Upper bound for the decay rate of the joint queue-length distribution in a two-node Markovian queueing system

Ken'ichi Katou[*], Naoki Makimoto[†], Yukio Takahashi[*]

Abstract

This paper studies the geometric decay property of the joint queue-length distribution $\{p(n_1, n_2)\}$ of a two-node Markovian queueing system in the steady state. For arbitrarily given positive integers c_1, c_2, d_1 and d_2, an upper bound $\overline{\eta}(c_1, c_2)$ of the decay rate is derived in the sense

$$\exp\left\{\limsup_{n \to \infty} n^{-1} \log p(c_1 n + d_1, c_2 n + d_2)\right\} \leq \overline{\eta}(c_1, c_2) < 1.$$

It is shown that the upper bound coincides with the exact decay rate in most systems for which the exact decay rate is known. Moreover, as a function of c_1 and c_2, $\overline{\eta}(c_1, c_2)$ takes one of eight types, and the types explain some curious properties reported in the paper [2].

Keywords: Markovian queueing network, geometric decay property, joint queue-length distribution

AMS subject classification: 60K25 Queueing theory

1 Introduction

This paper studies the geometric decay property of the tail of the joint queue-length distribution of a two-node open queueing network with MAP inputs, PH services and random routings. In a previous paper [5], the authors derived an upper bound for the decay rate of the *marginal* queue-length distribution for the same model. Using the result, here we derive an upper bound of the decay rate for the *joint* queue-length distribution.

We refer the nodes of the network as Node 1 and Node 2, and denote by $p(n_1, n_2)$ the stationary probability that there are n_1 customers in Node 1 and n_2 customers in Node 2. One might expect that $p(n_1, n_2)$ decays geometrically, namely for some constants η_1, η_2 and C

$$p(n_1, n_2) \approx C\,\eta_1^{n_1}\eta_2^{n_2} \quad \text{for large } n_1 \text{ and } n_2. \tag{1.1}$$

[*]Department of Mathematical and Computing Sciences, Tokyo Institute of Technology, Ookayama 2-12-1, Meguro-ku, Tokyo 152-8552, Japan, E-mail: {katou.k.ab@m.titech.ac.jp[*], yukio@is.titech.ac.jp[‡]}, Tel, FAX: +81-3-5734-3851

[†]Graduate School of Systems Management, The University of Tsukuba, Otsuka 3-29-1, Bunkyo-ku, Tokyo 112-0012, Japan, E-mail: makimoto@gssm.otsuka.tsukuba.ac.jp

In case of a Jackson network, (1.1) holds with equality for any n_1 and n_2 if we take η_k as the traffic intensity ρ_k at Node k $(k = 1, 2)$. However, numerical results for tandem queueing systems PH/PH/1→/PH/1 in Fujimoto et al. [2] showed that the situation is not so simple. In the paper [2] the asymptotic behavior of $p(n_1, n_2)$ was examined numerically when $n_1 = c_1 n + d_1$ and $n_2 = c_2 n + d_2$ go to infinity as $n \to \infty$ for arbitrarily given positive integers c_1, c_2, d_1 and d_2. Based on the results of an extensive experiment, a conjecture was proposed. There exists a threshold $\tilde{\rho}_2$, and the behavior of $p(n_1, n_2)$ is different between the cases that the traffic intensity of the second stage $\rho_2 < \tilde{\rho}_2$ and $\rho_2 > \tilde{\rho}_2$.

The aim of this paper is to study this curious property theoretically rather than numerically. And we will do it in a more general setting. Our model here is two-node Markovian queueing systems for which the geometric decay property of the tail of the marginal queue-length distribution was studied in [5]. Instead of (1.1), we consider here a weaker decay property

$$\eta^*(c_1, c_2, d_1, d_2) = \exp\left\{\limsup_{n \to \infty} \frac{1}{n} \log p(c_1 n + d_1, c_2 n + d_2)\right\} < 1, \tag{1.2}$$

for arbitrarily chosen positive integers c_1, c_2, d_1 and d_2. We will refer $\eta^*(c_1, c_2, d_1, d_2)$ above as the *decay rate* of $p(n_1, n_2)$ along line $l(c_1, c_2, d_1, d_2)$, where $l(c_1, c_2, d_1, d_2)$ is the set $\{(n_1, n_2) : n_1 = c_1 n + d_1, \ n_2 = c_2 n + d_2, \ n = 1, 2, \dots\}$ on the (n_1, n_2)–plane.

In one of our main theorems, Theorem 5.1, we will prove that the inequality (1.2) holds for any choice of integers c_1, c_2, d_1 and d_2 by showing the existence of a function $\overline{\eta}(c_1, c_2)$ such that

$$\eta^*(c_1, c_2, d_1, d_2) \leq \overline{\eta}(c_1, c_2) < 1. \tag{1.3}$$

The upper bound $\overline{\eta}(c_1, c_2)$ as a function of c_1 and c_2 takes one of eight forms depending on the position of the point (η_1^*, η_2^*) in a 2-dimension plane, where η_k^* is the decay rate of the marginal queue-length distribution of Node k $(k = 1, 2)$ defined by (3.3).

In the theorem, $\overline{\eta}(c_1, c_2)$ is derived by using η_1^* and η_2^*, but we usually don't know the values of them except for some special systems. Fortunately we know their upper bounds $\overline{\eta}_1^*$ and $\overline{\eta}_2^*$, which were derived in the previous paper [5]. Using these $\overline{\eta}_1^*$ and $\overline{\eta}_2^*$ instead of η_1^* and η_2^*, we can derive another upper bound $\overline{\eta}(c_1, c_2)$ which is calculable for any two-node Markovian queueing system. We will do this in another main theorem, Theorem 6.2. It is shown that the upper bound $\overline{\eta}(c_1, c_2)$ coincides with the exact decay rate $\eta^*(c_1, c_2, d_1, d_2)$ in some systems for which the exact decay rate is known. Jackson type queueing networks are among them. The theorem shows that the upper bound $\overline{\eta}(c_1, c_2)$ takes one of eight types as a function of c_1 and c_2. In one of them, $\overline{\eta}(c_1, c_2)$ is given by $(\eta_1^{h_1})^{c_1}(\hat{\eta}_2^{h_2})^{c_2}$ (see (6.13)). In another type, $\overline{\eta}(c_1, c_2)$ is given by $(\eta_1^{h_2})^{c_1}(\hat{\eta}_2^{h_2})^{c_2}$ if $-c_1/c_2 \leq (\hat{b}_2^{h_1} - b_2^{h_2})/(b_1^{h_1} - \hat{b}_1^{h_2})$ and by $(\hat{\eta}_1^{h_1})^{c_1}(\eta_2^{h_1})^{c_2}$ if $-c_1/c_2 \geq (\hat{b}_2^{h_1} - b_2^{h_2})/(b_1^{h_1} - \hat{b}_1^{h_2})$ (see (6.11)). Each of the types seems to correspond to the two cases above conjectured in [2]. There are other types of $\overline{\eta}(c_1, c_2)$ where the function takes more complicated forms as in (6.12), (6.15), (6.16) and (6.17). The systems reported in [2] for which the convergence speed of the ratio $p(n_1, n_2)/(\eta_1)^{n_1}(\eta_2)^{n_2}$ is very slow are of these types. Hence the convergence speed might relate to the form of the function $\overline{\eta}(c_1, c_2)$.
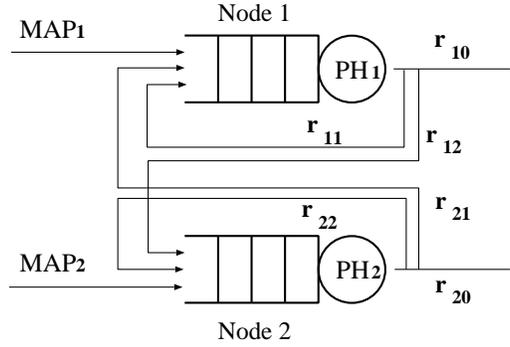
Figure 2.1: Two-node Markovian queueing system

The rest of the paper is organized as follows. In Section 2 we describe our model and introduce notations. We prepare some properties of the Markov chain which describes the stochastic behavior of the model in Section 3. In Section 4 we prove our fundamental lemma that gives an upper bound of the joint queue-length probability, and using the lemma we derive our main result in Section 5. The result on the marginal queue-length distributions in [5] is applied to the upper bound in Section 6. In Section 7 we discuss examples with some numerical results.

2 Model description and notations

In this section, we introduce our model and some notations. Our model here is the same as the one used in [5].

Model: The system is 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 $(\boldsymbol{T}_k, \boldsymbol{U}_k)$ [6]. 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 $(\boldsymbol{b}_k, \boldsymbol{S}_k)$ [9]. After being served, each customer proceeds to Node j ($j=1,2$) with probability r_{kj}, and leaves the system with probability $r_{k0} = 1 - r_{k1} - r_{k2}$. Without loss of generality, we assume that the exogeneous arrival rate λ_1 to Node 1 is positive and the routing probability r_{12} is positive. Exogenous arrival processes, service times and routings in both nodes are all stochastically independent. We will refer this model as a *two-node Markovian queueing system*. We use symbol "k" to refer to the node number of Node k. And for brevity of exposition, when symbol "k'" is used with a "k", it refers to the other node number, namely $k' = 2$ if $k = 1$, and $k' = 1$ if $k = 2$.

Vector and matrix notations: 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 $\boldsymbol{0}$ a row zero vector and by

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. They may be finite or infinite. Inequalities between vectors or matrices are considered elementwise.

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* of $\boldsymbol{A}$ 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.

Markov chain representations: 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 $(\boldsymbol{T}_k + \boldsymbol{U}_k)e^\top = \boldsymbol{0}^\top$. We denote the state space of the Markov chain by $\mathcal{I}_k$ and refer to the state of the chain as the *phase* of MAP_k. We assume that $\mathcal{I}_k$ is finite and $\boldsymbol{T}_k + \boldsymbol{U}_k$ is irreducible. The stationary probability vector of the chain is denoted by $\boldsymbol{a}_k$. The exogenous arrival rate to Node k is given by $\lambda_k = \left(-\boldsymbol{a}_k \boldsymbol{T}_k^{-1} e^\top\right)^{-1}$. When there exist no exogenous arrivals to Node 2, we consider both $\boldsymbol{T}_2$ and $\boldsymbol{U}_2$ is a scalar equal to 0, and $\lambda_2 = 0$.

The service time distribution PH_k also has an underlying finite absorbing Markov chain with transition rate matrix $\begin{pmatrix} \boldsymbol{S}_k & \boldsymbol{\sigma}_k^\top \\ \boldsymbol{0} & 0 \end{pmatrix}$ and an initial probability vector $(\boldsymbol{b}_k \; 0)$. Here $\boldsymbol{\sigma}_k^\top = -\boldsymbol{S}_k e^\top$. The state space of the 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. We refer to 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 $\boldsymbol{b}_k(-\boldsymbol{S}_k)^{-1} > \boldsymbol{0}$. The service rate at Node k is given by $\mu_k = \left(-\boldsymbol{b}_k \boldsymbol{S}_k^{-1} e^\top\right)^{-1}$. Of course, $\mu_k > 0$.

