数理解析研究所講究録 1351 巻 2004 年 63-79

On the asymptotic analysis of discrete-time queues

南山大学・数理情報学部情報通信学科　石崎　文雄 (Fumio Ishizaki)
Department of Information and Telecommunication Engineering,
Faculty of Mathematical Sciences and Information Engineering,
Nanzan University

Abstract

This paper derives a formula to calculate the asymptotic loss probability in a discrete-time finite-buffer queue with correlated arrivals and service interruptions. To derive the formula, we use an exact relation which holds under some assumptions between the exact loss probability in the finite-buffer queue and the queue length distribution in the corresponding infinite-buffer queue. The exact relation shown in this paper is considered as a generalization of the exact relations which have been established.

1 Introduction

In packet/cell networks, the estimation of packet/cell loss probability has been considered as one of the most important issues in connection admission and congestion controls. For this reason, considerable attentions have been paid to the analysis of queueing systems such as DBMAP/D/1/K queues (see, e.g., [3, 8, 19, 23] and references therein). In packet/cell networks, the arrival process at statistical multiplexer is usually a superposition of sources which typically generate time-correlated and bursty traffic due to their origin (e.g., periodic sampling of voice traffic or MPEG encoded real-time video traffic) or traffic shaping. On the other hand, the service process at statistical multiplexer may be subject to a scheduling mechanism (e.g., round-robin scheduling). For instance, the service for packet/cell transmission may be available every R slots, where R is a positive integer. The dynamics of such a multiplexer is modeled as a discrete-time single-server finite-buffer queue with correlated arrivals and service interruptions. Further, the dynamics of such a finite-buffer queue with correlated arrivals and service interruptions may be described as a finite-state Markov chain. The loss probability is then obtained from the stationary distribution of the Markov chain. Although we can directly apply the standard algorithm (e.g., as shown in [7, 18, 24]) to compute the stationary distribution of the Markov chain, the following difficulty arises in the computation. Usually, the number of states to describe the dynamics of multiplexer becomes prohibitively large. This makes the computation with enough accuracy very difficult. In addition, the standard algorithms such as block Gaussian elimination include subtractions and this often makes the algorithms unstable, especially, when the size of the matrices is large. Thus numerical algorithms to estimate the loss probability efficiently and stably should be developed. For this requirement, we derive a formula which calculates the asymptotic loss probability in a finite-buffer queue. This formula can estimate the loss probability more easily than the standard algorithm when the size of the matrices is large.

In this paper, discrete-time single-server queueing systems with correlated arrivals and service interruptions are studied. Queueing systems with service interruptions have wide applications to manufacturing, computer and telecommunication systems where the server is subject to breakdown (see, e.g., [16, 4, 26] and references therein). In the queueing systems, the arrival process is governed by a Markov chain and also the service process is governed by a Markov chain. More precisely, the number of arrivals in a slot depends on the state of the underlying Markov chain for the arrival process in the slot and the availability of the server is determined by the state of the underlying Markov chain for the service process. The service time of customer is geometrically distributed.

To derive the formula to compute the asymptotic loss probability, we use an exact relation holding between the exact loss probability in the finite-buffer queue and the queue length distribution in the corresponding infinite-buffer queue. Further, utilizing the fact that the queue length distribution in the infinite-buffer queue has a simple geometric asymptotic form, the formula estimates the *asymptotic* loss probability [14, 11] based on the asymptotic queueing analysis of infinite-buffer queues [1, 6, 13, 23]. Several researchers have studied similar exact relations holding between the loss probability in a finite-buffer queue and the queue length distribution in the corresponding infinite-buffer queue. Kang et al. [17] have considered discrete-time queueing systems where the arrival process is a superposition of Bernoulli sources and the service is available every R slots where R is a positive integer. For such queueing systems, they have established an exact relation holding between the loss probability in a finite-buffer queue and the queue length distribution in the corresponding infinite-buffer queue. Ishizaki and Takine [14] have considered discrete-time queueing systems where the arrival process is similar to (but a little restrictive compared to) the arrival process considered in this paper and the service is always available. For such queueing systems, they have established a proportional relation between the stationary queue length distribution in the finite-buffer queue and that in the corresponding infinite-buffer queue. Using this proportional relation, they have obtained an exact relation holding between the loss probability in a finite-buffer queue and the queue length distribution in the corresponding infinite-buffer queue. Ishizaki [11] has considered discrete-time queueing systems where the arrival process is similar to the arrival process considered in this paper and the service is available every R slots. A similar exact relation has been obtained for such queueing systems. In this paper, we study an exact relation in more general setting than the settings in [11, 14, 17]. The exact relation established in this paper is considered as a generalization and integration of those exact relations shown in [11, 14, 17].

The remainder of this paper is organized as follows. In Section 2, we consider an M/G/1-type Markov chain with some regenerative structure and a corresponding truncated Markov chain which is obtained from the M/G/1-type Markov chain by limiting its maximum level to K where K is a nonnegative integer. Under some assumptions, we derive a preliminary result (Theorem 1) for a proportional relation holding between the stationary distribution of the M/G/1-type Markov chain and that of the corresponding truncated Markov chain. The result is interpreted as a generalization of the proportional relation [14] between the stationary queue length distribution in the finite-buffer queue and that in the corresponding infinite-buffer queue. Section 3 shows a discrete-time single-server infinite-buffer queue with correlated arrivals and service interruptions whose dynamics is described as the M/G/1-type Markov chain considered in Section 2 and its corresponding finite-buffer queue whose dynamics is described as the corresponding truncated Markov chain. In Section 4, using the preliminary result derived in Section3, we establish an exact relation (Theorem 2) holding between the loss probability in a finite-buffer queue and the stationary queue length distribution in the corresponding infinite-buffer queue. Using Theorem 2 in the geometric expression of the asymptotic tail distribution in the infinite-buffer queue, Section 5 derives a formula (Theorem 3) to compute the asymptotic loss probability in the finite-buffer queue.

2 Preliminary result

In this section, we consider an M/G/1-type Markov chain with some regenerative structure and a corresponding truncated Markov chain which is obtained from the M/G/1-type Markov chain by limiting its maximum level to K where K is a nonnegative integer. We then provide a preliminary result for a proportional relation holding between the stationary distribution of the M/G/1-type Markov chain and that of the corresponding truncated Markov chain.

Throughout this paper, we use the following notation. For any matrix $\boldsymbol{C}$, $\left[\boldsymbol{C}\right]_{i,j}$ denotes the (i,j)th element of the matrix $\boldsymbol{C}$, and the row and column index numbers of any matrix are labeled from 0. Further, for any positive integer i and j, $\boldsymbol{O}_i$ and $\boldsymbol{I}_i$ denote the $i \times i$ zero matrix and the $i \times i$ identity matrix, respectively. Similarly, for any vector $\boldsymbol{c}$, $\left[\boldsymbol{c}\right]_i$ denotes the ith element of the vector $\boldsymbol{c}$, and the row or column index numbers of any vector are labeled from 0.

