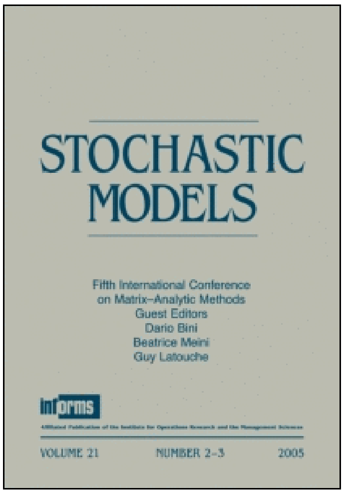

Stochastic Models
Publication details, including instructions for authors and subscription information:
http://www.informaworld.com/smpp/title~content=t713597301

Light-Tailed Behavior in QBD Processes with Countably Many Phases

Qi-Ming He [a]; Hui Li [b]; Yiqiang Q. Zhao [c]
[a] Department of Industrial Engineering, Dalhousie University, Halifax, Nova Scotia, Canada [b] Department of Mathematics, Mount Saint Vincent University, Halifax, Nova Scotia, Canada [c] School of Mathematics and Statistics, Carleton University, Ottawa, Ontario, Canada

Online Publication Date: 01 January 2009

To cite this Article He, Qi-Ming, Li, Hui and Zhao, Yiqiang Q.(2009)'Light-Tailed Behavior in QBD Processes with Countably Many Phases',Stochastic Models,25:1,50 — 75

To link to this Article: DOI: 10.1080/15326340802640974

URL: http://dx.doi.org/10.1080/15326340802640974

Stochastic Models, 25:50–75, 2009

ISSN: 1532-6349 print/1532-4214 online
DOI: 10.1080/15326340802640974

Taylor & Francis
Taylor & Francis Group

LIGHT-TAILED BEHAVIOR IN QBD PROCESSES WITH COUNTABLY MANY PHASES

Qi-Ming He[1], Hui Li[2], and Yiqiang Q. Zhao[3]

[1] *Department of Industrial Engineering, Dalhousie University, Halifax, Nova Scotia, Canada*
[2] *Department of Mathematics, Mount Saint Vincent University, Halifax, Nova Scotia, Canada*
[3] *School of Mathematics and Statistics, Carleton University, Ottawa, Ontario, Canada*

□ *Generally speaking, analysis of tail asymptotics in two-dimensional queueing systems is very challenging. Earlier work based on complex analysis led to determinations of exact forms of tail asymptotics. Ideas of large deviations, a powerful tool for characterizing light-tailed decay rates or analysis of rough tail asymptotics, have been utilized recently to develop probabilistic methods to do exact tail asymptotic analysis. Another promising approach to do tail asymptotics analysis, both exact and rough, is the matrix-analytic method. In this article, we combine the matrix-analytic method with techniques from probability and analysis to characterize tail asymptotics in a QBD process with infinitely many phases. The main results include conditions on:* (1) *exact geometric decay;* (2) *light-tailed behavior without an exact geometric decay, which in general is not the focus of the large deviations method; and* (3) *upper and lower bounds for stationary probabilities. We apply the main results to two two-dimensional queueing systems, including a polling system and a gated random-order server queue to characterize their light-tailed behavior of the queue length processes.*

Keywords Bounds; Decay rate; Gated random order service queue; Geometric decay; Infinitely many phases; Light tail; Matrix-analytic method; Polling system; QBD process.

Mathematics Subject Classification Primary 60K25; Secondary 60J99.

Received October 3, 2007; Accepted August 21, 2008
Address correspondence to Qi-Ming He, Department of Industrial Engineering, Dalhousie University, Halifax, Nova Scotia B3J 2X4, Canada; E-mail: Qi-Ming.he@dal.ca

1. INTRODUCTION

Analysis of queueing systems with a multi-dimensional state space is very challenging, including characterization of tail asymptotics. For two-dimensional systems, tail asymptotics of the stationary probability distribution are often characterized through methods based on complex analysis, large deviations techniques, probabilistic arguments, or matrix-analytic approaches.

In terms of the uniformization method, analytical continuation and analysis of singularities, exact tail aymptotics in a few two-dimensional queueing systems have been reported, such as Flatto and McKean[1,2] for the join-the-shortest-queue model and the parallel queues fed by arrivals with two types of demands, and recently Leeuwaarden[12] for the tandem queue with coupled processors. Tauberian theorems, which have been applied to queueing models such as the generalized processor sharing (GPS) networks for exact tail asymptotics, including both light-tailed and heavy-tailed behavior, for example, in van Uitert[31] and references therein, constitute another example in this category.

The large deviations technique is a powerful tool for studying rare events. Additional work has been done recently to address nontrivial issues in applying the general principle to queueing systems, such as Ridder and Shwartz[24] on the join-the-shortest-queue model. The classical large deviations principle does not intend to characterize the exact form of tail asymptotics. McDonald[16], and Foley and McDonald[3–5] modified the large deviations theory to extend the study to the exact tail asymptotics for two-dimensional queueing systems including a generalized join-the-shortest-queue model, and a modified Jackson network. In their studies, not only has the exact geometric decay been characterized, but also some forms of exact light-tailed decays have been recognized. Another probabilistic approach is the probability sample path argument, which was used to characterize heavy-tailed phenomena, for example, in Jelenković and Lazar[8], but it requires intuitive knowledge about the tail behavior.

Another promising method, the matrix-analytic method, is a relatively new one, which has been proven an efficient one for exact tail asymptotics for Markov chains of $GI/G/1$ type with finitely many phases, and also for exact geometric tail asymptotics for this type of Markov chain with infinitely many phases. For models with infinitely many phases, studies include Takahashi et al.[28], in which a sufficient condition was provided for the quasi-birth-and-death (QBD) process; this condition was reformulated in Haque[6], and Haque et al.[7] in a more favorable form for applications; a sufficient condition was provided in Miyazawa[18] for Markov chains of $M/G/1$ type; a slightly improved sufficient condition for Markov chains of $GI/G/1$ type was reported in Miyazawa and Zhao[19]. All of the above conditions were obtained under the positivity assumption on the generating function of the underlying renewal blocks. Without this assumption, Kroese et al.[11] provided a sufficient condition for exact geometric decay, which has been improved in Motyer and Taylor[22], and Li et al.[14]. Based on the above studies, Liu et al.[15] obtained two sufficient conditions for the level-expanding QBD model. Applying these conditions, researchers are able to characterize exact geometric tail asymptotics in several two-dimensional queueing systems, such as join-the-shortest-queue models (Takahashi et al.[28], Haque[6], Sakuma et al.[26], and Li et al.[14]), priority systems (Haque[6], Miyazawa and Zhao[19], and Xue and Alfa[32]), the parallel queues with two types of demands (Haque[6]), tandem queues (Haque [6], Kroese et al.[11], and Tang and Zhao[29]), a retrial queue (Li and Zhao[13]), among possible others.

In this article, we consider an irreducible, positive recurrent, and aperiodic QBD process, in discrete-time, with infinitely many phase (background) states. More specifically, we assume that the state space is given by

$$S = \{(0, j) : j \in S_0\} \cup \{(i, j) : i = 1, 2, \ldots, j = 0, 1, 2, \ldots\},$$

where S_0 is a countable set. For a state $(i, j) \in S$, we refer i and j as to the level and phase (background) variables, respectively. Partition the transition matrix of the QBD process according to the level as

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

Denote the unique positive stationary vector of P, partitioned according to the level too, by $\boldsymbol{\pi} = (\boldsymbol{\pi}_0, \boldsymbol{\pi}_1, \boldsymbol{\pi}_2, \ldots) = (\pi_{i,j})$, where $\boldsymbol{\pi}_0 = (\pi_{0,j})_{j \in S_0}$ and for $i = 1, 2, \ldots$, $\boldsymbol{\pi}_i = (\pi_{i,0}, \pi_{i,1}, \pi_{i,2} \ldots)$. Then, according to Tweedie [30], we have the following (operator-) matrix-geometric form solution for $\boldsymbol{\pi}$:

$$\boldsymbol{\pi}_{i+1} = \boldsymbol{\pi}_1 R^i, \quad \text{for } i \geq 1, \tag{2}$$

and $\boldsymbol{\pi}_0$ and $\boldsymbol{\pi}_1$ are the unique solution to

$$\boldsymbol{\pi}_0 = \boldsymbol{\pi}_0 B_0 + \boldsymbol{\pi}_1 C_0,$$

$$\boldsymbol{\pi}_1 = \boldsymbol{\pi}_0 A_0 + \boldsymbol{\pi}_1 (B + RC),$$

$$\boldsymbol{\pi}_0 \boldsymbol{e} + \boldsymbol{\pi}_1 \left(\sum_{i=0}^{\infty} R^i \right) \boldsymbol{e} = 1,$$

where $\boldsymbol{e}$ is the column vector of ones with a proper size, and R is the minimal nonnegative solution to $R = A + RB + R^2 C$, often referred to as the rate matrix.

Our focus of this article is to characterize the tail behavior in the stationary probabilities $\pi_{i,j}$. To be specific, light-tailed properties will be studied here, including the decay rate (large deviations type of results) along the level direction and for the marginal distribution, exact geometric decay, bounds for tails of the stationary probabilities, and applications to two interesting queueing models, a polling system, and a gated random-order server queue. The method employed here is based on the matrix-analytic method, combined with techniques used in probability and analysis. This research was also motivated by the challenge of how to use the matrix-analytic method to characterize exact but nongeometric tail asymptotics. As the first step of this study, we are able to identify some conditions, under which a light-tailed, but not exact geometric, decay reveals. In a very recent symposium report by Miyazawa[20], determination of the decay rate was addressed for a simpler class of QBD processes based on connections between the level process and the background process. An extended version of Ref.[20] is given in Miyazawa[21].

