Sufficient Conditions for a Geometric Tail in a QBD Process with Countably Many Levels and Phases

(Third revision)

Lani Haque[1], Liming Liu[2] and Yiqiang Q. Zhao[1]

[1]School of Mathematics and Statistics

Carleton University, Ottawa, Ontario

Canada K1S 5B6

[2] Industrial Engineering and Engineering Management

The Hong Kong University of Science and Technology

Clear Water Bay, Kowloon Hong Kong

August 25, 2004

Abstract

In this paper, we present sufficient conditions, under which the stationary probability vector of a QBD process with both infinite many levels and phases decays geometrically, characterized by the convergence norm η and the $1/\eta$-left-invariant vector $\boldsymbol{x}$ of the rate matrix R. We also present a method to compute η and $\boldsymbol{x}$ based on spectral properties of the censored matrix of a matrix function constructed with the repeating blocks of the transition matrix of the QBD process. What makes this method attractive is its simplicity; finding η reduces to determining the zeros of a polynomial. We demonstrate the application of our method through a few interesting examples.

Keywords: Two-dimensional system, tail asymptotics, geometric tail, radius of convergence, β-invariant measure, α-positivity, two-demand model, shortest queue model, priority queue, inventory queue.

1 Introduction

For two-dimensional models with both infinitely many levels and phases, the computation of the exact stationary probability distributions is usually very difficulty. Even when the computation is possible, one often has to develop a special mothod for each specific model. For example, the compensation method introduced by Adan [1] is used to compute the joint probability distribution of a specific class of two-dimensional processes. However, this method does not work for the two-stage inventory-queue considered in Haque, Liu and Zhao [5]. The analysis of tail behaviors to characterize the stationary probability distributions along either the level or the phase direction (or some other fixed direction) has been partially motivated by such difficulties.

Takahashi, Fujimoto and Makimoto [17] provide a set of sufficient conditions under which the invariant probability measure of a QBD process with countably many levels and phases has a geometric tail along the level direction, i.e. $\pi_{m,i} \sim c\eta^m x_i$ for each i as $m \to \infty$, or

$$\lim_{m\to\infty} \frac{\pi_{m,i}}{c\eta^m x_i} = 1.$$

Similar conditions are obtained by Foley and McDonald [4] (for a specific queueing model) and Miyazawa [12]. To determine the tail behavior in a specific model, one needs η and $\boldsymbol{x} = (x_0, x_1, \ldots)$ to verify whether these conditions are satisfied. However, neither Takahashi *et al.* [17], nor Foley and McDonald [4], nor Miyazawa [12] provides a method to compute η and $\boldsymbol{x}$. This shortcoming makes it difficult to apply these conditions in practical analysis. In this paper, we propose a set of sufficient conditions equivalent to those of Takahaski *et al.* [17] for the existence of a geometric tail of the stationary distribution of a QBD process. Based on these conditions, we develop a practical method for finding η and $\boldsymbol{x}$, which can be used to verify whether the stationary distribution for a specific process has a geometric tail. Specifically, we construct the so-called spectral equation $\phi(\eta) - 1 = 0$ by censoring the matrix $D(\eta)$, as defined in equation (3) for the given Markov chain. The root of $\phi(\eta) - 1 = 0$, which lies in the interval $(0, 1)$ and satisfies the conditions specified in Theorem 3 of Li and Zhao [8] (see, also [9]), is then the desired η. With η, we can compute the vector $\boldsymbol{x}$ using the RG-factorization of $\eta I - D(\eta)$ according to [19]. In fact, the reciprocal of η is the radius of convergence of both $D(\eta)$ and the rate matrix R for the Markov chain, and $\boldsymbol{x}$ is the $\frac{1}{\eta}$-invariant measure of $D(\eta)$ and R.

The main contributions of this paper are: improving the results in Takahashki *et al.* [17]

and Miyazawa [12]; providing a method for computing η, the reciprocal of the radius of convergence for the R measure; and demonstrating explicitly that the tail asymptotics of the stationary distribution of the Markov chain for some interesting models are geometric along the level direction. We will define the Markov chain in Section 2 and then provide new sufficient conditions for the existence of a geometric tail along its level direction in Section 3. Section 4 shows how to verify $1/\eta$-positivity, while Section 5 provides a method for determining the $1/\eta$-invariant measure $\boldsymbol{x}$ and $1/\eta$-invariant vector $\boldsymbol{y}$. In Section 6, we illustrate the above results through a few examples and show how to verify the boundary condition for each of them.

2 Mathematical Model

Consider an ergodic discrete-time Markov chain $\{X(n); n \geq 0\}$ on a two-dimensional state space $S = \{(m, i); m, i = 0, 1, 2, \ldots\}$ with transition probability matrix

$$P = \begin{bmatrix} B_0 & A_0 & & & & \\ C_0 & B & A & & & \\ & C & B & A & & \\ & & C & B & A & \\ & & & \ddots & \ddots & \ddots \end{bmatrix} \tag{1}$$

after partitioning the state space into levels $\mathcal{L}_m = \{(m, i); i = 0, 1, 2, \ldots\}$ for $m = 0, 1, 2, \ldots$. Let

$$Q = \begin{bmatrix} B & A & & & \\ C & B & A & & \\ & C & B & A & \\ & & \ddots & \ddots & \ddots \end{bmatrix},$$

which is the matrix constructed from P by deleting the boundaries. In this paper, we assume that the repeating blocks A, B and C have a tridiagonal structure, e.g.

$$A = \begin{bmatrix} \overline{a_1} & \overline{a_0} & & & & \\ \overline{a_2} & a_1 & a_0 & & & \\ & a_2 & a_1 & a_0 & & \\ & & a_2 & a_1 & a_0 & \\ & & & \ddots & \ddots & \ddots \end{bmatrix}.$$

Blocks A_0, B_0, and C_0 also have the same tridiagonal structure.

Let $\boldsymbol{\pi} = (\boldsymbol{\pi_0}, \boldsymbol{\pi_1}, \boldsymbol{\pi_2}, \ldots)$ denote the stationary probability vector or the invariant probability measure of P, where $\boldsymbol{\pi_m} = (\pi_{m,0}, \pi_{m,1}, \pi_{m,2}, \ldots)$. We are interested in the conditions under which $\boldsymbol{\pi_m}$ has an exact geometric tail asymptotic, i.e.

$$\boldsymbol{\pi_m} \sim c\eta^m \boldsymbol{x} \text{ as } m \to \infty, \tag{2}$$

where $0 < \eta < 1$, $\boldsymbol{x} = (x_0, x_1, \ldots)$ is a row vector to be determined, and c is a constant. By $\boldsymbol{\pi_m} \sim c\eta^m \boldsymbol{x}$ as $m \to \infty$ we mean that for each n,

$$\lim_{m\to\infty} \frac{\pi_{m,n}}{c\eta^m x_n} = 1.$$

In particular, our goal is to provide a practical method to verify the conditions for specific processes and compute the corresponding value of the decay rate η and vector $\boldsymbol{x}$. To this end, we consider, for $0 < \eta < 1$, the matrix

$$\begin{aligned} D(\eta) &= A + \eta B + \eta^2 C \\ &= \begin{bmatrix} \overline{\gamma}_\eta & \overline{\lambda}_\eta & 0 & 0 & 0 & \cdots \\ \overline{\mu}_\eta & \gamma_\eta & \lambda_\eta & 0 & 0 & \cdots \\ 0 & \mu_\eta & \gamma_\eta & \lambda_\eta & 0 & \cdots \\ 0 & 0 & \mu_\eta & \gamma_\eta & \lambda_\eta & \cdots \\ \vdots & \vdots & & \ddots & \ddots & \end{bmatrix}, \end{aligned} \tag{3}$$

where

$$\begin{aligned} \overline{\gamma}_\eta &= \overline{a_1} + \eta\overline{b_1} + \eta^2\overline{c_1}, \quad \gamma_\eta = a_1 + \eta b_1 + \eta^2 c_1, \\ \overline{\lambda}_\eta &= \overline{a_0} + \eta\overline{b_0} + \eta^2\overline{c_0}, \quad \lambda_\eta = a_0 + \eta b_0 + \eta^2 c_0, \\ \overline{\mu}_\eta &= \overline{a_2} + \eta\overline{b_2} + \eta^2\overline{c_2}, \quad \mu_\eta = a_2 + \eta b_2 + \eta^2 c_2. \end{aligned}$$