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

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

is a time-continuous Markov chain representing the stochastic behavior of the whole system. Its state is represented by a sextuple $(n_1, n_2, i_1, i_2, j_1, j_2)$, and the state space is given as

$$\mathcal{S} = \{\{0\} \times \{0\} \times \mathcal{I}_1 \times \mathcal{I}_2 \times \{0\} \times \{0\}\} \cup \{\{0\} \times \mathcal{N} \times \mathcal{I}_1 \times \mathcal{I}_2 \times \{0\} \times \mathcal{J}_2\}$$
$$\cup \{\mathcal{N} \times \{0\} \times \mathcal{I}_1 \times \mathcal{I}_2 \times \mathcal{J}_1 \times \{0\}\} \cup \{\mathcal{N} \times \mathcal{N} \times \mathcal{I}_1 \times \mathcal{I}_2 \times \mathcal{J}_1 \times \mathcal{J}_2\}, \tag{2.2}$$

where $\mathcal{N} = \{1, 2, 3, \cdots\}$. 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: Hereafter we assume the traffic intensity of Node k is strictly less than 1,

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

This assumption implies that the Markov chain $\{\boldsymbol{X}(t)\}$ is stable (see [5, 10]).

To make our discussion simpler, hereafter we assume there exists no direct feedbacks to the same node, namely $r_{kk} = 0$ for $k = 1, 2$. This does not restrict any generality as long as we are concerned only about the numbers of customers in Nodes 1 and 2. Because, when $r_{kk} > 0$, we may change the routing probabilities to $\tilde{r}_{k0} = r_{k0}/(1 - r_{kk})$, $\tilde{r}_{kk} = 0$ and $\tilde{r}_{kk'} = r_{kk'}/(1 - r_{kk})$, and use the service time distribution $(\tilde{\boldsymbol{b}}_k, \tilde{\boldsymbol{S}}_k) = (\boldsymbol{b}_k, \boldsymbol{S}_k + r_{kk}\boldsymbol{\sigma}_k^\top \boldsymbol{b}_k)$. The new model has the same $\{\boldsymbol{X}(t)\}$ process as the original one.

3   Balance equations and doubly geometric solution

For further discussion, here we prepare some notations related to stationary probabilities 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

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

Joint queue-length probabilities and marginal queue-length probabilities of Node k are written as

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

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

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

In Theorem 4.1 of [5], an upper bound $\overline{\eta}_k^*$ of η_k^* was derived and proved to be less than 1. This implies that the decay rate itself is strictly less than 1, i.e. $\eta_k^* < 1$.

Balance equations: For $n_1, n_2 \geq 1$, we let $\mathcal{C}(n_1, n_2)$ be the set of states at which there are n_1 customers in Node 1 and n_2 customers in Node 2, namely

$$
\mathcal{C}(n_1, n_2) \; = \; \{n_1\} \times \{n_2\} \times \mathcal{I}_1 \times \mathcal{I}_2 \times \mathcal{J}_1 \times \mathcal{J}_2.
\tag{3.4}
$$

We call $\mathcal{C}(n_1, n_2)$ as a *cell*. When $n_1 = 0$ and/or $n_2 = 0$, we define cell $\mathcal{C}(n_1, n_2)$ 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 $\mathcal{C}(n_1, n_2)$ can be 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 a vector form as

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

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 takes a slightly different form.

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

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

It is defined for y in the interval $\mathcal{D}[g_k] = (\delta_k^g, \infty)$, where $\delta_k^g(< 0)$ is its abscissa of convergence. The service rate is given by $\mu_k = -1/g_k'(0)$, where the prime $(\prime)$ 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

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

The 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 known that $f_k(x)$ is the Perron-Frobenius eigenvalue of the matrix $\boldsymbol{U}_k(x\boldsymbol{I} - \boldsymbol{T}_k)^{-1}$ for $x \in \mathcal{D}[f_k]$.

For a monotone function h, we denote its inverse function by $\mathrm{inv}[h]$. Let ϕ_k be the inverse function of $\log f_k$, and ψ_k be that of $\log g_k$, i.e.

$$
\phi_k(a) = \mathrm{inv}[\log f_k](a) \quad \text{and} \quad \psi_k(a) = \mathrm{inv}[\log g_k](a).
\tag{3.9}
$$

These functions are defined on the whole real line $(-\infty, +\infty)$. If $\lambda_2 = 0$, we consider $\phi_2(a) \equiv 0$. Functions ϕ_k and ψ_k can be interpreted probabilistically using LSTs of the number of exogenous arrivals and the number of (fictitious) customers served at Node k during time interval $(0, t]$. See [4, 5] for a detailed interpretation.

Doubly geometric form solution: Using functions introduced above, we construct a solution to the local balance equations (3.6) for $n_1, n_2 \geq 2$. Arbitrarily choose a pair of real numbers (a_1, a_2) and let

$$
\eta_k = e^{a_k}.
\tag{3.10}
$$

And let

$$
\kappa(a_1, a_2) = \phi_1(a_1) + \phi_2(a_2) + \psi_1(-a_1 + h_2(a_2)) + \psi_2(-a_2 + h_1(a_1)),
\tag{3.11}
$$

where

$$
h_k(a_k) = -\log\left(r_{k'k}\, e^{-a_k} + r_{k'0} \right).
\tag{3.12}
$$

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

We consider matrices $\boldsymbol{T}_k + \eta_k^{-1} \boldsymbol{U}_k$ and $\boldsymbol{S}_k + \left(r_{k0} \eta_k + r_{kk'} \eta_k \eta_{k'}^{-1} \right) \boldsymbol{\sigma}_k^\top \boldsymbol{b}_k$. From the irreducibility of the MAP_k and PH_k representations, these matrices are irreducible and have simple Perron-Frobenius eigenvalues $x_k = \phi_k(a_k)$ and $-y_k = \psi_k(-a_k + h_{k'}(a_{k'}))$, respectively. We denote by $\overline{\boldsymbol{\nu}}_k$ and $\boldsymbol{\nu}_k$ the unique (up to multiplicative constants) positive left eigenvectors associated with them. Now we let

$$\boldsymbol{\nu} = \overline{\boldsymbol{\nu}}_1 \otimes \overline{\boldsymbol{\nu}}_2 \otimes \boldsymbol{\nu}_1 \otimes \boldsymbol{\nu}_2, \tag{3.13}$$

and for arbitrarily given $n_1, n_2 \geq 2$, let

$$\boldsymbol{p}^\dagger(m_1, m_2) = \eta_1^{m_1} \eta_2^{m_2} \boldsymbol{\nu}, \qquad m_1 = n_1, n_1 \pm 1 \quad \text{and} \quad m_2 = n_2, n_2 \pm 1. \tag{3.14}$$

Substituting $\boldsymbol{p}(m_1, m_2) = \boldsymbol{p}^\dagger(m_1, m_2)$, a direct calculation shows the right hand side of (3.6) becomes $\kappa(a_1, a_2)\boldsymbol{\nu}$. If $\kappa(a_1, a_2) = 0$ then $\boldsymbol{p}^\dagger(m_1, m_2)$ given by (3.14) satisfies the local balance equations (3.6) around cell $\mathcal{C}(n_1, n_2)$. So the function $\kappa(a_1, a_2)$ is crucial in our discussion. Related to the function, we introduce a set of pairs (a_1, a_2) as

$$\mathcal{K} = \{(a_1, a_2) \ : \ \kappa(a_1, a_2) \leq 0\}. \tag{3.15}$$

This set is bounded and convex on the (a_1, a_2)-plane. Its periphery

$$\mathcal{K}_{loop} = \{(a_1, a_2) \ : \ \kappa(a_1, a_2) = 0\} \tag{3.16}$$

is a loop passing the origin. See Figure 5.1 for an example of $\mathcal{K}_{loop}$. It is round in the sense that any tangential line is tangent to $\mathcal{K}_{loop}$ at a single point. A segment of the loop passes the third quadrant of the (a_1, a_2)-plane, i.e. $\mathcal{K}_{loop} \cap \{a_1 < 0\} \cap \{a_2 < 0\} \neq \phi$. For other properties see Section 7 of [5].

4   Fundamental lemma

For the decay rate η_k^* of the marginal queue-length distribution of Node k we put

$$b_k^* = \log \eta_k^* \quad (< 0) \tag{4.1}$$

and introduce two subsets of $\mathcal{K}$:

$$\begin{aligned}
\mathcal{H}^+ &= \{(a_1, a_2) \ : \ \kappa(a_1, a_2) \leq 0, \ b_1^* < a_1 < 0, \ a_2 < 0\}, \quad \text{and} \\
\mathcal{H}^- &= \{(a_1, a_2) \ : \ \kappa(a_1, a_2) \leq 0, \ a_1 < 0, \ b_2^* < a_2 < 0\}.
\end{aligned} \tag{4.2}$$

These sets are nonempty. The following is a key lemma for our discussion. It is proved at the end of this section after preparing a series of lemmas. Note that η_k^+ and η_k^- defined in the lemma are strictly less than 1 from the definition of $\mathcal{H}^+$ and $\mathcal{H}^-$ in (4.2). Hence this lemma shows a geometric decay property of the joint queue-length distribution $\{p(n_1, n_2)\}$.

Lemma 4.1 (Fundamental lemma) *For arbitrarily given* $(a_1^+, a_2^+) \in \mathcal{H}^+$ *and* $(a_1^-, a_2^-) \in \mathcal{H}^-$, *we put* $\eta_k^+ = e^{a_k^+}$ *and* $\eta_k^- = e^{a_k^-}$. *Then there exist positive constants* C^+ *and* C^- *such that for any nonnegative integers* n_1 *and* n_2

$$p(n_1, n_2) < C^+ (\eta_1^+)^{n_1} (\eta_2^+)^{n_2} + C^- (\eta_1^-)^{n_1} (\eta_2^-)^{n_2}. \tag{4.3}$$

To prove (4.3), we exploit properties of rate matrix of a quasi-birth-and-death process having infinite number of states in each level [5, 7]. We consider a general 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* (QBD) *process having infinite number of states in each level* if, after partitioned into levels, its transition rate matrix is of a block tri-diagonal form

$$\boldsymbol{Q} = \begin{pmatrix} \overline{\boldsymbol{Q}}_1 & \overline{\boldsymbol{Q}}_0 & & \\ \overline{\boldsymbol{Q}}_2 & \boldsymbol{Q}_1 & \boldsymbol{Q}_0 & \\ & \boldsymbol{Q}_2 & \boldsymbol{Q}_1 & \ddots \\ & & \ddots & \ddots \end{pmatrix}. \tag{4.4}$$

Note 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 partition it into subvectors as $\boldsymbol{\pi} = (\boldsymbol{\pi}(0)\ \boldsymbol{\pi}(1)\ \boldsymbol{\pi}(2)\ \cdots)$ according to the levels. It is known [8, 9] that $\boldsymbol{\pi}$ takes a matrix geometric form as

$$\boldsymbol{\pi}(n) = \boldsymbol{\pi}(1)\boldsymbol{R}^{n-1}, \qquad n = 1, 2, 3, \cdots, \tag{4.5}$$

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

$$\boldsymbol{Q}_0 + \boldsymbol{R}\boldsymbol{Q}_1 + \boldsymbol{R}^2\boldsymbol{Q}_2 = \boldsymbol{O}. \tag{4.6}$$

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 the Perron-Frobenius eigenvalue of $\boldsymbol{R}$. However, in our case, the dimension of $\boldsymbol{R}$ is infinite and we cannot use the concept "eigenvalue". Lemma 5.1 of [5] provides an alternative tool for evaluating powers of $\boldsymbol{R}$ in such a case. Using the lemma together with (4.5), we can easily obtain an useful inequality for state probabilities as in Lemma 5.2 of [5].

