Practical Time-Scale Fitting of Self-Similar Traffic with Markov-Modulated Poisson Process

Tadafumi Yoshihara [a], Shoji Kasahara [b] and Yutaka Takahashi [c]

[a] *NTT Information Sharing Platform Laboratories*
3-9-11 Midori-cho, Musashino-shi, Tokyo 180-8585, Japan
E-mail: yoshihara.tadafumi@na.tnl.ntt.co.jp
[b] *Graduate School of Information Science, Nara Institute of Science and Technology*
8916-5 Takayama, Ikoma, Nara 630-0101, Japan
E-mail: kasahara@is.aist-nara.ac.jp
[c] *Graduate School of Informatics, Kyoto University*
Kyoto 606-8501, Japan
E-mail: takahashi@i.kyoto-u.ac.jp

Recent measurements of packet/cell streams in multimedia communication networks have revealed that they have the self-similar property and are of different characteristics from traditional traffic streams. 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 resulting MMPPs. We also examine the queueing behavior of the resulting MMPP/D/1 queueing systems. We compare the analytical results of MMPP/D/1 with the simulation ones of the queueing system with self-similar input.

Keywords: Self-similar traffic, MMPP, Fitting, FBM, Queueing performance.

AMS Subject classification: 60K25, 68M20

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* [4,12]. 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 [5], fractional Brownian motion (FBM) [14] and fractional autoregressive integrated moving average (FARIMA) model [4]. 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. [1,2] have proposed a model based on Markov-modulated Poisson process (MMPP) as a superposition of two state Markov processes. In [16], a discrete-time 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 [13], they have constructed a self-similar process from an infinity of on-off sources with Pareto service demands.

As for the relating works in terms of describing self-similar processes, [18] considered the fractal point processes (FPPs) and proposed four models based on FPPs. [6] proposed the approximation method for long-tail distributions using hyperexponential distributions.

Another significant and practical issue of the self-similarity is the effect of the time-scaling. [9] discussed the impact of the long-range dependence (LRD) on the buffer occupancy and indicated that LRD does not affect the buffer occupancy when the busy periods of the system are not large. In [17], the authors considered the critical time scale (CTS) and showed that the buffer behavior at the time-scale beyond the CTS is not significantly affected.

In this paper, we first give some definitions of self-similarity which are equivalent to those in [3]. In particular, we propose an equivalent definition of asymptotically self-similar process. Then, using the definition of the strictly self-similar process, 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. [1,2], where traffic is modeled by the superposition of several two state MMPPs.

In [1,2], the authors proposed the fitting method which is mainly focused on the covariance structure of the second-order self-similar process. More precisely, 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.

It is well known that the queueing performance deteriorates with the self-similar traffic whose Hurst parameter is between 0.5 and 1, i.e., LRD traffic [4]. Hence it is

important to predict the queueing behavior under the self-similar traffic with LRD. In general, the fitting methods based on the second-order statistics of counts for the arrival process are not sufficient for predicting the queueing performance. (See [1,2] and references therein for details.) Here, we have two goals for developing the fitting method using MMPP. First one is that the resulting MMPP will have the same statistical characteristics as the original self-similar traffic. Second goal is to predict the queueing behavior under any traffic conditions. That is, the performance measures such as the mean waiting time and the tail distribution of queue length are the same as those of simulation results driven by the self-similar traffic. For judging whether our fitting method works well or not, we consider the following criteria:

1. Statistical Characteristics
   If the variance-time curve of the resulting MMPP agrees with that of self-similar process over specified time-scales, we can say that the fitting method works well in the sense of statistical characteristics of traffic.

2. Queueing Performance
   We consider the mean waiting time and the tail distribution of queue length. If those of the queueing system with resulting MMPP input agree with the simulation results under any traffic conditions, we can say that the fitting method works well in the sense of queueing performance.

   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, conclusions and discussions are presented in section 8.

2.   Markov-Modulated Poisson Process

   In this section, we summarize some main characteristics of MMPP [7,8]. 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 λ_i $(1 \leq i \leq m)$. 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

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

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

$$\begin{cases} \dfrac{d}{dt} P(n, t) = P(n, t)(Q - \Lambda) + P(n - 1, t)\Lambda, & n \geq 1, \ t \geq 0, \\ P(0, 0) = I, \end{cases} \tag{1}$$

where I is the identity matrix. Multiplying (1) by z^n and summing over $n = 0, 1, \ldots$, we obtain

$$\begin{cases} \dfrac{d}{dt} P^*(z, t) = P^*(z, t)(Q - \Lambda) + z P^*(z, t)\Lambda, \\ P^*(z, 0) = I, \end{cases} \tag{2}$$