The matrix $D(\eta)$ is substochastic due to our choice of η. Also, denote by $E(\eta)$ the submatrix of $D(\eta)$ without the boundary row and column,

$$E(\eta) = \begin{bmatrix} \gamma_\eta & \lambda_\eta & 0 & 0 & 0 & \cdots \\ \mu_\eta & \gamma_\eta & \lambda_\eta & 0 & 0 & \cdots \\ 0 & \mu_\eta & \gamma_\eta & \lambda_\eta & 0 & \cdots \\ 0 & 0 & \mu_\eta & \gamma_\eta & \lambda_\eta & \cdots \\ \vdots & \vdots & & \ddots & \ddots & \end{bmatrix}. \tag{4}$$

We will assume that both $D(\eta)$ and $E(\eta)$ are irreducible.

3 Tail asymptotics

In this section, we provide conditions under which (2) holds. The following two conditions are obtained by Takahashi, Fujimoto and Makimoto [17].

Condition 1 *There exists a positive constant η with $0 < \eta < 1$, a positive row vector $\boldsymbol{x}$ and a positive column vector $\boldsymbol{y}$ such that*

a. $\boldsymbol{x}(\eta^{-1}A + B + \eta C) = \boldsymbol{x}$;

b. $(\eta^{-1}A + B + \eta C)\boldsymbol{y} = \boldsymbol{y}$;

c. $\boldsymbol{x}\boldsymbol{e} < \infty$ *where $\boldsymbol{e}$ is a column vector of ones,* $\boldsymbol{x}\boldsymbol{y} < \infty$;

d. $\eta^{-1}\boldsymbol{x}A\boldsymbol{y} \neq \eta\boldsymbol{x}C\boldsymbol{y}$.

Let $\boldsymbol{z} = A(\eta^{-1}I - G)\boldsymbol{y}$, where $\boldsymbol{y}$ is the column vector in Condition 1 and G is the minimal non-negative solution of

$$G = AG^2 + BG + C.$$

Then, $\boldsymbol{x}$ and $\boldsymbol{z}$ are the row and column $1/\eta$-invariant vectors of R, i.e.

$$\boldsymbol{x} = \frac{1}{\eta}\boldsymbol{x}R \quad \text{and} \quad \boldsymbol{z} = \frac{1}{\eta}R\boldsymbol{z},$$

where R is the minimal non-negative solution of

$$R = A + RB + R^2C.$$

Condition 2 *For the column vector $\boldsymbol{z}$ given above, $\boldsymbol{\pi_1}\boldsymbol{z} < \infty$.*

Under Condition 1 and Condition 2, Takahashi, Fujimoto and Makimoto [17] showed that the tail asymptotic result in (2) holds. However, verifying these two conditions directly is very difficult if not impossible. This has restricted the application of the above result. In the following we construct new conditions which are equivalent to Condition 1 and Condition 2 and can be easily verified.

Condition 3 *Let $\alpha(\eta)$ be the radius of convergence of the matrix $D(\eta)$ (referring to [16] for a definition of the radius of convergence).*

a. *There exists an η with $0 < \eta < 1$ such that $D(\eta)$ is $\alpha(\eta)$-positive with $\alpha(\eta)$-invariant measure and vector $\boldsymbol{x}$ and $\boldsymbol{y}$, respectively, and $\alpha(\eta) = 1/\eta$, and $\boldsymbol{x}\boldsymbol{e} < \infty$.*

b. $\eta^{-1}\boldsymbol{x}A\boldsymbol{y} \neq \eta\boldsymbol{x}C\boldsymbol{y}$.

Condition 4 *The column vector* $\boldsymbol{y}$ *is such that* $\boldsymbol{\pi_1}\boldsymbol{y} < \infty$.

It follows from Theorem 6.4 of Seneta [16] that Condition 3a implies Condition 1a, b and c. Therefore, Condition 3 implies Condition 1. Condition 4 implies Condition 2 noticing that

$$\boldsymbol{z} = A(\eta^{-1}I - G)\boldsymbol{y} \leq \eta^{-1}A\boldsymbol{y} \leq \boldsymbol{y}.$$

Now, a comment on the proof provided by Takahashi, Fujimoto and Makimoto [17] is in order. Recall that the stationary probability vector $\boldsymbol{\pi} = (\boldsymbol{\pi_0}, \boldsymbol{\pi_1}, \boldsymbol{\pi_2}, \ldots)$ for the Markov chain defined in the previous section can be expressed in matrix (operator)-geometric form $\boldsymbol{\pi_{i+1}} = \boldsymbol{\pi_1}R^i$ for $i \geq 0$. Therefore, if the tail behavior of the infinite matrix R can be characterized, then we can characterize the tail asymptotics of $\boldsymbol{\pi_i}$. Theorem 6.5 of Seneta [16], essentially says that if R is $1/\eta$-positive then $\boldsymbol{\pi_i}$ has a geometric tail along the level direction, i.e. $\boldsymbol{\pi_i} \sim c\eta^i\boldsymbol{x}$, where $1/\eta$ is the radius of convergence of R and $\boldsymbol{x}$ is the $1/\eta$-invariant measure of R. In order to use Theorem 6.5 of [16], R must be irreducible and aperiodic, which cannot be verified since R is usually not available. From the following discussion, we will see that this condition can be dropped.

We consider the submatrix $R(A)$ of R consisting of all non-zero rows of R. We may also interprete $R(A)$ as follows. Let $\mathcal{L}_m(A) = \{(m,i) \in \mathcal{L}_m; (A\boldsymbol{e})_i \geq \boldsymbol{0} \text{ and } (A\boldsymbol{e})_i \neq \boldsymbol{0}\}$. According to Lemma 4(a) of [18], the R measure for P may be written as $R = A\hat{Q}_{11}$, where $\hat{Q}_{11}$ is the $(1,1)^{st}$ block entry of the fundamental matrix of Q. The i^{th} row of R is zero if $(m,i) \notin \mathcal{L}_m(A)$. By reordering the levels, we may rewrite R as

$$R = \begin{array}{cc} & \begin{array}{cc} \mathcal{L}_m(A) & \overline{\mathcal{L}_m}(A) \end{array} \\ \begin{array}{c} \mathcal{L}_m(A) \\ \overline{\mathcal{L}_m}(A) \end{array} & \left(\begin{array}{cc} R(A) & \overline{R}(A) \\ 0 & 0 \end{array} \right) \end{array},$$

where $\overline{\mathcal{L}_m}(A) = \mathcal{L}_m^c(A)$. Thus $R(A)$ is the submatrix of R on the index set $\mathcal{L}_m(A) \times \mathcal{L}_m(A)$, and in order to use Theorem 6.5 of [16] it is sufficient to only require $R(A)$ to be irreducible and aperiodic.

A sufficient condition for $R(A)$ to be irreducible is for Q to be irreducible. If this is so, then $\hat{Q}_{11} \geq 0$ but $\hat{Q}_{11} \neq 0$, thus $R(A) > 0$. Once $R(A)$ is irreducible and Condition 3 holds then $R(A)$ is $1/\eta$-positive. If it is further assumed that $R(A)$ is aperiodic, which is true when $D(\eta)$ is aperiodic, then Theorem 6.5 of [16] implies the following theorem, which is equivalent to Theorem 2 of [17]. The aperiodic condition is in fact not crucial according to results in the periodic case.

