INSTITUTE OF POLICY AND PLANNING SCIENCES

Discussion Paper Series

No. 958

A Queue with semi-Markovian Batch plus Poissom Arrivals with Application to the MPEG Frame Sequence

by

Dean Wu and Hideaki Takagi

November 2001

UNIVERSITY OF TSUKUBA
Tsukuba, Ibaraki 305-8573
JAPAN

A Queue with semi-Markovian Batch plus Poisson Arrivals with Application to the MPEG Frame Sequence

Dean Wu
Doctoral Program in Policy and Planning Sciences, University of Tsukuba
1-1-1 Tennoudai, Tsukuba-shi, Ibaraki 305-8573, Japan

and

Hideaki Takagi (Corresponding author)
Institute of Policy and Planning Sciences, University of Tsukuba
1-1-1 Tennoudai, Tsukuba-shi, Ibaraki 305-8573, Japan
phone: +81-298-53-5003; fax: +81-298-55-3849
e-mail: takagi@shako.sk.tsukuba.ac.jp

Abstract

We consider a queueing system with a single server having a mixture of a semi-Markov process (SMP) and a Poisson process as the arrival process, where each SMP arrival contains a batch of customers. The service times are exponentially distributed. We derive the distributions of the queue length and the waiting times of both SMP and Poisson customers. The results are applied to the case in which the SMP arrivals correspond to the exact sequence of Motion Picture Experts Group (MPEG) frames. Poisson arrivals are regarded as interfering traffic. In the numerical examples, the mean and variance of the waiting time of the ATM cells generated from the MPEG frames of real video data are evaluated.

Key words: Semi-Markov process; Batch arrival; Queue; Waiting time; MPEG; Group of pictures (GOP); Rouché's theorem

1 Introduction

Çinlar [1] considered a queueing system with a single server with semi-Markovian arrivals. There are a finite number of types of customers, and the types of successively arriving customers form a Markov chain. Further, the nth interarrival time has a distribution function which may depend on the types of both nth and $n-1$st arrivals. Separately, Kuczura [6] analyzed a GI + M/M/1 queue in which the arrival process is a mixture of a renewal process and a Poisson process. His analysis has been extended by Yagyu and Takagi [13] to an SSMP$^{[X]}$ + M/M/1 queue, where the SSMP is a special semi-Markov process in that the nth interarrival time distribution depends only on the type of $n-1$st arrival, and each SSMP arrival contains a batch of customers. The result has been applied to the Motion Picture Experts Group (MPEG) frame sequence as the SSMP$^{[X]}$ arrival process. The Markov chain underlying the SSMP has three states corresponding to the I-, B-, and P-frames. The state transition probabilities are determined in proportion to the frequency of appearance of these frames in a Group of Pictures (GOP). The effects

of interfering Poisson traffic on the waiting time of the ATM cells generated from the MPEG frames of some real video data have been evaluated.

In this paper, we first consider an $SMP^{[X]}$ + M/M/1 queue by extending the analysis of an $SSMP^{[X]}$ + M/M/1 queue in Yagyu and Takagi [13]. The semi-Markov arrival process is the same as that of Çinlar [1]. We derive the distributions of the queue length and the waiting times of both SMP and Poisson customers. The results are then applied to the case in which the SMP arrivals correspond to the exact MPEG sequence of frames, namely "IBBPBBPBBPBB," in a GOP. Thus the Markov chain underlying the SMP has twelve states with cyclic transitions. Poisson arrivals are regarded as interfering traffic. Taking the numerical data for the sizes of the three types of MPEG frames from the Jurassic Park video, we evaluate the mean and variance of the waiting time of the ATM cells generated from the frames.

2 SMP Batch Arrival Process

Consider an arrival process with L types of customers, each type arriving in batches of random size, and the state of the system is determined by the type of arriving customers. We say that the system enters state l when a batch of type l arrives. Let $g_l(k)$ denote the probability of batch size being k for type l customers, $l = 1, ..., L$. Suppose that the arrival points of each type of customers are Markov points by which the system passes through L states with transition probability matrix $\mathbf{P} = (p_{lm})$, $l, m = 1, ..., L$. The nth interarrival time may depend on the types of nth and $n-1$st arrivals. Let $A_{lm}(t)$ be the distribution function of interarrival time in state l, given that the next state is m. For a given sequence of arrival points, all interarrival times are mutually independent. This arrival process is referred to as a semi-Markov batch arrival process ($SMP^{[X]}$).

Clearly, the probability that SMP moves from state l to state m in time t is given by $p_{lm}A_{lm}(t)$. Since $\mathbf{P}$ is a stochastic matrix, we have

$$\sum_{m=1}^{L} p_{lm} = 1; \quad l = 1, \ldots, L.$$

Let $[\pi_1, \ldots, \pi_L]$ be the stationary distribution of the Markov chain with transition probability matrix $\mathbf{P} = (p_{lm})$. Then we have a set of the balance equations and the normalizing condition as follows:

$$\pi_m = \sum_{l=1}^{L} \pi_l p_{lm}; \quad m = 1, \ldots L \quad ; \quad \sum_{l=1}^{L} \pi_l = 1. \tag{1}$$

In Figure 1, we illustrate this semi-Markov arrival process, where A_{lm} represents the interarrival time between the arrivals of type l and type m customers. For convenience' sake, A_{lm} is also referred to as the sojourn time in state l when the next state is m in this paper. Note that there are two characteristics for this arrival process: (i) the state of the underlying Markov chain is determined by the type of arriving customers, and (ii) the interarrival time depends on both the current and next states of the Markov chain.

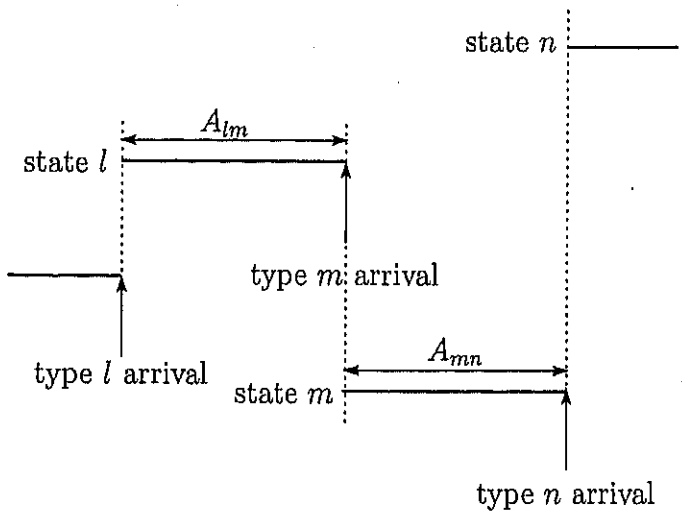

Figure 1: Semi-Markov arrival process.

3 Queue Length in SMP$^{[X]}$+M/M/1

In an SMP$^{[X]}$ + M/M/1 queueing system, the arrival process is a mixture of an SMP and a Poisson process. The arrival rate from the Poisson process is denoted by λ. The service times for the SMP and Poisson customers are assumed to have common exponential distribution with mean $1/\mu$. Finally, it has a single server and an infinite-capacity waiting room.

We analyze the queue length in the SMP$^{[X]}$ + M/M/1 system. The queue length $X(t)$ at time t is the number of both SMP and Poisson customers, including those waiting and in service, in the system at time t. We extend the approach proposed by Yagyu and Takagi [13] for an SSMP$^{[X]}$ + M/M/1 system in order to analyze our SMP$^{[X]}$ + M/M/1 system. Notice that, between the successive batch arrival epochs of SMP customers, the process $X(t)$ behaves exactly like the queue length in an M/M/1 system. By the method similar to the one in [13], we study the bivariate Markovian sequence $\{(X^{(n)}, S^{(n)}); n = 0, 1, 2, \dots\}$ embedded at the points of SMP arrivals, where $X^{(n)}$ denotes the number of both the SMP and Poisson customers found in the system by the first customer in the nth arriving batch of SMP customers, and $S^{(n)}$ denotes the state of the underlying Markov chain immediately *after* the nth SMP arrival (Figure 2).

Recall that the transition probability

$$P_{i,j}(t) := P\{X(t) = j | X(0) = i\}; \qquad t > 0$$

in the birth-and-death process for the queue length of an M/M/1 system with arrival rate λ and service rate μ is given by [9, p.93]

$$\begin{aligned} P_{i,j}(t) &= \rho^{\frac{1}{2}(j-i)} e^{-(\lambda+\mu)t} \left[I_{i-j}\left(2t\sqrt{\lambda\mu}\right) + \rho^{-\frac{1}{2}} I_{i+j+1}\left(2t\sqrt{\lambda\mu}\right) \right. \\ &+ \left. (1-\rho) \sum_{k=1}^{\infty} \rho^{-\frac{1}{2}(k+1)} I_{i+j+k+1}\left(2t\sqrt{\lambda\mu}\right) \right], \end{aligned} \tag{2}$$

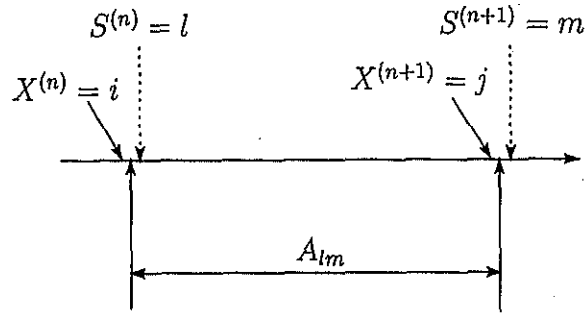

Figure 2: State transition in the Markov chain $\{(X^{(n)}, S^{(n)}); n = 0, 1, 2, \dots\}$.

where $\rho := \lambda/\mu$, and $I_i(t)$ is the modified Bessel function of the first kind of index i. For a nonnegative integer i, it is defined as

$$I_i(t) \equiv I_{-i}(t) := \left(\frac{t}{2}\right)^i \sum_{j=0}^{\infty} \frac{1}{j!(i+j)!} \left(\frac{t}{2}\right)^{2j}; \quad t \geq 0.$$

For the time-homogeneous Markov chain $\{(X^{(n)}, S^{(n)}); n = 0, 1, 2, \dots\}$, the state transition probability is given by

$$P\{X^{(n+1)} = j, S^{(n+1)} = m | X^{(n)} = i, S^{(n)} = l\} = p_{lm} \sum_{k=1}^{\infty} g_l(k) \int_0^{\infty} P_{i+k,j}(t) dA_{lm}(t)$$
$$i, j = 0, 1, 2, \dots; \; l, m = 1, \dots, L. \quad (3)$$

