a)^{k\theta+j} \frac{j}{\Gamma(k\theta + 1 + j)}. \end{aligned} \tag{4.16}$$

An expression for the factorial cumulants can now be obtained by substituting (4.16) into (2.6)–(2.7). For example, for $r \geq 3$ and as $a \rightarrow 0^+$, the leading term in thus obtained relation leads to

$$\kappa_{[r]}(a) \sim \frac{r! \lambda (\beta a)^{(r-1)\theta+1} \overline{R}_r}{\beta \Gamma((r-1)\theta + 2)} = \frac{r! \lambda \beta^{(r-1)\theta} \overline{R}_r}{\Gamma((r-1)\theta + 2)} a^{1+(r-1)\theta}, \tag{4.17}$$

which is consistent with (4.3)–(4.5).

Remark 4.6. An asymptotic behavior of the density f not captured by Assumption 4.1 is when $f(t)$ decays faster than any power as $t \rightarrow 0^+$. This could be expressed, for example, by the assumption that

$$f(t) \sim C t^\delta e^{-|\log t|^\beta}, \quad \text{as } t \rightarrow 0^+, \tag{4.18}$$

where $\delta \in \mathbb{R}$, $\beta > 1$ and $C > 0$. Results analogous to (4.3) and (4.9)–(4.10) could be obtained, we believe, under the assumption (4.18). However, we shall not pursue this direction here for the following reason. A prototypical example of the density satisfying (4.18) is that of a lognormal distribution. When working with the Internet traffic data for Section 7, we found the behaviour at small scale to be difficult to capture using the lognormal distribution. This is consistent with the observations in Mandelbrot [25], who found the moments of a lognormal distribution to be “localized” (in other words, different moments to be determined by different portions of the density of the lognormal distribution), making the use of the distribution quite delicate in practice.

5. MOMENT AND CUMULANT BEHAVIOR AT LARGE SCALES IN THE SLOW AND FAST GROWTH REGIMES

The results of Proposition 3.2 and Corollaries 3.3 and 3.4, are valid when $a \rightarrow \infty$ and thus are asymptotic in nature. In fact, depending on the magnitude of the arrival rate λ , a different scaling behavior could be observed for some quantities of interest over a range of large scales a . We shall focus here and refine the behavior of central moments, not just in terms of a but also λ . The cases of other moments and cumulants will be discussed briefly in Remark 5.2 below.

As in some related work to be discussed below, we shall distinguish between the so-called *slow growth regime*, defined as

$$\frac{\lambda}{a^{\alpha-1}} \rightarrow 0, \quad (5.1)$$

and the *fast growth regime*, defined as

$$\frac{\lambda}{a^{\alpha-1}} \rightarrow \infty, \quad (5.2)$$

where α is the power-law exponent appearing in (3.1). From a practical perspective, the relations (5.1) and (5.2) are mathematical idealizations for the conditions that $\lambda/a^{\alpha-1}$ is small and large, respectively. What “small” and “large” mean from a practical perspective will be discussed in Section 6. Since $a \rightarrow \infty$, note that (5.2) is possible only when $\lambda \rightarrow \infty$ as well. But to reiterate, this does not mean that λ changes with a – the condition (5.2) stands for $\lambda/a^{\alpha-1}$ being large. Under the regime (5.1), λ can be “constant” or “increase” slower than $a^{\alpha-1}$. The reader unfamiliar with these regimes from the literature will be able to follow the arguments below without much difficulty. As will be shown, the regimes (5.1) and (5.2) are natural in the simple asymptotic analysis of central moments to be carried out below.

We do not have a general result for the asymptotics of central moments of PCP under the slow and fast growth conditions. The key difficulty is seemingly the lack of a direct general formula relating central moments to factorial cumulants (see also Rem. 5.1 below). But such formula can be derived for a number of first moments of interest, as done for the first seven central moments in Appendix A, and then be used to obtain the asymptotics of central moments under the slow and fast growths.

By using the formula (A.1) and substituting the factorial cumulants $\kappa_{[1]}(t)$ and $\kappa_{[2]}(t)$ from (3.2) (valid for any λ , possibly $\lambda \rightarrow \infty$), we can write

$$m_2^0(a) \sim C_{\kappa,[1]}\lambda a + C_{\kappa,[2]}\lambda a^{3-\alpha}, \quad (5.3)$$

as $a \rightarrow \infty$. In view of (5.1) and (5.2),

$$m_2^0(a) \sim C_{\kappa,[2]}\lambda a^{3-\alpha}, \quad (5.4)$$

for the slow and fast growth regimes. From (A.2) and proceeding as above,

$$m_3^0(a) \sim C_{\kappa,[1]}\lambda a + 3C_{\kappa,[2]}\lambda a^{3-\alpha} + C_{\kappa,[3]}\lambda a^{4-\alpha}, \quad (5.5)$$

as $a \rightarrow \infty$ and

$$m_3^0(a) \sim C_{\kappa,[3]}\lambda a^{4-\alpha}, \quad (5.6)$$

in both growth regimes. From the relationship between the fourth central moment and factorial cumulants in (A.3), we can write the asymptotic relation

$$\begin{aligned} m_4^0(a) &\sim C_{\kappa,[1]}\lambda a + 3C_{\kappa,[1]}^2\lambda^2 a^2 + 7C_{\kappa,[2]}\lambda a^{3-\alpha} + 6C_{\kappa,[3]}\lambda a^{4-\alpha} \\ &\quad + C_{\kappa,[4]}\lambda a^{5-\alpha} + 6C_{\kappa,[1]}C_{\kappa,[2]}\lambda^2 a^{4-\alpha} + 3C_{\kappa,[2]}^2\lambda^2 a^{6-2\alpha}, \end{aligned} \quad (5.7)$$

as $a \rightarrow \infty$. Now, the behavior of $m_4^0(a)$ is different depending on the slow or fast growth regime, yielding

$$m_4^0(a) \sim \begin{cases} C_{\kappa,[4]} \lambda a^{5-\alpha}, & \text{slow growth,} \\ 3C_{\kappa,[2]}^2 \lambda^2 a^{6-2\alpha}, & \text{fast growth.} \end{cases} \tag{5.8}$$

The reader is encouraged to check that the terms $\lambda a^{5-\alpha}$ and $\lambda^2 a^{6-2\alpha}$ are indeed dominant in (5.7) in the slow and fast growth regimes, respectively.