The outcome of the above discussion can be summarized as the following theorem.

Theorem 1 *Assume Condition 3 and Condition 4 hold. If Q and $D(\eta)$ are irreducible and $D(\eta)$ is aperiodic, then the stationary probability vector $\boldsymbol{\pi}$ of P is geometric along the level variable, i.e.*

$$\boldsymbol{\pi_m} \sim c\eta^m \boldsymbol{x} \quad \text{as } m \to \infty, \tag{5}$$

where $c = \frac{\boldsymbol{\pi_1 z}}{\boldsymbol{xz}}$.■

Remark: Condintion 4 can be replaced by $\boldsymbol{\pi_0 y} < \infty$ in the case that $A_0 = A$ in P and c in Theorem 1 will be given by $c = \frac{\boldsymbol{\pi_0 z}}{\boldsymbol{xz}}$. As discussed earlier, (5) also holds in the case when $D(\eta)$ is periodic. Finally, we noticed (during the revision of this paper) that the requirement $\boldsymbol{xe} < \infty$ in Condition 3a is a natural outcome of other conditions and Condition 3b is not necessary according to [13].

It is also noticed that there may be many different light and heavy tailed asymptotics for the stationary probability vector of a QBD process with countably many levels and phases. In general, the tail asymptotics depends on both non-boundary and boundary transition properties in A, B and C, and A_0, B_0 and C_0, respectively. In this paper, a set of sufficient conditions are provided, under which the stationary probability vector decay geometrically and the decay rate is characterized by the non-boundary transition properties, or the convergence norm of $D(\eta)$. Under other conditions, the boundary transition properties may determine the geometric decay rate, but the geometric decay rate may not coincide with the convergence norm as characterized in this paper.

4 $1/\eta$-positivity

In this section and the next, we will address key issues on how to practically verify Condition 3 given in Theorem 1. The verification of Condition 4 will be addressed through various examples in later sections.

The radius of convergence $1/\overline{\eta}$, or the convergence norm $\overline{\eta}$, for $E(\eta)$ will be of interest to us shortly. For η to be the decay rate, we require that the convergence norm of $D(\eta)$ coincides with the this η, and $\eta > \overline{\eta}$.

The convergence of $\overline{\eta}$ may be found using the result of Kijima in Theorem 2.1 in [7]. This result says that $\overline{\eta}$ is the convergence norm of $E(\eta)$, where $\overline{\eta}$ is given in Lemma 2.1 of [7]. We use the method outlined by Kijima (Lemma 2.1 in [7]) to find $\overline{\eta}$. We need to determine the Perron-Frobenius eigenvalue $\chi(z)$ for the generating function $C^\star(z)$ for the blocks in $D(\eta)$, or $C^\star(z) = \mu_\eta + \gamma_\eta z + \lambda_\eta z^2$ in our case. Since the blocks in $E(\eta)$ in our case are scalars, $C^\star(z)$ is a scalar function. The Perron-Frobenius eigenvalue of $C^\star(z)$ is simply the generating function itself, i.e. $\chi(z) = C^\star(z)$. Now we need to solve

$$\chi(\theta) = \overline{\eta}\theta \text{ and } \chi'(\theta) = \overline{\eta},$$

or

$$\chi(\theta) = \mu_\eta + \gamma_\eta\theta + \lambda_\eta\theta^2 = \overline{\eta}\theta \tag{6}$$

and

$$\chi'(\theta) = \gamma_\eta + 2\lambda_\eta\theta = \overline{\eta} \tag{7}$$

for $\overline{\eta}$. Substituting $\overline{\eta}$ in (7) into (6) and solving for θ, we obtain

$$\theta = \sqrt{\frac{\mu_\eta}{\lambda_\eta}}.$$

The positive square root is taken since $\theta > 0$. Substituting this value for θ into (7) we have the convergence norm of $E(\eta)$

$$\overline{\eta} = \gamma_\eta + 2\sqrt{\lambda_\eta\mu_\eta} = a_1 + \eta b_1 + \eta^2 c_1 + 2\sqrt{(a_0 + \eta b_0 + \eta^2 c_0)(a_2 + \eta b_2 + \eta^2 c_2)}. \tag{8}$$

Next, based on censoring and spectral properties of $D(\eta)$, we compute the roots of the spectral equation $\phi(\eta) - 1 = 0$. Then, the determination of $1/\eta$-positivity of $D(\eta)$ or Condition 3 will be done by comparing the convergence norm η of $D(\eta)$ to $\overline{\eta}$.

For matrix $D(\eta)$, consider the censored entry $\phi(\eta)$ with a single state phase zero as the censoring set, which is given by

$$\phi(\eta) = \overline{\gamma}_\eta/\eta + \frac{\overline{\lambda}_\eta\overline{\mu}_\eta}{2\lambda_\eta\mu_\eta}\left(1 - \gamma_\eta/\eta - \sqrt{(\gamma_\eta/\eta - 1)^2 - 4\lambda_\eta\mu_\eta/\eta^2}\right).$$

Consider the roots of the spectral equation $\phi(\eta) - 1 = 0$, which are also zeros of

$$[(\overline{\lambda}_\eta\overline{\mu}_\eta)^2 - \overline{\lambda}_\eta\overline{\mu}_\eta\overline{\gamma}_\eta\gamma_\eta + \lambda_\eta\mu_\eta\overline{\gamma}_\eta^2] + \eta[-2\overline{\gamma}_\eta\lambda_\eta\mu_\eta + \overline{\lambda}_\eta\overline{\mu}_\eta(\overline{\gamma}_\eta + \gamma_\eta)] + \eta^2[\lambda_\eta\mu_\eta - \overline{\lambda}_\eta\overline{\mu}_\eta]. \tag{9}$$

Notice that (9) is a polynomial of at most 8 degrees evaluated at η and will henceforth be denoted by

$$q(\eta) = D_0 + D_1\eta + \ldots + D_8\eta^8, \tag{10}$$

where the coefficients D_i are functions of the entries of A, B and C. Moreover, a zero η of $q(\eta)$ is a zero of $\phi(\eta) - 1$ if and only if

$$\frac{\overline{\gamma}_\eta}{\eta} + \frac{\overline{\lambda}_\eta \overline{\mu}_\eta}{2\lambda_\eta \mu_\eta}\left(1 - \frac{\gamma_\eta}{\eta}\right) - 1 \geq 0. \tag{11}$$

This provides us with a practical method to determine the value of η such that $D(\eta)$ is possibly $1/\eta$-positive. It also provides us with a region in which the tail asymptotic property given in Theorem 1 may hold (see the priority queueing model given in Section 6.2). Together with the following theorem, one can check the $1/\eta$-positivity of $D(\eta)$.

Theorem 2 *If η is a zero of $\phi(\eta) - 1$ in the interval $(0,1)$ satisfying $\eta \geq \overline{\eta}$, then the convergence norm of $D(\eta)$ is given by η and $D(\eta)$ is $\frac{1}{\eta}$-recurrent. In this case, η is the only zero of $\phi(\eta) - 1$ satisfying $0 < \eta < 1$. Moreover, if $\eta > \overline{\eta}$ then $D(\eta)$ is $\frac{1}{\eta}$-positive. In this case, the $1/\eta$-invariant measure $\boldsymbol{x}$ and $1/\eta$-invariant vector $\boldsymbol{y}$ of $D(\eta)$ exist and $\boldsymbol{xy} < \infty$.*

Proof: We want to show that $D(\eta)$ is $1/\eta$-recurrent. If we can show that $1/\eta$ is the $\beta^\star$ in Theorem 3(ii) of [8], then $D(\eta)$ is $\frac{1}{\eta}$-recurrent. This is obvious since $0 < \eta \geq \overline{\eta}$ and $\phi(\eta) = 1$. The uniqueness follows from the uniqueness of the radius of convergence of $D(\eta)$. If the radius of convergence $1/\eta$ of $D(\eta)$ satisfies $\frac{1}{\eta} < \overline{\alpha}$ and $D(\eta)$ is $1/\eta$-recurrent, by Theorem 4 and Remark 8 of [8], $D(\eta)$ is $1/\eta$-positive.