Assuming that this Markov chain is ergodic, the limiting distribution

$$P(i, l) := \lim_{n \to \infty} P\{X^{(n)} = i, S^{(n)} = l\}; \quad i = 0, 1, 2, \dots; \; l = 1, \dots, L \quad (4)$$

satisfies the balance equations

$$P(j, m) = \sum_{i=0}^{\infty} \sum_{l=1}^{L} \sum_{k=1}^{\infty} p_{lm} g_l(k) P(i, l) \int_0^{\infty} P_{i+k,j}(t) dA_{lm}(t); \quad j = 0, 1, 2, \dots; m = 1, \dots, L \quad (5)$$

and the normalization condition

$$\sum_{i=0}^{\infty} \sum_{l=1}^{L} P(i, l) = 1. \quad (6)$$

Let us introduce the generating function for $\{P(i, l); i = 0, 1, 2, \dots\}$ by

$$\Phi_l(z) := \sum_{i=0}^{\infty} P(i, l) z^i; \quad l = 1, \dots, L.$$

By definition, we must have

$$\Phi_l(1) = \pi_l; \quad l = 1, \ldots, L. \tag{7}$$

Multiplying (5) by z^j and summing over $j = 0, 1, 2, \ldots$, we obtain

$$\Phi_m(z) = \sum_{i=0}^{\infty} \sum_{l=1}^{L} \sum_{k=1}^{\infty} p_{lm} g_l(k) P(i,l) \int_0^{\infty} \Gamma_{i+k}(z,t) dA_{lm}(t); \quad m = 1, \ldots, L, \tag{8}$$

where

$$\Gamma_i(z,t) := \sum_{j=0}^{\infty} P_{i,j}(t) z^j; \quad i = 0, 1, 2, \ldots .$$

While this function is not simple, its Laplace transform is given by [9, p.89]

$$\gamma_i(z,s) := \int_0^{\infty} e^{-st} \Gamma_i(z,t) dt = \frac{z^{i+1} - (1-z)[\eta(s)]^{i+1}/[1-\eta(s)]}{zs - (1-z)(\mu - \lambda z)}, \tag{9}$$

where

$$\eta(s) := \frac{\lambda + \mu + s - \sqrt{(\lambda + \mu + s)^2 - 4\lambda\mu}}{2\lambda}.$$

Let us transform the real integral

$$\int_0^{\infty} \Gamma_{i+k}(z,t) dA_{lm}(t)$$

appearing in (8) into a complex integral involving $\gamma_{i+k}(z,s)$ and $\alpha_{lm}(s)$, the Laplace-Stieltjes transform (LST) of $A_{lm}(t)$. To do so, note the inverse transform

$$\Gamma_{i+k}(z,t) = \frac{1}{2\pi \mathbf{i}} \int_{c-\mathbf{i}\infty}^{c+\mathbf{i}\infty} e^{st} \gamma_{i+k}(z,s) ds,$$

where $c > 0$, $\mathbf{i} := \sqrt{-1}$, and the integration path $\int_{c-\mathbf{i}\infty}^{c+\mathbf{i}\infty}$ is the *Bromwich integral*, being written as $\int_{Br}$ hereafter. Furthermore, if $\alpha_{lm}(t)$ denotes the LST of $A_{lm}(t)$, we have

$$\int_0^{\infty} e^{st} dA_{lm}(t) = \alpha_{lm}(-s).$$

Thus we get

$$\int_0^{\infty} \Gamma_{i+k}(z,t) dA_{lm}(t) = \frac{1}{2\pi \mathbf{i}} \int_{Br} \gamma_{i+k}(z,s) \alpha_{lm}(-s) ds. \tag{10}$$

Substituting (10) into (8), we obtain

$$\Phi_m(z) = \sum_{l=1}^{L} p_{lm} \sum_{k=1}^{\infty} g_l(k) \sum_{i=0}^{\infty} P(i,l) \frac{1}{2\pi \mathbf{i}} \int_{Br} \gamma_{i+k}(z,s) \alpha_{lm}(-s) ds. \tag{11}$$

Changing the order of summation and integration, we get the following set of simultaneous equations for $\{\Phi_l(z); l = 1, \ldots, L\}$:

$$\Phi_m(z) = \sum_{l=1}^{L} p_{lm} \frac{1}{2\pi \mathbf{i}} \int_{Br} \left[\frac{z\Phi_l(z)G_l(z) - (1-z)H_l(s)}{zs - (1-z)(\mu - \lambda z)} \right] \alpha_{lm}(-s)ds; \quad m = 1, \ldots, L, \tag{12}$$

where

$$H_l(s) := \frac{\eta(s)G_l[\eta(s)]\Phi_l[\eta(s)]}{1-\eta(s)}, \qquad l = 1, \ldots, L. \tag{13}$$

Note that letting $z = 1$ in (12) recovers (1), because

$$\frac{1}{2\pi \mathbf{i}} \int_{Br} \frac{\alpha_{lm}(-s)}{s} ds = 1.$$

Following Kuczura [6], we may comment on the Bromwich integral in (12) as follows. Since $P_{i+k,j}(t)$ is the probability, its generating function $\Gamma_{i+k}(z,t)$ is uniformly convergent for $|z| \leq 1$, and $\gamma_{i+k}(z,s)$ is analytic for $|z| \leq 1$ and $\Re(s) > 0$. Hence the bracketed part of the integrand in (12) is analytic for $|z| \leq 1$ and $\Re(s) > 0$, since it is the convergent series of $\sum_{i=0}^{\infty} \sum_{k=1}^{\infty} P(i,l) g_l(k) \gamma_{i+k}(z,s)$. On the other hand, since $A_{lm}(t)$ is the distribution function, $\alpha_{lm}(s)$ is analytic for $\Re(s) > 0$. For $\Re(s) < 0$, $\alpha_{lm}(s)$ may or may not be analytic. However, $\alpha_{lm}(s)$ is meromorphic for $\Re(s) < 0$ in many cases, including the cases in which the distribution of A_{lm} is exponential, Erlang, and a linear combination thereof.

If we assume that $\alpha_{lm}(s)$ is meromorphic for the left-half plane $\Re(s) < 0$, all the poles of $\alpha_{lm}(-s)$ are in the right-half plane $\Re(s) > 0$. Hence the integrand in (12) is meromorphic in the right-half plane. Thus we can use the residue theorem to evaluate the integrand over the contour consisting of the line $(c+\mathbf{i}R, c-\mathbf{i}R)$ and a semicircle of radius R in the right-half plane which connects $c-\mathbf{i}R$ with $c+\mathbf{i}R$ counterclockwise. We can choose c and R such that all the poles of $\alpha_{lm}(-s)$ are interior to this contour for all $l = 1, \ldots, L$. Then the Bromwich integrals in (12) are evaluated only at the poles of $\alpha_{lm}(-s)$'s. Here the terms resulting from $H_l(s)$ are simply constants. Therefore, (12) is not a set of integral equations but simply a set of linear equations for $\{\Phi_l(z); l = 1, \ldots, L\}$ albeit containing unknown constants as coefficients. These unknown constants are determined from the condition that the generating function $\Phi_l(z)$ is analytic for $|z| \leq 1$ and that $\Phi_l(1) = \pi_l$ for $l = 1, \ldots, L$.

The marginal distribution for the number of SMP and Poisson customers found in the system by the first customer of an arriving SMP batch is denoted by

$$P(i) := \lim_{n \to \infty} P\{X^{(n)} = i\} = \sum_{l=1}^{L} P(i,l); \quad i = 0, 1, 2, \ldots . \tag{14}$$

The generating function for $\{P(i); i = 0, 1, 2, \ldots\}$ is then given by

$$\Phi(z) := \sum_{i=0}^{\infty} P(i) z^i = \sum_{m=1}^{L} \Phi_m(z). \tag{15}$$

Substituting (12) into (15) and rearranging terms yields

$$\Phi(z) = \sum_{m=1}^{L} \frac{1}{2\pi \mathrm{i}} \int_{Br} \left[\frac{z\Phi_l(z)G_l(z) - (1-z)H_l(s)}{zs - (1-z)(\mu - \lambda z)} \right] \alpha_l(-s)ds.$$

Here, we define

$$A_l(t) := \sum_{m=1}^{L} p_{lm} A_{lm}(t),$$

which is the distribution function of the sojourn time in state l, whose LST is given by $\alpha_l(s)$.

In particular, if the sojourn time A_{lm} follows an exponential distribution with mean $1/\alpha_{lm}$, equation (12) is free from the Bromwich integral, and it is reduced to

$$\Phi_m(z) = \sum_{l=1}^{L} \frac{p_{lm}}{q_{lm}(z)} [z\Phi_l(z)G_l(z) - (1-z)H_{lm}]; \quad m = 1, \ldots, L, \tag{16}$$

where

$$H_{lm} := \frac{\eta(\alpha_{lm})\Phi_l[\eta(\alpha_{lm})]G_l[\eta(\alpha_{lm})]}{1 - \eta(\alpha_{lm})}, \tag{17}$$

and

$$q_{lm}(z) := z - \frac{1}{\alpha_{lm}}(1-z)(\mu - \lambda z); \quad l, m = 1, \ldots, L. \tag{18}$$