Similarly, the relation (A.4) gives that

$$m_5^0(a) \sim C_{\kappa,[1]} \lambda a + 15C_{\kappa,[2]} \lambda a^{3-\alpha} + 25C_{\kappa,[3]} \lambda a^{4-\alpha} + 10C_{\kappa,[4]} \lambda a^{5-\alpha} + C_{\kappa,[5]} \lambda a^{6-\alpha} + 10C_{\kappa,[1]}^2 \lambda^2 a^2 + 40C_{\kappa,[1]} C_{\kappa,[2]} \lambda^2 a^{4-\alpha} + 10C_{\kappa,[1]} C_{\kappa,[3]} \lambda^2 a^{5-\alpha} + 30C_{\kappa,[2]}^2 \lambda^2 a^{6-2\alpha} + 10C_{\kappa,[2]} C_{\kappa,[3]} \lambda^2 a^{7-2\alpha}, \tag{5.9}$$

as $a \rightarrow \infty$ and therefore,

$$m_5^0(a) \sim \begin{cases} C_{\kappa,[5]} \lambda a^{6-\alpha}, & \text{slow growth,} \\ 10C_{\kappa,[2]} C_{\kappa,[3]} \lambda^2 a^{7-2\alpha}, & \text{fast growth.} \end{cases} \tag{5.10}$$

We also get from (A.5) that

$$m_6^0(a) \sim C_{\kappa,[1]} \lambda a + 31C_{\kappa,[2]} \lambda a^{3-\alpha} + 90C_{\kappa,[3]} \lambda a^{4-\alpha} + 65C_{\kappa,[4]} \lambda a^{5-\alpha} + 15C_{\kappa,[5]} \lambda a^{6-\alpha} + C_{\kappa,[6]} \lambda a^{7-\alpha} + 25C_{\kappa,[1]}^2 \lambda^2 a^2 + 180C_{\kappa,[1]} C_{\kappa,[2]} \lambda^2 a^{4-\alpha} + 110C_{\kappa,[1]} C_{\kappa,[3]} \lambda^2 a^{5-\alpha} + 15C_{\kappa,[1]} C_{\kappa,[4]} \lambda^2 a^{6-\alpha} + 195C_{\kappa,[2]}^2 \lambda^2 a^{6-2\alpha} + 150C_{\kappa,[2]} C_{\kappa,[3]} \lambda^2 a^{7-2\alpha} + 25C_{\kappa,[1]}^2 \lambda^2 a^{8-2\alpha} + 10C_{\kappa,[2]} C_{\kappa,[4]} \lambda^2 a^{8-2\alpha} + 15C_{\kappa,[1]}^3 \lambda^3 a^3 + 45C_{\kappa,[1]}^2 C_{\kappa,[2]} \lambda^3 a^{5-\alpha} + 45C_{\kappa,[1]} C_{\kappa,[2]}^2 \lambda^3 a^{7-2\alpha} + 15C_{\kappa,[2]}^2 \lambda^3 a^{9-3\alpha}, \tag{5.11}$$

as $a \rightarrow \infty$, which yields

$$m_6^0(a) \sim \begin{cases} C_{\kappa,[6]} \lambda a^{7-\alpha}, & \text{slow growth,} \\ 15C_{\kappa,[2]}^3 \lambda^3 a^{9-3\alpha}, & \text{fast growth.} \end{cases} \tag{5.12}$$

Similarly, from (A.6),

$$m_7^0(a) \sim \begin{cases} C_{\kappa,[7]} \lambda a^{8-\alpha}, & \text{slow growth,} \\ 105C_{\kappa,[2]}^2 C_{\kappa,[3]} \lambda^3 a^{10-3\alpha}, & \text{fast growth.} \end{cases} \tag{5.13}$$

The relations above lead us to conjecture that, for $r \geq 2$ and $a \rightarrow \infty$,

$$m_r^0(a) \sim \begin{cases} C_{\kappa,[r]} \lambda a^{r-\alpha+1}, & \text{slow growth,} \\ \frac{r! C_{\kappa,[2]}^{r/2}}{(r/2)! 2^{r/2}} \lambda^{r/2} a^{(3-\alpha)r/2}, & \text{fast growth and even } r, \\ \frac{r! C_{\kappa,[2]}^{(r-1)/2-1} C_{\kappa,[3]}^2}{((r-1)/2-1)! 2^{(r-1)/2-1} 3!} \lambda^{(r-1)/2} a^{(3-\alpha)(r-1)/2+1}, & \text{fast growth and odd } r, \end{cases} \tag{5.14}$$

In fact, we have checked the conjecture (5.14) not only up to the seventh central moment but up to the tenth central moment. (The formulae relating the central moments and factorial moments naturally get quite lengthy for larger r and are therefore not included in Appendix A.)

Remark 5.1. The difficulty in proving (5.14) in general was indicated above but it is instructive to provide some further insight. First, we note that the proof of Corollary 3.3 cannot be used directly to show (5.14).

Indeed, the terms associated with (3.15) in the proof of the corollary lead to $\kappa_{[r]}(a) \sim c\lambda a^{r-(\alpha+1)}$ but this term is no longer necessarily dominant in the fast regime. For example, note the presence of $\lambda a^{r-(\alpha-1)} = \lambda a^{7-\alpha}$ in (5.11) when $r = 6$. But this term is indeed dominated by $\lambda^3 a^{9-3\alpha}$ in the fast regime. Second, the actual difficulty is in tracking the dominant term. For example, when $r = 6$, the dominant term arises from $\kappa_{[2]}(a)^3 \sim c\lambda^3 a^{6-3(\alpha-1)}$ which enters into the moment $m_6(a)$ through (3.12). But, for example, $m_6(a)$ also contains the term $\kappa_{[1]}(a)^3 \kappa_{[3]}(a) \sim c\lambda^4 a^{6-4(\alpha-1)}$. Though this term dominates $\kappa_{[2]}(a)^3$, it does not appear in (5.11) since it gets canceled once substituted into (2.17).

Several interesting observations can be made concerning (5.14). First, note that the conjectured behavior in the slow growth regime is exactly the same as in (3.10) for a fixed λ . Second, it is interesting to compare (5.14) with the available results concerning the large scale behavior of PCPs at large scales. In the slow growth regime, one has

$$\left\{ \frac{N_e(0, au) - \mathbb{E}N_e(0, au)}{a^{1/\alpha}} \right\}_{u \in [0,1]} \xrightarrow{fdd} \{L_\alpha(u)\}_{u \in [0,1]}, \quad \text{as } a \rightarrow \infty, \tag{5.15}$$

where the convergence is in the sense of finite-dimensional distributions and L_α is an α -stable Lévy motion (Mikosch and Samorodnitsky [27], Prop. 5.11). In the fast growth regime, on the other hand,

$$\left\{ \frac{N_e(0, au) - \mathbb{E}N_e(0, au)}{\lambda^{1/2} a^{(3-\alpha)/2}} \right\}_{u \in [0,1]} \xrightarrow{fdd} \{B_H(u)\}_{u \in [0,1]}, \quad \text{as } a \rightarrow \infty, \tag{5.16}$$

where B_H is fractional Brownian motion with the self-similarity parameter $H = (3 - \alpha)/2$ (Mikosch and Samorodnitsky [27], Prop. 4.7). The use of a in (5.15) and (5.16) are somewhat misleading since for these results, a is usually thought as the length of the observation window $(0, a)$, whereas a is the analysis scale throughout this work. We nevertheless use a for simplicity, as well as to indicate that the behavior (5.14) is consistent with the normalization used in (5.16) when r is even. Interestingly, this is not the case for odd r , showing that some moments of the left-hand side of (5.16) do not converge to those of the limiting process.