The rest of the article is organized into four sections. The main results obtained in this article are stated in the next section, which are proved in Section 3. Details of applications of the main results to two queueing models are provided in Section 4.

2. MAIN RESULTS

To state the main results reported in this article, we adopt the following definition.

Definition 2.1. We say that $\pi_{i,j}$ has a light tail with the decay rate $0 \le \eta < 1$ along the level direction if for each fixed j,

$$\overline{\lim_{n\to\infty}} \frac{\log \pi_{n,j}}{n} = \log \eta. \tag{3}$$

We say that $\pi_{i,j}$ decays exactly geometrically along the level direction if for each fixed j, there exist an η independent of j with $0 < \eta < 1$, and $x_j > 0$ such that

$$\lim_{n\to\infty} \frac{\pi_{n,j}}{\eta^n} = cx_j \tag{4}$$

for a constant c with $0 < c < \infty$. In an obvious way, we can define the light-tailed decay rate and exact geometric decay rate for a marginal distribution of $\pi_{i,j}$.

Remark 2.1. (i) When $\eta = 0$ in equation (3), the decay of the probabilities to zero is faster than any exponential one; (ii) the η in the definition for exact geometric decay along the level direction is also the light-tailed decay rate.

To state our main results, we also need the following concepts.

Definition 2.2. Let $T = (t_{i,j})$ be a countable nonnegative square matrix and let $T^n = (t_{i,j}^{(n)})$. For a fixed pair of i and j, the radius of convergence or the convergence parameter $\alpha_{i,j}^{(T)}$ (or simply $\alpha_{i,j}$) of $t_{i,j}$ is defined by

$$\alpha_{i,j}^{(T)} = \sup_{z\geq 0}\left\{z : \sum_{n=0}^{\infty} t_{i,j}^{(n)} z^n < \infty\right\}. \tag{5}$$

The radius of convergence or the convergence parameter α_T (or simply α) and the convergence norm ζ_T (or simply ζ) of T are defined by $\alpha_T = \inf_{i,j}\{\alpha_{i,j}^{(T)}\}$ and $\zeta_T = 1/\alpha_T$, respectively.

Remark 2.2. (i) If T is irreducible, then $\alpha_{i,j}$ is independent of i and j; (ii) if $\alpha = 0$, then the convergence norm ζ is defined as ∞.

Definition 2.3. Let $T = (t_{i,j})$ be a countable nonnegative square matrix. For $\beta > 0$, a nonnegative nonzero row vector $\boldsymbol{x}$ is called a β-subinvariant (β-superinvariant) measure if $\beta\boldsymbol{x}T \leq (\geq)\boldsymbol{x}$ and a β-invariant measure if $\beta\boldsymbol{x}T = \boldsymbol{x}$; a nonnegative nonzero column vector $\boldsymbol{y}$ is called a β-subinvariant (β-superinvariant) vector (or function) if $\beta T\boldsymbol{y} \leq (\geq)\boldsymbol{y}$ and a β-invariant vector (or function) if $\beta T\boldsymbol{y} = y$. T is called α-positive if there exist α-invariant measure $\boldsymbol{x}$ and vector $\boldsymbol{y}$ such that $\boldsymbol{x}\boldsymbol{y} < \infty$.

For convenience, we define

$$D(\eta) = A + \eta B + \eta^2 C. \tag{6}$$

It is important to mention the following relationship between $D(\eta)$ and the matrix R, in terms of which conditions imposed on R stated in the main results can be easily replaced by the counterpart conditions on $D(\eta)$. This has an obvious advantage since $D(\eta)$ is given, but R is usually unknown.

Lemma 2.1. *For a nonnegative, nonzero square matrix T, if either of the following conditions is satisfied:*

(1) *T has only one irreducible class,*
(2) *For each i, $\alpha_T = \inf_j\{\alpha_{i,j}^{(T)}\}$,*

then the convergence norm ζ_T is given by:

$$\zeta_T = \inf\{\eta > 0 : \textit{there exists a nonnegative nonzero } \boldsymbol{x} \textit{ such that } \boldsymbol{x}T \le \eta\boldsymbol{x}\}. \quad (7)$$

Remark 2.3. This lemma is an extension of the result for the irreducible case, which is needed to deal with some applications, including those in this article.

For $D(\eta)$ with $\eta > 0$, define

$$\eta_{\min} = \inf\{\eta > 0 : \text{there exists a nonnegative nonzero } \boldsymbol{x} \text{ such that } \boldsymbol{x}D(\eta) \le \eta\boldsymbol{x}\}. \quad (8)$$

Lemma 2.2. *Let R be the rate matrix of the QBD process P defined in* (1). *If either* (1) *or* (2) *in Lemma* 2.1 *is satisfied by R then we have $\zeta_R = \eta_{\min}$.*

Lemma 2.3. *Assume $0 < \eta < 1$ and $\boldsymbol{x} \ge 0$ is a nonnegative nonzero row vector. Matrix R satisfies* (1) *or* (2) *given in Lemma* 2.1. *Then,*

(a) *$\boldsymbol{x}$ is a $1/\eta$-invariant vector of R if and only if $\boldsymbol{x}$ is a $1/\eta$-invariant vector of $D(\eta)$;*
(b) *R has a $1/\eta$-subinvariant measure if and only if $D(\eta)$ has a $1/\eta$-subinvariant measure;*
(c) *R has a $1/\eta$-supinvariant measure if and only if $D(\eta)$ has a $1/\eta$-supinvariant measure.*

Remark 2.4. Based on the aforementioned lemmas, matrix R can be replaced by matrix $D(\eta)$ and ζ_R by $\eta_{\min}$, respectively, in every relevant property of the following main results if either (1) or (2) in Lemma 2.1 is satisfied.

The main results on tail asymptotics in this article include the following.

2.1. Conditions for (Exact) Geometric Decay

Our first theorem gives a necessary condition for exact geometric decay. This condition can be used to show light-tailed behavior without an exact geometric decay.

Theorem 2.1.1. *For the QBD process defined in* (1)*, if* $\pi_{i,j}$ *decays exactly geometrically along the level direction with decay rate* η*, then* $\boldsymbol{x} = (x_0, x_1, \ldots)$ *(defined in equation* (4)*) is a* $1/\eta$*-subinvariant measure of both* $D(\eta)$ *and* R*. The vector* $\boldsymbol{x}$ *is* $1/\eta$*-invariant if each column of* A*,* B*, and* C *contains only finitely many nonzero entries.*

Remark 2.1.1. In Takahashi et al.[28], a sufficient condition based on α-positivity was reported. As an extension of their study, in Li et al.[14], two sufficient conditions (Theorems 2.1.1 and 2.2.1) are provided for exact geometric decay, one for the α-positive case, and the other for non-α-positive case in terms of a positive invariant measure $\boldsymbol{x}$. Another sufficient condition without requiring α-positivity is obtained in Kroese et al.[11]. A sufficient condition in terms of a positive invariant vector $\boldsymbol{y}$ is also possible, which is provided below.

Theorem 2.1.2. *For the QBD process defined in* (1) *and for* $0 < \eta < 1$*, if*

$$\lim_{n\to\infty} \frac{R^n}{\eta^n} = 0, \tag{9}$$

and R *has a positive* $1/\eta$*-invariant vector* $\boldsymbol{y} = (y_0, y_1, y_2, \ldots)$*, i.e.,* $R\boldsymbol{y} = \eta\boldsymbol{y}$*, such that* $\lim_{n\to\infty} 1/y_n = c$ *for some constant* $0 \le c \le \infty$*, and* $\sum_{j=0}^{\infty} \pi_{1,j} y_j < \infty$*, then*

$$\lim_{n\to\infty} \frac{\boldsymbol{\pi}_n \boldsymbol{e}}{\eta^n} = \frac{c}{\eta} \sum_{j=0}^{\infty} \pi_{1,j} y_j.$$

In the case that $0 < c < \infty$*, the marginal distribution for the level variable decays exactly geometrically with the decay rate* η*.*

Remark 2.1.2. (i) For any $\eta > \zeta_R$, (9) is always true; (ii) If (9) holds for an irreducible R, then R is $1/\eta$-null or $1/\eta$-transient; (iii) in the case that $c = 0$, the condition $R\boldsymbol{y} = \eta\boldsymbol{y}$ can be replaced by $R\boldsymbol{y} \le \eta\boldsymbol{y}$.

Remark 2.1.3. Under the conditions of Theorem 2.1.2, it is easy to see that if the marginal distribution $\boldsymbol{\pi}_k \boldsymbol{e}$ has a light tail, then so does the joint distribution along the level direction; if ζ_R is the decay rate of the marginal distribution, then it is also the decay rate for the joint distribution along the level direction.

A relationship between the exact geometric decay along the level direction and in the marginal distribution for the level variable is described in the following theorem.

Theorem 2.1.3. *If* $\lim_{n\to\infty} \pi_{n,j}/\eta^n = cx_j$ *for* $0 \le c \le \infty$, *then* $\lim_{n\to\infty} \boldsymbol{\pi}_n \boldsymbol{e}/\eta^n = c\sum_{j=0}^{\infty} x_j$, *whenever* $\pi_{1,j}/x_j < M < \infty$ *uniformly and* $\sum_{j=0}^{\infty} x_j < \infty$. *Furthermore,* $\boldsymbol{\pi}_n \boldsymbol{e}$ *has an exact geometric decay if* $0 < c < \infty$.

