Small and large scale behavior of moments of Poisson cluster processes

Nelson Antunes
CEMAT/University of Lisbon

Vladas Pipiras
University of North Carolina

Patrice Abry
CNRS and École Normale Supérieure de Lyon

Darryl Veitch
University of Technology Sydney

March 2, 2016

Abstract

The usual, factorial and central moments of Poisson cluster processes are studied at small and large scales, leading to various scaling relations. At large scales, it is assumed that the distribution of cluster sizes is heavy-tailed. The scaling relations are also examined from the numerical standpoint, especially at small scales in connection to the so-called multifractal analysis. Other issues addressed include the transition between small and large scales, the special cases of Poisson cluster processes of interest, the advantages of using factorial moments over the usual moments, among others. The obtained results are motivated and applied to a real data trace from Internet traffic, where Poisson cluster processes have been used as physical models of choice.

1 Introduction

A Poisson cluster process (PCP, for short) consists of points 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, $j \geq 0$ (with $0 < S_0 < S_1 < \ldots$). The clusters are i.i.d. copies with a random but 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 the arrival time of the cluster. Such PCPs are also known as the Bartlett-Lewis processes after Bartlett (1963) and Lewis (1964) (see, for example, Cox and Isham (1980), Daley and Vere-Jones (2003)).

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

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

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 $T \to \infty$. 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 (1980), Karr (1991), Daley and Vere-Jones (2003)) and used successfully in applications (e.g. computer failure patterns in Lewis (1964), software reliability in Zeephongsekul et al. (1994), neuronal spike trains in Grüneis et al. (1989, 1990), physics in Saleh and Teich (1982), Lowen and Teich (2005), rainfall in Onof et al. (2000)). The motivating application in this work is the computer traffic observed on a network link, where points are data packets and clusters are packet flows (essentially document files, website or other application contents). The use of PCPs in modeling the data packet traffic was popularized by Hohn et al. (2003). See also Faÿ et al. (2006), Mikosch and Samorodnitsky (2007), Fasen and Samorodnitsky (2009), González-Arévalo and Roy (2010), Antunes and Pipiras (2015). The variants of PCPs for modeling Internet traffic are the ON/OFF model (e.g. Leland et al. (1994)), the infinite source Poisson arrival process (e.g. Mikosch et al. (2002), Guerin et al. (2003)), or the renewal point process (e.g. Kaj (2002), Gaigalas and Kaj (2003)).

In this work, we focus on the moments and cumulants of PCPs. On one hand, moments and cumulants are most basic and fundamental to any random quantity and, in fact, have already been studied for PCPs to some extent (see references in Section 2 below). We are particularly interested here in their scaling behavior at large (course) and small (fine) scales, especially in connection to the use of PCP models for Internet traffic.

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

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

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

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

where $n^{[r]} = n(n-1)\dots(n-r+1)$ for a nonnegative integer n. Central moments are natural to consider in view of some of the large scale limiting results available for centered PCPs (see (7.3) and (7.4) below). Fractional moments are considered because, as will be shown, they may be more informative about PCPs than the usual moments.

The quantities most convenient to work with in the context of point processes 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 $r \geq 1$,

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

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

where $P_a(z)$ is the probability generating function of the equilibrium PCP on the interval $(0,a)$ (see Section 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)–(1.4) and (1.5).

1.1 Description of the main results and other contributions

Small and large scale behaviors of the quantities (1.2)–(1.4) and (1.5)–(1.6) refer, respectively, to $a \to 0^+$ and $a \to \infty$. As we will show, any of the moments (1.2)–(1.4) satisfy the scaling relation

$$m_{gen,r}(a) \sim C_{gen,r} a^{\zeta_{gen}(r)}, \quad \text{as } a \to 0^+ \text{ or } a \to \infty, \tag{1.7}$$

	Exponents functions	
moments	large scales ($a \to \infty$)	small scales ($a \to 0^+$)
$m_r(a)$	$\zeta(r) = r$	$\zeta(r) = 1$
$m_{[r]}(a)$	$\zeta_{[]}(r) = r$	$\zeta_{[]}(r) = 1 + (r-1)\theta$
$m_r^0(a)$	$\zeta_0(r) = r - \alpha + 1$ ($r \geq 2$)	$\zeta_0(r) = 1$
cumulants	large scales ($a \to \infty$)	small scales ($a \to 0^+$)
$\kappa_r(a)$	$\eta(r) = r - \alpha + 1$	$\eta(r) = 1$
$\kappa_{[r]}(a)$	$\eta_{[]}(r) = r - \alpha + 1$ ($r \geq 2$)	$\eta_{[]}(r) = 1 + (r-1)\theta$

Table 1: The exponents functions $\zeta_{gen}(r)$ in (1.7) and $\eta_{gen}(r)$ in (1.8) for the various moments and cumulants, together with the notation for each case.

where $m_{gen,r}(a)$ stands for one of the moments (1.2)–(1.4), $C_{gen,r} > 0$ is a constant and $\zeta_{gen}(r)$ is an exponents function (the abbreviation "gen" stands for "general", "generic"). The form of the constant $C_{gen,r}$ and the function $\zeta_{gen}(r)$ depends on which of the moments (1.2)–(1.4) is considered, and whether $a \to 0^+$ or $a \to \infty$. Likewise, we will also show that

$$\kappa_{gen,r}(a) \sim c_{gen,r} a^{\eta_{gen}(r)}, \quad \text{as } a \to 0^+ \text{ or } a \to \infty, \tag{1.8}$$

where $\kappa_{gen,r}(a)$ stands for one of the cumulants (1.5)–(1.6), $c_{gen,r} > 0$ is a constant and $\eta_{gen}(r)$ is an exponents function. The form of the constant $c_{gen,r}$ and the function $\eta_{gen}(r)$ depends on which of the cumulants (1.5)–(1.6) is considered, and whether $a \to 0^+$ or $a \to \infty$.

In showing (1.7) and (1.8), we shall make the following assumptions motivated by the applications to Internet traffic. At large scales ($a \to \infty$), we shall assume, in particular, that the cluster size distribution is heavy-tailed with exponent $\alpha \in (1, 2)$ in the sense that

$$\mathbb{P}(W_j > w) \sim C_W w^{-\alpha}, \quad \text{as } w \to \infty, \tag{1.9}$$

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. (2010)). At small scales ($t \to 0^+$), we shall assume that the cumulative distribution F_1 of interarrival times $A_{j,m}$ has a density f_1 satisfying: for $\theta > 0$ and $C_f > 0$,

$$f_1(t) \sim C_f t^{\theta - 1}, \quad \text{as } t \to 0^+. \tag{1.10}$$

In the applications to Internet traffic, F_1 is often taken as a gamma distribution, thus satisfying (1.10).

Under the assumptions (1.9) and (1.10), our scaling results (1.7) and (1.8) are summarized in Table 1, whose entries are the exponents functions $\zeta_{gen}(r)$ and $\eta_{gen}(r)$ for the moments and cumulants of interest, including the notation of the functions for each specific case. (The constants in (1.7) and (1.8) will also be derived but are not reported in Table 1.) Several interesting conclusions can be drawn from Table 1. For example, on the moments side, note that the exponent α is captured by the central moments only, while at small scales, the exponent θ is captured by the factorial moments only. The factorial cumulants, on the other hand, have these exponents at both large and small scales. One obvious interest in the obtained scaling relations is that they could be used to estimate the underlying model parameters, such as α and θ.

The theoretical results summarized in Table 1 aside, our paper considers several other aspects of the problem and is related to other work found in the literature. We characterize the transition between small and large scales, study several special cases of PCPs of interest, and also examine our results on a real data trace from Internet traffic. In the application, 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.

We also note that the results of Table 1 hold for a fixed arrival rate λ. In the analysis of PCPs and related models at large scales under the assumption (1.9) (and thus for Internet traffic), it is common to consider the rate λ as a function of T, that is, $\lambda = \lambda(T)$, where T is the length of the observation window $(0, T)$ of PCP. The connections of this work to the case $\lambda = \lambda(T)$ and some scaling results when taking $a = T$ will be discussed below (see Section 7).

Our results on the large scale asymptotic behavior of the various moments are closest in the spirit to those of Dombry and Kaj (2013) who considered moments 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 (2013). Further comparison with the work of Dombry and Kaj (2013) will be provided below (see Remark 7.2 below).

Our results at small scales are somewhat connected to the so-called multifractality, which refers to the scaling behavior (1.7) of the usual moments $m_r(a)$ as $a \to 0^+$ with nonlinear function ζ_r. According to Table 1, the exponents function $\zeta(r) = 1$ for the PCP is linear, corresponding to monofractal behavior. From a practical perspective, however, $\zeta(r)$ is estimated over a range of small scales and we shall argue through numerical experiments that the scaling (1.7) with $\zeta(r) = 1$ occurs for a range of too small a's to be observable in practice. For observable ranges of small a's, nonlinear $\zeta(r)$ could be estimated, suggesting a multifractal behavior which is spurious. These findings are consistent with the viewpoint expressed in Veitch et al. (2005). Multifractal-like features of PCPs were also observed and discussed in Hohn et al. (2003), and even modeled through multifractals in Ribeiro et al. (2005), Krishna et al. (2012).

