Queueing Syst (2011) 69:29–44
DOI 10.1007/s11134-011-9231-4

The on–off network traffic model under intermediate scaling

Clément Dombry · Ingemar Kaj

Received: 24 September 2010 / Published online: 18 June 2011

Abstract The result provided in this paper helps complete a unified picture of the scaling behavior in heavy-tailed stochastic models for transmission of packet traffic on high-speed communication links. Popular models include infinite source Poisson models, models based on aggregated renewal sequences, and models built from aggregated on–off sources. The versions of these models with finite variance transmission rate share the following pattern: if the sources connect at a fast rate over time the cumulative statistical fluctuations are fractional Brownian motion, if the connection rate is slow the traffic fluctuations are described by a stable Lévy motion, while the limiting fluctuations for the intermediate scaling regime are given by fractional Poisson motion. In this paper, we prove an invariance principle for the normalized cumulative workload of a network with m on–off sources and time rescaled by a factor a. When both the number of sources m and the time scale a tend to infinity with a relative growth given by the so-called 'intermediate connection rate' condition, the limit process is the fractional Poisson motion. The proof is based on a coupling between the on–off model and the renewal type model.

Keywords On–off process · Workload process · Renewal process · Intermediate scaling · Fractional Poisson motion · Fractional Brownian motion · Lévy motion · Heavy tails · Long-range dependence

Mathematics Subject Classification (2000) 60G22 · 60F05

C. Dombry
Laboratoire LMA, Université de Poitiers, 86962 Futuroscope-Chasseneuil cedex, France
e-mail: clement.dombry@math.univ-poitiers.fr

I. Kaj (✉)
Department of Mathematics, Uppsala University, Box 480, SE 751 06 Uppsala, Sweden
e-mail: ikaj@math.uu.se

1 Introduction

It is well-known that packet traffic on high-speed links exhibit data characteristics consistent with long-range dependence and self-similarity. To explain the possible mechanisms behind this behavior, various network traffic models have been developed where these features arise as heavy-tail phenomena; see Resnick [20]. A natural basis for modeling such systems, applied early on during these developments, is the view of packet traffic composed of a large number of aggregated streams where each source alternates between an active on-state transmitting data and an inactive off-state. The traffic streams generate on average a given mean-rate traffic, they have stationary increments and they are considered statistically independent. In particular, the transmission channel is able to accommodate peak-rate traffic corresponding to all sources being in the on-state. To capture in this model the strong positive dependence manifest in empirical trace data measurements, it is assumed that the duration of on-periods and/or off-periods are subject to heavy-tailed probability distributions. It is then interesting to analyze the workload of total traffic over time and understand the random fluctuations around its cumulative average. Our continued interest in these questions comes from the finding that several scaling regimes exist with disparate asymptotic limits.

The first result of the type we have in mind is Taqqu, Willinger and Sherman [21], which introduces a double limit technique. In this sequential scheme, if the on–off model is averaged first over the level of aggregation and then over time the resulting limit process is fractional Brownian motion. As the fundamental example of a Gaussian self-similar process with long-range dependence, this limit preserves the inherent long-range dependence of the original workload fluctuations. On the other hand, averaging first over time and then over the number of traffic sources the limit process is a stable Lévy motion. This alternative scaling limit is again self-similar but lacks long-range dependence since the increments are independent. Moreover, having infinite variance the limiting workload is itself heavy-tailed. In Mikosch, Resnick, Rootzén and Stegeman, [7], the double limits are replaced by a single scheme where instead the number of sources grows at a rate which is relative to time. Two limit regimes of fast growth and slow growth are identified and two limit results corresponding to these are established, where again fractional Brownian motion and stable Lévy motion appear as scaled limit processes of the centered on–off workload. The purpose of this paper is to show that an additional limit process, fractional Poisson motion, arises under an intermediate regime which can be viewed as a balanced scaling between slow and fast growth. In this case the scale of time grows essentially as a power function of the number of traffic sources. As will be recalled, fractional Poisson motion does indeed provide a bridge between fractional Brownian motion and stable Lévy motion.

The intermediate limit regime discussed here is indicated in Kaj [12], and introduced in Gaigalas and Kaj [10], where limit results are given for a different but related class of traffic models under three scaling regimes referred to as slow, intermediate and fast connection rates. The workload process is again the superposition of independent traffic streams with stationary increments but now each source generates packets according to a finite mean renewal counting process with heavy-tailed inter-renewal-cycle lengths. The link to the class of on–off models is that each pair of an on-period and a successive off-period forms a renewal cycle and the number of such on–off cycles generate a heavy-tailed renewal counting process. Moreover, if we associate with each renewal cycle a reward given by the length of its on-period and apply a suitable interpretation of partial rewards, then the corresponding renewal-reward process coincides with the on–off workload process.

To explain briefly the limit result in [10] under intermediate connection rate, let $(N^i(t))_i$ be i.i.d. copies of a stationary renewal counting process associated with a sequence of inter-renewal times of finite mean μ and a regularly varying tail function $\bar{F}(t) \sim L(t)t^{-\gamma}$, characterized by an index γ, $1 < \gamma < 2$, and a slowly varying function L. Let $m \to \infty$ and $a \to \infty$ in such a way that $mL(a)/a^{\gamma-1} \to \mu c^{\gamma-1}$ for some constant $c > 0$. Then the weak convergence holds,

$$\frac{1}{a}\sum_{i=1}^{m}\left(N^i(at) - \frac{at}{\mu}\right) \quad \Longrightarrow \quad -\frac{1}{\mu} c Y_\gamma(t/c),$$

where $Y_\gamma(t)$ is an almost surely continuous, positively skewed, non-Gaussian and non-stable random process, which is defined by a particular representation of the characteristic function of its finite-dimensional distributions. Additional properties of the limit process are obtained in Kaj [13] and Gaigalas [9], where it is shown with two different methods that Y_γ can be represented as a stochastic integral with respect to a Poisson measure $N(dx, du)$ on $\mathbb{R} \times \mathbb{R}^+$ with intensity measure $n(dx, du) = \gamma\, dx\, u^{-\gamma-1}\, du$. Indeed,

$$Y_\gamma(t) = \int_{\mathbb{R}\times\mathbb{R}^+} \int_0^t 1_{[x,x+u]}(y)\, dy \left(N(dx, du) - \gamma\, dx\, u^{-\gamma-1}\, du\right), \quad t \geq 0,$$