We consider a discrete-time Markov chain $\{(X_n, V_n)\}_{n=0}^{\infty}$ whose state space is $\mathcal{S} = \{(k,l) \mid k = 0,1,\ldots; l = 0,\ldots,L\}$, where L is a nonnegative integer. We assume that the Markov chain $\{(X_n, V_n)\}_{n=0}^{\infty}$ is an M/G/1-type Markov chain and its down-shift matrices have some special structures shown in the assumption below. Given that sequences of $(L+1) \times (L+1)$ matrices $\{\boldsymbol{A}_i\}$ $(i = 0,1,\ldots)$ and $\{\boldsymbol{B}_i\}$ $(i = 0,1,\ldots)$, we consider the Markov chain whose transition probability matrix $\boldsymbol{Q}^{(\infty)}$ has the following block structure:

$$\boldsymbol{Q}^{(\infty)} = \begin{bmatrix} \boldsymbol{B}_0 & \boldsymbol{B}_1 & \boldsymbol{B}_2 & \boldsymbol{B}_3 & \cdots \\ \boldsymbol{A}_0 & \boldsymbol{A}_1 & \boldsymbol{A}_2 & \boldsymbol{A}_3 & \cdots \\ \boldsymbol{O} & \boldsymbol{A}_0 & \boldsymbol{A}_1 & \boldsymbol{A}_2 & \cdots \\ \boldsymbol{O} & \boldsymbol{O} & \boldsymbol{A}_0 & \boldsymbol{A}_1 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix}, \tag{1}$$

where $\boldsymbol{O}$ denotes the $(L+1) \times (L+1)$ zero matrix. We also consider a discrete-time Markov chain $\{(Y_n, V_n)\}_{n=0}^{\infty}$ which is obtained from the Markov chain $\{(X_n, V_n)\}_{n=0}^{\infty}$ by limiting its maximum level to K where K is a nonnegative integer. In other words, its state space is $\mathcal{S} = \{(k, m) \mid k = 0, \ldots, K; m = 0, \ldots, L\}$ and its transition probability matrix $\boldsymbol{Q}^{(K)}$ has the following block structure:

$$\boldsymbol{Q}^{(K)} = \begin{bmatrix} \boldsymbol{B}_0 & \boldsymbol{B}_1 & \boldsymbol{B}_2 & \boldsymbol{B}_3 & \cdots & \boldsymbol{B}_{K-1} & \boldsymbol{B}_K^* \\ \boldsymbol{A}_0 & \boldsymbol{A}_1 & \boldsymbol{A}_2 & \boldsymbol{A}_3 & \cdots & \boldsymbol{A}_{K-1} & \boldsymbol{A}_K^* \\ \boldsymbol{O} & \boldsymbol{A}_0 & \boldsymbol{A}_1 & \boldsymbol{A}_2 & \cdots & \boldsymbol{A}_{K-2} & \boldsymbol{A}_{K-1}^* \\ \boldsymbol{O} & \boldsymbol{O} & \boldsymbol{A}_0 & \boldsymbol{A}_1 & \cdots & \boldsymbol{A}_{K-3} & \boldsymbol{A}_{K-2}^* \\ \vdots & & \ddots & \ddots & & & \vdots \\ \vdots & & & \ddots & \ddots & & \vdots \\ \boldsymbol{O} & \cdots & \cdots & \cdots & \boldsymbol{O} & \boldsymbol{A}_0 & \boldsymbol{A}_1^* \end{bmatrix}, \tag{2}$$

where $\boldsymbol{A}_k^* = \sum_{m=k}^{\infty} \boldsymbol{A}_m$, $\boldsymbol{B}_k^* = \sum_{m=k}^{\infty} \boldsymbol{B}_m$, and $\boldsymbol{O}$ denotes the $(L+1) \times (L+1)$ zero matrix.

For the structure of the down-shift matrix $\boldsymbol{A}_0$, the following assumption is made.

Assumption 1 There exists a $1 \times (L+1)$ probability vector $\boldsymbol{a}$ such that

$$\boldsymbol{A}_0 = \boldsymbol{A}_0 \boldsymbol{e} \boldsymbol{a},$$

where $\boldsymbol{e}$ is an $(L+1) \times 1$ column vector with unit elements.

Note that Assumption 1 is equivalent to the following statement: For $i = 1, \ldots, K$, $\mathrm{P}(V_n = k \mid Y_{n-1} = i, Y_n = i-1, V_{n-1} = j)$ is independent of j (or $\mathrm{P}(V_n = k \mid X_{n-1} = i, X_n = i-1, V_{n-1} = j)$ is independent of j for $i = 1, 2, \ldots$). For $i = 1, \ldots, K$, we then have $[\boldsymbol{a}]_k = \mathrm{P}(V_n = k \mid Y_{n-1} = i, Y_n = i-1, V_{n-1} = j)$ (or we then have $[\boldsymbol{a}]_k = \mathrm{P}(V_n = k \mid X_{n-1} = i, X_n = i-1, V_{n-1} = j)$ for $i = 1, 2, \ldots$) where $[\boldsymbol{a}]_k$ denotes the kth element of $\boldsymbol{a}$. In other words, Assumption 1 means that when the down-ward shift of the level $\{Y_n\}$ (or $\{X_n\}$) occurs, $\{V_n\}$ regenerates. We also made the following assumption.

Assumption 2 The Markov chains $\{(X_n, V_n)\}$ and $\{(Y_n, V_n)\}$ are irreducible and positive recurrent.

Under Assumption 2, the stationary distribution of the Markov chain $\{(X_n, V_n)\}$ and the stationary distribution of the Markov chain $\{(Y_n, V_n)\}$ exist, and they are uniquely determined [2].

Let $\boldsymbol{x}$ and $\boldsymbol{y}$ denote the stationary distribution of the Markov chain $\{(X_n, V_n)\}$ and that of the Markov chain $\{(Y_n, V_n)\}$, respectively. We then have

$$\boldsymbol{x} = \boldsymbol{x}\boldsymbol{Q}^{(\infty)}, \qquad \boldsymbol{y} = \boldsymbol{y}\boldsymbol{Q}^{(K)}, \tag{3}$$

where

$$\boldsymbol{x} = (\boldsymbol{x}_0, \boldsymbol{x}_1, \ldots), \qquad \boldsymbol{y} = (\boldsymbol{y}_0, \boldsymbol{y}_1, \ldots, \boldsymbol{y}_K),$$

$\boldsymbol{x}_j$ is a $1 \times (L+1)$ vector whose lth element $[\boldsymbol{x}_j]_l$ is given by $[\boldsymbol{x}_j]_l = \mathrm{P}(X_n = j, V_n = l)$, and $\boldsymbol{y}_j$ is a $1 \times (L+1)$ vector whose lth element $[\boldsymbol{y}_j]_l$ is given by $[\boldsymbol{y}_j]_l = \mathrm{P}(Y_n = j, V_n = l)$.

The following lemma for the stationary distribution $\boldsymbol{x}$ is readily obtained from the results shown in [21, 22].

Lemma 1 *Under Assumptions 1 and 2, the stationary distribution $\boldsymbol{x}$ satisfies the equations*

$$\boldsymbol{x}_0 = \boldsymbol{x}_0 \overline{\boldsymbol{B}}_0, \tag{4}$$

$$\boldsymbol{x}_i = \left(\boldsymbol{x}_0 \overline{\boldsymbol{B}}_i + \sum_{k=1}^{i-1} \boldsymbol{x}_k \overline{\boldsymbol{A}}_{i-k+1} \right) \left(\boldsymbol{I} - \overline{\boldsymbol{A}}_1 \right)^{-1} \qquad i = 1, 2, \ldots, \tag{5}$$

where for $\nu = 0, \ldots, K-1$, we define $\overline{\boldsymbol{A}}_\nu$ as

$$\overline{\boldsymbol{A}}_\nu = \boldsymbol{A}_\nu + \boldsymbol{A}^*_{\nu+1}\boldsymbol{e}\boldsymbol{a}, \qquad \overline{\boldsymbol{B}}_\nu = \boldsymbol{B}_\nu + \boldsymbol{B}^*_{\nu+1}\boldsymbol{e}\boldsymbol{a},$$

and $\boldsymbol{e}$ is a column vector with unit elements.