By Theorem 6.2 of [16] an $1/\eta$-invariant measure $\boldsymbol{x}$ of $D(\eta)$ exists, i.e. $\boldsymbol{x}D(\eta) = \eta\boldsymbol{x}$. Similarly, we can show that $D(\eta)$ has an $1/\eta$-invariant vector $\boldsymbol{y}$. If $D(\eta)$ is $1/\eta$-positive, by Theorem 6.4 of [16], $\boldsymbol{xy} < \infty$. ■

5 $1/\eta$-invariant measures and vectors

In this section, we show how to explicitly determine the $1/\eta$-invariant measure $\boldsymbol{x}$ and $1/\eta$-invariant vector $\boldsymbol{y}$ when $D(\eta)$ is $1/\eta$-positive. We also reformulate Condition 3b in terms of the calculated $1/\eta$-invariant measure and $1/\eta$-invariant vector.

By Theorem 10 in [9], the $1/\eta$-invariant measure $\boldsymbol{x} = (x_0, x_1, x_2, \ldots)$ of $D(\eta)$ is given by

$$x_0 = z_0, \;\; x_k = z_0 R_{01}(1/\eta) R^{k-1}(1/\eta), \;\; k \geq 1, \tag{12}$$

where $z_0 > 0$ is any positive number. $R(1/\eta)$ is given by

$$R(1/\eta) = \frac{(\eta - \gamma_\eta) - \sqrt{(\gamma_\eta - \eta)^2 - 4\lambda_\eta\mu_\eta}}{2\mu_\eta}$$

and $R_{01}(1/\eta)$ by

$$R_{01}(1/\eta) = \frac{\overline{\lambda}_\eta}{\lambda_\eta} R(1/\eta).$$

In order to find the $1/\eta$-invariant vector $\boldsymbol{y} = (y_0, y_1, y_2, \ldots)^T$ of $D(\eta)$, we use the *RG*-factorization given by equation (25) in [9]. That is using

$$I - \frac{1}{\eta} D(\eta) = (I - R_U(1/\eta))(I - U_D(1/\eta))(I - G_L(1/\eta))$$

where

$$I - R_U(1/\eta) = \begin{bmatrix} 1 & -R_{01}(1/\eta) & & & \\ & 1 & -R(1/\eta) & & \\ & & 1 & -R(1/\eta) & \\ & & & \ddots & \ddots \end{bmatrix},$$

$$I - U_D(1/\eta) = \begin{bmatrix} 1 - U_0(1/\eta) & & & \\ & 1 - U(1/\eta) & & \\ & & 1 - U(1/\eta) & \\ & & & \ddots \quad \ddots \end{bmatrix},$$

and

$$1 - G_L(1/\eta) = \begin{bmatrix} 1 & & & & \\ -G_{10}(1/\eta) & 1 & & & \\ & -G(1/\eta) & 1 & & \\ & & -G(1/\eta) & 1 & \\ & & & \ddots & \ddots \end{bmatrix}.$$

We now solve

$$0 = \left(I - \frac{1}{\eta} D(\eta)\right) \boldsymbol{y} = (I - R_U(1/\eta))(I - U_D(1/\eta))(I - G_L(1/\eta))\boldsymbol{y}.$$

Letting $\boldsymbol{w} = (I - G_L(1/\eta))\boldsymbol{y}$, we have

$$0 = (I - R_U(1/\eta))(I - U_D(1/\eta))\boldsymbol{w}. \tag{13}$$

The product $(I - R_U(1/\eta))(I - U_D(1/\eta))$ expands to

$$\begin{bmatrix} 1 - U_0(1/\eta) & -R_{01}(1/\eta)(1 - U(1/\eta)) & & & \\ & 1 - U(1/\eta) & -R(1/\eta)(1 - U(1/\eta)) & & \\ & & 1 - U(1/\eta) & -R(1/\eta)(1 - U(1/\eta)) & \\ & & & \ddots & \ddots \end{bmatrix}.$$

Therefore, (13) becomes

$$0 = (1 - U_0(1/\eta))w_0 - R_{01}(1/\eta)(1 - U(1/\eta))w_1, \tag{14}$$

$$0 = (1 - U(1/\eta))w_n - R(1/\eta)(1 - U(1/\eta))w_{n+1}, \ n \geq 1. \tag{15}$$

Since $U_0(1/\eta) = 1$, we may take $\boldsymbol{w} = (w_0, 0, 0, \ldots)$ with $w_0 > 0$ to satisfy (14) and (15). Now we need to solve

$$\boldsymbol{w} = (I - G_L(1/\eta))\boldsymbol{y}. \tag{16}$$

Simplifying (16) gives us

$$\begin{aligned}
0 &= w_1 = -G_{10}(1/\eta)y_0 + y_1 \quad \text{or} \quad y_1 = G_{10}(1/\eta)y_0, \\
0 &= w_2 = -G(1/\eta)y_1 + y_2 \quad \text{or} \quad y_2 = G_{10}(1/\eta)G(1/\eta)y_0, \\
&\vdots \\
0 &= w_n = -G(1/\eta)y_{n-1} + y_n \quad \text{or} \quad y_n = G^{n-1}(1/\eta)G_{10}(1/\eta)y_0.
\end{aligned}$$

Therefore, the $1/\eta$-invariant vector $\boldsymbol{y} = (y_0, y_1, y_2, \ldots)^T$ of $D(\eta)$ is given by

$$y_0 = w_0 > 0, \ y_n = G_{10}(1/\eta)G^{n-1}(1/\eta)w_0, \ n \geq 1, \tag{17}$$

where $G(1/\eta)$ is given by

$$G(1/\eta) = \frac{(\eta - \gamma_\eta) - \sqrt{(\gamma_\eta - \eta)^2 - 4\lambda_\eta \mu_\eta}}{2\lambda_\eta}$$

and $G_{10}(1/\eta)$ by

$$G_{10}(1/\eta) = \frac{\overline{\mu}_\eta}{\mu_\eta} G(1/\eta).$$

Using the above expressions for the $1/\eta$-invariant measure $\boldsymbol{x}$ and vector $\boldsymbol{y}$ the result $\boldsymbol{xy} < \infty$ reduces to

$$\begin{aligned}
\boldsymbol{xy} &= \sum_{i=0}^{\infty} x_i y_i = z_0 w_0 + \sum_{n=1}^{\infty} x_n y_n \\
&= z_0 w_0 \left(1 + R_{10}(1/\eta) G_{10}(1/\eta) \sum_{n=1}^{\infty} R^{n-1}(1/\eta) G^{n-1}(1/\eta) \right) \\
&= z_0 w_0 \left(1 + \frac{\overline{\lambda}_\eta \overline{\mu}_\eta}{\lambda_\eta \mu_\eta} \sum_{n=1}^{\infty} \left(\frac{\mu_\eta}{\lambda_\eta} \right)^n R^{2n}(1/\eta) \right) < \infty,
\end{aligned}$$

where $G(1/\eta) = \frac{\mu_\eta}{\lambda_\eta} R(1/\eta)$. So,

$$\boldsymbol{x}\boldsymbol{y} < \infty \quad \text{if and only if} \quad \sum_{n=1}^{\infty} \left(\frac{\mu_\eta}{\lambda_\eta} R^2(1/\eta) \right)^n < \infty.$$

