Monografías del Seminario Matemático García de Galdeano **33**, 215–222 (2006)

A REFLECTED FBM LIMIT FOR ASYMPTOTICALLY BALANCED FLUID MODELS WITH HEAVY TAILED ON/OFF SOURCES

Rosario Delgado

Abstract. We consider a family of non-deterministic fluid models similar to that introduced by Harrison in [3] as the deterministic fluid analog for an open multiclass network, but with the difference that we suppose the process of external arrivals to be a non-deterministic aggregated cumulative packet process generated by a large enough number of heavy tailed ON/OFF sources, N. Scaling in time by a factor r and in state space conveniently, and letting N and r approach infinity (in this order) we prove that the scaled *immediate workload process* converges to a *reflected fractional Brownian motion (rfBm)* under heavy traffic.

Keywords: Reflected fractional Brownian motion, asymptotically balanced fluid model, heavy traffic, heavy tailed ON/OFF source, workload process.

AMS classification: 60K25, 60F05, 68F17, 60G15, 90B22.

§1. Introduction

The presence of long-range dependence in broadband network traffic and that of self-similarity in modern high-speed network traffic lead to the question of finding adequate traffic models for these situations. One simple physical explanation for this kind of phenomenon is the superposition of many ON/OFF sources with strictly alternating ON- and OFF-periods and whose ON- or OFF-periods have high variability. Taqqu et al. prove in [4, Theorem 1] the convergence of the aggregate cumulative packet process to the fBm (a self-similar and long-range dependent process). It is known (see [2]) that this convergence carries over to the stationary buffer content process: the scaled workload process converges to the fBm, reflected appropriately to be non-negative, for fluid models with only one server or station and only one fluid class, without reentering.

In this work we deal with the same question in a more general setting. Specifically, we consider a fluid model for a network with J stations and K fluid classes (with $K \geq J$), with a single server and an infinite buffer at each station, feedback and FIFO (first-in-first-out) discipline. We suppose (and this gives the difference with the model considered by Harrison) that the process of external arrivals is a non-deterministic aggregated cumulative packet process generated by a large enough number of heavy tailed ON/OFF sources. We prove, by following the methodology of Williams in [6], that after adequate scaling, the immediate workload process converges to a *reflected fractional Brownian process* W (rather than to a reflected Brownian model), by generalizing the result of [2].

The paper is organized as follows. In Section 2 some preliminares, notations and definitions are introduced. The fluid model that we consider in this work is introduced in Section 3. In Section 4 the main result of the paper is given. That result is a statement about the asymptotic behavior of the fluid model in which we give sufficient conditions for the convergence of the scaled immediate workload process, under asymptotic balancing (heavy-traffic), to a reflected fractional Brownian motion.

§2. Preliminares, notations and definitions

$\mathscr{D}-\lim_{n\to\infty} X^n = X$ denotes the *convergence in distribution* on $\mathscr{C}^d$ of stochastic d-dim. processes X^n to X and $\mathrm{P}-\lim_{n\to\infty} X^n = X$ its *convergence in probability* (unif. on compacts). We will use $\tilde{\lim}$ to denote the limit in the sense of the convergence of the finite-dimensional distributions. We now introduce a process known as *reflected fractional Brownian motion* (rfBm), that starts and behaves like a fractional brownian motion inside the positive orthant, but that is not allowed to exit it because of instantaneous "reflection" at the boundary given by faces. Its precise definition is as follows:

Definition 1. A *rfBm* on $S = \mathbb{R}_+^J$ associated with data $(x, H, \theta, \Gamma, R)$, where $x, \theta \in S$, $H \in (1/2, 1)$ and Γ and R are $J \times J$ matrices, being Γ a positive definite one, is a $J-$dimensional process $W = \{W(t), t \geq 0\}$ such that

(i) W has continuous paths and $W(t) \in S = \mathbb{R}_+^J$ for all $t \geq 0$, a.s.,

(ii) $W = X + RY$, a.s., with X and Y two $J-$dimensional processes defined on the same probability space as X, verifying:

(iii) X is a fBm with associated data (x, H, θ, Γ) (that is, it is a Gaussian process with $E\big(X(t)\big) = x + \theta\, t$ and $\mathrm{Cov}\big(X(t), X(s)\big) = \frac{1}{2}\left(t^{2H} + s^{2H} - |t-s|^{2H}\right)\Gamma$),

(iv) Y has continuous and non-decreasing paths, and for each j a.s., $Y_j(0) = 0$ and Y_j only increases when W is on face $F_j = \{y \in S = \mathbb{R}_+^J : y_j = 0\}$ (i.e. $\int_0^\infty 1_{\{W_j(s)>0\}}\, dY_j(s) = 0$).