The practical implications of the slow and fast regimes, and of the scaling relations in (5.14) will be discussed in greater detail in Sections 6 and 7 below.

Remark 5.2. We have analyzed above the behavior of central moments under the slow and fast growth conditions. In the case of the factorial cumulants $\kappa_{[r]}(a)$, an inspection of the proof of Proposition 3.2 reveals that the asymptotic result in (3.2) holds irrespective of the slow and fast growth regimes. The same conclusion can be reached for the factorial moments $m_{[r]}(a)$ and the usual moments $m_r(a)$, with the respective asymptotics in (3.8) and (3.9) valid for both regimes. The case of the usual cumulants $\kappa_r(a)$ is more delicate. But an examination of the first ten cumulants leads us to conjecture that the asymptotics in (3.16) holds for both regimes as well.

Remark 5.3. As mentioned in Section 1, our analysis of the moments of PCP at large scales is closest in the spirit to that of the moments of renewal point processes carried out by Dombry and Kaj [9]. In contrast to the approach taken here, Dombry and Kaj [9] work with somewhat more general moment measures. The different growth regimes in superimposing renewal point processes are considered by Dombry and Kaj [9] but *not* for the behavior of the moment measures. We have not aimed specifically to be different from Dombry and Kaj [9] but just became aware of their work towards the end of this project.

6. TRANSITIONS BETWEEN DIFFERENT SCALING REGIMES

The analysis carried out in Sections 3 and 4 shows the existence of biscaling for all the moments and cumulants considered, that is, the different scaling behaviors at large and small scales. Moreover, the study of Section 5 revealed a further biscaling of central moments at large scales, that is, the different scaling behaviors in the slow and fast growth regimes. Since the scaling behaviors are different depending on the situation, one could expect

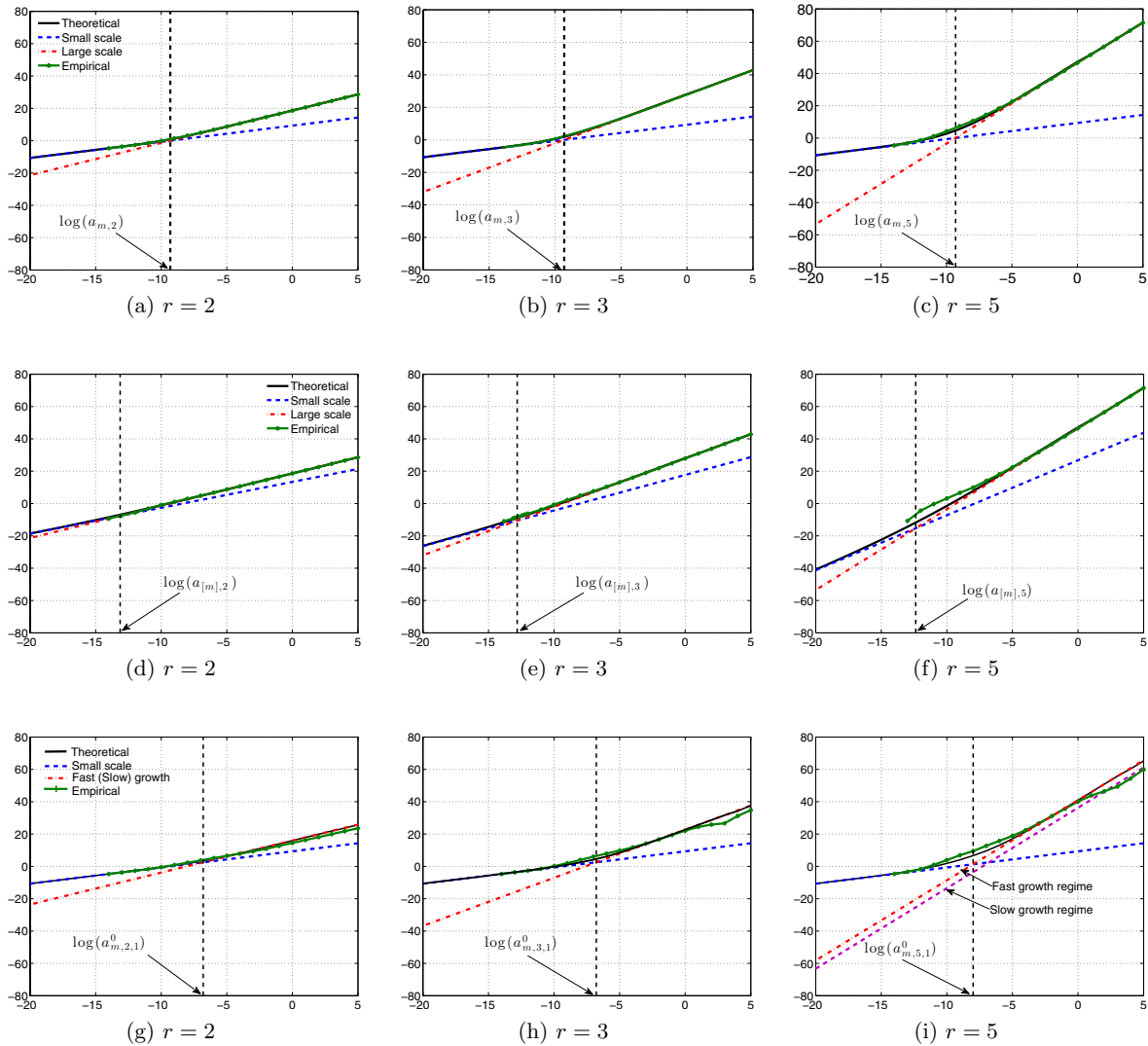


FIGURE 1. *Top*: moments (log-log scale); *Middle*: factorial moments (log-log scale); *Bottom*: central moments (log-log scale).

them to be separated by a “knee,” the transition scale (or the range of such) where the behavior changes as one moves from one scaling behavior to another. The biscaling and the “knee” are clearly seen in Figure 1 (to be discussed in more detail in Sect. 7) where the factorial moments, moments and central moments are plotted on the log scale for both axes.