Note that $\Phi_m(1)$ is equal to π_m, $m = 1, \ldots, L$, which is the stationary distribution of the underlying Markov chain. Therefore, we have the following set of balance equations and the normalizing condition:

$$\Phi_m(1) = \sum_{l=1}^{L} p_{lm}\Phi_l(1); \quad m = 1, \ldots, L,$$

$$\sum_{m=1}^{L} \Phi_m(1) = 1.$$

Now, equation (16) can be written in matrix form as

$$\mathbf{\Phi}(z)\mathbf{V}(z) = z\mathbf{\Phi}(z)\mathbf{G}(z)\mathbf{Q}(z) - (1-z)\mathbf{1}\mathrm{diag}[\mathbf{H}^{\mathrm{t}}\mathbf{Q}(z)], \tag{19}$$

where $\mathbf{\Phi}(z) := [\Phi_1(z), \ldots, \Phi_L(z)]$, $\mathbf{1} := [1, \ldots, 1]$,

$$\mathbf{G}(z) := \begin{bmatrix} G_1(z) & 0 & \ldots & 0 \\ 0 & G_2(z) & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & G_L(z) \end{bmatrix}, \tag{20}$$

$$\mathbf{V}(z) := \begin{bmatrix} \prod_{j=1}^{L} q_{j1}(z) & 0 & \dots & 0 \\ 0 & \prod_{j=1}^{L} q_{j2}(z) & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \prod_{j=1}^{L} q_{jL}(z) \end{bmatrix}, \tag{21}$$

$$\mathbf{Q}(z) := \begin{bmatrix} p_{11}\prod_{j\neq 1} q_{j1}(z) & p_{12}\prod_{j\neq 1} q_{j2}(z) & \dots & p_{1L}\prod_{j\neq 1} q_{jL}(z) \\ p_{21}\prod_{j\neq 2} q_{j1}(z) & p_{22}\prod_{j\neq 2} q_{j2}(z) & \dots & p_{2L}\prod_{j\neq 2} q_{jL}(z) \\ \vdots & \vdots & \ddots & \vdots \\ p_{L1}\prod_{j\neq L} q_{j1}(z) & p_{L2}\prod_{j\neq L} q_{j2}(z) & \dots & p_{LL}\prod_{j\neq L} q_{jL}(z) \end{bmatrix}, \tag{22}$$

and

$$\mathbf{H}(z) := \begin{bmatrix} H_{11} & H_{12} & \dots & H_{1L} \\ H_{21} & H_{22} & \dots & H_{2L} \\ \vdots & \vdots & \ddots & \vdots \\ H_{L1} & H_{L2} & \dots & H_{LL} \end{bmatrix}. \tag{23}$$

In equation (19), diag$\mathbf{X}$ is a diagonal matrix whose elements are taken from the corresponding elements of $\mathbf{X}$, and $\mathbf{H}^t$ is the transpose of $\mathbf{H}$. We may write (19) as

$$\Phi(z)\mathbf{F}(z) = (z-1)1\text{diag}[\mathbf{H}^t\mathbf{Q}(z)], \tag{24}$$

where

$$\mathbf{F}(z) := \mathbf{V}(z) - z\mathbf{G}(z)\mathbf{Q}(z). \tag{25}$$

Let adj$\mathbf{F}(z)$ denote the adjoint matrix of $\mathbf{F}(z)$. Multiplying (24) on the right by adj$\mathbf{F}(z)$, we have

$$\Phi(z) = \frac{(z-1)1\text{diag}[\mathbf{H}^t\mathbf{Q}(z)]\text{adj}\mathbf{F}(z)}{\det \mathbf{F}(z)}. \tag{26}$$

It is shown in Appendix 1 that there are L^2 zeros for $\det \mathbf{F}(z)$ in the unit disk $|z| \leq 1$ if the condition

$$\alpha g + \lambda < \mu \tag{27}$$

is satisfied. Here

$$\alpha := \frac{1}{\sum_{l=1}^{L} \pi_l \sum_{m=1}^{L} \frac{p_{lm}}{\alpha_{lm}}} \tag{28}$$

is the arrival rate of the batches of SMP customers, and

$$g := \sum_{l=1}^{L} \pi_l g_l \tag{29}$$

is the average batch size. The condition in (27) means that the sum of the arrival rates of SMP and Poisson customers is less than the service rate. Therefore, it is a sufficient condition for the stability of our system. Thus the set of L^2 unknown parameters $\{H_{lm}; l, m = 1, 2, \ldots, L\}$ can be determined by solving the same number of linear equations corresponding to the zeros of $\det \mathbf{F}(z)$ in $|z| \leq 1$.

4 Waiting Times in SMP$^{[X]}$+M/M/1

Let us investigate the waiting time for an arbitrary customer in an SMP$^{[X]}$+M/M/1 system. In section 4.1, the waiting time distribution for an arbitrary SMP customer in a batch is derived. In section 4.2, based on the theory of Markov renewal processes, the waiting time distribution for an arbitrary Poisson customer is given.

4.1 Waiting Time of SMP Customers

We first consider the waiting time W of an SMP customer. Let us focus on a randomly chosen *tagged* SMP customer included in a batch that arrives to bring state l. Recall that the probability generating function for the number of customers placed before the tagged customer in this batch is given by [11, p.45]

$$\hat{G}_l(z) = \frac{1 - G_l(z)}{g_l(1 - z)}, \tag{30}$$

where g_l is the mean batch size. Thus the LST $D_l(s)$ of the distribution function for the sum of the service times for those customers before the tagged customer in the batch is given by

$$D_l(s) = \hat{G}_l[B(s)] = \frac{1 - G_l[B(s)]}{g_l[1 - B(s)]}, \tag{31}$$

where $B(s) := \mu/(s + \mu)$.

If the service is given in the order of arrival, the waiting time of an arbitrary SMP customer (tagged) in a batch consists of the waiting time of the first customer of that batch and the service times for the customers placed before the tagged customer in

the batch. Therefore, the LST of the distribution function for the waiting time of an arbitrary SMP customer included in a batch that brings state l is given by

$$\Phi_l[B(s)]D_l(s).$$

Finally we get the LST $\Omega(s)$ of the distribution function for the waiting time W of an arbitrary SMP customer as

$$\Omega(s) = \frac{1}{g}\sum_{l=1}^{L} g_l\Phi_l[B(s)]D_l(s) = \frac{1}{g[1-B(s)]}\sum_{l=1}^{L}\Phi_l[B(s)]\{1-G_l[B(s)]\}, \tag{32}$$

where

$$g := \sum_{l=1}^{L}\pi_l g_l$$

is the overall mean batch size. The mean $E[W]$ and the second moment $E[W^2]$ of the waiting time are then given by

$$E[W] = \frac{1}{g\mu}\left(\sum_{l=1}^{L} E_l[X]g_l + \frac{g^{(2)}}{2}\right), \tag{33}$$

$$E[W^2] = \frac{1}{g\mu^2}\left(\sum_{l=1}^{L}\left\{(E_l[X]+E_l[X^2])g_l + E_l[X]g_l^{(2)}\right\} + g^{(2)} + \frac{g^{(3)}}{3}\right), \tag{34}$$

where

$$g_l^{(i)} = G_l^{(i)}(1), \quad g^{(i)} = \sum_{l=1}^{L}\pi_l g_l^{(i)}; \quad i = 2,3,$$

$$E_l[X] = \Phi_l^{(1)}(1), \quad E_l[X^2] = \Phi_l^{(2)}(1) + E_l[X]; \quad l = 1,\ldots,L.$$

4.2 Waiting Time of Poisson Customers

We proceed to consider the waiting time W^* of a Poisson customer. According to the PASTA (Poisson arrivals see time averages) property, the number of customers that an arriving Poisson customer finds in the system has the same distribution as the number X^* of customers present in the system at an arbitrary time in steady state. Thus we will find the generating function $\Phi^*(z)$ for the probability distribution of X^*.

To do so, note that the interval between an arbitrary time and the preceding SMP arrival time corresponds to the backward recurrence time in the Markov renewal process that counts the number of state transitions in the SMP. The joint distribution for the

backward recurrence time in state l and the probability that the next state is m is given by

$$\hat{A}_{lm}(t) = \frac{p_{lm}}{E[A_l]} \int_0^t [1 - A_{lm}(x)]dx; \quad t \geq 0, \tag{35}$$

where

$$E[A_l] := \sum_{m=1}^{L} p_{lm} E[A_{lm}]$$

is the mean sojourn time in state l.

Conditioning on the number of customers and the states of the SMP at the preceding and the next arrival points, and integrating with the backward recurrence time distribution in (35), the steady-state distribution of X^* is given by

$$P(X^* = j) = \sum_{i=0}^{\infty} \sum_{l=1}^{L} P(i,l) \sum_{k=1}^{\infty} g_l(k) \int_0^{\infty} P_{i+k,j}(t) d\hat{A}_l(t); \quad j = 0, 1, 2, \ldots, \tag{36}$$

where

$$\hat{A}_l(t) := \frac{E[A_l]}{E[A]} \sum_{m=1}^{L} \hat{A}_{lm}(t) = \frac{1}{E[A]} \sum_{m=1}^{L} p_{lm} \int_0^t [1 - A_{lm}(x)]dx; \quad t \geq 0$$

is the conditional distribution function for the backward recurrence time in state l. The mean interarrival time $E[A]$ between the batches of SMP customers is given by

$$E[A] := \sum_{l=1}^{L} \pi_l E[A_l].$$

From (36), the generating function $\Phi^*(z)$ for X^* is given by

$$\Phi^*(z) := \sum_{j=0}^{\infty} P(X^* = j) z^j = \sum_{i=0}^{\infty} \sum_{l=1}^{L} P(i,l) \sum_{k=1}^{\infty} g_l(k) \int_0^{\infty} \Gamma_{i+k}(z,t) d\hat{A}_l(t). \tag{37}$$

Using the relation similar to (10), we obtain

$$\Phi^*(z) = \sum_{l=1}^{L} \frac{1}{2\pi \mathrm{i}} \int_{Br} \left[\frac{z\Phi_l(z)G_l(z) - (1-z)H_l(s)}{zs - (1-z)(\mu - \lambda z)} \right] \hat{\alpha}_l(-s) ds, \tag{38}$$

where $H_l(s)$ is given in (13), and $\hat{\alpha}_l(s)$ is the LST of $\hat{A}_l(t)$. Again, the Bromwich integrals are evaluated only at the poles of $\hat{\alpha}_l(-s)$'s in the right-half plane $\Re(s) > 0$ in most cases.

The LST $\Omega^*(s)$ of the distribution function for the waiting time W^* of an arbitrary Poisson customer is expressed as

$$\Omega^*(s) = \Phi^*[B(s)]. \tag{39}$$

The mean $E[W^*]$ and the second moment $E[(W^*)^2]$ are then given by

$$E[W^*] = \frac{1}{\mu} E[X^*], \quad E[(W^*)^2] = \frac{E[X^*] + E[(X^*)^2]}{\mu^2}, \tag{40}$$

respectively, where $E[X^*]$ and $E[(X^*)^2]$ are obtained from $\Phi^*(z)$.

5 Application to the MPEG Frame Sequence

Let us use the $SMP^{[X]}$ + M/M/1 system to model the traffic in the ATM network in which the transmission of MPEG frames is interfered by other traffic. The waiting time of an arbitrary ATM cell generated from MPEG frames is studied. In Section 5.1, a brief description of MPEG coding scheme is given. In section 5.2, the transmission of MPEG frame sequence with interfering traffic is modeled by an $SMP^{[X]}$ + M/M/1 system. Assuming that the MPEG frame arrival process is also Poisson, we obtain the formula for evaluating the waiting time of an arbitrary ATM cell. In Section 5.3, some numerical results using the statistics of a real video film are presented.