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

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

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 (see Appendix A and [8])

$$\mathrm{Var}(N_t) = \frac{\sigma_2 \lambda_1 + \sigma_1 \lambda_2}{\sigma_1 + \sigma_2} t + 2 A_1 t - \frac{2 A_1}{\sigma_1 + \sigma_2} (1 - e^{-(\sigma_1 + \sigma_2)t}), \tag{3}$$

where

$$A_1 = \frac{\sigma_1 \sigma_2 (\lambda_1 - \lambda_2)^2}{(\sigma_1 + \sigma_2)^3}.$$

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 [3] 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,

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

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

$$r(k) = \frac{1}{2}\delta^2(k^{2-\beta}). \tag{4}$$

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

$$r^{(m)}(k) \rightarrow \frac{1}{2}\delta^2(k^{2-\beta}), \quad \text{as } m \rightarrow \infty. \tag{5}$$

Note that (4) implies that for all $m = 1, 2, \ldots,$

$$r^{(m)}(k) = r(k). \tag{6}$$

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 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 1 and 2. For the case of exact second-order self-similarity, what we discuss in the following is shown in [19].

Theorem 3. [19] X satisfies (4) if and only if for all $m = 1, 2, \ldots$,

$$\mathrm{Var}(X^{(m)}) = \sigma^2 m^{-\beta}. \tag{7}$$

Proof.   See Appendix B.   □

Next theorem implies that (4), or equivalently (7) implies (6).

Theorem 4. [19]   If X satisfies (4), for all $m = 1, 2, \ldots$,

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

Proof.   See Appendix B.   □

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 5. X satisfies (5) if and only if for

$$\mathrm{Var}(X^{(m)}) \sim L(m)m^{-\beta}, \quad \text{as } m \to \infty, \tag{8}$$

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 (8). We consider the averaged process $X' = X^{(m)}$. In a similar way to derive (28), from (25),

$$\mathrm{Var}(X^{(hm)}) = \frac{\mathrm{Var}(X^{(m)})}{h} + \frac{2}{h^2} \sum_{k=1}^{h} (h-k)\mathrm{Cov}(X_t^{(m)}, X_{t+k}^{(m)}). \tag{9}$$

Dividing (9) 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

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

$$= h^{-\beta}. \tag{10}$$

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

$$\begin{aligned}
\lim_{m' \to \infty} \frac{L(nm')}{L(m')} &= \lim_{m' \to \infty} \frac{\left(\frac{nm'}{h}\right)^{\beta} \mathrm{Var}(X^{(\frac{nm'}{h})})}{\left(\frac{m'}{h}\right)^{\beta} \mathrm{Var}(X^{(\frac{m'}{h})})} \\
&= \lim_{m' \to \infty} n^{\beta} \frac{\mathrm{Var}(X^{(\frac{nm'}{h})})}{\mathrm{Var}(X^{(\frac{m'}{h})})} \\
&= n^{\beta}n^{-\beta} \qquad (\text{From}(10)) \\
&= 1.
\end{aligned}$$

This proves necessity. For the converse, suppose that X satisfies (8). Then, from (26), we obtain

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

This proves sufficiency. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\square$

From Theorems 3 and 5, we can define the self-similar process with the variance of the averaged process.

Definition 6. X is called exactly second-order self-similar with the Hurst parameter $H = 1 - \beta/2$ if

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

Definition 7. X is called *asymptotically* second-order self-similar with the Hurst parameter $H = 1 - \beta/2$ if

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

In our fitting method, we consider the self-similarity under Definition 6, that is, we develop the fitting method using (7).

4.   Superposition Technique of MMPPs

We use a continuous-time MMPP for modeling the self-similar traffic. We construct an 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 [4]. 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

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

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

$$\lambda = \lambda_p + \sum_{i=1}^{d} \frac{\lambda_i}{2}. \tag{11}$$

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 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 (3), 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

$$\begin{aligned}
\mathrm{Var}(X_t^{(m)}) &= \frac{\lambda_p}{m} + \sum_{i=1}^{d} \mathrm{Var}(N_{t|i}^{(m)}) \\
&= \frac{\lambda}{m} + \frac{1}{4} \sum_{i=1}^{d} \left\{ \frac{1}{m\sigma_i} - \frac{1}{2m^2\sigma_i^2}(1 - e^{-2m\sigma_i}) \right\} \lambda_i^2 \\
&\equiv \frac{\lambda}{m} + \frac{1}{4} \sum_{i=1}^{d} \eta_i \lambda_i^2.
\end{aligned} \tag{12}$$

where we used (11). Using (7) and (12), we match the variance at d different points m_i $(1 \leq i \leq d)$. Suppose the range of time-scales over which we want the process to express self-similarity of the original process is $m_{\min} \leq m \leq m_{\max}$, then m_i is defined by