Proof: Since $\{(X_n, V_n)\}$ is a Markov chain of M/G/1 type, we have [22]

$$\boldsymbol{x}_k = \left(\boldsymbol{x}_0 \tilde{\boldsymbol{B}}_k + \sum_{j=1}^{k-1} \boldsymbol{x}_j \tilde{\boldsymbol{A}}_{k+1-j} \right) \left(\boldsymbol{I} - \tilde{\boldsymbol{A}}_1 \right)^{-1}, \qquad k = 1, 2, \ldots, \tag{6}$$

where $\tilde{\boldsymbol{A}}_k$ $(k = 1, 2, \ldots)$ and $\tilde{\boldsymbol{B}}_k$ $(k = 1, 2, \ldots)$ are substochastic matrices, which are given by

$$\tilde{\boldsymbol{A}}_k = \sum_{j=k}^{\infty} \boldsymbol{A}_j \boldsymbol{G}^{j-k}, \qquad \tilde{\boldsymbol{B}}_k = \sum_{j=k}^{\infty} \boldsymbol{B}_j \boldsymbol{G}^{j-k},$$

for an $(L+1) \times (L+1)$ stochastic matrix $\boldsymbol{G}$ whose (i,j)th element denotes the conditional probability that the Markov chain $\{(X_n, V_n)\}$ starting in state $(l+1, i)$ (for any level l) will reach level l eventually and end up in phase j when it reaches level l. On the other hand, from Assumption 1, we see that

$$\boldsymbol{G} = \boldsymbol{e}\boldsymbol{a}. \tag{7}$$

From (6) and (7), we see that (5) holds.

Let $\boldsymbol{K}$ denote an $(L+1)\times(L+1)$ stochastic matrix whose (i,j)th element denotes the conditional probability that the Markov chain $\{(X_n, V_n)\}$ starting in state $(0,i)$ will reach level 0 eventually and end up in phase j when it reaches level 0. From (7), we then have [21]

$$\boldsymbol{K} = \sum_{k=0}^{\infty} \boldsymbol{B}_k \boldsymbol{G}^k = \boldsymbol{B}_0 + \boldsymbol{B}_1^* \boldsymbol{e}\boldsymbol{a} = \overline{\boldsymbol{B}}_0. \tag{8}$$

Since $\boldsymbol{x}_0$ is an invariant vector for $\boldsymbol{K}$ [21], we obtain

$$\boldsymbol{x}_0 = \boldsymbol{x}_0 \boldsymbol{K}. \tag{9}$$

From (8) and (9), we see that (4) holds. ■

The following lemma for the stationary distribution $\boldsymbol{y}$ is readily obtained as a special case of the result shown in [10].

Lemma 2 *Under Assumptions 1 and 2, the stationary distribution $\boldsymbol{y}$ is determined by the equations*

$$\boldsymbol{y}_0 = \boldsymbol{y}_0 \overline{\boldsymbol{B}}_0, \tag{10}$$

$$\boldsymbol{y}_i = \left(\boldsymbol{y}_0 \overline{\boldsymbol{B}}_i + \sum_{k=1}^{i-1} \boldsymbol{y}_k \overline{\boldsymbol{A}}_{i-k+1}\right)\left(\boldsymbol{I} - \overline{\boldsymbol{A}}_1\right)^{-1} \qquad i = 1, \ldots, K-1, \tag{11}$$

$$\boldsymbol{y}_K = \left(\boldsymbol{y}_0 \boldsymbol{B}_K + \sum_{k=1}^{K-1} \boldsymbol{y}_k \boldsymbol{A}^*_{K-k+1}\right)\left(\boldsymbol{I} - \boldsymbol{A}^*_1\right)^{-1}, \tag{12}$$

$$\sum_{n=0}^{K} \boldsymbol{y}_n \boldsymbol{e} = 1,$$

where for $\nu = 0, \ldots, K-1$, we define $\overline{\boldsymbol{A}}_\nu$ as

$$\overline{\boldsymbol{A}}_\nu = \boldsymbol{A}_\nu + \boldsymbol{A}^*_{\nu+1}\boldsymbol{e}\boldsymbol{a}, \qquad \overline{\boldsymbol{B}}_\nu = \boldsymbol{B}_\nu + \boldsymbol{B}^*_{\nu+1}\boldsymbol{e}\boldsymbol{a},$$

and $\boldsymbol{e}$ is a column vector with unit elements.

The following theorem, which establishes a proportional relation holding between the stationary distribution of the M/G/1-type Markov chain and that of the corresponding truncated Markov chain, is a direct conclusion of Lemmas 1 and 2.

Theorem 1 *Under Assumptions 1 and 2, there exists a constant c such that*

$$\boldsymbol{y}_i = c\boldsymbol{x}_i, \qquad i = 0, 1, \ldots, K-1. \tag{13}$$

In addition, the proportional constant c in (13) can be expressed as

$$c = \frac{\boldsymbol{\pi} \boldsymbol{A}_0 \boldsymbol{e}}{\sum_{i=0}^{K} \boldsymbol{x}_i \boldsymbol{A}_0 \boldsymbol{e}}, \tag{14}$$

where $\boldsymbol{\pi}$ is a $1 \times (M+1)$ vector whose lth element is given by $[\boldsymbol{\pi}]_l = \mathrm{P}(V_n = j)$.

Proof: First we recursively show that (13) holds. Recall that from (4) and (11), both $\boldsymbol{x}_0$ and $\boldsymbol{y}_0$ are invariant vectors for the stochastic matrix $\boldsymbol{K} = \overline{\boldsymbol{B}}_0$. Since the Markov chains $\{(X_n, V_n)\}$ and $\{(Y_n, V_n)\}$ are irreducible and positive recurrent from Assumption 1, there exists some constant satisfying $\boldsymbol{y}_0 = c\boldsymbol{x}_0$. We thus see that (13) holds for $i = 0$. Suppose that (13) holds for some $k-1$ ($k \in \{1, \ldots, K-1\}$) and $i = 0, \ldots, k-1$. Then, since (5) and (11) are identical recursions, we see that $\boldsymbol{y}_k = c\boldsymbol{x}_k$ is satisfied. We therefore show that there exists a constant c such that $\boldsymbol{y}_i = c\boldsymbol{x}_i$ for $i = 0, 1, \ldots, K-1$.

Next we show (14) along with similar lines of the proof shown in [14]. From (3), we have

$$\boldsymbol{y}_K = \boldsymbol{y}_0 \boldsymbol{B}^*_{K_1} + \sum_{i=1}^{K} \boldsymbol{y}_i \boldsymbol{A}^*_{K-i+1}. \tag{15}$$

Using (4) and noting $\sum_{i=0}^{K} \boldsymbol{y}_i = \boldsymbol{\pi}$, we rewrite (15) to

$$(\boldsymbol{\pi} - c\sum_{i=1}^{K} \boldsymbol{x}_i)\boldsymbol{e} = c\boldsymbol{x}_0 \boldsymbol{B}^*_K \boldsymbol{e} + c\sum_{i=1}^{K-1} \boldsymbol{x}_i \boldsymbol{A}^*_{K-i+1}\boldsymbol{e} + (\boldsymbol{\pi} - c\sum_{i=0}^{K-1} \boldsymbol{x}_i)\boldsymbol{A}^*_1 \boldsymbol{e}, \tag{16}$$

where c is a constant appearing in (13). From (16), it follows that

$$\boldsymbol{\pi}(\boldsymbol{I} - \boldsymbol{A}^*_1)\boldsymbol{e} = c\left[\boldsymbol{x}_0 \boldsymbol{B}^*_K + \sum_{i=1}^{K-1} \boldsymbol{x}_i \boldsymbol{A}^*_{K-i+1} + \sum_{i=0}^{K-1} \boldsymbol{x}_i (\boldsymbol{I} - \boldsymbol{A}^*_1)\right] \boldsymbol{e}. \tag{17}$$