The location of the “knee” (*i.e.*, the transition scale a) can be approximated by equating the scaling relations for different scaling behaviors and quantities of interest, and solving with respect to a . For the factorial cumulants, by equating the relations (3.2) and (4.3), we obtain

$$C_{\kappa,[r]} \lambda a^{r-\alpha+1} = c_{\kappa,[r]} \lambda a^{1+(r-1)\theta}, \quad r \geq 2,$$

and solving for a leads to the approximate location of the “knee” (the transition scale from small to large scales) given by $a_{\kappa,[r]} = (c_{\kappa,[r]}/C_{\kappa,[r]})^{1/(r-(r-1)\theta-\alpha)}$ or

$$a_{\kappa,[r]} = \left(\frac{(r-2)!C_{F,r-1}(\mathbb{E}A)^{r-\alpha}(\alpha-1)(r-\alpha)(r+1-\alpha)\bar{R}_r}{((r-1)\theta+1)C_W} \right)^{1/(r-(r-1)\theta-\alpha)}, \quad r \geq 2, \tag{6.1}$$

by using the forms of the coefficients $C_{\kappa,[r]}$ and $c_{\kappa,[r]}$. Similarly, from Corollaries 3.3 and 4.3, the locations of the “knee” (the transition scales) in the transition from small to large scales for the factorial moments and the usual moments are approximated by, respectively,

$$a_{m,[r]} = \left(\frac{r!C_{F,r-1}\bar{R}_r}{\lambda^{r-1}((r-1)\theta+1)(\mathbb{E}W)^r} \right)^{1/(r-(r-1)\theta-1)}, \quad \theta \in (0, 1), \tag{6.2}$$

$$a_{m,r} = \frac{1}{\lambda\mathbb{E}W}, \quad r \geq 2. \tag{6.3}$$

We note that the location of the “knee” for the moments does not depend on the order r and the distribution of the interarrival times A . When $\theta > 1$, the factorial moments have the same asymptotics for both large and small scales. When $\theta = 1$, the asymptotics are the same in a but differ in the multiplicative constants, and our approach does not yield a transition scale.

The case of the central moments is more delicate due to the effect of the slow and fast growth regimes discussed in Section 5. We note first that for fixed λ and as a increases, the fast growth regime comes before the slow growth regime; indeed, as $a \rightarrow \infty$, $\lambda/a^{\alpha-1} \rightarrow 0$ for fixed λ . Thus, for the central moments, we shall distinguish between two transition scales $a_{m,r,1}^0$ and $a_{m,r,2}^0$: $a_{m,r,1}^0$ being the transition scale (“knee”) from small scales into the fast growth, and $a_{m,r,2}^0$ being the transition scale (“knee”) from the fast growth into the slow growth. By equating (4.11) and (5.14) (in the fast regime) and solving for a leads to

$$a_{m,r,1}^0 = \begin{cases} \left(\frac{r!C_{\kappa,[2]}^{r/2}}{\mathbb{E}W(r/2)!2^{r/2}} \lambda^{r/2-1} \right)^{-2/((3-\alpha)r-2)}, & \text{even } r \geq 2, \\ \left(\frac{r!C_{\kappa,[2]}^{(r-1)/2-1}C_{\kappa,[3]}}{\mathbb{E}W((r-1)/2-1)!2^{(r-1)/2-1}3!} \lambda^{(r-1)/2-1} \right)^{-2/((r-1)(3-\alpha))}, & \text{odd } r \geq 2. \end{cases} \tag{6.4}$$

Similarly, equating (5.14) in the fast and slow regimes leads to

$$a_{m,r,2}^0 = \begin{cases} \left(\frac{r!C_{\kappa,[2]}^{r/2}}{C_{\kappa,[r]}(r/2)!2^{r/2}} \lambda^{r/2-1} \right)^{2/((r-2)(\alpha-1))}, & \text{even } r \geq 4, \\ \left(\frac{r!C_{\kappa,[2]}^{(r-1)/2-1}C_{\kappa,[3]}}{C_{\kappa,[r]}((r-1)/2-1)!2^{(r-1)/2-1}3!} \lambda^{(r-1)/2-1} \right)^{2/((r-3)(\alpha-1))}, & \text{odd } r \geq 4. \end{cases} \tag{6.5}$$

(Recall from Sect. 5 that there is difference in the two regimes only when $r \geq 4$). As will be seen in Section 7, the transition scale $a_{m,r,2}^0$ may be too large to observe in practice.

For the usual cumulants, the location of the transition scale $a_{\kappa,r}$, $r \geq 2$, from small to large scales follows from (3.16) and (4.14), to yield

$$a_{\kappa,r} = \left(\frac{(\alpha-1)(r-\alpha)(r+1-\alpha)\mathbb{E}W(\mathbb{E}A)^{r-\alpha}}{r(r-1)C_W} \right)^{1/(r-\alpha)}, \quad r \geq 2. \tag{6.6}$$

Finally, the various regimes and scaling relations are summarized in Table 1 of Section 1. They will be examined through a numerical study and an application to real data in the next section.

7. NUMERICAL STUDY AND APPLICATION TO INTERNET TRAFFIC

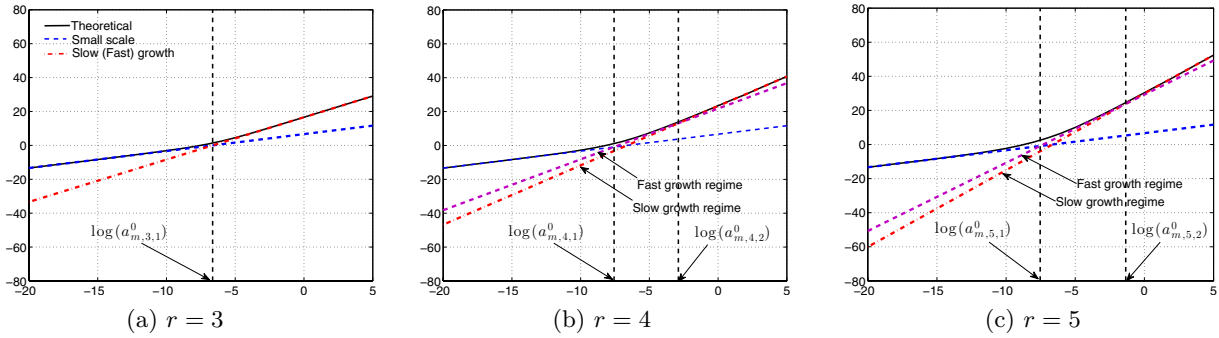
In this section, we illustrate the scaling relations of the various moments of PCP through a numerical study by using an Internet traffic data set (which we find more interesting and illuminating than using synthetic data). We exclude cumulants from the discussion for shortness sake. We consider a publicly available Internet trace, Auckland⁵, which is one hour long and consists of 38 308 012 packets which make 1 371 756 flows.