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

where

$$a = \left(\frac{m_{\max}}{m_{\min}}\right)^{\frac{1}{d-1}}, \qquad d > 1. \tag{13}$$

Here, 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 $x > 0$,

$$1 - 2x < e^{-2x} < 1 - 2x + 2x^2. \tag{14}$$

Using (14), we obtain

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

That is, for all i,

$$0 < \eta_i < 1. \tag{15}$$

From (12) and (15), we obtain

$$\frac{\lambda}{m} < \mathrm{Var}(X_t^{(m)}) < \frac{\lambda}{m} + \frac{1}{4}\sum_{i=1}^{d}\lambda_i^2$$

$$= \frac{\lambda}{m} + \sum_{i=1}^{d}\left(\frac{\lambda}{2}\right)^2$$

$$\leq \frac{\lambda}{m} + \left(\sum_{i=1}^{d}\frac{\lambda}{2}\right)^2$$

$$\leq \frac{\lambda}{m} + \lambda^2, \qquad \text{(from (11))}$$

i.e.,

$$\frac{\lambda}{m} < \mathrm{Var}(X_t^{(m)}) < \frac{\lambda}{m} + \lambda^2. \tag{16}$$

We must choose m_i such that (16) is satisfied at any m_i. This condition comes from that we use a simple IPP as a sub-process. Practically, this condition never matters 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 large enough to make $\mathrm{Var}(X_t^{(m_1)})$ larger than $\lambda/m_{\min}$.

Next, we assume the following relation between σ_i and m_i

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

That is, σ_i can be described as

$$\sigma_i = \frac{m_1}{m_i}\sigma_1 \quad (1 \leq i \leq d). \tag{17}$$

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 $(2 \leq i \leq d)$ by using (17). Furthermore, we can obtain λ_p from (11) 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. We show the parameters preliminarily required for our fitting procedure in Table 1.

Procedure of Parameter Fitting

Step 0. Find the range of σ_1 heuristically and fix σ_1.

Step 1. Determine λ_i as the function of σ_i. From (7) and (12), we have

$$\sigma^2 \begin{bmatrix} m_1^{-\beta} \\ m_2^{-\beta} \\ \vdots \\ m_d^{-\beta} \end{bmatrix} = \lambda \begin{bmatrix} m_1^{-1} \\ m_2^{-1} \\ \vdots \\ m_d^{-1} \end{bmatrix} + B \begin{bmatrix} \lambda_1^2 \\ \lambda_2^2 \\ \vdots \\ \lambda_d^2 \end{bmatrix}, \tag{18}$$

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

$$B_{ij} = \frac{1}{4 m_i \sigma_j} - \frac{1}{8 m_i^2 \sigma_j^2}(1 - e^{-2 m_i \sigma_j}). \tag{19}$$

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

Step 2. Determine the value of σ_1 from the range found in Step 0. Here we consider the integral of the difference between the log-scales variance curve of the process given by (12) and that of the self-similar process given by (7) 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 (18).

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.

Remark 8. In Table 2, we compare our method with that of [1]. Here, we call the procedure of [1] covariance method and ours variance method. In our method, the fitting procedure is exactly constructed while that of [1] contains some approximations. However, our method has a constraint of (16) which implies that our method does not holds when $\mathrm{Var}(X_t^{(m)}) \leq \lambda/m$ or $\mathrm{Var}(X_t^{(m)}) \geq \lambda/m + \lambda^2$. That is, our method is not

Table 1
Preliminarily Required Parameters for 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

enough to fit the process whose variance is too small or too large in comparison with the mean arrival rate. The method of [1] does not contain the specified constraints.

Remark 9. [1] proposed the refined method of [2], where the type of MMPP components is the switched Poisson Process (SPP). Since our method is based on using IPPs, the accuracy of the resulting MMPP of [1] is expected to be better than ours. However, the parameters of resulting IPPs can be changed into SPPs keeping first- and second-order properties unchanged (see [1] and the references therein for details). In addition, this modification is performed by matching the autocorrelation of each SPP with the observed process. Thus it is possible to refine our method using SPPs. However, we do not treat this refinement in this paper.

Table 2
Comparisons between Variance and Covariance Methods

	Variance Method	Covariance Method
Required Parameters	λ, H, d, σ^2, Time Scale	$\lambda, H, d, r(1)$, Time Scale
Type of Component MMPPs	IPP	SPP
Parameter Fitting	Exact	Approximation
Constraint	Eq. (16)	None

6. Sufficient Condition for Non-Singularity of B

In this section, we consider the sufficient condition under which the matrix B is invertible. We use the following lemma ([15] 2.3.2. Perturbation Lemma).

Lemma 10. 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

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

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

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

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 [10]. Note that we obtain the following inequality in a similar way to (15)

$$0 < \gamma_k < \frac{1}{4}. \tag{20}$$

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

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

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 (13).

