Journal of the Operations Research
Society of Japan
2004, Vol. 47, No. 4, 314-338

UPPER BOUND FOR THE DECAY RATE OF THE MARGINAL QUEUE-LENGTH DISTRIBUTION IN A TWO-NODE MARKOVIAN QUEUEING SYSTEM

Ken'ichi Katou
The University of Electro-Communications

Naoki Makimoto
The University of Tsukuba

Yukio Takahashi
Tokyo Institute of Technology

(Received September 30, 2003; Revised June 14, 2004)

Abstract We study a geometric decay property for two-node queueing networks, not restricted to ones having acyclic configuration. We take a matrix-analytic approach, and prove the geometric decay property of the marginal queue-length distributions by giving an upper bound of the exact decay rate for each node. The upper bound coincides with the exact decay rate for Jackson networks and MAP/M/1→/M/1 tandem queues.

Keywords: Queue, matrix-analytic approach, marginal queue length, decay rate, tail probability

1. Introduction

This paper studies geometric decay of marginal queue-length distributions in a two-node Markovian queueing system.

For single queues in a broad class, it is well known that the stationary queue-length distribution $\{p(n)\}$ has a tail decaying geometrically, that is, there exist positive constants $\eta^* < 1$ and $C < \infty$ such that

$$\lim_{n\to\infty} \frac{p(n)}{(\eta^*)^n} = C. \tag{1.1}$$

Many authors proved this property for various queueing models. For example, Kendall [7] proved it for GI/M/1 queues, Takahashi [16] for PH/PH/c queues, Neuts and Takahashi [13] for GI/PH/c queues with heterogeneous servers, and Falkenberg [3] for M/G/1-type queues. Further, Glynn and Whitt [6] and others proved a weaker property $\lim_{n\to\infty} n^{-1}\log p(n) = \log \eta^*$ using the large deviation principle for a wide class of queues.

In a single queue, the *decay rate* η^* is related to the Laplace-Stieltjes transforms $f(x)$ and $g(x)$ of the interarrival and service time distributions. For a PH/PH/1 queue, it was shown in [16] that η^* is given by $f(x^*)$, where x^* is a unique positive root of the equation

$$f(x^*)g(-x^*) = 1. \tag{1.2}$$

For the marginal queue-length distributions in queueing networks, several authors proved the geometric decay property in some special models. Chang [2], Ganesh and Anantharam [4] and Bertsimas et al. [1] among others proved the geometric decay property in a weak sense in some tandem queues or in some tree-type queueing networks. These studies mostly based on the acyclic or feedforward configurations of the networks and their proofs exploited

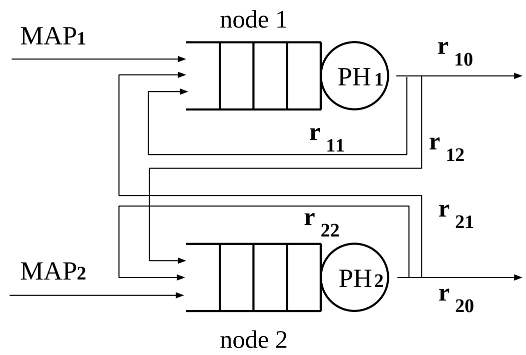

Figure 2.1: Two-node Markovian queueing system

the fact that, for example, the behavior of the first node is not affected by the second one. Miyazawa [11] discussed the decay rate of the marginal queue length distribution in generalized Jackson networks. His conjecture was derived in part from our prework result, an extension of it is reported here.

This paper studies the geometric decay property in two-node networks, not restricted to those having acyclic configuration. Our model, which will be called a two-node Markovian queueing system, is a two-node open queueing network having single servers, buffers with infinite capacity, MAP inputs, PH service distributions and random routings with arbitrary configuration. We take a matrix-analytic approach and by using a lemma for a quasi-birth-and-death process with infinite number of states in each level we prove the geometric decay property of the marginal queue-length distributions by giving an upper bound of the exact decay rate for each node. Our upper bound coincides with the exact decay rate in Jackson networks and MAP/M/1→/M/1 tandem queues studied in [4].

The rest of the paper is organized as follows. We introduce our model and some notation in Section 2. The stochastic behavior of the model is represented by a two-dimensional Markov chain. In Section 3, we discuss a doubly geometric form solution to the balance equations of the Markov chain. Its solution plays an important role for obtaining our upper bound. In Section 4, we state our main theorem. In section 5, we introduce a lemma given in Makimoto et al. [9] for geometric decay of level probabilities in a quasi-birth-and-death process having infinite number of states in each level. The lemma is our main tool to derive our upper bound. The main theorem is proved in Sections 6 and 7, and the propositions in Section 4 are proved in Section 8.

2. Model Description and Notations

In this section we introduce our model and some notation.

Model: We consider an open queueing network with two nodes, node 1 and node 2 (Figure 2.1). At node k ($k=1,2$), customers arrive from outside of the system via a Markovian arrival process MAP$_k$ with representation $(\boldsymbol{T}_k, \boldsymbol{U}_k)$ [8]. There is a single server and a buffer of infinite capacity. Customers are served in a usual FCFS (First Come First Served) manner. Service times are subject to a common phase-type distribution PH$_k$ with representation $(\boldsymbol{b}_k, \boldsymbol{S}_k)$ [12]. After being served, customers proceed to node j ($j=1,2$) with probability r_{kj}, and leave the system with probability $r_{k0} = 1 - r_{k1} - r_{k2}$. Exogenous arrival processes, service times and routing are all stochastically independent. We will refer this model as a *two-node Markovian queueing system.*

To include the tandem queue case, we allow node 2 to have no exogenous arrivals.

Assume that node 1 does have exogenous arrivals (later this will be represented as $\lambda_1 > 0$) and $r_{12} > 0$. The latter assumption causes that nodes are dependent on each other.

Node representation: For brevity of exposition, in this paper, we use the following convention for representing node numbers. Symbol "k" is used to refer the node number of node k. It stands for 1 or 2. If a "k" appears in equations, inequalities, definitions, theorems, etc., then these should be understood for $k = 1$ *and* 2 unless stated otherwise. If symbol "k'" is used with a "k", then it refers the other node number, namely $k' = 2$ if $k = 1$, and $k' = 1$ if $k = 2$.

Vector and matrix notation: Row vectors are represented by bold lower case letters (except for the Markov chain $\boldsymbol{X}(t)$ representing the system behavior). To represent a column vector we attach a superscript $\top$ to the corresponding row vector. We denote by $\mathbf{0}$ a row zero vector and by $\boldsymbol{e}$ a row vector with all elements equal to 1. Matrices are represented with bold upper case letters. We denote by $\boldsymbol{O}$ a zero matrix and by $\boldsymbol{I}$ an identity matrix. Dimensions of vectors and matrices should be understood from the context. Vectors and matrices $\mathbf{0}$, $\boldsymbol{e}$, $\boldsymbol{O}$ and $\boldsymbol{I}$ are used for both cases with finite dimension and infinite dimension. Inequalities between vectors or matrices are considered elementwise. Limits of sequences of matrices or vectors are also considered elementwise. For a vector $\boldsymbol{x}$, we denote by diag$[\boldsymbol{x}]$ a diagonal matrix having i-th element of $\boldsymbol{x}$ in its i-th diagonal element.

We extend our use of terminology "Perron-Frobenius eigenvalue" to an eigenvalue of a finite-dimensional square matrix having nonnegative off-diagonal elements and possibly negative diagonal elements. Let $\boldsymbol{A}$ be such a matrix. We will say that a real number x is the *Perron-Frobenius eigenvalue* (PFE) of $\boldsymbol{A}$ and denote it by $\mathrm{pf}\,[\boldsymbol{A}]$ if $x+s$ is the Perron-Frobenius eigenvalue in the usual sense (i.e. the maximal eigenvalue) of the nonnegative matrix $\boldsymbol{A} + s\boldsymbol{I}$ for a sufficiently large s. Obviously, x does not depend on the choice of s. We note that, if $\boldsymbol{A}$ is irreducible, $\mathrm{pf}\,[\boldsymbol{A}]$ is a simple root of the equation $|x\boldsymbol{I} - \boldsymbol{A}| = 0$ and the eigenvector associated with it is positive and unique up to a multiplicative constant.

Markov chain representation: The exogenous arrival process MAP$_k$ has an underlying finite Markov chain with transition rate matrix $\boldsymbol{T}_k + \boldsymbol{U}_k$. Elements of $\boldsymbol{U}_k$ govern state transitions accompanied by arrivals, and off-diagonal elements of $\boldsymbol{T}_k$ govern those without arrivals. Diagonal elements of $\boldsymbol{T}_k$ are negative so that $(\boldsymbol{T}_k + \boldsymbol{U}_k)\boldsymbol{e}^\top = \mathbf{0}^\top$. We denote the state space of the Markov chain by $\mathcal{I}_k$ and refer to the state of the Markov chain as the *phase* of MAP$_k$. We assume that $\mathcal{I}_k$ is finite and $\boldsymbol{T}_k + \boldsymbol{U}_k$ is irreducible. We write $\boldsymbol{a}_k$ the stationary probability vector of the matrix $\boldsymbol{T}_k + \boldsymbol{U}_k$. The exogenous arrival rate is given by $\lambda_k = \left(-\boldsymbol{a}_k\boldsymbol{T}_k^{-1}\boldsymbol{e}^\top\right)^{-1}$. From the model assumption, $\lambda_1 > 0$. When there exist no exogenous arrivals to node 2, both $\boldsymbol{T}_2$ and $\boldsymbol{U}_2$ as a scalar and equal to 0, and set $\lambda_2 = 0$.

The service time distribution PH$_k$ also has an underlying finite absorbing Markov chain with transition rate matrix $\begin{pmatrix} \boldsymbol{S}_k & \boldsymbol{\sigma}_k^\top \\ \mathbf{0} & 0 \end{pmatrix}$ and an initial probability vector $(\,\boldsymbol{b}_k\ 0\,)$. Here $\boldsymbol{\sigma}_k^\top = -\boldsymbol{S}_k\boldsymbol{e}^\top$. The state space of the Markov chain is represented as $\mathcal{J}_k \cup \{0\}$, where $\mathcal{J}_k$ is a finite set of transient states and 0 is a single absorbing state. When a new service starts at node k, the Markov chain starts from a transient state chosen according to the distribution $\boldsymbol{b}_k$, and the service lasts until the chain is absorbed in the absorbing state. The service time distribution has a density function $\boldsymbol{b}_k \exp\{t\boldsymbol{S}_k\}\boldsymbol{\sigma}_k^\top$ for $t \geq 0$. We will refer the state of the chain as the *phase* of PH$_k$. We assume the representation $(\boldsymbol{b}_k, \boldsymbol{S}_k)$ is irreducible in the sense that $\boldsymbol{b}_k(-\boldsymbol{S}_k)^{-1} > \mathbf{0}$. The service rate is given by $\mu_k = \left(-\boldsymbol{b}_k\boldsymbol{S}_k^{-1}\boldsymbol{e}^\top\right)^{-1}$. Of course, $\mu_k > 0$.

Using these underlying Markov chains, we construct a time-continuous Markov chain that represents the stochastic behavior of the whole system. Let $N_k(t)$ be the number of customers in node k at time t, $I_k(t)$ the phase of MAP_k, and $J_k(t)$ the phase of PH_k. We put $J_k(t) = 0$ when $N_k(t) = 0$. Then, the vector

$$\boldsymbol{X}(t) = (N_1(t), N_2(t), I_1(t), I_2(t), J_1(t), J_2(t)) \tag{2.1}$$

constitutes a Markov chain. A typical state can be represented as a sextuple $(n_1, n_2, i_1, i_2, j_1, j_2)$, and the state space is given by

$$\begin{aligned} \mathcal{S} \;=\; & \{\{0\} \times \{0\} \times \mathcal{I}_1 \times \mathcal{I}_2 \times \{0\} \times \{0\}\} \cup \{\{0\} \times \mathcal{N} \times \mathcal{I}_1 \times \mathcal{I}_2 \times \{0\} \times \mathcal{J}_2\} \\ & \cup \{\mathcal{N} \times \{0\} \times \mathcal{I}_1 \times \mathcal{I}_2 \times \mathcal{J}_1 \times \{0\}\} \cup \{\mathcal{N} \times \mathcal{N} \times \mathcal{I}_1 \times \mathcal{I}_2 \times \mathcal{J}_1 \times \mathcal{J}_2\}\ , \end{aligned} \tag{2.2}$$

where $\mathcal{N} = \{1, 2, \cdots\}$. From the irreducibility assumptions of the MAP_k and PH_k representations and from the model assumption that $\lambda_1 > 0$ and $r_{12} > 0$, the chain $\{\boldsymbol{X}(t)\}$ is irreducible.

Stability condition: The chain $\{\boldsymbol{X}(t)\}$ is stable if and only if

$$\rho_k = \frac{(1 - r_{k'k'})\lambda_k + r_{k'k}\lambda_{k'}}{\{(1 - r_{kk})(1 - r_{k'k'}) - r_{kk'}r_{k'k}\}\mu_k} < 1 \qquad \text{for } k = 1, 2. \tag{2.3}$$

This can be proved as in Sigman [15] where a similar stability condition was proved for a general K-node queueing network with a single MMPP (Markov modulated Poisson process) as an input. The proof is not difficult and we omit it here. Hereafter we assume that the condition (2.3) is satisfied. Note that (2.3) implies $r_{kk} < 1$.

3. Balance Equations and Doubly Geometric Form Solution

In order to describe our main result, we shall prepare some notation related to the stationary distribution of the Markov chain $\{\boldsymbol{X}(t)\}$.

Stationary probabilities: Assuming the chain $\{\boldsymbol{X}(t)\}$ is in the steady state, we denote its state probabilities as

$$\begin{aligned} p(n_1, n_2)_{i_1, i_2, j_1, j_2} \;=\; & \mathrm{P}\{(N_1(t), N_2(t), I_1(t), I_2(t), J_1(t), J_2(t)) = (n_1, n_2, i_1, i_2, j_1, j_2)\}, \\ & (n_1, n_2, i_1, i_2, j_1, j_2) \in \mathcal{S}. \end{aligned} \tag{3.1}$$

The joint queue-length probabilities and the marginal queue-length probabilities of node k are written as

$$\begin{aligned} & p(n_1, n_2) = \mathrm{P}\{N_1(t) = n_1, N_2(t) = n_2\}, \quad n_1, n_2 = 0, 1, 2, \ldots\ , \quad \text{and} \\ & p_k(n_k) = \mathrm{P}\{N_k(t) = n_k\}, \quad n_k = 0, 1, 2, \ldots\ . \end{aligned} \tag{3.2}$$

The *decay rate* η_k^* of the marginal queue-length distribution $\{p_k(n_k)\}$ is defined by

$$\log \eta_k^* = \limsup_{n_k \to \infty} \frac{1}{n_k} \log p_k(n_k). \tag{3.3}$$

Obviously, $\eta_k^* \le 1$.

Balance equations: For $n_1, n_2 \ge 1$, we let

$$\begin{aligned} \mathcal{C}(n_1, n_2) \;&=\; \{n_1\} \times \{n_2\} \times \mathcal{I}_1 \times \mathcal{I}_2 \times \mathcal{J}_1 \times \mathcal{J}_2 \\ &=\; \{(n_1, n_2, i_1, i_2, j_1, j_2) \,|\, (i_1, i_2, j_1, j_2) \in \mathcal{I}_1 \times \mathcal{I}_2 \times \mathcal{J}_1 \times \mathcal{J}_2\}\ , \end{aligned} \tag{3.4}$$

and call it *Cell* (n_1, n_2). This is a set of states at which there are n_1 customers in node 1 and n_2 customers in node 2. When $n_1 = 0$ and/or $n_2 = 0$, we define $\mathcal{C}(n_1, n_2)$ in a similar manner by replacing $\mathcal{J}_1$ and/or $\mathcal{J}_2$ above with $\{0\}$. Clearly $p(n_1, n_2) = \mathrm{P}\{\boldsymbol{X}(t) \in \mathcal{C}(n_1, n_2)\}$. The vector of state probabilities corresponding to states in $\mathcal{C}(n_1, n_2)$ is denoted by

$$\boldsymbol{p}(n_1, n_2) = (p(n_1, n_2)_{i_1,i_2,j_1,j_2}\,;\ (n_1, n_2, i_1, i_2, j_1, j_2) \in \mathcal{C}(n_1, n_2))\ . \tag{3.5}$$

For $n_1, n_2 \geq 2$, the set of balance equations around $\mathcal{C}(n_1, n_2)$ is written in vector form as

$$\begin{aligned} \boldsymbol{0} \;=\; & \boldsymbol{p}(n_1, n_2)(\boldsymbol{T}_1 \oplus \boldsymbol{T}_2 \oplus (\boldsymbol{S}_1 + r_{11}\,\boldsymbol{\sigma}_1^\top \boldsymbol{b}_1) \oplus (\boldsymbol{S}_2 + r_{22}\,\boldsymbol{\sigma}_2^\top \boldsymbol{b}_2)) \\ & + \boldsymbol{p}(n_1 - 1, n_2)(\boldsymbol{U}_1 \otimes \boldsymbol{I} \otimes \boldsymbol{I} \otimes \boldsymbol{I}) + \boldsymbol{p}(n_1, n_2 - 1)(\boldsymbol{I} \otimes \boldsymbol{U}_2 \otimes \boldsymbol{I} \otimes \boldsymbol{I}) \\ & + \{r_{10}\,\boldsymbol{p}(n_1 + 1, n_2) + r_{12}\,\boldsymbol{p}(n_1 + 1, n_2 - 1)\}\,(\boldsymbol{I} \otimes \boldsymbol{I} \otimes \boldsymbol{\sigma}_1^\top \boldsymbol{b}_1 \otimes \boldsymbol{I}) \\ & + \{r_{20}\,\boldsymbol{p}(n_1, n_2 + 1) + r_{21}\,\boldsymbol{p}(n_1 - 1, n_2 + 1)\}\,(\boldsymbol{I} \otimes \boldsymbol{I} \otimes \boldsymbol{I} \otimes \boldsymbol{\sigma}_2^\top \boldsymbol{b}_2), \end{aligned} \tag{3.6}$$

where $\otimes$ indicates a Kronecker product operation and $\oplus$ a Kronecker sum operation. If $n_1 \leq 1$ or $n_2 \leq 1$, the equation has to be changed slightly.

