No. 985

Multiserver Queue with Semi-Markovian Batch Arrivals
with Application to the MPEG Frame Sequence

by

De-An Wu and Hideaki Takagi

May 2002

Multiserver Queue with Semi-Markovian Batch Arrivals with Application to the MPEG Frame Sequence

De-An Wu

Doctoral Program in Policy and Planning Sciences, University of Tsukuba

1-1-1 Tennoudai, Tsukuba-shi, Ibaraki 305-8573, Japan

phone: +81-298-53-5167;   fax: +81-298-53-5167

e-mail: tbu@sk.tsukuba.ac.jp

and

Hideaki Takagi

Vice President, University of Tsukuba

1-1-1 Tennoudai, Tsukuba-shi, Ibaraki 305-8577, Japan

phone: +81-298-53-2005;   fax: +81-298-53-6310

e-mail: takagi@sk.tsukuba.ac.jp

Abstract

We consider a queueing system consisting of multiple identical servers and a common queue. The service time follows an exponential distribution and the arrival process is governed by a semi-Markov process (SMP). The motivation to study the queueing system with SMP arrivals lies in that it can model the auto-correlated traffic on the high speed network generated by a real time communication, for example, the MPEG-encoded VBR video. Our analysis is based on the theory of piecewise Markov process. We first derive the distributions of the queue size and the waiting time. The stability condition of the system is also discussed. When the sojourn time of SMP follows an exponential distribution all the unknown constants contained in the generating function of queue size can be determined through the zeros of the denominator for this generating function. Based on the result of the analysis, we propose a model to evaluate the waiting time of MPEG video traffic on an ATM network with multiple channels. Here, the SMP corresponds to the exact MPEG sequence of frames. Finally, a numerical example using a real video data is shown.

Key words: Semi-Markov process; batch arrivals; multiserver queue; waiting time; MPEG; Rouché's theorem

1   Introduction

We analyze a queueing system with multiple identical servers and a common queue which is fed by a semi-Markov batch arrival process. The service time follows an exponential distribution, and the capacity of waiting room is infinite. This system is denoted by $\text{SMP}^{[X]}/\text{M}/c$ throughout the paper.

Queueing systems with an arrival process governed by a semi-Markov process (SMP) have been studied extensively since Çinlar [2] first analyzed an SMP/M/1 queue. The motivation to study the queueing system with an SMP arrival lies in that it can model

the auto-correlated traffic on the high speed network generated by a real time communication, for example, Motion Picture Experts Group (MPEG) -encoded variable bit rate (VBR) video. Arjas [1] and de Smit [14] apply a matrix factorization method to an SMP/G/1 queue. Sengupta [13] provides a matrix-exponential solution to an SMP/PH/1 queue. Haβlinger [5] analyzes a discrete time SMP/G/1 queue using Wiener-Hopf factorization approach. To the best of our knowledge, there are few studies on multiple-server queues with SMP arrival. In [10, p.164], Neuts analyzes an SMP/M/c queue by the well-known matrix-geometric method. While this method may handle the situation in which the customers arrive in batches of random size, it is difficult to conduct numerical computation, because it typically leads to blocks of very large dimension in partitioning the transition matrix.

On the other hand, Kuczura [8] studies a piecewise Markov process. Based on this theory, Yagyu and Takagi [16] consider an SSMP$^{[X]}$+M/M/1 queue, where the SSMP stands for a special semi-Markov process in which the sojourn time (interarrival time) in the state of SMP is determined only by the current state. Wu and Takagi [17] extend this model to a more general case, namely, an SMP$^{[X]}$+M/M/1 queue. The method for dealing with an SMP$^{[X]}$/M/c queue in this paper is also the theory of piecewise Markov process.

The rest of paper is organized as follows. Section 2 develops the formulas for a pure death process as preliminaries. The semi-Markov batch arrival process is described in Section 3. In Section 4, the generating function for the queue size in the SMP$^{[X]}$/M/c system is derived. The stability condition of this system is discussed in Section 5. In Appendix 3, we prove that the unknown constants contained in the generating function for the queue size can be determined by the zeros of the denominator for this generating function, when the sojourn time in the state of SMP follows an exponential distribution. The waiting time distribution is studied in Section 6. In Section 7 we use an SMP$^{[X]}$/M/c queue to model the transmission of MPEG frames in multiple channels in an ATM network. A numerical example using a real video data taken from the Jurassic Park video is also given.

2   Transient Behavior of a Pure Death Process

We first consider a pure death process $\{X(t); t \geq 0\}$ for a population which evolves in the following way. At $t = 0$ the population size equals i (> 0). Each individual dies in the time interval $[t, t + \Delta t)$ with probability $\mu \Delta t$. If the population size i is less than c (a constant positive integer), the death rate is $i\mu$, otherwise it is $c\mu$. We define the transition probability for the population size

$$P_{i,j}(t) := P\{X(t) = j | X(0) = i\}; \quad 0 \leq j \leq i, \tag{1}$$

and the generating function of its Laplace transform

$$P_i^*(s, z) := \sum_{j=0}^{i} P_{i,j}^*(s) z^j, \tag{2}$$

where

$$P_{i,j}^*(s) := \int_0^\infty e^{-st} P_{i,j}(t) dt$$

is the Laplace transform of $P_{i,j}(t)$.

Since the transient behavior of the pure death process $\{X(t); t \geq 0\}$ depends on the initial population size i, we derive $P_i^*(s, z)$ for two cases, $0 < i \leq c$ and $i > c$, separately. When $0 < i \leq c$, the corresponding Kolmogorov forward equations are

$$
\begin{aligned}
P_{i,0}'(t) &= \mu P_{i,1}(t), \\
P_{i,j}'(t) &= -j\mu P_{i,j}(t) + (j+1)\mu P_{i,j+1}(t); \quad 1 \leq j \leq i-1, \\
P_{i,i}'(t) &= -i\mu P_{i,i}(t),
\end{aligned}
\tag{3}
$$

with the initial condition

$$P_{i,i}(0) = 1.$$

Taking the Laplace transform of (3) gives

$$
\begin{aligned}
sP_{i,0}^*(s) &= \mu P_{i,1}^*(s), \\
sP_{i,j}^*(s) &= -j\mu P_{i,j}^*(s) + (j+1)\mu P_{i,j+1}^*(s); \quad 1 < j \leq i-1, \\
sP_{i,i}^*(s) &= 1 - i\mu P_{i,i}^*(s),
\end{aligned}
$$

Solving this set of equations, we have

$$P_{i,j}^*(s) = \frac{i!\mu^{i-j} B(s; j-1)}{j! B(s; i)}; \quad 0 \leq j \leq i \leq c, \tag{4}$$

where $B(s; j)$ is defined by

$$
\begin{aligned}
B(s; j) &:= s(s+\mu)(s+2\mu)\cdots(s+j\mu); \quad j \geq 0, \\
B(s; -1) &:= 1.
\end{aligned}
\tag{5}
$$

Substituting (4) into (2) yields

$$P_i^*(s, z) = \tau_i(s, z); \quad 0 < i \leq c, \tag{6}$$

where

$$\tau_k(s, z) := \frac{k!\mu^k}{B(s; k)} \sum_{j=0}^k \frac{B(s; j-1)}{j!} \left(\frac{z}{\mu}\right)^j; \quad k \geq 1. \tag{7}$$

When the initial population size i is greater than c, the Kolmogorov forward equations become

$$
\begin{aligned}
P_{i,0}'(t) &= \mu P_{i,1}(t), \\
P_{i,j}'(t) &= -j\mu P_{i,j}(t) + (j+1)\mu P_{i,j+1}(t); \quad 1 \leq j \leq c-1, \\
P_{i,j}'(t) &= -c\mu P_{i,j}(t) + c\mu P_{i,j+1}(t); \quad c \leq j \leq i-1, \\
P_{i,i}'(t) &= -c\mu P_{i,i}(t).
\end{aligned}
\tag{8}
$$

3

Taking the Laplace transform and solving the resulting equations, we obtain

$$P_{i,j}^*(s) = \begin{cases} \dfrac{c!\mu^{c-j}B(s;j-1)}{j!B(s;c)}\left(\dfrac{c\mu}{s+c\mu}\right)^{i-c} ; & 0 \le j \le c, \\[4mm] \dfrac{1}{s+c\mu}\left(\dfrac{c\mu}{s+c\mu}\right)^{i-j} ; & c < j \le i. \end{cases} \tag{9}$$

Multiplying (9) by z^j and summing from 0 to i, we get

$$P_i^*(s,z) = \frac{z^{i+1}}{(s+c\mu)z - c\mu} + \xi_c(s,z)\left(\frac{c\mu}{s+c\mu}\right)^i ; \quad i > c, \tag{10}$$

where

$$\xi_k(s,z) := \left(\frac{s+k\mu}{k\mu}\right)^k \left[\frac{k!\mu^k}{B(s;k)}\sum_{j=0}^{k-1}\frac{B(s;j-1)}{j!}\left(\frac{z}{\mu}\right)^j - \frac{k\mu z^k}{(s+k\mu)[(s+k\mu)z - k\mu]}\right];$$
$$k \ge 1. \tag{11}$$

In Appendix 1 we derive several expressions for $\tau_k(s,z)$ and $\xi_k(s,z)$ which will be used in the analysis of an $\text{SMP}^{[X]}/\text{M}/c$ system in Section 4.

3   SMP Batch Arrival Process

We next consider an arrival process with L types of customers, each type arriving in batches of random size. Customers arrive at time epochs $T_0, T_1,\ldots$, with $T_0 := 0$. Then $A_n := T_n - T_{n-1}$, $n \ge 1$, is the interarrival time, and we set $A_0 := 0$. The type of customers arriving at epoch T_n is denoted by $S^{(n)}$. Let $g_l(k)$ denote the probability of batch size being k for type l customers, $l = 1,\ldots,L$. The arrival process is referred to as a semi-Markov process (SMP) if the following condition holds:

$$P\{S^{(n+1)} = l, A_{n+1} \le t | S^{(0)},\ldots,S^{(n)}, A_1,\ldots,A_n\} = P\{S^{(n+1)} = l, A_{n+1} \le t | S^{(n)}\};$$
$$l = 1,\ldots,L ;\ t \ge 0. \tag{12}$$

For this SMP arrival process, we say that the underlying process enters state l when a batch of type l arrives. Let $Q_{lm}(t)$ be the probability that the arrival process moves from state l to state m in time t, i.e.

$$Q_{lm}(t) := P\{S^{(n+1)} = m, A_{n+1} \le t | S^{(n)} = l\}.$$

Let p_{lm} denote the probability that the arrival of type l is followed by the arrival of type m, and let $A_{lm}(t)$ be the distribution function of the time interval between those successive arrivals. Thus

$$Q_{lm}(t) = p_{lm}A_{lm}(t),$$

where

$$Q_{lm}(\infty) = p_{lm}.$$

Hereafter we use a matrix $\mathbf{P} = (p_{lm})$, which is a stochastic matrix as

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

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

$$\pi_m = \sum_{l=1}^{L} \pi_l p_{lm}; \quad m = 1, \ldots, L \quad ; \quad \sum_{l=1}^{L} \pi_l = 1. \tag{13}$$

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.
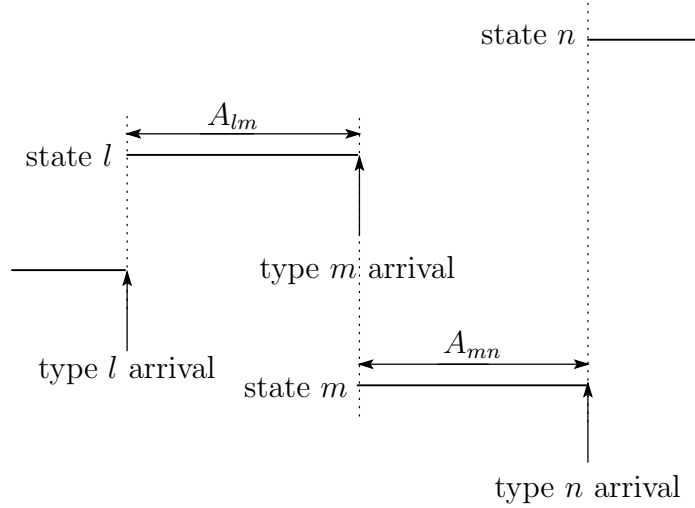

Figure 1: Semi-Markov arrival process.

