Practical Time-scale Fitting of Self-similar Traffic with Markov-modulated Poisson Process

Tadafumi Yoshihara†, Shoji Kasahara†‡ and Yutaka Takahashi†‡

†Graduate School of Information Science, Nara Institute of Science and Technology
8916-5 Takayama, Ikoma, Nara 630-01, Japan
‡Telecommunications Advancement Organization of Japan(TAO)
Shiba 2-31-19, Minato-Ku, Tokyo 105, Japan

Abstract

Recent measurements of packet/cell streams in multimedia communication networks have revealed that they have the self-similar property and are of different character istics from traditional traffic streams. Therefore, a number of studies of modeling the self-similar traffic have been performed. In this paper, we first give some definitions of self-similarity. Then, we propose a fitting method for the self-similar traffic in terms of Markov-modulated Poisson process (MMPP). We construct an MMPP as the superposition of two-state MMPPs and fit it so as to match the variance function over several time-scales. Numerical examples show that the variance function of the self-similar process can be well represented by that of four-state MMPP. We also examine the queueing behavior of the resulting MMPP/D/1 queueing systems. We compare the results with the simulation for the queueing systems with the self-similar process as an input.

1 Introduction

Recently a number of high-quality, high-resolution measurements of multimedia traffic in high-speed networks such as packet streams in local area networks (LAN), cell streams from variable bit rate (VBR) video streams in ATM networks, etc., have been carried out and analyzed. They have shown that the traffic from those networks appears to be *self-similar*. The self-similar traffic is characterized by that the correlation never vanishes in a large time-scale. Its traffic looks the same regardless of time-scales over a long range interval. This fractal behavior makes the traffic very bursty. These properties of the self-similar traffic are quite different from those of traditional traffic models such as Poisson process, Markovian arrival process (MAP), etc.

The above observation has initiated studies of new models such as chaotic maps [4], fractional Brownian motion (FBM) [5] and ARIMA [6]. They can describe the self-similar behavior in a relatively simple manner. However, queueing theoretical techniques developed in the past are hardly applicable for these models.

On the other hand, a number of models based on traditional traffic models have been proposed. One approach is to emulate self-similarity over a certain range of time-scales with finite state Markovian models. [1] propose a model based on Markov-modulated Poisson process (MMPP) as a superposition of two state Markov processes. In [12], a Markov chain emulating self-similarity which is quite easy to manipulate and depends on only two parameters has been analyzed. Another approach is to model self-similarity through superposition of infinite Markovian sources. In [10], they have constructed a self-similar process from an infinity of on-off sources with Pareto service demands.

In this paper, we first give some definitions of self-similarity which are equivalent to those in [2]. Then, using them, we propose a fitting method for self-similar traffic in terms of MMPP. Our fitting method is based on the model by Anderson et al. [1], where traffic is modeled by the superposition of several two state MMPPs. In [1], the parameters of MMPP are determined so as to match the autocorrelation function which is approximately evaluated. In our method, however, the parameters are determined so as to match the variance of the measured traffic which is exactly evaluated.

The paper is organized as follows. In section 2, we summarize some important characteristics of MMPP. In section 3, we overview the concept of self-similarity and give some definitions. In section 4, we explain the idea of superposing two-state MMPPs to model self-similar traffic. In section 5, a fitting procedure of MMPP is given. In section 6, we consider the condition on the preliminarily required parameters for the fitting. In section 7, a number of numerical results are shown in order to verify the usefulness of our fitting method. Finally, some conclusions are drawn in the last section.

2 MMPP

In this section, we summarize some main characteristics of MMPP. MMPP is a doubly stochastic Poisson process. In the case of m-state MMPP, its arrival rate is determined by the state of a continuous-time Markov chain with infinitesimal generator Q and Poisson arrival rates $\lambda_i \ (1 \le i \le m)$. That is, the arrival rate is λ_i when the Markov chain is in state i. Matrix Λ which describes Poisson arrival rates is called the arrival rate matrix. In the two-state case, Q and Λ are given by

$$Q = \begin{bmatrix} -\sigma_1 & \sigma_1 \\ \sigma_2 & -\sigma_2 \end{bmatrix}, \quad \Lambda = \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix}.$$

In the following, we derive the mean and the variance of the number of arrivals in MMPP. Although we consider the case of two-state MMPP, the results presented below also apply to the general case. Let N_t be the number of arrivals in $(0, t]$ and J_t be the state of the Markov chain at time t. We consider a matrix $P(n,t)$ whose (i,j)-th element is defined as

$$P_{ij}(n,t) = \Pr\{N_t = n, J_t = j | N_0 = 0, J_0 = i\}, \qquad 1 \le i,j \le 2.$$

The matrices $P(n,t)$ satisfy the following forward Chapman-Kolmogorov equations

$$\left\{\begin{array}{rcl} \dfrac{d}{dt}P(n,t) & = & P(n,t)(Q-\Lambda)+P(n-1,t)\Lambda, \quad n\geq 1,\ t\geq 0, \\ P(0,0) & = & I. \end{array}\right. \tag{1}$$

Multiplying (1) by z^n and summing over $n=0,1,\ldots$, we obtain

$$\left\{\begin{array}{rcl} \dfrac{d}{dt}P^*(z,t) & = & P^*(z,t)(Q-\Lambda)+zP^*(z,t)\Lambda, \\ P^*(z,0) & = & I, \end{array}\right. \tag{2}$$

where $P^*(z,t)$ is the generating function of $P(n,t)$. Solving (2) for $P^*(z,t)$, we obtain

$$P^*(z,t)=\exp\{[Q+(z-1)\Lambda]t\}.$$

For the time-stationary MMPP, the mean of N_t is given by

$$\begin{aligned} \mathrm{E}(N_t) &= \boldsymbol{\pi}\left.\frac{\partial P^*(z,t)}{\partial z}\right|_{z=1}\boldsymbol{e} \\ &= \boldsymbol{\pi}\Lambda\boldsymbol{e}t \\ &= \frac{\sigma_2\lambda_1+\sigma_1\lambda_2}{\sigma_1+\sigma_2}t, \end{aligned}$$

where $\boldsymbol{e}=(1,1)$ and $\boldsymbol{\pi}$ is the steady state vector of the Markov chain such that

$$\boldsymbol{\pi}Q=0, \quad \boldsymbol{\pi}\boldsymbol{e}=1.$$

The variance of N_t is given by

$$\mathrm{Var}(N_t)=\boldsymbol{\pi}\left.\frac{\partial^2 P^*(z,t)}{\partial z^2}\right|_{z=1}\boldsymbol{e}+\mathrm{E}(N_t)-\{\mathrm{E}(N_t)\}^2. \tag{3}$$

To evaluate the first term of (3) we use the following equation

$$\left.\frac{\partial^2 P^*(z,t)}{\partial z^2}\right|_{z=1}=\boldsymbol{L}^{-1}\left[\left.\frac{\partial^2\tilde{P}(z,s)}{\partial z^2}\right|_{z=1}\right],$$

where $\tilde{P}(z,s)$ is the Laplace transform of $P^*(z,t)$ and $\boldsymbol{L}^{-1}$ denotes the inverse Laplace transform. $\tilde{P}(z,s)$ is given by

$$\tilde{P}(z,s)=[sI-Q-(z-1)\Lambda]^{-1}.$$

The second derivative of $\tilde{P}(z,s)$ at $z=1$ is given by

$$\left.\frac{\partial^2\tilde{P}(z,s)}{\partial z^2}\right|_{z=1}=2[sI-Q]^{-1}\{\Lambda[sI-Q]^{-1}\}^2.$$

Since $\boldsymbol{\pi}[sI-Q]=s\boldsymbol{\pi}$, we have $\boldsymbol{\pi}/s=\boldsymbol{\pi}[sI-Q]^{-1}$. Similarly, since $[sI-Q]\boldsymbol{e}=s\boldsymbol{e}$, we have $\boldsymbol{e}/s=[sI-Q]^{-1}\boldsymbol{e}$. Using these, we obtain

$$\begin{aligned} \boldsymbol{\pi}\left.\frac{\partial^2\tilde{P}(z,s)}{\partial z^2}\right|_{z=1}\boldsymbol{e} &= \frac{2}{s^2}\boldsymbol{\pi}\Lambda(sI-Q)^{-1}\Lambda\boldsymbol{e} \\ &= \frac{2}{\sigma_1+\sigma_2}\left[\frac{A_1}{s+\sigma_1+\sigma_2}+\frac{A_2}{s^3}+\frac{A_3}{s^2}+\frac{A_4}{s}\right], \end{aligned} \tag{4}$$

where

$$\begin{aligned}
A_1 &= \frac{\sigma_1\sigma_2(\lambda_1-\lambda_2)^2}{(\sigma_1+\sigma_2)^3},\\
A_2 &= \frac{(\sigma_1\lambda_2+\sigma_2\lambda_1)^2}{\sigma_1+\sigma_2},\\
A_3 &= (\sigma_1+\sigma_2)A_1,\\
A_4 &= -A_1.
\end{aligned}$$

Inverting (4) yields

$$\boldsymbol{\pi}\left.\frac{\partial^2 P^*(z,t)}{\partial z^2}\right|_{z=1}\boldsymbol{e} = \frac{2}{\sigma_1+\sigma_2}\left[A_1 e^{-(\sigma_1+\sigma_2)t} + \frac{A_2}{2}t^2 + A_3 t + A_4\right].$$

Thus, we obtain

$$\mathrm{Var}(N_t) = \frac{\sigma_2\lambda_1+\sigma_1\lambda_2}{\sigma_1+\sigma_2}t + 2A_1 t - \frac{2A_1}{\sigma_1+\sigma_2}(1-e^{-(\sigma_1+\sigma_2)t}). \tag{5}$$

3 Self-similarity

In this section, we overview the concept of self-similarity of the stochastic process. First, we summarize Cox's definitions of self-similarity [2] and then, we show the equivalent definitions to those of Cox.