Laplace-Stieltjes transforms: The Laplace-Stieltjes transform (LST) of the service time distribution PH_k is given by

$$g_k(y) = \boldsymbol{b}_k(y\boldsymbol{I} - \boldsymbol{S}_k)^{-1}\boldsymbol{\sigma}_k^\top. \tag{3.7}$$

Its domain can be extended to the interval $\mathcal{D}[g_k] = (\delta_k^g, \infty)$, where $\delta_k^g (< 0)$ is its abscissa of convergence. It is easily seen that g_k is strictly decreasing, infinitely differentiable and strictly log-convex on $\mathcal{D}[g_k]$. Further $\lim_{y\to\infty} g_k(y) = 0$ and $\lim_{y\downarrow\delta_k^g} g_k(y) = \infty$. The service rate is given by $\mu_k = -1/g_k'(0)$, where the prime ($'$) indicates a derivative.

For MAP_k, if $\lambda_k > 0$, we let $T_k^A(n)$ be the n-th exogenous arrival epoch at node k, and define the *asymptotic LST* of the exogenous interarrival times by

$$\log f_k(x) = \lim_{n\to\infty} \frac{1}{n} \log E[e^{-xT_k^A(n)}]. \tag{3.8}$$

If the exogenous arrival process to node k is a renewal process, f_k reduces to the ordinary LST of the interarrival time distribution. This function f_k is defined on the interval $\mathcal{D}[f_k] = (\delta_k^f, \infty)$, where $\delta_k^f (< 0)$ is its abscissa of convergence. It is easily seen that $f_k(x)$ is the PFE of the matrix $\boldsymbol{U}_k(x\boldsymbol{I} - \boldsymbol{T}_k)^{-1}$ for $x \in \mathcal{D}[f_k]$. Similar to g_k, the function f_k is strictly decreasing, infinitely differentiable and strictly log-convex on $\mathcal{D}[f_k]$, and $\lim_{x\to\infty} f_k(x) = 0$ and $\lim_{x\downarrow\delta_k^f} f_k(x) = \infty$. The exogenous arrival rate is given by $\lambda_k = -1/f_k'(0)$. If there exist no exogenous arrivals to node 2, the function f_2 is not defined and we set $\lambda_2 = 0$.

Doubly geometric form solution: Our main result relates in a deep manner to a doubly geometric form solution to the balance equations around a cell. We will say that a solution $\{\boldsymbol{p}^\dagger(m_1, m_2)\,;\ \ m_1 = n_1, n_1 \pm 1 \text{ and } m_2 = n_2, n_2 \pm 1\}$ to the balance equations (3.6) around $\mathcal{C}(n_1, n_2)$ is of a *doubly geometric form* if there exist positive numbers η_1 and η_2 and a positive vector $\boldsymbol{\nu}$ such that

$$\boldsymbol{p}^\dagger(m_1, m_2) = \eta_1^{m_1}\eta_2^{m_2}\boldsymbol{\nu}. \tag{3.9}$$

Substituting (3.9) into (3.6), we have

$$\begin{aligned} \boldsymbol{0} \;=\; & \eta_1^{n_1}\eta_2^{n_2}\,\boldsymbol{\nu}\left[\left(\boldsymbol{T}_1 + \eta_1^{-1}\boldsymbol{U}_1\right) \oplus \left(\boldsymbol{T}_2 + \eta_2^{-1}\boldsymbol{U}_2\right)\right. \\ & \oplus \left(\boldsymbol{S}_1 + \left(r_{10}\,\eta_1 + r_{11} + r_{12}\,\eta_1\eta_2^{-1}\right)\boldsymbol{\sigma}_1^\top \boldsymbol{b}_1\right) \\ & \left.\oplus \left(\boldsymbol{S}_2 + \left(r_{20}\,\eta_2 + r_{21}\,\eta_1^{-1}\eta_2 + r_{22}\right)\boldsymbol{\sigma}_2^\top \boldsymbol{b}_2\right)\right]. \end{aligned} \tag{3.10}$$

This equation indicates that $\boldsymbol{\nu}$ is a left eigenvector of the matrix in the brackets associated with eigenvalue 0. The matrix is represented as a Kronecker sum of four smaller matrices. So, the eigenvalue 0 is given by a sum of eigenvalues of these four smaller matrices and the eigenvector $\boldsymbol{\nu}$ is a Kronecker product of eigenvectors associated with those eigenvalues. Let us denote these eigenvalues as x_1, x_2, $-y_1$ and $-y_2$, and corresponding eigenvectors as $\overline{\boldsymbol{\nu}}_1$, $\overline{\boldsymbol{\nu}}_2$, $\boldsymbol{\nu}_1$ and $\boldsymbol{\nu}_2$. Then

$$x_1 + x_2 - y_1 - y_2 = 0, \tag{3.11}$$
$$\boldsymbol{\nu} = \overline{\boldsymbol{\nu}}_1 \otimes \overline{\boldsymbol{\nu}}_2 \otimes \boldsymbol{\nu}_1 \otimes \boldsymbol{\nu}_2, \tag{3.12}$$
$$\overline{\boldsymbol{\nu}}_k \left(\boldsymbol{T}_k + \eta_k^{-1}\boldsymbol{U}_k\right) = x_k \overline{\boldsymbol{\nu}}_k\,, \qquad \text{and} \tag{3.13}$$
$$\boldsymbol{\nu}_k \left(\boldsymbol{S}_k + \left(r_{k0}\,\eta_k + r_{kk} + r_{kk'}\,\eta_k\eta_{k'}^{-1}\right)\boldsymbol{\sigma}_k^\top \boldsymbol{b}_k\right) = -y_k\,\boldsymbol{\nu}_k\,. \tag{3.14}$$

The matrices in (3.13) and (3.14) have nonnegative off-diagonal elements, and we can speak of Perron-Frobenius eigenvalues for them. For fixed η_1 and η_2, from the irreducibility assumptions of the representations for MAP$_k$ and PH$_k$, these matrices are irreducible and have simple PFE's and unique positive eigenvectors associated with them. Here we are interested in positive $\boldsymbol{\nu}$. So, x_k and $-y_k$ have to be PFE's and $\overline{\boldsymbol{\nu}}_k$ and $\boldsymbol{\nu}_k$ are associated positive eigenvectors. If eigenvalues x_k and $-y_k$ satisfy (3.11) then (3.10) holds.

Using the functions f_k and g_k, we can rewrite relations among values η_k, x_k and y_k. From (3.13), a simple calculation shows that

$$\overline{\boldsymbol{\nu}}_k\,\boldsymbol{U}_k(x_k\boldsymbol{I} - \boldsymbol{T}_k)^{-1} = \eta_k\,\overline{\boldsymbol{\nu}}_k. \tag{3.15}$$

This equation implies that η_k is the PFE of the matrix $\boldsymbol{U}_k(x_k\boldsymbol{I} - \boldsymbol{T}_k)^{-1}$, and hence, if $\lambda_k > 0$,

$$\eta_k = f_k(x_k). \tag{3.16}$$

If $\lambda_k = 0$, we have assumed $\boldsymbol{T}_k = \boldsymbol{U}_k = 0$ (scalar). Hence (3.13) implies that $x_k = 0$ and $\overline{\boldsymbol{\nu}}_k$ is arbitrary (so we let $\overline{\boldsymbol{\nu}}_k = 1$ (scalar)). The equation (3.14) can be rewritten as

$$\boldsymbol{\nu}_k = (\boldsymbol{\nu}_k\boldsymbol{\sigma}_k^\top)\,\left(r_{k0}\,\eta_k + r_{kk} + r_{kk'}\,\eta_k\eta_{k'}^{-1}\right)\boldsymbol{b}_k\,(-y_k\boldsymbol{I} - \boldsymbol{S}_k)^{-1}\,. \tag{3.17}$$

Postmultiplied by $\boldsymbol{\sigma}_k^\top$, we have from (3.7)

$$\left(r_{k0}\,\eta_k + r_{kk} + r_{kk'}\,\eta_k\eta_{k'}^{-1}\right)\,g_k(-y_k) = 1. \tag{3.18}$$

Thus, under the doubly geometric form assumption (3.9), if $\lambda_2 > 0$, the six variables η_1, η_2, x_1, x_2, y_1 and y_2 satisfy five equations (3.16), (3.18) (two equations each) and (3.11). If $\lambda_2 = 0$, the six variables satisfy $x_2 = 0$, (3.16) for $k = 1$, (3.18) for $k = 1, 2$, and (3.11).

4. Main Theorem

To make our discussions simpler, hereafter we assume there exists no direct feedback to the same node, that is $r_{kk} = 0$ for $k = 1, 2$. This does not reduce any generality as long as we are concerned about the numbers of customers in nodes 1 and 2, because, if $r_{kk} > 0$, we may change the routing probabilities to

$$\tilde{r}_{k0} = r_{k0}/(1 - r_{kk}), \quad \tilde{r}_{kk} = 0 \quad \text{and} \qquad \tilde{r}_{kk'} = r_{kk'}/(1 - r_{kk}) \tag{4.1}$$

and the service time distribution so that

$$(\tilde{\boldsymbol{b}}_k, \tilde{\boldsymbol{S}}_k) = (\boldsymbol{b}_k, \boldsymbol{S}_k + r_{kk}\boldsymbol{\sigma}_k^\top\boldsymbol{b}_k) \quad \text{and} \quad \tilde{\boldsymbol{\sigma}}_k = (1 - r_{kk})\boldsymbol{\sigma}_k. \tag{4.2}$$

The new model with these modified routing probabilities and service time distribution has the same $\{\boldsymbol{X}(t)\}$ process as the original one. Thus, the decay rate of the new model coincides with the original one.

Two-variable representations and inverse functions: To represent our main result in a simpler form, we write the six variables η_1, η_2, x_1, x_2, y_1 and y_2 as functions of two variables $a_1 = \log \eta_1$ and $a_2 = \log \eta_2$. Clearly,

$$\eta_k = e^{a_k}. \tag{4.3}$$

To represent x_k and y_k as functions of a_1 and a_2, we need to introduce inverse functions. For an arbitrary monotone function h, we denote its inverse function by $\mathrm{inv}[h]$. For the moment, we assume that $\lambda_k > 0$. Let ϕ_k be the inverse function of $\log f_k$, and ψ_k be that of $\log g_k$, i.e.

$$\phi_k(a) = \mathrm{inv}[\log f_k](a) \quad \text{and} \quad \psi_k(a) = \mathrm{inv}[\log g_k](a). \tag{4.4}$$

These functions are defined on the whole real line $(-\infty, +\infty)$. From Theorem 1 of Glynn and Whitt [5], the functions ϕ_k and ψ_k can be interpreted probabilistically in the following manner. Let $N_k^A(t)$ be the number of exogenous arrivals at node k during time interval $(0, t]$, and $N_k^S(t)$ the number of (fictitious) customers served at node k during $(0, t]$ provided that the server continues processing. Then we have

$$\phi_k(a) = \lim_{t\to\infty} \frac{1}{t} \log E[e^{-aN_k^A(t)}] \qquad \text{and} \qquad \psi_k(a) = \lim_{t\to\infty} \frac{1}{t} \log E[e^{-aN_k^S(t)}]. \tag{4.5}$$

Since f_k and g_k are decreasing, infinitely differentiable and log-convex, ϕ_k and ψ_k are also decreasing, infinitely differentiable and convex. They satisfy the following properties:

$$\begin{aligned} &\phi_k(0) = \psi_k(0) = 0, \quad \phi_k'(0) = -\lambda_k, \quad \psi_k'(0) = -\mu_k, \quad \lim_{a\to+\infty} \phi_k(a) = \delta_k^f < 0, \\ &\lim_{a\to+\infty} \psi_k(a) = \delta_k^g < 0 \quad \text{and} \quad \lim_{a\to-\infty} \phi_k(a) = \lim_{a\to-\infty} \psi_k(a) = +\infty. \end{aligned} \tag{4.6}$$

From (3.16) and (3.18), the variables x_k and y_k can be represented as

$$x_k = \phi_k(a_k) \quad \text{and} \quad -y_k = \psi_k(-a_k + h_{k'}(a_{k'})), \tag{4.7}$$

$$\text{where} \quad h_k(a) = -\log\left(r_{k'k}\, e^{-a} + r_{k'0}\right). \tag{4.8}$$

If $r_{k'k} > 0$, the function h_k is strictly increasing and concave, and satisfies the following properties:

$$h_k(0) = 0, \quad h_k'(0) = r_{k'k}, \quad \lim_{a\to+\infty} h_k(a) = -\log r_{k'0} \quad \text{and} \quad \lim_{a\to-\infty} h_k(a) = -\infty. \tag{4.9}$$

If $r_{k'k} = 0$ (this may occur only for $k = 1$ from the assumption $r_{12} > 0$), then $h_k(a_k) \equiv 0$. From (4.7), the equation (3.11) is rewritten as

$$\kappa(a_1, a_2) \equiv \phi_1(a_1) + \phi_2(a_2) + \psi_1(-a_1 + h_2(a_2)) + \psi_2(-a_2 + h_1(a_1)) = 0. \tag{4.10}$$

This function κ plays a very important role in our discussion.

So far we have assumed that $\lambda_2 > 0$. If $\lambda_2 = 0$, we cannot define f_2 in (3.8) since the random variable $T_2^A(n)$ does not exist. So in this case we set

$$\phi_2(a) \equiv 0, \quad a \in (-\infty, +\infty) \qquad (\text{when } \lambda_2 = 0). \tag{4.11}$$

By this rule, we need not distinguish the cases $\lambda_2 > 0$ and $\lambda_2 = 0$ in most of the subsequent discussions.

To summarize, for any pair (a_1, a_2) satisfying $\kappa(a_1, a_2) = 0$, the corresponding values of variables η_1, η_2, x_1, x_2, y_1 and y_2 are given by (4.3) and (4.7), and the corresponding positive vectors $\overline{\boldsymbol{\nu}}_1$, $\overline{\boldsymbol{\nu}}_2$, $\boldsymbol{\nu}_1$, $\boldsymbol{\nu}_2$ and $\boldsymbol{\nu}$ are derived from (3.13), (3.14) and (3.12). Then a doubly geometric form solution (3.9) formed with η_1, η_2 and $\boldsymbol{\nu}$ given above satisfies the set of balance equations around $\mathcal{C}(n_1, n_2)$ as in (3.10).

In order to describe our main theorem we introduce sets on the (a_1, a_2)-plane and numbers related to them. First we note that, as will be shown in Lemma 7.1, the set

$$\mathcal{K}_{loop} = \{(a_1, a_2) \ : \ \kappa(a_1, a_2) = 0\} \tag{4.12}$$

is a loop passing through the origin, i.e. $\kappa(0, 0) = 0$ (see Figure 4.1 for an example). We let

$$\mathcal{E}_k = \{(a_1, a_2) \in \mathcal{K}_{loop} \ : \ a_k < 0 \text{ and } h_k(a_k) \le a_{k'} \le 0\} \quad \text{and} \tag{4.13}$$

$$b_k^{\mathcal{E}_k} = \inf\{a_k \ : \ \exists\, a_{k'} \text{ such that } (a_1, a_2) \in \mathcal{E}_k\}. \tag{4.14}$$

Further we define

$$\begin{aligned} \mathcal{F}_k &= \mathcal{E}_k \cap \{(a_1, a_2) \ : \ a_{k'} \ge b_{k'}^{\mathcal{E}_{k'}}\} \\ &= \{(a_1, a_2) \in \mathcal{K}_{loop} \ : \ a_k < 0 \text{ and } \max\{h_k(a_k),\ b_{k'}^{\mathcal{E}_{k'}}\} \le a_{k'} \le 0\}. \end{aligned} \tag{4.15}$$

and let

$$\overline{\eta}_k = \exp\{b_k^{\mathcal{F}_k}\}, \quad \text{where} \quad b_k^{\mathcal{F}_k} = \inf\{a_k \ : \ \exists\, a_{k'} \text{ such that } (a_1, a_2) \in \mathcal{F}_k\}. \tag{4.16}$$

As will be shown in Lemma 7.2, $\mathcal{E}_k$ and $\mathcal{F}_k$ are nonempty and the numbers $b_k^{\mathcal{E}_k}$, $b_k^{\mathcal{F}_k}$ and $\overline{\eta}_k$ are all well defined. Our main theorem is stated as follows.

Theorem 4.1 $\overline{\eta}_k = \exp\{b_k^{\mathcal{F}_k}\}$ *is less than 1 and is greater than or equal to the decay rate* η_k^* *of the marginal queue-length distribution* $\{p_k(n)\}$ *defined in* (3.3), *namely*

$$\eta_k^* \le \overline{\eta}_k < 1. \tag{4.17}$$

The proof of this theorem needs a long discussion using a series of lemmas. So we postpone it to Section 6.

