

ON THE EFFECT OF FINITE BUFFER TRUNCATION IN A TWO-NODE JACKSON NETWORK

YUTAKA SAKUMA * AND

MASAKIYO MIYAZAWA, ** *Tokyo University of Science*

Abstract

We consider a two-node Jackson network in which the buffer of node 1 is truncated. Our interest is in the limit of the tail decay rate of the queue-length distribution of node 2 when the buffer size of node 1 goes to infinity, provided that the stability condition of the unlimited network is satisfied. We show that there can be three different cases for the limit. This generalizes some recent results obtained for the tandem Jackson network. Special cases and some numerical examples are also presented.

Keywords: Decay rate; Jackson network; finite truncation; stationary distribution

2000 Mathematics Subject Classification: Primary 60K25; 60F10
Secondary 60J10; 60K20

1. Introduction

We are interested in computing the tail decay rates of the stationary queue-length distributions of queueing networks by the finite truncation of some buffers, provided that the stability of the networks holds. In general, the decay rates are hard to compute for queueing networks unless they have product-form stationary distributions, and are more easily computed for queues with single buffers (see, e.g. [3], [9], [10], and references therein). So, it is natural to approximate the decay rates for queueing networks by truncating the buffers of all queues except for the ones in which we are interested. This is referred to as a finite truncation. We expect that such a truncation approximates the original decay rate well as the truncation level becomes large.

Surprisingly, however, Kroese *et al.* [8] have recently reported that finite truncation may not work well for a tandem Jackson network of two nodes. Suppose that the first buffer is truncated: they proved that the tail decay rate for the queue at the second node converges to a value smaller than the original one as the truncation level goes to infinity, given that the first node has higher traffic intensity than the second node. If this is not the case, then the original decay rate is well approximated.

In this paper, we consider this truncation problem for a more general network, i.e. a two-node Jackson network with arbitrary routing topology. As is well known, the joint queue lengths of this network have a product-form stationary distribution. Our aim is to see what conditions guarantee that the tail decay rate of the queue at one node is well approximated by the finite buffer truncation of the other. We also consider the value to which the decay rate of the truncated model converges if the approximation fails. This value will indicate when we must be careful in approximating queueing networks with infinite buffers by those with finite buffers.

Received 10 May 2004; revision received 8 September 2004.

* Postal address: Department of Information Sciences, Tokyo University of Science, Noda City, Chiba 278-8510, Japan.

** Email address: miyazawa@is.noda.tus.ac.jp

Since the Jackson network may have feedback loops and exogenous arrivals at both nodes, the situation looks much more complicated than the tandem queue of [8]. Nevertheless, we can identify the conditions under which the finite truncation yields good approximations of the original decay rates, and to which values the rates converge if this is not the case. For the tandem queue, there are two possible values for the decay rates: the original and another, smaller, value. For the general Jackson network of two nodes, we have three possible values, and some of the necessary conditions are nonintuitive. To prove our results, we use a formulation in terms of a quasi-birth–death process, as in [8]. However, we employ an approach different from that in [8], where orthogonal polynomials are used in the proofs. This technique seems to be highly dependent on the structure of the transition-rate matrix of the tandem Jackson network. Our approach, on the other hand, is direct and elementary, but requires results of [12].

In this way, we generalize the aforementioned results of [8]. However, there is another noteworthy finding in [8], namely that the decay rate can be arbitrarily changed in some range if we appropriately modify the behaviour of the queue at the first node when the second node is empty. This implies that the decay rate also strongly depends on the boundary behaviour of the queueing network and of the quasi-birth–death process, in general. We do not address this problem here.

The paper comprises four sections. In Section 2, we describe the Jackson network of two nodes in which the buffer of node 1 is truncated, and present basic results for computing the decay rate. A complete characterization of the limit of the decay rate is given in Section 3. In Section 4, we discuss special cases, including the tandem Jackson network studied in [8]. We also illustrate the convergence of the decay rate with numerical examples.

2. Two-node Jackson network with a truncated buffer

We consider a version of the two-node Jackson network. The nodes of this network are numbered 1 and 2. Node 1 has a buffer with capacity $m \leq \infty$, which includes the customer being served, so the total number of customers in node 1 is limited by m . Node 2 has an infinite buffer. Customers arrive at node 1 or node 2 according to a Poisson process of rate λ_1 or λ_2 , respectively. When a customer arriving at node 1 finds the buffer full, he is rejected and leaves the network. Customers at each node are served in a first-come–first-served manner. The service times of customers at node 1 and node 2 are independently and exponentially distributed with means μ_1^{-1} and μ_2^{-1} , respectively. After completing service at node 1 or node 2, customers enter node 2 or node 1 with probability p or q , or leave the system with probability $1 - p$ or $1 - q$, respectively, where $0 \leq p, q \leq 1$. To exclude a trivial case, we assume that $p > 0$. This queueing network is referred to as a *two-node Jackson network with a truncated buffer*, and is a standard (two-node) Jackson network if m is infinite.

Let $L_{i(m)}(t)$ be the number of customers in node i at time t , for $i = 1, 2$. Clearly, $(L_{1(m)}(t), L_{2(m)}(t))$ is a two-dimensional continuous-time Markov chain with state space $\mathcal{M} \times \mathcal{N}$, where $\mathcal{M} := \{0, 1, \dots, m\}$ and $\mathcal{N} := \{0, 1, \dots\}$. The state of this Markov chain is referred to as a network state. When $m = \infty$, $L_{i(\infty)}(t)$ is denoted by $L_i(t)$, for simplicity.

Define ρ_1 and ρ_2 as

$$\rho_1 = \frac{\lambda_1 + q\lambda_2}{(1 - pq)\mu_1}, \quad \rho_2 = \frac{p\lambda_1 + \lambda_2}{(1 - pq)\mu_2}.$$

It is easy to show that ρ_1 and ρ_2 are respectively the traffic intensities of nodes 1 and 2 when m is infinite. In this case, if

$$\rho_1 < 1 \quad \text{and} \quad \rho_2 < 1, \tag{2.1}$$

then the Markov chain $(L_1(t), L_2(t))$ has the stationary distribution π , given by

$$\pi(n_1, n_2) = (1 - \rho_1)(1 - \rho_2)\rho_1^{n_1}\rho_2^{n_2}, \quad n_1, n_2 \geq 0, \quad (2.2)$$

as is well known (see, e.g. [6]). Here, (n_1, n_2) denotes the network state.

Throughout the paper, we assume that (2.1) holds. Then, the stationary distribution of the network state always exists for $1 \leq m \leq \infty$, since the process $(L_{1(m)}(t), L_{2(m)}(t))$ is stochastically dominated by the case of $m = \infty$, as we will see in Lemma 2.1. We denote this (finite- m) stationary distribution by $\pi_{(m)}$, and denote random vectors subject to $\pi_{(m)}$ by $(L_{1(m)}, L_{2(m)})$. Thus,

$$\pi_{(m)}(n_1, n_2) = P(L_{1(m)} = n_1, L_{2(m)} = n_2).$$

Our primary concern is with the tail decay rate of the stationary distribution of node 2, which is defined as the positive number $z_{(m)}$ such that

$$\lim_{n \rightarrow \infty} z_{(m)}^{-n} P(L_{2(m)} = n) = c_{(m)},$$

for some positive constant $c_{(m)}$. The existence of this limit will be verified in Lemma 3.1. This definition is slightly stronger than

$$z_{(m)} = \lim_{n \rightarrow \infty} \frac{P(L_{2(m)} = n + 1)}{P(L_{2(m)} = n)}.$$

Note that the value of $z_{(m)}$ is also determined by

$$z_{(m)} = \inf \left\{ z \geq 0: \sum_{n=0}^{\infty} z^{-n} P(L_{2(m)} \geq n) < \infty \right\}, \quad (2.3)$$

since

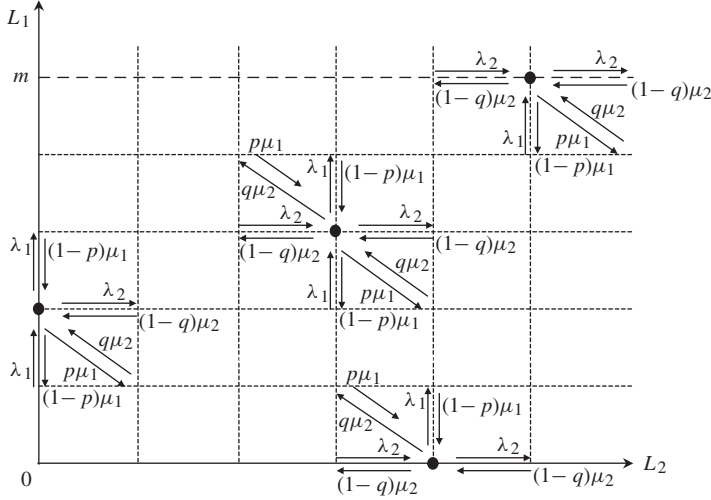
$$(1 - z) \sum_{n=0}^{\infty} z^{-n} P(L_{2(m)} \geq n) = \sum_{n=0}^{\infty} z^{-n} P(L_{2(m)} = n) - z.$$

We will consider the limit of $z_{(m)}$ as m goes to infinity. To this end, we first consider the stationary distribution $\pi_{(m)}$ for each $m < \infty$. Viewing $L_{1(m)}(t)$ as the phase and $L_{2(m)}(t)$ as the level, the continuous-time Markov chain $(L_{1(m)}(t), L_{2(m)}(t))$ can be considered a quasi-birth-death process with $m + 1$ phases in each level. After uniformization of this quasi-birth-death process, its stationary distribution is identical to that of the discrete-time Markov chain with the transition diagram shown in Figure 1, where the constants $\lambda_1, \lambda_2, \mu_1$, and μ_2 are normalized in such a way that

$$\lambda_1 + \lambda_2 + \mu_1 + \mu_2 = 1. \quad (2.4)$$

Hence, the transition matrix $P_{(m)}$ of this discrete-time Markov chain is block tridiagonal and given by

$$P_{(m)} = \begin{pmatrix} \mathbf{B}_{0(m)} & \mathbf{A}_{1(m)} & & & \\ \mathbf{B}_{-1(m)} & \mathbf{A}_{0(m)} & \mathbf{A}_{1(m)} & & \\ & \mathbf{A}_{-1(m)} & \mathbf{A}_{0(m)} & \ddots & \\ & & \mathbf{A}_{-1(m)} & \ddots & \\ & & & \ddots & \ddots \end{pmatrix},$$

