Journal of the Operations Research
Society of Japan
Vol. 39, No. 2, June 1996

A NONPREEMPTIVE PRIORITY MAP/G/1 QUEUE WITH TWO CLASSES OF CUSTOMERS

Tetsuya Takine
Osaka University

(Received August 8, 1994)

Abstract This paper considers a nonpreemptive priority queue with two classes of customers. Customers in each priority class arrive to the system according to a Markovian arrival process (MAP). Since the MAP is weakly dense in the class of stationary point processes, it is a fairly general arrival process. The service times of customers in each priority class are independent and identically distributed according to a general distribution function which may differ among two priority classes. Using both the generating function technique and the matrix analytic method, we derive various formulas for the queue length and waiting time distributions. We also discuss the algorithmic implementation of the analytical results along with numerical examples.

1. Introduction

In many situations, customers impose different requirements on the system. In order to provide a solution which satisfies the particular requirement for each customer, a priority mechanism must be employed. Thus, the priority queue is one of the fundamental models in queueing theory and there exist a number of papers which study priority queues, e.g., [10, 18, 24, 25]. Recently, broadband integrated services digital network (BISDN) has emerged as an important field in which the priority mechanism is expected to be employed. Namely, BISDN is expected to provide services to such diverse traffic as video, voice and data. These traffic types have very different needs as far as quality of service is concerned. One of the promising ways to provide services is to implement a priority mechanism among traffic classes [27].

Most of the existing works on priority queues have assumed that the arrival process of customers in each priority class follows a Poisson process. The Poisson process, however, may not be suitable to describe bursty traffic such as video and voice, where there exists a fair amount of correlation and variation [6, 9]. This paper studies a priority queue under the assumption that customers in each priority class arrive to the system according to a Markovian arrival process (MAP) which was introduced in [14]. The MAP includes as special cases the Markov modulated Poisson process ($MMPP$) and the superposition of phase-type renewal processes. Recently, Asmussen and Koole have shown that the MAP is weakly dense in the class of stationary point processes [2]. Therefore, the MAP is a fairly general process and has a capability of representing the correlation inherent in bursty traffic such as video and voice.

More specifically, we consider a nonpreemptive priority $MAP/G/1$ queue with two priority classes of customers. The system consists of a single-server and a buffer of infinite capacity to accommodate arriving customers from both priority classes. The service times of customers in each priority class are independent and identically distributed (i.i.d.) according to a general distribution function (DF) which may differ from one another. Customers are served under the nonpreemptive priority discipline [25]. The nonpreemptive priority discipline is also called the head-of-the-line priority [10].

Let us review the related works. Machihara [17] and Sugahara et al. [23] have studied priority queues with two priority classes assuming correlated arrivals of high priority customers and Poisson arrivals of low priority customers. Takine et al. [28] have studied a nonpreemptive priority $MAP/G/1$ queue with many priority classes, where a common service time distribution among different priority classes is assumed. Also Takine and Hasegawa have studied the workload process in the $MAP/G/1$ queue with state-dependent service times and the results have been applied to characterize the waiting time distribution in the preemptive resume priority $MAP/G/1$ queue [26]. Discrete-time priority queues with correlated arrivals have also been studied in [8, 11, 22, 27]. In all of those four papers, the service times of customers in all classes are assumed to be constant and equal to the slot size. As a result, the performance of high priority customers can be evaluated without considering lower priority customers.

Queueing systems with MAP arrivals have been analyzed through the matrix analytic method developed by Neuts [19]. Readers are also referred to [20, 15, 16]. Note that the matrix analytic method allows no more than one random variable to have the countably infinite space for the description of system dynamics. However, the model considered in this paper requires two mutually dependent random variables, each of which is defined in the countably infinite space (i.e., the number of customers in each priority class). Thus, a straightforward application of the matrix analytic method does not help solve the model. In this paper, using both the generating function technique and the matrix analytic method, we derive various formulas for the queue length and waiting time distributions in each priority class. We also discusses the algorithmic implementation of the analytical results and provide our experience in computing various quantities of interest.

The remainder of this paper is organized as follows. In section 1, the queueing model and some preliminary results are described. In sections 2 and 3, we analyze the queue length and waiting time distributions of high priority customers and of low priority customers, respectively. In section 4, the algorithmic implementation of the analytical results is discussed, along with numerical examples. Finally, in section 5, we provide concluding remarks.

2. MATHEMATICAL MODEL AND PRELIMINARY RESULTS

There are two priority classes of customers. High priority customers arrive to the system according to a MAP (Markovian Arrival Process) with representation $(\widetilde{\boldsymbol{C}}_H, \widetilde{\boldsymbol{D}}_H)$, where $\widetilde{\boldsymbol{C}}_H$ and $\widetilde{\boldsymbol{D}}_H$ are $M_H \times M_H$ matrices. M_H denotes the number of states in the underlying Markov chain which governs high priority arrivals. Also, low priority customers arrive to the system according to a MAP with representation $(\widetilde{\boldsymbol{C}}_L, \widetilde{\boldsymbol{D}}_L)$, where $\widetilde{\boldsymbol{C}}_L$ and $\widetilde{\boldsymbol{D}}_L$ are $M_L \times M_L$ matrices. As for the details of the MAP, readers are referred to section 2.1 of [14]. In Appendix, we reproduce the definition of the MAP from [14]. Let λ_H (resp. λ_L) denote the mean arrival rate of high (resp. low) priority customers. We then have

$$\lambda_H = \tilde{\boldsymbol{\pi}}_H \widetilde{\boldsymbol{D}}_H \boldsymbol{e}, \qquad \lambda_L = \tilde{\boldsymbol{\pi}}_L \widetilde{\boldsymbol{D}}_L \boldsymbol{e}, \tag{2.1}$$

where $\boldsymbol{e}$ denotes a column vector with all elements equal to one and $\tilde{\boldsymbol{\pi}}_H$ (resp. $\tilde{\boldsymbol{\pi}}_L$) denotes a $1 \times M_H$ (resp. $1 \times M_L$) vector which satisfies

$$\begin{aligned} \tilde{\boldsymbol{\pi}}_H(\widetilde{\boldsymbol{C}}_H + \widetilde{\boldsymbol{D}}_H) &= \boldsymbol{0}, & \tilde{\boldsymbol{\pi}}_H \boldsymbol{e} &= 1, \\ \tilde{\boldsymbol{\pi}}_L(\widetilde{\boldsymbol{C}}_L + \widetilde{\boldsymbol{D}}_L) &= \boldsymbol{0}, & \tilde{\boldsymbol{\pi}}_L \boldsymbol{e} &= 1. \end{aligned}$$

We assume that all customers arriving to the system are accommodated in the buffer of infinite capacity.

There is a single server who serves customers according to the nonpreemptive priority discipline. Thus once the service for a customer starts, it will proceed to completion. The service times of high (resp. low) priority customers are independent and identically distributed according to a DF $H_H(x)$ with mean h_H (resp. $H_L(x)$ with mean h_L). Let ρ_H (resp. ρ_L) denote the utilization factor of high (resp. low) priority customers, i.e., $\rho_H = \lambda_H h_H$ and $\rho_L = \lambda_L h_L$. Furthermore, let λ denote the overall arrival rate $\lambda_H + \lambda_L$ and ρ denote the overall utilization factor $\rho_H + \rho_L$. We assume that all customers arriving to the system are eventually served, i.e., $\rho < 1$. The service times and the arrival processes are assumed to be mutually independent.

Before proceeding to the analysis, we consider the superposed arrival process of two independent *MAP*s with representations $(\widetilde{\boldsymbol{C}}_H, \widetilde{\boldsymbol{D}}_H)$ and $(\widetilde{\boldsymbol{C}}_L, \widetilde{\boldsymbol{D}}_L)$ [28]. Let M be $M_H M_L$. In order to distinguish high priority arrivals from low priority arrivals, we introduce the following $M \times M$ matrices:

$$\boldsymbol{D}_H = \widetilde{\boldsymbol{D}}_H \otimes \boldsymbol{I}_L, \quad \boldsymbol{D}_L = \boldsymbol{I}_H \otimes \widetilde{\boldsymbol{D}}_L,$$

$$\boldsymbol{C} = \widetilde{\boldsymbol{C}}_H \oplus \widetilde{\boldsymbol{C}}_L, \quad \boldsymbol{D} = \boldsymbol{D}_H + \boldsymbol{D}_L.$$

where $\otimes$ (resp. $\oplus$) denotes the Kronecker product (resp. the Kronecker sum) [5], and $\boldsymbol{I}_H$ (resp. $\boldsymbol{I}_L$) denotes the identity matrix of the same order as $\widetilde{\boldsymbol{D}}_H$ (resp. $\widetilde{\boldsymbol{D}}_L$).

We denote by $N_{H,t}$ (resp. $N_{L,t}$) the number of high (resp. low) priority arrivals in $(0, t]$ and by S_t the state of the underlying Markov chain of the superposed arrival process at time t. Let $\boldsymbol{N}(n_1, n_2, t)$ denote an $M \times M$ matrix whose (i, j)th element represents $\Pr\{N_{H,t} = n_1, N_{L,t} = n_2, S_t = j \mid S_0 = i\}$. The matrices $\boldsymbol{N}(n_1, n_2, t)$ $(n_1, n_2 \geq 0,\ t \geq 0)$ satisfy the forward Chapman-Kolmogorov equations:

$$\frac{d}{dt}\boldsymbol{N}(n_1, n_2, t) = \boldsymbol{N}(n_1, n_2, t)\boldsymbol{C} + \boldsymbol{N}(n_1 - 1, n_2, t)\boldsymbol{D}_H + \boldsymbol{N}(n_1, n_2 - 1, t)\boldsymbol{D}_L,$$

where $\boldsymbol{N}(0, 0, 0) = \boldsymbol{I}$, $\boldsymbol{N}(-1, n_2, t) = \boldsymbol{0}$ and $\boldsymbol{N}(n_1, -1, t) = \boldsymbol{0}$. Let $\boldsymbol{N}^*(z, \omega, t)$ denote the matrix generating function (GF) of $\boldsymbol{N}(n_1, n_2, t)$. For $|z| \leq 1$, $|\omega| \leq 1$ and $t \geq 0$, we have

$$\boldsymbol{N}^*(z, \omega, t) = \sum_{n_1=0}^{\infty} \sum_{n_2=0}^{\infty} \boldsymbol{N}(n_1, n_2, t) z^{n_1} \omega^{n_2} = e^{\left(\boldsymbol{C} + z\boldsymbol{D}_H + \omega\boldsymbol{D}_L\right)t}.$$

The mean arrival rate λ_H (resp. λ_L) of high (resp. low) priority customers given in (2.1) is now expressed as

$$\lambda_H = \boldsymbol{\pi}\boldsymbol{D}_H\boldsymbol{e}, \qquad \lambda_L = \boldsymbol{\pi}\boldsymbol{D}_L\boldsymbol{e},$$

where $\boldsymbol{\pi}$ denotes a $1 \times M$ vector whose jth element represents the stationary probability of the underlying Markov chain of the superposed arrival process being in state j. Note that $\boldsymbol{\pi}$ is given by

$$\boldsymbol{\pi} = \widetilde{\boldsymbol{\pi}}_H \otimes \widetilde{\boldsymbol{\pi}}_L,$$

and it satisfies

$$\boldsymbol{\pi}\left(\boldsymbol{C} + \boldsymbol{D}\right) = 0, \qquad \boldsymbol{\pi}\boldsymbol{e} = 1.$$

In what follows, we shall provide some preliminary results given in [26], which have analyzed the workload process in the *MAP*/*G*/1 queue with state-dependent service times. As explained in [26], the arrival process considered in this paper is characterized by a *MAP* with state-dependent service times. Therefore, in the rest of this section, we summarize results in [26], which will be used later in this paper.

We define $\boldsymbol{Q}$ as an $M \times M$ matrix which represents the infinitesimal generator of an underlying Markov chain obtained by excising busy periods. Namely, if we observe the system only when the server is idle, the dynamics of the observed underlying Markov chain is governed by the generator $\boldsymbol{Q}$. The matrix $\boldsymbol{Q}$ satisfies

$$\boldsymbol{Q} = \boldsymbol{C} + \boldsymbol{D}_H \int_0^\infty e^{\boldsymbol{Q}x} dH_H(x) + \boldsymbol{D}_L \int_0^\infty e^{\boldsymbol{Q}x} dH_L(x). \tag{2.2}$$

Note here that $e^{\boldsymbol{Q}x}$ denotes the transition probability matrix of the underlying Markov chain during the first passage time to the idle state of the server given that the first passage time starts with the amount x of the initial work. Let $\boldsymbol{\kappa}$ denote a $1 \times M$ vector which satisfies

$$\boldsymbol{\kappa Q} = \boldsymbol{0}, \qquad \boldsymbol{\kappa e} = 1.$$

Furthermore, let $\boldsymbol{V}(s)$ denote a $1 \times M$ vector whose jth element represents the Laplace-Stieltjes transform (LST) for the amount of work in the system (i.e., the sum of the unfinished work of both priority classes) when the underlying Markov chain is in state j. We then have

$$\boldsymbol{V}(s) = (1-\rho)s\boldsymbol{\kappa}\left[s\boldsymbol{I} + \boldsymbol{C} + H_H^*(s)\boldsymbol{D}_H + H_L^*(s)\boldsymbol{D}_L\right]^{-1}, \qquad Re(s) > 0, \tag{2.3}$$

where $H_H^*(s)$ (resp. $H_L^*(s)$) denotes the LST for service times of high (resp. low) priority customers. The recursive formula for the derivatives of $\boldsymbol{V}(s)$ evaluated at $s = 0+$ is found in [26]. In the analysis presented below, we assume the system is in equilibrium.