4 Queue Size in an SMP$^{[X]}$/M/c System

We proceed to analyze the SMP$^{[X]}$/M/c queueing system which is described as follows. The arrival process in batches of customers is a semi-Markov process defined in Section 3. There are c parallel identical servers and a common queue. The service time follows an exponential distribution with mean $1/\mu$. Finally, the capacity of the waiting room is infinite, and the service discipline is first-come first-served (FCFS).

Let $X(t)$ denote the number of customers present in the system (queue size) at time t, including the customers both in service and waiting. It is noted that the system behaves like a pure death process, analyzed in Section 2, between the successive batch arrival epochs. We analyze the queue size in the $\text{SMP}^{[X]}/\text{M}/c$ system by means of a discrete-time Markov chain of two random variables $\{(X^{(n)}, S^{(n)}); n = 0, 1, 2, \dots\}$, where $X^{(n)}$ denotes the number of customers in the system seen by nth arrival, and $S^{(n)}$ denotes the state of the underlying process immediately after the nth arrival.

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

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

(14)

where $P_{i,j}(t)$ is transition probability of the pure death process defined by (1). Assuming that this Markov chain is ergodic, the limiting distribution

$$P(i, l) := \lim_{n \to \infty} P\{X^{(n)} = i, S^{(n)} = l\}; \quad i = 0, 1, 2, \dots ; \ l = 1, \dots, L \quad (15)$$

satisfies the balance equations

$$P(j, m) = \sum_{i=0}^{\infty} \sum_{l=1}^{L} \sum_{k=\max(1,j-i)}^{\infty} p_{lm} g_l(k) P(i, l) \int_0^\infty P_{i+k,j}(t) dA_{lm}(t);$$
$$j = 0, 1, 2, \dots ; m = 1, \dots, L, \quad (16)$$

and the normalization condition

$$\sum_{i=0}^{\infty} \sum_{l=1}^{L} P(i, l) = 1. \quad (17)$$

We transform (16) to a complex integral, since we want to utilize (6) and (10) to convert (16) in terms of the generating function for the queue size. Since $P_{i,j}^*(s)$ is the Laplace transform of $P_{i,j}(t)$, we have the inversion

$$P_{i,j}(t) = \frac{1}{2\pi \mathbf{i}} \int_{b-\mathbf{i}\infty}^{b+\mathbf{i}\infty} e^{st} P_{i,j}^*(s) ds, \quad (18)$$

where $b > 0$, $\mathbf{i} := \sqrt{-1}$, and the integration $\int_{b-\mathbf{i}\infty}^{b+\mathbf{i}\infty}$ denotes the *Bromwich integral*, being written as $\int_{Br}$ hereafter. Substituting (18) into (16) yields

$$P(j, m) = \sum_{i=0}^{\infty} \sum_{l=1}^{L} \sum_{k=\max(1,j-i)}^{\infty} p_{lm} g_l(k) P(i, l) \frac{1}{2\pi \mathbf{i}} \int_{Br} P_{i+k,j}^*(s) \alpha_{lm}(-s) ds;$$
$$j = 0, 1, 2, \dots ; m = 1, \dots, L, \quad (19)$$

6

where

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

is the Laplace-Stieltjes transform (LST) of $A_{lm}(t)$. 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

$$\Phi_l(1) = \pi_l; \quad l = 1, \dots, L. \tag{20}$$

Multiplying (19) by z^j and summing over $j = 0, 1, 2, \dots, i + k$, we obtain

$$\Phi_m(z) = \sum_{l=1}^L p_{lm} \sum_{i=0}^\infty P(i, l) \sum_{k=1}^\infty g_l(k) \frac{1}{2\pi\mathbf{i}} \int_{Br} \sum_{j=0}^{i+k} P_{i+k,j}^*(s) z^j \alpha_{lm}(-s) ds. \tag{21}$$

As shown in Appendix 2, this can be expressed as

$$\Phi_m(z) = \sum_{l=1}^L p_{lm} \frac{1}{2\pi\mathbf{i}} \int_{Br} \left[\frac{z G_l(z) \Phi_l(z) + (z-1) \Big[\lambda(s, z) H_l(s) + \eta_l(s, z) \Big]}{(s + c\mu) z - c\mu} \right] \alpha_{lm}(-s) ds, \tag{22}$$

where $G_l(z)$ is the generating function of $g_l(k)$,

$$\lambda(s, z) := \left(\frac{s + c\mu}{c\mu} \right)^{c-1} \frac{\gamma_c(s, z) - \delta_c(s, z)}{B(s; c - 1)}, \tag{23}$$

$$H_l(s) := G_l \left(\frac{c\mu}{s + c\mu} \right) \Phi_l \left(\frac{c\mu}{s + c\mu} \right); \quad l = 1, 2, \dots, L, \tag{24}$$

$$\eta_l(s, z) := \sum_{i=0}^{c-1} P(i, l) \sum_{k=1}^{c-i} g_l(k)$$
$$\times \left\{ \frac{\mu \Big[c\gamma_{i+k}(s, z) - (i + k)\delta_{i+k}(s, z) + (c - i - k) B(s; i + k - 1) z^{i+k} \Big]}{B(s; i + k)} \right.$$
$$\left. - \frac{\gamma_c(s, z) - \delta_c(s, z)}{B(s; c - 1)} \left(\frac{s + c\mu}{c\mu} \right)^{c-i-k-1} \right\}; \quad l = 1, 2, \dots, L, \tag{25}$$

$$\gamma_k(s, z) := \sum_{j=0}^{k-1} \frac{k!}{j!} B(s; j - 1) \mu^{k-j} z^j; \quad k \geq 1, \tag{26}$$

and

$$\delta_k(s, z) := \sum_{j=1}^{k-1} \frac{(k-1)!}{(j-1)!} B(s; j-1) \mu^{k-j} z^j; \quad k \geq 2,$$

$$\delta_1(s, z) := 0. \tag{27}$$

Following Kuczura [7], we may comment on the Bromwich integral in (22) as follows. Since $P_{i+k,j}(t)$ is the probability, the generating function $P^*_{i+k}(s, z)$ of $P^*_{i+k,j}(s)$ is analytic for $|z| \leq 1$ and $\Re(s) > 0$. Hence the bracketed part of the integrand in (22) is analytic for $|z| \leq 1$ and $\Re(s) > 0$, since it is the convergent series of $\sum_{i=0}^{\infty} \sum_{k=1}^{\infty} P(i, l) g_l(k) P^*_{i+k}(s, z)$. 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 (22) 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 $(b + \mathbf{i}R, b - \mathbf{i}R)$ and a semicircle of radius R in the right-half plane which connects $b - \mathbf{i}R$ with $b + \mathbf{i}R$ counterclockwise. We can choose b 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 (22) are evaluated only at the poles of $\alpha_{lm}(-s)$'s. Therefore, (22) is not a set of integral equations but simply a set of linear equations for $\{\Phi_l(z); \ l = 1, \ldots, L\}$ 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 other relations for $\{P(i, l); \ i = 0, \ldots, c-1\}; \ l = 1, \ldots, L$. In Appendix 3, we show explicitly how to determine $\Phi_l(z)$ when the sojourn time follows an exponential distribution.

5 Stability Condition for an SMP$^{[\mathrm{X}]}$/M/c System

Let us discuss the stability condition for the SMP$^{[\mathrm{X}]}$/M/c system. To do so, we rewrite (22) as

$$\Phi_m(z) = \sum_{l=1}^{L} p_{lm} \left[z G_l(z) \Phi_l(z) \Psi_{lm}(z) + (z-1) \hat{B}_{lm}(z) \right], \tag{28}$$

where

$$\Psi_{lm}(z) := \frac{1}{2\pi \mathbf{i}} \int_{Br} \frac{\alpha_{lm}(-s)}{(s + c\mu)z - c\mu} ds, \tag{29}$$

and

$$\hat{B}_{lm}(z) := \frac{1}{2\pi \mathbf{i}} \int_{Br} \frac{\lambda(s, z) H_l(s) + \eta_l(s, z)}{(s + c\mu)z - c\mu} \alpha_{lm}(-s) ds. \tag{30}$$

Evaluating $\Psi_{lm}(z)$ at $z = 1$ yields

$$\Psi_{lm}(1) = \frac{1}{2\pi\mathbf{i}} \int_{Br} \frac{\alpha_{lm}(-s)}{s} ds = 1. \tag{31}$$

Differentiating $\Psi_{lm}(z)$ and evaluating at $z = 1$, we obtain

$$\begin{aligned}
\left. \frac{d\Psi_{lm}(z)}{dz} \right|_{z=1} &= \frac{1}{2\pi\mathbf{i}} \int_{Br} \frac{-s - c\mu}{s^2} ds \int_0^\infty e^{st} dA_{lm}(t) \\
&= \int_0^\infty dA_{lm}(t) \frac{1}{2\pi\mathbf{i}} \int_{Br} \frac{-s - c\mu}{s^2} e^{st} ds \\
&= -1 - c\mu\hat{a}_{lm},
\end{aligned} \tag{32}$$

where $\hat{a}_{lm}$ is mean of the distribution function $A_{lm}(t)$. These results are used later.

Now, equation (28) can be written in matrix form as

$$\mathbf{\Phi}(z)\hat{\mathbf{F}}(z) = (z - 1)\mathbf{1}\mathrm{diag}[\hat{\mathbf{B}}^{\mathrm{t}}(z)\mathbf{P}], \tag{33}$$

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

$$\hat{\mathbf{F}}(z) := \mathbf{I}_L - z\mathbf{G}(z)\hat{\mathbf{Q}}(z), \tag{34}$$

$$\mathbf{G}(z) := \begin{bmatrix} G_1(z) & 0 & \dots & 0 \\ 0 & G_2(z) & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & G_L(z) \end{bmatrix}, \tag{35}$$

$$\hat{\mathbf{Q}}(z) := \begin{bmatrix} p_{11}\Psi_{11}(z) & p_{12}\Psi_{12}(z) & \dots & p_{1L}\Psi_{1L}(z) \\ p_{21}\Psi_{21}(z) & p_{22}\Psi_{22}(z) & \dots & p_{2L}\Psi_{2L}(z) \\ \vdots & \vdots & \ddots & \vdots \\ p_{L1}\Psi_{L1}(z) & p_{L2}\Psi_{L2}(z) & \dots & p_{LL}\Psi_{LL}(z) \end{bmatrix}, \tag{36}$$

$$\hat{\mathbf{B}}(z) := \begin{bmatrix} \hat{B}_{11}(z) & \hat{B}_{12}(z) & \dots & \hat{B}_{12}(z) \\ \hat{B}_{21}(z) & \hat{B}_{22}(z) & \dots & \hat{B}_{2L}(z) \\ \vdots & \vdots & \ddots & \vdots \\ \hat{B}_{L1}(z) & \hat{B}_{L2}(z) & \dots & \hat{B}_{LL}(z) \end{bmatrix}, \tag{37}$$

and $\mathbf{I}_L$ denotes an $L \times L$ identity matrix. In equation (33), diag$\mathbf{X}$ is a diagonal matrix whose elements are taken from the corresponding elements of $\mathbf{X}$, and $\hat{\mathbf{B}}^{\mathrm{t}}(z)$ is the transpose of $\hat{\mathbf{B}}(z)$.