([10] uses $\bar{F}(t) \sim L(t)\gamma^{-1}t^{-\gamma}$, consequently $n(dx, du) = dx\, u^{-\gamma-1}\, du$). We call this process fractional Poisson motion with Hurst index $H = (3-\gamma)/2 \in (1/2, 1)$. With

$$\sigma_\gamma^2 = \frac{2}{(\gamma-1)(2-\gamma)(3-\gamma)} = \frac{1}{2H(1-H)(2H-1)} = \sigma_H^2, \tag{1}$$

we may put $Y_\gamma(t) = \sigma_\gamma P_H(t)$ and obtain the standard fractional Poisson motion P_H. A calculation reveals

$$\mathrm{Cov}(P_H(s), P_H(t)) = \frac{1}{2}\left(|s|^{2H} + |t|^{2H} - |t-s|^{2H}\right).$$

For comparison, standard fractional Brownian motion of index H has the representation

$$B_H(t) = \frac{1}{\sigma_H}\int_{\mathbb{R}\times\mathbb{R}^+} \int_0^t 1_{[x,x+u]}(y)\, dy\, M(dx, du),$$

where $M(dx, du)$ is a Gaussian random measure on $\mathbb{R} \times \mathbb{R}^+$ which is characterized by the control measure $(3-2H)\, dx\, u^{-2(2-H)}\, du$. The covariance functions of B_H and P_H coincide. The fast connection rate limit for the model of aggregated renewal processes applies if $mL(a)/a^{\gamma-1} \to \infty$ and the slow connection rate limit if $mL(a)/a^{\gamma-1} \to 0$. For suitable normalizing sequences, the limit processes under these assumptions are fractional Brownian motion with Hurst index $H = (3 - \gamma)/2$ in the case of fast growth and a stable Lévy motion with self-similarity index $1/\gamma$ in the slow growth situation, see [10].

A number of other models have been suggested for the flow of traffic in communication networks. The superposition of independent renewal-reward processes applies more generally to sources which attain random transmission rates at random times, and not merely switch between on and off. For a model where the length of a transmission cycle as well as the transmission rate during the cycle are allowed to be heavy-tailed, Levy and Taqqu [15], Pipiras and Taqqu [18], and Pipiras, Taqqu and Levy [19], established results for slow and fast growth scaling analogous to those for the on–off model. In addition, they obtained as a fast growth scaling limit a stable, self-similar process with stationary but not independent increments, coined the telecom process. A further category of models for network traffic with long-range dependence over time starts from the assumption that long-lived traffic sessions arrive according to a Poisson process. The sessions carry workload which is transmitted either at fixed rate, at a random rate throughout the session, or at a randomly varying rate over the session length. Such models, called infinite source Poisson models, are widely accepted as realistic workload processes for Internet traffic. Indeed, it is natural to assume that web flows on a non-congested backbone link are initiated according to a Poisson process while the duration of sessions and transmission rates are highly variable. The conditions under which slow, intermediate and fast scaling results exist and fractional Brownian motion, fractional Poisson motion, stable Lévy motion and telecom processes arise in the asymptotic limits are known in great detail for variants of the infinite source Poisson model, see Kaj and Taqqu [14]. In [14], Y_γ is called the intermediate telecom process. Mikosch and Samorodnitsky [8] consider scaling limits for a general class of input processes, which includes as special cases the models already mentioned as well as other cumulative cluster-type processes. It is shown that fractional Brownian motion is a robust limit for a variety of models under fast growth conditions, whereas the slow growth behavior is more variable with a number of different stable processes arising in the limit.

Our current result completes the picture for the intermediate scaling regime, where neither of the mechanisms of fast or slow growth are predominant. In this case, we show that the fluctuations which build up in the on–off model are robust and again described by the fractional Poisson motion, parallel to what is known to be valid for infinite source Poisson and renewal-based traffic models. In Sect. 2 we introduce properly both the on–off model and the renewal-based model to be used as an approximation and we state the relevant background results for these models. In Sect. 3 we state the main result and give the structure of the proof. Section 4 is devoted to remaining and technical aspects of the proof.

To conclude this introduction, we add some comments on the term *fractional Poisson motion*, which is also used in [4]. The process P_H appears naturally as a limit process, it has the same covariance function as fractional Brownian motion and a similar integral representation, with respect to a Poisson measure rather than an Gaussian measure. In [4] it is shown that the analogous limit process P_H for the case $0 < H < 1/2$ is such that its marginals $P_H(t)$ have the same distributions as $(X_t - X'_t)/\sqrt{2}$, where X_t, X'_t are independent random variables both with a Poisson distribution of mean t^{2H}. All in all, the phrase fractional Poisson motion seems natural. However, we should stress that the closely related term *fractional Poisson process* appears in several recent papers to denote other objects. Wang et al. [22–24], construct processes called fractional Poisson processes as stochastic integrals with respect to Poisson random measures and study their properties; the resulting processes are different from P_H considered here, since different kernels are used. Another approach, entirely different from the present one, is based on fractional differential equations obtained by replacing the standard derivative in the equations governing the probability distribution of the homogeneous Poisson process by a fractional derivative, see for example the papers by Jumarie [11], by Mainardi et al. [16] or by Beghin and Orsingher [2, 3]. These fractional Poisson processes are different from the process P_H considered in the present paper.