We shall first describe how the parameters are fitted to the data trace. The flow (Poisson) arrival parameter λ is estimated directly from the sample mean of flow interarrival times, yielding $\hat{\lambda} = 396$ (flows/sec). We choose the flow (cluster) size W to be zeta distributed, which is taken to be heavy-tailed and may be thought of as a discrete counterpart of the Pareto distribution, with p.m.f. $p_W(w) = 1/(w^\alpha \zeta(\alpha))$, $w \geq 1$, $\alpha > 1$, where $\zeta(\alpha)$ is the Riemann zeta function. Note that the mean of W is $\alpha/(\alpha - 1)$. Calculating the empirical value of the mean size of flows in the trace results in $\hat{\alpha} = 1.02$. The distribution of interarrival times between packets of a flow (between points of a cluster) is often modeled by a gamma distribution with parameters θ and β (see (4.2)). However, determining the appropriate parameters is not trivial as pointed out by Hohn *et al.* [18] and a similar approach as in their work is considered here. One of the quantities of interest is the packet arrival rate within a flow $1/\mathbb{E}A = \beta/\theta$. An estimate for this in-flow packet arrival rate using the median or the mean of the empirical rates of all flows performs poorly. Since PCP represents the overall packet arrival process, it is essential to capture the impact of each value of the rate of a flow in terms of its packets. Therefore, the rate is weighed by the number of interarrival times in each flow. This results in an estimate for β/θ that is generally considerably above a simple mean. The parameter θ is tuned to fit the estimation of the scaling exponent function of the factorial moments over small scales in (4.9). The fitting procedure using the second empirical factorial moment yields $\hat{\theta} = 0.60$ and the in-flow rate then results in $\hat{\beta} = 526.32$.

Plots (a)–(c) in Figure 1 show the usual moments $m_r(a)$ against scale a , for $r = 2, 3, 5$, using the natural logarithmic scale for the two axis. (The first moment yields a straight line which is not very informative.) We compute the *theoretical* values of the moments first using the formula (2.7) and then the relations (2.12) and (2.15) based on the estimated parameters. The *empirical* values of the moments have been computed through the number of packet arrivals on contiguous non-overlapping intervals of size a over all trace duration, with the smallest value for a being 10^{-6} sec (≈ -13.8 in the log scale; the packet arrivals were extracted with increments of 1 microsecond). We also include the dashed lines corresponding to the scaling relations of the moments at *small scales* and *large scales* using (3.9) and (4.10), respectively. Note from the plots that these relations hold for a wide range of values. The vertical dotted line depicts with a good accuracy the transition between the small and large scales computed through the log of (6.3). Plots (a)–(c) show that the log-moments (theoretical line) of the PCP fit well the empirical values with a small deviation around the transition between scales for the fifth (and higher) moments. (This is also observed for the factorial moments and central moments below.) It is in this region of the transition between time scales where potential differences between PCP and the data are more pronounced. The discrepancy might be due to the fact that in the trace considered here, large flows tend to have shorter interarrival times. In fact, the more general PCP considered by Westcott [34], allows the interarrival times A to be non-identically distributed and dependent on the cluster (flow) size. Another potential cause is the impact of the flow size distribution. We plan to examine these possibilities further elsewhere. However, the fits in the plots of Figure 1 are already quite acceptable so that any payoff might be minimal at the expense of oversophistication of the model.

Plots (d)–(f) in Figure 1 show the analogous plots for the factorial moments $m_{[r]}(a)$, for $r = 2, 3, 5$. We compute the empirical factorial moments using the number of packet arrivals over contiguous non-overlapping intervals of size a over all trace duration. The quality of the fit is good with the exception around the transition between scales for the higher moment possibly due to the shorter interarrival times of packets of large flows as mentioned above. We also note that the empirical line does not extend as far for small values of a as for the usual moments. It is especially difficult to estimate the factorial moments for small a because of the order of

⁵Auckland IX, file 20080327-080000-0, Available: <http://wand.net.nz/wits/auck/9/>

FIGURE 2. Central moments (log-log scale) with $\alpha = 1.5$.

the theoretical values. In this case, a trace with a longer duration is needed for more intervals of size a to be used in the estimation.

The central moments $m_r^0(a)$ for the same order r values are depicted in plots (g)–(i) of Figure 1. We point out that in computing the theoretical values for large a using the factorial cumulants (2.7) and then the relations to central moments, the truncation parameter in the first sum of (2.7) has to be much larger compared with the other moments. This is also reflected in the empirical central moments, where the variation in these moments for larger a means that the number of contiguous non-overlapping intervals of length a used in the estimation is insufficient.

Note that the plot (i) of Figure 1 also includes the lines corresponding to the slow and fast growth regimes according to (5.14). (As discussed in Sect. 5, the scaling behavior for $r = 2, 3$ is the same for the two regimes.) From (5.14), the slopes of the two lines are $r - \alpha + 1 = 6 - \alpha$ (slow growth) and $(3 - \alpha)(r - 1)/2 + 1 = 7 - 2\alpha$ (fast growth). For the choice of $\alpha = 1.02$ used above, these slopes are nearly identical, which can also be seen from the plot. Although note that the intercepts of the two lines are quite different. The nearly identical slopes of the two lines also translate into the fact that they cross at $a_{m,5,2}^0$ in (6.5) which is too large to be observable for this trace (the actual value of $a_{m,5,2}^0$ is approximately 1.8761×10^{91}). Thus, the slow growth regime is not observed either, and the theoretical values follow closely the line corresponding to the fast growth for all observable large scales. The transition scale $a_{m,5,1}^0$ computed from (6.4) captures well the change from the small scales to the large scales in the fast growth regime. Finally, we also note that observing the fast growth scaling only is consistent with the fast growth condition itself: for the chosen $\lambda = 396$ flows/sec and $\alpha = 1.02$, the ratio $\lambda/a^{\alpha-1}$ in (5.1) and (5.2) is relatively large for the largest a considered in plot (i) of Figure 1.

In plot (i) of Figure 1 only the scaling in the fast growth regime was observable since the value $\alpha = 1.02$ is very close to 1 (see the discussion above). In order to illustrate the possibility of observing the different scaling behaviors of the fast and slow growth regimes, we repeated the numerical study with the same parameters but changing $\alpha = 1.02$ to $\alpha = 1.5$. The results are reported in Figure 2 for $r = 3, 4$ and 5 (with no distinction between the slow and fast growth scalings for $r = 3$). For $r = 5$, especially, we can now see the transition of the scaling behavior at $a_{m,5,1}^0$ from small scales into large scales associated with the fast growth, and then at $a_{m,5,2}^0$ from the fast growth scaling into the slow growth scaling. The findings are also consistent with the fact that $\lambda/a^{\alpha-1} = 396/a^{1/2}$ is now smaller for large observable scales a .

8. CONCLUSIONS AND DIRECTIONS FOR FUTURE WORK

In this work, we derived the asymptotic behaviors at large and small scales of the various moments and cumulants of the PCP, and compared them through the information carried in terms of the model parameters. At large scales, the slow and fast growth regimes were introduced in the literature to establish the convergence

of an appropriately normalized PCP. We analyzed the behavior of the central moments under both regimes and showed through numerical results how the two regimes could be observed depending on the model parameters and the observation window. We identified the transition scales where the behavior changes as one moves from one scaling behavior to another. The derived results can contribute in developing more robust estimation procedures, avoiding misleading inference while studying the scaling behavior of the current Internet traffic, including the controversial multifractality at small scales, a direction that we shall pursue in a future work.