Remark 4.1 (On the sets $\mathcal{E}_k$ and $\mathcal{F}_k$) The set

$$\mathcal{E}'_k = \{(a_1, a_2) \in \mathcal{K}_{loop} \ : \ a_k < 0 \text{ and } h_k(a_k) < a_{k'} < 0\} \tag{4.18}$$

consists of points (a_1, a_2) satisfying the condition of Lemma 6.1 given in Section 6 together with constraints $a_k < 0$ and $a_{k'} < 0$. $\mathcal{E}_k$ is almost the same as this set. However $\mathcal{E}'_k$ becomes empty if $h_k(a_k) \equiv 0$ and hence if $r_{k'k} = 0$. To avoid this inconvenience and to make the exposition of the theorem general, $\mathcal{E}_k$ is slightly changed from $\mathcal{E}'_k$ so that it is nonempty in any case. $\mathcal{F}_k$ is defined so that it provides the coordinate of the limit point of the converging sequence of points given by (6.31) in Section 6. ♢

Remark 4.2 (An LST version) Theorem 4.1 can be restated in terms of f_k and g_k directly. For brevity of exposition, here we state the result only for the case $\lambda_2 > 0$. If $\lambda_2 = 0$, we have to change equations (4.19) and definitions (4.20) below slightly. From (3.16), (3.18) and (3.11), we have three equations for four variables x_1, x_2, y_1 and y_2 as

$$\begin{aligned}&\left(r_{k0}\, f_k(x_k) + r_{kk} + r_{kk'}\,\frac{f_k(x_k)}{f_{k'}(x_{k'})}\right) g_k(-y_k) = 1 \quad \text{for} \quad k = 1, 2, \quad \text{and} \\ &x_1 + x_2 - y_1 - y_2 = 0.\end{aligned} \tag{4.19}$$

We define positive numbers $\hat{x}_k$ and $\overline{\eta}'_k$ as

$$\begin{aligned}\hat{x}_k &= \sup\{x_k \ : \ \exists(x_1, x_2, y_1, y_2) \text{ satisfying } (4.19), x_{k'} \geq 0, x_k > 0 \text{ and } y_{k'} \leq 0\}, \\ \overline{\eta}'_k &= \inf\{f_k(x_k) \ : \ \exists(x_1, x_2, y_1, y_2) \text{ satisfying } (4.19), x_k > 0, x_{k'} \in [0, \hat{x}_{k'}] \text{ and } y_{k'} \leq 0\}.\end{aligned} \tag{4.20}$$

Then from relations in (4.7) we see $\overline{\eta}'_k = \overline{\eta}_k$. ♢

Remark 4.3 (Corollaries on procedure) To calculate $b_1^{\mathcal{F}_1}$ and $b_2^{\mathcal{F}_2}$ in individual models, we shall write them more explicitly. We consider the curves and straight lines

$$\kappa(a_1, a_2) = 0 \quad (\text{the set } \mathcal{K}_{loop}), \quad a_{k'} = h_k(a_k), \quad a_k = 0 \quad \text{and} \quad a_k = b_k^{\mathcal{E}_k} \tag{4.21}$$

on the (a_1, a_2)-plane (see Figure 4.1 for an example). Properties of these curves and lines will be examined in Section 7. We start with introducing notations for some sets and coordinates of intersections of the curves and lines. We let

$$\mathcal{K} = \{(a_1, a_2) \ : \ \kappa(a_1, a_2) \leq 0\} \quad \text{and} \quad \mathcal{K}_k = \{(a_1, a_2) \in \mathcal{K} \ : \ a_{k'} \leq 0\}. \tag{4.22}$$

The set $\mathcal{K}$ is convex (Lemma 7.1) and its periphery is $\mathcal{K}_{loop}$. For a given number b, the straight line $a_k = b$ either intersects with $\mathcal{K}_{loop}$ twice, is tangent to $\mathcal{K}_{loop}$ at a single point, or never meets $\mathcal{K}_{loop}$. In the following discussions, for brevity of exposition, sometimes we say the line intersects with $\mathcal{K}_{loop}$ at two points even when the line is tangent to the loop.

The straight line $a_{k'} = 0$ intersects with $\mathcal{K}_{loop}$ at two points, one of which is the origin (Lemma 7.1). Let b_k^0 be the smaller k-th coordinate of the two intersections, namely

$$b_1^0 = \min\{a_1 \ : \ \kappa(a_1, 0) = 0\} \quad \text{and} \quad b_2^0 = \min\{a_2 \ : \ \kappa(0, a_2) = 0\}. \tag{4.23}$$

The curve $a_{k'} = h_k(a_k)$ intersects $\mathcal{K}_{loop}$ at two points, at the origin and at a point having negative coordinates (Lemma 7.1). Let $(b_1^{h_k}, b_2^{h_k})$ be the coordinates of the latter point, namely, for example, if $k = 1$,

$$b_1^{h_1} = \{\text{unique negative root of equation } \kappa(a_1, h_1(a_1)) = 0 \text{ for } a_1\} \quad \text{and} \quad b_2^{h_1} = h_1(b_1^{h_1}). \tag{4.24}$$

The straight line $a_k = b_k^{h_k}$ intersects $\mathcal{K}_{loop}$ at two points, one of which is $(b_1^{h_k}, b_2^{h_k})$. Let $b_{k'}^{h_k,c}$ be the k'-th coordinate of the other intersection. Namely, for example, if $k = 1$,

$$b_2^{h_1,c} = \left\{\text{root other than } b_2^{h_1} \text{ of equation } \kappa(b_1^{h_1}, a_2) = 0 \text{ for } a_2\right\}. \tag{4.25}$$

We let

$$\begin{aligned}b_k^{\mathcal{K}(k)} &= \inf\{a_k \ : \ \exists\, a_{k'} \text{ such that } (a_1, a_2) \in \mathcal{K}\}, \quad \text{and} \\ b_k^{\mathcal{K}_k} &= \inf\{a_k \ : \ \exists\, a_{k'} \text{ such that } (a_1, a_2) \in \mathcal{K}_k\}.\end{aligned} \tag{4.26}$$

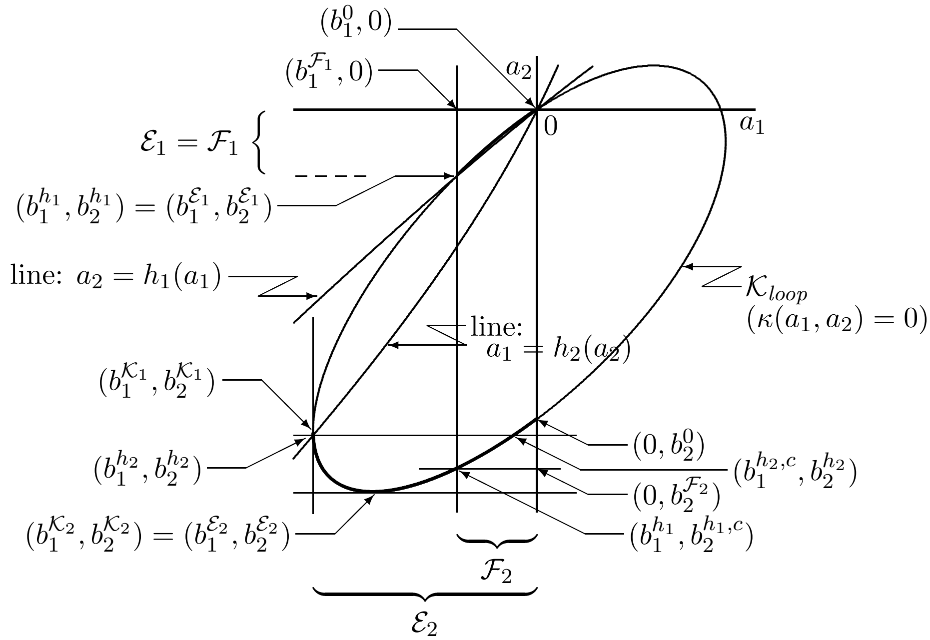

Figure 4.1: Curves and lines on the (a_1, a_2)-plane for the model in Example 4.1

Since these numbers are less than or equal to $b_k^{h_k}$ (< 0), they are negative. Each of infimums in (4.26) is attained by a single point on $\mathcal{K}_{loop}$ (Lemma 7.2). We denote by $(b_1^{\mathcal{K}(k)}, b_2^{\mathcal{K}(k)})$ the coordinates of the point attaining the first infimum and by $(b_1^{\mathcal{K}_k}, b_2^{\mathcal{K}_k})$ those attaining the second infimum. If $b_2^{\mathcal{K}(1)} \leq 0$, $(b_1^{\mathcal{K}_1}, b_2^{\mathcal{K}_1}) = (b_1^{\mathcal{K}(1)}, b_2^{\mathcal{K}(1)})$, and if $b_2^{\mathcal{K}(1)} \geq 0$, $(b_1^{\mathcal{K}_1}, b_2^{\mathcal{K}_1}) = (b_1^0, 0)$.

The set $\mathcal{E}_1$ defined in (4.13) is the upper left arc (segment) of $\mathcal{K}_{loop}$ between points $(b_1^0, 0)$ and $(b_1^{h_1}, b_2^{h_1})$ in the third quadrant (Lemma 7.2, see Figure 4.1 for an example). The latter end point belongs to the set, but the former does only when $b_1^0 < 0$. We let $(b_1^{\mathcal{E}_1}, b_2^{\mathcal{E}_1})$ be the coordinates of the point that attains the infimum of a_1 in $\mathcal{E}_1$ as in (4.14). Similarly the set $\mathcal{E}_2$ is the lower right arc of $\mathcal{K}_{loop}$ between $(0, b_2^0)$ and $(b_1^{h_2}, b_2^{h_2})$, and we denote by $(b_1^{\mathcal{E}_2}, b_2^{\mathcal{E}_2})$ the coordinates of the point at which the infimum of a_2 in $\mathcal{E}_2$ is attained. Then, by tracing the arc $\mathcal{E}_k$ from the end point near the origin, we see that coordinates $(b_1^{\mathcal{E}_k}, b_2^{\mathcal{E}_k})$ are given as follows (see Lemma 7.2 and its proof).

$$(b_1^{\mathcal{E}_k}, b_2^{\mathcal{E}_k}) = \begin{cases} (b_1^{\mathcal{K}_k}, b_2^{\mathcal{K}_k}) & \text{if } b_{k'}^{h_k,c} \geq b_{k'}^{h_k}, \\ (b_1^{h_k}, b_2^{h_k}) & \text{if } b_{k'}^{h_k,c} \leq b_{k'}^{h_k}. \end{cases} \tag{4.27}$$

Since $\mathcal{F}_k$ is a subset of $\mathcal{E}_k$ restricted by $a_{k'} \geq b_{k'}^{\mathcal{E}_{k'}}$, by comparing $b_{k'}^{\mathcal{E}_k}$ and $b_{k'}^{\mathcal{E}_{k'}}$ (see Lemma 7.2), we see that the infimum $b_k^{\mathcal{F}_k}$ in (4.16) is given by

$$b_k^{\mathcal{F}_k} = \begin{cases} b_k^{\mathcal{E}_k} & \text{if } b_{k'}^{\mathcal{E}_k} \geq b_{k'}^{\mathcal{E}_{k'}}, \\ b_k^{h_{k'},c} & \text{if } b_{k'}^{\mathcal{E}_k} \leq b_{k'}^{\mathcal{E}_{k'}}. \end{cases} \tag{4.28}$$

Using these relations, we can calculate $\overline{\eta}_k$ concretely as stated in the following corollary.

Corollary 4.1 *The upper bound $\overline{\eta}_k$ given in Theorem 4.1 can be calculated through* (4.27) *and* (4.28) *as* $\overline{\eta}_k = \exp\{b_k^{\mathcal{F}_k}\}$.

The procedure presented in Corollary 4.1 can be written without using $b_k^{\mathcal{E}_k}$ and $b_{k'}^{\mathcal{E}_k}$. Note that from Lemma 7.1 the case where $b_2^{h_1} \leq b_2^{h_2}$ and $b_1^{h_2} \leq b_1^{h_1}$ cannot occur.

Corollary 4.2 *The pair* $(b_1^{\mathcal{F}_1}, b_2^{\mathcal{F}_2})$ *which derives the upper bounds* $\overline{\eta}_1$ *and* $\overline{\eta}_2$ *in Theorem 4.1 is given by*

$$(b_1^{\mathcal{F}_1}, b_2^{\mathcal{F}_2}) = \begin{cases} (b_1^{h_1}, b_2^{h_2}) & \text{if } b_2^{h_1,c} \le b_2^{h_1},\ b_1^{h_2,c} \le b_1^{h_2},\ b_2^{h_1} \ge b_2^{h_2} \text{ and } b_1^{h_2} \ge b_1^{h_1}, \\ (b_1^{h_1}, b_2^{h_1,c}) & \text{if } b_2^{h_1,c} \le b_2^{h_1},\ b_1^{h_2,c} \le b_1^{h_2},\ b_2^{h_1} \ge b_2^{h_2} \text{ and } b_1^{h_2} \le b_1^{h_1}, \\ & \text{or if } b_2^{h_1,c} \le b_2^{h_1},\ b_1^{h_2,c} \ge b_1^{h_2} \text{ and } b_1^{\mathcal{K}_2} \le b_1^{h_1}, \\ (b_1^{h_2,c}, b_2^{h_2}) & \text{if } b_2^{h_1,c} \le b_2^{h_1},\ b_1^{h_2,c} \le b_1^{h_2},\ b_2^{h_1} \le b_2^{h_2} \text{ and } b_1^{h_2} \ge b_1^{h_1}, \\ & \text{or if } b_2^{h_1,c} \ge b_2^{h_1},\ b_1^{h_2,c} \le b_1^{h_2} \text{ and } b_2^{\mathcal{K}_1} \le b_2^{h_2}, \\ (b_1^{h_1}, b_2^{\mathcal{K}_2}) & \text{if } b_2^{h_1,c} \le b_2^{h_1},\ b_1^{h_2,c} \ge b_1^{h_2} \text{ and } b_1^{\mathcal{K}_2} \ge b_1^{h_1}, \\ (b_1^{\mathcal{K}_1}, b_2^{h_2}) & \text{if } b_2^{h_1,c} \ge b_2^{h_1},\ b_1^{h_2,c} \le b_1^{h_2} \text{ and } b_2^{\mathcal{K}_1} \ge b_2^{h_2}, \\ (b_1^{\mathcal{K}_1}, b_2^{\mathcal{K}_2}) & \text{if } b_2^{h_1,c} \ge b_2^{h_1} \text{ and } b_1^{h_2,c} \ge b_1^{h_2}. \end{cases} \tag{4.29}$$

$\diamondsuit$

Example 4.1 Figure 4.1 shows curves and coordinates of intersections on the (a_1, a_2)-plane for the model with two-phase Erlang renewal arrival processes and two-phase Erlang service distributions, where $\lambda_1 = 1.0$, $\lambda_2 = 0.5$, $\mu_1 = 3.0$, $\mu_2 = 5.0$, $r_{12} = r_{10} = 0.5$, $r_{21} = 0.8$ and $r_{20} = 0.2$. The procedure presented in Corollary 4.1 is applied in the following manner.

$$\left.\begin{array}{l}(b_1^{h_1}, b_2^{h_1}) = (-0.4073, -0.3381) \\ (b_1^{\mathcal{K}_1}, b_2^{\mathcal{K}_1}) = (-1.1407, -1.6380)\end{array}\right\} \implies (b_1^{\mathcal{E}_1}, b_2^{\mathcal{E}_1}) = (b_1^{h_1}, b_2^{h_1}) = (-0.4073, -0.3381)$$

$$\left.\begin{array}{l}(b_1^{h_2}, b_2^{h_2}) = (-1.1403, -1.6592) \\ (b_1^{\mathcal{K}_2}, b_2^{\mathcal{K}_2}) = (-0.8390, -1.9508)\end{array}\right\} \implies (b_1^{\mathcal{E}_2}, b_2^{\mathcal{E}_2}) = (b_1^{\mathcal{K}_2}, b_2^{\mathcal{K}_2}) = (-0.8390, -1.9508)$$

$$b_2^{\mathcal{E}_1} > b_2^{\mathcal{E}_2} \implies b_1^{\mathcal{F}_1} = b_1^{\mathcal{E}_1} = b_1^{h_1} = -0.4073, \qquad \overline{\eta}_1 = \exp\{b_1^{\mathcal{F}_1}\} = 0.6654$$

$$b_1^{\mathcal{E}_2} < b_1^{\mathcal{E}_1} \implies b_2^{\mathcal{F}_2} = b_2^{h_1,c} = -1.8301, \qquad \overline{\eta}_2 = \exp\{b_2^{\mathcal{F}_2}\} = 0.1604.$$

$\diamondsuit$

Remark 4.4 (Special models) In some special cases where exact decay rates are known, we can check whether our upper bound coincides with the exact decay rate or not. The followings are propositions for such special cases. Proofs will be given in Section 8.

Proposition 4.1 *In a Jackson type two-node queueing system, the upper bound* $\overline{\eta}_k$ *coincides with the exact decay rate* $\eta_k^* = \rho_k$ *given in* (2.3).

Proposition 4.2 *If* $r_{21} = 0$, *we have* $\overline{\eta}_1 = \eta_1^*$ (< 1).

Proposition 4.3 *In a tandem queueing system MAP/M/1* $\to$ */M/1, the upper bound* $\overline{\eta}_k$ *coincides with the exact decay rate* η_k^*. $\diamondsuit$

5. Matrix Geometric Form of a QBD Process Having Infinite Number of States in Each Level

Our proof of the main theorem, Theorem 4.1, is based on a lemma concerning to the rate matrix in a quasi-birth-and-death process having infinite number of states in each level [9]. The lemma has been presented only in a technical paper, we introduce it with a brief proof here.