Differentiating (33) and evaluating the result at $z = 1$, we obtain

$$\mathbf{\Phi}'(1)(\mathbf{I}_L - \mathbf{P}) + \boldsymbol{\pi}\hat{\mathbf{F}}'(1) = \mathbf{1}\mathrm{diag}[\hat{\mathbf{B}}^{\mathrm{t}}(1)\mathbf{P}], \tag{38}$$

9

Here we have used $\hat{\mathbf{F}}(1) = \mathbf{I}_L - \mathbf{P}$ since $\mathbf{G}(1) = \mathbf{I}_L$ and $\hat{\mathbf{Q}}(1) = \mathbf{P}$. Note also that $\boldsymbol{\Phi}(1) = \boldsymbol{\pi}$. Multiplying (38) on the right by $\mathbf{1}^t := [1, \ldots, 1]^t$ and noting that $(\mathbf{I}_L - \mathbf{P})\mathbf{1}^t = 0$, we get

$$\boldsymbol{\pi}\hat{\mathbf{F}}'(1)\mathbf{1}^t = \mathbf{1}\mathrm{diag}[\hat{\mathbf{B}}^t(1)\mathbf{P}]\mathbf{1}^t. \tag{39}$$

To determine the left-hand side of (39), we differentiate (34) and evaluate the result at $z = 1$, where we use (31) and (32). Then we have

$$\hat{\mathbf{F}}'(1) = -\hat{\mathbf{Q}}'(1) - \mathbf{G}'(1)\mathbf{P} - \mathbf{P}, \tag{40}$$

where

$$\mathbf{G}'(1) = \begin{bmatrix} g_1 & 0 & \ldots & 0 \\ 0 & g_2 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & g_L \end{bmatrix}, \tag{41}$$

and

$$\hat{\mathbf{Q}}'(1) = \begin{bmatrix} p_{11}(-1 - c\mu\hat{a}_{11}) & p_{12}(-1 - c\mu\hat{a}_{12}) & \ldots & p_{1L}(-1 - c\mu\hat{a}_{1L}) \\ p_{21}(-1 - c\mu\hat{a}_{21}) & p_{22}(-1 - c\mu\hat{a}_{22}) & \ldots & p_{2L}(-1 - c\mu\hat{a}_{2L}) \\ \vdots & \vdots & \ddots & \vdots \\ p_{L1}(-1 - c\mu\hat{a}_{L1}) & p_{L2}(-1 - c\mu\hat{a}_{L1}) & \ldots & p_{LL}(-1 - c\mu\hat{a}_{LL}) \end{bmatrix}. \tag{42}$$

Here g_l is the mean batch size of type l customers. Multiplying (39) on the right by $\mathbf{1}^t$ and substituting (41) and (42) yields

$$\begin{aligned} \hat{\mathbf{F}}'(1)\mathbf{1}^t &= -\hat{\mathbf{Q}}'(1)\mathbf{1}^t - \mathbf{G}'(1)\mathbf{1}^t - \mathbf{1}^t \\ &= \begin{bmatrix} \sum_{m=1}^{L} p_{1m}(1 + c\mu\hat{a}_{1m}) \\ \sum_{m=1}^{L} p_{2m}(1 + c\mu\hat{a}_{2m}) \\ \vdots \\ \sum_{m=1}^{L} p_{Lm}(1 + c\mu\hat{a}_{Lm}) \end{bmatrix} - \begin{bmatrix} g_1 \\ g_2 \\ \vdots \\ g_L \end{bmatrix} - \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix}. \end{aligned} \tag{43}$$

Finally, multiplying (43) on the left by $\boldsymbol{\pi}$, we obtain

$$\boldsymbol{\pi}\hat{\mathbf{F}}'(1)\mathbf{1}^t = \sum_{l=1}^{L} \pi_l \sum_{m=1}^{L} p_{lm}(1 + c\mu\hat{a}_{lm}) - g - 1 = c\mu\hat{a} - g, \tag{44}$$

10

where $\hat{a}$ is the mean of interarrival time for SMP arrivals defined by

$$\hat{a} := \sum_{l=1}^{L} \pi_l \sum_{m=1}^{L} p_{lm} \hat{a}_{lm}, \tag{45}$$

and g is the mean batch size given by

$$g := \sum_{l=1}^{L} \pi_l g_l. \tag{46}$$

The expression in (44) is the left-hand side of (39). Thus we have

$$c\mu\hat{a} - g = \sum_{l=1}^{L} \sum_{m=1}^{L} \hat{B}_{lm}(1)p_{lm}. \tag{47}$$

From Lemma 5 in Appendix 1, we see that the right-hand side of this equation is positive. Hence we have

$$\frac{g}{\hat{a}} < c\mu. \tag{48}$$

The condition in (48) means that the customer arrival rate is less than the total service rate. Thus it assures the stability of our system.

6 Waiting time of an Arbitrary Customer

We proceed to determine the waiting time W of an arbitrary customer. Let us focus on a randomly chosen *tagged* customer included in a batch that arrives to bring state l. Let $\hat{G}_l$ denote the number of customers placed before the tagged customer in this batch, and $W_l(t)$ be the waiting time distribution of this tagged customer. The distribution of $\hat{G}_l$ is given by [15, p.46]

$$\hat{g}_l(k) = \frac{\displaystyle\sum_{j=k+1}^{\infty} g_l(k)}{g_l}; \quad k = 0, 1, 2, \ldots, \tag{49}$$

and its generating function is

$$\hat{G}_l(z) = \frac{1 - G_l(z)}{g_l(1 - z)}. \tag{50}$$

The probability that the tagged customer need not wait is

$$W_l(0) = \sum_{k=0}^{c-1} \sum_{i=0}^{c-1-k} \hat{g}_l(k) P(i, l).$$

11

On the other hand, if the tagged customer in a batch of size k arrives and finds that the number $i+k$ of customers in the system is great than c, he must wait until $i+k+1-c$ customers depart before he enters service. Therefore, his waiting time has $(i+k+1-c)$-stage Erlang distribution. Thus, $W_l(t)$ is given by

$$
\begin{aligned}
W_l(t) = & \sum_{k=0}^{c-1}\sum_{i=0}^{c-1-k} \hat{g}_l(k)P(i,l) + \sum_{k=1}^{c-1}\sum_{i=c-k}^{c-1} \hat{g}_l(k)P(i,l)\int_0^t \frac{c\mu(c\mu y)^{i+k-c}e^{-c\mu y}}{(i+k-c)!}dy \\
& + \sum_{k=0}^{c-1}\sum_{i=c}^{\infty} \hat{g}_l(k)P(i,l)\int_0^t \frac{c\mu(c\mu y)^{i+k-c}e^{-c\mu y}}{(i+k-c)!}dy \\
& + \sum_{k=c}^{\infty}\sum_{i=0}^{\infty} \hat{g}_l(k)P(i,l)\int_0^t \frac{c\mu(c\mu y)^{i+k-c}e^{-c\mu y}}{(i+k-c)!}dy.
\end{aligned}
\tag{51}
$$

Taking the LST of $W_l(t)$, we obtain

$$
\begin{aligned}
\Omega_l(s) = & \frac{[B(s)]^{1-c}}{g_l[1-B(s)]}\Big(1 - G_l[B(s)]\Big)\Phi_l[B(s)] \\
& + \sum_{k=0}^{c-1}\sum_{i=0}^{c-1-k} \hat{g}_l(k)P(i,l)\Big(1 - [B(s)]^{i+k+1-c}\Big),
\end{aligned}
\tag{52}
$$

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

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

$$
\begin{aligned}
\Omega(s) = & \frac{1}{g}\sum_{l=1}^{L} g_l\Omega_l(s) \\
= & \frac{1}{g}\left\{ \frac{[B(s)]^{1-c}}{1-B(s)}\sum_{l=1}^{L}\Big(1 - G_l[B(s)]\Big)\Phi_l[B(s)] \right. \\
& \left. + \sum_{l=1}^{L}\sum_{k=0}^{c-1}\sum_{i=0}^{c-1-k} g_l\hat{g}_l(k)P(i,l)\Big(1 - [B(s)]^{i+k+1-c}\Big) \right\}.
\end{aligned}
\tag{53}
$$

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

$$
E[W] = \frac{1}{gc\mu}\left[\sum_{l=1}^{L} g_l E_l[X] + \frac{g^{(2)}}{2} - \sum_{l=1}^{L}\sum_{k=0}^{c-1}\sum_{i=0}^{c-1-k} g_l\hat{g}_l(k)P(i,l)(i+k+1-c)\right] - \frac{c-1}{c\mu},
\tag{54}
$$

$$
\begin{aligned}
E[W^2] = & \frac{1}{g(c\mu)^2}\left\{ \sum_{l=1}^{L}\left[g_l\Big(E_l[X^2] + (3-2c)E_l[X]\Big) + g_l^{(2)}E_l[X]\right] + (2-c)g^{(2)} + \frac{g^{(3)}}{3} \right. \\
& \left. - \sum_{l=1}^{L}\sum_{k=0}^{c-1}\sum_{i=0}^{c-1-k} g_l\hat{g}_l(k)P(i,l)(i+k+1-c)(i+k+2-c) \right\} + \frac{(c-1)(c-2)}{(c\mu)^2},
\end{aligned}
\tag{55}
$$

12

where

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

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

7 Application to the MPEG Frame Sequence

Let us use the $\mathrm{SMP}^{[X]}/\mathrm{M}/c$ system to model the traffic in the ATM network in which the MPEG frames are transmitted through multiple channels. The waiting time of an arbitrary ATM cell generated from MPEG frames is studied. In Section 7.1, a brief description of MPEG coding scheme is given. In section 7.2, the transmission of MPEG frame sequence is modeled by an $\mathrm{SMP}^{[X]}/\mathrm{M}/c$ system. Assuming that the MPEG frame arrival process is Poisson, we obtain the formula for evaluating the waiting time of an arbitrary ATM cell. In Section 7.3, some numerical results using the statistics of a real video film are presented.
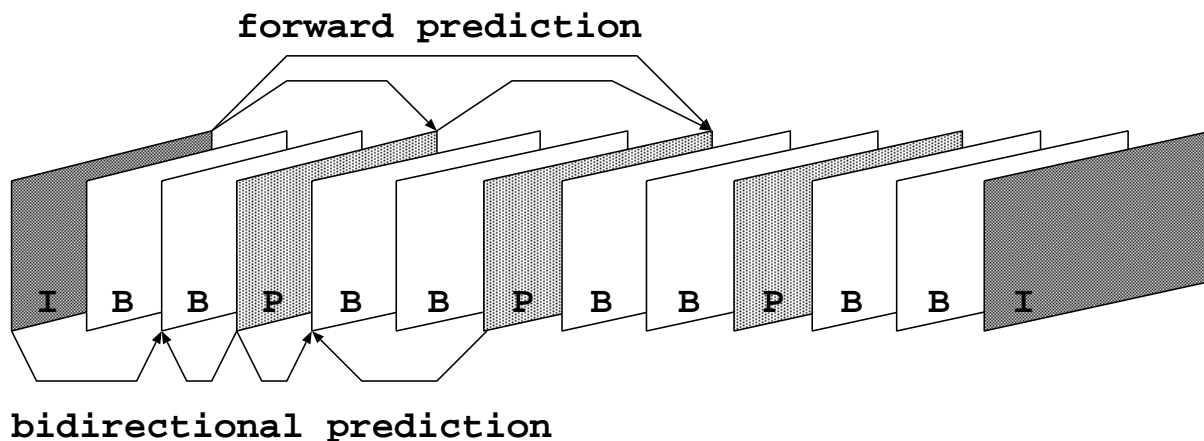
7.1 MPEG Video Coding Scheme

Figure 2: Group of pictures (GOP) of an MPEG stream [9].