In other future work, it would be interesting to establish the same small scale behavior of the (central) moments and cumulants under more general assumptions on the distribution of the interarrival times between the points in a cluster - we assumed above a particular type of the distribution motivated by the Internet traffic but the resulting scaling relations do not involve the specific model parameters. As another extension, the interarrival times could be assumed to be non-identically distributed and dependent on the cluster size.

APPENDIX A. FORMULAE RELATING CENTRAL MOMENTS AND FACTORIAL CUMULANTS

For Section 5, the relations between the first seven central moments and factorial cumulants using (2.17) along with (2.15) and (2.12) are listed below. We use the variable t instead of a as in Section 5 to emphasize that these are general relations between central moments and factorial cumulants.

$$m_2^0(t) = \kappa_{[1]}(t) + \kappa_{[2]}(t), \quad (\text{A.1})$$

$$m_3^0(t) = \kappa_{[1]}(t) + 3\kappa_{[1]}(t) + \kappa_{[3]}(t), \quad (\text{A.2})$$

$$m_4^0(t) = \kappa_{[1]}(t) + 3\kappa_{[1]}^2(t) + 7\kappa_{[2]}(t) + 6\kappa_{[1]}(t)\kappa_{[2]}(t) + 3\kappa_{[2]}^2(t) + 6\kappa_{[3]}(t) + \kappa_{[4]}(t), \quad (\text{A.3})$$

$$m_5^0(t) = \kappa_{[1]}(t) + 10\kappa_{[1]}^2(t) + 15\kappa_{[2]}(t) + 40\kappa_{[1]}\kappa_{[2]}(t) + 30\kappa_{[2]}^2(t) + 25\kappa_{[3]}(t) \\ + 10\kappa_{[1]}(t)\kappa_{[3]}(t) + 10\kappa_{[2]}(t)\kappa_{[3]}(t) + 10\kappa_{[4]}(t) + \kappa_{[5]}(t), \quad (\text{A.4})$$

$$m_6^0(t) = \kappa_{[1]}(t) + 25\kappa_{[1]}^2(t) + 15\kappa_{[1]}^3(t) + 31\kappa_{[2]}(t) + 180\kappa_{[1]}(t)\kappa_{[2]}(t) + 45\kappa_{[1]}^2(t)\kappa_{[2]}(t) \\ + 195\kappa_{[2]}^2(t) + 45\kappa_{[1]}(t)\kappa_{[2]}^2(t) + 15\kappa_{[2]}^3(t) + 90\kappa_{[3]}(t) + 110\kappa_{[1]}(t)\kappa_{[3]}(t) + 150\kappa_{[2]}(t)\kappa_{[3]}(t) \\ + 10\kappa_{[3]}^2(t) + 65\kappa_{[4]}(t) + 15\kappa_{[1]}(t)\kappa_{[4]}(t) + 15\kappa_{[2]}(t)\kappa_{[4]}(t) + 15\kappa_{[5]}(t) + \kappa_{[6]}(t), \quad (\text{A.5})$$

$$m_7^0(t) = \kappa_{[1]}(t) + 56\kappa_{[1]}^2(t) + 105\kappa_{[1]}^3(t) + 63\kappa_{[2]}(t) + 686\kappa_{[1]}(t)\kappa_{[2]}(t) + 525\kappa_{[1]}^2(t)\kappa_{[2]}(t) + 1050\kappa_{[2]}^2(t) \\ + 735\kappa_{[1]}(t)\kappa_{[2]}^2(t) + 315\kappa_{[2]}^3(t) + 301\kappa_{[3]}(t) + 770\kappa_{[1]}(t)\kappa_{[3]}(t) + 105\kappa_{[1]}^2(t)\kappa_{[3]}(t) + 1400\kappa_{[2]}(t)\kappa_{[3]}(t) \\ + 210\kappa_{[1]}(t)\kappa_{[2]}(t)\kappa_{[3]}(t) + 105\kappa_{[2]}^2(t)\kappa_{[3]}(t) + 210\kappa_{[3]}^2(t) + 350\kappa_{[4]}(t) + 245\kappa_{[1]}(t)\kappa_{[4]}(t) + 315\kappa_{[2]}(t)\kappa_{[4]}(t) \\ + 35\kappa_{[3]}(t)\kappa_{[4]}(t) + 140\kappa_{[5]}(t) + 21\kappa_{[1]}(t)\kappa_{[5]}(t) + 21\kappa_{[2]}(t)\kappa_{[5]}(t) + 21(t)\kappa_{[6]}(t) + \kappa_{[7]}(t). \quad (\text{A.6})$$

APPENDIX B. FACTORIAL CUMULANTS OF PCP

We derive here the formulae (2.6)–(2.7) for the factorial cumulants of PCP. As noted following the formulae, they appears in Westcott [34] when the first points in the clusters are excluded. We shall modify slightly the argument of Westcott [34] to include the first points in the clusters, leading to the exact same formulae (2.6)–(2.7).

The factorial cumulants are obtained through the formula (2.5) based on the probability generating function $P_t(z)$ in (2.4). For the equilibrium process, $P_t(z)$ is defined as the limit

$$P_t(z) = \lim_{x \rightarrow \infty} \mathbb{E} z^{N(x, x+t)},$$

where N is the transient PCP. We need first to introduce some notation. Let $S_j, j \geq 1$, be the Poisson arrivals of the first points of the clusters, and $X_{j,k}, k \geq 0$, be the distances of the cluster points from the first point in cluster j , with $X_{j,0} = 0$. Let $X_k, k \geq 0$, be the generic distances of cluster points, with $X_0 = 0$ and the understanding that there is a finite but random number of X_k 's in a cluster. The point process consisting of the points X_k is referred to as a subsidiary point process in Westcott [34]. To make connection to Westcott [34], we shall denote the subsidiary process by $N^{(s)}$ when the first point $X_0 = 0$ is excluded, and also let

$$P(z; a, b) = \mathbb{E} z^{N^{(s)}(a,b)}, \quad P(z; b) = P(z; 0, b) = \mathbb{E} z^{N^{(s)}(0,b)}, \quad P(z) = P(z; \infty),$$

that is, the probability generating functions associated with the subsidiary process $N^{(s)}$ (excluding $X_0 = 0$).