Note that we have

$$(\boldsymbol{I} - \boldsymbol{A}^*_1)\boldsymbol{e} = \boldsymbol{A}_0 \boldsymbol{e}. \tag{18}$$

Also note that the following equilibrium equation holds:

$$\boldsymbol{x}_0 \boldsymbol{B}^*_K + \sum_{i=1}^{K-1} \boldsymbol{x}_i \boldsymbol{A}^*_{K-i+1} = \boldsymbol{x}_K \boldsymbol{A}_0, \tag{19}$$

where the left hand side of (19) denotes the total flow from a macro-state which composes of the states that the level is less than K into a macro-state which composes of the states that the level is greater than or equal to K, and the right hand side of (19) denotes the total flow from the latter macro-state into the former macro-state. Using (18) and (19) in (17), we obtain

$$\boldsymbol{\pi}\boldsymbol{A}_0 \boldsymbol{e} = c\sum_{i=0}^{K} \boldsymbol{x}_i \boldsymbol{A}_0 \boldsymbol{e}. \tag{20}$$

Note here that $\sum_{i=0}^{K} \boldsymbol{x}_i \boldsymbol{A}_0 \boldsymbol{e} > 0$ under Assumption 1. Thus, from (20), we derive (14). ∎

3 Queueing model

In this section, we consider two queueing models, i.e., a discrete-time finite-buffer queue with correlated arrivals and service interruptions and the corresponding infinite-buffer queue. The queueing models are considered as an extension of ones considered in [14, 11]. Both queueing systems are identical except for the buffer capacity. The dynamics of the infinite-buffer queue and that of the finite-buffer queue are described by the Markov chains $\{(X_n, V_n)\}$ and $\{(Y_n, V_n)\}$ which are defined in the previous section, respectively.

We briefly describe the queueing systems below. Time is slotted and the slot length is equal to a unit time. The arrival of batch occurs at the beginning of slots immediately after departures (i.e., early arrival model [5, 25]). The service time of customer is i.i.d. (independent and identically distributed) and it follows a geometric distribution with mean $1/\gamma$ ($\gamma > 0$). The service is sometimes interrupted according to a stochastic sequence. The service of customer starts at the beginning of a slot and ends at the end of the slot (i.e., on slot boundaries). the finite-buffer queue accommodates at most K customers including the one in service. Thus, if m ($m \geq K - k + 1$) customers arrive to find k customers (including the one in service) in the system, only $K - k$ customers are accommodated in the system, and the remaining $m - (K - k)$ customers are discarded. On the other hand, the infinite-buffer queue accommodates all arriving customers and no customers are discarded.

We now describe the queueing systems in more detail. We begin with the description of the service process. To describe the service process, we introduce a Markov chain. Let $\{S_n\}_{n\in\mathbf{Z}_+}$ denote a Markov chain on $\mathcal{R} = \{0, \ldots, R\}$ where R is a positive integer. The service is available in the nth slot if and only if $S_n = 0$. We call the Markov chain $\{S_n\}$ the underlying Markov chain for the service process. We assume that the underlying Markov chain for the service process is stationary and ergodic. We next describe the arrival process. Let $\{A_n\}_{n\in\mathbf{Z}}$ denote a stochastic sequence where A_n represents the number of arrivals in the nth slot. We assume that $\{A_n\}_{n\in\mathbf{Z}}$ is governed by a Markov chain $\{P_n\}_{n\in\mathbf{Z}_+}$ on $\mathcal{M} = \{0, \ldots, M\}$ where M denotes a nonnegative integer. More precisely, we assume that given P_n, A_n is conditionally independent of all other random variables. We call the Markov chain $\{P_n\}$ the underlying Markov chain for the arrival process. We assume that the underlying Markov chain for the arrival process is stationary and ergodic. We now consider the queueing processes. Let $\{X_n\}_{n\in\mathbf{Z}_+}$ and $\{Y_n\}_{n\in\mathbf{Z}_+}$ denote a stochastic sequence representing the queue length (including a customer in service) in the infinite-buffer queue and that in the finite-buffer queue, respectively. Let $\{D_n\}_{n\in\mathbf{Z}_+}$ denote a Bernoulli sequence on $\{0, 1\}$ where $\mathrm{P}(D_n = 1) = \gamma$ and $\mathrm{P}(D_n = 0) = 1 - \gamma$ for $n \in \mathbb{Z}_+$. We assume that $\{S_n\}$, $\{P_n\}$ and $\{D_n\}$ are independent with each other. The queueing processes $\{X_n\}$ and $\{Y_n\}$ evolve according to the following recursions with initial queue length X_0 and Y_0:

$$X_{n+1} = (X_n - 1_{\{S_n=0\}} D_n)^+ + A_{n+1},$$

$$Y_{n+1} = \min[(Y_n - 1_{\{S_n=0\}} D_n)^+ + A_{n+1}, K],$$

where $(\cdot)^+ = \min(\cdot, 0)$ and 1 denotes the indicator function. Let Z_n ($n \in \mathbb{Z}$) denote a random variable representing the number of lost customers in the nth slot in the finite-buffer queue. Z_n ($n \in \mathbb{Z}$) is given by

$$Z_n = ((Y_{n-1} - 1_{\{S_{n-1}=0\}} D_{n-1})^+ + A_n - K)^+.$$

Now we describe a stochastic setting for the arrival and service processes. First, we describe the stochastic setting for the arrival process. Let $\hat{\boldsymbol{U}}$ denote the transition matrix of the underlying Markov chain for the arrival process, i.e., $[\hat{\boldsymbol{U}}]_{i,j} = \mathrm{P}(P_{n+1} = j \mid P_n = i)$ for $i, j \in \mathcal{M}$. Let $\hat{\boldsymbol{\pi}}$ denote the stationary vector of the underlying Markov chain for the arrival process, i.e., $[\hat{\boldsymbol{\pi}}]_i = \mathrm{P}(P_n = i)$. $\hat{\boldsymbol{\pi}}$ then satisfies $\hat{\boldsymbol{\pi}} = \hat{\boldsymbol{\pi}}\hat{\boldsymbol{U}}$ and $\hat{\boldsymbol{\pi}}\boldsymbol{e} = 1$. We denote by $\hat{a}_j(k)$ the conditional probability that k customers arrive given that the underlying Markov chain is in state j:

$$\hat{a}_j(k) = \mathrm{P}(A_n = k \mid P_n = j), \qquad j \in \mathcal{M}, \quad k = 0, 1, 2, \ldots.$$

Let $\hat{A}_{i,j}(k)$ denote the conditional joint probability of 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 underlying Markov chain was in state i in the nth slot. Namely,

$$\hat{A}_{i,j}(k) = \mathrm{P}(A_{n+1} = k, P_{n+1} = j \mid P_n = i) = \hat{a}_j(k)[\hat{\boldsymbol{U}}]_{i,j}, \qquad i, j \in \mathcal{M}. \tag{21}$$

Let $\hat{\boldsymbol{A}}_k$ and $\hat{\boldsymbol{A}}_k^*$ ($k = 0, 1, \ldots$) denote $(M+1) \times (M+1)$ matrices whose (i,j)th elements are given by $\hat{A}_{i,j}(k)$ and $\sum_{m=k}^{\infty} \hat{A}_{i,j}(m)$, respectively. Note that $\hat{\boldsymbol{A}}_k$ (resp. $\hat{\boldsymbol{A}}_k^*$) represents the transition matrix of the underlying Markov chain when k customers (resp. more than or equal to k customers) arrive at the system.