2 The on–off model and background results

We begin by introducing the on–off model using similar notations as in [7]. Let X_{on} and $X_1, X_2, \ldots$ be i.i.d. non-negative random variables with distribution F_{on} representing the lengths of on-periods. Similarly let $Y_{\text{off}}, Y_1, Y_2, \ldots$ be i.i.d. non-negative random variables with distribution F_{off} representing the lengths of off-periods. The X- and Y-sequences are supposed to be independent. For any distribution function F we write $\bar{F} = 1 - F$ for the right tail. We fix two parameters, α_{on} and α_{off}, such that

$$1 < \alpha_{\text{on}} < \alpha_{\text{off}} < 2, \tag{2}$$

and assume that

$$\bar{F}_{\text{on}}(x) = x^{-\alpha_{\text{on}}} L_{\text{on}}(x) \quad \text{and} \quad \bar{F}_{\text{off}}(x) = x^{-\alpha_{\text{off}}} L_{\text{off}}(x), \quad x \to \infty, \tag{3}$$

with $L_{\text{on}}, L_{\text{off}}$ arbitrary functions slowly varying at infinity. Hence both distributions F_{on} and F_{off} have finite mean values μ_{on} and μ_{off} but their variances are infinite. Assumption (2) agrees with that of [7]. However, thanks to a simple symmetry argument, we can also cover the case $\alpha_{\text{on}} > \alpha_{\text{off}}$. The case $\alpha_{\text{on}} = \alpha_{\text{off}}$, for which the on–off process is an alternating renewal process, falls outside of the class of processes we are able to study within the methodology developed here.

We consider the renewal sequence generated by alternating on- and off-periods. For the purpose of stationarity we introduce random variables (X_0, Y_0) representing the initial on- and off-periods as follows: let B, $X_{\text{on}}^{\text{eq}}$, $Y_{\text{off}}^{\text{eq}}$ be independent random variables, independent of $\{X_{\text{on}}, (X_n), Y_{\text{off}}, (Y_n)\}$, and such that B is Bernoulli with

$$\mathbb{P}(B = 1) = 1 - \mathbb{P}(B = 0) = \mu_{\text{on}}/\mu,$$

and $X_{\text{on}}^{\text{eq}}$ and $Y_{\text{off}}^{\text{eq}}$ have distribution functions

$$F_{\text{on}}^{\text{eq}}(x) = \frac{1}{\mu_{\text{on}}} \int_0^x \bar{F}_{\text{on}}(s)\, ds \quad \text{and} \quad F_{\text{off}}^{\text{eq}}(x) = \frac{1}{\mu_{\text{off}}} \int_0^x \bar{F}_{\text{off}}(s)\, ds,$$

respectively. Here the superscript ‘eq’ stands for ‘equilibrium’. Now, let

$$X_0 = B X_{\text{on}}^{\text{eq}} \quad \text{and} \quad Y_0 = B Y_{\text{off}} + (1 - B) Y_{\text{off}}^{\text{eq}}.$$

Note that X_0 and Y_0 are conditionally independent given B but not independent. At time $t = 0$ the system starts in the on-state if $B = 1$ and in the off-state if $B = 0$. With this initial distribution, the alternating renewal sequence is stationary and the probability that the system is in the on-state at any time t is μ_{on}/μ. Renewal events occur at the start of each on-period. Inter-renewal times are given by the independent sequence $Z_i = X_i + Y_i, i \geq 0$, where Z_i has distribution $F = F_{\mathrm{on}} * F_{\mathrm{off}}$ and mean $\mu = \mu_{\mathrm{on}} + \mu_{\mathrm{off}}$ for $i \geq 1$, and Z_0 has distribution function

$$F^{\mathrm{eq}}(x) = \frac{1}{\mu} \int_0^x \bar{F}(s)\, ds.$$

The renewal sequence $(T_n)_{n \geq 1}$ with delay T_0 is defined by

$$T_n = \sum_{i=0}^{n} Z_i,$$

and we denote by $N(t)$ the associated counting process

$$N(t) = \sum_{n \geq 0} 1_{(0,t]}(T_n).$$

Note that $N(t)$ has stationary increments and expectation $\mathbb{E}[N(t)] = t/\mu$. Moreover, because of (2), the tail behavior of the inter-renewal times is given by

$$\bar{F}(x) \sim L_{\mathrm{on}}(x) x^{-\alpha_{\mathrm{on}}}, \quad x \to \infty, \tag{4}$$

see Asmussen [1], Chap. IX, Corollary 1.11. The on–off input process is the indicator process for the on-state defined by

$$I(t) = 1_{[0,X_0)}(t) + \sum_{n \geq 0} 1_{[T_n, T_n + X_{n+1})}(t), \quad t \geq 0.$$

The source is in the on-state if $I(t) = 1$ and in the off-state if $I(t) = 0$. The input process $I(t)$ is strictly stationary with mean

$$\mathbb{E}\big[I(t)\big] = \mathbb{P}\big(I(t) = 1\big) = \mu_{\mathrm{on}}/\mu.$$

The associated cumulative workload defined by

$$W(t) = \int_0^t I(s)\, ds, \quad t \geq 0$$

is a stationary increment process with mean $\mathbb{E}[W(t)] = t\mu_{\mathrm{on}}/\mu$.

Let $(I^j, W^j, N^j)_{j \geq 1}$ denote i.i.d. copies of the input process I, the cumulative workload process W, and the renewal counting process N for the stationary on–off model. For $m \geq 1$, consider a server fed by m independent on–off sources. We define the cumulative workload of the m-server system as the superposition process

$$W_m(t) = \sum_{j=1}^{m} W^j(t), \quad t \geq 0,\ m \geq 1,$$

and the renewal-cycle counting process for m aggregated traffic sources by

$$N_m(t) = \sum_{j=1}^{m} N^j(t), \quad t \geq 0,\ m \geq 1.$$

In this paper, we are mainly concerned with the asymptotic properties of the cumulative workload when the number of sources, m, increases and time t is rescaled by a factor $a > 0$. Thus, we consider the centered and rescaled process

$$\frac{W_m(at) - mat\mu_{\text{on}}/\mu}{b(a,m)} = \frac{1}{b(a,m)} \sum_{j=1}^{m} \int_0^{at} \left(I^j(s) - \frac{\mu_{\text{on}}}{\mu} \right) ds, \quad t \geq 0,$$

where the renormalization $b(a,m)$ will be made precise in the sequel. The asymptotic is considered when both $m \to \infty$ and $a \to \infty$. To deal with the simultaneous limits, one usually consider the limits either when $a = a(m) \to \infty$ as $a \to \infty$ or when $m = m(a) \to \infty$ as $a \to \infty$. Preference of one over the other asymptotic is a matter of interpretation of the model, and both cases are used in the literature. We consider here the slightly more general case when $m = m_n$ and $a = a_n$ and we will always suppose that

$$m = m_n \to \infty \quad \text{and} \quad a = a_n \to \infty \quad \text{as } n \to \infty.$$