In the MPEG coding [9], 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.

These frames are arranged in a deterministic sequence "IBBPBBPBBPBB," which is called a Group of Pictures (GOP) as shown in Figure 2. The length of the GOP 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).

7.2 Traffic Model for MPEG Frame Sequence

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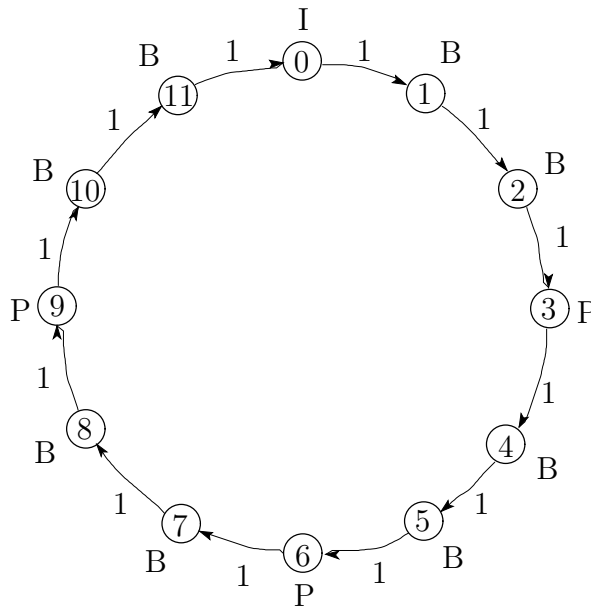

Figure 3: State transition diagram of the MPEG frame pattern.

The stationary distribution of this Markov chain is given by

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

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, $l = 0, \ldots, 11$, and let c be the number of channels for transmitting the MPEG frames. Equations in (22) then become

$$\Phi_m(z) = \frac{1}{q(z)} \Big\{ z G_{m-1}(z) \Phi_{m-1}(z) + (z-1)[\lambda(\alpha, z) H_{m-1} + \eta_{m-1}(\alpha, z)] \Big\};$$
$$m = 0, \ldots, 11, \qquad (56)$$

where

$$q(z) := \frac{1}{\alpha}[(\alpha + c\mu)z - c\mu],$$

and

$$H_m := G_m \left(\frac{c\mu}{\alpha + c\mu} \right) \Phi_m \left(\frac{c\mu}{\alpha + c\mu} \right).$$

We note that $\{H_m;\ m = 0, \ldots, 11\}$, and $\{P(i, m);\ i = 0, \ldots, c-1;\ m = 0, \ldots, 11\}$, included in $\eta_m(\alpha, z)$, are constants to be determined. Hereafter state "$-m$" should read state "$12 - m$". Solving the set of equations in (56), we get

$$\Phi_m(z) = \frac{(z-1) \sum_{k=0}^{11} z^k [q(z)]^{11-k} [\lambda(\alpha, z) H_{m+k-1} + \eta_{m+k-1}(\alpha, z)] \prod_{l=m-k}^{m-1} G_l(z)}{T(z)};$$
$$m = 0, \ldots, 11, \quad (57)$$

where

$$T(z) := [q(z)]^{12} - z^{12} \prod_{l=0}^{11} G_l(z). \qquad (58)$$

It is shown in Appendix 3 that there are twelve zeros of $T(z)$ in $|z| \leq 1$ under the condition

$$\alpha g < c\mu,$$

where

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

is the mean size of an MPEG frame. Therefore, by using the twelve zeros of $T(z)$ in $|z| \leq 1$ together with relations

$$P(i, m) = \frac{1}{i!} \frac{d^i \Phi_m(z)}{dz^i} \bigg|_{z=0}; \quad i = 0, \ldots, c-1; \quad m = 0, \ldots, 11,$$

we can solve the set of $12+12c$ linear equations for $\{H_m,\ P(i, m);\ m = 0, \ldots, 11;\ i = 0, \ldots, c-1\}$. This completes the determination of parameters in the model.

15

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

I-frame			B-frame			P-frame		
mean	var	c.v.	mean	var	c.v.	mean	var	c.v.
143.427	918.704	0.211	19.033	135.021	0.612	37.659	632.568	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
19	0.132	2	0.105	2	0.053

7.3   Numerical Examples

Let us evaluate the waiting time of an arbitrary ATM cell in the model. The real video film data for the Jurassic Park is prepared by Rose [11], and it can be downloaded from the web site *http://nero.informatik.uni-wuerzburg.de/MPEG/*. In addition, we need to assume some distribution for the number of cells in each frame (frame size) so that we can calculate the distribution of the waiting time numerically.

Frey and Nguen-Quang [3] and Sarkar et al. [12] 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

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

where, we set

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

Table 1 shows the statistics for the number of ATM cells in each frame type for the Jurassic Park video, which have been calculated by assuming that every frame is divided into a group of cells each with a payload of 48 bytes. The fitting parameters for the negative binominal distributions determined from the mean and variance of the data in Table 1 are given in Table 2.

To compare the influence of the number of channels, we keep the total transmission rate at 20Mbps, which corresponds to 2×2350 cells/sec, and vary c from 2 to 3. For

16

each set of parameters we have exactly twelve zeros in the unit disk. The zeros of $T(z)$ for $c = 2$ and $\mu = 2350$ cells/sec are plotted in the complex z-plane in Figure 4.
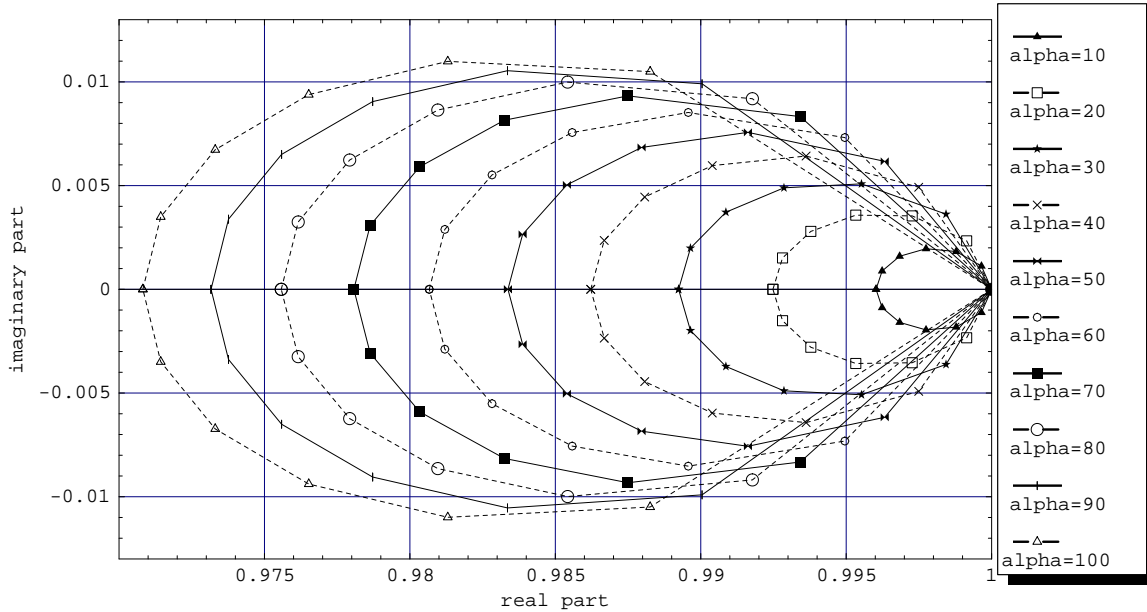

Figure 4: Zeros of $T(z)$ in the unit disk when $c = 2$, $\mu = 2350$ cells/sec.

Figures 5 and 6 show the mean and the variance, respectively, of the waiting time of an arbitrary ATM cell in the MPEG frames for $c = 2$, $\mu = 2350$ cells/sec and for $c = 3$, $\mu = (2/3) \times 2350$ cells/sec. It is observed that at low arrival rate α (frames/sec) both the mean and variance are relatively flat, but they increase rapidly with respect to α at heavy load. It is also observed that the difference (for both the mean and variance) between $c = 2$ and $c = 3$ is very small if we keep the total transmission rate at a constant $c\mu = 2 \times 2350$. In other words, the number of channels does not influence much on the mean and variance of the waiting time as far as the total transmission rate is a constant. This is just like the situation in an M/M/c queueing system.

8 Summary

In this paper, we have first analyzed the queue size in an SMP[X]/M/c system, where the underlying SMP has L states. The formulas of the mean and variance of the waiting time for an arbitrary customer have been presented. When the state sojourn times are exponentially distributed, we have proved that there exist L^2 zeros in the unit disk in the denominator of the generating function for the queue size if the arrival rate is less than the total service rate. 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
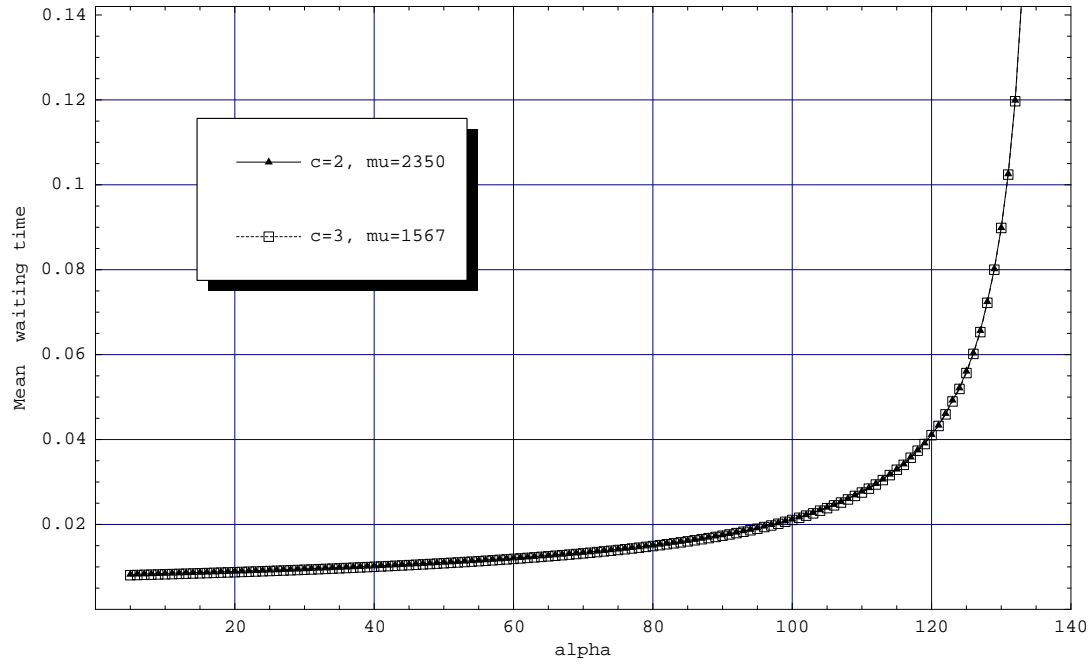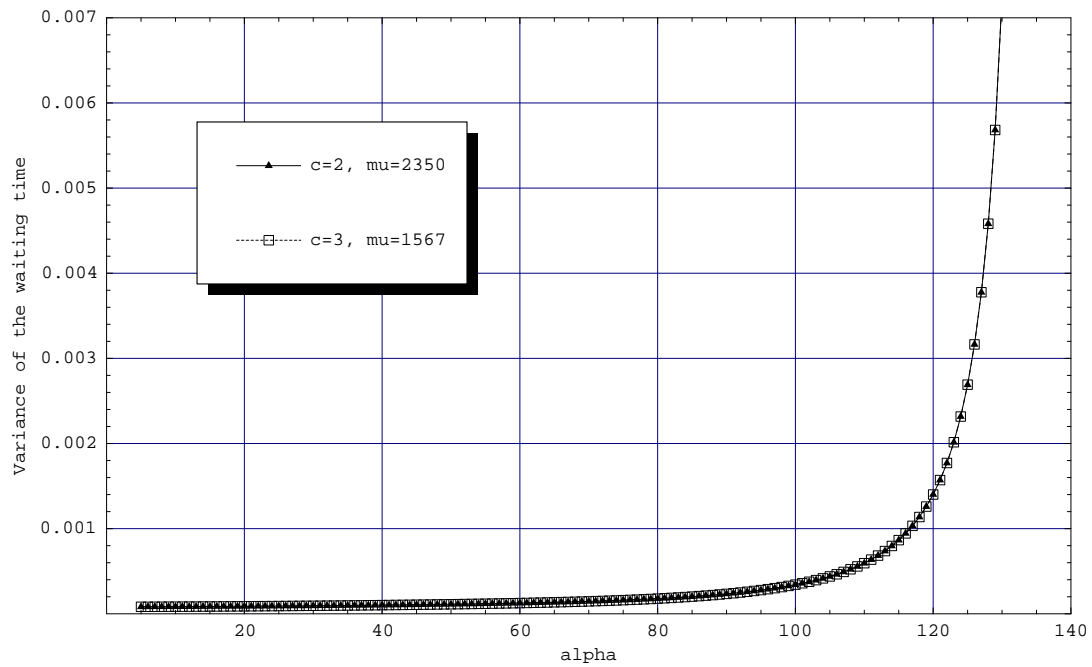
Figure 5: Mean waiting time for an arbitrary cell.