FIGURE 1: The transition diagram of $(L_{1(m)}(t), L_{2(m)}(t))$.

where $\mathbf{B}_{-1(m)} = \mathbf{A}_{-1(m)}$, $\mathbf{B}_{0(m)}$, $\mathbf{A}_{0(m)}$, and $\mathbf{A}_{1(m)}$ are $(m+1)$ -dimensional square matrices given by

$$\mathbf{B}_{0(m)} = \begin{pmatrix} \mu_1 + \mu_2 & \lambda_1 & & & \\ (1-p)\mu_1 & \mu_2 & \lambda_1 & & \\ & \ddots & \ddots & \ddots & \\ & & (1-p)\mu_1 & \mu_2 & \lambda_1 \\ & & & (1-p)\mu_1 & \lambda_1 + \mu_2 \end{pmatrix},$$

$$\mathbf{A}_{1(m)} = \begin{pmatrix} \lambda_2 & & & & \\ p\mu_1 & \lambda_2 & & & \\ & \ddots & \ddots & & \\ & & p\mu_1 & \lambda_2 & \\ & & & p\mu_1 & \lambda_2 \end{pmatrix},$$

$$\mathbf{A}_{0(m)} = \begin{pmatrix} \mu_1 & \lambda_1 & & & \\ (1-p)\mu_1 & 0 & \lambda_1 & & \\ & \ddots & \ddots & \ddots & \\ & & (1-p)\mu_1 & 0 & \lambda_1 \\ & & & (1-p)\mu_1 & \lambda_1 + q\mu_2 \end{pmatrix},$$

$$\mathbf{A}_{-1(m)} = \begin{pmatrix} (1-q)\mu_2 & q\mu_2 & & & \\ & (1-q)\mu_2 & q\mu_2 & & \\ & & \ddots & \ddots & \\ & & & (1-q)\mu_2 & q\mu_2 \\ & & & & (1-q)\mu_2 \end{pmatrix}.$$

It is easy to show that $\mathbf{P}_{(m)}$ is irreducible and aperiodic, meaning that the Markov chain can have at most one stationary distribution. The results of the following lemma are intuitively obvious, but we state the proof, as it plays an important role in this paper.

Lemma 2.1. *For each $t \geq 0$, $(L_{1(m)}(t), L_{2(m)}(t))$ is stochastically increasing in m , i.e.*

$$P(L_{1(m)}(t) \geq n_1, L_{2(m)}(t) \geq n_2)$$

is increasing in m for each $n_1, n_2 \geq 0$. Therefore, $(L_{1(m)}(t), L_{2(m)}(t))$ is stochastically bounded by $(L_1(t), L_2(t))$.

Proof. We alter the service discipline at node 1 from first-come–first-served to last-come–first-served, and adopt a push-out rejection policy, whereby the longest-staying customer is forced to leave the node when an arriving customer finds the buffer full. For this network, we denote the number of customers in node i by $\tilde{L}_{i(m)}(t)$ for $i = 1, 2$. By the memoryless property of the exponential distribution, process $(L_{1(m)}(t), L_{2(m)}(t))$ is stochastically identical to process $(\tilde{L}_{1(m)}(t), \tilde{L}_{2(m)}(t))$. Since new customers are always accepted for service at node 1, the customers there remain when m is increased. Hence, $\tilde{L}_{1(m)}(t)$ and $\tilde{L}_{1(m+1)}(t)$ decrease at the same instants, and we have

$$\tilde{L}_{1(m)}(t) \leq \tilde{L}_{1(m+1)}(t), \quad t \geq 0. \quad (2.5)$$

As an example, see Figure 2, which depicts how the longest-staying customers are replaced by those newly arriving, where ‘ $\uparrow$ ’ and ‘ $\Downarrow$ ’ represent the arrival and the departure of a customer, respectively. For each $m \geq 1$, let $\tilde{N}_{1(m)}(t)$ be the number of customers who complete service at node 1 by time t . Then, the above arguments imply that

$$\tilde{N}_{1(m)}(t) \leq \tilde{N}_{1(m+1)}(t), \quad t \geq 0.$$

Hence, we have

$$\tilde{L}_{2(m)}(t) \leq \tilde{L}_{2(m+1)}(t), \quad t \geq 0.$$

Combining this with (2.5) concludes the proof.

By Lemma 2.1, the process $(L_{1(m)}(t), L_{2(m)}(t))$ is stochastically bounded by $(L_1(t), L_2(t))$, and the condition (2.1) implies the existence of the stationary distribution (2.2) of $(L_1(t), L_2(t))$. Therefore, the process $(L_{1(m)}(t), L_{2(m)}(t))$ has a unique stationary distribution; that is, it is ergodic.

We now review basic results for a general block-tridiagonal transition matrix $\mathbf{P}$ given by

$$\mathbf{P} = \begin{pmatrix} \mathbf{B}_0 & \mathbf{A}_1 & \mathbf{O} & \cdots \\ \mathbf{B}_{-1} & \mathbf{A}_0 & \mathbf{A}_1 & \ddots \\ \mathbf{O} & \mathbf{A}_{-1} & \mathbf{A}_0 & \ddots \\ \mathbf{O} & \mathbf{O} & \mathbf{A}_{-1} & \ddots \\ \vdots & \ddots & \ddots & \ddots \end{pmatrix}, \quad (2.6)$$

where $\mathbf{A}_{-1}$, $\mathbf{A}_0$, and $\mathbf{A}_1$ are square matrices of dimension $m+1$, $\mathbf{B}_0$ and $\mathbf{B}_{-1}$ are $(m+1) \times (m_0+1)$ matrices, and $\mathbf{O}$ stands for a null matrix of the appropriate size. Here, $0 \leq m_0, m \leq \infty$: note that the blocks $\mathbf{A}_i$ and $\mathbf{B}_i$ may be of infinite dimension.

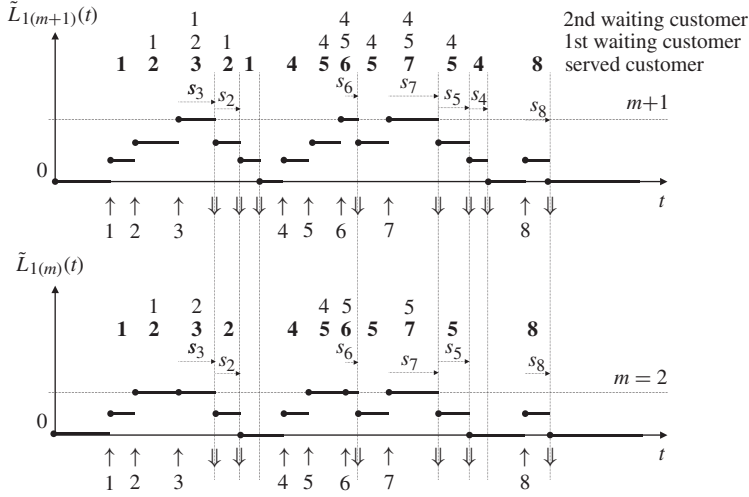


FIGURE 2: Sample paths of $\tilde{L}_{1(m)}(t)$ and $\tilde{L}_{1(m+1)}(t)$, where $m = 2$. The service of customer i is denoted by s_i .

Also note that $\{A_\ell, \ell = -1, 0, 1\}$ defines a discrete-time Markov additive process

$$\{(X_n, Y_n)\}_{n \geq 0}$$

such that

$$P(X_{n+1} = j, Y_{n+1} - Y_n = \ell \mid X_n = i) = A_\ell(i, j),$$

where $A_\ell(i, j)$ is the (i, j) th entry of A_ℓ . In what follows, we will consider the periodicity of this Markov additive process. In this section, we assume that the transition matrix

$$\mathbf{A} := \mathbf{A}_{-1} + \mathbf{A}_0 + \mathbf{A}_1$$

is irreducible. We next define the periodicity of the Markov additive kernel $\{A_\ell, \ell = -1, 0, 1\}$, extending the corresponding definition of the Markov renewal process in [2, Chapter 10, Section 2]. For each $i \in \mathcal{M}$, let $d(i)$ be the greatest common divisor of

$$\{\ell_1 + \ell_2 + \cdots + \ell_k : A_{\ell_1}(i, j_1)A_{\ell_2}(j_1, j_2) \times \cdots \times A_{\ell_k}(j_{k-1}, i) > 0, \text{ where} \\ k \geq 1; \ell_m = -1, 0, 1; m = 1, 2, \dots, k; j_n \in \mathcal{M}; n = 1, 2, \dots, k-1\}.$$

It is easy to show that, by the irreducibility of $\mathbf{A}$, $d(i)$ does not depend on i . Therefore we simply denote it by d , and call the Markov additive process d -arithmetic. For the two-node Jackson network, $\{A_{\ell(m)}, \ell = -1, 0, 1\}$ is clearly 1-arithmetic except in trivial cases.

Remark 2.1. In the literature, d -arithmeticity means that d is the largest positive number such that there exists a function δ from $\mathcal{M}$ to $[0, d)$, and $A_\ell(i, j) > 0$ implies that $\ell - \delta(i) + \delta(j)$ is a multiple of d (see, e.g. [1] and [15]). It is clear that this definition is equivalent to the above definition (see, e.g. [12] for more details). Since we are only concerned with the integer-valued additive component, 1-arithmeticity may be said to be *nonarithmeticity*. However,

this terminology is also applied to a Markov additive process that has a real-valued additive component (see, e.g. [1]), which includes the present model as a special case. To avoid this possible confusion, we prefer ‘1-arithmetic’ to ‘nonarithmetic’.

Since $\mathbf{P}_{(m)}$ is block-tridiagonal and ergodic and since $\{\mathbf{A}_{\ell(m)}, \ell = -1, 0, 1\}$ is 1-arithmetic, the following results will be of use to us.