Remark 2.1. *$\boldsymbol{\kappa}$ is the stationary probability vector of the underlying Markov chain given that the server is idle [26].*

3. ANALYSIS OF HIGH PRIORITY CLASS

In this section, we consider various quantities of interest with respect to high priority customers. In section 3.1, we study the distribution of the number of high priority customers in the system immediately after departures of customers in any priority class. In section 3.2, we study the number of high priority customers at a random point in time. Finally, in section 3.3, we study the waiting time distribution of high priority customers.

3.1. Number of High Priority Customers at Departures

In this subsection, we first consider the joint distribution of the numbers of high and low priority customers immediately after departures of customers in any priority class. We choose the time instants immediately after departures as imbedded points. Let N_H (resp. N_L) denote the number of high (resp. low) priority customers at imbedded points and S denote the state of the underlying Markov chain. We define $\boldsymbol{P}^*(z,\omega)$ ($|z| \leq 1$, $|\omega| \leq 1$) as a $1 \times M$ vector whose jth element $P_j^*(z,\omega)$ is given by $P_j^*(z,\omega) = E\left[z^{N_H}\omega^{N_L}I_{\{S=j\}}\right]$, where $I_{\{\chi\}}$ denotes the indicator function of the event χ. We then have

$$\begin{aligned}\boldsymbol{P}^*(z,\omega) = {} & \left[\boldsymbol{P}^*(z,\omega) - \boldsymbol{P}^*(0,\omega)\right]\boldsymbol{A}^*(z,\omega)/z + \left[\boldsymbol{P}^*(0,\omega) - \boldsymbol{P}^*(0,0)\right]\boldsymbol{B}^*(z,\omega)/\omega \\ & + \boldsymbol{P}^*(0,0)(-\boldsymbol{C})^{-1}\left[\boldsymbol{D}_H\boldsymbol{A}^*(z,\omega) + \boldsymbol{D}_L\boldsymbol{B}^*(z,\omega)\right],\end{aligned} \tag{3.1}$$

where, for $|z| \leq 1$ and $|\omega| \leq 1$,

$$\boldsymbol{A}^*(z,\omega) = \int_0^\infty e^{(\boldsymbol{C}+z\boldsymbol{D}_H+\omega\boldsymbol{D}_L)x} dH_H(x), \tag{3.2}$$

$$\boldsymbol{B}^*(z,\omega) = \int_0^\infty e^{(\boldsymbol{C}+z\boldsymbol{D}_H+\omega\boldsymbol{D}_L)x} dH_L(x). \tag{3.3}$$

Note that the (i,j)th element of matrix $\boldsymbol{A}^*(z,\omega)$ (resp. $\boldsymbol{B}^*(z,\omega)$) represents the double GF for the numbers of high and low priority arrivals during a service time of a high (resp. low) priority customer when the underlying Markov chain is in state j at the end of the service, given that the underlying Markov chain is in state i at the beginning of the service. (3.1) is interpreted as follows. The factor $[\boldsymbol{P}^*(z,\omega) - \boldsymbol{P}^*(0,\omega)]$ represents the vector GF of the joint distribution of the numbers of high and low priority customers when a high priority service starts at an imbedded point, while the factor $[\boldsymbol{P}^*(0,\omega) - \boldsymbol{P}^*(0,0)]$ represents the vector GF for the number of low priority customers when a low priority service starts at an imbedded point. Note that in the latter case, there are no high priority customers in the system. The factor $\boldsymbol{P}^*(0,0)(-\boldsymbol{C})^{-1}\boldsymbol{D}_H$ (resp. $\boldsymbol{P}^*(0,0)(-\boldsymbol{C})^{-1}\boldsymbol{D}_L$) represents the stationary probability vector of the state of the underlying Markov chain at the beginning of a service of a high (resp. low) priority customer following an idle period which starts from an imbedded point with no customer. Taking the above facts into account, we obtain (3.1).

We now consider the marginal distribution of the number of high priority customers in the system at imbedded points. Let $\boldsymbol{P}_H(z)$ denote a $1 \times M$ vector whose jth element represents the GF for the number of high priority customers in the system at imbedded points when the underlying Markov chain is in state j. By definition, we have

$$\boldsymbol{P}_H(z) = \boldsymbol{P}^*(z,1), \qquad |z| \le 1,$$

and therefore, it follows from (3.1) that

$$\begin{aligned}(3.4)\qquad \boldsymbol{P}_H(z) = [\boldsymbol{P}_H(z) - \boldsymbol{P}^*(0,1)]\,\boldsymbol{A}_H(z)/z + [\boldsymbol{P}^*(0,1) - \boldsymbol{P}^*(0,0)]\boldsymbol{B}_H(z) \\ + \boldsymbol{P}^*(0,0)(-\boldsymbol{C})^{-1}[\boldsymbol{D}_H\boldsymbol{A}_H(z) + \boldsymbol{D}_L\boldsymbol{B}_H(z)],\end{aligned}$$

where, for $|z| \le 1$,

$$\boldsymbol{A}_H(z) = \boldsymbol{A}^*(z,1), \qquad \boldsymbol{B}_H(z) = \boldsymbol{B}^*(z,1).$$

Note here that $\boldsymbol{A}_H(z)$ and $\boldsymbol{B}_H(z)$ are given by (see (3.2) and (3.3))

$$\boldsymbol{A}_H(z) = \int_0^\infty e^{(\boldsymbol{C}_H + z\boldsymbol{D}_H)x} dH_H(x), \qquad \boldsymbol{B}_H(z) = \int_0^\infty e^{(\boldsymbol{C}_H + z\boldsymbol{D}_H)x} dH_L(x),$$

where

$$\boldsymbol{C}_H = \boldsymbol{C} + \boldsymbol{D}_L.$$

We now rewrite (3.4) as

$$\begin{aligned}(3.5)\qquad \boldsymbol{P}_H(z)\,[z\boldsymbol{I} - \boldsymbol{A}_H(z)] = \boldsymbol{P}^*(0,1)[z\boldsymbol{B}_H(z) - \boldsymbol{A}_H(z)] \\ + z\boldsymbol{P}^*(0,0)(-\boldsymbol{C})^{-1}\,[\boldsymbol{D}_H\boldsymbol{A}_H(z) + \boldsymbol{C}_H\boldsymbol{B}_H(z)]\,.\end{aligned}$$

Note that (3.5) contains two unknown vectors $\boldsymbol{P}^*(0,0)$ and $\boldsymbol{P}^*(0,1)$.

Theorem 3.1. *The vector* $\boldsymbol{P}^*(0,0)$ *is given by*

$$(3.6)\qquad \boldsymbol{P}^*(0,0) = \frac{(1-\rho)\boldsymbol{\kappa}(-\boldsymbol{C})}{\lambda}.$$

Proof. Note that the total number of customers in the system is a step function with unit jumps. Thus, the distribution of the total number of customers immediately before arrivals is the same as that immediately after departures [4]. Since the server is idle with probability $1-\rho$, the jth element of the vector $(1-\rho)\boldsymbol{\kappa}$ represents the stationary probability that the server is idle and the underlying Markov chain is in state j at a random point in time (see Remark 1). We then have

$$P^*(0,0)e = \frac{(1-\rho)\kappa De}{\pi De} = \frac{(1-\rho)\kappa(-C)e}{\lambda}. \tag{3.7}$$

On the other hand, κ is expressed as

$$\frac{P^*(0,0)(-C)^{-1}}{P^*(0,0)(-C)^{-1}e} = \kappa. \tag{3.8}$$

The key observation in (3.8) is as follows. Given that the server is idle, the probability that the underlying Markov chain is in state j at a random point in time is given by the ratio of the mean amount of time spent in state j during an idle period to the mean length of idle periods. Note that (i,j)th element of $(-C)^{-1}$ is the mean amount of time spent in state j, starting from state i, before an arrival [12]. Taking these into consideration, we obtain (3.8).

It follows from (3.8) that

$$P^*(0,0) = P^*(0,0)(-C)^{-1}e\kappa(-C), \tag{3.9}$$

from which, we obtain

$$P^*(0,0)e = P^*(0,0)(-C)^{-1}e\kappa(-C)e. \tag{3.10}$$

Comparing (3.10) with (3.7), we have

$$P^*(0,0)(-C)^{-1}e = \frac{1-\rho}{\lambda}. \tag{3.11}$$

(3.6) follows from (3.9) and (3.11). □

Next we consider $P^*(0,1)$. To obtain $P^*(0,1)$, we need the following lemma.

Lemma 3.2. *$P^*(0,1)e$ is given by*

$$P^*(0,1)e = \frac{\lambda_L}{\lambda} + \frac{1-\rho}{\lambda}\kappa D_H e. \tag{3.12}$$

Proof. The probability that an arbitrary departuring customer is of high priority is equal to λ_H/λ. Thus we have

$$1 - P^*(0,1)e + P^*(0,0)(-C)^{-1}D_H e = \lambda_H/\lambda, \tag{3.13}$$

from which, it follows that

$$\lambda P^*(0,1)e = \lambda_L + \lambda P^*(0,0)(-C)^{-1}D_H e. \tag{3.14}$$

Finally, we obtain (3.12) from (3.6) and (3.14). □

In order to derive the unknown vector $P^*(0,1)$, we construct another imbedded Markov chain. Among departure points of all customers, we remove points at which high priority customers exist. As new imbedded points, we select only departure points at which there are no high priority customers in the system.

Let $\psi = (\psi_0, \psi_1, \ldots)$ denote the stationary probability vector of the newly constructed imbedded Markov chain, where ψ_i is a $1 \times M$ vector whose jth element represents the stationary joint probability of having i low priority customers in the system and the underlying Markov chain being in state j at imbedded points. Furthermore, let $\boldsymbol{\Psi}(\omega)$ denote the vector GF of the stationary vector $\boldsymbol{\psi}_i$:

$$\boldsymbol{\Psi}(\omega)=\sum_{i=0}^{\infty}\boldsymbol{\psi}_i\omega^i, \qquad |\omega|\le 1.$$

From the definition of the newly constructed imbedded Markov chain, we have

$$\boldsymbol{\Psi}(\omega)=\boldsymbol{P}^*(0,\omega)/\boldsymbol{P}^*(0,1)\boldsymbol{e},$$

from which, it follows that

$$\text{(3.15)} \qquad \boldsymbol{\Psi}(0)=\boldsymbol{P}^*(0,0)/\boldsymbol{P}^*(0,1)\boldsymbol{e}, \qquad \boldsymbol{\Psi}(1)=\boldsymbol{P}^*(0,1)/\boldsymbol{P}^*(0,1)\boldsymbol{e}.$$

Note that $\boldsymbol{\Psi}(0)$ has already been obtained (see (3.6) and (3.12)).

We now consider the state transition from an imbedded point to the next imbedded point. Let $\boldsymbol{G}_H(\omega)$ ($|\omega|\le 1$) denote an $M\times M$ matrix whose (i,j)th element represents the GF for the number of low priority arrivals in a busy period of high priority customers when the underlying Markov chain is in state j at the end of the busy period given that the underlying Markov chain is in state i at the beginning of the busy period. To obtain $\boldsymbol{G}_H(\omega)$, we introduce $\boldsymbol{A}_k(\omega)$ ($|\omega|\le 1$) which satisfies

$$\text{(3.16)} \qquad \sum_{k=0}^{\infty}\boldsymbol{A}_k(\omega)z^k=\boldsymbol{A}^*(z,\omega).$$

That is, $\boldsymbol{A}_k(\omega)$ denotes the matrix GF for the number of low priority arrivals in the service time of a high priority customer when k high priority customers arrive in the service time. Using $\boldsymbol{A}_k(\omega)$, we have

$$\text{(3.17)} \qquad \boldsymbol{G}_H(\omega)=\sum_{k=0}^{\infty}\boldsymbol{A}_k(\omega)\boldsymbol{G}_H(\omega)^k.$$

The key observation in (3.17) is that the number of low priority arrivals in a busy period of high priority customers is given by the sum of the number of low priority arrivals in the first service of the busy period (which is represented by $\boldsymbol{A}_k(\omega)$) and the number of low priority arrivals in k busy periods following the first service time (which is represented by $\boldsymbol{G}_H(\omega)^k$) when k high priority customers arrive in the first service time. Similarly, we define $\boldsymbol{B}_k(\omega)$ ($|\omega|\le 1$) which satisfies

$$\sum_{k=0}^{\infty}\boldsymbol{B}_k(\omega)z^k=\boldsymbol{B}^*(z,\omega).$$

Note that $\boldsymbol{B}_k(\omega)$ represents the matrix GF for the number of low priority arrivals in the service time of a low priority customer when k high priority customers arrive in the service time.