Level partition for the two-node Markovian queueing system: We partition $\mathcal{S}$ in (2.2) into levels by the smaller number of customers in Node 1 and Node 2 as

$$\mathcal{S} = \bigcup_{n=0}^{\infty} \mathcal{L}(n), \quad \text{where} \quad \mathcal{L}(n) = \bigcup_{\min\{n_1, n_2\}=n} \mathcal{C}(n_1, n_2). \tag{4.7}$$

By this partition, the transition rate matrix can be written in the form (4.4). To describe submatrices $\boldsymbol{Q}_0$, $\boldsymbol{Q}_1$ and $\boldsymbol{Q}_2$ explicitly, we further partition $\mathcal{L}(n)$ into cells as

$$\mathcal{L}(n) = \bigcup_{m=-\infty}^{+\infty} \mathcal{L}(n, m), \quad \text{where} \quad \mathcal{L}(n, m) = \begin{cases} \mathcal{C}(n, n-m) & \text{if } m < 0, \\ \mathcal{C}(n+m, n) & \text{if } m \geq 0. \end{cases} \tag{4.8}$$

8

We refer to $\mathcal{L}(n)$ as the nth *level* and to $\mathcal{L}(n,m)$ as the mth *sublevel* of the nth level. By suitably arranging the order of states, $\boldsymbol{Q}_i$, $i = 0, 1, 2$, are written as

$$
\boldsymbol{Q}_0 = \begin{pmatrix}
\ddots & \ddots & \ddots & & & & & & \\
 & \boldsymbol{O} & \boldsymbol{A}_1 & \boldsymbol{M}_{21} & & & & & \\
 & & \boldsymbol{O} & \boldsymbol{A}_1 & \boldsymbol{M}_{21} & & & & \\
 & & & \boldsymbol{O} & \boldsymbol{A}_1 & & & & \\
 & & & & [\boldsymbol{O}] & & & & \\
 & & & & \boldsymbol{A}_2 & \boldsymbol{O} & & & \\
 & & & & \boldsymbol{M}_{12} & \boldsymbol{A}_2 & \boldsymbol{O} & & \\
 & & & & & \boldsymbol{M}_{12} & \boldsymbol{A}_2 & \boldsymbol{O} & \\
 & & & & & & \ddots & \ddots & \ddots
\end{pmatrix}, \tag{4.9}
$$

$$
\boldsymbol{Q}_1 = \begin{pmatrix}
\ddots & \ddots & \ddots & & & & & & \\
 & \boldsymbol{A}_2 & \boldsymbol{D} & \boldsymbol{M}_{20} & & & & & \\
 & & \boldsymbol{A}_2 & \boldsymbol{D} & \boldsymbol{M}_{20} & & & & \\
 & & & \boldsymbol{A}_2 & \boldsymbol{D} & \boldsymbol{M}_{20} & \boldsymbol{M}_{21} & & \\
 & & & & \boldsymbol{A}_2 & [\boldsymbol{D}] & \boldsymbol{A}_1 & & \\
 & & & & \boldsymbol{M}_{12} & \boldsymbol{M}_{10} & \boldsymbol{D} & \boldsymbol{A}_1 & \\
 & & & & & \boldsymbol{M}_{10} & \boldsymbol{D} & \boldsymbol{A}_1 & \\
 & & & & & & \boldsymbol{M}_{10} & \boldsymbol{D} & \boldsymbol{A}_1 \\
 & & & & & & \ddots & \ddots & \ddots
\end{pmatrix}, \tag{4.10}
$$

$$
\boldsymbol{Q}_2 = \begin{pmatrix}
\ddots & \ddots & \ddots & & & & & \\
 & \boldsymbol{M}_{12} & \boldsymbol{M}_{10} & \boldsymbol{O} & & & & \\
 & & \boldsymbol{M}_{12} & \boldsymbol{M}_{10} & \boldsymbol{O} & & & \\
 & & & \boldsymbol{M}_{12} & \boldsymbol{M}_{10} & [\boldsymbol{O}] & \boldsymbol{M}_{20} & \boldsymbol{M}_{21} \\
 & & & & & \boldsymbol{O} & \boldsymbol{M}_{20} & \boldsymbol{M}_{21} \\
 & & & & & \boldsymbol{O} & \boldsymbol{M}_{20} & \boldsymbol{M}_{21} \\
 & & & & & & \ddots & \ddots & \ddots
\end{pmatrix}, \tag{4.11}
$$

where

$$
\begin{aligned}
\boldsymbol{A}_1 &= \boldsymbol{U}_1 \otimes \boldsymbol{I} \otimes \boldsymbol{I} \otimes \boldsymbol{I}, & \boldsymbol{A}_2 &= \boldsymbol{I} \otimes \boldsymbol{U}_2 \otimes \boldsymbol{I} \otimes \boldsymbol{I}, \\
\boldsymbol{M}_{10} &= r_{10}\boldsymbol{I} \otimes \boldsymbol{I} \otimes \boldsymbol{\sigma}_1^\top \boldsymbol{b}_1 \otimes \boldsymbol{I}, & \boldsymbol{M}_{20} &= r_{20}\boldsymbol{I} \otimes \boldsymbol{I} \otimes \boldsymbol{I} \otimes \boldsymbol{\sigma}_2^\top \boldsymbol{b}_2, \\
\boldsymbol{M}_{12} &= r_{12}\boldsymbol{I} \otimes \boldsymbol{I} \otimes \boldsymbol{\sigma}_1^\top \boldsymbol{b}_1 \otimes \boldsymbol{I}, & \boldsymbol{M}_{21} &= r_{21}\boldsymbol{I} \otimes \boldsymbol{I} \otimes \boldsymbol{I} \otimes \boldsymbol{\sigma}_2^\top \boldsymbol{b}_2,
\end{aligned}
$$
$$
\text{and} \quad \boldsymbol{D} = \boldsymbol{T}_1 \oplus \boldsymbol{T}_2 \oplus \boldsymbol{S}_1 \oplus \boldsymbol{S}_2. \tag{4.12}
$$

Here the submatrices with brackets indicate the position corresponding to the 0th sublevel. The peripheral submatrices $\overline{\boldsymbol{Q}}_i$, $i = 0, 1, 2$, are more complicated. However they don't appear in our proof of the fundamental lemma and we omit their explicit description.

The stationary state probability vector $\boldsymbol{\pi}$ is also partitioned according to levels and sublevels.

9

Using $\boldsymbol{p}(n_1, n_2)$ in (3.5), the nth subvector $\boldsymbol{\pi}(n)$ of $\boldsymbol{\pi} = (\boldsymbol{\pi}(0), \boldsymbol{\pi}(1), \boldsymbol{\pi}(2), \cdots)$ is represented as

$$\boldsymbol{\pi}(n) = \left(\begin{array}{ccccc} \cdots & \boldsymbol{p}(n, n+2) & \boldsymbol{p}(n, n+1) & [\boldsymbol{p}(n,n)] & \boldsymbol{p}(n+1, n) & \boldsymbol{p}(n+2, n) & \cdots \end{array} \right). \tag{4.13}$$

Here the subvector with brackets indicates the position of the 0th sublevel.

For an arbitrarily given pair of real numbers (a_1, a_2), we construct a vector $\boldsymbol{q}$ using η_k in (3.10) and $\boldsymbol{\nu}$ in (3.13) as

$$\boldsymbol{q} = \left(\begin{array}{ccccc} \cdots & \eta_2^2 \boldsymbol{\nu} & \eta_2 \boldsymbol{\nu} & [\boldsymbol{\nu}] & \eta_1 \boldsymbol{\nu} & \eta_1^2 \boldsymbol{\nu} & \cdots \end{array} \right). \tag{4.14}$$

Lemma 4.2 *If the pair (a_1, a_2) satisfies the condition $\kappa(a_1, a_2) \leq 0$, then the vector $\boldsymbol{q}$ given by (4.14) satisfies $\boldsymbol{q}\boldsymbol{R} \leq \eta_1 \eta_2 \boldsymbol{q}$.*

Proof From Lemma 5.1 of [5], if there exists some constant ξ satisfying $\boldsymbol{q}(\xi^{-1}\boldsymbol{Q}_0 + \boldsymbol{Q}_1 + \xi \boldsymbol{Q}_2) \leq \boldsymbol{0}$ then $\boldsymbol{q}\boldsymbol{R} \leq \xi \boldsymbol{q}$. From the matrix representations of $\boldsymbol{Q}_i$'s (4.10)$\sim$(4.12) above, if we set $\xi = \eta_1 \eta_2$, the mth subvector of $\boldsymbol{q}(\xi^{-1}\boldsymbol{Q}_0 + \boldsymbol{Q}_1 + \xi \boldsymbol{Q}_2)$ for $m = 2, 3, \cdots$ is given by

$$\frac{1}{\eta_1 \eta_2} \left(\eta_1^{m+1} \boldsymbol{\nu} \boldsymbol{A}_2 + \eta_1^{m+2} \boldsymbol{\nu} \boldsymbol{M}_{12} \right) + \left(\eta_1^{m-1} \boldsymbol{\nu} \boldsymbol{A}_1 + \eta_1^m \boldsymbol{\nu} \boldsymbol{D} + \eta_1^{m+1} \boldsymbol{\nu} \boldsymbol{D}_1 \right)$$

$$+ \eta_1 \eta_2 \left(\eta_1^{m-1} \boldsymbol{\nu} \boldsymbol{D}_2 + \eta_1^{m-2} \boldsymbol{\nu} \boldsymbol{M}_{21} \right)$$

$$= \eta_1^m \boldsymbol{\nu} \left[\left(\boldsymbol{T}_1 + \frac{1}{\eta_1} \boldsymbol{U}_2 \right) \oplus \left(\boldsymbol{T}_2 + \frac{1}{\eta_2} \boldsymbol{U}_2 \right) \right.$$

$$\left. \oplus \left(\boldsymbol{S}_1 + \eta_1 \left(r_{10} + \frac{r_{12}}{\eta_2} \right) \boldsymbol{\sigma}_1^\top \boldsymbol{b}_1 \right) \oplus \left(\boldsymbol{S}_2 + \eta_2 \left(r_{20} + \frac{r_{21}}{\eta_1} \right) \boldsymbol{\sigma}_2^\top \boldsymbol{b}_2 \right) \right]. \tag{4.15}$$

A direct calculation shows the right hand side of the above equation becomes $\kappa(a_1, a_2) \eta_1^m \boldsymbol{\nu}$. Similarly, for $m = -2, -3, \cdots$ the mth subvector of $\boldsymbol{q}(\xi^{-1}\boldsymbol{Q}_0 + \boldsymbol{Q}_1 + \xi \boldsymbol{Q}_2)$ is given by $\kappa(a_1, a_2) \eta_2^{-m} \boldsymbol{\nu}$. When $m = -1, 0, 1$, the equation for the mth subvector takes a slightly different form from the one above, but we can easily check that it is also given by $\kappa(a_1, a_2) \eta_2 \boldsymbol{\nu}$, $\kappa(a_1, a_2) \boldsymbol{\nu}$ or $\kappa(a_1, a_2) \eta_1 \boldsymbol{\nu}$ according to the value of m. Hence $\boldsymbol{q}(\xi^{-1}\boldsymbol{Q}_0 + \boldsymbol{Q}_1 + \xi \boldsymbol{Q}_2) = \kappa(a_1, a_2) \boldsymbol{q}$. Then the condition $\kappa(a_1, a_2) \leq 0$ implies that $\boldsymbol{q}(\xi^{-1}\boldsymbol{Q}_0 + \boldsymbol{Q}_1 + \xi \boldsymbol{Q}_2) \leq \boldsymbol{0}$, and Lemma 5.1 of [5] assures that $\boldsymbol{q}$ satisfies $\boldsymbol{q}\boldsymbol{R} \leq \xi \boldsymbol{q}$. $\diamond$

Proof of fundamental lemma: The key idea of the proof of Lemma 4.1 is to split the state probability vector $\boldsymbol{\pi}(1)$ into two parts. Let

$$\boldsymbol{\pi}^-(1) = \left(\begin{array}{cccccc} \cdots & \boldsymbol{p}(1,3) & \boldsymbol{p}(1,2) & \boldsymbol{0}, & \boldsymbol{0} & \boldsymbol{0} & \cdots \end{array} \right), \quad \text{and}$$
$$\boldsymbol{\pi}^+(1) = \left(\begin{array}{cccccc} \cdots & \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{p}(1,1), & \boldsymbol{p}(2,1) & \boldsymbol{p}(3,1) & \cdots \end{array} \right). \tag{4.16}$$

Then, obviously $\boldsymbol{\pi}^-(1) + \boldsymbol{\pi}^+(1) = \boldsymbol{\pi}(1)$.