Figure 6: Variance of the waiting time for an arbitrary cell.

18

evaluated. It is observed that, just like the situation in an M/M/c queueing system, the number of channels does not influence much on the mean and variance of the waiting time as far as the total transmission rate $c\mu$ is a constant.

References

[1] E. Arjas, "On a fundamental identity in the theory of semi-Markov processes," *Advances in Applied Probability*, Vol.4, No.2, pp.258–270, August 1972.

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

[3] 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.

[4] 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.

[5] G. Haβlinger, "Waiting time, busy periods and output models of a server analyzed via Wiener-Hopf factorization," *Performance Evaluation*, Vol.40, No.1–3, pp.3–26, March 2000.

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

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

[8] A. Kuczura, "Piecewise Markov processes," *SIAM Journal on Applied Mathematics*, Vol.24, No.2, pp.169–181, March 1973

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

[10] M. F. Neuts, *Matrix-Geometric Solutions in Stochastic Models: An Algorithmic Approach*, The Johns Hopkins University Press, Baltimore, 1981.

[11] 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.

[12] 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.

[13] B. Sengupta, "The semi-Markov queue: theory and applications," *Communications in Statistics–Stochastic Models*, Vol.6, No.3, pp.383–413, September 1990.

[14] J. H. de Smit, "The single server semi-Markov queue," *Stochastic Processes and their Applications*, Vol.22, No.1, pp.37–50, May 1986.

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

[16] 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.

[17] D. Wu and H. Takagi, "A queue with semi-Markovian batch plus Poisson arrivals with application to the MPEG frame sequence," Discussion paper No.958, Institute of Policy and Planning Sciences, University of Tsukuba, November 2001.

Appendix 1: Derivation of Preliminary Formulas

Lemma 1 *For any positive integer $k \geq 1$ the following relation holds:*

$$\gamma_k(s, z) = \frac{k\mu}{sz}[B(s; k-1)z^k - (z-1)\delta_k(s, z)], \qquad (A.1)$$

where $\gamma_k(s, z)$ and $\delta_k(s, z)$ are defined in (26) and (27), respectively.

Proof. Obviously, (A.1) holds for $k = 1$ as both sides equal μ. Thus we prove it for $k \geq 2$. Equation (A.1) can be written as

$$k\mu[B(s; k-1)z^k - (z-1)\delta_k(s, z)] = sz\gamma_k(s, z). \qquad (A.2)$$

Substituting (27) into the left-hand side of (A.2) and expanding yields

$$k\mu\left(B(s; k-1)z^k - (z-1)\sum_{j=1}^{k-1} \frac{(k-1)!}{(j-1)!} B(s; j-1)\mu^{k-j}z^j\right)$$

$$= \sum_{j=1}^{k-1} \frac{k!}{(j-1)!} B(s; j-1)\mu^{k+1-j}z^j - \sum_{j=2}^{k} \frac{k!}{(j-2)!} B(s; j-2)\mu^{k+2-j}z^j + k\mu z^k B(s; k-1)$$

$$= \sum_{j=2}^{k-1} \frac{k!}{(j-1)!} B(s; j-2)\mu^{k+1-j}sz^j - \frac{k!}{(k-2)!} B(s; k-2)\mu^2 z^k + k!\mu^k sz + k\mu z^k B(s; k-1)$$

$$= sz\sum_{j=1}^{k-2} \frac{k!}{j!} B(s; j-1)\mu^{k-j}z^j + sz[k!\mu^k + k\mu B(s; k-2)z^{k-1}]$$

$$= sz\gamma_k(s, z), \qquad (A.3)$$

where we have used $B(s; j-1) = [s + (j-1)\mu]B(s; j-2)$, $j \geq 1$. □

Lemma 2 *For any positive integer $k \geq 1$ the following relation holds:*

$$\xi_k(s, z) = \frac{z-1}{(s+k\mu)z - k\mu}\left(\frac{s+k\mu}{k\mu}\right)^{k-1}\frac{\gamma_k(s, z) - \delta_k(s, z)}{B(s; k-1)}, \qquad (A.4)$$

where $\xi_k(s, z)$ is defined in (11).

Proof. Obviously, relation (A.4) holds for $k = 1$ as both sides equal

$$\frac{(z-1)\mu}{[(s+\mu)z - \mu]s}.$$

We next prove the case for $k \geq 2$. From (11), we have

$$\xi_k(s, z) = \frac{1}{[(s+k\mu)z - k\mu]B(s; k-1)} \left(\frac{s+k\mu}{k\mu}\right)^{k-1}$$
$$\times \left([sz + k\mu(z-1)](k-1)!\mu^{k-1} \sum_{j=0}^{k-1} \frac{B(s; j-1)}{j!} \left(\frac{z}{\mu}\right)^j - B(s; k-1)z^k\right)$$
$$= \frac{1}{[(s+k\mu)z - k\mu]B(s; k-1)} \left(\frac{s+k\mu}{k\mu}\right)^{k-1}$$
$$\times \left((z-1)\gamma_k(s, z) + \frac{sz}{k\mu}\gamma_k(s, z) - B(s; k-1)z^k\right). \tag{A.5}$$

However, from (A.2) we see that

$$\frac{sz}{k\mu}\gamma_k(s, z) - B(s; k-1)z^k = (1-z)\delta_k(s, z). \tag{A.6}$$

Using relation (A.6) in (A.5) gives (A.4). □

Lemma 3 *For any positive integer $1 \leq k \leq c$ the following relation holds:*

$$\tau_k(s, z) - \frac{z^{k+1}}{(s+c\mu)z - c\mu} = \frac{(z-1)\mu\left[c\gamma_k(s, z) - k\delta_k(s, z) + (c-k)B(s; k-1)z^k\right]}{[(s+c\mu)z - c\mu]B(s; k)}, \tag{A.7}$$

where $\tau_k(s, z)$ is defined in (7).

Proof. It is easy to verify that (A.7) holds for $k = 1$ as both sides equal

$$\frac{(z-1)\mu[(c-1)sz + c\mu]}{s(s+\mu)[(s+c\mu)z - c\mu]}.$$

We next consider the case for $k \geq 2$. Rewriting $\xi_k(s, z)$ in terms of $\tau_k(s, z)$ in (11), we have

$$\xi_k(s, z) = \left(\frac{s+k\mu}{k\mu}\right)^k \left(\tau_k(s, z) - \frac{z^k}{s+k\mu} - \frac{k\mu z^k}{(s+k\mu)[(s+k\mu)z - k\mu]}\right). \tag{A.8}$$

Together with (A.4), we obtain

$$\tau_k(s, z) - \frac{z^{k+1}}{(s+k\mu)z - k\mu} = \frac{(z-1)k\mu[\gamma_k(s, z) - \delta_k(s, z)]}{[(s+k\mu)z - k\mu]B(s; k)}. \tag{A.9}$$

21

However, we may write

$$\tau_k(s,z) - \frac{z^{k+1}}{(s+k\mu)z - k\mu}$$
$$= \left(\tau_k(s,z) - \frac{z^{k+1}}{(s+c\mu)z - c\mu}\right) + \frac{z^{k+1}}{(s+c\mu)z - c\mu} - \frac{z^{k+1}}{(s+k\mu)z - k\mu}. \tag{A.10}$$

From (A.9) and (A.10), we have

$$\tau_k(s,z) - \frac{z^{k+1}}{(s+c\mu)z - c\mu}$$
$$= \frac{(z-1)\mu\Big[[sz + c\mu(z-1)]k[\gamma_k(s,z) - \delta_k(s,z)] + (c-k)B(s;k)z^{k+1}\Big]}{[(s+c\mu)z - c\mu][(s+k\mu)z - k\mu]B(s;k)}$$
$$= \frac{(z-1)\mu\Big[[sz + k\mu(z-1) + (c-k)\mu(z-1)]k[\gamma_k(s,z) - \delta_k(s,z)]\Big]}{[(s+c\mu)z - c\mu][(s+k\mu)z - k\mu]B(s;k)}$$
$$+ \frac{(z-1)\mu\Big[(c-k)B(s;k-1)z^k[sz + k\mu(z-1) + k\mu]\Big]}{[(s+c\mu)z - c\mu][(s+k\mu)z - k\mu]B(s;k)}$$
$$= \frac{(z-1)\mu[sz + k\mu(z-1)]\Big[k[\gamma_k(s,z) - \delta_k(s,z)] + (c-k)B(s;k-1)z^k\Big]}{[(s+c\mu)z - c\mu][(s+k\mu)z - k\mu]B(s;k)}$$
$$+ \frac{(z-1)\mu(c-k)\Big[[k\mu(z-1)\gamma_k(s,z) - k\mu(z-1)\delta_k(s,z) + k\mu B(s;k-1)z^k\Big]}{[(s+c\mu)z - c\mu][(s+k\mu)z - k\mu]B(s;k)}$$

$$\tag{A.11}$$

However, from (A.2), we have

$$k\mu[B(s;k-1)z^k - (z-1)\delta_k(s,z)] + k\mu(z-1)\gamma_k(s,z) = [(s+k\mu)z - k\mu]\gamma_k(s,z) \tag{A.12}$$

Substituting (A.12) into (A.11) yields (A.7). □

Lemma 4 *The following relations hold for $\gamma_k(s,z)$ and $\delta_k(s,z)$ at $z = 1$.*

$$\gamma_k(s,1) = \frac{k\mu}{s}B(s;k-1); \quad k \geq 1, \tag{A.13}$$

$$\delta_k(s,1) = \frac{(k-1)\mu}{s+\mu}B(s;k-1); \quad k \geq 1. \tag{A.14}$$

Proof. Evaluating the value of (A.1) at $z = 1$ yields (A.13). We next prove (A.14) by induction.

