NUMERICAL IMPLEMENTATION OF THE AUGMENTED TRUNCATION APPROXIMATION TO SINGLE-SERVER QUEUES WITH LEVEL-DEPENDENT ARRIVALS AND DISASTERS

Masatoshi Kimura Tetsuya Takine
Osaka University

(Received April 3, 2020; Revised September 3, 2020)

Abstract This paper considers the computation of the stationary queue length distribution in single-server queues with level-dependent arrivals and disasters. We assume that service times follow a general distribution and therefore, we consider the stationary queue length distribution via an imbedded Markov chain. Because this imbedded Markov chain has infinitely many states, level dependence, and bidirectional jumps of levels, it is hard to compute the solution of the global balance equation exactly. We thus consider the augmented truncation approximation. In particular, we focus on the computation of the truncated state transition probability matrix of the imbedded Markov chain, assuming that the underlying continuous-time absorbing Markov chain during a service time is not uniformizable. Under some stability conditions, we develop a computational procedure for the truncated transition probability matrix, where the upper bound of errors owing to truncation can be set in advance. We also provide some numerical examples and demonstrate that our procedure works well.

Keywords: Queue, disaster, level dependence, augmented truncation approximation, uniformization technique, error bound

1. Introduction

In queueing systems, a disaster is the event that all customers are removed from the system. For instance, resets in computer systems, which arise from overloads or virus infections, can be modeled by disasters. Such events are also called clearing [2], (total) catastrophe [4, 12], or mass exodus [5] in the literature. The stationary workload [2, 9, 10] and the stationary queue length [6, 7, 13, 18] in single-server queues with disasters and generally distributed service times have been studied in the literature.

Queueing systems whose arrival rates depend on the queue length are termed state-dependent queues or level-dependent queues. Such queueing systems naturally arise in practical modeling. In many situations, a long waiting line discourages prospective customers from joining the queue, so that the net arrival rate depends on the queue length. The level-dependent batch Markovian arrival process (LD-BMAP) is known as a generalization of the Markovian arrival process (MAP), and the stationary distribution in LD-BMAP/G/1 queues has been discussed in [8].

It would be natural that the occurrence of disasters also depends on the state of the system. For instance, a huge backlog may cause a fatal fault or a compulsory reset. This paper considers a single-server queue, where arrivals and disasters occur according to a level-dependent marked Markovian arrival process (LD-MMAP) described in later. We assume that customers are served on a FIFO basis and that their service times are independent and identically distributed (i.i.d.) according to a general distribution $H(x)$ ($x \geq 0$) with finite
mean $E[H]$. To the best of our knowledge, single-server queues with generally distributed service times and level-dependent disasters have not been studied so far.

We define $L(t)$ $(t \geq 0)$ as the number of customers in the system at time $t$ and $S(t)$ $(t \geq 0)$ as the state of the auxiliary variable at time $t$. We call $L(t)$ level at time $t$ and $S(t)$ phase at time $t$. We also define level $k$ as the subset $\{k\} \times \mathcal{M}$ ($k = 0, 1, \ldots$) of the state space $\{0, 1, \ldots\} \times \mathcal{M}$ of $\{(L(t), S(t))\}_{t \geq 0}$. Let $N(s, t]$ ($s < t$) denote the number of customers entering the system in time interval $(s, t]$ under the assumption that no service completions occur in $(s, t]$. We then define $C_k$, $D_k$, and $\Gamma_k$ $(k = 0, 1, \ldots)$ as $M \times M$ matrices whose elements are denoted by capital bold letters, lowercase bold Greek letters, and lowercase bold Roman letters, respectively.

For any pair of states $(k, i)$ and $(k, j)$, respectively.

\[ \frac{\Pr(N(t, t + \Delta t] = 0, S(t + \Delta t) = j \mid (L(t), S(t)) = (k, i))}{\Delta t}, \quad i, j \in \mathcal{M}, \quad i \neq j, \]

\[ \frac{\Pr(N(t, t + \Delta t] = 1, S(t + \Delta t) = j \mid (L(t), S(t)) = (k, i))}{\Delta t}, \quad i, j \in \mathcal{M}, \]

\[ \frac{\Pr(\text{a disaster occurs in } (t, t + \Delta t], S(t + \Delta t) = j \mid (L(t), S(t)) = (k, i))}{\Delta t}, \quad i, j \in \mathcal{M}, \]

and

\[ [C_k]_{i,j} = \lim_{\Delta t \to 0} \frac{\Pr(N(t, t + \Delta t] = 0, S(t + \Delta t) = j \mid (L(t), S(t)) = (k, i))}{\Delta t}, \quad i \in \mathcal{M}. \]

We assume $\Gamma_0 = O$ without loss of generality. We also assume $D_k \neq O$ $(k = 0, 1, \ldots)$ and $0 < [-C_k]_{i,i} < \infty$ $(k = 0, 1, \ldots; i \in \mathcal{M})$. By definition, $\{(C_k, D_k, \Gamma_k); k = 0, 1, \ldots\}$ describes the arrival/disaster process. Roughly speaking, when $k$ $(k = 0, 1, \ldots)$ customers stay in the system (i.e., $L(t) = k$), an arrival (resp. a disaster) occurs along with a transition driven by $D_k$ (resp. $\Gamma_k$). By definition, $(C_k + D_k + \Gamma_k)$ $(k = 0, 1, \ldots)$ represents the infinitesimal generator of the underlying phase process $\{S(t)\}_{t \geq 0}$ when $L(t) = k$. Note that

\[(C_k + D_k + \Gamma_k)e = 0, \quad k = 0, 1, \ldots, \tag{1.1}\]

where $e$ denotes a column vector with an appropriate dimension, whose elements are all equal to one. Note that when a service completes, $L(t)$ decreases by one, while $S(t)$ remains unchanged. Throughout this paper, matrices, row vectors, and column vectors are denoted by capital bold letters, lowercase bold Greek letters, and lowercase bold Roman letters, respectively.

We are interested in the stationary distribution of $\{(L(t), S(t))\}_{t \geq 0}$. We thus assume that for any pair of states $(k_1, j_1)$ and $(k_2, j_2)$ $(k_1, k_2 = 0, 1, \ldots; j_1, j_2 \in \mathcal{M})$, there exists a sample path from $(k_1, j_1)$ to $(k_2, j_2)$. Unfortunately, the necessary and sufficient stability condition for queueing systems with level-dependent arrivals and disasters has not been clarified. In this paper, we assume the following sufficient condition for the stability.

**Assumption 1.1.** $\{(L(t), S(t))\}_{t \geq 0}$ satisfies either of the following two conditions.

(i) There exists $K_1^\dagger \in \{0, 1, \ldots\}$ and $\alpha > 0$ such that $D_k e - \kappa \Gamma_k e \leq -\alpha (k \geq K_1^\dagger)$, or

(ii) there exists $K_2^\dagger \in \{0, 1, \ldots\}$ and $\alpha > 0$ such that $D_k e - \mu e \leq -\alpha (k \geq K_2^\dagger)$, where $\mu = 1/E[H]$.

Assumption 1.1 (i) is a drift condition ignoring service completions, while Assumption 1.1 (ii) is a drift condition ignoring disasters. Note that under Assumption 1.1, the arrival/disaster process characterized by $\{(C_k, D_k, \Gamma_k); k = 0, 1, \ldots\}$ is non-explosive, i.e.,
for an arbitrary finite \( T > 0 \) and an arbitrary initial state \((L(0), S(0))\), the number of arrivals in time interval \([0, T]\) is finite with probability 1.

**Remark 1.1.** The stability under Assumption 1.1 (i) can be shown as follows. For a while, we assume that services never complete. In this case, \(\{(L(t), S(t))\}_{t \geq 0}\) forms a continuous-time Markov chain and it is readily shown by Pakes’s Lemma [3, Corollary 1.1 of Chap. 5] that \(\{(L(t), S(t))\}_{t \geq 0}\) is positive-recurrent under Assumption 1.1 (i). Therefore, in the model with a finite \(E[\tau]\) described above, the mean first passage time from any state to the empty system is finite, so that \(\{(L(t), S(t))\}_{t \geq 0}\) is stable. On the other hand, the stability under Assumption 1.1 (ii) can be shown by considering a modified model in which all customers remain in the system at the occurrence of disasters (i.e., the ordinary LD-MAP/G/1 queueing model without disasters), and applying Theorem 3.3-3 in [8].

Let \(\pi = (\pi_0, \pi_1, \ldots)\) denote the stationary distribution of \(\{(L(t), S(t))\}_{t \geq 0}\), where \(\pi_t\) denotes a \(1 \times M\) vector whose \(j\)th \((j \in M)\) element is given by \(\lim_{t \to \infty} \Pr((L(t), S(t) = (\ell, j))\). Since service times are generally distributed, we consider an embedded Markov chain associated with \(\{(L(t), S(t))\}_{t \geq 0}\). Specifically, the standard approach to the analysis of the stationary distribution \(\pi\) is outlined as follows.

**Step 1:** Construct an imbedded Markov chain: We choose time instants at which services complete and disasters occur as imbedded Markov points. We define \(\tau_n (n = 0, 1, \ldots)\) as the \(n\)th imbedded Markov point, where \(\tau_0 = 0\). Let \((L_n, S_n) = (L(\tau_n), S(\tau_n)) (n = 0, 1, \ldots)\). It is clear that \(\{(L_n, S_n)\}_{n=0,1,\ldots}\) forms a discrete-time Markov chain whose state transition probability matrix \(P\) takes the following form:

\[
P = \begin{pmatrix}
A_{0,0} & A_{0,1} & A_{0,2} & A_{0,3} & A_{0,4} & A_{0,5} & \cdots \\
A_{1,0} & A_{1,1} & A_{1,2} & A_{1,3} & A_{1,4} & A_{1,5} & \cdots \\
A_{2,0} & A_{2,1} & A_{2,2} & A_{2,3} & A_{2,4} & A_{2,5} & \cdots \\
A_{3,0} & O & A_{3,2} & A_{3,3} & A_{3,4} & A_{3,5} & \cdots \\
A_{4,0} & O & O & A_{4,3} & A_{4,4} & A_{4,5} & \cdots \\
A_{5,0} & O & O & O & A_{5,4} & A_{5,5} & \cdots \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots
\end{pmatrix},
\]

(1.2)

where \(A_{k,0} (k = 1, 2, \ldots)\) and \(A_{k,\ell} (k = 0, 1, \ldots, \ell = k - 1, k, \ldots)\) are \(M \times M\) matrices.

**Step 2:** Compute the stationary distribution of the imbedded Markov chain: We define \(\kappa = (\kappa_0, \kappa_1, \ldots)\) as the stationary distribution of the imbedded Markov chain \(\{(L_n, S_n)\}_{n=0,1,\ldots}\), where \(\kappa_\ell\) denotes a \(1 \times M\) vector whose \(j\)th \((j \in M)\) element is given by \(\lim_{n \to \infty} \Pr((L_n, S_n) = (\ell, j))\). By definition, \(\kappa\) satisfies \(\kappa = \kappa P\) and \(\kappa e = 1\).

**Step 3:** Compute the stationary distribution \(\pi = (\pi_0, \pi_1, \ldots)\) of \(\{(L(t), S(t))\}_{t \geq 0}\): \(\pi\) is given in terms of \(\kappa_k (k = 0, 1, \ldots)\) as follows.

\[
\pi_0 = \frac{\kappa_0(-C_0)^{-1}}{E[T]}, \quad \pi_\ell = \frac{\kappa_0(-C_0)^{-1}D_0W_{1,\ell}}{E[T]} + \sum_{k=1}^\ell \frac{\kappa_k W_{k,\ell}}{E[T]}, \quad \ell = 1, 2, \ldots,
\]

(1.3)

where \(E[T]\) denotes the average of \(\tau_{n+1} - \tau_n (n = 0, 1, \ldots)\) and \(W_{k,\ell} (k = 1, 2, \ldots, \ell = k, k+1, \ldots)\) denotes an \(M \times M\) matrix whose \((i,j)\)th \((i,j \in M)\) element represents the mean total sojourn time in state \((\ell, j)\) in \((\tau_n, \tau_{n+1}]\), starting from \((L_n, S_n) = (k, i)\).

The purpose of this paper is to establish a method for numerical implementation of the above three steps in the LD-MMAP/G/1 queue with disasters. In what follows, we identify the problems to be resolved in each step.
In Step 1, we have to compute \( A_{k,\ell} \)'s \((k = 0, 1, \ldots, \ell = 0, k - 1, k, \ldots)\) in \( P \). As we will see in (2.1), \( A_{k,\ell} \)'s are given in terms of \( M \times M \) matrices \( B_{A,k,\ell} \)'s in \( B_A \):

\[
B_A = \begin{pmatrix}
I & O & O & O & \cdots \\
B_{A,1,0} & B_{A,1,1} & B_{A,1,2} & B_{A,1,3} & \cdots \\
B_{A,2,0} & O & B_{A,2,2} & B_{A,2,3} & \cdots \\
B_{A,3,0} & O & O & B_{A,3,3} & \cdots \\
\vdots & \vdots & \vdots & \vdots & \ddots
\end{pmatrix} = \int_0^\infty \exp[Ut]dH(t), \quad (1.4)
\]

where \( U \) denotes the infinitesimal generator of an absorbing Markov chain that represents the arrival process until a disaster occurs.

\[
U = \begin{pmatrix}
O & O & O & O & \cdots \\
\Gamma_1 & C_1 & D_1 & O & \cdots \\
\Gamma_2 & O & C_2 & D_2 & O & \cdots \\
\Gamma_3 & O & O & C_3 & D_3 & \cdots \\
\Gamma_4 & O & O & O & C_4 & \cdots \\
\vdots & \vdots & \vdots & \vdots & \vdots & \ddots
\end{pmatrix}. \quad (1.5)
\]

In what follows, we use the following convention, as with \( B_A \) in (1.4). For any matrix \( X \) composed of \( M \times M \) block matrices, let \( X_{k,\ell} \) \((k, \ell \geq 0)\) denote the \((k, \ell)\)th block matrix of \( X \). For example, for \( U \) in (1.5), \( U_{k,k} = C_k \) \((k = 1, 2, \ldots)\).