With the introduced notation and letting $h_x(y) = z^{1(x, x+t)}(y)$, we get that

$$\log P_t(z) = \lim_{x \rightarrow \infty} \log \mathbb{E} \prod_{j=1}^{\infty} \prod_{k=0}^{\infty} h_x(S_j + X_{j,k}) = \lim_{x \rightarrow \infty} \left(-\lambda \int_0^{\infty} \left(1 - \mathbb{E} \prod_{k=0}^{\infty} h_x(u + X_k) \right) du \right),$$

where we used the fact that the probability generating functional of a Poisson process is given by $\mathbb{E} \prod_{j=1}^{\infty} g(S_j) = \exp\{-\lambda \int_0^{\infty} (1 - g(u)) du\}$ for suitable deterministic functions g . Splitting the integral \int_0^{∞} into \int_x^{x+t} and \int_0^x , it follows that

$$\begin{aligned} \log P_t(z) &= \lim_{x \rightarrow \infty} \left(-\lambda \int_x^{x+t} \left(1 - \mathbb{E} \prod_{k=0}^{\infty} z^{1(x, x+t)}(u+X_k) \right) du - \lambda \int_0^x \left(1 - \mathbb{E} \prod_{k=0}^{\infty} z^{1(x, x+t)}(u+X_k) \right) du \right) \\ &= -\lambda \int_0^t \left(1 - \mathbb{E} \prod_{k=0}^{\infty} z^{1(0,t)}(v+X_k) \right) dv - \lambda \int_0^{\infty} \left(1 - \mathbb{E} \prod_{k=0}^{\infty} z^{1(v, v+t)}(X_k) \right) dv, \end{aligned}$$

after the change of variables $v = u - x$ for the first integral, and $v = x - u$ for the second integral and letting $x \rightarrow \infty$. Since $v + X_0 = v \in (0, t)$ for $v \in (0, t)$ (for the first integral above) and $X_0 = 0 \notin (v, v + t)$ for $v > 0$, we get further that

$$\begin{aligned} \log P_t(z) &= -\lambda \int_0^t \left(1 - z \mathbb{E} \prod_{k=1}^{\infty} z^{1(0,t)}(v+X_k) \right) dv - \lambda \int_0^{\infty} \left(1 - \mathbb{E} \prod_{k=1}^{\infty} z^{1(v, v+t)}(X_k) \right) dv \\ &= -\lambda \left(\int_0^t (1 - zP(z; t - v)) dv + \int_0^{\infty} (1 - P(z; v, v + t)) dv \right) \\ &= -\lambda \left(\int_0^t (1 - zP(z; v)) dv + \int_0^{\infty} (1 - P(z; v, v + t)) dv \right), \end{aligned} \tag{B.1}$$

by using the notation above and another change of variables ($t - v$ to v in the first integral). We shall next evaluate the two integrals in (B.1).

Let now $W^{(0)} = W - 1$ be the number of points in a cluster excluding the first point, and $R_w^{(0)} = \mathbb{P}(W^{(0)} \geq w) = \mathbb{P}(W - 1 \geq w) = \mathbb{P}(W \geq w + 1) = R_{w+1}$. As shown in Westcott [34], equation (4),

$$P(z; u) = P(z) + (1 - z) \sum_{j=0}^{\infty} z^j R_{j+1}^{(0)} (1 - F_{j+1}(u)).$$

Note that

$$P(z) = \mathbb{E}z^{N^{(s)}(0,\infty)} = \mathbb{E}z^{W^{(0)}} = \sum_{j=0}^{\infty} z^j \mathbb{P}(W^{(0)} = j) = \sum_{j=0}^{\infty} z^j (R_j^{(0)} - R_{j+1}^{(0)})$$

(with $R_0^{(0)} = 1$). After simple algebraic manipulations, this leads to

$$P(z; u) = 1 + (z - 1) \sum_{j=0}^{\infty} z^j R_{j+1}^{(0)} F_{j+1}(u).$$

Then, for the first integral in (B.1),

$$\int_0^t (1 - zP(z; v))dv = -(z - 1)t - (z - 1) \sum_{j=0}^{\infty} z^{j+1} R_{j+1}^{(0)} \int_0^t F_{j+1}(u)du. \tag{B.2}$$

As shown in Westcott [34] (see the arguments following Eq. (19) and terminating with Thm. 4), the second integral in (B.1) can be written as

$$\int_0^{\infty} (1 - P(z; v, v + t))dv = -(z - 1) \sum_{k=1}^{\infty} z^{k-1} J_k \sum_{j=0}^{\infty} R_{j+k}^{(0)},$$

where

$$J_k = \int_0^t (F_{k-1}(x) - F_k(x))dx,$$

with $F_0(x) = 1$. Basic algebraic manipulations lead to

$$\begin{aligned} \int_0^{\infty} (1 - P(z; v, v + t))dv &= -(z - 1)t\mathbb{E}W^{(0)} \\ &- (z - 1)^2 \sum_{k=1}^{\infty} z^{k-1} \int_0^t F_k(x)dx \sum_{j=1}^{\infty} R_{j+k}^{(0)} + (z - 1) \sum_{k=1}^{\infty} z^{k-1} \int_0^t F_k(x)dx R_k^{(0)}. \end{aligned} \tag{B.3}$$

By using (B.2) and (B.3), we can express (B.1) as

$$\begin{aligned} \log P_t(z) &= (z - 1)\lambda t + (z - 1)^2 \lambda \sum_{j=0}^{\infty} z^j R_{j+1}^{(0)} \int_0^t F_{j+1}(x)dx \\ &+ (z - 1)\lambda t\mathbb{E}W^{(0)} + (z - 1)^2 \lambda \sum_{k=1}^{\infty} z^{k-1} \int_0^t F_k(x)dx \sum_{j=1}^{\infty} R_{j+k}^{(0)} \\ &= \lambda(z - 1) \left(t + t\mathbb{E}W^{(0)} + (z - 1) \sum_{k=1}^{\infty} z^{k-1} \int_0^t F_k(x)dx \sum_{j=0}^{\infty} R_{j+k}^{(0)} \right). \end{aligned}$$

By noting that $1 + \mathbb{E}W^{(0)} = \mathbb{E}W$ and that

$$\sum_{j=0}^{\infty} R_{j+k}^{(0)} = \sum_{j=0}^{\infty} \mathbb{P}(W^{(0)} \geq j + k) = \sum_{j=0}^{\infty} \mathbb{P}(W \geq j + k + 1) = \sum_{j=1}^{\infty} R_{j+k},$$

we get further that

$$\log P_t(z) = \lambda(z - 1) \left(t\mathbb{E}W + (z - 1) \sum_{k=1}^{\infty} z^{k-1} \int_0^t F_k(x)dx \sum_{j=1}^{\infty} R_{j+k} \right).$$

By using (2.5), this yields the formulae (2.6)–(2.7).

Acknowledgements. The authors would like to thank the anonymous reviewer for carefully reading the manuscript and for giving very detailed and helpful comments and suggestions. The second author was supported in part by the NSA grant H98230-13-1-0220.