Table 1 also suggests that the factorial moments, and not the usual moments, are more informative about the PCP at small scales, since their exponents function $\zeta_{[]}(r) = 1 + \theta(r - 1)$ involves θ. We also find that this exponents function can be estimated more reliably, compared to the exponents function $\zeta(r) = 1$ of the usual moments as noted above. This suggests that the factorial moments should be used instead of the usual moments in the analysis of PCPs at small scales, though a more thorough study of these observations is left for future work. We should also note that 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. (1989), de Wolf et al. (1996), 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 and for fixed λ is studied in Section 3. Section 4 concerns the behavior of the moments at small scales. The transition between small and large scales is discussed in Section 5. The special case of PCPs when the interarrival times follow a normal distribution (shifted so that the probability of it being negative is negligible) is treated in Section 6. The large scale behavior in the case of $\lambda = \lambda(T)$ mentioned above is discussed in Section 7. The application to Internet traffic can be found in Section 8. Finally, in Appendix A, we provide the formulae relating the first 7 central moments and factorial cumulants, which are used in Section 7, and in Appendix B, we derive the formula for the factorial cumulants of PCP, adapting the approach of Westcott (1973).

2 Moments and cumulants of Poisson cluster processes

We gather here several formulae for the moments and related quantities of some 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 identically distributed as A having distribution function

$$F_1(t) = \mathbb{P}(A \leq t), \quad t > 0, \tag{2.1}$$

and its kth convolution will be denoted $F_k = F_1 * \ldots * F_1$, $k \geq 1$. The number of points W_j in a cluster is identically distributed as W with probability mass function

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

and the 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

$$\kappa_{[r]}(t) = \frac{\partial^r \log P_t(z)}{\partial z^r}\Big|_{z=1}, \quad r \geq 1,\ t > 0. \tag{2.5}$$

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 $r \geq 2$,

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

The formulas (2.6)–(2.7) appear in Westcott (1973) when the first points of the clusters are excluded. When the first points are included, the formulas are derived in Appendix B below.

The factorial cumulants are related to *factorial moments*

$$m_{[r]}(t) = \mathbb{E}N_e(0, t)^{[r]}, \tag{2.8}$$

where $n^{[r]} = n(n-1)\ldots(n-r+1)$. The relationship is the same as that between the usual cumulants and moments, that is,

$$m_{[1]}(t) = \kappa_{[1]}(t), \quad m_{[2]}(t) = \kappa_{[2]}(t) + \kappa_{[1]}(t)^2, \quad m_{[3]}(t) = \kappa_{[3]}(t) + 3\kappa_{[2]}(t)\kappa_{[1]}(t) + \kappa_{[1]}(t)^3 \tag{2.9}$$

and, in general,

$$m_{[r]}(t) = \sum_{k=0}^{r-1} \binom{r-1}{k} \kappa_{[r-k]}(t) m_{[k]}(t) \tag{2.10}$$

and also

$$m_{[r]}(t) = \sum_{k=1}^{r} B_{r,k}\Big(\kappa_{[1]}(t), \kappa_{[2]}(t), \dots, \kappa_{[r-k+1]}(t)\Big), \tag{2.11}$$

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

$$\begin{aligned} B_{r,k}&(x_1, x_2, \dots, x_{r-k+1}) \\ &= \sum_{(n_1,n_2,\dots,n_{r-k+1})\in S_{r,k}} \frac{r!}{n_1! n_2! \dots n_{r-k+1}!} \Big(\frac{x_1}{1!}\Big)^{n_1} \Big(\frac{x_2}{2!}\Big)^{n_2} \dots \Big(\frac{x_{r-k+1}}{(r-k+1)!}\Big)^{n_{r-k+1}} \end{aligned} \tag{2.12}$$

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 $n_1 + 2n_2 + \dots + (r-k+1)n_{r-k+1} = r$.

Factorial moments, on the other hand, can be related back to the usual moments

$$m_r(t) = \mathbb{E}N_e(0,t)^r, \quad r \geq 1, \tag{2.13}$$

through the relation

$$m_r(t) = \sum_{j=1}^{r} \Delta_{j,r} m_{[j]}(t), \tag{2.14}$$

where $\Delta_{j,r}$ are the Stirling numbers of the second kind (e.g. Daley and Vere-Jones (2003), 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

$$m_r^0(t) = \mathbb{E}(N_e(0,t) - \mathbb{E}N_e(0,t))^r, \quad r \geq 1, \tag{2.15}$$

which are related to the usual moments through

$$m_r^0(t) = \sum_{j=0}^{r} \binom{r}{j} (-1)^{r-j} m_j(t) m_1(t)^{r-j}. \tag{2.16}$$

Finally, for convenience we use the following recursion formula relating cumulants and central moments (e.g. Willink (2003))

$$\kappa_r(t) = m_r^0(t) - \sum_{j=1}^{r-2} m_j^0(t) \kappa_{r-j}(t), \quad r \geq 2, \tag{2.17}$$

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

3 Moment behavior at large scales

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

Assumption W: The distribution of W_j is heavy-tailed, that is,

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

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 given below. 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 3.1 *Suppose that the distribution of the number of points in a cluster of PCP satisfies Assumption W above. The factorial cumulants $\kappa_{[r]}(a)$, $r \geq 1$, of PCP then satisfy:*

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

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: 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 W yield

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

where $a \vee b = \max\{a, b\}$ and

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

Similarly,

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

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

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

and

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

where

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

By Theorem 5.1, (ii), in Gut (2009), Chapter 2, $\mathbb{E}((N(u)/u)^{2-\alpha}) \to (1/\mathbb{E}A_{j,m})^{2-\alpha}$ as $u \to \infty$ and hence

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

The relation (3.2) with $r=2$ now follows from (3.4)–(3.7). The relation (3.2) with $r>2$ can be proved similarly by arguing that

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

and again using the same result of Gut (2009). □

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

Corollary 3.1 *Under the assumptions of Proposition 3.1, 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:*

$$m_{[r]}(a) \sim C_{m,[r]} \lambda^r a^r, \tag{3.8}$$
$$m_r(a) \sim C_{m,r} \lambda^r a^r, \tag{3.9}$$
$$m_r^0(a) \sim C_{m,r}^0 \lambda a^{r-\alpha+1}, \quad \text{as } a \to \infty, \tag{3.10}$$

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 (2.9)–(2.10). Indeed, when $r=2$, it follows from the second relation in (2.9) 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 (2.10) is of the order $a \cdot a^{r-1} = a^r$ (with the constant $C_{m,[r]}$) when $k = r-1$, and of the smaller order $a^{r-k-\alpha+1} \cdot a^k = a^{r-\alpha+1}$ when $k < r$. The relation (3.9) follows immediately from (2.14) and (3.8).

The relation (3.10) is slightly more difficult to deal with. We shall use the relation (2.16) to express $m_r^0(a)$ in terms of the moments $m_j(a)$, $j = 1,\dots,r$, and the relations (2.14) and (2.11) 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 (2.14) and (2.11) yield

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

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

$$a^{n_1} a^{n_2(2-\alpha+1)} \ldots a^{n_{p-q+1}(p-q+1-\alpha+1)} = \begin{cases} a^{p-(q-n_1)(\alpha-1)}, & q \leq p-1, \\ a^q, & q = p = n_1. \end{cases} \tag{3.13}$$

This order is largest when

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

which corresponds to

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

But, when substituted into (2.16), this term yields

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

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

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

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

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

When substituted into (2.16), this yields

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

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). $\square$

Corollary 3.2 *Under the assumptions of Proposition 3.1, the cumulants $\kappa_r(a)$, $r \geq 2$, of PCP satisfy:*

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

where

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

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

Proof: To show (3.16), we shall use first the relationship between cumulants and central moments (2.17) and then the asymptotic behavior of the central moments at large scales (3.10). The relation is trivial for the first cumlants since $\kappa_2(a) = m_2^0(a)$ and $\kappa_3(a) = m_3^0(a)$. By induction, if (3.16) holds for $2, \ldots, r-1$, then it also holds for r since in (2.17) 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 order $a^{r-\alpha+1}$. □

4 Moment 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 $a \to 0^+$. We focus on the following class of distributions of the interarrival times between points in clusters.

Assumption A: Suppose that the cumulative distribution F_1 of interarrival times has a density f_1 satisfying: for $\theta > 0$ and $C_f > 0$,

$$f_1(t) \sim C_f t^{\theta-1}, \quad \text{as } t \to 0^+. \tag{4.1}$$

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

$$f_1(t) = \frac{\beta(\beta t)^{\theta-1}e^{-\beta t}}{\Gamma(\theta)}, \quad t > 0, \tag{4.2}$$

where $\Gamma(\cdot)$ denotes the usual gamma function. The gamma distribution will be used in the application to Internet traffic in Section 8 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 4.1 *Suppose that the distribution of the interarrival times of PCP satisfies Assumption A above. The factorial cumulants $\kappa_{[r]}(a)$, $r \geq 1$, of PCP then satisfy:*

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

where

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

with $\overline{R}_r = \sum_{w=r}^\infty R_w = \sum_{w=r}^\infty \mathbb{P}(W \geq r)$ 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 $B(\cdot,\cdot)$.