We consider a time-continuous ergodic Markov chain on a two-dimensional state space $\mathcal{S} = \{(n, i); n, i = 0, 1, 2, \cdots\}$. Let $\mathcal{L}(n)$ be the set of states $\{(n, i); i = 0, 1, 2, \cdots\}$ with common n and call it *Level* n. The whole state space $\mathcal{S}$ is partitioned into levels as $\mathcal{S} = \bigcup_{n=0}^{\infty} \mathcal{L}(n)$. The Markov chain is called a *quasi-birth-and-death process having infinite number of states in each level* if, after partitioned into levels, its transition rate matrix becomes of a block tri-diagonal form

$$\boldsymbol{Q} = \begin{pmatrix} \overline{\boldsymbol{Q}}_1 & \overline{\boldsymbol{Q}}_0 & & & \\ \overline{\boldsymbol{Q}}_2 & \boldsymbol{Q}_1 & \boldsymbol{Q}_0 & & \\ & \boldsymbol{Q}_2 & \boldsymbol{Q}_1 & \boldsymbol{Q}_0 & \\ & & \boldsymbol{Q}_2 & \boldsymbol{Q}_1 & \ddots \\ & & & \ddots & \ddots \end{pmatrix}. \tag{5.1}$$

Notice that $\overline{\boldsymbol{Q}}_i$ and $\boldsymbol{Q}_i$ have infinite dimension.

Let $\boldsymbol{\pi}$ be the stationary state probability vector of $\boldsymbol{Q}$ and divide it into subvectors as $\boldsymbol{\pi} = (\boldsymbol{\pi}(0)\ \boldsymbol{\pi}(1)\ \boldsymbol{\pi}(2)\ \cdots)$ according to the partition of $\mathcal{S}$ into levels. It is known [10] that $\boldsymbol{\pi}$ takes a matrix geometric form as

$$\boldsymbol{\pi}(n) = \boldsymbol{\pi}(1)\boldsymbol{R}^{n-1}, \qquad n = 1, 2, \cdots \tag{5.2}$$

where $\boldsymbol{R}$, called the *rate matrix*, is the minimal nonnegative solution of the matrix quadratic equation

$$\boldsymbol{Q}_0 + \boldsymbol{R}\boldsymbol{Q}_1 + \boldsymbol{R}^2\boldsymbol{Q}_2 = \boldsymbol{O}. \tag{5.3}$$

If the dimension of $\boldsymbol{R}$ were finite, the level distribution $\{\boldsymbol{\pi}(n)\,\boldsymbol{e}^\top\}$ would decay geometrically fast with rate equal to pf $[\boldsymbol{R}]$. However, in our case, the dimension of $\boldsymbol{R}$ is infinite and we cannot use the concept "eigenvalue". A sufficient condition for geometric decay with a given decay rate was obtained in Takahashi et al. [17]. Here we use another lemma, Lemma 5.1 below, to evaluate powers of $\boldsymbol{R}$. A similar result was obtained in Theorem 5.4 of Ramaswami and Taylor [14] where (5.5) below was derived with equality but under slightly stronger conditions.

Lemma 5.1 *Assume that diagonal elements of $\boldsymbol{Q}_1$ are bounded. If there exists a positive number ξ and a positive vector $\boldsymbol{q}$ satisfying*

$$\boldsymbol{q}\left(\frac{1}{\xi}\boldsymbol{Q}_0 + \boldsymbol{Q}_1 + \xi\boldsymbol{Q}_2\right) \le \boldsymbol{0}, \tag{5.4}$$

then

$$\boldsymbol{q}\boldsymbol{R} \le \xi\boldsymbol{q}. \tag{5.5}$$

Proof Choose a sufficiently large positive number ν so that $\boldsymbol{I} + \frac{1}{\nu}\boldsymbol{Q}_1$ is nonnegative, and put

$$\boldsymbol{P}_0 = \frac{1}{\nu}\boldsymbol{Q}_0, \quad \boldsymbol{P}_1 = \boldsymbol{I} + \frac{1}{\nu}\boldsymbol{Q}_1, \quad \boldsymbol{P}_2 = \frac{1}{\nu}\boldsymbol{Q}_2. \tag{5.6}$$

Starting with $\boldsymbol{R}(0) = \boldsymbol{O}$, we recursively define matrices $\boldsymbol{R}(n)$ by the relation

$$\boldsymbol{R}(n) = \boldsymbol{P}_0 + \boldsymbol{R}(n-1)\boldsymbol{P}_1 + \boldsymbol{R}^2(n-1)\boldsymbol{P}_2, \quad n = 1, 2, \cdots. \tag{5.7}$$

Then $\boldsymbol{R}(n)$ is nonnegative, nondecreasing and bounded from above by $\boldsymbol{R}$. It is easily checked that $\boldsymbol{R}(n)$ converges to $\boldsymbol{R}$ as $n \to \infty$. By mathematical induction, we can show that $\boldsymbol{q}\boldsymbol{R}(n) \le \xi\boldsymbol{q}$ for all $n \ge 0$. Since $\boldsymbol{q}$ and $\boldsymbol{R}(n)$ are nonnegative, by letting n to infinity we have $\boldsymbol{q}\boldsymbol{R} \le \xi\boldsymbol{q}$ from the monotone convergence theorem. ♢

From this lemma, we can derive an important inequality for the discussion of decay rates.

Lemma 5.2 *In addition to the conditions of Lemma 5.1, if* $\boldsymbol{q}$ *satisfies*

$$\boldsymbol{\pi}(1) \leq c\boldsymbol{q} \tag{5.8}$$

for some positive constant c, then we have

$$\boldsymbol{\pi}(n) \leq c\xi^{n-1}\boldsymbol{q}, \qquad n = 1, 2, \cdots. \tag{5.9}$$

Moreover, if $\boldsymbol{q}\boldsymbol{e}^\top < +\infty$, *then* ξ *is an upper bound of the decay rate of the level distribution in the sense*

$$\limsup_{n\to\infty} \frac{1}{n} \log \boldsymbol{\pi}(n)\boldsymbol{e}^\top \leq \log \xi. \tag{5.10}$$

Proof The lemma is clear from (5.2) and (5.5). ♢

6. Proof of the Main Theorem

Now we shall prove Theorem 4.1. We start with some preparatory discussion. Through this section, if a pair $(a_1, a_2) \in \mathcal{K}_{loop}$ is once chosen, we consider the values of the associated variables η_1, η_2, x_1, x_2, y_1 and y_2 are given by (4.3) and (4.7), and the associated vectors $\overline{\boldsymbol{\nu}}_1$, $\overline{\boldsymbol{\nu}}_2$, $\boldsymbol{\nu}_1$, $\boldsymbol{\nu}_2$ and $\boldsymbol{\nu}$ are given by (3.13), (3.14) and (3.12).

Our main tool here is Lemma 5.1 given in the preceding section. To apply the lemma, we reformulate the Markov chain $\{\boldsymbol{X}(t)\} = \{(N_1(t), N_2(t), I_1(t), I_2(t), J_1(t), J_2(t))\}$ defined in (2.1) to make it a quasi-birth-and-death process on the partition $\{\mathcal{L}_2(n_2),\ n_2 = 0, 1, 2, \ldots\}$ of the state space $\mathcal{S}$ into levels according to the number of customers in node 2. Here the n_2-th *level* is the set of states with n_2 customers in node 2, $\mathcal{L}_2(n_2) = \bigcup_{n_1=0}^{\infty} \mathcal{C}(n_1, n_2)$, $n_2 = 0, 1, 2, \ldots$. The level $\mathcal{L}_2(n_2)$ is further partitioned into cells $\{\mathcal{C}(n_1, n_2),\ n_1 = 0, 1, 2, \ldots\}$ according to the number of customers in node 1. If we rearrange states in a lexicographical order of $(n_2, n_1, i_1, i_2, j_1, j_2)$ (here n_1 and n_2 are interchanged from the original ordering), then the transition rate matrix $\boldsymbol{Q}$ of $\{\boldsymbol{X}(t)\}$ takes of the form (5.1) with submatrices $\boldsymbol{Q}_0$, $\boldsymbol{Q}_1$ and $\boldsymbol{Q}_2$ given by

$$\boldsymbol{Q}_0 = \begin{pmatrix} \boldsymbol{I}\otimes\boldsymbol{U}_2\otimes\boldsymbol{I} & \boldsymbol{O} & & \\ \boldsymbol{I}\otimes\boldsymbol{I}\otimes r_{12}\boldsymbol{\sigma}_1^\top\otimes\boldsymbol{I} & \boldsymbol{I}\otimes\boldsymbol{U}_2\otimes\boldsymbol{I}\otimes\boldsymbol{I} & \boldsymbol{O} & \\ & \boldsymbol{I}\otimes\boldsymbol{I}\otimes r_{12}\boldsymbol{\sigma}_1^\top\boldsymbol{b}_1\otimes\boldsymbol{I} & \boldsymbol{I}\otimes\boldsymbol{U}_2\otimes\boldsymbol{I}\otimes\boldsymbol{I} & \ddots \\ & & \ddots & \ddots \end{pmatrix},$$

$$\boldsymbol{Q}_1 = \begin{pmatrix} \boldsymbol{T}_1\oplus\boldsymbol{T}_2\oplus\boldsymbol{S}_2 & \boldsymbol{U}_1\otimes\boldsymbol{I}\otimes\boldsymbol{b}_1\otimes\boldsymbol{I} & & \\ \boldsymbol{I}\otimes\boldsymbol{I}\otimes r_{10}\boldsymbol{\sigma}_1^\top\otimes\boldsymbol{I} & \boldsymbol{T}_1\oplus\boldsymbol{T}_2\oplus\boldsymbol{S}_1\oplus\boldsymbol{S}_2 & \boldsymbol{U}_1\otimes\boldsymbol{I}\otimes\boldsymbol{I}\otimes\boldsymbol{I} & \\ & \boldsymbol{I}\otimes\boldsymbol{I}\otimes r_{10}\boldsymbol{\sigma}_1^\top\boldsymbol{b}_1\otimes\boldsymbol{I} & \boldsymbol{T}_1\oplus\boldsymbol{T}_2\oplus\boldsymbol{S}_1\oplus\boldsymbol{S}_2 & \ddots \\ & & \ddots & \ddots \end{pmatrix},$$

$$\boldsymbol{Q}_2 = \begin{pmatrix} \boldsymbol{I}\otimes\boldsymbol{I}\otimes r_{20}\boldsymbol{\sigma}_2^\top\boldsymbol{b}_2 & \boldsymbol{I}\otimes\boldsymbol{I}\otimes\boldsymbol{b}_1\otimes r_{21}\boldsymbol{\sigma}_2^\top\boldsymbol{b}_2 & & \\ \boldsymbol{O} & \boldsymbol{I}\otimes\boldsymbol{I}\otimes\boldsymbol{I}\otimes r_{20}\boldsymbol{\sigma}_2^\top\boldsymbol{b}_2 & \boldsymbol{I}\otimes\boldsymbol{I}\otimes\boldsymbol{I}\otimes r_{21}\boldsymbol{\sigma}_2^\top\boldsymbol{b}_2 & \\ & \boldsymbol{O} & \boldsymbol{I}\otimes\boldsymbol{I}\otimes\boldsymbol{I}\otimes r_{20}\boldsymbol{\sigma}_2^\top\boldsymbol{b}_2 & \ddots \\ & & \ddots & \ddots \end{pmatrix}.$$

Since $\overline{\boldsymbol{Q}}_i$, $i = 0, 1, 2$, do not appear explicitly in our proof, we omit their exact description. The state probability vector $\boldsymbol{\pi}_2$ corresponding to our ordering of states is also partitioned into subvectors as

$$\boldsymbol{\pi}_2 = \begin{pmatrix} \boldsymbol{\pi}_2(0) & \boldsymbol{\pi}_2(1) & \boldsymbol{\pi}_2(2) & \cdots \end{pmatrix}. \tag{6.1}$$

The subvector $\boldsymbol{\pi}_2(n_2)$ is partitioned into state probability vectors $\boldsymbol{p}(n_1, n_2)$ defined in (3.5) for cells $\mathcal{C}(n_1, n_2)$, $n_1 = 0, 1, 2, \dots,$ as

$$\boldsymbol{\pi}_2(n_2) = (\, \boldsymbol{p}(0, n_2) \quad \boldsymbol{p}(1, n_2) \quad \boldsymbol{p}(2, n_2) \quad \cdots \,), \qquad n_2 = 0, 1, 2, \dots. \tag{6.2}$$

We choose a pair $(a_1, a_2) \in \mathcal{K}_{loop}$ arbitrarily and apply Lemma 5.1 to the partition $\{\mathcal{L}_2(n_2)\}$ with $\xi = \eta_2$. Then the key matrix $\xi^{-1}\boldsymbol{Q}_0 + \boldsymbol{Q}_1 + \xi\boldsymbol{Q}_2$ has the following block tri-diagonal structure:

$$\boldsymbol{D} \;=\; \frac{1}{\eta_2}\boldsymbol{Q}_0 + \boldsymbol{Q}_1 + \eta_2\boldsymbol{Q}_2 = \begin{pmatrix} \overline{\boldsymbol{B}}_1 & \overline{\boldsymbol{B}}_0 & & \\ \overline{\boldsymbol{B}}_2 & \boldsymbol{B}_1 & \boldsymbol{B}_0 & \\ & \boldsymbol{B}_2 & \ddots & \ddots \\ & & \ddots & \end{pmatrix}, \tag{6.3}$$

$$\begin{aligned}
\overline{\boldsymbol{B}}_0 &= \boldsymbol{U}_1 \otimes \boldsymbol{I} \otimes \boldsymbol{b}_1 \otimes \boldsymbol{I} + \boldsymbol{I} \otimes \boldsymbol{I} \otimes \boldsymbol{b}_1 \otimes r_{21}\eta_2\boldsymbol{\sigma}_2^\top\boldsymbol{b}_2,\\
\overline{\boldsymbol{B}}_1 &= \boldsymbol{T}_1 \oplus \left(\boldsymbol{T}_2 + \frac{1}{\eta_2}\boldsymbol{U}_2\right) \oplus \left(\boldsymbol{S}_2 + r_{20}\eta_2\boldsymbol{\sigma}_2^\top\boldsymbol{b}_2\right),\\
\overline{\boldsymbol{B}}_2 &= \boldsymbol{I} \otimes \boldsymbol{I} \otimes \left(r_{10} + \frac{r_{12}}{\eta_2}\right)\boldsymbol{\sigma}_1^\top \otimes \boldsymbol{I},\\
\boldsymbol{B}_0 &= \boldsymbol{U}_1 \otimes \boldsymbol{I} \otimes \boldsymbol{I} \otimes \boldsymbol{I} + \boldsymbol{I} \otimes \boldsymbol{I} \otimes \boldsymbol{I} \otimes r_{21}\eta_2\boldsymbol{\sigma}_2^\top\boldsymbol{b}_2, &(6.4)\\
\boldsymbol{B}_1 &= \boldsymbol{T}_1 \oplus \left(\boldsymbol{T}_2 + \frac{1}{\eta_2}\boldsymbol{U}_2\right) \oplus \boldsymbol{S}_1 \oplus \left(\boldsymbol{S}_2 + r_{20}\eta_2\boldsymbol{\sigma}_2^\top\boldsymbol{b}_2\right),\\
\boldsymbol{B}_2 &= \boldsymbol{I} \otimes \boldsymbol{I} \otimes \left(r_{10} + \frac{r_{12}}{\eta_2}\right)\boldsymbol{\sigma}_1^\top\boldsymbol{b}_1 \otimes \boldsymbol{I}.
\end{aligned}$$

We introduce some notation. Let $\mathcal{L}_2$ denote a replica of $\mathcal{L}_2(n_2)$ for some $n_2 > 0$, and $\mathcal{C}_2(n_1)$ a replica of $\mathcal{C}(n_1, n_2)$. Then $\boldsymbol{D}$ can be regarded as a matrix whose elements are indexed in $\mathcal{L}_2 \times \mathcal{L}_2$. For a row vector $\boldsymbol{y}$ with elements indexed in $\mathcal{L}_2$ we partition it into subvectors according to the partition $\{\mathcal{C}_2(n_1),\ n_1 = 0, 1, 2, \dots\}$ as $\boldsymbol{y} = (\, \boldsymbol{y}(0) \quad \boldsymbol{y}(1) \quad \boldsymbol{y}(2) \quad \cdots \,)$, and define its *decay rate* $\mathrm{dr}[\boldsymbol{y}]$ by

$$\log \mathrm{dr}[\boldsymbol{y}] = \limsup_{n_1 \to \infty} \frac{1}{n_1} \log \boldsymbol{y}(n_1)\boldsymbol{e}^\top. \tag{6.5}$$

For $r = \mathrm{dr}[\boldsymbol{y}]$, if there exists a positive constant C such that

$$\lim_{n \to \infty} \frac{\boldsymbol{y}(n_1)\boldsymbol{e}^\top}{r^{n_1}} = C, \tag{6.6}$$

then we will say $\boldsymbol{y}$ decays geometrically *in strong sense* with rate r. We use this terminology even if $r \geq 1$.