By Theorem 5.2 of [15], $\boldsymbol{x}$ is the $\frac{1}{\eta}$-invariant measure of R. However, we require an $1/\eta$-invariant vector $\boldsymbol{z}$ for R as well in order to explicitly determine the constant c in Theorem 1. Condition 3b is the remaining condition required by Takahashi *et al.* [17], which reduces to

$$\frac{1}{\eta} \left(y_0(x_0\overline{a_1} + x_1\overline{a_2}) + y_1(\overline{a_0} + x_1a_1 + x_2a_2) + \sum_{n=2}^{\infty} y_n(x_{n-1}a_0 + x_na_1 + x_{n+1}a_2) \right)$$
$$\neq \eta \left(y_0(x_0\overline{c_1} + x_1\overline{c_2}) + y_1(\overline{c_0} + x_1c_1 + x_2c_2) + \sum_{n=2}^{\infty} y_n(x_{n-1}c_0 + x_nc_1 + x_{n+1}c_2) \right)$$

upon expansion. From the above we see that, if

$$\frac{1}{\eta}\overline{a_i} = \eta\overline{c_i}, \quad \text{and } \frac{1}{\eta}a_i = \eta c_i, \quad \text{for } i = 0, 1, 2,$$

then

$$\frac{1}{\eta}\boldsymbol{x}A\boldsymbol{y} = \eta\boldsymbol{x}C\boldsymbol{y}.$$

Otherwise, $\frac{1}{\eta}\boldsymbol{x}A\boldsymbol{y} \neq \eta\boldsymbol{x}C\boldsymbol{y}$. However, as remarked earlier this condition is not required.

6 Examples

In this section we will be looking at the application of Theorem 1 through a few examples. The first example is the two-demand model studied by Flatto and Hahn [3]. It illustrats the case that $D(\eta)$ is $1/\eta$-positive only if $\mu_1 \neq \mu_2$. The second example is a priority queueing model, by which we demonstrate how Condition 4 can be verified and when Theorem 1 can be applied. The third example is the classic symmetric shortest queue, which has been studied by many authors. This example is used to show when Theorem 1 can be applied, in which the verification of Condition 4 is not trivial. The last example is a two-stage inventory queue under an echelon base-stock control policy studied in Haque, Liu and Zhao [5]. This example again illustrates when Theorem 1 can be applied. We notice that in the first and the last examples, Condition 4 is obviously satisfied.

6.1 Two-demand model

A double queue [3] arises when arriving customers to the system simultaneously place two demands to two different servers working independently. In this queue, customer arrivals form a Poisson process with rate 1 and the service times at the two servers are exponential with rates μ_1 and μ_2, respectively. Let $X_1(t)$ and $X_2(t)$ represent the numbers of customers waiting or in service at time t in the queues with service rates μ_1 and μ_2, respectively. Then $X(t) = (X_1(t), X_2(t))$ is a Markov chain with state space $S = \{(i,j); i,j = 0,1,2,\ldots\}$. This Markov chain is stable if and only if $\mu_1 > 1$ and $\mu_2 > 1$. Without loss of generality, we assume that $1 < \mu_1 \le \mu_2$. For convenience, the servers with rate μ_1 and μ_2 are referred to as the slower and faster servers, respectively, and their queues as slower and faster queues, respectively.

The equilibrium equations are given as follows,

$$\begin{aligned}
(1+\mu_2)\pi_{0,j} &= \mu_1\pi_{1,j} + \mu_2\pi_{0,j+1}, \;\; i=0, j\ge 1, \\
(1+\mu_1)\pi_{i,0} &= \mu_1\pi_{i+1,0} + \mu_2\pi_{i,1}, \;\; i\ge 1, j=0, \\
(1+\mu_1+\mu_2)\pi_{i,j} &= \mu_1\pi_{i+1,j} + \mu_2\pi_{i,j+1} + \pi_{i-1,j-1}, \;\; i\ge 1, j\ge 1, \\
\pi_{0,0} &= \mu_1\pi_{1,0} + \mu_2\pi_{0,1},
\end{aligned}$$

from which we have an infinitesimal generator, which upon uniformization gives the transition matrix P in equation (1). Specifically, let $\theta = 1/(1+\mu_1+\mu_2)$, then

$$A = \begin{bmatrix} 0 & \theta & & & \\ & 0 & \theta & & \\ & & 0 & \theta & \\ & & & \ddots & \ddots \end{bmatrix},$$

$$B = \begin{bmatrix} 1-\theta(1+\mu_1) & & & & \\ \theta\mu_2 & 0 & & & \\ & \theta\mu_2 & 0 & & \\ & & \ddots & \ddots & \end{bmatrix},$$

and $C = \theta\mu_1 I$. In order to apply Theorem 2 we need the matrix in (3), where

$$\begin{aligned}
\overline{\gamma}_\eta &= \eta(1-\theta(1+\mu_1)) + \eta^2\theta\mu_1, & \gamma_\eta &= \eta^2\theta\mu_1, \\
\overline{\lambda}_\eta &= \theta, & \lambda_\eta &= \overline{\lambda}_\eta, \\
\overline{\mu}_\eta &= \eta\theta\mu_2, & \mu_\eta &= \overline{\mu}_\eta.
\end{aligned}$$

Using (8), the convergence norm $\overline{\eta}$ of $E(\eta)$ is given by

$$\overline{\eta} = \frac{\eta^2\mu_1 + 2\sqrt{\eta\mu_2}}{1+\mu_1+\mu_2}. \tag{18}$$

The convergence norm of $D(\eta)$ is given by η, where η is a zero of the function in (10). Simplifying the equation from (9) gives

$$\frac{\eta^2\mu_2^2}{(1+\mu_1+\mu_2)^4}(1-\mu_1\eta)(1-\eta) = 0,$$

which is a polynomial equation of degree two and yields two roots $\eta = \frac{1}{\mu_1}$ and 1. Since $\mu_1 > 1$, we have $0 < \frac{1}{\mu_1} < 1$, and hence the convergence norm of $D(\eta)$ or of the R-measure for P is $\eta = \mu_1$.

Next we need to determine if $D(\eta)$ is $1/\eta$-positive, which is equivalent to determining if $\eta > \overline{\eta}$. Substituting $\eta = \frac{1}{\mu_1}$ into (18), the convergence norm for $E(\eta)$ is reduced to

$$\overline{\eta} = \frac{1}{\mu_1}\frac{1+2\sqrt{\mu_1\mu_2}}{1+\mu_1+\mu_2}.$$

We need to determine if $\frac{1+2\sqrt{\mu_1\mu_2}}{1+\mu_1+\mu_2} < 1$. For $\mu_1 \neq \mu_2$,

$$\begin{aligned}
0 &< (\sqrt{\mu_1} - \sqrt{\mu_2})^2, \\
0 &< \mu_1 - 2\sqrt{\mu_1\mu_2} + \mu_2, \\
1 &< 1 + \mu_1 + \mu_2 - 2\sqrt{\mu_1\mu_2}, \\
1 + 2\sqrt{\mu_1\mu_2} &< 1 + \mu_1 + \mu_2, \\
\frac{1+2\sqrt{\mu_1\mu_2}}{1+\mu_1+\mu_2} &< 1
\end{aligned}$$

as required. Therefore, $1/\mu_1 = \eta > \overline{\eta}$. By theorem 2, the $1/\eta$-invariant measure $\boldsymbol{x}$ and vector $\boldsymbol{y}$ for $D(\eta)$ exist, for $\mu_1 \neq \mu_2$. Next, we need to determine if $\frac{1}{\eta}\boldsymbol{x}A\boldsymbol{y} \neq \eta\boldsymbol{x}C\boldsymbol{y}$. We may use the expressions for the $1/\eta$-invariant measure $\boldsymbol{x}$ and vector $\boldsymbol{y}$ given in equations (12) and (17), respectively. The R-measure of $D(\eta)$ is given by

$$R(1/\eta) = \frac{\mu_1}{\mu_2}$$

since $1 < \mu_1 < \mu_2$ and $R(1/\eta)$ is non-negative. We have

$$R_{01}(1/\eta) = \frac{\mu_1}{\mu_2}, \ \ G(1/\eta) = 1, \quad \text{and } G_{10}(\alpha) = 1.$$