Next, we describe the stochastic setting for the service process. Let $\tilde{\boldsymbol{U}}$ denote the one-step state transition matrix of the underlying Markov chain for the service process, i.e., $[\tilde{\boldsymbol{U}}]_{i,j} = \mathrm{P}(S_{n+1} = j \mid S_n = i)$ for $i, j \in \mathcal{R}$. Further, we define $\tilde{\boldsymbol{U}}_0$ and $\tilde{\boldsymbol{U}}_1$ as

$$[\tilde{\boldsymbol{U}}_0]_{i,j} = \begin{cases} [\tilde{\boldsymbol{U}}]_{i,j} & (j = 0), \\ 0 & (j \neq 0), \end{cases}$$

$$[\tilde{\boldsymbol{U}}_1]_{i,j} = \begin{cases} [\tilde{\boldsymbol{U}}]_{i,j} & (j \neq 0), \\ 0 & (j = 0), \end{cases}$$

Note here that we have $\tilde{\boldsymbol{U}} = \tilde{\boldsymbol{U}}_0 + \tilde{\boldsymbol{U}}_1$. Let $\tilde{\boldsymbol{\pi}}$ denote the stationary vector of the underlying Markov chain for the service process, i.e., $[\tilde{\boldsymbol{\pi}}]_i = \mathrm{P}(S_n = i)$. $\tilde{\boldsymbol{\pi}}$ then satisfies $\tilde{\boldsymbol{\pi}} = \tilde{\boldsymbol{\pi}}\tilde{\boldsymbol{U}}$ and $\tilde{\boldsymbol{\pi}}\boldsymbol{e} = 1$.

Finally we will give a series of definitions and assumptions, which make the Markov chains $\{(X_n, V_n)\}$ and $\{(Y_n, V_n)\}$ described in this section have the transition probability matrices (1) and (2), respectively, and satisfy Assumptions 1 and 2. For this purpose, we begin with the definition of random variables V_n ($n = 0, 1, \ldots$). We first consider a mapping $f : \mathcal{S} \times \mathcal{M} \to \mathcal{V}$ defined as

$$f(x, y) = (M+1)x + y,$$

where $\mathcal{V} = \{0, \ldots, (M+1)(R+1)-1\}$. We then define random variables V_n $(n = 0, 1, \ldots)$ on $\mathcal{V}$ by

$$V_n = f(S_n, P_n) = (M+1)S_n + P_n.$$

We next define an $(M+1)(R+1) \times (M+1)(R+1)$ matrix $\boldsymbol{A}_i$ $(i = 0, 1, \ldots)$ by

$$\boldsymbol{A}_i = \gamma \tilde{\boldsymbol{U}}_0 \otimes \hat{\boldsymbol{A}}_i + \left((1-\gamma)\tilde{\boldsymbol{U}}_0 + \tilde{\boldsymbol{U}}_1\right) \otimes \hat{\boldsymbol{A}}_{i-1}, \tag{22}$$

where for notational convenience, we define $\hat{\boldsymbol{A}}_{-1}$ as $\hat{\boldsymbol{A}}_{-1} = \boldsymbol{O}$. Similarly, we define an $(M+1)(R+1) \times (M+1)(R+1)$ matrix $\boldsymbol{B}_i$ $(i = 0, 1, \ldots)$ by

$$\boldsymbol{B}_i = \tilde{\boldsymbol{U}} \otimes \hat{\boldsymbol{A}}_i. \tag{23}$$

Note that obviously $\{(X_n, V_n)\}$ and $\{(Y_n, V_n)\}$ defined in this section become Markov chains whose transition matrices are given by (1) and (2), respectively, with $L = (M+1)(R+1) - 1$. Let $\boldsymbol{\pi}$ denote the stationary vector of the Markov chain $\{V_n\}$, i.e., $[\boldsymbol{\pi}]_i = \mathrm{P}(V_n = i)$. Since the underlying Markov chain $\{P_n\}$ for the arrival process and the underlying Markov chain $\{S_n\}$ for the service process are independent, $\boldsymbol{\pi}$ is given by

$$\boldsymbol{\pi} = \tilde{\boldsymbol{\pi}} \otimes \hat{\boldsymbol{\pi}}, \tag{24}$$

where $\otimes$ denotes the Kronecker product. We assume that the Markov chains $\{(X_n, V_n)\}$ and $\{(Y_n, V_n)\}$ defined in this section satisfy Assumptions 1 and 2. We can replace Assumptions 1 with the following two assumptions, which is more directly associated with the queueing models.

Assumption 3 There exists a $1 \times (M+1)$ probability vector $\hat{\boldsymbol{a}}$ such that

$$\hat{\boldsymbol{A}}_0 = \hat{\boldsymbol{A}}_0 \boldsymbol{e} \hat{\boldsymbol{a}}.$$

Assumption 4 There exists a $1 \times (R+1)$ probability vector $\tilde{\boldsymbol{a}}$ such that

$$\tilde{\boldsymbol{U}}_0 = \tilde{\boldsymbol{U}}_0 \boldsymbol{e} \tilde{\boldsymbol{a}}.$$

We then define a $1 \times (M+1)(R+1)$ probability vector $\boldsymbol{a}$ by $\boldsymbol{a} = \tilde{\boldsymbol{a}} \otimes \hat{\boldsymbol{a}}$. In fact, if Assumptions 3 and 4 are satisfied, Assumptions 1 is satisfied.

Proposition 1 *Under Assumptions 3 and 4,* $\boldsymbol{A}_0 = \boldsymbol{A}_0 \boldsymbol{e} \boldsymbol{a}$.

Proof: From the definition (22) of $\boldsymbol{A}_0$ and Assumptions 3 and 4, we have

$$\boldsymbol{A}_0\boldsymbol{ea} = \gamma(\tilde{\boldsymbol{U}}_0 \otimes \hat{\boldsymbol{A}}_0)\boldsymbol{e}(\tilde{\boldsymbol{a}} \otimes \hat{\boldsymbol{a}}) = \gamma(\tilde{\boldsymbol{U}}_0\boldsymbol{e}\tilde{\boldsymbol{a}}) \otimes (\hat{\boldsymbol{A}}_0\boldsymbol{e}\hat{\boldsymbol{a}}) = \gamma\tilde{\boldsymbol{U}}_0 \otimes \hat{\boldsymbol{A}}_0 = \boldsymbol{A}_0.$$

■

In the setting made in this section and under Assumptions 2, 3 and 4 (or Assumptions 1 and 2), Theorem 1 holds for the Markov chains $\{(X_n, V_n)\}$ and $\{(Y_n, V_n)\}$ defined in this section, and it establishes a proportional relation between the stationary queue length distribution in the finite-buffer queue and that in the corresponding infinite-buffer queue.

4 Exact relation between loss probability and queue length

In this section, we will establish an exact relation holding between the loss probability in the finite-buffer queue and the queue length distribution in the infinite-buffer queue. The exact relation is directly derived from the proportional relation (Theorem 1).

We define the loss probability P_{loss} in the finite-buffer queue as

$$P_{loss} \triangleq \frac{\mathrm{E}[Z_n]}{\mathrm{E}[A_n]}. \tag{25}$$

Let ρ denote the traffic intensity which is given by

$$\rho \triangleq \frac{1}{\gamma}\mathrm{E}[A_n] = \frac{1}{\gamma}\hat{\boldsymbol{\pi}}\sum_{k=1}^{\infty} k\hat{\boldsymbol{A}}_k\boldsymbol{e}. \tag{26}$$

We assume that $\rho < [\tilde{\boldsymbol{\pi}}]_0$. This assumption guarantees that the infinite-buffer queue is stable and the Markov chain $\{(X_n, V_n)\}$ is positive recurrent.

The following formula for the loss probability immediately follows.