Now we shall construct a positive vector $\boldsymbol{q}$ that satisfies the inequality $\boldsymbol{q}\boldsymbol{D} \leq \boldsymbol{0}$ and has a known decay rate η_1. We shall use the doubly geometric form solution (3.9) to do this. We choose a pair (a_1, a_2) on the loop $\mathcal{K}_{loop}$ arbitrarily. The associated positive vector $\boldsymbol{\nu}$ satisfies the balance equation (3.10), which is rewritten as

$$\boldsymbol{\nu}\left(\frac{1}{\eta_1}\boldsymbol{B}_0 + \boldsymbol{B}_1 + \eta_1\boldsymbol{B}_2\right) = \boldsymbol{0}. \tag{6.7}$$

As a candidate for $\boldsymbol{q}$ in (5.4), consider a row vector

$$\boldsymbol{\gamma} = (\, \boldsymbol{0} \quad \eta_1\boldsymbol{\nu} \quad \eta_1^2\boldsymbol{\nu} \quad \eta_1^3\boldsymbol{\nu} \quad \cdots \,). \tag{6.8}$$

Clearly, $\boldsymbol{\gamma}$ decays geometrically in strong sense with rate η_1. A direct calculation shows that

$$\boldsymbol{\gamma}\boldsymbol{D} = \boldsymbol{\gamma}\left(\frac{1}{\eta_2}\boldsymbol{Q}_0 + \boldsymbol{Q}_1 + \eta_2\boldsymbol{Q}_2\right) = (\,\boldsymbol{\tau}_0 \quad \boldsymbol{\tau}_1 \quad \boldsymbol{0} \quad \boldsymbol{0} \quad \cdots\,), \quad \text{where} \tag{6.9}$$

$$\begin{aligned} \boldsymbol{\tau}_0 &= \overline{\boldsymbol{\nu}}_1 \otimes \overline{\boldsymbol{\nu}}_2 \otimes \boldsymbol{\nu}_2, \\ \boldsymbol{\tau}_1 &= -\overline{\boldsymbol{\nu}}_1\boldsymbol{U}_1 \otimes \overline{\boldsymbol{\nu}}_2 \otimes \boldsymbol{\nu}_1 \otimes \boldsymbol{\nu}_2 - r_{21}\eta_2 g_2(-y_2)\,\overline{\boldsymbol{\nu}}_1 \otimes \overline{\boldsymbol{\nu}}_2 \otimes \boldsymbol{\nu}_1 \otimes \boldsymbol{b}_2. \end{aligned} \tag{6.10}$$

Here we assume, without loss of generality, that $\boldsymbol{\nu}_k$ is normalized so that $\boldsymbol{\nu}_k\boldsymbol{\sigma}_k = g_k(-y_k)$. Unfortunately, $\boldsymbol{\tau}_0 > \boldsymbol{0}$ and the vector $\boldsymbol{\gamma}$ does not satisfy condition (5.4) for $\boldsymbol{q}$. However the vector $\boldsymbol{\gamma}$ will turn out to be a component of a vector $\boldsymbol{q}$ having a known decay rate.

We will derive another candidate vector $\boldsymbol{\ell}$ for $\boldsymbol{q}$ satisfying $\boldsymbol{\ell}D \leq \boldsymbol{0}$ as in (6.20) below. We need some preparation. To use the Markov chain theory, we convert $\boldsymbol{D}$ to a defective transition rate matrix $\boldsymbol{D}'$. Let $\overline{\boldsymbol{\zeta}}_k^\top$ be the positive right eigenvector of the matrix $\left(\boldsymbol{T}_k + \eta_k^{-1}\boldsymbol{U}_k\right)$ associated with eigenvalue x_k, and $\boldsymbol{\zeta}_k^\top$ be that of $\left(\boldsymbol{S}_k + \eta_k\left(r_{k0} + \eta_{k'}^{-1}r_{kk'}\right)\boldsymbol{\sigma}_k^\top\boldsymbol{b}_k\right)$ associated with $-y_k$:

$$\begin{aligned} \left(\boldsymbol{T}_k + \tfrac{1}{\eta_k}\boldsymbol{U}_k\right)\overline{\boldsymbol{\zeta}}_k^\top &= x_k\,\overline{\boldsymbol{\zeta}}_k^\top, \\ \left(\boldsymbol{S}_k + \eta_k\left(r_{k0} + \tfrac{r_{kk'}}{\eta_{k'}}\right)\boldsymbol{\sigma}_k^\top\boldsymbol{b}_k\right)\boldsymbol{\zeta}_k^\top &= -y_k\,\boldsymbol{\zeta}_k^\top. \end{aligned} \tag{6.11}$$

If we set $\boldsymbol{\zeta}^\top = \overline{\boldsymbol{\zeta}}_1^\top \otimes \overline{\boldsymbol{\zeta}}_2^\top \otimes \boldsymbol{\zeta}_1^\top \otimes \boldsymbol{\zeta}_2^\top$, then

$$\left(\frac{1}{\eta_1}\boldsymbol{B}_0 + \boldsymbol{B}_1 + \eta_1\boldsymbol{B}_2\right)\boldsymbol{\zeta}^\top = \boldsymbol{0}^\top. \tag{6.12}$$

We let

$$\boldsymbol{\chi}^\top = \begin{pmatrix} \eta_1\, g_1(-y_1)\,\overline{\boldsymbol{\zeta}}_1^\top \otimes \overline{\boldsymbol{\zeta}}_2^\top \otimes \boldsymbol{\zeta}_2^\top \\ \boldsymbol{\zeta}^\top \\ \eta_1^{-1}\boldsymbol{\zeta}^\top \\ \eta_1^{-2}\boldsymbol{\zeta}^\top \\ \vdots \end{pmatrix}. \tag{6.13}$$

A direct calculation shows that

$$\boldsymbol{D}\boldsymbol{\chi}^\top = \left(\frac{1}{\eta_2}\boldsymbol{Q}_0 + \boldsymbol{Q}_1 + \eta_2\boldsymbol{Q}_2\right)\boldsymbol{\chi}^\top = \begin{pmatrix} \eta_1\, g_1(-y_1)\, y_1\,\overline{\boldsymbol{\zeta}}_1^\top \otimes \overline{\boldsymbol{\zeta}}_2^\top \otimes \boldsymbol{\zeta}_2^\top \\ \boldsymbol{0} \\ \boldsymbol{0} \\ \vdots \end{pmatrix}. \tag{6.14}$$

Here we assume $\boldsymbol{\zeta}_1^\top$ is normalized so that $\boldsymbol{b}_1\boldsymbol{\zeta}_1^\top = g_1(-y_1)$. If $y_1 \leq 0$, then the vector on the right hand side of the above equation is nonpositive. Hereafter we assume that $y_1 < 0$ (for later convenience, we don't include the case $y_1 = 0$ in our assumption). We introduce diagonal matrices $\boldsymbol{Z} = \mathrm{diag}[\boldsymbol{\chi}^\top]$, $\boldsymbol{Z}_0 = \mathrm{diag}[\overline{\boldsymbol{\zeta}}_1^\top \otimes \overline{\boldsymbol{\zeta}}_2^\top \otimes \boldsymbol{\zeta}_2^\top]$ and $\boldsymbol{Z}_1 = \mathrm{diag}[\boldsymbol{\zeta}^\top]$, and set

$$\boldsymbol{D}' = \boldsymbol{Z}^{-1}\boldsymbol{D}\boldsymbol{Z} = \boldsymbol{Z}^{-1}\left(\frac{1}{\eta_2}\boldsymbol{Q}_0 + \boldsymbol{Q}_1 + \eta_2\boldsymbol{Q}_2\right)\boldsymbol{Z}. \tag{6.15}$$

Then $\boldsymbol{D}'$ takes of the form

$$\boldsymbol{D}' = \begin{pmatrix} \overline{\boldsymbol{B}}_1' & \overline{\boldsymbol{B}}_0' & & \\ \overline{\boldsymbol{B}}_2' & \boldsymbol{B}_1' & \boldsymbol{B}_0' & \\ & \boldsymbol{B}_2' & \ddots & \ddots \\ & & \ddots & \end{pmatrix}, \quad \text{where} \tag{6.16}$$

$$\overline{\boldsymbol{B}}'_0 = (\eta_1 g_1(-y_1))^{-1}\boldsymbol{Z}_0^{-1}\overline{\boldsymbol{B}}_0\boldsymbol{Z}_1, \quad \overline{\boldsymbol{B}}'_1 = \boldsymbol{Z}_0^{-1}\overline{\boldsymbol{B}}_1\boldsymbol{Z}_0, \quad \overline{\boldsymbol{B}}'_2 = \eta_1 g_1(-y_1)\boldsymbol{Z}_1^{-1}\overline{\boldsymbol{B}}_2\boldsymbol{Z}_0,$$
$$\boldsymbol{B}'_0 = \eta_1^{-1}\boldsymbol{Z}_1^{-1}\boldsymbol{B}_0\boldsymbol{Z}_1, \quad \boldsymbol{B}'_1 = \boldsymbol{Z}_1^{-1}\boldsymbol{B}_1\boldsymbol{Z}_1 \quad \text{and} \quad \boldsymbol{B}'_2 = \eta_1\boldsymbol{Z}_1^{-1}\boldsymbol{B}_2\boldsymbol{Z}_1.$$

We note that $(\boldsymbol{B}'_0 + \boldsymbol{B}'_1 + \boldsymbol{B}'_2)\boldsymbol{e}^\top = \boldsymbol{0}^\top$ from (6.12), and hence from (6.14)

$$\boldsymbol{D}'\boldsymbol{e}^\top = \eta_1\, g_1(-y_1)\, y_1 \begin{pmatrix} \boldsymbol{e}^\top \\ \boldsymbol{0}^\top \\ \boldsymbol{0}^\top \\ \vdots \end{pmatrix} \leq \boldsymbol{0}^\top. \tag{6.17}$$

This implies that we may consider $\boldsymbol{D}'$ to be a defective transition rate matrix of a Markov chain $\{\boldsymbol{Y}(t)\}$ on the state space $\mathcal{L}_2 \cup \{0\}$ with $\{0\}$ being an absorbing state. Here "defective" means that $\boldsymbol{D}'$ is a matrix formed from a complete transition rate matrix by deleting the row and column corresponding to the absorbing state. Since we have assumed $y_1 < 0$, the transition rate $-\eta_1\, g_1(-y_1)\, y_1$ from a state in $\mathcal{C}_2(0) \subset \mathcal{L}_2$ to the absorbing state is strictly positive. Now we assume $\{\boldsymbol{Y}(t)\}$ starts from a state in $\mathcal{C}_2(0)$ selected randomly, i.e. $\mathrm{P}\{\boldsymbol{Y}(0) = (0, i_1, i_2, 0, j_2)\} = 1/v$ for $(0, i_1, i_2, 0, j_2) \in \mathcal{C}_2(0)$, where $v = |\mathcal{C}_2(0)|$ is the number of states in $\mathcal{C}_2(0)$, and we let

$$\ell'(n_1, i_1, i_2, j_1, j_2) = \int_0^\infty \mathrm{P}\{\boldsymbol{Y}(t) = (n_1, i_1, i_2, j_1, j_2)\} dt. \tag{6.18}$$

This quantity can be interpreted as the expected time spent in state $(n_1, i_1, i_2, j_1, j_2)$ until absorption or drift to ∞ occurs. From the irreducibility of $\{\boldsymbol{X}(t)\}$ we see that $\ell'(n_1, i_1, i_2, j_1, j_2)$ is positive, and it is finite for any state $(n_1, i_1, i_2, j_1, j_2) \in \mathcal{L}_2$ since the chain has an absorbing state. From the Kolmogorov's forward equation, the row vector $\boldsymbol{\ell}' = (\, \ell'(n_1, i_1, i_2, j_1, j_2)\,)$ satisfies the equation

$$\boldsymbol{\ell}'\boldsymbol{D}' = (\, -v^{-1}\boldsymbol{e} \quad \boldsymbol{0} \quad \boldsymbol{0} \quad \cdots \,). \tag{6.19}$$

If we put $\boldsymbol{\ell} = \boldsymbol{\ell}'\boldsymbol{Z}^{-1}$, we have

$$\boldsymbol{\ell}\boldsymbol{D} = (\, -(v\eta_1 g_1(-y_1))^{-1}\boldsymbol{e}\boldsymbol{Z}_0^{-1} \quad \boldsymbol{0} \quad \boldsymbol{0} \quad \cdots \,), \tag{6.20}$$

and this shows that $\boldsymbol{\ell}$ can serve as a candidate vector for $\boldsymbol{q}$ in Lemma 5.1 since it satisfies inequality (5.4) with $\xi = \eta_2$.

The next task is to examine the decay rate of $\boldsymbol{\ell}$. We partition $\boldsymbol{\ell}'$ and $\boldsymbol{\ell}$ into subvectors corresponding to cells as $\boldsymbol{\ell}' = (\, \boldsymbol{\ell}'(0) \quad \boldsymbol{\ell}'(1) \quad \boldsymbol{\ell}'(2) \quad \cdots \,)$ and $\boldsymbol{\ell} = (\, \boldsymbol{\ell}(0) \quad \boldsymbol{\ell}(1) \quad \boldsymbol{\ell}(2) \quad \cdots \,)$. Then from the block-tri-diagonal form (6.16) of $\boldsymbol{D}'$ and from the definition (6.18), we see that $\boldsymbol{\ell}'$ takes a matrix-geometric form. To see this, for $m > 0$ and $\tau > 0$, we denote by $\hat{\boldsymbol{R}}'^{(m)}$ a matrix of order $\mathcal{C}_2(1) \times \mathcal{C}_2(1)$ with

$$\int_0^\infty \mathrm{P}\{\boldsymbol{Y}(\tau + t) = (m+1, i'_1, i'_2, j'_1, j'_2),\ \boldsymbol{Y}(s) \text{ does not visit level 1 during } (\tau, \tau + t) \text{ after it leaves level } 1 \,|\, \boldsymbol{Y}(\tau) = (1, i_1, i_2, j_1, j_2)\}\, dt \tag{6.21}$$

in its $(\,(i_1, i_2, j_1, j_2),\ (i'_1, i'_2, j'_1, j'_2)\,)$th element. Clearly the quantity does not depend on τ and is finite from the same reason as $\ell'(n_1, i_1, i_2, j_1, j_2)$ in (6.18). Then following the standard probabilistic discussion on sample paths (see Neuts [12]), we see that, if we put $\hat{\boldsymbol{R}}' = \hat{\boldsymbol{R}}'^{(1)}$,

$$\boldsymbol{\ell}'(m+1) = \boldsymbol{\ell}'(1)\, \hat{\boldsymbol{R}}'^{(m)} = \boldsymbol{\ell}'(1)\, \left(\hat{\boldsymbol{R}}'\right)^m, \tag{6.22}$$

and that $\hat{\boldsymbol{R}}'$ coincides with the nonnegative minimal solution of the matrix quadratic equation

$$\boldsymbol{B}_0' + \hat{\boldsymbol{R}}'\boldsymbol{B}_1' + \hat{\boldsymbol{R}}'^2\boldsymbol{B}_2' = \boldsymbol{O}. \tag{6.23}$$

Since $m > 0$ is arbitrary, (6.22) shows that $\boldsymbol{\ell}'$ decays geometrically in strong sense with rate equal to the PFE of rate matrix $\hat{\boldsymbol{R}}'$, i.e. $\mathrm{dr}\,[\boldsymbol{\ell}'] = \mathrm{pf}\left[\hat{\boldsymbol{R}}'\right]$. From Lemmas 1.3.2 and 1.3.4 of [12], we know that the equation $\mathrm{pf}\left[\frac{1}{z}\boldsymbol{B}_0' + \boldsymbol{B}_1' + z\boldsymbol{B}_2'\right] = 0$ for z has at most two positive roots and $\mathrm{pf}[\hat{\boldsymbol{R}}']$ is the minimal one of the two. Since $z = 1$ is a root of the equation, $\mathrm{pf}[\hat{\boldsymbol{R}}'] \le 1$ and hence $\mathrm{dr}\,[\boldsymbol{\ell}'] \le 1$. Since $\boldsymbol{\ell} = \boldsymbol{\ell}'\boldsymbol{Z}^{-1}$,

$$\boldsymbol{\ell}(n_1) = \eta_1^{n_1-1}\,\boldsymbol{\ell}'(n_1)\,\boldsymbol{Z}_1^{-1}, \qquad n_1 = 1, 2, 3, \ldots. \tag{6.24}$$

This shows that $\boldsymbol{\ell}$ also decays geometrically in strong sense with rate less than or equal to η_1, i.e., $\mathrm{dr}[\boldsymbol{\ell}] \le \eta_1$.

This $\boldsymbol{\ell}$, however, is not a satisfactory candidate for $\boldsymbol{q}$, because to check the condition (5.8) in Lemma 5.2 we need to know the decay rate of $\boldsymbol{q}$ exactly. So, instead of $\boldsymbol{\ell}$ alone, we use $\boldsymbol{\gamma}$ defined in (6.8) together. Recall that $\boldsymbol{\gamma}$ is nonnegative, $\boldsymbol{\gamma D}$ is nonpositive except for the first subvector $\boldsymbol{\tau}_0$ in (6.9), and $\boldsymbol{\gamma}$ decays geometrically in strong sense with rate η_1. On the other hand, $\boldsymbol{\ell}$ is positive, $\boldsymbol{\ell D}$ is nonpositive with the first subvector being strictly negative, and $\boldsymbol{\ell}$ decays geometrically in strong sense with rate less than or equal to η_1. So if we choose a sufficiently small positive number ε, the vector

$$\boldsymbol{q} = \varepsilon\boldsymbol{\gamma} + \boldsymbol{\ell}, \tag{6.25}$$

is positive, satisfies the inequality $\boldsymbol{qD} \le \boldsymbol{0}$, and decays geometrically in strong sense with rate exactly equal to η_1 as we have requested.

