1

# STUDIES ON DISCRETE-TIME QUEUES WITH CORRELATED ARRIVALS

Fumio Ishizaki

March 1996

# Preface

Discrete-time queues have been studied by many researchers, one of the earliest investigations being made by Meisling [Meis58]. Kobayashi and Konheim [Koba77] have provided surveys of applications of discrete-time queues to the performance evaluation of communication systems. Takagi [Taka93] has presented the analysis of many discrete-time queues including discrete-time queues with vacations, discrete-time priority queues, discrete-time queues with finite buffer. Bruneel and Kim [Brun93] have analyzed many discrete-time queues including discrete-time queues with service interruptions, and multiple servers. In most of the analyses, arrivals have been assumed to occur independently from one time slot to another. Also, the arrival and the service processes have been assumed to be independent each other. Since these simplified assumptions are not realistic in many situations, several mechanisms have been proposed to bring correlations into the arrival and the service processes.

The main contribution of this dissertation is that the development of the analytical methods for the various types of discrete-time queues with correlated arrivals. Chapters 2 and 3 are devoted to the analysis of discrete-time queues with Markov modulation. In chapter 2, we consider discrete-time queues with a generalized switched batch Bernoulli arrival and a general service time processes. The arrival and the service processes in this system are semi-Markovian in the sense that their distributions depend not only on the state of the alternating renewal process in the current slot but also on the state in the next slot. Furthermore, sojourn times in each state are generally distributed. In chapter 3, we consider DBMAP/D/1/K queues. We develop the approximate formulas for the loss probability. The accuracy of the approximations are extensively examined through numerical experiments. The results are readily applied to call admission control in high speed networks. Chapters 4 and 5 are devoted to the analysis of discrete-time queues with a gate. In chapter 4, we consider discrete-time BBP/G/1 queues with a gate, where the intervals between successive openings of the gate are geometrically distributed. On the other hand, in chapter 5, we consider discrete-time BBP/G/1 queues with a gate, where the intervals between successive openings of the gate are bounded, and independent and identically distributed. In chapters 4 and 5, we derive complete sets of the analytical results for various performance measures.

The author believes that the analytical methods developed in this dissertation posses the high applicability to performance evaluation of communication systems. Finally, the author would like to hope that the studies in this dissertation will be helpful for future research in this field.

March 1996

Fumio Ishizaki



# Acknowledgment

I would like to express my sincere appreciation to Professor Toshiharu Hasegawa of Kyoto University for his encouragement and supervision through the course of this research. Talking with him about both technical and non-technical topics was very stimulating and interesting.

I am highly indebted to Associate Professor Tetsuya Takine of Osaka University for his insightful suggestions and stimulating discussions. In my graduate course, it was my pleasure to have the opportunity to attend his seminar on queueing theory and stochastic process. I was so excited to study queueing theory, stochastic process and their applications in many informal seminars directed by him. Without his support, none of this work would have been done. Special thank also goes to him for his intensive reading of this manuscript and accurate comments.

I am heartily grateful to Associate Professor Yutaka Takahashi of Kyoto University for his enlightening suggestions and enthusiastic discussions. Moreover, I am thankful that he furnished a great amount of information about the state of the art in the performance evaluation of communication systems.

Thanks are due to Professor Hideo Miyahara of Osaka University, Professor Tatsuya Suda of University of California, Irvine, Professor Yuji Oie of Nara Institute of Science and Technology, Professor Hiroaki Terada of Osaka University, Professor Ryosaku Shimada of the University of Tokushima, Professor Koji Kajitani of Kinki University, Associate Professor Masayuki Murata of Osaka University, Associate Professor Kozo Mayumi of the University of Tokushima, Assistant Professor Hiroyuki Kawano of Kyoto University, Assistant Professor Shoji Kasahara of Kyoto University, Assistant Professor Naoto Miyoshi of Kyoto University and all my friends for their encouragement and enthusiastic discussions.

# Contents

# List of Figures

# List of Tables





# Chapter 1

# Introduction

Discrete-time queues have been extensively studied by many researchers (see [Taka93], [Brun93] and references therein). They have not only their theoretical interests but also their rich applicability, for example, to the performance evaluation of communication systems which are operated based on a time-slot basis. In most of the analyses, arrivals have been assumed to occur independently from one time slot to another. Also, the arrival and the service processes have been assumed to be independent each other. However, these simplified assumptions are not realistic in many situations. For instance, in a communication system, the arrival process to the system is in some sense *correlated* because data sources usually generate data in a *bursty* fashion, i.e., a source which is active in a given slot is likely to remain inactive for a large number of consecutive slots. Thus, the arrivals occur back to back. On the other hand, in high-speed local and metropolitan area networks, modern reservation protocols for the use of transmission slots give rise to customer collection. If we regard the collected customers as a *supercustomer*, the service time of a supercustomer depends on the previous interarrival time. Indeed, if the previous interarrival time of a supercustomer is relatively long (short), then the service time of the supercustomer is likely to be relatively long (short) because the supercustomer consists of relatively many (few) customers. Thus, there exists positive correlation between the interarrival time and the service time. In order to deal with these kinds of situations, discrete-time queues *with correlated arrivals* are required. In this chapter, a brief survey of discrete-time queues with correlated arrivals is provided.

## 1.1 Discrete-Time Queues with Correlated Arrivals

In the present dissertation, we consider discrete-time queues where time axis is divided into intervals (called *slot*) which are equal in length. Through the dissertation, it is always assumed that all events start and end only at slot boundaries. Thus, their durations are integer multiples of a slot. For example, customers are assumed to arrive to the system at the beginning of a slot. The service times are positive integer multiples of a slot. Any departure from the system is assumed to take place at the end of a slot.

To bring correlations into the arrival and the service processes, several mechanisms have been proposed. Among them, one of the most well-known mechanisms is Markov modulation, which gives a doubly stochastic process driven by an underlying Markov chain. The mechanism includes a switched batch Bernoulli process (SBBP), a discrete-time batch Markovian arrival process (DBMAP) and their variants. Note that sojourn times in each state of a Markov chain are geometrically distributed (whereas the sojourn time in a state can be deterministic and equal to one slot). An SBBP is defined as a doubly stochastic batch Bernoulli process with batch size governed by a two-state Markov chain. It can be viewed as the discrete-time analogue of a switched batch Poisson process. In other words, the batch size arriving in a slot, which may take zero, depends on the state of the underlying Markov chain in the current slot. Furthermore, a generalized switched batch Bernoulli process has been proposed. It can be viewed as the discrete-time analogue of a generalized switched batch Poisson process. The term *generalized* means that the sojourn times in each state is not necessary to be geometrically distributed. Thus, the arrival process is governed not by a two state Markov chain but by an alternating renewal process. On the other hand, a DBMAP is defined as a doubly stochastic batch Bernoulli process with batch size governed by a finite-state Markov chain. It can be viewed as the discrete-time analogue of a batch Markovian arrival process (BMAP). In other words, the batch size arriving in a slot, which may take zero, depends not only on the state of the underlying Markov chain in the current slot but also on the state in the next slot.

Another type of the mechanisms is *gating*, which gives rise to customer collection. The system has two queues and a gate (see Fig. 1.1). In this system, customers arriving at the system are accommodated in the first queue at the gate. When the gate opens, all the customers who are waiting in the first queue move to the second queue at the server. The gate closes immediately after all the customers in the first queue move to the second queue. The server serves only the customers present in the second queue. Hereafter, we call the system *queues with a gate*.

In queues with a gate, we can observe several types of correlation. It is easy to see that, when a gate opening interval is (relatively) long, (relatively) many customers are likely to wait in the first queue, while (relatively) few customers are likely to wait in the second queue. Thus, the waiting times in the first queue and the second queue would be negatively correlated (i.e., a long waiting in the first queue leads to a short waiting in the second queue). Also, if we consider the second queue as an isolated system, the interarrival time of batches arriving at the second queue (i.e., the gate opening interval) and the number of customers in each batch (i.e., the number of customers who move to



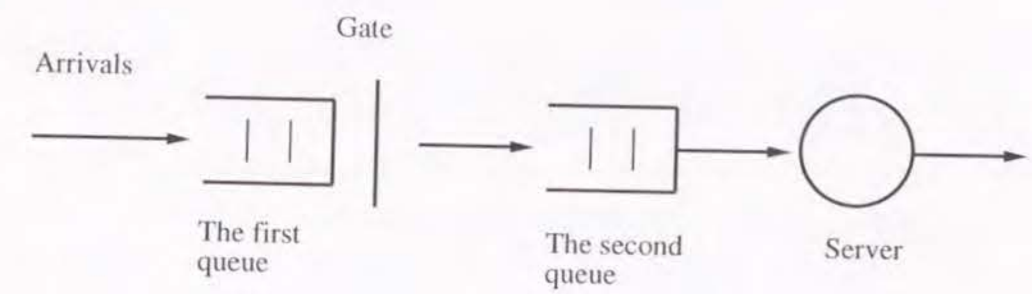


Figure 1.1: Queues with a Gate

the second queue at the same time) are positively correlated. Yet another view of this feature is that there exists the correlation between the interarrival time and the service time if each batch moving to the second queue is considered as a *supercustomer*.

From a theoretical viewpoint, queues with a gate falls into the category of queues with generalized vacations [Boxm89], [Fuhr85]. In the queue with generalized vacations, a server takes vacations even when waiting customers are present in the system. Note that, in the queue with a gate, there is a possibility that a server becomes idle, while customers are waiting outside the gate. Thus the idle periods of the server when waiting customers exist outside the gate are considered as vacations of the server. It is well known that in the queue with generalized vacations, the queue length, the amount of work in the system and the waiting time under the FCFS discipline have the so-called decomposition properties (see [Dosh90] and references therein). Note that the queue with generalized vacations has been studied mainly in the continuous-time model. However, very similar decomposition properties hold for the discrete-time counterpart, too. See, for example, [Boxm88].

## 1.2 Previous Works

Several types of discrete-time queues with correlated arrivals have been studied. In this section, we will review the previous works which concerns with the results given in this dissertation.

### 1.2.1 SBBP/G/1 Queues

Hashida et al. [Hash91] have proposed and analyzed SBBP as a modeling tool for a bursty and correlated input process to discrete-time queues. They have assumed that the service times of customers are independent and identically distributed, and the batch size arriving in a slot depends only on the state of the underlying Markov chain in the current slot. They have derived the probability generating functions (PGFs) for the number of customers in the system, the amount of the stationary work in the system and the waiting time of a customer.

### 1.2.2 Generalized SBBP/D/1 Queues

Liao and Mason [Liao89] have studied generalized SBBP/D/1 queues. They have assumed that the service times of customers are deterministic and equal to one slot, and the batch size depends only on the state of the underlying alternating renewal process in the current slot. Sojourn times in each state are generally distributed. The restriction they place on their mathematical model is that the PGFs of the sojourn time distributions are represented as rational functions. They have derived the PGF for the number of customers in the system.

### 1.2.3 DBMAP/D/1/K Queues

Blondia and Casals [Blon92] have studied DBMAP/D/1/K queues by a matrix analytical method. They have assumed that the service times of customers are deterministic and equal to one slot, and the batch size depends not only on the state of the underlying Markov chain in the current slot but also on the state in the next slot. They have derived the customer loss probability.

Takine et al. [Taki95] have studied DBMAP/D/1/K queues. They have presented analysis for various loss characteristics such as the loss probability, the consecutive loss probability and the distribution of the loss period length. Furthermore, they have presented the output process analysis and derived expressions for various statistics of the output process, including the joint distribution of the successive interdeparture times.

### 1.2.4 BBP + DBMAP/D/1 Queues

Takine et al. [Taki94] have analyzed BBP + DBMAP/D/1 queues, where BBP denotes a batch Bernoulli process. They have assumed that DBMAP customers have priority over BBP customers, and the service times of both DBMAP and BBP customers are deterministic and equal to one slot. Furthermore, they have assumed that the batch size of DBMAP customers depends not only on the state of the underlying Markov chain in the current slot but also on the state in the next slot, and the batch size of BBP customers is independent and identically distributed. They have derived the PGFs for the waiting times of both customers and the numbers of both customers in the system.



### 1.2.5 Queues with a Gate

Takahashi [Taka71] has studied continuous-time queues with a gate where the service times of customers are exponentially distributed and the gate opening intervals are deterministic or exponentially distributed. Borst et al. [Bors92], [Bors93] have studied the continuous-time queue with exponential gate opening intervals where the service times of customers are generally distributed. They were mainly concerned with the second queue, and discussed the effect of the correlation between the interarrival time and the number of customers in a batch on the performance of the second queue. Boxma and Combé [Boxm93] have studied an M/G/1 queue with a rather general dependency between the interarrival time and the service time. Kawata [Kawa93] has studied a discrete-time queue with geometrically distributed gate opening intervals and derived the probability generating function (PGF) for the sojourn times of a supercustomer.

## 1.3 Overview of the Dissertation

The purpose of the present dissertation is to provide a complete set of the analytical results on various types of discrete-time queues with correlated arrivals. They have a rich applicability of the mathematical model to important fundamental queueing systems and the performance evaluation of communication systems. We also provide numerical examples to show the computational feasibility of the analytical results. In view of the performance evaluation, one of the most interesting subject is the effect of correlations on the performance. This subject will be observed repeatedly through the numerical examples.

The organization of this dissertation is as follows. In the first half of this dissertation, we consider discrete-time queues with Markov modulation. In the latter half of this dissertation, we consider discrete-time queues with a gate.

In chapter 2, we consider a discrete-time queue with a generalized switched batch Bernoulli arrival and a general service time processes. The arrival and the service processes in this system are semi-Markovian in the sense that their distributions depend not only on the state of the alternating renewal process in the current slot but also on the state in the next slot. We derive the PGFs for the amount of work in the system and the waiting time of a customer under the FCFS (first-come, first-served) discipline. We also provide numerical examples to show the computational feasibility of the analytical results. Furthermore, we show a rich applicability to important queueing systems such as a discrete-time $GI^{[X]}/G/1$ queue and a discrete-time queue with service interruptions.

In chapter 3, we study the loss probability approximation in DBMAP/D/1/K queues. We propose the approximate formulas which are given in terms of the tail distribution of the queue length in the corresponding infinite-buffer queue. The approximate formulas become exact for any independent arrival process. We provide numerical experiments to show the accuracy of the approximations.

In chapter 4, we consider discrete-time BBP/G/1 queues with a gate, where BBP denotes a batch Bernoulli process. In this model, we assume that the intervals between successive openings of the gate are geometrically distributed. We provide a complete set of the analytical results for various performance measures. We also show some numerical examples, where we discuss the effect of three kinds of correlations in the model on the performance measures.

In chapter 5, we consider discrete-time BBP/G/1 queues with a gate. Contrary to the model considered in chapter 4, we assume that the intervals between successive openings of the gate are bounded, and independent and identically distributed (i.i.d.). We provide a complete set of the analytical results for various performance measures. We also show numerical examples and observe the effect of the distribution of the gate opening intervals on the performance measures.

Finally, in chapter 6, we provide concluding remarks.

The results discussed in chapter 2 is mainly taken from [Ishi93a] and [Ishi94c], chapter 3 from [Ishi94b] and [Ishi95b], chapter 4 from [Ishi96], chapter 5 from [Ishi93b], [Ishi94a] and [Ishi95a].



# Chapter 2

# Generalized SBBP/G/1 Queues

## 2.1 Introduction

In this chapter, we consider a discrete-time queue with a generalized switched batch Bernoulli arrival and a general service time processes. In this system, customers arrive to the system in batches and the service times of customers are generally distributed. The batch size and the service time distributions are governed by a discrete-time alternating renewal process with states 1 and 2. The arrival and the service processes in this system are semi-Markovian in the sense that their distributions depend not only on the state of the alternating renewal process in the current slot but also on the state in the next slot. Sojourn times in each state are generally distributed. The only restriction we place on our mathematical model is that the probability generating functions (PGFs) for the sojourn time are represented as rational functions.

The purpose of this chapter is two-fold. The first is to provide the analytical results in a fairly general assumption on discrete-time queues with two-state Markov modulation. The analytical results are readily applied to the performance evaluation of various communication systems [Ishi93a]. The second is to show a rich applicability of the mathematical model to important queueing systems such as a discrete-time $GI^{[X]}/G/1$ queue and a discrete-time queue with service interruptions. The discrete-time $GI^{[X]}/G/1$ queue is a fundamental queueing model and it has enormous potentialities to study queueing phenomena in general. On the other hand, queueing systems with service interruptions have wide applications to manufacturing, computer and communication systems where the server is subject to breakdown. Using the semi-Markovian nature of the arrival and the service mechanisms in our model, our model can be readily applied to analysis of those systems.

The organization of this chapter is as follows. In section 2.2, we describe the mathematical model of the generalized SBBP/G/1 queue. Note that, since the batch size and the service time distributions are semi-Markovian, we need four distinct notations to describe each of these distributions. In section 2.3, we analyze the generalized SBBP/G/1 queue and derive the PGFs for the amount of work in the system and the waiting time of a customer under the FCFS (first-come, first-served) discipline. We also provide numerical examples to show the computational feasibility of the analytical results. In sections 2.4–2.6, we show an application of the analytical results to important queueing systems such as a discrete-time GI$^{[X]}$/G/1 queue and a discrete-time queue with service interruptions.

## 2.2 Model

We consider the queueing model with the following characteristics:

- The system operates in a random environment defined by an alternating renewal process with state 1 and state 2. Each state starts from a slot boundary and ends immediately before a slot boundary. We assume that the alternating renewal process is stationary.
- Customers arrive to the system in a batch immediately before slot boundaries. The batch size may possibly be zero. The arrival and the service time processes are semi-Markovian in the sense that the batch size and the service time distributions of customers arriving in a slot depend not only on the state of the alternating renewal process in the current slot but also on the next state.
- There is a single server and the service discipline is work-conserving. Namely, when the server finds some amount of work immediately after a slot boundary, he serves exactly one unit of work in the current slot.

We now introduce random variables and notations to describe the above model. Let $\hat{T}_n$ denote a random variable representing the $n$th state transition epoch of the alternating renewal process taking a integer value. We assume that the sequence $\{\hat{T}_n\}_{n=-\infty}^{+\infty}$ satisfies

$$\cdots < \hat{T}_{-1} < \hat{T}_0 \le 0 < \hat{T}_1 < \hat{T}_2 < \cdots. \tag{2.1}$$

Let $P_n$ denote a random variable representing the state of the alternating renewal process at time $n$. We define the inter-event sequence $\{G_n\}_{n=-\infty}^{+\infty}$ as

$$G_n = \hat{T}_{n+1} - \hat{T}_n. \tag{2.2}$$

Let $B_n$ and $C_{n,m}$ $(m = 1, \ldots, B_n)$ denote random variables representing the batch size arriving at time $n$, which may take zero, and the service time of $m$th customer within the batch, respectively. Also, let $A_n$ denote a random variable representing the amount of



work brought into the system by the batch (i.e., the sum of the service times of customers who belongs to the batch). We define the following PGFs:

$$A_{ij}(z) \triangleq E\left[z^{A_n}\middle| P_n = i, P_{n+1} = j\right], \quad B_{ij}(z) \triangleq E\left[z^{B_n}\middle| P_n = i, P_{n+1} = j\right],$$
$$C_{ij}(z) \triangleq E\left[z^{C_{n,m}}\middle| P_n = i, P_{n+1} = j\right] \qquad (i, j = 1, 2). \tag{2.3}$$

We then have the following relationship:

$$A_{ij}(z) = B_{ij}(C_{ij}(z)) \qquad (i, j = 1, 2). \tag{2.4}$$

Sojourn times in state $i$ $(i = 1, 2)$ are distributed in accordance with a general distribution function. Let $g_i(k)$ denote the probability mass function of sojourn times in state $i$, i.e.,

$$g_i(k) = \Pr\left\{G_n = k \middle| P_{\hat{T}_n} = i\right\} \qquad (k = 1, 2, \ldots, n \ne 0, i = 1, 2). \tag{2.5}$$

We denote the PGF of the $g_i(k)$ by $G_i(z)$ $(i = 1, 2)$:

$$G_i(z) = \sum_{k=1}^{\infty} g_i(k) z^k \qquad (i = 1, 2). \tag{2.6}$$

The overall traffic intensity $\rho$ is then given by

$$\rho = \frac{A'_{11}(1)(G'_1(1) - 1) + A'_{22}(1)(G'_2(1) - 1) + A'_{12}(1) + A'_{21}(1)}{G'_1(1) + G'_2(1)}. \tag{2.7}$$

In the above and following equations, we use the symbol $f'(1)$ and $f''(1)$ to denote $\lim_{z\to 1} df(z)/dz$ and $\lim_{z\to 1} d^2 f(z)/dz^2$ for any function $f(z)$, respectively.

In the remainder of this chapter, we assume that $\rho < 1$ and the system is in equilibrium. Furthermore, we assume that the PGF $G_i(z)$ $(i = 1, 2)$ is represented as a rational function of $z$, which is the only restriction in our model. Namely, we assume that $G_i(z)$ can be written as

$$G_i(z) = G_{i1}(z) + G_{i2}(z) \qquad (i = 1, 2), \tag{2.8}$$

where $G_{i1}(z)$ is given by a polynomial:

$$G_{i1}(z) = \sum_{l=1}^{M_i} m_{il} z^l, \tag{2.9}$$

and $G_{i2}(z)$ is given by the fraction of two polynomials:

$$G_{i2}(z) = \frac{\displaystyle\sum_{j=1}^{N_i} n_{ij} z^j}{\displaystyle\prod_{k=1}^{K_i} (1 - \alpha_{ik} z)^{w_{ik}}}. \tag{2.10}$$

The degree of the numerator of (2.10) is not higher than the degree of the denominator; however, without loss of generality, we assume that

$$\sum_{k=1}^{K_i} w_{ik} = N_i. \tag{2.11}$$

## 2.3 Analysis of Generalized SBBP/G/1 Queues

In this section, we analyze the generalized SBBP/G/1 queues. For convenience of the analysis, we assign non-negative integer values $k \in \{0, 1, 2, \ldots\}$ sequentially to individual slot boundaries in each state. The time interval $[k-1, k)$ $(k = 1, 2, \ldots)$ in state $i$ $(i = 1, 2)$ is referred to as the $k$th slot of state $i$. In what follows, we refer to the amount of work at the beginning of the $k$th slot as that in the $k$th slot. We first observe the amount of work in the system at the beginning of the $(k+1)$st slot and relate it with that at the beginning of the $k$th slot. We then consider the amount of the stationary work in the system. Finally we consider the waiting time of a randomly chosen customer.

### 2.3.1 Work in the First Slot of Each State

In this subsection, we first observe the amount of work in the system immediately after the beginning of the $(k+1)$st slot and relate it with that immediately after the beginning of the $k$th slot. Let $U_n$ denote a random variable representing the amount of work in the system at time $n$. We define $L_i(k, z)$ as the PGF for the amount of work in the system given that the alternating renewal process is in the $(k+1)$st slot of state $i$ at time 0:

$$L_i(k, z) \triangleq E\left[z^{U_0} \middle| P_0 = i, -\hat{T}_0 = k\right] \qquad (i = 1, 2, k = 0, 1, 2, \ldots). \tag{2.12}$$

Relating $L_i(k, z)$ with $L_i(k-1, z)$, we have

$$L_i(k, z) = \left[\frac{L_i(k-1, z) - L_i(k-1, 0)}{z} + L_i(k-1, 0)\right] A_{ii}(z). \tag{2.13}$$

By applying (2.13) recursively, we obtain $L_i(k, z)$ in terms of $L_i(0, z)$ which denotes the PGF for the amount of work at the beginning of state $i$:

$$L_i(k, z) = \left(\frac{A_{ii}(z)}{z}\right)^k L_i(0, z) + (z-1)\sum_{j=1}^{k}\left(\frac{A_{ii}(z)}{z}\right)^j L_i(k-j, 0). \tag{2.14}$$

Since the system is in equilibrium, it is clear that the PGFs $L_i(0, z)$ are given by

$$L_1(0, z) = \sum_{k=1}^{\infty} L_2(k, z) g_2(k) \frac{A_{21}(z)}{A_{22}(z)}, \qquad L_2(0, z) = \sum_{k=1}^{\infty} L_1(k, z) g_1(k) \frac{A_{12}(z)}{A_{11}(z)}. \tag{2.15}$$



Substituting (2.14) into (2.15), we have the following expressions for $L_i(0, z)$ $(i = 1, 2)$:

$$L_1(0, z) = \frac{A_{21}(z)}{A_{22}(z)}\left[G_2\left(\frac{A_{22}(z)}{z}\right) L_2(0, z) + (z-1) X_2\left(\frac{A_{22}(z)}{z}\right)\right], \tag{2.16}$$

$$L_2(0, z) = \frac{A_{12}(z)}{A_{11}(z)}\left[G_1\left(\frac{A_{11}(z)}{z}\right) L_1(0, z) + (z-1) X_1\left(\frac{A_{11}(z)}{z}\right)\right], \tag{2.17}$$

where

$$X_i(z) \triangleq \sum_{m=1}^{\infty}\sum_{k=0}^{\infty} L_i(k, 0) g_i(k+m) z^m \qquad (i = 1, 2). \tag{2.18}$$

We solve (2.16) and (2.17) with respect to $L_i(0, z)$ $(i = 1, 2)$ and obtain

$$L_1(0, z) = \frac{(z-1)[A_{12}(z)A_{21}(z)G_2(A_{22}(z)/z)X_1(A_{11}(z)/z) + A_{11}(z)A_{21}(z)X_2(A_{22}(z)/z)]}{A_{11}(z)A_{22}(z) - A_{12}(z)A_{21}(z)G_1(A_{11}(z)/z)G_2(A_{22}(z)/z)}, \tag{2.19}$$

$$L_2(0, z) = \frac{(z-1)[A_{12}(z)A_{21}(z)G_1(A_{11}(z)/z)X_2(A_{22}(z)/z) + A_{22}(z)A_{12}(z)X_1(A_{11}(z)/z)]}{A_{11}(z)A_{22}(z) - A_{12}(z)A_{21}(z)G_1(A_{11}(z)/z)G_2(A_{22}(z)/z)}. \tag{2.20}$$

Thus, once $X_1(z)$ and $X_2(z)$ are known, the $L_i(0, z)$'s are obtained. In order to determine $X_i(z)$, we need the following lemma.

*Lemma 2.3.1:* The unknown function $X_i(z)$ can be rewritten as

$$X_i(z) = X_{i1}(z) + X_{i2}(z) \quad (i = 1, 2), \tag{2.21}$$

where

$$X_{i1}(z) = \sum_{j=1}^{M_i} x_i^*(j) z^j, \qquad X_{i2}(z) = \frac{\sum_{j=1}^{N_i} x_i^{**}(j) z^j}{\prod_{k=1}^{K_i}(1 - \alpha_{ik} z)^{w_{ik}}} \qquad (i = 1, 2). \tag{2.22}$$

PROOF: See [Brun84]. ∎

From (2.22), it is clear that the unknown functions $X_i(z)$'s contain only a finite number of unknown constants $x_i^*(j)$'s and $x_i^{**}(j)$'s when $G_i(z)$ $(i = 1, 2)$ is a rational function of $z$. Furthermore, $X_i(z)$ $(i = 1, 2)$ is a rational function of $z$, having the same denominator as $G_i(z)$ and a numerator of degree $M_i + N_i$. This formal similarity between the $G_i(z)$ and the $X_i(z)$ allows us to determine $X_i(z)$ (see Appendix A and also Appendix D).

### 2.3.2 Stationary Work

In this subsection, we consider the amount of the stationary work in the system. We define $U_i(z)$ $(i = 1, 2)$ as the PGF for the amount of work in the system in a randomly chosen slot given that the state of the alternating renewal process in the slot is $i$:

$$U_i(z) \triangleq E\left[z^{U_0}\middle| P_0 = i\right] \qquad (i = 1, 2). \tag{2.23}$$

We also define $U(z)$ as the PGF for the amount of the stationary work in the system:

$$U(z) \triangleq E\left[z^{U_0}\right]. \tag{2.24}$$

We then have the following theorem.

THEOREM 2.3.1: *The PGFs $U_i(z)$ and $U(z)$ are given by*

$$\begin{aligned} U_i(z) &= \frac{1}{G_i'(1)\{1 - A_{ii}(z)/z\}} \left[L_i(0, z)\left\{1 - G_i\left(\frac{A_{ii}(z)}{z}\right)\right\}\right. \\ &\left. +(z-1)\left\{\frac{A_{ii}(z)}{z}X_i(1) - X_i\left(\frac{A_{ii}(z)}{z}\right)\right\}\right]. \end{aligned} \tag{2.25}$$

$$U(z) = \pi_1 U_1(z) + \pi_2 U_2(z), \tag{2.26}$$

*respectively, where $\pi_i$ is given by*

$$\pi_i = \frac{G_i'(1)}{G_1'(1) + G_2'(1)} \qquad (i = 1, 2). \tag{2.27}$$

PROOF: We first consider $U_i(z)$. Let $p_i(k)$ $(i = 1, 2, k = 0, 1, 2, \ldots)$ denote the conditional probability that a randomly chosen slot of state $i$ is the $(k+1)$st slot of state $i$, given that the slot is in state $i$:

$$p_i(k) \triangleq \Pr\left\{-\hat{T}_0 = k \middle| P_0 = i\right\} \qquad (k = 1, 2, \ldots, i = 1, 2). \tag{2.28}$$

We then have [Burk75]

$$p_i(k) = \frac{1}{G_i'(1)} \sum_{n=k+1}^{\infty} g_i(n) \quad (k = 0, 1, 2, \ldots). \tag{2.29}$$

By definition, we have

$$U_i(z) = \sum_{k=0}^{\infty} p_i(k) L_i(k, z) \qquad (i = 1, 2). \tag{2.30}$$

From (2.14), (2.18), (2.29) and (2.30), we obtain (2.25).

Next we consider $U(z)$. Since

$$\Pr\{P_0 = i\} = \frac{G_i'(1)}{G_1'(1) + G_2'(1)} \qquad (i = 1, 2), \tag{2.31}$$

(2.26) immediately follows. ■



### 2.3.3 Waiting Time

In this subsection, we consider the waiting time of a randomly chosen customer. We assume here that the service discipline is the FCFS and as for customers who arrived in the same batch, the next customer for service is randomly selected among them.

Since our system is a single-server queue with batch arrivals, a tagged customer suffers from two components of delay. One is the waiting time of the tagged batch to which the tagged customer belongs. The other is the waiting time due to service times of customers in the tagged batch, who are served before the tagged customer. For the former, we define $F_{ij}(z)$ as the PGF for the waiting time of the tagged batch to which a randomly chosen customer belongs arriving in a slot given that the alternating renewal process is in state $i$ in the current slot and in state $j$ in the next slot. For the latter, we define $D_{ij}(z)$ as the PGF for a time interval from the beginning of the service of the first customer in the tagged batch to the beginning of the service of the customer given that he arrives to the system when the alternating renewal process is in state $i$ in the current slot and in state $j$ in the next slot. Further, we define $W(z)$ as the PGF for the waiting time of a randomly chosen customer. We then have the following theorem.

THEOREM 2.3.2: *The PGF $W(z)$ is given by*

$$W(z) = \sum_{i=1}^{2}\sum_{j=1}^{2} \nu_{ij} F_{ij}(z) D_{ij}(z), \tag{2.32}$$

*where*

$$\nu_{ij} = \begin{cases} \dfrac{B_{ii}'(1)(G_i'(1) - 1)}{B_{11}'(1)(G_1'(1) - 1) + B_{22}'(1)(G_2'(1) - 1) + B_{12}'(1) + B_{21}'(1)} & (i = j), \\[2ex] \dfrac{B_{ij}'(1)}{B_{11}'(1)(G_1'(1) - 1) + B_{22}'(1)(G_2'(1) - 1) + B_{12}'(1) + B_{21}'(1)} & (i \neq j), \end{cases} \tag{2.33}$$

$$F_{ii}(z) = \frac{1}{A_{ii}(z)} \frac{G_i'(1)U_i(z) - L_i(0, z)}{G_i'(1) - 1} \qquad (i = 1, 2), \tag{2.34}$$

$$F_{ij}(z) = \frac{L_j(0, z)}{A_{ij}(z)} \qquad (i \neq j, i, j = 1, 2), \tag{2.35}$$

$$D_{ij}(z) = \frac{1 - B_{ij}(C_{ij}(z))}{B_{ij}'(1)(1 - C_{ij}(z))} \qquad (i, j = 1, 2). \tag{2.36}$$

PROOF: Note that $\nu_{ij}$ denotes the probability that the tagged customer arrives when the alternating renewal process is in state $i$ in the current slot and in state $j$ in the next slot. We then obtain the expression (2.32) for $W(z)$.