5.1 MPEG Video Coding Scheme

In the MPEG coding [7], a video traffic is compressed using the following three types of frames.

- I-frames are generated independently of B- or P-frames and inserted periodically.
- P-frames are encoded for the motion compensation with respect to the previous I- or P-frame.
- B-frames are similar to P-frames, except that the motion compensation can be done with respect to the previous I- or P-frame, the next I- or P-frame, or the interpolation between them.

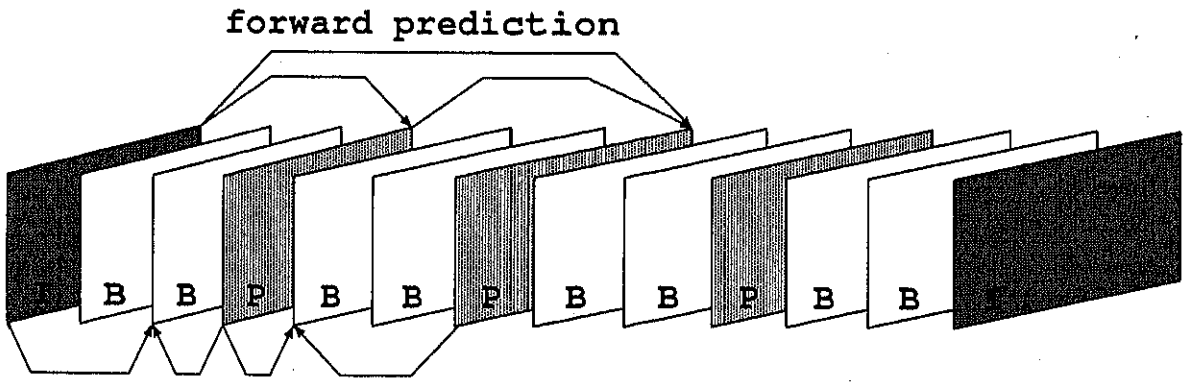

Figure 3: Group of pictures (GOP) of an MPEG stream [7].

These frames are arranged in a deterministic sequence "IBBPBBPBBPBB" as shown in Figure 3, which is called a Group of Pictures (GOP). The length of the GOP in Figure 3 is 12 frames. The traffic stream generated by the MPEG coding is characterized by two features, namely (i) the deterministic frame pattern in the GOP, and (ii) the distinguishable frame size distributions for the three types of frames (I, B and P).

5.2 Traffic Model for MPEG Frame Sequence

We are now in a position to apply the analysis results of an SMP$^{[X]}$ + M/M/1 system to the queueing model with MPEG frame sequence and interfering traffic. In this model, the Markov chain underlying the SMP has twelve states corresponding to the frame pattern "IBBPBBPBBPBB" in Figure 3. We index this sequence which represents the states in the Markov chain as 0 through 11. As shown in Figure 4, for any given state, the transition probability to the next state is unity, since the frame pattern is deterministic.

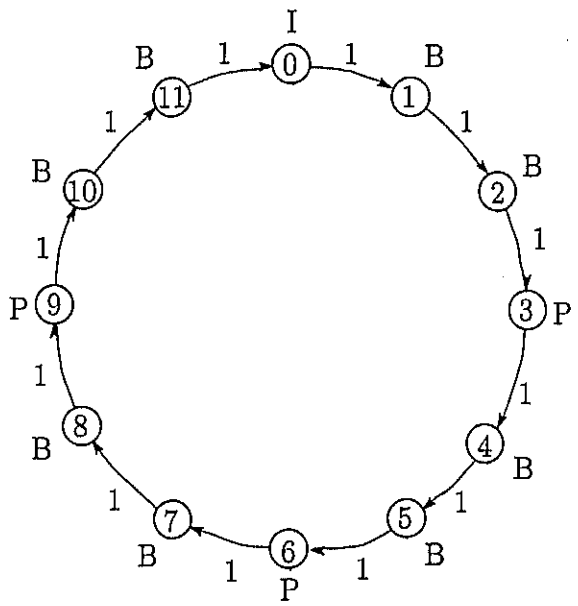

Figure 4: State transition diagram of the MPEG frame pattern.

The stationary distribution of this Markov chain is given by

$$\pi_l = \tfrac{1}{12} \quad ; \quad l = 0, \ldots, 11.$$

For the sake of simplicity in the expressions, we assume that the arrival process of the frames is Poisson with rate α as a (very) special case of the SMP. Let $G_l(z)$ denote the probability generating function for the number of ATM cells generated from the lth frame, $l = 0, \ldots, 11$. Equations in (12) become

$$\Phi_m(z) = \frac{1}{q(z)}[zG_{m-1}(z)\Phi_{m-1}(z) - (1-z)H_{m-1}]; \quad m = 0, \ldots, 11, \tag{41}$$

where

$$q(z) := z - \frac{1}{\alpha}(1-z)(\mu - \lambda z),$$

and H_m, $m = 0, \ldots, 11$, are constants to be determined. Hereafter state "$-m$" should

read state "$12-m$". Solving the set of equations in (41), we get

$$\Phi_m(z) = \frac{(z-1)\sum_{k=0}^{11} z^k [q(z)]^{11-k} H_{m-k-1} \prod_{l=m-k}^{m-1} G_l(z)}{T(z)}; \quad m = 0, \ldots, 11, \tag{42}$$

where

$$T(z) := [q(z)]^{12} - z^{12} \prod_{l=0}^{11} G_l(z). \tag{43}$$

Thus the following relations are established:

$$\Phi(z) := \sum_{m=0}^{11} \Phi_m(z) = \frac{(z-1)\sum_{k=0}^{11} z^k [q(z)]^{11-k} \sum_{j=0}^{11} H_j \sum_{l=j+1}^{j+k} G_l(z)}{T(z)}. \tag{44}$$

It is shown in Appendix 2 that there are twelve zeros of $T(z)$ in $|z| \leq 1$ under the condition

$$\alpha g + \lambda < \mu. \tag{45}$$

Here

$$g := \frac{1}{12} \sum_{l=0}^{11} g_l$$

is the mean size of an MPEG frame. Therefore, by using the twelve zeros of $T(z)$ in $|z| \leq 1$, we can solve the set of twelve linear equations for $\{H_m;\ m = 0, \ldots, 11\}$. This completes the determination of parameters in the model.

5.3 Numerical Examples

Let us evaluate the waiting time of an arbitrary ATM cell in the model with MPEG frame sequence and interfering traffic. The real video film data for the Jurassic Park (dino) is downloaded from the web site *http://nero.informatik.uni-wuerzburg.de/MPEG/* prepared by Rose [8]. We need to assume some distribution for the number of cells in each frame (frame size) so that we can calculate the value of waiting times numerically.

Frey and Nguen-Quang [2] and Sarkar et al. [10] propose the gamma distribution for the frame size. As a discrete version of the gamma distribution, let us assume that the distribution of the frame size is *negative binomial*. Thus the probability generating functions for the frame size are given by

$$G_l(z) = \left(\frac{p_l}{1 - q_l z}\right)^{n_l}; \quad q_l := 1 - p_l; \quad l = 0, \ldots, 11.$$

Table 1: Statistics for the frame size in ATM cells calculated from the MPEG traces for the Jurassic Park video.

I-frame			B-frame			P-frame		
mean	var	c.v.	mean	var	c.v.	mean	var	c.v.
143.4	918.7	0.211	19.0	135.0	0.612	37.7	632.6	0.667

Table 2: Parameters of the negative binomial distributions for the frame size of the Jurassic Park video.

I-frame		B-frame		P-frame	
n_{I}	p_{I}	n_{B}	p_{B}	n_{P}	p_{P}
26.534	0.156	3.123	0.141	2.384	0.060

where, referring to Figure 4, we set

$$\begin{aligned} p_l = p_{\mathrm{I}}, \quad n_l = n_{\mathrm{I}}; \quad & l = 0, \\ p_l = p_{\mathrm{B}}, \quad n_l = n_{\mathrm{B}}; \quad & l = 1, 2, 4, 5, 7, 8, 10, 11, \\ p_l = p_{\mathrm{P}}, \quad n_l = n_{\mathrm{P}}; \quad & l = 3, 6, 9. \end{aligned}$$

In Table 1 the statistics for the number of ATM cells in each frame type for the dino, which have been calculated by assuming that every frame is divided into a group of cells each with a payload of 48 bytes, are presented. The fitted parameters determined from the mean and variance of the actual data are given in Table 2. Figure 5 compares the histogram of the frame sizes with fitted negative binomial distribution.

Let us assume that cells are transmitted on a 10 Mbps channel, which corresponds to μ = 2,350 cells/sec. Substituting these parameters in (43), we have exactly twelve zeros in the unit disk. The zeros of $T(z)$ are plotted in the complex z-plane in Figure 6.

Figures 7 and 8 show the mean and the variance of the waiting times of an arbitrary ATM cell in the MPEG frames and an arbitrary Poisson arriving cell. It is observed that at low arrival rate α (frames/sec) the difference (for both the mean and variance) between SMP cell and Poisson cell is relatively large, while it becomes small as α increases. In other words, the influence of batch arrival is small when α is relatively large. It is also observed that MPEG cell always receive slightly worse treatment, i.e., bigger mean values of the waiting time, than Poisson arriving cell. This is because the s.c.v. of the interarrival times for the SMP arrival process is bigger than that of the Poisson arrival process which is unity. Kuczura [6] reports that the arrival process having bigger s.c.v. receives worse treatment than that with smaller s.c.v., which agrees with the present result.