3.1 Cox's Definitions of Self-similarity

We consider a second-order stationary process $X = \{X_t : t = 1, 2, ...\}$ with the variance σ^2 and the autocorrelation function $r(k)$, where $r(k)$ is given as

$$r(k) = \frac{\mathrm{Cov}(X_t, X_{t+k})}{\mathrm{Var}(X_t)}.$$

In the context of the packet traffic, X_t corresponds to the number of packets that arrive during the t-th time slot. We also consider a new sequence of $X_t^{(m)}$ which is obtained by averaging the original sequence in non-overlapping blocks. That is,

$$X_t^{(m)} = \frac{1}{m}\sum_{i=1}^{m} X_{(t-1)m+i}, \quad t = 1, 2, \ldots\ .$$

The new sequence is also a second-order stationary process with the autocorrelation function $r^{(m)}(k)$.

Let δ^2 denote the central second difference operator defined by that for any function $f(x)$,

$$\delta^2(f(x)) = \{f(x+1) - f(x)\} - \{f(x) - f(x-1)\}.$$

Then, definitions of self-similar process are given by the following:

Definition 3.1 *X is called exactly second-order self-similar with the Hurst parameter $H = 1 - \beta/2$ if*

$$r(k) = \frac{1}{2}\delta^2(k^{2-\beta}). \tag{6}$$

Definition 3.2 *X is called asymptotically second-order self-similar with the Hurst parameter $H = 1 - \beta/2$ if*

$$r^{(m)}(k) \to \frac{1}{2}\delta^2(k^{2-\beta}), \quad \text{as } m \to \infty. \tag{7}$$

Note that (6) implies that for all $m = 1, 2, \ldots$,

$$r^{(m)}(k) = r(k). \tag{8}$$

We show this in the next subsection. We are interested in the range $0.5 < H < 1$ because the process has the long-range dependence in this range. In the case that $H = 0.5$, X is a second-order pure noise with $\mathrm{Var}(X^{(m)}) = \mathrm{Var}(X)/m$.

3.2 Equivalent Definitions of Self-Similarity

In this subsection, we give equivalent definitions to Definition 3.1 and 3.2. For the case of exactly second-order self-similarity, what we discuss in the following is shown in [14].

Theorem 3.1 *X satisfies (6) if and only if for all $m = 1, 2, \ldots$,*

$$Var(X^{(m)}) = \sigma^2 m^{-\beta}. \tag{9}$$

Before we prove the theorem, note that $\mathrm{Var}(X_t)$ and $\mathrm{Cov}(X_t, X_{t+k})$ have the following relations.

$$\begin{aligned} \mathrm{Var}(X^{(m)}) &= \frac{\sigma^2}{m} + \frac{2}{m^2}\sum_{k=1}^{m}(m-k)\mathrm{Cov}(X_t, X_{t+k}), && (10)\\ \mathrm{Cov}(X_t, X_{t+k}) &= \frac{1}{2}\delta^2(k^2\mathrm{Var}(X^{(k)})). && (11) \end{aligned}$$

Proof of Theorem 3.1. If X satisfies (6), then from (10),

$$\begin{aligned} \mathrm{Var}(X^{(m)}) &= \frac{\sigma^2}{m} + \frac{2\sigma^2}{m^2}\sum_{k=1}^{m}(m-h)r(k) \\ &= \frac{\sigma^2}{m} + \frac{2\sigma^2}{m^2}\sum_{s=1}^{m-1}\sum_{k=1}^{s} r(k) \\ &= \frac{\sigma^2}{m}\left\{1 + \frac{1}{m}\sum_{s=1}^{m-1}\sum_{k=1}^{s}\delta^2(k^{2-\beta})\right\} \\ &= \frac{\sigma^2}{m}\left[1 + \frac{1}{m}\sum_{s=1}^{m-1}\sum_{k=1}^{s}\left\{(k+1)^{2-\beta} - k^{2-\beta}\right\} - \left\{k^{2-\beta} - (k-1)^{2-\beta}\right\}\right] \end{aligned}$$

$$
\begin{aligned}
&= \frac{\sigma^2}{m}\left\{1+\frac{1}{m}\sum_{s=1}^{m-1}(s+1)^{2-\beta}-1-s^{2-\beta}\right\} \\
&= \frac{\sigma^2}{m}\left\{1+\frac{1}{m}\left(m^{2-\beta}-(m-1)-1\right)\right\} \\
&= \sigma^2 m^{-\beta}.
\end{aligned}
$$

Conversely, assume that X satisfies (9), then from (11),

$$
\begin{aligned}
r(k) &= \frac{1}{2}\frac{\delta^2(k^2\mathrm{Var}(X^{(k)}))}{\sigma^2} \\
&= \frac{1}{2}\frac{\delta^2(k^2\sigma^2 k^{-\beta})}{\sigma^2} \\
&= \frac{1}{2}\delta^2(k^{2-\beta}).
\end{aligned}
$$

Hence the theorem follows. □

Here we show that (6), or equivalently (9) implies (8).

Theorem 3.2 *If X satisfies (6), for all $m = 1, 2, \ldots$,*

$$
r^{(m)}(k) = r(k).
$$

Proof. Assume that X satisfies (6). Now, we consider the averaged process $X' = X^{(m)}$. Applying (11) to X' yields

$$
\mathrm{Cov}(X'_t, X'_{t+k}) = \frac{1}{2}\delta^2(k^2\mathrm{Var}(X'^{(k)})). \tag{12}
$$

Note that

$$
X'^{(k)} = X^{(km)}.
$$

Then, from (12),

$$
\mathrm{Cov}(X_t^{(m)}, X_{t+k}^{(m)}) = \frac{1}{2}\delta^2(k^2\mathrm{Var}(X^{(km)})). \tag{13}
$$

Dividing by $\mathrm{Var}(X^{(m)})$, we obtain

$$
\begin{aligned}
r^{(m)}(k) &= \frac{1}{2}\frac{\delta^2(k^2\mathrm{Var}(X^{(km)}))}{\mathrm{Var}(X^{(m)})} \\
&= \frac{1}{2}\frac{\delta^2(k^2\sigma^2(km)^{-\beta})}{\sigma^2 m^{-\beta}} \\
&= \frac{1}{2}\delta^2(k^{2-\beta}) \\
&= r(k). \ \square
\end{aligned}
$$

Let $L(x)$ denote the slowly varying function at infinity, i.e. for any $n > 0$,

$$
\lim_{x\to\infty}\frac{L(nx)}{L(x)} = 1.
$$

Then, we have the following theorem about the asymptotically self-similar process.

Theorem 3.3 *X satisfies (7) if and only if for*

$$Var(X^{(m)}) \sim L(m)m^{-\beta}, \quad as\ m \to \infty, \tag{14}$$

where $a(x) \sim b(x)$ *means*

$$\lim_{x\to\infty} \frac{a(x)}{b(x)} = 1.$$

Proof. Let X be the process with autocorrelation function $r(k)$ satisfying (14). We consider the averaged process $X' = X^{(m)}$. In a similar way to derive (13), from (10),

$$\mathrm{Var}(X^{(hm)}) = \frac{\mathrm{Var}(X^{(m)})}{h} + \frac{2}{h^2}\sum_{k=1}^{h}(h-k)\mathrm{Cov}(X_t^{(m)}, X_{t+k}^{(m)}). \tag{15}$$

Dividing (15) by $\mathrm{Var}(X^{(m)})$ yields

$$\frac{\mathrm{Var}(X^{(hm)})}{\mathrm{Var}(X^{(m)})} = \frac{\mathrm{Var}(X^{(m)})}{h} + \frac{2}{h^2}\sum_{k=1}^{h}(h-k)\mathrm{Cov}(X_t^{(m)}, X_{t+k}^{(m)}).$$

Letting $m \to \infty$, we have

$$\begin{aligned}
\lim_{m\to\infty}\frac{\mathrm{Var}(X^{(hm)})}{\mathrm{Var}(X^{(m)})} &= \lim_{m\to\infty}\frac{1}{h} + \frac{2}{h^2}\sum_{s=1}^{h-1}\sum_{k=1}^{s} r^{(m)}(k) \\
&= \frac{1}{h}\left\{1 + \frac{1}{h}\sum_{s=1}^{h-1}\sum_{k=1}^{s}\delta^2(k^{2-\beta})\right\} \\
&= h^{-\beta}. \qquad (16)
\end{aligned}$$

Hence, as $m \to \infty$,

$$\mathrm{Var}(X^{(hm)}) \sim \mathrm{Var}(X^{(m)})h^{-\beta}.$$

Let $m' = hm$. Then ,as $m \to \infty$, we obtain

$$Var(X^{(m')}) \sim L(m')m'^{-\beta},$$

where

$$L(m') = \left(\frac{m'}{h}\right)^{\beta}\mathrm{Var}(X^{(\frac{m'}{h})}),$$

and

$$\begin{aligned}
\lim_{m'\to\infty}\frac{L(nm')}{L(m')} &= \lim_{m'\to\infty}\frac{\left(\frac{nm'}{h}\right)^{\beta}\mathrm{Var}(X^{(\frac{nm'}{h})})}{\left(\frac{m'}{h}\right)^{\beta}\mathrm{Var}(X^{(\frac{m'}{h})})} \\
&= \lim_{m'\to\infty} n^{\beta}\frac{\mathrm{Var}(X^{(\frac{nm'}{h})})}{\mathrm{Var}(X^{(\frac{m'}{h})})} \\
&= n^{\beta}n^{-\beta} \qquad (\mathrm{From}(16)) \\
&= 1.
\end{aligned}$$

This proves necessity. For the converse, suppose that X satisfies (14). Then, from (11), we obtain

$$\begin{aligned}
\lim_{m\to\infty} r^{(m)}(k) &= \frac{1}{2}\lim_{m\to\infty}\frac{\delta^2(k^2\mathrm{Var}(X_t^{(km)}))}{\mathrm{Var}(X^m)} \\
&= \frac{1}{2}\lim_{m\to\infty}\frac{\delta^2(k^2L(km)(km)^{-\beta})}{L(m)m^{-\beta}} \\
&= \frac{1}{2}\delta^2(k^{2-\beta})\lim_{m\to\infty}\frac{L(km)}{L(m)} \\
&= \frac{1}{2}\delta^2(k^{2-\beta}).
\end{aligned}$$

This proves sufficiency. □

From Theorem 3.1, and Theorem 3.3, we can define the self-similar process with the variance of the averaged process.