We now consider $F_{ij}(z)$. Recall that we observe the amount of work in the system immediately after slot boundaries, while customers arrive immediately before slot boundaries. Thus, we have

$$U_i^+(z) = F_{ii}(z)A_{ii}(z) \qquad (i = 1, 2), \tag{2.37}$$

where $U_i^+(z)$ $(i = 1, 2)$ denotes the PGF for the amount of work in the system in a randomly chosen slot of state $i$ given that it is not the first slot of state $i$. By definition, $U_i^+(z)$ is given by

$$U_i^+(z) = \frac{1}{1 - p_i(0)} \sum_{k=1}^{\infty} p_i(k) L_i(k, z) = \frac{G_i'(1)U_i(z) - L_i(0, z)}{G_i'(1) - 1} \qquad (i = 1, 2). \tag{2.38}$$

From (2.37) and (2.38), we obtain (2.34). According to a similar reasoning, we obtain (2.35). We next consider $D_{ij}(z)$. By using the results in batch arrivals [Burk75], we have (2.36). ∎

### 2.3.4 Numerical Example

In this subsection, we provide a numerical example in order to show the computational feasibility of the analytical results. In this example, we assume that the service times of customers are deterministic and equal to one slot regardless of the state, and the batch size follows a Poisson distribution. Moreover we assume that $A_{12}(z) = A_{11}(z)$ and $A_{21}(z) = A_{22}(z)$. We denote the mean sojourn time in state $i$ by $r_i^{-1}$. We assume that $r_1 = r_2 = r$. The traffic intensities in states 1 and 2 are fixed to 0.5 and 1.3, respectively, and the overall traffic intensity is given by 0.9. Thus, the arrival process consists of underload periods (state 1) and overload periods (state 2).

Fig.2.1 shows that the mean waiting time of customers for two cases: the deterministic and the geometric sojourn times. The mean waiting time is plotted as a function of the mean sojourn time $r^{-1}$. We first observe that the increase of the mean sojourn time leads to the increase of the mean waiting time of customers even when the overall traffic intensity is fixed. The second observation is that the mean waiting time in the case of the geometric sojourn times becomes longer than that in the case of the deterministic sojourn times. Thus not only the mean sojourn time but also the sojourn time distribution affects the mean waiting time.

## 2.4 Discrete-Time $\mathrm{GI}^{[X]}/\mathrm{G}/1$ Queues

In this section, we show an application of the analytical results to a discrete-time $\mathrm{GI}^{[X]}/\mathrm{G}/1$ queue. The discrete-time $\mathrm{GI}^{[X]}/\mathrm{G}/1$ queue has been extensively analyzed in the literature (see [Ackr80], [Mura91] and references therein). Among those, Murata and Miyahara



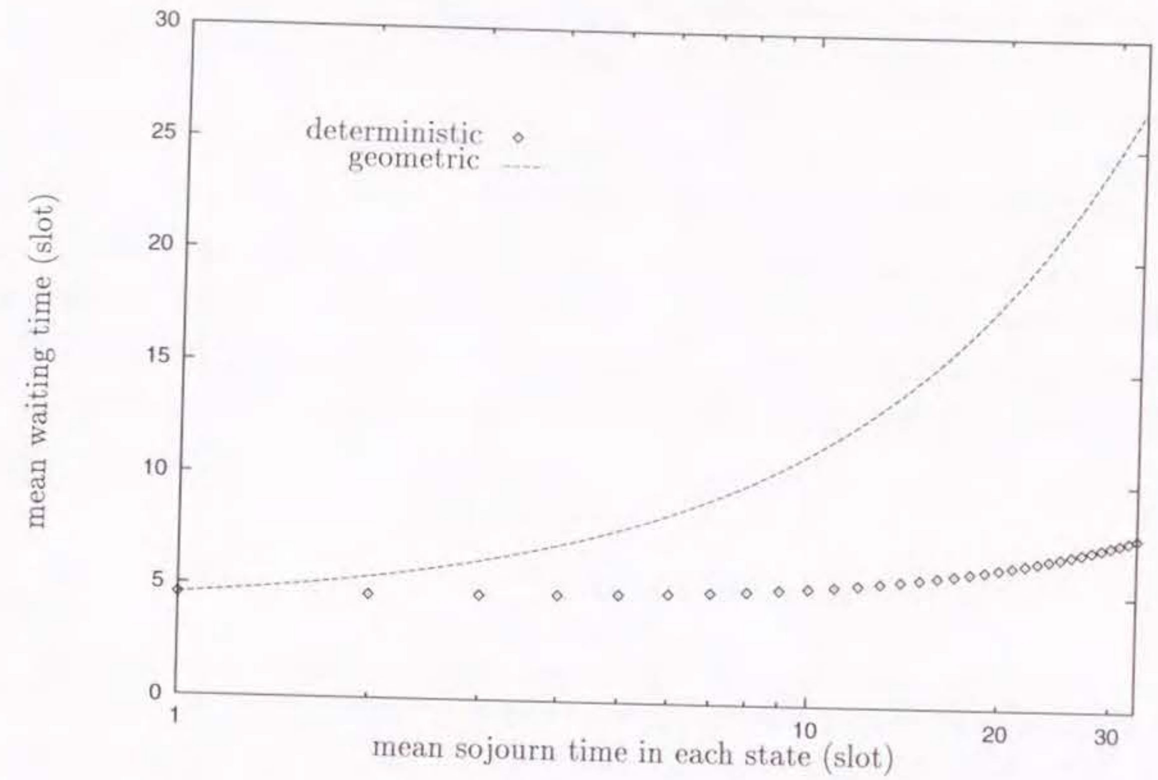


Figure 2.1: Mean sojourn time in each state vs. mean waiting time

[Mura91] have studied the discrete-time $\mathrm{GI}^{[X]}/\mathrm{G}/1$ queue under the most general assumption. We show that the queueing model considered in [Mura91] is a special case of our model and the analytical results given in [Mura91] are readily obtained from the results in section 2.3.

### 2.4.1 Analysis of Discrete-Time $\mathrm{GI}^{[X]}/\mathrm{G}/1$ Queues

Now we apply our model to a discrete-time $\mathrm{GI}^{[X]}/\mathrm{G}/1$ queue considered in [Mura91], which is characterized by the following PGFs. Let $G(z)$ denote the PGF of the interarrival time distribution of batches of positive size. We assume here that the PGF $G(z)$ is represented as a rational function. Let $B_g(z)$, $C_g(z)$ and $A_g(z)$ denote the PGFs for the batch size, the service time of a customer and the amount of work brought into the system by a batch, respectively. In order to analyze the above queueing model, we set the PGFs in our model in section 2.2 as follows:

$$\begin{aligned} &G_i(z) = G(z), \qquad B_{ii}(z) = 1, \qquad B_{12}(z) = B_{21}(z) = B_g(z), \\ &C_{ii}(z) = 1, \quad C_{12}(z) = C_{21}(z) = C_g(z) \quad (i = 1, 2), \end{aligned} \tag{2.39}$$

and therefore $A_{ii}(z) = 1$ and $A_{12}(z) = A_{21}(z) = A_g(z)$. In other words, we consider the model where batches of positive size arrive only when the state transitions from state 1 to state 2 or from state 2 to state 1 occur and sojourn times in each state correspond to interarrival times of batches of positive size. By definition, the traffic intensity is given by

$$\rho_g = A'_g(1)/G'(1). \tag{2.40}$$

In the following analysis, we assume that $\rho_g < 1$ and the system is in equilibrium.

We first consider the PGF $L_g(0, z)$ for the amount of work in the system immediately after arrivals of batches. Note that $L_g(0, z)$ corresponds to $L_i(0, z)$ in our original model. It then follows from (2.19) or (2.20) that

$$L_g(0, z) = \frac{(z-1)A_g(z)X_g(1/z)}{1 - A_g(z)G(1/z)}, \tag{2.41}$$

where

$$X_g(z) = \sum_{m=1}^{\infty}\sum_{k=0}^{\infty} L_g(k, 0)g(k+m)z^m, \tag{2.42}$$

and $g(m)$ is the probability that an interarrival time is equal to $m$ slots. Furthermore, from the normalizing equation $L_g(0, 1) = 1$, we have

$$X_g(1) = (1 - \rho_g)G'(1). \tag{2.43}$$

Recall that $X_g(z)$ is readily determined as shown in section 2.3.

*Remark 2.4.1:* Note that if interarrival times of batches have a geometric distribution, we have the following equation:

$$X_g(z) = X_g(1)G(z). \tag{2.44}$$

Thus the unknown function $X_g(z)$ is explicitly determined when the interarrival times are geometrically distributed.

Next, we consider the PGF $U_g(z)$ for the amount of the stationary work in the system. Note that $U_g(z)$ corresponds to $U_i(z)$ in the original model in section 2.3. It then follows from (2.25) that

$$U_g(z) = \frac{1}{G'(1)(1-1/z)}\left[L_g(0, z)\left\{1 - G\left(\frac{1}{z}\right)\right\} + (z-1)\left\{\frac{1}{z}X_g(1) - X_g\left(\frac{1}{z}\right)\right\}\right]. \tag{2.45}$$

Finally, we consider the PGF $W_g(z)$ for the waiting time of a randomly chosen customer. From the results in section 2.3, we obtain

$$W_g(z) = F_g(z)D_g(z), \tag{2.46}$$



where

$$F_g(z) = \frac{L_g(0, z)}{A_g(z)} = \frac{(z-1)X_g(1/z)}{1 - A_g(z)G(1/z)}, \qquad D_g(z) = \frac{1 - A_g(z)}{B'_g(1)(1 - C_g(z))}. \tag{2.47}$$

*Remark 2.4.2:* After some algebra with (2.45) and (2.46), we have the following relationship between work in the system and the waiting time in the $\mathrm{GI}^{[X]}/\mathrm{G}/1$ queue:

$$U_g(z) = 1 - \rho_g + \rho_g z W_g(z)\frac{1 - C_g(z)}{C'_g(1)(1-z)}. \tag{2.48}$$

Note that (2.48) can also be derived from the invariant relationship or the equality of the virtual delay and attained waiting time distribution (see, for example, [Miya83], [Miya92], [Saka90], [Seng89]).

## 2.5 Queues with Two Independent Inputs

In this section, we consider an application of our model to a single-server queue with two independent input streams: $\mathrm{GI}^{[X]}/\mathrm{G}$ and BBP/G input streams. The analytical results are directly applied to the performance evaluation of ATM multiplexers [Mura90]. The continuous-time single-server queue with independent GI/G and M/G input streams has previously been studied in [Hook72], [Ott84], [Ott87] (see also references therein). Those papers have shown that, roughly speaking, the amount of work in the system is decomposed into two independent components, one of which is the amount of the stationary work in the M/G/1 queue. We show that the amounts of not only the stationary work but also work immediately after arrivals of $\mathrm{GI}^{[X]}/\mathrm{G}$ customers is decomposed into two independent components, one of which is the amount of work in the BBP/G/1 queue. Though the service times of $\mathrm{GI}^{[X]}/\mathrm{G}$ customers are i.i.d. in the model considered in this chapter, the arrival process of the G/G stream need not be homogeneous in time and the service times of the customers need not be independent (see [Ott84]). In [Ishi95a], Ishizaki et al. have considered a single-server queue with two independent inputs where the interarrival times of $\mathrm{GI}^{[X]}/\mathrm{G}$ customers and the service time of the customers are dependent, and analytically shown that the amount of work in the system is decomposed into the two independent components.

### 2.5.1 Work in the System

We consider work in a single-server queue with two independent streams: $\mathrm{GI}^{[X]}/\mathrm{G}$ and BBP/G input streams. For $\mathrm{GI}^{[X]}/\mathrm{G}$ customers, let $G(z)$ denote the PGF for the interarrival time of batches of $\mathrm{GI}^{[X]}/\mathrm{G}$ customers. We assume here that the PGF $G(z)$ is represented as a rational function. Let $B_g(z)$ and $C_g(z)$ denote the PGFs for the batch size and the service time, respectively, of $\mathrm{GI}^{[\mathrm{X}]}/\mathrm{G}$ customers. Thus we have the PGF $A_g(z)$ for the amount of work brought into the system by a batch of $\mathrm{GI}^{[\mathrm{X}]}/\mathrm{G}$ customers:

$$A_g(z) = B_g(C_g(z)). \tag{2.49}$$

On the other hand, we denote by $B_b(z)$ and $C_b(z)$ the PGF of the batch size and the service time, respectively, of BBP/G customers. We then have the PGF $A_b(z)$ for the amount of work brought into the system by BBP/G customers in a slot:

$$A_b(z) = B_b(C_b(z)). \tag{2.50}$$

In order to analyze the queueing system, we set the PGFs in the original model in section 2.2 as follows:

$$\begin{aligned} &A_{ii}(z)=A_b(z), B_{ii}(z)=B_b(z), C_{ii}(z)=C_b(z), \\ &\qquad A_{12}(z)=A_{21}(z)=A_b(z)A_g(z) \quad (i=1,2). \end{aligned} \tag{2.51}$$

It is easy to see that the resulting queueing model corresponds to the single-server queue with two independent input streams described above. In the following analysis, we assume that $A_b'(1) + A_g'(1)/G'(1) < 1$ and the system is in equilibrium.

We first consider the amount of work in the system immediately after arrivals of $\mathrm{GI}^{[\mathrm{X}]}/\mathrm{G}$ customers. Note that this corresponds to the amount of work at the beginning of each state in the original model. Let $L(0, z)$ denote the PGF for the amount of work immediately after arrivals of $\mathrm{GI}^{[\mathrm{X}]}/\mathrm{G}$ customers. It then follows from (2.19) or (2.20) that

$$L(0,z) = \frac{(z-1)A_g(z)X(A_b(z)/z)}{1 - A_g(z)G(A_b(z)/z)}, \tag{2.52}$$

where

$$X(z) = \sum_{m=1}^{\infty}\sum_{k=0}^{\infty} L(k,0)g(k+m)z^m, \tag{2.53}$$

and $g(m)$ is the probability that an interarrival time of batches of $\mathrm{GI}^{[\mathrm{X}]}/\mathrm{G}$ customers is equal to $m$ slots. Note that (2.52) is rewritten to be

$$L(0,z) = U_b(z)L^*(0,z), \tag{2.54}$$

where $U_b(z)$ denotes the PGF for the amount of the stationary work in the BBP/G/1 queue:

$$U_b(z) = (1 - A_b'(1))\frac{(z-1)A_b(z)}{z - A_b(z)}, \tag{2.55}$$



which is obtained from (2.41), (2.43), (2.44) and (2.45), and $L^*(0,z)$ is given by

$$L^*(0,z) = \frac{1}{1-A_b'(1)}\frac{z-A_b(z)}{A_b(z)}\frac{A_g(z)X(A_b(z)/z)}{1-A_g(z)G(A_b(z)/z)}. \tag{2.56}$$

Next, we consider the PGF $U(z)$ for the amount of the stationary work in the system. Note that $U(z)$ corresponds to $U_i(z)$ in our original model. It then follows from (2.25) that

$$\begin{aligned} U(z) &= \frac{1}{G'(1)\{1-A_b(z)/z\}}\left[L(0,z)\left\{1 - G\left(\frac{A_b(z)}{z}\right)\right\}\right. \\ &\qquad \left. +(z-1)\left\{\frac{A_b(z)}{z}X(1) - X\left(\frac{A_b(z)}{z}\right)\right\}\right] \\ &= U_b(z)U^*(z), \end{aligned} \tag{2.57}$$

where $U^*(z)$ is given by

$$U^*(z) = \frac{1}{G'(1)(1-A_b'(1))}\frac{z}{A_b(z)}\left[\frac{(A_g(z)-1)X(A_b(z)/z)}{1-A_g(z)G(A_b(z)/z)} + \frac{A_b(z)}{z}X(1)\right]. \tag{2.58}$$

From (2.54) and (2.57), we see that the amount of work immediately after arrivals of $\mathrm{GI}^{[\mathrm{X}]}/\mathrm{G}$ customers and the amount of the stationary work are decomposed into two independent components, one of which is the amount of the stationary work in the BBP/G/1 queue. In the next subsection, we relate the other components, represented by $L^*(0,z)$ and $U^*(z)$, to the amount of work immediately after arrivals and the amount of the stationary work, respectively, in a special $\mathrm{GI}^{[\mathrm{X}]}/\mathrm{G}/1$ queue.

### 2.5.2 Special $\mathrm{GI}^{[\mathrm{X}]}/\mathrm{G}/1$ queue

We consider a special $\mathrm{GI}^{[\mathrm{X}]}/\mathrm{G}/1$ queue with the same interarrival time and the same batch size distributions as in the $\mathrm{GI}^{[\mathrm{X}]}/\mathrm{GI}$ input stream. We now assume that the PGF $\hat{C}_g(z)$ for the service time of a special $\mathrm{GI}^{[\mathrm{X}]}/\mathrm{G}$ customer is given by

$$\hat{C}_g(z) = C_g(z\Theta(z)), \tag{2.59}$$

where $\Theta(z)$ denotes the PGF of the delayed busy period distribution in the BBP/G/1 queue and $\Theta(z)$ satisfies

$$\Theta(z) = A_b(z\Theta(z)). \tag{2.60}$$

Note that (2.59) implies that the service times in the special $\mathrm{GI}^{[\mathrm{X}]}/\mathrm{G}/1$ queue are given by the delay cycle of BBP/G customers with the initial delay corresponding to the service time of a $\mathrm{GI}^{[\mathrm{X}]}/\mathrm{G}$ customer. Setting $\omega = z\Theta(z)$, we have

$$z = \frac{z\Theta(z)}{\Theta(z)} = \frac{z\Theta(z)}{A_b(z\Theta(z))} = \frac{\omega}{A_b(\omega)}. \tag{2.61}$$

Thus we obtain

$$\hat{C}_g\left(\frac{z}{A_b(z)}\right) = C_g(z). \tag{2.62}$$

We denote by $\hat{A}_g(z)$ the PGF for the amount of work brought into the system by a batch, namely,

$$\hat{A}_g(z) = B_g(\hat{C}_g(z)). \tag{2.63}$$

It then follows from (2.62) that

$$\hat{A}_g\left(\frac{z}{A_b(z)}\right) = A_g(z). \tag{2.64}$$

For the special $GI^{[X]}/G/1$ queue, we first consider the PGF $\hat{L}(0,z)$ for the amount of work in the system immediately after arrivals of batches. From (2.19) or (2.20), we obtain $\hat{L}(0,z)$:

$$\hat{L}(0,z) = \frac{(z-1)\hat{A}_g(z)\hat{X}(1/z)}{1-\hat{A}_g(z)G(1/z)}, \tag{2.65}$$

where

$$\hat{X}(z) = \sum_{m=1}^{\infty}\sum_{k=0}^{\infty} \hat{L}(k,0)g(k+m)z^m, \tag{2.66}$$

and $g(m)$ is the probability that an interarrival time of batches is equal to $m$ slots.

Next, we consider the PGF $\hat{U}(z)$ for the amount of the stationary work in the system. From (2.25), we obtain $\hat{U}(z)$:

$$\hat{U}(z) = \frac{1}{G'(1)(1-1/z)}\left[\hat{L}(0,z)\left\{1-G\left(\frac{1}{z}\right)\right\} + (z-1)\left\{\frac{1}{z}\hat{X}(1) - \hat{X}\left(\frac{1}{z}\right)\right\}\right]. \tag{2.67}$$

Substituting $z = z/A_b(z)$ in (2.65) and (2.67), we have

$$\hat{L}\left(0,\frac{z}{A_b(z)}\right) = \frac{(z/A_b(z)-1)A_g(z)\hat{X}(A_b(z)/z)}{1-A_g(z)G(A_b(z)/z)}, \tag{2.68}$$

$$\begin{aligned}\hat{U}\left(\frac{z}{A_b(z)}\right) &= \frac{1}{G'(1)}\frac{z}{A_b(z)}\left[\frac{A_g(z)\hat{X}(A_b(z)/z)}{1-A_g(z)G(A_b(z)/z)}\left\{1-G\left(\frac{A_b(z)}{z}\right)\right\}\right.\\ &\left.\quad+\frac{A_b(z)}{z}\hat{X}(1) - \hat{X}\left(\frac{A_b(z)}{z}\right)\right].\end{aligned} \tag{2.69}$$

We now define $\tilde{X}(z) \triangleq (1-A_b'(1))\hat{X}(z)$ and rewrite $\hat{L}(0,z/A_b(z))$ and $\hat{U}(z/A_b(z))$ as follows:

$$\hat{L}\left(0,\frac{z}{A_b(z)}\right) = \frac{1}{1-A_b'(1)}\frac{z-A_b(z)}{A_b(z)}\frac{A_g(z)\tilde{X}(A_b(z)/z)}{1-A_g(z)G(A_b(z)/z)}, \tag{2.70}$$

$$\hat{U}\left(\frac{z}{A_b(z)}\right) = \frac{1}{G'(1)(1-A_b'(1))}\frac{z}{A_b(z)}\left[\frac{(A_g(z)-1)\tilde{X}(A_b(z)/z)}{1-A_g(z)G(A_b(z)/z)} + \frac{A_b(z)}{z}\tilde{X}(1)\right]. \tag{2.71}$$

Note that, from (2.56), (2.58), (2.70) and (2.71), if $X(z)$ is identical to $\tilde{X}(z)$, then $L^*(0,z)$ and $U^*(z)$ are identical to $\hat{L}(0,z/A_b(z))$ and $\hat{U}(z/A_b(z))$, respectively. Hence we shall show that $X(z)$ is identical to $\tilde{X}(z)$ in the next subsection.



### 2.5.3 Relationship between $X(z)$ and $\tilde{X}(z)$

Recall that $X(z)$ is expressed by a linear combination of unknown values and a coefficient of each unknown value is determined only by the function $G(z)$. Since $\tilde{X}(z)$ in (2.66) has the same form as $X(z)$ in (2.53), $\tilde{X}(z)$ is also expressed by a linear combination of unknown values and a coefficient of each unknown is exactly the same as in $X(z)$. Furthermore the unknown values in $X(z)$ and $\tilde{X}(z)$ can be obtained by examining the zeros in the denominator of $U^*(z)$ and $\hat{U}(z)$ respectively.

Since both $U^*(z)$ and $\hat{U}(z)$ are PGFs, the terms

$$\frac{(A_g(z)-1)X(A_b(z)/z)}{1-A_g(z)G(A_b(z)/z)} \quad \text{and} \quad \frac{(A_g(z)-1)\tilde{X}(A_b(z)/z)}{1-A_g(z)G(A_b(z)/z)} \tag{2.72}$$

in $U^*(z)$ and $\hat{U}(z)$, respectively, have no poles inside the unit disk. Note here that the denominators in both terms are identical. Thus, the numerators in these terms becomes zero at some value of $z$, at which the denominators in both terms becomes zero. Hence it is clear that these conditions provide us with the same linear equations for the unknown constants in both terms. Furthermore, the normalizing equations $U(1) = 1$ and $\hat{U}(1) = 1$ provide us with the following equation:

$$X(1) = \tilde{X}(1) = G'(1)(1-A_b'(1)) - A_g'(1). \tag{2.73}$$

From these observations, we conclude that $X(z)$ is identical to $\tilde{X}(z)$, so that $L^*(0,z)$ and $U^*(z)$ are identical to $\hat{L}(0,z/A_b(z))$ and $\hat{U}(z/A_b(z))$, respectively. Thus the other factors are related to the amount of work in the special $GI^{[X]}/G/1$ queue. Note that, with (2.54), (2.57), (2.60) and (2.61), the decomposition results (2.54) and (2.57) are also given in the following forms:

$$L(0,z\Theta(z)) = U_b(z\Theta(z))\hat{L}(0,z), \qquad U(z\Theta(z)) = U_b(z\Theta(z))\hat{U}(z). \tag{2.74}$$

## 2.6 Queues with Service Interruptions

In this section, we consider an application of our model to a discrete-time queueing system with service interruptions. Queues with service interruptions have been extensively studied in the literature (see [Brun84], [Seng90] and references therein). In the context of the discrete-time queue, Bruneel [Brun84] has studied queues with service interruptions where the arrivals of customers are time-homogeneous and the service times of a customer are deterministic (equal to one slot). The lengths of on-periods are assumed to have a general distribution whose PGF is represented as a rational function. He has derived the PGF for the amount of the stationary work in the system. The model considered in section 2.2 enables us to have a more general model than the one in [Brun84]. Towsley [Tows80] has studied a breakdown model as a special case of queue with two-state Markov modulated input where each state is characterized by its own Bernoulli arrival process and independent error process, and derived the PGF for the queue length distribution. In his model, on-period lengths have a geometric distribution, while off-period lengths have a general distribution. Rubin and Zhang [Rubi92] have analyzed queues with deterministic on- and off-periods to obtain the performance measures in the TDMA scheme. Note that their model is a special case of our model. In our model, we assume that the batch size and the service time distributions of customers which arrive to the system during off-periods may differ from those when the server is working. These phenomena are naturally arising in real situations and a very similar model in continuous-time has been studied by Sengupta [Seng90]. We show that the PGF for the amount of work in the system is readily obtained from the results in section 2.3, and the analytical results include those in [Brun84] as a special case. Furthermore, we characterize the waiting time distribution of customers under the FCFS discipline.

### 2.6.1 Model

We consider a single-server queue in a random environment governed by a discrete-time alternating renewal process. We call the states of the alternating renewal process states 1 and 2. In state 1, the server works and therefore the server serves exactly one unit of work in a slot of state 1. State 2 denotes the breakdown of the server. We will also refer to states 1 and 2 as on- and off-periods, respectively. Let $g_i(n)$ $(i = 1, 2,\ n = 0, 1, 2, \ldots)$ denote the probability that time spent in state $i$ is equal to $n$. We denote by $G_i(z)$ $(i = 1, 2)$ the PGF of the $g_i(n)$. We assume here that the on-period PGF $G_1(z)$ is represented as a rational function. However, we do not make any assumptions regarding the form of the off-period PGF $G_2(z)$. Customers arrive to the system in batches. The batch size and the service time distributions may differ among the two states. Let $B_i(z)$ and $C_i(z)$ denote the PGFs for the batch size and the service time of a customer arriving in a slot of state $i$ $(i = 1, 2)$. Also, let $A_i(z)$ denote the PGF for the amount of work brought into the system in a slot of state $i$. In the following analysis, we assume that $(G_1'(1)A_1'(1)+G_2'(1)A_2'(1))/G_1'(1) < 1$ and the system is in equilibrium.

### 2.6.2 Work in the System

We first consider the imbedded workload process only during on-periods. We excise all slots during off-periods, gather all arrivals during each off-period, and put them in the last slot of the on-period preceding each off-period. As a result, the imbedded process



behaves like the workload in the queueing system with two independent input streams: one is a BBP/G input stream with the batch size PGF $B_1(z)$ and the service time PGF $C_1(z)$, and the other is a $\mathrm{GI}^{[X]}/\mathrm{G}$ input stream with the interarrival time PGF $G_1(z)$, the batch size PGF $G_2(B_2(z))$ and the service time PGF $C_2(z)$. Note that we have already analyzed the workload process in such a queue in section 2.5.

Let $\tilde{U}_1(z)$ be the stationary PGF for the amount of work in the imbedded process. From the results in the previous section, we obtain

$$\begin{aligned}\tilde{U}_1(z) &= \frac{1}{G_1'(1)\{1 - A_1(z)/z\}}\left[L_1(0,z)\left\{1 - G_1\left(\frac{A_1(z)}{z}\right)\right\}\right. \\ &\quad \left. +(z-1)\left\{\frac{A_1(z)}{z}X(1) - X\left(\frac{A_1(z)}{z}\right)\right\}\right], \end{aligned} \tag{2.75}$$

where

$$L_1(0,z) = \frac{(z-1)G_2(A_2(z))X(A_1(z)/z)}{1 - G_2(A_2(z))G_1(A_1(z)/z)}, \qquad X(z) = \sum_{m=1}^{\infty}\sum_{k=0}^{\infty} L_1(k,0)g_1(k+m)z^m. \tag{2.76}$$

Note that $L_1(0, z)$ in (2.76) is identical to $C_0(z)$ in [Brun84] if $A_1(z) = A_2(z)$.

The next step is to obtain the PGF for the amount of work in the system in a randomly chosen slot during off-periods. To do so, we excise all slots during on-periods. We then notice that the resulting workload process is a pure birth process except at some renewal epochs, where the distribution of time intervals between successive renewals has the PGF $G_2(z)$. Let $L_2(0, z)$ denote the PGF of the distribution of the amount of work immediately after renewal epochs. It is easy to see that $L_2(0, z)$ is equivalent to the PGF of the distribution of the amount of work immediately after arrivals of $\mathrm{GI}^{[X]}/\mathrm{G}$ customers. Thus, from (2.52), we obtain

$$L_2(0,z) = \frac{(z-1)X(A_1(z)/z)}{1 - G_2(A_2(z))G_1(A_1(z)/z)}, \tag{2.77}$$

where $X(z)$ is given by (2.76). Note that $L_2(0, z)$ in (2.77) is identical to $D_0(z)$ in [Brun84] if $A_1(z) = A_2(z)$. Let $L_2(k, z)$ denote the PGF for the amount of work in the system in the $(k+1)$st slot of state 2. We then have

$$L_2(k,z) = \{A_2(z)\}^k L_2(0,z). \tag{2.78}$$

We denote by $\tilde{U}_2(z)$ the PGF of the distribution of the amount of the stationary work during off periods. By definition, we obtain

$$\tilde{U}_2(z) = \sum_{k=0}^{\infty} p_2(k)L_2(k,z) = \frac{1 - G_2(A_2(z))}{G_2'(1)(1 - A_2(z))}L_2(0,z), \tag{2.79}$$

where $p_2(k)$ $(k = 0, 1, 2, \ldots)$ is the conditional probability that a randomly chosen slot of state 2 is the $(k+1)$st slot given the slot is in state 2 and given in (2.29).

Let $\tilde{U}(z)$ denote the PGF for the amount of work in the single-server queue with service interruptions. We then have

$$\begin{aligned}
\tilde{U}(z) &= \pi_1 \tilde{U}_1(z) + \pi_2 \tilde{U}_2(z) \\
&= \pi_1 \frac{1}{G_1'(1)\{1 - A_1(z)/z\}} \left[ L_1(0,z) \left\{ 1 - G_1\left(\frac{A_1(z)}{z}\right) \right\} \right. \\
&\quad \left. +(z-1) \left\{ \frac{A_1(z)}{z} X(1) - X\left(\frac{A_1(z)}{z}\right) \right\} \right] + \pi_2 \frac{1 - G_2(A_2(z))}{G_2'(1)(1 - A_2(z))} L_2(0,z), \quad (2.80)
\end{aligned}$$

where the $\pi_i$'s $(i = 1, 2)$ are given in (2.27).