With no possible confusion, the parameter n will generally be omitted, and all asymptotics will be implicitely considered as $n \to \infty$, unless stated otherwise. The relative growth of m and a have a major impact on the limit. Following the notation in [10], we consider the following three scaling regimes:

- fast connection rate

$$m_n \to \infty, \quad a_n \to \infty, \quad m_n L_{\text{on}}(a_n)/a_n^{\alpha_{\text{on}}-1} \to \infty; \tag{FCR}$$

- slow connection rate

$$m_n \to \infty, \quad a_n \to \infty, \quad m_n L_{\text{on}}(a_n)/a_n^{\alpha_{\text{on}}-1} \to 0; \tag{SCR}$$

- intermediate connection rate

$$m_n \to \infty, \quad a_n \to \infty, \quad m_n L_{\text{on}}(a_n)/a_n^{\alpha_{\text{on}}-1} \to \mu c^{\alpha_{\text{on}}-1}, \quad 0 < c < \infty. \tag{ICR}$$

In [7], the asymptotic behavior of the cumulative total workload is investigated under conditions (FCR) and (SCR).

Theorem 1 (Mikosch et al. [7]) *Recall assumptions* (2) *and* (3).

- *Under condition* (FCR) *and with the normalization* $b(a,m) = (a^{3-\alpha_{\text{on}}} L_{\text{on}}(a) m)^{1/2}$, *the following weak convergence of processes holds in the space of continuous functions on* $\mathbb{R}^+$ *endowed with the topology of uniform convergence on compact sets*:

$$\frac{W_m(at) - mat\mu_{\text{on}}/\mu}{b(a,m)} \quad \Longrightarrow \quad \sigma_{\alpha_{\text{on}}} \frac{\mu_{\text{on}}}{\mu^{3/2}} B_H(t), \quad t \geq 0$$

where $\sigma_{\alpha_{\mathrm{on}}}$ is given in (1) and $B_H(t)$ is a standard fractional Brownian motion with index $H = (3 - \alpha_{\mathrm{on}})/2$.

- *Under condition (SCR) and with the normalization*

$$b(a,m) = \inf\{x \geq 0 : \bar{F}_{\mathrm{on}}(x) \leq 1/am\},$$

the following convergence of finite-dimensional distributions holds:

$$\frac{W_m(at) - mat\mu_{\mathrm{on}}/\mu}{b(a,m)} \xrightarrow{fdd} \sigma_0 \frac{\mu_{\mathrm{off}}}{\mu^{1+1/\alpha_{\mathrm{on}}}} X_{\alpha_{\mathrm{on}},1,1}(t), \quad t \geq 0,$$

where $X_{\alpha_{\mathrm{on}},1,1}(t)$ is a standard α_{on}-stable Lévy motion totally skewed to the right, i.e. such that

$$X_{\alpha_{\mathrm{on}},1,1}(1) \sim S_{\alpha_{\mathrm{on}}}(1,1,0),$$

and

$$\sigma_0 = \frac{\Gamma(2-\alpha_{\mathrm{on}})\cos(\pi\alpha_{\mathrm{on}}/2)}{1-\alpha_{\mathrm{on}}}.$$

The intermediate regime for renewal processes was investigated in [10]. The formulation adopted here is given in [13], and is an immediate consequence of (4).

Theorem 2 (Gaigalas and Kaj [10]) *Under condition (ICR) and with the normalization $b(a,m) = a$, the following convergence of processes holds in the space of càd-làg functions on $\mathbb{R}^+$ endowed with the Skorokhod J_1-topology:*

$$\frac{N_m(at) - mat/\mu}{a} \quad \Longrightarrow \quad -\frac{1}{\mu}\sigma_{\alpha_{\mathrm{on}}} c P_H(t/c),$$

where $\sigma_{\alpha_{\mathrm{on}}}$ is given in (1), $P_H(t)$ is the standard fractional Poisson motion with Hurst index $H = (3 - \alpha_{\mathrm{on}})/2$

$$P_H(t) = \frac{1}{\sigma_{\alpha_{\mathrm{on}}}} \int_{\mathbb{R}\times\mathbb{R}^+} \int_0^t 1_{[x,x+u]}(y)\,dy\,\big(N(dx,du) - dx\,\alpha_{\mathrm{on}} u^{-\alpha_{\mathrm{on}}-1}\,du\big), \qquad (5)$$

and $N(dx,du)$ is a Poisson random measure on $\mathbb{R} \times \mathbb{R}^+$ with intensity $dx\,\alpha_{\mathrm{on}} \times u^{-\alpha_{\mathrm{on}}-1}\,du$.

3 Intermediate limit for the on–off model

In this section we investigate the intermediate scaling limit for the on–off model. The following is our main result.