Proposition 2.1. *Suppose that $\mathbf{P}$ of (2.6) is ergodic, and recall that m and m_0 may be infinite. Let $\mathbf{u} := \{\mathbf{u}_n, n \geq 0\}$ be its stationary distribution, where $\mathbf{u}_n$ is the (row) vector whose entries u_{ni} ($i \in \mathcal{M}$) are the stationary probabilities corresponding to level n , i.e. the n th block of $\mathbf{P}$.*

(i) *There then exists a minimal nonnegative solution $\mathbf{R}$ of the quadratic matrix equation*

$$\mathbf{X} = \mathbf{A}_1 + \mathbf{X}\mathbf{A}_0 + \mathbf{X}^2\mathbf{A}_{-1}, \quad (2.7)$$

and we have

$$\mathbf{u}_n = \mathbf{u}_0 \mathbf{R}^n, \quad n = 0, 1, 2, \dots,$$

where $\mathbf{u}_0$ is obtained as the left-invariant measure of $\mathbf{B}[\mathbf{R}] := \mathbf{B}_0 + \mathbf{R}\mathbf{B}_{-1}$ such that

$$\mathbf{u}_0(\mathbf{I} - \mathbf{R})^{-1} \mathbf{e} = 1,$$

where $\mathbf{I}$ is the identity matrix and $\mathbf{e}$ is the column vector all of whose entries are 1.

(ii) *Define the matrix generating function $\mathbf{A}(s)$ by $\mathbf{A}_1 + s\mathbf{A}_0 + s^2\mathbf{A}_{-1}$. If the Markov additive kernel $\{\mathbf{A}_\ell, \ell = -1, 0, 1\}$ is 1-arithmetic and if there exist a positive row vector $\mathbf{x}$, a positive column vector $\mathbf{y}$, and a real number $z \in (0, 1)$ satisfying the conditions*

$$\mathbf{x}\mathbf{A}(z) = z\mathbf{x}, \quad (2.8)$$

$$\mathbf{A}(z)\mathbf{y} = z\mathbf{y}, \quad (2.9)$$

$$\mathbf{x}\mathbf{y} < \infty, \quad (2.10)$$

then z^{-1} is the convergence radius of the matrix $\mathbf{R}$. Moreover, if

$$\mathbf{u}_0(z\mathbf{B}_0 + \mathbf{A}_1)\mathbf{y} < \infty \quad (2.11)$$

is satisfied, then z is the decay rate of the stationary probability u_{ni} with respect to n for each fixed $i \in \mathcal{M}$. That is,

$$\lim_{n \rightarrow \infty} z^{-n} u_{ni} = c_i$$

for some positive constant c_i .

Note that if m and m_0 are finite, i.e. if all the block sizes are finite, (i) is the well-known result due to Neuts [13], which is called the matrix-geometric form of a distribution. It is obvious how to extend this result to the case of infinite block sizes (see, e.g. [4], [8], and references therein). For finite m and m_0 , (ii) is known for a more general class of transition matrices (see, e.g. [5]), for which (2.10) and (2.11) are automatically satisfied. However, for infinite m and m_0 , these two conditions are crucial, as discussed in [12]. A different set of sufficient conditions for the geometric decay is given in [8]. In both situations, the marginal distribution at level 0 is important (see Section 1). For the infinite case, (ii) has only recently been proved. Fujimoto *et al.* [4] did so first, but they assumed certain extra conditions. These conditions were removed in [11] and [12], in which, respectively, the M/GI/1-type and GI/G/1-type queues with infinite phase spaces were studied.

Remark 2.2. The matrix $\mathbf{R}$ in Proposition 2.1 has the following probabilistic interpretation. Let $\tau_\ell = \inf\{n \geq 1 : Y_n \leq \ell\}$, where Y_n is the additive component of the Markov additive process $\{(X_n, Y_n)\}_{n \geq 0}$. That is, τ_ℓ is the first return time of the additive component to the set $(-\infty, \ell]$. Then the matrix $\mathbf{R}$ can be defined (in terms of its entries $R(i, j)$) by

$$R(i, j) := \mathbb{E} \left[\sum_{n \geq 0} \mathbf{1}(X_n = j, Y_n = \ell + 1, n < \tau_\ell) \mid X_0 = i, Y_0 = \ell \right], \quad i, j \in \mathcal{M}.$$

Similarly, if we define the matrix $\mathbf{U}$ by

$$U(i, j) := \mathbb{E} \left[\sum_{n \geq 0} \mathbf{1}(X_n = j, Y_n = \ell, n < \tau_{\ell-1}) \mid X_0 = i, Y_0 = \ell \right], \quad i, j \in \mathcal{M},$$

then it is easy to show that

$$\mathbf{R} = \mathbf{A}_1 \mathbf{U}.$$

It is also easily shown that all the entries of $\mathbf{U}$ are positive for the two-node Jackson network with a truncated buffer.

3. Limit of the decay rate

In this section, we consider the limit of $z_{(m)}$ for the two-node Jackson network with a truncated buffer of size m at node 1, as m goes to infinity. We first note the following facts, which are immediate consequences of (2.3), Lemma 2.1, and Proposition 2.1.

Lemma 3.1. *The decay rate $z_{(m)}$ exists, increases in m , and is bounded by ρ_2 .*

By this lemma, the limit of $z_{(m)}$ exists. We denote it by z_∞ :

$$z_\infty = \lim_{m \rightarrow \infty} z_{(m)}.$$

Note that z_∞ may not be the decay rate when $m = \infty$. This rate is actually ρ_2 , by (2.2). In fact, there can be three different cases for the limit z_∞ . These are presented in the following theorems.

Theorem 3.1. *Consider the two-node Jackson network with a truncated buffer satisfying the stability condition (2.1). If the condition*

$$\frac{p\rho_1}{1 - \rho_1} \leq \frac{\rho_2}{1 - \rho_2} \tag{3.1}$$

holds, then $z_\infty = \rho_2$.

For the next theorem and its proof, we need some extra notation. Define the function $f(r, z)$ as

$$f(r, z) = \mu_1 \{p + (1 - p)z\}r^2 + \{\lambda_2 - z + (1 - q)\mu_2 z^2\}r + (\lambda_1 + q\mu_2 z)z.$$

For each fixed z , $f(r, z) = 0$ is a quadratic equation in r . Let $D(z)$ be its discriminant, namely

$$D(z) = \{\lambda_2 - z + (1 - q)\mu_2 z^2\}^2 - 4\mu_1 z \{p + (1 - p)z\}(\lambda_1 + q\mu_2 z).$$

Theorem 3.2. For the model considered in Theorem 3.1, the equation $D(z) = 0$ has solutions in $(0, \rho_2]$. Let z_0 be the largest of them, and let

$$z_1 = \frac{p\lambda_1\mu_1 + \lambda_2(\lambda_1 + q\mu_2)}{\{(1-p)q\mu_1 + (1-q)(\lambda_1 + q\mu_2)\}\mu_2}.$$

If (3.1) does not hold, i.e. if

$$\frac{p\rho_1}{1-\rho_1} > \frac{\rho_2}{1-\rho_2}, \quad (3.2)$$

then either one of the following two cases holds:

(a) $z_\infty = z_0 < \rho_2$ if

$$z_1 > \rho_2 \quad \text{or} \quad \frac{(\lambda_1 + q\mu_2)z_1}{\mu_1\{p + (1-p)z_1\}} \leq \frac{\lambda_1 + q\mu_2 z_1}{\lambda_1 + q\mu_2}; \quad (3.3)$$

(b) $z_\infty = z_1 > z_0$ otherwise.

Remark 3.1. We have the following interpretation for condition (3.1) of Theorem 3.1. Since $\rho_i/(1-\rho_i)$ is the mean number of customers at node i for $i = 1, 2$, (3.1) can be written as

$$p E(L_1) \leq E(L_2), \quad (3.4)$$

where (L_1, L_2) is subject to the stationary distribution π (see (2.2)). Suppose that the congestion at a node is measured by the mean number of customers. Then we can say that, if p times the congestion of node 1 is not greater than the congestion of node 2, the finite truncation approximates the unlimited network well. Here, the factor of p is important, since node 2 is less affected by node 1 if p is small. However, it should be noted that condition (3.4) may not be necessary in order to have $z_\infty = \rho_2$ (see Remark 4.1, below). In Theorem 3.2, we have another condition, (3.3), but this seems to be hard to interpret.

In what follows, we prove these theorems using a series of lemmas, applying Proposition 2.1 for $\mathbf{P} = \mathbf{P}_{(m)}$. We first consider the case in which part (ii) of Proposition 2.1 is applicable for $m = \infty$. For this, we need a slightly stronger condition than (3.1), given by the strict inequality

$$\frac{p\rho_1}{1-\rho_1} < \frac{\rho_2}{1-\rho_2}. \quad (3.5)$$

Lemma 3.2. If $m = \infty$ and (3.5) holds, then the convergence radius of the minimal nonnegative solution $\mathbf{R}$ of (2.7) is ρ_2^{-1} .

Proof. In view of part (ii) of Proposition 2.1, it is sufficient to find positive vectors $\mathbf{x}$ and $\mathbf{y}$ satisfying the conditions (2.8), (2.9), and (2.10) for $z = \rho_2$ when (3.5) holds. We show that suitable vectors are given by

$$x_n = \rho_1^n, \quad y_n = \left(\frac{p + (1-p)\rho_2}{\rho_2} \right)^n, \quad n \geq 0, \quad (3.6)$$

where x_n and y_n are the n th entries of $\mathbf{x}$ and $\mathbf{y}$, respectively. (This convention for entries of vectors is used throughout the paper.) We need only verify (2.8), (2.9), and (2.10) for these vectors but, instead, we will derive them by solving (2.8) and (2.9). This constructive derivation will be useful in subsequent discussions.

Assume that (3.5) holds. For $m = \infty$, (2.8) is written as

$$\begin{aligned} \{\lambda_2 + \mu_1 z + (1 - q)\mu_2 z^2\}x_0 + \mu_1\{p + (1 - p)z\}x_1 &= zx_0, \\ (\lambda_1 + q\mu_2 z)zx_{n-1} + \{\lambda_2 + (1 - q)\mu_2 z^2\}x_n + \mu_1\{p + (1 - p)z\}x_{n+1} &= zx_n, \quad n \geq 1. \end{aligned} \quad (3.7)$$

Consider the quadratic equation $f(s, z) = 0$ for a fixed z , i.e.

$$\mu_1\{p + (1 - p)z\}s^2 + \{\lambda_2 - z + (1 - q)\mu_2 z^2\}s + (\lambda_1 + q\mu_2 z)z = 0, \quad (3.8)$$

and suppose that this equation has two positive roots s_1 and s_2 such that $0 < s_1 < s_2$. Then, $\mathbf{x}$ is given by

$$x_n = C_1 s_1^n + C_2 s_2^n, \quad n = 0, 1, \dots,$$

for some constants C_1 and C_2 , where $x_0 = 1$.