### 2.6.3 Waiting Time Distribution

In this subsection, we consider the waiting time distribution of customers under the FCFS discipline. Before we analyze the waiting time distribution, we characterize the amount of work in the system in a slightly different way from that in the previous subsection. We denote by $\psi_i(k,x)$ $(i = 1, 2,\ k = 0, 1, 2, \ldots,\ x = 0, 1, 2, \ldots)$ the conditional joint probability that a randomly chosen customer arrives in the $(k+1)$st slot of state $i$ and the amount of work seen by him upon arrival is equal to $x$ (including that brought by those customers served before him within the same batch) given that the tagged customer arrives in state $i$. Let $\Psi_i(k,z)$ denote the $z$-transform of the $\psi_i(k,x)$ with respect to $x$. It is clear that

$$\begin{aligned}
\Psi_1(k,z) &= p_1(k) \left[ \frac{L_1(k,z) - L_1(k,0)}{z} + L_1(k,0) \right] D_{11}(z), \\
\Psi_2(k,z) &= p_2(k) L_2(k,z) D_{22}(z), \qquad (2.81)
\end{aligned}$$

where $D_{ii}(z)$ $(i = 1, 2)$ is given in (2.36). By definition, we have

$$\sum_{k=0}^{\infty} \Psi_1(k,z) = \left[ \frac{\tilde{U}_1(z) - \tilde{U}_1(0)}{z} + \tilde{U}_1(0) \right] D_{11}(z), \qquad \sum_{k=0}^{\infty} \Psi_2(k,z) = \tilde{U}_2(z) D_{22}(z). \quad (2.82)$$

In analyzing the waiting time distribution of customers, we separately treat two types of customers, those who arrive during on-periods and those during off-periods. Let $w_i(y|k,x)$ $(i = 1, 2,\ k = 0, 1, 2, \ldots)$ denote the probability that the waiting time of a customer is equal to $y$ slots given that the customer arrives in the $(k+1)$st slot of state $i$ and the amount of work found by him upon arrival is $x$ $(x = 0, 1, 2, \ldots)$. Clearly, the waiting time of a customer who arrives during an on-period and finds $x$ units of work in front of him upon arrival is given by the sum of $x$ slots and the total length of off-periods that occur before the start of his service following the depletion of $x$ units of work. Let $N(k,x)$ denote the number of those off-periods before the start of his service given that the tagged customer arrives in the $(k+1)$st slot of state 1 and finds $x$ units of work in front of him upon arrival. Note that $N(k,x)$ is considered as a delayed renewal process whose first renewal time is equal to $n$ $(n = 0, 1, 2, \ldots)$ with probability $g_1(n+k+1)/(1 - \sum_{m=1}^{k} g_1(m))$ and the subsequent interrenewal time is equal to $n$ with probability $g_1(n)$. We then have

$$\Pr\{N(k,x) = l\} = \begin{cases}
\displaystyle\sum_{j=0}^{x} \frac{g_1(j+k+1)}{1 - \sum_{m=1}^{k} g_1(m)} \sum_{n=l-1}^{x-j} g_1^{*(l-1)}(n) \left\{ 1 - \sum_{m=1}^{x-j-n} g_1(m) \right\} & (l \geq 2), \\
\displaystyle\sum_{j=0}^{x} \frac{g_1(j+k+1)}{1 - \sum_{m=1}^{k} g_1(m)} \left\{ 1 - \sum_{n=1}^{x-j} g_1(n) \right\} & (l = 1), \\
\displaystyle 1 - \sum_{j=0}^{x} \frac{g_1(j+k+1)}{1 - \sum_{m=1}^{k} g_1(m)} & (l = 0),
\end{cases} \quad (2.83)$$

where $g_i^{*(k)}(x)$ $(i = 1, 2)$ is the $k$-fold convolution of $g_i(x)$ with itself. In the above and the following equations, the summation taken in decreasing order is defined to be zero. Using $\Pr\{N(k,x) = l\}$, we have the following expression for $w_i(y|k,x)$:

$$w_1(y|k,x) = \begin{cases}
0 & (y < x), \\
\displaystyle\sum_{l=0}^{x+1} g_2^{*(l)}(y-x) \Pr\{N(k,x) = l\} & (x \leq y),
\end{cases} \quad (2.84)$$

$$w_2(y|k,x) = \begin{cases}
0 & (y < x), \\
\displaystyle\sum_{n=0}^{y-x} w_1(y-n|0,x) \frac{g_2(n+k+1)}{1 - \sum_{m=1}^{k} g_2(m)} & (x \leq y).
\end{cases} \quad (2.85)$$

By combining $\psi_i(k,x)$ introduced in the beginning of this subsection with $w_i(y|k,x)$, we obtain the waiting time distribution of a randomly chosen customer. Let $W_c$ denote a random variable for the waiting time of a randomly chosen customer. We then have

$$\Pr\{W_c = y\} = \sum_{i=1}^{2} \eta_i \sum_{k=0}^{\infty} \sum_{x=0}^{y} \psi_i(k,x) w_i(y|k,x), \quad (2.86)$$

where $\eta_i$ is given by

$$\eta_i = \frac{B_i'(1) G_i'(1)}{B_1'(1) G_1'(1) + B_2'(1) G_2'(1)} \qquad (i = 1, 2). \quad (2.87)$$

Finally, we consider the mean waiting time $E[W_c]$. Let $w_i^{(1)}(k,x)$ $(i = 1, 2)$ denote the expectation of $w_i(y|k,x)$. It then follows from (2.84) that

$$w_1^{(1)}(k,x) = x + G_2'(1) m(k,x), \quad (2.88)$$

where $m(k,x) = E[N(k,x)]$. Note that $m(k,x)$ is the renewal function of the delayed renewal process, which starts at age $k$. Furthermore, from (2.85), we have

$$w_2^{(1)}(k,x) = w_1^{(1)}(0,x) + \frac{\sum_{n=0}^{\infty} n g_2(n+k+1)}{1 - \sum_{m=1}^{k} g_2(m)}. \tag{2.89}$$

Taking the expectation of both sides of (2.86), we have

$$\begin{aligned} E[W_c] &= \sum_{i=1}^{2} \eta_i \sum_{k=0}^{\infty}\sum_{x=0}^{\infty} \psi_i(k,x) w_i^{(1)}(k,x) \\ &= S_0 + \eta_1 G_2'(1) S_1 + \eta_2 G_2'(1) S_2, \end{aligned} \tag{2.90}$$

where

$$S_0 = \eta_1(\tilde{U}_1'(1) + D_{11}'(1) - 1 + \tilde{U}_1(0)) + \eta_2(\tilde{U}_2'(1) + D_{22}'(1) + \frac{G_2''(1)}{2G_2'(1)}), \tag{2.91}$$

$$S_1 = \sum_{k=0}^{\infty}\sum_{x=0}^{\infty} \psi_1(k,x) m(k,x), \qquad S_2 = \sum_{k=0}^{\infty}\sum_{x=0}^{\infty} \psi_2(k,x) m(0,x). \tag{2.92}$$

## 2.7 Conclusion

In this chapter, we analyze the generalized SBBP/G/1 queues, where the arrival and the service processes are semi-Markovian in the sense that their distributions depend not only on the state of the alternating renewal process in the current slot but also on the state in the next slot. We derive the PGFs for the amount of work in the system and the waiting time of a customer under the FCFS discipline. We also provide numerical examples to show the computational feasibility of the analytical results. Furthermore, we show an application of the analytical results to important queueing systems such as a discrete-time $\mathrm{GI}^{[X]}/\mathrm{G}/1$ queue and a discrete-time queue with service interruptions.



# Chapter 3

# DBMAP/D/1 Queues with Finite and Infinite Buffers

## 3.1 Introduction

In this chapter, we study the customer loss probability approximation in DBMAP/D/1/K queues. The approximate formulas are given in terms of the tail distribution of the queue length in the corresponding infinite-buffer queue. Though, as we mentioned in chapter 1, the formulas for the *exact* loss probability in DBMAP/D/1/K queue have been already derived (see, for example, [Blon92], [Taki95]), the reasons that we consider the *approximate* formula of the loss probability in this chapter are as follows. When the buffer size is large, the number of states and the size of the transition matrix representing the imbedded Markov chain to describe the dynamics of the system become prohibitively large. When the loss probability is very small, we have some difficulties in its computation. Those facts make the computation of the exact loss probability with enough accuracy very difficult, even if the exact analytical framework is available. Furthermore such an exact computation is time consuming. In application, we are mainly interested in those cases [Ishi95b]. Thus an efficiently computable yet accurate approximate formula of the loss probability should be developed.

The organization of this chapter is as follows. In section 3.2, we describe the mathematical model. In section 3.3, we propose simple approximations to the loss probability which are given in terms of the tail distribution of the queue length in the corresponding infinite-buffer queue. The approximate formulas are constructed in such a way that they become exact for any independent arrival process. The accuracy of the approximations is extensively examined through numerical experiments.

## 3.2 Model

We consider the queueing model with the following characteristics:

- Customers arrive in a batch. The arrivals are governed by an underlying $M$-state Markov chain. This Markov chain changes its state on slot boundaries. The number of customers arriving in a slot depends not only on the state of the underlying Markov chain in the current slot but also on the state in the next slot.
- The service times of customers are assumed to be constant and equal to one slot. The service of a customer starts at the beginning of a slot and ends at the end of the slot (i.e., on slot boundaries). Customers depart from the system at slot boundaries.
- The queueing system has finite buffer and accommodates at most $N$ customers including the one in service. Thus, when $m$ ($m \geq N - k + 1$) customers arrive to find $k$ customers (including the one in service) in the system, only $N - k$ customers are accommodated in the system, and the remaining $m - (N - k)$ customers are lost.

As for timings of arrivals, two queueing models have been explored in the past: the early arrival model and the late arrival model (see p.5 of [Taka93]). In the early arrival model, an arrival of a batch in the $n$th slot occurs immediately after the beginning of the $n$th slot. On the other hand, in the late arrival model, an arrival of a batch in the $n$th slot occurs immediately before the end of the $n$th slot. In what follows, we consider both queueing models in parallel.

Before proceeding to the analysis, we introduce some notations. The state transition matrix for the underlying Markov chain is denoted by $\boldsymbol{U} = \{U_{ij}\}$ $(i, j = 1, \ldots, M)$, where we assume $\boldsymbol{U}$ is irreducible. Let $P_n$ denote the state of the underlying Markov chain in the $n$th slot. Let $\boldsymbol{\pi} = (\pi_1, \ldots, \pi_M)$ denote the stationary state vector of this Markov chain. Note that $\boldsymbol{\pi}$ satisfies $\boldsymbol{\pi} = \boldsymbol{\pi}\boldsymbol{U}$ and $\boldsymbol{\pi}\boldsymbol{e} = 1$, where $\boldsymbol{e}$ is an $M \times 1$ vector with all elements equal to one. Let $A_n$ denote the number of customers arriving in the $n$th slot (i.e., in this model, the amount of work brought in the system by a batch arriving in the $n$th slot). We assume that $A_{n+1}$ depends on both $P_n$ and $P_{n+1}$ (see [Blon92], [Taki94] and [Taki95]). We denote by $a_{ij}(k)$ the probability of $k$ ($k \geq 0$) customers arriving in the current slot given that the underlying Markov chain was in state $i$ in the previous slot and is in state $j$ in the current slot:

$$a_{ij}(k) = \Pr\{A_{n+1} = k \mid P_n = i, P_{n+1} = j\} \qquad (i, j = 1, \ldots, M, \quad k = 0, 1, 2, \ldots). \quad (3.1)$$

Note that we assume $a_{ij}(k)$ is time homogeneous and is independent of $n$. Also, let $\hat{a}_{ij}(k)$ denote the conditional probability for the following events: $k$ customers arrive in the $(n+1)$st slot, and the underlying Markov chain is in state $j$ in the $(n+1)$st slot, given that the Markov chain was in state $i$ in the $n$th slot. Namely,

$$\hat{a}_{ij}(k) = \Pr\{A_{n+1} = k, P_{n+1} = j \mid P_n = i\} = a_{ij}(k)U_{ij} \qquad (i, j = 1, \ldots, M). \quad (3.2)$$

Let $\boldsymbol{A}_k$ and $\boldsymbol{B}_k$ denote $M \times M$ matrices whose $(i,j)$th elements are given by $\hat{a}_{ij}(k)$ and $\sum_{m=k}^{\infty} \hat{a}_{ij}(m)$, respectively, where $k \geq 0$. Note that $\boldsymbol{A}_k$ (resp. $\boldsymbol{B}_k$) represents the transition matrix of the underlying Markov chain when $k$ customers (resp. more than or equal to $k$ customers) arrive to the system. By definition, $\boldsymbol{B}_0 = \boldsymbol{U}$, where $\boldsymbol{U}$ denotes the transition probability matrix of the underlying Markov chain. Let $\rho$ denote the traffic intensity which is given by

$$\rho = \boldsymbol{\pi} \sum_{k=1}^{\infty} k \boldsymbol{A}_k \boldsymbol{e}. \quad (3.3)$$

## 3.3 Loss Probability Approximation

In this section, we propose simple approximate formulas for the loss probability in both the early and the late arrival models. First we consider the distribution of the number of customers in the system. Next we propose a heuristic approximation of the loss probability. Finally we examine the accuracy of the approximation through numerical experiments.

### 3.3.1 Distribution of the Number of Customers in the System

We observe the system immediately after all possible events (i.e., a departure of a customer and customer arrivals) happen around slot boundaries. For convenience of the analysis, we introduce slightly different definitions of the number of customers in the system for the early and the late arrival models. Let $Y_n^{(e)}$ denote a random variable which represents the number of customers in the system immediately after the beginning of $n$th slot in the early arrival model. On the other hand, let $Y_n^{(l)}$ denote a random variable which represents the number of customers in the system immediately after the end of the $n$th slot in the late arrival model. We note that both $Y_n^{(e)}$ and $Y_n^{(l)}$ include customers arriving and accommodated in the system in the $n$th slot. Then $\{Y_n^{(e)}, P_n; n = 0, 1, \ldots\}$ and $\{Y_n^{(l)}, P_n; n = 0, 1, \ldots\}$ constitute the the bivariate Markov chains, whose transition matrices are given by (see [Blon92] and [Taki95]):

The early arrival model:

$$\begin{bmatrix} A_0 & A_1 & A_2 & \cdots & A_{N-2} & A_{N-1} & B_N \\ A_0 & A_1 & A_2 & \cdots & A_{N-2} & A_{N-1} & B_N \\ 0 & A_0 & A_1 & \cdots & A_{N-3} & A_{N-2} & B_{N-1} \\ 0 & 0 & A_0 & \cdots & A_{N-4} & A_{N-3} & B_{N-2} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & A_0 & A_1 & B_2 \\ 0 & 0 & 0 & \cdots & 0 & A_0 & B_1 \end{bmatrix}. \tag{3.4}$$

The late arrival model:

$$\begin{bmatrix} A_0 & A_1 & A_2 & \cdots & A_{N-2} & A_{N-1} & B_N \\ A_0 & A_1 & A_2 & \cdots & A_{N-2} & B_{N-1} & 0 \\ 0 & A_0 & A_1 & \cdots & A_{N-3} & B_{N-2} & 0 \\ 0 & 0 & A_0 & \cdots & A_{N-4} & B_{N-3} & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & A_0 & B_1 & 0 \\ 0 & 0 & 0 & \cdots & 0 & B_0 & 0 \end{bmatrix}. \tag{3.5}$$

Let $\boldsymbol{y}_k^{(e)}$ (resp. $\boldsymbol{y}_k^{(l)}$) denote an $1 \times M$ vector whose $j$th element represents the joint stationary probability of $k$ customers in the system and the underlying Markov chain being in state $j$ in the early (resp. late) arrival model. Note that $\boldsymbol{y}_k^{(e)}$ and $\boldsymbol{y}_k^{(l)}$ satisfy the following equations:

The early arrival model:

$$\boldsymbol{y}_k^{(e)} = \boldsymbol{y}_0^{(e)} \boldsymbol{A}_k + \sum_{i=1}^{k+1} \boldsymbol{y}_i^{(e)} \boldsymbol{A}_{k+1-i} \qquad (0 \le k \le N-1), \tag{3.6}$$

$$\boldsymbol{y}_N^{(e)} = \boldsymbol{y}_0^{(e)} \boldsymbol{B}_N + \sum_{i=1}^{N} \boldsymbol{y}_i^{(e)} \boldsymbol{B}_{N+1-i}. \tag{3.7}$$

The late arrival model:

$$\boldsymbol{y}_k^{(l)} = \boldsymbol{y}_0^{(l)} \boldsymbol{A}_k + \sum_{i=1}^{k+1} \boldsymbol{y}_i^{(l)} \boldsymbol{A}_{k+1-i} \qquad (0 \le k \le N-2), \tag{3.8}$$

$$\boldsymbol{y}_{N-1}^{(l)} = \boldsymbol{y}_0^{(l)} \boldsymbol{A}_{N-1} + \sum_{i=1}^{N} \boldsymbol{y}_i^{(l)} \boldsymbol{B}_{N-i}, \tag{3.9}$$

$$\boldsymbol{y}_N^{(l)} = \boldsymbol{y}_0^{(l)} \boldsymbol{B}_N. \tag{3.10}$$

The above equations completely determine $\boldsymbol{y}_k^{(e)}$ and $\boldsymbol{y}_k^{(l)}$ $(0 \le k \le N)$ with the normalizing equations $\sum_{k=0}^{N} \boldsymbol{y}_k^{(e)} \boldsymbol{e} = 1$ and $\sum_{k=0}^{N} \boldsymbol{y}_k^{(l)} \boldsymbol{e} = 1$, where $\boldsymbol{e}$ denotes an $M \times 1$ vector whose all elements are equal to one. As for the algorithms to solve the above equations, readers are referred to [LeBo91] and [Taki95].



Let $P_{loss}^{(e)}$ (resp. $P_{loss}^{(l)}$) denote the loss probability in the early (resp. late) arrival model. It can be shown that the loss probabilities in the two models are given by (see [Taki95] as for the derivation):

$$P_{loss}^{(e)} = \frac{\rho - (1 - \boldsymbol{y}_0^{(e)} \boldsymbol{e})}{\rho}, \qquad P_{loss}^{(l)} = \frac{\rho - (1 - \boldsymbol{y}_0^{(l)} \boldsymbol{e})}{\rho}. \tag{3.11}$$

### 3.3.2 Heuristic Approximation of the Loss Probability

In this subsection, we propose approximate formulas of the loss probability which are given in terms of the tail distribution in the corresponding infinite-buffer queue. We first consider the distribution of the number of customers in the corresponding discrete-time queue with buffer of infinite capacity, which will be used to construct the approximate formulas. Let $\boldsymbol{x}_k$ denote an $1 \times M$ vector whose $j$th element represents the joint stationary probability of $k$ customers in the system and the underlying Markov chain being in state $j$ in the corresponding discrete-time queue with infinite buffer. Note that both the early and the late arrival models have the same distribution of the number of customers in the system when the buffer has infinite capacity. The $\boldsymbol{x}_k$ $(k \ge 0)$ satisfies [Taki94]

$$\boldsymbol{x}_k = \boldsymbol{x}_0 \boldsymbol{A}_k + \sum_{i=1}^{k+1} \boldsymbol{x}_i \boldsymbol{A}_{k+1-i} \qquad (k \ge 0), \tag{3.12}$$

with the normalizing equation $\sum_{k=0}^{\infty} \boldsymbol{x}_k \boldsymbol{e} = 1$. We define $\boldsymbol{X}(z)$ and $\boldsymbol{A}(z)$:

$$\boldsymbol{X}(z) = \sum_{k=0}^{\infty} \boldsymbol{x}_k z^k, \tag{3.13}$$

$$\boldsymbol{A}(z) = \sum_{k=0}^{\infty} \boldsymbol{A}_k z^k. \tag{3.14}$$

We then have [Taki94]

$$\boldsymbol{X}(z)[z\boldsymbol{I} - \boldsymbol{A}(z)] = (1-\rho)(z-1)\boldsymbol{g}\boldsymbol{A}(z), \tag{3.15}$$

where $\boldsymbol{g}$ denotes an $1 \times M$ vector. Note that $\boldsymbol{x}_0 = (1-\rho)\boldsymbol{g}$ can be obtained by solving a set of $M$ linear equations and $\boldsymbol{x}_k$ $(k \ge 1)$ can be recursively computed by the matrix analytic method (see [Neut89] and [Taki94]).

To construct the approximate formulas of the loss probability, we employ the following heuristic idea. We assume that the joint stationary probabilities $\boldsymbol{y}_k^{(e)}$ and $\boldsymbol{y}_k^{(l)}$ are related to $\boldsymbol{x}_k$ by truncating and renormalizing the $\boldsymbol{x}_k$ if the loss probability is very small. We claim that if the loss probability is very small, the traffic intensity $\rho$ should be less than 1, so that the distribution of the number of customers in the system in the corresponding infinite-buffer queue is well defined. Thus, we suggest a conditional approximation when the loss probability is very small:

$$\frac{\Pr(Y^{(e)}=k, P=j)}{\Pr(Y^{(e)} \le N)} \approx \frac{\Pr(X=k, P=j)}{\Pr(X \le N)} \quad (0 \le k \le N), \tag{3.16}$$

$$\frac{\Pr(Y^{(l)}=k, P=j)}{\Pr(Y^{(l)} \le N-1)} \approx \frac{\Pr(X=k, P=j)}{\Pr(X \le N-1)} \quad (0 \le k \le N-1), \tag{3.17}$$

where $Y^{(e)}$, $Y^{(l)}$ and $X$ denote generic random variables representing the number of customers in the system in the early arrival model, in the late arrival model and in the corresponding discrete-time infinite-buffer queue, respectively. Also $P$ denotes a generic random variable representing the state of the underlying Markov chain. The idea behind the above approximation is that the stationary probabilities in finite-buffer queues would not be affected so much by losses when the loss seldom happens (i.e., the loss probability is less than, say, $10^{-7}$). In other words, finite-buffer queues with a very small loss probability behave as if they would be the corresponding infinite-buffer queue given that the number of customers in the system is not greater than the buffer size.

The approximation implies that $\boldsymbol{y}_k^{(e)}$ and $\boldsymbol{y}_k^{(l)}$ are approximately expressed to be

$$\boldsymbol{y}_k^{(e)} \approx c^{(e)}\boldsymbol{x}_k \quad (0 \le k \le N), \tag{3.18}$$

$$\boldsymbol{y}_k^{(l)} \approx c^{(l)}\boldsymbol{x}_k \quad (0 \le k \le N-1), \qquad \boldsymbol{y}_N^{(l)} \approx c^{(l)}\boldsymbol{x}_0\boldsymbol{B}_N, \tag{3.19}$$

where $c^{(e)}$ and $c^{(l)}$ are given by

$$c^{(e)} = \left(\sum_{k=0}^{N} \boldsymbol{x}_k\boldsymbol{e}\right)^{-1}, \qquad c^{(l)} = \left(\sum_{k=0}^{N-1} \boldsymbol{x}_k\boldsymbol{e} + \boldsymbol{x}_0\boldsymbol{B}_N\boldsymbol{e}\right)^{-1}, \tag{3.20}$$

which come from the normalizing equations.

*Remark 3.3.1:* Similar conditional approximations have been studied by several researchers in the context of continuous-time queues. Readers are referred to [Gouw94], [Miya93], [Saka93], [Tijm92] and references therein.

THEOREM 3.3.1: *The approximations given in (3.18) and (3.19) become exact when the number of customers arriving to the system in a slot is i.i.d. (independent and identically distributed).*

PROOF: See Appendix B. ∎

From (3.18), (3.19) and (3.20), we have

$$\boldsymbol{y}_0^{(e)}\boldsymbol{e} \approx \frac{\boldsymbol{x}_0\boldsymbol{e}}{\sum_{k=0}^{N}\boldsymbol{x}_k\boldsymbol{e}}, \qquad \boldsymbol{y}_0^{(l)}\boldsymbol{e} \approx \frac{\boldsymbol{x}_0\boldsymbol{e}}{\sum_{k=0}^{N-1}\boldsymbol{x}_k\boldsymbol{e} + \boldsymbol{x}_0\boldsymbol{B}_N\boldsymbol{e}}. \tag{3.21}$$



We define the tail distribution $T_k$ $(k \ge 0)$ as

$$T_k = \sum_{m=k+1}^{\infty} \boldsymbol{x}_m\boldsymbol{e}. \tag{3.22}$$

Let $\widetilde{P}_{loss}^{(e)}$ and $\widetilde{P}_{loss}^{(l)}$ denote the approximate loss probabilities in the early and the late arrival models, respectively. Noting the equalities $\sum_{m=0}^{k}\boldsymbol{x}_m\boldsymbol{e} = 1 - T_k$ and $\boldsymbol{x}_0\boldsymbol{e} = 1-\rho$, and using the approximation (3.21) in (3.11), we have

$$\widetilde{P}_{loss}^{(e)} = \frac{(1-\rho)T_N}{\rho(1-T_N)}, \qquad \widetilde{P}_{loss}^{(l)} = \frac{(1-\rho)(T_{N-1} - \boldsymbol{x}_0\boldsymbol{B}_N\boldsymbol{e})}{\rho(1-T_{N-1}+\boldsymbol{x}_0\boldsymbol{B}_N\boldsymbol{e})}. \tag{3.23}$$

*Remark 3.3.2:* When the number of customers arriving to the system is i.i.d., the above approximate formulas become exact.

*Remark 3.3.3:* When the probability that the number of customers arriving in a slot is greater than or equal to the buffer size $N$ is zero, $\boldsymbol{B}_N = \boldsymbol{0}$. In such a case, (3.23) for the late arrival model is reduced to

$$\widetilde{P}_{loss}^{(l)} = \frac{(1-\rho)T_{N-1}}{\rho(1-T_{N-1})}, \tag{3.24}$$

which is given only in terms of the tail distribution $T_N$.

*Remark 3.3.4:* In any traffic condition, we have $\widetilde{P}_{loss}^{(e)} \le \widetilde{P}_{loss}^{(l)}$, which coincides with intuition. This inequality can be shown by noting the fact that $f(x) = x/(1-x)$ is an increasing function of $x$ $(0 \le x < 1)$ and

$$(T_{N-1} - \boldsymbol{x}_0\boldsymbol{B}_N\boldsymbol{e}) - T_N = \boldsymbol{x}_N\boldsymbol{e} - \boldsymbol{x}_0\boldsymbol{B}_N\boldsymbol{e} = \sum_{i=1}^{N}\boldsymbol{x}_i\boldsymbol{B}_{N-i+1}\boldsymbol{e} \ge 0, \tag{3.25}$$

where the second equality can be verified by summing up both sides of (3.12) from $k=0$ to $N-1$.

### 3.3.3 Accuracy of the Approximations

In this subsection, we provide the results of our numerical experiments to show the accuracy of the proposed approximations. In particular, we focus on the impacts of the correlation in the arrival process on the accuracy of the approximations, since the formulas become exact for i.i.d. arrival processes. For this purpose, we use the following simple arrival process in all numerical experiments in this subsection.

We assume that the arrival process is modulated by a two-state Markov chain with states 1 and 2, where the state transition probabilities $U_{ij}$ are given by $U_{11} = U_{22} = \alpha$ and $U_{12} = U_{21} = 1 - \alpha$ $(0 < \alpha < 1)$. The conditional probabilities $\tilde{a}_{1,j}(k)$ and $\tilde{a}_{2,j}(k)$ $(j = 1, 2)$ for the sizes of the arrival batches are given by

$$\tilde{a}_{1,j}(k) = \left(1 - \frac{(1+c)\rho}{1+(1+c)\rho}\right)\left(\frac{(1+c)\rho}{1+(1+c)\rho}\right)^k, \tag{3.26}$$

$$\tilde{a}_{2,j}(k) = \left(1 - \frac{(1-c)\rho}{1+(1-c)\rho}\right)\left(\frac{(1-c)\rho}{1+(1-c)\rho}\right)^k. \tag{3.27}$$

In other words, if the Markov chain was in state 1 (resp. state 2) in the current slot, the number of customers arriving in the current slot is geometrically distributed with the mean $(1+c)\rho$ (resp. $(1-c)\rho$). Note that $\rho$ denotes the overall traffic intensity, and $c$ $(0 \le c \le 1)$ is a parameter.

Through numerical examples, we investigate the impact of the variation and the correlation in arrivals. For our arrival model, the squared coefficient of variation $C_V^2$ of the number of customers arriving in a slot is found to be

$$C_V^2 = 1 + \rho^{-1} + 2c^2. \tag{3.28}$$

For a fixed value of the traffic intensity $\rho$, the squared coefficient of variation $C_V^2$ increases as the parameter $c$ does. Also the correlation coefficient $C_C(n)$ of the number of arrivals at lag $n$ for our arrival process is found to be

$$C_C(n) = \frac{c^2\rho}{1+(1+2c^2)\rho} \cdot (2\alpha - 1)^n. \tag{3.29}$$

Note that, by keeping $\rho$ and $c$ constant (which means keeping $C_V^2$ constant), the correlation coefficient $C_C(n)$ depends only on the term $2\alpha - 1$. When $\alpha = 0.5$, the arrival process is i.i.d., and by varying $\alpha$ from 0.5 to 1, we achieve varying degrees of non-negative correlations of arrivals. In the rest of this subsection, the tail distributions in the approximate formulas are computed by the matrix-analytic method [Neut89], [Taki94].

Tables 3.1, 3.2, and 3.3 show the loss probability obtained by the approximate formulas, the loss probability obtained by the exact analysis, and the relative error of the approximations to the exact results for various values of the buffer size $N$, where the three parameters $\rho$, $c$ and $\alpha$ are fixed. It is quite interesting to observe that the accuracy of the approximations is less sensitive to the buffer size $N$, especially in the range of $10^{-10}$ to $10^{-5}$ of the loss probability. To confirm this observation, we provide Table 3.4, which shows the ranges of the relative error and the buffer size when the loss probability falls in the range of $10^{-10}$ to $10^{-5}$, where $\rho = 0.2, 0.5, 0.8$, $c = 0.2, 0.5, 0.8$ and $\alpha = 0.6, 0.8, 0.95$.



This table again leads to the above observation; the accuracy of the approximation is almost insensitive to the buffer size when the loss probability is very small.

We now investigate the impacts of the correlation in the arrival process on the accuracy of the approximations. We also fix the traffic intensity $\rho$ to one of the values 0.2 (Light), 0.5 (Medium) or 0.8 (Heavy). We fix the buffer size $N$ in such a way that the loss probability falls in the range of $10^{-10}$ to $10^{-5}$, depending on the value of $\rho$. Further, we fix the parameter $c$ to one of the values 0.2 (Low), 0.5 (Moderate) or 0.8 (High). Figures 3.1 through 3.6 show the loss probability obtained by both the approximate formulas (indicated by a) and the exact analysis (indicated by e) as a function of parameter $\alpha$. For example, HMe indicates the loss probability obtained by the exact analysis (e) in the case of heavy traffic (H: $\rho = 0.8$) and moderate variation (M: $c = 0.5$). We observe that when the correlation in arrivals is not so strong (i.e., $0.5 < c < 0.7$), the approximations are surprisingly accurate. We also observe that the error of the approximations becomes large according to the increase of correlation. Even in those cases, we can use the approximations to estimate the order of magnitude of the loss probability.

*Remark 3.3.5:* In all our numerical experiments, the loss probability obtained by the approximate formulas is conservative, i.e., the approximate results are larger than the exact ones.

*Remark 3.3.6:* Since the tail distribution has a simple asymptotic expression in many situations [Abat94], [Falk94], we can use the asymptotic expression of the tail distribution in the approximation formulas. Ishizaki et al. [Ishi95b] have considered the loss probability approximation using the asymptotic expression of the tail distribution when the arrival process comes from the superposition of many independent sources which are particularly important in practice. Furthermore, in general case, they have shown an intuitive and simple derivation of the exact asymptotic formula, which would help understanding why such a simple formula comes out.

## 3.4 Conclusion

In this chapter, we study the loss probability approximation in DBMAP/D/1/K queues. We propose the approximate formulas which are given in terms of the tail distribution of the queue length in the corresponding infinite-buffer queue. The approximate formulas are constructed in such a way that they become exact for any independent arrival process. The accuracy of the approximations is extensively examined through numerical experiments. We observe that when the correlation in arrivals is not so strong, the approximations are surprisingly accurate. We also observe that the error of the approximations becomes large according to the increase of correlation. Even in those cases, we can use the approximations to estimate the order of magnitude of the loss probability.