Theorem 3 *Under condition (ICR) and with the normalization $b(a,m) = a$, the following convergence of processes holds in the space of continuous functions on $\mathbb{R}^+$ endowed with the topology of uniform convergence on compact sets:*

$$\frac{W_m(at) - mat\mu_{\mathrm{on}}/\mu}{a} \quad \Longrightarrow \quad \sigma_{\alpha_{\mathrm{on}}} \frac{\mu_{\mathrm{off}}}{\mu} c P_H(t/c),$$

with $P_H(t)$ *the standard fractional Poisson motion defined in* (5) *and* $\sigma_{\alpha_{\text{on}}} > 0$ *given by*

$$\sigma^2_{\alpha_{\text{on}}} = \frac{2}{(\alpha_{\text{on}} - 1)(2 - \alpha_{\text{on}})(3 - \alpha_{\text{on}})}.$$

Remarks 4 Basic properties of the process P_H, here and in other recent work called fractional Poisson motion, are discussed in the original papers [10] and [9]. Explicit formulas are given for moments and cumulants, and for the cumulant generating function of the marginal distributions. In particular, the one-dimensional marginals are characterized by their cumulant generating function

$$\log \mathbb{E} e^{\theta\sigma_{\alpha_{\text{on}}} P_H(t)} = \frac{1}{\alpha_{\text{on}} - 1} \int_0^t \int_0^v \theta^2 e^{\theta u} u^{-(\alpha_{\text{on}}-1)}\, du\, dv.$$

Other forms of these generating functions, and characteristic functions, which explain the Poisson representation (5), are given in [9, 13, 14]. However, to our knowledge there are no useful representation results known for the density function of $P_H(t)$. The trajectories of fractional Poisson motion P_H are known to be Hölder continuous of order δ for any $\delta \in (0, H)$. The fractional Poisson motion has stationary increments; it is not self-similar but does have a property of aggregate-similarity, introduced in [13], which allows for an interpretation of the scaling parameter c. Consider for each integer $m \geq 1$ the sequence $c_m = m^{1/(\alpha_{\text{on}}-1)}$. Then

$$c_m P_H(t/c_m) \stackrel{\text{fdd}}{=} \sum_{i=1}^m P_H^i(t),$$

where $P_H^1, P_H^2, \ldots$ are i.i.d. copies of P_H. Consider also the sequence $c'_m = m^{-1/(\alpha_{\text{on}}-1)}$. For any m,

$$\sum_{i=1}^m c'_m P_H^i(t/c'_m) \stackrel{\text{fdd}}{=} P_H(t).$$

Hence, by tracing the limit process in Theorem 3 as $c_m \to \infty$, we recover in distribution the succession of all aggregates $\sum_{1\leq i\leq m} P_H^i$, $m \geq 1$. Also, by letting $c'_m \to 0$ we find that the limit process represents successively smaller fractions which sum up to recover fractional Poisson motion.

These relations explain the fact that fractional Poisson motion acts as a bridge between the stable Lévy motion and fractional Brownian motion. First, $\{c^H P_H(t/c)\}$ converges weakly to $\{B_H(t)\}$, as $c \to \infty$. Indeed, $c_m^H P_H(t/c_m) \stackrel{\text{fdd}}{=} \frac{1}{\sqrt{m}} \sum_{1\leq i\leq m} P_H^i(t)$ and the Central Limit Theorem yields the Gaussian limit as $m \to \infty$. The required tightness property is shown in [9]. Moreover, it is shown in [9] that $c^{1/\alpha_{\text{on}}} P_H(t/c)$ converges in the sense of the finite-dimensional distributions as $c \to 0$ to the α_{on}-stable Lévy motion. To see that the limit must be α_{on}-stable, take

$d = c \cdot c'_m$ for any $c > 0$. Then

$$c^{1/\alpha_{\mathrm{on}}} P_H(t/c) \stackrel{\mathrm{fdd}}{=} \frac{1}{m^{1/\alpha_{\mathrm{on}}}} \sum_{i=1}^{m} d^{1/\alpha_{\mathrm{on}}} P_H^i(t/d), \quad m \geq 1,$$

and, assuming that the rescaled process $(c^{1/\alpha_{\mathrm{on}}} P_H(t/c))_{t\geq 0}$ converges to some non-trivial limit process L, we must have as $c \to 0$ (and hence $d \to 0$)

$$L(t) \stackrel{\mathrm{fdd}}{=} \frac{1}{m^{1/\alpha_{\mathrm{on}}}} \sum_{i=1}^{m} L^i(t), \quad m \geq 1.$$

This indicates that the limit L must be α_{on}-stable.

Heuristics of the proof of Theorem 3 To motivate that the limit process under intermediate connection rate appears naturally, we discuss a decomposition of the centered on–off process based on its representation as a renewal-reward model. We first note that the single source cumulative workload has the form

$$W(t) = X_0 \wedge t + \sum_{i=1}^{N(t)} X_i - (T_{N(t)-1} + X_{N(t)} - t)_+.$$

Similarly, focusing on off-periods rather than on-periods, we have

$$t - W(t) = Y_0 \wedge t + \sum_{i=1}^{N(t)} Y_i - (T_{N(t)} - t) \wedge Y_{N(t)}.$$

The centered single source workload is therefore

$$\begin{aligned} W(t) - \frac{\mu_{\mathrm{on}}}{\mu} t &= -\bigl(t - W(t)\bigr) + \frac{\mu_{\mathrm{off}}}{\mu} t \\ &= -\mu_{\mathrm{off}}\bigl(N(t) - t/\mu\bigr) - \sum_{i=1}^{N(t)} (Y_i - \mu_{\mathrm{off}}) + R(t) \end{aligned}$$

with

$$R(t) = (T_{N(t)} - t) \wedge Y_{N(t)} - Y_0 \wedge t.$$

Thus, for the workload of m sources,

$$\begin{aligned} W_m(t) - \frac{\mu_{\mathrm{on}}}{\mu} mt &= -\mu_{\mathrm{off}}\bigl(N_m(t) - mt/\mu\bigr) \\ &\quad - \sum_{j=1}^{m} \sum_{i=1}^{N^j(t)} \bigl(Y_i^j - \mu_{\mathrm{off}}\bigr) + \sum_{j=1}^{m} R^j(t) \end{aligned} \tag{6}$$

using obvious notations. The balancing of terms under the scaling relation (ICR), makes it plausible that both terms

$$\frac{1}{a}\sum_{j=1}^{m}\sum_{i=1}^{N^j(at)}\big(Y_i^j-\mu_{\text{off}}\big),\qquad \frac{1}{a}\sum_{j=1}^{m}R^j(at)$$

vanish in the limit. This suggests asymptotically,

$$\frac{W_m(at)-\mu_{\text{on}}mat/\mu}{a}\sim -\mu_{\text{off}}\frac{N_m(at)-mat/\mu}{a},\tag{7}$$

and so Theorem 2 would imply Theorem 3. In the next final section, we will compare rigorously the two processes in (7).