Definition 3.3 *X is called exactly second-order self-similar with the Hurst parameter $H = 1-\beta/2$ if*

$$Var(X^{(m)}) = \sigma^2 m^{-\beta}.$$

Definition 3.4 *X is called asymptotically second-order self-similar with the Hurst parameter $H = 1-\beta/2$ if*

$$Var(X^{(m)}) \sim L(m)m^{-\beta}, \quad as\ m \to \infty.$$

In our fitting method, we consider the self-similarity under Definition 3.3. we use self-similarity of (9).

4 Superposition Technique of MMPPs

We use a continuous-time MMPP for modeling the self-similar traffic. We construct MMPP with apparently self-similar behavior over several time-scales by superposing several MMPPs. First, consider two-state MMPPs with different time-scales. That is, the mean sojourn time of each process is in accordance with the different time-scale. Let us superpose them to make a new MMPP. When we see this process in a large time-scale, it looks like an ordinary two-state MMPP. If we look in a smaller time-scale, we find that each state is composed of a smaller MMPP. If we look in a still smaller time-scale, we find that each state of a smaller MMPP is again composed of a still smaller MMPP. This can be repeated only a finite number of times. Therefore the MMPP is not self-similar from the definitions in the previous section, because it looks constant when time-scale is larger than the time constant in itself. However it can emulate self-similarity over several time-scales. It is impossible to measure a given traffic during an infinite amount of time. It also has been observed that LAN traffic loses self-similarity in the order of days[3]. Thus, it is practically sufficient to use the process which has self-similarity over only several time-scales to model the real traffic.

Now we assume that the number of states of every underlying MMPP is two. So the MMPP obtained by the above method is also described by the superposition of several interrupted Poisson processes (IPP) and one Poisson process. We assume that the MMPP under consideration consists of $d(>1)$ IPPs and a Poisson process. We also assume that two modulating parameters of each IPP are equal. Then we can describe ith IPP as follows:

$$Q_i = \begin{bmatrix} -\sigma_i & \sigma_i \\ \sigma_i & -\sigma_i \end{bmatrix}, \quad \Lambda_i = \begin{bmatrix} \lambda_i & 0 \\ 0 & 0 \end{bmatrix}, \quad (1 \leq i \leq d).$$

Hence the superposition can be described as follows:

$$\begin{aligned} Q &= Q_1 \bigoplus Q_2 \bigoplus \cdots \bigoplus Q_d, \\ \Lambda &= \Lambda_1 \bigoplus \Lambda_2 \bigoplus \cdots \bigoplus \Lambda_d \bigoplus \lambda_p, \end{aligned}$$

where $\bigoplus$ means the Kronecker's sum and λ_p is the arrival rate of the Poisson process to be superposed. The whole arrival rate of the process λ is given by

$$\lambda = \lambda_p + \sum_{i=1}^{d} \frac{\lambda_i}{2}. \tag{17}$$

In the next section, we consider how to determine the parameters of these IPPs and the Poisson process.

5 Fitting Procedure

In this section we describe the process of determining the parameters of MMPPs. Their values are obtained so as to to match the variance over several time-scales. Parameters which we have to determine are σ_i, $\lambda_i (1 \leq i \leq d)$, and λ_p.

First, as preliminary we define the notations used in the procedure and describe some assumptions. Let $N_{t|i}$ be the number of arrivals in the i-th IPP during the tth time slot and $N_{t|p}$ be the number of arrivals in the Poisson process, and let $N_{t|i}^{(m)}$ and $N_{t|p}^{(m)}$ be respectively the averaged processes of $N_{t|i}$ and $N_{t|p}$. We assume that

$$\mathrm{Var}(X_t^{(m)}) = \mathrm{Var}(\sum_{i=1}^{d} N_{t|i}^{(m)} + N_{t|p}^{(m)}).$$

Using (5), we obtain the variance of the i-th IPP as

$$\mathrm{Var}(N_{t|i}^{(m)}) = \frac{\lambda_i}{2m} + \left\{ \frac{1}{4m\sigma_i} - \frac{1}{8m^2\sigma_i^2}(1 - e^{-2m\sigma_i}) \right\} \lambda_i^2.$$

The corresponding variance of the Poisson process is λ_p/m. Because the variance of a process which is a superposition of independent subprocesses equals the sum of individual variances, the variance of the whole process is given by

$$\begin{aligned}
\mathrm{Var}(X_t^{(m)}) &= \frac{\lambda_p}{m} + \sum_{i=1}^{d} \mathrm{Var}(N_{t|i}^{(m)}) \\
&= \frac{\lambda}{m} + \frac{1}{4}\sum_{i=1}^{d}\left\{\frac{1}{m\sigma_i} - \frac{1}{2m^2\sigma_i^2}(1-e^{-2m\sigma_i})\right\}\lambda_i^2 \\
&\equiv \frac{\lambda}{m} + \frac{1}{4}\sum_{i=1}^{d}\eta_i\lambda_i^2. \qquad (18)
\end{aligned}$$

where we used (17). Using (9) and (18), we match the variance at d different points m_i $(1 \le i \le d)$. Suppose time-scales over which we want the process to express self-similarity of the original process is $m_{\min} \le m \le m_{\max}$, then m_i is defined by

$$m_i = m_{\min} a^{i-1} \quad (1 \le i \le d),$$

where

$$a = \left(\frac{m_{\min}}{m_{\max}}\right)^{\frac{1}{d-1}}, \qquad d > 1. \qquad (19)$$

We investigate the property of η_i. Let $x = m\sigma_i$, then we have

$$\eta_i = \frac{1}{x} - \frac{1}{2x^2}\left(1 - e^{-2x}\right) \equiv f(x).$$

It is easily seen that for $x > 0$,

$$1 - 2x < e^{-2x} < 1 - 2x + 2x^2. \qquad (20)$$

Using (20), we obtain

$$0 < f(x) < 1.$$

That is, for all i,

$$0 < \eta_i < 1. \qquad (21)$$

From (18) and (21), we obtain

$$\begin{aligned}
\frac{\lambda}{m} < \mathrm{Var}(X_t^{(m)}) &< \frac{\lambda}{m} + \frac{1}{4}\sum_{i=1}^{d}\lambda_i^2 \\
&= \frac{\lambda}{m} + \sum_{i=1}^{d}\left(\frac{\lambda}{2}\right)^2 \\
&\le \frac{\lambda}{m} + \left(\sum_{i=1}^{d}\frac{\lambda}{2}\right)^2 \\
&\le \frac{\lambda}{m} + \lambda^2, \qquad \text{(from (17))}
\end{aligned}$$

i.e.,

$$\frac{\lambda}{m} < \mathrm{Var}(X_t^{(m)}) < \frac{\lambda}{m} + \lambda^2. \qquad (22)$$

We must choose such m_i that (22) is satisfied at any m_i. This condition comes from that we use a simple IPP as a sub-process. Practically, this condition never matter when m is large, but sometimes $\mathrm{Var}(X_t^{(m)})$ is too small when m is small. Therefore, we should be careful to choose m_1, which is enough large that $\mathrm{Var}(X_t^{(m_1)})$ is larger than λ/m_{min}.

Furthermore, we assume the following relation of σ_i and m_i

$$m_i\sigma_i = const \quad (1 \le i \le d).$$

That is, σ_i can be described as

$$\sigma_i = \frac{m_1}{m_i}\sigma_1 \quad (1 \le i \le d). \tag{23}$$

This assumption comes from the intuitive understanding that a self-similar process looks the same in any time-scale. By this assumption, we can reduce the number of the parameters to be determined. That is, if we determine σ_1, we can obtain the values of σ_i $(2 \le i \le d)$ by using (23). Furthermore, we can obtain λ_p from (17) if we determine λ_i. Now the parameters we need to find are only σ_1 and λ_i.

In the following, we describe the procedure of determining these parameters.

Procedure of Parameter Fitting

Step 1. Determine λ_i as the function of σ_i. From (9) and (18), we have

$$\sigma^2 \begin{bmatrix} m_1^{-\beta} \\ m_2^{-\beta} \\ \vdots \\ m_d^{-\beta} \end{bmatrix} = \lambda \begin{bmatrix} m_1^{-1} \\ m_2^{-1} \\ \vdots \\ m_d^{-1} \end{bmatrix} + B \begin{bmatrix} \lambda_1^2 \\ \lambda_2^2 \\ \vdots \\ \lambda_d^2 \end{bmatrix}, \tag{24}$$

where B is the $d \times d$ matrix whose (i, j) element is

$$B_{ij} = \frac{1}{4m_i\sigma_j} - \frac{1}{8m_i^2\sigma_j^2}(1 - e^{-2m_i\sigma_j}). \tag{25}$$

Solving this, we determine λ_i as the function of σ_i.