2.2. Bounds for the Joint Distribution $\pi_{i,j}$ and for Marginal Distribution $\pi_n e$

Theorem 2.2.1. *For the QBD process defined in* (1) *and for* $0 < \eta < 1$,

(a) *If R has a positive $1/\eta$-subinvariant measure $\boldsymbol{x} = (x_0, x_1, x_2, \ldots)$, or $\boldsymbol{x}R \le \eta\boldsymbol{x}$, such that $\pi_{1,j}/x_j \le M$, then*

$$\frac{\pi_{n,j}}{\eta^n} \le \frac{M}{\eta} x_j, \quad \textit{for all } j = 0, 1, \ldots.$$

If in addition $\sum_i x_i < \infty$, then

$$\frac{\boldsymbol{\pi}_n \boldsymbol{e}}{\eta^n} \le U < \infty$$

for some constant U.

(b) *If R has a positive $1/\eta$-superinvariant measure $\boldsymbol{x} = (x_0, x_1, x_2, \ldots)$, or $\boldsymbol{x}R \ge \eta\boldsymbol{x}$ such that $m \le \pi_{1,j}/x_j$, then*

$$\frac{m}{\eta} x_j \le \frac{\pi_{n,j}}{\eta^n}, \quad \textit{for all } j = 0, 1, \ldots,$$

and

$$\frac{\boldsymbol{\pi}_n \boldsymbol{e}}{\eta^n} \ge L > 0$$

for some constant L.

(c) *If R has a positive $1/\eta$-invariant measure $\boldsymbol{x} = (x_0, x_1, x_2, \ldots)$, or $\boldsymbol{x}R = \eta\boldsymbol{x}$ such that $m \le \pi_{1,j}/x_j \le M$, then*

$$\frac{m}{\eta} x_j \le \frac{\pi_{n,j}}{\eta^n} \le \frac{M}{\eta} x_j, \quad \textit{for all } j = 0, 1, \ldots.$$

If in addition $\sum_i x_i < \infty$, then

$$0 < L \le \frac{\boldsymbol{\pi}_n \boldsymbol{e}}{\eta^n} \le U < \infty$$

for some constants L and U.

Bounds can be obtained for the marginal distribution $\boldsymbol{\pi}_n \boldsymbol{e}$ in terms of conditions on subvariant or superinvariant vectors.

Theorem 2.2.2. *For the QBD process defined in* (1) *and for* $0 < \eta < 1$,

(a) *If* R *has a positive* $1/\eta$*-subinvariant vector* $\boldsymbol{y} = (y_0, y_1, y_2, \dots)$, *or* $R\boldsymbol{y} \le \eta \boldsymbol{y}$ *such that* $\sum_i \pi_{1,i} y_i < \infty$ *and* $1/y_j \le M < \infty$ *uniformly, then*

$$\frac{\boldsymbol{\pi}_n \boldsymbol{e}}{\eta^n} \le U < \infty$$

for some constant U.

(b) *If* R *has a positive* $1/\eta$*-superinvariant vector* $\boldsymbol{y} = (y_0, y_1, y_2, \dots)$, *or* $R\boldsymbol{y} \ge \eta \boldsymbol{y}$ *such that* $1/y_j \ge m > 0$ *uniformly (which implies that* $\sum_i \pi_{1,i} y_i < \infty$*), then*

$$\frac{\boldsymbol{\pi}_n \boldsymbol{e}}{\eta^n} \ge L > 0$$

for some constant L.

(c) *If* R *has a positive* $1/\eta$*-invariant vector* $\boldsymbol{y} = (y_0, y_1, y_2, \dots)$, *or* $R\boldsymbol{y} = \eta \boldsymbol{y}$ *such that* $\sum_i \pi_{1,i} y_i < \infty$ *and* $0 < m \le 1/y_j \le M < \infty$ *uniformly, then*

$$0 < L \le \frac{\boldsymbol{\pi}_n \boldsymbol{e}}{\eta^n} \le U < \infty$$

for some constants L *and* U.

Remark 2.2.1. It is worthwhile to point out that the convergence norm ζ_R for R always provides a lower bound for the decay rate (referring to Lemma 3.2).

2.3. Light-Tailed Behavior Without an Exact Geometric Decay

The results given here came from a preliminary study intended to characterize exact light-tailed asymptotics other than an exact geometric tail.

Theorem 2.3.1. *For the QBD process defined in* (1), *the joint distribution* $\pi_{n,j}$ *has a light tail along the level direction with the decay rate* η, *but without an exact geometric decay, if either of the following two conditions is satisfied:*

(1) $\lim_{n\to\infty} \pi_{n,j}/\theta^n = 0$ *for all* $\theta \ge \eta$ *and* $\overline{\lim}_{n\to\infty} \pi_{n,j}/\theta^n = \infty$ *for all* $\theta < \eta$;

(2) $\lim_{n\to\infty} \pi_{n,j}/\theta^n = 0$ *for all* $\theta > \eta$ *and* $\overline{\lim}_{n\to\infty} \pi_{n,j}/\theta^n = \infty$ *for all* $\theta \le \eta$.

Lemma 2.3.1. *If*

$$\lim_{n\to\infty} \frac{\pi_{n,j}}{\zeta_R^n} = 0,$$

and for each j, there exists an i_j such that $\alpha_{i_j,j} = 1/\zeta_R$, where $\alpha_{i,j}$ is defined by (5), then $\boldsymbol{\pi}$ does not have an exact geometric decay along the level direction. Instead, it has a light tail with decay rate ζ_R, or

$$\overline{\lim_{n\to\infty}} \frac{\log \pi_{n,j}}{n} = \log \zeta_R.$$

Theorem 2.3.2. *For the QBD process defined in (1) and for $0 < \eta < 1$, assume that*

$$\lim_{n\to\infty} R^n/\eta^n = 0. \tag{10}$$

If R has a $1/\eta$-subinvariant measure $\boldsymbol{x} = (x_0, x_1, x_2, \dots)$ and

$$\lim_{j\to\infty} \frac{\pi_{1,j}}{x_j} = 0, \tag{11}$$

then

$$\lim_{n\to\infty} \pi_{n,j}/\eta^n = 0, \tag{12}$$

which implies that η^n is an asymptotic bound for $\pi_{n,j}$. If in addition, the above-mentioned conditions hold for $\eta = \zeta_R$, then $\boldsymbol{\pi}$ does not have an exact geometric decay along the level direction. Instead, it has a light tail with decay rate ζ_R.

Remark 2.3.1. When R is irreducible, condition $\lim_{n\to\infty} R^n/\eta^n = 0$ implies that R is either $1/\eta$-null or $1/\eta$-transient.

We can similarly obtain the following characterization for the marginal distribution.

Theorem 2.3.3. *For the QBD process defined in (1), assume*

$$\lim_{n\to\infty} R^n/\zeta_R^n = 0. \tag{13}$$

If either of the following two sets of conditions are satisfied:

Set I of conditions.

(1) *R has a positive $1/\zeta_R$-subinvariant measure $\boldsymbol{x}$ or $\boldsymbol{x}R \le \zeta_R \boldsymbol{x}$;*
(2) $\lim_{j\to\infty} \pi_{1,j}/x_j = 0$ *and* $\sum_{j=0}^{\infty} x_j < \infty$.

Set II of conditions.

(1) *R has a positive $1/\zeta_R$-subinvariant vector $\boldsymbol{y}$ or $R\boldsymbol{y} \le \zeta_R \boldsymbol{y}$;*
(2) $\lim_{j\to\infty} 1/y_j = 0$ *and* $\sum_{j=0}^{\infty} \pi_{1,j} y_j < \infty$.

then the marginal distribution $\boldsymbol{\pi}_n \boldsymbol{e}$ *has a light tail with decay rate* ζ_R*, or*

$$\lim_{n\to\infty} \frac{\boldsymbol{\pi}_n \boldsymbol{e}}{\zeta_R^n} = 0 \quad \textit{and} \quad \overline{\lim_{n\to\infty}} \frac{\log \boldsymbol{\pi}_n \boldsymbol{e}}{n} = \log \zeta_R, \tag{14}$$

but $\boldsymbol{\pi}_n \boldsymbol{e}$ *does not have an exact geometric decay.*

2.4. Applications to Queueing Models

We can apply the above-mentioned results to many queueing models, for example, to the join-the-shortest-queue model, the tandem queue, the priority queueing system, the queueing system with two types of demand, etc. to obtain various tail asymptotics results, which have been reported in the literature. Instead of doing so, we apply the results to the polling system and the gated random order service queue, respectively, to show a light-tailed behavior in these two models, determine the decay rate, and confirm the nongeometric decay for both cases, which have not been reported before.

3. PROOFS

In this section, we provide proofs to all the main results stated in Section 2.

Since basic properties on subinvariant and superinvariant measures (vectors) are important to our proofs, and in the literature these properties are stated only for irreducible matrices, we need to verify them for reducible matrices due to the fact that the R matrix is often reducible even when $D(\eta)$ is irreducible. Instead of providing detailed verification, we only state relevant results and offer necessary comments.