Table 3.1: Accuracy of Approximations ($\rho = 0.2$, $c = 0.8$, $\alpha = 0.95$)

| | Early arrival model | | | Late arrival model | | |
|---|---|---|---|---|---|---|
| N | exact | approximate | error(%) | exact | approximate | error(%) |
| 5 | 3.176E-03 | 3.842E-03 | 20.98 | 7.507E-03 | 9.237E-03 | 23.05 |
| 6 | 1.113E-03 | 1.347E-03 | 21.01 | 2.775E-03 | 3.385E-03 | 21.98 |
| 7 | 3.904E-04 | 4.725E-04 | 21.03 | 1.012E-03 | 1.226E-03 | 21.19 |
| 8 | 1.369E-04 | 1.657E-04 | 21.03 | 3.651E-04 | 4.404E-04 | 20.63 |
| 9 | 4.803E-05 | 5.813E-05 | 21.04 | 1.308E-04 | 1.572E-04 | 20.22 |
| 10 | 1.685E-05 | 2.039E-05 | 21.04 | 4.661E-05 | 5.589E-05 | 19.92 |
| 11 | 5.911E-06 | 7.154E-06 | 21.04 | 1.654E-05 | 1.980E-05 | 19.69 |
| 12 | 2.073E-06 | 2.510E-06 | 21.04 | 5.854E-06 | 6.997E-06 | 19.53 |
| 13 | 7.274E-07 | 8.804E-07 | 21.04 | 2.067E-06 | 2.468E-06 | 19.41 |
| 14 | 2.552E-07 | 3.088E-07 | 21.04 | 7.286E-07 | 8.694E-07 | 19.32 |
| 15 | 8.951E-08 | 1.083E-07 | 21.04 | 2.565E-07 | 3.059E-07 | 19.25 |
| 16 | 3.140E-08 | 3.801E-08 | 21.04 | 9.025E-08 | 1.076E-07 | 19.20 |
| 17 | 1.102E-08 | 1.333E-08 | 21.04 | 3.172E-08 | 3.780E-08 | 19.16 |
| 18 | 3.864E-09 | 4.677E-09 | 21.04 | 1.115E-08 | 1.328E-08 | 19.13 |
| 19 | 1.356E-09 | 1.641E-09 | 21.04 | 3.915E-09 | 4.663E-09 | 19.11 |
| 20 | 4.755E-10 | 5.756E-10 | 21.04 | 1.375E-09 | 1.637E-09 | 19.09 |
| 21 | 1.668E-10 | 2.019E-10 | 21.04 | 4.825E-10 | 5.746E-10 | 19.08 |
| 22 | 5.853E-11 | 7.084E-11 | 21.03 | 1.694E-10 | 2.017E-10 | 19.06 |
| 23 | 2.054E-11 | 2.485E-11 | 21.01 | 5.946E-11 | 7.077E-11 | 19.02 |
| 24 | 7.209E-12 | 8.718E-12 | 20.94 | 2.088E-11 | 2.483E-11 | 18.93 |
| 25 | 2.534E-12 | 3.059E-12 | 20.73 | 7.346E-12 | 8.714E-12 | 18.62 |
| 26 | 8.941E-13 | 1.074E-12 | 20.14 | 2.595E-12 | 3.058E-12 | 17.85 |

Table 3.2: Accuracy of Approximations ($\rho = 0.5$, $c = 0.5$, $\alpha = 0.8$)

| | Early arrival model | | | Late arrival model | | |
|---|---|---|---|---|---|---|
| N | exact | approximate | error(%) | exact | approximate | error(%) |
| 10 | 3.222E-03 | 3.567E-03 | 10.73 | 5.308E-03 | 5.726E-03 | 7.887 |
| 12 | 1.234E-03 | 1.366E-03 | 10.72 | 2.039E-03 | 2.199E-03 | 7.854 |
| 14 | 4.732E-04 | 5.239E-04 | 10.71 | 7.830E-04 | 8.444E-04 | 7.841 |
| 16 | 1.815E-04 | 2.010E-04 | 10.71 | 3.006E-04 | 3.242E-04 | 7.836 |
| 18 | 6.966E-05 | 7.713E-05 | 10.71 | 1.154E-04 | 1.245E-04 | 7.834 |
| 20 | 2.674E-05 | 2.960E-05 | 10.71 | 4.430E-05 | 4.777E-05 | 7.833 |
| 22 | 1.026E-05 | 1.136E-05 | 10.71 | 1.700E-05 | 1.833E-05 | 7.832 |
| 24 | 3.938E-06 | 4.360E-06 | 10.71 | 6.526E-06 | 7.037E-06 | 7.832 |
| 26 | 1.511E-06 | 1.673E-06 | 10.71 | 2.505E-06 | 2.701E-06 | 7.832 |
| 28 | 5.800E-07 | 6.422E-07 | 10.71 | 9.613E-07 | 1.037E-06 | 7.832 |
| 30 | 2.226E-07 | 2.465E-07 | 10.71 | 3.689E-07 | 3.978E-07 | 7.832 |
| 32 | 8.543E-08 | 9.459E-08 | 10.71 | 1.416E-07 | 1.527E-07 | 7.832 |
| 34 | 3.279E-08 | 3.630E-08 | 10.71 | 5.434E-08 | 5.860E-08 | 7.832 |
| 36 | 1.258E-08 | 1.393E-08 | 10.71 | 2.086E-08 | 2.249E-08 | 7.832 |
| 38 | 4.830E-09 | 5.347E-09 | 10.71 | 8.004E-09 | 8.631E-09 | 7.832 |
| 40 | 1.854E-09 | 2.052E-09 | 10.71 | 3.072E-09 | 3.313E-09 | 7.832 |
| 42 | 7.114E-10 | 7.876E-10 | 10.71 | 1.179E-09 | 1.271E-09 | 7.832 |
| 44 | 2.730E-10 | 3.023E-10 | 10.71 | 4.525E-10 | 4.879E-10 | 7.831 |
| 46 | 1.048E-10 | 1.160E-10 | 10.71 | 1.737E-10 | 1.873E-10 | 7.829 |
| 48 | 4.022E-11 | 4.452E-11 | 10.71 | 6.665E-11 | 7.187E-11 | 7.823 |
| 50 | 1.544E-11 | 1.709E-11 | 10.71 | 2.559E-11 | 2.758E-11 | 7.804 |
| 52 | 5.925E-12 | 6.559E-12 | 10.69 | 9.824E-12 | 1.059E-11 | 7.760 |
| 54 | 2.276E-12 | 2.518E-12 | 10.64 | 3.776E-12 | 4.063E-12 | 7.605 |
| 56 | 8.744E-13 | 9.667E-13 | 10.55 | 1.454E-12 | 1.560E-12 | 7.290 |



Table 3.3: Accuracy of Approximations ($\rho = 0.8$, $c = 0.8$, $\alpha = 0.6$)

| | Early arrival model | | | Late arrival model | | |
|---|---|---|---|---|---|---|
| N | exact | approximate | error(%) | exact | approximate | error(%) |
| 25 | 7.945E-03 | 8.042E-03 | 1.228 | 9.096E-03 | 9.190E-03 | 1.037 |
| 30 | 4.106E-03 | 4.156E-03 | 1.209 | 4.691E-03 | 4.739E-03 | 1.020 |
| 35 | 2.138E-03 | 2.163E-03 | 1.200 | 2.440E-03 | 2.464E-03 | 1.011 |
| 40 | 1.117E-03 | 1.130E-03 | 1.195 | 1.274E-03 | 1.287E-03 | 1.006 |
| 45 | 5.849E-04 | 5.919E-04 | 1.192 | 6.669E-04 | 6.736E-04 | 1.003 |
| 50 | 3.065E-04 | 3.102E-04 | 1.191 | 3.494E-04 | 3.529E-04 | 1.002 |
| 55 | 1.607E-04 | 1.626E-04 | 1.190 | 1.832E-04 | 1.851E-04 | 1.001 |
| 60 | 8.431E-05 | 8.531E-05 | 1.190 | 9.610E-05 | 9.706E-05 | 1.001 |
| 65 | 4.423E-05 | 4.475E-05 | 1.190 | 5.041E-05 | 5.092E-05 | 1.001 |
| 70 | 2.320E-05 | 2.348E-05 | 1.189 | 2.645E-05 | 2.671E-05 | 1.001 |
| 75 | 1.217E-05 | 1.232E-05 | 1.189 | 1.388E-05 | 1.401E-05 | 1.001 |
| 80 | 6.387E-06 | 6.463E-06 | 1.189 | 7.280E-06 | 7.353E-06 | 1.001 |
| 85 | 3.351E-06 | 3.391E-06 | 1.189 | 3.820E-06 | 3.858E-06 | 1.001 |
| 90 | 1.758E-06 | 1.779E-06 | 1.189 | 2.004E-06 | 2.024E-06 | 1.001 |
| 95 | 9.225E-07 | 9.335E-07 | 1.189 | 1.051E-06 | 1.062E-06 | 1.001 |
| 100 | 4.840E-07 | 4.898E-07 | 1.189 | 5.517E-07 | 5.572E-07 | 1.001 |
| 105 | 2.540E-07 | 2.570E-07 | 1.189 | 2.895E-07 | 2.924E-07 | 1.001 |
| 110 | 1.332E-07 | 1.348E-07 | 1.189 | 1.519E-07 | 1.534E-07 | 1.001 |
| 115 | 6.991E-08 | 7.074E-08 | 1.189 | 7.968E-08 | 8.048E-08 | 1.001 |
| 120 | 3.668E-08 | 3.712E-08 | 1.189 | 4.181E-08 | 4.223E-08 | 1.001 |
| 125 | 1.925E-08 | 1.947E-08 | 1.189 | 2.194E-08 | 2.216E-08 | 1.000 |
| 130 | 1.010E-08 | 1.022E-08 | 1.189 | 1.151E-08 | 1.162E-08 | 1.000 |
| 135 | 5.298E-09 | 5.361E-09 | 1.190 | 6.039E-09 | 6.099E-09 | 1.000 |
| 140 | 2.780E-09 | 2.813E-09 | 1.190 | 3.168E-09 | 3.200E-09 | 0.999 |
| 145 | 1.458E-09 | 1.476E-09 | 1.190 | 1.662E-09 | 1.679E-09 | 0.998 |
| 150 | 7.652E-10 | 7.743E-10 | 1.191 | 8.722E-10 | 8.809E-10 | 0.995 |
| 155 | 4.015E-10 | 4.063E-10 | 1.192 | 4.577E-10 | 4.622E-10 | 0.989 |
| 160 | 2.106E-10 | 2.132E-10 | 1.194 | 2.402E-10 | 2.425E-10 | 0.977 |
| 165 | 1.105E-10 | 1.118E-10 | 1.199 | 1.260E-10 | 1.272E-10 | 0.955 |
| 170 | 5.798E-11 | 5.868E-11 | 1.209 | 6.616E-11 | 6.676E-11 | 0.910 |

Table 3.4: Ranges of Buffer Size and Relative Error

$\alpha = 0.6$

| | | Early arrival model | | Late arrival model | |
|---|---|---|---|---|---|
| $\lambda$ | c | range of error | range of buffer | range of error | range of buffer |
| 0.2 | 0.2 | $0.644 \sim 0.753$ | $8 \sim 14$ | $0.400 \sim 0.404$ | $9 \sim 15$ |
| 0.2 | 0.5 | $1.774 \sim 1.807$ | $9 \sim 17$ | $1.261 \sim 1.387$ | $10 \sim 18$ |
| 0.2 | 0.8 | $2.320 \sim 2.322$ | $10 \sim 19$ | $2.078 \sim 2.268$ | $11 \sim 20$ |
| 0.5 | 0.2 | $0.628 \sim 0.633$ | $17 \sim 34$ | $0.317 \sim 0.322$ | $18 \sim 35$ |
| 0.5 | 0.5 | $1.933 \sim 1.935$ | $20 \sim 41$ | $1.259 \sim 1.276$ | $21 \sim 42$ |
| 0.5 | 0.8 | $2.575 \sim 2.578$ | $25 \sim 51$ | $2.216 \sim 2.223$ | $26 \sim 52$ |
| 0.8 | 0.2 | $0.220 \sim 0.220$ | $47 \sim 1018$ | $0.114 \sim 0.120$ | $48 \sim 102$ |
| 0.8 | 0.5 | $0.854 \sim 0.859$ | $58 \sim 125$ | $0.528 \sim 0.559$ | $59 \sim 126$ |
| 0.8 | 0.8 | $1.189 \sim 1.199$ | $77 \sim 165$ | $0.948 \sim 1.001$ | $78 \sim 166$ |

$\alpha = 0.8$

| | | Early arrival model | | Late arrival model | |
|---|---|---|---|---|---|
| $\lambda$ | c | range of error | range of buffer | range of error | range of buffer |
| 0.2 | 0.2 | $2.825 \sim 3.299$ | $8 \sim 15$ | $1.833 \sim 1.848$ | $97 \sim 16$ |
| 0.2 | 0.5 | $7.275 \sim 7.359$ | $9 \sim 17$ | $5.311 \sim 5.689$ | $10 \sim 18$ |
| 0.2 | 0.8 | $9.763 \sim 9.768$ | $11 \sim 20$ | $8.776 \sim 9.254$ | $11 \sim 21$ |
| 0.5 | 0.2 | $3.738 \sim 3.749$ | $18 \sim 35$ | $2.236 \sim 2.249$ | $19 \sim 36$ |
| 0.5 | 0.5 | $10.712 \sim 10.713$ | $23 \sim 46$ | $7.826 \sim 7.832$ | $24 \sim 47$ |
| 0.5 | 0.8 | $14.41 \sim 14.42$ | $30 \sim 60$ | $12.81 \sim 12.82$ | $31 \sim 61$ |
| 0.8 | 0.2 | $1.506 \sim 1.513$ | $49 \sim 105$ | $1.010 \sim 1.042$ | $59 \sim 106$ |
| 0.8 | 0.5 | $5.278 \sim 5.283$ | $70 \sim 150$ | $3.945 \sim 3.988$ | $71 \sim 151$ |
| 0.8 | 0.8 | $6.782 \sim 6.789$ | $104 \sim 224$ | $5.961 \sim 5.976$ | $105 \sim 225$ |

$\alpha = 0.95$

| | | Early arrival model | | Late arrival model | |
|---|---|---|---|---|---|
| $\lambda$ | c | range of error | range of buffer | range of error | range of buffer |
| 0.2 | 0.2 | $6.105 \sim 6.946$ | $8 \sim 15$ | $4.082 \sim 4.177$ | $9 \sim 16$ |
| 0.2 | 0.5 | $14.94 \sim 15.08$ | $9 \sim 18$ | $11.19 \sim 11.79$ | $10 \sim 19$ |
| 0.2 | 0.8 | $21.04 \sim 21.04$ | $10 \sim 21$ | $19.06 \sim 19.53$ | $12 \sim 22$ |
| 0.5 | 0.2 | $14.85 \sim 14.91$ | $19 \sim 38$ | $10.79 \sim 10.80$ | $20 \sim 39$ |
| 0.5 | 0.5 | $41.66 \sim 41.67$ | $28 \sim 57$ | $33.91 \sim 33.92$ | $29 \sim 58$ |
| 0.5 | 0.8 | $61.23 \sim 61.24$ | $42 \sim 88$ | $56.44 \sim 56.45$ | $44 \sim 90$ |
| 0.8 | 0.2 | $10.50 \sim 10.52$ | $59 \sim 129$ | $9.542 \sim 9.561$ | $60 \sim 130$ |
| 0.8 | 0.5 | $25.39 \sim 25.41$ | $124 \sim 272$ | $22.32 \sim 22.35$ | $125 \sim 273$ |
| 0.8 | 0.8 | $24.38 \sim 24.38$ | $244 \sim 534$ | $22.41 \sim 22.41$ | $246 \sim 535$ |



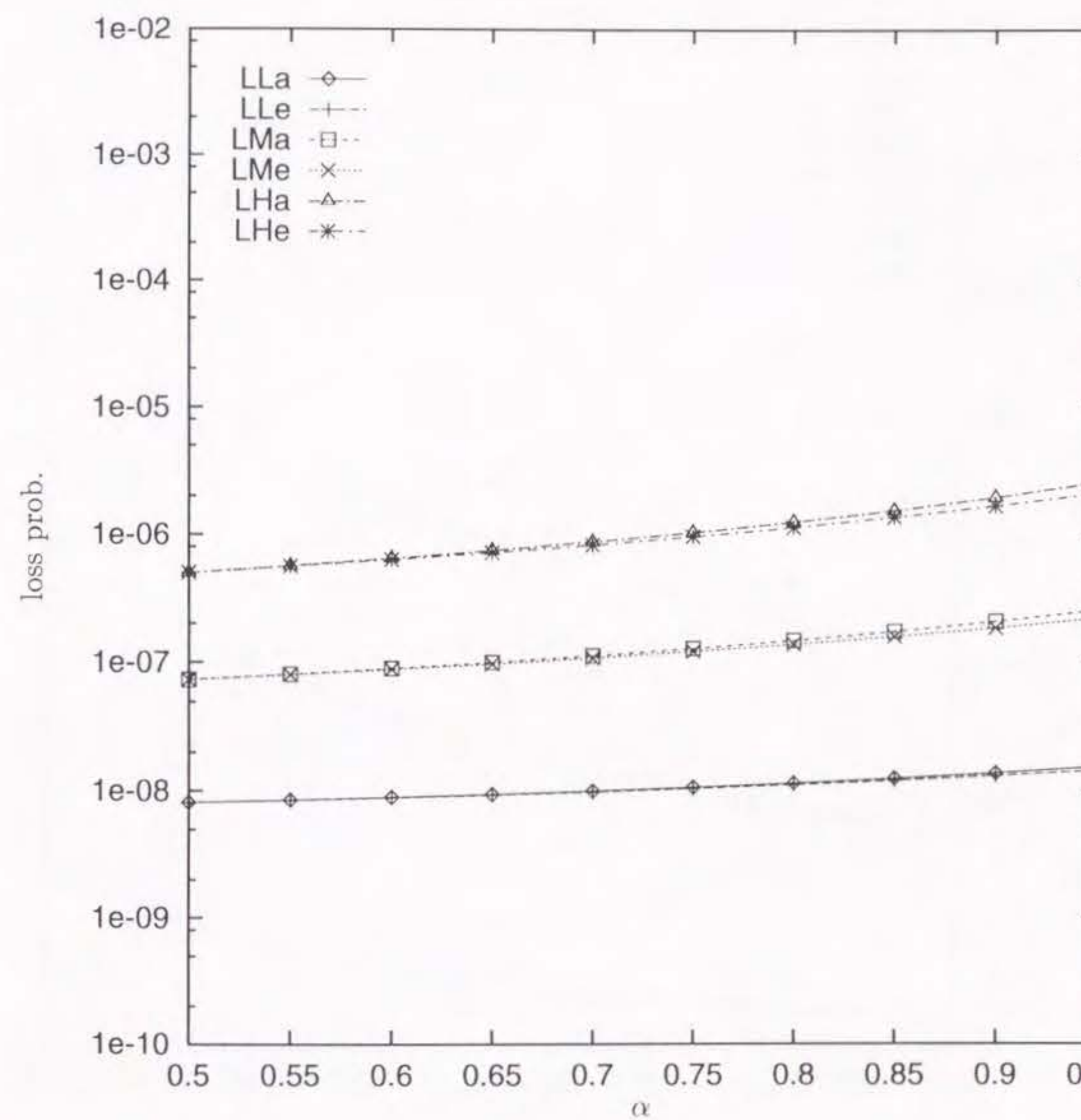


Figure 3.1: Loss Probability Approximation as a Function of $\alpha$

(Early Arrival Model, $N = 12$, $\rho = 0.2$)

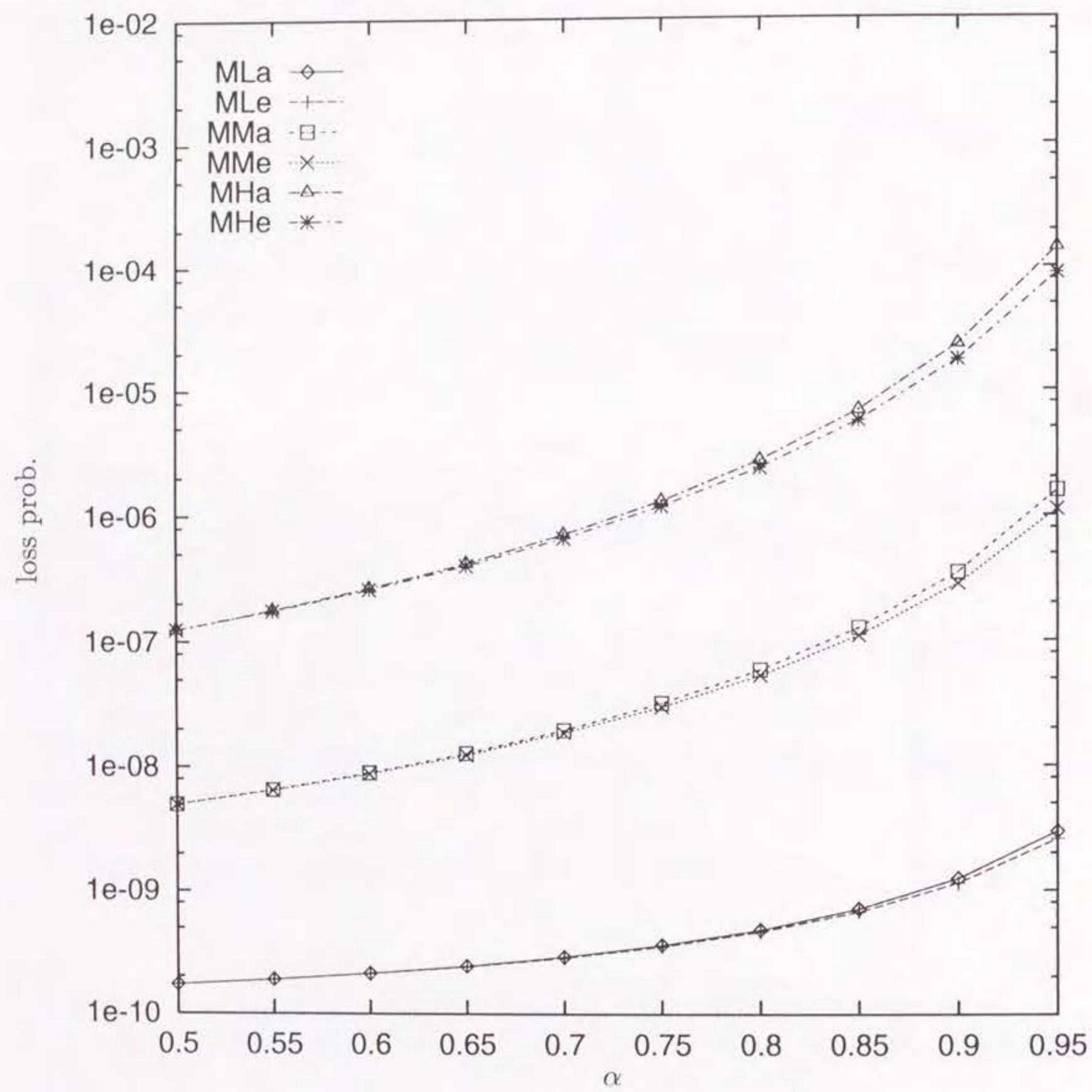


Figure 3.2: Loss Probability Approximation as a Function of $\alpha$

(Early Arrival Model, $N = 33$, $\rho = 0.5$)



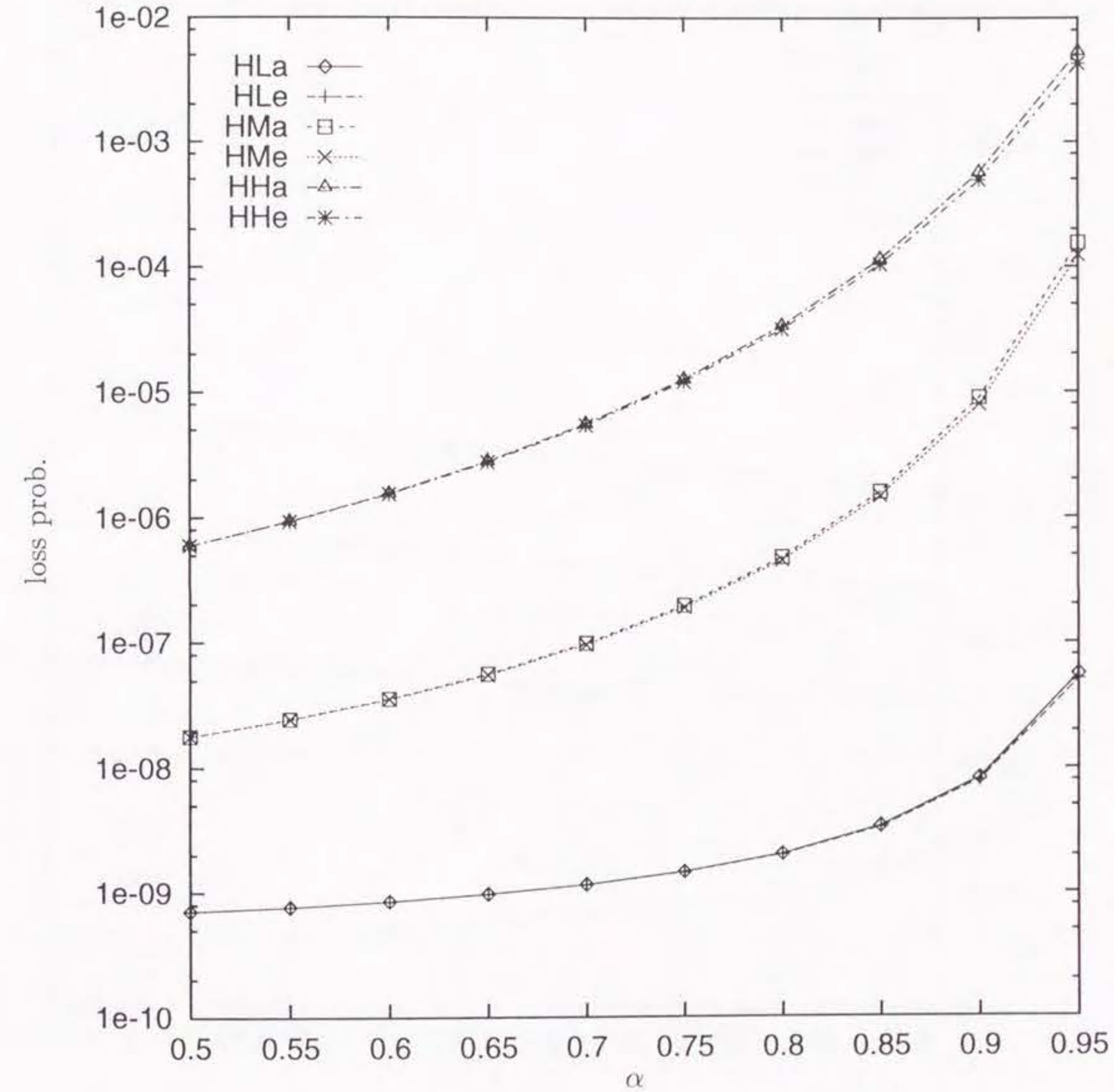


Figure 3.3: Loss Probability Approximation as a Function of $\alpha$

(Early Arrival Model, $N = 91$, $\rho = 0.8$)

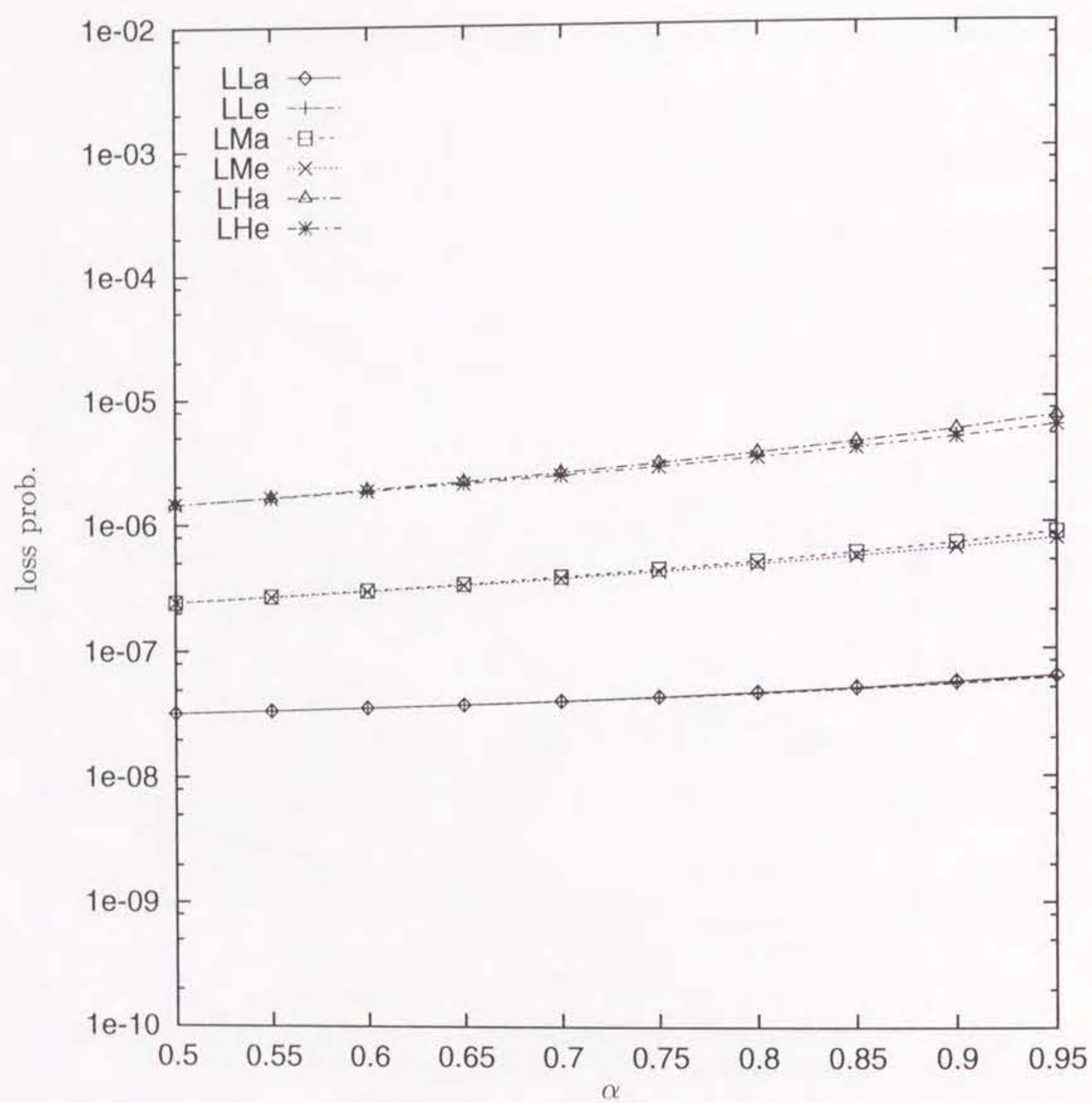


Figure 3.4: Loss Probability Approximation as a Function of $\alpha$

(Late Arrival Model, $N = 12$, $\rho = 0.2$)



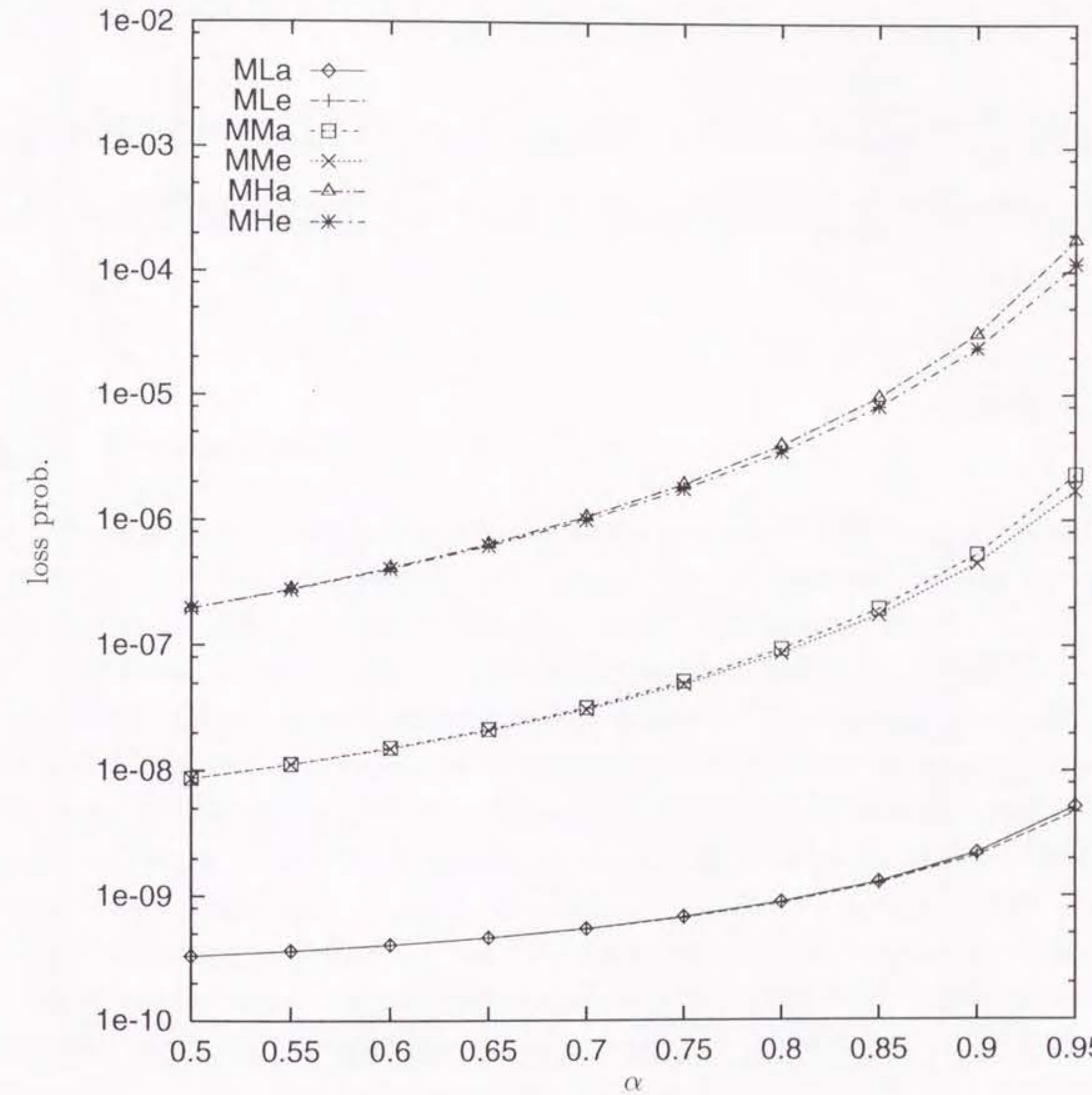


Figure 3.5: Loss Probability Approximation as a Function of $\alpha$

(Late Arrival Model, $N = 33$, $\rho = 0.5$)

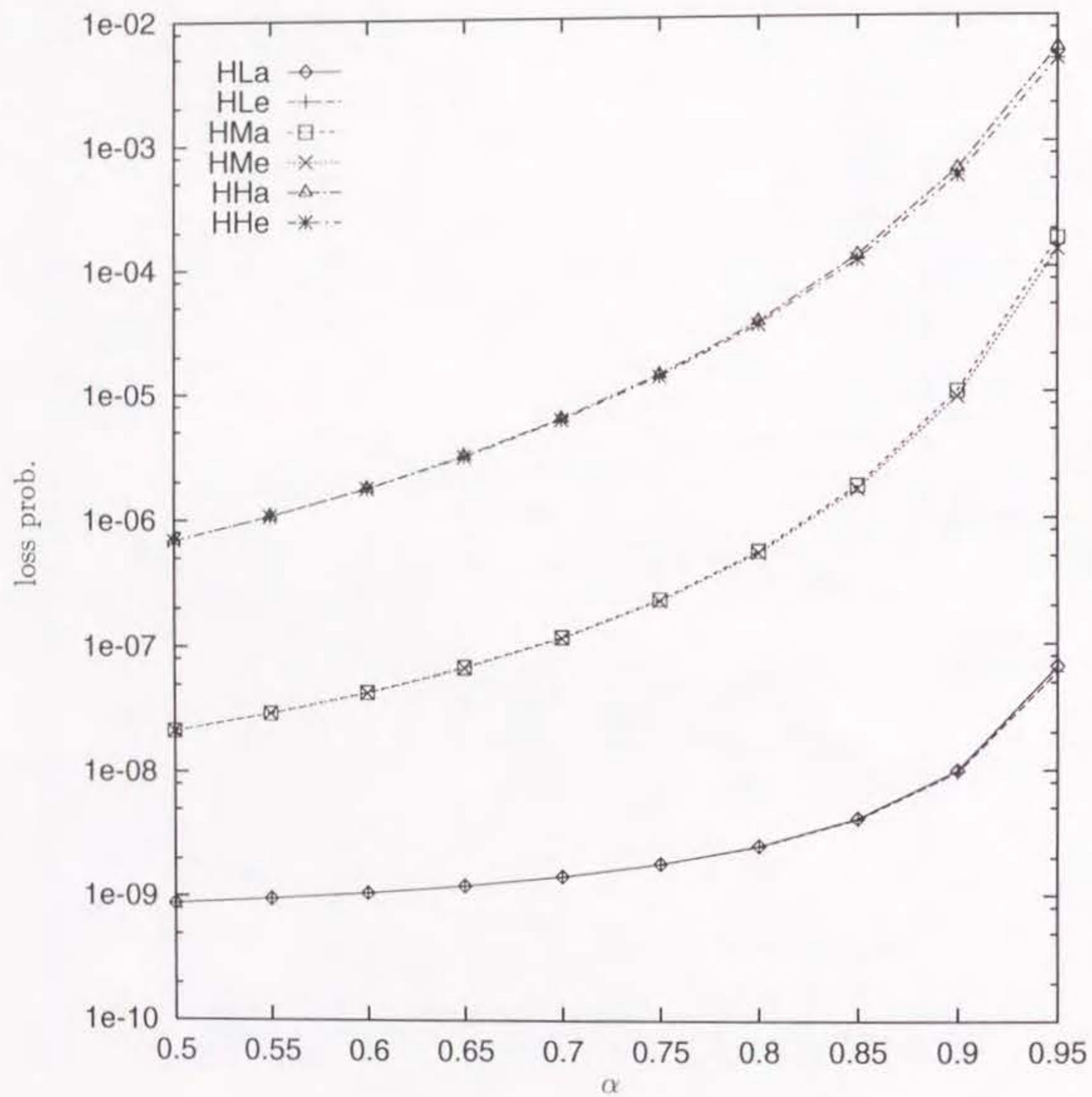


Figure 3.6: Loss Probability Approximation as a Function of $\alpha$

(Late Arrival Model, $N = 91$, $\rho = 0.8$)



# Chapter 4

# Queues with a Gate - Geometrically Distributed Gate Opening Intervals -

## 4.1 Introduction

In this chapter, we consider discrete-time BBP/G/1 queues with a gate, where BBP denotes a batch Bernoulli process. We assume that the intervals between successive openings of the gate are geometrically distributed. The purpose of this chapter is to provide a complete set of the analytical results for various performance measures.