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 $F_k(u) \sim C_f u^\theta/\theta =: C_{F,1}u^\theta$, as $u \to 0^+$. In fact, for any $k \geq 1$,

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

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 (4.8) hold for k, note that

$$F_{k+1}(u) = \int_0^u F_k(y)f_1(u-y)dy \sim C_{F,k}C_f \int_0^u y^{k\theta}(u-y)^{\theta-1}dy$$

$$= C_{F,k}C_f \int_0^1 z^{k\theta}(1-z)^{\theta-1}dz\, u^{(k+1)\theta} = C_{F,k}C_f B(k\theta+1, \theta-1)\, u^{(k+1)\theta} = C_{F,k+1}u^{(k+1)\theta}.$$

The relation (4.8) now implies that

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

In view of (4.8), the leading term for $\kappa_{[r]}(a)$ in (2.6)–(2.7) is of the desired order $a^{(r-1)\theta+1}$. To show that the sum of the remaining terms in negligible, one can use the argument above to conclude that, for any $\epsilon > 0$, $f_1(a) \leq Ca^{\theta-\epsilon-1}$, $a \in (0, a_0)$, and hence

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

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 $a^{(r-1)\theta+1}$) are thus bounded by a function of the order $a^{r(\theta-\epsilon)+1}$, which is negligible compared to $a^{(r-1)\theta+1}$ 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 $B(a, b) = \Gamma(a)\Gamma(b)/\Gamma(a+b)$. □

Corollary 4.1 *Suppose the distribution of the interarrival times of PCP satisfies Assumption A above with $\theta \in (0, 1]$. The moments $m_r(a)$, factorial moments $m_{[r]}(a)$ and central moments $m^0_r(a)$, $r \geq 2$, of PCP then satisfy:*

$$m_{[r]}(a) \sim c_{m,[r]}a^{1+(r-1)\theta}, \tag{4.9}$$

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

$$m^0_r(a) \sim c^0_{m,r}a, \quad \textit{as } a \to 0^+, \tag{4.11}$$

where

$$c_{m,[r]} = c_{\kappa,[r]}, \quad c_{m,r} = \lambda\mathbb{E}W, \quad c^0_{m,r} = \lambda\mathbb{E}W \tag{4.12}$$

with $c_{\kappa,[r]}$ appearing in (4.4). (When $r = 1$, $m_{[1]}(a) = m_1(a) = \kappa_{[1]}(a) = \lambda\mathbb{E}Wa$ and $m^0_1(a) = 0$.)

PROOF: To show (4.9), we argue by induction. The relation (4.9) with $r = 2$ holds in view of (2.9) and (4.3). Supposing it holds for $1, \dots, r-1$, it also holds for r by using (2.10) and (4.3). Indeed, the term $\kappa_{[r-k]}(a)m_{[k]}(a)$ in the sum (2.10) is of the order $a^{(r-1)\theta+1}$ when $k = 0$, and is of the smaller or equal order $a^{(r-k-1)\theta+1}a^{(k-1)\theta+1} = a^{(r-1)\theta+2-\theta}$ when $k \geq 1$.

The relation (4.10) follows from (4.9) and (2.14), since the term $m_{[1]}(a)$ dominates in the latter sum. Similarly, the relation (4.11) follows from (2.16) and (4.10) since the term $m_r(a)$ when $j = r$ is dominant in (2.16). □

Corollary 4.2 *Suppose the distribution of the interarrival times of PCP satisfies Assumption A above with $\theta \in (0, 1]$. The cumulants $\kappa_r(t)$, $r \geq 2$, of PCP then satisfy:*

$$\kappa_r(a) \sim c_{\kappa,r} a, \quad \text{as } a \to 0^+, \tag{4.13}$$

where

$$c_{\kappa,r} = \lambda \mathbb{E}W. \tag{4.14}$$

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

PROOF: The relation (4.13) can be shown by using (2.17) and (4.11). From (2.17) 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 (4.13) holds for $2, 3, \dots, r-1$, then it also holds for r since in (2.17) 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 $m_j^0(a)$ has order a. □

Example 4.1 Proposition 4.1 describes the asymptotic behavior of the factorial cumulants of PCP under Assumption A. A more explicit, non-asymptotic expression of the factorial cumulants can be obtained in the special case of the gamma distribution (4.2), 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

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

and hence

$$\begin{aligned}
\int_0^a F_k(u)du &= \frac{1}{\Gamma(k\theta)\beta}\Big(\beta a \gamma(k\theta, \beta a) - \gamma(k\theta + 1, \beta a)\Big) \\
&= \frac{1}{\Gamma(k\theta)\beta}\Big(\beta a (\beta a)^{k\theta}\Gamma(k\theta)e^{-\beta a}\sum_{j=0}^{\infty}\frac{(\beta a)^j}{\Gamma(k\theta + j + 1)} - (\beta a)^{k\theta+1}\Gamma(k\theta+1)e^{-\beta a}\sum_{j=0}^{\infty}\frac{(\beta a)^j}{\Gamma(k\theta + 1 + j + 1)}\Big) \\
&= \frac{e^{-\beta a}}{\beta}\sum_{j=0}^{\infty}(\beta a)^{k\theta+j+1}\Big(\frac{1}{\Gamma(k\theta + j + 1)} - \frac{k\theta}{\Gamma(k\theta + 1 + j + 1)}\Big) \\
&= \frac{e^{-\beta a}}{\beta}\sum_{j=0}^{\infty}(\beta a)^{k\theta+j+1}\frac{j+1}{\Gamma(k\theta + 1 + j + 1)} = \frac{e^{-\beta a}}{\beta}\sum_{j=1}^{\infty}(\beta a)^{k\theta+j}\frac{j}{\Gamma(k\theta + 1 + j)}
\end{aligned} \tag{4.15}$$

An expression for the factorial cumulants can now be obtained by substituting (4.15) into (2.6)–(2.7). For example, for $r \geq 3$ and as $a \to 0^+$, the leading term in thus obtained relation leads to

$$\kappa_{[r]}(a) \sim \frac{r!\lambda(\beta a)^{(r-1)\theta+1}\overline{R}_r}{\beta\Gamma((r-1)\theta+2)} = \frac{r!\lambda\beta^{(r-1)\theta}\overline{R}_r}{\Gamma((r-1)\theta+2)} a^{1+(r-1)\theta}, \tag{4.16}$$

which is consistent with (4.3)–(4.5).

The asymptotic results (4.9) and (4.10) show that using regular moments of PCP at small scales does not reveal the underlying interarrival distribution, since the dominating behavior is governed by t for all the moments. In contrast, the behavior of the factorial moments is much more informative. This is further discussed in Section 8 in an application context.

Remark 4.1 An asymptotic behavior of the density f_1 not captured by Assumption A is when $f_1(t)$ decays faster than any power as $t \to 0^+$. This could be expressed, for example, by the assumption that

$$f_1(t) \sim Ct^{\alpha}e^{-|\log t|^{\beta}}, \quad \text{as } t \to 0^+, \tag{4.17}$$

where $\alpha \in \mathbb{R}$, $\beta > 1$ and $C > 0$. Results analogous to (4.3) and (4.9)–(4.10) could be obtained, we believe, under the assumption (4.17). However, we shall not pursue this direction here for the following reason. A prototypical example of the density satisfying (4.17) is that of a lognormal distribution. When working with the Internet traffic data for Section 8, we found the lognormal distribution difficult to capture the left tail behavior. Indeed, it is known that the moments of log normal distributions are "localized", making the use of the distribution quite delicate in practice (Mandelbrot (1997)).

5 Transition from small to large scales

The analysis carried out in Sections 3 and 4 (for fixed Poisson arrival rate λ) shows the existence of biscaling, that is, the different scaling behaviors at large and small scales. Moreover, since the scaling behaviors are different, one could expect them to be separate by a "knee", the point (or the range) where the behavior changes as one moves from small to large scales. The biscaling and the "knee" are clearly seen in Figure 1 (to be discussed in more detail in Section 8) where the factorial moments, moments and central moments are plotted.

The location of the "knee" (i.e., the time t) can be approximated by equating the scaling relations at large and small scales for the several measures and solving with respect to a. For the factorial cumulants, by using the relations (3.2) and (4.3), the approximate location of the "knee" is given by

$$t_{\kappa,[r]} = \left(\frac{(r-2)!C_{F,r-1}(\mathbb{E}A)^{r-\alpha}(\alpha-1)(r-\alpha)(r+1-\alpha)\overline{R}_r}{((r-1)\theta+1)C_W}\right)^{1/(r-(r-1)\theta-\alpha)}, \qquad r \geq 2. \tag{5.1}$$

From Corollaries 3.1 and 4.1, the location of the "knee" for the factorial moments, the moments and central moments are approximated by, respectively,

$$t_{m,[r]} = \left(\frac{r!C_{F,r-1}\overline{R}_r}{\lambda^{r-1}((r-1)\theta+1)(\mathbb{E}W)^r}\right)^{1/(r-(r-1)\theta-1)}, \qquad \theta \neq 1, \tag{5.2}$$

$$t_{m,r} = \frac{1}{\lambda\mathbb{E}W_j}, \tag{5.3}$$

$$t^0_{m,r} = \left(\frac{(\alpha-1)(r-\alpha)(r+1-\alpha)\mathbb{E}W(\mathbb{E}A)^{r-\alpha}}{r(r-1)C_W}\right)^{1/(r-\alpha)}, \qquad r \geq 2. \tag{5.4}$$

If A is exponential, then $\theta = 1$ and equating (3.8) and (4.9) we obtain $t_{m,[r]} = 0$ for $r \geq 2$. This means that in this case there is no helpful intersection for the different scaling behaviors. We note that the location of the "knee" for the moments does not dependent on the order r and the distribution of the interarrival times A.