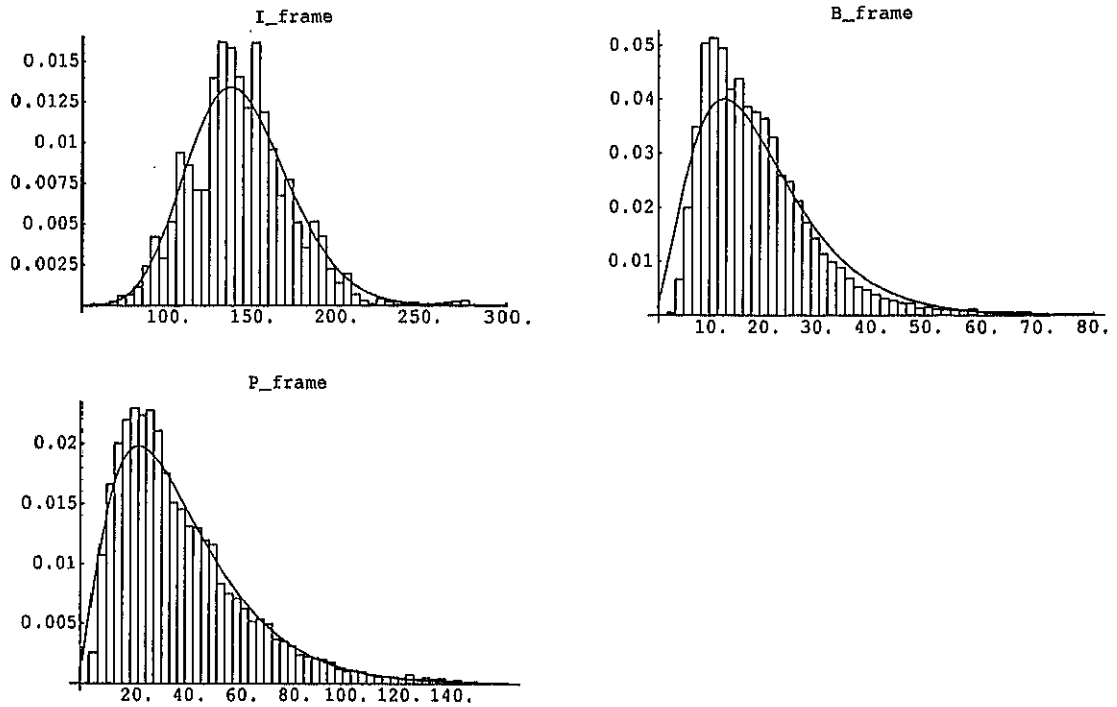

Figure 5: Histogram of the frame sizes and the fitted negative binominal distribution for the Jurassic Park video.

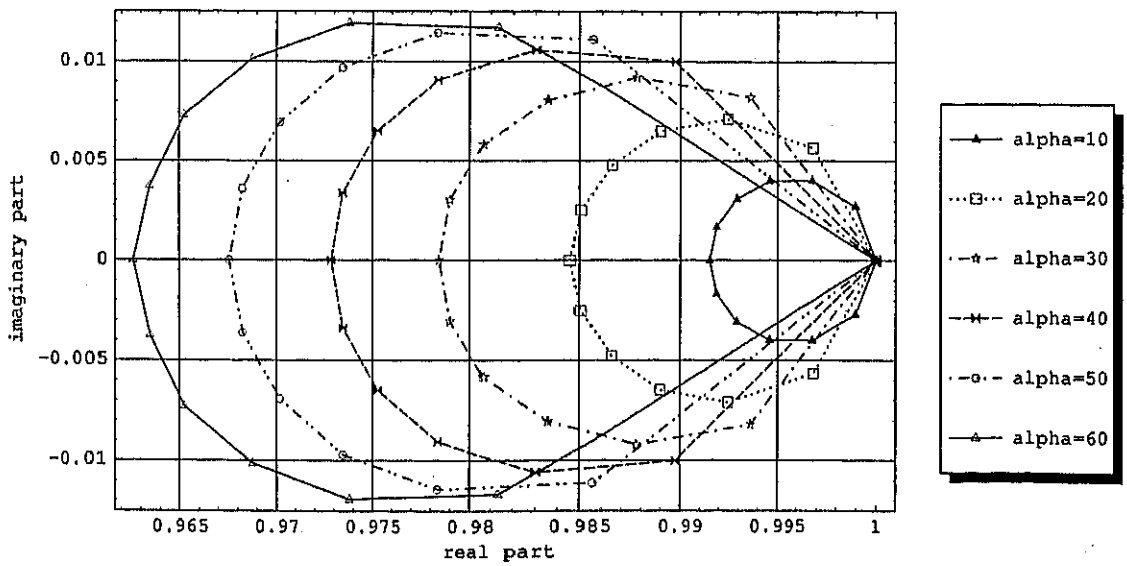

Figure 6: Zeros of $T(z)$ in the unit disk when $\lambda = 300$.

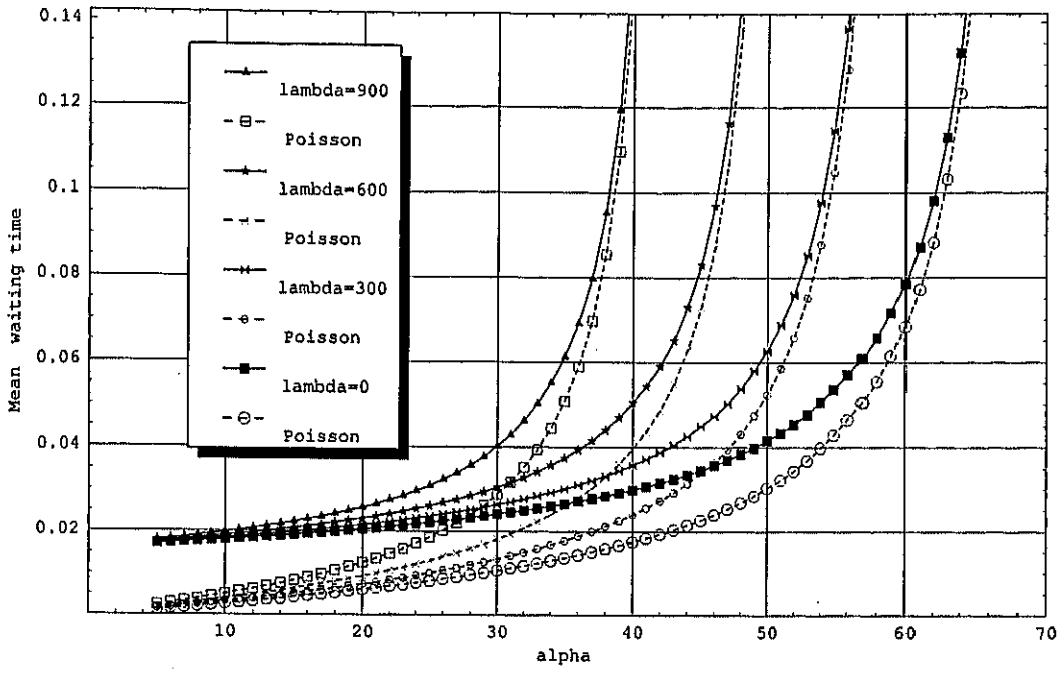

Figure 7: Mean waiting time for an arbitrary cell (Jurassic Park).

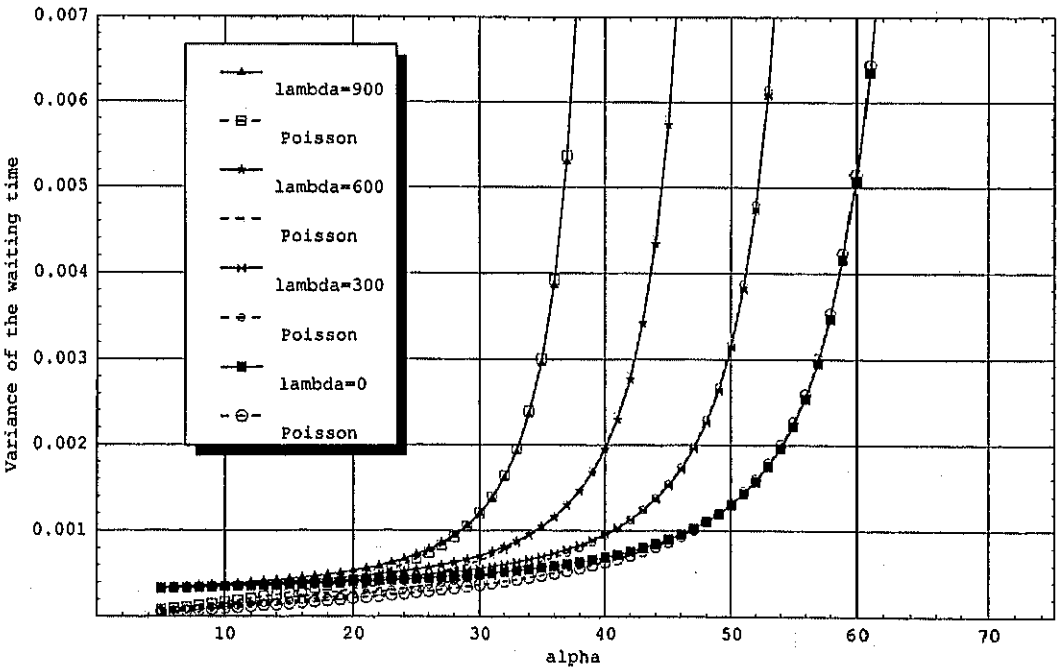

Figure 8: Variance of the waiting time for an arbitrary cell (Jurassic Park).

6 Summary

In this paper, we have first analyzed a queueing system having a mixture of an SMP in batch and a Poisson process as the arrival process, where the Poisson arrival is regarded as interfering traffic. Then we have modeled the arrival of the MPEG frame sequence as an SMP batch arrival process. This model captures two major features of the MPEG coding scheme: (i) the deterministic frame pattern and (ii) the distinct distributions for the size of the three types of frames. The waiting time of each ATM cell has been evaluated. It is observed that at low arrival rate of MPEG frames, the difference in the waiting times between the MPEG and Poisson cells is relatively large. It is also revealed that the MPEG cells receive slightly worse treatment than Poisson cells.

References

[1] E. Çinlar, "Queues with semi-Markovian arrivals," *Journal of Applied Probability*, Vol.4, No.2, pp.365-379, August 1967.

[2] M. Frey and S. Nguyen-Quang, "A gamma-based framework for modeling variable-rate MPEG video source: The GOP GBAR model," *IEEE/ACM Transactions on Networking*, Vol.8, No.6, pp. 710–719, December 2000.

[3] H. R. Gail, S. L. Hantler and B. A. Taylor, " Use of characteristic roots for solving infinite state Markov chains," *Computational Probability*, W. K. Grassmann (editor), pp.205–255, Kluwer Academic Publishers, 2000.

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

[5] F. Ishizaki, T. Takine, and T. Hasegawa, "Analysis of a discrete-time queue with gated priority," *Performance Evaluation*, Vol.23, No.2, pp. 121–143, August 1995.