The organization of this chapter is as follows. In section 4.2, we describe the mathematical model. Our model is considered as a discrete-time version of the model of [Bors93]. The model in this chapter, however, allows batch arrivals, while [Bors93] considers only single arrivals. In the next three sections, we provide various formulas of the performance measures of interest. In section 4.3, we study the number of customers in the system. We first derive the joint PGF for the numbers of customers in the first queue and the second queue immediately after departures of customers. The PGF is given in terms of a function which is represented by an infinite product. Next we derive the joint PGF for the numbers of customers in the first queue and the second queue at the beginning of a randomly chosen slot. Note that [Bors93] did not provide any results on the joint queue length distribution at a random point in time. Furthermore, we analytically show the decomposition properties for the total number of customers in the system at departures and at a randomly chosen slot. In section 4.4, we analyze the amount of work in the system. Using the joint distribution of the queue lengths and the remaining service time, we first derive the joint PGF for the amounts of work in the first queue and the second queue at the beginning of a randomly chosen slot. Next we derive the PGF for the amount of total work in the system. Furthermore, we show the decomposition property for the amount of total work in the system. Note that the PGF for the amount of work in the system is identical to the PGF for the sojourn time of a supercustomer [Kawa93]. In section 4.5, we consider the waiting times of customers. We derive the joint PGF for the waiting times of individual customers in the first queue and the second queue, and the PGF for the waiting time of a supercustomer. Also we analytically show the decomposition property for the total waiting time of individual customers. Finally, in section 4.6, we provide some numerical examples, where we discuss three kinds of correlations in the model: the effect of the correlation between the interarrival time and the service time of supercustomers on the mean waiting time of supercustomers, the effect of the correlation between the interarrival time of each batch composed of customers who move to the second queue at the same time and the number of the customers in the batch on the mean waiting time of individual customers in the second queue, and the correlation between the waiting times in the first queue and the second queue.

## 4.2 Model

We consider a discrete-time queueing model with the following characteristics:

- Customers arrive at the system in a batch immediately before slot boundaries. The batch sizes and the service times of individual customers are independent and identically distributed. Customers arriving at the system are accommodated in the first queue at the gate.

- The gate opens immediately before slot boundaries. When the gate opens, all the customers waiting in the first queue move to the second queue at the server. The travel times of customers to the second queue are assumed to be zero. We assume that customers arriving in a slot also move to the second queue when the gate opens in the slot, so that the waiting times of such customers in the first queue become zero. The gate closes immediately after all the customers in the first queue move to the second queue. The intervals between successive openings of the gate are geometrically distributed.

- There is a single server who serves the customers only in the second queue. When the server finds some amount of work in the second queue immediately after a slot boundary, he serves exactly one unit of work in the current slot. We assume that customers are served on an FCFS basis. Furthermore, as for customers who arrive in the same slot, the next customer for service is randomly chosen among those customers.



We now introduce random variables and notations to describe the above model. Let $B$ and $C$ denote random variables representing the number of individual customers who arrive at the system in a slot and the service time of an individual customer, respectively. Further, let $A$ denote a random variable representing the amount of work brought into the system in a slot (i.e., the sum of the service times of customers arriving in a slot). We define the following PGFs:

$$A(z) \triangleq E\left[z^A\right], \qquad B(z) \triangleq E\left[z^B\right], \qquad C(z) \triangleq E\left[z^C\right]. \tag{4.1}$$

By definition, we have

$$A(z) = B(C(z)). \tag{4.2}$$

Let $G$ denote a random variable representing the length of an interval between successive openings of the gate. Let $g(n) = \Pr\{G = n\}$ $(n \geq 1)$. We then have for a parameter $\gamma$

$$g(n) = (1-\gamma)\gamma^{n-1} \qquad (0 \leq \gamma \leq 1). \tag{4.3}$$

We denote the PGF of the $g(n)$ by $G(z)$:

$$G(z) = \sum_{n=1}^{\infty} g(n)z^n = \frac{(1-\gamma)z}{1-\gamma z}. \tag{4.4}$$

We assume that $B$, $C$ and $G$ are independent, identically distributed random variables, and those are independent each other. Throughout this chapter, for any PGF $f(z)$, we use the symbol $f'(1)$ to denote $\lim_{z\to 1-} df(z)/dz$. Furthermore, we assume $A'(1) = B'(1)C'(1) < 1$ and the system is in equilibrium.

## 4.3 Number of Individual Customers

In this section, we consider the numbers of individual customers in the first queue and the second queue. First we observe an imbedded Markov chain which is composed of two types of imbedded points. Next we derive the PGF for the number of customers immediately after departures of customers. Finally, we obtain the PGF for the number of customers at the beginning of a randomly chosen slot in terms of the PGF for the number of customers immediately after departures of customers.

### 4.3.1 Number of Customers Immediately after Departures

In this subsection, we derive the formula for the number of customers immediately after departures. To do so, we introduce an imbedded Markov chain which is composed of two types of imbedded points:

- type 1: immediately after departures of individual customers,
- type 2: immediately after gate opening instants during idle periods.

Let $X^{(1)}$ and $X^{(2)}$ denote random variables representing the numbers of individual customers in the first queue and the second queue, respectively, at a randomly chosen imbedded point. Note here that $X^{(1)}$ and $X^{(2)}$ are dependent. Moreover, let $T_P$ denote a random variable representing the type of a randomly chosen imbedded point, i.e., $T_P = i$ $(i = 1, 2)$ when a randomly chosen imbedded point is of type $i$. We define $P(z_1, z_2)$ and $Q(z_2)$ as

$$P(z_1, z_2) \triangleq E\left[z_1^{X^{(1)}} z_2^{X^{(2)}} \mathbf{1}_{\{T_P=1\}}\right], \tag{4.5}$$

$$Q(z_2) \triangleq E\left[z_2^{X^{(2)}} \mathbf{1}_{\{T_P=2\}}\right], \tag{4.6}$$

where $\mathbf{1}_T$ denotes the indicator function of a set $T$. Note here that $X^{(1)} = 0$ if $T_P = 2$ because all the customers waiting in the first queue move to the second queue when the gate opens.

Let $X_D^{(1)}$ and $X_D^{(2)}$ denote random variables representing the numbers of individual customers in the first queue and the second queue, respectively, immediately after the departure of a randomly chosen customer. Note here that $X_D^{(1)}$ and $X_D^{(2)}$ are dependent. We define $Q_D(z_1, z_2)$ as the joint PGF associated with $X_D^{(1)}$ and $X_D^{(2)}$:

$$Q_D(z_1, z_2) \triangleq E\left[z_1^{X_D^{(1)}} z_2^{X_D^{(2)}}\right]. \tag{4.7}$$

We then have the following theorem.

THEOREM 4.3.1: *$Q_D(z_1, z_2)$ satisfies*

$$\begin{aligned}[z_2 - C(\gamma B(z_1))]Q_D(z_1, z_2) &= C(\gamma B(z_1))[\Psi(z_2) - \Psi(z_1)] \\ &\quad + \left[\frac{z_2}{C(B(z_2))}\theta(z_1, z_2) - C(\gamma B(z_1))\right] \\ &\quad \cdot \frac{C(B(z_2))}{z_2 - C(B(z_2))} \frac{B(z_2) - 1}{1 - \gamma B(z_2)} \Psi(z_2),\end{aligned} \tag{4.8}$$

*where*

$$\Psi(z) = \frac{P(z, 0) + Q(0)}{P(1, 1)}, \tag{4.9}$$

$$\theta(z_1, z_2) = C(\gamma B(z_1)) + H(z_1, z_2) \tag{4.10}$$

*with*

$$H(z_1, z_2) = \frac{(1-\gamma)B(z_2)}{B(z_2) - \gamma B(z_1)} \left[C(B(z_2)) - C(\gamma B(z_1))\right]. \tag{4.11}$$

PROOF: See Appendix C.1. ∎



Note that (4.8) is rewritten to be

$$\begin{aligned}Q_D(z_1, z_2) &= \frac{1}{z_2 - C(\gamma B(z_1))}\Big[C(\gamma B(z_1))\{Q_D(z_2, 0) - Q_D(z_1, 0)\} \\ &\quad + \left\{\frac{z_2}{C(B(z_2))}\theta(z_1, z_2) - C(\gamma B(z_1))\right\} Q_D(z_2, z_2)\Big],\end{aligned} \tag{4.12}$$

The equation (4.12) may be useful because it can replace $Q_D(z_1, z_2)$ by $Q_D(z_1, 0)$, $Q_D(z_2, 0)$ and $Q_D(z_2, z_2)$. Indeed, we will use (4.12) in proofs.

(4.8) shows that $Q_D(z_1, z_2)$ is expressed in terms of $\Psi(z)$ $(z = z_1, z_2)$. We shall therefore consider $\Psi(z)$. Define for $|z_1| \le 1$:

$$\delta(z_1) \triangleq C(\gamma B(z_1)), \tag{4.13}$$

$$\delta^{(0)}(z_1) \triangleq z_1, \tag{4.14}$$

$$\delta^{(i)}(z_1) \triangleq \delta(\delta^{(i-1)}(z_1)) \qquad (i \ge 1). \tag{4.15}$$

Since the system is stable, $Q_D(z_1, z_2)$ is bounded and analytic for $|z_1| \le 1$ and $|z_2| \le 1$. Thus, for $z_2 = \delta(z_1)$, the left-hand side of (4.8) becomes zero, so that the right-hand side of (4.8) must become zero, too. We then have

$$\Psi(z_1) = \frac{(1-\gamma)B(\delta(z_1))(1 - \gamma B(z_1))}{(B(\delta(z_1)) - \gamma B(z_1))(1 - \gamma B(\delta(z_1)))} \Psi(\delta(z_1)). \tag{4.16}$$

To simplify notations, we introduce for $|z_1| \le 1$,

$$\phi(z_1) \triangleq \frac{(1-\gamma)B(\delta(z_1))(1 - \gamma B(z_1))}{(B(\delta(z_1)) - \gamma B(z_1))(1 - \gamma B(\delta(z_1)))}. \tag{4.17}$$

Then (4.16) becomes

$$\Psi(z_1) = \phi(z_1)\Psi(\delta(z_1)). \tag{4.18}$$

Using (4.18), we have the following theorem.

THEOREM 4.3.2: *$\Psi(z) = (P(z, 0) + Q(0))/P(1, 1)$ is given by*

$$\Psi(z) = \frac{1 - B'(1)C'(1)}{B'(1)}(1-\gamma)\frac{\alpha(z)}{\alpha(1)}, \tag{4.19}$$

*where for $|z_1| \le 1$,*

$$\alpha(z_1) = \prod_{h=0}^{\infty} \phi(\delta^{(h)}(z_1)). \tag{4.20}$$

PROOF: See Appendix C.2. ∎

Next we present a corollary which immediately follows from Theorem 4.3.1. Let $X_A^{(1)}$ and $X_A^{(2)}$ denote random variables representing the numbers of individual customers in the first queue and the second queue, respectively, at the beginning of a randomly chosen slot. Note here that $X_A^{(1)}$ and $X_A^{(2)}$ are dependent. We define $Q_{A|busy}(z_1, z_2)$ as the joint PGF for the numbers of individual customers in the first queue and the second queue at the beginning of a randomly chosen slot given that the server is busy. Also, we define $Q_{A|idle}(z_1)$ as the PGF for the number of individual customers in the first queue at the beginning of a randomly chosen slot given that the server is idle:

$$Q_{A|busy}(z_1, z_2) \triangleq E\left[z_1^{X_A^{(1)}} z_2^{X_A^{(2)}} \middle| T_S = 1\right], \quad Q_{A|idle}(z_1) \triangleq E\left[z_1^{X_A^{(1)}} \middle| T_S = 0\right], \tag{4.21}$$

where $T_S$ denotes a random variable defined as

$$T_S \triangleq \begin{cases} 1 & \text{if the server is busy,} \\ 0 & \text{if the server is idle,} \end{cases} \tag{4.22}$$

at a randomly chosen slot. Note that, by definition, $Q_{A|idle}(z)$ is given by

$$Q_{A|idle}(z) = \frac{1-\gamma}{1-\gamma B(z)} \frac{P(z,0)+Q(0)}{P(1,0)+Q(0)}. \tag{4.23}$$

Corollary 4.3.1: *The PGF $Q_D(z,z)$ for the total number of customers in the system immediately after departures is given by*

$$Q_D(z,z) = Q_{Db}(z) Q_{A|idle}(z), \tag{4.24}$$

*where*

$$Q_{Db}(z) = (1 - B'(1)C'(1)) \frac{(z-1)C(B(z))}{z - C(B(z))} \cdot \frac{B(z)-1}{B'(1)(z-1)} \tag{4.25}$$

*and $Q_{A|idle}(z)$ is given in (4.23).*

Proof: See Appendix C.4. ∎

*Remark 4.3.1:* Note that $Q_{Db}(z)$ denotes the PGF for the number of customers immediately after departures of customers corresponding BBP/G/1 queue without gates and $Q_{A|idle}(z)$ denotes the PGF for the number of individual customers in the system given that the server is idle. Thus, the total number of customers in the system immediately after departures are decomposed into the two independent factors.



### 4.3.2 Number of Customers in a Randomly Chosen Slot

In this subsection, we derive the formula for the number of customers at the beginning of a randomly chosen slot. To do so, we first consider the number of customers in the first queue and that in the second queue at the start of the service of a randomly chosen customer. Let $\hat{X}^{(1)}$ and $\hat{X}^{(2)}$ denote random variables representing the numbers of individual customers in the first queue and the second queue, respectively, at the start of the service of a randomly chosen customer. We define $\hat{Q}(z_1, z_2)$ as the joint PGF for the numbers of individual customers in the first queue and the second queue at the start of the service of a randomly chosen customer:

$$\hat{Q}(z_1, z_2) \triangleq E\left[z_1^{\hat{X}^{(1)}} z_2^{\hat{X}^{(2)}}\right]. \tag{4.26}$$

We then have the following lemma.

*Lemma 4.3.1:* $\hat{Q}(z_1, z_2)$ is given by

$$\begin{aligned} \hat{Q}(z_1, z_2) &= \frac{z_2}{z_2 - C(\gamma B(z_1))}\{Q_D(z_2, 0) - Q_D(z_1, 0)\} \\ &\quad + \frac{z_2}{C(B(z_2))} S(z_1, z_2) Q_D(z_2, z_2), \end{aligned} \tag{4.27}$$

where

$$S(z_1, z_2) = \frac{z_2 + H(z_1, z_2) - C(B(z_2))}{z_2 - C(\gamma B(z_1))} \tag{4.28}$$

and $H(z_1, z_2)$ is given in (4.11).

Proof: See Appendix C.5. ∎

Taking $z_1 = z_2 = z$ in (4.27) and noting

$$S(z,z) = 1, \tag{4.29}$$

we can easily confirm that the following equation holds:

$$\hat{Q}(z,z) = \frac{z}{C(B(z))} Q_D(z,z). \tag{4.30}$$

This equation implies that the number of customers in the system left behind by a randomly chosen customer is equal to the number of customers in the system at the start of his service (including himself) plus the number of customers arriving to the system during his service time minus one (himself).

We now consider the number of customers in the first queue and that in the second queue at the beginning of a randomly chosen slot. We define $Q_A(z_1, z_2)$ as the joint PGF associated with $X_A^{(1)}$ and $X_A^{(2)}$:

$$Q_A(z_1, z_2) \triangleq E\left[z_1^{X_A^{(1)}} z_2^{X_A^{(2)}}\right]. \tag{4.31}$$

We then have the following theorem.

THEOREM 4.3.3: *$Q_A(z_1, z_2)$ is given by*

$$\begin{aligned} Q_A(z_1,z_2) &= B'(1)C'(1)\hat{\theta}(z_1,z_2)\frac{z_2}{C(B(z_2))}Q_D(z_2,z_2) \\ &+B'(1)C'(1)\frac{z_2}{z_2-C(\gamma B(z_1))}\frac{1-C(\gamma B(z_1))}{C'(1)(1-\gamma B(z_1))} \\ &\cdot\{Q_D(z_2,0)-Q_D(z_1,0)\}+(1-B'(1)C'(1))\frac{Q_D(z_1,z_1)}{Q_{Db}(z_1)}, \end{aligned} \quad (4.32)$$

*where*

$$\hat{\theta}(z_1,z_2)=\hat{H}(z_1,z_2)+\frac{1-C(\gamma B(z_1))}{C'(1)(1-\gamma B(z_1))}S(z_1,z_2) \quad (4.33)$$

*with*

$$\hat{H}(z_1,z_2)=\frac{(1-\gamma)B(z_2)}{C'(1)(B(z_2)-\gamma B(z_1))}\left[\frac{B(z_2)-C(B(z_2))}{1-B(z_2)}-\frac{\gamma B(z_1)-C(\gamma B(z_1))}{1-\gamma B(z_1)}\right]. \quad (4.34)$$

PROOF: See Appendix C.6, which Lemma 4.3.1 is used in. ∎

We present a corollary which immediately follows from Theorem 4.3.2.

COROLLARY *4.3.2: The PGF $Q_A(z,z)$ for the total number of customers at the beginning of a randomly chosen slot is given by*

$$Q_A(z,z)=Q_{Ab}(z)Q_{A|idle}(z), \quad (4.35)$$

*where*

$$Q_{Ab}(z)=(1-B'(1)C'(1))\frac{(z-1)C(B(z))}{z-C(B(z))} \quad (4.36)$$

*and $Q_{A|idle}(z)$ is given in (4.23).*

PROOF: Setting $z_1=z_2=z$ in (4.32), noting

$$\hat{\theta}(z,z)=\frac{C(B(z))-1}{C'(1)(B(z)-1)}, \quad (4.37)$$

and using Corollary 4.3.1, it follows that

$$\begin{aligned} Q_A(z,z) &= B'(1)C'(1)\frac{z}{C(B(z))}\frac{C(B(z))-1}{C'(1)(B(z)-1)}Q_D(z,z) \\ &+(1-B'(1)C'(1))\frac{Q_D(z,z)}{Q_{Db}(z)} \\ &= \frac{B'(1)(z-1)}{B(z)-1}Q_D(z,z) \\ &= \frac{(z-1)C(B(z))}{z-C(B(z))}(1-B'(1)C'(1))\frac{1-\gamma}{1-\gamma B(z)}\frac{P(z,0)+Q(0)}{P(1,0)+Q(0)}, \end{aligned} \quad (4.38)$$



from which, (4.35) follows. ∎

*Remark 4.3.2:* Note that $Q_{Ab}(z)$ denotes the PGF for the number of customers at the beginning of a randomly chosen slot in the corresponding BBP/G/1 queue without gates. This decomposition result is a discrete-time example of the general result for the continuous-time queue given in [Fuhr85].

## 4.4 Work in the System

In this section, we consider the amounts of work in the first queue and the second queue. To obtain the formulas for the amount of work in the system, we first derive the formula for the joint PGF for the numbers of customers and the remaining service time at the beginning of a randomly chosen slot.

Let $\tilde{X}^{(1)}$ (resp. $\tilde{X}^{(2)}$) denote a random variable representing the number of individual customers who arrive and remain in the first queue (resp. arrive and move to the second queue) during the backward recurrence time of the service time of a customer who is served in a randomly chosen slot. Also, let $\hat{C}$ denote a random variable representing the forward recurrence time of the service time of a customer who is served in a randomly chosen slot. We define $Q_{A|busy}(z_1,z_2,w)$ as the joint PGF for the numbers of customers who arrive and remain in the first queue and customers who arrive and move to the second queue during the backward recurrence time of the service time of a customer who is served in a randomly chosen slot, and the forward recurrence time of the service time of the customer given that the server is busy:

$$Q_{A|busy}(z_1,z_2,w)\triangleq E\left[z_1^{\tilde{X}^{(1)}}z_2^{\tilde{X}^{(2)}}w^{\hat{C}}\middle|T_S=1\right]. \quad (4.39)$$

We then have the following lemma.

*Lemma 4.4.1:* The joint PGF $Q_{A|busy}(z_1,z_2,w)$ is given by

$$\begin{aligned} Q_{A|busy}(z_1,z_2,w) &= \hat{\theta}(z_1,z_2,w)\frac{z_2}{C(B(z_2))}Q_D(z_2,z_2) \\ &+\frac{z_2}{z_2-C(\gamma B(z_1))}\frac{w(C(w)-C(\gamma B(z_1)))}{C'(1)(w-\gamma B(z_1))} \\ &\cdot\{Q_D(z_2,0)-Q_D(z_1,0)\}, \end{aligned} \quad (4.40)$$

where

$$\hat{\theta}(z_1,z_2,w)=\hat{H}(z_1,z_2,w)+S(z_1,z_2)\frac{w(C(w)-C(\gamma B(z_1)))}{C'(1)(w-\gamma B(z_1))} \quad (4.41)$$

with

$$\begin{aligned}\hat{H}(z_1,z_2,w) &= \frac{(1-\gamma)B(z_2)}{C'(1)(B(z_2)-\gamma B(z_1))}\left[\frac{B(z_2)C(w)-wC(B(z_2))}{w-B(z_2)}\right.\\ &\quad \left.-\frac{\gamma B(z_1)C(w)-wC(\gamma B(z_1))}{w-\gamma B(z_1)}\right], \qquad (4.42)\end{aligned}$$

and $S(z_1,z_2)$ is given in (4.28).

PROOF: See Appendix C.7. ∎

We now consider the amount of work at the beginning of a randomly chosen slot. Let $U^{(1)}$ (resp. $U^{(2)}$) denote a random variable representing the amount of work in the first queue (resp. the second queue) at the beginning of a randomly chosen slot. Note here that $U^{(1)}$ and $U^{(2)}$ are dependent. We define the joint PGF $U(z_1,z_2)$ associated with $U^{(1)}$ and $U^{(2)}$:

$$U(z_1,z_2) \triangleq E\left[z_1^{U^{(1)}}z_2^{U^{(2)}}\right]. \qquad (4.43)$$

We then have the following theorem.

THEOREM 4.4.1: *The joint PGF $U(z_1,z_2)$ is given by*

$$\begin{aligned}U(z_1,z_2) &= (1-B'(1)C'(1))\frac{z_2(A(z_2)-1)}{z_2-A(z_2)}\frac{z_2-\gamma A(z_2)}{z_2-\gamma A(z_1)}\frac{Q_D(C(z_2),C(z_2))}{Q_{Db}(C(z_2))}\\ &\quad +(1-B'(1)C'(1))\frac{Q_D(C(z_1),C(z_1))}{Q_{Db}(C(z_1))}\\ &\quad +B'(1)\frac{z_2}{z_2-\gamma A(z_1)}\{Q_D(C(z_2),0)-Q_D(C(z_1),0)\}. \qquad (4.44)\end{aligned}$$

PROOF: See Appendix C.8, which Lemma 4.4.1 is used in. ∎

We present a corollary which immediately follows from Theorem 4.4.1. Let $U=U^{(1)}+U^{(2)}$ denote the amount of total work in the system in a randomly chosen slot and we define $U(z)$ as the PGF for $U$.

COROLLARY *4.4.1: The PGF $U(z)$ is given by*

$$U(z)=U_b(z)U_{idle}(z), \qquad (4.45)$$

*where*

$$\begin{aligned}U_b(z) &= (1-B'(1)C'(1))\frac{(z-1)A(z)}{z-A(z)}, \qquad (4.46)\\ U_{idle}(z) &= Q_{A|idle}(C(z))\\ &= \frac{1-\gamma}{1-\gamma A(z)}\frac{P(C(z),0)+Q(0)}{P(1,0)+Q(0)}. \qquad (4.47)\end{aligned}$$



PROOF: Using (C.35) in (4.44), it follows that

$$\begin{aligned}U(z) &= U(z,z)\\ &= \frac{(z-1)A(z)}{z-A(z)}(1-B'(1)C'(1))\frac{1-\gamma}{1-\gamma A(z)}\frac{P(C(z),0)+Q(0)}{P(1,0)+Q(0)}, \qquad (4.48)\end{aligned}$$

from which, (4.45) immediately follows. ∎

*Remark 4.4.1:*

1. Note that $U(z)$ is identical to the PGF for the sojourn time of supercustomers and coincides with the result in [Kawa93].
2. $U_b(z)$ denotes the PGF for the amount of work in the corresponding BBP/G/1 queue without gates, and $U_{idle}(z)$ denotes the PGF for the amount of work in the first queue given that the server is idle. Thus (4.45) shows that the amount of the total work in the system is decomposed into the two independent factors. This is a discrete-time example for the work decomposition property in the queue with the generalized vacations [Boxm89].

Note that, with (4.45) and noting $(P(1,0)+Q(0))/P(1,1)=(1-\gamma)(1-B'(1)C'(1))/B'(1)$ from Theorem 4.3.1, (4.44) is rewritten to be

$$U(z_1,z_2)=\frac{z_2}{z_2-\gamma A(z_1)}(1-\gamma)U(z_2)+\frac{\gamma(z_1-A(z_1))}{z_2-\gamma A(z_1)}\frac{z_2-1}{z_1-1}U(z_1). \qquad (4.49)$$

Th equation (4.49) may be useful because it can replace $U(z_1,z_2)$ by $U(z_1)$ and $U(z_2)$. Indeed, we will use (4.49) in order to derive the PGFs for the waiting times in the next section.

## 4.5 Waiting Times

In this section, we consider the waiting times of a supercustomer and an individual customer. We first derive the PGF for the waiting time of a randomly chosen supercustomer in terms of the PGF for the amount of work in the system. Next we obtain the PGFs for the waiting times of a randomly chosen individual customer in terms of the PGF for the amount of work in the system.

### 4.5.1 Waiting Time of a Supercustomer

In this subsection, we consider the waiting time of a randomly chosen supercustomer. We define a supercustomer as a batch composed of individual customers moving to the second queue at the same time when the gate opens. Note here that it is possible that there is no individual customer in the first queue when the gate opens. We regard such a case as an arrival of a supercustomer with zero service time at the second queue. Let $W_s$ denote a random variable representing the waiting time of a randomly chosen supercustomer. We define $W_s(z)$ as the PGF for $W_s$. We then have the following theorem.

THEOREM 4.5.1: *The PGF $W_s(z)$ is given by*

$$W_s(z) = \frac{1-\gamma}{z-\gamma}U(z) + \frac{z-1}{z-\gamma}(1 - B'(1)C'(1)), \tag{4.50}$$

*where $U(z)$ is given in (4.45).*

PROOF: It follows that

$$W_s(z) = \frac{U(1,z) - U(1,0)}{z} + U(1,0). \tag{4.51}$$

Using (4.49) in (4.51), (4.50) immediately follows. ■

*Remark 4.5.1:* After some algebra with (4.49) and (4.50), we have the following relationship between the work in the second queue and the waiting time and the sojourn time of a supercustomer:

$$U(1,z) = 1 - A'(1) + A'(1)z\frac{W_s(z) - U(z)}{G'(1)A'(1)(1-z)} \tag{4.52}$$

Note that (4.52) can also be derived from the equality of the virtual delay and attained waiting time distribution (see, for example, [Miya92, Saka90, Seng89]).

### 4.5.2 Waiting Time of a Customer

In this subsection, we consider the waiting time of a randomly chosen individual customer. Let $W_c^{(1)}$ (resp. $W_c^{(2)}$) denote a random variable representing the waiting time of a randomly chosen customer in the first queue (resp. the second queue). Note here that $W_c^{(1)}$ and $W_c^{(2)}$ are dependent. We define $W_c(z_1, z_2)$ as the joint PGF for $W_c^{(1)}$ and $W_c^{(2)}$:

$$W_c(z_1, z_2) \triangleq E\left[z_1^{W_c^{(1)}} z_2^{W_c^{(2)}}\right]. \tag{4.53}$$

We then have the following theorem.

THEOREM 4.5.2: *The joint PGF $W_c(z_1, z_2)$ is given by*

$$\begin{aligned} W_c(z_1, z_2) &= \frac{1-\gamma}{z_2 - \gamma z_1}\frac{1}{z_1 - A(z_2)}\left[(z_1 - z_2)U(z_2) + \frac{1-\gamma}{1-\gamma z_1}z_1(z_2 - 1)U(\gamma z_1)\right] \\ &\quad \cdot\frac{1 - A(z_2)}{B'(1)(1 - C(z_2))}, \end{aligned} \tag{4.54}$$



*where $U(z)$ is given in (4.45).*

PROOF: We divide the waiting time of a randomly chosen customer into three parts:

$$W_c^{(1)} = F^{(1)}, \quad W_c^{(2)} = F^{(2)} + D, \tag{4.55}$$

where $F^{(1)}$ (resp. $F^{(2)}$ ) denotes a random variable representing the waiting time of the batch which includes the randomly chosen customer in the first queue (resp. in the second queue), and $D$ denotes a random variable representing the sum of the service times of customers who arrive in the same batch as the randomly chosen customer and are served before the randomly chosen customer. Note here that $F^{(1)} + F^{(2)}$ and $D$ are independent.

Now we define the following PGFs:

$$F(z_1, z_2) \triangleq E\left[z_1^{F^{(1)}} z_2^{F^{(2)}}\right], \qquad D(z) \triangleq E\left[z^D\right]. \tag{4.56}$$

First we consider $F(z_1, z_2)$. Let $\hat{G}_A$ denote a random variable representing the remaining gate opening interval. Also, let $W_A^{(1)}$ and $W_A^{(2)}$ denote random variables representing the amounts of work in the first queue and the second queue, respectively, immediately before the arrival of a randomly chosen customer. Note that the joint distribution of the amount of work immediately before arrivals is identical to that at the beginning of a randomly chosen slot, since customers arrive to the system according to a batch Bernoulli process [Boxm88]. Thus, it follows that

$$\begin{aligned} F(z_1, z_2) &= E\left[z_1^{\hat{G}_A} z_2^{W_A^{(1)} + W_A^{(2)} - \hat{G}_A}\mathbf{1}_{\{W_A^{(2)} \geq \hat{G}_A\}}\right] + E\left[z_1^{\hat{G}_A} z_2^{W_A^{(1)}}\mathbf{1}_{\{W_A^{(2)} < \hat{G}_A\}}\right] \\ &= \sum_{k=0}^{\infty}\sum_{n=1}^{\infty}\sum_{m=n}^{\infty}(1-\gamma)\gamma^{n-1}z_1^{n-1}z_2^{k+m-1-(n-1)}\Pr\{U^{(1)} = k, U^{(2)} = m\} \\ &\quad + \sum_{k=0}^{\infty}\sum_{m=0}^{\infty}\sum_{n=m+1}^{\infty}(1-\gamma)\gamma^{n-1}z_1^{n-1}z_2^{k}\Pr\{U^{(1)} = k, U^{(2)} = m\} \\ &= \frac{1-\gamma}{z_2 - \gamma z_1}\{U(z_2, z_2) - U(z_2, \gamma z_1)\} + \frac{1-\gamma}{1-\gamma z_1}U(z_2, \gamma z_1). \end{aligned} \tag{4.57}$$

On the other hand, $D(z)$ is given by

$$D(z) = \frac{1 - A(z)}{B'(1)(1 - C(z))}. \tag{4.58}$$

We then have

$$W_c(z_1, z_2) = F(z_1, z_2)D(z_2). \tag{4.59}$$

Using (4.49) in (4.59), (4.54) immediately follows. ■

Let $W_c = W_c^{(1)} + W_c^{(2)}$ denote the total waiting time of a randomly chosen customer in the system. We then define $W_c(z)$, $W_{c1}(z)$ and $W_{c2}(z)$ as the PGFs for $W_c$, $W_c^{(1)}$ and

$W_c^{(2)}$, respectively. Now we present a corollary which immediately follows from Theorem 4.5.2.