A pair (W, Y) verifying (i), (ii) and (iv) is called a *solution of the R-regularization problem of X*. Proposition 4.2 of [5] shows that under condition (HR) on R (that implies the *completely-$\mathscr{S}$* condition) we have strong pathwise uniqueness of the solution of the R-regularization problem of X, being

(HR) **Assumption on matrix R:** R can be expressed as $I_J + \Theta$ with Θ a $J \times J$ matrix such that $|\Theta|$, that is the matrix obtained from Θ by replacing all its entries by their absolute value, has spectral radius less than 1.

§3. The fluid model

Let a network composed by J stations with a single server and an infinite buffer at each one, that processes continuous fluid. We distinguish among fluid of classes $1, \ldots, K$, with $K \geq J$. The many-to-one mapping for fluid classes to stations is described by the $J \times K$ *constituency matrix* C. For each k let $s(k)$ denote the station at which class k fluid is served. By following the ideas of [4], first of all suppose that there is only one external source of class k fluid that arrives to the network, and that the source can be ON or OFF. This source generates a stationary binary time series $\{U_k(t), t \geq 0\}$ where $U_k(t) = 1$ means that at time t the source is ON (and it is sending fluid to the network, at a traffic rate $\alpha_k > 0$), and $U_k(t) = 0$ means that it is OFF. Assume that, independently of k, the length of the ON-periods are i.i.d., those of the OFF-periods are i.i.d., and the lengths of ON- and OFF-periods are independent. The ON- and OFF-periods lengths may have different distributions. Denote the distribution functions, the mean values and the variances of the ON- and OFF-periods, respectively, by F_1 and F_2, $\tilde{\mu}_1$ and $\tilde{\mu}_2$, σ_1^2 and σ_2^2. Assume that as $x \to \infty$, $1 - F_1(x) \sim x^{-\beta_1} L_1(x)$ and $1 - F_2(x) \sim x^{-\beta_2} L_2(x)$ for some β_1 and β_2 in $(1, 2)$, and $L_1, L_2 > 0$ slowly varying functions at infinity.

Suppose now that for each class k fluid there are N i.i.d. sources, each one with its own binary time series $\{U_k^{(n)}(t), t \geq 0\}$, $n = 1, \ldots, N$, and that they are all independent. If all sources where ON, class k fluid would arrive at deterministic rate $\alpha_k^N > 0$, and the cumulative *external fluid traffic* up to time t would be deterministic and equal to $\alpha_k^N t$ (this is the case for the fluid model introduced by Harrison in [3]). Let

$$E_k^N(t) \stackrel{\text{def}}{=} \alpha_k^N \int_0^t \frac{1}{N} \left(\sum_{n=1}^{N} U_k^{(n)}(u) \right) du \tag{1}$$

be the cumulative external class k fluid generated up to time t (by the N sources). Let $\alpha^N = (\alpha_1^N, \ldots, \alpha_K^N)^T$. We assume that fluid at each station is processed in a firs-in-first-out (FIFO) basis and that our service discipline is *non-idling*.

Suppose that class k fluid is processed at a constant rate $\mu_k > 0$ (independent of N) if station $s(k)$ were never idle and the server devoted all of its attention to class k. Let $m_k = 1/\mu_k$ be the *mean service rate* for class k fluid, $m = (m_1, \ldots, m_K)^T$ and $\mu = (\mu_1, \ldots, \mu_K)^T$. Let $P_{k\ell}$ be the proportion of class k fluid that leaving station $s(k)$ goes next to station $s(\ell)$ as class ℓ fluid. We assume $P_{kk} = 0$ and that $P = (P_{k\ell})_{k,\ell=1}^K$ is a substochastic matrix with spectral radius less than one. Let $Q = (I_K - P^T)^{-1}$.