4 Proof of Theorem 3

The proof of Theorem 3 relies on the following three lemmas. Here, $\xrightarrow{P}$ denotes convergence in probability.

Lemma 1 *Under* (ICR) *regime, for all* $t\geq 0$,

$$\frac{1}{a}\sum_{j=1}^{m}\sum_{i=1}^{N^j(at)}\big(Y_i^j-\mu_{\text{off}}\big)\xrightarrow{P}0.\tag{8}$$

Lemma 2 *Under* (ICR) *regime, for all* $t\geq 0$,

$$\frac{1}{a}\sum_{j=1}^{m}R^j(at)\xrightarrow{P}0.$$

Lemma 3 *Under* (ICR) *regime, the sequence of processes*

$$\frac{W_m(at)-mat\mu_{\text{on}}/\mu}{a},\quad t\geq 0,\ n\geq 1$$

is tight in the space of continuous functions on $\mathbb{R}^+$.

Proof of Theorem 3 By Theorem 2, Lemma 1 and Lemma 2, the convergence of finite-dimensional distributions,

$$\frac{W_m(at)-mat\mu_{\text{on}}/\mu}{a}\longrightarrow\sigma_{\alpha_{\text{on}}}\frac{\mu_{\text{off}}}{\mu}cP_H(t/c),$$

is a consequence of the decomposition given in (6). By Lemma 3 the sequence is tight in the space of continuous functions on $\mathbb{R}^+$. Hence weak convergence holds in the space of continuous functions and Theorem 3 is proved. □

Proof of Lemma 1 We construct an alternative representation of the random variable in the left hand side of (8). Define

$$\widetilde{N}^1(at) = \inf\left\{k \geq 0;\, X_0^1 + \sum_{i=1}^{k} Z_i^1 \geq at\right\}$$

and for $j \geq 2$

$$\widetilde{N}^j(at) = \inf\left\{k \geq 0;\, X_0^j + \sum_{i=1}^{k} Z^1_{\widetilde{N}^{j-1}(at)+i} \geq at\right\}.$$

For $m \geq 1$, let $\widetilde{N}_m(at) = \sum_{j=1}^m \widetilde{N}^j(at)$. The random variables $\widetilde{N}^j(at)$, $j \geq 1$ are i.i.d. and for each fixed $t \geq 0$

$$\frac{1}{a}\sum_{j=1}^{m}\sum_{i=1}^{N^j(at)}\left(Y_i^j - \mu_{\text{off}}\right) \quad \text{and} \quad \frac{1}{a}\sum_{i=1}^{\widetilde{N}_m(at)}\left(Y_i^1 - \mu_{\text{off}}\right),$$

have the same distribution (note that the uni-dimensional marginal distributions are equal but not the multidimensional distributions). This representation will enable us to prove that under assumption (ICR), in the space of cád-lág functions on $\mathbb{R}^+$ endowed with the Skorokhod J_1-topology, we have the convergence

$$\left(\frac{1}{a}\sum_{i=1}^{[amu]}\left(Y_i^1 - \mu_{\text{off}}\right)\right)_{u\geq 0} \quad \Longrightarrow \quad 0, \tag{9}$$

where $[x]$ denotes the largest integer less or equal to the real x. Moreover,

$$\frac{1}{am}\tilde{N}_m(at) \xrightarrow{P} \frac{t}{\mu}, \qquad t \geq 0. \tag{10}$$

Equations (9) and (10) together imply that, for each $t \geq 0$,

$$\frac{1}{a}\sum_{i=1}^{\tilde{N}_m(at)}\left(Y_i^1 - \mu_{\text{off}}\right) \xrightarrow{P} 0$$

and this proves the lemma. Thus, it remains to prove (9) and (10).

To this aim, recall that the random variables $Y_i^1, i \geq 1$, are i.i.d. with distribution such that the tail function $\bar{F}_{\text{off}}$ is regularly varying with index $-\alpha_{\text{off}}$. Hence there exists a regularly varying function L such that the centered and rescaled sum

$$\left(\frac{1}{(am)^{1/\alpha_{\text{off}}}L(am)}\sum_{i=1}^{[amu]}\left(Y_i^1 - \mu_{\text{off}}\right)\right)_{u\geq 0}$$

converges in the space of càd-làg functions to some α_{off}-stable Lévy motion (see [17], the exact form of L or of the limit process are not needed here). This implies the convergence property (9), since the scaling assumption (ICR) with $\alpha_{\mathrm{on}} < \alpha_{\mathrm{off}}$ implies $a \ll (am)^{1/\alpha_{\mathrm{off}}} L(am)$, in the sense that the ratio $a/(am)^{1/\alpha_{\mathrm{off}}} L(am)$ has limit $+\infty$.