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

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

From $A_k A_k^{-1} = I$, 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 11. For all $k \geq 1$,

$$
|\xi_k| < \frac{1}{4\gamma_0^2} \left(1 + \frac{1}{4\gamma_0} \right)^{k-1}. \tag{21}
$$

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 (21) 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,

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

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

$$\leq \frac{1}{\gamma_0} \sum_{l=0}^{m} |\gamma_{m+1-l}| \, |\xi_l|$$

$$< \frac{1}{4\gamma_0} \sum_{l=0}^{m} |\xi_l| \qquad \text{(from (20))}$$

$$< \frac{1}{4\gamma_0} \left\{ \frac{1}{\gamma_0} + \sum_{l=1}^{m} \frac{1}{4\gamma_0^2} \left(1 + \frac{1}{4\gamma_0} \right)^{l-1} \right\}$$

$$= \frac{1}{4\gamma_0^2} \left(1 + \frac{1}{4\gamma_0} \right)^{m}.$$

From (i) and (ii), (21) is satisfied for all $k \geq 1$. $\qquad\qquad\square$

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

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

Using Lemma 11, 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 10, 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 12. If a satisfies the following inequality

$$a>1+\frac{1}{4c\gamma_0}\left(1+\frac{1}{4\gamma_0}\right)^{n-1}, \tag{22}$$

then the matrix B is invertible.

Remark 13. Note that the value of right hand side of (22) 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 show the variance-time curves of the resulting MMPPs. Then, we examine the queueing behavior of the resulting MMPP/D/1 comparing with the simulation for the queueing system with the self-similar traffic. We compute performance measures such as waiting time using analytical results presented in [7].

7.1. FBM and RMD Algorithm

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 [11] and use it as self-similar traffic. In this subsection, we summarize the FBM, FBT and RMD algorithm. The readers are referred to [11] and [14] 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 $A(t)$ [14]

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

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

$$E[A(t)] = \mu t, \quad \text{Var}[A(t)] = \zeta\mu|t|^{2H}.$$

We obtain the following

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

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 resulting sample sequences of fractional Gaussian noise (FGN), which is an increment process of FBM, are shown in Figure 1 and Figure 2. In those figures, the horizontal axis is time while the vertical axis represents the number of arrivals per unit interval.

7.2. Generated Sequences and Fitting Results

Using RMD method, we generated Samples 1 to 3, three sequences of arrival time whose time-scale is up to 10^6. Samples 1, 2 and 3 were generated changing $H = 0.6, 0.7$ and 0.8, respectively, with $\lambda = 1.0$ and $\sigma^2 = 0.6$. Then we estimated mean arrival rate λ, variance σ^2, and Hurst parameter H for each sequence using aggregated variance method (for details, see [12] and references therein). The results of estimation are presented in Table 3.

In Figures 3 and 4, we show the variance-time curves of the MMPPs obtained from Sample 2 according to our fitting method. In both figure, dotted lines illustrate the variance curves of resulting MMPPs and solid lines correspond to $\sigma^2 m^{-\beta}$.

Figure 3 represents the effect of the number of component MMPPs on the variance-time curves. We fit MMPP parameters to Sample 2 with $m_{\min} = 10^2$ and $m_{\max} = 10^5$,
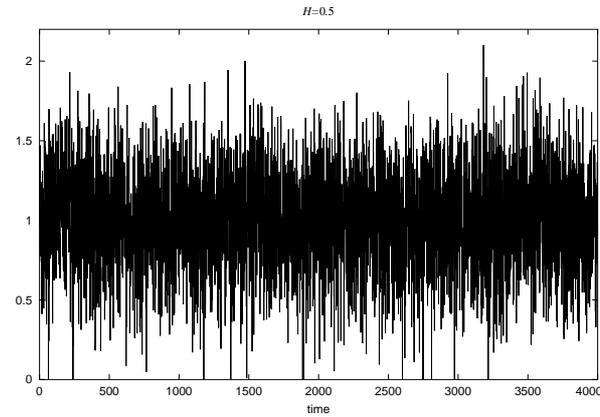
Figure 1. Sample Sequence of FGN Based on RMD Algorithm, $H = 0.5, \mu = 1.0, \zeta = 0.1$.
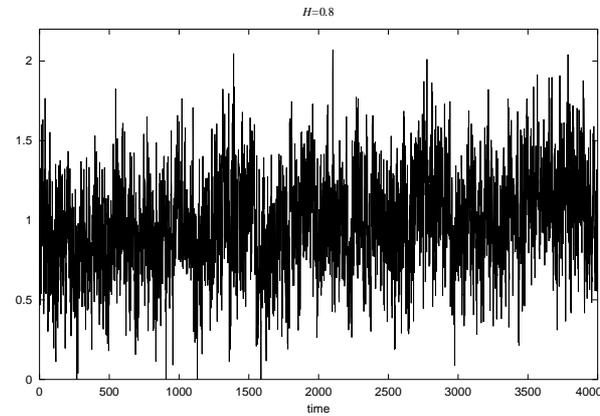
Figure 2. Sample Sequence of FGN Based on RMD Algorithm, $H = 0.8, \mu = 1.0, \zeta = 0.1$.