The following descriptive processes will be used to measure the performance of the queueing network:

The $K-$dimensional *fluid queue* Z^N, defined by: $Z_k^N(t)$ is the amount of class k fluid in queue or being processed at time t. The immediate *workload* W^N, a $J-$dimensional process defined by: $W_j^N(t)$ is the amount of time required for server j to complete processing of all fluids in queue (or being served) at station j at time t. Assume $W^N(0) = Z^N(0) = 0$. The $J-$dimensional *cumulative idletime process* Y^N, defined by: $Y_j^N(t)$ is the cumulative amount of time that server at station j has been idle in the time interval $[0, t]$, that is, $Y_j^N(t) = \int_0^t 1_{\{W_j^N(s)=0\}}\, ds$. Other auxiliar processes are the following: let A^N, defined by $A_k^N(t)$ is the total class k fluid arriving to station $s(k)$ up to time t, including both feedback flow and external input, and D^N defined by $D_k^N(t)$ is the total amount of class k fluid departing station $s(k)$ (both being routing to other station or leaving the network), up to time t. Assume $A^N(0) = D^N(0) = 0$. Let $F^N \stackrel{\text{def}}{=} P^T D^N$, and $L^N \stackrel{\text{def}}{=} CMA^N$, where $M \stackrel{\text{def}}{=} \text{diag}(m_1, \ldots, m_K)$.

These processes are related by the next *model equations*. For any $t \geq 0$, if $e = 1 \in \mathbb{R}^J$,

$$A^N(t) = E^N(t) + F^N(t),\ W^N(t) = L^N(t) - et + Y^N(t),\ Z^N(t) = A^N(t) - D^N(t),$$
$$D^N(t + C^T W^N(t)) = A^N(t),\ W^N(t) = CMZ^N(t).$$

We now consider a double sequence of fluid models having the same basic structure as described before. The associated processes and parameters are denoted by append r and N, where N is the number of sources, that tends to infinity, and r is the factor of scaling in time and tends to infinity through a strictly increasing sequence of strictly positive real numbers. Suppose that P, $\mu = (\mu_1, \ldots, \mu_K)^T$ and $m = (m_1, \ldots, m_K)^T$ are allowed to change with r but not with N. We use the notation P^r, μ^r, m^r, M^r and Q^r. We define $\lambda^{r,N}$ to be the unique $K-$dimensional vector solution to the *traffic equation*

$$\lambda^{r,N} = \alpha^{r,N} \frac{\tilde{\mu}_1}{\tilde{\mu}_1 + \tilde{\mu}_2} + (P^r)^T \lambda^{r,N}, \tag{2}$$

$\lambda_k^{r,N}$ is interpreted as the *class k fluid arrival rate due both to external and internal flow traffic*. We also define the *fluid traffic intensity* for station j as

$$\rho_j^{r,N} \stackrel{\text{def}}{=} \sum_{k \text{ served at station } j} \lambda_k^{r,N} m_k^r, \text{ for all } j \quad \big(\text{in matricial form, } \rho^{r,N} = CM^r \lambda^{r,N}\big).$$

In order to define the *scaled processes* associated to the fluid model, we must introduce previously some notation, by following [4]. For any $j = 1, 2$, set $a_j = \Gamma(2 - \beta_j)/(\beta_j - 1)$ and $b = \lim_{t \to \infty} t^{\beta_2 - \beta_1} L_1(t)/L_2(t)$. If $0 < b < \infty$ set $\beta_{min} = \beta_1 = \beta_2$, $L = L_2$ and we define

$$\sigma_{lim}^2 = \frac{2\left(\tilde{\mu}_2^2 a_1 b + \tilde{\mu}_1^2 a_2\right)}{\left(\tilde{\mu}_1 + \tilde{\mu}_2\right)^3 \Gamma(4 - \beta_{min})}.$$

If $b = 0$ or $b = \infty$, set $L = L_{min}$, where *min* is the index 1 if $b = \infty$ and is the index 2 if $b = 0$; *max* denoting the other index, and

$$\sigma_{lim}^2 = \frac{2\,\tilde{\mu}_{max}^2 a_{min}}{\left(\tilde{\mu}_1 + \tilde{\mu}_2\right)^3 \Gamma(4 - \beta_{min})}.$$

In either case, $\beta_{min} \in (1, 2)$.