Similarly, (2.9) is written as

$$\begin{aligned} \{\lambda_2 + \mu_1 z + (1 - q)\mu_2 z^2\}y_0 + (\lambda_1 + q\mu_2 z)zy_1 &= zy_0, \\ \mu_1\{p + (1 - p)z\}y_{n-1} + \{\lambda_2 + (1 - q)\mu_2 z^2\}y_n + z(\lambda_1 + q\mu_2 z)y_{n+1} &= zy_n, \quad n \geq 1. \end{aligned}$$

Consider the following quadratic equation in variable t :

$$(\lambda_1 + q\mu_2 z)zt^2 + \{\lambda_2 - z + (1 - q)\mu_2 z^2\}t + \mu_1\{p + (1 - p)z\} = 0.$$

Comparing this equation with (3.8), it clear that $t_1 = 1/s_2$ and $t_2 = 1/s_1$ are its solutions. Hence, $\mathbf{y}$ is given by

$$y_n = D_1 t_1^n + D_2 t_2^n, \quad n = 0, 1, \dots,$$

for some constants D_1 and D_2 , where $y_0 = 1$. Note that, if s_1 and s_2 are complex or identical, it is difficult to find vectors $\mathbf{x}$ and $\mathbf{y}$ satisfying (2.10) because then $x_n y_n$ may not vanish as n goes to infinity. For this reason, we have assumed that s_1 and s_2 are positive and not identical, where positivity of s_1 and s_2 comes from the fact that the vectors $\mathbf{x}$ and $\mathbf{y}$ must be positive. Since

$$s_1 t_2 = 1, \quad s_2 t_1 = 1, \quad s_1 t_1 < 1, \quad s_2 t_2 > 1,$$

and

$$x_n y_n = C_1 D_1 (s_1 t_1)^n + C_1 D_2 (s_1 t_2)^n + C_2 D_1 (s_2 t_1)^n + C_2 D_2 (s_2 t_2)^n,$$

we must have

$$C_2 D_2 = 0 \quad \text{and} \quad C_1 D_2 + C_2 D_1 = 0,$$

meaning that both C_2 and D_2 must vanish for condition (2.10) to hold. From $C_2 = 0$, we have $s_1 = x_1$, and substituting this into (3.7) yields

$$\{\lambda_2 - z + \mu_1 z + (1 - q)\mu_2 z^2\} + \mu_1\{p + (1 - p)z\}s_1 = 0. \quad (3.9)$$

Subtracting (3.9) times s_1 from $f(s_1, z) = 0$, we have

$$s_1 = \frac{\lambda_1 + q\mu_2 z}{\mu_1} \quad (3.10)$$

and, by substituting this back into (3.9), we find the solution $z \in (0, 1)$ to be

$$z = \rho_2.$$

Then, using this value for z in (3.10), and taking into account the fact that s_1 and s_2 are the solutions of the quadratic equation (3.8), we have

$$s_1 = \rho_1, \quad s_2 = \frac{\rho_2}{p + (1-p)\rho_2}. \quad (3.11)$$

It is easy to show that (3.5) is equivalent to $s_1 < s_2$ if s_1 and s_2 are given by (3.11). Thus, the vectors given by (3.6) satisfy conditions (2.8), (2.9), and (2.10) for $z = \rho_2$. Hence, the lemma is proved.

Lemma 3.3. *If (3.5) holds then $z_\infty = \rho_2$.*

Proof. Let $\mathbf{R}_{(m)}$ be the minimum nonnegative solution of (2.7) for $\mathbf{A}_i = \mathbf{A}_{i(m)}$, $i = 1, 0, -1$, i.e. such that

$$\mathbf{R}_{(m)} = \mathbf{A}_{1(m)} + \mathbf{R}_{(m)}\mathbf{A}_{0(m)} + \mathbf{R}_{(m)}^2\mathbf{A}_{-1(m)}.$$

Let $\underline{\mathbf{R}} := \liminf_m \mathbf{R}_{(m)}$ entrywise. By Fatou's lemma, we have

$$\underline{\mathbf{R}} \geq \mathbf{A}_1 + \underline{\mathbf{R}}\mathbf{A}_0 + \underline{\mathbf{R}}^2\mathbf{A}_{-1}.$$

Note that the matrix $\mathbf{R}$ can be obtained as the limit of $\mathbf{X}^{(n)}$, defined by

$$\begin{aligned} \mathbf{X}^{(0)} &:= \mathbf{O}, \\ \mathbf{X}^{(n+1)} &:= \mathbf{A}_1 + \mathbf{X}^{(n)}\mathbf{A}_0 + (\mathbf{X}^{(n)})^2\mathbf{A}_{-1}, \quad n \geq 0. \end{aligned}$$

Since $\underline{\mathbf{R}} \geq \mathbf{X}^{(0)} = \mathbf{O}$, we have, by induction

$$\underline{\mathbf{R}} \geq \mathbf{X}^{(n)}, \quad n \geq 0;$$

letting n go to infinity, we then have

$$\underline{\mathbf{R}} \geq \mathbf{R}. \quad (3.12)$$

Suppose that $\lambda_2 \neq 0$. Then $\mathbf{R}_{(m)}$ is irreducible, by Remark 2.2, since $\mathbf{U}$ is positive. Hence, by the Perron–Frobenius theorem, there exist a positive column vector $\mathbf{y}_{(m)}$ and a positive number $z_{(m)}$ such that $\mathbf{y}_{(m)}\mathbf{0} = 1$ and

$$\mathbf{R}_{(m)}\mathbf{y}_{(m)} = z_{(m)}\mathbf{y}_{(m)}.$$

Let $\underline{\mathbf{y}} := \liminf_{m \rightarrow \infty} \mathbf{y}_{(m)}$. Since $z_{(m)}$ monotonically converges, (3.12) and Fatou's lemma yield

$$\underline{\mathbf{R}}\underline{\mathbf{y}} \leq z_\infty \underline{\mathbf{y}}$$

(where inequalities between nonscalar quantities are componentwise). Note that $\underline{y}$ is a positive vector with finite entries since $y_0 = 1$ and $\mathbf{R}$ is irreducible. By Lemma 3.2, the convergence radius of $\mathbf{R}$ is ρ_2^{-1} . Hence, Theorem 6.3 of [14] implies that $z_\infty^{-1} \leq \rho_2^{-1}$, or

$$z_\infty \geq \rho_2.$$

As the inequality $z_\infty \leq \rho_2$ is implied by Lemma 3.1, we conclude that $z_\infty = \rho_2$.

It remains to consider the case in which $\lambda_2 = 0$. In this case, by Remark 2.2, $\mathbf{R}_{(m)}$ has the form

$$\mathbf{R}_{(m)} = \begin{pmatrix} 0 & \mathbf{0} \\ \mathbf{g} & \mathbf{R}_{(m)}(\mathcal{A}) \end{pmatrix},$$

where $\mathbf{0}$ is a row vector with all entries 0, $\mathbf{g}$ is a positive column vector, and $\mathbf{R}_{(m)}(\mathcal{A})$ is an irreducible matrix whose entries have indices in $\mathcal{A} := \{1, \dots, m\}$. Since we can apply the Perron–Frobenius theorem to $\mathbf{R}_{(m)}(\mathcal{A})$, there exist a positive right-eigenvector $\mathbf{y}_{(m)}(\mathcal{A})$, whose entries have indices in $\mathcal{A}$, and a positive eigenvalue $z_{(m)}$ such that

$$\mathbf{R}_{(m)} \begin{pmatrix} 0 \\ \mathbf{y}_{(m)}(\mathcal{A}) \end{pmatrix} = z_{(m)} \begin{pmatrix} 0 \\ \mathbf{y}_{(m)}(\mathcal{A}) \end{pmatrix}.$$

Let $\underline{\mathbf{y}}(\mathcal{A}) := \liminf_{m \rightarrow \infty} \mathbf{y}_{(m)}(\mathcal{A})$. As for $\underline{\mathbf{y}}$, Fatou's lemma yields

$$\mathbf{R} \begin{pmatrix} 0 \\ \underline{\mathbf{y}}(\mathcal{A}) \end{pmatrix} \leq z_\infty \begin{pmatrix} 0 \\ \underline{\mathbf{y}}(\mathcal{A}) \end{pmatrix},$$

and we have $z_\infty = \rho_2$.

Suppose that m is finite. In this case, we denote the matrix generating function $\mathbf{A}(z)$ by $\mathbf{A}_{(m)}(z)$. Let $\chi(z)$ be the Perron–Frobenius eigenvalue of $\mathbf{A}_{(m)}(z)$. Since $\mathbf{A}_{(m)}(z)$ is irreducible and aperiodic, it has a positive left-eigenvector $\mathbf{x}$ and a positive right-eigenvector $\mathbf{y}$ corresponding to $\chi(z)$, by the Perron–Frobenius theorem. It is known that $\chi(z)$ is a convex function (see, e.g. [7]). From the structure of $\mathbf{A}_{(m)}(z)$, it is clear that $\chi(0) > 0$ and $\chi(1) = 1$. These and the stability condition (2.1) imply that the equation $z = \chi(z)$ has one solution in $(0, 1)$. This solution must be $z_{(m)}$, by Proposition 2.1, since m is finite and conditions (2.8) and (2.9) are satisfied.