We now prove relation (10). The stationary renewal process $N(t)$ has mean t/μ and variance given asymptotically by

$$\operatorname{Var}\bigl[N(t)\bigr] \sim \sigma^2_{\alpha_{\mathrm{on}}} \frac{1}{\mu^3} t^{3-\alpha_{\mathrm{on}}} L_{\mathrm{on}}(t), \quad t \to \infty,$$

see [10], (30), and references therein. Hence, $\frac{1}{am}\widetilde{N}_m(at)$ has mean t/μ and variance under scaling (ICR), such that

$$\begin{aligned}\operatorname{Var}\biggl[\frac{1}{am}\widetilde{N}_m(at)\biggr] &= \frac{1}{a^2 m}\operatorname{Var}\bigl[N(at)\bigr] \\ &\sim \frac{a^{1-\alpha_{\mathrm{on}}} L_{\mathrm{on}}(at)}{m} \sigma^2_{\alpha_{\mathrm{on}}} \frac{1}{\mu^3} t^{3-\alpha_{\mathrm{on}}} \to 0, \quad a \to \infty.\end{aligned}$$

This shows that $\frac{1}{am}\widetilde{N}_m(at)$ converges in the mean square sense to t/μ, which implies (10). This ends the proof of Lemma 1. □

Proof of Lemma 2 Since $|R^j(t)| \le Y_0^j + (T^j_{N^j(t)} - t) \wedge Y^j_{N^j(t)}$, it is enough to prove

$$\frac{1}{a}\sum_{j=1}^m Y_0^j \xrightarrow{P} 0 \quad \text{and} \quad \frac{1}{a}\sum_{j=1}^m \bigl(T^j_{N_t^j} - t\bigr) \wedge Y^j_{N^j(t)} \xrightarrow{P} 0.$$

By stationarity, the random variables Y_0^j and $(T^j_{N_t^j} - t) \wedge Y^j_{N^j(t)}$ have the same distribution; they represent the remaining time after 0 and t, respectively, of the first off-period. Since both sums have the same distribution, we only consider the first one.

Using Karamata's Theorem (see [6]), the tail function $\bar{F}^{\mathrm{eq}}_{\mathrm{off}}$ satisfies

$$\bar{F}^{\mathrm{eq}}_{\mathrm{off}}(x) = \frac{1}{\mu_{\mathrm{off}}}\int_x^\infty \bar{F}_{\mathrm{off}}(s)\,ds \sim \frac{1}{\mu_{\mathrm{off}}} \frac{x^{-(\alpha_{\mathrm{off}}-1)}}{\alpha_{\mathrm{off}}-1} L_{\mathrm{off}}(x)$$

as $x \to \infty$. This implies that the random variable Y_0 has a regularly varying tail with index $-(\alpha_{\mathrm{off}}-1)$ and hence belongs to the domain of attraction of an $(\alpha_{\mathrm{off}}-1)$-stable distribution. Therefore there exists a slowly varying function L, such that

$$\frac{1}{m^{1/(\alpha_{\mathrm{off}}-1)} L(m)} \sum_{j=1}^m Y_0^j$$

converges in distribution to a stable law of index $\alpha_{\mathrm{off}} - 1$ (see [6]). Under scaling (ICR) with $\alpha_{\mathrm{on}} < \alpha_{\mathrm{off}}$, we have $a \gg m^{1/(\alpha_{\mathrm{off}}-1)} L(m)$ and hence

$$\frac{1}{a}\sum_{j=1}^m Y_0^j \quad \Longrightarrow \quad 0,$$

which is equivalent to the desired convergence in probability. □

Proof of Lemma 3 The proof of tightness given in [7] for fast scaling (FCR) can be adapted to our settings. We recall only the main lines. According to Billingsley [5], Theorem 12.3, it is enough to prove that for any t_1, t_2 with $|t_1 - t_2| \le 1$ and for some $\varepsilon > 0$, there exists a constant $C > 0$ and an $n_0 > 0$, such that for all $n \ge n_0$

$$\mathbb{E}\left[\frac{1}{a}\big|\big(W_m(at_2) - mat_2\mu_{\mathrm{on}}/\mu\big) - \big(W_m(at_1) - mat_1\mu_{\mathrm{on}}/\mu\big)\big|^2\right] \le C|t_2 - t_1|^{1+\varepsilon}.$$

Here and in the sequel, please recall that both $a = a_n$ and $m = m_n$ depend on n and tend to $+\infty$ as $n \to \infty$. Using the definition of W_m, centering and stationarity of increments, it is enough to prove that for all $t \in [0, 1]$ and $n \ge n_0$,

$$\frac{m}{a^2}\mathrm{Var}\big[W(at)\big] \le Ct^{1+\varepsilon} \tag{11}$$

(the constant C may change from one appearance to another). We shall prove this by using the following two estimates: On the one hand, according to [7], (7.1), we have as $t \to \infty$,

$$\mathrm{Var}\big[W(t)\big] \sim \sigma^2_{\alpha_{\mathrm{on}}}\frac{\mu^2_{\mathrm{on}}}{\mu^3}t^{3-\alpha_{\mathrm{on}}}L_{\mathrm{on}}(t). \tag{12}$$

On the other hand, since $|W(t)| \le t$ almost surely, the following global estimate holds for all $t \ge 0$,

$$\mathrm{Var}\big[W(t)\big] \le t^2. \tag{13}$$

The relation (12) and the scaling (ICR) together imply, as $n \to \infty$,

$$\frac{a^2}{m} \sim \frac{c^{1-\alpha_{\mathrm{on}}}}{\mu}a^{3-\alpha_{\mathrm{on}}}L_{\mathrm{on}}(a) \sim \frac{c^{1-\alpha_{\mathrm{on}}}}{\sigma^2_{\alpha_{\mathrm{on}}}}\frac{\mu^2}{\mu^2_{\mathrm{on}}}\mathrm{Var}\big[W(a)\big],$$

and so there is $C > 0$, such that for n large enough

$$\frac{m}{a^2}\mathrm{Var}\big[W(at)\big] \le C\frac{\mathrm{Var}[W(at)]}{\mathrm{Var}[W(a)]}.$$

By (12), the function $a \mapsto \mathrm{Var}[W(a)]$ is regularly varying with index $3 - \alpha_{\mathrm{on}}$. Then, using Potter bounds (see [6]), we conclude that there exist $a_0 > 0$ and $\varepsilon < 1 - \alpha_{\mathrm{on}}/2$, such that for all $t \in (0, 1)$ and $a \ge a_0/t$,

$$\frac{\mathrm{Var}[W(at)]}{\mathrm{Var}[W(a)]} \le \frac{1}{1-\varepsilon}t^{3-\alpha_{\mathrm{on}}-\varepsilon}$$