Proposition 2 *Under Assumption 2, P_{loss} is given by*

$$P_{loss} = 1 - \frac{1}{\rho}\Big[[\tilde{\boldsymbol{\pi}}]_0 - \boldsymbol{y}_0[(\tilde{\boldsymbol{U}}_0\boldsymbol{e}) \otimes \boldsymbol{e}]\Big]. \tag{27}$$

Proof: From Rate Conservation Law (see, e.g., [20]) or Little's formula, it immediately follows that

$$\mathrm{E}[A_n] - \mathrm{E}[Z_n] = \gamma\sum_{k=1}^{K} \boldsymbol{y}_k(\tilde{\boldsymbol{U}}_0 \otimes \hat{\boldsymbol{U}})\boldsymbol{e}, \tag{28}$$

where the left hand side is the expected upward drift of the queue length and the right hand side is the expected downward drift of the queue length. From (25), (26) and (28), we have

$$P_{loss} = 1 - \frac{1}{\rho}\sum_{k=1}^{K_1} \boldsymbol{x}_k^{(K_1)}(\tilde{\boldsymbol{U}}_0 \otimes \hat{\boldsymbol{U}})\boldsymbol{e}_{(M+1)R}. \tag{29}$$

Noting $\sum_{k=0}^{K} \boldsymbol{y}_k = \boldsymbol{\pi}$ and (24), we rewrite $\sum_{k=1}^{K} \boldsymbol{y}_k(\tilde{\boldsymbol{U}}_0 \otimes \hat{\boldsymbol{U}})\boldsymbol{e}$ as

$$\begin{aligned}
\sum_{k=1}^{K} \boldsymbol{y}_k(\tilde{\boldsymbol{U}}_0 \otimes \hat{\boldsymbol{U}})\boldsymbol{e} &= \sum_{k=1}^{K} \boldsymbol{y}_k \left[(\tilde{\boldsymbol{U}}_0\boldsymbol{e}) \otimes (\hat{\boldsymbol{U}}\boldsymbol{e})\right] \\
&= \sum_{k=1}^{K} \boldsymbol{y}_k \left[(\tilde{\boldsymbol{U}}_0\boldsymbol{e}) \otimes \boldsymbol{e}\right] \\
&= \sum_{k=0}^{K} \boldsymbol{y}_k \left[(\tilde{\boldsymbol{U}}_0\boldsymbol{e}) \otimes \boldsymbol{e}\right] - \boldsymbol{y}_0 \left[(\tilde{\boldsymbol{U}}_0\boldsymbol{e}) \otimes \boldsymbol{e}\right] \\
&= \boldsymbol{\pi} \left[(\tilde{\boldsymbol{U}}_0\boldsymbol{e}) \otimes \boldsymbol{e}\right] - \boldsymbol{y}_0 \left[(\tilde{\boldsymbol{U}}_0\boldsymbol{e}) \otimes \boldsymbol{e}\right] \\
&= (\tilde{\boldsymbol{\pi}} \otimes \hat{\boldsymbol{\pi}}) \left[(\tilde{\boldsymbol{U}}_0\boldsymbol{e}) \otimes \boldsymbol{e}\right] - \boldsymbol{y}_0 \left[(\tilde{\boldsymbol{U}}_0\boldsymbol{e}) \otimes \boldsymbol{e}\right] \\
&= \tilde{\boldsymbol{\pi}}\tilde{\boldsymbol{U}}_0\boldsymbol{e} - \boldsymbol{y}_0 \left[(\tilde{\boldsymbol{U}}_0\boldsymbol{e}) \otimes \boldsymbol{e}\right] \\
&= [\tilde{\boldsymbol{\pi}}]_0 - \boldsymbol{y}_0 \left[(\tilde{\boldsymbol{U}}_0\boldsymbol{e}) \otimes \boldsymbol{e}\right].
\end{aligned} \tag{30}$$

Substituting (30) into (29), we obtain (27) ∎

The following theorem establishes an exact relation holding between the loss probability in the finite-buffer queue and the stationary queue length distribution in the infinite-buffer queue, and the exact relation expresses the loss probability in the finite-buffer queue as a function of the stationary queue length distribution in the infinite-buffer queue.

Theorem 2 *Under Assumptions 2, 3 and 4, the loss probability P_{loss} is given in terms of the stationary distribution $\boldsymbol{x}$ as follows:*

$$P_{loss} = \frac{([\tilde{\boldsymbol{\pi}}]_0 - \rho) \sum_{i=K+1}^{\infty} \boldsymbol{x}_i \boldsymbol{A}_0 \boldsymbol{e}}{\rho \sum_{i=0}^{K} \boldsymbol{x}_i \boldsymbol{A}_0 \boldsymbol{e}}.$$

Proof: By similar argument when we derived (28), we obtain

$$\mathrm{E}[A_n] = \sum_{k=1}^{\infty} \boldsymbol{x}_k^{(K_0)} (\tilde{\boldsymbol{U}}_0 \otimes \hat{\boldsymbol{U}}) \boldsymbol{e}_{(M+1)R}. \tag{31}$$

By similar argument when we derived (30), the right hand side of (31) can be rewritten as

$$\sum_{k=1}^{\infty} \boldsymbol{x}_k (\tilde{\boldsymbol{U}}_0 \otimes \hat{\boldsymbol{U}}) \boldsymbol{e} = [\tilde{\boldsymbol{\pi}}]_0 - \boldsymbol{x}_0 \left[(\tilde{\boldsymbol{U}}_0\boldsymbol{e}) \otimes \boldsymbol{e}\right]. \tag{32}$$

Using (26) and (32) in (31), we derive

$$\boldsymbol{x}_0 \left[(\tilde{\boldsymbol{U}}_0\boldsymbol{e}) \otimes \boldsymbol{e}\right] = [\tilde{\boldsymbol{\pi}}]_0 - \rho. \tag{33}$$

From Theorem 1 and Proposition 2, using (33) and noting $\sum_{i=0}^{\infty} \boldsymbol{x}_i^{(K_0)} = \boldsymbol{\pi}$, we have

$$
\begin{aligned}
P_{loss} &= 1 - \frac{1}{\rho}\left[[\tilde{\boldsymbol{\pi}}]_0 - \frac{\boldsymbol{\pi}\boldsymbol{A}_0\boldsymbol{e}\boldsymbol{x}_0[(\tilde{\boldsymbol{U}}_0\boldsymbol{e}) \otimes \boldsymbol{e}]}{\sum_{i=0}^{K} \boldsymbol{x}_i\boldsymbol{A}_0\boldsymbol{e}}\right] \\
&= 1 - \frac{1}{\rho}\left[[\tilde{\boldsymbol{\pi}}]_0 - \frac{\boldsymbol{\pi}\boldsymbol{A}_0\boldsymbol{e}([\tilde{\boldsymbol{\pi}}]_0 - \rho)}{\sum_{i=0}^{K} \boldsymbol{x}_i\boldsymbol{A}_0\boldsymbol{e}}\right] \\
&= 1 + \frac{[\tilde{\boldsymbol{\pi}}]_0 \sum_{i=K+1}^{\infty} \boldsymbol{x}_i\boldsymbol{A}_0\boldsymbol{e} - \rho\boldsymbol{\pi}\boldsymbol{A}_0\boldsymbol{e}}{\rho\sum_{i=0}^{K} \boldsymbol{x}_i\boldsymbol{A}_0\boldsymbol{e}} \\
&= \frac{[\tilde{\boldsymbol{\pi}}]_0 \sum_{i=K+1}^{\infty} \boldsymbol{x}_i\boldsymbol{A}_0\boldsymbol{e} - \rho\sum_{i=K+1}^{\infty} \boldsymbol{x}_i\boldsymbol{A}_0\boldsymbol{e}}{\rho\sum_{i=0}^{K} \boldsymbol{x}_i\boldsymbol{A}_0\boldsymbol{e}} \\
&= \frac{([\tilde{\boldsymbol{\pi}}]_0 - \rho)\sum_{i=K_1+1}^{\infty} \boldsymbol{x}_i\boldsymbol{A}_0\boldsymbol{e}}{\rho\sum_{i=0}^{K} \boldsymbol{x}_i\boldsymbol{A}_0\boldsymbol{e}}.
\end{aligned}
$$

■

Remark 1 The exact relation (Theorem 2) is considered as a generalization and integration of the exact relations shown in [11, 12, 14, 17], and Theorem 2 includes those exact relations as special cases.