[6] A. Kuczura, "Queues with mixed renewal and Poisson inputs," *The Bell System Technical Journal*, Vol.51, No.6, pp.1305–1326, July–August 1972.

[7] D. Le Gall, "MPEG: A video compression standard for multimedia applications," *Communications of the ACM*, Vol.34, No.4, pp.46–58, April 1991.

[8] O. Rose, "Simple and efficient models for variable bit rate MPEG video traffic," *Performance Evaluation*, Vol.30, No.1–2, pp.69–85, July 1997.

[9] T. L. Saaty, *Elements of Queueing Theory with Applications*, McGraw-Hill Book Company, New York, 1961. Republished by Dover Publications, New York, 1983.

[10] U. K. Sarkar, S. Ramakrishnan, and D. Sarkar, "Segmenting full-length VBR video into slots for modeling with Markov-modulated gamma-based framework," *Proceedings of the SPIE*, Vol.4519, 2001.

[11] H. Takagi, *Queueing Analysis: A Foundation of Performance Evaluation, Volume 1: Vacation and Priority Systems, Part 1*, Elsevier, 1991.

[12] E. C. Titchmarsh, *The Theory of Functions*, Second edition, Oxford University Press, London, 1939.

[13] S. Yagyu and H. Takagi, "A queueing model with input of MPEG fame sequence and interfering traffic (a revised version)," Discussion paper No.951, Institute of Policy and Planning Sciences, University of Tsukuba, September 2001.

Appendix 1: Number of Zeros of detF(z) in (26) in $|z| \leq 1$

We first derive the stability condition in (27), and then prove that there are L^2 zeros of $\det \mathbf{F}(z)$ in (26) in $|z| \leq 1$.

In Section 3 we have derived the relation

$$\mathbf{\Phi}(z)\mathbf{F}(z) = (z-1)\mathbf{1}\mathrm{diag}[\mathbf{H}^{\mathrm{t}}\mathbf{Q}(z)], \tag{A.1}$$

where

$$\mathbf{F}(z) := \mathbf{V}(z) - z\mathbf{G}(z)\mathbf{Q}(z). \tag{A.2}$$

These are equations (24) and (25), respectively.

Let us first derive the stability condition in (27). Differentiating (A.1) and evaluating the result at $z = 1$, we obtain

$$\mathbf{\Phi}'(1)(\mathbf{I}_L - \mathbf{P}) + \boldsymbol{\pi}\mathbf{F}'(1) = \mathbf{1}\mathrm{diag}[\mathbf{H}^{\mathrm{t}}\mathbf{P}], \tag{A.3}$$

where $\mathbf{I}_L$ denotes an $L \times L$ identity matrix. Here we have used $\mathbf{F}(1) = \mathbf{I}_L - \mathbf{P}$ since $\mathbf{V}(1) = \mathbf{G}(1) = \mathbf{I}_L$ and $\mathbf{Q}(1) = \mathbf{P}$. Note also that $\mathbf{\Phi}(1) = \boldsymbol{\pi} := [\pi_1, \ldots, \pi_L]$. Multiplying (A.3) on the right by $\mathbf{1}^{\mathrm{t}} := [1, \ldots, 1]^{\mathrm{t}}$ and noting that $(\mathbf{I}_L - \mathbf{P})\mathbf{1}^{\mathrm{t}} = 0$, we get

$$\boldsymbol{\pi}\mathbf{F}'(1)\mathbf{1}^{\mathrm{t}} = \mathbf{1}\mathrm{diag}[\mathbf{H}^{\mathrm{t}}\mathbf{P}]\mathbf{1}^{\mathrm{t}}. \tag{A.4}$$

To determine the right-hand side of this equation, we see from (23) that

$$\mathbf{1}\mathrm{diag}[\mathbf{H}^{\mathrm{t}}\mathbf{P}]\mathbf{1}^{\mathrm{t}} = \sum_{j=1}^{L}\sum_{k=1}^{L} H_{jk}p_{jk}. \tag{A.5}$$

To determine the left-hand side of (A.4), we differentiate (A.2) and evaluate the result at $z = 1$. Then we have

$$\begin{aligned} \mathbf{F}'(1) &= \mathbf{V}'(1) - \mathbf{G}(1)\mathbf{Q}(1) - \mathbf{G}'(1)\mathbf{Q}(1) - \mathbf{G}(1)\mathbf{Q}'(1) \\ &= \mathbf{V}'(1) - \mathbf{P} - \mathbf{G}'(1)\mathbf{P} - \mathbf{Q}'(1), \end{aligned} \tag{A.6}$$

where

$$\mathbf{V}'(1) = \begin{bmatrix} \displaystyle\sum_{j=1}^{L} \frac{\alpha_{j1}+\mu-\lambda}{\alpha_{j1}} & 0 & \cdots & 0 \\ 0 & \displaystyle\sum_{j=1}^{L} \frac{\alpha_{j2}+\mu-\lambda}{\alpha_{j2}} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \displaystyle\sum_{j=1}^{L} \frac{\alpha_{jL}+\mu-\lambda}{\alpha_{jL}} \end{bmatrix}, \tag{A.7}$$

$$\mathbf{G}'(1) = \begin{bmatrix} g_1 & 0 & \dots & 0 \\ 0 & g_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & g_L \end{bmatrix}, \tag{A.8}$$

and

$$\mathbf{Q}'(1) = \begin{bmatrix} p_{11}\sum_{j\neq 1}\frac{\alpha_{j1}+\mu-\lambda}{\alpha_{j1}} & p_{12}\sum_{j\neq 1}\frac{\alpha_{j2}+\mu-\lambda}{\alpha_{j2}} & \dots & p_{1L}\sum_{j\neq 1}\frac{\alpha_{jL}+\mu-\lambda}{\alpha_{jL}} \\ p_{21}\sum_{j\neq 2}\frac{\alpha_{j1}+\mu-\lambda}{\alpha_{j1}} & p_{22}\sum_{j\neq 2}\frac{\alpha_{j2}+\mu-\lambda}{\alpha_{j2}} & \dots & p_{2L}\sum_{j\neq 2}\frac{\alpha_{jL}+\mu-\lambda}{\alpha_{jL}} \\ \vdots & \vdots & \ddots & \vdots \\ p_{L1}\sum_{j\neq L}\frac{\alpha_{j1}+\mu-\lambda}{\alpha_{j1}} & p_{L2}\sum_{j\neq L}\frac{\alpha_{j2}+\mu-\lambda}{\alpha_{j2}} & \dots & p_{LL}\sum_{j\neq L}\frac{\alpha_{jL}+\mu-\lambda}{\alpha_{jL}} \end{bmatrix}, \tag{A.9}$$

Multiplying (A.6) on the right by 1^{t} and substituting (A.7), (A.8), and (A.9) yields

$$\begin{aligned} \mathbf{F}'(1)1^{\mathrm{t}} &= \mathbf{V}'(1)1^{\mathrm{t}} - 1^{\mathrm{t}} - \mathbf{G}'(1)1^{\mathrm{t}} - \mathbf{Q}'(1)1^{\mathrm{t}} \\ &= \begin{bmatrix} \sum_{j=1}^{L}\frac{\alpha_{j1}+\mu-\lambda}{\alpha_{j1}} \\ \sum_{j=1}^{L}\frac{\alpha_{j2}+\mu-\lambda}{\alpha_{j2}} \\ \vdots \\ \sum_{j=1}^{L}\frac{\alpha_{jL}+\mu-\lambda}{\alpha_{jL}} \end{bmatrix} - \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix} - \begin{bmatrix} g_1 \\ g_2 \\ \vdots \\ g_L \end{bmatrix} - \begin{bmatrix} \sum_{k=1}^{L} p_{1k}\sum_{j\neq 1}\frac{\alpha_{jk}+\mu-\lambda}{\alpha_{jk}} \\ \sum_{k=1}^{L} p_{2k}\sum_{j\neq 2}\frac{\alpha_{jk}+\mu-\lambda}{\alpha_{jk}} \\ \vdots \\ \sum_{k=1}^{L} p_{Lk}\sum_{j\neq L}\frac{\alpha_{jk}+\mu-\lambda}{\alpha_{jk}} \end{bmatrix}. \end{aligned} \tag{A.10}$$

Finally, multiplying (A.10) on the left by $\boldsymbol{\pi}$, we obtain

$$\begin{aligned} \boldsymbol{\pi}\mathbf{F}'(1)1^{\mathrm{t}} &= \sum_{l=1}^{L}\pi_l\sum_{k=1}^{L}\frac{\alpha_{kl}+\mu-\lambda}{\alpha_{kl}} - \sum_{l=1}^{L}\pi_l - \sum_{l=1}^{L}\pi_l g_l \\ &\quad - \sum_{l=1}^{L}\pi_l\sum_{k=1}^{L}p_{lk}\sum_{j\neq l}\frac{\alpha_{jk}+\mu-\lambda}{\alpha_{jk}} \\ &= \sum_{l=1}^{L}\pi_l\sum_{k=1}^{L}\frac{\alpha_{kl}+\mu-\lambda}{\alpha_{kl}} - 1 - g \\ &\quad - \sum_{l=1}^{L}\pi_l\sum_{k=1}^{L}p_{lk}\sum_{j=1}^{L}\frac{\alpha_{jk}+\mu-\lambda}{\alpha_{jk}} + \sum_{l=1}^{L}\pi_l\sum_{k=1}^{L}p_{lk}\frac{\alpha_{lk}+\mu-\lambda}{\alpha_{lk}}. \end{aligned}$$

However, from the relations $\sum_{l=1}^{L} \pi_l p_{lk} = \pi_k$, $k = 1, \ldots, L$, we have

$$\begin{aligned}\sum_{l=1}^{L} \pi_l \sum_{k=1}^{L} p_{lk} \sum_{j=1}^{L} \frac{\alpha_{jk} + \mu - \lambda}{\alpha_{jk}} &= \sum_{k=1}^{L} \sum_{j=1}^{L} \frac{\alpha_{jk} + \mu - \lambda}{\alpha_{jk}} \sum_{l=1}^{L} \pi_l p_{lk} \\ &= \sum_{k=1}^{L} \pi_k \sum_{j=1}^{L} \frac{\alpha_{jk} + \mu - \lambda}{\alpha_{jk}}.\end{aligned}$$

Thus we get

$$\boldsymbol{\pi}\mathbf{F}'(1)\mathbf{1}^{\mathrm{t}} = (\mu - \lambda) \sum_{l=1}^{L} \pi_l \sum_{k=1}^{L} \frac{p_{lk}}{\alpha_{lk}} - g = \frac{\mu - \lambda}{\alpha} - g. \tag{A.11}$$

Here α is the arrival rate of SMP batches defined in (28), and g is the average batch size given in (29). This is the left-hand side of (A.4). Thus we have

$$\frac{\mu - \lambda}{\alpha} - g = \sum_{j=1}^{L} \sum_{k=1}^{L} H_{jk} p_{jk}.$$

Since the right-hand side of this equation is positive, we must have

$$\alpha g + \lambda < \mu, \tag{A.12}$$

which is the condition in (27). □

Multiplying (A.1) on the right by $\mathrm{adj}\mathbf{F}(z)$, we have

$$\boldsymbol{\Phi}(z) = \frac{(z-1)\mathbf{1}\mathrm{diag}[\mathbf{H}^{\mathrm{t}}\mathbf{Q}(z)]\mathrm{adj}\mathbf{F}(z)}{\det \mathbf{F}(z)}, \tag{A.13}$$

which is (26). This is analytic in $|z| < 1$ and continuous in $|z| \leq 1$.