For the pair $(a_1^+, a_2^+) \in \mathcal{H}^+$, we construct a positive vector $\boldsymbol{q}^+$ as in (4.14) using corresponding numbers η_1^+, η_2^+ and vector $\boldsymbol{\nu}^+$. Since $\kappa(a_1^+, a_2^+) \leq 0$ from the condition of $\mathcal{H}^+$ in (4.2), Lemma 4.2 assures that $\boldsymbol{q}^+$ satisfies

$$\boldsymbol{q}^+ \boldsymbol{R} \leq \eta_1^+ \eta_2^+ \boldsymbol{q}^+. \tag{4.17}$$

Further, since $\eta_1^+ > \eta_1^*$ from the condition $a_1^+ > b_1^*$ of $\mathcal{H}^+$, there exists a positive constant C_π^+ such that

$$\boldsymbol{p}(m+1,1) < C_\pi^+ (\eta_1^+)^m \, \boldsymbol{\nu}^+ \qquad \text{for} \quad m = 0, 1, 2, \cdots. \tag{4.18}$$

The inequality $\boldsymbol{0} < C_\pi^+ (\eta_2^+)^m \, \boldsymbol{\nu}^+$ trivially holds for $m = 1, 2, 3, \cdots$. Combining these inequalities we have

$$\boldsymbol{\pi}^+(1) \le C_\pi^+ \boldsymbol{q}^+. \tag{4.19}$$

For the pair $(a_1^-, a_2^-) \in \mathcal{H}^-$, we can also construct a positive vector $\boldsymbol{q}^-$ as in (4.14) using corresponding numbers η_1^-, η_2^- and vector $\boldsymbol{\nu}^-$. It satisfies similar inequalities to (4.17) and (4.19) as $\boldsymbol{q}^- \boldsymbol{R} \le \eta_1^- \eta_2^- \boldsymbol{q}^-$ and $\boldsymbol{\pi}^-(1) \le C_\pi^- \boldsymbol{q}^-$ for some positive constant C_π^-.

Since $\boldsymbol{\pi}(1) = \boldsymbol{\pi}^+(1) + \boldsymbol{\pi}^-(1) \le C_\pi^+ \boldsymbol{q}^+ + C_\pi^- \boldsymbol{q}^-$, from (4.5) we have

$$\begin{aligned}
\boldsymbol{\pi}(n) &= \boldsymbol{\pi}(1) \boldsymbol{R}^{n-1} \le C_\pi^+ \boldsymbol{q}^+ \boldsymbol{R}^{n-1} + C_\pi^- \boldsymbol{q}^- \boldsymbol{R}^{n-1} \\
&\le C_\pi^+ (\eta_1^+ \eta_2^+)^{n-1} \boldsymbol{q}^+ + C_\pi^- (\eta_1^- \eta_2^-)^{n-1} \boldsymbol{q}^- \qquad \text{for} \quad n = 1, 2, 3, \cdots.
\end{aligned} \tag{4.20}$$

Rewriting this inequality in subvector-wise, we have

$$\boldsymbol{p}(n_1, n_2) \le C_\pi^+ (\eta_1^+)^{n_1-1} (\eta_2^+)^{n_2-1} \boldsymbol{\nu}^+ + C_\pi^- (\eta_1^-)^{n_1-1} (\eta_2^-)^{n_2-1} \boldsymbol{\nu}^- \quad \text{for} \quad n_1, n_2 = 1, 2, 3, \cdots. \tag{4.21}$$

It is easily checked that this inequality also holds for the cases with $n_1 = 0$ or $n_2 = 0$. Post-multiplying (4.21) with $\boldsymbol{e}^\top$ and choosing C^+ and C^- so that $C^+ > (\eta_1^+ \eta_2^+)^{-1} C_\pi^+ \boldsymbol{\nu}^+ \boldsymbol{e}^\top$ and $C^- > (\eta_1^- \eta_2^-)^{-1} C_\pi^- \boldsymbol{\nu}^- \boldsymbol{e}^\top$, we get the inequality (4.3). $\diamondsuit$

5 Upper bound for the decay rate

Using Lemma 4.1, our fundamental lemma, we shall derive an upper bound for the decay rate $\eta^*(c_1, c_2, d_1, d_2)$ of the joint queue-length distribution $\{p(n_1, n_2)\}$ when n_1 and n_2 get large along line $l(c_1, c_2, d_1, d_2)$. Remind that $\eta^*(c_1, c_2, d_1, d_2)$ is defined by (1.2) and $l(c_1, c_2, d_1, d_2)$ is defined just below the equation.

Upper bound for decay rate: From Lemma 4.1, for any $(a_1^+, a_2^+) \in \mathcal{H}^+$ and $(a_1^-, a_2^-) \in \mathcal{H}^-$, there exist constants C^+ and C^- for which (4.3) holds. Hence we have

$$\begin{aligned}
p(c_1 n + d_1, c_2 n + d_1) &< D^+ \exp\left\{ n \left(c_1 a_1^+ + c_2 a_2^+ \right) \right\} + D^- \exp\left\{ n \left(c_1 a_1^- + c_2 a_2^- \right) \right\} \\
&\le D \exp\left\{ n \max\left[c_1 a_1^+ + c_2 a_2^+, \, c_1 a_1^- + c_2 a_2^- \right] \right\},
\end{aligned} \tag{5.1}$$

where $D^+ = C^+ \exp\left\{ d_1 a_1^+ + d_2 a_2^+ \right\}$, $D^- = C^- \exp\left\{ d_1 a_1^- + d_2 a_2^- \right\}$ and $D = D^+ + D^-$. It follows that

$$\limsup_{n \to \infty} \frac{1}{n} \log p(c_1 n + d_1, c_2 n + d_1) \le \max\left[c_1 a_1^+ + c_2 a_2^+, \, c_1 a_1^- + c_2 a_2^- \right]. \tag{5.2}$$

11

We note that the right hand side of the inequality does not depend on d_1 and d_2. Taking the infimum of the right hand side of (5.2) over possible pairs (a_1^+, a_2^+) and (a_1^-, a_2^-), we get an upper bound of the limes superior. We let

$$
\begin{aligned}
w^+(c_1, c_2) &= \inf_{(a_1^+, a_2^+) \in \mathcal{H}^+} \left(c_1 a_1^+ + c_2 a_2^+ \right), \\
w^-(c_1, c_2) &= \inf_{(a_1^-, a_2^-) \in \mathcal{H}^-} \left(c_1 a_1^- + c_2 a_2^- \right), \quad \text{and} \\
w(c_1, c_2) &= \max \left[w^+(c_1, c_2), w^-(c_1, c_2) \right].
\end{aligned}
\tag{5.3}
$$

Then we have the following.

Theorem 5.1 *For arbitrarily positive integers c_1, c_2 and nonnegative integers d_1 and d_2, the decay rate $\eta^*(c_1, c_2, d_1, d_2)$ of $p(n_1, n_2)$ along $l(c_1, c_2, d_1, d_2)$ is bounded from above by $\exp\{w(c_1, c_2)\}$, namely*

$$
\eta^*(c_1, c_2, d_1, d_2) \leq \exp\{w(c_1, c_2)\}.
\tag{5.4}
$$

Notations: To derive an explicit expression for $w(c_1, c_2)$, we prepare some more notations. As stated before, the closed set $\mathcal{K}$ is convex and its periphery $\mathcal{K}_{loop}$ is a round loop. So any straight line on the (a_1, a_2)-plane, if it meets $\mathcal{K}_{loop}$, intersects with $\mathcal{K}_{loop}$ at two points or is tangent to $\mathcal{K}_{loop}$ at a single point (see Lemma 7.2 of [5] for details). Let $b_1^{\mathcal{K}(1)}$ be the minimum of a_1 on $\mathcal{K}_{loop}$, i.e.

$$
b_1^{\mathcal{K}(1)} = \min\{a_1 \,:\, \exists a_2 \text{ such that } \kappa(a_1, a_2) = 0\},
\tag{5.5}
$$

and denote the coordinates of the point attaining the minimum as $(b_1^{\mathcal{K}(1)}, b_2^{\mathcal{K}(1)})$. Similarly we denote the coordinates of the point attaining the minimum of a_2 on $\mathcal{K}_{loop}$ as $(b_1^{\mathcal{K}(2)}, b_2^{\mathcal{K}(2)})$. It is known that $b_k^{\mathcal{K}(k)} < 0$. On the contrary, $b_{k'}^{\mathcal{K}(k)}$ may be positive or negative, or equal to 0.

For a given a_1° such that $b_1^{\mathcal{K}(1)} \leq a_1^\circ \leq 0$, we let $\theta_2(a_1^\circ)$ be the second coordinate of the lower intersection of the straight line $a_1 = a_1^\circ$ with $\mathcal{K}_{loop}$, namely $\theta_2(a_1^\circ)$ is the smaller root of the equation $\kappa(a_1^\circ, a_2) = 0$ for a_2. Similarly, for a given a_2° such that $b_2^{\mathcal{K}(2)} \leq a_2^\circ \leq 0$, we let $\theta_1(a_2^\circ)$ be the smaller root of the equation $\kappa(a_1, a_2^\circ) = 0$ for a_1. Especially we write

$$
b_{k'}^0 = \theta_{k'}(0), \quad \text{and} \quad \hat{b}_{k'}^* = \theta_{k'}(b_k^*) \quad \text{if } b_k^* \geq b_k^{\mathcal{K}(k)}.
\tag{5.6}
$$

Note that b_k^* is the logarithm of η_k^*, the decay rate of the marginal queue-length distribution at Node k as defined in (4.1).