Proof of Lemma 2.1. Under Condition (1), the proof for the irreducible case (for example, Theorem 6.3 of Seneta[27]) is also valid. Under Condition (2), we prove the result by showing the following two facts: (a) for every $\eta > \zeta_T$, there exists a $1/\eta$-subinvariant measure $\boldsymbol{x}$ of T; and (b) if T has a $1/\eta$-subinvariant measure, then $\eta \geq \zeta_T$. Fact (a) is true from the well-known fact that such a $1/\eta$-subinvariant measure exists for any nonnegative nonzero square matrix T. We show fact (b) by contradiction. Suppose that there were a $1/\eta$-subinvariant measure of T with $\eta < \zeta_T$. Let $\eta < \eta^* < \zeta_T$ and $0 < \epsilon = 1/\eta^* - 1/\eta$. By the definition of ζ_T and Condition (2), there exists a pair (i, j) such that $\alpha_{i,j}^{(T)} < \epsilon + \alpha_T$. Therefore, we would have

$$\infty = \sum_{n=0}^{\infty} \frac{x_i t_{i,j}^{(n)}}{(\eta^*)^n} \leq \sum_{n=0}^{\infty} \frac{\sum_{k=0}^{\infty} x_k t_{k,j}^{(n)}}{(\eta^*)^n} \leq \sum_{n=0}^{\infty} \frac{x_j \eta^n}{(\eta^*)^n} < \infty,$$

which is a contradiction. □

Proof of Lemma 2.2. First we show $\zeta_R \leq \eta_{\min}$. It is enough to show that for each $0 < \eta < 1$, if $D(\eta)$ has a $1/\eta$-subinvariant measure, then so does R. Consider the RG-factorization (for example, referring to Theorem 11 in Ref.[33]). In our case, we specifically have

$$\begin{aligned} \eta I - D(\eta) &= [(\eta I - R)(I - \Phi)](I - \eta G) \\ &= (\eta I - R)[(I - \Phi)(I - \eta G)] \\ &= (\eta I - R)(I - \Phi)(I - \eta G), \end{aligned} \tag{15}$$

where both $I - \Phi$ and $I - \eta G$ for $0 < \eta < 1$ are invertible for $0 < \eta < 1$. The above factorization is also associative with any nonnegative vector. Therefore, if $\boldsymbol{x}$ is a $1/\eta$-subinvariant measure of $D(\eta)$, then

$$\boldsymbol{x}[\eta I - D(\eta)] = \boldsymbol{x}(\eta I - R)(I - \Phi)(I - \eta G) \geq 0.$$

Multiplying from the right on both sides by the nonnegative inverses $\sum_k \Phi^k$ and $\sum_k (\eta G)^k$ leads to $x(\eta I - R) \geq 0$.

Next, we show that for $0 < \eta < 1$, if $\sum_k (R/\eta)^k < \infty$, or $(\eta I - R)$ is invertible, then $\sum_k (D(\eta)/\eta)^k < \infty$. Once again, consider the RG-factorization in (15). Since all states in the process corresponding to the substochastic matrix Φ are transient (for example, in Zhao[33]), it follows from Proposition 5.3 in Kemeny et al.[9] that $\sum_k \Phi^k < \infty$, or $I - \Phi$ is invertible. We also know that for $0 < \eta < 1$, $\sum_k (\eta G)^k < \infty$, or $I - \eta G$ is invertible (for example, in Li et al.[14]). Therefore, the minimum nonnegative inverse of $(\eta I - R)(I - \Phi)(I - \eta G)$ exists and is equal to $[\sum_k (\eta G)^k][\sum_k \Phi^k][\sum_k (R/\eta)^k]$ due to the fact that all the three above mentioned inverses are minimum nonnegative. This proves that the minimum nonnegative inverse of $\eta I - D(\eta)$ exists, or $\sum_k (D(\eta)/\eta)^k < \infty$. □

Remark 3.1. The idea in this proof was motivated by the comments by Miyazawa, who also proved this lemma in Ref.[20].

We can further prove the following relationship between $\boldsymbol{x}R$ and $\boldsymbol{x}D(\eta)$.

Lemma 3.1. *If $\boldsymbol{x}$ is a $1/\eta$-subinvariant measure of $D(\eta)$, then $\boldsymbol{x}R \leq \boldsymbol{x}D(\eta) \leq \eta\boldsymbol{x}$. Furthermore, if $\boldsymbol{x} > 0$, then $\eta \geq \zeta_R$, and therefore $\eta_{\min} \geq \zeta_R$.*

Proof of Lemma 3.1. By the same argument as used by Neuts[23], we can show that R is the limit of the matrix sequence $\{R[k]; k \geq 0\}$, where $R[0] = 0$ and $R[k+1] = A + R[k] + R^2[k]C$. If $\boldsymbol{x}D(\eta) \leq \eta\boldsymbol{x}$, then

$$\boldsymbol{x}R[k+1] = \boldsymbol{x}[A + R[k] + R^2[k]C] \leq \boldsymbol{x}D(\eta) \leq \eta x.$$

Therefore, $\boldsymbol{x}R[k] \leq \boldsymbol{x}D(\eta) \leq \eta\boldsymbol{x}$ for all k, which implies $\boldsymbol{x}R \leq \boldsymbol{x}D(\eta) \leq \eta\boldsymbol{x}$. The inequality $\eta \geq \zeta_R$ directly follows from Theorem 6.3 in Seneta[27] if R is irreducible. The proof can be modified to nonirreducible case. $\square$

Proof of Lemma 2.3. (a) It is clear from the factorization (15); (b) follows from Lemma 2.2, while (c) can be similarly proved. $\square$

Proof of Theorem 2.1.1. It follows from $\boldsymbol{\pi}P = \boldsymbol{\pi}$ that $\boldsymbol{\pi}_{n-1}A + \boldsymbol{\pi}_n B + \boldsymbol{\pi}_{n+1}C = \boldsymbol{\pi}_n$ for $n \geq 2$. Dividing the both sides of the equation by η^n yields

$$\frac{1}{\eta}\frac{\boldsymbol{\pi}_{n-1}}{\eta^{n-1}}A + \frac{\boldsymbol{\pi}_n}{\eta^n}B + \eta\frac{\boldsymbol{\pi}_{n+1}}{\eta^{n+1}}C = \frac{\boldsymbol{\pi}_n}{\eta^n}. \tag{16}$$

If each column of A, B, and C contains only finitely many nonzero entries, then taking the limit (or the upper limit) on the both sides of (16) leads to the conclusion that $\boldsymbol{x}$ is an $1/\eta$-invariant measure of $D(\eta)$. In general, by Fatou's lemma (16) leads to $\boldsymbol{x}D(\eta) \leq \boldsymbol{x}$. $\square$

Proof of Theorem 2.1.2. Based on the matrix-geometric form solution given in (2), we have

$$\frac{\pi_{n+1,j}}{\eta^{n+1}} = \frac{1}{\eta}\sum_{i=0}^{\infty}\pi_{1,i}\frac{r_{i,j}^{(n)}}{\eta^n}, \tag{17}$$

where $r_{i,j}^{(n)}$ are entries of R^n. Since $\boldsymbol{y}$ is the $1/\eta$-invariant vector of R,

$$\sum_{j=0}^{\infty}\frac{r_{i,j}^{(n)}y_j}{y_i\eta^n} = 1, \quad \text{for all } i \geq 0. \tag{18}$$

Let

$$a_{i,n} = \sum_{j=0}^{\infty}\frac{1}{y_j}\frac{r_{i,j}^{(n)}y_j}{y_i\eta^n}.$$

It follows from (18) and $\lim_{j\to\infty} 1/y_j = c$ that $a_{i,n}$ is uniformly bounded in both i and n. Notice also the assumption that $\lim_{n\to\infty} R^n/\eta^n = 0$, then for each i, we can apply the Toeplitz limit theorem (e.g., referring to Lemma 2.2.1 in Li et al.[14]) to $a_{i,n}$ to have $\lim_{n\to\infty} a_{i,n} = c$. Now, by using the dominated convergence theorem, we obtain

$$\lim_{n\to\infty}\frac{\boldsymbol{\pi}_{n+1}\boldsymbol{e}}{\eta^{n+1}} = \frac{1}{\eta}\sum_{i=0}^{\infty}\pi_{1,i}y_i \lim_{n\to\infty} a_{i,n} = \frac{c}{\eta}\sum_{i=0}^{\infty}\pi_{1,i}y_i.$$

In the case of $c = \infty$, we have $\lim_{n\to\infty} a_{i,n} = \infty$ by the Toeplitz limit theorem again, which completes the proof. □

Proof of Theorem 2.1.3. First assume that $0 \le c < \infty$. Since $\lim_{n\to\infty} \pi_{n,j}/\eta^n = cx_j$, $\boldsymbol{x}$ is an $1/\eta$-subinvariant measure of R by Theorem 2.1.1 and Lemma 2.3, which implies that

$$\sum_{i=0}^{\infty} \frac{x_i r_{i,j}^{(n)}}{x_j \eta^n} \le 1.$$

Let

$$b_{j,n} = \sum_{i=0}^{\infty} \frac{\pi_{1,i}}{x_i} \frac{x_i r_{i,j}^{(n)}}{x_j \eta^n}.$$

Then,

$$\frac{\boldsymbol{\pi}_{n+1}\boldsymbol{e}}{\eta^{n+1}} = \frac{1}{\eta} \sum_{j=0}^{\infty} x_j \sum_{i=0}^{\infty} \frac{\pi_{1,i}}{x_i} \frac{x_i r_{i,j}^{(n)}}{x_j \eta^n} = \frac{1}{\eta} \sum_{j=0}^{\infty} x_j b_{j,n}.$$