Table 3
Estimated Parameters of Sample Sequence

	Sample 1	Sample 2	Sample 3
λ	1.00441	1.00071	0.98349
σ^2	0.59710	0.60290	0.60008
H	0.59923	0.69952	0.79841

changing $d = 2, 3$ and 4. It is observed that we can imitate the variance curve of the exact self-similar process as d increases.

Figure 4 shows the variance-time curves with $d = 4$ changing the maximum time-

scale $m_{\max} = 10^4$, 10^6, and 10^8, respectively. From this figure, it is observed that the accuracy of MMPP fitting within the specified time-scale becomes worse when $m_{\max}$ increases.
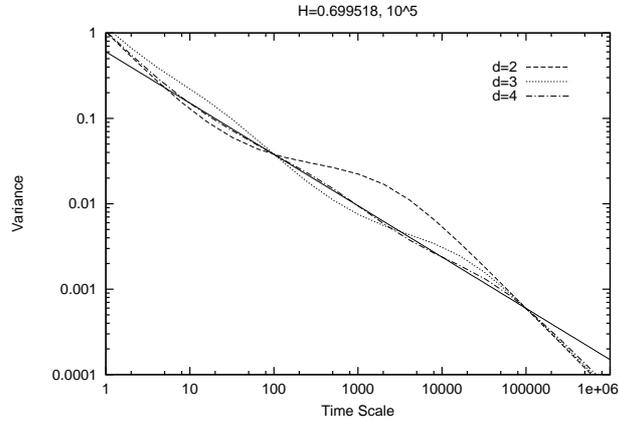
Figure 3. Variance-Time Curves of MMPPs for Sample 2, Changing d.
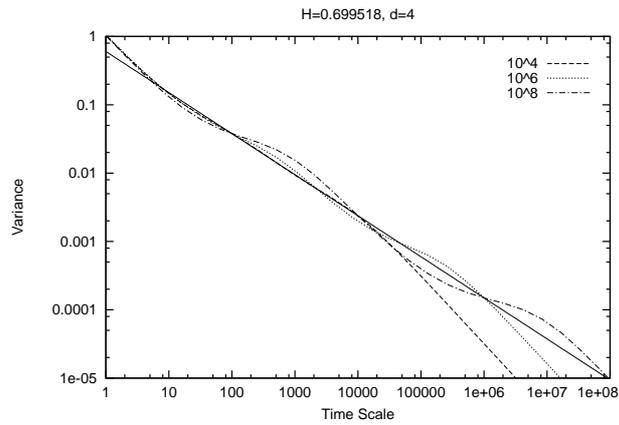
Figure 4. Variance-Time Curves of MMPPs for Sample 2, Changing Time-Scales.

7.3. Queueing Behavior

In this subsection, we show the queueing performance of the resulting MMPP/D/1 in comparison with the simulation. As the performance measure, we calculated mean

waiting time and complement distribution of queue length. Each simulation in the following graphs was driven by corresponding sample data generated by RMD (see Table 3).

Figures 5 to 7 show the waiting time curves of the resulting MMPP/D/1 in comparison with the simulation. In those figures, we set $m_{\min} = 10^2$ and $m_{\max} = 10^5$. The waiting times of MMPP/D/1 are calculated changing the number of component MMPPs. It is observed that our fitting method is performed well in the case of $H = 0.59923$ (Figure 5) while not good when $H = 0.79841$ (Figure 7).

In Figure 5, the waiting time of MMPP/D/1 exhibits the good agreement with the simulation result irrespective of the number of d. In Figure 6, the waiting times of MMPP/D/1 are getting different from the simulation result as the traffic intensity becomes large. In particular, the discrepancy between simulation and MMPP/D/1 under $d = 2$ is the largest. We also observe that the waiting times under $d = 3$, 4 and 5 are almost the same. These tendencies are remarkable in Figure 7.