5 Asymptotic loss probability

When we use Theorem 2 to calculate the loss probability in the finite-buffer queue, we need to compute $\boldsymbol{x}$. The computation of the $\boldsymbol{x}$ is not a easy task when M or R are large. In this section, we develop a formula which can estimate the loss probability more easily even when M or R are large. For this purpose, we exploit the property that the tail distribution of the queue length in infinite-buffer queues has a rather simple asymptotic form in many cases. In particular, since $\{(X_n, V_n)\}$ is an M/G/1 type Markov chain, the tail distribution $\sum_{k=N+1}^{\infty} \boldsymbol{x}_k$ has a simple geometric asymptotic form under some conditions. Exploiting this property, the formula derived in this section computes the *asymptotic loss probability* [11, 14].

We begin with the definition of notations which will appear in the formula. We define an $(M+1)(R+1) \times (M+1)(R+1)$ matrix generating function $\boldsymbol{A}(z)$, an $(M+1) \times (M+1)$ matrix generating function $\hat{\boldsymbol{A}}(z)$ and an $(R+1) \times (R+1)$ matrix generating function $\tilde{\boldsymbol{U}}_\gamma(z)$ as

$$
\boldsymbol{A}(z) = \sum_{k=0}^{\infty} \boldsymbol{A}_k z^k, \qquad \hat{\boldsymbol{A}}(z) = \sum_{k=0}^{\infty} \hat{\boldsymbol{A}}_k z^k, \qquad \tilde{\boldsymbol{U}}_\gamma(z) = (\gamma + (1-\gamma)z)\tilde{\boldsymbol{U}}_0 + z\tilde{\boldsymbol{U}}_1. \tag{34}
$$

Note here that from (22), we have

$$
\boldsymbol{A}(z) = \tilde{\boldsymbol{U}}_\gamma(z) \otimes \hat{\boldsymbol{A}}(z). \tag{35}
$$

Let $\delta(z)$ denote the Perron-Frobenius eigenvalue of $\boldsymbol{A}(z)$, and $\boldsymbol{u}(z)$ and $\boldsymbol{v}(z)$ denote its left and right eigenvectors which satisfy the normalizing conditions: $\boldsymbol{u}(z)\boldsymbol{v}(z) = 1$ and $\boldsymbol{u}(z)\boldsymbol{e} = 1$.

Also, let $\hat{\delta}(z)$ denote the Perron-Frobenius eigenvalue of $\hat{\boldsymbol{A}}(z)$, and $\hat{\boldsymbol{u}}(z)$ and $\hat{\boldsymbol{v}}(z)$ denote its left and right eigenvectors which satisfy the normalizing conditions: $\hat{\boldsymbol{u}}(z)\hat{\boldsymbol{v}}(z) = 1$ and $\hat{\boldsymbol{u}}(z)\boldsymbol{e} = 1$. Similarly, let $\tilde{\delta}(z)$ denote the Perron-Frobenius eigenvalue of $\tilde{\boldsymbol{U}}_\gamma(z)$, and $\tilde{\boldsymbol{u}}(z)$ and $\tilde{\boldsymbol{v}}(z)$ denote its left and right eigenvectors which satisfy the normalizing conditions: $\tilde{\boldsymbol{u}}(z)\tilde{\boldsymbol{v}}(z) = 1$ and $\tilde{\boldsymbol{u}}(z)\boldsymbol{e} = 1$. Note here that from (35), we have [21]

$$\delta(z) = \tilde{\delta}(z)\hat{\delta}(z), \quad \boldsymbol{u}(z) = \tilde{\boldsymbol{u}}(z) \otimes \hat{\boldsymbol{u}}(z), \quad \boldsymbol{v}(z) = \tilde{\boldsymbol{v}}(z) \otimes \hat{\boldsymbol{v}}(z). \tag{36}$$

Now we make several assumptions [1, 6, 13] to ensure that the queue length has a simple asymptotic expression.

Assumption 5

- There exists at least one zero of $\det[z\boldsymbol{I} - \boldsymbol{A}(z)]$ outside the unit disk.
- Among those, there exists a real and positive zero z^*, and the absolute value of z^* is strictly smaller than those of other zeros.
- $\boldsymbol{O} < \boldsymbol{A}(z) \ll +\infty, \quad 1 \le z \le z^*, \quad z \in \mathbb{R}$, where $\mathbb{R}$ denotes the set of all real numbers.

The following proposition shows that the tail distribution of the queue length in the infinite-buffer queue has a simple geometric expression. The proof is provided in [9].

Proposition 3 *Under Assumptions 2 and 5, $\sum_{n=N+1}^{\infty} \boldsymbol{x}_n$ is expressed as*

$$\sum_{n=N+1}^{\infty} \boldsymbol{x}_n = \frac{\gamma \boldsymbol{x}_0 \left[(\tilde{\boldsymbol{U}}_0 \tilde{\boldsymbol{v}}(z^*)) \otimes \hat{\boldsymbol{v}}(z^*) \right]}{\hat{\delta}(z^*)(\delta'(z^*) - 1)} (z^*)^{-N} \boldsymbol{u}(z^*) + o\left((z^*)^{-N}\right) \boldsymbol{e}', \qquad N \ge 0, \tag{37}$$

where $\boldsymbol{e}'$ denotes the $1 \times (M+1)(R+1)$ vector whose elements are all equal to one, and z^* is the minimum real solution of $z = \delta(z)$ for $z \in (1, \infty)$.

The following corollary is directly obtained from Theorem 2.

Corollary 1 *Under Assumptions 1, 2 and 3, the loss probability P_{loss} is expressed as*

$$P_{loss} = \frac{([\tilde{\boldsymbol{\pi}}]_0 - \rho) \sum_{i=K+1}^{\infty} \boldsymbol{x}_i \boldsymbol{A}_0 \boldsymbol{e}}{\rho \left(\gamma [\tilde{\boldsymbol{\pi}}]_0 \tilde{\boldsymbol{\pi}} \hat{\boldsymbol{A}}_0 \boldsymbol{e} - \sum_{i=K+1}^{\infty} \boldsymbol{x}_i \boldsymbol{A}_0 \boldsymbol{e} \right)}. \tag{38}$$

Proof: First, note that from the definition (22) of $\boldsymbol{A}_i$, $\boldsymbol{A}_0$ is expressed as

$$\boldsymbol{A}_0 = \gamma \tilde{\boldsymbol{U}}_0 \otimes \hat{\boldsymbol{A}}_0.$$

$\boldsymbol{A}_0 \boldsymbol{e}$ is thus expressed as

$$\boldsymbol{A}_0 \boldsymbol{e} = \gamma(\tilde{\boldsymbol{U}}_0 \otimes \hat{\boldsymbol{A}}_0)\boldsymbol{e} = \gamma(\tilde{\boldsymbol{U}}_0 \boldsymbol{e}) \otimes (\hat{\boldsymbol{A}}_0 \boldsymbol{e}). \tag{39}$$

From (24) and (39), we have

$$\boldsymbol{\pi} \boldsymbol{A}_0 \boldsymbol{e} = \gamma(\tilde{\boldsymbol{\pi}} \otimes \hat{\boldsymbol{\pi}}) \left[(\tilde{\boldsymbol{U}}_0 \boldsymbol{e}) \otimes (\hat{\boldsymbol{A}}_0 \boldsymbol{e}) \right] = \gamma(\tilde{\boldsymbol{\pi}} \tilde{\boldsymbol{U}}_0 \boldsymbol{e})(\hat{\boldsymbol{\pi}} \hat{\boldsymbol{A}}_0 \boldsymbol{e}). = \gamma[\tilde{\boldsymbol{\pi}}]_0 \hat{\boldsymbol{\pi}} \hat{\boldsymbol{A}}_0 \boldsymbol{e}. \tag{40}$$

Using $\sum_{i=0}^{\infty} \boldsymbol{x}_i^{(K_0)} = \boldsymbol{\pi}$ and (40) in Theorem 2, we obtain (38) ∎

Using Proposition 3 in Corollary 1, the following formula to compute the asymptotic loss probability is immediately obtained.