Let we define $H \stackrel{\text{def}}{=} (3 - \beta_{min})/2$. We have that $H \in (1/2, 1)$. Now we can define the *scaled processes,* denoted with a hat:

$$\hat{E}^{r,N}(t) \stackrel{\text{def}}{=} \sqrt{N}\,\frac{E^{r,N}(rt) - \alpha^{r,N} rt\,\frac{\tilde{\mu}_1}{\tilde{\mu}_1 + \tilde{\mu}_2}}{r^H L^{1/2}(r)}, \quad \hat{A}^{r,N}(t) \stackrel{\text{def}}{=} \sqrt{N}\,\frac{A^{r,N}(rt) - \lambda^{r,N} rt}{r^H L^{1/2}(r)},$$
$$\hat{D}^{r,N}(t) \stackrel{\text{def}}{=} \sqrt{N}\,\frac{D^{r,N}(rt) - \lambda^{r,N} rt}{r^H L^{1/2}(r)}, \quad \hat{F}^{r,N}(t) \stackrel{\text{def}}{=} \sqrt{N}\,\frac{F^{r,N}(rt) - (P^r)^T \lambda^{r,N} rt}{r^H L^{1/2}(r)},$$
$$\hat{W}^{r,N}(t) \stackrel{\text{def}}{=} \sqrt{N}\,\frac{W^{r,N}(rt)}{r^H L^{1/2}(r)}, \quad \hat{Z}^{r,N}(t) \stackrel{\text{def}}{=} \sqrt{N}\,\frac{Z^{r,N}(rt)}{r^H L^{1/2}(r)}$$
$$\hat{Y}^{r,N}(t) \stackrel{\text{def}}{=} \sqrt{N}\,\frac{Y^{r,N}(rt)}{r^H L^{1/2}(r)}, \quad \hat{L}^{r,N}(t) \stackrel{\text{def}}{=} \sqrt{N}\,\frac{L^{r,N}(rt) - \rho^{r,N} rt}{r^H L^{1/2}(r)}$$

Scaled equations will be used to determine the limit behavior of the normalized immediate workload $\hat{W}^{r,N}$ process. These equations are obtained by substituting scaled processes into model equations:

$$\hat{A}^{r,N}(t) = \hat{E}^{r,N}(t) + \hat{F}^{r,N}(t) \tag{3}$$

$$\hat{L}^{r,N}(t) = CM^r \hat{A}^{r,N}(t) \tag{4}$$

$$\hat{W}^{r,N}(t) = \hat{L}^{r,N}(t) + \hat{Y}^{r,N}(t) + \hat{\gamma}^{r,N} t\,, \text{ where } \hat{\gamma}^{r,N} \stackrel{\text{def}}{=} \sqrt{N}\, \frac{(\rho^{r,N} - e)\, r}{r^H L^{1/2}(r)} \tag{5}$$

$$\hat{W}^{r,N}(t) = CM^r \hat{Z}^{r,N}(t) \tag{6}$$

$$\hat{Z}^{r,N}(t) = \hat{A}^{r,N}(t) - \hat{D}^{r,N}(t) \tag{7}$$

$$\hat{F}^{r,N}(t) = (P^r)^T \hat{D}^{r,N}(t) \tag{8}$$

Now we reduce the system of scaled equations as follows: by substituting $\hat{D}^{r,N}$ from (7) into (8), and in turn substituting this expression into (3) we obtain

$$\hat{A}^{r,N}(t) = Q^r \Big(\hat{E}^{r,N}(t) - (P^r)^T \hat{Z}^{r,N}(t)\Big). \tag{9}$$