Finally, for the cumulants the location of the "knee" $t_{\kappa,[r]}$, $r \geq 2$, is the same as for the central moments since they have the same asymptotic behaviour at small and large scales.

6 The special case of normal interarrival times

The formulae (2.6)–(2.7) for the factorial cumulants involve the integrals $\int_0^t F_k(u)du$, where $F_k(u)$ is the distribution function of the sum of k interarrival times. By the central limit theorem, for large k, one can expect the sum to be approximated by a normal distribution and similarly $F_k(u)$ to be approximated well by the distribution function of the corresponding normal distribution. This approximation could be useful when evaluating the factorial cumulants through (2.6)–(2.7) numerically.

In a slightly different approach we could start by supposing that the intearrival times themselves are normal (shifted to the right so that they can be assumed positive for practical purposes). The integrals $\int_0^t F_k(u)du$ can then be evaluated in this special case as done in Proposition 6.2 below, and the relevance of the normal approximations be assessed numerically for larger moments (see Section 8).

We thus suppose that μ and σ are the mean and standard deviation of interarrival times $T_{j,m}$, respectively. Consider the rescaled normal random variable

$$X = \mu + \sigma Z, \tag{6.1}$$

where Z has a standard normal distribution. Letting $X_1, \ldots, X_k$, be independent random variables identically distributed as the random variable X, we have

$$X_1 + \ldots + X_k \stackrel{d}{=} k\mu + \sqrt{k}\sigma Z \tag{6.2}$$

and denote the respective distribution function by F_k with a slight abuse of notation. The next result provides an expansion of the integral $\int_0^t F_k(u)du$ in power series of t being computationally more amenable to handle. It is used below to derive the behavior of the moment measures for small scales.

Proposition 6.1 *For $k \geq 1$, let F_k be the distribution function of the sum (6.2). Then,*

$$\int_0^t F_k(u)du = \frac{1}{2\sqrt{2\pi k}\sigma} e^{-(t^2+k^2\mu^2)/(2k\sigma^2)} \sum_{i=0}^{\infty} \sum_{j=0}^{\lfloor i/2 \rfloor} t^{2+i} \frac{(\mu/\sigma^2)^{i-2j}}{(i-2j)!(2j\sigma^2)^j} \times \left(\frac{\Gamma(i/2-j+1/2)}{\Gamma(i/2+3/2)} - \frac{\Gamma(i/2-j+1)}{\Gamma(i/2+2)} \right), \tag{6.3}$$

where $\lfloor . \rfloor$ is the floor function.

Proof: Let $f_k(v) = 1/(\sqrt{2\pi k}\sigma)e^{-(v-k\mu)^2/(2k\sigma^2)}$ be the p.d.f. of $X_1 + \ldots + X_k$ in (6.2). Then,

$$\int_0^t F_k(u)du = \int_0^t \int_0^u f_k(v)dvdu = t\int_0^t f_k(v)dv - \int_0^t kf_k(v)dv. \tag{6.4}$$

Considering the two last terms in (6.4) separately, we have

$$\begin{aligned}\int_0^t f_k(v)dv &= \frac{1}{\sqrt{2\pi k}\sigma} e^{-k\mu^2/(2\sigma^2)} \int_0^t e^{v\mu/\sigma^2} e^{-v^2/(2k\sigma^2)} dv \\ &= \frac{1}{\sqrt{2\pi k}\sigma} e^{-k\mu^2/(2\sigma^2)} \sum_{j=0}^{\infty} \left(\frac{\mu}{\sigma^2}\right)^j \int_0^t v^j e^{-v^2/(2k\sigma^2)} dv. \end{aligned} \quad (6.5)$$

For the integrals in (6.5), by making the change of variables $x = v^2/(2k\sigma^2)$, we obtain that

$$\int_0^t v^j e^{-v^2/(2k\sigma^2)} dv = 2^{(j-1)/2} (k\sigma^2)^{(j+1)/2} \gamma\left(\frac{j+1}{2}, \frac{t^2}{2k\sigma^2}\right), \quad (6.6)$$

where $\gamma(.,.)$ is the lower incomplete gamma function defined and considered in Example 4.1. Now, by using the power series expansion of the lower incomplete gamma function in (6.6) and replacing in (6.5), we find after some algebra that

$$t \int_0^t f_k(u)du = \frac{1}{2\sqrt{2\pi k}\sigma} e^{-(t^2+k^2\mu^2)/(2k\sigma^2)} \sum_{i=0}^{\infty} t^{2+i} \sum_{j=0}^{\lfloor i/2 \rfloor} \frac{(\mu/\sigma^2)^{i-2j} \Gamma(i/2 - j + 1/2)}{(i-2j)!(2k\sigma^2)^j \Gamma(i/2 + 3/2)}.$$

Making the expansion of the last term in (6.4) through the same steps as above, the result (6.3) follows. □

As a result of independent interest, we state the asymptotic relations for the moments at small scales when the distribution of the interarrival times is normal. We omit the proof which follows similar lines as those of Proposition 4.1 and Corollary 4.1. Indeed, the result can be checked easily for factorial cumulants by substituting (6.3) into (2.6)–(2.7), letting a become small and finally using the relations between moments. We note that at large scales, the results of Section 3 are still valid in this case.

Proposition 6.2 *Suppose that $T_{j,m}$ is normally distributed with mean μ and standard deviation σ so that the integrals $\int_0^a F_k(u)du$ are given by (6.3). The factorial cumulants $\kappa_{[r]}(a)$, $r \geq 2$, of PCP then satisfy:*

$$\kappa_{[r]}(a) \sim \frac{\lambda r! e^{-(r-1)^2\mu^2/(2(r-1)\sigma^2)} \overline{R}_r}{2\sqrt{2\pi(r-1)}\sigma} a^2, \quad \text{as } t \to 0. \quad (6.7)$$

The moments $m_r(a)$, the factorial moments $m_{[r]}(a)$ and the central moments $m_r^0(a)$ satisfy:

$$m_{[2]}(a) \sim \lambda \left(\frac{e^{-\mu^2/(2\sigma^2)} \overline{R}_2}{\sqrt{2\pi}\sigma} + \lambda(\mathbb{E}W_j)^2 \right) a^2, \quad m_{[r]}(a) \sim \frac{\lambda r! e^{-(r-1)^2\mu^2/(2(r-1)\sigma^2)} \overline{R}_r}{2\sqrt{2\pi(r-1)}\sigma} a^2, \quad (6.8)$$

$$m_r(a) \sim \lambda \mathbb{E}W_j a, \quad (6.9)$$

$$m_r^0(a) \sim \lambda \mathbb{E}W_j a, \quad r \geq 2, \quad \text{as } a \to 0. \quad (6.10)$$

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

The proposition above shows that at small time scale the asymptotic relations for the moments and central moments are the same as in (4.10) and (4.11), respectively.

7 Moment behavior at large scales in the slow and fast growth regimes

As noted in Section 1, in the analysis of PCPs and related models at large scales under the assumption (1.9), it is common to consider the rate λ as a function of T, that is, $\lambda = \lambda(T)$, where T is the length of the observation window $(0,T)$ of PCP. Note that this does not mean that λ varies in the observation window but rather that, once the observation window is fixed, λ is viewed as a function of T. Two different regimes are distinguished: letting $\lambda = \lambda(T)$, the *slow growth regime* is defined as

$$\frac{\lambda(T)}{T^{\alpha-1}} \to 0 \tag{7.1}$$

and the *fast growth regime* as

$$\frac{\lambda(T)}{T^{\alpha-1}} \to \infty. \tag{7.2}$$

One thinks of $\lambda = \lambda(T)$ as the connection rate in the context of Internet traffic (Mikosch et al. (2002)).

In the slow growth regime, one then has

$$\Big\{\frac{N_e(0,Tu) - \mathbb{E}N_e(0,Tu)}{T^{1/\alpha}}\Big\}_{u\in[0,1]} \stackrel{fdd}{\to} \{L_\alpha(u)\}_{u\in[0,1]}, \quad \text{as } T\to\infty, \tag{7.3}$$

where the convergence is in the sense of finite-dimensional distributions and L_α is an α–stable Lévy motion (Mikosch and Samorodnitsky (2007), Proposition 5.11). In the fast growth regime, on the other hand,

$$\Big\{\frac{N_e(0,Tu) - \mathbb{E}N_e(0,Tu)}{\lambda(T)^{1/2}T^{(3-\alpha)/2}}\Big\}_{u\in[0,1]} \stackrel{fdd}{\to} \{B_H(u)\}_{u\in[0,1]}, \quad \text{as } T\to\infty, \tag{7.4}$$

where B_H is fractional Brownian motion (FBM) with the self-similarity parameter $H = (3-\alpha)/2$ (Mikosch and Samorodnitsky (2007), Proposition 4.7). So, for example, if the connection rate λ is large compared to $T^{\alpha-1}$, one expects the PCP to be approximated by FBM.