**Remark 1.2.** Note that \( \exp[Ut] \) \((t \geq 0)\) can be regarded as a transition probability matrix, i.e., \( \exp[Ut] \geq O \) and \( \exp[Ut]e = e \). Therefore, all elements in \( B_A \) are finite because \( \int_0^\infty \exp[Ut]dH(t)e = \int_0^\infty dH(t)e = e \).

The stable and accurate numerical implementation of block matrices \( B_{A,k,\ell} \) in \( B_A \) is crucial and it is one of the main topics of this paper. We define \( \theta \) as

\[
\theta = \sup_{k=0,1,\ldots} \left( \max_{i \in \mathcal{M}} [-C_k]_{i,i} \right).
\]

If \( \theta \) is finite, the uniformization technique is directly applicable and \( B_{A,k,\ell} \) \((k = 0, 1, \ldots, \ell = 0, k - 1, k, \ldots)\) is given by the weighted sum of infinitely many \( M \times M \) non-negative matrices, where the weights are given in terms of \( \theta \) and the service time distribution \( H(x) \). The advantages of the uniformization technique are that (i) the procedure is numerically stable because it does not involve subtractions and that (ii) for a given error bound \( \epsilon \), we can set an appropriate truncation point \( n := n^*(\epsilon) \) of the infinite sum representing \( B_{A,k,\ell} \).

Note that \( \theta \) can be infinite and in this case, there are no standard methods for computing block matrices \( B_{A,k,\ell} \) in \( B_A \) in (1.4), which guarantee error bounds. In this paper, we consider computing an approximation to \( B_{A,k,\ell} \) by using only the \((m + 1)M \times (m + 1)M\) north-west corner submatrix \( U(m) \) of \( U \) because \( \exp[Ut(m)t] \) is always uniformizable since \([-C_k]_{i,i} < \infty \) \((i \in \mathcal{M})\) for all \( k = 1, 2, \ldots \).

The truncation of \( U \) at level \( m \) does not affect the accuracy of computed \( B_{A,k,\ell} \) \((\ell = k, k + 1, \ldots, m)\) because it can be represented completely in terms of \( U(m) \). Contrarily, the accuracy of computed \( B_{A,k,0} \) is affected by the truncation of \( U \) by the following reason. Note that \( B_{A,k,0} \) includes the probability that a disaster occurs before a service completion given that the service starts with \( L_n = k \leq m \). We then consider the level \( L(\tau_{n+1}-) \) immediately before the occurrence of a disaster. Because \( \Pr(L(\tau_{n+1}-) = \ell \mid L(\tau_n) = k) > 0 \) for all


\[ \ell > m, \] the exact expression of \(B_{A,k,0}\) is given in terms of the whole \(U\). Therefore, if we compute \(B_{A,k,0}\) using the truncated \(U(m)\) of \(U\) at level \(m\), we will obtain an approximation \(B_{A,k,0}(m)\) to \(B_{A,k,0}\), which ignores all sample paths with \(L(\tau_{n+1}+) > m\). Consequently, in terms of an approximation \(A_{k,0}(m)\) to \(A_{k,0}\), we have for \(k = 0, 1, \ldots, m\) and \(i, j \in M\),

\[ [A_{k,0}(m)]_{i,j} = \Pr((L_{n+1}, S_{n+1}) = (0, j), L(\tau_{n+1}-) \leq m \mid (L_n, S_n) = (k, i)). \]

The above discussion makes it clear that our approach has two sources of errors: one is the truncation of \(U\) at level \(m\) and the other is the truncation of the weighted infinite sum obtained by uniformization. While the latter is controllable, the error control of the former is not straightforward because it requires information beyond level \(m\).

To control errors inherent in our approach, we adopt a cross-layer design of the computational procedure. Specifically, we assume that the augmented truncation approximation (ATA) is employed in Step 2, as we will explain below, where the truncation level \(K\) in the ATA is assumed to be given in advance. We then utilize this information in Step 1, i.e., \(K(ATA)\) is employed in Step 2, as we will explain below, where the truncation level 

\[ \text{Step 1:} \quad \text{Choose an appropriate truncation level} \quad K, \]

\[ \text{Step 2:} \quad \text{Compute the stationary distribution} \quad \kappa. \]

In Step 2, we compute the stationary distribution \(\kappa\) of the imbedded Markov chain \(\{(L_n, S_n)\}_{n=0,1,...}\). Owing to the level dependence of arrivals and disasters, however, it is hard to obtain the analytical expression of the stationary distribution \(\kappa\). We thus adopt the ATA, which attempts to compute an approximation \(\kappa^{\text{approx}}\) to the stationary distribution \(\kappa\) as follows. We first choose an appropriate level \(K\) and partition the stationary distribution \(\kappa\) as \(\kappa = (\kappa^{(1)}(K) \kappa^{(2)}(K))\), where

\[ \kappa^{(1)}(K) = (\kappa_0 \kappa_1 \ldots \kappa_K), \quad \kappa^{(2)}(K) = (\kappa_{K+1} \kappa_{K+2} \ldots). \]

In the ATA, we approximate \(\kappa^{(2)}(K)\) to be \(0\) and attempt to compute \(\kappa^{(1)}(K)\) based on the north-west corner submatrix \(P(K)\) of the transition probability matrix \(P\) in (1.2), where

\[ P(K) = \begin{pmatrix}
A_{0,0} & A_{0,1} & A_{0,2} & A_{0,3} & \cdots & A_{0,K-1} & A_{0,K} \\
A_{1,0} & A_{1,1} & A_{1,2} & A_{1,3} & \cdots & A_{1,K-1} & A_{1,K} \\
A_{2,0} & A_{2,1} & A_{2,2} & A_{2,3} & \cdots & A_{2,K-1} & A_{2,K} \\
A_{3,0} & O & A_{3,2} & A_{3,3} & \cdots & A_{3,K-1} & A_{3,K} \\
\vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\
A_{K-1,0} & O & O & O & \cdots & A_{K-1,K-1} & A_{K-1,K} \\
A_{K,0} & O & O & O & \cdots & A_{K,K-1} & A_{K,K} 
\end{pmatrix}. \]

Specifically, we compute an approximation \(\kappa^{\text{approx}}(K)\) to \(\kappa^{(1)}(K)\) by solving

\[ \kappa^{\text{approx}}(K) = \kappa^{\text{approx}}(K)[P(K) + P_A(K)], \quad \kappa^{\text{approx}}(K)e = 1, \]

(1.6)

where \(P_A(K)\) denotes an augmentation matrix such that \(P_A(K) \geq O\) and \([P(K) + P_A(K)]e = e\). We then adopt \(\kappa^{\text{approx}} = (\kappa^{\text{approx}}(K) 0)\) as an approximation to \(\kappa\), which is known to converge to \(\kappa\) as \(K\) goes to infinity under some technical conditions [14,16]. In the literature, upper bounds of the error \(\|\kappa^{\text{approx}} - \kappa\|_1\) for \(\kappa^{\text{approx}}\) are also studied under various drift conditions [14,15].

Let \(\kappa(K) = (\kappa_0(K) \kappa_1(K) \cdots \kappa_K(K))\) denote the conditional stationary distribution given that the level is not greater than \(K\), where \(\kappa_\ell(K) (\ell = 0, 1, \ldots, K)\) denotes a \(1 \times M\) vector whose \(j\)th \((j \in M)\) element is given by \(\lim_{n \to \infty} \Pr((L_n, S_n) = (\ell, j) \mid L_n \leq K)\).
Note here that $\kappa^{\text{approx}}(K)$ in (1.6) can be regarded as an approximation to the conditional stationary distribution $\kappa(K)$. If numerical errors in elements of $P(K)$ are negligible, the error in an approximation $\kappa^{\text{approx}} = (\kappa^{\text{approx}}(K) \mathbf{0})$ to $\kappa$ can be evaluated by the tail probability and the difference between the ATA solution $\kappa^{\text{approx}}(K)$ and $\kappa(K)$ [11, 19].

$$2\zeta(K) \leq \| \kappa - \kappa^{\text{approx}} \|_1 \leq 2\zeta(K) + \epsilon(K),$$

(1.7)

where $\| \cdot \|_1$ stands for the total variation norm or $\ell_1$-norm and

$$\zeta(K) = \lim_{n \to \infty} \Pr(L_n > K), \quad \epsilon(K) = \| \kappa^{\text{approx}}(K) - \kappa(K) \|_1.$$

The tail probability $\zeta(K)$ decreases monotonically as $K$ increases. Unfortunately, however, it is hard to analyze the tail probability in our model due to the level-dependence. We thus evaluate it numerically after obtaining the ATA solution. On the other hand, the error in an approximation $\kappa^{\text{approx}} = (\kappa^{\text{approx}}(K) \mathbf{0})$ to $\kappa$ can be evaluated by the tail probability and the difference between the ATA solution $\kappa^{\text{approx}}(K)$ and $\kappa(K)$ [11, 19]. In this paper, we discuss reasonable selections of the augmentation matrix $P_A(K)$, based on the results in [11, Section 2.4].

In Step 3, we compute the approximation $\pi^{\text{approx}} = (\pi_0^{\text{approx}} \pi_1^{\text{approx}} \cdots)$ to the stationary distribution $\pi$ using $\kappa^{\text{approx}} = (\kappa^{\text{approx}}(K) \mathbf{0})$. It follows from (1.3) that $\pi^{\text{approx}}(\ell = 0, 1, \ldots)$ is given in terms of $\kappa^{\text{approx}}(K)$:

$$\pi_0^{\text{approx}} = \frac{\kappa_0^{\text{approx}}(K)(-C_0)^{-1}}{c},$$

(1.8)

$$\pi_\ell^{\text{approx}} = \frac{\kappa_0^{\text{approx}}(K)(-C_0)^{-1}D_0W_{1,\ell}}{c} + \sum_{k=1}^{\min(\ell, K)} \frac{\kappa_k^{\text{approx}}(K)W_{k,\ell}}{c}, \quad \ell = 1, 2, \ldots,$$

(1.9)

where $c$ denotes the normalizing constant so that $\pi^{\text{approx}}c = 1$. Note here that $W_{k,\ell}$'s in (1.9) are given by block matrices in $W$ defined as

$$W = \begin{pmatrix}
I & O & O & O & O & \cdots \\
W_{1,0} & W_{1,1} & W_{1,2} & W_{1,3} & W_{1,4} & \cdots \\
W_{2,0} & O & W_{2,2} & W_{2,3} & W_{2,4} & \cdots \\
W_{3,0} & O & O & W_{3,3} & W_{3,4} & \cdots \\
W_{4,0} & O & O & O & W_{4,4} & \cdots \\
\vdots & \vdots & \vdots & \vdots & \vdots & \ddots
\end{pmatrix} = \int_0^\infty \exp[Ut](1 - H(t)) dt.$$

It then follows that

$$W = E[H]B_W,$$

where

$$B_W = \int_0^\infty \exp[Ut] \frac{1 - H(t)}{E[H]} dt.$$ (1.10)

Because $(1 - H(t))/E[H]$ is the probability density function of elapsed service times, $B_W$ can be computed in the same way as $B_A$ in (1.4).

If the numerical errors in $W_{k,\ell}$ ($k = 1, 2, \ldots, K, \ell = k, k+1, \ldots$) are negligible, the error in $\pi^{\text{approx}}$ can be evaluated as follows.

$$\| \pi - \pi^{\text{approx}} \|_1 \leq c_p \| \kappa - \kappa^{\text{approx}} \|_1 + o(\| \kappa - \kappa^{\text{approx}} \|_1),$$

(1.11)
Queues with LD Arrivals and Disasters

where \( c_p \) is a finite coefficient. (1.11) shows that we can control the error in \( \pi^{\text{approx}} \) through the error in \( \kappa^{\text{approx}} \).

In summary, to guarantee a sufficient accuracy of \( \pi^{\text{approx}} \), we have to pay attention to the followings.

1. \( K \) should be set appropriately so as to make the tail probability \( \zeta(K) \) negligible,
2. the numerical errors in \( A_{k,\ell} \)'s and \( W_{k,\ell} \)'s should be sufficiently small, and
3. the error \( \epsilon(K) \) in \( \kappa^{\text{approx}}(K) \) due to the selection of \( P_A(K) \) should be small as much as possible.

In this paper, we mainly consider the last two points. Through numerical examples, we also discuss the upper bound of the tail probability and an appropriate truncation level \( K \) in the ATA.

The rest of this paper is organized as follows. In Section 2, we consider the numerical implementation of \( A_{k,\ell} \)'s and \( W_{k,\ell} \)'s. In Section 3, we briefly discuss the selection of the augmentation matrix and the error bound in the ATA solution. In section 4, we discuss the error in the approximation to the stationary distribution. In Section 5, we show some numerical examples. Finally, we conclude this paper in Section 6.