Step 2. Determine the values of σ_1. Here we consider the integral of the difference between the log-scales variance curve of the process given by (18) and that of the self-similar process given by (9) over defined time-scales. We determine the value of σ_1 so as to minimize that integral.

Step 3. Determine the values of λ_i from (24).

In step 1, it is necessary that B is non-singular. It is difficult to prove the non-singularity of B for any positive integer of d, however, we can show that if a is sufficiently large, B is non-singular for any σ_1. We discuss the non-singularity of the matrix B in the next section. We can solve this problem by choosing such a.

When we minimize the integral in step 2, we must be careful to keep the values of λ_i and λ_p larger than zero. We consider the minimum at the log-scale because we can treat smaller time-scales more carefully.

At the end of this section, we show the parameters preliminarily required for our fitting procedure in Table 1.

Table 1: Preliminarily required parameters for the fitting

parameter	meaning
λ	Arrival rate of the whole process
m_{min}, m_{max}	Minimum and maximum of time-scales over which self-similarity is taken into consideration
σ^2	Variance
H	Hurst parameter
d	Number of IPPs

6 Sufficient Condition for Existence of B^{-1}

In this section, we consider the sufficient condition under which the matrix B is invertible. We use the following lemma ([13] 2.3.2. Perturbation Lemma).

Lemma 6.1 *Let A and C denote $n\times n$ matrices with real elements. Suppose that A is invertible with $\|A^{-1}\| \leq \alpha$, where $\|\cdot\|$ is an arbitrary matrix norm. If $\|A - C\| \leq \beta$ and $\alpha\beta < 1$, then C is invertible, and*

$$\left\|C^{-1}\right\| \leq \frac{\alpha}{1-\alpha\beta}.$$

Let $c = m_1\sigma_1$, then the (i,j)-th element of B becomes

$$B_{ij} = \frac{1}{4}\left\{\frac{1}{ca^{i-j}} - \frac{1}{2c^2a^{2(i-j)}}(1 - e^{-2ca^{i-j}})\right\}.$$

We define

$$\gamma_k = \frac{1}{4}\left\{\frac{a^k}{c} - \frac{a^{2k}}{2c^2}(1 - e^{-2ca^{-k}})\right\}.$$

Then we have

$$B = \begin{pmatrix} \gamma_0 & \gamma_1 & \cdots & \gamma_{n-1} \\ \gamma_{-1} & \gamma_0 & \cdots & \gamma_{n-2} \\ \vdots & \vdots & \ddots & \vdots \\ \gamma_{-n+1} & \gamma_{-n+2} & \cdots & \gamma_0 \end{pmatrix}.$$

This type of the matrix is called Toeplitz [7]. Note that we obtain the following inequality in a similar way to (21)

$$0 < \gamma_k < \frac{1}{4}. \tag{26}$$

We define the $n \times n$ matrix A as

$$A = \begin{pmatrix} \gamma_0 & \gamma_1 & \cdots & \gamma_{n-1} \\ 0 & \gamma_0 & \cdots & \gamma_{n-2} \\ \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & \gamma_0 \end{pmatrix}.$$

Then, $A - B$ is given by

$$A - B = -\begin{pmatrix} 0 & \cdots & 0 & 0 \\ \gamma_{-1} & \ddots & 0 & 0 \\ \vdots & \ddots & \ddots & \vdots \\ \gamma_{-n+1} & \cdots & \gamma_{-1} & 0 \end{pmatrix}.$$

For the matrix norm, we consider the l_1-norm. The l_1-norm of A is defined by

$$\|A\|_1 = \max_{1 \le j \le n} \sum_{i=1}^{n} |a_{ij}|,$$

where a_{ij} is the (i,j)-th element of A. We obtain $\|A - B\|_1$ as

$$\begin{aligned} \|A - B\|_1 &= \sum_{k=-n+1}^{-1} |\gamma_k| \\ &= \sum_{k=-n+1}^{-1} \frac{1}{4} \left\{ \frac{a^k}{c} - \frac{a^{2k}}{2c^2} \left(1 - e^{-2ca^{-k}} \right) \right\} \\ &< \sum_{k=-n+1}^{-1} \frac{1}{4c} a^k \\ &< \frac{1}{4c(a-1)}. \end{aligned}$$

Note that $a > 1$ from (19).

Next, we consider $\|A^{-1}\|_1$. We define the submatrix of A as

$$A_k = \begin{pmatrix} \gamma_0 & \gamma_1 & \cdots & \gamma_{k-1} \\ 0 & \gamma_0 & \cdots & \gamma_{k-2} \\ \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & \gamma_0 \end{pmatrix}.$$

Since A_k is an upper triangular matrix, A_k^{-1} must be in the form as

$$A_k^{-1} = \begin{pmatrix} \xi_0 & \xi_1 & \cdots & \xi_{k-1} \\ 0 & \xi_0 & \cdots & \xi_{k-2} \\ \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & \xi_0 \end{pmatrix}.$$

Because $A_k A_k^{-1} = I$ where I is the identity matrix, we have

$$\xi_0 = \frac{1}{\gamma_0},$$

and

$$\gamma_0\xi_{k-1} + \gamma_1\xi_{k-2} + \cdots + \gamma_{k-1}\xi_0 = 0, \qquad k \geq 2.$$

Now we state the following lemma.

Lemma 6.2 *For all $k \geq 1$,*

$$|\xi_k| < \frac{1}{4\gamma_0^2}\left(1 + \frac{1}{4\gamma_0}\right)^{k-1}. \tag{27}$$

Proof. We prove the lemma by induction.

(i) For $k = 1$, we have

$$\xi_1 = -\frac{\gamma_1}{\gamma_0^2}.$$

Hence,

$$|\xi_1| = \frac{\gamma_1}{\gamma_0^2} < \frac{1}{4\gamma_0^2}.$$

(ii) Assume that (27) is satisfied for all $k \leq m$. For $k = m + 1$,

$$\xi_{m+1} = -\frac{1}{\gamma_0}\sum_{l=0}^{m}\gamma_{m+1-l}\xi_l.$$

Hence,

$$\begin{aligned}
|\xi_{m+1}| &= \frac{1}{\gamma_0}\left|\sum_{l=0}^{m}\gamma_{m+1-l}\xi_l\right| \\
&\leq \frac{1}{\gamma_0}\sum_{l=0}^{m}|\gamma_{m+1-l}\xi_l| \\
&\leq \frac{1}{\gamma_0}\sum_{l=0}^{m}|\gamma_{m+1-l}|\,|\xi_l| \\
&< \frac{1}{4\gamma_0}\sum_{l=0}^{m}|\xi_l| \qquad \text{(from (26))} \\
&< \frac{1}{4\gamma_0}\left\{\frac{1}{\gamma_0} + \sum_{l=1}^{m}\frac{1}{4\gamma_0^2}\left(1 + \frac{1}{4\gamma_0}\right)^{l-1}\right\} \\
&= \frac{1}{4\gamma_0^2}\left(1 + \frac{1}{4\gamma_0}\right)^{m}.
\end{aligned}$$

From (i) and (ii), (27) is satisfied for all $k \geq 1$. □

We obtain $\left\|A^{-1}\right\|_1$ as

$$\left\|A^{-1}\right\|_1 = \sum_{k=0}^{n-1}|\xi_k|.$$

Using Lemma 6.2, we obtain the following inequality

$$\left\|A^{-1}\right\|_1 < \frac{1}{\gamma_0} + \sum_{k=1}^{n-1} \frac{1}{4\gamma_0^2}\left(1+\frac{1}{4\gamma_0}\right)^{k-1} = \frac{1}{\gamma_0}\left(1+\frac{1}{4\gamma_0}\right)^{n-1}.$$

From Lemma 6.1, the sufficient condition for the matrix B to be invertible is given by

$$\frac{1}{\gamma_0}\left(1+\frac{1}{4\gamma_0}\right)^{n-1}\frac{1}{4c(a-1)} < 1.$$

Solving for a, we obtain

$$a > 1 + \frac{1}{4c\gamma_0}\left(1+\frac{1}{4\gamma_0}\right)^{n-1}.$$

We state the following theorem as the result of this section.

Theorem 6.1 *If a satisfies the following inequality*

$$a > 1 + \frac{1}{4c\gamma_0}\left(1+\frac{1}{4\gamma_0}\right)^{n-1}, \tag{28}$$

then the matrix B is invertible.

Remark. Note that the value of right hand side of (28) depends on c and n, i.e., m_1, σ_1 and n.

7 Numerical Results

In this section, we present some numerical results obtained from our fitting method. We consider only the case that the number of IPPs d equals 2, that is, two IPPs are superposed. First, we show the variance-time curves of the resulting MMPPs. Next, we examine the queueing behavior of the resulting MMPP comparing with the simulation result for the queueing system with the self-similar traffic. For our simulation, simulated self-similar traffic trace is needed. We generate the fractional Brownian traffic (FBT) based on the FBM with the random midpoint displacement (RMD) algorithm and use it as the self-similar traffic. We summarize the FBT and the RMD algorithm and show some numerical examples of the simulation and the resulting MMPP/D/1.