Using $\boldsymbol{G}_H(\omega)$, we have the following equation for $\boldsymbol{\Psi}(\omega)$:

$$\text{(3.18)} \; \boldsymbol{\Psi}(\omega)=\left\{\frac{\boldsymbol{\Psi}(\omega)-\boldsymbol{\Psi}(0)}{\omega}\right\}\sum_{k=0}^{\infty}\boldsymbol{B}_k(\omega)\boldsymbol{G}_H(\omega)^k + \boldsymbol{\Psi}(0)(-\boldsymbol{C})^{-1}\left(\boldsymbol{D}_H\sum_{k=0}^{\infty}\boldsymbol{A}_k(\omega)\boldsymbol{G}_H(\omega)^k+\boldsymbol{D}_L\sum_{k=0}^{\infty}\boldsymbol{B}_k(\omega)\boldsymbol{G}_H(\omega)^k\right).$$

With (3.17), (3.18) becomes

$$(3.19)\quad \boldsymbol{\Psi}(\omega)[\omega \boldsymbol{I} - \boldsymbol{G}_L(\omega)] = \boldsymbol{\Psi}(0)\left[\omega(-\boldsymbol{C})^{-1}\{\boldsymbol{D}_H\boldsymbol{G}_H(\omega) + \boldsymbol{D}_L\boldsymbol{G}_L(\omega)\} - \boldsymbol{G}_L(\omega)\right],$$

where $\boldsymbol{G}_L(\omega)$ is defined as

$$(3.20)\qquad \boldsymbol{G}_L(\omega) = \sum_{k=0}^{\infty} \boldsymbol{B}_k(\omega)\boldsymbol{G}_H(\omega)^k, \qquad |\omega| \le 1.$$

Setting $\omega = 1$ in (3.19) yields

$$(3.21)\qquad \boldsymbol{\Psi}(1)(\boldsymbol{I} - \boldsymbol{G}_L) = \boldsymbol{\Psi}(0)(-\boldsymbol{C})^{-1}[\boldsymbol{D}_H\boldsymbol{G}_H + \boldsymbol{C}_H\boldsymbol{G}_L],$$

where $\boldsymbol{G}_H = \boldsymbol{G}_H(1)$ and $\boldsymbol{G}_L = \boldsymbol{G}_L(1)$. Note that $\boldsymbol{G}_H$ corresponds to the state transition matrix of the underlying Markov chain during a busy period of high priority customers. Thus, $\boldsymbol{G}_H$ satisfies (see (10) of [14])

$$\boldsymbol{G}_H = \int_0^{\infty} e^{(\boldsymbol{C}_H + \boldsymbol{D}_H\boldsymbol{G}_H)x} dH_H(x).$$

Also, $\boldsymbol{G}_L$ satisfies

$$\boldsymbol{G}_L = \int_0^{\infty} e^{(\boldsymbol{C}_H + \boldsymbol{D}_H\boldsymbol{G}_H)x} dH_L(x).$$

Let $\boldsymbol{g}_H$ denote the invariant probability vector of the transition matrix $\boldsymbol{G}_H$, namely,

$$\boldsymbol{g}_H = \boldsymbol{g}_H\boldsymbol{G}_H, \quad \boldsymbol{g}_H\boldsymbol{e} = 1.$$

Note that $\boldsymbol{g}_H\boldsymbol{G}_L = \boldsymbol{g}_H$ [14]. Adding $\boldsymbol{\Psi}(1)\boldsymbol{e}\boldsymbol{g}_H$ to both sides of (3.21), and observing that $\boldsymbol{I} - \boldsymbol{G}_L + \boldsymbol{e}\boldsymbol{g}_H$ is non-singular, we obtain

$$(3.22)\qquad \boldsymbol{\Psi}(1) = \boldsymbol{g}_H + \boldsymbol{\Psi}(0)(-\boldsymbol{C})^{-1}[\boldsymbol{D}_H\boldsymbol{G}_H + \boldsymbol{C}_H\boldsymbol{G}_L][\boldsymbol{I} - \boldsymbol{G}_L + \boldsymbol{e}\boldsymbol{g}_H]^{-1},$$

where we use the equality $\boldsymbol{\Psi}(1)\boldsymbol{e} = 1$. Finally, using (3.15) and (3.22), we have the following theorem.

Theorem 3.3. *The vector $\boldsymbol{P}^*(0,1)$ is given by*

$$(3.23)\quad \boldsymbol{P}^*(0,1) = \boldsymbol{P}^*(0,1)\boldsymbol{e}\boldsymbol{g}_H + \boldsymbol{P}^*(0,0)(-\boldsymbol{C})^{-1}[\boldsymbol{D}_H\boldsymbol{G}_H + \boldsymbol{C}_H\boldsymbol{G}_L][\boldsymbol{I} - \boldsymbol{G}_L + \boldsymbol{e}\boldsymbol{g}_H]^{-1},$$

where $\boldsymbol{P}^*(0,0)$ and $\boldsymbol{P}^*(0,1)\boldsymbol{e}$ are given in (3.6) and (3.12), respectively.

Thus $\boldsymbol{P}_H(z)$ is completely determined by (3.5), (3.6) and (3.23). In what follows, we provide the recursive formula for the derivatives of $\boldsymbol{P}_H(z)$ evaluated at $z = 1$. Since the derivation of this formula is routine (see [19]), we omit the proof.

Corollary 3.4. *We define for $n \ge 0$*

$$\boldsymbol{P}_H^{(n)} = \lim_{z\to 1}\frac{d^n}{dz^n}\boldsymbol{P}_H(z), \qquad \boldsymbol{A}_H^{(n)} = \lim_{z\to 1}\frac{d^n}{dz^n}\boldsymbol{A}_H(z), \qquad \boldsymbol{B}_H^{(n)} = \lim_{z\to 1}\frac{d^n}{dz^n}\boldsymbol{B}_H(z),$$

where $\boldsymbol{P}_H^{(0)} = \boldsymbol{P}_H(1)$, $\boldsymbol{A}_H^{(0)} = \boldsymbol{A}_H(1)$ and $\boldsymbol{B}_H^{(0)} = \boldsymbol{B}_H(1)$. Then $\boldsymbol{P}_H^{(n)}$ is recursively obtained by

$$\begin{aligned}
\boldsymbol{Z}_H^{(0)} &= \boldsymbol{P}^*(0,1)(\boldsymbol{B}_H^{(0)} - \boldsymbol{A}_H^{(0)}) + \boldsymbol{P}^*(0,0)(-\boldsymbol{C})^{-1}\left[\boldsymbol{D}_H\boldsymbol{A}_H^{(0)} + \boldsymbol{C}_H\boldsymbol{B}_H^{(0)}\right],\\
\boldsymbol{P}_H^{(0)} &= \boldsymbol{\pi} + \boldsymbol{Z}_H^{(0)}[\boldsymbol{I} - \boldsymbol{A}_H^{(0)} + \boldsymbol{e}\boldsymbol{\pi}]^{-1},\\
\boldsymbol{Z}_H^{(1)} &= \boldsymbol{P}^*(0,1)(\boldsymbol{B}_H^{(0)} + \boldsymbol{B}_H^{(1)} - \boldsymbol{A}_H^{(1)})\\
&\quad + \boldsymbol{P}^*(0,0)(-\boldsymbol{C})^{-1}\left[\boldsymbol{D}_H(\boldsymbol{A}_H^{(0)} + \boldsymbol{A}_H^{(1)}) + \boldsymbol{C}_H(\boldsymbol{B}_H^{(0)} + \boldsymbol{B}_H^{(1)})\right],
\end{aligned}$$

and for $n \geq 1$,

$$Z_H^{(n+1)} = \sum_{m=0}^{n-1} \binom{n+1}{m} P_H^{(m)} A_H^{(n+1-m)} + P^*(0,1)(B_H^{(n+1)} + (n+1)B_H^{(n)} - A_H^{(n+1)})$$
$$+ P^*(0,0)(-C)^{-1}\left[D_H(A_H^{(n+1)} + (n+1)A_H^{(n)}) + C_H(B_H^{(n+1)} + (n+1)B_H^{(n)})\right],$$
$$P_H^{(n)} e = \frac{Z_H^{(n+1)} e}{(n+1)(1-\rho_H)}$$
$$+ \frac{1}{1-\rho_H}\left[Z_H^{(n)} - nP_H^{(n-1)}(I - A_H^{(1)})\right][I - A_H^{(0)} + e\pi]^{-1} A_H^{(1)} e,$$
$$P_H^{(n)} = P_H^{(n)} e\pi + \left[Z_H^{(n)} - nP_H^{(n-1)}(I - A_H^{(1)})\right][I - A_H^{(0)} + e\pi]^{-1}.$$

Next we consider the probability mass function of the number of high priority customers immediately after departures of any priority class. Let p_k^H $(k = 0, 1, \ldots)$ denote a $1 \times M$ vector which satisfies

$$\sum_{k=0}^{\infty} p_k^H z^k = P_H(z).$$

Note that the jth element of p_k^H denotes the joint probability of having k high priority customers and the underlying Markov chain being in state j immediately after departures of any priority class. Furthermore, let A_k^H and B_k^H denote $M \times M$ matrices which satisfies

$$\sum_{k=0}^{\infty} A_k^H z^k = A_H(z), \qquad \sum_{k=0}^{\infty} B_k^H z^k = B_H(z).$$

We then have the following theorem.

Theorem 3.5. *The vector* p_k^H $(k = 1, 2, \ldots)$ *is obtained by the following recursion:*

$$p_k^H = \left[P^*(0,1)\overline{B}_k^H + P^*(0,0)(-C)^{-1}\left(D_H\overline{A}_k^H + C_H\overline{B}_k^H\right) + \sum_{n=1}^{k-1} p_n^H \overline{A}_{k+1-n}^H\right]$$
$$\cdot \left(I - \overline{A}_1^H\right)^{-1}, \qquad k \geq 1,$$

with $p_0^H = P^*(0,1)$, *where* $\overline{A}_k^H$ *and* $\overline{B}_k^H$ *are given by*

$$\overline{A}_k^H = \sum_{n=k}^{\infty} A_n^H (G_H)^{n-k}, \qquad \overline{B}_k^H = \sum_{n=k}^{\infty} B_n^H (G_H)^{n-k}.$$

Proof. It is clear, by definition, that $p_0^H = P^*(0,1)$. The recursion for p_k^H $(k \geq 1)$ is obtained by observing the system only when the number of high priority customers is equal to or less than k at departures and by considering the transition to the state having k high priority customers. Since a very similar recursion is well known (see [21]), we omit the details of the proof. □

3.2. Number of High Priority Customers at a Random Point in Time

In this subsection, we derive various formulas for the number of high priority customers at a random point in time. To do so, we first consider the number of high priority customers immediately after departures of high priority customers. Let $x^H = (x_0^H, x_1^H, \ldots)$ denote the stationary probability vector of the number of high priority customers immediately after departures of high priority class, where $\boldsymbol{x}_k^H$ denotes a $1 \times M$ vector whose jth element represents the joint probability of having k high priority customers in the system and the underlying Markov chain being in state j immediately after departures of high priority class. Furthermore, let $\boldsymbol{X}_H(z)$ denote the vector GF of the stationary vector $\boldsymbol{x}_k^H$:

$$\boldsymbol{X}_H(z) = \sum_{k=0}^{\infty} \boldsymbol{x}_k^H z^k, \qquad |z| \leq 1.$$

We then have the following theorem.

Theorem 3.6. *$\boldsymbol{X}_H(z)$ is given in terms of $\boldsymbol{P}_H(z)$ by*

$$\boldsymbol{X}_H(z) = \frac{\lambda}{\lambda_H}\left[\boldsymbol{P}_H(z) - \left\{\boldsymbol{P}^*(0,1) + \boldsymbol{P}^*(0,0)(-\boldsymbol{C})^{-1}\boldsymbol{C}_H\right\}\boldsymbol{B}_H(z)\right]. \tag{3.24}$$

Proof. By the definitions of $\boldsymbol{X}_H(z)$ and $\boldsymbol{P}_H(z)$, we have

$$\boldsymbol{X}_H(z) = \frac{[\boldsymbol{P}_H(z) - \boldsymbol{P}^*(0,1)]\boldsymbol{A}_H(z)/z + \boldsymbol{P}^*(0,0)(-\boldsymbol{C})^{-1}\boldsymbol{D}_H\boldsymbol{A}_H(z)}{[\boldsymbol{P}_H(1) - \boldsymbol{P}^*(0,1)]\boldsymbol{A}_H(1)\boldsymbol{e} + \boldsymbol{P}^*(0,0)(-\boldsymbol{C})^{-1}\boldsymbol{D}_H\boldsymbol{A}_H(1)\boldsymbol{e}}.$$

Noting $\boldsymbol{A}_H(1)\boldsymbol{e} = \boldsymbol{e}$, $\boldsymbol{P}_H(1)\boldsymbol{e} = 1$ and (3.13), we have

$$\boldsymbol{X}_H(z) = \frac{\lambda}{\lambda_H}\left[\{\boldsymbol{P}_H(z) - \boldsymbol{P}^*(0,1)\}\boldsymbol{A}_H(z)/z + \boldsymbol{P}^*(0,0)(-\boldsymbol{C})^{-1}\boldsymbol{D}_H\boldsymbol{A}_H(z)\right].$$

Finally, using (3.5), we obtain (3.24). □

The followings are direct conclusions of Theorem 3.6.