COROLLARY *4.5.1: The PGFs $W_{c1}(z)$,$W_{c2}(z)$ and $W_c(z)$ are given by*

$$W_{c1}(z) = \frac{1-\gamma}{1-\gamma z}, \tag{4.60}$$

$$W_{c2}(z) = \frac{1-\gamma}{z-\gamma}\frac{1-z}{1-A(z)}\left[U(z) - (1-B'(1)C'(1))\right]\frac{1-A(z)}{B'(1)(1-C(z))} \tag{4.61}$$

$$W_c(z) = W_b(z)U_{back}(z), \tag{4.62}$$

*respectively, where*

$$W_b(z) = (1-B'(1)C'(1))\frac{z-1}{z-A(z)}\cdot\frac{1-A(z)}{B'(1)(1-C(z))}, \tag{4.63}$$

$$U_{back}(z) = \frac{1-\gamma}{1-\gamma z}\frac{U(\gamma z)}{1-B'(1)C'(1)}, \tag{4.64}$$

*and $U(z)$ is given in (4.45).*

PROOF: Letting $z_2 = 1$, $z_1 = 1$ and $z_1 = z_2 = z$ in (4.54), we obtain (4.60), (4.61) and (4.62), respectively. ∎

*Remark 4.5.2:*

1. $W_b(z)$ denotes the PGF for the waiting times of customers in the corresponding BBP/G/1 queue without gates, and $U_{back}(z)$ denotes the PGF for the backward recurrence time of the gate opening interval given that the server is idle. This is a discrete-time example of the waiting-time decomposition property in the queue with generalized vacations [Fuhr85].

2. After some algebra with (4.49) and (4.61), we have the following relationship between the work in the second queue and the waiting time and the sojourn time of an individual customer in the second queue:

$$U(1,z) = 1 - A'(1) + A'(1)zW_{c2}(z)\frac{1-C(z)}{C'(1)(1-z)} \tag{4.65}$$

Note that (4.65) can also be derived from the equality of the virtual delay and attained waiting time distribution (see, for example, [Miya92, Saka90, Seng89]).



## 4.6 Numerical Examples

In this section, we provide some numerical examples. First we regard the second queue as an isolated system and observe the effect of the gate opening interval on the mean waiting time. More precisely, we consider the gate opening interval in terms of the covariances and the correlation coefficients. In section 4.1, we mentioned two types of correlations:

- type 1: correlation between the interarrival time $G$ and the service time $C_S$ of each supercustomer,

- type 2: correlation between the interarrival time $G$ of each batch composed of customers who move to the second queue at the same time and the number $B_G$ of the customers.

Then, the covariances and the correlation coefficients for the two types of correlation are given by

$$Cov[G, C_S] = \frac{\gamma}{(1-\gamma)^2}A'(1), \tag{4.66}$$

$$correl[G, C_S] = \left[\frac{\gamma(A'(1))^2}{(1-\gamma)(A'(1)+A''(1)) + (2\gamma-1)(A'(1))^2}\right]^{1/2}, \tag{4.67}$$

$$Cov[G, B_G] = \frac{\gamma}{(1-\gamma)^2}B'(1), \tag{4.68}$$

$$correl[G, B_G] = \left[\frac{\gamma(B'(1))^2}{(1-\gamma)(B'(1)+B''(1)) + (2\gamma-1)(B'(1))^2}\right]^{1/2}, \tag{4.69}$$

respectively. Note here that the correlation coefficients are increasing functions of the mean gate opening interval.

Now we observe the effect of the gate opening interval on the mean waiting time of supercustomers. We show the formula for the mean waiting time of supercustomers in Appendix C.9. To compare the result, we also consider a corresponding Geo/G/1 queue where the PGF for the service time of a customer is $G(A(z))$ and the PGF for the interarrival time of customers is $G(z)$. Fig. 4.1 shows the mean waiting time of (super)customers in the second queue as a function of the parameter $\gamma$ in the following settings: (1) the number of individual customers arriving to the system in a slot is geometrically distributed with mean 0.6, (2) the service times of individual customers are deterministic and equal to one slot. Note here that the increase of the parameter $\gamma$ implies the increase of the correlation coefficient between the interarrival time and the service time of each supercustomer. In Fig. 4.1, we observe that the correlation, which is positive, leads to the reduction of the mean waiting time of supercustomers, whereas the mean waiting time increases with the increase of the correlation coefficient. A similar observation has been shown in [Bors93].

Next, we observe the effect of the gate opening interval on the mean waiting time of individual customers in the second queue. We show the formula for the mean waiting time of individual customers in Appendix C.10. To compare the result, we also consider a corresponding BBP/G/1 queue where the PGF for the service time of a customer is $C(z)$ and the PGF for the batch size (arriving to the second queue) is $(1-\gamma)G(B(z))+\gamma$. Fig. 4.2 shows the mean waiting time of individual customers in the second queue as a function of the parameter $\gamma$ in the same settings as those in Fig. 4.1. Note here that the increase of the parameter $\gamma$ implies the increase of the correlation coefficient between the interarrival time of the batches and the number of customers in each batch. In Fig. 4.2, we also observe that the positive correlation leads to the reduction of the mean waiting time of individual customers in the second queue, whereas the mean waiting time increases with the increase of the correlation coefficient.

Finally we observe the correlation between the waiting times of a randomly chosen individual customer in the first queue and in the second queue. Fig. 4.3 shows the correlation coefficients between the waiting times of a randomly chosen individual customer in the first queue and in the second queue, which are obtained by using numerical differentiation as a function of the parameter $\gamma$ in the following settings: (1) the number of individual customers arriving to the system in a slot is geometrically distributed with mean 0.4, 0.6 and 0.8, (2) the service times of individual customers are deterministic and equal to one slot. In Fig. 4.3, we observe that, as expected, the correlation is negative and the correlation coefficient decreases with the increase of the parameter $\gamma$. Further, we observe that the increase of the traffic intensity leads to the increase of the correlation coefficient.

## 4.7 Conclusion

In this chapter, we consider discrete-time BBP/G/1 queues with a gate, where the intervals between successive openings of the gate are geometrically distributed. We derive the following joint PGFs:

- joint PGF for the numbers of customers in the first queue and the second queue at the beginning of a randomly chosen slot
- joint PGF for the amounts of work in the first queue and the second queue at the beginning of a randomly chosen slot
- joint PGF for the waiting times of individual customers in the first queue and the second queue

We also derive the PGF for the sojourn time and the waiting time of a supercustomer. Furthermore, we provide some numerical examples and observe the effect of the correlations on the performance measures through the numerical examples.



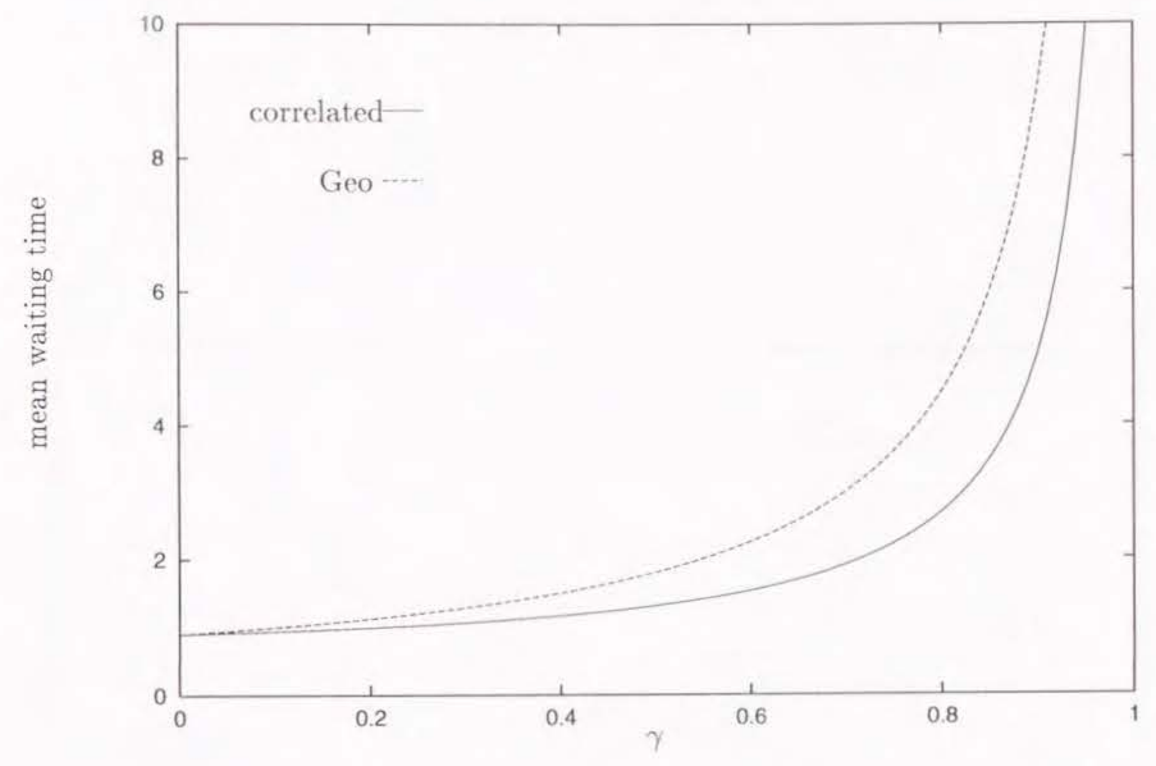


Figure 4.1: Effect of correlation (1)

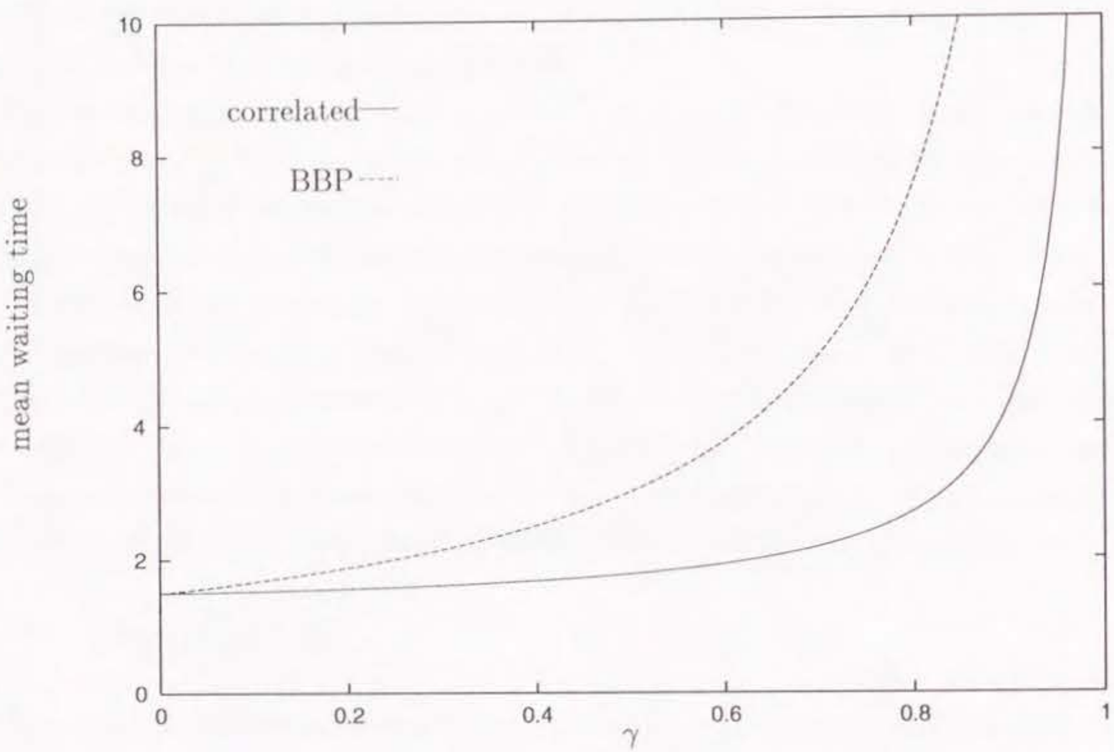


Figure 4.2: Effect of correlation (2)



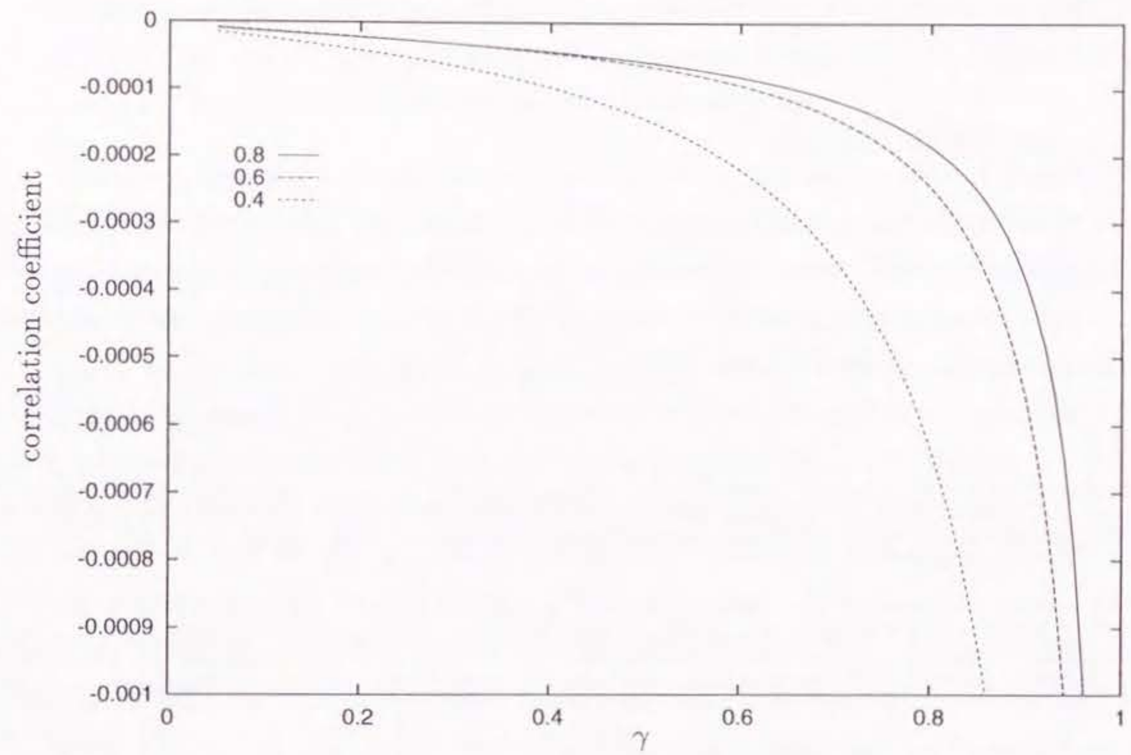


Figure 4.3: Correlation between the waiting times in each queue

# Chapter 5

# Queues with a Gate - Bounded Gate Opening Intervals -

## 5.1 Introduction

In this chapter, we consider discrete-time BBP/G/1 queues with a gate, where BBP denotes a batch Bernoulli process. Contrary to the model in the previous chapter, we assume that the intervals between successive openings of the gate are bounded, and independent and identically distributed (i.i.d.).

The organization of the chapter is as follows. In section 5.2, we describe the mathematical model. In section 5.3, we derive the joint PGF for the amounts of the stationary work in the first queue and the second queue, and analytically show the work decomposition property for the amount of work in the system. In section 5.4, we first derive the PGFs for the sojourn time and the waiting time of a supercustomer, and analytically show the relationship among the PGFs for the amount of the work in the second queue, the sojourn time and the waiting time of a supercustomer. Next, we derive the joint PGF for the waiting times of an individual customer in the first queue and the second queue, and analytically show the decomposition property for the total waiting time. Furthermore, we show the relationship between the PGFs for the amount of the work in the second queue and the waiting time in the second queue. In section 5.5, we discuss the number of individual customers in the system. Using the results in section 5.4, we derive the PGF for the number of individual customers in the system, and analytically show the queue length decomposition property. In section 5.6, we provide numerical examples to show the computational feasibility of the analytical results.

## 5.2 Model

We consider the queueing model with the following characteristics:



- Customers arrive at the system in a batch immediately before slot boundaries. The batch sizes and the service times of individual customers are independent and identically distributed. Customers arriving at the system are accommodated in the first queue at the gate.
- The gate opens immediately before the slot boundaries. When the gate opens, all the customers waiting in the first queue move to the second queue at the server. The travel times of customers to the second queue are assumed to be zero. We assume that customers arriving in a slot also move to the second queue when the gate opens in the slot, so that the waiting times of such customers in the first queue become zero. The gate closes immediately after all the customers in the first queue move to the second queue. The intervals between successive openings of the gate are i.i.d.. We assume that the gating process is stationary.
- There is a single server who serves the customers only in the second queue. When the server finds some amount of the work in the second queue immediately after a slot boundary, he serves exactly one unit of the work in the current slot. We assume that customers are served on an FCFS basis. Furthermore, as for customers who arrive in the same slot, the next customer for service is randomly chosen among those customers.

Note that when the gate opens in every slot (i.e., all the gate opening intervals are equal to one slot), the model described above is reduced to the ordinary BBP/G/1 queue without gates.

We now introduce random variables and notations to describe the above model. Let $\hat{T}_n$ denote a random variable representing the $n$th gate opening epoch taking a integer value. We assume that the sequence $\{\hat{T}_n\}_{n=-\infty}^{+\infty}$ satisfies

$$\cdots < \hat{T}_{-1} < \hat{T}_0 \leq 0 < \hat{T}_1 < \hat{T}_2 < \cdots. \tag{5.1}$$

We define the inter-event sequence $\{G_n\}_{n=-\infty}^{+\infty}$ as

$$G_n = \hat{T}_{n+1} - \hat{T}_n. \tag{5.2}$$

Let $g(k)$ $(k \geq 1)$ denote the probability that an interval between successive openings of the gate is equal to $k$ slots:

$$g(k) = \Pr\{G_n = k\} \qquad (n \neq 0). \tag{5.3}$$

In what follows, we assume that the intervals between successive openings of the gate are bounded by $M$ slots with $g(M) \neq 0$. Thus, we have $g(k) = 0$ for $k > M$. We denote the

PGF of the $g(k)$ by $G(z)$:

$$G(z) = \sum_{k=1}^{M} g(k) z^k. \tag{5.4}$$

Let $B$ and $C$ denote random variables representing the number of individual customers who arrive at the system in a slot and the service time of an individual customer, respectively. Further, let $A$ denote a random variable representing the amount of work brought into the system in a slot (i.e., the sum of the service times of customers arriving in a slot). We define the following PGFs:

$$A(z) \triangleq E\left[z^A\right], \qquad B(z) \triangleq E\left[z^B\right], \qquad C(z) \triangleq E\left[z^C\right]. \tag{5.5}$$

By definition, we have

$$A(z) = B(C(z)). \tag{5.6}$$

Throughout the chapter, for any PGF $f(z)$, we use the symbol $f'(1)$ to denote $\lim_{z\to 1-} df(z)/dz$. Further, we assume $A'(1) < 1$ and the system is in equilibrium.

## 5.3 Work in the System

In this section, we consider the amount of the work in the system. For convenience of the analysis, we assign non-negative integer values $k \in \{0, 1, 2, ...\}$ sequentially to individual slot boundaries as 0 is assigned to the slot boundaries immediately after the gate closes. Time interval $[k-1, k)$ $(k = 1, 2, ...)$ is referred to as the $k$th slot. First, we derive the PGF for the amount of the work in the second queue immediately after the beginning of the $(k+1)$st slot. Next, we derive the PGF for the amount of the stationary work in the system as well as in each queue.

### 5.3.1 Work in the Second Queue

In this subsection, we first observe the amount of the work in the second queue immediately after the beginning of the $(k+1)$st slot and relate it with that immediately after the beginning of the $k$th slot. In what follows, we refer to the amount of the work immediately after the beginning of the $k$th slot as that in the $k$th slot. Let $U_n^{(2)}$ denote a random variable representing the amount of work in the second queue at time $n$. We define $L(k, z)$ as the PGF for the amount of the work in the second queue in the $(k+1)$st slot is equal to $n$, given that the gate does not open in the $k$th slot:

$$L(k, z) \triangleq E\left[z^{U_0^{(2)}} \middle| -\hat{T}_0 = k\right] \qquad (i = 1, 2, k = 0, 1, 2, \ldots). \tag{5.7}$$



Relating $L(k, z)$ with $L(k-1, z)$, we have

$$L(k, z) = \frac{1}{z} L(k-1, z) + \frac{z-1}{z} L(k-1, 0) \qquad (k = 1, 2, \ldots). \tag{5.8}$$

By applying (5.8) recursively, we obtain

$$L(k, z) = \left(\frac{1}{z}\right)^k L(0, z) + (z-1) \sum_{j=1}^{k} \left(\frac{1}{z}\right)^j L(k-j, 0). \tag{5.9}$$

Since the system is in equilibrium, it is clear that the PGF $L(0, z)$ is given by

$$L(0, z) = \sum_{k=1}^{M} g(k) L(k, z) [A(z)]^k. \tag{5.10}$$

Substituting (5.9) into (5.10), we have the following expression for $L(0, z)$:

$$L(0, z) = G\left(\frac{A(z)}{z}\right) L(0, z) + (z-1) \sum_{k=0}^{M-1} L(k, 0) z^k \sum_{j=k+1}^{M} g(j) \left(\frac{A(z)}{z}\right)^j. \tag{5.11}$$

We solve (5.11) with respect to $L(0, z)$ and obtain

$$L(0, z) = \frac{(z-1) X(A(z)/z, z)}{1 - G(A(z)/z)}, \tag{5.12}$$

where $X(z, w)$ is given by

$$X(z, w) = \sum_{k=0}^{M-1} L(k, 0) w^k \sum_{j=k+1}^{M} g(j) z^j. \tag{5.13}$$

We then have the $M$ unknown values $L(k, 0)$ $(k = 0, .., M-1)$ in (5.13), which can be determined (see Appendix D). Note here that, from the normalizing condition $L(0, 1) = 1$ in (5.11), we obtain

$$1 = \frac{X(1, 1)}{G'(1)(1 - A'(1))}. \tag{5.14}$$

The equation (5.14) is used when we determine the unknown values.

### 5.3.2 Stationary Work

In this subsection, we first consider the amount of the stationary work in each queue. Let $U_n^{(1)}$ denote a random variable representing the amount of work in the first queue at time $n$. Note that $U_n^{(1)}$ and $U_n^{(2)}$ are dependent. We define the joint PGF $U(z_1, z_2)$ for the amounts of the stationary work in the first queue and in the second queue:

$$U(z_1, z_2) \triangleq E\left[z_1^{U_0^{(1)}} z_2^{U_0^{(2)}}\right]. \tag{5.15}$$

We then have the following theorem.

THEOREM 5.3.1: *The joint PGF $U(z_1, z_2)$ is given by*

$$U(z_1, z_2) = \frac{1}{G'(1)\{1 - A(z_1)/z_2\}} \left[ L(0, z_2) \left\{ 1 - G\left(\frac{A(z_1)}{z_2}\right) \right\} + (z_2 - 1) \left\{ \frac{A(z_1)}{z_2} X(1, A(z_1)) - X\left(\frac{A(z_1)}{z_2}, z_2\right) \right\} \right]. \tag{5.16}$$

PROOF: Let $p(k)$ denote where $p(k)$ denotes the probability that a randomly chosen slot is the $(k+1)$st slot:

$$p(k) \triangleq \Pr\{-\hat{T}_0 = k\} \qquad (k = 0, 1, 2, \ldots, M-1). \tag{5.17}$$

We then have [Burk75]

$$p(k) = \frac{1}{G'(1)} \sum_{n=k+1}^{M} g(n) \qquad (k = 0, 1, 2, \ldots, M-1). \tag{5.18}$$

By definition, we have

$$U(z_1, z_2) = \sum_{k=0}^{M-1} p(k) L(k, z_2) (A(z_1))^k. \tag{5.19}$$

From (5.9), (5.13), (5.18) and (5.19), (5.16) immediately follows. ∎

We present a corollary which immediately follows from Theorem 5.3.1. Let $U = U_0^{(1)} + U_0^{(2)}$ denote the amount of the stationary total work in the system. We then define $U_1(z)$, $U_2(z)$ and $U(z)$ as the PGFs for $U_0^{(1)}$, $U_0^{(2)}$ and $U$, respectively.

COROLLARY *5.3.1: The PGFs $U_1(z)$, $U_2(z)$ and $U(z)$ are given by*

$$U_1(z) = \frac{1 - G(A(z))}{G'(1)(1 - A(z))}, \tag{5.20}$$

$$U_2(z) = \frac{1}{G'(1)\{1 - 1/z\}} \left[ L(0, z) \left\{ 1 - G\left(\frac{1}{z}\right) \right\} + (z-1) \left\{ \frac{1}{z} X(1, 1) - X\left(\frac{1}{z}, z\right) \right\} \right], \tag{5.21}$$

$$U(z) = U_b(z) \hat{G}_{idle}(A(z)), \tag{5.22}$$

*respectively, where*

$$U_b(z) = \frac{(z-1)A(z)}{z - A(z)} (1 - A'(1)), \tag{5.23}$$

$$\hat{G}_{idle}(z) = \frac{X(1, z)}{(1 - A'(1))G'(1)} = \sum_{k=0}^{M-1} p(k) \frac{L(k, 0)}{1 - A'(1)} z^k, \tag{5.24}$$



PROOF: Letting $z_2 = 1$, $z_1 = 1$ and $z_1 = z_2 = z$ in (5.16) and using (5.13), we obtain (5.20), (5.21) and (5.22), respectively. ∎

*Remark 5.3.1:* $U_b(z)$ denotes the PGF for the amount of the stationary work in the corresponding BBP/G/1 queue without gates, and $\hat{G}_{idle}(A(z))$ denotes the PGF for the amount of the work in the first queue given that the server is idle. Thus (5.22) shows that the amount of the total work in the system is decomposed into the two independent factors. This is a discrete-time example for the work decomposition property in the queue with generalized vacations [Boxm89].

## 5.4 Waiting Times

In this section, we first consider the waiting time and the sojourn time of a supercustomer, where each batch of customers who move to the second queue at the same time is called a supercustomer. Next we consider the waiting time of an individual customer.

### 5.4.1 Waiting Time of a Supercustomer

In this subsection, we consider the waiting time and the sojourn time of a supercustomer in the second queue. Note here that it is possible that there is no individual customer in the first queue when the gate opens. We regard such a case as an arrival of a supercustomer with zero service time at the second queue. Let $W_s$ denote a random variable representing the waiting time of a supercustomer. We define $W_s(z)$ as the PGF for $W_s$. We then have the following theorem.

THEOREM 5.4.1: *The PGF $W_s(z)$ is given by*

$$W_s(z) = G\left(\frac{1}{z}\right) L(0, z) + (z-1) X\left(\frac{1}{z}, z\right). \tag{5.25}$$

PROOF: It is clear that

$$W_s(z) = \sum_{k=1}^{M} g(k) L(k, z). \tag{5.26}$$

From (5.9), (5.13) and (5.26), (5.25) immediately follows. ∎

Next we consider the sojourn time of a supercustomer. Let $R_s$ denote a random variable representing the sojourn time of a supercustomer. We define $R_s(z)$ as the PGF for $R_s$. We then have the following theorem.

THEOREM 5.4.2: *The PGF $R_s(z)$ is given by*

$$R_s(z) = L(0, z). \tag{5.27}$$

PROOF: Observe that the sojourn time of a supercustomer is identical to the amount of the work in the first slot. (5.27) immediately follows. ■

*Remark 5.4.1:* Note here that $R_s(z)$ is *not* given by the product of $W_s(z)$ and the PGF $G(A(z))$ for the service times of supercustomers, since the waiting times and the service times of supercustomers are dependent.

*Remark 5.4.2:* After some algebra with (5.14), (5.21), (5.25) and (5.27), we have the following relationship among the amount of the work in the second queue, the waiting time and the sojourn time of a supercustomer:

$$U_2(z) = 1 - A'(1) + A'(1)z\frac{W_s(z) - R_s(z)}{G'(1)A'(1)(1-z)}. \tag{5.28}$$

Note that (5.28) can also be derived from the equality of the virtual delay and attained waiting time distribution (see, for example, [Miya92, Saka90, Seng89]).

### 5.4.2 Waiting Time of a Customer

In this subsection, we consider the waiting time of an individual customers. We first derive the PGF for the total waiting time of a randomly chosen individual customer in the system. Let $\gamma_{nk}$ denote the joint probability that a randomly chosen customer belongs to the supercustomer who arrives to the second queue in the $k$th slot ($k = 1, 2, \ldots, M$) and belongs to the batch which arrives in the $n$th ($n = 1, \ldots, k$) slot. We call this customer (resp. this supercustomer) a tagged customer (resp. a tagged supercustomer). It follows that

$$\gamma_{nk} = \frac{g(k)}{G'(1)} \quad (k = 1, \ldots, M, n = 1, \ldots, k). \tag{5.29}$$

We define $F(k, z)$ as the PGF for the amount of the work in the second queue seen by the tagged supercustomer arriving in the $k$th slot (i.e., the waiting time of the supercustomer). It follows that

$$F(k, z) = \frac{L(k-1, z) - L(k-1, 0)}{z} + L(k-1, 0) = L(k, z) \qquad (k = 1, \ldots, M). \tag{5.30}$$

Furthermore, we define $D(n, z)$ as the PGF for a time interval from the beginning of the service of the first customer of the tagged supercustomer arriving in the $k$th slot to the beginning of the service of the tagged customer (who has arrived in the $n$th slot). By using the results in batch arrivals, we have

$$D(n, z) = (A(z))^{n-1}\frac{1 - A(z)}{B'(1)(1 - C(z))} \quad (n = 1, \ldots, k). \tag{5.31}$$

Let $W_c^{(1)}$ (resp. $W_c^{(2)}$) denote a random variable representing the waiting time of a randomly chosen customer in the first queue (resp. in the second queue). Note here that $W_c^{(1)}$ and $W_c^{(2)}$ are dependent. We define $W(z_1, z_2)$ as the joint PGF for the waiting times of a randomly chosen customer in the first queue and in the second queue:

$$W(z_1, z_2) \triangleq E\left[z_1^{W_c^{(1)}} z_2^{W_c^{(2)}}\right]. \tag{5.32}$$

We then have the following theorem.

THEOREM 5.4.3: *The joint PGF $W(z_1, z_2)$ is given by*

$$\begin{aligned} W(z_1, z_2) &= \frac{1 - A(z_2)}{G'(1)B'(1)(1 - C(z_2))(z_1 - A(z_2))}\left[L(0, z_2)\left\{G\left(\frac{z_1}{z_2}\right) - 1\right\}\right. \\ &\quad \left. + (z_2 - 1)X\left(\frac{z_1}{z_2}, z_2\right)\right]. \end{aligned} \tag{5.33}$$

PROOF: It follows that

$$\begin{aligned} W_c(z_1, z_2) &= \sum_{k=1}^{M}\sum_{n=1}^{k}\gamma_{nk}F(k, z_2)D(n, z_2)z_1^{k-n} \\ &= \frac{1 - A(z_2)}{G'(1)B'(1)(1 - C(z_2))(z_1 - A(z_2))} \\ &\quad \cdot\left[\sum_{k=1}^{M} g(k)L(k, z_2)z_1^k - \sum_{k=1}^{M} g(k)L(k, z_2)[A(z_2)]^k\right]. \end{aligned} \tag{5.34}$$

From (5.9), (5.10), (5.13) and (5.34), (5.33) immediately follows. ■

We present a corollary which immediately follows from Theorem 5.4.3. Let $W_c = W_c^{(1)} + W_c^{(2)}$ denote a random variable representing the total waiting time of a randomly chosen customer in the system. We define $W_c(z)$, $W_{c1}(z)$ and $W_{c2}(z)$ as the PGFs for $W_c$, $W_c^{(1)}$ and $W_c^{(2)}$, respectively.

COROLLARY 5.4.1: *The PGFs $W_{c1}(z)$, $W_{c2}(z)$ and $W_c(z)$ are given by*

$$W_{c1}(z) = \frac{1 - G(z)}{G'(1)(1-z)}, \tag{5.35}$$

$$W_{c2}(z) = \frac{1}{G'(1)B'(1)(1 - C(z))}\left[W_s(z) - R_s(z)\right], \tag{5.36}$$

$$W_c(z) = W_b(z)\hat{G}_{idle}(z), \tag{5.37}$$

*respectively, where*

$$W_b(z) = (1 - B'(1)C'(1))\frac{z-1}{z-A(z)} \cdot \frac{1-A(z)}{B'(1)(1-C(z))}, \tag{5.38}$$

*and* $\hat{G}_{idle}(z)$ *is given in (5.24).*

PROOF: Letting $z_2 = 1$, $z_1 = 1$ and $z_1 = z_2 = z$ in (5.34) and using (5.27) and (5.25), we obtain (5.35), (5.36) and (5.37), respectively. ■

*Remark 5.4.3:*

1. $W_b(z)$ denotes the PGF for the waiting times of customers in the corresponding BBP/G/1 queue without gates, and $\hat{G}_{idle}(z)$ denotes the PGF for the backward recurrence time of the gate opening interval given that the server is idle. This is a discrete-time example of the waiting-time decomposition property in the queue with generalized vacations [Fuhr85].

2. After some algebra with (5.28) and (5.36), we have the following relationship between the amount of the work in the second queue and the waiting time of an individual customer in the second queue:

$$U(1, z) = 1 - A'(1) + A'(1)zW_{c2}(z)\frac{1-C(z)}{C'(1)(1-z)}. \tag{5.39}$$

Note that (5.39) can also be derived from the equality of the virtual delay and attained waiting time distribution (see, for example, [Miya92, Saka90, Seng89]).