7.1 Variance

In Figures 1-4, we show the variance-time curves of the MMPPs obtained from our fitting method. Solid lines illustrate the variance of obtained MMPPs and dotted lines correspond to $\sigma^2 m^{-\beta}$. We set $d = 2$, $m_1 = 1$, $m_2 = 10^4$, $\lambda = 1.0$ and changed values of H and σ^2 as follows: $(H, \sigma^2) = (0.6, 1.1), (0.8, 1.5), (0.8, 1.9)$ and $(0.9, 1.5)$. We can imitate accurately the variance curve of the self-similar process with the four state MMPP when the value of H is not so large and the value of σ^2 is moderate. However, when the Hurst parameter is large, or when σ^2 is near the upper bound or lower bound of (22), the difference is quite large. One of the reasons is that we must keep the values of λ_i and λ_p larger than zero when we minimize the difference of variances for the self-similar traffic and for the resulting MMPP.

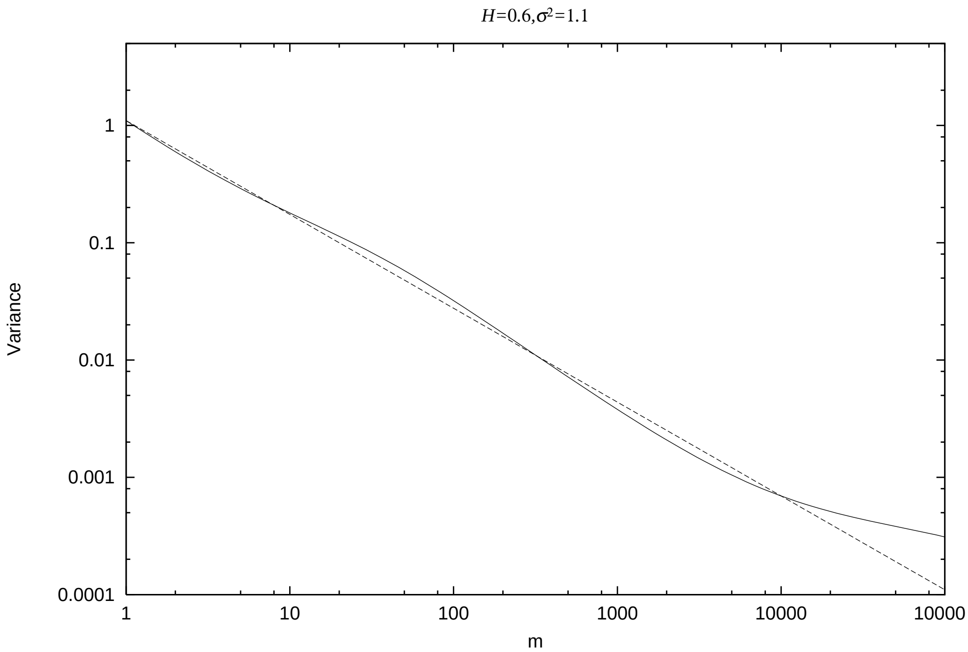

Figure 1: Variance-time curves of MMPP, $H = 0.6, \sigma^2 = 1.1$

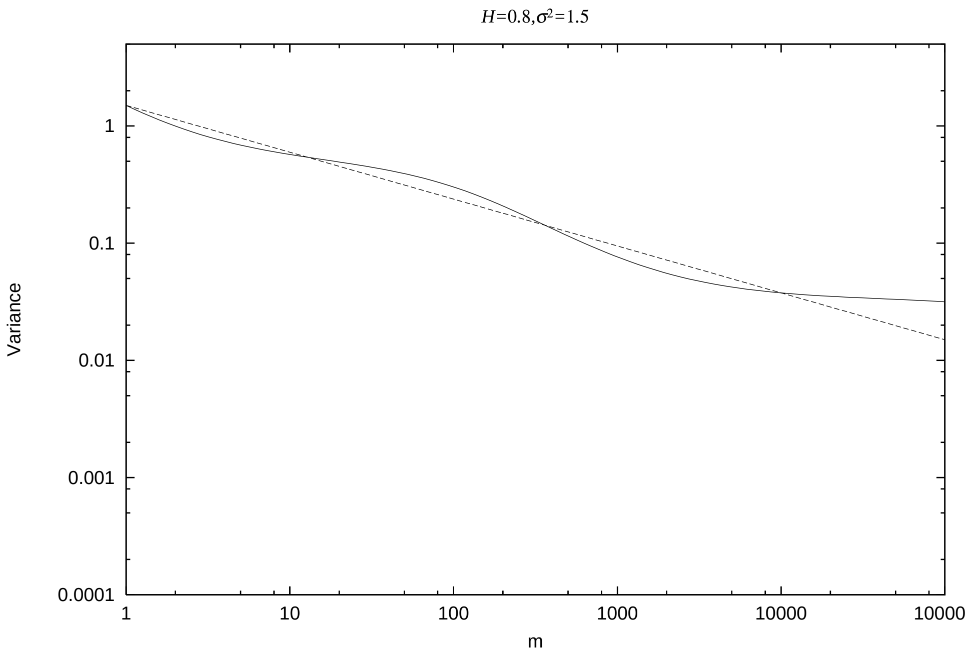

Figure 2: Variance-time curves of MMPP, $H = 0.8, \sigma^2 = 1.5$

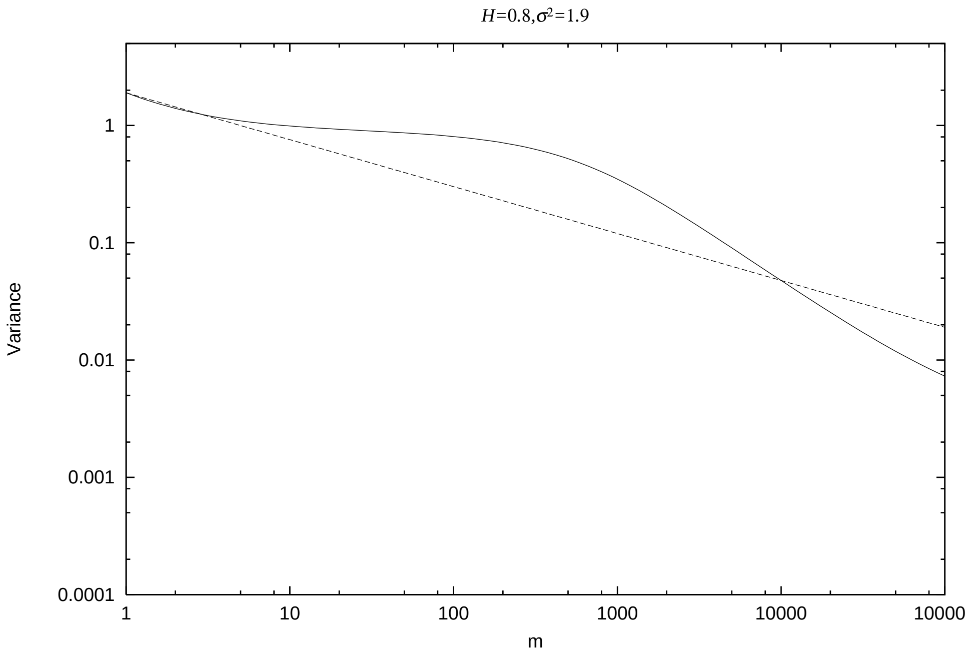

Figure 3: Variance-time curves of MMPP, $H = 0.8, \sigma^2 = 1.9$

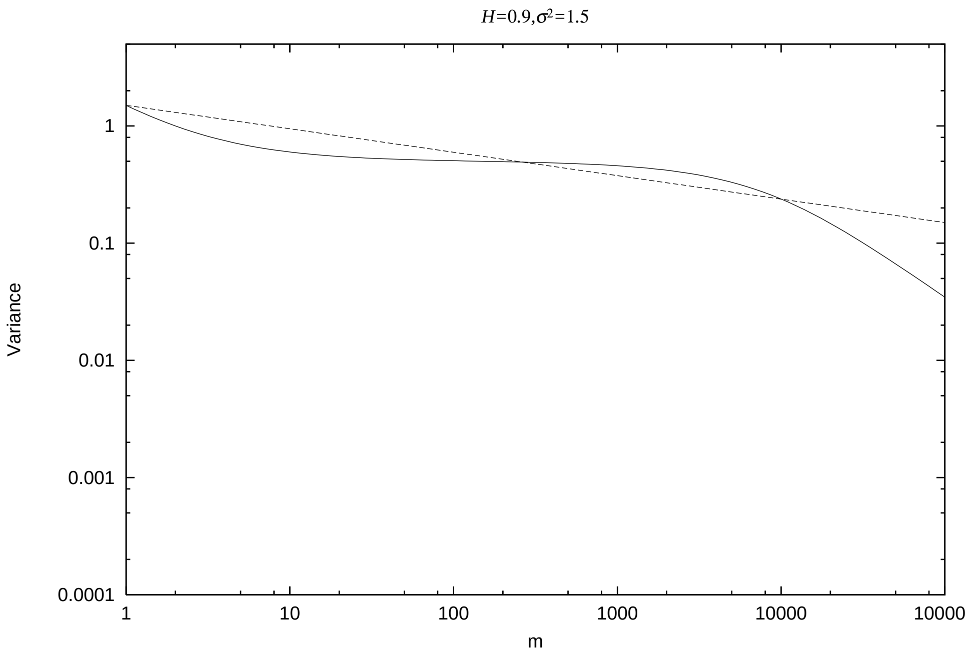

Figure 4: Variance-time curves of MMPP, $H = 0.9, \sigma^2 = 1.5$

7.2 FBM and RMD algorithm

In the following, we summarize the FBM, FBT and RMD algorithm. The readers are referred to [11] and [8] for details. The FBM $Z(t)$ is a continuous zero mean Gaussian process. It has stationary increments and

$$E[Z(t)^2] = |t|^{2H}.$$

In the case of $H = 0.5$, $Z(t)$ is the standard Brownian motion. Using FBM, the fractional Brownian traffic (FBT) is defined as the cumulative arrival process $A(t)$ [11]:

$$A(t) = \mu t + \sqrt{\zeta\mu} Z(t), \qquad -\infty < t < \infty,$$

where $Z(t)$ is the FBM, μ the mean rate, and ζ the variance coefficient. The mean and the variance of the FBT are as follows:

$$\begin{aligned} E[A(t)] &= \mu t, \\ \mathrm{Var}[A(t)] &= \zeta\mu|t|^{2H}. \end{aligned}$$

We obtain the following

$$\mathrm{Var}(A(\xi t)) = \xi^{2H}\mathrm{Var}(A(t)),$$

which shows $A(t)$ is self-similar.