We next compute $\mathbf{x} = (x_0, x_1, \dots, x_m)$ in closed form. Equation (2.8), i.e.

$$\mathbf{x} \mathbf{A}_{(m)}(z_{(m)}) = z_{(m)} \mathbf{x},$$

can be written as

$$\{\lambda_2 + \mu_1 z_{(m)} + (1 - q)\mu_2 z_{(m)}^2\}x_0 + \mu_1\{p + (1 - p)z_{(m)}\}x_1 = z_{(m)}x_0, \quad (3.13)$$

$$\begin{aligned} z_{(m)}(\lambda_1 + q\mu_2 z_{(m)})x_{n-1} + \{\lambda_2 + (1 - q)\mu_2 z_{(m)}^2\}x_n \\ + \mu_1\{p + (1 - p)z_{(m)}\}x_{n+1} = z_{(m)}x_n, \\ 1 \leq n \leq m - 1, \end{aligned} \quad (3.14)$$

$$z_{(m)}(\lambda_1 + q\mu_2 z_{(m)})x_{m-1} + \{\lambda_2 + (\lambda_1 + q\mu_2)z_{(m)} + (1 - q)\mu_2 z_{(m)}^2\}x_m = z_{(m)}x_m. \quad (3.15)$$

These equations inductively define the series $\{x_n, 0 \leq n \leq m\}$.

As is well known, the solutions x_n can be obtained by using solutions of the quadratic equation (in r)

$$f(r, z_{(m)}) = 0. \quad (3.16)$$

Denote these (latter) solutions by $r_{1(m)}$ and $r_{2(m)}$ ($|r_{1(m)}| \leq |r_{2(m)}|$), and let $x_0 = 1$. From (3.14) and the quadratic equation (3.16), we have

$$(r_{2(m)} - r_{1(m)})x_n = (r_{2(m)} - x_1)r_{1(m)}^n + (x_1 - r_{1(m)})r_{2(m)}^n, \quad 0 \leq n \leq m-1. \quad (3.17)$$

(See Appendix A for this derivation.) Using (3.16) and the boundary conditions (3.13) and (3.15), lengthy but straightforward computations yield

$$\begin{aligned} & \left(r_{2(m)} - \frac{z_{(m)}}{p + (1-p)z_{(m)}} \right) \{ (\lambda_1 + q\mu_2)r_{1(m)} - (\lambda_1 + q\mu_2z_{(m)}) \} \\ &= \left(r_{1(m)} - \frac{z_{(m)}}{p + (1-p)z_{(m)}} \right) \{ (\lambda_1 + q\mu_2)r_{2(m)} - (\lambda_1 + q\mu_2z_{(m)}) \} \left(\frac{r_{1(m)}}{r_{2(m)}} \right)^{m+1}. \end{aligned} \quad (3.18)$$

(Appendix A also contains this computation.) Note that if $r_{1(m)} = r_{2(m)}$ then (3.18) is just an identity. Thus, if $r_{1(m)} \neq r_{2(m)}$, $z_{(m)}$ is expected to be obtained as a solution of (3.18) when $r_{1(m)}$ and $r_{2(m)}$ satisfy (3.16). Since $z_{(m)}$ monotonically converges to z_∞ , $r_{1(m)}$ and $r_{2(m)}$ can be chosen so that they each converge to a limit as m goes to infinity. Denote these limits by $r_{1\infty}$ and $r_{2\infty}$, respectively. They are obtained as the roots of the quadratic equation (in r)

$$f(r, z_\infty) = 0.$$

Lemma 3.4. *The root z_0 of the discriminant $D(z)$ (see Theorem 3.2) exists, and we have*

$$z_0 \leq z_\infty \leq \rho_2. \quad (3.19)$$

In particular, $z_0 = z_\infty = \rho_2$ only if

$$\frac{p\rho_1}{1-\rho_1} = \frac{\rho_2}{1-\rho_2}. \quad (3.20)$$

Proof. We first prove the existence of $z_0 \in (0, \rho_2]$. For this, decompose $D(z)$ as

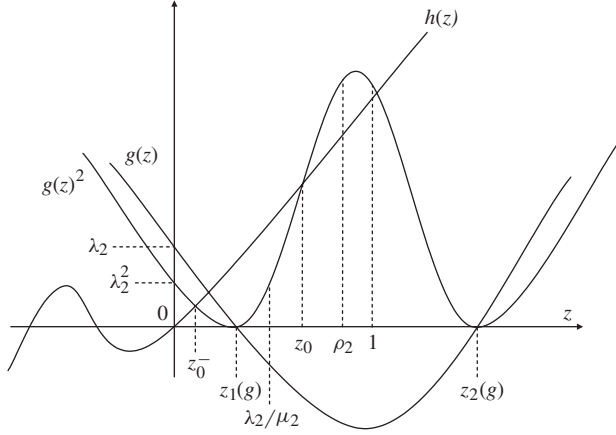
$$D(z) = g(z)^2 - h(z),$$

where

$$\begin{aligned} g(z) &= \lambda_2 - z + (1-q)\mu_2z^2, \\ h(z) &= 4\mu_1z\{p + (1-p)z\}(\lambda_1 + q\mu_2z). \end{aligned}$$

It is easy to show that $g(z) = 0$ has two nonnegative solutions $z_1(g)$ and $z_2(g)$ such that

$$\begin{aligned} 0 \leq z_1(g) &= \frac{1 - \sqrt{1 - 4(1-q)\lambda_2\mu_2}}{2(1-q)\mu_2} \\ &< 1 < z_2(g) = \frac{1 + \sqrt{1 - 4(1-q)\lambda_2\mu_2}}{2(1-q)\mu_2}, \end{aligned}$$

FIGURE 3: The relation between $g(z)$, $g(z)^2$, and $h(z)$.

where, since

$$1 - 4(1 - q)\lambda_2\mu_2 = (\mu_2 - \lambda_2)^2 + 2(\lambda_2 + \mu_2)(\lambda_1 + \mu_1) + (\lambda_1 + \mu_1)^2 + 4q\lambda_2\mu_2 > 0$$

(where we have used (2.4)), the argument of the square root is positive. Obviously, $h(z)$ is increasing in $z \geq 0$ and $h(0) = 0 \leq g(0)^2$. We also note that

$$D(\rho_2) = g(\rho_2)^2 - h(\rho_2) = \left[\mu_1(1 - \rho_1)(1 - \rho_2) \left(\frac{\rho_2}{1 - \rho_2} - \frac{p\rho_1}{1 - \rho_1} \right) \right]^2 \geq 0. \quad (3.21)$$

(The verification of (3.21) requires some complicated computations, so we defer them to Appendix B.) Furthermore, we have

$$D(1) = g(1)^2 - h(1) = (\lambda_1 + q\mu_2 - \mu_1)^2 \geq 0.$$

Hence, we conclude that $D(z) = 0$ has two solutions in $[0, \rho_2]$ and that the larger solution is $z_0 \in (z_1(g), \rho_2]$. For convenience, we denote the smaller solution by z_0^- (see Figure 3).

We are now ready to prove (3.19). Assume that $z_\infty < z_0$. Since node 2 has exogenous arrivals subject to the Poisson process with rate λ_2 , $z_{(m)}$ is greater than λ_2/μ_2 for all $m \geq 1$. Hence, $\lambda_2/\mu_2 < z_\infty$ since $p > 0$. It is easy to show that

$$\begin{aligned} g(0) &= \lambda_2 \geq 0, \\ g\left(\frac{\lambda_2}{\mu_2}\right) &= \frac{\lambda_2}{\mu_2} \{\mu_2 - 1 + (1 - q)\lambda_2\} \\ &= \frac{\lambda_2}{\mu_2} \{-(\lambda_1 + \mu_1) - q\lambda_2\} < 0. \end{aligned}$$

Hence, $g(z) = 0$ has a unique solution $z \in [0, \lambda_2/\mu_2)$, which must be $z_1(g)$. Thus, we find that

$$z_1(g) < \frac{\lambda_2}{\mu_2} < z_\infty,$$

which implies that $D(z_{(m)}) < 0$ for sufficiently large $m \geq 1$, since $z_\infty < z_0$. Assuming that m is such a large integer, (3.16) has complex roots and, so, $r_{1(m)}$ and $r_{2(m)}$ can be expressed as

$$r_{1(m)} = \xi_{(m)} e^{-i\theta_{(m)}}, \quad r_{2(m)} = \xi_{(m)} e^{i\theta_{(m)}},$$

for real numbers $\xi_{(m)}$ and $\theta_{(m)}$, where $i = (-1)^{1/2}$. Applying these values in (3.18) yields

$$\begin{aligned} & \left(r_{2(m)} - \frac{z_{(m)}}{p + (1-p)z_{(m)}} \right) \{(\lambda_1 + q\mu_2)r_{1(m)} - (\lambda_1 + q\mu_2z_{(m)})\} \\ &= \left(r_{1(m)} - \frac{z_{(m)}}{p + (1-p)z_{(m)}} \right) \{(\lambda_1 + q\mu_2)r_{2(m)} - (\lambda_1 + q\mu_2z_{(m)})\} e^{-2i(m+1)\theta_{(m)}}. \end{aligned} \quad (3.22)$$

Since $z_{(m)} \uparrow z_\infty < z_0$, we have

$$\begin{aligned} \lim_{m \rightarrow \infty} r_{1(m)} &= r_{1\infty} = \xi_\infty e^{-i\theta_\infty}, \\ \lim_{m \rightarrow \infty} r_{2(m)} &= r_{2\infty} = \xi_\infty e^{i\theta_\infty}, \end{aligned}$$

where $r_{1\infty}$ and $r_{2\infty}$ are the solutions of $f(r, z_\infty) = 0$. Since $D(z_\infty) < 0$, we have $\xi_\infty, \theta_\infty \neq 0$, which implies that the left-hand side of (3.22) converges while the right-hand side does not, which is a contradiction. Thus, having $z_\infty < z_0$ is impossible, and we obtain (3.19). To complete the proof, we note that, by (3.21), $D(\rho_2) = 0$ only if (3.20) holds. Since $D(\rho_2) = 0$ is equivalent to $z_0 = \rho_2$, the proof is complete.

3.1. Proofs of Theorems 3.1 and 3.2

3.1.1. *Proof of Theorem 3.1.* By Lemma 3.4, (3.20) implies that $z_\infty = \rho_2$. Combining this with Lemma 3.3 gives Theorem 3.1.