From above observations, our fitting method performs well for evaluating the mean waiting time when the Hurst parameter is small. In this case, the number of component MMPPs does not affect the waiting time so much. However, when the Hurst parameter is large, our fitting method does not work well. In addition, the mean waiting time is affected by the number of component MMPPs, and increasing the number of components does not always make mean waiting time improved.
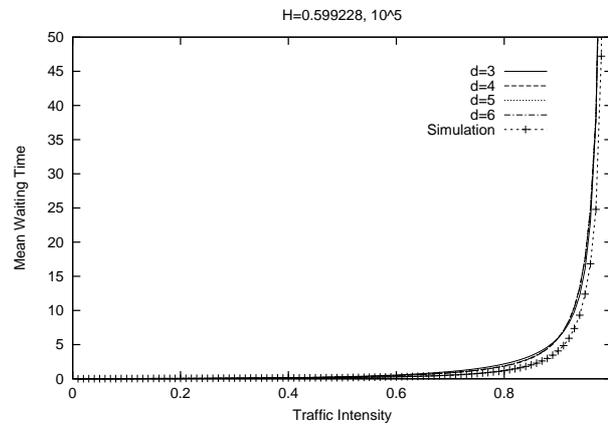
Figure 5. Mean Waiting Time under Sample 1.

Next, we investigate how the time-scale chosen in our fitting procedure affects the buffer behavior. Let L denote the number of customers in MMPP/D/1 system at an arbitrary point of time. Here we consider the probability $\Pr(L > N)$ as the loss
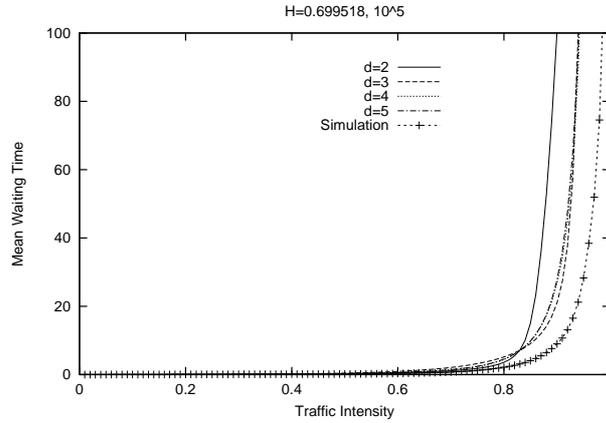
Figure 6. Mean Waiting Time under Sample 2.
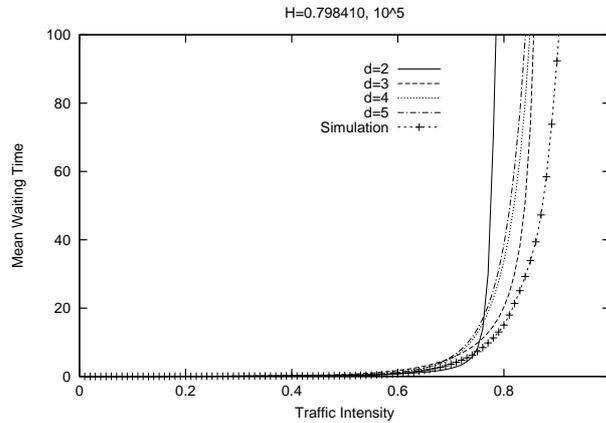
Figure 7. Mean Waiting Time under Sample 3.

probability of the system with size N. Figures 8 to 10 present the curves of $\Pr(L > 10)$ changing $m_{\max}$, while $m_{\min}$ is fixed to 10^2. In these figures, Sample 2 is used for fitting and simulation.

In Figure 8, we set the number of component MMPPs equal to two. It is observed that $\Pr(L > 10)$ is not affected by $m_{\max}$ so much and that it gives good approximation in comparison with the simulation. However, Figure 9 illustrates that $m_{\max}$ have a large effect on $\Pr(L > 10)$ in the case of $d = 3$. In this figure, $\Pr(L > 10)$ becomes large as $m_{\max}$ is getting large. Figure 10 is the case of $d = 4$ and we observe the same tendency as Figure 8.

From these results, it seems that our fitting method is not robust in terms of
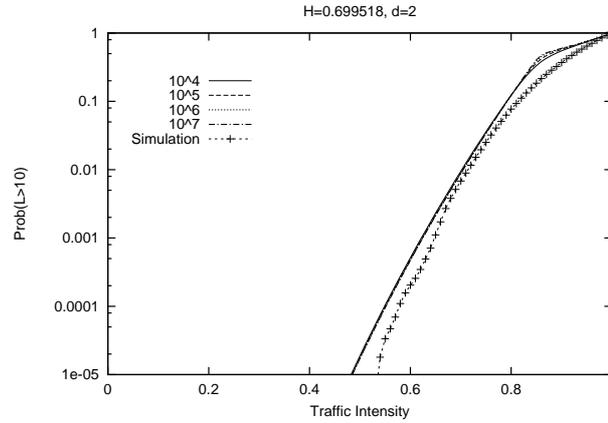
Figure 8. $\Pr(L > 10)$ of Sample 2, $d = 2$.
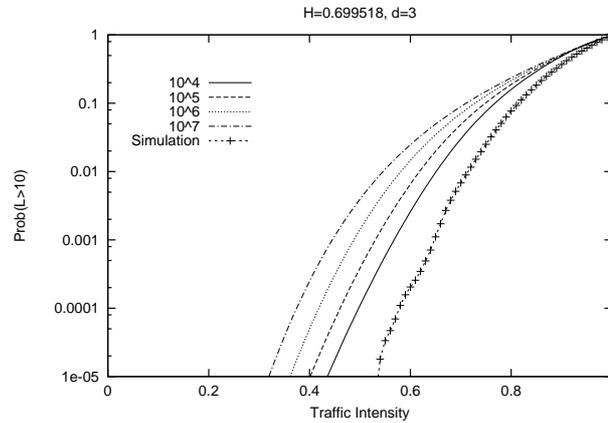
Figure 9. $\Pr(L > 10)$ of Sample 2, $d = 3$.