The assumptions we have made for the above discussion are that (i) $(a_1, a_2) \in \mathcal{K}_{loop}$ (or equivalently $\kappa(a_1, a_2) = 0$) and (ii) $y_1 < 0$ (or equivalently $a_1 > h_2(a_2)$ from (4.6) and (4.7)). Hence we have the following

Lemma 6.1 *Suppose that $(a_1, a_2) \in \mathcal{K}_{loop}$ satisfies the condition $a_1 > h_2(a_2)$. Then for the partition $\{\mathcal{L}_2(n_2)\}$, there exists a positive vector $\boldsymbol{q}$ that satisfies condition (5.4) in Lemma 5.1 with $\xi = \eta_2 = e^{a_2}$ and decays geometrically in strong sense with rate $\eta_1 = e^{a_1}$.*

Using Lemma 5.2 and this lemma, we can derive two key lemmas for the proof of our main theorem. Before stating the lemmas, we introduce another partition of the state space by the number of customers in node 1. Let $\{\mathcal{L}_1(n_1),\ n_1 = 0, 1, 2, \ldots\}$ be the partition of $\mathcal{S}$, where $\mathcal{L}_1(n_1) = \bigcup_{n_2=0}^{\infty} \mathcal{C}(n_1, n_2)$, and $\{\mathcal{C}(n_1, n_2),\ n_2 = 0, 1, 2, \ldots\}$ is a partition of $\mathcal{L}_1(n_1)$ into cells. We denote by $\mathcal{L}_1$ a replica of $\mathcal{L}_1(n_1)$. We also denote the state probability vector by $\boldsymbol{\pi}_1$ subjecting to the lexicographical order of $(n_1, n_2, i_1, i_2, j_1, j_2)$, and according to the partition $\{\mathcal{L}_1(n_1)\}$ we divide it into subvectors as $\boldsymbol{\pi}_1 = (\ \boldsymbol{\pi}_1(0)\ \ \boldsymbol{\pi}_1(1)\ \ \boldsymbol{\pi}_1(2)\ \ \cdots\ )$, where $\boldsymbol{\pi}_1(n_1) = (\ \boldsymbol{p}(n_1, 0)\ \ \boldsymbol{p}(n_1, 1)\ \ \boldsymbol{p}(n_1, 2)\ \ \cdots\ )$, $n_1 = 0, 1, 2, \ldots$. The vector $\boldsymbol{\pi}_1$ is essentially the same as $\boldsymbol{\pi}_2$ in (6.1) except for ordering of elements. For row vectors with elements indexed in $\mathcal{L}_1$, the decay rate is defined in a similar manner to (6.5).

The first lemma we shall prove is an application of Lemma 6.1 to the case $a_1 = 0$. Note that we have defined b_2^0 in (4.23). We let $\eta_2^0 = \exp\{b_2^0\}$.

Lemma 6.2 *If $b_2^0 < 0$, then $\log \mathrm{dr}\,[\boldsymbol{\pi}_1(1)] \le b_2^0$, or equivalently, $\mathrm{dr}\,[\boldsymbol{\pi}_1(1)] \le \eta_2^0$.*

Proof Since $r_{12} > 0$ and $b_2^0 < 0$, we have $h_2(b_2^0) < 0$. Hence the pair $(0, b_2^0)$ satisfies the condition of Lemma 6.1, and there exists a positive vector $\boldsymbol{q}$ that satisfies (5.4) with $\xi = \eta_2^0$

and decays geometrically in strong sense with rate $e^0 = 1$. So there exists a positive constant c such that $\boldsymbol{q} > c^{-1}\boldsymbol{e}$, and since $\boldsymbol{\pi}_2(1) \le \boldsymbol{e}$, the condition (5.8) in Lemma 5.2 is satisfied. Then from (5.9) we have

$$\boldsymbol{\pi}_2(n_2) \le c \left(\eta_2^0\right)^{n_2-1} \boldsymbol{q}, \qquad n_2 = 1, 2, \dots . \tag{6.26}$$

Considering the second subvector $\boldsymbol{p}(1, n_2)$ of $\boldsymbol{\pi}_2(n_2)$, we see that $\boldsymbol{p}(1, n_2) \le c\,(\eta_2^0)^{n_2-1}\,\boldsymbol{q}(1)$, where $\boldsymbol{q}(1)$ is the second subvector of $\boldsymbol{q}$ in the representation $\boldsymbol{q} = (\ \boldsymbol{q}(0)\ \ \boldsymbol{q}(1)\ \ \boldsymbol{q}(2)\ \ \cdots\ )$ according to the partition $\{\mathcal{C}_2(n_1)\}$ of $\mathcal{L}_2$. Since n_2 is arbitrary, we have $\mathrm{dr}\,[\boldsymbol{\pi}_1(1)] \le \eta_2^0$. $\diamondsuit$

The next lemma is to get a new upper bound for η_2^* when an upper bound of η_1^* is known.

Lemma 6.3 *Suppose that* $(a_1, a_2) \in \mathcal{K}_{loop}$ *satisfies the condition of Lemma 6.1, namely* $a_1 > h_2(a_2)$. *If* $a_1 < 0$ *and* $\mathrm{dr}\,[\boldsymbol{\pi}_2(1)] < \eta_1$ $(= e^{a_1} < 1)$, *then* $\eta_2^* \le \eta_2$ $(= e^{a_2} < 1)$. *If* $\eta_1^* < \eta_1$ $(= e^{a_1} < 1)$, *then the same inequality holds.*

Proof The vector $\boldsymbol{q}$ in (6.25) has decay rate η_1. So, if $\mathrm{dr}\,[\boldsymbol{\pi}_2(1)] < \eta_1$, the condition (5.8) of Lemma 5.2 is clearly satisfied with $\boldsymbol{\pi}_2(1)$ in place of $\boldsymbol{\pi}(1)$. From the assumption $a_1 < 0$, $\boldsymbol{q}\boldsymbol{e}^\top$ is finite since $\eta_1 < 1$. Hence we have the desired result from Lemmas 5.1, 5.2 and 6.1. If $\eta_1^* < \eta_1 < 1$, the condition $\mathrm{dr}\,[\boldsymbol{\pi}_2(1)] < \eta_1$ is trivially satisfied. $\diamondsuit$

Interchanging roles of node 1 and node 2 in Lemmas 6.2 and 6.3, we obtain similar results for $\mathrm{dr}\,[\boldsymbol{\pi}_2(1)]$ and η_1^*. Notice that in Lemmas 6.2 and 6.3 the assumption $r_{12} > 0$ is crucial, because, if $r_{12} = 0$, then $h_2(a_2) \equiv 0$ and we cannot choose any a_2 such that $h_2(a_2) < 0$.

Lemma 6.4 *Assume that* $r_{21} > 0$. (i) *If* $b_1^0 < 0$, *then* $\mathrm{dr}\,[\boldsymbol{\pi}_2(1)] \le \eta_1^0 = \exp\{b_1^0\}$. (ii) *For* $(a_1, a_2) \in \mathcal{K}_{loop}$ *such that* $h_1(a_1) < a_2 < 0$, *the inequality* $\eta_1^* \le \eta_1$ $(= e^{a_1} < 1)$ *holds if* $\mathrm{dr}\,[\boldsymbol{\pi}_1(1)] < \eta_2$ $(= e^{a_2} < 1)$ *or if* $\eta_2^* < \eta_2$ $(= e^{a_2} < 1)$.

Using these lemmas, we prove Theorem 4.1 first for the case $r_{21} = 0$ and then for the case $r_{21} > 0$. In the proofs, we use some properties of sets $\mathcal{E}_k$, $\mathcal{F}_k$, $\mathcal{F}'_k$ and $\mathcal{G}_k(b_{k'})$, the latter two are defined afterwards. These properties are proved in the next section.

Proof of Theorem 4.1 for the case $\boldsymbol{r_{21} = 0}$ Proposition 4.2 says $\overline{\eta}_1 = \eta_1^* < 1$. Hence (4.17) trivially holds for $k = 1$. We shall prove it for $k = 2$. From Lemma 7.3, $\log \overline{\eta}_1 = b_1^{\mathcal{E}_1} = b_1^0 < 0$. Hence, the set $\mathcal{F}_2$ in (4.15) reduces to

$$\mathcal{F}_2 = \{(a_1, a_2) \in \mathcal{K}_{loop}\ :\ a_2 < 0 \text{ and } \max\{h_2(a_2),\ b_1^0\} \le a_1 \le 0\}. \tag{6.27}$$

Now we consider a slightly different set

$$\mathcal{F}'_2 = \{(a_1, a_2) \in \mathcal{K}_{loop}\ :\ \max\{h_2(a_2),\ b_1^0\} < a_1 < 0\}. \tag{6.28}$$

We see that, for any point $(a_1, a_2) \in \mathcal{F}'_2$, the conditions of Lemma 6.3 are satisfied, and hence

$$\eta_2^* \le e^{a_2} < 1. \tag{6.29}$$

Taking infimum over such points, we get a bound for η_2^*:

$$\eta_2^* \le \exp\left\{b_2^{\mathcal{F}'_2}\right\}, \quad \text{where} \quad b_2^{\mathcal{F}'_2} = \inf\{a_2\ :\ \exists\, a_1 \text{ such that } (a_1, a_2) \in \mathcal{F}'_2\}. \tag{6.30}$$

As will be shown in Lemma 7.3, $\mathcal{F}_2$ (and also $\mathcal{F}'_2$) is an arc of $\mathcal{K}_{loop}$, and $\mathcal{F}_2$ and $\mathcal{F}'_2$ only differ in their one or two end points. Hence the infimum $b_2^{\mathcal{F}'_2}$ in (6.30) is equal to the infimum $b_2^{\mathcal{F}_2}$ in (4.16), and the inequality in (6.30) is equivalent to $\eta_2^* \le \overline{\eta}_2 = \exp\{b_2^{\mathcal{F}_2}\}$. $\diamondsuit$

The case $r_{21} > 0$ can be proved using a similar idea. However in this case we know neither η_1^* nor η_2^*. So, this time, we have to construct an alternating sequence of a_1 and a_2 converging to $b_1^{\mathcal{F}_1}$ and $b_2^{\mathcal{F}_2}$, respectively. To make our exposition clearer, we introduce a set $\mathcal{G}_k(b_{k'})$ and a function $\chi_k(b_{k'})$ for $b_{k'} < 0$. The set $\mathcal{G}_k(b_{k'})$ plays a similar role to $\mathcal{F}_2'$ in the proof of the case $r_{21} = 0$ above.

Assume that $r_{21} > 0$. For an arbitrary number $b_{k'} < 0$ we let

$$\mathcal{G}_k(b_{k'}) = \{(a_1, a_2) \in \mathcal{K}_{loop} : \max\{h_k(a_k), b_{k'}\} < a_{k'} < 0\}, \quad \text{and} \tag{6.31}$$

$$\chi_k(b_{k'}) = \inf\{a_k : \exists a_{k'} \text{ such that } (a_1, a_2) \in \mathcal{G}_k(b_{k'})\}. \tag{6.32}$$

Explanative descriptions of $\mathcal{G}_k(b_{k'})$ and $\chi_k(b_{k'})$ will be given in Lemma 7.4. The set $\mathcal{G}_k(b_{k'})$ is a subset of $\mathcal{E}_k$ and nonempty. Hence $\chi_k(b_{k'})$ is well defined for any $b_{k'} < 0$. We regard this χ_k as a function of $b_{k'}$. It is negative and continuous. Further, it is strictly increasing in the interval $[b_{k'}^{\mathcal{E}_k}, 0)$ and constant equal to $b_k^{\mathcal{E}_k}$ in the interval $(-\infty, b_{k'}^{\mathcal{E}_k}]$ (see Figure 6.1 for an example). Since $\mathcal{E}_k$ differs from $\mathcal{G}_k(b_{k'}^{\mathcal{E}_k})$ in only one or two end points and $\mathcal{F}_k$ differs from $\mathcal{G}_k(b_{k'}^{\mathcal{E}_{k'}})$ in one or two end points as stated in Lemma 7.4, it is easily seen that $\chi_k(b_{k'}^{\mathcal{E}_k}) = b_k^{\mathcal{E}_k}$ and $\chi_k(b_{k'}^{\mathcal{E}_{k'}}) = b_k^{\mathcal{F}_k}$. Using this function $\chi_k(b_{k'})$, Lemma 6.3 and (ii) of Lemma 6.4 above are restated as follows.

Lemma 6.5 *Assume that* $r_{kk'} > 0$. *For* $b_{k'} < 0$, *if either* $\log \mathrm{dr}\,[\boldsymbol{\pi}_k(1)] \le b_{k'}$ *or* $\log \eta_{k'}^* \le b_{k'}$, *then* $\log \eta_k^* \le \chi_k(b_{k'})\,(< 0)$ *or equivalently* $\eta_k^* \le \exp\{\chi_k(b_{k'})\}\,(< 1)$.

Proof Any point $(a_1, a_2) \in \mathcal{G}_k(b_{k'})$ satisfies the condition of (ii) of Lemma 6.4. Hence $\eta_k^* \le e^{a_k}$. Taking the infimum in $\mathcal{G}_k(b_{k'})$ we obtain the inequality to be proved. (To apply the lemmas here, we cannot include a point with $\max\{h_k(a_k), b_{k'}\} = a_{k'}$ in the set $\mathcal{G}_k(b_{k'})$. This is the reason why we have introduced slightly different sets from $\mathcal{E}_k$ and $\mathcal{F}_k$.) ♢

Now we proceed to the proof of our main theorem for the case $r_{21} > 0$.

Proof of Theorem 4.1 for the case $r_{21} > 0$ Lemma 7.1 ensures that either $b_1^0 < 0$ or $b_2^0 < 0$. Without loss of generality, we assume that $b_2^0 < 0$. If $b_2^0 = 0$, we may exchange the roles of node 1 and node 2 in the following argument. We introduce an alternating sequence $\{b_2^{(0)}, b_1^{(1)}, b_2^{(1)}, b_1^{(2)}, b_2^{(2)}, b_1^{(3)}, \cdots\}$ of negative numbers by relations

$$\begin{cases} b_2^{(0)} = b_2^0, \\ b_1^{(m)} = \chi_1(b_2^{(m-1)}), \qquad m = 1, 2, \cdots, \\ b_2^{(m)} = \chi_2(b_1^{(m)}), \qquad m = 1, 2, \cdots. \end{cases} \tag{6.33}$$

Since $b_2^0 < 0$, this sequence is well-defined (see Figure 6.1 for an example). From Lemma 6.2 we have $\log \mathrm{dr}\,[\boldsymbol{\pi}_1(1)] \le b_2^0 = b_2^{(0)}$. Then applying Lemma 6.5 with $k = 1$ and $b_{k'} = b_2^{(0)}$, we have $\log \eta_1^* \le \chi_1(b_2^{(0)}) = b_1^{(1)}$. Applying Lemma 6.5 again with $k = 2$ and $b_{k'} = b_1^{(1)}$, we have $\log \eta_2^* \le \chi_2(b_1^{(1)}) = b_2^{(1)}$. Applying Lemma 6.5 iteratively in this manner, we have

$$\log \eta_k^* \le b_k^{(m)}, \qquad m = 1, 2, \cdots. \tag{6.34}$$

Now we show that the subsequence $\{b_k^{(m)}, m = 1, 2, \ldots\}$ converges to $b_k^{\mathcal{F}_k}$. Consider graphs $a_1 = \chi_1(a_2)$ and $a_2 = \chi_2(a_1)$ on the (a_1, a_2)-plane (for an example, see Figure 6.1). The graph $a_1 = \chi_1(a_2)$ consists of the upper left arc of the loop $\mathcal{K}_{loop}$ between $(b_1^0, 0)$ and $(b_1^{\mathcal{E}_1}, b_2^{\mathcal{E}_1})$ and a semi-infinite segment of a straight line $a_1 = b_1^{\mathcal{E}_k}$ starting from $(b_1^{\mathcal{E}_1}, b_2^{\mathcal{E}_1})$.