Clearly,

$$\lim_{n\to\infty} \frac{\pi_{n+1,j}}{\eta^{n+1}} = cx_j = \frac{x_j}{\eta} \lim_{n\to\infty} \sum_{i=0}^{\infty} \frac{\pi_{1,i}}{x_i} \frac{x_i r_{i,j}^{(n)}}{x_j \eta^n} = \frac{x_j}{\eta} \lim_{n\to\infty} b_{j,n},$$

or $\lim_{n\to\infty} b_{j,n} = c\eta$. From $\pi_{1,i}/x_i < M$, we also have $b_{j,n} \le M < \infty$. It follows from the dominated convergence theorem that

$$\lim_{n\to\infty} \frac{\boldsymbol{\pi}_{n+1}\boldsymbol{e}}{\eta^{n+1}} = \frac{1}{\eta} \sum_{j=0}^{\infty} x_j \lim_{n\to\infty} b_{j,n} = c \sum_{j=0}^{\infty} x_j.$$

If $c = \infty$, we have

$$\frac{\boldsymbol{\pi}_n \boldsymbol{e}}{\eta^{n+1}} > \frac{\pi_{n,1}}{\eta^n} \to \infty \quad \text{as } n \to \infty,$$

which completes the proof. □

Proof of Theorem 2.2.1. Based on the matrix-geometric form solution given in (2), we have (17), or

$$\frac{\pi_{n+1,j}}{\eta^{n+1}} = \frac{x_j}{\eta} \sum_{i=0}^{\infty} \frac{\pi_{1,i}}{x_i} \frac{x_i r_{i,j}^{(n)}}{x_j \eta^n}, \tag{19}$$

where $r_{i,j}^{(n)}$ are entries of R^n.

(a) It follows from that $\boldsymbol{x}R \leq \eta \boldsymbol{x}$ that $\boldsymbol{x}R^n \leq \eta^n \boldsymbol{x}$, which implies that

$$\sum_{i=0}^{\infty} \frac{x_i r_{i,j}^{(n)}}{x_j \eta^n} \leq 1. \tag{20}$$

Replacing the condition $\pi_{1,j}/x_j \leq M$ and (20) in the equality (19) leads to the conclusion.

Part (b) can be proved similarly, and part (c) can be obtained from (a) and (b). □

Proof of Theorem 2.2.2. (a) It follows from that $R\boldsymbol{y} \leq \eta \boldsymbol{y}$ that $R^n \boldsymbol{y} \leq \eta^n \boldsymbol{y}$, which implies that

$$\sum_{j=0}^{\infty} \frac{r_{i,j}^{(n)} y_j}{y_i \eta^n} \leq 1. \tag{21}$$

Replacing the conditions $\sum_i \pi_{1,i} y_i < \infty$ and $1/y_j \leq M < \infty$ uniformly, and (21) in the equality (17) leads to

$$\begin{aligned}\frac{\boldsymbol{\pi}_{n+1}\boldsymbol{e}}{\eta^{n+1}} &= \frac{1}{\eta}\sum_{i=0}^{\infty}\pi_{1,i}\left(\sum_{j=0}^{\infty}\frac{r_{i,j}^{(n)}}{\eta^n}\right)\\ &= \frac{1}{\eta}\sum_{i=0}^{\infty}\pi_{1,i}\left(\sum_{j=0}^{\infty}\frac{1}{y_j}\frac{r_{i,j}^{(n)}y_j}{y_i\eta^n}\right)y_i \leq \frac{M}{\eta}\sum_{i=0}^{\infty}\pi_{1,i}y_i < \infty.\end{aligned}$$

Part (b) can be proved similarly, and part (c) can be obtained from (a) and (b). □

Proof of Theorem 2.3.1. First, we assume that the conditions in (1) hold. Write

$$\frac{\pi_{n,j}}{\theta^n} = e^{\log\frac{\pi_{n,j}}{\theta^n}} = e^{n\left(\frac{\log \pi_{n,j}}{n} - \log\theta\right)}.$$

It follows from the assumption

$$\lim_{n\to\infty} \frac{\pi_{n,j}}{\theta^n} = 0, \quad \text{for all } \theta \geq \eta \tag{22}$$

that we must have

$$\overline{\lim_{n\to\infty}} \frac{\log \pi_{n,j}}{n} \stackrel{\text{def}}{=} \log \eta_j \leq \log \eta,$$

or $\eta_j \leq \eta$, since otherwise there would exist $\eta_j > \theta > \eta$ and a subsequence $\{\pi_{n_k,j}/\theta^{n_k}\}$ such that $\lim_{k\to\infty} \log \pi_{n_k,j}/n_k > \log \theta$, which implies

$$\lim_{k\to\infty} \frac{\pi_{n_k,j}}{\theta_k^n} = \lim_{k\to\infty} e^{\log \frac{\pi_{n_k,j}}{\theta^{n_k}}} = \lim_{k\to\infty} e^{n_k\left(\frac{\log \pi_{n_k,j}}{n_k} - \log\theta\right)} = \infty,$$

which contradicts to the assumption (22) for all $\theta \geq \eta$. Suppose that $\eta_j < \eta$, or

$$\overline{\lim_{n\to\infty}} \frac{\log \pi_{n,j}}{n} < \log \eta.$$

Then, there would exist a $\theta < \eta$ such that $\pi_{n,j}/\theta^n \leq 1$, when n is sufficiently large, which contradicts to the second condition in (a). Hence, $\eta_j = \eta$ must hold. That is, $\pi_{n,j}$ has a light tail along level direction with the decay rate η. Clearly, $\pi_{n,j}$ does not have an exact geometric decay since $\lim_{n\to\infty} \frac{\pi_{n,j}}{\eta^n} = 0$.

We can similarly prove the result under assumption (2). $\square$

Notice that ζ_R always provides a lower bound for the decay rate as stated below, which is also used in the proof to Theorem 2.3.2.

Lemma 3.2. *For the QBD process defined in* (1), *the convergence norm ζ_R serves as a lower bound for the decay rate of the joint probabilities $\boldsymbol{\pi}_n$ along the level direction.*

Proof of Lemma 3.2. It is a direct consequence of the definition of ζ_R and the matrix-geometric form solution for $\boldsymbol{\pi}_n$. $\square$

Proof of Lemma 2.3.1. By Theorem 2.3.1, it is sufficient to show that for $\eta < \zeta_R$, $\overline{\lim}_{n\to\infty} \pi_{n,j}/\eta^n = \infty$. By the definition of $\alpha_{i,j}^{(R)}$ and the assumption that there exists an i_j such that $\alpha_{i_j,j} = 1/\zeta_R$, for $\eta < \zeta_R$, $r_{i_j,j}^{(n)}/\eta^n$ is unbounded; or $\overline{\lim}_{n\to\infty} r_{i_j,j}^{(n)}/\eta^n = \infty$. The theorem follows now from

$$\overline{\lim_{n\to\infty}} \pi_{n+1,j}/\eta^{n+1} = \overline{\lim_{n\to\infty}} \frac{1}{\eta} \sum_{k=0}^{\infty} \pi_{1,k} \frac{r_{k,j}^{(n)}}{\eta^n} \geq \overline{\lim_{n\to\infty}} \frac{1}{\eta} \pi_{1,i_j} \frac{r_{i_j,j}^{(n)}}{\eta^n} = \infty. \qquad \square$$

Proof of Theorem 2.3.2. By $\boldsymbol{\pi}_{n+1} = \boldsymbol{\pi}_1 R^n$ and the condition in (11), we obtain

$$\frac{\pi_{n+1,j}}{\eta^{n+1}} = \sum_{i=0}^{\infty} \frac{\pi_{1,i} r_{i,j}^{(n)}}{\eta^{n+1}} = \sum_{i=0}^{N} \pi_{1,i} \frac{r_{i,j}^{(n)}}{\eta^{n+1}} + \sum_{i=N+1}^{\infty} \frac{\pi_{1,i}}{x_i} \frac{x_i r_{i,j}^{(n)}}{\eta^{n+1}},$$

where N is an integer such that for any $i > N$, $\pi_{1,i}/x_i \leq \epsilon$, where ϵ is a positive number.

By the condition in (10), for fixed j and $1 \leq i \leq N$, $r_{i,j}^{(n)}/\eta^n \leq \epsilon$ holds for all large enough n. Since $\boldsymbol{x}R \leq \eta\boldsymbol{x}$, we have $\boldsymbol{x}R^n \leq \eta^n\boldsymbol{x}$ for all n, which implies

$$\sum_{i=0}^{\infty} \frac{x_i r_{i,j}^{(n)}}{\eta^n} \leq x_j.$$

Therefore, for fixed j and all large enough n, we obtain

$$\frac{\pi_{n+1,j}}{\eta^{n+1}} \leq \frac{\epsilon}{\eta}\sum_{i=0}^{N}\pi_{1,i} + \epsilon\sum_{i=N+1}^{\infty}\frac{x_i r_{i,j}^{(n)}}{\eta^{n+1}} \leq \left(\frac{1+x_j}{\eta}\right)\epsilon.$$

Since ϵ can be chosen arbitrarily small for fixed j, equation (12) is proved.