3.1.2. *Proof of Theorem 3.2.* We first prove case (a), by contradiction. Suppose that $z_\infty \neq z_0$, which is equivalent to the inequalities $z_0 < z_\infty \leq \rho_2$, by Lemma 3.4. Then, $D(z_\infty) > 0$ and, so, $r_{1\infty} < r_{2\infty}$. If we let m go to infinity in (3.18), this implies that

$$\left(r_{2\infty} - \frac{z_\infty}{p + (1-p)z_\infty} \right) \{(\lambda_1 + q\mu_2)r_{1\infty} - (\lambda_1 + q\mu_2z_\infty)\} = 0.$$

Hence, either

$$r_{1\infty} = \frac{\lambda_1 + q\mu_2z_\infty}{\lambda_1 + q\mu_2} \quad (3.23)$$

or

$$r_{2\infty} = \frac{z_\infty}{p + (1-p)z_\infty}. \quad (3.24)$$

Suppose that (3.23) holds. Then, since $r_{1\infty}$ and $r_{2\infty}$ are the solutions of $f(r, z_\infty) = 0$, we have

$$r_{1\infty} + r_{2\infty} = \frac{z_\infty - \lambda_2 - (1-q)\mu_2z_\infty^2}{\mu_1\{p + (1-p)z_\infty\}}, \quad (3.25)$$

$$r_{1\infty}r_{2\infty} = \frac{(\lambda_1 + q\mu_2z_\infty)z_\infty}{\mu_1\{p + (1-p)z_\infty\}}, \quad (3.26)$$

and (3.26) and (3.23) imply that

$$r_{2\infty} = \frac{(\lambda_1 + q\mu_2)z_\infty}{\mu_1\{p + (1-p)z_\infty\}}. \quad (3.27)$$

Substituting this into (3.25) yields

$$z_\infty = z_1.$$

(See Appendix C for the details of this computation.) Hence, neither statement in (3.3) is true, since $r_{1\infty} < r_{2\infty}$ and $z_\infty \leq \rho_2$. On the other hand, if (3.24) holds, then we have

$$z_\infty = \rho_2 \quad \text{and} \quad r_{1\infty} = \rho_1.$$

These contradict (3.2), since $r_{1\infty} < r_{2\infty}$, and, so, both statements in (3.3) are again false. Hence, we have proved part (a) of Theorem 3.2.

We next prove part (b). Suppose that (3.3) does not hold. If we denote the two solutions of $f(r, z_1) = 0$ by r_{1*} and r_{2*} , it is easy to show that they are given by

$$r_{1*} = \frac{\lambda_1 + q\mu_2 z_1}{\lambda_1 + q\mu_2}, \quad r_{2*} = \frac{(\lambda_1 + q\mu_2)z_1}{\mu_1\{p + (1-p)z_1\}},$$

which must differ from one another, since (3.3) is false. We will first show that

$$z_0 < z_1. \quad (3.28)$$

Note that $D(z_1) > 0$ and $z_1 \leq \rho_2$. From the proof of Lemma 3.4, the equation $D(z) = 0$ has two solutions z_0^- and z_0 in $[0, \rho_2]$. If $\lambda_2 = 0$ then $z_0^- = 0$ and, so, $D(z_1) > 0$ implies (3.28). If $\lambda_2 > 0$ then $D(z_1) > 0$ is equivalent to either $0 < z_1 < z_0^-$ or $z_0 < z_1 \leq \rho_2$. Suppose that $0 < z_1 < z_0^-$. Then $g(z_1) > 0$ since $g(z) > 0$ for $0 < z < z_0^-$ (see Figure 3). However, this implies that

$$r_{1*} + r_{2*} = -\frac{g(z_1)}{\mu_1\{p + (1-p)z_1\}} < 0,$$

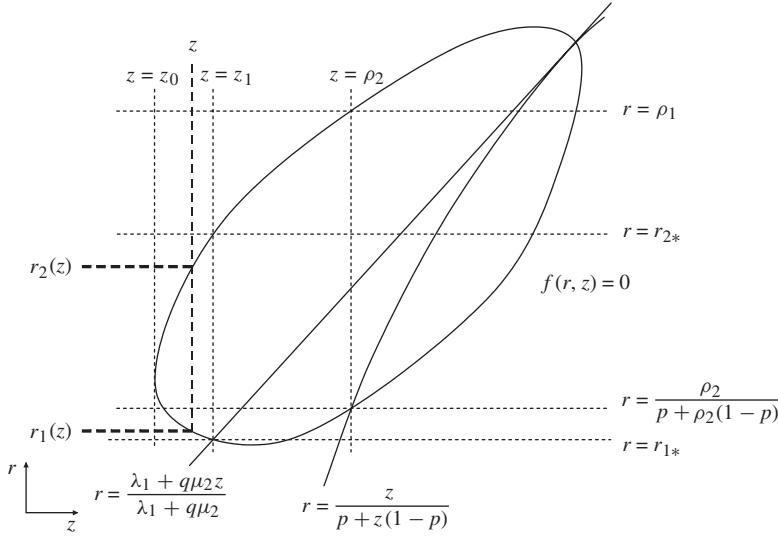
since r_{1*} and r_{2*} are the solutions of the quadratic equation $f(r, z_1) = 0$. This contradicts the fact that r_{1*} and r_{2*} are positive and, so, we conclude that (3.28) holds.

Now suppose that $z_\infty > z_0$. We then have either (3.23) or (3.24) again. However, (3.24) contradicts (3.2). Therefore, we must have (3.23), from which part (b) follows. To prove the theorem, we must then show that $z_\infty > z_0$ indeed holds. To do so, it is sufficient to find a positive integer m , for any $\varepsilon > 0$, such that $r_{1(m)}$, $r_{2(m)}$, and $z_{(m)}$ satisfy (3.16) and (3.18), and that $|z_1 - z_{(m)}| < \varepsilon$. By the assumption that (3.3) does not hold, we have $r_{1*} < r_{2*}$. Note that $z_0 < z_1$ holds, by (3.28), and that (3.16) has two positive solutions r_{1*} and r_{2*} for $z_{(m)} = z_1$. Suppose that

$$r_{2*} = \frac{z_1}{p + (1-p)z_1}.$$

Then we have

$$z_1 = \rho_2 \quad \text{and} \quad r_{1*} = \rho_1.$$

FIGURE 4: The solution z of (3.29).

These contradict (3.2), since $r_{1*} < r_{2*}$, implying that

$$r_{2*} \neq \frac{z_1}{p + (1-p)z_1}.$$

Let z be a positive number such that $|z_1 - z| < \varepsilon$, and let $r_1(z)$ and $r_2(z)$ be the solutions of $f(r, z) = 0$ such that $r_1(z) < r_2(z)$. These solutions will be positive, by the continuity of the quadratic solution, if ε is sufficiently small. We further take ε to be small enough that, for some $\eta > 0$,

$$r_2(z) \neq \frac{z}{p + (1-p)z}, \quad \frac{r_1(z)}{r_2(z)} < 1 - \eta,$$

for all z satisfying $|z_1 - z| < \varepsilon$. We now consider (3.18) with $z_{(m)} = z$, i.e.

$$\begin{aligned} (\lambda_1 + q\mu_2)r_1(z) &= (\lambda_1 + q\mu_2z) \\ &+ \frac{r_1(z) - z/(p + (1-p)z)}{r_2(z) - z/(p + (1-p)z)} \{(\lambda_1 + q\mu_2)r_2(z) - (\lambda_1 + q\mu_2z)\} \left(\frac{r_1(z)}{r_2(z)} \right)^{m+1}. \end{aligned} \quad (3.29)$$

The final term on the right-hand side of this equation becomes arbitrarily small when m becomes large. For m large enough, we can then find a solution z of (3.29) that satisfies $|z_1 - z| < \varepsilon$ (see, e.g. Figure 4). This gives $z_{(m)}$, and we obtain part (b).

4. Special cases and numerical examples

In this section, we consider special cases of Theorems 3.1 and 3.2 and present some numerical examples. The condition (3.3) of Theorem 3.2 is complicated and not intuitive. Hence, we first consider some cases in which this condition is not needed, i.e. part (a) holds automatically

(and, therefore, part (b) is not required). To ensure this, we assume that $q = 0$, i.e. that there is no feedback route. In this case, we have

$$\begin{aligned}
 C(z_1 - \rho_2) &= (1 - pq)\{p\lambda_1\mu_1 + (\lambda_1 + q\mu_2)\lambda_2\} \\
 &\quad - (p\lambda_1 + \lambda_2)\{(1 - p)q\mu_1 + (1 - q)(\lambda_1 + q\mu_2)\} \\
 &= \{(1 - pq)p\lambda_1 - (1 - p)q(p\lambda_1 + \lambda_2)\}\mu_1 \\
 &\quad + (\lambda_1 + q\mu_2)\{(1 - pq)\lambda_2 - (1 - q)(p\lambda_1 + \lambda_2)\} \\
 &= \{p(1 - q)\lambda_1 - (1 - p)q\lambda_2\}\mu_1 + (\lambda_1 + q\mu_2)\{(1 - p)q\lambda_2 - (1 - q)p\lambda_1\} \\
 &= \{(1 - q)p\lambda_1 - (1 - p)q\lambda_2\}\{\mu_1 - (\lambda_1 + q\mu_2)\}, \tag{4.1}
 \end{aligned}$$

where $C = (1 - pq)\mu_2\{(1 - p)q\mu_1 + (1 - q)(\lambda_1 + q\mu_2)\} > 0$. In particular, if $q = 0$ then

$$C(z_1 - \rho_2) = p\lambda_1(\mu_1 - \lambda_1) > 0,$$

by the stability condition for node 1. The following results are immediate consequences of these computations and Theorem 3.2.

Corollary 4.1. *If $q = 0$, condition (3.2) of Theorem 3.2 implies that $z_\infty = z_0$.*

In the tandem Jackson network of [8], $p = 1$ and $q = \lambda_2 = 0$. Furthermore, (3.1) is equivalent to $\rho_1 \leq \rho_2$ for $p = 1$. So, Corollary 4.1 shows that the decay rate results of [8] are special cases of Theorems 3.1 and 3.2.