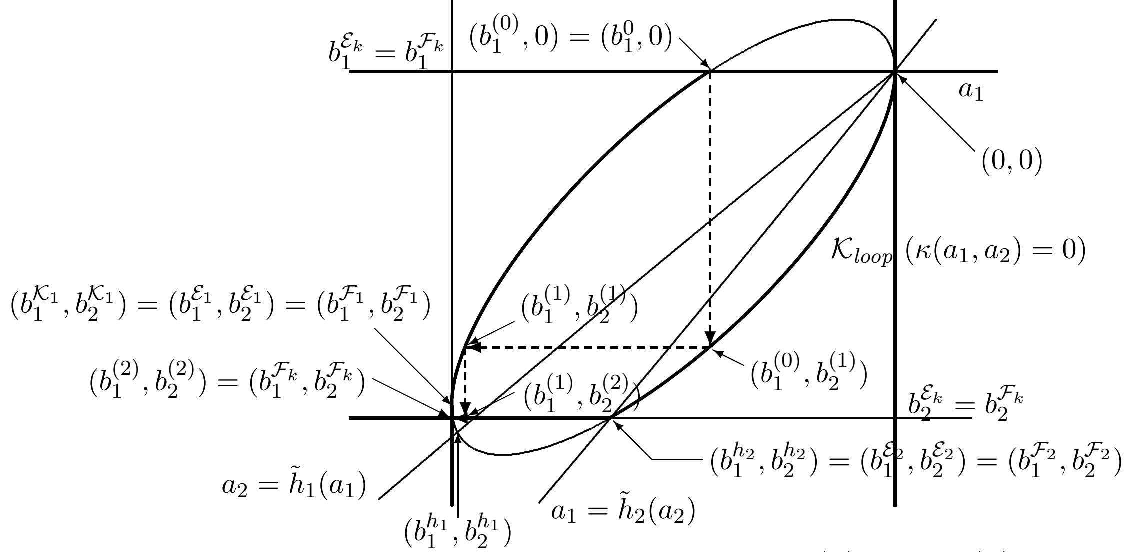

Figure 6.1: Convergence of subsequences $\{b_1^{(m)}\}$ and $\{b_2^{(m)}\}$

The graph $a_2 = \chi_2(a_1)$ also consists of a lower right arc and a segment of a straight line. Hence the graphs intersect once, and from (4.28) the coordinates of the intersection are given by $(b_1^{\mathcal{F}_k}, b_2^{\mathcal{F}_k})$. Thus the subsequences $\{b_1^{(m)}\}$ and $\{b_2^{(m)}\}$ converge to the corresponding coordinates of the intersection. In this case, the convergence is taken in finite steps, because at least one of graphs is a straight line in a neighborhood of the intersection. $\diamondsuit$

Remark 6.1 *Our proof of Theorem 4.1 largely depends on the model structure. One may think it would be possible to make a similar discussion using the matrix structure of the Markov chain* $\{\boldsymbol{X}(t)\}$ *only. It should be, however, noted that our construction of the vector* $\boldsymbol{q}$ *exploits transition structures of not only nonboundary cells* $\mathcal{C}(n_1, n_2)$, $n_1, n_2 = 1, 2, \ldots$, *but also boundary cells* $\mathcal{C}(n, 0)$ *and* $\mathcal{C}(0, n)$, $n = 1, 2, \ldots$. *If, for example, the service time distribution is different from* PH_k *when the number of customers in node* k *is equal to 1 (requiring a special service, etc.), then our proof has to be changed, though it is expected that the same upper bound* $\overline{\eta}_k$ *is obtained. In this sense, the study on the effects of transition structures of boundary cells is remained for future work.*

7. Properties of $\mathcal{K}$-, $\mathcal{E}$-, $\mathcal{F}$- and $\mathcal{G}$-sets

In the proofs and discussions of Theorem 4.1 we have postponed the proof of some properties of sets $\mathcal{E}_k$, $\mathcal{F}_k$, $\mathcal{F}'_2$ and $\mathcal{G}_k(b_{k'})$. We shall prove them here. We start with examining precise properties of the curves $\kappa(a_1, a_2) = 0$ ($\mathcal{K}_{loop}$) and $h_k(a_k) = a_{k'}$ defined in (4.10) and (4.8). The first lemma is related to the curves and straight lines in (4.21) on the (a_1, a_2)-plane. Note that the sets $\mathcal{K}$ and $\mathcal{K}_k$ are defined in (4.22).

Lemma 7.1 *The curves and straight lines in* (4.21) *intersect at the origin* $(a_1, a_2) = (0, 0)$ *and satisfy the following properties:*

(i) *The region* $\mathcal{K}$ *is convex. The curve* $\kappa(a_1, a_2) = 0$ ($\mathcal{K}_{loop}$) *is a loop passing through the origin.*

(ii) *Any tangential line of* $\mathcal{K}_{loop}$ *is tangent to it at a single point. Therefore, for a given number b, the straight line* $a_k = b$ *either intersects* $\mathcal{K}_{loop}$ *twice, is tangent to* $\mathcal{K}_{loop}$ *at a single point, or does not meet* $\mathcal{K}_{loop}$.

(iii) *The straight line* $a_{k'} = 0$ *either intersects with* $\mathcal{K}_{loop}$ *at two points, one of which is the origin, or is tangent to* $\mathcal{K}_{loop}$ *at the origin.* (*The smaller k-th coordinate of the two intersections is denoted as* b_k^0.)

(iv) *Either* $b_1^0 < 0$ *or* $b_2^0 < 0$.

(v) *The curve* $h_k(a_k) = a_{k'}$ *intersects with* $\mathcal{K}_{loop}$ *twice, at the origin and at a point* $(b_1^{h_k}, b_2^{h_k})$. *If* $r_{k'k} > 0$ *then both* $b_1^{h_k}$ *and* $b_2^{h_k}$ *are negative.*

(vi) $a_k \le h_k(a_k)$ *for any* $a_k < 0$*, and hence* $b_k^{h_k} \le b_{k'}^{h_k}$. *The equality holds only when* $r_{k'k} = 1$.

(vii) $h_k(h_{k'}(a_k)) < a_k$ *for any* $a_k < 0$. *Either* $a_1 < h_1(a_1)$ *for any* $a_1 < 0$ *or* $a_2 < h_2(a_2)$ *for any* $a_2 < 0$. *Especially,* $b_1^{h_1} < b_1^{h_2}$ *or* $b_2^{h_2} < b_2^{h_1}$.

Proof It is clear from definitions (4.8) and (4.10) that the curves and straight lines in (4.21) intersect at the origin. The other properties are proved in the following manner.

(i) From the monotonicity and the convexity/concavity of functions ϕ_k, ψ_k and h_k

$$\frac{\partial^2}{\partial a_1^2}\kappa(a_1,a_2) > 0 \quad \text{and} \quad \left\{\frac{\partial^2}{\partial a_1^2}\kappa(a_1,a_2)\right\}\cdot\left\{\frac{\partial^2}{\partial a_2^2}\kappa(a_1,a_2)\right\} - \left\{\frac{\partial^2}{\partial a_1 \partial a_2}\kappa(a_1,a_2)\right\}^2 > 0 \quad (7.1)$$

on the whole plane. Hence the function κ is strictly convex, and the set $\mathcal{K}$ is convex. The boundedness of $\mathcal{K}$ can be proved using (4.6) and (4.9). Since $\mathcal{K}$ is convex and bounded, $\kappa(a_1, a_2) = 0$ forms a loop.

(ii) From the strict convexity of the function κ, the statement follows.

(iii) This is a direct consequence of (ii) with the fact $\kappa(0,0) = 0$.

(iv) Since $\frac{d}{da_1}\kappa(a_1,0)\Big|_{a_1=0} = -\lambda_1 + \mu_1 - r_{21}\mu_2$, $b_1^0 < 0$ if and only if $\mu_1 > \lambda_1 + r_{21}\mu_2$. Similarly, $b_2^0 < 0$ if and only if $\mu_2 > \lambda_2 + r_{12}\mu_1$. Suppose that $b_1^0 \ge 0$. Then from the stability condition $\rho_2 < 1$ in (2.3) we can easily see that $\mu_2 > \lambda_2 + r_{12}\mu_1$, and this implies that $b_2^0 < 0$.

(v) It is easily seen that the function $\kappa(a_1, h_1(a_1))$ of a_1 is strictly convex, and that $\kappa(0, h_1(0)) = 0$ and $\lim_{a_1 \to \pm\infty} \kappa(a_1, h_1(a_1)) = +\infty$. Since $\rho_1 < 1$, $\frac{d}{da_1}\kappa(a_1, h_1(a_1))\Big|_{a_1=0} = -(\lambda_1 + r_{21}\lambda_2) + (1 - r_{12}r_{21})\mu_1 > 0$. Thus the equation $\kappa(a_1, h_1(a_1)) = 0$ has a negative root $b_1^{h_1}$ other than 0. When $r_{k'k} > 0$, from the monotonicity of h, it is clear that $b_2^{h_1} = h(b_1^{h_1}) < 0$.

(vi) The statements are trivial from the definition of h_k in (4.8).

(vii) Since $r_{12}r_{21} < 1$ from the stability condition, the second statement is a direct consequence of (vi). Then the first and the third statements are trivial. ◇

The next lemma is related to the sets $\mathcal{K}$, $\mathcal{K}_k$, $\mathcal{E}_k$ and $\mathcal{F}_k$.

Lemma 7.2 $\mathcal{E}_1$ *and* $\mathcal{E}_2$ *are disjoint. For the sets* $\mathcal{K}_1$, $\mathcal{E}_1$ *and* $\mathcal{F}_1$*, the following properties hold. Similar properties hold for sets* $\mathcal{K}_2$, $\mathcal{E}_2$ *and* $\mathcal{F}_2$.

(i) *The infimum* $b_1^{\mathcal{K}(1)}$ *of* a_1 *in* $\mathcal{K}$ *is attained at a single point* $(b_1^{\mathcal{K}(1)}, b_2^{\mathcal{K}(1)})$. *The infimum* $b_1^{\mathcal{K}_1}$ *of* a_1 *in* $\mathcal{K}_1$ *is attained at a single point* $(b_1^{\mathcal{K}_1}, b_2^{\mathcal{K}_1})$. *If* $b_2^{\mathcal{K}(1)} \le 0$*, then* $(b_1^{\mathcal{K}_1}, b_2^{\mathcal{K}_1}) = (b_1^{\mathcal{K}(1)}, b_2^{\mathcal{K}(1)})$. *If* $b_2^{\mathcal{K}(1)} \ge 0$*, then* $(b_1^{\mathcal{K}_1}, b_2^{\mathcal{K}_1}) = (b_1^0, 0)$.

(ii) $\mathcal{E}_1$ *is the arc of* $\mathcal{K}_{loop}$ *between points* $(b_1^0, 0)$ *and* $(b_1^{h_1}, b_2^{h_1})$ *in the region* $\{(a_1, a_2) : a_1 \le a_2 \le 0\}$. *The end point* $(b_1^{h_1}, b_2^{h_1})$ *belongs to the set. The other end point* $(b_1^0, 0)$ *belongs to the set if and only if* $b_1^0 < 0$.

(iii) *The infimum* $b_1^{\mathcal{E}_1}$ *of* a_1 *in* $\mathcal{E}_1$ *is attained at* $(b_1^{h_1}, b_2^{h_1})$ *or at* $(b_1^{\mathcal{K}_1}, b_2^{\mathcal{K}_1})$. *If* $b_2^{h_1,c} \le b_2^{h_1}$*, then* $(b_1^{\mathcal{E}_1}, b_2^{\mathcal{E}_1}) = (b_1^{h_1}, b_2^{h_1})$. *If* $b_2^{h_1,c} \ge b_2^{h_1}$*, then* $(b_1^{\mathcal{E}_1}, b_2^{\mathcal{E}_1}) = (b_1^{\mathcal{K}_1}, b_2^{\mathcal{K}_1})$.

(iv) *If* $b_2^{\mathcal{E}_1} \ge b_2^{h_2}$*, then* $\mathcal{F}_1 = \mathcal{E}_1$. *If* $b_2^{\mathcal{E}_1} \le b_2^{h_2}$*, then* $\mathcal{F}_1$ *is the arc of* $\mathcal{K}_{loop}$ *between points* $(b_1^0, 0)$ *and* $(b_1^{h_2,c}, b_2^{h_2})$. *In the latter case, the end point* $(b_1^{h_2,c}, b_2^{h_2})$ *belongs to* $\mathcal{F}_1$*, but the other end point* $(b_1^0, 0)$ *belongs to the set if and only if* $b_1^0 < 0$.

(v) *The infimum* $b_1^{\mathcal{F}_1}$ *of* a_1 *in* $\mathcal{F}_1$ *is given by* $b_1^{\mathcal{E}_1}$ *if* $b_2^{\mathcal{E}_1} \ge b_2^{\mathcal{E}_2}$*, and given by* $b_1^{h_2,c}$ *if* $b_2^{\mathcal{E}_1} \le b_2^{\mathcal{E}_2}$.

Proof (i) The first statement is a direct consequence of (ii) of Lemma 7.1. If $b_2^{\mathcal{K}(1)} \le 0$, then $(b_1^{\mathcal{K}(1)}, b_2^{\mathcal{K}(1)}) \in \mathcal{K}_1$ and hence $(b_1^{\mathcal{K}_1}, b_2^{\mathcal{K}_1}) = (b_1^{\mathcal{K}(1)}, b_2^{\mathcal{K}(1)})$. If $b_2^{\mathcal{K}(1)} \ge 0$, then $(b_1^{\mathcal{K}_1}, b_2^{\mathcal{K}_1}) = (b_1^0, 0)$ from the convexity of $\mathcal{K}$. Thus the infimum $b_1^{\mathcal{K}_1}$ is attained at a single point.

(ii) The statements are obvious from the definition (4.13) and (vi) of Lemma 7.1.

(iii) We trace the arc $\mathcal{E}_1$ from the end point $(b_1^0, 0)$. If $b_2^{\mathcal{K}(1)} \geq 0$, the coordinate a_1 increases monotonically and the minimum of a_1 in $\mathcal{E}_1$ is attained at the starting point $(b_1^0, 0) = (b_1^{\mathcal{K}_1}, b_2^{\mathcal{K}_1})$. If $b_2^{\mathcal{K}_1} \leq 0$, from the convexity of $\mathcal{K}$, the point $(b_1^{\mathcal{K}(1)}, b_2^{\mathcal{K}(1)})$ lies on the left arc of $\mathcal{K}_{loop}$ between points $(b_1^{h_1}, b_2^{h_1})$ and $(b_1^{h_1}, b_2^{h_1,c})$. If we trace $\mathcal{E}_1$ from $(b_1^0, 0)$, the coordinate a_1 decreases first and we eventually reach the minimum point $(b_1^{\mathcal{K}_1}, b_2^{\mathcal{K}_1})$ or the other end point $(b_1^{h_1}, b_2^{h_1})$. If we reach $(b_1^{\mathcal{K}_1}, b_2^{\mathcal{K}_1})$ first (this occurs when $b_2^{h_1,c} \geq b_2^{h_1}$), the coordinate a_1 begins to increase then. Hence the minimum is attained at $(b_1^{\mathcal{K}_1}, b_2^{\mathcal{K}_1})$. If we reach $(b_1^{h_1}, b_2^{h_1})$ first (this occurs when $b_2^{h_1,c} \leq b_2^{h_1}$), the minimum is attained at $(b_1^{h_1}, b_2^{h_1})$.

(iv) The statements are easily led from the definition of $\mathcal{F}_1$.

(v) If $(b_1^{\mathcal{E}_2}, b_2^{\mathcal{E}_2}) = (b_1^{\mathcal{K}_2}, b_2^{\mathcal{K}_2})$, then clearly $b_2^{\mathcal{E}_1} \geq b_2^{\mathcal{K}_2} = b_2^{\mathcal{E}_2}$. Hence $\mathcal{F}_1 = \mathcal{E}_1$ and $b_1^{\mathcal{F}_1} = b_1^{\mathcal{E}_1}$. On the other hand, if $(b_1^{\mathcal{E}_2}, b_2^{\mathcal{E}_2}) = (b_1^{h_2}, b_2^{h_2})$, we have to compare $b_2^{\mathcal{E}_1}$ with $b_2^{\mathcal{E}_2} = b_2^{h_2}$. From the definition, if $b_2^{\mathcal{E}_1} \geq b_2^{h_2}$, then $\mathcal{F}_1 = \mathcal{E}_1$ and $b_1^{\mathcal{F}_1} = b_1^{\mathcal{E}_1}$. If $b_2^{\mathcal{E}_1} \leq b_2^{h_2}$, then $\mathcal{F}_1$ is the arc between $(b_1^0, 0)$ and $(b_1^{h_2,c}, b_2^{h_2})$, and the infimum is attained at the latter end point.

The fact that $\mathcal{E}_1 \bigcap \mathcal{E}_2 = \phi$ is proved as follows. From (ii) above, $\mathcal{E}_k$ is in the region $\{(a_1, a_2) : a_k \leq a_{k'} \leq 0\}$. Further from the definition (4.13) and (vi) of Lemma 7.1, $\mathcal{E}_k$ contains a point (a_1, a_2) on the line $a_1 = a_2$ only when $a_k = h_k(a_k) = a_{k'}$, and this may happen only when $r_{k'k} = 1$. The stability condition requires that $r_{12}r_{21} < 1$. Hence either $\mathcal{E}_1$ or $\mathcal{E}_2$ does not contain any points on the line $a_1 = a_2$, and they are disjoint. ♢

Now we consider the case $r_{21} = 0$. The set $\mathcal{F}_2'$ was defined in (6.28).

Lemma 7.3 *Assume that* $r_{21} = 0$. *Then the following properties hold.*

(i) $b_1^{h_1} = b_1^0 < 0$ *and* $b_2^{h_1} = 0$.

(ii) $\mathcal{E}_1 = \mathcal{F}_1 = \{(b_1^0, 0)\}$, *and hence* $b_1^{\mathcal{E}_1} = b_1^0$ *and* $b_1^{\mathcal{F}_1} = b_1^0$.

(iii) $\mathcal{F}_2'$ *consists of points of* $\mathcal{F}_2$ *other than two end points.* $\mathcal{F}_2'$ *is nonempty.*

Proof (i) We can prove $b_1^0 < 0$ in a similar manner to the proof of (iv) of the previous lemma. The equalities $b_1^{h_1} = b_1^0$ and $b_2^{h_1} = 0$ are direct consequences of the fact $h_1(a_1) \equiv 0$.

(ii) The statement is trivial from the definitions (4.13) and (4.15) and from (i) above.

(iii) The statement is clear from the definitions (4.15) and (6.28). Both $b_1^{h_2}$ and $b_1^0 < 0$ are negative. Hence the two end points of $\mathcal{F}_2$ cannot be identical. ♢

Now we consider the case $r_{21} > 0$. To describe the set $\mathcal{G}_k(b_{k'})$ defined in (6.31) concretely, we introduce notations for coordinates of intersections of the straight line $a_{k'} = b_{k'}$ with $\mathcal{K}_{loop}$. For a variable $b_{k'}$ such that $b_{k'}^{h_k} \leq b_{k'} < 0$, let $\theta_k(b_{k'})$ and $\theta_k^c(b_{k'})$ be the k-th coordinates of the intersections such that $\theta_k(b_{k'}) \leq \theta_k^c(b_{k'})$.