2. Computation of \( A_{k,\ell} \)'s and \( W_{k,\ell} \)'s

As stated in the preceding section, we will adopt the ATA in Step 2. We thus consider the computation of the north-west corner submatrix \( P(K) \) of \( P \) for a given \( K \). By definition, block matrices \( A_{k,\ell} \)'s \((k, \ell = 0, 1, \ldots, K)\) in \( P(K) \) are given in terms of \( M \times M \) block matrices in \( B_A \) defined by (1.4).

\[
A_{k,\ell} = \begin{cases} 
(-C_0)^{-1}D_0A_{1,\ell}, & k = 0, \ell = 0, 1, \ldots, K, \\
B_{A,1,0} + B_{A,1,1}, & k = 1, \ell = 0, \\
B_{A,1,\ell+1}, & k = 1, \ell = 1, 2, \ldots, K, \\
B_{A,k,0}, & k = 2, 3, \ldots, K, \ell = 0, \\
B_{A,k,\ell+1}, & k = 2, 3, \ldots, K, \ell = k - 1, k, \ldots, K, \\
O, & \text{otherwise}.
\end{cases}
\]  

(2.1)

Therefore, we consider the computation of \( B_{A,k,\ell} \)'s below. We also discuss the computation of \( W_{k,\ell} \) in this section because \( W = E[H]B_W \) and block matrices in \( B_W \) can be computed in the same way as those in \( B_A \).

2.1. The uniformizable case

We first consider the case that \( \exp[Ut] \) is uniformizable, i.e., there exists a finite \( \theta \) such that \([-C_k]_{i,i} \leq \theta \) \((k = 1, 2, \ldots; i \in \mathcal{M})\). In this case, by applying the uniformization technique to \( B_A \) in (1.4) and \( B_W \) in (1.10), we can rewrite them to be

\[
B_A = \int_0^\infty e^{-\theta t} \exp[(I + \theta^{-1}U)\theta t]dH(t) = \sum_{n=0}^\infty \gamma_n V^n,
\]

\[
B_W = \int_0^\infty e^{-\theta t} \exp[(I + \theta^{-1}U)\theta t] \frac{1 - H(t)}{E[H]} dt = \sum_{n=0}^\infty \eta_n V^n,
\]

where \( \gamma_n \) and \( \eta_n \) \((n = 0, 1, \ldots)\) denote probability functions given by

\[
\gamma_n = \int_0^\infty e^{-\theta t} \frac{(\theta t)^n}{n!} dH(t), \quad \eta_n = \int_0^\infty e^{-\theta t} \frac{(\theta t)^n}{n!} \frac{1 - H(t)}{E[H]} dt,
\]
and \( V = I + \theta^{-1}U \geq O \). Note here that \( V^0 = I \) and \( V\mathbf{e} = \mathbf{e} \). Block matrices \( V^{(n)}_{k,\ell} \) in \( V^n \) \((n = 2, 3, \ldots)\) can be computed recursively by \( V^{(1)}_{k,\ell} = V_{k,\ell} \) and for \( n = 2, 3, \ldots \),

\[
V^{(n)}_{k,\ell} = \begin{cases} 
V^{(n-1)}_{k,\ell} + \sum_{r=k}^{k+n-1} V^{(n-1)}_{k,r} \theta^{-1} \Gamma_r, & \ell = 0, \\
V^{(n-1)}_{k,k} (I + \theta^{-1} C_k), & \ell = k, \\
V^{(n-1)}_{k,k} \theta^{-1} D_{\ell-1} + V^{(n-1)}_{k,\ell} \theta^{-1} (I + \theta^{-1} C_\ell), & \ell = k + 1, k + 2, \ldots, K + 1, \\
O, & \text{otherwise.}
\end{cases}
\]

In numerical implementation, we assume that the recursion of \( V^{(n)}_{k,\ell} \) for \( B_{A,k,\ell} \) stops at \( n = n_A \) and that for \( B_{W,k,\ell} \) at \( n = n_W \).

\[
B_{A}^{\text{comp},n_A} = \sum_{n=0}^{n_A} \gamma_n V^n, \quad B_{W}^{\text{comp},n_W} = \sum_{n=0}^{n_W} \eta_n V^n.
\]

We then adopt \( B_{A}^{\text{comp},n_A} \) as an approximation to \( B_{A} \) and \( B_{W}^{\text{comp},n_W} \) as an approximation to \( B_{W} \), where \( O \leq B_{A}^{\text{comp},n_A} \leq B_{A} \) and \( O \leq B_{W}^{\text{comp},n_W} \leq B_{W} \). In what follows, we determine \( n_A = n_A^* \) and \( n_W = n_W^* \) separately in such a way that for \( k = 0, 1, \ldots, K \),

\[
\sum_{\ell=0}^{K} (A_{k,\ell} - A_{k,\ell}^{\text{comp},n_A^*}) \mathbf{e} \leq \epsilon_A \mathbf{e}, \quad \sum_{\ell=k}^{\infty} (W_{k,\ell} - W_{k,\ell}^{\text{comp},n_W^*}) \mathbf{e} \leq \epsilon_W \mathbf{e}, \tag{2.2}
\]

where \( \epsilon_A \) and \( \epsilon_W \) are sufficiently small positive constants.

It follows from (2.1) that the inequality for \( A_{k,\ell}^{\text{comp},n_A^*} \) \((k = 0, 1, \ldots, K)\) in (2.2) is equivalent to

\[
\sum_{\ell=0}^{K+1} (B_{A,k,\ell} - B_{A,k,\ell}^{\text{comp},n_A^*}) \mathbf{e} \leq \epsilon_A \mathbf{e}, \quad k = 1, 2, \ldots, K. \tag{2.3}
\]

Note here that

\[
\sum_{\ell=0}^{K+1} (B_{A,k,\ell} - B_{A,k,\ell}^{\text{comp},n_A^*}) \mathbf{e} \leq \sum_{\ell=0}^{\infty} (B_{A,k,\ell} - B_{A,k,\ell}^{\text{comp},n_A^*}) \mathbf{e} = \sum_{n=n_A+1}^{\infty} \gamma_n \sum_{\ell=0}^{\infty} V^{(n)}_{k,\ell} \mathbf{e} = \sum_{n=n_A+1}^{\infty} \gamma_n \mathbf{e}.
\]

Therefore, we set \( n_A = n_A^* \) in such a way that

\[
\sum_{n=0}^{n_A^*} \gamma_n \geq 1 - \epsilon_A,
\]

which guarantees (2.3).

Similarly, we have for \( W_{k,\ell}^{\text{comp},n_W^*} \),

\[
\sum_{\ell=k}^{\infty} (W_{k,\ell} - W_{k,\ell}^{\text{comp},n_W^*}) \mathbf{e} \leq \sum_{\ell=0}^{\infty} (W_{k,\ell} - W_{k,\ell}^{\text{comp},n_W^*}) \mathbf{e}
\]

\[
= \mathbb{E}[H] \sum_{n=n_W+1}^{\infty} \eta_n \sum_{\ell=0}^{\infty} V^{(n)}_{k,\ell} \mathbf{e} = \mathbb{E}[H] \sum_{n=n_W+1}^{\infty} \eta_n \mathbf{e}.
\]
Therefore, we set \( n_W = n_W^* \) in such a way that
\[
\sum_{n=0}^{n_W^*} \eta_n \geq 1 - \frac{\epsilon_W}{E[H]},
\]
which guarantees the inequality for \( W_{k,k+\ell}^{\text{comp},n_W^*} \) in (2.2). Note that for each \( k = 1, 2, \ldots, K \), the computation of \( B_{W,k,\ell}^{\text{comp},n_W^*} \) automatically stops at \( \ell = k + n_W^* \) and \( B_{W,k,\ell}^{\text{comp},n_W^*} = 0 \) for all \( \ell > k + n_W^* \).

### 2.2. The non-uniformizable case

In this subsection, we consider the computation of approximations to \( A_{k,\ell} \)'s and \( W_{k,\ell} \)'s under the assumption that \( \exp[U] \) is non-uniformizable, i.e.,
\[
\sup_{k=0,1,\ldots} \left( \max_{i\in M} [-C_k]_{i,i} \right) = \infty. \tag{2.4}
\]

We define \( U(m) \) \( (m = 1, 2, \ldots) \) as
\[
U(m) = \begin{pmatrix}
O & O & O & O & \cdots & O & O \\
\Gamma_1 & C_1 & D_1 & O & \cdots & O & O \\
\Gamma_2 & O & C_2 & D_2 & \cdots & O & O \\
\Gamma_3 & O & O & C_3 & \cdots & O & O \\
\vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\
\Gamma_{m-1} & O & O & O & \cdots & C_{m-1} & D_{m-1} \\
\Gamma_m & O & O & O & \cdots & O & C_m
\end{pmatrix}, \quad m = 1, 2, \ldots.
\]

By definition, \( U(m) \) \( (m = 1, 2, \ldots) \) is a defective infinitesimal generator. Let \( \theta_m \) \( (m = 1, 2, \ldots) \) denote the maximum of the absolute values of the diagonal elements of \( U(m) \).
\[
\theta_m = \max_{k=1,2,\ldots,m} [-C_k]_{i,i}, \quad m = 1, 2, \ldots.
\]

In what follows, we first consider \( A_{k,\ell} \)'s and then consider \( W_{k,\ell} \)'s.

We assume that \( U \) is truncated at level \( m_\Lambda \) in computing \( A_{k,\ell} \)'s, where \( m_\Lambda \geq K + 1 \). We then define \( B_A(m_\Lambda) \) as
\[
B_A(m_\Lambda) = \int_0^\infty \exp[U(m_\Lambda)t]dH(t) = \sum_{n=0}^\infty \gamma_n(m_\Lambda)V^n(m_\Lambda), \tag{2.5}
\]
where \( \gamma_n(m_\Lambda) \) \( (n = 0, 1, \ldots) \) denotes a probability function given by
\[
\gamma_n(m_\Lambda) = \int_0^\infty e^{-\theta_m t}\frac{\theta_m^nt^n}{n!}dH(t),
\]
and \( V(m_\Lambda) = I + \theta_m^{-1}U(m_\Lambda) \). Note that \( V(m_\Lambda) \) is substochastic, i.e., \( V(m_\Lambda)e \leq e \).

For any \( m = 1, 2, \ldots, \) block matrices \( V_{k,\ell}^{(n)} \) in \( V^n(m) \) \( (n = 1, 2, \ldots) \) can be computed recursively by \( V_{k,\ell}^{(1)}(m) = V_{k,\ell}(m) \) and for \( n = 2, 3, \ldots, \)
\[
V_{k,\ell}^{(n)}(m) = \begin{cases}
V_{k,\ell}^{(n-1)}(m) + \sum_{r=k}^{\min(m,k+n-1)} V_{k,r}^{(n-1)}(m)\theta_m^{-1}\Gamma_r, & \ell = 0, \\
V_{k,k}^{(n-1)}(m)(I + \theta_m^{-1}C_k), & \ell = k, \\
V_{k,k-1}^{(n-1)}(m)\theta_m^{-1}D_{\ell-1} + V_{k,\ell}^{(n-1)}(m)(I + \theta_m^{-1}C_\ell), & \ell = k + 1, k + 2, \ldots, \min(m,k+n), \\
O, \quad \text{otherwise},
\end{cases} \tag{2.6}
\]
In numerical implementation, we have to truncate the infinite sum in (2.5), i.e.,

\[
B_{A,\ell}^{\text{comp},n_{A}}(m_{A}) = \sum_{n=0}^{n_{A}} \gamma_{n}(m_{A}) V_{k,0}^{(n)}(m_{A}).
\]

We adopt \(B_{A,k,\ell}^{\text{comp},n_{A}}(m_{A})\) as an approximation to \(B_{A,k,\ell}\). Specifically, for \(k = 1, 2, \ldots, K\), an approximation \(B_{A,k,\ell}^{\text{comp},n_{A}}(m_{A})\) to \(B_{A,k,\ell}\) is given by

\[
B_{A,k,\ell}^{\text{comp},n_{A}}(m_{A}) = \begin{cases} 
\sum_{n=0}^{n_{A}} \gamma_{n}(m_{A}) V_{k,0}^{(n)}(m_{A}), & \ell = 0, \\
\sum_{n=\ell-k}^{n_{A}} \gamma_{n}(m_{A}) V_{k,0}^{(n)}(m_{A}), & \ell = k, k+1, \ldots, \min(m_{A}, k + n_{A}), \\
O, & \text{otherwise.}
\end{cases}
\]

Furthermore, it follows from (2.1) that an approximation \(A_{k,\ell}^{\text{comp},n_{A}}(m_{A})\) to \(A_{k,\ell}\) is given by

\[
A_{k,\ell}^{\text{comp},n_{A}}(m_{A}) = \begin{cases} 
(-C_{0})^{-1} D_{0} A_{k,\ell}^{\text{comp},n_{A}}(m_{A}), & k = 0, \ell = 0, 1, \ldots, m_{A} - 1, \\
B_{A,1,1}^{\text{comp},n_{A}}(m_{A}), & k = 0, \ell = 0, \\
B_{A,1,\ell+1}^{\text{comp},n_{A}}(m_{A}), & k = 1, \ell = 1, 2, \ldots, m_{A} - 1, \\
B_{A,k,\ell}^{\text{comp},n_{A}}(m_{A}), & k = 2, 3, \ldots, K, \ell = 0, \\
B_{A,k,\ell+1}^{\text{comp},n_{A}}(m_{A}), & k = 2, 3, \ldots, K, \ell = k - 1, \ldots, m_{A} - 1, \\
O, & \text{otherwise.}
\end{cases}
\]

Note here that

\[A_{k,\ell}^{\text{comp},n_{A}}(m_{A}) \leq A_{k,\ell}, \quad k = 0, 1, \ldots, K, \ell = 0, 1, \ldots, m_{A} - 1.\]  

(2.9)

For a given \(\epsilon_{A} > 0\), we would like to find \(m_{A} = m_{A}^{*}\) and \(n_{A} = n_{A}^{*}\) such that

\[\sum_{k=0}^{K} (A_{k,\ell} - A_{k,\ell}^{\text{comp},n_{A}}(m_{A}^{*})) e \leq \epsilon_{A} e, \quad k = 0, 1, \ldots, K.\]  

(2.10)

Remark 2.1. It follows from (2.9) that for \(k = 0, 1, \ldots, K\),

\[
\sum_{\ell=0}^{K} (A_{k,\ell} - A_{k,\ell}^{\text{comp},n_{A}}(m_{A})) e \leq \sum_{\ell=0}^{m_{A}-1} (A_{k,\ell} - A_{k,\ell}^{\text{comp},n_{A}}(m_{A})) e \\
\leq \sum_{\ell=0}^{\infty} A_{k,\ell} e - \sum_{\ell=0}^{m_{A}-1} A_{k,\ell}^{\text{comp},n_{A}}(m_{A}) e = e - \sum_{\ell=0}^{m_{A}-1} A_{k,\ell}^{\text{comp},n_{A}}(m_{A}) e.
\]

(2.11)

Therefore, if \(m_{A}\) and \(n_{A}\) are fixed, we can evaluate the upper bound of the left-hand side of (2.10) by computing \(A_{k,\ell}^{\text{comp},n_{A}}(m_{A})\) \((k = 0, 1, \ldots, K, \ell = 0, k, k + 1, \ldots, m_{A})\).