Now, if $0 < q < 1$, then case (b) of Theorem 3.2 may be relevant. For example, let us consider the case in which $p = 1$. If we let $\alpha = \lambda_1 + q\mu_2$ then

$$\begin{aligned}
 z_1 &= \frac{\lambda_1\mu_1 + \alpha\lambda_2}{(1 - q)\alpha\mu_2} \\
 &= \rho_2 + \frac{\lambda_1(\mu_1 - \alpha)}{(1 - q)\alpha\mu_2}.
 \end{aligned}$$

Using this, r_{1*} , and r_{2*} in the proof of Theorem 3.2, we compute that

$$\begin{aligned}
 r_{2*} - r_{1*} &= \frac{\alpha}{\mu_1}z_1 - \frac{\lambda_1 + q\mu_2z_1}{\alpha} \\
 &= \frac{\alpha}{\mu_1}z_1 - 1 + \frac{q\mu_2}{\alpha}(1 - z_1) \\
 &= \frac{q\mu_2}{\alpha}(1 - \rho_2) + \frac{\lambda_1\mu_1 + \alpha\lambda_2}{(1 - q)\mu_1\mu_2} - 1 - \frac{q\lambda_1(\mu_1 - \alpha)}{(1 - q)\alpha^2} \\
 &= \frac{q\mu_2}{\alpha}(1 - \rho_2) + \frac{\alpha\lambda_2 + \mu_1(\alpha - \mu_2)}{(1 - q)\mu_1\mu_2} + \frac{q\lambda_1(\alpha - \mu_1)}{(1 - q)\alpha^2}.
 \end{aligned}$$

Note that $\alpha - \mu_2 = \lambda_1 - (1 - q)\mu_2 < 0$, by the stability of node 2. Also note that having $\alpha \geq \mu_1$ is equivalent to having $z_1 \leq \rho_2$. However, it is not possible to have $\alpha = \mu_1$ in case (b) of Theorem 3.2 since, if this relation were to hold, the right-hand side of the above equation

would become

$$\begin{aligned}
 & \frac{q\mu_2(\mu_2 - (\mu_1 + \lambda_2))}{(1-q)\mu_1\mu_2} + \frac{\alpha\lambda_2 + \mu_1(\alpha - \mu_2)}{(1-q)\mu_1\mu_2} \\
 &= \frac{q\mu_2(\mu_2 - (\mu_1 + \lambda_2))}{(1-q)\mu_1\mu_2} - \frac{\mu_1(\mu_2 - (\mu_1 + \lambda_2))}{(1-q)\mu_1\mu_2} \\
 &= \frac{\lambda_1(\mu_1 + \lambda_2 - \mu_2)}{(1-q)\mu_1\mu_2} \\
 &= \frac{\lambda_1(\lambda_1 + \lambda_2 - (1-q)\mu_2)}{(1-q)\mu_1\mu_2} \\
 &< 0
 \end{aligned} \tag{4.2}$$

(where we have set $\alpha = \mu_1 = \lambda_1 + q\mu_2$). Thus, we have the following result.

Corollary 4.2. *Suppose that $p = 1$ and $1 > \rho_1 > \rho_2$. Then case (b) of Theorem 3.2 holds if and only if $\alpha > \alpha_0 > \mu_1$, where α_0 is the solution of the equation*

$$\frac{q}{(1-q)x} \{(1-q)\mu_2(1-\rho_2) + \lambda_1\} = \frac{q\lambda_1\mu_1}{(1-q)x^2} + \frac{\mu_1\mu_2 - x(\mu_1 + \lambda_2)}{(1-q)\mu_1\mu_2}, \quad x > \mu_1. \tag{4.3}$$

Proof. Denote the left- and right-hand sides of (4.3) by $\phi(x)$ and $\psi(x)$, respectively. Both of these functions are convex and decreasing, and $\phi(\mu_1) < \psi(\mu_1)$ by (4.2). Since $\phi(x)$ is nonnegative but $\psi(x) \rightarrow -\infty$ as $x \rightarrow \infty$, we can find a unique solution of (4.3). Furthermore, $\phi(x) > \psi(x)$ for $x \geq \mu_1$ if and only if $x > \alpha_0$. Hence, $\alpha > \alpha_0$ implies that $r_{1*} < r_{2*}$. This, together with the fact that $\alpha > \mu_1$, implies that all the conditions on case (b) hold.

Remark 4.1. For $0 < p < 1$, it may also be possible to have $z_1 = \rho_2$. For this, we assume that $\alpha = \mu_1$, which implies that $z_1 = \rho_2$, by (4.1). Hence, we have $z_\infty = \rho_2$ if $r_{1*} < r_{2*}$. The latter condition holds if and only if

$$\begin{aligned}
 r_{1*} &= \frac{\lambda_1 + q\mu_2 z_1}{\alpha} \\
 &< \frac{z_1}{p + (1-p)z_1} = r_{2*},
 \end{aligned}$$

which is equivalent to

$$\begin{aligned}
 & \frac{1}{\alpha\{p + (1-p)z_1\}} [\alpha z_1 - \{p(1-z_1) + z_1\}\{\alpha - q\mu_2(1-z_1)\}] \\
 &= \frac{1-z_1}{\alpha(p + (1-p)z_1)} \{q\mu_2(1-p)z_1 - p\lambda_1\} > 0. \tag{4.4}
 \end{aligned}$$

Thus, if

$$p(1-q)\lambda_1 < (1-p)q\lambda_2$$

then we have $z_\infty = \rho_2$ in case (b) of Theorem 3.2. Note that $p < 1$, $q > 0$, and $\lambda_2 > 0$ are necessary conditions for this to hold.

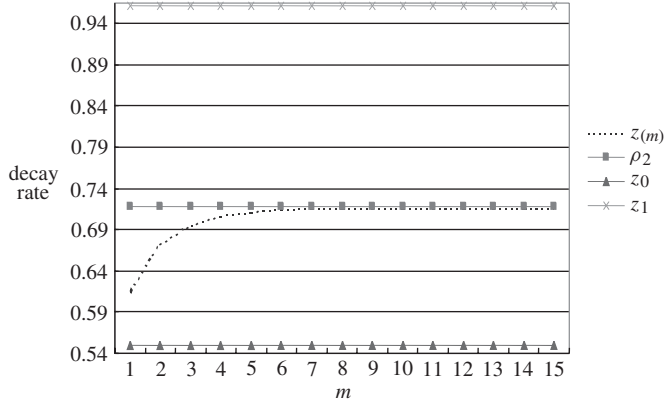


FIGURE 5: Example 1 (of Theorem 3.1).

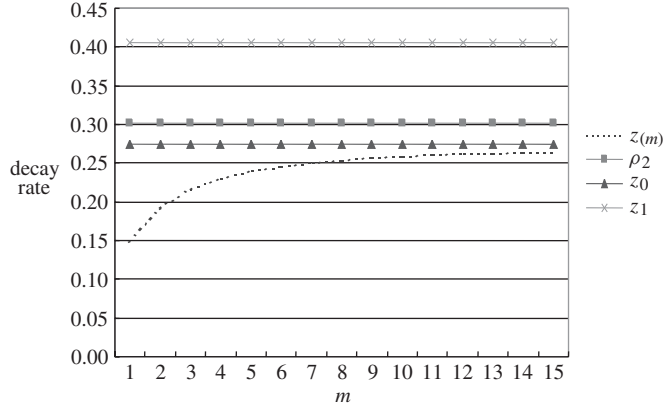


FIGURE 6: Example 2 (case (a) of Theorem 3.2).

In the rest of this section, we use numerical examples to illustrate how the decay rates converge in the three cases in Theorems 3.1 and 3.2. We consider the following examples:

1. $p = 0.85$, $q = 0.5$, $\mu_1 : \mu_2 : \lambda_1 : \lambda_2 = 3 : 2 : 0.5 : 0.4$;
2. $p = 0.9$, $q = 0.1$, $\mu_1 : \mu_2 : \lambda_1 : \lambda_2 = 2 : 4 : 1 : 0.2$;
3. $p = 0.85$, $q = 0.7$, $\mu_1 : \mu_2 : \lambda_1 : \lambda_2 = 2 : 3 : 0.5 : 0.4$.

Example 1 is the case of Theorem 3.1; example 2 is case (a) of Theorem 3.2, where the two conditions in (3.3) hold; and example 3 is case (b) of Theorem 3.2. In each case, we first compute the matrix $R_{(m)}$ as the minimal nonnegative solution of (2.7), and then obtain the Perron–Frobenius eigenvalues. These are the decay rates, and are plotted in Figures 5, 6, and 7 for examples 1, 2, and 3, respectively.

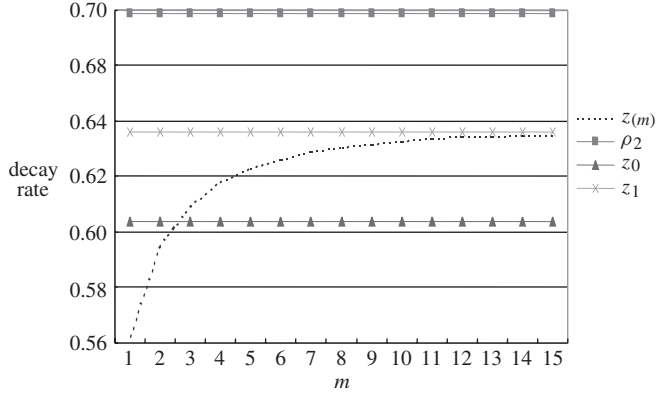


FIGURE 7: Example 3 (case (b) of Theorem 3.2).

Appendix A. The proofs of (3.17) and (3.18)

For convenience, define a , b , and c as

$$\begin{aligned} a &= \mu_1\{p + (1-p)z(m)\}, \\ b &= \lambda_2 - z(m) + (1-q)\mu_2 z(m)^2, \\ c &= z(m)(\lambda_1 + q\mu_2 z(m)). \end{aligned}$$

Note that $r_{1(m)}$ and $r_{2(m)}$ ($|r_{1(m)}| \leq |r_{2(m)}|$) are the two solutions of the quadratic equation (3.16), i.e.

$$ar^2 + br + c = 0. \quad (\text{A.1})$$

Since $x_0 = 1$, from (3.14) we have

$$\begin{aligned} x_{n+1} - r_{1(m)}x_n &= r_{2(m)}^n(x_1 - r_{1(m)}), \\ x_{n+1} - r_{2(m)}x_n &= r_{1(m)}^n(x_1 - r_{2(m)}), \end{aligned}$$

by iteration. Subtracting both sides yields (3.17).