By equation (12) the $1/\eta$-invariant measure of $D(\eta)$ is

$$\boldsymbol{x} = \left(1, \frac{\mu_1}{\mu_2}, \left(\frac{\mu_1}{\mu_2}\right)^2, \ldots, \left(\frac{\mu_1}{\mu_2}\right)^n, \ldots\right),$$

and by equation (17) the $1/\eta$-invariant vector for $D(\eta)$ is $\boldsymbol{y} = (1, 1, 1, \ldots)^T = \boldsymbol{e}$. Finally, since $\boldsymbol{y} = \boldsymbol{e}$ and $\sum_{m,n=0}^{\infty} \pi_{m,n} = 1$, $\boldsymbol{\pi_0 e} < \infty$, which means that Condition 4 is satisfied due to the remark following Theorem 1. Therefore, we may conclude that

$$\boldsymbol{\pi_m} \sim c\eta^m \boldsymbol{x} \quad = \quad \frac{c}{\mu_1^m} \left(1, \frac{\mu_1}{\mu_2}, \ldots, \left(\frac{\mu_1}{\mu_2}\right)^n, \ldots\right),$$

which is consistent with the result in Theorems 7.1 and 7.2 of [3].

When $\mu_1 = \mu_2$, the polynomial equation from (10) reduces to

$$\frac{\mu_1^2}{(1+2\mu_1)^4}(1-\mu_1\eta)(1-\eta) = 0,$$

which again implies $\eta = 1$ or $\eta = \frac{1}{\mu_1}$. Therefore $\eta = \frac{1}{\mu_1} < 1$ since $1 < \mu_1$. We now have

$$\overline{\eta} = \frac{\eta^2\mu_1 + 2\sqrt{\eta\mu_1}}{1+2\mu_1} = \frac{1}{\mu_1}\frac{(1+2\mu_1)}{(1+2\mu_1)} = \frac{1}{\mu_1} = \eta.$$

Therefore by Theorem 3 of [8], $D(\eta)$ is $\frac{1}{\eta}$-recurrent since $\phi(\eta) = 1$. In order to determine if $D(\eta)$ is $\frac{1}{\eta}$-positive, we need to check whether $\eta > \overline{\eta}$ holds. From the above discussion, we see that $\eta = \overline{\eta}$. So $D(\eta)$ is not $\frac{1}{\eta}$-positive, but $\frac{1}{\eta}$-null. In this case, Theorem 1 cannot be applied. It turns out that $\boldsymbol{\pi_m}$ does not have an exact geometric asymptotic tail, as shown in Theorem 7.1 and 7.2 of Flatto and Hahn [3].

6.2 Priority queue

Consider the classical preemptive priority queue (Miller [11]). Two classes of Poisson customers arrive independently at rate λ_l for the lower priority customers and at rate λ_h for higher priority customers. For simplicity, we assume that both classes of customers require the same exponential service time at rate μ_c from a single server. With the preemptive rule, a higher priority customer, upon arrival, passes all lower priority customers in the queue or takes over the service if a lower priority customer is currently being served. Let $X_l(t)$ and $X_h(t)$ be the number of lower and higher priority customers in the system at time t, respectively. Then, $X(t) = (X_l(t), X_h(t))$ is a continuous-time Markov chain with state space $S = \{(i,j); i, j = 0, 1, 2, \ldots\}$. Take $X_l(t)$ as the level. The infinitesimal generator upon uniformization gives the transition matrix P as in equation (1), where

$A = \lambda_l I$,

$$B = \begin{bmatrix} 0 & \lambda_h & & & \\ \mu_c & 0 & \lambda_h & & \\ & \mu_c & 0 & \lambda_h & \\ & & \ddots & \ddots & \ddots \end{bmatrix},$$

$$C = \begin{bmatrix} \mu_c & & & \\ & 0 & & \\ & & 0 & \\ & & & \ddots \end{bmatrix}.$$

In order to apply Theorem 2 we need matrix (3), where

$$\begin{aligned} \overline{\gamma}_\eta &= \lambda_l + \mu_c \eta^2, \quad \gamma_\eta = \lambda_l, \\ \overline{\lambda}_\eta &= \lambda_\eta = \lambda_h \eta, \\ \overline{\mu}_\eta &= \mu_\eta = \mu_c \eta. \end{aligned}$$

Using (8), the convergence norm $\overline{\eta}$ of $E(\eta)$ is given by

$$\overline{\eta} = \lambda_l + 2\eta\sqrt{\lambda_h \mu_c}. \tag{19}$$

The convergence norm of $D(\eta)$ is given by η, where η is a zero of function (10). (10) simplifies to

$$\lambda_h \mu_c^2 \eta^4 (\lambda_l + \lambda_h - \eta + \mu_c \eta^2),$$

which is reduced to a polynomial equation of degree two for our purpose. The only zero η in $(0, 1)$ is

$$\eta = \rho = \frac{\lambda_l + \lambda_h}{\mu_c}.$$

The condition $\eta < 1$ coincides with the stability condition of the system. Let

$$\rho_l = \frac{\lambda_l}{\mu_c} \quad \text{and} \quad \rho_h = \frac{\lambda_h}{\mu_c}.$$

To see if $\phi(\rho) = 1$, we need to verify inequality (11), which in this case is equivalent to

$$\lambda_l \mu_c + 2(\lambda_l + \lambda_h)^2 - (\lambda_l + \lambda_h) > 0.$$

Substituting $\lambda_l + \lambda_h + \mu_c = 1$, the above becomes $(\lambda_l + \lambda_h)^2 > \lambda_h \mu_c$ or $\rho^2 > \rho_h$.

For $D(\eta)$ to be $1/\eta$-positive, we need $\eta > \overline{\eta}$ or equivalently

$$\frac{\lambda_l + \lambda_h}{\mu_c} > \lambda_l + 2(\lambda_l + \lambda_h)\sqrt{\frac{\lambda_h}{\mu_c}}. \tag{20}$$

Inequality (20) provides us a region within which Thoerem 1 holds.

We use the expressions in (12) and (17) for the $1/\eta$-invariant measure $\boldsymbol{x}$ and vector $\boldsymbol{y}$ which become, after some calculations,

$$\boldsymbol{x} = \left(1, \frac{\rho_h}{\rho}, \left(\frac{\rho_h}{\rho}\right)^2, \dots\right)$$

and

$$\boldsymbol{y} = \left(1, \frac{1}{\rho}, \left(\frac{1}{\rho}\right)^2, \dots\right)^T.$$

It is clear now that the verification of Condition 4 is not trivial. We recall that the marginal distribution $\pi_{\cdot,n} = \sum_m \pi_{m,n}$ of the number of customers in the system with the higher priority is the same as the distribution of the number of customers in the system for the standard $M/M/1$ queue. Therefore,

$$\pi_{\cdot,n} = (1 - \rho_h)\rho_h^n, \quad n \geq 0.$$

Condition 4 then becomes

$$\boldsymbol{\pi_1 y} = \sum_n \pi_{1,n} y_n \leq \sum_n \pi_{\cdot,n} y_n = (1 - \rho_h) \sum_n \left(\frac{\rho_h}{\rho}\right)^n < \infty,$$

since $\rho_h/\rho \leq \sqrt{\rho_h}/\rho < 1$. In conclusion, when (20) is true, we have $\boldsymbol{\pi_m} \sim c\boldsymbol{x}\rho^m$ according to Theorem 1.