This approach has two potential sources of errors: the truncation of \(U\) at level \(m_{A}\) and the truncation of the infinite sum in (2.5) at \(n = n_{A}\). Note here that

\[
\sum_{\ell=0}^{m_{A}} V_{k,\ell}^{(n)}(m_{A}) e = e, \quad k = 1, 2, \ldots, K, \quad n = 0, 1, \ldots, m_{A} - k.
\]
Therefore, if $n_A \leq m_A - K$,

$$
\sum_{\ell=0}^{m_A-1} A_{k,\ell}^{\text{comp},n_A}(m_A) e = \sum_{\ell=0}^{m_A} B_{k,\ell}^{\text{comp},n_A} e = \sum_{\ell=0}^{m_A} \sum_{n=0}^{n_A} \gamma_n(m_A) V_{k,\ell}^{(n)}(m_A) e = \sum_{n=0}^{n_A} \gamma_n(m_A) e,
$$

for all $k = 1, 2, \ldots, K$. It then follows from (2.11) that

$$
\sum_{\ell=0}^{K} ( A_{k,\ell} - A_{k,\ell}^{\text{comp},n_A}(m_A)) e \leq \left( 1 - \sum_{n=0}^{n_A} \gamma_n(m_A) \right) e \quad \text{if} \quad n_A \leq m_A - K. \quad (2.12)
$$

Because the right-hand side of (2.12) takes the minimum value at $n_A = m_A - K$, (2.10) would hold if we could find $m_A = m_A^*$ such that

$$
\left( 1 - \sum_{n=0}^{m_A^*-K} \gamma_n(m_A^*) \right) e = \Pr(N_{\theta_{m_A^*}}(H) > m_A^* - K) \leq \epsilon_A,
$$

where $N_{\theta_{m_A^*}}(H)$ denotes the number of Poisson arrivals with rate $\theta_{m_A^*}$ during a randomly chosen service time $H$.

Note, however, that $m_A^*$ satisfying (2.13) does not necessary exist in general. The reason is that (i) for a fixed $x$, $\Pr(N_{\theta_{m_A}}(H) > x)$ is an increasing function of $\theta_{m_A}$ and (ii) $\theta_{m_A}$ is a nondecreasing function of $m_A$, which imply that $\Pr(N_{\theta_{m_A}}(H) > m_A - K)$ may or may not decrease as $m_A$ increases. Besides, $\Pr(N_{\theta_{m_A}}(H) > x)$ depends on the distribution of $H$. For example, if $\theta_{m_A}/m_A$ is equal to zero in the limit $m_A \to \infty$, it can be shown that $m_A - K > E[N_{\theta_{m_A}}(H)] = \theta_{m_A} E[H]$ for a sufficient large $m_A$ and in this case, the one-side Chebyshev's inequality (also called Cantelli's inequality) [20, pp.198–199] yields

$$
\Pr(N_{\theta_{m_A}}(H) > m_A - K) \leq \frac{\theta_{m_A}^2 \Var[H]}{\theta_{m_A}^2 \Var[H] + \theta_{m_A} E[H] + (m_A - K - \theta_{m_A} E[H])^2}, \quad (2.14)
$$

where the variance $\Var[H]$ of service times is assumed to be finite. Furthermore, it can be readily verified that the right-hand side of (2.14) is equal to zero in the limit $m_A \to \infty$ if $\theta_{m_A}/m_A$ is equal to zero in the limit $m_A \to \infty$. Another example is that if the service time distribution belongs to a certain class of long-tailed distributions, we have [1, Eq. (1.1)]

$$
\lim_{x \to \infty} \frac{\Pr(N_{\theta_{m_A}}(H) > x)}{\Pr(H > x/\theta_{m_A})} = 1.
$$

These examples suggest that for the existence of $m_A = m_A^*$ satisfying (2.13), we need some additional assumptions on the service time distribution and/or the asymptotic property of the sequence of $\theta_{m_A}$’s. In what follows, we develop a numerical procedure that always works under Assumption 1.1 introduced in Section 1.

**Theorem 2.1.** Consider $\{(L(t), S(t))\}_{t \geq 0}$ whose arrival/disaster process is non-explosive, i.e., $\sum_{t=0}^{\infty} A_{k,\ell} e = e$. We then have for $k = 0, 1, \ldots, K$,

$$
\sum_{\ell=0}^{K} (A_{k,\ell} - A_{k,\ell}^{\text{comp},n_A}(m_A)) e \leq \max_{i \in M} \left( b^+_{A,(k,i)}(m_A) \right) e + \left( 1 - \sum_{n=0}^{n_A} \gamma_n(m_A) \right) e, \quad (2.15)
$$

where $b^+_{A,(k,i)}(m_A)$ ($k = 0, 1, \ldots, K; i \in M$) is defined as

$$
b^+_{A,(k,i)}(m_A) = \Pr(L(\tau_{n+1}) > m_A \mid (L_n, S_n) = (k, i)). \quad (2.16)
$$
Proof. Note first that for an arbitrary $\epsilon > 0$, 
\[
\sum_{\ell=0}^{K} (A_{1,\ell} - A_{1,\ell}^{\text{comp},n \Lambda}(m_{\Lambda}))e \leq \epsilon e \Rightarrow \sum_{\ell=0}^{K} (A_{0,\ell} - A_{0,\ell}^{\text{comp},n \Lambda}(m_{\Lambda}))e \leq (-C_{0})^{-1}D_{0} \cdot \epsilon e = \epsilon e.
\]
We thus consider (2.15) for $k = 1, 2, \ldots, K$.

It follows from (2.1) and $B_{A,k,\ell}^{\text{comp},n \Lambda}(m_{\Lambda}) \leq B_{A,k,\ell}(k = 1, 2, \ldots, K, \ell = 0, 1, \ldots, m_{\Lambda})$ that for $k = 1, 2, \ldots, K$, 
\[
\sum_{\ell=0}^{K} (A_{k,\ell} - A_{k,\ell}^{\text{comp},n \Lambda}(m_{\Lambda}))e = \sum_{\ell=0}^{K} (B_{A,k,\ell} - B_{A,k,\ell}^{\text{comp},n \Lambda}(m_{\Lambda}))e \leq \sum_{\ell=0}^{\infty} (B_{A,k,\ell} - B_{A,k,\ell}^{\text{comp},n \Lambda}(m_{\Lambda}))e;
\]
where $B_{A,k,\ell}^{\text{comp},n \Lambda}(m_{\Lambda}) = O$ for $\ell > m_{\Lambda}$. In what follows, we will show for $k = 1, 2, \ldots, K$, 
\[
\sum_{\ell=0}^{\infty} (B_{A,k,\ell} - B_{A,k,\ell}^{\text{comp},n \Lambda}(m_{\Lambda}))e \leq \max_{i \in M} (b_{A,k,i}^{+}(m_{\Lambda}))e + \left(1 - \sum_{n=0}^{m_{\Lambda}} \gamma_{n}(m_{\Lambda})\right)e. \tag{2.17}
\]
By definition, $B_{A,k,\ell}(m_{\Lambda}) = B_{A,k,\ell}(k = 1, 2, \ldots, K, \ell = 1, 2, \ldots, m_{\Lambda})$, where $m_{\Lambda} \geq K + 1$. It then follows that for $k = 1, 2, \ldots, K$, 
\[
\sum_{\ell=0}^{\infty} (B_{A,k,\ell} - B_{A,k,\ell}^{\text{comp},n \Lambda}(m_{\Lambda}))e = (B_{A,k,0} - B_{A,k,0}(m_{\Lambda}))e + \sum_{\ell=0}^{m_{\Lambda}} (B_{A,k,\ell}(m_{\Lambda}) - B_{A,k,\ell}^{\text{comp},n \Lambda}(m_{\Lambda}))e + \sum_{\ell=m_{\Lambda}+1}^{\infty} B_{A,k,\ell}e. \tag{2.18}
\]
Note here that for $k = 0, 1, \ldots, K$ and $j \in M$, 
\[
\left[(B_{A,k,0} - B_{A,k,0}(m_{\Lambda}))e + \sum_{\ell=m_{\Lambda}+1}^{\infty} B_{A,k,\ell}e\right]_{j} = b_{A,k,i}^{+}(m_{\Lambda}) = \max_{i \in M} (b_{A,k,i}^{+}(m_{\Lambda})). \tag{2.19}
\]
On the other hand, we have for $k = 1, 2, \ldots, K$, 
\[
\sum_{\ell=0}^{m_{\Lambda}} (B_{A,k,\ell}(m_{\Lambda}) - B_{A,k,\ell}^{\text{comp},n \Lambda}(m_{\Lambda}))e = \sum_{\ell=0}^{m_{\Lambda}} \sum_{n=n_{\Lambda}+1}^{\infty} \gamma_{n}(m_{\Lambda})V_{k,\ell}^{(n)}(m_{\Lambda})e \leq \sum_{n=n_{\Lambda}+1}^{\infty} \gamma_{n}(m_{\Lambda})e. \tag{2.20}
\]
(2.17) now follows from (2.18), (2.19), and (2.20), which completes the proof. \hfill \square

By definition, $\max_{i \in M} (b_{A,k,i}^{+}(m_{\Lambda}))$ on the right-hand side of (2.15) monotonically converges to zero as $m_{\Lambda}$ goes to infinity and for a fixed $m_{\Lambda}$, the second term also converges to zero monotonically as $n_{\Lambda}$ goes to infinity. Under Assumption 1.1 (i), we have the following theorem, where 
\[
b_{A,k}^{+}(m_{\Lambda}) = (b_{A,k,k}^{+}(m_{\Lambda}) b_{A,k,k+1}^{+}(m_{\Lambda}) \cdots b_{A,k,M}^{+}(m_{\Lambda}))^{T}.
\]
**Theorem 2.2.** We consider \( \{(L(t), S(t))\}_{t \geq 0} \), where (2.4) is assumed to hold. For arbitrary positive integers \( K \) and \( m_A \) \( (m_A \geq K + 1) \), we have

\[
\mathbf{b}_{A,K}^+(m_A) \leq (-C_K)^{-1}D_K(-C_{K+1})^{-1}D_{K+1} \cdots (-C_{m_A})^{-1}D_{m_A}e. \tag{2.21}
\]

Furthermore, for arbitrary \( K \) and \( \epsilon > 0 \), there exists \( m_A = m_A^* \) such that

\[
(-C_K)^{-1}D_K(-C_{K+1})^{-1}D_{K+1} \cdots (-C_{m_A})^{-1}D_{m_A}e \leq \epsilon e, \tag{2.22}
\]

under Assumption 1.1 (i).

**Proof.** We first prove (2.21) by a probabilistic argument, even though it can also be shown algebraically. It follows from (2.16) that

\[
b_{A,(K,i)}^+(m_A) = \Pr(L(\tau_{n+1}) > m_A \mid (L_n, S_n) = (K, i))
= \Pr(L(t) = m_A + 1 \text{ for some } t \in (\tau_n, \tau_{n+1}) \mid (L_n, S_n) = (K, i)).
\]

Note here that \( \tau_{n+1} = \min(d_{n+1}, \tau_n + H_n) \), where \( d_{n+1} \) denotes the first occurrence time of a disaster after time \( \tau_n \) and \( H_n \) denotes the service time starting at time \( \tau_n \). Let \( \tilde{N}(s, t) \) \((s < t)\) denote the number of arrivals during an interval \((s, t)\) under the assumption that the service never completes. We then have

\[
b_{A,(K,i)}^+(m_A) = \Pr(L(t) = m_A + 1 \text{ for some } t \in (\tau_n, \tau_{n+1}) \mid (L_n, S_n) = (K, i))
\leq \Pr(L(t) = m_A + 1 \text{ for some } t \in (\tau_n, d_{n+1}) \mid (L_n, S_n) = (K, i)) = \Pr(\tilde{N}(\tau_n, d_{n+1}) > m_A - K \mid (L_n, S_n) = (K, i)) = \left[ (-C_K)^{-1}D_K(-C_{K+1})^{-1}D_{K+1} \cdots (-C_{m_A})^{-1}D_{m_A}e \right]_i,
\]

from which (2.21) follows.