Substituting (9) into (4) and the resulting into (5) yields

$$\hat{W}^{r,N}(t) = \hat{\xi}^{r,N}(t) - CM^r Q^r (P^r)^T \hat{Z}^{r,N}(t) + \hat{Y}^{r,N}(t), \tag{10}$$

where

$$\hat{\xi}^{r,N}(t) \stackrel{\text{def}}{=} CM^r Q^r \hat{E}^{r,N}(t) + \hat{\gamma}^{r,N} t. \tag{11}$$

We now introduce, like in [6], a $K \times J$ matrix $\Delta^{r,N}$ by $\Delta^{r,N}_{kj} = \lambda^{r,N}_k / \rho^{r,N}_j$ if $s(k) = j$ and 0 otherwise. We need to impose an assumption on $\Delta^{r,N}$:

(H$\Delta^{r,N}$) **Assumption on matrix $\Delta^{r,N}$:** (satisfied if $K = J$)

$$CM^r Q^r \Delta^{r,N} \quad \text{is invertible for all } N \text{ and } r \text{ (big enough)}. \tag{12}$$

By using matrix $\Delta^{r,N}$ we can define processes $\hat{\varepsilon}^{r,N}$ and $\hat{\eta}^{r,N}$ by

$$\hat{\varepsilon}^{r,N}(t) \stackrel{\text{def}}{=} \hat{Z}^{r,N}(t) - \Delta^{r,N} \hat{W}^{r,N}(t), \quad \hat{\eta}^{r,N}(t) \stackrel{\text{def}}{=} -CM^r Q^r (P^r)^T \hat{\varepsilon}^{r,N}(t), \tag{13}$$

and a $J \times J$ matrix $G^{r,N} \stackrel{\text{def}}{=} CM^r Q^r (P^r)^T \Delta^{r,N}$. Then, $I_J + G^{r,N} = CM^r Q^r \Delta^{r,N}$ and we can rewrite (10) as

$$\hat{W}^{r,N}(t) = \hat{\xi}^{r,N}(t) - G^{r,N} \hat{W}^{r,N}(t) + \hat{\eta}^{r,N}(t) + \hat{Y}^{r,N}(t), \text{ and with } R^{r,N} = (I_J + G^{r,N})^{-1},$$

$$\hat{W}^{r,N}(t) = R^{r,N}\Big(\hat{\xi}^{r,N}(t) + \hat{\eta}^{r,N}(t) + \hat{Y}^{r,N}(t)\Big) = \hat{X}^{r,N}(t) + R^{r,N} \hat{Y}^{r,N}(t), \tag{14}$$

$$\text{with } \hat{X}^{r,N}(t) \stackrel{\text{def}}{=} R^{r,N}\Big(\hat{\xi}^{r,N}(t) + \hat{\eta}^{r,N}(t)\Big). \tag{15}$$

§4. The main result

Before to state the main result of this work, we need to introduce some assumptions:

(HαmP) **Assumptions on $\alpha^{r,N}$, m^r and P^r:**
We will assume that there are two $K-$dimensional vectors $\alpha > 0$ and $m > 0$, and a $K \times K$ matrix P, substochastic and with spectral radius strictly less than one, which satisfy: for any r (big enough), $\exists \alpha^r = \lim_{N\to\infty} \alpha^{r,N}$, $\exists \alpha = \lim_{r\to\infty} \alpha^r$, $\exists m = \lim_{r\to\infty} m^r$, and $\exists P = \lim_{r\to\infty} P^r$. Therefore, there exists $\Delta^r = \lim_{N\to\infty} \Delta^{r,N}$, $\Delta = \lim_{r\to\infty} \Delta^r$, $Q = \lim_{r\to\infty} Q^r \big(= (I_K - P^T)^{-1}\big)$, $G^r = \lim_{N\to\infty} G^{r,N}$, $G = \lim_{r\to\infty} G^r$ and $M = \lim_{r\to\infty} M^r$.

(H$\Delta^r\Delta$) **Assumption on matrices Δ^r and Δ:** (satisfied if $K = J$)
Matrices $C\,M^r\,Q^r\,\Delta^r$, for r big enough, and $C\,M\,Q\,\Delta$ are invertible (therefore, matrices $R^r \stackrel{\text{def}}{=} \big(I_J + G^r\big)^{-1}$ and $R \stackrel{\text{def}}{=} \big(I_J + G\big)^{-1}$ are well defined).