6.3 Symmetric shortest queue

The SSQ model, i.e. the symmetric shortest queue may be described as follows: Customers arrive in a Poisson stream at rate λ to two identical exponential servers in parallel with common service rate μ, where $\rho = \lambda/2\mu < 1$ for the system stability. Each of the two parallel server has its own queue and, upon arrival, a job joins the shortest queue. If the queue lengths are equal, the customer joins either queue with probability 1/2. Let $Y_1(t)$ and $Y_2(t)$ be the lengths of the two queues at time t, then $Y(t) = (Y_1(t), Y_2(t))$ is a Markov chain with state space $S = \{(i,j); i,j = 0,1,2,\dots\}$, but its infinitesimal generator is level dependent. To eliminate this dependence, let $X_1(t) = \min(Y_1(t), Y_2(t))$

and $X_2(t) = |Y_2(t) - Y_1(t)|$. Now $X(t) = (X_1(t), X_2(t))$ is a Markov chain with state space $S = \{(i,j); i,j = 0,1,2,\ldots\}$ and a level independent infinitesimal generator, taking $X_1(t)$ as the level. Upon uniformization and assuming $\lambda + 2\mu = 1$, we obtain the transition probability matrix P as given in equation (1) with

$$A = \begin{bmatrix} 0 & & & & \\ \lambda & 0 & & & \\ & \lambda & 0 & & \\ & & & \ddots & \ddots \end{bmatrix},$$

$$B = \begin{bmatrix} 0 & \lambda & & & \\ \mu & 0 & 0 & & \\ & \mu & 0 & 0 & \\ & & \ddots & \ddots & \ddots \end{bmatrix},$$

and

$$C = \begin{bmatrix} 0 & 2\mu & & & \\ & 0 & \mu & & \\ & & 0 & \mu & \\ & & & \ddots & \ddots \end{bmatrix}.$$

For the matrix $D(\eta)$ defined by (3), we have

$$\begin{aligned} &\overline{\gamma}_\eta = \gamma_\eta = 0, \\ &\overline{\lambda}_\eta = 2\mu\eta^2 + \lambda\eta, \qquad \lambda_\eta = \mu\eta^2, \\ &\overline{\mu}_\eta = \mu_\eta = \lambda + \mu\eta. \end{aligned}$$

Using (8), the convergence norm $\overline{\eta}$ of $E(\eta)$ is then given by

$$\overline{\eta} = 2\sqrt{\lambda_\eta \mu_\eta} = 2\eta\sqrt{\mu(\lambda + \mu\eta)}. \tag{21}$$

To determine the convergence norm of $D(\eta)$, we first compute the zeros of (9), which can be simplified as follows

$$\eta^2(\lambda + \mu\eta)^2[(2\mu\eta + \lambda)^2 - \eta],$$

since $1 = (\lambda + 2\mu)^2 = \lambda^2 + 4\lambda\mu + 4\mu^2$. This leads to a polynomial equation of degree two and yields $\eta = 1$ or $\eta = \rho^2$. We note that $\eta = \rho^2$ is the only candidate for $\phi(\eta) = 1$. Also, it is clear that the convergence norm of $D(\eta)$ is $\eta = \rho^2 > \overline{\eta}$, and, based on which, we conclude that $D(\rho^2)$ is $1/\rho^2$-positive and, by Theorem 2, the $1/\rho^2$-invariant measure $\boldsymbol{x}$ and vector $\boldsymbol{y}$ exist. Equations (12) and (17) may be used to show, respectively, that

$$\boldsymbol{x} = \left(\mu \frac{2+\rho}{\rho}, 1, \frac{\rho^2}{2+\rho}, \left(\frac{\rho^2}{2+\rho}\right)^2, \left(\frac{\rho^2}{2+\rho}\right)^3, \dots \right)$$

and

$$\boldsymbol{y} = \left(1, \frac{1}{\rho}, \frac{1}{\rho^2}, \frac{1}{\rho^3}, \dots \right).$$

So far, all conditions required by Theorem 1 can be verified except Condition 4 or $\boldsymbol{\pi_0 y} < \infty$. This is not trivial and the method used for the priority queueing model cannot be applied here since we do not know the marginal distribution. Following the proof of (37) in Proposition 1 in Foley and McDonald [4], it can be shown that Theorem 14.3.7 of Meyn and Tweedie [10] guarantees our Condition 4.

In conclusion, we have $\boldsymbol{\pi_m} \sim c\eta^n \boldsymbol{x}$.

6.4 Two-stage inventory-queue

We consider the two-stage inventory-queue model studied in Haque, Liu and Zhao [5]. An inventory-queue is formed by a server (we only consider single server service stations) and an output store following the server. When a demand is received from the outside or a downstream station, it will be satisfied by an available unit from the store or backordered if the store is empty. In either case, the arrival of a demand will trigger the creation of a new job which will either be processed immediately by the server if it is free or join a job queue in front of the server. The server stops when there is no job in the job queue or when there is no material supply for the jobs. We assume that R_1 units are available initially in the output store. Then from the above description, the server will always try to restore the initial inventory level R_1 in the output store. We call R_1 the base stock level and such a system base-stock inventory-queue. Our two-stage system is formed by two base-stock inventory-queues in tandem. We assume that station 1 at the upstream has unlimited material supplies and hence the server will keep working as long as the job queue is not empty. Station 2 at the downstream satisfies the external demand from its output store. When an external demand arrives, it triggers the creation of a job at the upstream station as well as the downstream station. A job in the job queue of station 2 can only be processed when it has a matching unit from the output store. When server 2 is free and the first job of its job queue does not have a matching unit because the upstream output store is out of stock, the server will be idle until the service completion at the upstream station. This phenomenon is called starvation. We also note that probabilistically the inventory level at the upstream output store has no upper bound. One may refer to [5] for details on how this two stage system works.

Let the external demand process be Poisson with rate λ, and the service times at both stations be exponential with rates μ_1 and μ_2 for station 1 and station 2, respectively. The stability conditions for this system are $\rho_1 = \lambda/\mu_1 < 1$ and $\rho_2 = \lambda/\mu_2 < 1$. Let $N_i(t)$, $i = 1, 2$, be the number of jobs in the job queue of station i at time t, and $U(t)$ the number of jobs in the job queue of station 2 which have matching units from output store 1 at time t. When every job in the job queue of station 2 has a matching unit, $U(t)$ equals $N_2(t)$, the number of jobs in the job queue in station 2 at time t. It can be shown that $(N_1(t), U(t))$ is a continuous-time two-dimensional Markov chain. Upon uniformization, the transition matrix P is given by

$$P = \begin{array}{c} \text{level} \\ L_{\leq R_1} \\ R_1+1 \\ R_1+2 \\ R_1+3 \\ \vdots \\ \vdots \end{array} \begin{array}{c} \begin{array}{cccccc} L_{\leq R_1} & R_1+1 & R_1+2 & \cdots & \cdots & \cdots \end{array} \\ \left(\begin{array}{cccccc} B_0 & A_0 & & & & \\ C_0 & B & A & & & \\ & C & B & A & & \\ & & C & B & A & \\ & & & \ddots & \ddots & \ddots \\ & & & & \ddots & \ddots & \ddots \end{array} \right) \end{array},$$

where $L_{\leq R_1}$ represents levels 0, 1, ..., R_1, and $A = \theta\lambda I$,

$$B = \theta \begin{pmatrix} \mu_2 & & & & \\ \mu_2 & 0 & & & \\ & \mu_2 & 0 & & \\ & & \ddots & \ddots & \\ & & & \ddots & \ddots \end{pmatrix}$$

and

$$C = \theta \begin{pmatrix} 0 & \mu_1 & & & & \\ & 0 & \mu_1 & & & \\ & & 0 & \mu_1 & & \\ & & & \ddots & \ddots & \\ & & & & \ddots & \ddots \end{pmatrix}$$

with $\theta = 1/(\lambda + \mu_1 + \mu_2)$. Also recall that R_1 denotes the base-stock level at stations 1. We note that details about A_0, B_0 and C_0 are not provided here since they are not important in our discussion.