We use the RMD algorithm in order to generate FBT. The RMD algorithm generates FBM traces approximately. It never requires large amount of time to generate long traces. However, it must be applied carefully in quantitative studies because the parameters of the generated traces can differ from their target values.

The algorithm is summarized as follows. Assume that we want to generate an FBM trace in the time interval $[0, T]$. Roughly speaking, it works recursively; first subdivides the interval $[0, T]$, then determines the values of the process at the midpoints from the values at the endpoints. Let us consider the case of determining the values $Z(\frac{t_1+t_2}{2})$ at the midpoint of an interval $[t_1, t_2]$ from the values $Z(t_1)$ and $Z(t_2)$ at the endpoints. In this algorithm, it is assumed that the midpoint displacement $Z(\frac{t_1+t_2}{2}) - \frac{Z(t_1)+Z(t_2)}{2}$ is independent of the increment $Z(t_2) - Z(t_1)$ over the whole interval. This assumption is not valid except for the case of $H = 0.5$, but results in fast computation at the expense of exactness. The sample sequences of FBT generated by RMD algorithm are shown in Figure 5 and Figure 6. In those figures, the vertical axis is time while the horizontal axis represents the number of arrivals per unit interval.

7.3 Queueing Behavior

In this subsection, we examine the queueing performance of the resulting MMPP/D/1. As the performance measure, we calculate the mean waiting time. First, we generate a sample sequence of FBT with the RMD algorithm. Then, we fit MMPPs to this sample sequence. Finally, we compare the mean waiting time of the resulting MMPPs with simulation results with the sample

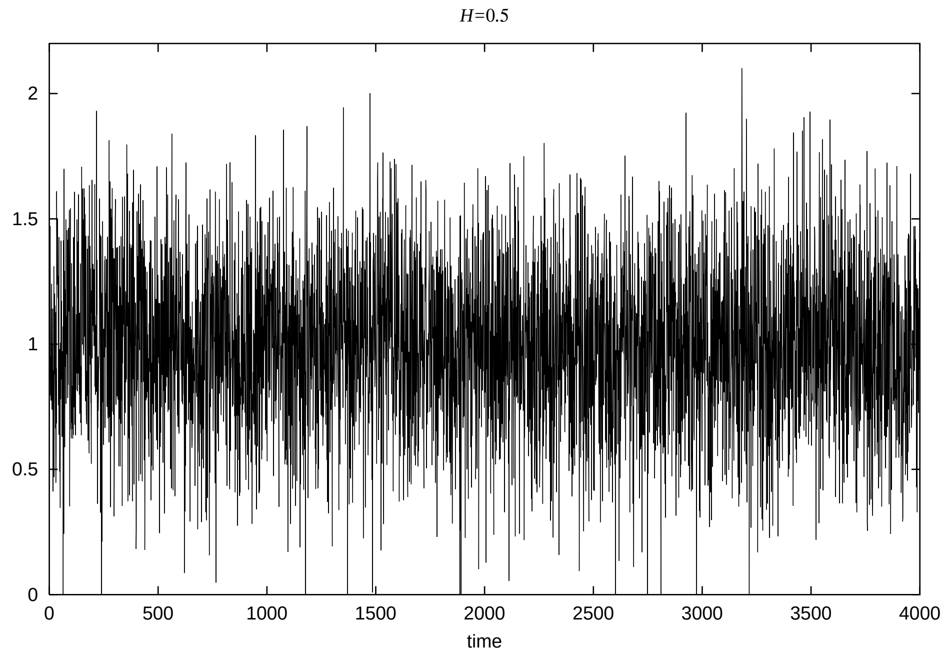

Figure 5: Sample sequence of FBT based on RMD algorithm, $H = 0.5, \mu = 1.0, \zeta = 0.1$

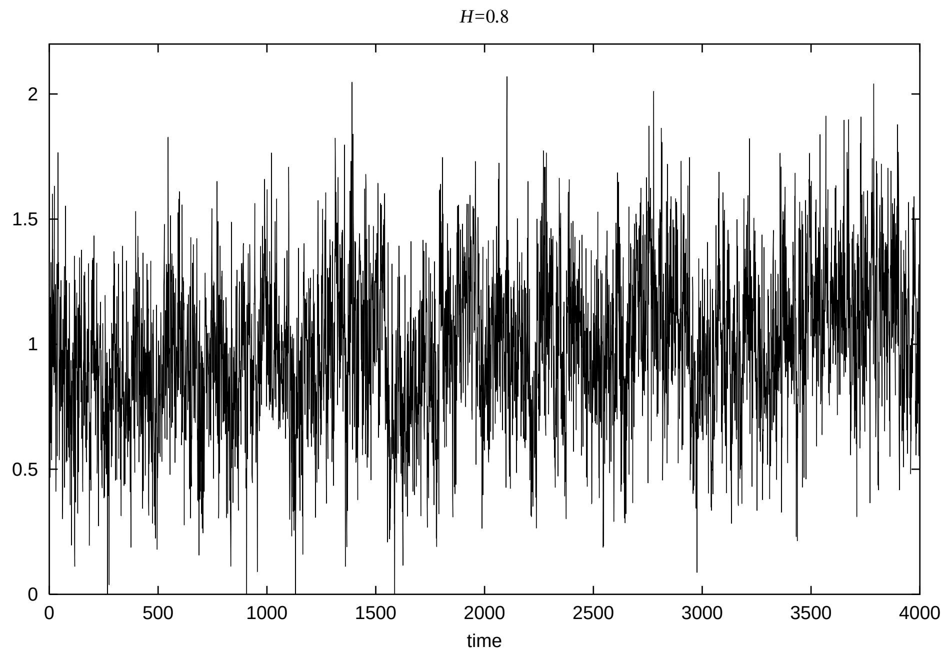

Figure 6: Sample sequence of FBT based on RMD algorithm, $H = 0.8, \mu = 1.0, \zeta = 0.1$

sequence. We consider four different samples (see Table 2). Note that in fitting, we need to estimate the parameters of FBT again. Table 2 shows estimated parameters of the sample sequence.

Figures 7 to 10 show the variance-time curves of the resulting MMPPs. From these graphs we see that the difference is very large when m is small. We need to set ζ small compared with μ when we generate sample sequences of FBT with the RMD algorithm in order to keep $Z(t)$ positive. On the other hand, the variance of the MMPP is conditioned by (22). Therefore it is impossible to fit the variance at small m. However, as m gets larger, the variance of the self-similar process gets smaller so as to satisfy (22) because it decays slowly.

Figures 11 to 14 present the mean waiting time of the MMPP/D/1 and the results of the simulations with the sample sequences. In Figure 12 and 13, we see that the mean waiting time of the resulting MMPPs is very close to the simulation results when the load is high while it is not close when the load is small. However, in Figure 12 and 13 the discrepancy is very large throughout a global range of the load. These results suggest that our method works enough for middle values of H but the difference becomes large when H is near 0.5 or 1.

Table 2: Estimated parameters of sample sequence

	sample1	sample2	sample3	sample4
λ	1.0	1.0	1.0	1.0
σ^2	0.18	0.19	0.18	0.19
H	0.59	0.68	0.78	0.86
$m_{\min}$	5×10^4	150	50	50
$m_{\max}$	10^5	10^4	10^4	10^4

8 Conclusions

In this paper, we first gave some definitions of self-similarity. We then introduced a fitting method for the self-similar traffic in terms of MMPP. We constructed an MMPP as the superposition of two-state MMPPs and fit it so as to match the variance function over several time-scales. Numerical results have shown that the variance function of the self-similar process can be well represented by that of four-state MMPP when the Hurst parameter is not so large and the variance is moderate. We also examined the queueing behavior of the resulting MMPP/D/1 queueing systems. The comparison of the mean waiting time of the MMPP/D/1 queues with the simulation results has shown that our method is accurate for large load when the value of the Hurst parameter is in the middle range.