When equation (12) is true for $\eta = \zeta_R$, the results directly follow from Lemma 2.3.1. □

Proof of Theorem 2.3.3. First, we assume that Set I of conditions holds. It follows from the matrix-geometric form solution that

$$\frac{\boldsymbol{\pi}_{n+1}\boldsymbol{e}}{\zeta_r^{n+1}} = \frac{\boldsymbol{\pi}_1 r^n e}{\zeta_r^{n+1}} = \frac{\sum_{j=0}^{\infty}\sum_{i=0}^{\infty}\pi_{1,i}r_{i,j}^{(n)}}{\zeta_r^{n+1}}.$$

For any given $\epsilon > 0$, we choose N such that $\pi_{1,i}/x_i < \epsilon$ for $i > N$, and K such that $\sum_{j=K+1}^{\infty} x_j < \epsilon$. Also, we choose L such that $r_{i,j}^{(n)}/\zeta_R^n < \epsilon/(K+1)$ for $0 \leq i \leq N$, $0 \leq j \leq K$, and $n > L$. Then, for $n > L$, we have

$$\begin{aligned}
\frac{\boldsymbol{\pi}_{n+1}\boldsymbol{e}}{\zeta_r^{n+1}} &= \sum_{j=0}^{K}\sum_{i=0}^{N}\pi_{1,i}\frac{r_{i,j}^{(n)}}{\zeta_r^{n+1}} + \sum_{j=0}^{\infty}\sum_{i=N+1}^{\infty}\frac{\pi_{1,i}}{x_i}\frac{x_i r_{i,j}^{(n)}}{\zeta_r^{n+1}} + \sum_{j=K+1}^{\infty}\sum_{i=0}^{N}\frac{\pi_{1,i}}{x_i}\frac{x_i r_{i,j}^{(n)}}{\zeta_r^{n+1}} \\
&\leq \left\{\epsilon\sum_{i=0}^{N}\pi_{1,i} + \epsilon\sum_{j=0}^{\infty}x_j + \max_{1\leq i\leq N}\left\{\frac{\pi_{1,i}}{x_i}\right\}\sum_{j=K+1}^{\infty}x_j\right\}\frac{1}{\zeta_R} \\
&\leq \left\{1 + \sum_{j=0}^{\infty}x_j + \max_{i\geq 1}\left\{\frac{\pi_{1,i}}{x_i}\right\}\right\}\frac{\epsilon}{\zeta_R}.
\end{aligned}$$

Since ϵ can be made arbitrarily small, the first result in (14) is proven. The second result is a direct consequence of the first result in (14) and Lemma 3.2.

Next, we assume that Set II of conditions holds. Similar to the above proof, the result can be obtained using the following expression:

$$\frac{\boldsymbol{\pi}_{n+1}\boldsymbol{e}}{\zeta_r^{n+1}} = \sum_{j=0}^{K}\sum_{i=0}^{N}\pi_{1,i}\frac{r_{i,j}^{(n)}}{\zeta_r^{n+1}} + \sum_{j=0}^{\infty}\sum_{i=N+1}^{\infty}\pi_{1,i}\frac{r_{i,j}^{(n)}y_j}{\zeta_r^{n+1}}\frac{1}{y_j}$$
$$+ \sum_{j=K+1}^{\infty}\sum_{i=0}^{N}\pi_{1,i}\frac{r_{i,j}^{(n)}y_j}{\zeta_r^{n+1}}\frac{1}{y_j}. \qquad \square$$

4. APPLICATIONS

The results obtained in previous sections can be applied to various queueing systems, including the following two examples.

4.1. Polling System

We consider an exhaustive polling system with one server switching between two waiting lines that contain type 1 and type 2 customers, respectively. At any time, if the server is serving a type k customer, $k = 1, 2$, it will keep serving type k customers, and switch over to serving the other type of customer only when the queue for type k customers becomes empty. The server goes into idle state only if there is no customer in the system; and it becomes activated immediately upon the arrival of a new arrival. We assume that there is no switchover time for the server between the two types of customers. The arrival processes for both types of customers are Poisson and the service times are exponential with rates λ_1, λ_2, μ_1, and μ_2, respectively. Let $q_1(t)$ and $q_2(t)$ be the queue length of type 1 and type 2 customers in the system at time t, and let $S(t)$ be the status of the server at time t, where

$$S(t) = \begin{cases} 0, & \text{when the server is idle,} \\ 1, & \text{when the server is serving a type 1 customer,} \\ 2, & \text{when the server is serving a type 2 customer.} \end{cases}$$

It is easy to see that $\{(q_1(t), (S(t), q_2(t))); t \geq 0\}$ is a Markov chain with a state space $\{(0, 0, 0)\} \cup \{(0, 2, j); j = 1, 2, \dots\} \cup \{(n, (1, j)); n = 1, 2, \dots,$ and $j = 0, 1, 2, \dots\} \cup \{(n, (2, j)); n = 1, 2, \dots,$ and $j = 1, 2, \dots\}$. It is well known that the Markov chain is ergodic if and only if $\lambda_1/\mu_1 + \lambda_2/\mu_2 < 1$, which is assumed to be true throughout this section. Let $\boldsymbol{\pi} = (\boldsymbol{\pi}_0, \boldsymbol{\pi}_1, \boldsymbol{\pi}_2, \dots)$ be the stationary probability vector of the Markov chain partitioned according to the level, where

$$\boldsymbol{\pi}_0 = (\pi_{0,0}, \pi_{0,2,1}, \pi_{0,2,2}, \pi_{0,2,3}, \dots),$$

$$\boldsymbol{\pi}_n = (\boldsymbol{\pi}_{n,1}, \boldsymbol{\pi}_{n,2}), \quad n \geq 1,$$

with

$$\boldsymbol{\pi}_{n,1} = (\pi_{n,1,0}, \pi_{n,1,1}, \pi_{n,1,2}, \dots),$$
$$\boldsymbol{\pi}_{n,2} = (\pi_{n,2,1}, \pi_{n,2,2}, \pi_{n,2,3}, \dots).$$

In the following, we first focus on the tail asymptotics of $\boldsymbol{\pi}_{n,2,j}$ for every fixed j, or the characterization of the tail asymptotics for the number of type-1 customers when the server is serving a type-2 customers. The analysis of tail asymptotics of $\boldsymbol{\pi}_{n,1,j}$ for every fixed j can be similarly done by interchange q_1 and q_2.

Let $P = I + Q/(\lambda_1 + \lambda_2 + \mu_1 + \mu_2)$, where I is the identity matrix. Then, P is a transition probability matrix having the same stationary probability vector $\boldsymbol{\pi}$ of Q. Without loss of generality, we assume that $\lambda_1 + \lambda_2 + \mu_1 + \mu_2 = 1$. Then, the detail of A, B, and C in the probability matrix P in (1) are given by

$$A = \lambda_1 I, \quad B = I + \begin{bmatrix} B_1 & 0 \\ B_3 & B_2 \end{bmatrix}, \quad C = \begin{bmatrix} \mu_1 I & \\ & 0 \end{bmatrix}$$

with

$$B_1 = \begin{bmatrix} \mu_2 & \lambda_2 & & \\ & \mu_2 & \lambda_2 & \\ & & \ddots & \ddots \\ & & & \ddots \end{bmatrix}, \quad B_2 = \begin{bmatrix} \mu_1 & \lambda_2 & & & \\ \mu_2 & \mu_1 & \lambda_2 & & \\ & \mu_2 & \mu_1 & \lambda_2 & \\ & & \ddots & \ddots & \ddots \end{bmatrix},$$

$$B_3 = \begin{bmatrix} \mu_2 & & & \\ & 0 & & \\ & & 0 & \\ & & & \ddots \end{bmatrix}.$$

The stationary distribution $\boldsymbol{\pi}$ has a matrix-geometric form solution $\boldsymbol{\pi}_n = \boldsymbol{\pi}_1 R^{n-1}$, $n = 1, 2, \dots$, where

$$R = \begin{bmatrix} R_1 & 0 \\ R_3 & R_2 \end{bmatrix}$$

according to the specific structures of A, B, and C. Notice that R_1, R_2, and R_3 are the minimal nonnegative solutions to

$$R_1 = \lambda_1 I + R_1 B_1 + R_1^2 \mu_1, \tag{23}$$

$$R_2 = \lambda_1 I + R_2 B_2, \tag{24}$$

$$R_3 = R_3 B_1 + R_2 B_3 + \mu_1 (R_3 R_1 + R_2 R_3). \tag{25}$$

For the tail asymptotics of $\boldsymbol{\pi}_{n,2,j}$ for every fixed j, we can concentrate on the analysis of the matrix R_2, which can be expressed as

$$R_2 = \lambda_1 (I - B_2)^{-1} = \lambda_1 \sum_{n=0}^{\infty} B_2^n. \tag{26}$$

Lemma 4.1.1. (a) *B_2 is irreducible;* (b) *the convergence norm of B_2 is $\zeta_{B_2} = \mu_1 + 2\sqrt{\lambda_2 \mu_2}$; and* (c) *$B_2$ is $1/\zeta_{B_2}$-transient and has the following $1/\zeta_{B_2}$-invariant measure $\boldsymbol{x} = (x_i)$ and vector $\boldsymbol{y} = (y_i)$:*

$$x_i = i\left(\sqrt{\frac{\lambda_2}{\mu_2}}\right)^i, \quad y_i = i\left(\sqrt{\frac{\mu_2}{\lambda_2}}\right)^i, \quad i \geq 1.$$

Proof of Lemma 4.1.1. (a) is obvious. By Lemma 2.1 in Kijima[10], (b) is true, and B_2 is $1/\zeta_{B_2}$-transient. (c) is obtained by a direct calculation. □

Based on the above lemma, we have the following properties on R_2.