For a point (a_1, a_2) on $\mathcal{K}_{loop}$ such that $b_1^{\mathcal{K}(1)} < a_1 < b_1^{\mathcal{K}(2)}$ and $b_2^{\mathcal{K}(2)} < a_2 < b_2^{\mathcal{K}(1)}$, let $\sigma(a_1, a_2)$ be the gradient of the tangential line of the loop at the point, namely

$$
\sigma(a_1, a_2) = -\frac{\frac{\partial}{\partial a_1} \kappa(a_1, a_2)}{\frac{\partial}{\partial a_2} \kappa(a_1, a_2)}.
\tag{5.7}
$$

As a convention, we regard $\sigma(b_1^{\mathcal{K}(1)}, b_2^{\mathcal{K}(1)}) = -\infty$ and $\sigma(b_1^{\mathcal{K}(2)}, b_2^{\mathcal{K}(2)}) = 0$. For an arbitrarily given negative number u, let $(\tau_1(u), \tau_2(u))$ be the coordinates of the point (a_1, a_2) on $\mathcal{K}_{loop}$ such that
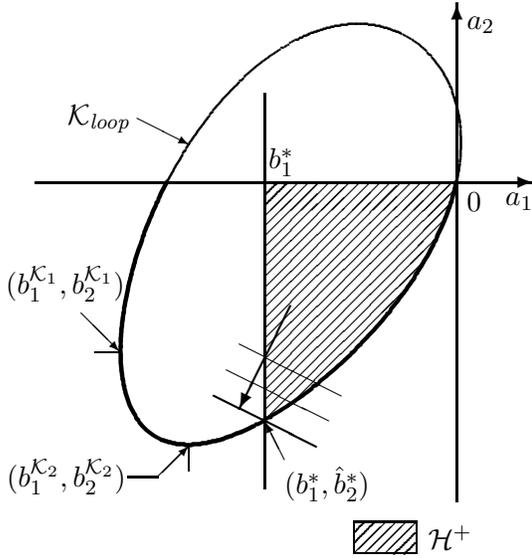
Figure 5.1: $\mathcal{H}^+$ and the line attaining $w^+(c_1, c_2)$ when $b_1^* \geq b_1^{\mathcal{K}_2}$
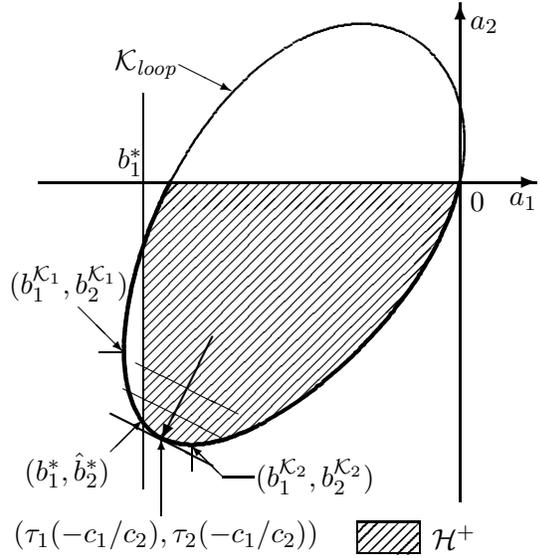
Figure 5.2: $\mathcal{H}^+$ and the line attaining $w^+(c_1, c_2)$ when $b_1^{\mathcal{K}_1} < b_1^* < b_1^{\mathcal{K}_2}$

$b_1^{\mathcal{K}(1)} < a_1 < b_1^{\mathcal{K}(2)}$, $b_2^{\mathcal{K}(2)} < a_2 < b_2^{\mathcal{K}(1)}$ and $\sigma(a_1, a_2) = u$. Note that a line which is tangent to $\mathcal{K}_{loop}$ at $(\tau_1(u), \tau_2(u))$ with gradient u is $-u a_1 + a_2 = -u\tau_1(u) + \tau_2(u)$ $(= const)$.

Let $b_1^{\mathcal{K}_1}$ be the minimum of a_1 on $\mathcal{K}_{loop}$ with $a_2 \leq 0$, i.e.

$$b_1^{\mathcal{K}_1} = \min\{a_1 \ : \ \exists a_2 \text{ such that } a_2 \leq 0 \text{ and } \kappa(a_1, a_2) = 0\}, \tag{5.8}$$

and we denote the coordinates of the point attaining the minimum as $(b_1^{\mathcal{K}_1}, b_2^{\mathcal{K}_1})$. The point is given by $(b_1^{\mathcal{K}(1)}, b_2^{\mathcal{K}(1)})$ if $b_2^{\mathcal{K}(1)} \leq 0$, and by $(b_1^0, 0)$ if $b_2^{\mathcal{K}(1)} > 0$. Similarly we denote the coordinates of the point attaining the minimum of a_2 on $\mathcal{K}_{loop}$ with $a_1 \leq 0$ as $(b_1^{\mathcal{K}_2}, b_2^{\mathcal{K}_2})$.

For positive integers c_1 and c_2, we put

$$v(c_1, c_2) = \min\{c_1 a_1 + c_2 a_2 \ : \ (a_1, a_2) \in \mathcal{K}, \ a_1 \leq 0, \ a_2 \leq 0\}. \tag{5.9}$$

Some considerations reveal that

$$v(c_1, c_2) = \begin{cases} c_1 \tau_1(-c_1/c_2) + c_2 \tau_2(-c_1/c_2), & \text{if } \tau_1(-c_1/c_2) < 0 \text{ and } \tau_2(-c_1/c_2) < 0, \\ c_1 b_1^0, & \text{if } \tau_2(-c_1/c_2) \geq 0, \\ c_2 b_2^0, & \text{if } \tau_1(-c_1/c_2) \geq 0. \end{cases} \tag{5.10}$$

Note that, the second case may occur only when $b_2^{\mathcal{K}(1)} > 0$ and the third case may occur only when $b_1^{\mathcal{K}(2)} > 0$. Hence, if $b_1^{\mathcal{K}(2)} < 0$ and $b_2^{\mathcal{K}(1)} < 0$, then the first case occurs.

Explicit expressions for $w^+(c_1, c_2)$ and $w^-(c_1, c_2)$: For a given pair of positive integers c_1 and c_2, we shall derive explicit expressions for $w^+(c_1, c_2)$ and $w^-(c_1, c_2)$. For $w^+(c_1, c_2)$, we check cases $b_1^* \geq b_1^{\mathcal{K}_2}$, $b_1^{\mathcal{K}_1} < b_1^* < b_1^{\mathcal{K}_2}$ and $b_1^* \leq b_1^{\mathcal{K}_1}$ separately.

13

When $b_1^* \geq b_1^{\mathcal{K}_2}$, the infimum of $c_1 a_1^+ + c_2 a_2^+$ over points $(a_1^+, a_2^+) \in \mathcal{H}^+$ is attained at $(b_1^*, \hat{b}_2^*)$ as illustrated in Figure 5.1. Hence $w^+(c_1, c_2) = c_1 b_1^* + c_2 \hat{b}_2^*$.

When $b_1^{\mathcal{K}_1} < b_1^* < b_1^{\mathcal{K}_2}$, the line $c_1 a_1 + c_2 a_2 = v(c_1, c_2)$ is tangent to $\mathcal{K}_{loop}$ at $(\tau_1(-c_1/c_2), \tau_2(-c_1/c_2))$. If $\tau_1(-c_1/c_2) \geq b_1^*$, this point is included in $\mathcal{H}^+$ as illustrated in Figure 5.2. Hence the infimum is attained at the point, and $w^+(c_1, c_2) = v(c_1, c_2)$. If $\tau_1(-c_1/c_2) < b_1^*$, the point is not included in $\mathcal{H}^+$, and the infimum is attained at the extreme point $(b_1^*, \hat{b}_2^*)$ with value $w^+(c_1, c_2) = c_1 b_1^* + c_2 \hat{b}_2^*$.

When $b_1^* \leq b_1^{\mathcal{K}_1}$, the set $\mathcal{H}^+$ consists of all the points of $\mathcal{K}$ in the third quadrant (except for the point $(b_1^*, \hat{b}_2^*)$ when $b_1^* = b_1^{\mathcal{K}_1}$). Hence the infimum $w^+(c_1, c_2)$ is given by $v(c_1, c_2)$.

Summarizing the above arguments, we have an explicit expression for $w^+(c_1, c_2)$ as in (5.11) below. Note that here we use the fact that $b_1^* \geq \tau_1(-c_1/c_2)$ when $b_1^* \geq b_1^{\mathcal{K}_2}$. For $w^-(c_1, c_2)$ we have a similar expression as in (5.12).

Lemma 5.2 *For a given pair of positive integers c_1 and c_2, $w^+(c_1, c_2)$ and $w^-(c_1, c_2)$ defined in (5.3) are expressed as follows.*

$$w^+(c_1, c_2) \;=\; \begin{cases} c_1 b_1^* + c_2 \hat{b}_2^*, & \text{if } b_1^* \geq \tau_1(-c_1/c_2) \ \text{ and } \ b_1^* \geq b_1^{\mathcal{K}_1}, \\ v(c_1, c_2), & \text{if } b_1^* \leq \tau_1(-c_1/c_2) \ \text{ or } \ b_1^* \leq b_1^{\mathcal{K}_1}, \end{cases} \qquad (5.11)$$

$$w^-(c_1, c_2) \;=\; \begin{cases} c_1 \hat{b}_1^* + c_2 b_2^*, & \text{if } b_2^* \geq \tau_2(-c_1/c_2) \ \text{ and } \ b_2^* \geq b_2^{\mathcal{K}_2}, \\ v(c_1, c_2), & \text{if } b_2^* \leq \tau_2(-c_1/c_2) \ \text{ or } \ b_2^* \leq b_2^{\mathcal{K}_2}. \end{cases} \qquad (5.12)$$

Explicit expression for $w(c_1, c_2)$: Using Lemma 5.2, we shall derive an explicit expression for $w(c_1, c_2) = \max\{w^+(c_1, c_2), \ w^-(c_1, c_2)\}$. We consider three cases separately: (i) $\tau_1(-c_1/c_2) < 0$ and $\tau_2(-c_1/c_2) < 0$, (ii) $\tau_1(-c_1/c_2) \geq 0$ (and hence $\tau_2(-c_1/c_2) < 0$), and (iii) $\tau_2(-c_1/c_2) \geq 0$ (and hence $\tau_1(-c_1/c_2) < 0$).