Recall that $\boldsymbol{\Phi}(1) = \boldsymbol{\pi}$. Since $\det \mathbf{F}(1) = \det[\mathbf{I}_L - \mathbf{P}] = 0$, the point $z = 1$ is the common zero of the denominator and the numerator for the right-hand side of (A.13). Thus we investigate the value of the derivative of $\det \mathbf{F}(z)$ at $z = 1$:

$$\gamma = \frac{d}{dz} \det \mathbf{F}(z) \Big|_{z=1}.$$

Theorem 1 *If $\alpha g + \lambda < \mu$, then $\gamma > 0$.*

Proof. To determine γ, we use the well-known relations in linear algebra:

$$\mathbf{F}(z)\mathrm{adj}\mathbf{F}(z) = \det \mathbf{F}(z)\mathbf{I}_L = \mathrm{adj}\mathbf{F}(z)\mathbf{F}(z). \tag{A.14}$$

Differentiating the second equality, evaluating the value at $z = 1$, and multiplying on the right by $\mathbf{1}^{\mathrm{t}}$, we obtain

$$\gamma \mathbf{1}^{\mathrm{t}} = \mathrm{adj}\mathbf{F}(1)\mathbf{F}'(1)\mathbf{1}^{\mathrm{t}}. \tag{A.15}$$

An expression for adj$\mathbf{F}(1)$ may be found as follows. Evaluating (A.14) at $z = 1$ and using $\det \mathbf{F}(1) = 0$, we have

$$\mathbf{P}\text{adj}\mathbf{F}(1) = \text{adj}\mathbf{F}(1) = \text{adj}\mathbf{F}(1)\mathbf{P}.$$

Since $\mathbf{P}$ is an irreducible stochastic matrix, the first equality implies that each column of adj$\mathbf{F}(1)$ is a multiple of 1^t (recall that $\mathbf{P}1^t = 1^t$). Similarly, the second equality implies that each row of adj$\mathbf{F}(1)$ is a multiple of $\boldsymbol{\pi}$ (recall that $\boldsymbol{\pi}\mathbf{P} = \boldsymbol{\pi}$). It follows that there is a constant c such that

$$\text{adj}\mathbf{F}(1) = c \begin{bmatrix} \boldsymbol{\pi} \\ \vdots \\ \boldsymbol{\pi} \end{bmatrix}. \tag{A.16}$$

We claim that adj$\mathbf{F}(1)$ is a positive matrix [4, p.359]. From the form of (A.16), it is enough to show that the diagonal elements, say, κ_l, $l = 1, \ldots, L$, of adj$\mathbf{F}(1)$ are positive. To see this, note that

$$\kappa_l = (-1)^{l+l} \det[\mathbf{F}_{(l,l)}(1)] = \det[\mathbf{I}_{L-1} - \mathbf{P}_{(l,l)}],$$

where $\mathbf{P}_{(l,l)}$ is the matrix $\mathbf{P}$ with its lth row and lth column removed. Since $\mathbf{P}$ is irreducible, the spectral radius of $\mathbf{P}_{(l,l)}$ is strictly less than unity. This implies that $\det[\mathbf{I}_{L-1} - t\mathbf{P}_{(l,l)}] \neq 0$ for real t satisfying $0 \leq t \leq 1$. Since this determinant function of t is positive for $t = 0$ and never zero, by continuity it is also positive for $t = 1$, i.e., $\kappa_l > 0$. Thus adj$\mathbf{F}(1)$ is positive, and we conclude that $c > 0$ in (A.16).

Substituting (A.16) into (A.15) and noting (A.11) yields

$$\gamma = c\Big(\frac{\mu - \lambda}{\alpha} - g\Big). \tag{A.17}$$

Using $c > 0$ and the condition (A.12), we see that γ is positive. □

We next show that there are L^2 zeros for $\det \mathbf{F}(z)$ in the unit disk. To do so, we use a lemma in [3, p.239]: *Let $f(z,t)$ be a function analytic for z within and on a closed contour C, and continuous for t in some interval $\mathcal{I}$. If $f(z,t) \neq 0$ for $z \in C$ and $t \in \mathcal{I}$, then the number of zeros of $f(z,t)$ inside C is the same for all $t \in \mathcal{I}$.*

For our purpose, let

$$f(z,t) := \det \mathbf{F}(z,t),$$

where

$$\mathbf{F}(z,t) := \mathbf{V}(z) - zt\mathbf{G}(z)\mathbf{Q}(z).$$

We choose a closed contour $C := \{z; |z| = 1\}$ and an interval $\mathcal{I} := \{t; t \in [0,1)\}$. Obviously, $f(z,t)$ is analytic in C and continuous for $t \in \mathcal{I}$. We first prove that $f(z,t) \neq 0$ for $z \in C$ and $t \in \mathcal{I}$, and then prove that there are L^2 zeros for $f(z,1) = \det \mathbf{F}(z)$ in C using the above lemma.

Theorem 2

(a) $\det \mathbf{F}(z,t) \neq 0$ *for* $|z| = 1$ *and* $t \in [0,1)$.

(b) $\det \mathbf{F}(z) \neq 0$ *for* $|z| = 1,\ z \neq 1$.

Proof. We consider $\det \mathbf{F}(z,t)$ for $|z| = 1$ and $t \in [0,1]$. Note that $\det \mathbf{F}(z) = \det \mathbf{F}(z,1)$. Then $\mathbf{F}(z,t)$ can be written as

$$\begin{aligned} \mathbf{F}(z,t) &= \mathbf{V}(z) - zt\mathbf{G}(z)\mathbf{Q}(z) \\ &= \mathbf{V}(z) - zt\mathbf{G}(z)\mathbf{L}(z)\mathbf{V}(z) \\ &= [\mathbf{I}_L - zt\mathbf{G}(z)\mathbf{L}(z)]\mathbf{V}(z), \end{aligned} \tag{A.18}$$

where

$$\mathbf{L}(z) := \begin{bmatrix} \dfrac{p_{11}}{q_{11}(z)} & \dfrac{p_{12}}{q_{12}(z)} & \cdots & \dfrac{p_{1L}}{q_{1L}(z)} \\ \dfrac{p_{21}}{q_{21}(z)} & \dfrac{p_{22}}{q_{22}(z)} & \cdots & \dfrac{p_{2L}}{q_{2L}(z)} \\ \vdots & \vdots & \ddots & \vdots \\ \dfrac{p_{L1}}{q_{L1}(z)} & \dfrac{p_{L2}}{q_{L2}(z)} & \cdots & \dfrac{p_{LL}}{q_{LL}(z)} \end{bmatrix}. \tag{A.19}$$

Therefore we have

$$\det \mathbf{F}(z,t) = \det[\mathbf{I}_L - zt\mathbf{G}(z)\mathbf{L}(z)] \cdot \det \mathbf{V}(z). \tag{A.20}$$

Since

$$|q_{jk}(z)| = \left| \frac{1}{\alpha_{jk}}[(\alpha_{jk} + \lambda + \mu)z - (\lambda z^2 + \mu)] \right| \geq \frac{1}{\alpha_{jk}}[\alpha_{jk} + \lambda + \mu - (\lambda + \mu)] = 1$$

for $|z| = 1$, we see that

$$|\det \mathbf{V}(z)| = \left| \prod_{k=1}^{L} \prod_{j=1}^{L} q_{jk}(z) \right| \geq 1, \quad \text{for } |z| = 1.$$

It follows that $\det \mathbf{V}(z) \neq 0$ for $|z| = 1$.

We next prove that $\mathbf{I}_L - zt\mathbf{G}(z)\mathbf{L}(z)$ is nonsingular for $|z| = 1$ and $t \in [0,1)$ and that $\mathbf{I}_L - z\mathbf{G}(z)\mathbf{L}(z)$ is nonsingular for $|z| = 1,\ z \neq 1$. These are equivalent to $\det[\mathbf{I}_L - zt\mathbf{G}(z)\mathbf{L}(z)] \neq 0$ and $\det[\mathbf{I}_L - z\mathbf{G}(z)\mathbf{L}(z)] \neq 0$, respectively. To do this, we use the notion of strictly diagonally dominant: *A square matrix* $X = (x_{ij})$ *is (row) strictly diagonally dominant if* $|x_{ii}| > \sum_{j \neq i} |x_{ij}|$ *for every row* i, and the Levy-Desplanques theorem: *A strictly diagonally dominant matrix is nonsingular* [4, p.349].