Lemma 4.1.2. (a) *R_2 is irreducible;* (b) *the convergence norm of R_2 is given by*

$$\zeta_{R_2} = \frac{\lambda_1}{\lambda_1 + \left(\sqrt{\mu_2} - \sqrt{\lambda_2}\right)^2};$$

and (c) *R_2 is $1/\zeta_{R_2}$-transient, and $\boldsymbol{x}$ and $\boldsymbol{y}$ given in Lemma* 4.1.1 *are also a $1/\zeta_{R_2}$-invariant measure and vector of R_2, respectively.*

Proof of Lemma 4.1.2. The proof is obvious by noticing (24) and (26). □

Theorem 4.1.1. *Both $\boldsymbol{\pi}_{n,2}$ and the marginal distribution $\boldsymbol{\pi}_{n,2}\boldsymbol{e}$ have a light-tailed decay as n the number of type-*1 *customers in the system goes to infinity with the same ζ_{R_2} as their decay rate, but neither of them has an exact geometric decay.*

Proof of Theorem 4.1.1. Since R_2 is $1/\zeta_{R_2}$-transient, equation (10) is satisfied for $\eta = \zeta_{R_2}$. Therefore, according to Theorem 2.3.2, $\boldsymbol{\pi}_{n,2}$ has the nongeometric light-tailed decay rate ζ_{R_2}. The proof is complete by a similar argument by applying Theorem 2.3.3 to the marginal vector $\boldsymbol{\pi}_{n,2}\boldsymbol{e}$. In order to apply Theorem 2.3.3, we introduce $\tilde{x} = (\tilde{x}_i)$ as

$$\tilde{x}_i = \left(\sqrt{\frac{\lambda_2}{\mu_2}}\right)^i.$$

Then $\tilde{\boldsymbol{x}}R \le \zeta_{R_2}\tilde{\boldsymbol{x}}$, i.e., $\tilde{\boldsymbol{x}}$ is a $1/\zeta_{R_2}$-subinvariant measure of R_2, which satisfies $\sum_i \tilde{x}_i < \infty$. Now, all conditions of Theorem 2.3.3 are satisfied. □

If we interchange the role of $q_1(t)$ and $q_2(t)$ in the above analysis, the following conclusion can be reached.

Corollary 4.1.1. *Both $\boldsymbol{\pi}_{n,1}$ and the marginal distribution $\boldsymbol{\pi}_{n,1}\boldsymbol{e}$ have a light-tailed decay as n the number of type-2 customers in the system goes to infinity with the same decay rate*

$$\frac{\lambda_2}{\lambda_2 + \left(\sqrt{\mu_1} - \sqrt{\lambda_1}\right)^2},$$

but neither of them has an exact geometric decay.

4.2. Gated Random Order Service Queue

Consider an $M/M/1$ queue with a service room and a waiting room, which was considered by Resing and Rietman[25] using the compensation procedure to compute the joint stationary probability distribution. Our focus is on the analysis of tail asymptotics.

This model is a variation of the standard $M/M/1$ queue by adding a gated waiting room to the system. The original buffer (queue) is then called the service room. Specifically, upon arrival of a customer, if the service room is nonempty, the arriving customer enters the gated waiting room and waits until all customers in the service room complete their services. When the service room becomes empty, all customers in the waiting room are instantaneously transferred into the service room in random order, in which they will receive their services. However, if upon arrival, the service room is empty, the arriving customer goes directly into the service room and receives its service immediately. Assume that customers arrive according to a Poisson process with rate λ and the service times of the customers are exponential with the same rate μ, satisfying the stability condition: $\rho = \lambda/\mu < 1$. Let $X_1(t)$ and $X_2(t)$ be the number of customers in the waiting room and in the service room including the customer receiving the service at time t, respectively. Then, $\{(X_1(t), (X_2(t))); t \ge 0\}$ is a Markov chain with a state space $\{(0, j);$

$j = 0, 1, 2, \ldots\} \cup \{(i, j); i, j = 1, 2, \ldots\}$. Let the stationary probability vector $\boldsymbol{\pi}$ of the Markov chain be partitioned as $\boldsymbol{\pi} = (\boldsymbol{\pi}_0, \boldsymbol{\pi}_1, \boldsymbol{\pi}_2, \ldots)$, where

$$\boldsymbol{\pi}_0 = (\pi_{0,0}, \pi_{0,1}, \pi_{0,2}, \ldots) \quad \text{and} \quad \boldsymbol{\pi}_n = (\pi_{n,1}, \pi_{n,2}, \ldots), \quad n \geq 1,$$

which satisfy the following stationary equations:

$$\begin{aligned}
(\lambda + \mu)\pi_{0,0} &= \mu\pi_{0,1} + \mu\pi_{0,0}, \\
(\lambda + \mu)\pi_{0,1} &= \lambda\pi_{0,0} + \mu\pi_{1,1} + \mu\pi_{0,2}, \\
(\lambda + \mu)\pi_{0,n} &= \mu\pi_{n,1} + \mu\pi_{0,n+1}, \quad n \geq 2, \\
(\lambda + \mu)\pi_{k,n} &= \lambda\pi_{k-1,n} + \mu\pi_{k,n+1}, \quad k, n \geq 1.
\end{aligned}$$

Upon uniformization, we can convert the infinitesimal generator to a transition probability matrix P given as

$$P = \begin{bmatrix} B_0 & A_0 & & & & \\ C_1 & B & A & & & \\ C_2 & & B & A & & \\ \vdots & & & \ddots & \ddots & \\ \vdots & & & & \ddots & \end{bmatrix},$$

where

$$B = \frac{1}{1+\rho}E \quad \text{and} \quad A = \frac{\rho}{1+\rho}I$$

with

$$E = \begin{bmatrix} 0 & & & \\ 1 & 0 & & \\ & 1 & 0 & \\ & & \ddots & \ddots \end{bmatrix}.$$

The details of the other matrices are not of interest in our analysis. The stationary distribution $\boldsymbol{\pi}$ has a matrix geometric form solution: $\boldsymbol{\pi}_n = \boldsymbol{\pi}_1 R^{n-1}$, $n = 1, 2, \ldots$, where R is the minimal nonnegative solution to $R = A + RB$.

Lemma 4.2.1. *For the gated random order service queue, we have the following:*

(a) $R = \frac{\rho}{1+\rho} \sum_{k=0}^{\infty} \left(\frac{1}{1+\rho}\right)^k E^k$.
(b) $\lim_{n\to\infty} \frac{R^n}{\eta^n} = 0$ *for* $\frac{\rho}{1+\rho} < \eta \leq 1$.

(c) *For* $\frac{\rho}{1+\rho} < \eta \leq 1$, *$R$ has a $1/\eta$-invariant measure* $\boldsymbol{w} = (w_1, w_2, \dots)$ *given by* $w_{k+1} = \theta_\eta^k w_1$, *where*

$$\theta_\eta = \frac{\eta(1+\rho) - \rho}{\eta}.$$

In particular, when $\eta = \rho$, *we have* $\theta_\rho = \rho$. *That is,* $\boldsymbol{x} = (1, \rho, \rho^2, \dots)$ *satisfies* $\boldsymbol{x}R = \rho\boldsymbol{x}$. *In the case that* $0 < \eta \leq \frac{\rho}{1+\rho}$, *$R$ has no positive $1/\eta$-subinvariant measure.*

(d) $\lim_{k\to\infty} \frac{\pi_{1,k}}{x_k} = \lim_{k\to\infty} \frac{\pi_{1,k}}{\rho^{k-1}} = 0$ *and* $\lim_{k\to\infty} \frac{\pi_{1,k}}{w_k} = \lim_{k\to\infty} \frac{\pi_{1,k}}{\theta_\eta^{k-1} w_1} = \infty$ *for* $\frac{\rho}{1+\rho} < \eta < \rho$.

Proof of Lemma 4.2.1. (a) Let $R^* = \frac{\rho}{1+\rho}\sum_{k=0}^{\infty}(\frac{1}{1+\rho})^k E^k$, the summation in (i). Then, R^* is finite and nonnegative by noticing that $E\boldsymbol{e} \leq \boldsymbol{e}$ and therefore $E^n\boldsymbol{e} \leq \boldsymbol{e}$, for all $n \geq 0$. By direct calculations, we can easily check that R^* is a solution to

$$R = A + RB = \frac{\rho}{1+\rho} I + \frac{1}{1+\rho} RE, \tag{27}$$

which implies that

$$R^*\left(I - \frac{1}{1+\rho}E\right) = \frac{\rho}{1+\rho} I. \tag{28}$$

To see that R^* is the minimal nonnegative solution to $R = A + RB$, notice that by equation (28), any nonnegative solution R must satisfy (after N iterations)

$$R = \frac{\rho}{1+\rho}\sum_{k=0}^{N}\left(\frac{1}{1+\rho}\right)^k E^k + \left(\frac{1}{1+\rho}\right)^N R^{N+1}E.$$

Taking N to infinity on the both sides, we obtain

$$R = R^* + \lim_{N\to\infty}\left(\frac{1}{1+\rho}\right)^N R^{N+1}E \geq R^*,$$

since R is nonnegative. Thus, R^* is the minimal nonnegative solution.