It is clear that both sides of (A.14) equal 0 for $k = 1$. Assume that the inductive hypothesis is true for $k = n$, that is

$$\delta_n(s,1) = \frac{(n-1)\mu}{s+\mu}B(s;n-1). \tag{A.15}$$

22

Let us consider the case for $k = n + 1$. Form (27), we have

$$\delta_{n+1}(s, 1) = \sum_{j=1}^{n} \frac{n!}{(j-1)!} B(s; j-1)\mu^{n+1-j} = n\mu\delta_n(s, 1) + n\mu B(s; n-1). \quad \text{(A.16)}$$

By the inductive hypothesis (A.15), we have

$$n\mu\delta_n(s, 1) + n\mu B(s; n-1) = \frac{n\mu}{s+\mu} B(s; n). \quad \text{(A.17)}$$

This completes the induction. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\square$

Lemma 5 *For $l, m = 1, 2, \ldots, L$,*

$$\hat{B}_{lm}(1) = \frac{1}{2\pi\mathbf{i}} \int_{Br} \frac{\lambda(s, 1)H_l(s) + \eta_l(s, 1)}{s} \alpha_{lm}(-s)ds \quad \text{(A.18)}$$

is positive, where $H_l(s)$ is defined in (24).

Proof. From Lemma 4, (23) and (25), we have

$$\lambda(s, 1) = \frac{c\mu^2}{s(s+\mu)} \left(\frac{s+c\mu}{c\mu}\right)^c, \quad \text{(A.19)}$$

and

$$\eta_l(s, 1) = \sum_{i=0}^{c-1} P(i, l) \sum_{k=1}^{c-i} g_l(k) \frac{\mu[(c-i-k)s + c\mu]}{s(s+\mu)}$$

$$- \frac{c\mu^2}{s(s+\mu)} \left(\frac{s+c\mu}{c\mu}\right)^c \sum_{i=0}^{c-1} P(i, l) \sum_{k=1}^{c-i} g_l(k) \left(\frac{c\mu}{s+c\mu}\right)^{i+k}. \quad \text{(A.20)}$$

Substituting (A.19) and (A.20) into (A.18), we obtain

$$\hat{B}_{lm}(1) = \frac{1}{2\pi\mathbf{i}} \int_{Br} \left\{ \sum_{i=0}^{c-1} P(i, l) \sum_{k=1}^{c-i} g_l(k) \frac{\mu[(c-i-k)s + c\mu]}{s^2(s+\mu)} + \frac{c\mu^2}{s^2(s+\mu)} \left(\frac{s+c\mu}{c\mu}\right)^c \right.$$

$$\times \left\{ \sum_{i=0}^{c-1} \sum_{k=c-i+1}^{\infty} P(i, l)g_l(k) \left(\frac{c\mu}{s+c\mu}\right)^{i+k} \right.$$

$$\left.\left. + \sum_{i=c}^{\infty} P(i, l) \sum_{k=1}^{\infty} g_l(k) \left(\frac{c\mu}{s+c\mu}\right)^{i+k} \right\} \right\} \alpha_{lm}(-s)ds$$

$$= \int_0^{\infty} dA_{lm}(t) \left\{ \sum_{i=0}^{c-1} \sum_{k=1}^{c-i} P(i, l)g_l(k) \left((c-i-k) \int_0^t f_1^{(1)}(y)dy \right.\right.$$

$$\left.\left. + c\mu \int_0^t \int_0^t f_1^{(1)}(y)(dy)^2\right) + \sum_{i=0}^{c-1} \sum_{k=c-i+1}^{\infty} P(i, l)g_l(k)c\mu \int_0^t \int_0^t f_1^{(1)}(\cdot) * f_{i+k-c}^{(c)}(y)(dy)^2 \right.$$

23

$$+ \sum_{i=c}^{\infty} \sum_{k=1}^{\infty} P(i,l) g_l(k) c\mu \int_0^t \int_0^t f_1^{(1)}(\cdot) * f_{i+k-c}^{(c)}(y)(dy)^2 \Bigg\}, \tag{A.21}$$

where $f_k^{(n)}(y)$ is the density function of the gamma distribution defined by

$$f_k^{(n)}(y) := \frac{(n\mu)^k y^{k-1}}{(k-1)!} e^{-n\mu y}; \quad k, n = 1, 2, \ldots,$$

and $f_i^{(m)}(\cdot) * f_k^{(n)}(y)$ is the convolution of the $f_i^{(m)}(y)$ and $f_k^{(n)}(y)$. Thus, $\hat{B}_{lm}(1)$ is positive.
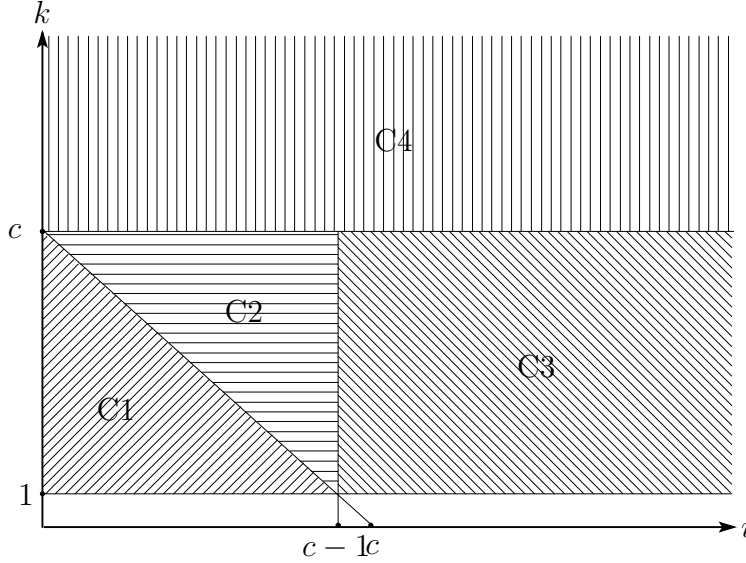□

Appendix 2: Derivation of (22)

Figure 7: Four regions for evaluating the r.h.s. of (21).

In order to derive (22) from (21), we consider the right-hand side of (21) for the four regions in the (i, k) plane as shown in Figure 7. The region labeled C1 corresponds to the case in which some of c servers are idle throughout interarrival interval of the customers. Using the relation (6), we see that the right-hand side of (21) in region C1 equals

$$\sum_{l=1}^{L} p_{lm} \frac{1}{2\pi\mathbf{i}} \int_{Br} \left[\sum_{i=0}^{c-1} P(i,l) \sum_{k=1}^{c-i} g_l(k) \tau_{i+k}(s, z) \right] \alpha_{lm}(-s) ds. \tag{A.22}$$

In region C2, all servers are busy throughout the time interval between two successive batch arrivals. According to (10), the right-hand side of (21) then equals

$$\sum_{l=1}^{L} p_{lm} \frac{1}{2\pi\mathbf{i}} \int_{Br} \sum_{i=1}^{c-1} P(i,l) \sum_{k=c+1-i}^{c} g_l(k) \left[\frac{z^{i+k+1}}{(s+c\mu)z - c\mu} + \xi_c(s, z) \left(\frac{c\mu}{s+c\mu} \right)^{i+k} \right] \alpha_{lm}(-s) ds$$

24

$$= \sum_{l=1}^{L} p_{lm} \frac{1}{2\pi \mathbf{i}} \int_{Br} \sum_{i=1}^{c-1} P(i,l) \left\{ \sum_{k=1}^{c} g_l(k) \left[\frac{z^{i+k+1}}{(s+c\mu)z - c\mu} + \xi_c(s,z) \left(\frac{c\mu}{s+c\mu} \right)^{i+k} \right] \right.$$

$$\left. - \sum_{k=1}^{c-i} g_l(k) \left[\frac{z^{i+k+1}}{(s+c\mu)z - c\mu} + \xi_c(s,z) \left(\frac{c\mu}{s+c\mu} \right)^{i+k} \right] \right\} \alpha_{lm}(-s) ds$$

$$= \sum_{l=1}^{L} p_{lm} \frac{1}{2\pi \mathbf{i}} \int_{Br} \left\{ \frac{z}{(s+c\mu)z - c\mu} G_l^c(z) \Phi_l^{c-1}(z) + \xi_c(s,z) G_l^c \left(\frac{c\mu}{s+c\mu} \right) \Phi_l^{c-1} \left(\frac{c\mu}{s+c\mu} \right) \right.$$

$$\left. - \sum_{i=1}^{c-1} P(i,l) \sum_{k=1}^{c-i} g_l(k) \left[\frac{z^{i+k+1}}{(s+c\mu)z - c\mu} + \xi_c(s,z) \left(\frac{c\mu}{s+c\mu} \right)^{i+k} \right] \right\} \alpha_{lm}(-s) ds, \qquad \text{(A.23)}$$

where

$$\Phi_l^c(z) := \sum_{j=1}^{c} P(j,l) z^j,$$

and

$$G_l^c(z) := \sum_{k=1}^{c} g_l(k) z^k.$$

Next we consider the right-hand side of (21) in region C3. It equals

$$\sum_{l=1}^{L} p_{lm} \frac{1}{2\pi \mathbf{i}} \int_{Br} \sum_{k=1}^{c} g_l(k) \sum_{i=c}^{\infty} P(i,l) \left[\frac{z^{i+k+1}}{(s+c\mu)z - c\mu} + \xi_c(s,z) \left(\frac{c\mu}{s+c\mu} \right)^{i+k} \right] \alpha_{lm}(-s) ds$$

$$= \sum_{l=1}^{L} p_{lm} \frac{1}{2\pi \mathbf{i}} \int_{Br} \sum_{k=1}^{c} g_l(k) \left\{ \frac{z^{k+1}}{(s+c\mu)z - c\mu} \left[\sum_{i=0}^{\infty} P(i,l) z^i - \sum_{i=0}^{c-1} P(i,l) z^i \right] \right.$$

$$\left. + \xi_c(s,z) \left(\frac{c\mu}{s+c\mu} \right)^k \left[\sum_{i=0}^{\infty} P(i,l) \left(\frac{c\mu}{s+c\mu} \right)^i - \sum_{i=0}^{c-1} P(i,l) \left(\frac{c\mu}{s+c\mu} \right)^i \right] \right\} \alpha_{lm}(-s) ds$$

$$= \sum_{l=1}^{L} p_{lm} \frac{1}{2\pi \mathbf{i}} \int_{Br} \left[\frac{z}{(s+c\mu)z - c\mu} G_l^c(z) \Phi_l(z) - \frac{z}{(s+c\mu)z - c\mu} G_l^c(z) \Phi_l^{c-1}(z) \right.$$

$$\left. + \xi_c(s,z) G_l^c \left(\frac{c\mu}{s+c\mu} \right) \Phi_l \left(\frac{c\mu}{s+c\mu} \right) - \xi_c(s,z) G_l^c \left(\frac{c\mu}{s+c\mu} \right) \Phi_l^{c-1} \left(\frac{c\mu}{s+c\mu} \right) \right] \alpha_{lm}(-s) ds.$$

$$\text{(A.24)}$$

Finally we investigate the right-hand side of (21) in region C4. In a way similar to

(A.24), we obtain that it equals

$$\sum_{l=1}^{L} p_{lm} \sum_{k=c+1}^{c} g_l(k) \sum_{i=0}^{\infty} P(i,l) \frac{1}{2\pi \mathbf{i}} \int_{Br} \left[\frac{z^{i+k+1}}{(s+c\mu)z - c\mu} + \xi_c(s,z) \left(\frac{c\mu}{s+c\mu} \right)^{i+k} \right] \alpha_{lm}(-s)ds$$

$$= \sum_{l=1}^{L} p_{lm} \frac{1}{2\pi \mathbf{i}} \int_{Br} \left[\frac{z}{(s+c\mu)z - c\mu} G_l(z) \Phi_l(z) - \frac{z}{(s+c\mu)z - c\mu} G_l^c(z) \Phi_l(z) \right.$$

$$\left. + \xi_c(s,z) G_l \left(\frac{c\mu}{s+c\mu} \right) \Phi_l \left(\frac{c\mu}{s+c\mu} \right) - \xi_c(s,z) G_l^c \left(\frac{c\mu}{s+c\mu} \right) \Phi_l \left(\frac{c\mu}{s+c\mu} \right) \right] \alpha_{lm}(-s)ds.$$

$$\text{(A.25)}$$

Summing over the four regions, we get

$$\Phi_m(z) = \sum_{l=1}^{L} p_{lm} \frac{1}{2\pi \mathbf{i}} \int_{Br} \left\{ \frac{z}{(s+c\mu)z - c\mu} G_l(z) \Phi_l(z) + \xi_c(s,z) G_l \left(\frac{c\mu}{s+c\mu} \right) \Phi_l \left(\frac{c\mu}{s+c\mu} \right) \right.$$

$$+ \sum_{i=0}^{c-1} P(i,l) \sum_{k=1}^{c-i} g_l(k) \left[\tau_{i+k}(s,z) - \frac{z^{i+k+1}}{(s+c\mu)z - c\mu} \right.$$

$$\left. \left. - \xi_c(s,z) \left(\frac{c\mu}{s+c\mu} \right)^{i+k} \right] \right\} \alpha_{lm}(-s)ds.$$

$$\text{(A.26)}$$

Using Lemma 2 and Lemma 3 in (A.26) yields (22).