(see the proof of Theorem 4 in [7] for details). This implies that for all $t \in (0, 1)$ and all a such that $at \ge a_0$,

$$\frac{m}{a^2}\mathrm{Var}\big[W(at)\big] \le \frac{C}{1-\varepsilon}t^{3-\alpha_{\mathrm{on}}-\varepsilon} \le Ct^{1+\varepsilon}.$$

On the other hand, if $t \le a_0/a$, we use the estimate (13) and obtain, for a large enough,

$$\frac{m}{a^2}\mathrm{Var}\big[W(at)\big] \le \frac{Ca^2t^2}{\mathrm{Var}[W(a)]} \le C\frac{a^2t^2}{a^{3-\alpha_{\mathrm{on}}}L_{\mathrm{on}}(a)} \le C\frac{(at)^{1+\varepsilon}a_0^{1-\varepsilon}}{a^{3-\alpha_{\mathrm{on}}}L_{\mathrm{on}}(a)}$$

and so

$$\frac{m}{a^2}\mathrm{Var}\big[W(at)\big] \le C\frac{a_0^{1-\varepsilon}t^{1+\varepsilon}}{a^{2-\alpha_{\mathrm{on}}-\varepsilon}L_{\mathrm{on}}(a)} \le Ct^{1+\varepsilon}.$$

In the last inequality, we use the fact that $2-\alpha_{\mathrm{on}}-\varepsilon>0$ and so $a^{2-\alpha_{\mathrm{on}}-\varepsilon}L_{\mathrm{on}}(a)\to\infty$ as $a\to\infty$; taking a_0 large enough, we can suppose that for $a\ge a_0$, $a^{2-\alpha_{\mathrm{on}}-\varepsilon}L_{\mathrm{on}}(a)$ remains bounded away from zero.

By combining the estimates for the cases $at\ge a_0$ and $at\le a_0$ we obtain (11), which completes the proof. □

References

1. Asmussen, S.: Ruin Probabilities. World Scientific, Singapore (2000)
2. Beghin, L., Orsingher, E.: Fractional Poisson processes and related planar random motions. Electron. J. Probab. **14**, 1790–1826 (2009)
3. Beghin, L., Orsingher, E.: Poisson-type processes governed by fractional and higher-order recursive differential equations. Electron. J. Probab. **15**, 684–709 (2010)
4. Biermé, H., Estrade, A., Kaj, I.: Self-similar random fields and rescaled random balls models. J. Theor. Probab. **23**, 1110–1141 (2010)
5. Billingsley, P.: Convergence of Probability Measures, 2nd edn. Wiley, New York (1968)
6. Bingham, N.H., Goldie, C.M., Teugels, J.H.: Regular Variation. Cambridge University Press, Cambridge (1987)
7. Mikosch, T., Resnick, S., Rootzen, H., Stegeman, A.: Is network traffic approximated by stable Lévy motion or fractional Brownian motion. Ann. Appl. Probab. **12**(1), 23–68 (2002)
8. Mikosch, T., Samorodnitsky, G.: Scaling limits for cumulative input processes. Math. Oper. Res. **32**(4), 890–918 (2007)
9. Gaigalas, R.: A Poisson bridge between fractional Brownian motion and stable Lévy motion. Stoch. Process. Appl. **116**, 447–462 (2006)
10. Gaigalas, R., Kaj, I.: Convergence of scaled renewal processes and a packet arrival model. Bernoulli **9**(4), 671–703 (2003)
11. Jumarie, G.J.: Fractional master equation: non-standard analysis and Liouville–Rieman derivative. Chaos Solitons Fractals **12**, 2577–2587 (2001)
12. Kaj, I.: Stochastic Modeling in Broadband Communications Systems. SIAM Monographs on Mathematical Modeling and Computation, vol. 8. SIAM, Philadelphia (2002)
13. Kaj, I.: Limiting fractal random processes in heavy-tailed systems. In: Levy-Lehel, J., Lutton, E. (eds.) Fractals in Engineering, New Trends in Theory and Applications, pp. 199–218. Springer, London (2005)
14. Kaj, I., Taqqu, M.S.: Convergence to fractional Brownian motion and to the Telecom process: the integral representation approach. In: Vares, M.E., Sidoravicius, V. (eds.) An Out of Equilibrium 2. Progress in Probability, vol. 60, pp. 383–427. Birkhäuser, Basel (2008)
15. Lévy, J.B., Taqqu, M.S.: Renewal reward processes with heavy-tailed interrenewal times and heavy-tailed rewards. Bernoulli **6**, 23–44 (2000)
16. Mainardi, F., Gorenflo, R., Scalas, E.: A fractional generalization of the Poisson processes. Vietnam J. Math. **32**, 53–64 (2005)
17. Petrov, V.V.: Sums of Independent Random Variables. Springer, New York (1975)
18. Pipiras, V., Taqqu, M.S.: The limit of a renewal-reward process with heavy-tailed rewards is not a linear fractional stable motion. Bernoulli **6**, 607–614 (2000)

19. Pipiras, V., Taqqu, M.S., Lévy, L.B.: Slow, fast and arbitrary growth conditions for renewal reward processes when the renewals and the rewards are heavy-tailed. Bernoulli **10**, 121–163 (2004)
20. Resnick, S.I.: Heavy-Tail Phenomena, Probabilistic and Statistical modeling. Springer Series in Operations Research and Financial Engineering. Springer, New York (2007)
21. Taqqu, M.S., Willinger, W., Sherman, R.: Proof of a fundamental result in self-similar traffic modeling. Comput. Commun. Rev. **27**, 5–23 (1997)
22. Wang, X.-T., Wen, Z.-X.: Poisson fractional process. Chaos Solitons Fractals **18**, 169–177 (2003)
23. Wang, X.-T., Wen, Z.-X., Zhang, S.-Y.: Fractional Poisson processes (II). Chaos Solitons Fractals **28**, 143–147 (2006)
24. Wang, X.-T., Wen, Z.-X., Fan, S.: Nonhomogeneous fractional Poisson processes. Chaos Solitons Fractals **31**, 236–141 (2007)