REFERENCES

- [1] P. Abry, R. Baraniuk, P. Flandrin, R. Riedi and D. Veitch, Multiscale nature of network traffic. *IEEE Signal Proc. Magazine* **3** (2002) 28–46.
- [2] P. Abry, P. Borgnat, F. Ricciato, A. Scherrer and D. Veitch, Revisiting an old friend: on the observability of the relation between long range dependence and heavy tail. *Telecommun. Syst.* **43** (2010) 147–165.
- [3] N. Antunes and V. Pipiras, Estimation of Flow Distributions from Sampled Traffic. *ACM Trans. Model. Perform. Eval. Comput. Syst.* **1** (2016) 11:1–11:28.
- [4] M.S. Bartlett, The spectral analysis of point processes. *J. Roy. Stat. Soc. Series B. Methodological* **25** (1963) 264–296.
- [5] P. Carruthers, E.M. Friedlander, C.C. Shih and R.M. Weiner, Multiplicity fluctuations in finite rapidity windows. Intermittency or quantum statistical correlation? *Phys. Lett. B* **222** (1989) 487–492.
- [6] D.R. Cox and V. Isham, Point Processes. *Monographs on Applied Probability and Statistics*. Chapman and Hall, London-New York (1980)
- [7] D.J. Daley and D. Vere-Jones, An Introduction to the Theory of Point Processes. Elementary Theory and Methods. Vol. I, *Probability and its Applications* (New York), 2nd edn. Springer-Verlag, New York (2003).
- [8] E.A. De Wolf, I.M. Dremin and W. Kittel, Scaling laws for density correlations and fluctuations in multiparticle dynamics. *Phys. Rep.* **270** (1996) 1–141.
- [9] C. Dombry and I. Kaj, Moment measures of heavy-tailed renewal point processes: asymptotics and applications. *ESAIM. PS* **17** (2013) 567–591.
- [10] V. Fasen and G. Samorodnitsky, A fluid cluster Poisson input process can look like a fractional Brownian motion even in the slow growth aggregation regime. *Adv. Appl. Prob.* **41** (2009) 393–427.
- [11] G. Faÿ, B. González-Arévalo, T. Mikosch and G. Samorodnitsky, Modeling teletraffic arrivals by a Poisson cluster process. *Queueing Syst.* **54** (2006) 121–140.
- [12] R. Gaigalas and I. Kaj, Convergence of scaled renewal processes and a packet arrival model. *Bernoulli* **9** (2003) 671–703.
- [13] B. González-Arévalo and J. Roy, Simulating a Poisson cluster process for Internet traffic packet arrivals. *Comput. Commun.* **33** (2010) 612–618.
- [14] F. Grüneis, M. Nakao and M. Yamamoto, Counting statistics of 1/f fluctuations in neuronal spike trains. *Biol. Cybernet.* **62** (1990) 407–413.
- [15] F. Grüneis, M. Nakao, M. Yamamoto, T. Musha and H. Nakahama, An interpretation of 1/f fluctuations in neuronal spike trains during dream sleep. *Biol. Cybernet.* **60** (1989) 161–169.
- [16] C.A. Guerin, H. Nyberg, O. Perrin, S. Resnick, H. Rootzén and C. Stărică, Empirical testing of the infinite source Poisson data traffic model. *Stoch. Models* **19** (2003) 56–199.
- [17] A. Gut, Stopped Random Walks, *Springer Series in Operations Research and Financial Engineering*, 2nd edn. Springer, New York. Limit theorems and applications (2009)
- [18] N. Hohn, D. Veitch and P. Abry, Cluster processes: a natural language for network traffic. *IEEE Trans. Signal Process.* **51** (2003) 2229–2244.
- [19] I. Kaj, Stochastic Modeling in Broadband Communications Systems, *SIAM Monographs on Mathematical Modeling and Comput.*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA. (2002)
- [20] A.F. Karr, Point Processes and their Statistical Inference. Vol. 7 of *Probability: Pure and Applied*, 2nd edn. Marcel Dekker, Inc., New York (1991).
- [21] P.M. Krishna, V. Gadre and U.B. Desai, Multifractal Based Network Traffic Modeling. Springer Science and Business Media (2012).
- [22] W.E. Leland, M.S. Taqqu, W. Willinger and D.V. Wilson, On the self-similar nature of Ethernet traffic (Extended version). *IEEE/ACM Trans. Netw.* **2** (1994) 1–15.
- [23] A.W.P. Lewis, A branching Poisson process model for the analysis of computer failure patterns (with discussion). *J. Roy. Stat. Soc. Series B. Methodological* **26** (1964) 398–456.
- [24] S.B. Lowen and M.C. Teich, Fractal-Based Point Processes. *Wiley Series in Probability and Statistics*. Wiley-Interscience [John Wiley and Sons], Hoboken, NJ (2005)
- [25] B. Mandelbrot, A case against the lognormal distribution. *Fractals and Scaling in Finance*, Springer New York (1997) 252–269.
- [26] T. Mikosch, S. Resnick, H. Rootzén and A. Stegeman, Is network traffic approximated by stable Lévy motion or fractional Brownian motion?. *Ann. Appl. Prob.* **12** (2002) 23–68.
- [27] T. Mikosch and G. Samorodnitsky, Scaling limits for cumulative input processes. *Math. Oper. Res.* **32** (2007) 890–918.
- [28] C. Onof, R.E. Chandler, A. Kakou, P. Northrop, H.S. Wheeler and V. Isham, Rainfall modelling using Poisson-cluster processes: a review of developments. *Stoch. Environ. Res. Risk Assess.* **14** (2000) 384–411.
- [29] G. Peccati and M.S. Taqqu, Wiener Chaos: Moments, Cumulants and Diagrams of Bocconi. Vol. 1 *Springer Series*, Springer (2011)

- [30] V.J. Ribeiro, Z.-L. Zhang, S. Moon and C. Diot, Small-time scaling behavior of Internet backbone traffic. *Comput. Netw.* **48** (2005) 315–334.
- [31] E.A.B. Saleh, and M.C. Teich, Multiplied-Poisson noise in pulse, particle, and photon detection. *Proc. IEEE* **70** (1982) 229–245.
- [32] D. Veitch, N. Hohn and P. Abry, Multifractality in TCP/IP traffic: the case against. *Comput. Netw.* **48** (2005) 293–313.
- [33] H. Wendt, P. Abry and S. Jaffard, Bootstrap for empirical multifractal analysis. *IEEE Signal Process. Magazine* **24** (2007) 38–48.
- [34] M. Westcott, Results in the asymptotic and equilibrium theory of Poisson cluster processes. *J. Appl. Probl.* **10** (1973) 807–823.
- [35] R. Willink, Relationships between central moments and cumulants, with formulae for the central moments of gamma distributions. *Commun. Statist. Theory Methods* **32** (2003) 701–704.
- [36] P. Zeephongsekul, G. Xia and S. Kumar, Software reliability growth models based on cluster point processes. *Int. J. Sys. Sci.* **25** (1994) 737–751.