Appendix 3: Queue Size in an $\text{SMP}^{[X]}/\text{M}/c$ System When the State Sojourn Time Follows an Exponential Distribution

In this section, we show that the unknown constants contained in the generating function (22) can be determined through the zeros of the denominator for this generating function in the unit disk, when the sojourn time in the state of SMP follows an exponential distribution.

If the sojourn time A_{lm} follows an exponential distribution with mean $1/\alpha_{lm}$, equation (22) is free from the Bromwich integral, and it is reduced to

$$\Phi_m(z) = \sum_{l=1}^{L} \frac{p_{lm}}{q_{lm}(z)} \left[z G_l(z) \Phi_l(z) + (z-1) B_{lm}(z) \right]; \quad m = 1, \ldots, L, \quad \text{(A.27)}$$

where

$$q_{lm}(z) := \frac{1}{\alpha_{lm}} [(\alpha_{lm} + c\mu)z - c\mu], \quad \text{(A.28)}$$

$$B_{lm}(z) := \lambda(\alpha_{lm}, z)H_{lm} + \eta_l(\alpha_{lm}, z), \tag{A.29}$$

and

$$H_{lm} := G_l\left(\frac{c\mu}{\alpha_{lm} + c\mu}\right)\Phi_l\left(\frac{c\mu}{\alpha_{lm} + c\mu}\right). \tag{A.30}$$

We also have the relations

$$P(i, l) = \frac{1}{i!}\frac{d^i\Phi_l(z)}{dz^i}\bigg|_{z=0}; \quad l = 1, \ldots, L, \tag{A.31}$$

for $\{P(i, l); \; i = 0, \ldots, c - 1\}$ that appear in $\eta_l(\alpha_{lm}, z)$. Note that each $B_{lm}(z)$ is a polynomial in z. The set $\{B_{lm}(z); \; l, m = 1, \ldots, L\}$ contains L^2 unknown constants $\{H_{lm}; \; l, m = 1, \ldots, L\}$ and cL unknown constants $\{P(i, l); \; i = 0, \ldots, c-1; \; l = 1, \ldots, L\}$.

Now, equation (A.27) can be written in matrix form as

$$\mathbf{\Phi}(z)\mathbf{V}(z) = z\mathbf{\Phi}(z)\mathbf{G}(z)\mathbf{Q}(z) + (z - 1)\mathbf{1}\mathrm{diag}[\mathbf{B}^t(z)\mathbf{Q}(z)], \tag{A.32}$$

where

$$\mathbf{V}(z) := \begin{bmatrix} \displaystyle\prod_{j=1}^{L} q_{j1}(z) & 0 & \ldots & 0 \\ 0 & \displaystyle\prod_{j=1}^{L} q_{j2}(z) & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & \displaystyle\prod_{j=1}^{L} q_{jL}(z) \end{bmatrix}, \tag{A.33}$$

$$\mathbf{Q}(z) := \begin{bmatrix} p_{11}\displaystyle\prod_{j\neq 1} q_{j1}(z) & p_{12}\displaystyle\prod_{j\neq 1} q_{j2}(z) & \ldots & p_{1L}\displaystyle\prod_{j\neq 1} q_{jL}(z) \\ p_{21}\displaystyle\prod_{j\neq 2} q_{j1}(z) & p_{22}\displaystyle\prod_{j\neq 2} q_{j2}(z) & \ldots & p_{2L}\displaystyle\prod_{j\neq 2} q_{jL}(z) \\ \vdots & \vdots & \ddots & \vdots \\ p_{L1}\displaystyle\prod_{j\neq L} q_{j1}(z) & p_{L2}\displaystyle\prod_{j\neq L} q_{j2}(z) & \ldots & p_{LL}\displaystyle\prod_{j\neq L} q_{jL}(z) \end{bmatrix}, \tag{A.34}$$

$$\mathbf{B}(z) := \begin{bmatrix} B_{11}(z) & B_{12}(z) & \ldots & B_{1L}(z) \\ B_{21}(z) & B_{21}(z) & \ldots & B_{2L}(z) \\ \vdots & \vdots & \ddots & \vdots \\ B_{L1}(z) & B_{L2}(z) & \ldots & B_{LL}(z) \end{bmatrix}, \tag{A.35}$$

and $\mathbf{G}(z)$ is given by (35). We may rewrite (A.32) as

$$\mathbf{\Phi}(z)\mathbf{F}(z) = (z - 1)\mathbf{1}\mathrm{diag}[\mathbf{B}^t(z)\mathbf{Q}(z)], \tag{A.36}$$

where

$$\mathbf{F}(z) := \mathbf{V}(z) - z\mathbf{G}(z)\mathbf{Q}(z). \tag{A.37}$$

Let $\mathrm{adj}\mathbf{F}(z)$ denote the adjoint matrix of $\mathbf{F}(z)$. Multiplying (A.36) on the right by $\mathrm{adj}\mathbf{F}(z)$, we have

$$\mathbf{\Phi}(z) = \frac{(z-1)\mathbf{1}\mathrm{diag}[\mathbf{B}^{\mathrm{t}}(z)\mathbf{Q}(z)]\mathrm{adj}\mathbf{F}(z)}{\det \mathbf{F}(z)}. \tag{A.38}$$

Note that (A.38) contains the set of $L^2 + cL$ unknown constants $\{H_{lm}, P(i,l); i = 0, \ldots, c-1; l, m = 1, \ldots, L\}$. We will show that there are L^2 zeros for $\det \mathbf{F}(z)$ in the unit disk $|z| \leq 1$ if the condition

$$\alpha g < c\mu \tag{A.39}$$

is satisfied, which is the special case of the stability condition in (48). Here

$$\alpha := \frac{1}{\displaystyle\sum_{l=1}^{L} \pi_l \sum_{m=1}^{L} \frac{p_{lm}}{\alpha_{lm}}} \tag{A.40}$$

is the arrival rate of the batches, and g is the average batch size given in (46). Those unknown constants can be determined by solving the same number of linear equations which consist of the set of equations yielded from the zeros of $\det \mathbf{F}(z)$ in $|z| \leq 1$ and relations (A.31).

Differentiating (A.36) and evaluating the result at $z = 1$, we obtain

$$\mathbf{\Phi}'(1)(\mathbf{I}_L - \mathbf{P}) + \boldsymbol{\pi}\mathbf{F}'(1) = \mathbf{1}\mathrm{diag}[\mathbf{B}^{\mathrm{t}}(1)\mathbf{P}]. \tag{A.41}$$

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 $\mathbf{\Phi}(1) = \boldsymbol{\pi}$. Multiplying (A.41) on the right by $\mathbf{1}^{\mathrm{t}}$ and noting that $(\mathbf{I}_L - \mathbf{P})\mathbf{1}^{\mathrm{t}} = 0$, we get

$$\boldsymbol{\pi}\mathbf{F}'(1)\mathbf{1}^{\mathrm{t}} = \mathbf{1}\mathrm{diag}[\mathbf{B}^{\mathrm{t}}(1)\mathbf{P}]\mathbf{1}^{\mathrm{t}}. \tag{A.42}$$

To determine the left-hand side of (A.42), we differentiate (A.37) and evaluate the result at $z = 1$. Then we have

$$\begin{aligned}
\mathbf{F}'(1) &= \mathbf{V}'(1) - \mathbf{G}(1)\mathbf{Q}(1) - \mathbf{G}'(1)\mathbf{Q}(1) - \mathbf{G}(1)\mathbf{Q}'(1) \\
&= \mathbf{V}'(1) - \mathbf{P} - \mathbf{G}'(1)\mathbf{P} - \mathbf{Q}'(1),
\end{aligned} \tag{A.43}$$

where

$$\mathbf{V}'(1) = \begin{bmatrix} \displaystyle\sum_{j=1}^{L} \frac{\alpha_{j1} + c\mu}{\alpha_{j1}} & 0 & \cdots & 0 \\ 0 & \displaystyle\sum_{j=1}^{L} \frac{\alpha_{j2} + c\mu}{\alpha_{j2}} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \displaystyle\sum_{j=1}^{L} \frac{\alpha_{jL} + c\mu}{\alpha_{jL}} \end{bmatrix}, \tag{A.44}$$

28

$$\mathbf{Q}'(1) = \begin{bmatrix} p_{11} \displaystyle\sum_{j\neq 1} \frac{\alpha_{j1}+c\mu}{\alpha_{j1}} & p_{12} \displaystyle\sum_{j\neq 1} \frac{\alpha_{j2}+c\mu}{\alpha_{j2}} & \cdots & p_{1L} \displaystyle\sum_{j\neq 1} \frac{\alpha_{jL}+c\mu}{\alpha_{jL}} \\ p_{21} \displaystyle\sum_{j\neq 2} \frac{\alpha_{j1}+c\mu}{\alpha_{j1}} & p_{22} \displaystyle\sum_{j\neq 2} \frac{\alpha_{j2}+c\mu}{\alpha_{j2}} & \cdots & p_{2L} \displaystyle\sum_{j\neq 2} \frac{\alpha_{jL}+c\mu}{\alpha_{jL}} \\ \vdots & \vdots & \ddots & \vdots \\ p_{L1} \displaystyle\sum_{j\neq L} \frac{\alpha_{j1}+c\mu}{\alpha_{j1}} & p_{L2} \displaystyle\sum_{j\neq L} \frac{\alpha_{j2}+c\mu}{\alpha_{j2}} & \cdots & p_{LL} \displaystyle\sum_{j\neq L} \frac{\alpha_{jL}+c\mu}{\alpha_{jL}} \end{bmatrix}, \qquad (A.45)$$

and $\mathbf{G}'(1)$ is given by (41). Multiplying (A.43) on the right by $\mathbf{1}^{\mathrm{t}}$ and substituting (A.44), (41), and (A.45) yields

$$\mathbf{F}'(1)\mathbf{1}^{\mathrm{t}} = \mathbf{V}'(1)\mathbf{1}^{\mathrm{t}} - \mathbf{1}^{\mathrm{t}} - \mathbf{G}'(1)\mathbf{1}^{\mathrm{t}} - \mathbf{Q}'(1)\mathbf{1}^{\mathrm{t}}$$