The digression to the slow and fast regimes might be confusing to a number of readers, and this is for a good reason. There seem to be two conflicting implications of (7.3)–(7.4) in relation to the results of Section 3. On one hand, suppose that over the observation window $(0,T)$, λ is much bigger than $T^{\alpha-1}$, suggesting the fast regime. Then, within this observation window and with the normalization $\sigma_T = \lambda(T)^{1/2}T^{(3-\alpha)/2}$ in (7.4),

$$\begin{aligned} N_e(0,a) - \mathbb{E}N_e(0,a) &= N_e(0,T\frac{a}{T}) - \mathbb{E}N_e(0,T\frac{a}{T}) \\ &\stackrel{d}{\approx} \sigma_T B_H(\frac{a}{T}) \stackrel{d}{=} \frac{\sigma_T}{T^H}a^H B_H(1), \end{aligned}$$

by using (7.4) and the H–self-similarity of FBM, where $\stackrel{d}{\approx}$ denotes the approximation in distribution. Raising the left- and right-hand sides of the expression above to the power r and taking expectations, this suggests that the central moments scale up to a constant as a^{rH}, and hence the exponents function is

$$\widetilde{\zeta}_0(r) = rH = r(3-\alpha)/2.$$

On the other hand, over the observation window, the arrival rate is usually quite constant and hence from Corollary 3.1, the central moments scale with the exponents

$$\zeta_0(r) = r - \alpha + 1.$$

Note that the two exponents functions $\widetilde{\zeta}_0(r)$ and $\zeta_0(r)$ are the same only for $r = 2$. The numerical study carried out in Section 8 on a real Internet trace associated with the fast regime is, in fact, in line with the exponents $\zeta_0(r)$ and not $\widetilde{\zeta}_0(r)$. The real issues behind these observations and findings are not understood well yet, and are left for future work.

In a direction of independent interest, we also note that the results obtained in Section 3 can be used to study the moments and cumulants of PCPs under the slow and fast growth conditions taking $a = T$, that is, the moments of the expressions in the numerators of (7.3)–(7.4). We first note that the behavior of the factorial cumulants in (3.2), the factorial moments in (3.8) and the usual moments in (3.9) holds for any $\lambda \to \infty$, and in particular is the same in the slow and fast growth regimes. It is the case of central moments which is more delicate, as perhaps suggested by the different asymptotic behaviors in (7.3) and (7.4), which involve centering.

In fact, we do not have a general result for the asymptotics of centered moments of PCP under the slow and fast growth conditions when $a = T$. The key difficulty is seemingly the lack of a direct general formula relating central moments to factorial cumulants (see also Remark 7.1 below). Such formula for the first seven central moments is given in Appendix A, and can be used to derive the asymptotics of central moments under the slow and fast growth. By using the formula (A.1) and substituting the factorial cumulants $\kappa_{[1]}(t)$ and $\kappa_{[2]}(t)$ from (3.2), we can write

$$m_2^0(T) \sim C_{\kappa,[1]}\lambda T + C_{\kappa,[2]}\lambda T^{3-\alpha}, \tag{7.5}$$

as $T \to \infty$. In view of (7.1) and (7.2),

$$m_2^0(T) \sim C_{\kappa,[2]}\lambda T^{3-\alpha}, \tag{7.6}$$

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

$$m_3^0(T) \sim C_{\kappa,[1]}\lambda T + 3C_{\kappa,[2]}\lambda T^{3-\alpha} + C_{\kappa,[3]}\lambda T^{4-\alpha}, \tag{7.7}$$

as $T \to \infty$ and

$$m_3^0(T) \sim C_{\kappa,[3]}\lambda T^{4-\alpha}, \tag{7.8}$$

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

$$\begin{aligned} m_4^0(T) \sim C_{\kappa,[1]}\lambda T + 3C_{\kappa,[1]}^2\lambda^2T^2 + 7C_{\kappa,[2]}\lambda T^{3-\alpha} + 6C_{\kappa,[3]}\lambda T^{4-\alpha} \\ + C_{\kappa,[4]}\lambda T^{5-\alpha} + 6C_{\kappa,[1]}C_{\kappa,[2]}\lambda^2T^{4-\alpha} + 3C_{\kappa,[2]}^2\lambda^2T^{6-2\alpha}, \end{aligned} \tag{7.9}$$

as $T \to \infty$, with the two growth regimes now being distinct,

$$m_4^0(T) \sim \begin{cases} C_{\kappa,[4]}\lambda T^{5-\alpha}, & \text{slow growth,} \\ 3C_{\kappa,[2]}^2\lambda^2T^{6-2\alpha}, & \text{fast growth.} \end{cases} \tag{7.10}$$

Similarly, the relation (A.4) gives that

$$\begin{aligned} m_5^0(T) \sim C_{\kappa,[1]}\lambda T + 15C_{\kappa,[2]}\lambda T^{3-\alpha} + 25C_{\kappa,[3]}\lambda T^{4-\alpha} + 10C_{\kappa,[4]}\lambda T^{5-\alpha} + C_{\kappa,[5]}\lambda T^{6-\alpha} + 10C_{\kappa,[1]}^2\lambda^2T^2 \\ + 40C_{\kappa,[1]}C_{\kappa,[2]}\lambda^2T^{4-\alpha} + 10C_{\kappa,[1]}C_{\kappa,[3]}\lambda^2T^{5-\alpha} + 30C_{\kappa,[2]}^2\lambda^2T^{6-2\alpha} + 10C_{\kappa,[2]}C_{\kappa,[3]}\lambda^2T^{7-2\alpha}, \end{aligned} \tag{7.11}$$

as $T \to \infty$ and therefore,

$$m_5^0(T) \sim \begin{cases} C_{\kappa,[5]}\lambda T^{6-\alpha}, & \text{slow growth,} \\ 10C_{\kappa,[2]}C_{\kappa,[3]}\lambda^2T^{7-2\alpha}, & \text{fast growth.} \end{cases} \tag{7.12}$$

We also get from (A.5) that

$$\begin{aligned}m_6^0(T) \sim\ & C_{\kappa,[1]}\lambda T + 31C_{\kappa,[2]}\lambda T^{3-\alpha} + 90C_{\kappa,[3]}\lambda T^{4-\alpha} + 65C_{\kappa,[4]}\lambda T^{5-\alpha} + 15C_{\kappa,[5]}\lambda T^{6-\alpha} + C_{\kappa,[6]}\lambda T^{7-\alpha}\\ & + 25C_{\kappa,[1]}^2\lambda^2T^2 + 180C_{\kappa,[1]}C_{\kappa,[2]}\lambda^2T^{4-\alpha} + 110C_{\kappa,[1]}C_{\kappa,[3]}\lambda^2T^{5-\alpha} + 15C_{\kappa,[1]}C_{\kappa,[4]}\lambda^2T^{6-\alpha}\\ & + 195C_{\kappa,[2]}^2\lambda^2T^{6-2\alpha} + 150C_{\kappa,[2]}C_{\kappa,[3]}\lambda^2T^{7-2\alpha} + 25C_{\kappa,[1]}^2\lambda^2T^{8-2\alpha} + 10C_{\kappa,[2]}C_{\kappa,[4]}\lambda^2T^{8-2\alpha}\\ & + 15C_{\kappa,[1]}^3\lambda^3T^3 + 45C_{\kappa,[1]}^2C_{\kappa,[2]}\lambda^3T^{5-\alpha} + 45C_{\kappa,[1]}C_{\kappa,[2]}^2\lambda^3T^{7-2\alpha} + 15C_{\kappa,[2]}^2\lambda^3T^{9-3\alpha}, \quad (7.13)\end{aligned}$$

as $T \to \infty$, which yields

$$m_6^0(T) \sim \begin{cases} C_{\kappa,[6]}\lambda T^{7-\alpha}, & \text{slow growth,}\\ 15C_{\kappa,[2]}^3\lambda^3T^{9-3\alpha}, & \text{fast growth.}\end{cases} \quad (7.14)$$

Similarly, from (A.6),

$$m_7^0(T) \sim \begin{cases} C_{\kappa,[7]}\lambda T^{8-\alpha}, & \text{slow growth,}\\ 105C_{\kappa,[2]}^2C_{\kappa,[3]}\lambda^3T^{10-3\alpha}, & \text{fast growth.}\end{cases} \quad (7.15)$$

The relations above lead us to conjecture that, for $r \geq 2$,

$$m_r^0(T) \sim \begin{cases} C_{\kappa,[r]}\lambda(T)T^{r-\alpha+1}, & \text{slow growth,}\\ \frac{r!C_{\kappa,[2]}^{r/2}}{(r/2)!2!^{r/2}}\lambda(T)^{r/2}T^{(3-\alpha)r/2}, & \text{fast growth and even } r,\\ \frac{r!C_{\kappa,[2]}^{(r-1)/2-1}C_{\kappa,[3]}}{((r-1)/2-1)!2!^{(r-1)/2-1}3!}\lambda(T)^{(r-1)/2}T^{(3-\alpha)(r-1)/2+1}, & \text{fast growth and odd } r,\end{cases} \quad (7.16)$$

In fact, we checked the conjecture (7.16) 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.)

Several interesting observations can be made concerning (7.16). First, note that the conjectured behavior in the slow growth regime is exactly the same as in (3.10) for fixed λ. Second, note that the behavior (7.16) in the fast regime depends on whether r is even or odd. Moreover, the behavior is seemingly consistent with the normalization used in (7.4) only when r is even. That is, the odd (higher than 3) moments of the left-hand side of (7.4) may not converge to those of the limiting process.