(i) case $\tau_1(-c_1/c_2) < 0$ and $\tau_2(-c_1/c_2) < 0$: Since the point $(\tau_1(-c_1/c_2), \tau_2(-c_1/c_2))$ is on $\mathcal{K}_{loop}$, we have $\tau_1(-c_1/c_2) \geq b_1^{\mathcal{K}_1}$ and $\tau_2(-c_1/c_2) \geq b_2^{\mathcal{K}_2}$. Hence (5.11) and (5.12) are rewritten as

$$w^+(c_1, c_2) \;=\; \begin{cases} c_1 b_1^* + c_2 \hat{b}_2^*, & \text{if } b_1^* \geq \tau_1(-c_1/c_2), \\ v(c_1, c_2), & \text{if } b_1^* \leq \tau_1(-c_1/c_2), \end{cases}$$

$$w^-(c_1, c_2) \;=\; \begin{cases} c_1 \hat{b}_1^* + c_2 b_2^*, & \text{if } b_2^* \geq \tau_2(-c_1/c_2), \\ v(c_1, c_2), & \text{if } b_2^* \leq \tau_2(-c_1/c_2). \end{cases}$$

Further we note that $c_1 b_1^* + c_2 \hat{b}_2^* \geq v(c_1, c_2)$ and $c_1 \hat{b}_1^* + c_2 b_2^* \geq v(c_1, c_2)$ since the points $(b_1^*, \hat{b}_2^*)$ and $(\hat{b}_1^*, b_2^*)$ are on $\mathcal{K}_{loop}$ and $v(c_1, c_2) = c_1 \tau_1(-c_1/c_2) + c_2 \tau_2(-c_1/c_2)$. So if we define regions in the third quadrant of the (a_1, a_2)-plane as

(when $\tau_1(-c_1/c_2) < 0$ and $\tau_2(-c_1/c_2) < 0$)

$$\mathcal{W}^0 = \{(a_1, a_2) \ : \ a_1 \leq \tau_1(-c_1/c_2) \text{ and } a_2 \leq \tau_2(-c_1/c_2)\},$$

$$\mathcal{W}^+ = \{(a_1, a_2) \ : \ a_1 \geq \tau_1(-c_1/c_2), \ c_1 a_1 + c_2 \theta_2(a_1) \geq c_1 \theta_1(a_2) + c_2 a_2 \text{ and } a_1, a_2 < 0\},$$

$$\mathcal{W}^- = \{(a_1, a_2) \ : \ a_2 \geq \tau_2(-c_1/c_2), \ c_1 a_1 + c_2 \theta_2(a_1) \leq c_1 \theta_1(a_2) + c_2 a_2 \text{ and } a_1, a_2 < 0\},$$
$$(5.13)$$
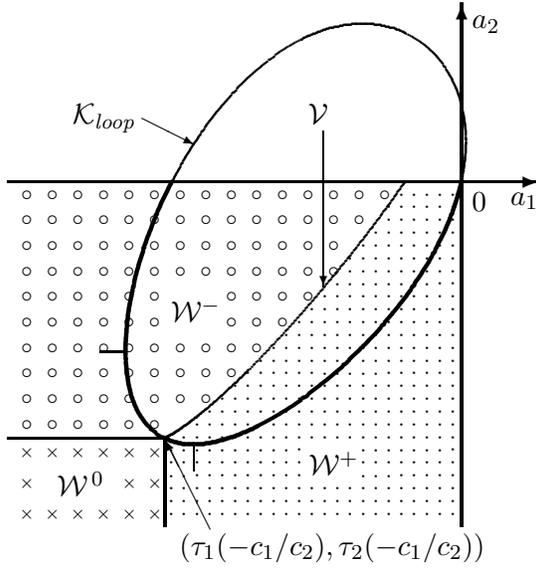
Figure 5.3: Curve $\mathcal{V}$ and regions $\mathcal{W}^+$, $\mathcal{W}^-$ and $\mathcal{W}^0$
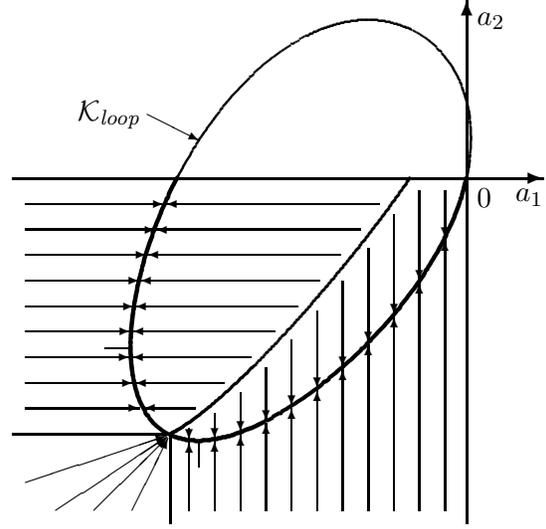

Figure 5.4: Relation between (b_1^*, b_2^*) and the point attaining $w(c_1, c_2)$

then we have Lemma 5.3 below.

(ii) case $\tau_1(-c_1/c_2) \geq 0$: In this case $b_1^* < 0 \leq \tau_1(-c_1/c_2)$, and hence $w^+(c_1, c_2) = v(c_1, c_2)$. On the other hand, $b_2^{\mathcal{K}_2} = b_2^0 \geq \tau_2(-c_1/c_2)$. Hence $w^-(c_1, c_2) = c_1\hat{b}_1^* + c_2 b_2^*$ if $b_2^* \geq b_2^0$, and $w^-(c_1, c_2) = v(c_1, c_2)$ if $b_2^* \leq b_2^0$. Since always $c_1\hat{b}_1^* + c_2 b_2^* \geq v(c_1, c_2)$, we have $w(c_1, c_2) = w^-(c_1, c_2)$. So the three regions for Lemma 5.3 below are given as follows.

$$(\text{when } \tau_1(-c_1/c_2) \geq 0)$$
$$\mathcal{W}^0 = \{(a_1, a_2) \ : \ a_1 < 0 \text{ and } a_2 \leq b_2^0\}, \tag{5.14}$$
$$\mathcal{W}^+ = \phi, \qquad \mathcal{W}^- = \{(a_1, a_2) \ : \ a_1 < 0 \text{ and } b_2^0 \leq a_2 < 0\}.$$

(iii) case $\tau_2(-c_1/c_2) \geq 0$: Similarly the three regions for Lemma 5.3 are given as follows.

$$(\text{when } \tau_2(-c_1/c_2) \geq 0)$$
$$\mathcal{W}^0 = \{(a_1, a_2) \ : \ a_1 \leq b_1^0 \text{ and } a_2 < 0\}, \tag{5.15}$$
$$\mathcal{W}^+ = \{(a_1, a_2) \ : \ b_1^0 \leq a_1 < 0 \text{ and } a_2 < 0\}, \qquad \mathcal{W}^- = \phi.$$

Lemma 5.3 *For a given pair of positive integers c_1 and c_2, $w(c_1, c_2)$ defined in (5.3) is expressed as follows.*

$$w(c_1, c_2) = \begin{cases} c_1 b_1^* + c_2 \hat{b}_2^* & \text{if } (b_1^*, b_2^*) \text{ in } \mathcal{W}^+, \\ c_1 \hat{b}_1^* + c_2 b_2^* & \text{if } (b_1^*, b_2^*) \text{ in } \mathcal{W}^-, \\ v(c_1, c_2) & \text{if } (b_1^*, b_2^*) \text{ in } \mathcal{W}^0, \end{cases} \tag{5.16}$$

where $\mathcal{W}^0$, $\mathcal{W}^+$ and $\mathcal{W}^-$ are regions defined in (5.13), (5.14) or (5.15).

Note that these expressions for $\mathcal{W}^+$, $\mathcal{W}^-$ and $\mathcal{W}^0$ depend only on the ratio $-c_1/c_2$, and any pair (c_1, c_2) having a common ratio leads to same regions. Figure 5.3 shows an example of the

15
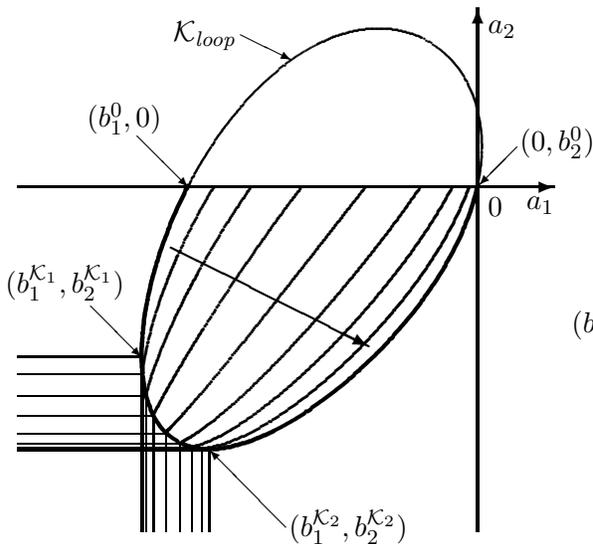
Figure 5.5: Movement of $\mathcal{W}$'s when $-c_1/c_2$ increases from $-\infty$ to 0
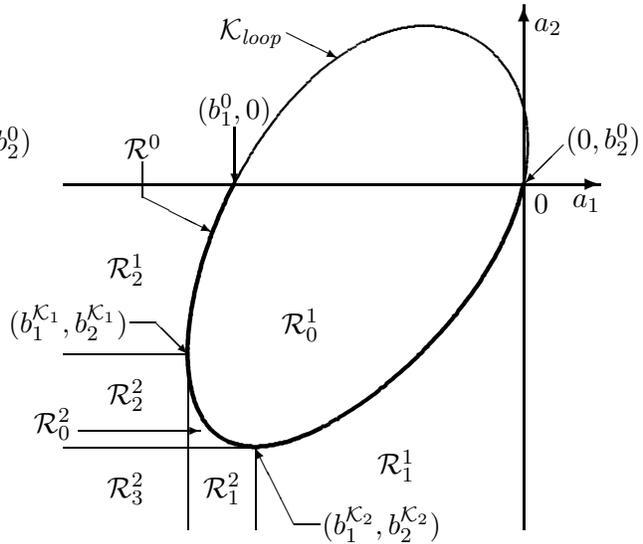
Figure 5.6: Partition by the form of the function $w(c_1, c_2)$

division of the third quadrant of the (a_1, a_2)-plane into sets $\mathcal{W}^+$, $\mathcal{W}^-$ and $\mathcal{W}^0$ in the case (i). The boundary between $\mathcal{W}^+$ and $\mathcal{W}^-$ is denoted by $\mathcal{V}$. Figure 5.4 shows how (b_1^*, b_2^*) is related to a point on $\mathcal{K}_{loop}$ which gives the value of $w(c_1, c_2)$.

Remark 5.1 From Theorem 5.1, an upper bound for the decay rate $\eta^*(c_1, c_2, d_1, d_2)$ of $p(n_1, n_2)$ along $l(c_1, c_2, d_1, d_2)$ is given by $e^{w(c_1,c_2)}$. Hence, the lemma above shows that the upper bound takes the form either

$$(\eta_1^*)^{c_1}(\hat{\eta}_2^*)^{c_2}, \ (\hat{\eta}_1^*)^{c_1}(\eta_2^*)^{c_2} \quad \text{or} \quad e^{v(c_1,c_2)}, \tag{5.17}$$

where $\hat{\eta}_k^* = \exp\{\hat{b}_k^*\}$. Let $\eta_k^0 = \exp\{b_k^0\}$. From (5.10), $e^{v(c_1,c_2)}$ is given by $\exp\{c_1\tau_1(-c_1/c_2) + c_2\tau_2(-c_1/c_2)\}$ when $\tau_1(-c_1/c_2) < 0$ and $\tau_2(-c_1/c_2) < 0$, by $(\eta_1^0)^{c_1}$ when $\tau_2(-c_1/c_2) \geq 0$, and by $(\eta_2^0)^{c_2}$ when if $\tau_1(-c_1/c_2) \geq 0$. Remind that η_k^* is the exact decay rate of the marginal queue-length distribution of Node k. In some systems (such as Jackson type networks), $\hat{\eta}_k^*$ coincides with η_k^*. However, the bound in general may differ from $(\eta_1^*)^{c_1}(\eta_2^*)^{c_2}$. It might be smaller or even larger than $(\eta_1^*)^{c_1}(\eta_2^*)^{c_2}$. $\diamondsuit$

$w(c_1, c_2)$ as a function of c_1 and c_2: So far we have treated c_1 and c_2 being fixed. Now we vary c_1 and c_2 and examine the behavior of $w(c_1, c_2)$ as a function of them.