From (20) and (A.19) we have

$$\mathbf{I}_L - zt\mathbf{G}(z)\mathbf{L}(z)$$
$$= \begin{bmatrix} 1 - ztG_1(z)\frac{p_{11}}{q_{11}(z)} & -ztG_1(z)\frac{p_{12}}{q_{12}(z)} & \dots & -ztG_1(z)\frac{p_{1L}}{q_{1L}(z)} \\ -ztG_2(z)\frac{p_{21}}{q_{21}(z)} & 1 - ztG_2(z)\frac{p_{22}}{q_{22}(z)} & \dots & -ztG_2(z)\frac{p_{2L}}{q_{2L}(z)} \\ \vdots & \vdots & \ddots & \vdots \\ -ztG_L(z)\frac{p_{L1}}{q_{L1}(z)} & -ztG_L(z)\frac{p_{L2}}{q_{L2}(z)} & \dots & 1 - ztG_L(z)\frac{p_{LL}}{q_{LL}(z)} \end{bmatrix}. \quad \text{(A.21)}$$

For case (a) in which $|z| = 1$ and $t \in [0, 1)$, we see that

$$\sum_{j \neq i} \left| -ztG_i(z)\frac{p_{ij}}{q_{ij}(z)} \right| < \sum_{j \neq i} \left| zG_i(z)\frac{p_{ij}}{q_{ij}(z)} \right|$$
$$\leq \sum_{j \neq i} \frac{p_{ij}}{|q_{ij}(z)|} \leq \sum_{j \neq i} p_{ij} = 1 - p_{ii} \leq \left| 1 - ztG_i(z)\frac{p_{ii}}{q_{ii}(z)} \right|.$$

Thus $\mathbf{I}_L - zt\mathbf{G}(z)\mathbf{L}(z)$ is a strictly diagonally dominant matrix. It follows from the Levy-Desplanques theorem that $\mathbf{I}_L - zt\mathbf{G}(z)\mathbf{L}(z)$ is nonsingular. From (A.20), we conclude that $\det \mathbf{F}(z, t) \neq 0$ for $|z| = 1$ and $t \in [0, 1)$.

For case (b) in which $|z| = 1$, $z \neq 1$, since $|G_i(z)| < 1$, we see that

$$\sum_{j \neq i} \left| -zG_i(z)\frac{p_{ij}}{q_{ij}(z)} \right| < \sum_{j \neq i} \frac{p_{ij}}{|q_{ij}(z)|} \leq \sum_{j \neq i} p_{ij} = 1 - p_{ii} < \left| 1 - zG_i(z)\frac{p_{ii}}{q_{ii}(z)} \right|.$$

Thus $\mathbf{I}_L - z\mathbf{G}(z)\mathbf{L}(z)$ is also a strictly diagonally dominant matrix. It follows again that $\mathbf{I}_L - z\mathbf{G}(z)\mathbf{L}(z)$ is nonsingular. From (A.20), we conclude that $\det \mathbf{F}(z) \neq 0$ for $|z| = 1, z \neq 1$. □

Theorem 3 *If $\gamma > 0$, $\det \mathbf{F}(z)$ has $L^2 - 1$ zeros in $|z| < 1$, and it has a simple zero at $z = 1$.*

Proof. Our proof follows [3, p.241]. We first observe that $\det \mathbf{F}(z, 0) = \det \mathbf{V}(z)$ has L^2 zeros in $|z| \leq 1$, because each element $q_{ij}(z)$ in $\mathbf{V}(z)$ has a single zero at

$$z_{ij} = \frac{\lambda + \mu + \alpha_{ij} - \sqrt{(\lambda + \mu + \alpha_{ij})^2 - 4\lambda\mu}}{2\lambda}$$

in $|z| \leq 1$. From Theorem 2(a), we have $\det \mathbf{F}(z, t) \neq 0$ for $|z| = 1$ and $t \in [0, 1)$. Thus, according to the above lemma, there are L^2 zeros of $\det \mathbf{F}(z, t)$ in $|z| < 1$ for all $t \in [0, 1)$.

We next investigate $\det \mathbf{F}(z, t)$ at $t = 1$. Note that

$$\det \mathbf{F}(1, 1) = \det \mathbf{F}(1) = \det[\mathbf{I}_L - \mathbf{P}] = 0.$$

If $\gamma > 0$, the point $z = 1$ is a simple zero of the function $\det \mathbf{F}(z, 1) = \det \mathbf{F}(z)$. Since $\det \mathbf{F}(1, 1) = 0$, then $\det \mathbf{F}(1 - \varepsilon, 1) < 0$ for small $\varepsilon > 0$. By continuity in $t \in [0, 1)$, there is small τ so that $\det \mathbf{F}(1 - \varepsilon, 1 - \tau) < 0$. However, $\det \mathbf{F}(1, 0) = \det \mathbf{V}(1) = 1$ and $\det \mathbf{F}(1, t) \neq 0$ for $0 \leq t < 1$ as shown above. By continuity, $\det \mathbf{F}(1, t) > 0$ for $0 \leq t < 1$, so in particular, $\det \mathbf{F}(1, 1 - \tau) > 0$. Therefore, $\det \mathbf{F}(1 - \varepsilon_1, 1 - \tau) = 0$ for some $0 < \varepsilon_1 < \varepsilon$. The same argument holds for $\tau \to 0$, so the simple zero at $z = 1$ is the limit of zeros from inside the unit disk. It follows that $\det \mathbf{F}(z, 1) = \det \mathbf{F}(z)$ has L^2 zeros in $|z| \leq 1$. From Theorem 2(b), $\det \mathbf{F}(z)$ has $L^2 - 1$ zeros in $|z| < 1$. □

Appendix 2: Number of Zeros of T(z) in (43) in $|z| \leq 1$

We claim that $T(z)$ in (43) has exactly twelve zeros in the unit disk $|z| \leq 1$ if the condition

$$\alpha g + \lambda < \mu \tag{A.22}$$

is satisfied. Equivalently, we consider

$$\hat{T}(z) := \alpha^{12} T(z) = [\hat{q}(z)]^{12} - \alpha^{12} z^{12} \prod_{l=0}^{11} G_l(z), \tag{A.23}$$

where

$$\hat{q}(z) := \alpha z - (1 - z)(\mu - \lambda z). \tag{A.24}$$

Our proof is based on Rouché's theorem [12, p.116]: *If $f(z)$ and $h(z)$ are analytic functions of z inside and on a closed contour C, and $|h(z)| < |f(z)|$ on C, then $f(z)$ and $f(z) + h(z)$ have the same number of zeros inside C.*

We prove the above claim in a way similar to those in [5] and [13]. Let

$$f(z) := [\hat{q}(z)]^{12}, \qquad h(z) := -\alpha^{12} z^{12} \prod_{l=0}^{11} G_l(z). \tag{A.25}$$

Then $\hat{T}(z) = f(z) + h(z)$.

Let us choose a closed contour C so as to include $z = 1$ as an internal point, which is obviously a zero of $\hat{T}(z)$. In particular, we choose C as

$$C := \left\{ z = e^{i\theta}; 0 < \theta < 2\pi \right\} \bigcup \lim_{\varepsilon \to 0} C_\varepsilon, \tag{A.26}$$

where

$$C_\varepsilon := \left\{ z = 1 + \varepsilon e^{i\varphi}; -\frac{\pi}{2} \leq \varphi \leq \frac{\pi}{2} \right\} \tag{A.27}$$

is a semicircle centered at $z = 1$ with radius $\varepsilon > 0$. The functions $f(z)$ and $h(z)$ are analytic inside and on the contour C.

We now compare $|f(z)|$ and $|h(z)|$ on $\mathcal{C}$. First, we look at z on the unit circle $|z| = 1$. Since $\hat{q}(z) = (\alpha + \lambda + \mu)z - (\lambda z^2 + \mu)$, we see that

$$|\hat{q}(z)| \geq \alpha + \lambda + \mu - (\lambda + \mu) = \alpha$$

on $|z| = 1$. Hence, for $|z| = 1$, $z \neq 1$, it holds that

$$|f(z)| \geq \alpha^{12}, \qquad |h(z)| < \alpha^{12},$$

because $G_l(z) < 1$ since $G_l(z)$'s are probability generating functions for $l = 0, \ldots, 11$. Thus, $|h(z)| < |f(z)|$ for $|z| = 1$, $z \neq 1$.

We next look at $z = 1 + \varepsilon e^{i\varphi}$ on $\mathcal{C}_\varepsilon$, for which

$$\begin{aligned} \hat{q}(z) &= (\alpha + \lambda + \mu)(1 + \varepsilon e^{i\varphi}) - \lambda(1 + \varepsilon e^{i\varphi})^2 - \mu \\ &= \alpha + (\mu + \alpha - \lambda)\varepsilon e^{i\varphi} + o(\varepsilon). \end{aligned} \tag{A.28}$$

It follows that

$$\begin{aligned} |f(z)|^2 &= \left|[\alpha + (\mu + \alpha - \lambda)\varepsilon e^{i\varphi} + o(\varepsilon)]^{12}\right|^2 \\ &= \alpha^{24} + 24\alpha^{23}(\mu + \alpha - \lambda)\varepsilon \cos\varphi + o(\varepsilon). \end{aligned} \tag{A.29}$$

We also have

$$\begin{aligned} |h(z)|^2 &= \left|\alpha^{12}(1 + \varepsilon e^{i\varphi})^{12} \prod_{l=0}^{11}[1 + g_l \varepsilon e^{i\varphi} + o(\varepsilon)]\right|^2 \\ &= \alpha^{24}\left[1 + 2\sum_{l=0}^{11} g_l \varepsilon \cos\varphi + 24\varepsilon \cos\varphi + o(\varepsilon)\right]. \end{aligned} \tag{A.30}$$

Hence

$$\begin{aligned} |f(z)|^2 - |h(z)|^2 &= 24\alpha^{23}\varepsilon \cos\varphi \left[\mu - \lambda - \frac{\alpha}{12}\sum_{l=0}^{11} g_l\right] \\ &= 24\alpha^{23}\varepsilon \cos\varphi(\mu - \lambda - \alpha g). \end{aligned} \tag{A.31}$$

Therefore, if the condition in (A.22) holds, we see that $|h(z)|^2 < |f(z)|^2$ (thus $|h(z)| < |f(z)|$) on $\mathcal{C}_\varepsilon$ for a sufficiently small value of ε. Hence we have shown that $|h(z)| < |f(z)|$ on the entire contour $\mathcal{C}$. Thus the functions $f(z)$ and $h(z)$ satisfy the condition of Rouché's theorem with contour $\mathcal{C}$. It follows that $f(z)$ and $f(z) + h(z) = \hat{T}(z)$ have the same number of zeros inside $\mathcal{C}$.

Finally, we consider the number of zeros of $f(z) = [\hat{q}(z)]^{12}$ inside $\mathcal{C}$. Clearly, there is a single zero of $\hat{q}(z)$ inside $\mathcal{C}$, which is

$$z = \frac{\lambda + \mu + \alpha - \sqrt{(\lambda + \mu + \alpha)^2 - 4\lambda\mu}}{2\lambda}.$$

Thus $f(z)$ has a zero with twelve-fold multiplicity inside $\mathcal{C}$. Hence we conclude that $\hat{T}(z)$ has twelve zeros inside $\mathcal{C}$. □