Corollary 3.7. *Let $\boldsymbol{X}_H^{(n)}$ $(n \geq 0)$ denote the nth derivative of $\boldsymbol{X}_H(z)$ evaluated at $z = 1$, where $\boldsymbol{X}_H^{(0)} = \boldsymbol{X}_H(1)$. Then, for $n \geq 0$,*

$$\boldsymbol{X}_H^{(n)} = \frac{\lambda}{\lambda_H}\left[\boldsymbol{P}_H^{(n)} - \left\{\boldsymbol{P}^*(0,1) + \boldsymbol{P}^*(0,0)(-\boldsymbol{C})^{-1}\boldsymbol{C}_H\right\}\boldsymbol{B}_H^{(n)}\right].$$

Corollary 3.8. *The vector $\boldsymbol{x}_k^H$ $(k \geq 0)$ is given by*

$$\boldsymbol{x}_k^H = \frac{\lambda}{\lambda_H}\left[\boldsymbol{p}_k^H - \left\{\boldsymbol{P}^*(0,1) + \boldsymbol{P}^*(0,0)(-\boldsymbol{C})^{-1}\boldsymbol{C}_H\right\}\boldsymbol{B}_{k\cdot}^H\right].$$

Let $\boldsymbol{Y}_H(z)$ $(|z| \leq 1)$ denote a $1 \times M$ vector whose ith element represents the GF for the number of high priority customers in the system at a random point in time when the underlying Markov chain is in state i.

Theorem 3.9. *$\boldsymbol{Y}_H(z)$ and $\boldsymbol{X}_H(z)$ are related by*

$$\boldsymbol{Y}_H(z)(\boldsymbol{C}_H + z\boldsymbol{D}_H) = \lambda_H(z-1)\boldsymbol{X}_H(z). \tag{3.25}$$

Proof. (3.25) is a special case of the result in [29], which has shown the relationship between the stationary queue length distributions at a random point in time and at departures in the stationary queue with batch *MAP* arrivals. Besides, we can show (3.25) by using an approach similar to [28]. □

According to the discussion in [14], we have the following results.

Corollary 3.10. *Let $\boldsymbol{Y}_H^{(n)}$ $(n \geq 0)$ denote the nth derivative of $\boldsymbol{Y}_H(z)$ evaluated at $z = 1$, where $\boldsymbol{Y}_H^{(0)} = \boldsymbol{Y}_H(1) = \boldsymbol{\pi}$. Then, $\boldsymbol{Y}_H^{(n)}$ $(n \geq 1)$ is obtained by the following recursion:*

$$\begin{aligned}\boldsymbol{Y}_H^{(n)}\boldsymbol{e} &= \boldsymbol{X}_H^{(n)}\boldsymbol{e} + n\left(\boldsymbol{X}_H^{(n-1)} - \boldsymbol{Y}_H^{(n-1)}\boldsymbol{D}_H/\lambda_H\right)\left(\boldsymbol{e\pi} - \boldsymbol{C}_H - \boldsymbol{D}_H\right)^{-1}\boldsymbol{D}_H\boldsymbol{e},\\ \boldsymbol{Y}_H^{(n)} &= \boldsymbol{Y}_H^{(n)}\boldsymbol{e\pi} + n\left(\boldsymbol{Y}_H^{(n-1)}\boldsymbol{D}_H - \lambda_H\boldsymbol{X}_H^{(n-1)}\right)\left(\boldsymbol{e\pi} - \boldsymbol{C}_H - \boldsymbol{D}_H\right)^{-1}.\end{aligned}$$

Corollary 3.11. *Let $\boldsymbol{y}_k^H$ $(k \geq 0)$ denote a $1 \times M$ vector which satisfies*

$$\sum_{k=0}^{\infty}\boldsymbol{y}_k^H z^k = \boldsymbol{Y}_H(z).$$

Then, $\boldsymbol{y}_k^H$ $(k \geq 1)$ is recursively obtain by

$$\begin{aligned}\boldsymbol{y}_0^H &= \lambda_H\boldsymbol{x}_0^H(-\boldsymbol{C}_H)^{-1},\\ \boldsymbol{y}_k^H &= \left[\boldsymbol{y}_{k-1}^H\boldsymbol{D}_H + \boldsymbol{\lambda}_H\left(\boldsymbol{x}_k^H - \boldsymbol{x}_{k-1}^H\right)\right](-\boldsymbol{C}_H)^{-1}, \qquad k \geq 1.\end{aligned}$$

3.3. Waiting Time of High Priority Customers

In this subsection, we consider the waiting time of high priority customers. For $|z| \leq 1$ and $Re(s) > 0$, let $\boldsymbol{A}_H(z,s)$ (resp. $\boldsymbol{B}_H(z,s)$) denote an $M \times M$ matrix which represents the matrix GF/LST of the joint distribution of the number of high priority arrivals during the backward recurrence time of a high (resp. low) priority service and the forward recurrence time of the service time. By definition, we have

$$\boldsymbol{A}_H(z,s) = \int_0^\infty \frac{x dH_H(x)}{h_H}\int_0^x \frac{dy}{x} e^{(\boldsymbol{C}_H + z\boldsymbol{D}_H)y} e^{-s(x-y)}.$$

After some calculations, we obtain

$$\boldsymbol{A}_H(z,s) = \frac{\boldsymbol{A}_H(z) - H_H^*(s)\boldsymbol{I}}{h_H}[s\boldsymbol{I} + \boldsymbol{C}_H + z\boldsymbol{D}_H]^{-1},$$

where $H_H^*(s)$ denotes the LST of the DF $H_H(x)$. Similarly,

$$\boldsymbol{B}_H(z,s) = \frac{\boldsymbol{B}_H(z) - H_L^*(s)\boldsymbol{I}}{h_L}[s\boldsymbol{I} + \boldsymbol{C}_H + z\boldsymbol{D}_H]^{-1},$$

where $H_L^*(s)$ denotes the LST of the DF $H_L(x)$.

For $|z| \leq 1$ and $Re(s) > 0$, let $\boldsymbol{\Pi}_H(z,s)$ (resp. $\boldsymbol{\Pi}_L(z,s)$) denote a $1 \times M$ vector which represents the vector GF/LST of the number of high priority customers and the forward recurrence time of the current service when a high (resp. low) priority customer is being served.

Lemma 3.12. *$\boldsymbol{\Pi}_H(z,s)$ and $\boldsymbol{\Pi}_L(z,s)$ are given by*

$$(3.26)\quad \boldsymbol{\Pi}_H(z,s) = \rho_H \cdot \frac{\lambda}{\lambda_H}\left[\boldsymbol{P}_H(z) - \boldsymbol{P}^*(0,1) + z\boldsymbol{P}^*(0,0)(-\boldsymbol{C})^{-1}\boldsymbol{D}_H\right]\boldsymbol{A}_H(z,s),$$

$$(3.27)\quad \boldsymbol{\Pi}_L(z,s) = \rho_L \cdot \frac{\lambda}{\lambda_L}\left[\boldsymbol{P}^*(0,1) - \boldsymbol{P}^*(0,0) + z\boldsymbol{P}^*(0,0)(-\boldsymbol{C})^{-1}\boldsymbol{D}_L\right]\boldsymbol{B}_H(z,s).$$

Proof. The above equations are derived from the following observation. The probability that the server is busy for a high (resp. low) priority service at a random point in time is given by ρ_H (resp. ρ_L). Given that the server is busy for a high priority service, the joint GF/LST for the number of high priority customers and the forward recurrence time of the current service is given by the product of the vector GF for the number of high priority customers immediately after the beginning of the current service (which is given by $\lambda\{\boldsymbol{P}_H(z)-\boldsymbol{P}^*(0,1)+z\boldsymbol{P}^*(0,0)(-\boldsymbol{C})^{-1}\boldsymbol{D}_H\}/\lambda_H$) and the matrix GF/LST $\boldsymbol{A}_H(z,s)$ for the number of high priority arrivals during the backward recurrence time of the current service and the forward recurrence time of the service time. These observations yield (3.26). (3.27) is also derived from very similar observations. □

Remark 3.13. *The vector GF $\boldsymbol{Y}_H(z)$ of the number of high priority customers at a random point in time is given by*

$$\boldsymbol{Y}_H(z) = (1-\rho)\boldsymbol{\kappa} + \lim_{s\to 0+}\boldsymbol{\Pi}_H(z,s) + \lim_{s\to 0+}\boldsymbol{\Pi}_L(z,s). \tag{3.28}$$

After some calculations with (3.28), we can verify (3.25).

Let $\boldsymbol{V}_H(s)$ $(Re(s)>0)$ denote a $1\times M$ vector whose jth element represents the LST of the virtual waiting time of high priority customers when the underlying Markov chain is in state j.

Theorem 3.14. *$\boldsymbol{V}_H(s)$ is given by*

$$\begin{aligned}\boldsymbol{V}_H(s) = {} & [(1-\rho)s\boldsymbol{\kappa} + \{1-H_L^*(s)\}\{\lambda\boldsymbol{P}^*(0,1)+(1-\rho)\boldsymbol{\kappa}\boldsymbol{C}_H\}] \\ & \cdot [s\boldsymbol{I}+\boldsymbol{C}_H+H_H^*(s)\boldsymbol{D}_H]^{-1}.\end{aligned} \tag{3.29}$$

Proof. By definition, we have

$$\boldsymbol{V}_H(s) = (1-\rho)\boldsymbol{\kappa} + \boldsymbol{\Pi}_H(H_H^*(s),s)/H_H^*(s) + \boldsymbol{\Pi}_L(H_L^*(s),s). \tag{3.30}$$

The substitution of (3.26) and (3.27) into (3.30) and some straightforward calculations yield

$$\begin{aligned}\boldsymbol{V}_H(s) = {} & (1-\rho)\boldsymbol{\kappa} + \lambda\left[\{1-H_L^*(s)\}\boldsymbol{P}^*(0,1) - \boldsymbol{P}^*(0,0)(-\boldsymbol{C})^{-1}\{H_H^*(s)\boldsymbol{D}_H\right. \\ & + H_L^*(s)\boldsymbol{C}_H\}][s\boldsymbol{I}+\boldsymbol{C}_H+H_H^*(s)\boldsymbol{D}_H]^{-1}.\end{aligned}$$

Finally, using (3.6), we obtain (3.29). □

We define for $n\geq 0$

$$\boldsymbol{V}_H^{(n)} = (-1)^n \lim_{s\to 0+}\frac{d^n}{ds^n}\boldsymbol{V}_H(s),$$

with $\boldsymbol{V}_H^{(0)} = \boldsymbol{V}_H(0+) = \boldsymbol{\pi}$ and for $n\geq 1$,

$$h_H^{(n)} = (-1)^n \lim_{s\to 0+}\frac{d^n}{ds^n}H_H^*(s), \qquad h_L^{(n)} = (-1)^n \lim_{s\to 0+}\frac{d^n}{ds^n}H_L^*(s).$$

We provide the recursive formula for $\boldsymbol{V}_H^{(n)}$ without proof.

Corollary 3.15. *$V_H^{(n)}$ $(n \geq 1)$ is obtained by the following recursion.*

$$\begin{aligned}
Z^{(1)} &= h_L^{(1)} \{\lambda P^*(0,1) + (1-\rho)\kappa C_H\} + (1-\rho)\kappa, \\
Z^{(n)} &= \sum_{m=2}^{n} \binom{n}{m} V_H^{(n-m)} h_H^{(m)} D_H + h_L^{(n)} \{\lambda P^*(0,1) + (1-\rho)\kappa C_H\}, \qquad n \geq 2, \\
V_H^{(n)} e &= \frac{Z^{(n+1)} e}{(n+1)(1-\rho_H)} \\
&\quad + \frac{h_H^{(1)}}{1-\rho_H} \left[n V_H^{(n-1)} \left(h_H^{(1)} D_H - I \right) + Z^{(n)} \right] (e\pi - C - D)^{-1} D_H e, \\
V_H^{(n)} &= V_H^{(n)} e\pi + \left[n V_H^{(n-1)} \left(h_H^{(1)} D_H - I \right) + Z^{(n)} \right] (e\pi - C - D)^{-1}.
\end{aligned}$$

Let $W_H(s)$ $(Re(s) > 0)$ denote a $1 \times M$ vector whose jth element represents the LST of the actual waiting time of high priority customers when the underlying Markov chain is in state j immediately after arrivals. By considering the state of the underlying Markov chain upon high priority arrivals, we have the following theorem.