We note that, as $-c_1/c_2 \to -\infty$, the curve $\mathcal{V}$ in Figure 5.3 tends to the upper left arc of $\mathcal{K}_{loop}$ from $(b_1^0, 0)$ to $(b_1^{\mathcal{K}_1}, b_2^{\mathcal{K}_1})$, and as $-c_1/c_2 \to 0$, $\mathcal{V}$ tends to the lower right arc of $\mathcal{K}_{loop}$ from $(0, b_2^0)$ to $(b_1^{\mathcal{K}_2}, b_2^{\mathcal{K}_2})$. Two half-lines, which are boundaries between $\mathcal{W}^+$ and $\mathcal{W}^0$ and between $\mathcal{W}^-$ and $\mathcal{W}^0$, also move as $\mathcal{V}$ moves. Hence, when $-c_1/c_2$ increases from $-\infty$ to 0, the boundaries of the regions move as illustrated in Figure 5.5. So, if we partition the third quadrant of the (a_1, a_2)-plane into

eight regions as illustrated in Figure 5.6, the function $w(c_1, c_2)$ takes different forms according to the region in which (b_1^*, b_2^*) falls. Formally these regions are defined as follows.

$$
\begin{aligned}
\mathcal{R}^0 &= \{(a_1, a_2) \ : \ \kappa(a_1, a_2) = 0, \ a_1 < 0 \text{ and } a_2 < 0\}, \\
\mathcal{R}_0^1 &= \{(a_1, a_2) \ : \ \kappa(a_1, a_2) < 0, \ a_1 < 0 \text{ and } a_2 < 0\}, \\
\mathcal{R}_1^1 &= \{(a_1, a_2) \ : \ \kappa(a_1, a_2) > 0, \ b_1^{\mathcal{K}_2} \le a_1 < 0 \text{ and } b_2^{\mathcal{K}_2} a_1 - b_1^{\mathcal{K}_2} a_2 \ge 0\}, \\
\mathcal{R}_2^1 &= \{(a_1, a_2) \ : \ \kappa(a_1, a_2) > 0, \ b_2^{\mathcal{K}_1} \le a_2 < 0 \text{ and } b_2^{\mathcal{K}_1} a_1 - b_1^{\mathcal{K}_1} a_2 \le 0\}, \\
\mathcal{R}_0^2 &= \{(a_1, a_2) \ : \ \kappa(a_1, a_2) > 0, \ b_1^{\mathcal{K}_1} < a_1 < b_1^{\mathcal{K}_2} \text{ and } b_2^{\mathcal{K}_2} < a_2 < b_2^{\mathcal{K}_1}\}, \\
\mathcal{R}_1^2 &= \{(a_1, a_2) \ : \ \kappa(a_1, a_2) > 0, \ b_1^{\mathcal{K}_1} < a_1 < b_1^{\mathcal{K}_2} \text{ and } a_2 \le b_2^{\mathcal{K}_2}\}, \\
\mathcal{R}_2^2 &= \{(a_1, a_2) \ : \ \kappa(a_1, a_2) > 0, \ a_1 \le b_1^{\mathcal{K}_1} \text{ and } b_2^{\mathcal{K}_2} < a_2 < b_2^{\mathcal{K}_1}\}, \\
\mathcal{R}_3^2 &= \{(a_1, a_2) \ : \ \kappa(a_1, a_2) > 0, \ a_1 \le b_1^{\mathcal{K}_1} \text{ and } a_2 \le b_2^{\mathcal{K}_2}\}.
\end{aligned}
\tag{5.18}
$$

Note that $\kappa(a_1, a_2) = 0$ implies that the point (a_1, a_2) is on $\mathcal{K}_{loop}$, $\kappa(a_1, a_2) < 0$ implies that (a_1, a_2) is inside of $\mathcal{K}_{loop}$, and $\kappa(a_1, a_2) > 0$ implies that (a_1, a_2) is outside of $\mathcal{K}_{loop}$. Then we have the following lemma. Remind that, from Theorem 5.1, $\exp\{w(c_1, c_2)\}$ is our upper bound of the decay rate $\eta^*(c_1, c_2, d_1, d_2)$ of $p(n_1, n_2)$ along line $l((c_1, c_2, d_1, d_2)$.

Lemma 5.4 *The function $w(c_1, c_2)$ is given as follows.*

$$\text{If } (b_1^*, b_2^*) \in \mathcal{R}^0, \ \text{then} \ w(c_1, c_2) = c_1 b_1^* + c_2 b_2^*. \tag{5.19}$$

$$\text{If } (b_1^*, b_2^*) \in \mathcal{R}_0^1, \ \text{then} \ w(c_1, c_2) = \begin{cases} c_1 b_1^* + c_2 \hat{b}_2^* & for \ -c_1/c_2 \le (\hat{b}_2^* - b_2^*)/(b_1^* - \hat{b}_1^*), \\ c_1 \hat{b}_1^* + c_2 b_2^* & for \ -c_1/c_2 \ge (\hat{b}_2^* - b_2^*)/(b_1^* - \hat{b}_1^*). \end{cases} \tag{5.20}$$

$$\text{If } (b_1^*, b_2^*) \in \mathcal{R}_1^1, \ \text{then} \ w(c_1, c_2) = c_1 b_1^* + c_2 \hat{b}_2^*. \tag{5.21}$$

$$\text{If } (b_1^*, b_2^*) \in \mathcal{R}_2^1, \ \text{then} \ w(c_1, c_2) = c_1 \hat{b}_1^* + c_2 b_2^*. \tag{5.22}$$

$$\text{If } (b_1^*, b_2^*) \in \mathcal{R}_0^2, \ \text{then} \ w(c_1, c_2) = \begin{cases} c_1 b_1^* + c_2 \hat{b}_2^* & for \ -c_1/c_2 \le \sigma(b_1^*, \hat{b}_2^*), \\ v(c_1, c_2) & for \ \sigma(b_1^*, \hat{b}_2^*) \le -c_1/c_2 \le \sigma(\hat{b}_1^*, b_2^*), \\ c_1 \hat{b}_1^* + c_2 b_2^* & for \ -c_1/c_2 \ge \sigma(\hat{b}_1^*, b_2^*). \end{cases} \tag{5.23}$$

$$\text{If } (b_1^*, b_2^*) \in \mathcal{R}_1^2, \ \text{then} \ w(c_1, c_2) = \begin{cases} c_1 b_1^* + c_2 \hat{b}_2^* & for \ -c_1/c_2 \le \sigma(b_1^*, \hat{b}_2^*), \\ v(c_1, c_2) & for \ -c_1/c_2 \ge \sigma(b_1^*, \hat{b}_2^*). \end{cases} \tag{5.24}$$

$$\text{If } (b_1^*, b_2^*) \in \mathcal{R}_2^2, \ \text{then} \ w(c_1, c_2) = \begin{cases} v(c_1, c_2) & for \ -c_1/c_2 \le \sigma(\hat{b}_1^*, b_2^*), \\ c_1 \hat{b}_1^* + c_2 b_2^* & for \ -c_1/c_2 \ge \sigma(\hat{b}_1^*, b_2^*). \end{cases} \tag{5.25}$$

$$\text{If } (b_1^*, b_2^*) \in \mathcal{R}_3^2, \ \text{then} \ w(c_1, c_2) = v(c_1, c_2). \tag{5.26}$$

Proof The case $(b_1^*, b_2^*) \in \mathcal{R}^0$ is a special case where (b_1^*, b_2^*) is on $\mathcal{K}_{loop}$ and hence $\hat{b}_k^* = b_k^*$ from the definition (5.6). Since (b_1^*, b_2^*) is in $\mathcal{W}^+$ or $\mathcal{W}^-$, $w(c_1, c_2)$ is always given by $c_1 b_1^* + c_2 b_2^*$ from (5.16). This proves (5.19).

If $(b_1^*, b_2^*) \in \mathcal{R}_0^1$, then $(b_1^*, b_2^*) \in \mathcal{W}^+$ for small $-c_1/c_2$, and $(b_1^*, b_2^*) \in \mathcal{W}^-$ for large $-c_1/c_2$. The changing point is determined by the equation $c_1 b_1^* + c_2 \hat{b}_2^* = c_1 \hat{b}_1^* + c_2 b_2^*$. A trite calculation proves (5.20). If $(b_1^*, b_2^*) \in \mathcal{R}_1^1$, then (b_1^*, b_2^*) stays in $\mathcal{W}^+$ for any $-c_1/c_2$. Hence, from (5.16), $w(c_1, c_2) = c_1 b_1^* + c_2 \hat{b}_2^*$, and this proves (5.21). (5.22) is proved in a similar manner.

If $(b_1^*, b_2^*) \in \mathcal{R}_0^2$, then $(b_1^*, b_2^*) \in \mathcal{W}^-$ for large $-c_1/c_2$, $(b_1^*, b_2^*) \in \mathcal{W}^0$ for medium $-c_1/c_2$, and $(b_1^*, b_2^*) \in \mathcal{W}^+$ for small $-c_1/c_2$. The changing points are given by equations $b_2^* = \tau_2(-c_1/c_2)$ and $b_1^* = \tau_1(-c_1/c_2)$. Since these equations are equivalent to $-c_1/c_2 = \sigma(\hat{b}_1^*, b_2^*)$ and to $-c_1/c_2 = \sigma(b_1^*, \hat{b}_2^*)$ respectively, we have (5.23). When $(b_1^*, b_2^*) \in \mathcal{R}_1^2$, the case $(b_1^*, b_2^*) \in \mathcal{W}^-$ cannot occur. When $(b_1^*, b_2^*) \in \mathcal{R}_2^2$, the case $(b_1^*, b_2^*) \in \mathcal{W}^+$ cannot occur. When $(b_1^*, b_2^*) \in \mathcal{R}_3^2$, always $(b_1^*, b_2^*) \in \mathcal{W}^0$. Hence (5.24)$\sim$(5.26) are proved immediately from (5.16). $\diamondsuit$

Combining the above lemma with Theorem 5.1, we have an upper bound for the decay rate $\eta^*(c_1, c_2, d_1, d_2)$ of the joint queue-length probability $p(n_1, n_2)$ along line $l(c_1, c_2, d_1, d_2)$.

6 Use of upper bound $\overline{\eta}_k^*$

In the preceding section, the upper bound $\exp\{w(c_1, c_2)\}$ of $\eta^*(c_1, c_2, d_1, d_2)$ was derived from the exact decay rates η_1^* and η_2^* of the marginal queue-length distributions. However, these marginal decay rates are usually unknown. By scrutinizing the deriving process, we see that another upper bound can be derived in the same way by using arbitrary upper bounds for the marginal decay rates. Here we shall apply the upper bound $\overline{\eta}_k^*$ proposed in [5].

For arbitrarily given positive numbers η_k (< 1), we let

$$\begin{aligned}
\mathcal{H}^+(\eta_1) &= \{(a_1, a_2) \ : \ \kappa(a_1, a_2) \le 0, \ \log \eta_1 < a_1 < 0, \ a_2 < 0\}, \quad \text{and} \\
\mathcal{H}^-(\eta_2) &= \{(a_1, a_2) \ : \ \kappa(a_1, a_2) \le 0, \ a_1 < 0, \ \log \eta_2 < a_2 < 0\}.
\end{aligned} \tag{6.1}$$

We choose $\eta_k^\sharp$ so that $\eta_k^* \le \eta_k^\sharp < 1$, then $\mathcal{H}^+(\eta_1^\sharp)$ is a subset of $\mathcal{H}^+$ and $\mathcal{H}^-(\eta_2^\sharp)$ is a subset of $\mathcal{H}^-$. So, our fundamental lemma, Lemma 4.1, still holds if we use $\mathcal{H}^+(\eta_1^\sharp)$ instead of $\mathcal{H}^+$ and $\mathcal{H}^-(\eta_2^\sharp)$ instead of $\mathcal{H}^-$. We let $w^\sharp(c_1, c_2; \eta_1^\sharp, \eta_2^\sharp)$ be the function defined by (5.3) using $\mathcal{H}^+(\eta_1^\sharp)$ and $\mathcal{H}^-(\eta_2^\sharp)$ instead of $\mathcal{H}^+$ and $\mathcal{H}^-$. Then it is represented as in (5.19) $\sim$ (5.26) where b_k^* is replaced with $b_k^\sharp = \log \eta_k^\sharp$ and $\hat{b}_{k'}^*$ with $\hat{b}_{k'}^\sharp = \theta_{k'}(b_k^\sharp)$.