We need to consider (3), which now has $\overline{\gamma}_\eta = \theta(\lambda + \eta\mu_2)$, $\overline{\lambda}_\eta = \lambda_\eta = \theta\eta^2\mu_1$, $\gamma_\eta = \theta\lambda$ and $\overline{\mu}_\eta = \mu_\eta = \theta\eta\mu_2$. The convergence norm $\overline{\eta}$ of $E(\eta)$ in (4) is given by

$$\overline{\eta} = \frac{\lambda + 2\eta\sqrt{\eta\mu_1\mu_2}}{\lambda + \mu_1 + \mu_2}. \tag{22}$$

The convergence norm of $D(\eta)$ is given by η, where η is a zero of (9) given by

$$\theta^4\eta^4\mu_1\mu_2^2[\eta^2\mu_1 + \lambda - \eta(\lambda + \mu_1)]. \tag{23}$$

This leads to a polynomial equation of degree two and its only root in $(0, 1)$ is λ/μ_1. To be $1/\eta$-positive, we need $\lambda/\mu_1 > \overline{\eta}$. To verify Condition 3b, we need the $1/\eta$-invariant measure $\boldsymbol{x}$ and vector $\boldsymbol{y}$, respectively, of $D(\eta)$. Notice that upon substitution of $\eta = \lambda/\mu_1$ into (3), we have

$$D(\eta) = \rho_1\theta \begin{pmatrix} \mu_1 + \mu_2 & \lambda & & & & \\ \mu_2 & \mu_1 & \lambda & & & \\ & \mu_2 & \mu_1 & \lambda & & \\ & & \ddots & \ddots & \ddots & \\ & & & \ddots & \ddots & \ddots \end{pmatrix}. \tag{24}$$

This is a product of ρ_1 and a positive recurrent stochastic matrix. Therefore, $\boldsymbol{x}$ is simply equal to the stationary probability vector of this positive recurrent matrix, which is given by $\boldsymbol{x} = (1 - \rho_2)(1, \rho_2, \rho_2^2, \ldots)$, and $\boldsymbol{y}$ is simply $\boldsymbol{e}$, the column vector of ones. Some simple calculations will then justify Condition 3b. Finally, Condition 4 is satisfied since $\boldsymbol{\pi_1 y} = \boldsymbol{\pi_1 e} < \infty$. Therefore, we have $\boldsymbol{\pi_m} \sim c\rho_1^m \boldsymbol{x}$.

7 Conclusions

We have provided a set of verifiable new conditions for the existence of exact geometric light tails for a large class of Markov chains with infinitely many levels and phases. Four models are discussed as examples to illustrate how to use our conditions to obtain the exact geometric tails.

There are some opportunities for further research along this direction. For example, the entries of the infinite blocks A, B and etc. may be finite matrices, and/or the structure of the infinite blocks is more general, say M/G/1 type, GI/M/1 type or $GI/G/1$ type. Substantial research is needed to address the light tail issue in these kinds of systems. We have only considered processes in which the transition matrix has a QBD structure. Another way to generalize our method would be to consider more general transition matrix structures, say those of the $M/G/1$ type, the $GI/M/1$ type or the $GI/G/1$ type. Finally, we may also consider the case when $D(\eta)$ is reducible or when P has a level dependent structure.

Acknowledgments

The authors acknowledge the supports of the Natural Sciences and Engineering Research Council of Canada (NSERC), Carleton University, Hong Kong RGC (HKUST6198/01E) and Hong Kong University of Science and Technology. The authors thank the referee for providing valuable comments and suggestions to improve the quality of this paper. The quality of the paper has been further improved based on a number of editorial corrections according to the list made by a student of the Editor-in-Chief.

References

[1] Adan, I. *A compensation Approach for Queueing Problems*; Ph.D. Thesis; Eindhoven University of Technology, 1991.

[2] Adan, I.J.B.F.; Boxma, O.J.; Resing, J.A.C. Queueing Models with Multiple Waiting Lines. Queueing Systems **2001** *37*, 65–98.

[3] Flatto, L.; Hahn, S. Two Parallel Queues Created by Arrivals with Two Demands. SIAM J. Appl. Math. **1984**, *44* (5), 1041–1053.

[4] Foley, R.D.; McDonald, D.R. Join the Shortest Queue: Stability and Exact Asymptotics. Annals of Applied Probability **2001** *11* (3), 569–607.

[5] Haque, L.; Liu, L.; Zhao, Y.Q. Approximating Starvation Probability of a Two-Stage Inventory-Queue Model. **2003**, submitted for publication.

[6] Kemeny, J.G.; Snell, J.L.; Knapp, A.W. *Denumerable Markov Chains*; Springer-Verlag: New York, 1976.

[7] Kijima, M. Quasi-Stationary Distributions of Single-Server Phase Type Queues. Mathematics of Operations Research **1993** *18* (2), 423–437.

[8] Li, Quan-Lin; Zhao, Yiqiang Q. A Constructive Method for Finding β Invariant Measure for Transition Matrices of $M/G/1$ Type. *In Matrix-Analytic Methods: Theory and Applications*; Latouche, G., Taylor, P., Eds.; World Scientific, 2002; 237–263.

[9] Li, Quan-Lin; Zhao, Yiqiang Q. β-Invariant Measure for Transition Matrices of $GI/M/1$ Type. Stochastic Models **2002** *19* (2), 201–233.

[10] Meyn, S.P.; Tweedie, R.C. *Markov Chains and Stochastic Stability*; Springer: New York, 1993.

[11] Miller, Douglas R. Computation of Steady-State Probabilites for $M/M/1$ Priority Queues. *Operations Research* **1981**, **29**, (5), 945-958.

[12] Miyazawa, M. The Markov Renewal Approach for the Stationary Distributions in the $M/G/1$ Type Queue with Countable Many Background States. **2002**, submitted for publication **need check QUESTA for update**.

[13] Miyazawa, M.; Zhao, Y.Q. The Stationary Tail Asymptotics in the $GI/G/1$ Type Queue with Countably Many Background States. **2003**, submitted for publication.

[14] Neuts, Marcel F. *Matrix-Geometric Solutions in Stochastic Models*; The Johns Hopkins University Press: Baltimore, Maryland, 1981.

[15] Ramaswami, V.; Taylor, P.G. Some Properties of the Rate Operators in level dependent Quasi Birth and Death Processes with a Countable Number of Phases. Stochastic Models **1996** *12* (1), 143–164.

[16] Seneta, E. *Non-Negative Matrices and Markov Chains*; Springer-Verlag: New York, 1981.

[17] Takahashi, Y; Fujuimoto, K; Makimoto, N. Geometric Decay of the Steady-State Probabilities in a Quasi-Birth-and-Death Process with a Countable Number of Phases. Stochastic Models **2001** *17* (1), 1–24.

[18] Zhao, Yiqiang, Q. Censoring Technique in Studying Block-Structured Markov Chains. *In Advances in Algorithmic Methods for Stochastic Models*; Latouche, G., Taylor, P., Eds.; Notable Publications, 2000; 417–433.

[19] Zhao, Y.Q.; Li, W.; Braun, W.J. Censoring, Factorizations and Spectral Analysis for Transition Matrices with Block-Repeating Entries. Methodology and Computing in Applied Probability **2003** *5*, 35–58.