Next we consider (2.22) under Assumption 1.1 (i). It follows from (1.1) and Assumption 1.1 (i), i.e., \( \Gamma_k e > \{1/k\} \cdot D_k e \) \((k = K_1, K_1^* + 1, \ldots)\), that \((-C_k)^{-1}D_k e < \{k/(k + 1)\} \cdot e \) \((k = K_1, K_1^* + 1, \ldots)\). We then have

\[
\lim_{m_A \to \infty} (-C_k^*)^{-1}D_k^*(-C_{k+1}^*)^{-1}D_{k+1}^* \cdots (-C_{m_A}^*)^{-1}D_{m_A}^* e
< \lim_{m_A \to \infty} \frac{K_1^*}{K_1^* + 1} \cdot \frac{K_1^* + 1}{K_1^* + 2} \cdots \frac{m_A}{m_A + 1} \cdot e = \lim_{m_A \to \infty} \frac{K_1^*}{m_A + 1} \cdot e = 0,
\]

so that for an arbitrary \( \epsilon \) \((0 < \epsilon < 1)\), \( m_A = m_A^* \) satisfying (2.22) exists. \( \square \)

Next we consider Assumption 1.1 (ii). Preliminary to it, we show the following lemma, where for any vector \( \mathbf{x} \), we define \( \text{diag}(\mathbf{x}) \) as a diagonal matrix whose ith diagonal element is given by the ith element of \( \mathbf{x} \).

**Lemma 2.1.** Let \( C_k^2 = C_k + \text{diag}(\Gamma_k e) \).

(i) For any \( t \geq 0 \), we have

\[
\exp[C_k t] \leq \exp[C_k^2 t], \quad k = 1, 2, \ldots. \tag{2.23}
\]

(ii) For any \( M \times 1 \) vector \( \mathbf{x} \geq 0 \), we have

\[
(-C_k^2)^{-1}D_k = [-(C_k^2 - \text{diag}(\mathbf{x}))]^{-1}[D_k + \text{diag}(\mathbf{x})(-C_k^2)^{-1}D_k]. \tag{2.24}
\]
(iii) For any $M \times 1$ vector $x \geq 0$ and $\tau \geq 0$, we have

$$\int_0^\tau \exp[C_k^t] D_k dt \leq \int_0^\tau \exp [(C_k^t - \text{diag}(x)) t] [D_k + \text{diag}(x)(-C_k^t)^{-1} D_k] dt.$$  \hspace{1cm} (2.25)

**Proof.** We first consider (2.23). Since $\text{diag}(\Gamma_k e) \geq O$, we have $C_k \leq C_k + \text{diag}(\Gamma e) = C_k^e$. It then follows that $[I + \theta_m^{-1} C_k]^n \leq [I + \theta_m^{-1} C_k^e]^n$ and therefore

$$\exp[C_k t] = e^{-\theta_m t} \sum_{n=0}^\infty \frac{(\theta_m t)^n}{n!} [I + \theta_m^{-1} C_k]^n \leq e^{-\theta_m t} \sum_{n=0}^\infty \frac{(\theta_m t)^n}{n!} [I + \theta_m^{-1} C_k^e]^n = \exp[C_k^e t].$$

Furthermore, (2.24) follows from $[-(C_k^e - \text{diag}(x))](C_k^e)^{-1} D_k = D_k + \text{diag}(x)(-C_k^e)^{-1} D_k$. Finally, (2.25) follows from

$$\int_0^\tau \exp [(C_k^t - \text{diag}(x)) t] [D_k + \text{diag}(x)(-C_k^t)^{-1} D_k] dt$$

$$= (I - \exp[(C_k^e - \text{diag}(x)) \tau]) [- (C_k^e - \text{diag}(x))]^{-1} [D_k + \text{diag}(x)(-C_k^e)^{-1} D_k]$$

$$= (I - \exp[(C_k^e - \text{diag}(x)) \tau]) (-C_k^e)^{-1} D_k$$

$$\geq (I - \exp[C_k^e \tau])(-C_k^e)^{-1} D_k = \int_0^\tau \exp[C_k^e t] D_k dt,$$

where (2.24) is used in the second equality and the inequality can be shown in the same way as (2.23) because $(-C_k^e)^{-1} D_k \geq O$. \hfill \Box

**Theorem 2.3.** We consider $\{(L(t), S(t))\}_{t \geq 0}$ under Assumption 1.1 (ii), where (2.4) is assumed to hold. For an arbitrary integer $K > 0$, we have

$$\max_{i \in M} (b_{A, (K,i)}^+(m_A)) \leq 1 - \sum_{n=0}^{m_A - \max(K, K_2^e)} \int_0^\infty e^{-\mu t} \frac{(\mu t)^n}{n!} dH(t), \quad m_A \geq \max(K, K_2^e),$$

where $\mu = 1/E[H]$ and $K_2^e$ is given in Assumption 1.1 (ii).

Because the proof of Theorem 2.3 is a bit lengthy, it is given in Appendix A. Theorems 2.1–2.3 enable us to compute the approximation $A_{k, \ell}^{\text{approx}, m_A}(k, \ell \leq K)$ to $A_{k, \ell}$, which satisfies (2.10) under Assumption 1.1. We summarize the procedures in Figure 1.

Next, we consider the computation of an approximation to $W_{k, \ell}$'s using $U(m_W)$. We first define $W(m_W)$ ($m_W > K$) as

$$W(m_W) = \int_0^\infty \exp[U(m_W) t](1 - H(t)) dt = E[H] \cdot B_W(m_W),$$

where

$$B_W(m_W) = \int_0^\infty \exp[U(m_W) t] \frac{1 - H(t)}{E[H]} dt.$$  \hspace{1cm} (2.26)

Note here that $(1 - H(t))/E[H]$ is the probability density function of the equilibrium random variable for service times and therefore, we have

$$B_W(m_W) = \sum_{n=0}^\infty \eta_n(m_W) V^n(m_W).$$
where \( \eta(m_W) \) (\( n = 0, 1, \ldots \)) denotes a probability function given by

\[
\eta(m_W) = \int_{0}^{\infty} e^{-\theta m_W t} (\theta m_W t)^n \frac{1 - H(t)}{n!} dH(t) dH(t),
\]

and \( V(m_W) = I + \theta^{-1} m_W U(m_W) \). In numerical implementation, we truncate the infinite sum in (2.26) at \( n = n_W \). Then

\[
B^{\text{comp}, n_W}_{n_W} = \sum_{n=0}^{n_W} \eta(m_W) V^n(m_W).
\]

It then follows that for \( k = 1, 2, \ldots, K \), an approximation \( W^{\text{comp}, n_W}_{k, \ell}(m_W) \) to \( W_{k, \ell} \) is given by

\[
W^{\text{comp}, n_W}_{k, \ell}(m_W) = \begin{cases} 
\text{E}[H] \sum_{n=1}^{n_W} \eta(m_W) V^{(n)}_{k,0}(m_W), & \ell = 0, \\
\text{E}[H] \sum_{n=\ell-k}^{n_W} \eta(m_W) V^{(n)}_{k,\ell}(m_W), & \ell = k, k+1, \ldots, \min(m_W, k+n_W), \\
O, & \text{otherwise}.
\end{cases}
\]

(2.27)
We would like to find \( m_W = m_W^* \) and \( n_W = n_W^* \) such that for a given \( \epsilon_W > 0 \),

\[
\sum_{\ell=k}^{\infty} (W_{k,\ell} - W_{k,\ell}^{\text{comp},n_W}(m_W)) e = E[H] \sum_{\ell=k}^{\infty} (B_{W,k,\ell} - B_{W,k,\ell}^{\text{comp},n_W}(m_W)) e \leq \epsilon_W e, \tag{2.28}
\]

for all \( k = 1, 2, \ldots, K \). Note here that \( B_{W,k,\ell}^{\text{comp},n_W}(m_W) = O \) for all \( \ell > \min(m_W, k + n_W) \).

Since \( B_{W,k,\ell}^{\text{comp},n_W}(m_W) \leq B_{W,k,\ell} \), we have

\[
\sum_{\ell=0}^{\infty} (B_{W,k,\ell} - B_{W,k,\ell}^{\text{comp},n_W}(m_W)) e \leq \sum_{\ell=0}^{\infty} (B_{W,k,\ell} - B_{W,k,\ell}^{\text{comp},n_W}(m_W)) e.
\]

Therefore, (2.28) holds if

\[
\sum_{\ell=0}^{\infty} (B_{W,k,\ell} - B_{W,k,\ell}^{\text{comp},n_W}(m_W)) e \leq \frac{\epsilon_W}{E[H]} e, \quad k = 1, 2, \ldots, K. \tag{2.29}
\]

Because the left-hand side of (2.29) takes the same form as that of (2.17), we can obtain the following corollary according to the same line of discussions as for Theorems 2.1–2.3.

**Corollary 2.1.** For arbitrary \( K, m_W (0 < K < m_W) \) and \( n_W > 0 \), we have

\[
\sum_{\ell=0}^{\infty} (B_{W,k,\ell} - B_{W,k,\ell}^{\text{comp},n_W}(m_W)) e \leq \max_{i \in M} (b^+_W(K,i) (m_A)) e + \left( 1 - \sum_{n=0}^{n_W} \eta_n(m_W) \right) e,
\]

where \( b^+_W(K) = (b^+_W(K,1)(m_W) b^+_W(K,2)(m_W) \cdots b^+_W(K,M)(m_W))^T \) is given by

\[
b^+_W(K)(m_W) = (B_{W,k,0} - B_{W,k,0}(m_W)) e + \sum_{\ell=m_W+1}^{\infty} B_{W,k,\ell} e.
\]

Under Assumption 1.1 (i),

\[
b^+_W(K)(m_W) \leq \epsilon e, \tag{2.30}
\]

holds for a given \( \epsilon > 0 \) if \( m_W \) satisfies \((-C_K)^{-1}D_K (-C_{K+1})^{-1}D_{K+1} \cdots (-C_{m_W})^{-1}D_{m_W} e \leq \epsilon e \). On the other hand, under Assumption 1.1 (ii), (2.30) holds for a given \( \epsilon > 0 \) if \( m_W \) satisfies \( m_W \geq \max(K, K_2') \) and

\[
1 - \sum_{n=0}^{m_W - \max(K,K_2')} \frac{e^{-\mu t} (\mu t)^n}{n!} \cdot \frac{1 - H(t)}{E[H]} dt \leq \epsilon,
\]

where \( \mu = 1/E[H] \) and \( K_2' \) is given in Assumption 1.1 (ii).

Figure 2 shows the computational procedures for \( W_{k,\ell} \) \( (k = 1, 2, \ldots, K, \ell = 0, k, k + 1, \ldots, \min(m_W, k + n_W)) \) satisfying (2.28). Once we compute \( W_{k,\ell}^{\text{comp},n_W}(m_W) \)'s, we can estimate the error contained in those by

\[
\sum_{\ell=k}^{\infty} (W_{k,\ell} - W_{k,\ell}^{\text{comp},n_W}(m_W)) e \leq E[H] e - \sum_{\ell=0}^{m_W} W_{k,\ell}^{\text{comp},n_W}(m_W) e, \tag{2.31}
\]

which can be derived in the same way as (2.11).
where \( J(\mathcal{K} \xi) := \{ (C_k, D_k, \Gamma_k); k = 0, 1, \ldots \} \), and \( \epsilon_W \).

**Output:** \( m_W^* \) and \( W_{k,\ell}^{\text{comp}, n_W} \) (\( k = 1, 2, \ldots, K, \ell = 0, k, k + 1, \ldots, m_W^* \)).

Let \( m_W^* := \min \left( m_W > K; \left( -C_K \right)^{-1} D_K (\left( -C_K \right)^{-1} D_K + \cdots \right) \right) \leq \frac{\epsilon_W}{2E[H]} e \).

Let \( n_W^* := \min \left( n_W \geq 0; \sum_{n=0}^{n_W} \eta_n(m_W^*) \geq 1 - \frac{\epsilon_W}{2E[H]} \right) \).

Compute \( V_{k,\ell}^{(n)}(m_W^*) \)'s (\( n = 0, 1, \ldots, n_W^* \)) by (2.6).

Compute \( W_{k,\ell}^{\text{comp}, n_W}(m_W^*) \) (\( k = 1, 2, \ldots, K, \ell = 0, k, k + 1, \ldots, \min(m_W^*, k + n_W^*) \)).

(a) Under Assumption 1.1 (i)

**Input:** \( \{(C_k, D_k, \Gamma_k); k = 0, 1, \ldots \} \), \( K \), and \( \epsilon_W \).

**Output:** \( m_W^* \) and \( W_{k,\ell}^{\text{comp}, n_W} \) (\( k = 1, 2, \ldots, K, \ell = 0, k, k + 1, \ldots, m_W^* \)).