Remark 7.1 The difficulty in proving (7.16) in general was indicated above but it is instructive to provide some further insight. First, we note that the proof of Corollary 3.1 cannot be used directly to show (7.16). Indeed, the terms associated with (3.15) in the proof of the corollary lead to $\kappa_{[r]}(T) \sim c\lambda T^{r-(\alpha+1)}$ but this term is no longer necessarily dominant in the fast regime. For example, note the presence of $\lambda T^{r-(\alpha-1)} = \lambda T^{7-\alpha}$ in (7.13) when $r = 6$. But this term is indeed dominated by $\lambda^3T^{9-3\alpha}$ in the fast regime. Second, the actual difficulty is in tracking the dominant term. For example, when $r = 6$, the dominant term arises from $\kappa_{[2]}(T)^3 \sim c\lambda^3T^{6-3(\alpha-1)}$ which enters into the moment $m_6(T)$ through (3.12). But, for example, $m_6(T)$ also contains the term $\kappa_{[1]}(T)^3\kappa_{[3]}(T) \sim c\lambda^4T^{6-4(\alpha-1)}$. Though this term dominates $\kappa_{[2]}(T)^3$, it does not appear in (7.13) since it gets canceled once substituted into (2.16).

Remark 7.2 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 (2013). In contrast to the approach taken here, Dombry and Kaj (2013) work with somewhat more general moment measures. The different growth regimes in superimposing renewal point processes are considered by Dombry and Kaj (2013) but *not* for the behavior of the moment measures. We have not aimed specifically to be different from Dombry and Kaj (2013) but just became aware of their work towards the end of this project.

8 Numerical study and application to Internet traffic

In this section, we illustrate the results of Sections 2–6 through a numerical study by using an Internet traffic data set (which we find more interesting and illuminating than using synthetic data). We consider the publicly available Internet trace, Auckland[1], which is one hour long and consists of 38,308,012 packets which make 1,371,756 flows. In Section 8.1, we examine the usual, central and factorial moments. Section 8.2 concerns the scaling exponents functions, especially in connection to multifractal analysis.

8.1 Moments, factorial moments and central moments

We illustrate here the scaling relations of the various moments of PCP in the parameters setting suggested by the Auckland data trace. 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 $\widehat{\lambda} = 396$ (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 $\widehat{\alpha} = 1.02$. The distribution of interarrival times between packets of a flow (between points of a cluster) is often modeled by a gamma distribution (see (4.2)). However, determining the appropriate parameters is not trivial as pointed out by Hohn et al. (2003) 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 $1/\mathbb{E}T_{j,m} = \beta/\theta$. An estimate for this in-flow packet arrival rate using the median or the mean of the empirical rate of each flow 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 $\widehat{\theta} = 0.6$ and the in-flow rate then results in $\widehat{\beta} = 1/0.002$.

Plots (a)–(c) in Figure 1 show the usual moments $m_r(t)$ against time t, 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 (2.11) and (2.14) 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 t over all trace duration, with the smallest value for t being 10^{-6} sec (≈ -13.8 in the log scale; the packet arrivals were extracted with increments of 1 microsecond). We also include the straight dashed lines corresponding to the scaling relations of the moments at *small scales* and *large scales* using (3.9) and (4.10), respectively. Note 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 (5.3). 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 potencial 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

[1]Auckland IX, file 20080327-080000-0, Available: http://wand.net.nz/wits/auck/9/

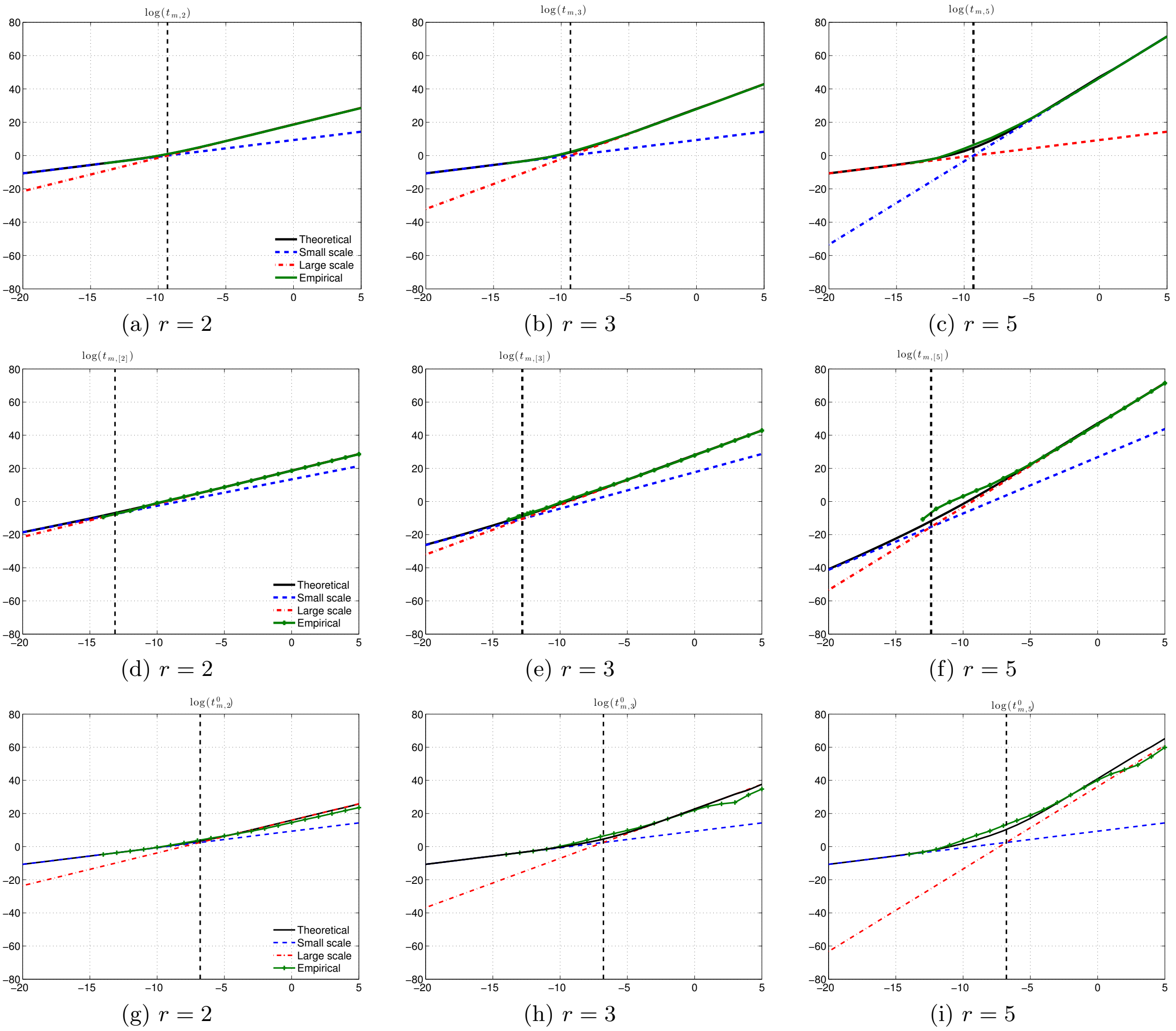
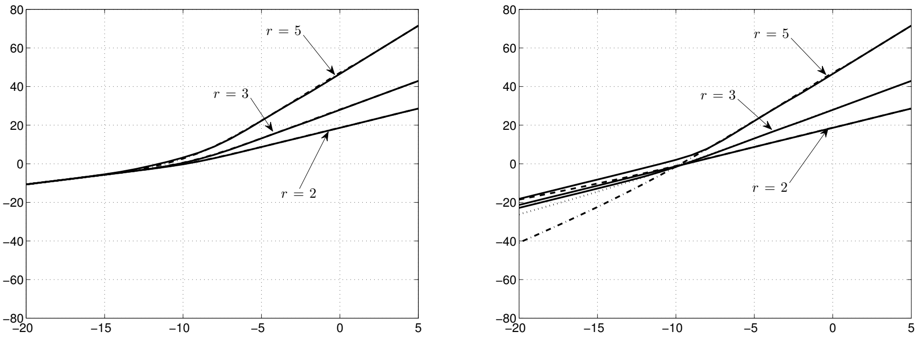


Figure 1: Top: moments (log-log scale); Middle: factorial moments (log-log scale); Bottom: central moments (log-log scale).

shorter interarrival times. In fact, the more general PCP considered by Westcott (1973), allows the interarrival times $T_{j,k}$ to be non-identically distributed and dependent on the cluster (flow) size. We have started exploring this direction but also feel that it may go beyond the scope of this article. A related concern is that 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 $m_{[r]}(t)$, for $r = 2, 3, 5$. We compute the empirical factorial moments using the number of packet arrivals over contiguous non-overlapping intervals of size t 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 t as for the usual moments. It is especially difficult to estimate the factorial moments for small t because of the order of the theoretical values. In this case, a trace with a longer duration is needed for more intervals of size



(a) Moments (b) Factorial moments

Figure 2: Rescaled normal distribution. The dashed lines correspond to assuming normal interarrival times.

t to be used in the estimation.

The central moments $m_r^0(t)$ 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 t 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 t means that the number of contiguous non-overlapping intervals of length t used in the calculation is insufficient.