References

[1] A. T. Anderson and B. F. Nielsen, "An application of superposition of two state Markovian

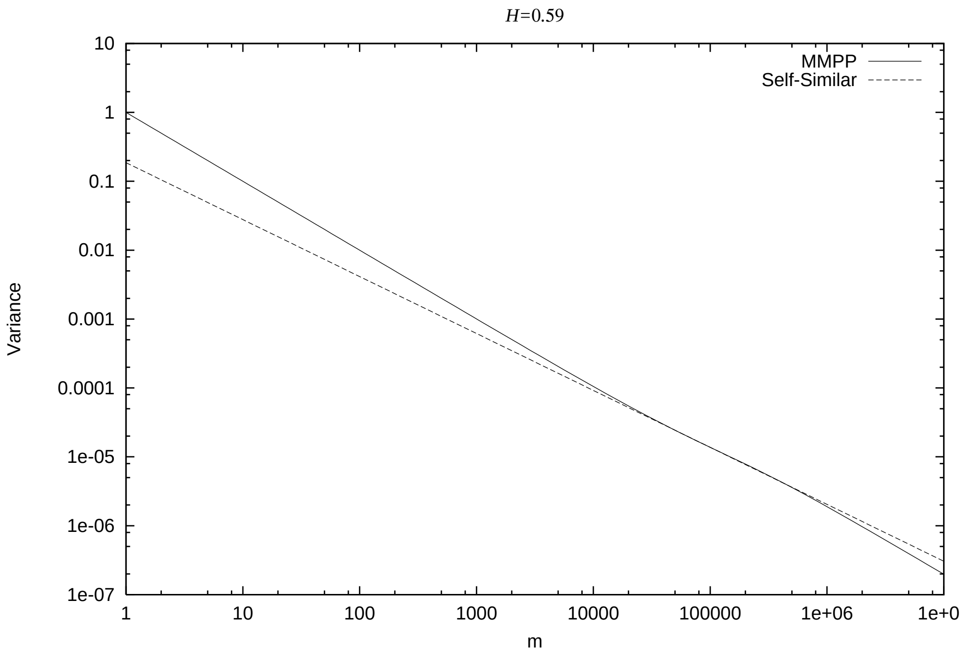

Figure 7: Variance-time curve, $H = 0.59$

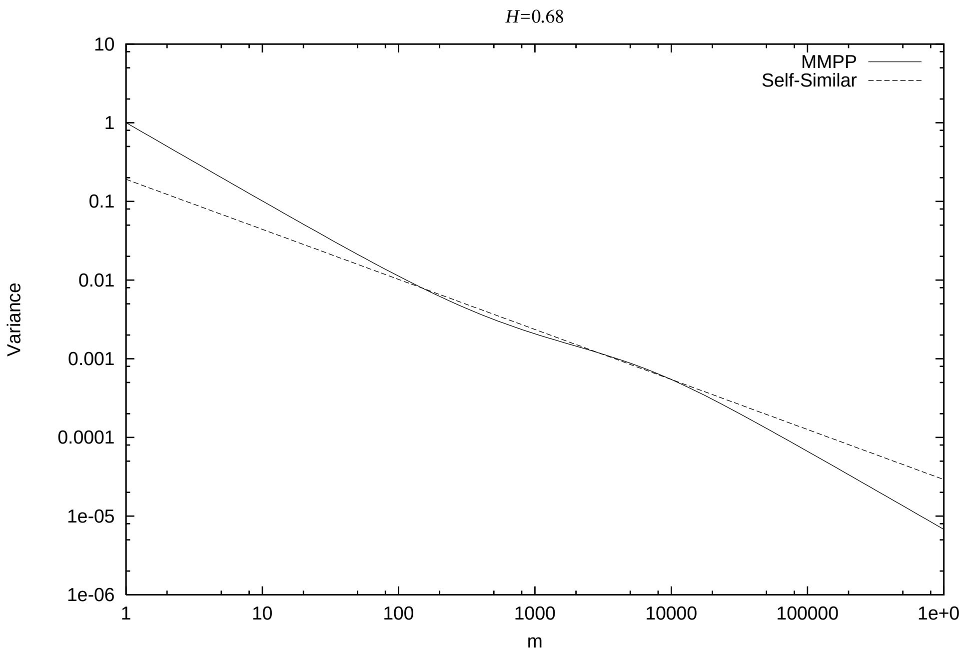

Figure 8: Variance-time curve, $H = 0.68$

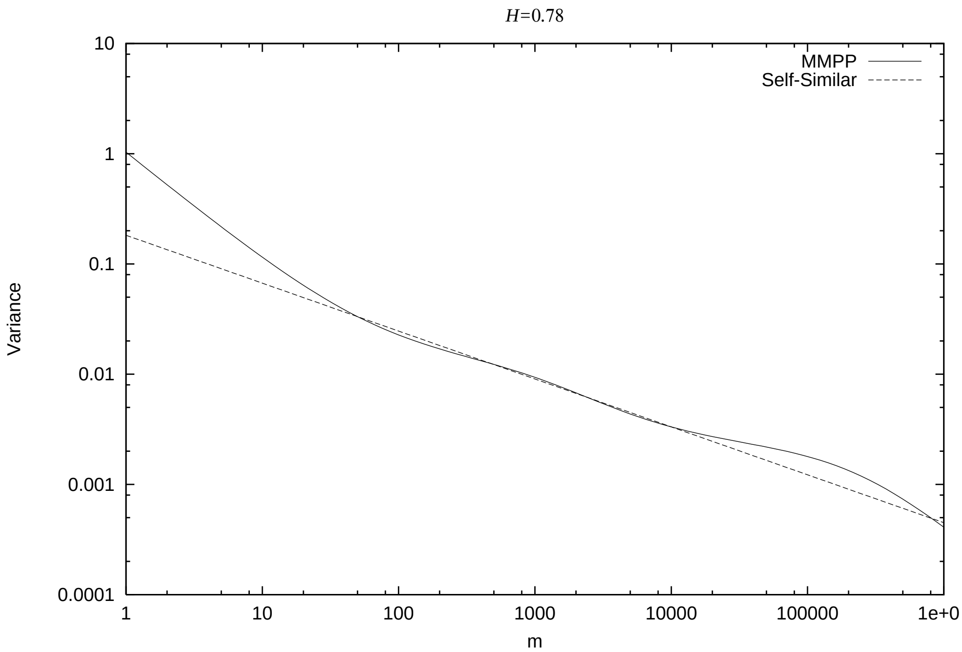

Figure 9: Variance-time curve, $H = 0.78$

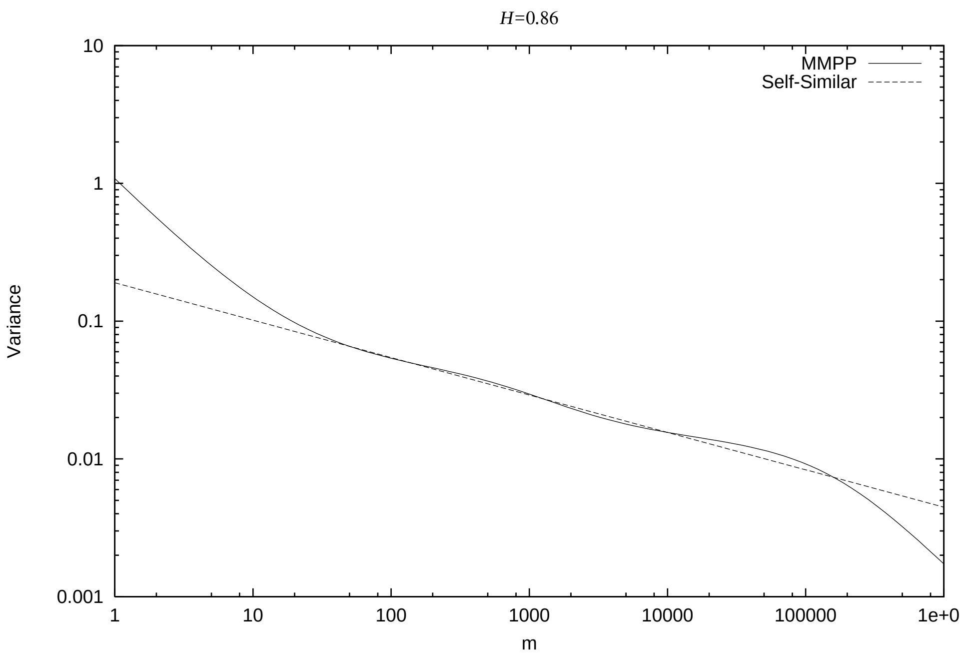

Figure 10: Variance-time curve, $H = 0.86$

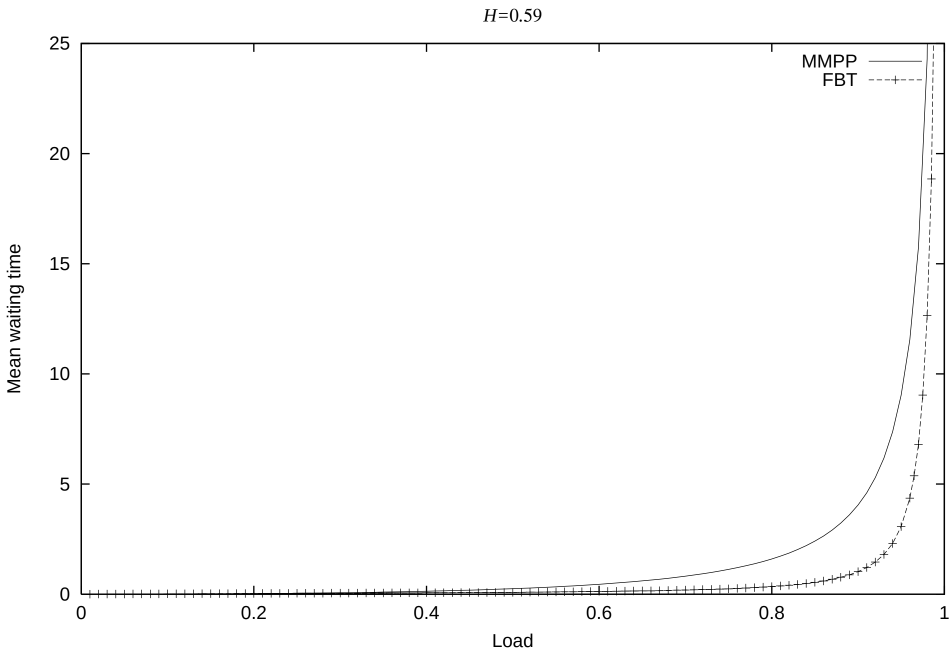

Figure 11: Mean waiting time vs Load, $H = 0.59$

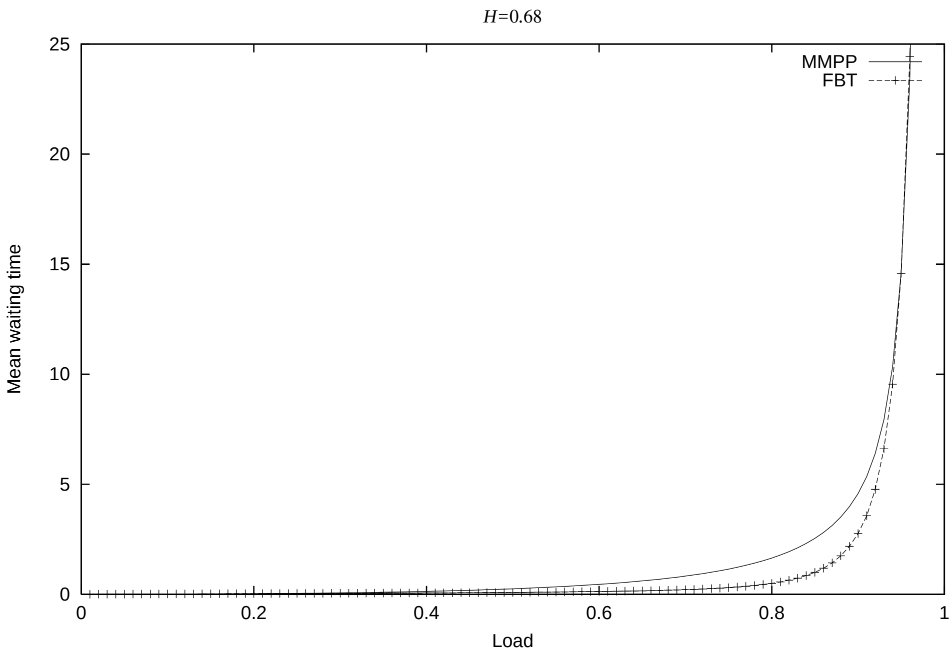

Figure 12: Mean waiting time vs Load, $H = 0.68$

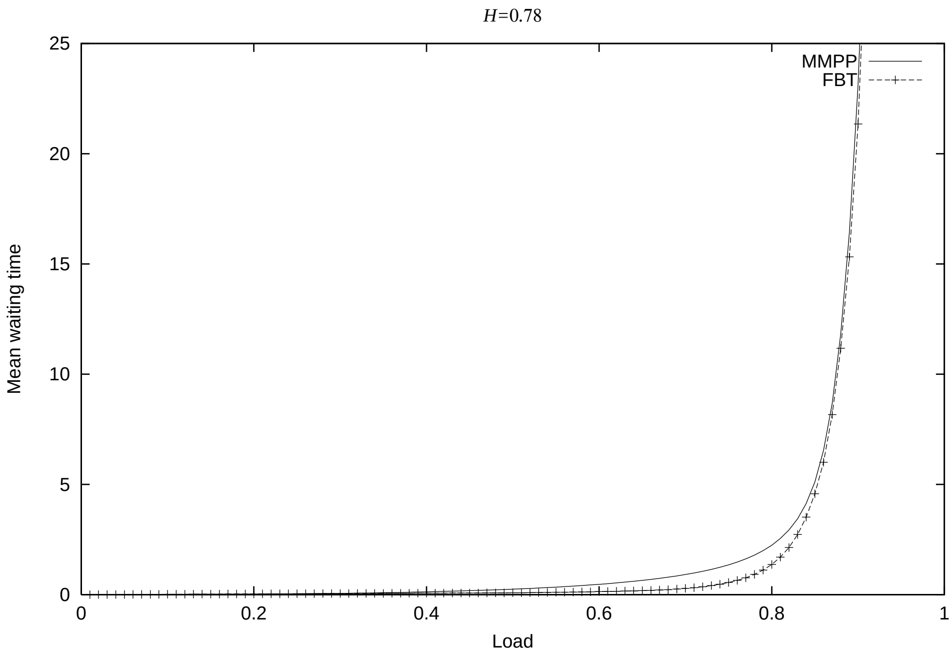

Figure 13: Mean waiting time vs Load, $H = 0.78$

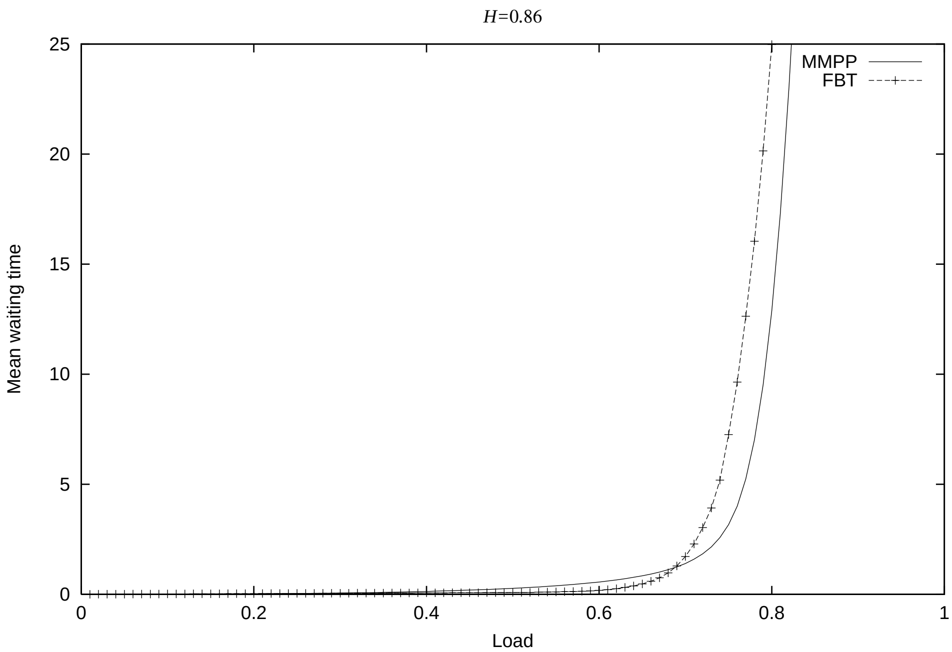

Figure 14: Mean waiting time vs Load, $H = 0.86$

source to the modeling of self-similar behavior," *Proc. INFOCOM'97,* Kobe, Japan, pp. 196-204, 1997.

[2] D. R. Cox, "Long-range dependence: A review," in *Statistics: An appraisal,* H. A. David and H. T. David, eds. Ames, IA: The Iowa State University Press, pp. 55-74, 1984.

[3] A. Erramilli, O. Narayan and W. Willinger, "Fractal queueing models," *Frontiers in Queueing,* ed. J. H. Dshalalow, CRC Press, Inc., pp. 245-269. 1997.

[4] A. Erramilli and R. P. Singh, "Chaotic maps as models of packet traffic," *Proc. 14th ITC,* Antibes, Juan-les-Pins, France, pp. 329-338, 1994.

[5] W. Fisher and K. S. Meier-Hellstern, "The Markov-modulated Poisson process (MMPP) cookbook," *Performance Evaluation,* vol. 18, pp.149-171, 1992.