choosing time-scale and hence it is not available for investigating the impact of the time-scale on the queueing behavior. Further research is needed to capture the effect of the time-scale using Markovian input model.

Remarks. 1. In our fitting procedure, we cannot choose any number of d. For example, the range of d is from three to seven under Sample 1 with $m_{\min} = 10^2$ and $m_{\max} = 10^5$, while it is from two to four under Sample 2 with $m_{\min} = 10^2$ and $m_{\max} = 10^4$. The number of component MMPPs is affected by both Hurst parameter and the range of time-scales.

2. For applying our fitting method to the practical situation, we must be careful to choose
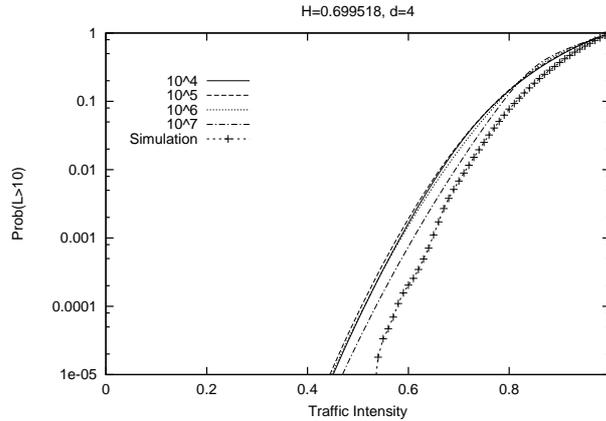
Figure 10. $\Pr(L > 10)$ of Sample 2, $d = 4$.

the number of component MMPPs. Under numerical calculations, we also observed that the resulting MMPP/D/1 becomes untractable when both Hurst parameter and the number of component MMPPs are large. For example, the probability matrix of fundamental period, namely G in [7], is computed in a iterative way and the convergence of G is largely affected by Hurst parameter and the number of states of MMPP.

8. Conclusions

In this paper, we first gave some definitions of self-similarity and proved the equivalent definition of second-order asymptotically self-similar process. We then proposed 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.

In the numerical examples, we presented the variance-time curves and queueing performance. As for the variance-time curves, the resulting MMPP mimics the ideal self-similar process well over the specified time-scales. In addition, the accuracy becomes improved as the number of component MMPPs is getting large. Therefore, we can say that our proposed method works well in the sense of statistical characteristics of traffic.

In terms of the queueing performance, we investigated the mean waiting time and the tail distribution of the queue length for the single-server queueing system with deterministic service. Numerical results of mean waiting time showed that our fitting method works well under light traffic regardless of the Hurst parameter. This is because the long-range effect of the arrival process tends to disappear under light traffic. However, under heavy traffic, the discrepancy between MMPP/D/1 and simulation becomes large

when H is around or greater than 0.7. This is due to the LRD of the self-similar traffic. From this result, our fitting method is not sufficient for predicting the mean waiting time accurately under the heavy traffic even when the value of H is moderate. It is the limitation of the fitting based on the second-order statistics [1,2,4].

Fortunately, it seems that our method succeeds in giving the upper bound of the mean waiting time even when the traffic intensity is high. For obtaining the tight upper bounds, we have to be careful for choosing the number of component MMPPs. Numerical results suggested that the large number of component MMPPs gives the tight upper bound of the mean waiting time.

As for the tail distribution of the queue length, our proposed fitting does not work accurately. However, as is the case of the mean waiting time, the numerical results showed that it gives the upper bounds regardless of the traffic intensity. In particular, the values under heavy load tend to be tight. Therefore, our fitting method is useful for evaluating the tail distribution under heavy traffic situation.