Theorem 3.16. *$W_H(s)$ is given by*

$$W_H(s) = V_H(s) D_H / \lambda_H.$$

Corollary 3.17. *$V_H(s)e$ and $W_H(s)e$ are related by*

$$V_H(s)e = 1 - \rho + \rho_L \frac{1 - H_L^*(s)}{sh_L} + \rho_H \frac{1 - H_H^*(s)}{sh_H} W_H(s)e. \tag{3.31}$$

Proof. Post-multiplying both sides of (3.29) by $[sI + C_H + H_H^*(s) D_H]e$ and noting (3.6), (3.12) and

$$\lambda [P^*(0,1) + (1-\rho)\kappa C_H] e = \lambda_L, \tag{3.32}$$

$$[sI + C_H + H_H^*(s) D_H] e = se - \{1 - H_H^*(s)\} D_H e,$$

we obtain (3.31). □

4. ANALYSIS OF LOW PRIORITY CLASS

In this section, we consider various quantities of interest with respect to low priority customers. In section 4.1, we study the distribution of the number of low priority customers in the system at imbedded points introduced in section 3.1. In section 4.2, we study the number of low priority customers at a random point in time. Finally, in section 4.3, we study the waiting time distribution of low priority customers.

4.1. Number of Low Priority Customers at Imbedded Points

To analyze the low priority class, we first consider the number of low priority customers immediately after departures given that there are no high priority customers at those instants. Recall that $\Psi(\omega)$ is defined as the vector GF for the number of low priority customers immediately after departures given that there are no high priority customers.

We rewrite (3.19) as

$$\Psi(\omega)[\omega I - G_L(\omega)] = \omega \Psi(0)(-C)^{-1} [D_H G_H(\omega) + C_H G_L(\omega)] + (\omega - 1)\Psi(0) G_L(\omega).$$

We define for $n \geq 0$

$$\boldsymbol{\Psi}^{(n)} = \lim_{\omega \to 1} \frac{d^n}{d\omega^n} \boldsymbol{\Psi}(\omega), \qquad \boldsymbol{G}_H^{(n)} = \lim_{\omega \to 1} \frac{d^n}{d\omega^n} \boldsymbol{G}_H(\omega), \qquad \boldsymbol{G}_L^{(n)} = \lim_{\omega \to 1} \frac{d^n}{d\omega^n} \boldsymbol{G}_L(\omega),$$

with $\boldsymbol{\Psi}^{(0)} = \boldsymbol{\Psi}(1)$, $\boldsymbol{G}_H^{(0)} = \boldsymbol{G}_H(1)$ and $\boldsymbol{G}_L^{(0)} = \boldsymbol{G}_L(1)$. We provide the recursive formula for $\boldsymbol{\Psi}^{(n)}$ without proof.

Corollary 4.1. *$\boldsymbol{\Psi}^{(n)}$ $(n \geq 0)$ are recursively obtained by*

$$\begin{aligned}
\boldsymbol{Z}_L^{(1)} &= \boldsymbol{\Psi}(0)(-\boldsymbol{C})^{-1}\left[\boldsymbol{D}_H\left(\boldsymbol{G}_H^{(0)} + \boldsymbol{G}_H^{(1)}\right) + \boldsymbol{C}_H\left(\boldsymbol{G}_L^{(0)} + \boldsymbol{G}_L^{(1)}\right)\right] + \boldsymbol{\Psi}(0)\boldsymbol{G}_L^{(0)}, \\
\boldsymbol{Z}_L^{(n)} &= \sum_{m=2}^{n} \binom{n}{m} \boldsymbol{\Psi}^{(n-m)}\boldsymbol{G}_L^{(m)} + n\boldsymbol{\Psi}(0)\boldsymbol{G}_L^{(n-1)} \\
&\qquad + \boldsymbol{\Psi}(0)(-\boldsymbol{C})^{-1}\left[\boldsymbol{D}_H\left(n\boldsymbol{G}_H^{(n-1)} + \boldsymbol{G}_H^{(n)}\right) + \boldsymbol{C}_H\left(n\boldsymbol{G}_L^{(n-1)} + \boldsymbol{G}_L^{(n)}\right)\right], \quad n \geq 2, \\
\boldsymbol{\Psi}^{(n)}\boldsymbol{e} &= \left[\frac{1}{n+1}\boldsymbol{Z}_L^{(n+1)}\boldsymbol{e} + \left\{\boldsymbol{Z}_L^{(n)} - n\boldsymbol{\Psi}^{(n-1)}\left(\boldsymbol{I} - \boldsymbol{G}_L^{(1)}\right)\right\}\left(\boldsymbol{I} - \boldsymbol{G}_L + \boldsymbol{e}\boldsymbol{g}_H\right)^{-1}\boldsymbol{G}_L^{(1)}\boldsymbol{e}\right] \\
&\qquad \Big/ \left(1 - \boldsymbol{g}_H\boldsymbol{G}_L^{(1)}\boldsymbol{e}\right), \\
\boldsymbol{\Psi}^{(n)} &= \boldsymbol{\Psi}^{(n)}\boldsymbol{e}\boldsymbol{g}_H + \left\{\boldsymbol{Z}_L^{(n)} - n\boldsymbol{\Psi}^{(n-1)}\left(\boldsymbol{I} - \boldsymbol{G}_L^{(1)}\right)\right\}\left(\boldsymbol{I} - \boldsymbol{G}_L + \boldsymbol{e}\boldsymbol{g}_H\right)^{-1},
\end{aligned}$$

where $\boldsymbol{\Psi}^{(0)} = \boldsymbol{\Psi}(0)$ is given in (3.22).

Let $\boldsymbol{G}_k^H$ and $\boldsymbol{G}_k^L$ denote $M \times M$ matrices which satisfy

$$\sum_{k=0}^{\infty} \boldsymbol{G}_k^H \omega^k = \boldsymbol{G}_H(\omega), \qquad \sum_{k=0}^{\infty} \boldsymbol{G}_k^L \omega^k = \boldsymbol{G}_L(\omega).$$

According to the consideration similar to the proof of Theorem 3.5, we have the following theorem for the recursive formula to compute the coefficient vector $\boldsymbol{\psi}_i$.

Theorem 4.2. *The vector $\boldsymbol{\psi}_k$ $(k \geq 1)$ is obtained by the following recursion:*

$$\boldsymbol{\psi}_k = \left[\boldsymbol{\psi}_0(-\boldsymbol{C})^{-1}\left(\boldsymbol{D}_H\overline{\boldsymbol{G}}_k^H + \boldsymbol{D}_L\overline{\boldsymbol{G}}_k^L\right) + \sum_{n=1}^{k-1} \boldsymbol{\psi}_n \overline{\boldsymbol{G}}_{k+1-n}^L\right]\left(\boldsymbol{I} - \overline{\boldsymbol{G}}_1^L\right)^{-1}, \qquad k \geq 1,$$

with $\boldsymbol{\psi}_0 = \boldsymbol{\Psi}(0)$, where

$$\overline{\boldsymbol{G}}_k^H = \sum_{n=k}^{\infty} \boldsymbol{G}_n^H \boldsymbol{G}^{n-k}, \qquad \overline{\boldsymbol{G}}_k^L = \sum_{n=k}^{\infty} \boldsymbol{G}_n^L \boldsymbol{G}^{n-k},$$

and

$$\boldsymbol{G} = \int_0^{\infty} e^{\boldsymbol{Q}x} dH_L(x).$$

4.2. Number of Low Priority Customers at a Random Point in Time

In this subsection, we derive various formulas for the number of low priority customers at a random point in time. To do so, we first consider the number of low priority customers immediately after departures of low priority customers. Let $\boldsymbol{x}^L = (\boldsymbol{x}_0^L, \boldsymbol{x}_1^L, \ldots)$ denote the stationary probability vector of the number of low priority customers immediately after departures of low priority class, where $\boldsymbol{x}_k^L$ is a $1 \times M$ vector whose jth element represents the joint probability of having k low priority customers in the system and the underlying Markov chain being in state j immediately after departures of low priority customers. Furthermore, let $\boldsymbol{X}_L(\omega)$ denote the vector GF of the stationary vector $\boldsymbol{x}_k^L$:

$$\boldsymbol{X}_L(\omega) = \sum_{k=0}^{\infty} \boldsymbol{x}_k^L \omega^k, \qquad |\omega| \leq 1.$$

Let $\boldsymbol{B}_L(\omega)$ ($|\omega| \leq 1$) denote an $M \times M$ matrix which represents the matrix GF for the number of low priority arrivals during a service time of a low priority customer. Note that

$$\boldsymbol{B}_L(\omega) = \int_0^{\infty} e^{(\boldsymbol{C}_L + \omega \boldsymbol{D}_L)x} dH_L(x),$$

where $\boldsymbol{C}_L$ is given by

$$\boldsymbol{C}_L = \boldsymbol{C} + \boldsymbol{D}_H.$$

Theorem 4.3. *$\boldsymbol{X}_L(\omega)$ is given in terms of $\boldsymbol{\Psi}(\omega)$ by*

$$\boldsymbol{X}_L(\omega) = \frac{\lambda}{\lambda_L} \boldsymbol{P}^*(0,1)\boldsymbol{e} \left[\{\boldsymbol{\Psi}(\omega) - \boldsymbol{\Psi}(0)\} / \omega + \boldsymbol{\Psi}(0)(-\boldsymbol{C})^{-1}\boldsymbol{D}_L \right] \boldsymbol{B}_L(\omega). \tag{4.1}$$

Proof. By definition, $\boldsymbol{X}_L(\omega)$ is given by

$$\boldsymbol{X}_L(\omega) = \frac{[\{\boldsymbol{\Psi}(\omega) - \boldsymbol{\Psi}(0)\} / \omega + \boldsymbol{\Psi}(0)(-\boldsymbol{C})^{-1}\boldsymbol{D}_L] \boldsymbol{B}_L(\omega)}{1 - \boldsymbol{\Psi}(0)\boldsymbol{e} + \boldsymbol{\Psi}(0)(-\boldsymbol{C})^{-1}\boldsymbol{D}_L \boldsymbol{e}}.$$

Multiplying the numerator and denominator by $\boldsymbol{P}^*(0,1)\boldsymbol{e}$ and noting (3.15) and (3.32), we obtain (4.1). ∎

We define for $n \geq 0$

$$\boldsymbol{X}_L^{(n)} = \lim_{\omega \to 1} \frac{d^n}{d\omega^n} \boldsymbol{X}_L(\omega), \qquad \boldsymbol{B}_L^{(n)} = \lim_{\omega \to 1} \frac{d^n}{d\omega^n} \boldsymbol{B}_L(\omega),$$

with $\boldsymbol{X}_L^{(0)} = \boldsymbol{X}_L(1)$ and $\boldsymbol{B}_L^{(0)} = \boldsymbol{B}_L(1)$. The followings are direct conclusions of Theorem 4.3.

Corollary 4.4. *$\boldsymbol{X}_L^{(n)}$ $(n \geq 1)$ is obtained by the following recursion.*

$$\boldsymbol{X}_L^{(n)} = -n\boldsymbol{X}_L^{(n-1)} + \frac{\lambda}{\lambda_L} \boldsymbol{P}^*(0,1)\boldsymbol{e} \cdot \left[\sum_{m=0}^{n} \binom{n}{m} \boldsymbol{\Psi}^{(m)} \boldsymbol{B}_L^{(n-m)} + \boldsymbol{\Psi}(0)(-\boldsymbol{C})^{-1} \left(\boldsymbol{C}_H \boldsymbol{B}_L^{(n)} + n\boldsymbol{D}_L \boldsymbol{B}_L^{(n-1)} \right) \right].$$

Corollary 4.5. *$\boldsymbol{x}_k^L$ $(k \geq 0)$ is obtained by*

$$\boldsymbol{x}_k^L = \frac{\lambda}{\lambda_L} \boldsymbol{P}^*(0,1)\boldsymbol{e} \left[\sum_{m=1}^{k+1} \boldsymbol{\psi}_m \boldsymbol{B}_{k+1-m}^L + \boldsymbol{\psi}_0(-\boldsymbol{C})^{-1}\boldsymbol{D}_L \boldsymbol{B}_k^L \right],$$

where $\boldsymbol{B}_k^L$ $(k \geq 0)$ is an $M \times M$ matrix which satisfies

$$\sum_{k=0}^{\infty} \boldsymbol{B}_k^L \omega^k = \boldsymbol{B}_L(\omega).$$

Let $\boldsymbol{Y}_L(\omega)$ ($|\omega| \le 1$) denote a $1 \times M$ vector whose jth element represents the GF for the number of low priority customers in the system at a random point in time when the underlying Markov chain is in state j. Applying the result in [29], we obtain the following theorem.

Theorem 4.6. *$\boldsymbol{Y}_L(\omega)$ and $\boldsymbol{X}_L(\omega)$ are related by*

$$\boldsymbol{Y}_L(\omega)(\boldsymbol{C}_L + \omega \boldsymbol{D}_L) = \lambda_L(\omega - 1)\boldsymbol{X}_L(\omega).$$

Since the above formula takes the same form as in Theorem 3.9, the derivatives of $\boldsymbol{Y}_L(\omega)$ evaluated at $\omega = 1$ in terms of $\boldsymbol{X}_L^{(n)}$ and the coefficient matrices of $\boldsymbol{Y}_L(\omega)$ in terms of $\boldsymbol{x}_k^L$ are obtained by the same recursion as in Corollaries 3.10 and 3.11, with appropriate changes of notations.