In Remark **??**, we raised some questions about the relevance of slow and fast growth regimes in practice. Indeed, in the trace considered here, the arrival rate of flows is quite constant over all trace duration ($\lambda \approx 396$ flows/sec). In spite of the ratio $\lambda(t)/t^{\alpha-1}$ in (7.1) and (7.2) being large for the trace length $t = 3600$ sec, the empirical central moments scale according to the slow growth rate (7.16) or fixed arrival rate (3.10) as represented in Figure 1. We also note that these asymptotic relations (7.16) and (3.10) compared with the exact values (theoretical line) for higher order central moments (see $r = 5$) start to hold only for larger values of t.

Finally, we also comment here on the relevance of the rescaled normal distribution considered in Section 6. Figure 2 shows the moments and factorial moments when the integral terms in (2.7) are computed through (6.1)–(6.3). We also plot the respective values shown in Figure 1 where the interarrival times are gamma distributed. Though the results for the moments may appear surprising, they are in fact consistent with and can be explained through the asymptotic relations. At small time scales, the asymptotic relations for the moments do not depend on the distribution of the interarrival times (cf. (4.10) and (6.9)). This is not the case for the factorial moments which depend on the underlying distribution (see (4.9) and (6.8)). On the other hand, at large scales, there is no significant diference between the gamma and the normal distributions for both the usual and factorial moments (which agrees with the relations (3.8) and (3.9)). This observation suggests that the normal approximation can indeed be used in computing the integrals $\int_0^t F_k(u)du$.

8.2 Scaling exponents functions

The results (4.9) and (4.10) established in Corollary 4.1 can be written as

$$m_r(t) \sim c_{m,r} t^{\zeta(r)}, \quad m_{[r]}(t) \sim c_{m,[r]} t^{\zeta_{[]}(r)}, \quad \text{as } t \to 0^+, \tag{8.1}$$

where $\zeta(r) \equiv 1$ and $\zeta_{[]}(r) = 1+(r-1)\theta$. The relations of the form (8.1) are of special interest in the so-called multifractal analysis where $\zeta(r)$ and $\zeta_{[]}(r)$ are known as *scaling exponents (functions).* In the multifractal analysis, the usual moments are typically considered for all $r > 0$, and other multiresolution quantities instead of $N_e(0,t)$ are often used, such as wavelet coefficients, wavelet leaders, and others (see e.g. Jaffard et al. (2007)). When a scaling exponents function is nonlinear, the underlying random process is referred to as being multifractal (and monofractal if the function is linear). From a practical perspective, the relation (8.1) is approximated over a range of small t's as

$$\widehat{m}_r(t) \approx \widehat{c}_{m,r} t^{\widehat{\zeta}(r)}, \quad \widehat{m}_{[r]}(t) \approx \widehat{c}_{m,[r]} t^{\widehat{\zeta}_{[]}(r)}, \quad \text{for } t_1 < t < t_2, \tag{8.2}$$

where $\widehat{m}_r(t)$ are the sample moments and $\widehat{m}_{[r]}(t)$ are the sample factorial moments. The estimated scaling exponents functions $\widehat{\zeta}(r)$ and $\widehat{\zeta}_{[]}(r)$ are then obtained as the slopes in the linear regressions of $\log \widehat{m}_r(t)$ and $\log \widehat{m}_{[r]}(t)$ on $\log t$ over the range $t_1 < t < t_2$. Note that analogous scaling exponents functions $\zeta(r)$ and $\zeta_{[]}(r)$ ($\widehat{\zeta}(r)$ and $\widehat{\zeta}_{[]}(r)$) can be defined at large scales as well, based on the relations (3.8) and (3.9).

Plots (a)–(b) in Figure 3 depict the estimated scaling exponents functions of the moments $\widehat{\zeta}(r)$ over different ranges at large and small scales. In order to provide accurate results for the higher order moments and to investigate the impact of different regression ranges, the estimates are obtained as the slopes of the linear regressions of $\log m_r(t)$ on $\log t$ (theoretical lines in Figure 1 and also computed for all $r = 1, 2, \ldots, 6$). At small scales, over the range $[-20, -15]$ of $\log t$, the scaling exponents function is approximately 1 (the true value) and would lead to the conclusion of monofractal behaviour. On the other hand, over the ranges $[-20, t_{m,r}]$ and $[-15, t_{m,r}]$, the scaling exponents function is nonlinear showing evidence of (spurious) multifractality of the random process. At large scales, the estimate of $\zeta(r)$ is a linear function in r over both ranges considered, $[t_{m,r}, 5]$ and $[0, 5]$, being close to $\zeta(r) = r$.

Plots (c)–(d) in Figure 3 show the estimated scaling exponents function of the factorial moments. In this case, the lines look quite linear at both scales for the different ranges. The true functions $\zeta_{[]}(r) = 1+(r-1)\theta$ and $\zeta_{[]}(r) = r$ at small and large scales, respectively, are close to the empirical estimates. For example, the fit of the lines at small scales gives θ approximately equal to 0.75, 0.76 and 0.81 for $[-20, -15]$, $[-20, \log(t_{m,r})]$ and $[-15, \log(t_{m,r})]$, respectively (recall that θ was estimated as 0.6 in Section 8.1). This shows that at small scales, the exponent function can be estimated more reliably for the factorial moments than for the usual moments.

A Formulae relating central moments and factorial cumulants

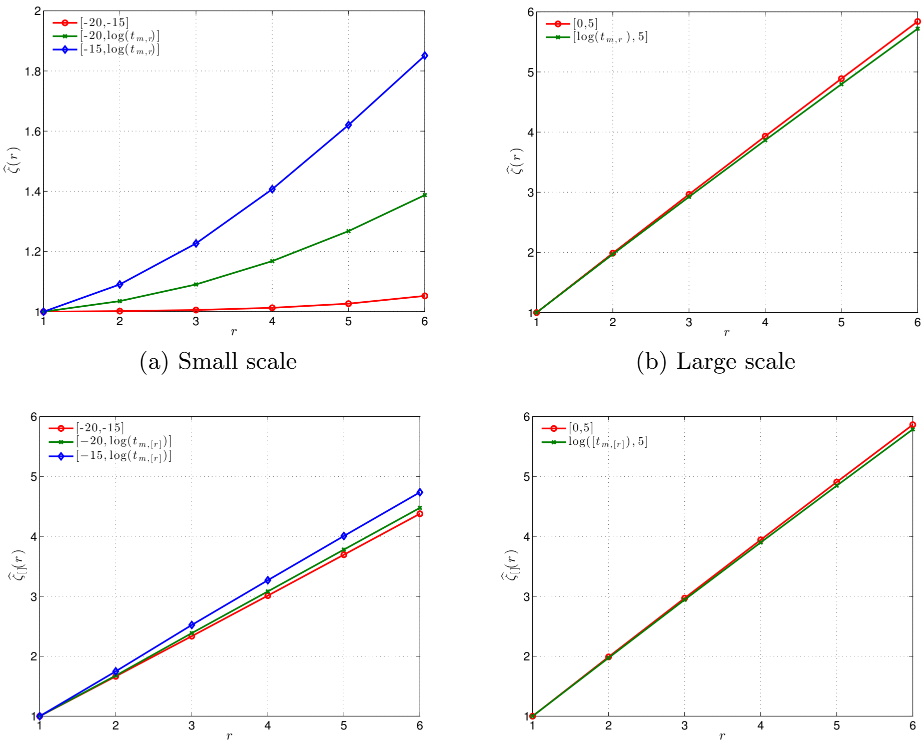
For Section 7, the relations between the first seven central moments and factorial cumulants using (2.16) along with (2.14) and (2.11) are listed below. We use the variable t instead of T in the observation as in Section 7 to emphasize that these are general relations between central moments and factorial cumulants.

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

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

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

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



(a) Small scale (b) Large scale

(c) Small scale (d) Large scale

Figure 3: Estimated scaling exponent functions: moments (top) and factorial moments (bottom).