## 5.5 Number of Individual Customers

In this section, we consider the PGF for the number of individual customers in the system. We first consider the number of individual customers in the system immediately after departures. Let $Q_D(z)$ denote the PGF for the number of customers immediately after the departure of the tagged customer. Those customers are classified into two types. One includes customers who arrive to the system in the same batch as the departing customer. The other includes customers who arrive to the system during the sojourn time of the departing customer. To obtain $Q_D(z)$, we need the following lemma. Let $R_c(z, w)$ denote the joint PGF for the sojourn time of a randomly chosen customer (we call this tagged



customer hereafter) and the number of individual customers who arrive in the same slot as the tagged customer and are served after the tagged customer.

*Lemma 5.5.1:* The joint PGF $R_c(z, w)$ is given by

$$R_c(z, w) = R_1(z)\frac{B(w) - A(z)}{B'(1)(w - C(z))}C(z), \tag{5.40}$$

where

$$R_1(z) = \frac{z-1}{z-A(z)}\frac{X(1, z)}{G'(1)}. \tag{5.41}$$

PROOF: To obtain $R_c(z, w)$, we divide the sojourn time of the tagged customer into three intervals: (1) the waiting time of the batch of customers who arrive to the system in the same slot as the tagged customer, (2) the service times of customers who arrive in the same slot as the tagged customer and are served before the tagged customer, and (3) the service time of the tagged customer, whose PGF is given by $C(z)$. Note here that those three intervals are mutually independent. Also note that the sum of the first two intervals is equivalent to the waiting time of the tagged customer, whose PGF is given in (5.37). It is easy to see from (5.37) that the PGF $R_1(z)$ for the first interval is given by

$$R_1(z) = \frac{z-1}{z-A(z)}\frac{X(1, z)}{G'(1)}, \tag{5.42}$$

since the PGF $R_2(z)$ for the second interval is given by

$$R_2(z) = \frac{1-A(z)}{B'(1)(1-C(z))}. \tag{5.43}$$

Now we are ready to derive $R_c(z, w)$. Note that the length of the second interval and the number of customers who arrive in the same slot as the tagged customer and are served after the tagged customer are dependent. The joint PGF for the second interval and the number of such customers is given by

$$\frac{B(w) - A(z)}{B'(1)(w - C(z))}. \tag{5.44}$$

Therefore we obtain (5.40). ■

Now we have the following theorem.

THEOREM 5.5.1: *The PGF* $Q_D(z)$ *is given by*

$$Q_D(z) = \frac{1-B(z)}{B'(1)(1-z)}Q_b(z)\hat{G}_{idle}(B(z)), \tag{5.45}$$

*where*

$$Q_b(z) = \frac{(1 - A'(1))(z-1)C(B(z))}{z - C(B(z))}, \tag{5.46}$$

*and* $\hat{G}_{idle}(z)$ *is given in (5.24).*

PROOF: Noting $Q_D(z) = R_c(B(z), z)$ and Lemma 5.5.1, (5.45) immediately follows. ∎

Next we consider the PGF $Q(z)$ for the number of customers in the system.

THEOREM 5.5.2: *The PGF* $Q(z)$ *is given by*

$$Q(z) = Q_b(z)\hat{G}_{idle}(B(z)). \tag{5.47}$$

PROOF: First we consider the PGF $Q_A(z)$ for the number of customers immediately before arrivals. Note that $Q_A(z)$ and $Q_D(z)$ are related by [Faki91]

$$\frac{1 - B(z)}{B'(1)(1-z)} Q_A(z) = Q_D(z). \tag{5.48}$$

Therefore we have

$$Q_A(z) = Q_b(z)\hat{G}_{idle}(B(z)). \tag{5.49}$$

Since customers arrive to the system according to the batch Bernoulli process, $Q(z)$ is identical to $Q_A(z)$. Therefore we obtain (5.47). ∎

*Remark 5.5.1:* $Q_b(z)$ denotes the PGF for the number of customers in the corresponding BBP/G/1 queue without gates and $\hat{G}_{idle}(B(z))$ denotes the PGF for the number of individual customers in the system given that the server is idle. This decomposition result is a discrete-time example of the general result for the continuous-time queue given in [Fuhr85].

## 5.6 Numerical Examples

In this section, we provide some numerical examples of the analytical results. Fig. 5.1 shows the mean waiting times of individual customers for various distributions of the gate opening intervals as a function of the mean number of arrivals in a slot in the following settings: (1) the number of arrivals in a slot is geometrically distributed, (2) the service times of individual customers are deterministic and equal to one slot, and (3) the mean gate opening interval is equal to five slots. In the deterministic distribution of the gate opening intervals, we set $g(5) = 1$, in the binomial distribution, we set $g(n) =$



$(8!/\{(n-1)!(9-n)!\})(0.5)^8$ for $n = 1, \ldots, 9$, in the uniform distribution, we set $g(n) = 1/9$ for $n = 1, \ldots, 9$, and in the bimodal distribution, we set $g(1) = g(9) = 0.5$. In Fig. 5.1, the variance $V$ of the gate opening intervals is also shown. We observe that the distribution of the gate opening intervals affects the mean waiting time of individual customers and the mean waiting time increases with the increase of the variance of the gate opening intervals.

## 5.7 Conclusion

In this chapter, we consider discrete-time BBP/G/1 queues with a gate, where the intervals between successive openings of the gate are bounded, and independent and identically distributed. We derive the joint PGFs for the amounts of the stationary work, and the waiting time of an individual customer in the first queue and the second queue. We also derived the PGFs for the sojourn time and the waiting time of a supercustomer, and the number of individual customers in the system. Furthermore, we provide numerical examples and observe the effect of the distribution of the gate opening interval on the mean waiting times of a customer.

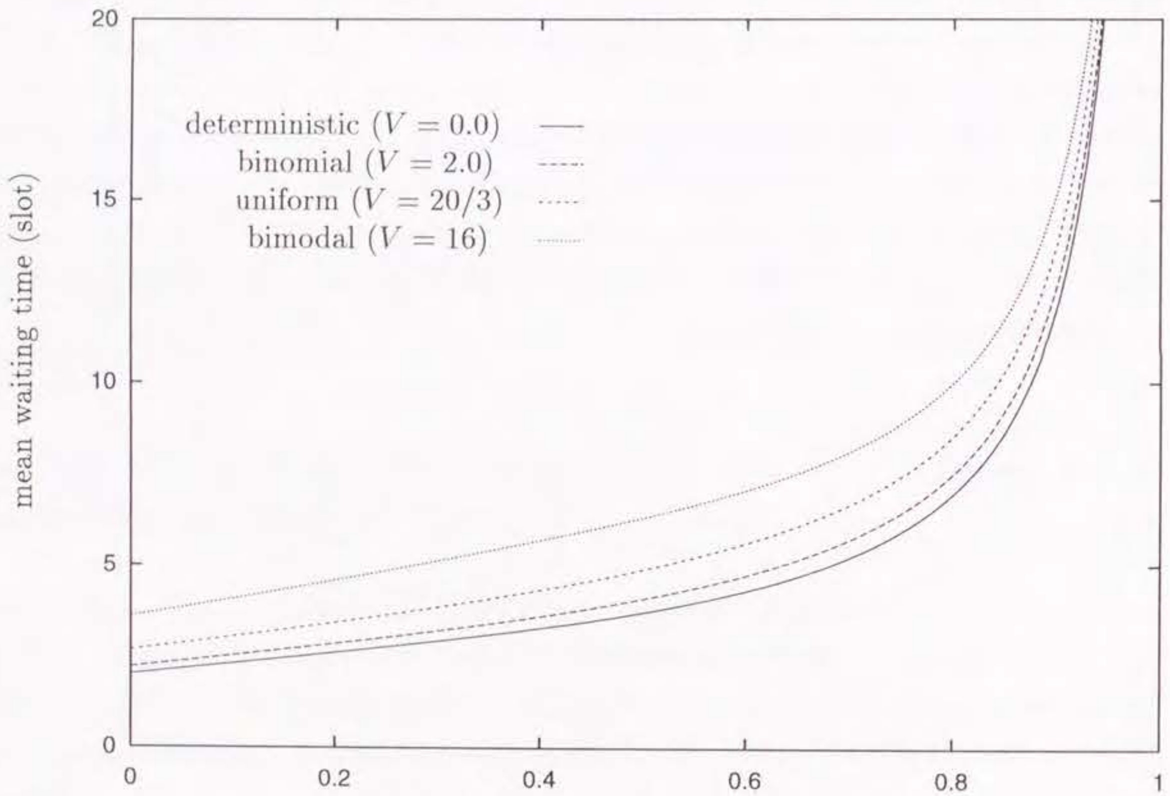


Figure 5.1: Mean Waiting Times



# Chapter 6

# Concluding Remarks

## 6.1 Summary of Results

In this dissertation, we extensively studied discrete-time queues with correlated arrivals. In particular, we considered two types of mechanisms to bring correlations into the arrival and the service processes: Markov modulation and gating. For the former, the summary of the results is described below.

1. We analyzed the generalized SBBP/G/1 queues. In the model, the arrival and the service processes were semi-Markovian in the sense that their distributions depended not only on the state of the alternating renewal process in the current slot but also on the state in the next slot. We derived the PGFs for the amount of work in the system and the waiting time of a customer. We also showed applications of the analytical results to important queueing systems.

2. We considered the loss probability approximations in DBMAP/D/1/K queues. We proposed the approximate formulas which were given in terms of the tail distribution of the queue length in the corresponding infinite-buffer queue. The approximate formulas were constructed in such a way that they became exact for any independent arrival process. We extensively examined the accuracy of the approximations through numerical experiments. We observed that, when the correlation in arrivals was not so strong, the approximations were surprisingly accurate.

For the latter, the summary of the results is described below.

1. We analyzed discrete-time BBP/G/1 queues with a gate, where the intervals between successive openings of the gate were geometrically distributed. We derived the following joint PGFs:

    - joint PGF for the numbers of customers in the first queue and the second queue at the beginning of a randomly chosen slot

- joint PGF for the amounts of work in the first queue and the second queue at the beginning of a randomly chosen slot
- joint PGF for the waiting times of individual customers in the first queue and the second queue

We also derived the PGF for the sojourn time and the waiting time of a super-customer. The effect of the correlations on the performance was discussed through numerical examples.

2. We analyzed discrete-time BBP/G/1 queues with a gate, where the intervals between successive openings of the gate were bounded, and independent and identically distributed. We derived the joint PGFs for the amounts of the stationary work, and the waiting times of an individual customer in the first queue and the second queue. We also obtained the PGFs for the sojourn time and the waiting time of a super-customer, and the number of individual customers in the system. We observed the effect of the distribution of the gate opening interval on the mean waiting times of a customer through numerical examples.

## 6.2 Future Research Topics

In this dissertation, the author considered the two types of specific mechanisms to bring correlation into the arrival and the service processes, i.e., Markov modulation and gating. Also, the studies of queues with correlated arrivals by other approaches should be considered. The author then suggests them as future research topics.

1. Recent measurement studies of the traffic data from networks and services such as ISDN packet networks, Ethernet LANs and Variable Bit Rate (VBR) video sources have indicated how complex the traffic patterns can be. The studies convincingly demonstrate the presence of features such as self similarity and fractal dimensions in the traffic patterns. For modeling the traffic which has those features, fractal processes may be more suitable than conventional stochastic processes. Thus, fractal queueing theory should be developed.

2. Contrary to the framework in this dissertation, more general framework without specifying the mechanisms has been studied. It uses the elements of the more general methods of point processes. For example, Palm-martingale framework tries to give a unified view when only very weak assumptions are made on the input to a queueing system. We need to show further usefulness of the framework.



# Appendix A

# Determination of the Unknown Constants

In order to determine the unknown constants $x_i^*(j)$'s and $x_i^{**}(j)$'s, we focus our attention on (2.19) and (2.20). They contain the functions $G_i(A_{ii}(z)/z)$ and $X_i(A_{ii}(z)/z)$. We shall derive expressions for these functions under the assumption that the $G_i(z)$'s are rational functions.

Using the formulas (2.9), (2.10) and (2.22), we obtain the following expressions for $i = 1, 2$:

$$G_i\left(\frac{A_{ii}(z)}{z}\right) = \frac{\Pi_i(z)P_i(z) + z^{M_i}Q_i(z)}{z^{M_i}\Pi_i(z)}, \qquad X_i\left(\frac{A_{ii}(z)}{z}\right) = \frac{\Pi_i(z)X_i^*(z) + z^{M_i}X_i^{**}(z)}{z^{M_i}\Pi_i(z)}, \tag{A.1}$$

where

$$P_i(z) = \sum_{l=1}^{M_i} m_{il}\{A_{ii}(z)\}^l z^{M_i-l}, \qquad Q_i(z) = \sum_{j=1}^{N_i} n_{ij}\{A_{ii}(z)\}^j z^{N_i-j}, \tag{A.2}$$

$$X_i^*(z) = \sum_{l=1}^{M_i} x_i^*(l)\{A_{ii}(z)\}^l z^{M_i-l}, \qquad X_i^{**}(z) = \sum_{j=1}^{N_i} x_i^{**}(j)\{A_{ii}(z)\}^j z^{N_i-j}, \tag{A.3}$$

$$\Pi_i(z) = \prod_{k=1}^{K_i}\{z - \alpha_{ik}A_{ii}(z)\}^{w_{ik}}. \tag{A.4}$$

Substituting (A.1) into (2.19), we obtain the following expression for $L_1(0, z)$:

$$L_1(0, z) = \frac{\tilde{L}_{N1}(z)}{\tilde{L}_D(z)}, \tag{A.5}$$

where

$$\begin{aligned}\tilde{L}_{N1}(z) &= (z-1)\Big[A_{12}(z)A_{21}(z)\{\Pi_2(z)P_2(z) + z^{M_2}Q_2(z)\}\{\Pi_1(z)X_1^*(z) + z^{M_1}X_1^{**}(z)\} \\ &\quad + z^{M_1}A_{11}(z)A_{21}(z)\Pi_1(z)\{\Pi_2(z)X_2^*(z) + z^{M_2}X_2^{**}(z)\}\Big], \end{aligned} \tag{A.6}$$

$$\begin{aligned}\tilde{L}_D(z) &= z^{M_1+M_2}A_{11}(z)A_{22}(z)\Pi_1(z)\Pi_2(z) \\ &\quad - A_{12}(z)A_{21}(z)\{\Pi_1(z)P_1(z) + z^{M_1}Q_1(z)\}\{\Pi_2(z)P_2(z) + z^{M_2}Q_2(z)\}. \end{aligned} \tag{A.7}$$

These expressions enable us to determine the unknown constants $x_i^*(j)$'s and $x_i^{**}(j)$'s in the functions $X_i^*(z)$ and $X_i^{**}(z)$. Whenever the condition $\rho < 1$ (see (2.7)) for the existence of the stochastic equilibrium is satisfied, the denominator $\tilde{L}_D(z)$ has exactly $M_1 + M_2 + N_1 + N_2$ zeros inside the unit disk of the complex plane, one of which equals unity. This can be shown by Rouché's theorem (See [Ishi93a] and Appendix D.1). Due to the shortage of space, we omit the proof. Since $L_1(0, z)$ is a PGF, it has no poles inside the unit disk. Thus the $M_1 + M_2 + N_1 + N_2$ zeros of the denominator must be zeros of the numerator $\tilde{L}_{N1}(z)$ as well. This condition provides us with $M_1 + M_2 + N_1 + N_2 - 1$ linear equations for the unknowns $x_i^*(j)$'s and $x_i^{**}(j)$'s, (no equation is obtained for the zero $z = 1$). Furthermore we have the following equation from the normalizing equation $L_1(0, 1) = 1$:

$$X_1(1) + X_2(1) = (1 - \rho)(G_1'(1) + G_2'(1)). \tag{A.8}$$

Thus, if the coefficient matrix associated with the $M_1 + M_2 + N_1 + N_2$ equations for the unknowns $x_i^*(j)$'s and $x_i^{**}(j)$'s has rank $M_1 + M_2 + N_1 + N_2$, the unknown constants $x_i^*(j)$'s and $x_i^{**}(j)$'s are determined, and hence the $X_i(z)$'s. It is very hard to show that. However, we claim that we can obtain the unknowns by solving $M_1 + M_2 + N_1 + N_2$ linear equations in most applications (See Appendix D).



# Appendix B

# PROOF OF THEOREM 3.3.1

First we consider the early arrival model. Summing up the both sides of (3.6) and (3.7), we obtain

$$\sum_{k=0}^{N} \boldsymbol{y}_k^{(e)} = \sum_{k=0}^{N} \boldsymbol{y}_k^{(e)} \boldsymbol{U}, \tag{B.1}$$

which implies

$$\sum_{k=0}^{N} \boldsymbol{y}_k^{(e)} = \boldsymbol{\pi}, \tag{B.2}$$

since $\sum_{k=0}^{N} \boldsymbol{y}_k^{(e)} \boldsymbol{e} = 1$. Thus (3.6) and (B.2) completely determine the $\boldsymbol{y}_k^{(e)}$. Suppose $\boldsymbol{y}_k^{(e)} = c^{(e)} \boldsymbol{x}_k$ $(0 \le k \le N)$. Since (3.6) and (3.12) take the same form, $\boldsymbol{y}_k^{(e)} = c^{(e)} \boldsymbol{x}_k$ satisfies (3.6). When the number of customers arriving to the system is i.i.d., $\boldsymbol{x}_k$ is expressed as $\boldsymbol{x}_k = x_k \boldsymbol{\pi}$ with some constant $x_k$. Therefore $\boldsymbol{y}_k^{(e)} = c^{(e)} \boldsymbol{x}_k$ satisfies (B.2), too. As a result, $\boldsymbol{y}_k^{(e)} = c^{(e)} \boldsymbol{x}_k$ $(0 \le k \le N)$ becomes exact when the number of customers arriving to the system is i.i.d..

Next we consider the late arrival model. Summing up the both sides of (3.8), (3.9) and (3.10), we obtain

$$\sum_{k=0}^{N} \boldsymbol{y}_k^{(l)} = \sum_{k=0}^{N} \boldsymbol{y}_k^{(l)} \boldsymbol{U}, \tag{B.3}$$

which again implies

$$\sum_{k=0}^{N} \boldsymbol{y}_k^{(l)} = \boldsymbol{\pi}, \tag{B.4}$$

since $\sum_{k=0}^{N} \boldsymbol{y}_k^{(l)} \boldsymbol{e} = 1$. Thus (3.8), (3.10) and (B.4) completely determine the $\boldsymbol{y}_k^{(l)}$. Suppose $\boldsymbol{y}_k^{(l)} = c^{(l)} \boldsymbol{x}_k$ $(0 \le k \le N-1)$. Since (3.8) and (3.12) take the same form, $\boldsymbol{y}_k^{(l)} = c^{(l)} \boldsymbol{x}_k$ satisfies (3.8). When the number of customers arriving to the system is i.i.d., $\boldsymbol{x}_k$ $(0 \le k \le N)$ is expressed as $\boldsymbol{x}_k = x_k \boldsymbol{\pi}$ with some constant $x_k$. Furthermore, for any independent arrival process, we have $\boldsymbol{B}_N = B_N \boldsymbol{U}$, where $B_N$ is some constant. Therefore we have

$$\sum_{k=0}^{N} \boldsymbol{y}_k^{(l)} = c^{(l)} \left( \sum_{k=0}^{N-1} x_k + x_0 B_N \right) \boldsymbol{\pi} = \boldsymbol{\pi}, \tag{B.5}$$

where we use $\boldsymbol{\pi U} = \boldsymbol{\pi}$. As a result, $\boldsymbol{y}_k^{(l)} = c^{(l)}\boldsymbol{x}_k$ $(0 \le k \le N-1)$ and $\boldsymbol{y}_N^{(l)} = c^{(l)}\boldsymbol{x}_0\boldsymbol{B}_N$ become exact when the number of customers arriving to the system is i.i.d..



# Appendix C

# Proofs in Chapter 4

## C.1 Proof of Theorem 4.3.1

By definition, $Q_D(z_1, z_2)$ satisfies

$$Q_D(z_1, z_2) = \frac{P(z_1, z_2)}{P(1,1)}. \tag{C.1}$$

Thus $Q_D(z_1, z_2)$ is obtained once we have $P(z_1, z_2)$. Using the memoryless property of the gate opening intervals, we have

$$Q(z_2) = G(B(z_2))\{P(z_2, 0) + Q(0)\}. \tag{C.2}$$

Let $Y^{(1)}$ denote a random variable representing the number of individual customers who arrive and remain in the first queue during the service time of a customer whose service starts immediately after the randomly chosen imbedded point. Also, let $Y^{(2)}$ denote a random variable representing the number of individual customers who arrive and move to the second queue during the service time. Note here that $Y^{(1)}$ and $Y^{(2)}$ are dependent. We define a random variable $T_G$ as

$$T_G \triangleq \begin{cases} 1 & \text{if the gate opens at least once during the service time,} \\ 0 & \text{if the gate does not open during the service time.} \end{cases} \tag{C.3}$$

Note here that $Y^{(2)} = 0$ if $T_G = 0$. We define $H(z_1, z_2)$ as

$$H(z_1, z_2) \triangleq E\left[z_1^{Y^{(1)}} z_2^{Y^{(2)}} \mathbf{1}_{\{T_G=1\}}\right]. \tag{C.4}$$

To derive an expression for $H(z_1, z_2)$, we suppose that the service time of a randomly chosen customer is $\tau$ $(1 \le \tau < \infty)$ and the gate last opens in the $k$th $(1 \le k \le \tau)$ slot during the service time of the customer. It follows that

$$\begin{aligned} H(z_1, z_2) &= \sum_{\tau=1}^{\infty} \Pr\{C = \tau\} \sum_{k=1}^{\tau} (1-\gamma)\gamma^{k-1}(B(z_1))^{k-1}(B(z_2))^{\tau-k+1} \\ &= \frac{(1-\gamma)B(z_2)}{B(z_2) - \gamma B(z_1)}\big[C(B(z_2)) - C(\gamma B(z_1))\big]. \end{aligned} \tag{C.5}$$

Furthermore, we define $\theta(z_1, z_2)$ as

$$\begin{aligned}\theta(z_1, z_2) &\triangleq E\left[z_1^{Y^{(1)}} z_2^{Y^{(2)}} \mathbf{1}_{\{T_G=0\}}\right] + E\left[z_1^{Y^{(1)}} z_2^{Y^{(2)}} \mathbf{1}_{\{T_G=1\}}\right] \\ &= C(\gamma B(z_1)) + H(z_1, z_2). \end{aligned} \tag{C.6}$$

To obtain $P(z_1, z_2)$, we now consider three exclusive events:

- The preceding imbedded point is of type 1, there exists at least one customer in the second queue at the preceding imbedded point and the gate does not open during the service time of the customer who is served immediately after the imbedded point, i.e., $\{T_P = 1, X^{(2)} > 0, T_G = 0\}$,
- The preceding imbedded point is of type 1, there exists at least one customer in the second queue at the preceding imbedded point and the gate opens during the service time of the customer who is served immediately after the imbedded point, i.e., $\{T_P = 1, X^{(2)} > 0, T_G = 1\}$,
- The preceding imbedded point is of type 2 and there exists at least one customer in the second queue, i.e., $\{T_P = 2, X^{(2)} > 0\}$.

From the above observation, we obtain

$$\begin{aligned}P(z_1, z_2) &= \big[P(z_1, z_2) - P(z_1, 0)\big]\frac{1}{z_2}C(\gamma B(z_1)) \\ &\quad + \big[P(z_2, z_2) - P(z_2, 0)\big]\frac{1}{z_2}H(z_1, z_2) + \big[Q(z_2) - Q(0)\big]\frac{1}{z_2}\theta(z_1, z_2). \end{aligned} \tag{C.7}$$

Setting $z_1 = z_2$ in (C.7) and using (C.2), $H(z_2, z_2) = C(B(z_2)) - C(\gamma B(z_2))$ and $\theta(z_2, z_2) = C(B(z_2))$, we have

$$P(z_2, z_2) = \frac{C(B(z_2))}{z_2 - C(B(z_2))}\big[G(B(z_2)) - 1\big]\big[P(z_2, 0) + Q(0)\big]. \tag{C.8}$$

Using (4.4), (C.2) and (C.8) in (C.7), we obtain

$$\begin{aligned}\big[z_2 - C(\gamma B(z_1))\big]P(z_1, z_2) &= C(\gamma B(z_1))\big[P(z_2, 0) - P(z_1, 0)\big] \\ &\quad + \left[\frac{z_2}{C(B(z_2))}\theta(z_1, z_2) - C(\gamma B(z_1))\right] \\ &\quad \cdot \frac{C(B(z_2))}{z_2 - C(B(z_2))}\frac{B(z_2) - 1}{1 - \gamma B(z_2)}\big[P(z_2, 0) + Q(0)\big], \end{aligned} \tag{C.9}$$

from which and (C.1), (4.8) immediately follows.



## C.2 Proof of Theorem 4.3.2

Iterating (4.18), we obtain for $|z_1| \le 1$ and $M \ge 0$

$$\Psi(z_1) = \prod_{h=0}^{M} \phi(\delta^{(h)}(z_1))\Psi(\delta^{(M+1)}(z_1)). \tag{C.10}$$

We now need the following lemma.

*Lemma C.2.1:*

1. The equation $\delta(z_1) = z_1$ ($|z_1| \le 1$) has a unique solution $z_1^*$ and $z_1^*$ is real.
2. $\lim_{M\to\infty} \delta^{(M)}(z_1) = z_1^*$ for all $z_1$ with $|z_1| \le 1$.
3. $\prod_{h=0}^{\infty} \phi(\delta^{(h)}(z_1))$ converges for all $z_1$ with $|z_1| \le 1$.

The proof of Lemma C.2.1 is given in Appendix C.3. Letting $M \to \infty$ in (C.10), then Lemma C.2.1 leads to the following expression for $\Psi(z_1)$:

$$\Psi(z_1) = \prod_{h=0}^{\infty} \phi(\delta^{(h)}(z_1))\Psi(z_1^*). \tag{C.11}$$

Thus (C.11) becomes

$$\Psi(z_1) = \alpha(z_1)\Psi(z_1^*). \tag{C.12}$$

Letting $z_1 = 1$ in (C.12), we obtain

$$\Psi(z_1^*) = \frac{1}{\alpha(1)}\Psi(1). \tag{C.13}$$

Substituting (C.13) into (C.12) leads to

$$\Psi(z_1) = \frac{\alpha(z_1)}{\alpha(1)}\Psi(1). \tag{C.14}$$

Also, letting $z_2 = 1$ in (C.8), we have

$$\Psi(1) = \frac{1 - B'(1)C'(1)}{B'(1)}(1 - \gamma). \tag{C.15}$$

(4.19) immediately follows from (C.14) and (C.15).

## C.3 Proof of Lemma C.2.1

Using (4.13), we then find that $\delta(z_1) = z_1$ if and only if $\xi(\gamma B(z_1)) = \gamma B(z_1)$ and that $\xi^{(h)}(\gamma B(z_1)) = \gamma B(\delta^{(h)}(z_1))$ ($|z_1| \le 1, h = 1, 2, ...$), where

$$\xi(z) \triangleq \gamma A(z), \tag{C.16}$$

$$\xi^{(0)}(z) \triangleq z, \tag{C.17}$$

$$\xi^{(h)}(z) \triangleq \xi(\xi^{(h-1)}(z)) \quad (h = 1, 2, ...). \tag{C.18}$$

From the results in [Kawa93], we know that the equation $\xi(w) = w$, $|w| \le 1$ has a unique solution $w^*$, that $0 \le w^* \le 1$ and that $\lim_{M\to\infty} \xi^{(M)}(w) = w^*$ for all $w$ with $|w| \le 1$. As $|\gamma B(z_1)| \le 1$ for all $z_1$ with $|z_1| \le 1$, we conclude that the equation $\delta(z_1) = z_1$, $|z_1| \le 1$ has a unique solution $z_1^* = C(w^*)$, $z_1^*$ is real, and that $\lim_{M\to\infty} \delta^{(M)}(z_1) = z_1^*$ for all $z_1$ with $|z_1| \le 1$.

Using (4.17), we have

$$\prod_{h=0}^{M} \phi(\delta^{(h)}(z_1)) = \frac{1 - \gamma B(z_1)}{1 - \gamma B(\delta^{(M+1)}(z_1))} \prod_{h=0}^{M} \frac{(1-\gamma)B(\delta^{(h+1)}(z_1))}{B(\delta^{(h+1)}(z_1)) - \gamma B(\delta^{(h)}(z_1))}. \tag{C.19}$$

Thus we obtain

$$\prod_{h=0}^{\infty} \phi(\delta^{(h)}(z_1)) = \frac{1 - \gamma B(z_1)}{1 - \gamma B(z_1^*)} \prod_{h=0}^{\infty} \frac{(1-\gamma)B(\delta^{(h+1)}(z_1))}{B(\delta^{(h+1)}(z_1)) - \gamma B(\delta^{(h)}(z_1))}. \tag{C.20}$$

From the theory of infinite products, the infinite product

$$\prod_{h=0}^{\infty} \frac{(1-\gamma)B(\delta^{(h+1)}(z_1))}{B(\delta^{(h+1)}(z_1)) - \gamma B(\delta^{(h)}(z_1))} \tag{C.21}$$

converges if and only if the infinite sum

$$\sum_{h=0}^{\infty}\left[1 - \frac{(1-\gamma)B(\delta^{(h+1)}(z_1))}{B(\delta^{(h+1)}(z_1)) - \gamma B(\delta^{(h)}(z_1))}\right] = \gamma \sum_{h=0}^{\infty} \frac{B(\delta^{(h+1)}(z_1)) - B(\delta^{(h)}(z_1))}{B(\delta^{(h+1)}(z_1)) - \gamma B(\delta^{(h)}(z_1))} \tag{C.22}$$

converges. For some real $\sigma_1$ ($0 \le \sigma_1 \le 1$), we have

$$\left|\delta^{(h+1)}(z_1) - \delta^{(h)}(z_1)\right| = \left|\delta^{(h)}(z_1) - \delta^{(h-1)}(z_1)\right|\left|\delta'(\sigma_1)\right|. \tag{C.23}$$

Since $|\delta'(\sigma_1)| \le |\delta'(1)| = \gamma B'(1)C'(\gamma)$, we obtain

$$\left|\delta^{(h+1)}(z_1) - \delta^{(h)}(z_1)\right| \le \gamma B'(1)C'(\gamma)\left|\delta^{(h)}(z_1) - \delta^{(h-1)}(z_1)\right|. \tag{C.24}$$

Similarly, we have for some real $\sigma_2$ ($0 \le \sigma_2 \le 1$),

$$\left|B(\delta^{(h+1)}(z_1)) - B(\delta^{(h)}(z_1))\right| = \left|\delta^{(h+1)}(z_1) - \delta^{(h)}(z_1)\right|\left|B'(\sigma_2)\right|. \tag{C.25}$$



Since $|B'(\sigma_2)| < 1$, we obtain

$$\left|B(\delta^{(h+1)}(z_1)) - B(\delta^{(h)}(z_1))\right| < \left|\delta^{(h+1)}(z_1) - \delta^{(h)}(z_1)\right|. \tag{C.26}$$

From (C.26), it follows that

$$\left|\frac{B(\delta^{(h+1)}(z_1)) - B(\delta^{(h)}(z_1))}{B(\delta^{(h+1)}(z_1)) - \gamma B(\delta^{(h)}(z_1))}\right| < \left|\frac{\delta^{(h+1)}(z_1) - \delta^{(h)}(z_1)}{B(\delta^{(h+1)}(z_1)) - \gamma B(\delta^{(h)}(z_1))}\right|. \tag{C.27}$$

We define $x_n$ as

$$x_n \triangleq \frac{\delta^{(n)}(z_1) - \delta^{(n-1)}(z_1)}{B(\delta^{(n)}(z_1)) - \gamma B(\delta^{(n-1)}(z_1))}. \tag{C.28}$$

Using (C.24), we have

$$\begin{aligned} \left|\frac{x_{n+1}}{x_n}\right| &= \left|\frac{\delta^{(n+1)}(z_1) - \delta^{(n)}(z_1)}{\delta^{(n)}(z_1) - \delta^{(n-1)}(z_1)} \frac{B(\delta^{(n)}(z_1)) - \gamma B(\delta^{(n-1)}(z_1))}{B(\delta^{(n+1)}(z_1)) - \gamma B(\delta^{(n+1)}(z_1))}\right| \\ &\le \left|\frac{B(\delta^{(n+1)}(z_1)) - \gamma B(\delta^{(n)}(z_1))}{B(\delta^{(n+2)}(z_1)) - \gamma B(\delta^{(n+1)}(z_1))}\right| . \end{aligned} \tag{C.29}$$

Since $\gamma B'(1)C'(\gamma) < 1$ and

$$\lim_{n\to\infty}\left|\frac{B(\delta^{(n+1)}(z_1)) - B(\delta^{(n)}(z_1))}{B(\delta^{(n+2)}(z_1)) - B(\delta^{(n+1)}(z_1))}\right| = 1, \tag{C.30}$$

the infinite sum

$$\sum_{h=0}^{\infty}\left|\frac{\delta^{(h+1)}(z_1) - \delta^{(h)}(z_1)}{B(\delta^{(h+1)}(z_1)) - \gamma B(\delta^{(h)}(z_1))}\right| \tag{C.31}$$

converges, and therefore the infinite product

$$\prod_{h=0}^{\infty} \phi(\delta^{(h)}(z_1)) \tag{C.32}$$

also converges.