4.3. Waiting Times of Low Priority Customers

The waiting time distribution of low priority customers in the nonpreemptive priority queue is identical to that in the counterpart of the preemptive resume priority queue. Since the waiting time distribution in the preemptive resume priority queue has been analyzed in more general settings than in this paper [26], we only provide the results from [26].

Let $\widehat{\boldsymbol{w}}_L(x)$ denote a $1 \times M$ vector whose jth element represents the probability that the amount of work in the system is equal to or less than x immediately before an arrival of an arbitrary low priority customer when the underlying Markov chain is in state j immediately after the arrival. We denote the LST of $\widehat{\boldsymbol{w}}_L(x)$ by $\widehat{\boldsymbol{w}}_L^*(s)$:

$$\widehat{\boldsymbol{w}}_L^*(s) = \int_0^\infty e^{-sx} d\widehat{\boldsymbol{w}}_L(x), \qquad Re(s) > 0.$$

Note that $\widehat{\boldsymbol{w}}_L^*(s)$ is given in terms of $\boldsymbol{V}^*(s)$:

$$\widehat{\boldsymbol{w}}_L^*(s) = \boldsymbol{V}^*(s)\boldsymbol{D}_L/\lambda_L,$$

where $\boldsymbol{V}(s)$ is given in (2.3). We then have the following theorem [26].

Theorem 4.7. *The LST $w_L^*(s)$ for the waiting time of low priority customers is given by*

$$w_L^*(s) = \int_0^\infty d\widehat{\boldsymbol{w}}_L(x) e^{\boldsymbol{Q}_H^*(s)x}\boldsymbol{e}, \qquad Re(s) > 0,$$

where $\boldsymbol{Q}_H^*(s)$ is an $M \times M$ matrix which satisfies

$$\boldsymbol{Q}_H^*(s) = \boldsymbol{C}_H - s\boldsymbol{I} + \boldsymbol{D}_H \int_0^\infty e^{\boldsymbol{Q}_H^*(s)x} dH_H(x), \qquad Re(s) > 0. \tag{4.2}$$

Let $\boldsymbol{\kappa}_H$ denote a $1 \times M$ vector which satisfies

$$\boldsymbol{\kappa}_H \boldsymbol{Q}_H = \boldsymbol{0}, \qquad \boldsymbol{\kappa}_H \boldsymbol{e} = 1,$$

where

$$\boldsymbol{Q}_H = \lim_{s \to 0+} \boldsymbol{Q}_H^*(s) = \boldsymbol{C}_H + \boldsymbol{D}_H \boldsymbol{G}_H.$$

Note here that

$$\boldsymbol{\kappa}_H = \boldsymbol{g}_H.$$

Using the result in [26], we have the following result.

Corollary 4.8. *The mean waiting time $E[w_L]$ of low priority customers is given by*

$$E[w_L] = \frac{E[\widehat{w}_L]}{1-\rho_H} + \left[\frac{\boldsymbol{\pi}\boldsymbol{D}_L}{\lambda_L} - \int_0^\infty d\widehat{\boldsymbol{w}}_L(x)e^{\boldsymbol{Q}_H x}\right]\left[(\boldsymbol{e}-\boldsymbol{\beta}_H)\boldsymbol{g}_H - \boldsymbol{C} - \boldsymbol{D}\right]^{-1}\boldsymbol{e},$$

where

$$E[\widehat{w}_L] = -\lim_{s\to 0+}\frac{d}{ds}\widehat{\boldsymbol{w}}_L^*(s)\boldsymbol{e}, \qquad \boldsymbol{\beta}_H = h_H\boldsymbol{D}_H\boldsymbol{e},$$

and

$$\int_0^\infty d\widehat{\boldsymbol{w}}_L(x)e^{\boldsymbol{Q}_H x} = \frac{\lambda}{\lambda_L}\left[\boldsymbol{P}^*(0,1) + \boldsymbol{P}^*(0,0)(-\boldsymbol{C})^{-1}\boldsymbol{C}_H\right].$$

5. ALGORITHMIC IMPLEMENTATION AND NUMERICAL EXAMPLES

In this section, we first consider the algorithms of the essential quantities in computing various performance measures obtained in this paper. Next we discuss the numerical implementation of these algorithms. Finally, we show some numerical examples.

5.1. Algorithms

We note that $\boldsymbol{Q}$, $\boldsymbol{G}_H$, $\boldsymbol{G}_L$, $\boldsymbol{A}_k^H$, $\boldsymbol{B}_k^H$, $\boldsymbol{G}_k^H$ and $\boldsymbol{G}_k^L$ are the essential quantities in computing the queue length distributions. In this subsection, we provide the algorithms to compute these quantities.

As for the computation of $\boldsymbol{Q}$, we combine the algorithm in [26] and the linear extrapolation proposed in section 3 of [7]. That is, the approximation $\boldsymbol{Q}^*[N]$ to matrix $\boldsymbol{Q}$ is computed in the following way. Starting with $\boldsymbol{Q}[0] = \boldsymbol{0}$, we compute for $N = 1, 2\ldots$

$$\begin{aligned}\boldsymbol{Q}[N] &= \boldsymbol{C} + \boldsymbol{D}_H\int_0^\infty e^{\boldsymbol{Q}[N-1]x}dH_H(x) + \boldsymbol{D}_L\int_0^\infty e^{\boldsymbol{Q}[N-1]x}dH_L(x),\\ \boldsymbol{Q}^*[N] &= \boldsymbol{Q}[N] + \boldsymbol{J}_N^{(1)}\left(\boldsymbol{Q}[N] - \boldsymbol{Q}[N-1]\right),\end{aligned}$$

where $\boldsymbol{J}_N^{(1)}$ is a diagonal matrix so that every row sum in $\boldsymbol{Q}^*[N]$ becomes zero.

As for the computation of $\boldsymbol{G}_H$, the detailed descriptions of the computational algorithms are provided in [15] and [28]. We follow the algorithm in [28]. Namely, the approximation $\boldsymbol{G}_H^*[N]$ to matrix $\boldsymbol{G}_H$ is obtained by starting $\boldsymbol{G}_H[0] = \boldsymbol{0}$ and computing for $N = 1, 2, \ldots,$

$$(5.1)\qquad \begin{aligned}\boldsymbol{G}_H[N] &= \sum_{n=0}^\infty \gamma_n^H\left[\boldsymbol{I} + \theta_H^{-1}(\boldsymbol{C}_H + \boldsymbol{D}_H\boldsymbol{G}_H[N-1])\right]^n,\\ \boldsymbol{G}_H^*[N] &= \boldsymbol{G}_H[N] + \boldsymbol{J}_N^{(2)}\left(\boldsymbol{G}_H[N] - \boldsymbol{G}_H[N-1]\right),\end{aligned}$$

where $\boldsymbol{J}_N^{(2)}$ is a diagonal matrix so that matrix $\boldsymbol{G}_H^*[N]$ is stochastic and

$$\theta_H = \max_i(-[\boldsymbol{C}_H]_{ii}), \qquad \gamma_n^H = \int_0^\infty e^{-\theta_H x}\frac{(\theta_H x)^n}{n!}dH_H(x).$$

Furthermore, the approximation $\boldsymbol{G}_L^*[N]$ to matrix $\boldsymbol{G}_L$ is computed by

$$(5.2)\qquad \begin{aligned}\boldsymbol{G}_L[N] &= \sum_{n=0}^\infty \gamma_n^L\left[\boldsymbol{I} + \theta_H^{-1}(\boldsymbol{C}_H + \boldsymbol{D}_H\boldsymbol{G}_H[N])\right]^n,\\ \boldsymbol{G}_L^*[N] &= \boldsymbol{G}_L[N] + \boldsymbol{J}_N^{(3)}\left(\boldsymbol{G}_L[N] - \boldsymbol{G}_L[N-1]\right),\end{aligned}$$

where $J_N^{(3)}$ is a diagonal matrix so that matrix $G_L^*[N]$ is stochastic and

$$\gamma_n^L = \int_0^\infty e^{-\theta_H x} \frac{(\theta_H x)^n}{n!} dH_L(x). \tag{5.3}$$

Next we consider the computation of matrices A_k^H and B_k^H. We follows the algorithm in [28] for these quantities. Namely, matrices A_k^H and B_k^H are computed by

$$A_k^H = \sum_{n=k}^{\infty} \gamma_n^H F_n^{(k)}, \qquad B_k^H = \sum_{n=k}^{\infty} \gamma_n^L F_n^{(k)}, \tag{5.4}$$

where $k \geq 0$ and matrices $F_n^{(k)}$ $(n = 0, 1, 2, \ldots, k = 0, 1, \ldots, n)$ are given by the following recursion with $F_0^{(0)} = I$:

$$F_{n+1}^{(k)} = \begin{cases} F_n^{(0)}(I + \theta_H^{-1} C_H) = (I + \theta_H^{-1} C_H)^{n+1}, & k = 0, \\ F_n^{(k)}(I + \theta_H^{-1} C_H) + F_n^{(k-1)}(\theta_H^{-1} D_H), & 1 \leq k \leq n, \\ F_n^{(n)}(\theta_H^{-1} D_H) = (\theta_H^{-1} D_H)^{n+1}, & k = n+1. \end{cases}$$

Once the matrices A_k^H and B_k^H are obtained, the moment matrices $A_H^{(n)}$ and $B_H^{(n)}$ are computed by

$$A_H^{(n)} = \sum_{k=n}^{\infty} \frac{k!}{(k-n)!} A_k^H, \qquad B_H^{(n)} = \sum_{k=n}^{\infty} \frac{k!}{(k-n)!} B_k^H.$$

Next we consider the computation of the coefficient matrices G_k^H and G_k^L of $G_H(\omega)$ and $G_L(\omega)$.

Lemma 5.1. *$G_H(\omega)$ and $G_L(\omega)$ satisfy the following equations.*

$$G_H(\omega) = \int_0^\infty e^{(C + D_H G_H(\omega) + \omega D_L)x} dH_H(x), \tag{5.5}$$

$$G_L(\omega) = \int_0^\infty e^{(C + D_H G_H(\omega) + \omega D_L)x} dH_L(x). \tag{5.6}$$

Proof. According to a discussion similar to section 2 of [26], we can show that the matrix GF for the number of low priority arrivals during the first passage time, (governed only by high priority arrivals) to the idle state with the initial work x is given by

$$e^{(C + D_H G_H(\omega) + \omega D_L)x}$$

from which, (5.5) and (5.6) follow. For brevity, we omit the details of the proof. □

We now consider the computation of G_k^H. We rewrite (5.5) as

$$G_H(\omega) = \sum_{k=0}^{\infty} \hat{\gamma}_n^H \left[I + \theta^{-1} \left(C + D_H G_H(\omega) + \omega D_L \right) \right]^n, \tag{5.7}$$

where

$$\theta = \max_i(-[C]_{i,i}), \qquad \hat{\gamma}_n^H = \int_0^\infty e^{-\theta x} \frac{(\theta x)^n}{n!} dH_H(x).$$

Let $\boldsymbol{\Phi}_n^{(k)}$ $(n \geq 0,\ k \geq 0)$ denote an $M \times M$ which satisfies

$$\sum_{k=0}^{\infty} \boldsymbol{\Phi}_n^{(k)} \omega^k = \left[\boldsymbol{I} + \theta^{-1}(\boldsymbol{C} + \boldsymbol{D}_H \boldsymbol{G}_H(\omega) + \omega \boldsymbol{D}_L)\right]^n. \tag{5.8}$$

We then have from (5.7) and (5.8)

$$\boldsymbol{G}_k^H = \sum_{n=0}^{\infty} \hat{\gamma}_n^H \boldsymbol{\Phi}_n^{(k)}, \qquad k \geq 0. \tag{5.9}$$

Note here that

$$\begin{aligned}\sum_{k=0}^{\infty} \boldsymbol{\Phi}_{n+1}^{(k)} \omega^k &= \sum_{k=0}^{\infty} \boldsymbol{\Phi}_n^{(k)} \omega^k \left[\boldsymbol{I} + \theta^{-1}(\boldsymbol{C} + \boldsymbol{D}_H \boldsymbol{G}_H(\omega) + \omega \boldsymbol{D}_L)\right] \\ &= \sum_{k=0}^{\infty} \boldsymbol{\Phi}_n^{(k)} \left[\boldsymbol{I} + \theta^{-1}\boldsymbol{C}\right] \omega^k + \sum_{k=0}^{\infty} \boldsymbol{\Phi}_n^{(k)} \theta^{-1} \boldsymbol{D}_H \boldsymbol{G}_H(\omega) \omega^k \\ &\quad + \sum_{k=0}^{\infty} \boldsymbol{\Phi}_n^{(k)} \theta^{-1} \boldsymbol{D}_L \omega^{k+1}.\end{aligned} \tag{5.10}$$

Comparing the coefficient matrices of ω^k in both sides of (5.10), we obtain for $n \geq 0$

$$\begin{aligned}\boldsymbol{\Phi}_{n+1}^{(0)} &= \boldsymbol{\Phi}_n^{(0)} \left[\boldsymbol{I} + \theta^{-1}\left(\boldsymbol{C} + \boldsymbol{D}_H \boldsymbol{G}_0^H\right)\right], \\ \boldsymbol{\Phi}_{n+1}^{(k)} &= \boldsymbol{\Phi}_n^{(k)} \left[\boldsymbol{I} + \theta^{-1}\boldsymbol{C}\right] + \sum_{m=0}^{k} \boldsymbol{\Phi}_n^{(m)} \theta^{-1} \boldsymbol{D}_H \boldsymbol{G}_{k-m}^H + \boldsymbol{\Phi}_n^{(k-1)} \theta^{-1} \boldsymbol{D}_L,\end{aligned}$$

with $\boldsymbol{\Phi}_0^{(0)} = \boldsymbol{I}$ and $\boldsymbol{\Phi}_0^{(k)} = \boldsymbol{0}$ for $k \geq 1$.