$$= \begin{bmatrix} \displaystyle\sum_{j=1}^{L} \frac{\alpha_{j1}+c\mu}{\alpha_{j1}} \\ \displaystyle\sum_{j=1}^{L} \frac{\alpha_{j2}+c\mu}{\alpha_{j2}} \\ \vdots \\ \displaystyle\sum_{j=1}^{L} \frac{\alpha_{jL}+c\mu}{\alpha_{jL}} \end{bmatrix} - \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix} - \begin{bmatrix} g_1 \\ g_2 \\ \vdots \\ g_L \end{bmatrix} - \begin{bmatrix} \displaystyle\sum_{k=1}^{L} p_{1k} \displaystyle\sum_{j\neq 1} \frac{\alpha_{jk}+c\mu}{\alpha_{jk}} \\ \displaystyle\sum_{k=1}^{L} p_{2k} \displaystyle\sum_{j\neq 2} \frac{\alpha_{jk}+c\mu}{\alpha_{jk}} \\ \vdots \\ \displaystyle\sum_{k=1}^{L} p_{Lk} \displaystyle\sum_{j\neq L} \frac{\alpha_{jk}+c\mu}{\alpha_{jk}} \end{bmatrix}.$$

$$(A.46)$$

Finally, multiplying (A.46) on the left by $\boldsymbol{\pi}$, we obtain

$$\boldsymbol{\pi}\mathbf{F}'(1)\mathbf{1}^{\mathrm{t}} = \sum_{l=1}^{L} \pi_l \sum_{k=1}^{L} \frac{\alpha_{kl}+c\mu}{\alpha_{kl}} - \sum_{l=1}^{L} \pi_l - \sum_{l=1}^{L} \pi_l g_l$$

$$- \sum_{l=1}^{L} \pi_l \sum_{k=1}^{L} p_{lk} \sum_{j\neq l} \frac{\alpha_{jk}+c\mu}{\alpha_{jk}}$$

$$= \sum_{l=1}^{L} \pi_l \sum_{k=1}^{L} \frac{\alpha_{kl}+c\mu}{\alpha_{kl}} - 1 - g$$

$$- \sum_{l=1}^{L} \pi_l \sum_{k=1}^{L} p_{lk} \sum_{j=1}^{L} \frac{\alpha_{jk}+c\mu}{\alpha_{jk}} + \sum_{l=1}^{L} \pi_l \sum_{k=1}^{L} p_{lk} \frac{\alpha_{lk}+c\mu}{\alpha_{lk}}.$$

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

$$\sum_{l=1}^{L} \pi_l \sum_{k=1}^{L} p_{lk} \sum_{j=1}^{L} \frac{\alpha_{jk}+c\mu}{\alpha_{jk}} = \sum_{k=1}^{L} \sum_{j=1}^{L} \frac{\alpha_{jk}+c\mu}{\alpha_{jk}} \sum_{l=1}^{L} \pi_l p_{lk}$$

$$= \sum_{k=1}^{L} \pi_k \sum_{j=1}^{L} \frac{\alpha_{jk}+c\mu}{\alpha_{jk}}.$$

29

Thus we get

$$\boldsymbol{\pi}\mathbf{F}'(1)\mathbf{1}^{\mathrm{t}} = c\mu \sum_{l=1}^{L} \pi_l \sum_{k=1}^{L} \frac{p_{lk}}{\alpha_{lk}} - g = \frac{c\mu}{\alpha} - g. \tag{A.47}$$

From the stability condition (48), we must have

$$\alpha g < c\mu, \tag{A.48}$$

which is the condition in (A.39).

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.38). Thus we investigate the value of the derivative of $\det\mathbf{F}(z)$ at $z = 1$:

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

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

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

$$\mathbf{F}(z)\mathrm{adj}\mathbf{F}(z) = \det\mathbf{F}(z)\mathbf{I}_L = \mathrm{adj}\mathbf{F}(z)\mathbf{F}(z). \tag{A.49}$$

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

$$\gamma\mathbf{1}^{\mathrm{t}} = \mathrm{adj}\mathbf{F}(1)\mathbf{F}'(1)\mathbf{1}^{\mathrm{t}}. \tag{A.50}$$

An expression for $\mathrm{adj}\mathbf{F}(1)$ may be found as follows. Evaluating (A.49) at $z = 1$ and using $\det\mathbf{F}(1) = 0$, we have

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

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

$$\mathrm{adj}\mathbf{F}(1) = h \begin{bmatrix} \boldsymbol{\pi} \\ \vdots \\ \boldsymbol{\pi} \end{bmatrix}. \tag{A.51}$$

We claim that $\mathrm{adj}\mathbf{F}(1)$ is a positive matrix [6, p.359]. From the form of (A.51), it is enough to show that the diagonal elements, say, κ_l, $l = 1, \ldots, L$, of $\mathrm{adj}\mathbf{F}(1)$ are positive. To see this, note that

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

30

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 $\mathrm{adj}\mathbf{F}(1)$ is positive, and we conclude that $h > 0$ in (A.51).

Substituting (A.51) into (A.50) and noting (A.47) yields

$$\gamma = h\left(\frac{c\mu}{\alpha} - g\right). \tag{A.52}$$

Using $h > 0$ and the condition (A.48), we see that γ is positive. $\qquad\square$

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 [4, p.239]: *Let $f(z,t)$ be a function analytic for z within and on a closed contour $\mathcal{C}$, and continuous for t in some interval $\mathcal{I}$. If $f(z,t) \neq 0$ for $z \in \mathcal{C}$ and $t \in \mathcal{I}$, then the number of zeros of $f(z,t)$ inside $\mathcal{C}$ 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 $\mathcal{C} := \{z; |z| = 1\}$ and an interval $\mathcal{I} := \{t; t \in [0,1)\}$. Obviously, $f(z,t)$ is analytic in $\mathcal{C}$ and continuous for $t \in \mathcal{I}$. We first prove that $f(z,t) \neq 0$ for $z \in \mathcal{C}$ and $t \in \mathcal{I}$, and then prove that there are L^2 zeros for $f(z,1) = \det \mathbf{F}(z)$ in $\mathcal{C}$ 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

$$\begin{aligned}
\mathbf{F}(z,t) &= \mathbf{V}(z) - zt\mathbf{G}(z)\mathbf{Q}(z) \\
&= \mathbf{V}(z) - zt\mathbf{G}(z)\mathbf{L}(z)\mathbf{V}(z) \\
&= [\mathbf{I}_L - zt\mathbf{G}(z)\mathbf{L}(z)]\mathbf{V}(z),
\end{aligned} \tag{A.53}$$

where

$$\mathbf{L}(z) := \begin{bmatrix}
\dfrac{p_{11}}{q_{11}(z)} & \dfrac{p_{12}}{q_{12}(z)} & \cdots & \dfrac{p_{1L}}{q_{1L}(z)} \\[2ex]
\dfrac{p_{21}}{q_{21}(z)} & \dfrac{p_{22}}{q_{22}(z)} & \cdots & \dfrac{p_{2L}}{q_{2L}(z)} \\[2ex]
\vdots & \vdots & \ddots & \vdots \\[2ex]
\dfrac{p_{L1}}{q_{L1}(z)} & \dfrac{p_{L2}}{q_{L2}(z)} & \cdots & \dfrac{p_{LL}}{q_{LL}(z)}
\end{bmatrix}. \tag{A.54}$$

31

Therefore we have

$$\det \mathbf{F}(z, t) = \det[\mathbf{I}_L - zt\mathbf{G}(z)\mathbf{L}(z)] \cdot \det \mathbf{V}(z). \tag{A.55}$$

Since

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

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 that $\det[\mathbf{I}_L - zt\mathbf{G}(z)\mathbf{L}(z)]$ and $\det[\mathbf{I}_L - z\mathbf{G}(z)\mathbf{L}(z)]$ are invertible, respectively. To do this, we use the notion of maximum row sum matrix norm [6, p.295], and a corollary in [6, p.301]. We state them as follows: *Suppose there is a $L \times L$ matrix $\mathbf{A}_L = \{a_{ij}\}$. The maximum row sum matrix norm of $\mathbf{A}_L$ is defined by*

$$|||\mathbf{A}_L|||_\infty := \max_{1 \leq i \leq L} \sum_{j=1}^{L} |a_{ij}|.$$

A matrix $\mathbf{A}_L$ is invertible if there is a matrix norm (e.g. $|||\cdot|||_\infty$) such that $|||\mathbf{I}_L - \mathbf{A}_L||| < 1$, if this condition is satisfied,

$$\mathbf{A}_L^{-1} = \sum_{k=0}^{\infty} (\mathbf{I}_L - \mathbf{A}_L)^k.$$

From (35) and (A.54) we have

$$
\mathbf{I}_L - zt\mathbf{G}(z)\mathbf{L}(z)
$$
$$
= \begin{bmatrix}
1 - ztG_1(z)\dfrac{p_{11}}{q_{11}(z)} & -ztG_1(z)\dfrac{p_{12}}{q_{12}(z)} & \cdots & -ztG_1(z)\dfrac{p_{1L}}{q_{1L}(z)} \\[2ex]
-ztG_2(z)\dfrac{p_{21}}{q_{21}(z)} & 1 - ztG_2(z)\dfrac{p_{22}}{q_{22}(z)} & \cdots & -ztG_2(z)\dfrac{p_{2L}}{q_{2L}(z)} \\[2ex]
\vdots & \vdots & \ddots & \vdots \\[2ex]
-ztG_L(z)\dfrac{p_{L1}}{q_{L1}(z)} & -ztG_L(z)\dfrac{p_{L2}}{q_{L2}(z)} & \cdots & 1 - ztG_L(z)\dfrac{p_{LL}}{q_{LL}(z)}
\end{bmatrix}. \tag{A.56}
$$

For our purpose, we define

$$\mathbf{A}(z, t) := \mathbf{I}_L - zt\mathbf{G}(z)\mathbf{L}(z), \tag{A.57}$$

32

and then

$$\mathbf{I}_L - \mathbf{A}(z, t) = zt\mathbf{G}(z)\mathbf{L}(z). \tag{A.58}$$

The absolute sum of ith row for $zt\mathbf{G}(z)\mathbf{L}(z)$ on $|z| = 1$ and $0 \le t \le 1$ is satisfied the following relations

$$\sum_{j=1}^{L} \left| ztG_i(z)\frac{p_{ij}}{q_{ij}(z)} \right| \le t|G_i(z)| \sum_{j=1}^{L} p_{ij} = t|G_i(z)|. \tag{A.59}$$

Thus we have

$$\left\| \left| \mathbf{I}_L - \mathbf{A}(z, t) \right| \right\|_\infty \le t \max_i |G_i(z)|. \tag{A.60}$$

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

$$\left\| \left| \mathbf{I}_L - \mathbf{A}(z, t) \right| \right\|_\infty = \left\| \left| tz\mathbf{G}(z)\mathbf{L}(z) \right| \right\|_\infty < \max_i |G_i(z)| \le 1.$$

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

$$\left\| \left| \mathbf{I}_L - \mathbf{A}(z, t) \right| \right\|_\infty = \left\| \left| z\mathbf{G}(z)\mathbf{L}(z) \right| \right\|_\infty \le \max_i |G_i(z)| < 1.$$

From the Corollary [6, p.301], it follows that $\mathbf{I}_L - zt\mathbf{G}(z)\mathbf{L}$ is nonsingular for both $|z| = 1$, $t \in [0, 1)$, and $|z| = 1$, $z \neq 1$, $t = 1$. From (A.55), we conclude that $\det\mathbf{F}(z, t) \neq 0$ for $|z| = 1$, $t \in [0, 1)$ and $\det\mathbf{F}(z) \neq 0$ for $|z| = 1$, $z \neq 1$. $\qquad \square$

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

$$z_{ij} = \frac{c\mu}{\alpha_{ij} + c\mu}$$

in $|z| \le 1$. From Theorem 2, 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 $0 \le t < 1$ as shown above. By continuity, $\det \mathbf{F}(1, t) > 0$ for $0 \le t < 1$, 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 $|z| \le 1$. From Theorem 2(b), $\det \mathbf{F}(z)$ has $L^2 - 1$ zeros in $|z| < 1$. $\qquad \square$