Let \( m_W^* := \min \left( m_W > K; \sum_{n=0}^{m_W-\max(K,K)} \int_0^{\infty} e^{-nt} \frac{(\mu t)^n}{n!} H(t) dt \geq 1 - \frac{\epsilon_W}{2E[H]} \right) \).

Let \( n_W^* := \min \left( n_W \geq 0; \sum_{n=0}^{n_W} \eta_n(m_W^*) \geq 1 - \frac{\epsilon_W}{2E[H]} \right) \).

Compute \( V_{k,\ell}^{(n)}(m_W^*) \)'s (\( n = 0, 1, \ldots, n_W^* \)) by (2.6).

Compute \( W_{k,\ell}^{\text{comp}, n_W}(m_W^*) \) (\( k = 1, 2, \ldots, K, \ell = 0, k, k + 1, \ldots, \min(m_W^*, k + n_W^*) \)).

(b) Under Assumption 1.1 (ii)

Figure 2: Computational procedures for \( W_{k,\ell} \)'s in the non-uniformizable case

3. The Augmented Truncation Approximation to the Imbedded Markov Chain

In this section, we discuss the augmented truncation approximation (ATA) to the imbedded Markov chain with the truncated transition probability matrix \( P(K) \). In the ATA, the selection of the augmentation matrix \( P_A(K) \) in (1.6) is crucial and we briefly summarize how to manage it according to [11]. First of all, we restrict our attention to the linear ATA solution obtained by (2.6). We then define \( \kappa^{\text{approx}}(K; \xi) \) as the linear ATA solution obtained by

\[
\kappa^{\text{approx}}(K; \xi) = \kappa^{\text{approx}}(K; \xi) \left[ P(K) + (I - P(K)) \xi \right], \quad \kappa^{\text{approx}}(K; \xi)e = 1.
\]

Let \( [\xi]_{(k,i)} \) \( (k = 0, 1, \ldots, K; i \in \mathcal{M}) \) denote the \( (kM+i) \)th element of \( \xi \) and let \( [P]_{(k,i),(\ell,j)} \) \( (k, \ell = 0, 1, \ldots; i, j \in \mathcal{M}) \) denote the \( (kM+i, \ell M+j) \)th element of \( P \). We then define \( \Xi(K) \) as

\[
\Xi(K) = \{ \xi \in \mathbb{R}^{(K+1)M}; \xi \geq 0, \xi e = 1, [\xi]_{(k,i)} > 0 \text{ if } (k,i) \in \mathcal{J}^+(K), [\xi]_{(k,i)} = 0 \text{ if } (k,i) \notin \mathcal{J}^+(K) \},
\]

where \( \mathcal{J}^+(K) \) denotes a subset of levels zero to \( K \), whose states are directly reachable from
Remark 3.1. Unfortunately, we cannot show \( \lim_{K \to \infty} \varepsilon(K) = 0 \) in general. Note, however, that if
\[
\Pr((L_{n+1}, S_{n+1}) = (\ell, j) \mid (L_n, S_n) = (k, i)) > 0,
\]
for all \((k, i) \in \{0, 1, \ldots, K_1^n\} \times \mathcal{M}, (\ell, j) \in \mathcal{J}^+(K_1^n)\), we can show \( \lim_{K \to \infty} \varepsilon(K) = 0 \) under Assumption 1.1 (i). Due to the shortage of space, we omit the details.

\[\begin{array}{|c|}
\hline
\text{Input:} & P(K) \text{ and } \mathcal{J}^+(K) . \\
\text{Output:} & \kappa_{\text{approx}} = (\kappa_{\text{approx}}(K) \ 0) \text{ and } \varepsilon(K) . \\
\hline
\end{array}\]

Compute \( \kappa_{\text{approx}}(K; e^T_{(k,i)}) = x_{(k,i)}/x_{(k,i)}e \) for all \((k, i) \in \mathcal{J}^+(K)\),
where \( x_{(k,i)} \) is the unique solution of \( x_{(k,i)}(I - P(K)) = e^T_{(k,i)} \).

Compute \( \kappa_{\text{approx}}(K) \) by (3.3) and set \( \kappa_{\text{approx}} = (\kappa_{\text{approx}}(K) \ 0) \).

Compute the error bound \( \varepsilon(K) \) by (3.4).

Figure 3: Computational procedure for \( \kappa_{\text{approx}} = (\kappa_{\text{approx}}(K) \ 0) \)

at least one state in levels \( K + 1 \) or higher.

\[
\mathcal{J}^+(K) = \left\{ (0, i) ; i \in \mathcal{M}, \sum_{k=K+1}^{\infty} [e^T \Gamma_k]_i > 0 \right\} \cup \{(K, i) ; i \in \mathcal{M}\} . \tag{3.1}
\]

Note that there exists \( \xi^* \in \Xi(K) \) such that \( \kappa_{\text{approx}}(K; \xi^*) = \kappa(K) \) [11, Remark 2.2].

Furthermore, for an arbitrary \( \xi \in \Xi(K) \), \( \kappa_{\text{approx}}(K; \xi) \) is given by a convex combination of \( \kappa_{\text{approx}}(K; e^T_{(k,i)}) \)'s \(( (k, i) \in \mathcal{J}^+(K) \) with positive weights [11, Eq. (2.33)], where \( e^T_{(k,i)} \) denotes the unit row vector whose \((k,M+i)\)th element is equal to one. Specifically,

\[
\kappa_{\text{approx}}(K; \xi) = \sum_{(k,i) \in \mathcal{J}^+(K)} [\alpha(K; \xi)]_{(k,i)} \kappa_{\text{approx}}(K; e^T_{(k,i)}),
\]

where \( \alpha(K; \xi) \in \Xi(K) \) and there is a one-to-one correspondence between \( \xi \) and \( \alpha(K; \xi) \) [11, Eq. (2.34)]. It then follows that [11, Eq. (2.37)]

\[
\| \kappa_{\text{approx}}(K; \xi) - \kappa(K) \|_1 \leq \varepsilon(K) = \max_{(k,i) \in \mathcal{J}^+(K)} \| \kappa_{\text{approx}}(K; \xi) - \kappa_{\text{approx}}(K; e^T_{(k,i)}) \|_1 . \tag{3.2}
\]

Based on this observation, the following approximation \( \kappa_{\text{approx}}(K) \) to the conditional stationary distribution \( \kappa(K) \) is suggested in [11, Implication 4]:

\[
\kappa_{\text{approx}}(K) = \sum_{(k,i) \in \mathcal{J}^+(K)} \frac{1}{|\mathcal{J}^+(K)|} \kappa_{\text{approx}}(K; e^T_{(k,i)}), \tag{3.3}
\]

which presumes that the transition structure in levels higher than \( K \) is unavailable, and the upper bound \( \varepsilon(K) \) of the error \( \varepsilon(K) \) in (3.2) is given by

\[
\| \kappa(K) - \kappa_{\text{approx}}(K) \|_1 \leq \varepsilon(K) = \max_{(k,i) \in \mathcal{J}^+(K)} \| \kappa_{\text{approx}}(K) - \kappa_{\text{approx}}(K; e^T_{(k,i)}) \|_1 . \tag{3.4}
\]

In summary, we compute the approximation \( \kappa_{\text{approx}}(K) \) and its error bound \( \varepsilon(K) \) by the procedure in Figure 3.

Remark 3.1. Unfortunately, we cannot show \( \lim_{K \to \infty} \varepsilon(K) = 0 \) in general. Note, however, that if
4. The Error Bound in the Approximation to the Stationary Distribution

If we obtain an approximation \( \kappa_{\text{approx}} \) to the stationary distribution \( \kappa \) of the imbedded Markov chain \( \{(L_n, S_n)\}_{n=0,1,...} \) and \( W_{k,\ell}^{\text{comp,nW}}(m_W) \)'s, we can compute an approximation \( \pi_{\text{approx}} = (\pi_{0,\text{approx}}^{\ell,\text{approx}} \pi_{1,\text{approx}}^{0,\text{approx}} \cdots) \) to the stationary distribution \( \pi \) by (1.8) and (1.9), where \( W_{k,\ell} \) is replaced by \( W_{k,\ell}^{\text{comp,nW}}(m_W) \). Since \( W_{k,\ell}^{\text{comp,nW}}(m_W) = O(\ell > \min(m_W, k + n_W)) \), the computation of \( \pi_{\text{approx}} \)'s stops at \( \ell = \min(m_W, K + n_W) \). Recall that for an arbitrary \( c_W > 0 \), \( W_{k,\ell}^{\text{comp,nW}}(m_W) \)'s satisfy (2.28). We thus discuss the error in \( \pi_{\text{approx}} \), assuming that errors in \( W_{k,\ell}^{\text{comp,nW}}(m_W) \)'s are negligible, i.e.,

\[
\pi = \frac{\kappa \hat{W}}{\kappa \hat{W} e} = \frac{\kappa \hat{W}}{E[T]} \cdot \pi_{\text{approx}} = \frac{\kappa_{\text{approx}} \hat{W}}{\kappa_{\text{approx}} \hat{W} e},
\]

where

\[
\hat{W} = \begin{pmatrix}
(-C_0)^{-1} & (C_0)^{-1}D_0 W_{1,1} & (C_0)^{-1}D_0 W_{1,2} & (C_0)^{-1}D_0 W_{1,3} & \cdots \\
O & W_{1,1} & W_{1,2} & W_{1,3} & \cdots \\
O & O & W_{2,2} & W_{2,3} & \cdots \\
O & O & O & W_{3,3} & \cdots \\
\vdots & \vdots & \vdots & \vdots & \ddots
\end{pmatrix},
\]

and \( E[T] = \kappa \hat{W} e \) denotes the average length of two consecutive imbedded points.

**Theorem 4.1.** If errors in \( W_{k,\ell}^{\text{comp,nW}}(m_W) \)'s are negligible, we have

\[
\|\pi - \pi_{\text{approx}}\|_1 \leq \frac{T_{\text{upper max}}}{\kappa_{\text{approx}} \hat{W} e} \|\kappa - \kappa_{\text{approx}}\|_1 \left(1 + \frac{T_{\text{upper max}}}{\kappa_{\text{approx}} \hat{W} e - T_{\text{upper max}}} \||\kappa - \kappa_{\text{approx}}\|_1\right). \tag{4.1}
\]

where \( T_{\text{upper max}} = E[H] + \max_{i \in M} \|(-C_0)^{-1} e_i\| \).

**Remark 4.1.** (1.11) comes from (4.1) and

\[
\frac{T_{\text{upper max}}}{\kappa_{\text{approx}} \hat{W} e} \|\kappa - \kappa_{\text{approx}}\|_1 \left(1 + \frac{T_{\text{upper max}}}{\kappa_{\text{approx}} \hat{W} e - T_{\text{upper max}}} \||\kappa - \kappa_{\text{approx}}\|_1\right) = \frac{T_{\text{upper max}}}{\kappa_{\text{approx}} \hat{W} e} \left(1 + \frac{T_{\text{upper max}}}{\kappa_{\text{approx}} \hat{W} e} \|\kappa - \kappa_{\text{approx}}\|_1 + o(\|\kappa - \kappa_{\text{approx}}\|_1)\right).
\]

**Proof.** Let \( T_{\text{max}} = \max_{(k,i) \in \{0,1,...\} \times M} \|\hat{W}_{(k,i)}\| \). Because \( \text{diag}^{-1}(\hat{W} e) \hat{W} e = e \), we have

\[
\|\pi - \pi_{\text{approx}}\|_1 = \left\| \frac{\kappa \text{diag}(\hat{W} e)}{\kappa \hat{W} e} \text{diag}^{-1}(\hat{W} e) \hat{W} - \frac{\kappa_{\text{approx}} \text{diag}(\hat{W} e)}{\kappa_{\text{approx}} \hat{W} e} \text{diag}^{-1}(\hat{W} e) \hat{W} \right\|_1
\]

\[
\leq \left\| \frac{\kappa \text{diag}(\hat{W} e) - \kappa_{\text{approx}} \text{diag}(\hat{W} e)}{\kappa_{\text{approx}} \hat{W} e} \right\|_1,
\]

\[
\leq \left\| \frac{(\kappa - \kappa_{\text{approx}}) \text{diag}(\hat{W} e)}{\kappa_{\text{approx}} \hat{W} e} \right\|_1 + \left\| \frac{1}{\kappa \hat{W} e} - \frac{1}{\kappa_{\text{approx}} \hat{W} e} \right\|_1 \|\kappa \text{diag}(\hat{W} e)\|_1
\]

\[
\leq \frac{T_{\text{max}}}{\kappa_{\text{approx}} \hat{W} e} \|\kappa - \kappa_{\text{approx}}\|_1 + \frac{\|\kappa \text{approx} - \kappa\| \hat{W} e}{\kappa \hat{W} e \cdot \kappa_{\text{approx}} \hat{W} e} \cdot T_{\text{max}}
\]

\[
\leq \frac{T_{\text{max}}}{\kappa_{\text{approx}} \hat{W} e} \|\kappa - \kappa_{\text{approx}}\|_1 + \frac{T_{\text{max}}^2}{\kappa \hat{W} e \cdot \kappa_{\text{approx}} \hat{W} e} \|\kappa - \kappa_{\text{approx}}\|_1. \tag{4.2}
\]
Note here that $T_{\text{max}}$ can be bounded from above.

$$T_{\text{max}} = \max_{(k,i) \in \{0, \ldots, m\} \times \mathcal{M}} E[\tau_{n+1} - \tau_n \mid L_n = k, S_n = i] \leq E[H] + \max_{i \in \mathcal{M}} [(\mathbf{C}_0)^{-1} \mathbf{e}]_i.$$ (4.3)

Note also that $\kappa \mathbf{W} \mathbf{e}$ can be bounded from below.

$$\kappa \mathbf{W} \mathbf{e} = \kappa^{\text{approx}} \mathbf{W} \mathbf{e} + (\kappa - \kappa^{\text{approx}}) \mathbf{W} \mathbf{e} \geq \kappa^{\text{approx}} \mathbf{W} \mathbf{e} - \|\kappa - \kappa^{\text{approx}}\|_1 \cdot T_{\text{max}}.$$ (4.4)

Theorem 4.1 now follows from (4.2), (4.3), and (4.4). \hfill $\square$

We can evaluate the error bound (4.1) for $\pi^{\text{approx}}$ if the upper bound of $\|\kappa - \kappa^{\text{approx}}\|_1$ is available. Recall that the upper bound of $\|\kappa - \kappa^{\text{approx}}\|_1$ is given by $2\zeta(K) + \varepsilon(K)$ (cf. (1.7)). While $\varepsilon(K)$ is given by (3.4), the tail distribution $\zeta(K)$ is hard to evaluate analytically. We will discuss a rough estimation of $\zeta(K)$ using numerical examples in the next section.