As we stated in the second remark of section 5, changing the parameters of resultant IPPs into SPPs will provide the further improvement of the bounds for performance measures. This modification can be performed not only with keeping first- and second-order properties unchanged, but also by matching the autocorrelation of each SPP with the observed process. Thus it is expected to refine the tightness of bounds for performance measures using the modified method based on the SPPs.

Since our fitting method is not robust in terms of choosing time-scale, it is not enough for investigating the impact of the time-scale on the queueing behavior, which is a growing subject on the practical use of the Markovian traffic model. Further research is needed to capture the effect of the time-scale using Markovian input model.

Acknowledgements

Appendix

A. Proof of (3)

Using $P^*(z,t)$, the variance of N_t is given by

$$\text{Var}(N_t) = \boldsymbol{\pi} \left. \frac{\partial^2 P^*(z,t)}{\partial z^2}\right|_{z=1} \boldsymbol{e} + \text{E}(N_t) - \{\text{E}(N_t)\}^2. \tag{23}$$

To evaluate the first term of (23), 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 $[sI - Q]e = se$, we have $\boldsymbol{e}/s = [sI - Q]^{-1}\boldsymbol{e}$. Using these, we obtain

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

$$= \frac{2}{\sigma_1 + \sigma_2}\left[\frac{A_1}{s + \sigma_1 + \sigma_2} + \frac{A_2}{s^3} + \frac{A_3}{s^2} + \frac{A_4}{s}\right], \qquad (24)$$

where

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

Inverting (24) yields

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

Substituting above into (23), we obtain (3).

B. Proofs of Theorems in Section 3

In this appendix, we summarize the proofs of Theorems 3 and 4 in Section 3. For details, the readers are referred to [19].

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

$$\mathrm{Var}(X^{(m)}) = \frac{\sigma^2}{m} + \frac{2}{m^2}\sum_{k=1}^{m}(m-k)\mathrm{Cov}(X_t, X_{t+k}), \qquad (25)$$

$$\mathrm{Cov}(X_t, X_{t+k}) = \frac{1}{2}\delta^2(k^2\mathrm{Var}(X^{(k)})). \qquad (26)$$

Proof of Theorem 3. If X satisfies (4), then from (25),

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

Conversely, assume that X satisfies (7), then from (26),

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

Hence the theorem follows. $\qquad\square$

Proof of Theorem 4. Assume that X satisfies (4). Now, we consider the averaged process $X' = X^{(m)}$. Applying (26) to X' yields

$$
\mathrm{Cov}(X'_t, X'_{t+k}) = \frac{1}{2} \delta^2(k^2 \mathrm{Var}(X'^{(k)})). \tag{27}
$$

Note that

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

Then, from (27),

$$
\mathrm{Cov}(X_t^{(m)}, X_{t+k}^{(m)}) = \frac{1}{2} \delta^2(k^2 \mathrm{Var}(X^{(km)})). \tag{28}
$$

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

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

$\qquad\square$

References

[1] A. T. Anderson and B. F. Nielsen, "A Markovian Approach for Modeling Packet Traffic with Long-Range Dependence," *IEEE Journal on Selected Areas in Communications*, vol. 16, no. 5, pp.719-732, June 1998.

[2] A. T. Anderson and B. F. Nielsen, "An Application of Superposition of Two State Markovian Source to the Modeling of Self-Similar Behavior," *Proc. INFOCOM'97*, Kobe, Japan, pp.196-204, 1997.

[3] 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.

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

[5] 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.

[6] A. Feldmann and W. Whitt, "Fitting Mixtures of Exponentials to Long-Tail Distributions to Analyze Network Performance Models," *Proc. INFOCOM'97*, Kobe, Japan, pp.1096-1104, 1997.

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

[8] 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.

[9] D. P. Heyman and T. V. Lakshman, "What Are the Implications of Long-Range Dependence for VBR-Video Traffic Engineering?," *IEEE/ACM Transactions on Networking*, vol. 4, no. 3, pp.301-317, June 1996.

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

[11] 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.

[12] 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.

[13] 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.

[14] 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.

[15] J. M. Ortega and W. C. Rheinboldt, *Iterative Solution of Nonlinear Equations in Several Variables*, Academic Press, 1970.

[16] S. Robert and J. Y. Le Boudec, "New Models for Pseudo Self-Similar Traffic," *Performance Evaluation*, vol. 30, pp.57-68, 1997.

[17] B. K. Ryu and A. Elwalid, "The Importance of Long-Range Dependence of VBR Video Traffic in ATM Traffic Engineering: Myths and Realities," *ACM SiGCOMM'96*, pp.3-14, 1996.

[18] B. Ryu and S. B. Lowen, "Point Process Models for Self-Similar Network Traffic, with Applications," *Commun. Statist.–Stochastic Models*, vol. 14, no. 3, pp.735-761, 1998.

[19] 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.