(b) Based on (i) and mathematical induction, we can show that

$$R^n = \left(\frac{\rho}{1+\rho}\right)^n \sum_{k=0}^{\infty}\binom{n-1+k}{k}\left(\frac{1}{1+\rho}\right)^k E^k, \quad \text{for all } n \geq 1.$$

Let $R^n = (r_{i,j}^{(n)})$. From the structure of E, we have

$$r_{j+k,j}^{(n)} = \left(\frac{\rho}{1+\rho}\right)^n \binom{n-1+k}{k} \left(\frac{1}{1+\rho}\right)^k, \quad k = 0, 1, 2, \ldots,$$

and $r_{i,j}^{(n)} = 0$ whenever $i < j$, from which the property in (ii) is immediate.

(c) This can be verified by direct calculations based on $w = \boldsymbol{w}R$, where R is given in (i).

(d) It follows from the result in Resing and Rietman[25] that

$$\pi_{1,k} = \rho^{k-1} \sum_{m=1}^{\infty} a_{m,k} b_m,$$

where

$$a_{m,k} = \left(\frac{1-\rho^m}{1-\rho^{m+1}}\right)^{k-1} \quad \text{and} \quad b_m = \left(\frac{\rho(1-\rho^m)}{1-\rho^{m+1}}\right)\left(\frac{\rho^{3+m}}{(1-\rho^{m+1})(1-\rho^m)}\right).$$

Clearly, $\sum_{m=1}^{\infty} b_m < \infty$ and for each m, $\lim_{k\to\infty} a_{m,k} = 0$, and $a_{m,k} \le 1$. Therefore, by the dominated convergence theorem,

$$\lim_{k\to\infty} \frac{\pi_{1,k}}{x_k} = \rho^{-1} \sum_{m=1}^{\infty} b_m \lim_{k\to\infty} a_{m,k} = 0.$$

For $\frac{\rho}{1+\rho} < \eta < \rho$, by Lemma 4.2.1(3), we have $w_{k+1} = \theta_\eta^k w_1$. Clearly, $\frac{\rho}{\theta_\eta} > 1$ since $\theta_\eta = \frac{\eta(1+\rho)-\rho}{\eta} < \rho$. It follows that there exists a positive integer m_0 such that

$$\frac{1-\rho^{m_0-1}}{1-\rho^{m_0}} \frac{\rho}{\theta_\eta} > 1$$

since

$$\lim_{m\to\infty} \frac{1-\rho^{m-1}}{1-\rho^m} = 1.$$

Hence, we have

$$\begin{aligned}
\frac{\pi_{1,k}}{\theta_\eta^{k-1}} &= \frac{\rho^{k-1}}{\theta_\eta^{k-1}} \sum_{m=1}^{\infty} b_m \left(\frac{1-\rho^{m-1}}{1-\rho^m}\right)^{k-1} \\
&\ge \frac{\rho^{k-1}}{\theta_\eta^{k-1}} b_{m_0} \left(\frac{1-\rho^{m_0-1}}{1-\rho^{m_0}}\right)^{k-1} \longrightarrow \infty, \quad \text{as } k \to \infty,
\end{aligned}$$

since $b_m > 0$ for every m. This completes the proof of the lemma. □

Theorem 4.2.1. *Both the joint distribution $\boldsymbol{\pi}_n$ and the marginal distribution $\boldsymbol{\pi}_n \boldsymbol{e}$ have a light-tailed decay with the same decay rate ρ, but neither of them has an exact geometric decay.*

Proof of Theorem 4.2.1. Lemma 4.2.1(b), (c) and (d) imply condition (10). Therefore, we can apply Theorem 2.3.2 to conclude the light-tailed behavior along the level direction. The conclusion for the marginal distribution can be similarly done by applying Theorem 2.3.3. □

As a conclusion, we make the following remarks. We expect that the main result or idea presented in this article can be used to analyze other interesting models. In combination with the large deviations method, it is promising to extend the aforementioned study to multi-dimensional systems.

ACKNOWLEDGMENTS

The authors thank Masakiyo Miyazawa of Tokyo University of Sciences for the discussions on Lemma 2.1 and Keith Oliver for proofreading the manuscript.

REFERENCES

1. Flatto, L.; McKean, H.P. Two queues in parallel. Communications on Pure and Applied Mathematics **1977**, *30*, 255–263.
2. Flatto, L.; McKean, H.P. Two parallel queues created by arrivals with two demands I. SIAM J. Appl. Math. **1984**, *44*, 1041–1053.
3. Foley, R.D.; McDonald, D.R. Join the shortest queue: stability and exact asymptotics. Annals of Applied Probability **2001**, *11*, 569–607.
4. Foley, R.D.; McDonald, D.R. Large deviations of a modified Jackson network: stability and rough asymptotics. Annals of Applied Probability **2004**, *15*, 519–541.
5. Foley, R.D.; McDonald, D.R. Bridges and networks: exact asymptotics. Annals of Applied Probability **2005**, *15*, 542–586.
6. Haque, L. Tail Behavior for Stationary Distributions for Two-Dimensional Stochastic Models. PhD thesis, Carleton University: Ottawa, Canada, 2003.
7. Haque, L.; Liu, L.; Zhao, Y.Q. Sufficient conditions for a geometric tail in a QBD process with countably many levels and phases. Stochastic Models **2005**, *21*, 77–99.
8. Jelenković, P.R.; Lazar, A.A. Asymptotic results for multiplexing subexponential on-off processes. Advances in Applied Probability **1999**, *31*, 394–421.
9. Kemeny, J.G.; Snell, J.L.; Knapp, A.W. *Denumerable Markov Chains*; 2nd Ed., Springer-Verlag, New York, 1976.
10. Kijima, M. Quasi-stationary distributions of single-server phase type queues. Mathematics of Operations Research **1993**, *18*, 423–437.
11. Kroese, D.P.; Scheinhardt, W.R.W.; Taylor, P.G. Spectral properties of the tandem Jackson network, seen as a quasi-birth-and-death process. Annals of Applied Probability **2004**, *14*, 2057–2089.
12. Leeuwaarden, Johan S.H. van Queueing Models for Cable Access Networks. PhD thesis, Technische Universiteit Eindhoven, The Netherlands, 2005.
13. Li, H.; Zhao, Y.Q. A retrial queue with a constant retrial rate, server break downs and impatient customers. Stochastic Models **2005**, *21*, 531–550.

14. Li, L.; Miyazawa, M.; Zhao, Y.Q. Geometric decay in a QBD process with countable background states with applications to a join-the-shortest-queue model. Stochastic Models **2007**, *23*, 413–438.
15. Liu, L.; Miyazawa, M.; Zhao, Y.Q. Geometric decay in level-expanding QBD models. Annals of Operations Research **2008**, *160*, 83–98.
16. McDonald, D.R. Asymptotic of first passage times for random walk in an orthant. Annals of Applied Probability **1999**, *9*, 110–145.
17. Miller, D.R. Computation of steady-state probabilities for $M/M/1$ priority queues. Operations Research **1981**, *29*, 945–958.
18. Miyazawa, M. The Markov renewal approach to $M/G/1$ type queues with countably many background states. Queueing Systems **2004**, *46*, 177–196.
19. Miyazawa, M.; Zhao, Y.Q. The stationary tail asymptotics in the $GI/G/1$ type queue with countably many background states. Adv. in Appl. Probab. **2004**, *36*, 1231–1251.
20. Miyazawa, M. Doubly *QBD* process and a solution to the tail decay rate problem. The Second Asia-Pacific Symposium on Queueing Theory and Network Applications (QTNA2007), August 1–4, 2007, Kobe, Japan.
21. Miyazawa, M. *Tail Decay Rates in Doubly QBD Processes*; 2007, submitted.
22. Motyer, A.J.; Taylor, P.G. Decay rates for quasi-birth-and-death process with countably many phases and tri-diagonal block generators. Advances in Applied Probability **2006**, *38*, 522–544.
23. Neuts, M.F. *Matrix-Geometric Solutions in Stochastic Models: An Algorithmic Approach*; The Johns Hopkins University Press, Baltimore, MD, 1981.
24. Ridder, A.; Shwartz, A. Large deviations without principle; Join the shortest queue. Mathematical Method in Operations Research **2005**, *62*, 467–483.
25. Resing, J.A.C.; Rietman, R. The $M/M/1$ queue with gated random order of service. Statistica Neerlandica **2004**, *58*, 97—110.
26. Sakuma, Y.; Miyazawa, M.; Zhao, Y.Q. Decay rate for a $PH/M/2$ queue with shortest queue discipline. Queueing Systems **2006**, *53*, 189–201.
27. Seneta, E. *Non-Negative Matrices and Markov Chains*; Springer-Verlag, New York, 1981.
28. Takahashi, Y.; Fujimoto, 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*, 1–24.
29. Tang, J.; Zhao, Y.Q. Stationary tail asymptotics of a tandem queue with feedback. Annals of Operations Research **2008**, *160*, 173–190.
30. Tweedie, R. Operator-geometric stationary distributions for Markov chains, with application to queueing models. Adv. Appl. Probab. **1982**, *14*, 368–391.
31. van Uitert, M.J.G. Generalized Processor Sharing Queues. PhD thesis, Technical University of Eindhoven, Eindhoven, The Netherlands, 2003.
32. Xue, J.; Alfa, A.S. Tail probability of low-priority queue length in a discrete-time priority $BMAP/PH/1$ queue. Stochastic Models **2005**, *21*, 799–820.
33. Zhao, Y.Q. Censoring technique in studying block-structured Markov chains, In *Advances in Algorithmic Methods for Stochastic Models*; Latouche, G., Taylor, P., Eds.; Notable Publications: New Jersey, USA, 2000, 417–433.