5. Numerical Examples and Discussions

The purpose of this section is two-fold: one is to demonstrate the soundness of our numerical implementation and the other is to discuss how to obtain a qualitatively decent approximation. For these purposes, we consider a single-server queue with balking and level-dependent disasters, where customers are served on a FIFO basis with deterministic service times with $E[H] = 1.0$. We assume that customers arrive according to a Markov-modulated Poisson process (MMPP), whose underlying Markov chain $\{S(t)\}_{t \geq 0}$ is assumed to be a two-state birth-and-death process with identical transition rate $\alpha = 0.01$. Note that the stationary distribution $\pi(t)$ of the underlying Markov chain $\{S(t)\}_{t \geq 0}$ is given by $\pi(t) = (0.5, 0.5)$. We assume that if $S(t) = 1$ (resp. $S(t) = 2$), customers arrive according to a Poisson process with rate $\lambda_1 = 1.2$ (resp. $\lambda_2 = 1.8$). If a customer finds $k$ ($k = 0, 1, \ldots$) customers on arrival, he/she joins the queue with probability $\beta_k = (2.0 + 0.99^k)/3$. Let $L(t)$ ($t \geq 0$) denote the number of customers in the system at time $t$. We assume that when $L(t) = k$ ($k = 1, 2, \ldots$), a disaster occurs at rate $r_k = \max(0, (k - 200)/100)$.

The net arrival/disaster process of $\{(L(t), S(t))\}_{t \geq 0}$ is characterized by $\{(C_k, D_k, \Gamma_k); k = 0, 1, \ldots\}$, where

$$C_k = \begin{pmatrix} c_{k,1} & \alpha \\ \alpha & c_{k,2} \end{pmatrix}, \quad D_k = \beta_k \cdot \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}, \quad \Gamma_k = r_k \mathbf{I}, \quad k = 0, 1, \ldots,$$

where $c_{k,i} = -\alpha - \beta_k \lambda_i - r_k$ ($k = 0, 1, \ldots, i = 1, 2$). It then follows that $\mathcal{J}^+(K) = \{0, K\} \times \{1, 2\}$ and $\theta_m$ is given by

$$\theta_m = \max_{k=1, 2, \ldots, m} \left(0.01 + \frac{2 + 0.99^k}{3} \cdot 1.8 + \max \left(0, \frac{k - 200}{100}\right)\right)$$

$$\theta_m = \begin{cases} 1.804, & m = 1, 2, \ldots, 254, \\ 1.21 + 0.6 \cdot 0.99^k + \frac{k - 200}{100}, & m = 255, 256, \ldots. \end{cases}$$

It is easy to see that $\{(L(t), S(t))\}_{t \geq 0}$ is non-explosive, $\exp[Ut]$ is non-uniformizable, Assumption 1.1 (i) holds with $K_1 = 208$, and the condition in Remark 3.1 holds. By definition, $\gamma_n(m_A)$ and $\eta_n(m_W)$ ($n = 0, 1, \ldots$) are given by

$$\gamma_n(m_A) = \frac{(\theta_m A_E[H])^n}{n!} e^{-\theta_m A_E[H]}, \quad \eta_n(m_W) = \frac{1}{\theta_m W_E[H]} (1 - \sum_{i=0}^n \gamma_i(m_W)).$$
In what follows, we set $\epsilon_A = \epsilon_W = 10^{-14}$.

In Step 1, we conduct the procedure in Figure 1 (a) to compute $P(K)$. Figure 4 shows the error bound $\varepsilon_A^{\text{max}}(K)$ for $A_{k,\ell}^{\text{comp},n_A}(m_A)$'s in $P(K)$ as a function of $K$ ($K > K_1^\ast = 208$), where

$$
\varepsilon_A^{\text{max}}(K) = \max_{k=0,1,\ldots,K} \left[ e - \sum_{\ell=0}^{m_A-1} A_{k,\ell}^{\text{comp},n_A}(m_A) e \right]_j \geq \max_{k=0,1,\ldots,K} \left[ \sum_{\ell=0}^{K} (A_{k,\ell} - A_{k,\ell}^{\text{comp},n_A}(m_A)) e \right]_j.
$$

Note that the inequality in the above comes from (2.11). We observe that the computational procedure for $A_{k,\ell}^{\text{comp},n_A}(m_A)$ works as designed because we set $\epsilon_A = 10^{-14}$. We make a comment on the ripple-like shape of $\varepsilon_A^{\text{max}}(K)$. As shown in Theorem 2.1, there are two sources of errors in computing an approximation to $\{A_{k,j}\}$: the truncation of $U$ (i.e., $m_A$ in (2.15)) and the truncation of infinite sums (i.e., $n_A$ in (2.15)). The selection of $m_A$ is conservative as shown in the proof of Theorem 2.2, so that the numerical error due to the truncation of $U$ seems to be negligible. On the other hand, the selection of $n_A$ is rigid and for a fixed $n_A$, the second term on the right-hand side of (2.15) increases with $K$ because $\theta_{m_A}$ increases with $K$. This causes a ripple-like shape of $\varepsilon_A^{\text{max}}(K)$, which is inherent in the non-uniformizable case.

Next, we consider the ATA in Step 2. According to the procedure in Figure 3, we compute an approximation $\kappa^{\text{approx}}(K)$ to the conditional stationary distribution $\kappa(K)$. Figure 5 shows the error bound $\varepsilon(K)$ of $\kappa^{\text{approx}}(K)$ in (3.4) as a function of $K$ ($K > K_1^\ast = 208$). We observe that (i) $\varepsilon(K)$ takes a large value when $k$ is close to $K_1^\ast$, (ii) it decreases as $K$ increases, and then, (iii) it takes a ripple-like shape as $\varepsilon_A^{\text{max}}(K)$ in Figure 4.

Note that the conditional stationary distribution of any ergodic Markov chain with the same north-west corner submatrix $P(K)$ and the same $\mathcal{J}^+(K)$ defined in (3.1) is given by a mixture of probability distributions $\nu_{k,i}(K) = \kappa^{\text{approx}}(K; e''_{(k,i)}) ((k, i) \in \mathcal{J}^+(K))$ with positive weights [11, Theorem 2.2]. Therefore, the degree of freedom (i.e., the extent to which the transition structure in level $K+1$ or higher affects the conditional stationary distribution $\kappa(K)$) is large for a moderate $K$, which leads to the above-mentioned observation (i). Observation (ii) conforms with the convergence property in Remark 3.1. Let $P^{\text{comp}}(K)$ denote the computed $P(K)$ with $A_{k,\ell}^{\text{comp},n_A}(m_A)$'s. When $K$ is large enough to suppress the degree of freedom, $\|P(K) - P^{\text{comp}}(K)\|$ becomes the dominant source of error, which leads to observation (iii).

Figure 6 shows approximations to the conditional stationary distribution $\kappa_k(K)e$ ($k \geq 200$) for $K = 220, 240, \ldots, 320$. Recall that the drift is negative if $L(t) > K_1^\ast = 208$. There-
Therefore, we expect that $\kappa_k(K)e$ ($k \geq 208$) is a decreasing function of $k$, whereas $\kappa_k^{\text{approx}}(K)e$ takes the minimum value at some $k = u_{\text{min}}(K)$ close to $K$ and it turns upward. This pathological phenomenon is caused by our selection of equal weights in (3.3) for $\nu_{k,i}(K) = \kappa^{\text{approx}}(K; e^T_{(k,i)})$ ($k = 0, K, i = 1, 2$). While $\nu_{0,i}(K)$ ($i = 1, 2$) has a strictly decreasing tail, $\nu_{0,i}(K)$ ($i = 1, 2$) takes the minimum value at some $k$ close to $K$ and it turns upward. The equally-weighted mixture of those distributions results in this pathological phenomenon. Note that similar phenomena can also occur in standard augmentation strategies such as last-column/last-column block augmentations. When $K$ is large enough (i.e., $K \geq 260$), however, $\kappa_k^{\text{approx}}(K)e$ is less than $\epsilon_A(K) = 10^{-14}$, which suggests that such a $K$ is large enough to obtain a reliable result in finite precision computation. Furthermore, in such a case, $\kappa_k^{\text{approx}}(K)e$ for $k$ less than and not close to $u_{\text{min}}(K)$ seems to have a decent accuracy as they are indistinguishable from $\kappa_k^{\text{approx}}(K)e$ for $K = 320$. In what follows, we restrict our attention to $K = 300$.

We attempt to eliminate an inaccurate tail in $\{\kappa_k^{\text{approx}}(K); k = 0, 1, \ldots, K\}$. Although $\kappa_k^{\text{approx}}(K)e \simeq \epsilon_A = 10^{-14}$ for $k = 257$ can be a guideline for it, Figure 6 suggests that it would be too conservative because $u_{\text{min}}(K) = 286$ for $K = 300$. We thus set $K^* = 280$ and obtain an approximation $\kappa_k^{\text{approx}}(K^*) = (\kappa_k^{\text{approx}}(K^*) \kappa_k^{\text{approx}}(K^*) \cdots \kappa_k^{\text{approx}}(K^*))$, where

$$\kappa_k^{\text{approx}}(K^*) = \kappa_k^{\text{approx}}(K) / \sum_{n=0}^{K^*} \kappa_n^{\text{approx}}(K)e, \quad k = 0, 1, \ldots, K^*.$$  

We adopt $\kappa_k^{\text{approx}} = (\kappa_k^{\text{approx}}(K^*) 0)$ instead of $\kappa_k^{\text{approx}} = (\kappa_k^{\text{approx}}(K) 0)$. Figure 7 shows $\kappa_k^{\text{approx}}(K)e$ for $K = 300$ and $\kappa_k^{\text{approx}}(K^*)e$ for $K^* = 280$.

We then consider the error bound (1.7) for $\kappa_k^{\text{approx}}$. Note here that $\kappa_k^{\text{approx}}(K^*)$ is also given by a convex combination of four probability vectors corresponding to states $(0, i)$ and $(K, i)$ ($i = 1, 2$) (cf. (3.3)) and $\|\kappa^{\text{approx}}(K^*) - \kappa(K^*)\|_1$ can be bounded in a way similar to (3.4). As a result, we have $\|\kappa^{\text{approx}}(K^*) - \kappa(K^*)\|_1 \leq \epsilon(K^*) \simeq 2.1084 \times 10^{-13}$, while $\epsilon(K) \simeq 2.1270 \times 10^{-13}$ for $\kappa_k^{\text{approx}}(K)$. Therefore, in terms of 1-norm, the difference between two approximations $\kappa_k^{\text{approx}}(K^*)$ and $\kappa_k^{\text{approx}}(K)$ is negligible.

Next we consider a rough estimation of the tail bound $\zeta(K^*)$ in (1.7). Because the magnitude of the negative drift increases with $k$ ($k \geq 208$), we expect that the tail probability decreases more rapidly than any exponential function, as suggested in Figure 7. Noting $\kappa_k^{257}(K^*)e \simeq \epsilon_A = 10^{-14}$, we assume that

$$\kappa_k e \leq \kappa_k^{\text{approx}}(K^*)e \cdot r^{k-K^*}, \quad k = K^* + 1, K^* + 2, \ldots,$$
where \( r^{K_\ast - 257} = \kappa_{K_\ast}^{\text{approx}}(K_\ast)e/\kappa_{257}^{\text{approx}}(K_\ast)e \simeq 1.2338 \times 10^{-8} \). We thus have \( r \simeq 0.4530 \) and 
\[
\zeta(K_\ast) \leq \kappa_{K_\ast}^{\text{approx}}(K_\ast)e \cdot r/(1 - r) \simeq 7.9031 \times 10^{-29},
\]
which indicates that \( K_\ast = 280 \) is large enough to ignore the error due to the truncation at \( K_\ast \).

Let \( \kappa_{K_\ast}^{\text{approx}} = (\kappa_{\text{approx}}(K_\ast), 0) \). We then have \( \| \kappa_{\text{approx}} - \kappa \|_1 \leq 2\zeta(K_\ast) + \varepsilon(K_\ast) \leq 2.1084 \times 10^{-13} \). In a similar way, we also obtain \( \| \kappa_{\text{approx}} - \kappa \|_1 \leq 3.2362 \times 10^{-13} \). Moreover, \( \| \kappa_{\text{approx}} - \kappa_{\text{approx}} \|_1 \simeq 2.2151 \times 10^{-15} \). We thus conclude that in terms of 1-norm, the difference between \( \kappa_{\text{approx}} \) and \( \kappa_{\text{approx}} \) is not significant.