Therefore, the approximation $\boldsymbol{G}_k^H[N]$ to matrix $\boldsymbol{G}_k^H$ is obtained by starting with $\boldsymbol{G}_k^H[0] = \boldsymbol{0}$ $(k \geq 0)$ and computing for $N = 1, 2, \ldots,$

$$\boldsymbol{G}_k^H[N] = \sum_{n=0}^{\infty} \hat{\gamma}_n^H \boldsymbol{\Phi}_n^{(k)}[N], \tag{5.11}$$

where $\boldsymbol{\Phi}_0^{(0)}[N] = \boldsymbol{I}$ $\boldsymbol{\Phi}_0^{(k)}[N] = \boldsymbol{0}$ $(k \geq 1)$ and $\boldsymbol{\Phi}_n^{(k)}[N]$ $(n \geq 1,\ k \geq 0)$ are recursively computed by

$$\begin{aligned}\boldsymbol{\Phi}_{n+1}^{(0)}[N] &= \boldsymbol{\Phi}_n^{(0)}[N] \left[\boldsymbol{I} + \theta^{-1}\left(\boldsymbol{C} + \boldsymbol{D}_H \boldsymbol{G}_0^H[N-1]\right)\right], \\ \boldsymbol{\Phi}_{n+1}^{(k)}[N] &= \boldsymbol{\Phi}_n^{(k)}[N] \left[\boldsymbol{I} + \theta^{-1}\boldsymbol{C}\right] + \sum_{m=0}^{k} \boldsymbol{\Phi}_n^{(m)}[N] \theta^{-1} \boldsymbol{D}_H \boldsymbol{G}_{k-m}^H[N-1] \\ &\quad + \boldsymbol{\Phi}_n^{(k-1)}[N] \theta^{-1} \boldsymbol{D}_L.\end{aligned} \tag{5.12}$$

Once the approximations $\boldsymbol{G}_k^H[N]$ to matrices $\boldsymbol{G}_k^H$ are obtained, the approximation $\boldsymbol{G}_k^L[N]$ to matrix $\boldsymbol{G}_k^L$ is computed by

$$\boldsymbol{G}_k^L[N] = \sum_{n=0}^{\infty} \hat{\gamma}_n^L \boldsymbol{\Phi}_n^{(k)}[N], \tag{5.13}$$

where

$$\hat{\gamma}_n^H = \int_0^{\infty} e^{-\theta x} \frac{(\theta x)^n}{n!} dH_L(x).$$

Furthermore, the moment matrices $\boldsymbol{G}_H^{(n)}$ and $\boldsymbol{G}_L^{(n)}$ are computed by

$$\boldsymbol{G}_H^{(n)} = \sum_{k=n}^{\infty} \frac{k!}{(k-n)!} \boldsymbol{G}_k^H[N], \qquad \boldsymbol{G}_L^{(n)} = \sum_{k=n}^{\infty} \frac{k!}{(k-n)!} \boldsymbol{G}_k^L[N].$$

5.2. Stopping Criteria

In computing matrix $\boldsymbol{Q}$ by the algorithm described in the preceding subsection, we need a criterion to stop the recursive computation of matrices $\boldsymbol{Q}[N]$. Since $\boldsymbol{Q}[N]$ is entrywise nondecreasing and $\boldsymbol{Q}[\infty]$ is the infinitesimal generator of a positive recurrent Markov chain [26], we adopt the following scheme in computing matrix $\boldsymbol{Q}$. First we set ϵ_Q such that $0 < \epsilon_Q \ll 1$. Then we stop the computation of $\boldsymbol{Q}[N]$ when the following stopping criterion is satisfied for some $N = N^*$

$$\max_i[-\boldsymbol{Q}[N^*]\boldsymbol{e}]_i < \epsilon_Q.$$

We then use the approximation $\boldsymbol{Q}^*[N^*]$ to matrix $\boldsymbol{Q}$.

Next we consider the computation of matrix $\boldsymbol{G}_H$. In the algorithm described in the preceding subsection, we need to truncate the infinite sum in (5.1) at some positive integer value T_G^H. Furthermore, we need a criterion to stop recursive computation of matrices $\boldsymbol{G}_H[N]$. We adopt the following schemes in computing the approximation to matrix $\boldsymbol{G}_H$. First we set two parameters ϵ_G and ϵ_G^* such that $0 < \epsilon_G \le \epsilon_G^* \ll 1$. Then we choose the truncation index T_G^H in such a way that T_G^H is the minimal integer value which satisfies

$$\sum_{n=0}^{T_G^H} \gamma_n^H \ge 1 - \epsilon_G.$$

Further we stop the recursive computation of $\boldsymbol{G}_H[N]$ when the following stopping criterion is satisfied for some $N = N^*$:

$$\min_i[\boldsymbol{G}_H[N^*]\boldsymbol{e}]_i > 1 - \epsilon_G^*.$$

We then use the approximation $\boldsymbol{G}_H^*[N^*]$ to matrix $\boldsymbol{G}_H$.

In computing $\boldsymbol{G}_L^*[N]$, we need to truncate the infinite sum in (5.2) at some positive integer value T_G^L. We choose the truncation index T_G^L in such a way that T_G^L is the minimal value which satisfies

$$\sum_{n=0}^{T_G^L} \gamma_n^L \ge 1 - \epsilon_G.$$

We then use the approximation $\boldsymbol{G}_L^*[N^*]$ to matrix $\boldsymbol{G}_L$.

Next, we consider the computation of matrices $\boldsymbol{A}_k^H$. We need to truncate the infinite sum in (5.4) for each k and stop computing matrices $\boldsymbol{A}_k^H$ at some k. We follows the scheme in [28]. Namely, we first set a parameter ϵ_A^* $(0 < \epsilon_A^* \ll 1)$ and we choose the truncation index T_A in such a way that T_A is the minimal value which satisfies

$$\sum_{n=0}^{T_A} \gamma_n^H \ge 1 - \epsilon_A^*.$$

Then the approximation $\boldsymbol{A}_k^{H*}$ to matrix $\boldsymbol{A}_k^H$ is obtained as

$$\boldsymbol{A}_k^{H*} = \sum_{n=k}^{T_A} \gamma_n^H \boldsymbol{F}_n^{(k)}, \qquad k = 0, 1, \ldots, T_A,$$

and $\boldsymbol{A}_k^{H*} = \boldsymbol{0}$ for $k \ge T_A + 1$. Note that the approximation $\boldsymbol{A}_k^{H*}$ satisfies the following inequality [28]:

$$\min_i \left[\sum_{k=0}^{T_A} \boldsymbol{A}_k^{H*}\boldsymbol{e}\right]_i \ge 1 - \epsilon_A^*. \tag{5.14}$$

Since $\boldsymbol{A}_H(1) = \sum_{k=0}^{\infty} \boldsymbol{A}_k^H$ is stochastic, we scale up each row of matrices $\boldsymbol{A}_k^{H*}$ in such a way that matrix $\boldsymbol{A}_H(1)$ becomes stochastic.

Similarly, in computing the approximation $\boldsymbol{B}_k^{H*}$ to matrix $\boldsymbol{B}_k^H$, we follow the above scheme with the truncation index T_B in such a way that T_B is the minimal value which satisfies

$$\sum_{n=0}^{T_B} \gamma_n^L \geq 1 - \epsilon_A^*.$$

The rest is the same as in computing $\boldsymbol{A}_k^{H*}$.

Finally, we consider the computation of $\boldsymbol{G}_k^H$ and $\boldsymbol{G}_k^L$. In the algorithm provided in the preceding subsection, we need to truncate the infinite sums in (5.11) and (5.13), and stop computing for some k as well as N. For convenience, we slightly modify the algorithm as follows. First we set $T_G = \max(T_G^H, T_G^L)$. Then we compute $\boldsymbol{\Phi}_n^{(0)}[N]$ $(0 \leq n \leq T_G)$ in (5.12) iteratively for $N = 1, 2, \ldots$, and stop the computation when the following criterion is satisfied for some $N = N_0^*$

$$\max_{i,j} \left[\boldsymbol{G}_0^H[N_0^*] - \boldsymbol{G}_0^H[N_0^* - 1]\right]_{i,j} < \epsilon_G.$$

We use the approximation $\boldsymbol{G}_0^H[N_0^*]$ to matrix $\boldsymbol{G}_0^H$. And then we compute the approximation $\boldsymbol{G}_0^L[N_0^*]$ to matrix $\boldsymbol{G}_0^L$ using (5.13), where the infinite sum is truncated at T_G.

Now we suppose we have computed the approximation $\boldsymbol{G}_k^H[N_k^*]$ and $\boldsymbol{G}_k^L[N_k^*]$ for $0 \leq k \leq K-1$ $(K \geq 1)$. Then, we compute $\boldsymbol{\Phi}_n^{(K)}[N]$ $(0 \leq n \leq T_G)$ iteratively for $N = 1, 2, \ldots$ using

$$\begin{aligned} \boldsymbol{\Phi}_{n+1}^{(K)}[N] = \boldsymbol{\Phi}_n^{(K)}[N]\left[\boldsymbol{I} + \theta^{-1}\left(\boldsymbol{C} + \boldsymbol{D}_H \boldsymbol{G}_0^H[N_0^*]\right)\right] + \boldsymbol{\Phi}_n^{(0)}[N_0^*]\theta^{-1}\boldsymbol{D}_H\boldsymbol{G}_K^H[N-1] \\ + \sum_{m=1}^{K-1} \boldsymbol{\Phi}_n^{(m)}[N_m^*]\theta^{-1}\boldsymbol{D}_H\boldsymbol{G}_{K-m}^H[N_{K-m}^*] + \boldsymbol{\Phi}_n^{(K-1)}[N_{K-1}^*]\theta^{-1}\boldsymbol{D}_L, \end{aligned}$$

and we stop the computation when the following criterion is satisfied for some $N = N_K^*$

$$\max_{i,j} \left[\boldsymbol{G}_K^H[N_K^*] - \boldsymbol{G}_K^H[N_K^* - 1]\right]_{i,j} < \epsilon_G.$$

We use the approximation $\boldsymbol{G}_K^H[N_K^*]$ to matrix $\boldsymbol{G}_K^H$. And then we compute the approximation $\boldsymbol{G}_K^L[N_K^*]$ to matrix $\boldsymbol{G}_K^L$ using (5.13), where the infinite sum is truncated at T_G. We stop the computation of $\boldsymbol{G}_k^H$ and $\boldsymbol{G}_k^L$ for some $k = K^*$ when the following criteria are simultaneously satisfied

$$\min_i \left[\sum_{k=0}^{K^*} \boldsymbol{G}_k^H[N_k^*]\boldsymbol{e}\right]_i > 1 - \epsilon_G^*, \qquad \min_i \left[\sum_{k=0}^{K^*} \boldsymbol{G}_k^L[N_k^*]\boldsymbol{e}\right]_i > 1 - \epsilon_G^*.$$

And we set $\boldsymbol{G}_k^{H*} = \boldsymbol{G}_k^{L*} = \boldsymbol{0}$ for $k \geq K^* + 1$.

5.3. Numerical Examples

In this subsection, we show some numerical examples assuming the following. The service times of high (resp. low) priority customers are constant and equal to one (resp. two). The arrival process of high (resp. low) priority customers is an *MMPP* with states 1 to M_H (resp. M_L) [19]. The underlying Markov chains for high and low priority arrivals transit from a given state only to its adjacent states. State transition rates for high (resp.