## C.4 Proof of Corollary 4.3.1

First we note that

$$\theta(z, z) = C(B(z)). \tag{C.33}$$

Furthermore, from (4.9) and (C.15), we have

$$\frac{P(1,0) + Q(0)}{P(1,1)} = \frac{1 - B'(1)C'(1)}{B'(1)}(1-\gamma). \tag{C.34}$$

Letting $z_1 = z_2 = z$, and using (C.33) and (C.34) in (4.8), we obtain

$$\begin{aligned} Q_D(z,z) &= \frac{(z-1)C(B(z))}{z - C(B(z))}(1 - B'(1)C'(1))\frac{B(z)-1}{B'(1)(z-1)} \\ &\cdot \frac{1-\gamma}{1-\gamma B(z)} \frac{P(z,0) + Q(0)}{P(1,0) + Q(0)}, \end{aligned} \tag{C.35}$$

from which, (4.24) immediately follows.

## C.5 Proof of Lemma 4.3.1

We observe the imbedded point preceding the service time of a randomly chosen customer. We then consider two events:

- The imbedded point is of type 1 and there exists at least one customer in the second queue at the imbedded point, i.e., $\{T_P = 1, X^{(2)} > 0\}$,
- The imbedded point is of type 2 and there exists at least one customer in the second queue at the imbedded point, i.e., $\{T_P = 2, X^{(2)} > 0\}$.

From the above observation and using (4.12), (C.1), (C.2) and (C.8), it follows that

$$\begin{aligned}\hat{Q}(z_1, z_2) &= \big[P(z_1, z_2) - P(z_1, 0) + Q(z_2) - Q(0)\big]/P(1,1) \\ &= \big[P(z_1, z_2) + P(z_2, 0) - P(z_1, 0) \\ &\quad + \{G(B(z_2)) - 1\}\{P(z_2, 0) + Q(0)\}\big]/P(1,1) \\ &= \frac{z_2}{z_2 - C(\gamma B(z_1))}\left[Q_D(z_2, 0) - Q_D(z_1, 0) + \{\frac{\theta(z_1, z_2)}{C(B(z_2))} - 1\}Q_D(z_2, z_2)\right] \\ &\quad + \frac{z_2}{C(B(z_2))}Q_D(z_2, z_2), \end{aligned} \tag{C.36}$$

from which and (C.6), (4.27) immediately follows.

## C.6 Proof of Theorem 4.3.3

Since the server is busy with probability $B'(1)C'(1)$, we have

$$Q_A(z_1, z_2) = B'(1)C'(1)Q_{A|busy}(z_1, z_2) + (1 - B'(1)C'(1))Q_{A|idle}(z_1). \tag{C.37}$$

We relate $Q_{A|busy}(z_1, z_2)$ with $\hat{Q}(z_1, z_2)$. To do so, we define $\tilde{C}$ and $\tilde{G}$ as random variables which represent the backward recurrence time of the service time of an individual customer and that of the gate opening interval, respectively. We then consider two events:

- The server is busy and the gate opened at least once during the backward recurrence time of the current service, i.e., $\{T_S = 1, \tilde{C} > \tilde{G}\}$,
- The server is busy and the gate did not open during the backward recurrence time of the current service, i.e., $\{T_S = 1, \tilde{C} \le \tilde{G}\}$.

Let $\tilde{X}^{(1)}$ denote a random variable representing the number of individual customers who arrive and remain in the first queue during the backward recurrence time of the service time of a customer who is served in a randomly chosen slot. Also, let $\tilde{X}^{(2)}$ denote a random variable representing the number of individual customers who arrive and move to the second queue during the backward recurrence time of the service time. We define $\hat{H}(z_1, z_2)$ as

$$\hat{H}(z_1, z_2) \triangleq E\left[z_1^{\tilde{X}^{(1)}} z_2^{\tilde{X}^{(2)}} \mathbf{1}_{\{\tilde{C}>\tilde{G}\}} \middle| T_S = 1\right]. \tag{C.38}$$

It then follows from an argument similar to the one for (C.5) that

$$\hat{H}(z_1, z_2) = \sum_{\tau=1}^{\infty} \frac{1}{C'(1)} \sum_{n=\tau+1}^{\infty} \Pr\{C = n\} \sum_{k=0}^{\tau-1} (1-\gamma)\gamma^k (B(z_1))^k (B(z_2))^{\tau-k}, \tag{C.39}$$

from which, we obtain (4.34). We then have

$$Q_{A|busy}(z_1, z_2) = \hat{Q}(z_2, z_2)\hat{H}(z_1, z_2) + \hat{Q}(z_1, z_2)\frac{1 - C(\gamma B(z_1))}{C'(1)(1 - \gamma B(z_1))}. \tag{C.40}$$

Finally, using Corollary 4.3.1, Lemma 4.3.2 and (4.30), we obtain

$$\begin{aligned} Q_A(z_1, z_2) &= B'(1)C'(1)Q_{A|busy}(z_1, z_2) + (1 - B'(1)C'(1))Q_{A|idle}(z_1) \\ &= B'(1)C'(1)\left[\hat{H}(z_1, z_2) + \frac{1 - C(\gamma B(z_1))}{C'(1)(1 - \gamma B(z_1))}S(z_1, z_2)\right] \\ &\quad \cdot \frac{z_2}{C(B(z_2))}Q_D(z_2, z_2) \\ &\quad + B'(1)C'(1)\frac{z_2}{z_2 - C(\gamma B(z_1))}\frac{1 - C(\gamma B(z_1))}{C'(1)(1 - \gamma B(z_1))} \\ &\quad \cdot \{Q_D(z_2, 0) - Q_D(z_1, 0)\} + (1 - B'(1)C'(1))\frac{Q_D(z_1, z_1)}{Q_{Db}(z_1)}, \end{aligned} \tag{C.41}$$

from which, (4.32) follows.

## C.7 Proof of Lemma 4.4.1

We define $\hat{H}(z_1, z_2, w)$:

$$\hat{H}(z_1, z_2, w) \triangleq E\left[z_1^{\tilde{X}^{(1)}} z_2^{\tilde{X}^{(2)}} w^{\tilde{C}} \mathbf{1}_{\{\tilde{C}>\tilde{G}\}} \middle| T_S = 1\right]. \tag{C.42}$$

From an argument similar to the one for (C.5), it follows that

$$\hat{H}(z_1, z_2, w) = \sum_{n=0}^{\infty}\sum_{\tau=1}^{n}\sum_{k=0}^{\tau-1} \frac{\Pr\{C = n+1\}}{C'(1)} w^{n-\tau}(1-\gamma)\gamma^k (B(z_1))^k (B(z_2))^{\tau-k}, \tag{C.43}$$

which yields (4.42). By definition and using (4.30), we have

$$Q_{A|busy}(z_1, z_2, w) = \hat{Q}(z_2, z_2)\hat{H}(z_1, z_2, w) + \hat{Q}(z_1, z_2)\frac{w(C(w) - C(\gamma B(z_1)))}{C'(1)(w - \gamma B(z_1))}$$

$$
\begin{aligned}
&= \left[\hat{H}(z_1, z_2, w) + S(z_1, z_2)\frac{w(C(w) - C(\gamma B(z_1)))}{C'(1)(w - \gamma B(z_1))}\right]\hat{Q}(z_2, z_2) \\
&\quad + \frac{z_2}{z_2 - C(\gamma B(z_1))}\frac{w(C(w) - C(\gamma B(z_1)))}{C'(1)(w - \gamma B(z_1))} \\
&\quad \cdot\{Q_D(z_2, 0) - Q_D(z_1, 0)\},
\end{aligned} \tag{C.44}
$$

from which, (4.40) immediately follows.

## C.8 Proof of Theorem 4.4.1

By definition and using (4.23) and Lemma 4.4.1, we have

$$
\begin{aligned}
U(z_1, z_2) &= B'(1)C'(1)\frac{1}{C(z_2)}Q_{A|busy}(C(z_1), C(z_2), z_2) \\
&\quad + (1 - B'(1)C'(1))Q_{A|idle}(C(z_1)) \\
&= B'(1)C'(1)\frac{1}{C(z_2)}\left[\hat{Q}(C(z_2), C(z_2))\hat{H}(C(z_1), C(z_2), z_2)\right. \\
&\quad \left. + \hat{Q}(C(z_1), C(z_2))\frac{z_2(C(z_2) - C(\gamma A(z_1)))}{C'(1)(z_2 - \gamma A(z_1))}\right] \\
&\quad + (1 - B'(1)C'(1))\frac{P(C(z_1), 0) + Q(0)}{P(1, 0) + Q(0)}\frac{1 - \gamma}{1 - \gamma A(z_1)} \\
&= B'(1)C'(1)\frac{1}{C(z_2)}\left[\hat{H}(C(z_1), C(z_2), z_2)\right. \\
&\quad \left. + \frac{z_2(C(z_2) - C(\gamma A(z_1)))}{C'(1)(z_2 - \gamma A(z_1))}S(C(z_1), C(z_2))\right]\hat{Q}((C(z_2), C(z_2)) \\
&\quad + B'(1)\frac{z_2}{z_2 - \gamma A(z_1)}\{Q_D(C(z_2), 0) - Q_D(C(z_1), 0)\} \\
&\quad + (1 - B'(1)C'(1))\frac{Q_D(C(z_1), C(z_1))}{Q_{Db}(C(z_1))} \\
&= B'(1)C'(1)\frac{1}{C(z_2)}\hat{\theta}(C(z_1), C(z_2), z_2)\hat{Q}((C(z_2), C(z_2)) \\
&\quad + B'(1)\frac{z_2}{z_2 - \gamma A(z_1)}\{Q_D(C(z_2), 0) - Q_D(C(z_1), 0)\} \\
&\quad + (1 - B'(1)C'(1))\frac{Q_D(C(z_1), C(z_1))}{Q_{Db}(C(z_1))}.
\end{aligned} \tag{C.45}
$$

Note that

$$
\hat{\theta}(C(z_1), C(z_2), z_2) = \frac{1}{C'(1)}\frac{z_2}{z_2 - A(z_2)}\frac{z_2 - \gamma A(z_2)}{z_2 - \gamma A(z_1)}\{C(z_2) - C(A(z_2))\}. \tag{C.46}
$$

(4.44) immediately follows from the above two equations.



## C.9 Mean Waiting Time of Supercustomers

We consider the mean waiting time of supercustomers $E[W_s]$. From (4.50), we have

$$
\begin{aligned}
E[W_s] &= \left.\frac{d}{dz}W_s(z)\right|_{z=1} \\
&= U'(1) - \frac{1}{1-\gamma}A'(1)
\end{aligned} \tag{C.47}
$$

Using Theorem 4.3.1, Corollary 4.4 and (C.34), we obtain

$$
\begin{aligned}
U'(1) &= A'(1) + \frac{A''(1)}{2(1 - A'(1))} \\
&\quad + C'(1)\sum_{n=0}^{\infty}\frac{\phi'(\delta^{(n)}(1))\delta^{(n)\prime}(1)}{\phi(\delta^{(n)}(1))} + \frac{\gamma}{1-\gamma}A'(1).
\end{aligned} \tag{C.48}
$$

Thus, we have

$$
E[W_s] = \frac{A''(1)}{2(1 - A'(1))} + C'(1)\sum_{n=0}^{\infty}\frac{\phi'(\delta^{(n)}(1))\delta^{(n)\prime}(1)}{\phi(\delta^{(n)}(1))}. \tag{C.49}
$$

## C.10 Mean Waiting Times of Customers

We consider the mean waiting times of individual customers $E[W_c^{(1)}]$, $E[W_c^{(2)}]$ and $E[W_c]$. From (4.60), it follows that

$$
\begin{aligned}
E[W_c^{(1)}] &= \left.\frac{d}{dz}W_{c1}(z)\right|_{z=1} \\
&= \frac{\gamma}{1-\gamma}.
\end{aligned} \tag{C.50}
$$

From (4.61), we obtain

$$
\begin{aligned}
E[W_c^{(2)}] &= \left.\frac{d}{dz}W_{c2}(z)\right|_{z=1} \\
&= \frac{1}{A'(1)}U'(1) - \frac{C''(1)}{2C'(1)} - \frac{1}{1-\gamma},
\end{aligned} \tag{C.51}
$$

where $U'(1)$ is given in (C.48). Moreover, we have

$$
\begin{aligned}
E[W_c] &= E[W_c^{(1)}] + E[W_c^{(2)}] \\
&= \frac{1}{A'(1)}U'(1) - \frac{C''(1)}{2C'(1)} - 1.
\end{aligned} \tag{C.52}
$$

# Appendix D

# Determination of the Unknown Values

In order to determine the unknown values $L(k,0)$ $(k=0,1,\ldots,M-1)$, we focus our attention on (5.12). We rewrite (5.12) as

$$L(0,z)=\frac{(z-1)\sum_{k=0}^{M-1}L(k,0)z^k\sum_{j=k+1}^{M}g(j)[A(z)]^j z^{M-j}}{z^M-\sum_{n=1}^{M}g(n)[A(z)]^n z^{M-n}}. \tag{D.1}$$

This expression enable us to determine $L(0,0)$, $L(1,0)$, ..., $L(M-1,0)$. Whenever the condition $A'(1)<1$ for the existence of the stochastic equilibrium is satisfied, the denominator of (D.1) has exactly $M$ zeros inside the unit disk of the complex plane, one of which equals unity. Furthermore, under some additional condition, those zeros are all distinct. The proof is given in Appendix D.1. Since $L(0,z)$ is a PGF, it has no poles inside the unit disk. Thus $M$ distinct zeros of the denominator must be zeros of the numerator as well. This condition provides us with $M-1$ linear equations with respect to $L(k,0)$ $(k=0,1,\ldots,M-1)$ (no equation is obtained for the zero $z=1$). Furthermore, we have the $M$th equation from the normalizing condition (5.14), which yields

$$\sum_{k=0}^{M-1}p(k)L(k,0)=1-A'(1), \tag{D.2}$$

where $p(k)$ denotes the probability that a randomly chosen slot is the $(k+1)$st slot and is given in (5.18). Note that (D.2) implies that the probability that the server is idle (i.e., $\sum_{k=0}^{M-1}p(k)L(k,0)$) is given by $1-A'(1)$.

In a summary, the unknown constants $L(k,0)$ satisfy

$$\begin{pmatrix} f_0(z_1) & f_1(z_1) & \cdots & f_{M-1}(z_1)\\ f_0(z_2) & f_1(z_2) & \cdots & f_{M-1}(z_2)\\ \vdots & \vdots & \vdots & \vdots\\ f_0(z_{M-1}) & f_1(z_{M-1}) & \cdots & f_{M-1}(z_{M-1})\\ f_0(1) & f_1(1) & \cdots & f_{M-1}(1)\end{pmatrix}\begin{pmatrix} L(0,0)\\ L(1,0)\\ \vdots\\ L(M-2,0)\\ L(M-1,0)\end{pmatrix}=\begin{pmatrix}0\\0\\ \vdots\\ 0\\ c\end{pmatrix}, \tag{D.3}$$



where $z_i$'s $(i=1,\ldots,M-1)$ denote roots of the denominator of (D.1) which are not equal to unity, $c=G'(1)\{1-A_b'(1)-A'(1)\}$ and $f_k(z)$ $(k=0,\ldots,M-1)$ are given by

$$f_k(z)=z^k\sum_{n=k+1}^{M}g(n)A(z)^n z^{M-n}. \tag{D.4}$$

It is very hard to show the uniqueness of the solution $L(k,0)$ of (D.3), which can be accomplished by showing the determinant of the coefficient matrix in (D.3) is not equal to zero (see [Gail92]). However, we can show the linear independency of $f_k(z)$ $(k=0,\ldots,M-1)$, which implies that the Lebesgue measure of the set of zeros of the determinant of the coefficient matrix in (D.3) is equal to zero [Lee94]. The proof of the linear independency is given in Appendix D.2. Thus we claim that we can obtain the $L(k,0)$ by solving (D.3) in most applications.

## D.1 Zeros of the Denominator

In this appendix, we show that the denominator of (D.1) i.e.,

$$S(z)\triangleq z^M-\sum_{n=1}^{M}g(n)[A(z)]^n z^{M-n} \tag{D.5}$$

has exactly $M$ zeros inside the unit disk of the complex plane, one of which equals unity when the inequality $A'(1)<1$ holds. The proof is based on Rouché's theorem: *If $d(z)$ and $h(z)$ are analytic functions of $z$ inside and on a closed contour $C$, and $|h(z)|<|d(z)|$ on $C$, then $d(z)$ and $d(z)-h(z)$ have the same number of zeros inside $C$.*

Let

$$d(z)=z^M, \tag{D.6}$$

$$h(z)=\sum_{n=1}^{M}g(n)(A(z))^n z^{M-n}. \tag{D.7}$$

Then $S(z)=d(z)-h(z)$. Obviously, the functions $d(z)$ and $h(z)$ are analytic in a part of the complex plane. Let us rewrite $h(z)$ as

$$h(z)=z^M G\left(\frac{A(z)}{z}\right), \tag{D.8}$$

by using (5.4). We set $z=e^{j\theta}$ $(0\le\theta<2\pi)$ for $z$ on the unit circle. This yields

$$|d(z)|=1, \tag{D.9}$$

$$|h(z)|=\left|G\left(e^{-j\theta}A\left(e^{j\theta}\right)\right)\right|<1 \quad \text{for all } \theta\neq 0, \tag{D.10}$$

because $G(z)$ and $A(z)$ are PGFs. Hence, $|h(z)| < |d(z)|$ for all $z \neq 1$ on the unit circle. Let us choose the contour $C$ so as to include $z = 1$ as an internal point, which is obviously a zero of $S(z)$. In particular, we choose the contour $C$ as

$$C \triangleq \{z = e^{j\theta}; 0 < \theta < 2\pi\} \cup \lim_{\varepsilon \to 0} C_\varepsilon, \tag{D.11}$$

where

$$C_\varepsilon \triangleq \left\{z = 1 + \varepsilon e^{j\alpha}; -\frac{\pi}{2} < \alpha < \frac{\pi}{2}\right\} \tag{D.12}$$

is a semicircle centered at $z = 1$ with radius $\varepsilon$, outside the unit circle. For $z \in C_\varepsilon$, let $z = 1 + \varepsilon e^{j\alpha}$. It follows that

$$\begin{aligned}
|h(z)|^2 &= \left|\sum_{n=1}^{M} g(n)\{A(1+\varepsilon e^{j\alpha})\}^n (1+\varepsilon e^{j\alpha})^{M-n}\right|^2 \\
&= \left|\sum_{n=1}^{M} g(n)\{(1 + A'(1)\varepsilon e^{j\alpha} + o(\varepsilon)\}^n\right. \\
&\qquad \left.\cdot\{(1 + (M-n)\varepsilon e^{j\alpha} + o(\varepsilon)\}\right|^2 \\
&= \left|1 + \{M + G'(1)(A'(1) - 1)\}\varepsilon e^{j\alpha} + o(\varepsilon)\right|^2 \\
&= 1 + 2\{M + G'(1)(A'(1) - 1)\}\varepsilon \cos\alpha + o(\varepsilon) \qquad (D.13)
\end{aligned}$$

On the other hand, we have

$$|d(z)|^2 = \left|(1+\varepsilon e^{j\alpha})^M\right|^2 = \left|1 + M\varepsilon e^{j\alpha} + o(\varepsilon)\right|^2 = 1 + 2M\varepsilon\cos\alpha + o(\varepsilon) \tag{D.14}$$

Since we assume $A'(1) < 1$, we have $|h(z)|^2 < |d(z)|^2$ (and therefore $|h(z)| < |d(z)|$) on $C_\varepsilon$ for a sufficiently small value of $\varepsilon$, and hence also on the entire contour $C$.

Thus the functions $d(z)$, $h(z)$ and the contour $C$ satisfy the conditions of Rouché's theorem. It follows that $d(z)$ and $d(z) - h(z) = S(z)$ have the same number of zeros inside $C$. Since the function $d(z)$ has $M$ zeros in the unit disk, $S(z)$ also has $M$ zeros in the unit disk, one of which equals unity.

When $h(z)$ has no zeros in the unit disk, $h(z)^{\frac{1}{M}}$ is analytic. Under this assumption, we can show in a similar manner that

$$z - e^{j2\pi i/M} h(z)^{\frac{1}{M}}, \qquad (i = 1, \ldots, M) \tag{D.15}$$

has exactly one zero $z_i$ in the unit disk, which is also the zero of $S(z)$. Furthermore, those $z_i$'s $(i = 1, \ldots, M)$ are all distinct. To prove this, we assume otherwise. Then, without loss of generality, let $u$ be a multiple zero that corresponds to $i = i_1$ and $i = i_2$ in (D.15), where $i_1 \neq i_2$. Since $h(z) \neq 0$ in the unit disk,

$$\frac{e^{j2\pi i_1/M} h(u)^{\frac{1}{M}}}{e^{j2\pi i_2/M} h(u)^{\frac{1}{M}}} = e^{j2\pi(i_1 - i_2)/M} = 1, \tag{D.16}$$

which is a contradiction. Thus, $S(z)$ has exactly $M$ zeros, and when $h(z)$ has no zeros in the unit disk, they are all distinct.



## D.2 Linear independency of the $f_k(z)$

In this appendix, we show the linear independency of the $f_k(z)$ $(k = 0, \ldots, M-1)$, following an approach in [Lee94]. To do so, we assume that $f_k(z)$ are linearly dependent, i.e., there exist complex numbers $\beta_0, \beta_1, \ldots, \beta_{M-1}$, not all

$$\sum_{k=0}^{M-1} \beta_k f_k(z) = 0, \tag{D.17}$$

for all $|z| \leq 1$. We now rewrite (D.4) to be

$$f_k(z) = \sum_{i=k}^{M-1} z^i J_{k,i}(z), \tag{D.18}$$

where

$$J_{k,i}(z) = g(M + k - i)A(z)^{M+k-i}. \tag{D.19}$$

Substituting (D.18) into (D.17) and rearranging terms yield

$$\sum_{k=0}^{M-1} \beta_k f_k(z) = \sum_{i=0}^{M-1} z^i \sum_{k=0}^{i} \beta_k J_{k,i}(z) = 0. \tag{D.20}$$

First we show $\beta_0 = 0$. Substituting $z = 0$ into (D.20) yields

$$\sum_{k=0}^{M-1} \beta_k f_k(0) = \beta_0 J_{0,0}(0) = \beta_0 g(M) A(0)^M = 0. \tag{D.21}$$

Note here that $g(M) > 0$ (by the assumption; see section 5.2). Also, $A(0) > 0$, since $\rho < 1$. These facts imply that $\beta_0 = 0$.

Next we consider $\beta_1$. Substituting $\beta_0 = 0$ into (D.20), dividing the resulting equation by $z$ and letting $z$ go to zero, we have

$$\lim_{z \to 0} \frac{1}{z} \sum_{k=1}^{M-1} \beta_k f_k(z) = \beta_1 J_{1,1}(0) = \beta_1 g(M) A_b(0)^{M-1} A(0)^M = 0. \tag{D.22}$$

Thus we have $\beta_1 = 0$. Substituting $\beta_1 = 0$ and repeating the same argument recursively, we conclude that $\beta_k = 0$ for $k = 0, \ldots, M-1$. This contradicts our assumption. Hence $f_k(z)$ $(k = 0, \ldots, M-1)$ are linearly independent.

# References


[Abat94] J.Abate, G.L.Choudhury and W.Whitt, "Asymptotics for steady-state tail probabilities in structured Markov chains of M/G/1-type," *Communications in Statistics - Stochastic Models*, **10** (1994) 99–144.

[Ackr80] M.H.Ackroyd, "Computing waiting time distribution for the G/G/1 queue by signal processing methods," *IEEE Transactions on Communications*, **28** (1980) 52-58.

[Blon92] C.Blondia and O.Casals, "Statistical multiplexing of VBR sources: a matrix-analytic approach," *Performance Evaluation*, **16** (1992) 5–20.

[Bors92] S.C.Borst, O.J.Boxma and M.B.Combé, "Collection of customers: a correlated M/G/1 queue," *Performance Evaluation Review*, **20** (1992) 47–59.

[Bors93] S.C.Borst, O.J.Boxma, and M.B.Combé, "An M/G/1 queue with customer collection, " *Communications in Statistics - Stochastic Models*, **9** (1993) 341–371.

[Boxm88] O.J. Boxma and W.P. Groenendijk, "Waiting times in discrete-time cyclic-service systems," *IEEE Transactions on Communications*, **36** (1988) 164–170.

[Boxm89] O.J. Boxma, "Workloads and waiting times in single-server systems with multiple customer classes," *Queueing Systems*, **5** (1989) 185–214.

[Boxm93] O.J. Boxma and M.B. Combé, "The correlated M/G/1 queue," *Archiv für Elektronik und Übertragungstechnik*, **47** (1993) 330-335.

[Brun84] H.Bruneel, "Analysis of discrete-time buffer with one single output channel, subject to a general interruption process," *Performance '84*, (ed. E.Gelenbe) North-Holland, Amsterdam, (1984), 103-115.

[Brun93] H.Bruneel and B.G.Kim, *Discrete-time models for communications systems including ATM*, Kluwer Academic Publishers, Boston, (1993).





[Burk75] P.J.Burke, "Delays in single-server queues with batch inputs," *Operations Research*, **26** (1975) 830–833.

[Dosh90] B.T. Doshi, "Single server queues with vacations," *Stochastic Analysis of Computer and Communication Systems* (ed. H. Takagi), North-Holland, Amsterdam (1990) 217–266.

[Falk94] E.Falkenberg, "On the asymptotic behavior of the stationary distribution of Markov chains of M/G/1-type," *Communications in Statistics - Stochastic Models*, **10** (1994) 75–98.

[Faki91] D.Fakinos, "The relationship between limiting queue size distributions at arrival and departure epochs in a bulk queue," *Stochastic Processes and Their Applications*, **37** (1991) 327–329.

[Fuhr85] S.W.Fuhrmann and R.B.Cooper, "Stochastic decompositions in the M/G/1 queue with generalized vacations," *Operations Research*, **33** (1985) 1117–1129.

[Gail92] H.R.Gail, S.L.Hantler and B.A. Taylor, "Spectral analysis of M/G/1 type Markov chain," IBM Research Report RC 17765 (#78182) (1992).

[Gouw94] F.N.Gouweleeuw, "The loss probability in finite-buffer queues with batch arrivals and complete rejection," *Probability in the Engineering and Informational Sciences*, **8** (1994) 221–227.

[Hash91] O.Hashida, Y.Takahashi and S.Shimogawa, "Switched batch Bernoulli process (SBBP) and the discrete-time SBBP/G/1 queue with application to statistical multiplexer performance," *IEEE Journal on Selected Areas in Communications*, **9** (1991) 394-401.

[Hook72] J.A.Hooke, "A priority queue with low-priority arrivals general," *Operations Research*, **20** (1972) 373-380.

[Ishi93a] F.Ishizaki, T.Takine, Y.Takahashi and T.Hasegawa, "Generalized SBBP/G/1 queue and its application to ATM networks," *Modelling and Performance Evaluation of ATM Technology*, (eds. H.Perros, G.Pujolle and Y.Takahashi), North-Holland, Amsterdam, (1993) 397–416.

[Ishi93b] F.Ishizaki, T.Takine and T.Hasegawa, "Analysis of a discrete-time queue with a gate, " Technical Report #93011, Department of Applied Mathematics and Physics, Kyoto University (1993).

[Ishi94a] F.Ishizaki, T.Takine and T.Hasegawa, "Analysis of a discrete-time queue with a gate," *The Fundamental Role of Teletraffic in the Evolution of Telecommunications Networks* (eds. J.Labetoulle and J.W.Roberts), Elsevier, Amsterdam, (1994) 169–178.

[Ishi94b] F.Ishizaki, T.Takine, H.Terada and T.Hasegawa, "Loss probability approximation of a statistical multiplexer and its application to call admission control in high-speed networks," Technical Report ISE-TR-94-009, Department of Information Systems Engineering, Faculty of Engineering, Osaka University (1994).

[Ishi94c] F.Ishizaki, T.Takine, Y.Takahashi and T.Hasegawa, "A generalized SBBP/G/1 queue and its applications," *Performance Evaluation*, **21** (1994) 163–181.

[Ishi95a] F.Ishizaki, T.Takine and T.Hasegawa, "Analysis of a discrete-time queue with gated priority," *Performance Evaluation*, **23** (1995) 121–143.

[Ishi95b] F.Ishizaki, T.Takine, H.Terada and T.Hasegawa, "Loss probability approximation of a statistical multiplexer and its application to call admission control in high-speed networks," *Proceedings of IEEE GLOBECOM '95*, Singapore, (1995) 417–421.

[Ishi96] F.Ishizaki, T.Takine and T.Hasegawa, "Analysis of a discrete-time queue with geometrically distributed gate opening intervals," to appear in *Journal of the Operations Research Society of Japan*, **39** (1996).

[Kawa93] T.Kawata, "Analysis of a discrete-time queue with dependence between inter-arrival and service times, " Bachelor thesis, Department of Applied Mathematics and Physics, Faculty of Engineering, Kyoto University (1993) (in Japanese).

[Koba77] H.Kobayashi and A.G.Konheim, "Queueing models for computer communications system analysis," *IEEE Transactions on Communications*, **25** (1977) 2–29.

[LeBo91] J.-Y.Le Boudec, "An efficient solution model for Markov models of ATM links with loss probabilities," *IEEE Journal on Selected Areas in Communications*, **9** (1991) 408–417.

[Lee94] D.S. Lee, "A two-queue model with exhaustive and limited service disciplines," submitted for publication (1994).

[Liao89] K.-Q.Liao and L.G.Mason, "A discrete-time single server queue with a two-level modulated input and its applications," *Proceedings of IEEE GLOBECOM '89*, Dallas, Texas, (1989) 913–918.



[Meis58] T.Meisling, "Discrete-time queueing theory," *Operations Research*, **6** (1958) 96–105.

[Miya83] M.Miyazawa, "The derivation of invariance relations in complex queueing systems with stationary inputs," *Advances in Applied Probability*, **15** (1983) 874–885.

[Miya92] M.Miyazawa and Y.Takahashi, "Rate conservation principle for discrete-time queues," *Queueing Systems*, **12** (1992) 215–230.

[Miya93] M.Miyazawa and H.Tijms, "Comparison of two approximations for the loss probability in finite-buffer queues," *Probability in the Engineering and Informational Sciences*, **7** (1993) 19–27.

[Mura90] M.Murata, Y.Oie, T.Suda and H.Miyahara, "Analysis of a discrete-time single-server queue with bursty inputs for traffic control in ATM networks," *IEEE Journal on Selected Areas in Communications*, **8** (1990) 447–458.

[Mura91] M.Murata and H.Miyahara, "An analytic solution of the waiting time distribution for the discrete-time GI/G/1 queue," *Performance Evaluation*, **13** (1991) 87-96.

[Neut89] M.F.Neuts, *Structured stochastic matrices of* M/G/1 *type and their applications*. Marcel Dekker, New York, (1989).

[Ott84] T.J.Ott, "On the M/G/1 queue with additional inputs," *Journal of Applied Probability*, **21** (1984) 129-142.

[Ott87] T.J.Ott, "The single-server queue with independent GI/G and M/G input streams," *Advances in Applied Probability*, **19** (1987) 266-286.

[Rubi92] I.Rubin and Z.Zhang, "Message delay analysis for TDMA schemes using contiguous-slot assignments," *IEEE Transactions on Communications*, **40** (1992) 730-737.

[Saka90] H.Sakasegawa and R.W.Wolff, "The equality of the virtual delay and attained waiting time distributions," *Advances in Applied Probability*, **22** (1990) 257–259.

[Saka93] H.Sakasegawa, M.Miyazawa and G.Yamazaki, "Evaluating the overflow probability using the infinite queue," *Management Science*, **39** (1993) 1238–1245.

[Seng89] B.Sengupta, "An invariance relationship for the G/G/1 queue," *Advances in Applied Probability*, **21** (1989) 956–957.

[Seng90] B.Sengupta, "A queue with service interruptions in an alternating random environment," *Operations Research*, **38** (1990) 308-318.

[Taka93] H.Takagi, *Queueing analysis, vol.3*, North-Holland, Amsterdam, (1993).

[Taka71] Y.Takahashi, "Queueing systems with gates," *Journal of the Operations Research Society of Japan*, **14** (1971) 103–126.

[Taki94] T.Takine, B.Sengupta and T.Hasegawa, "An analysis of a discrete-time queue for broadband ISDN with priority among traffic classes," *IEEE Transactions on Communications*, **42** (1994) 1837–1845.

[Taki95] T.Takine, T.Suda and T.Hasegawa, "Cell loss and output process analyses of a finite-buffer discrete-time ATM queueing system with correlated arrivals," *IEEE Transactions on Communications*, **43** (1995) 1022–1037.

[Tijm92] H.Tijms, "Heuristics for finite-buffer queues," *Probability in the Engineering and Informational Sciences*, **6** (1992) 227–285.

[Tows80] D.Towsley, "The analysis of a statistical multiplexer with nonindependent arrivals and errors," *IEEE Transactions on Communications*, **28** (1980) 65-72.