Lemma 7.4 *Assume that* $r_{21} > 0$. *Then the following properties hold.*

(i) *If* $b_2^{h_1,c} > b_2^{h_1}$ *and* $\theta_1^c(b_2^{h_1}) = b_1^{h_1}$ *(this is the case where, when we trace the arc* $\mathcal{E}_1$ *from the end point* $(b_1^0, 0)$*, we first meet the point* $(b_1^{\mathcal{K}_1}, b_2^{\mathcal{K}_1})$*, next meet* $(b_1^{\mathcal{K}(2)}, b_2^{\mathcal{K}(2)})$ *and then finally reach the other end point* $(b_1^{h_1}, b_2^{h_1})$*), for* b_2 *such that* $b_2^{h_1} < b_2 < 0$, $\mathcal{G}_1(b_2)$ *is the arc of* $\mathcal{K}_{loop}$ *between points* $(b_1^0, 0)$ *and* $(\theta_1(b_2), b_2)$ *contained in* $\mathcal{E}_1$*, for* b_2 *such that* $b_2^{\mathcal{K}(2)} \leq b_2 \leq b_2^{h_1}$, $\mathcal{G}_1(b_2)$ *is the union of two arcs, one between* $(b_1^0, 0)$ *and* $(\theta_1(b_2), b_2)$ *and one between* $(\theta_1^c(b_2), b_2)$ *and* $(b_1^{h_1}, b_2^{h_1})$*, and for* $b_2 \leq b_2^{\mathcal{K}(2)}$, $\mathcal{G}_1(b_2)$ *is the arc between* $(b_1^0, 0)$ *and* $(b_1^{h_1}, b_2^{h_1})$ *(the same one as* $\mathcal{E}_1$ *except for end points). If* $b_2^{h_1,c} \leq b_2^{h_1}$ *or* $\theta_1(b_2^{h_1}) = b_1^{h_1}$*, for* b_2 *such that* $b_2^{h_1} < b_2 < 0$, $\mathcal{G}_1(b_2)$ *is the arc of* $\mathcal{K}_{loop}$ *between points* $(b_1^0, 0)$ *and* $(\theta_1(b_2), b_2)$ *contained in* $\mathcal{E}_1$*, and for* $b_2 \leq b_2^{h_1}$, $\mathcal{G}_1(b_2) = \mathcal{G}_1(b_2^{h_1})$*. The set in the latter case is the arc of* $\mathcal{K}_{loop}$ *between points* $(b_1^0, 0)$ *and* $(b_1^{h_1}, b_2^{h_1})$ *and same as* $\mathcal{E}_1$ *except for end points. In any case,* $\mathcal{G}_1(b_2)$ *is nonempty but contains no end points.* $\mathcal{G}_2(b_1)$ *is given in a similar manner.*

(ii) $\chi_k(b_{k'}) = \theta_k(b_{k'})$ *if* $b_{k'} \geq b_{k'}^{\mathcal{E}_k}$, *and* $\chi_k(b_{k'}) = b_k^{\mathcal{E}_k}$ *if* $b_{k'} \leq b_{k'}^{\mathcal{E}_k}$. *Hence the function* χ_k *is negative and continuous. It is strictly increasing in the interval* $[b_{k'}^{\mathcal{E}_k}, o)$ *and is constant equal to* $b_k^{\mathcal{E}_k}$ *in the interval* $(-\infty, b_{k'}^{\mathcal{E}_k}]$. *Further* $\chi_k(b_{k'}^{\mathcal{E}_k}) = b_k^{\mathcal{E}_k}$ *and* $\chi_k(b_{k'}^{\mathcal{E}_{k'}}) = b_k^{\mathcal{F}_k}$.

Proof (i) The form of $\mathcal{G}_k$ given in the statement is clear from the definition (6.31).

(ii) This property is easily derived from (i) above. ♢

8. Proofs of Propositions

Proof of Proposition 4.1 For brevity of discussion, we prove the case $r_{21} > 0$. The case $r_{21} = 0$ can be proved in a similar manner. When the system is of Jackson type, i.e. arrivals are Poissonian and services are exponential, the functions take the following form:

$$\begin{aligned} &\phi_k(a_k) = \lambda_k\left(e^{-a_k} - 1\right), \qquad \psi(a_k) = \mu_k\left(e^{-a_k} - 1\right), \qquad \text{and} \\ &\psi(-a_k + h_{k'}(a_{k'})) = \mu_k\left(e^{a_k - h_{k'}(a_{k'})} - 1\right) = \mu_k e^{a_k}\left(r_{kk'}e^{-a_{k'}} + r_{k0}\right) - \mu_k\,. \end{aligned} \tag{8.1}$$

A straightforward calculation shows that

$$\kappa(a_1, a_2) \;=\; \mu_1\left(1 - \rho_1 e^{-a_1}\right)\left(e^{a_1 - h_2(a_2)} - 1\right) + \mu_2\left(1 - \rho_2 e^{-a_2}\right)\left(e^{a_2 - h_1(a_1)} - 1\right). \tag{8.2}$$

For the intersection $(b_1^{h_1}, b_2^{h_1})$, we substitute $h_1(a_1)$ for a_2 in (8.2). Then

$$\kappa(a_1, h_1(a_1)) = \mu_1\left(1 - \rho_1 e^{-a_1}\right)\left(e^{a_1 - h_2(h_1(a_1))} - 1\right) = 0. \tag{8.3}$$

As shown in (vi) of Lemma 7.1, $a_1 < h_2(h_1(a_1))$ when $a_1 < 0$. So the equation has a unique negative solution $a_1 = \log\rho_1$. Hence $b_1^{h_1} = \log\rho_1$ and $b_2^{h_1} = h_1(b_1^{h_1}) = -\log(r_{21}\rho_1^{-1} + r_{20})$. Similarly $b_2^{h_2} = \log\rho_2$ and $b_1^{h_2} = h_2(b_2^{h_2}) = -\log(r_{12}\rho_2^{-1} + r_{10})$. It is also seen from (8.2) that the point $(a_1, a_2) = (\log\rho_1, \log\rho_2)$ is on $\mathcal{K}_{loop}$. It follows that $b_2^{h_1,c} = \log\rho_2$ and $b_1^{h_2,c} = \log\rho_1$, and hence $b_2^{h_1,c} = b_2^{h_2}$ and $b_1^{h_2,c} = b_1^{h_1}$. We shall check four possible cases individually.

(i) When $b_2^{h_1,c} \leq b_2^{h_1}$ and $b_1^{h_2,c} \leq b_1^{h_2}$, we have $b_2^{h_1} \geq b_2^{h_1,c} = b_2^{h_2}$ and $b_1^{h_2} \geq b_1^{h_2,c} = b_1^{h_1}$. This case corresponds to the first line of the right hand side of (4.29), and $(b_1^{\mathcal{F}_1}, b_2^{\mathcal{F}_2}) = (b_1^{h_1}, b_2^{h_2}) = (\log\rho_1, \log\rho_2)$.

(ii) When $b_2^{h_1,c} \leq b_2^{h_1}$ and $b_1^{h_2,c} \geq b_1^{h_2}$, we have to compare $b_1^{\mathcal{K}_2}$ with $b_1^{h_1}$. From (i) of Lemma 7.1 and the condition $b_1^{h_2,c} \geq b_1^{h_2}$, we have $b_1^{h_2} \leq b_1^{\mathcal{K}_2} \leq b_1^{h_2,c} = b_1^{h_1}$. Hence this case corresponds to the third line of the right hand side of (4.29), and $(b_1^{\mathcal{F}_1}, b_2^{\mathcal{F}_2}) = (b_1^{h_1}, b_2^{h_1,c}) = (\log\rho_1, \log\rho_2)$.

(iii) When $b_2^{h_1,c} \geq b_2^{h_1}$ and $b_1^{h_2,c} \leq b_1^{h_2}$, we can use a similar argument to (ii) above by interchanging the roles of node 1 and node 2. This case corresponds to the fifth line of the right hand side of (4.29), and $(b_1^{\mathcal{F}_1}, b_2^{\mathcal{F}_2}) = (\log\rho_1, \log\rho_2)$ again.

(iv) The case $b_2^{h_1,c} \geq b_2^{h_1}$ and $b_1^{h_2,c} \geq b_1^{h_2}$ cannot occur since $b_2^{h_1,c} = b_1^{h_1}$ and $b_1^{h_2,c} = b_2^{h_2}$ and, as noted just above Corollary 4.1, the case $b_2^{h_1} \leq b_2^{h_2}$ and $b_1^{h_2} \leq b_1^{h_1}$ cannot occur.

Thus, in any of the four cases, we see that $(b_1^{\mathcal{F}_1}, b_2^{\mathcal{F}_2}) = (\log\rho_1, \log\rho_2)$. ♢

Proof of Proposition 4.2 When $r_{21} = 0$, node 1 reduces to an ordinary MAP/PH/1 queue satisfying the local stability condition $\rho_1 = \lambda_1/\mu_1 < 1$. In this case $h_1(a_1) \equiv 0$, and as shown in Lemma 7.3, $b_1^{h_1} = b_1^0 < 0$ and $\mathcal{F}_1$ consists of a single point $(b_1^0, 0)$. Hence $\overline{\eta}_1 = \exp\{b_1^0\}$. From (4.23), b_1^0 is a negative solution of the equation $\kappa(a_1, 0) = \phi_1(a_1) + \psi_1(-a_1) = 0$. This equation is equivalent to (1.2) if $f(x)$ and $g(y)$ are replaced with $f_1(x)$ and $g_1(y)$. By applying Proposition 9 of Glynn and Whitt [6] to the above equation, we see that $\lim_{n\to\infty} \frac{1}{n}\log p_1(n) = \exp\{b_1^0\}$, and $\eta_1^* = \exp\{b_1^0\} = \overline{\eta}_1$. ♢

Proof of Proposition 4.3 When the system is a two-stage tandem queueing system, $\lambda_2 = 0$, $r_{12} = 1$ and $r_{21} = 0$. Therefore for a tandem queueing system MAP/M/1 → /M/1, the functions are reduced to $h_1(a_1) = 0$, $h_2(a_2) = a_2$, $\phi_2(a_2) = 0$, $\psi_1(a_1) = \mu_1 (e^{-a_1} - 1)$, $\psi_2(a_2) = \mu_2 (e^{-a_2} - 1)$ and $\kappa(a_1, a_2) = \phi_1(a_1) + \mu_1 (e^{a_1 - a_2} - 1) + \mu_2 (e^{a_2} - 1)$. From Proposition 4.2, our upper bound $\overline{\eta}_1$ for node 1 coincides with the exact decay rate η_1^*. As shown in Lemma 7.3, the set $\mathcal{F}_1$ consists of a single point $(b_1^0, 0)$ and hence $b_1^{\mathcal{F}_1} = b_1^0$. From (4.29), candidates for $b_2^{\mathcal{F}_k}$ are the three numbers, $b_2^{h_2}$, $b_2^{\mathcal{K}_2}$ and $b_2^{h_1,c}$. First, $b_2^{h_2}$ is a solution of the equation $\phi_1(a_2) + \mu_2 (e^{a_2} - 1) = 0$. Next, $b_2^{\mathcal{K}_2}$ is the second coordinate of a solution (a_1, a_2) of the pair of equations $\kappa(a_1, a_2) = \phi_1(a_1) + \mu_1 (e^{a_1 - a_2} - 1) + \mu_2 (e^{a_2} - 1) = 0$ and $\frac{\partial}{\partial a_1} \kappa(a_1, a_2) = \phi'(a_1) + \mu_1 e^{a_1} = 0$. Finally, $b_2^{h_1,c}$ is given by $\log(\mu_1 \overline{\eta}_1 / \mu_2)$.

Ganesh and Anantharam [4] gave the exact decay rates as a solution of a set of equations for a tandem queueing system GI/M/1 → /M/1. The set of equations there is the same as the one given here. Hence, our upper bounds coincide with the exact decay rates for MAP/M/1→/M/1 tandem queue.

Acknowledgments: The authors express their sincere thanks to the editor and anonymous referees for their helpful comments, which contribute largely to improve the manuscript.

A. List of symbols

$(\boldsymbol{T}_k, \boldsymbol{U}_k)$, $(\boldsymbol{b}_k, \boldsymbol{S}_k)$: MAP_k and PH_k representation, and we write $\boldsymbol{\sigma}_k^\top = -\boldsymbol{S}_k \boldsymbol{e}^\top$
r_{kj} : routing probability ($r_{k0} + r_{k1} + r_{k2} = 1$, $r_{12} > 0$)
$N_k(t)$, $I_k(t)$, $J_k(t)$: number of customers, phase of MAP, and of PH in node k at time t
$\mathcal{N}$, $\mathcal{I}_k$, $\mathcal{J}_k$: set of positive integers, MAP_k's phase and PH_k's phase
$\boldsymbol{X}(t)$: Markov chain defined in (2.1)
$\mathcal{S}$: state space of chain $\{\boldsymbol{X}(t)\}$
λ_k, μ_k : rate of external arrivals to node k, and service rate of node k
ρ_k : traffic intensity of node k
$p(n_1, n_2)_{i_1,i_2,j_1,j_2}$: state probability of $\{\boldsymbol{X}(t)\}$ in the steady state
$p(n_1, n_2)$: joint queue length probability
$p_k(n_k)$: marginal queue length probability of node k
$\mathcal{C}(n_1, n_2)$: Cell(n_1, n_2), set of states defined in (3.4)
$\boldsymbol{p}(n_1, n_2)$: vector of state probabilities corresponding to states in $\mathcal{C}(n_1, n_2)$
η_k^* : decay rate of the marginal queue length distribution of node k defined in (3.3)
$\overline{\eta}_k$: upper bound for η_k^* defined in (4.16)
f_k : asymptotic LST of the external interarrival times defined in (3.8)
g_k : Laplace-Stieltjes transform of PH_k
ϕ_k, ψ_k : inverse function of $\log f_k$ and $\log g_k$
$\boldsymbol{\nu}$, $\overline{\boldsymbol{\nu}}_k$, $\boldsymbol{\nu}_k$: vectors defined in (3.12), (3.13) and (3.14) respectively
κ, h_k : function defined in (4.10) and (4.8) respectively
$\mathcal{K}_{loop}$, $\mathcal{E}_k$, $\mathcal{F}_k$, $\mathcal{K}$, $\mathcal{K}_k$, $b_k^{\mathcal{E}_k}$, $b_k^{\mathcal{F}_k}$: sets and numbers defined in (4.12)∼(4.16)
b_k^0, $b_{k'}^{h_k,c}$, $(b_1^{h_k}, b_2^{h_k})$, $(b_1^{\mathcal{K}(k)}, b_2^{\mathcal{K}(k)})$, $(b_1^{\mathcal{K}_k}, b_2^{\mathcal{K}_k})$: numbers and points used for a concrete representation of $\overline{\eta}_k$ in Corollary 4.2

References

[1] D. Berstimas, I. C. Paschalidis and J. N. Tsitsiklis: On the large deviations behavior of acyclic networks of G/G/1 queues. *The Annals of Applied Probability*, **8** (1998), 1027-1069.

[2] C. S. Chang: Sample path large deviations and intree networks. *Queueing Systems*, **20** (1995), 7-36.

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

[4] A. Ganesh and V. Anantharam: Stationary tail probabilities in exponential server tandems with renewal arrivals. *Queueing Systems*, **22** (1996), 203-249.

[5] P. W. Glynn and W. Whitt: Large deviations behavior of counting processes and their inverses. *Queueing Systems*, **17** (1994), 107-128.

[6] P. W. Glynn and W. Whitt: Logarithmic asymptotics for steady-state tail probabilities in a single-server queue. *Journal of Applied Probability*, **31A** (1994), 131-156.

[7] D. G. Kendall: Some problems in the theory of queues. *Journal of the Royal Statistical Society*, Ser. B, **13** (1951), 151-185.

[8] D. M. Lucantoni: New results on the single server queue with a batch Markovian arrival process. *Stochastic Models*, **7** (1991), 1-46.

[9] N. Makimoto, Y. Takahashi and K. Fujimoto: Upper bounds for the geometric decay rate of the stationary distribution in two-stage tandem queues. *Research Report on Mathematical and Computing Sciences B-326* (Tokyo Institute of Technology, 1997).

[10] D. R. Miller: Computation of steady-state probabilities for M/M/1 priority queues. *Operations Research*, **29** (1981), 945-958.

[11] M. Miyazawa: Conjectures on decay rates of tail probabilities in generalized Jackson and batch movement networks. *Journal of the Operations Research Society of Japan*, **46** (2003), 74-98.

[12] M. F. Neuts: *Matrix-Geometric Solutions in Stochastic Models : An Algorithmic Approach* (The Johns Hopkins University Press, Baltimore, 1981).

[13] M. F. Neuts and Y. Takahashi: Asymptotic behavior of the stationary distributions in the GI/PH/c queue with heterogeneous servers. *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete*, **57** (1981), 441-452.

[14] V. Ramaswami and P. G. Taylor: Some properties of the rate matrices in level dependent quasi-birth-and-death processes with a countable number of phases. *Stochastic Models*, **12** (1996), 143-165.

[15] K. Sigman: The stability of open queueing networks. *Stochastic Processes and their Applications*, **35** (1990), 11-25.

[16] Y. Takahashi: Asymptotic exponentiality of the tail of the waiting time distribution in a PH/PH/c queue. *Advances in Applied Probability*, **13** (1981), 619-630.

[17] Y. Takahashi, K. Fujimoto and N. Makimoto: Geometric decay of the steady-state probabilities in a quasi-birth-and-death process with a countable number of phases. *Stochastic Models*, **17** (2001), 1-24.

Ken'ichi Katou
Graduate School of Information Systems
The University of Electro-Communications
Chofugaoka 1-5-1, Chofu-shi,
Tokyo 182-8585, Japan
E-mail: kkatou@is.uec.ac.jp