(AB) **Asymptotically balanced (heavy-traffic) fluid model assumption:** there exists $\gamma \in \mathbb{R}^J$ such that

$$\lim_{r\to\infty}\left(\lim_{N\to\infty} \sqrt{N}\, \frac{(\rho^{r,N} - e)\, r}{r^H L^{1/2}(r)}\right) \Big(= \lim_{r\to\infty}\big(\lim_{N\to\infty} \hat{\gamma}^{r,N}\big)\Big) = \gamma.$$

The final assumption that we consider is a form of *state space collapse*, which relates scaled immediate workload and fluid queue length processes.

(SSC) **Assumption of state space collapse:** (satisfied if $K = J$)
For any r (big enough), there exists $\hat{\hat{\varepsilon}}^r = \mathrm{P}-\lim_{N\to\infty} \hat{\varepsilon}^{r,N} \in \mathbb{R}^K$, and $\lim_{r\to\infty} \hat{\hat{\varepsilon}}^r = 0$.

The main result uses the following lemma that lean on Theorem 1 of [4].

Lemma 1. *In our setting, under (HαmP) and (SSC), for any r (big enough), there exists*

$$(\hat{\hat{E}}^r, \hat{\hat{\varepsilon}}^r) = \widetilde{\lim_{N\to\infty}}\, (\hat{E}^{r,N}, \hat{\varepsilon}^{r,N}) \quad \textit{and} \quad \mathscr{D}-\lim_{r\to\infty} (\hat{\hat{E}}^r, \hat{\hat{\varepsilon}}^r) = (B_H, 0),$$

where B_H is K-dimensional fBm with associated data $(0,\, \frac{3-\beta_{min}}{2},\, 0,\, \sigma^2_{lim}\, \mathrm{diag}(\alpha)^2)$.

Remark 1. In Theorem 2 below we will impose condition (HR) for matrices R^r (for r big enough) and R. We note that if $K = J$ this assumption is trivially accomplished by R^r, and also by R if P has spectral radius less that one.

Theorem 2. *In our setting, assume (H$\Delta^{r,N}$), (HαmP), (H$\Delta^r\Delta$), (AB), (HR) for matrices R^r, for r big enough, and R, and (SSC). Then, for any r (big enough) there exist $\hat{\hat{W}}^r = \widetilde{\lim_{N\to\infty}}\, \hat{W}^{r,N}$, $\hat{\hat{X}}^r = \widetilde{\lim_{N\to\infty}}\, \hat{X}^{r,N}$, $\hat{\hat{Y}}^r = \widetilde{\lim_{N\to\infty}}\, \hat{Y}^{r,N}$, there also exist $W = \mathscr{D}-\lim_{r\to\infty} \hat{\hat{W}}^r$, $X = \mathscr{D}-\lim_{r\to\infty} \hat{\hat{X}}^r$, $Y = \mathscr{D}-\lim_{r\to\infty} \hat{\hat{Y}}^r$, $W = X + RY$, and it is a rfBm on $S = \mathbb{R}^J_+$ with associated data $(0, H = \frac{3-\beta_{min}}{2}, R\gamma, \Gamma, R)$, where*

$$\Gamma = \sigma^2_{lim}\, R\,C\,M\,Q\, \mathrm{diag}(\alpha)^2\, Q^T\, M\,C^T\, R^T.$$

Proof. We recall that $\hat{W}^{r,N}(t) = \hat{X}^{r,N}(t) + R^{r,N}\hat{Y}^{r,N}(t)$, with

$$\hat{X}^{r,N}(t) = R^{r,N} C M^r Q^r \left(\hat{E}^{r,N}(t) - (P^r)^T \hat{\varepsilon}^{r,N}(t)\right) + R^{r,N}\hat{\gamma}^{r,N} t,$$

by (14), (15), (11) and (13).