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

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

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 (1973) when the first points in the clusters are excluded. We shall modify slightly the argument of Westcott (1973) 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

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

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 clust6er 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 (1973). To make connection to Westcott (1973), 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

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

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 $\int_0^{\infty}$ into $\int_x^{x+t}$ and $\int_0^x$, it follows that

$$\log P_t(z) = \lim_{x\to\infty} \Big(-\lambda \int_x^{x+t} \Big(1 - \mathbb{E} \prod_{k=0}^{\infty} z^{1_{(x,x+t)}(u+X_k)}\Big) du - \lambda \int_0^{x} \Big(1 - \mathbb{E} \prod_{k=0}^{\infty} z^{1_{(x,x+t)}(u+X_k)}\Big) du \Big)$$

$$= -\lambda \int_0^{t} \Big(1 - \mathbb{E} \prod_{k=0}^{\infty} z^{1_{(0,t)}(v+X_k)}\Big) dv - \lambda \int_0^{\infty} \Big(1 - \mathbb{E} \prod_{k=0}^{\infty} z^{1_{(v,v+t)}(X_k)}\Big) dv,$$

after the change of variables $v = u - x$ for the first integral, and $v = x - u$ for the second integral and letting $x \to \infty$. 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

$$\log P_t(z) = -\lambda \int_0^{t} \Big(1 - z\mathbb{E} \prod_{k=1}^{\infty} z^{1_{(0,t)}(v+X_k)}\Big) dv - \lambda \int_0^{\infty} \Big(1 - \mathbb{E} \prod_{k=1}^{\infty} z^{1_{(v,v+t)}(X_k)}\Big) dv$$

$$= -\lambda \Big(\int_0^{t} (1 - zP(z; t - v)) dv + \int_0^{\infty} (1 - P(z; v, v + t)) dv \Big)$$

$$= -\lambda \Big(\int_0^{t} (1 - zP(z; v)) dv + \int_0^{\infty} (1 - P(z; v, v + t)) dv \Big), \tag{B.1}$$

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 (1973), Eq. (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

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

(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 (1973) (see the arguments following Eq. (19) and terminating with Theorem 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

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

Basic algebraic manipulations lead to

$$\int_0^{\infty} (1 - P(z; v, v+t))dv = -(z-1)t\mathbb{E}W^{(0)}$$

$$-(z-1)^2 \sum_{k=1}^{\infty} z^{k-1} \int_0^t F_k(x)dx \sum_{j=1}^{\infty} R_{j+k}^{(0)} - (z-1) \sum_{k=1}^{\infty} z^{k-1} \int_0^t F_k(x)dx R_k^{(0)}. \tag{B.3}$$

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

$$\log P_t(z) = (z-1)\lambda t + (z-1)^2 \lambda \sum_{j=0}^{\infty} z^j R_{j+1}^{(0)} \int_0^t F_{j+1}(x)dx$$

$$+(z-1)\lambda t\mathbb{E}W^{(0)} + (z-1)^2 \lambda \sum_{k=1}^{\infty} z^{k-1} \int_0^t F_k(x)dx \sum_{j=1}^{\infty} R_{j+k}^{(0)}$$

$$= \lambda(z-1)\Big(t + t\mathbb{E}W^{(0)} + (z-1) \sum_{k=1}^{\infty} z^{k-1} \int_0^t F_k(x)dx \sum_{j=0}^{\infty} R_{j+k}^{(0)}\Big).$$

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

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

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

References

Abry, P., Borgnat, P., Ricciato, F., Scherrer, A. & Veitch, D. (2010), 'Revisiting an old friend: on the observability of the relation between long range dependence and heavy tail', *Telecommunication Systems* **43**(3-4), 147–165.

Antunes, N. & Pipiras, V. (2015), Estimation of flow distributions from sampled traffic, Preprint.

Bartlett, M. S. (1963), 'The spectral analysis of point processes', *Journal of the Royal Statistical Society. Series B. Methodological* **25**, 264–296.

Carruthers, P., Friedlander, E. M., Shih, C. C. & Weiner, R. M. (1989), 'Multiplicity fluctuations in finite rapidity windows. Intermittency or quantum statistical correlation?', *Physics Letters B* **222**(3), 487–492.

Cox, D. R. & Isham, V. (1980), *Point Processes*, Chapman & Hall, London-New York. Monographs on Applied Probability and Statistics.

Daley, D. J. & Vere-Jones, D. (2003), *An Introduction to the Theory of Point Processes. Vol. I*, Probability and its Applications (New York), second edn, Springer-Verlag, New York. Elementary Theory and Methods.

De Wolf, E. A., Dremin, I. M. & Kittel, W. (1996), 'Scaling laws for density correlations and fluctuations in multiparticle dynamics', *Physics Reports* **270**(1), 1–141.

Dombry, C. & Kaj, I. (2013), 'Moment measures of heavy-tailed renewal point processes: asymptotics and applications', *ESAIM. Probability and Statistics* **17**, 567–591.

Fasen, V. & Samorodnitsky, G. (2009), 'A fluid cluster Poisson input process can look like a fractional Brownian motion even in the slow growth aggregation regime', *Advances in Applied Probability* **41**(2), 393–427.

Faÿ, G., González-Arévalo, B., Mikosch, T. & Samorodnitsky, G. (2006), 'Modeling teletraffic arrivals by a Poisson cluster process', *Queueing Systems* **54**(2), 121–140.

Gaigalas, R. & Kaj, I. (2003), 'Convergence of scaled renewal processes and a packet arrival model', *Bernoulli* **9**(4), 671–703.

González-Arévalo, B. & Roy, J. (2010), 'Simulating a Poisson cluster process for Internet traffic packet arrivals', *Computer Communications* **33**(5), 612–618.

Grüneis, F., Nakao, M. & Yamamoto, M. (1990), 'Counting statistics of 1/f fluctuations in neuronal spike trains', *Biological Cybernetics* **62**(5), 407–413.

Grüneis, F., Nakao, M., Yamamoto, M., Musha, T. & Nakahama, H. (1989), 'An interpretation of 1/f fluctuations in neuronal spike trains during dream sleep', *Biological Cybernetics* **60**(3), 161–169.

Guerin, C. A., Nyberg, H., Perrin, O., Resnick, S., Rootzén, H. & Stărică, C. (2003), 'Empirical testing of the infinite source Poisson data traffic model', *Stochastic Models* **19**, 56–199.

Gut, A. (2009), *Stopped Random Walks*, Springer Series in Operations Research and Financial Engineering, second edn, Springer, New York. Limit theorems and applications.

Hohn, N., Veitch, D. & Abry, P. (2003), 'Cluster processes: a natural language for network traffic', *IEEE Trans. Signal Process.* **51**(8), 2229–2244.

Jaffard, S., Lashermes, B. & Abry, P. (2007), *Wavelet Analysis and Applications*, Birkhäuser Basel, chapter Wavelet Leaders in Multifractal Analysis, pp. 201–246.

Kaj, I. (2002), *Stochastic Modeling in Broadband Communications Systems*, SIAM Monographs on Mathematical Modeling and Computation, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA.

Karr, A. F. (1991), *Point Processes and their Statistical Inference*, Vol. 7 of *Probability: Pure and Applied*, second edn, Marcel Dekker, Inc., New York.

Krishna, P. M., Gadre, V. & Desai, U. B. (2012), *Multifractal Based Network Traffic Modeling*, Springer Science & Business Media.

Leland, W. E., Taqqu, M. S., Willinger, W. & Wilson, D. V. (1994), 'On the self-similar nature of Ethernet traffic (Extended version)', *IEEE/ACM Transactions on Networking* **2**, 1–15.

Lewis, P. A. W. (1964), 'A branching Poisson process model for the analysis of computer failure patterns (with discussion)', *Journal of the Royal Statistical Society. Series B. Methodological* **26**, 398–456.

Lowen, S. B. & Teich, M. C. (2005), *Fractal-Based Point Processes*, Wiley Series in Probability and Statistics, Wiley-Interscience [John Wiley & Sons], Hoboken, NJ.

Mandelbrot, B. (1997), A case against the lognormal distribution, *in* 'Fractals and Scaling in Finance', Springer New York, pp. 252–269.

Mikosch, T. & Samorodnitsky, G. (2007), 'Scaling limits for cumulative input processes', *Mathematics of Operations Research* **32**(4), 890–918.

Mikosch, T., Resnick, S., Rootzén, H. & Stegeman, A. (2002), 'Is network traffic approximated by stable Lévy motion or fractional Brownian motion?', *The Annals of Applied Probability* **12**(1), 23–68.

Onof, C., Chandler, R. E., Kakou, A., Northrop, P., Wheater, H. S. & Isham, V. (2000), 'Rainfall modelling using Poisson-cluster processes: a review of developments', *Stochastic Environmental Research and Risk Assessment* **14**(6), 384–411.

Ribeiro, V. J., Zhang, Z.-L., Moon, S. & Diot, C. (2005), 'Small-time scaling behavior of Internet backbone traffic', *Computer Networks* **48**(3), 315–334.

Saleh, B. E. A. & Teich, M. C. (1982), 'Multiplied-Poisson noise in pulse, particle, and photon detection', *Proceedings of the IEEE* **70**(3), 229–245.

Veitch, D., Hohn, N. & Abry, P. (2005), 'Multifractality in TCP/IP traffic: the case against', *Computer Networks* **48**(3), 293–313.

Westcott, M. (1973), 'Results in the asymptotic and equilibrium theory of Poisson cluster processes', *Journal of Applied Probability* **10**(4), 807–823.

Willink, R. (2003), 'Relationships between central moments and cumulants, with formulae for the central moments of gamma distributions', *Communications in Statistics - Theory and Methods* **32**(4), 701–704.

Zeephongsekul, P., Xia, G. & Kumar, S. (1994), 'Software reliability growth models based on cluster point processes', *International Journal of Systems Science* **25**(4), 737–751.

(N. Antunes) Center for Computational and Stochastic Mathematics, University of Lisbon, Avenida Rovisco Pais 1049-001, Lisbon, Portugal.
E-mail address: nantunes@ualg.pt

(V. Pipiras) Department of Statistics and Operations Research, University of North Carolina, CB 3260, Chapel Hill, NC 27599, USA.
E-mail address: pipiras@email.unc.edu

(P. Abry) CNRS and École Normale Supérieure de Lyon, 46, allée d'Italie, F-69364 Lyon cedex 7, France.
E-mail address: patrice.abry@ens-lyon.fr

(D. Veitch) School of Computing and Communications, University of Technology Sydney, P.O. Box 123, Broadway, NSW 2007, Australia.
E-mail address: Darryl.Veitch@uts.edu.au