Equation (3.15) is equivalent to

$$cx_{m-1} + \{b + (\lambda_1 + q\mu_2)z(m)\}x_m = 0 \quad (\text{A.2})$$

and, by setting $n = m - 1$ in (3.14), we have

$$ax_m + bx_{m-1} + cx_{m-2} = 0. \quad (\text{A.3})$$

The elimination of x_m from (A.2) and (A.3) yields

$$\frac{c}{a}x_{m-1} + \left(\frac{b}{a} + \frac{(\lambda_1 + q\mu_2)z(m)}{a}\right)\left(-\frac{b}{a}x_{m-1} - \frac{c}{a}x_{m-2}\right) = 0. \quad (\text{A.4})$$

Since the constants a , b , and c are the coefficients of the quadratic equation (A.1), by multiplying both sides of (A.4) by $(r_{2(m)} - r_{1(m)})$ and by using (3.17), we have

$$\begin{aligned} r_{1(m)}r_{2(m)}\{(r_{2(m)} - x_1)r_{1(m)}^{m-1} + (x_1 - r_{1(m)})r_{2(m)}^{m-1}\} + \left\{-(r_{1(m)} + r_{2(m)}) + \frac{(\lambda_1 + q\mu_2)z(m)}{a}\right\} \\ \times [(r_{1(m)} + r_{2(m)})\{(r_{2(m)} - x_1)r_{1(m)}^{m-1} + (x_1 - r_{1(m)})r_{2(m)}^{m-1}\} \\ - r_{1(m)}r_{2(m)}\{(r_{2(m)} - x_1)r_{1(m)}^{m-2} + (x_1 - r_{1(m)})r_{2(m)}^{m-2}\}] = 0. \end{aligned} \quad (\text{A.5})$$

By (3.13), we also have

$$x_1 = \frac{-b - \mu_1 z(m)}{a};$$

substituting this into (A.5) yields

$$\begin{aligned} r_{2(m)}\left(r_{2(m)} + \frac{b + \mu_1 z(m)}{a}\right)r_{1(m)}^m + r_{1(m)}\left(\frac{-b - \mu_1 z(m)}{a} - r_{1(m)}\right)r_{2(m)}^m \\ + \left\{-(r_{1(m)} + r_{2(m)}) + \frac{(\lambda_1 + q\mu_2)z(m)}{a}\right\}\left(r_{2(m)} + \frac{b + \mu_1 z(m)}{a}\right)r_{1(m)}^m \\ + \left\{-(r_{1(m)} + r_{2(m)}) + \frac{(\lambda_1 + q\mu_2)z(m)}{a}\right\}\left(\frac{-b - \mu_1 z(m)}{a} - r_{1(m)}\right)r_{2(m)}^m = 0. \end{aligned}$$

Rearranging the above terms, we have

$$\begin{aligned} \left(-r_{1(m)} + \frac{(\lambda_1 + q\mu_2)z(m)}{a}\right)\left(-r_{1(m)} + \frac{\mu_1 z(m)}{a}\right)r_{1(m)}^m \\ + \left(-r_{2(m)} + \frac{(\lambda_1 + q\mu_2)z(m)}{a}\right)\left(r_{2(m)} - \frac{\mu_1 z(m)}{a}\right)r_{2(m)}^m = 0. \end{aligned} \quad (\text{A.6})$$

Multiplying both sides by both $r_{1(m)}r_{2(m)}$ and $a = \mu_1\{p + (1-p)z(m)\}$, we then have

$$\begin{aligned} \left(-r_{1(m)} + \frac{z(m)}{p + (1-p)z(m)}\right)\{-\lambda_1 - q\mu_2 z(m) + (\lambda_1 + q\mu_2)r_{2(m)}\}r_{1(m)}^{m+1} \\ + \left(r_{2(m)} - \frac{z(m)}{p + (1-p)z(m)}\right)\{-\lambda_1 - q\mu_2 z(m) + (\lambda_1 + q\mu_2)r_{1(m)}\}r_{2(m)}^{m+1} = 0, \end{aligned} \quad (\text{A.7})$$

which is equivalent to (3.18).

Appendix B. The proof of (3.21)

From the definitions of g and ρ_2 , we have

$$\begin{aligned} g(\rho_2)^2 &= \{\lambda_2 - \rho_2 + (1-q)\mu_2\rho_2^2\}^2 \\ &= \left(-\lambda_2 + \rho_2 - (1-q)\rho_2\frac{p\lambda_1 + \lambda_2}{1-pq}\right)^2 \\ &= \left((\lambda_1 + \lambda_2 + \mu_1 + \mu_2)\rho_2 - \lambda_2 - (1-q)\rho_2\frac{p\lambda_1 + \lambda_2}{1-pq}\right)^2 \\ &= \frac{1}{(1-pq)^2}\{(1-pq)(\lambda_1 + \lambda_2 + \mu_1 + \mu_2)\rho_2 \\ &\quad - (1-pq)\lambda_2 - (1-q)\rho_2(p\lambda_1 + \lambda_2)\}^2, \end{aligned}$$

where the third equality follows from the fact that $\lambda_1 + \lambda_2 + \mu_1 + \mu_2 = 1$. Since

$$\begin{aligned} & (1 - pq)\rho_2(\lambda_1 + \lambda_2 + \mu_2) - (1 - q)\rho_2(\lambda_2 + p\lambda_1) - (1 - pq)\lambda_2 \\ &= \rho_2\{(1 - pq)\mu_2 + (1 - p)(\lambda_1 + q\lambda_2)\} - (1 - pq)\lambda_2 \\ &= (\lambda_1 + q\lambda_2)\{p + \rho_2(1 - p)\}, \end{aligned}$$

we obtain

$$g(\rho_2)^2 = \frac{1}{(1 - pq)^2} \{(1 - pq)\mu_1\rho_2 + (p + \rho_2(1 - p))(\lambda_1 + q\lambda_2)\}^2.$$

On the other hand, using the fact that

$$\lambda_1 + q\mu_2\rho_2 = \frac{\lambda_1 + q\lambda_2}{1 - pq},$$

we have

$$h(\rho_2) = \frac{4}{1 - pq} \mu_1\rho_2\{p + \rho_2(1 - p)\}(\lambda_1 + q\lambda_2).$$

Hence, we arrive at

$$\begin{aligned} g(\rho_2)^2 - h(\rho_2) &= \frac{1}{(1 - pq)^2} \{(1 - pq)\mu_1\rho_2 - (p + \rho_2(1 - p))(\lambda_1 + q\lambda_2)\}^2 \\ &= \{\mu_1\rho_2 - (p + \rho_2(1 - p))\mu_1\rho_1\}^2 \\ &= \left(\mu_1(1 - \rho_1)(1 - \rho_2) \left(\frac{\rho_2}{1 - \rho_2} - \frac{p\rho_1}{1 - \rho_1} \right) \right)^2. \end{aligned}$$

Appendix C. The proof that $z_\infty = z_1$

Substituting (3.23) and (3.27) into (3.25), we have

$$\begin{aligned} & \{(1 - p)q\mu_1\mu_2 + (1 - q)\mu_2(\lambda_1 + q\mu_2)\}z_\infty^2 \\ &+ [\mu_1\{(1 - p)\lambda_1 + pq\mu_2\} + (\lambda_1 + q\mu_2)^2 - (\lambda_1 + q\mu_2)]z_\infty \\ &+ p\lambda_1\mu_1 + \lambda_2(\lambda_1 + q\mu_2) = 0. \end{aligned}$$

It is clear that $z_\infty = 1$ satisfies this equation. Since the equation is quadratic in z_∞ , we have

$$z_\infty = \frac{p\lambda_1\mu_1 + \lambda_2(\lambda_1 + q\mu_2)}{\{(1 - p)q\mu_1 + (1 - q)(\lambda_1 + q\mu_2)\}\mu_2}$$

for the other solution. Hence, $z_\infty = z_1$.

Acknowledgements

The authors are grateful to an anonymous referee for his comments and detailed suggestions, which helped to improve the presentation of the paper. This research was supported in part by the JSPS under grant no. 13680532.

References

- [1] ALSMEYER, G. (1994). On the Markov renewal theorem. *Stoch. Process. Appl.* **50**, 37–56.
- [2] ÇINLAR, E. (1975). *Introduction to Stochastic Processes*. Prentice-Hall, Englewood Cliffs, NJ.
- [3] FUJIMOTO, K., TAKAHASHI, Y. AND MAKIMOTO, N. (1998). Asymptotic properties of stationary distributions in two-stage queueing systems. *J. Operat. Res. Soc. Japan* **41**, 118–141.
- [4] FUJIMOTO, K., TAKAHASHI, Y. AND MAKIMOTO, N. (2001). Geometric decay of the steady-state probabilities in a quasi-birth–death process with a countable number of phases. *Stoch. Models* **17**, 1–24.
- [5] HÖGLUND, T. (1991). The ruin problem for finite Markov chains. *Ann. Prob.* **19**, 1298–1310.
- [6] JACKSON, J. R. (1957). Networks of waiting lines. *Operat. Res.* **5**, 518–521.
- [7] KINGMAN, J. F. C. (1961). A convexity property of positive matrices. *Quart. J. Math. Oxford* **12**, 283–284.
- [8] KROESE, D. P., SCHEINHARDT, W. R. W. AND TAYLOR, P. G. (2004). Spectral properties of the tandem Jackson network, seen as a quasi-birth-and-death process. *Ann. Appl. Prob.* **14**, 2057–2089.
- [9] McDONALD, D. R. (1999). Asymptotics of first passage times for random walk in an orthant. *Ann. Appl. Prob.* **9**, 110–145.
- [10] MIYAZAWA, M. (2003). Conjectures on decay rates of tail probabilities in generalized Jackson and batch movement networks. *J. Operat. Res. Soc. Japan* **46**, 74–98.
- [11] MIYAZAWA, M. (2004). The Markov renewal approach to M/G/1 type queues with countably many background states. *Queueing Systems* **46**, 177–196.
- [12] MIYAZAWA, M. AND ZHAO, Y. Q. (2004). The stationary tail asymptotics in the GI/G/1-type queue with countably many background states. *Adv. Appl. Prob.* **36**, 1231–1251.
- [13] NEUTS, M. F. (1981). *Matrix-Geometric Solutions in Stochastic Models* (Johns Hopkins Ser. Math. Sci. **2**). Johns Hopkins University Press, Baltimore, MD.
- [14] SENETA, E. (1981). *Nonnegative Matrices and Markov Chains*. Springer, New York.
- [15] SHURENKOV, V. M. (1984). On the theory of Markov renewal. *Theory Prob. Appl.* **29**, 247–265.