By Lemma 1 and the Continuous Mapping Theorem, we obtain the existence of $\hat{\hat{X}}^r = \tilde{\lim_{N\to\infty}} \hat{X}^{r,N}$, $X = \mathscr{D}-\lim_{r\to\infty} \hat{\hat{X}}^r$, and that $X(t) = RCMQB_H(t) + R\gamma t$, that is a J-dim. fBm with associated data $(0, H = \frac{3-\beta_{min}}{2}, R\gamma, \Gamma)$, $\Gamma = \sigma_{lim}^2 RCMQ\,\mathrm{diag}(\alpha)^2 Q^T MC^T R^T$.

By the other way, we can write

$$\hat{W}^{r,N}(t) = \left(\hat{X}^{r,N}(t) + \left(R^{r,N} - R^r\right)\hat{Y}^{r,N}(t)\right) + R^r\hat{Y}^{r,N}(t). \tag{16}$$

Since R^r is *completely-$\mathscr{S}$* by (HR), we can apply the *oscillation inequality* given in Lemma 1 of [1] to obtain that there is a constant $C_{R^r} > 0$, that only depends on R^r, such that for any $T \geq 0$, if define $\mathrm{Osc}\big(\omega(\cdot), [0, T]\big)$ as $\sup_{0\leq s<t\leq T} |\omega(t) - \omega(s)|$, we have that

$$\mathrm{Osc}\Big(\hat{Y}^{r,N}(\cdot), [0, T]\Big) \leq C_{R^r}\, \mathrm{Osc}\Big(\hat{X}^{r,N}(\cdot) + (R^{r,N} - R^r)\hat{Y}^{r,N}(\cdot), [0, T]\Big). \tag{17}$$

By hypothesis, $R^{r,N}$ converges to R^r as $N \to \infty$. Consequently, there exists N_0 such that for any $N \geq N_0$, $C_{R^r}\,|R^{r,N} - R^r| < 1/2$. Thus, (17) implies that for any r (big enough), if $N \geq N_0$,

$$\left\|\hat{Y}^{r,N}(\cdot)\right\|_T = \mathrm{Osc}\Big(\hat{Y}^{r,N}(\cdot), [0, T]\Big) \leq 2\, C_{R^r}\, \mathrm{Osc}\Big(\hat{X}^{r,N}(\cdot), [0, T]\Big) \leq 4C_{R^r}\left\|\hat{X}^{r,N}(\cdot)\right\|_T \tag{18}$$

Due to the continuity of $\hat{\hat{E}}^r$ (that implies the continuity of $\hat{\hat{X}}^r$), we have that, for any $T \geq 0$, for any r (big enough), for any $\varepsilon > 0$, there exist $K_\varepsilon > 0$ and N_1 such that, if $N \geq N_1$, $P\big(\left\|\hat{X}^{r,N}(\cdot)\right\|_T \leq K_\varepsilon/(4C_{R^r})\big) \geq 1 - \varepsilon$. By using this fact, (18) gives that if $N \geq N_1 \vee N_0$, $P\left(\left\|\hat{Y}^{r,N}(\cdot)\right\|_T \leq K_\varepsilon\right) \geq 1 - \varepsilon$. Furthermore, there exists N_2 such that for any $N \geq N_2$, $|R^{r,N} - R^r| < \varepsilon/K_\varepsilon$, and as a result, for any $N \geq \max\{N_0, N_1, N_2\}$,

$$P\big(|(R^{r,N} - R^r)|\,\left\|\hat{Y}^{r,N}(\cdot)\right\|_T \geq \varepsilon\big) \leq \varepsilon,$$

that is, $\mathrm{P}-\lim_{N\to\infty} (R^{r,N} - R^r)\hat{Y}^{r,N} = 0$.