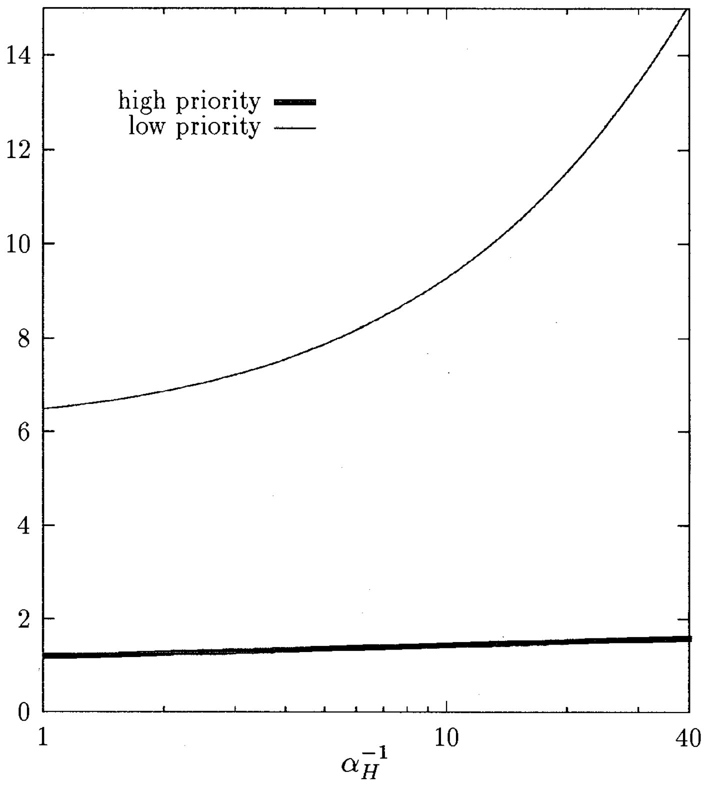

Fig.1 Mean Waiting Times as a Function of α_H^{-1}.

low) priority class do not depend on the current state of the underlying Markov chain and are denoted by α_H (resp. α_L). When the underlying Markov chain is in state j, high (resp. low) priority customers arrive to the system according to a Poisson process with density ja_H (resp. ja_L). In the remainder of this subsection, we assume $M_H = 3$, $M_L = 2$, $\alpha_L = 0.1$, $a_H = 0.25$, and $a_L = 0.1$. It is easy to show that the utilization factors are given by $\rho_H = 0.5$ and $\rho_L = 0.3$, and they are independent of the values of α_H and α_L. Note that there remains only one free parameter α_H. As α_H decreases, the sojourn time in each state of the underlying Markov chain for high priority class becomes longer, so that the correlation in high priority arrivals becomes higher.

Fig. 1 shows the mean waiting times as a function of α_H^{-1}. The mean waiting time of low priority customers increases with the increase of the correlation in high priority arrivals, while the mean waiting time of high priority customers is not affected so much. This phenomenon is explained as follows. In the above settings, when the underlying Markov chain of high priority arrivals is in state 3 and that of low priority arrivals is in state 2, the traffic intensities of high and low priority arrivals are 0.75 and 0.4, respectively. Thus in such a period, the overall traffic intensity is greater than one, i.e., the system is overloaded. As a result, arriving customers of low priority class are accumulated in the buffer during the overloaded period. Fig. 2 shows the probability mass function of the queue length distribution in the queue with the above settings. As we expect, the queue length tail of low priority customers is longer than that of high priority customers.

6. Concluding Remarks

In this paper we studied a nonpreemptive priority $MAP/G/1$ queue with two classes of customers, where the service times in each priority class may have a different distribution

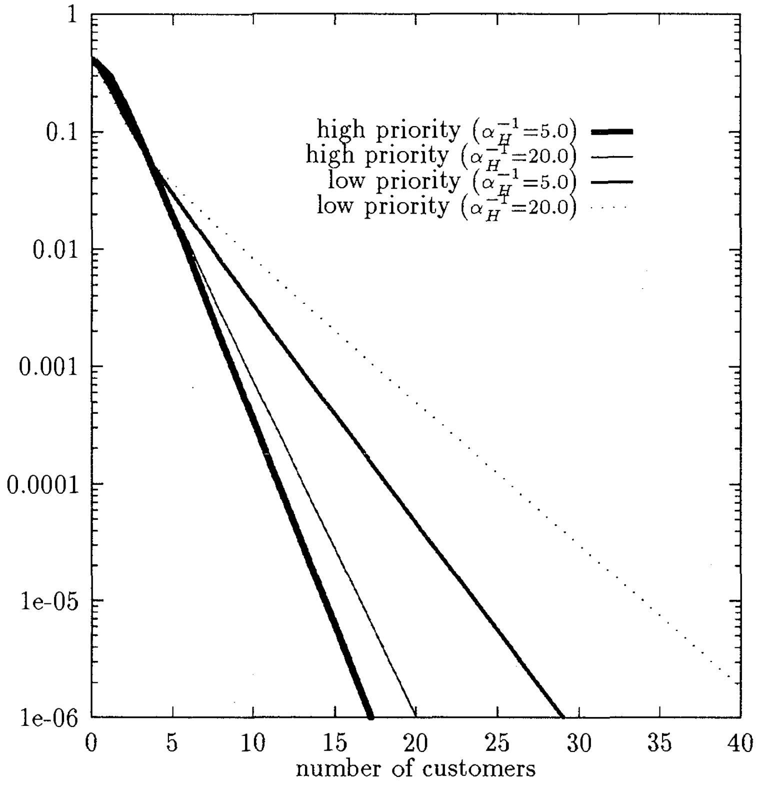

Fig.2 Queue Length Distribution

from one another. We derived the GF for the queue length and the LST for the waiting time in each class. Furthermore, as for the high priority class, we derived the recursive formulas to compute the mass function of the queue length, the moments of the queue length and the waiting time distributions. On the other hand, as for the low priority class, we derived the recursive formulas to compute the mass function of the queue length and the moments of the queue length distribution, as well as the mean waiting time. We also discussed the algorithmic implementation to compute the formulas in detail and provided some numerical results.

An algorithmic formula to compute the moments of the waiting time distribution of the low priority class still remains as an open research issue. The main difficulty lies in the computation of the high-order derivatives of $\boldsymbol{Q}_H^*(s)$ given in (4.2). A similar difficulty has already been recognized in computing the high-order derivatives of the fundamental matrix $\boldsymbol{G}(s)$ in the $MAP/G/1$ queue, which plays a central role in the matrix-analytic method [19]. Alternatively, we would utilize a numerical technique in [3], which enables us to obtain moments of a nonnegative distribution from its transform. Also, using a numerical inversion technique in [1], we would obtain the waiting time distribution in each priority class from its LST. Those numerical techniques, however, are beyond the scope of this paper.

Appendix (Definition of MAP [14])

We reproduce the definition of MAP from [14]. Consider a Markov chain on the state space $\{1, 2, \ldots, M+1\}$, where $\{1, 2, \ldots, M\}$ are transient states and $\{M+1\}$ is absorbing.

Absorption, starting from any state, is certain. The *MAP* is then defined as follows. Assume the Markov chain is in transient state i $(1 \le i \le M)$. The sojourn time in state i is exponentially distributed with parameter a_i. When the sojourn time has elapsed, there are two possibilities. With probability $p_{i,j}$ $(1 \le j \le M)$, the Markov chain enters the absorbing state with an arrival and is instantaneously restarted in the transient state j. With probability $q_{i,j}$ $(1 \le j \le M,\ j \ne i)$, the chain immediately enters the transient state j. Note that

$$\sum_{\substack{j=1\\ j\neq i}}^{M} q_{i,j} + \sum_{j=1}^{M} p_{i,j} = 1, \qquad \text{for } 1 \le i \le M.$$

Equivalently, if we define $\boldsymbol{C}$ and $\boldsymbol{D}$ as $M \times M$ matrices whose (i,j)th elements are denoted by $C_{i,j}$ and $D_{i,j}$, respectively, and for each i $(1 \le i \le M)$, we define $D_{i,j} = a_i p_{i,j}$ $(1 \le j \le M)$, $C_{i,j} = a_i q_{i,j}$ $(1 \le j \le M,\ j \ne i)$ and $C_{i,i} = -a_i$, then the elementary probability of an arrival in an infinitesimal interval of length dt which leaves the Markov chain in state j, given the Markov chain being in state i is $D_{i,j}dt$. Similarly, the elementary probability of a transition to state j without arrivals in an infinitesimal interval of length dt, given the Markov chain being in state i $(i \ne j)$ is $C_{i,j}dt$. Thus, the infinitesimal generator of the underlying Markov chain which governs the arrival process is given by $\boldsymbol{C} + \boldsymbol{D}$, and when a transition driven by $\boldsymbol{D}$ happens, an arrival occurs.

References

[1] Abate, J. and Whitt, W.: The Fourier-series method for inverting transforms of probability distributions. *Queueing Sys.*, vol.10 (1992) 5–88.

[2] Asmussen, S. and Koole, G.: Marked point processes as limits of Markovian arrival streams. *J. Appl. Prob.*, vol.30 (1993) 365–372.

[3] Choudhury, G.L. and Lucantoni, D.: Numerical computation of large number of moments with application to asymptotic analysis. *Opns. Res.*, to appear.

[4] Cooper, R.B.: *Introduction to Queueing Theory, 2nd ed.* Elsevier, New York (1981).

[5] Graham, A.: *Kronecker Products and Matrix Calculus with Applications.* Ellis Horwood, Chichester (1981).

[6] Grünenfelder, R., Cosmas, J.P., Manthrope, S. and Odinma-Okafor, A.: Characterization of video codecs as autoregressive moving average processes and related queueing performance. *IEEE J. Selected Areas in Comm.*, vol.9 (1991) 284–293.

[7] Gün, L.: Experimental results on matrix-analytical solution techniques – Extensions and comparisons. *Stoch. Mod.*, vol.5 (1989) 669–682.

[8] Hashida, O. and Takahashi, Y.: A discrete-time priority queue with switched batch Bernoulli process inputs and constant service time. *Proc. of ITC-13*, Copenhagen, Denmark (1991) 521–526.

[9] Heffes, H. and Lucantoni, D.M.: A Markov modulated characterization of packetized voice and data traffic and related statistical multiplexer performance. *IEEE J. Selected Areas in Comm.*, vol.4 (1986) 856–868.

[10] Jaiswal, N.K.: *Priority Queues.* Academic Press, New York (1968).

[11] Khamisy, A. and Sidi, M.: Discrete-time priority queueing systems with two-state Markov modulated arrival processes. *Proc. of IEEE INFOCOM '91*, Bal Harbour, Florida (1991) 1456–1463.

[12] Latouche, G., Jacobs, P.A. and Gaver, D.P.: Finite Markov chain models skip-free in one direction. *Naval Res. Logist. Quart.*, vol.31 (1984) 571–588.

[13] Lucantoni, D.M. and Ramaswami, V.: Efficient algorithms for solving the non-linear matrix equations arising in phase type queues. *Stoch. Mod.*, vol.1 (1985) 29–51.

[14] Lucantoni, D.M., Meier-Hellstern, K.S. and Neuts, M.F.: A single-server queue with server vacations and a class of non-renewal arrival processes. *Adv. Appl. Prob.*, vol.22 (1990) 676–705.

[15] Lucantoni, D.M.: New results on the single server queue with a batch Markovian arrival process. *Stoch. Mod.*, vol.7 (1991) 1–46.

[16] Lucantoni, D.M.: The BMAP/G/1 queue: A tutorial. *Models and Techniques for Performance Evaluation of Computer and Communication Systems*, L. Donatiello and R. Nelson (eds.) Springer Verlag (1993) 330–358.

[17] Machihara, F.: On the queue with PH-Markov renewal preemption. *J. Opns. Res. Soc. J.*, vol.36 (1993) 13–28.

[18] Miller, R.G.: Priority queues. *Ann. Math. Stat.*, vol.31 (1960) 86–103.

[19] Neuts, M.F.: *Structured Stochastic Matrices of $M/G/1$ Type and Their Applications.* Marcel Dekker, New York (1989).

[20] Ramaswami, V.: The N/G/1 queue and its detailed analysis. *Adv. Appl. Prob.*, vol.12 (1980) 222–261.

[21] Ramaswami, V.: Stable recursion for the steady state vector for Markov chains of $M/G/1$ type. *Stoch. Mod.*, vol.4 (1988) 183–188.

[22] Ren, J.F., Mark, J. and Wong, J.W.: Analysis of integrated services on a switch with output buffering. Preprint. Department of Electrical and Computer Engineering, University of Waterloo, Waterloo, Ontario, Canada (1991).

[23] Sugahara, A., Takine, T., Takahashi, Y. and Hasegawa, T.: Analysis of a nonpreemptive priority queue with *SPP* arrivals of high class. *Perfor. Eval.*, vol.21 (1995) 215–238.

[24] Takács, L.: Priority Queues. *Opns. Res.*, vol.12 (1964) 63–74.

[25] Takagi, H.: *Queueing Analysis: A Foundation of Performance Evaluation Volume 1 Vacation and Priority Systems, Part 1.* North-Holland, Amsterdam (1991).

[26] Takine, T. and Hasegawa, T.: The workload in the $MAP/G/1$ queue with state-dependent services: its application to a queue with preemptive resume priority. *Stoch. Mod.*, vol.10 (1994) 183–204.

[27] Takine, T., Sengupta, B. and Hasegawa, T.: An analysis of a discrete-time queue for broadband ISDN with priorities among traffic classes. *IEEE Trans. Comm.*, vol.42 (1994) 1837–1845.

[28] Takine, T., Matsumoto, Y., Suda, T. and Hasegawa, T.: Mean waiting times in nonpreemptive priority queues with Markovian arrival and i.i.d. service processes. *Perfor. Eval.*, vol.20 (1994) 131–149.

[29] Takine, T. and Takahashi, Y.: On the relationship between queue lengths at a random instant and at a departure in the stationary queue with BMAP arrivals. submitted for publication.

Tetsuya Takine

Department of Information Systems Engineering
Faculty of Engineering
Osaka University
2-1 Yamadaoka, Suita,
Osaka 565, Japan
e-mail: takine@ise.eng.osaka-u.ac.jp