Finally, in Step 3, we compute an approximation to the stationary distribution \( \pi \). We first compute \( W_{k,\ell}^{\text{comp},n_w}(m_w) \)'s according to the procedure in Figure 2 (a), where \( K \) is replaced by \( K = 280 \), \( m_w = 332 \), and \( n_w = 21 \). We then obtain the error bound in (2.31) for \( K = 280 \).

\[
\max_{j \in \mathcal{M}} \left[ \sum_{k=1}^{K_\ast} \left( W_{k,\ell} - W_{k,\ell}^{\text{comp},n_w}(m_w) \right)e \right]_j \leq \max_{j \in \mathcal{M}} \left[ E[H]e - \sum_{\ell=0}^{m_w} W_{k,\ell}^{\text{comp},n_w}(m_w)e \right]_j \\
\simeq 4.2188 \times 10^{-15},
\]
which shows that the procedure for \( W_{k,\ell}^{\text{comp},n_w}(m_w) \)'s works as designed because \( \epsilon_w = 10^{-14} \).

We then compute \( \{ \pi_k^{\text{approx}}(K_\ast); \ k = 0, 1, \ldots \} \) and construct the stationary distribution of the number of customers.

\[
\pi_{\text{approx}} = (\pi_0^{\text{approx}}(K_\ast), \pi_1^{\text{approx}}(K_\ast), \ldots, \pi_{301}^{\text{approx}}(K_\ast), 0, 0, \ldots),
\]
where we use \( \min(m_w, K_\ast + n_w) = 301 \) for \( K_\ast = 280 \). Figure 8 shows \( \pi_{\text{approx}} \), where the corresponding result for \( \pi_{\text{approx}} \) based on \( \kappa_{\text{approx}}(K) \) for \( K = 300 \) is also plotted for reference. The error bound (4.1) for \( \pi_{\text{approx}} \) is given by \( 1.092 \times 10^{-12} \), where we use \( T_{\text{upper}}^{\text{max}} \simeq 1.8311 \) and \( \kappa_{\text{approx}} \hat{W}e \simeq 1.0004 \). It thus seems that \( \pi_{\text{approx}} \) is not only a qualitatively but also quantitatively decent approximation.

6. Conclusion

This paper considered the augmented truncation approximation (ATA) to single-server queues with level-dependent arrivals and disasters and generally distributed service times. We first discussed how to compute the truncated transition probability matrix \( P(K) \) of the imbedded Markov chain. If the continuous-time Markov chain representing arrivals and disasters is not uniformizable, the computation of \( P(K) \) is not straightforward. In this
paper, we developed numerical procedures for computing $P(K)$ that satisfies a predefined error bound, under some stability conditions. We also briefly discussed how to construct the augmentation matrix in the ATA and how to obtain the error bound of the approximate solution to the stationary queue length distribution. Numerical examples demonstrated that our procedure works well as designed.

A. Proof of Theorem 2.3

Suppose that a service of length $H$ starts at time 0. Let $T_1$ denote the first occurrence time of arrivals after time 0 and we define $T_n$ ($n = 2, 3, \ldots$) as the interarrival time of the $(n - 1)$st and $n$th arrivals after time 0. We also define $T_1$ as the first occurrence time of disasters after time 0. We then have

$$b^+(A, K, i)(m_A) = \Pr(T_1 + T_2 + \cdots + T_{m_A-K+1} \leq \min(H, \Gamma_1) \mid (L(0), S(0)) = (K, i)), \quad i \in M.$$ 

Associated with the Markov chain $\{(L(t), S(t))\}_{t \geq 0}$, we consider a Markov chain $\{(L^*(t), S^*(t))\}_{t \geq 0}$ characterized by $\{(C^*_k, D^*_k) ; \quad k = 0, 1, \ldots\}$, where $C^*_k$ is introduced in Lemma 2.1. In what follows, we use the following convention: For any symbol $X$ related to the original Markov chain $\{(L(t), S(t))\}_{t \geq 0}$, $X^*$ denotes the corresponding quantity for the Markov chain $\{(L^*(t), S^*(t))\}_{t \geq 0}$. It then follows from Lemma 2.1 (i) that

$$b^+_A(K, i)(m_A) = \int \cdots \int \exp[C^*_K t_1]D_K \cdots \exp[C^*_{m_A} t_{m_A-K+1}]D_{m_A} e^H(y) \mathcal{H}(y)$$

so that

$$b^+_A(K, i)(m_A) \leq \Pr(T_1^* + T_2^* + \cdots + T_{m_A-K+1}^* \leq H \mid (L^*(0), S^*(0)) = (K, i)). \quad (A.1)$$

Moreover, associated with the Markov chain $\{(L^*(t), S^*(t))\}_{t \geq 0}$, we introduce a Markov chain $\{(L^*(t), S^*(t))\}_{t \geq 0}$ characterized by $\{(C^*_k(x_k), D^*_k(x_k)) ; \quad k = 0, 1, \ldots\}$, where

$$C^*_k(x_k) = C^*_k - \text{diag}(x_k), \quad D^*_k(x_k) = D_k + \text{diag}(x_k)(-C^*_k)^{-1}D_k, \quad k = 1, 2, \ldots,$$

for $x_k \geq 0$. We use the superscript $^*$ to represent quantities related to the Markov chain $\{(L^*(t), S^*(t))\}_{t \geq 0}$. It then follows from Lemma 2.1 (iii) that for $k = 1, 2, \ldots$,

$$\Pr(T_1^* \leq \tau, S^*(T_1^*) = j \mid (L^*(0), S^*(0)) = (k, i)) \quad = \quad \frac{1}{\tau} \int_0^\tau \exp[C^*_k t]D_k dt]_{i,j}$$

$$\leq \frac{1}{\tau} \int_0^\tau \exp[(C^*_k - \text{diag}(x_k)) t]D_k(x_k) dt]_{i,j} = \Pr(T_1^* \leq \tau, S^*(T_1^*) = j \mid (L^*(0), S^*(0)) = (k, i)).$$

Note here that Lemma 2.1 (ii) implies

$$\Pr(S^*(T_1^*) = j \mid (L^*(0), S^*(0)) = (k, i)) = \Pr(S^*(T_1^*) = j \mid (L^*(0), S^*(0)) = (k, i)). \quad (A.2)$$

It then follows that

$$\Pr(T_1^* \leq \tau \mid (L^*(0), S^*(0)) = (k, i), S^*(T_1^*) = j)$$
Furthermore, it follows from (A.2) that follows from [17, Theorem 1.A.3 (b)] that for \(L\) conditionally independent given \((a, 0, \ldots)\). Let \(S = \sum_{k=1}^{n} (L^k(0), S^k(0))\) and \(a_n^* = T_1^* + T_2^* + \cdots + T_n^*\). Because \(T_n^*\)'s (resp. \(T_n^\ast\)) are conditionally independent given \((L^k(0), S^k(0))\) and \(S_n^*\)'s (resp. \((L^*(0), S^*(0))\) and \(S_n^\ast\)'s), it follows from [17, Theorem 1.A.3 (b)] that for \(m = 1, 2, \ldots, \)

\[
\Pr(T_1^* + T_2^* + \cdots + T_m^* \leq x \mid (L^*(0), S^*(0)) = (K, i), S_n^* = j_n \quad (n = 1, 2, \ldots, m)) \\
\leq \Pr(T_1^* + T_2^* + \cdots + T_m^* \leq x \mid (L^*(0), S^*(0)) = (K, i))
\]

Furthermore, it follows from (A.2) that

\[
\Pr(S_n^* = j_n \quad (n = 1, 2, \ldots, m) \mid (L^*(0), S^*(0)) = (K, i)) = \Pr(S_n^* = j_n \quad (n = 1, 2, \ldots, m) \mid (L^*(0), S^*(0)) = (K, i))
\]

We thus have

\[
\Pr(T_1^* + T_2^* + \cdots + T_{m_A-K+1}^* \leq x \mid (L^*(0), S^*(0)) = (K, i)) \\
\leq \Pr(T_1^* + T_2^* + \cdots + T_{m_A-K+1}^* \leq x \mid (L^*(0), S^*(0)) = (K, i))
\]

It then follows from (A.1) and (A.3) that

\[
b_{(K,i)}^+(m_A) \leq \Pr(T_1^* + T_2^* + \cdots + T_{m_A-K+1}^* \leq H \mid (L^*(0), S^*(0)) = (K, i)) \\
\leq \Pr(T_{n(K,K^\dagger_2)}^* + T_{n(K,K^\dagger_2)+1}^* + \cdots + T_{m_A-K+1}^* \leq H \mid (L^*(0), S^*(0)) = (K, i))
\]

where \(n(K, K^\dagger_2) = \max(0, K^\dagger_2 - K) + 1 \geq 1\).

We now set \(x_k = 0\) for \(k = 1, 2, \ldots, K^\dagger_2 - 1\) and \(x_k = \mu e - D_k e\) for \(k = K^\dagger_2, K^\dagger_2 + 1, \ldots\). Note that \(x_k \geq 0\) \((k = 1, 2, \ldots)\) under Assumption 1.1 (ii). Furthermore, we have \(D_k^*(x_k)e = \mu e\) \((k = K^\dagger_2, K^\dagger_2 + 1, \ldots)\), which implies that if \(L^*(t) \geq K^\dagger_2\), arrivals after time \(t\) occur at constant rate \(\mu\), regardless of \(S^*(t)\). In other words, when \((L^*(0), S^*(0)) = (K, i)\), \(T_n^*\)'s \((n = n(K, K^\dagger_2), n(K, K^\dagger_2) + 1, \ldots)\) are i.i.d. exponential random variables. The theorem now follows from (A.4) and this observation.

Acknowledgments
This research was supported in part by JSPS KAKENHI Grant No. 19K11841.

References
[1] S. Asmussen, C. Klüppelberg, and K. Sigman: Sampling at subexponential times, with queueing applications. *Stochastic Processes and Their Applications*, **79** (1999), 265–286.
[2] O.J. Boxma, D. Perry, and W. Stadje: Clearing models for M/G/1 queues. *Queueing Systems*, **38** (2001), 287–306.
[3] P. Brémaud: *Markov Chains: Gibbs Fields, Monte Carlo Simulation, and Queues* (Springer, New York, NY, 1999)
[4] X. Chao and Y. Zheng: Transient analysis of immigration birth-death processes with total catastrophes. Probability in the Engineering and Information Sciences, 17 (2003), 83–106.

[5] A. Chen and E. Renshaw: The M/M/1 queue with mass exodus and mass arrivals when empty. Journal of Applied Probability, 34 (1997), 192–207.

[6] A. Dudin and S. Nishimura: A BMAP/SM/1 queueing system with Markovian arrival input of disasters. Journal of Applied Probability, 36 (1999), 868–881.

[7] A. Dudin and O. Semenova: A stable algorithm for stationary distribution calculation for a BMAP/SM/1 queueing system with Markovian arrival input of disasters. Journal of Applied Probability, 41 (2004), 547–556.

[8] J. Hofmann: The BMAP/G/1 queue with level-dependent arrivals - An overview. Telecommunication Systems, 16 (2001), 347–359.

[9] Y. Inoue and T. Takine: The FIFO single-server queue with disasters and multiple Markovian arrivals streams. Journal of Industrial and Management Optimization, 10 (2014), 57–87.

[10] G. Jain and K. Sigman: A Pollaczek-Khintchine formula for M/G/1 queues with disasters. Journal of Applied Probability, 33 (1996), 1191–1200.

[11] M. Kimura and T. Takine: Characterization of the conditional stationary distribution in Markov chains via systems of linear inequalities. to appear in Advances in Applied Probability, 52 (2020).

[12] B.K. Kumar, A. Vijayakumar, and S. Sophia: Transient analysis for state-dependent queues with catastrophes. Stochastic Analysis and Applications, 26 (2008), 1201–1217.

[13] Q.-L. Li and C. Lin: The M/G/1 processor-sharing queue with disasters. Computers and Mathematics with Applications, 51 (2006), 987–998.

[14] Y. Liu: Augmented truncation approximations of discrete-time Markov chains. Operations Research Letters, 38 (2010), 218–222.

[15] H. Masuyama: Error bounds for last-column-block-augmented truncations of block-structured Markov chains. Journal of the Operations Research Society of Japan, 60 (2017), 271–320.

[16] E. Seneta: Computing the stationary distribution for infinite Markov chain. Linear Algebra and its Applications, 34 (1980), 259–267.

[17] M. Shaked and J.G. Shanthikumar: Stochastic Order (Springer, New York, NY, 2007).

[18] Y.W. Shin: BMAP/G/1 queue with correlated arrivals of customers and disasters. Operations Research Letters, 32 (2004), 364–373.

[19] Y.Q. Zhao and D. Liu: The censored Markov chain and best augmentation. Journal of Applied Probability, 33 (1996), 623–629.

[20] J.V. Uspensky: Introduction to Mathematical Probability (McGraw-Hill, New York, NY, 1937).

Masatoshi Kimura
Department of Information and Communications Technology
Graduate School of Engineering
Osaka University
2-1 Yamadaoka, Suita, Osaka 565-0871, Japan
E-mail: m-kimura@post.comm.eng.osaka-u.ac.jp