Let we define $\hat{\Omega}^{r,N}$ as $\hat{W}^{r,N} - (R^{r,N} - R^r)\hat{Y}^{r,N}$. Then, if they exist, we have that $\tilde{\lim_{N\to\infty}} \hat{\Omega}^{r,N} = \tilde{\lim_{N\to\infty}} \hat{W}^{r,N}$, and by (16) we also have that $\hat{\Omega}^{r,N} = \hat{X}^{r,N} + R^r\hat{Y}^{r,N}$, with R^r verifying (HR) for any r (big enough), and $\hat{\hat{X}}^r = \tilde{\lim_{N\to\infty}} \hat{X}^{r,N}$ having continuous paths. Then, there exists a unique strong pathwise solution of the R^r-regularization problem of $\hat{\hat{X}}^r$, that coincides with $\big(\tilde{\lim_{N\to\infty}} \hat{\Omega}^{r,N}, \tilde{\lim_{N\to\infty}} \hat{Y}^{r,N}\big)$. If we denote $\tilde{\lim_{N\to\infty}} \hat{Y}^{r,N}$ by $\hat{\hat{Y}}^r$ and $\tilde{\lim_{N\to\infty}} \hat{\Omega}^{r,N} = \tilde{\lim_{N\to\infty}} \hat{W}^{r,N}$ by $\hat{\hat{W}}^r$, we have that the unique solution of the R^r-regularization problem of $\hat{\hat{X}}$ is $(\hat{\hat{W}}^r, \hat{\hat{Y}}^r)$, and then $\hat{\hat{W}}^r = \hat{\hat{X}}^r + R^r\hat{\hat{Y}}^r$. This fact implies that $\hat{\hat{W}}^r$, $\hat{\hat{X}}^r$ and $\hat{\hat{Y}}^r$ satisfy hypothesis of the *invariant principle* of Theorem 4.1 in [5] with matrix R^r, taking into account that $\mathscr{D}-\lim_{r\to\infty} \hat{\hat{X}}^r = X$,

$\lim_{r\to\infty} R^r = R$, and R is a *Completely-$\mathscr{S}$* matrix, by (HR). We have that $\left\{\left(\hat{\hat{W}}^r, \hat{\hat{X}}^r, \hat{\hat{Y}}^r\right)\right\}_r$ inherits tightness from sequence $\left\{\hat{\hat{X}}^r\right\}_r$, and consequently, by (HR) (see Corollary 4.3 of [5]), we obtain that there exists $\mathscr{D}-\lim_{r\to\infty}\left(\hat{\hat{W}}^r, \hat{\hat{X}}^r, \hat{\hat{Y}}^r\right) = (W,X,Y)$, where $W = X + RY$ and conditions of Definition 1 are satisfied. Therefore, W is a rfBm on $S = \mathbb{R}^J_+$ with associated data $(0, H = \frac{3-\beta_{min}}{2}, R\gamma, \Gamma, R)$. $\square$

Acknowledgements

Partially supported by projects BFM2003-00261, MTM2004-01175 and the research group 2001SGR-00174 (CIRIT))

References

[1] BERNARD, A., AND EL KHARROUBI, A. Régulations déterministes et stochastiques dans le premier orthant de $\mathbb{R}^n$. *Stochastics and Stochastics Reports 34* (1991), 149–167.

[2] DEBICKI, K., AND MANDJES, M. Traffic with an fBm limit: Convergence of the Stationary Workload process. *Queueing Systems 46* (2004), 113–127.

[3] HARRISON, J. M. Balanced fluid models of multiclass queueing networks: a heavy traffic conjecture. In *Stochastic Networks. IMA Volumes in Mathematics and its Applications, 71*, Frank P. Kelly and Ruth J. Williams, Eds. Springer-Verlag, New-York, 1995, pp. 1–20.

[4] TAQQU, M. S., WILLINGER, W., AND SHERMAN, R. Proof of a Fundamental Result in Self-Similar Traffic Modeling. *Comput. Commun. Rev. 27* (1997), 5–23.

[5] WILLIAMS, R. J. An invariance principle for semimartingale reflecting Brownian motions in an orthant. *Queueing Systems 30* (1998), 5–25.

[6] WILLIAMS, R. J. Diffusion approximations for open multiclass queueing networks: sufficient conditions involving state space collapse. *Queueing Systems 30* (1998), 27–88.

Rosario Delgado
Departament de Matemàtiques. Universitat Autònoma de Barcelona
Edifici C- Campus de la UAB
08193 Bellaterra (Cerdanyola del Vallès)-Barcelona (Spain)
delgado@mat.uab.es