Theorem 3 *Under Assumptions 2, 3, 4 and 5, the loss probability P_{loss} is asymptotically expressed as*

$$P_{loss} \approx \left(\frac{1}{\rho} - \frac{1}{[\tilde{\boldsymbol{\pi}}]_0} \right) \frac{\boldsymbol{x}_0 \left[(\tilde{\boldsymbol{U}}_0 \tilde{\boldsymbol{v}}(z^*)) \otimes \hat{\boldsymbol{v}}(z^*) \right]}{\tilde{\delta}(z^*)(\delta'(z^*) - 1)} \frac{\boldsymbol{u}(z^*) \boldsymbol{A}_0 \boldsymbol{e}}{\hat{\boldsymbol{\pi}} \hat{\boldsymbol{A}}_0 \boldsymbol{e}} (z^*)^{-K}. \tag{41}$$

Proof: Using (37) in (38), we obtain

$$\begin{aligned} P_{loss} &= \frac{[\tilde{\boldsymbol{\pi}}]_0 - \rho}{\rho} \sum_{l=1}^{\infty} \left(\frac{\sum_{i=K+1}^{\infty} \boldsymbol{x}_i \boldsymbol{A}_0 \boldsymbol{e}}{\gamma [\tilde{\boldsymbol{\pi}}]_0 \hat{\boldsymbol{\pi}} \hat{\boldsymbol{A}}_0 \boldsymbol{e}} \right)^l \\ &\approx \frac{[\tilde{\boldsymbol{\pi}}]_0 - \rho}{\rho} \frac{\sum_{i=K+1}^{\infty} \boldsymbol{x}_i \boldsymbol{A}_0 \boldsymbol{e}}{\gamma [\tilde{\boldsymbol{\pi}}]_0 \hat{\boldsymbol{\pi}} \hat{\boldsymbol{A}}_0 \boldsymbol{e}} \\ &\approx \left(\frac{1}{\rho} - \frac{1}{[\tilde{\boldsymbol{\pi}}]_0} \right) \frac{\boldsymbol{x}_0 \left[(\tilde{\boldsymbol{U}}_0 \tilde{\boldsymbol{v}}(z^*)) \otimes \hat{\boldsymbol{v}}(z^*) \right]}{\tilde{\delta}(z^*)(\delta'(z^*) - 1)} \frac{\boldsymbol{u}(z^*) \boldsymbol{A}_0 \boldsymbol{e}}{\hat{\boldsymbol{\pi}} \hat{\boldsymbol{A}}_0 \boldsymbol{e}} (z^*)^{-K} \end{aligned}$$

∎

Remark 2 The formula (41) to compute the asymptotic loss probability (Theorem 3) is a generalization of the formula (Corollary 5) derived in [14]. When $R = 1$, $\tilde{\boldsymbol{U}}_0 = 1$ and $\tilde{\boldsymbol{U}}_1 = 0$, the formula (41) is reduced to the formula in [14].

References

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

[2] Asmussen, S. *Applied probability and queues*, John Wiley & Sons, Chichester, 1987.

[3] C. Blondia and O. Casals, Statistical multiplexing of VBR sources: a matrix-analytic approach, Perform. Eval. 16 (1992) 5–20.

[4] H. Bruneel, Analysis of discrete-time buffer with one single output channel subject to a general interruption process, in; Proc. Performance '84, 1984, pp.103-115.

[5] M. L. Chaudhry, On numerical computations of some discrete-time queues, in: W. K. Grassmann, (Ed.), Computational Probability, Kluwer Academic Publishers, Boston, 2000, pp.81–111.

[6] E. Falkenberg, On the asymptotic behavior of the stationary distribution of Markov chains of $M/G/1$-type, Stochastic Models 10 (1994) 75–98.

[7] W, K. Grassman, M. L. Taksar and D. P. Heyman, Regenerative analysis and steady state distribution for Markov chains. Oper. Res. 33 (1985) 1107–1116.

[8] C. Herrmann, The complete analysis of the discrete time finite DBMAP/G/1/N queue, Perform. Eval. 43 (2001) 95–121.

[9] F. Ishizaki, Loss probability in a finite queue with service interruptions and queue length distribution in the corresponding infinite queue, submitted, 2002.

[10] F. Ishizaki, Numerical method for discrete-time finite-buffer queues with some regenerative structure, Stochastic Models 19 (2002) 25-39.

[11] F. Ishizaki, Exact relation between loss probability and queue length in an ATM multiplexer with correlated arrivals and periodic vacations, in; Proc. 14th ICOIN, 2000, pp.1A-3.1–1A-3.8.

[12] F. Ishizaki, G. C. Lin and T. Suda, On-line sensitivity analysis of feedback controlled queueing systems with respect to buffer capacity, in; Proc. CDC '99, 1999, pp.40–50.

[13] F. Ishizaki and T. Takine, Bounds for the tail distribution in a queue with the superposition of general periodic Markov sources, in; Proc. INFOCOM '97, 1997, pp.1088–1095.

[14] F. Ishizaki and T. Takine, Loss probability in a finite discrete-time queue in terms of the steady state distribution of an infinite queue, Queueing Systems 31 (1999) 317–326.

[15] F. Ishizaki and T. Takine, Bounds for the tail distribution in a queue with a superposition of general periodic Markov sources: theory and application, Queueing Systems 34 (2000) 67–100.

[16] F. Ishizaki, T. Takine, Y. Takahashi and T. Hasegawa, A generalized SBBP/G/1 queue and its applications, Perform. Eval. 21 (1994) 163–181.

[17] K. Kang, B. Steyaert and C. Kim, A simple relation between loss performance and buffer contents in a statistical multiplexer with periodic vacations, IEICE Trans. Commun. E80-B (1997) 1749–1752.

[18] G. Latouche, P. A. Jacobs and D. P. Gaver, Finite Markov chain models skip-free in one direction, Naval Res. Logistics Quarterly 31 (1984) 571–588.

[19] J.-Y. Le Boudec, An efficient solution method for Markov models of ATM links with loss probabilities, IEEE J. Sel. Areas Commun. 9 (1991) 408–417.

[20] M. Miyazawa and Y. Takahashi, Rate conservation principle for discrete-time queues, Queueing Systems 12 (1992) 215–230.

[21] M. F. Neuts, Structured stochastic matrices of M/G/1 type and their applications, Marcel Dekker, New York, 1989.

[22] V. Ramaswami, A stable recursion for steady state vector in Markov chains of M/G/1 type, Stochastic Models 4 (1988) 183–188.

[23] K. Sohraby, On the asymptotic behavior of heterogeneous statistical multiplexer with applications. in; Proc. INFOCOM '92, 1992, pp.839–847.

[24] W. J. Stewart, Introduction to the numerical solution of Markov chains, Princeton University Press, Princeton, 1994.

[25] H. Takagi, Queueing Analysis, vol.3, North-Holland, Amsterdam, 1993.

[26] D. Towsley, The analysis of a statistical multiplexer with nonindependent arrivals and errors, IEEE Trans. Commun. 28 (1980) 65-72.