[6] H. Heffes and D. M. Lucantoni, "A Markov modulated characterization of packetized voice and data traffic and related statistical multiplexer performance," *IEEE Journal on Selected Areas in Communications,* vol. 4, no. 6, pp. 856-868, 1986.

[7] R. A. Horn and C. R. Johnson, *Matrix analysis,* Cambridge University Press, 1985.

[8] W.-C. Lau, A. Erramilli, J. L. Wang, and W. Willinger, "Self-similar traffic generation: The random midpoint displacement algorithm and its properties," *Proc. ICC'95,* Seattle, WA, pp. 466-472, 1995.

[9] W. Leland, M. Taqqu, W. Willinger and D. Wilson, "On the self-similar nature of Ethernet traffic(extended version)," *IEEE/ACM Transactions on Networking,* vol. 2, no. 1, pp. 1-15, 1994.

[10] N. Likhanov, B. Tsybakov and N. D. Georganas, "Analysis of an ATM buffer with self-similar ('fractal') input traffic," *IEEE INFOCOM'95,* Boston, MA, pp. 985-992, 1995.

[11] I. Norros, "On the use of fractional Brownian motion in the theory of connectionless networks," *IEEE Journal on Selected Areas in Communications,* vol. 13, pp. 953-962, 1994.

[12] S. Robert and J. Y. Le Boudec, "New models for pseudo self-similar traffic," *Performance Evaluation,* vol. 30, pp. 57-68, 1997.

[13] J. M. Ortega and W. C. Rheinboldt, *Iterative solution of nonlinear equations in several variables,* Academic Press, 1970.

[14] B. Tsybakov and N. D. Georganas, "On self-similar traffic in ATM queues: definitions, overflow probability bound, and cell delay distribution," *IEEE/ACM Transactions on Networking,* vol. 5, no. 3, pp.397-409, 1997.