The following lemma is a trivial generalization of Theorem 5.1.

Lemma 6.1 *For positive numbers $\eta_1^\sharp$ and $\eta_2^\sharp$ such that $\eta_k^* \le \eta_k^\sharp < 1$, we let $\eta^\sharp(c_1, c_2; \eta_1^\sharp, \eta_2^\sharp) = \exp\{w^\sharp(c_1, c_2; \eta_1^\sharp, \eta_2^\sharp)\}$. Then it is an upper bound of the decay rate $\eta^*(c_1, c_2, d_1, d_2)$ of $p(n_1, n_2)$ along $l(c_1, c_2, d_1, d_2)$ for arbitrarily given positive integers c_1, c_2, d_1 and d_2. Namely*

$$\eta^*(c_1, c_2, d_1, d_2) \le \eta^\sharp(c_1, c_2; \eta_1^\sharp, \eta_2^\sharp). \tag{6.2}$$

In the previous paper [5], the authors have derived an upper bound $\overline{\eta}_k^*$ of the decay rate η_k^* of the *marginal* queue-length distribution for node k. The upper bound is defined as follows. For the function $h_k(a_k)$ defined in (3.12) we let

$$\mathcal{E}_k = \{(a_1, a_2) \in \mathcal{K}_{loop} \ : \ a_k < 0 \text{ and } h_k(a_k) \le a_{k'} \le 0\}, \quad \text{and} \tag{6.3}$$

$$b_k^{\mathcal{E}_k} = \inf\{a_k \ : \ \exists\, a_{k'} \text{ such that } (a_1, a_2) \in \mathcal{E}_k\}. \tag{6.4}$$

Further we restrict $\mathcal{E}_k$ for $a_{k'}$ by $b_{k'}^{\mathcal{E}_{k'}}$ as

$$\begin{aligned}
\mathcal{F}_k &= \mathcal{E}_k \cap \{(a_1, a_2) \ : \ a_{k'} \geq b_{k'}^{\mathcal{E}_{k'}}\} \\
&= \{(a_1, a_2) \in \mathcal{K}_{loop} \ : \ a_k < 0 \text{ and } \max\{h_k(a_k), b_{k'}^{\mathcal{E}_{k'}}\} \leq a_{k'} \leq 0\},
\end{aligned} \tag{6.5}$$

and let

$$b_k^{\mathcal{F}_k} = \inf\{a_k \ : \ \exists\, a_{k'} \text{ such that } (a_1, a_2) \in \mathcal{F}_k\}. \tag{6.6}$$

Then our upper bound is defined as

$$\overline{\eta}_k^* = \exp\{b_k^{\mathcal{F}_k}\}. \tag{6.7}$$

Theorem 4.1 in [5] proved that $\eta_k^* \leq \overline{\eta}_k^* < 1$. We will write as $\overline{\eta}(c_1, c_2)$ the upper bound $\eta^\sharp(c_1, c_2; \overline{\eta}_1^*, \overline{\eta}_2^*)$ in (6.2) using $\overline{\eta}_k^*$.

To describe $\overline{\eta}_k^*$ or $b_k^{\mathcal{F}_k}$ more concretely, we prepare some notations. The curve $a_{k'} = h_k(a_k)$ on the (a_1, a_2)-plane intersects with $\mathcal{K}_{loop}$ at two points, at the origin $(0,0)$ and at a point having negative k-th coordinate. Let $(b_1^{h_k}, b_2^{h_k})$ be the coordinates of the latter point, namely, for example, if $k = 1$,

$$b_1^{h_1} = \{\text{unique negative root of equation } \kappa(a_1, h_1(a_1)) = 0 \text{ for } a_1\} \quad \text{and} \quad b_2^{h_1} = h_1(b_1^{h_1}). \tag{6.8}$$

Further we let $\hat{b}_{k'}^{h_k} = \theta_{k'}(b_k^{h_k})$. (This quantity was represented by $b_{k'}^{h_k, c}$ in [5] with a slightly different definition.) Note that, for example, $\hat{b}_2^{h_1}$ is the second coordinate of the lower intersection of the straight line $a_1 = b_1^{h_1}$ with $\mathcal{K}_{loop}$. Further we denote as

$$\eta_k^{h_k} = \exp\{b_k^{h_k}\}, \quad \eta_{k'}^{h_k} = \exp\{b_{k'}^{h_k}\} \quad \text{and} \quad \hat{\eta}_{k'}^{h_k} = \exp\{\hat{b}_{k'}^{h_k}\}. \tag{6.9}$$

In Corollary 4.2 of [5], it was shown that the pair $(b_1^{\mathcal{F}_1}, b_2^{\mathcal{F}_2})$, which derives the upper bounds $\overline{\eta}_1^*$ and $\overline{\eta}_2^*$, takes one of the following six forms depending on conditions designated in Theorem 6.2 below:

(a) $(b_1^{h_1}, b_2^{h_2})$ $(\in \mathcal{R}_0^1 \cup \mathcal{R}^0 \cup \mathcal{R}_0^2)$

(b) $(b_1^{h_1}, \hat{b}_2^{h_1})$ $(\in \mathcal{R}^0)$

(c) $(\hat{b}_1^{h_2}, b_2^{h_2})$ $(\in \mathcal{R}^0)$

(d) $(b_1^{h_1}, b_2^{\mathcal{K}_2})$ (on the boundary between $\mathcal{R}_0^2$ and $\mathcal{R}_1^2$)

(e) $(b_1^{\mathcal{K}_1}, b_2^{h_2})$ (on the boundary between $\mathcal{R}_0^2$ and $\mathcal{R}_2^2$)

(f) $(b_1^{\mathcal{K}_1}, b_2^{\mathcal{K}_2})$ (at the corner of $\mathcal{R}_3^2$)

We will refer to them as "type" of the model. Then, from Lemma 5.4, we can get a concrete expression for $\overline{\eta}(c_1, c_2) = \eta^\sharp(c_1, c_2; \overline{\eta}_1^*, \overline{\eta}_2^*)$ in each type.

Theorem 6.2 *For positive integers c_1, c_2, d_1 and d_2, the decay rate $\eta^*(c_1, c_2, d_1, d_2)$ of the joint queue-length probability $p(n_1, n_2)$ along line $l(c_1, c_2, d_1, d_2)$ is bounded from above by $\overline{\eta}(c_1, c_2)$ given below.*

(a) *If* $\max\{b_1^{h_1}, b_1^{\mathcal{K}_2}\} < b_1^{h_2}$ *and* $\max\{b_2^{h_2}, b_2^{\mathcal{K}_1}\} < b_2^{h_1}$ $\left[(b_1^{\mathcal{F}_1}, b_2^{\mathcal{F}_2}) = (b_1^{h_1}, b_2^{h_2})\right]$, *then*

(a-0) *if* $\kappa(b_1^{h_1}, b_2^{h_2}) = 0$ $\left[(b_1^{h_1}, b_2^{h_2}) \in \mathcal{R}^0\right]$,

$$\overline{\eta}(c_1, c_2) = (\eta_1^{h_1})^{c_1}(\eta_2^{h_2})^{c_2}, \tag{6.10}$$

(a-1) *if* $\kappa(b_1^{h_1}, b_2^{h_2}) < 0$ $\left[(b_1^{h_1}, b_2^{h_2}) \in \mathcal{R}_0^1\right]$,

$$\overline{\eta}(c_1, c_2) = \begin{cases} (\eta_1^{h_1})^{c_1}(\hat{\eta}_2^{h_1})^{c_2} & for \ -c_1/c_2 \le (\hat{b}_2^{h_1} - b_2^{h_2})/(b_1^{h_1} - \hat{b}_1^{h_2}), \\ (\hat{\eta}_1^{h_2})^{c_1}(\eta_2^{h_2})^{c_2} & for \ -c_1/c_2 \ge (\hat{b}_2^{h_1} - b_2^{h_2})/(b_1^{h_1} - \hat{b}_1^{h_2}), \end{cases} \tag{6.11}$$

(a-2) *if* $\kappa(b_1^{h_1}, b_2^{h_2}) > 0$ $\left[(b_1^{h_1}, b_2^{h_2}) \in \mathcal{R}_0^2\right]$,

$$\overline{\eta}(c_1, c_2) = \begin{cases} (\eta_1^{h_1})^{c_1}(\hat{\eta}_2^{h_1})^{c_2} & for \ -c_1/c_2 \le \sigma(b_1^{h_1}, \hat{b}_2^{h_1}), \\ \exp\{v(c_1, c_2)\} \ (= \exp\{c_1\tau_1(-c_1/c_2) + c_2\tau_2(-c_1/c_2)\}) \\ \qquad for \ \sigma(b_1^{h_1}, \hat{b}_2^{h_1}) \le -c_1/c_2 \le \sigma(\hat{b}_1^{h_2}, b_2^{h_2}), \\ (\hat{\eta}_1^{h_2})^{c_1}(\eta_2^{h_2})^{c_2} & for \ -c_1/c_2 \ge \sigma(\hat{b}_1^{h_2}, b_2^{h_2}). \end{cases} \tag{6.12}$$

(b) *If* $\max\{b_1^{h_2}, b_1^{\mathcal{K}_2}\} \le b_1^{h_1}$ $\left[(b_1^{\mathcal{F}_1}, b_2^{\mathcal{F}_2}) = (b_1^{h_1}, \hat{b}_2^{h_1}) \in \mathcal{R}^0\right]$, *then*

$$\overline{\eta}(c_1, c_2) = (\eta_1^{h_1})^{c_1}(\hat{\eta}_2^{h_1})^{c_2}. \tag{6.13}$$

(c) *If* $\max\{b_2^{h_1}, b_2^{\mathcal{K}_1}\} \le b_2^{h_2}$ $\left[(b_1^{\mathcal{F}_1}, b_2^{\mathcal{F}_2}) = (\hat{b}_1^{h_2}, b_2^{h_2}) \in \mathcal{R}^0\right]$, *then*

$$\overline{\eta}(c_1, c_2) = (\hat{\eta}_1^{h_2})^{c_1}(\eta_2^{h_2})^{c_2}. \tag{6.14}$$

(d) *If* $\max\{b_1^{h_1}, b_1^{h_2}\} \le b_1^{\mathcal{K}_2}$ *and* $b_2^{h_1} > b_2^{\mathcal{K}_1}$ $\left[(b_1^{\mathcal{F}_1}, b_2^{\mathcal{F}_2}) = (b_1^{h_1}, b_2^{\mathcal{K}_2}) \in \mathcal{R}_1^2\right]$, *then*

$$\overline{\eta}(c_1, c_2) = \begin{cases} (\eta_1^{h_1})^{c_1}(\hat{\eta}_2^{h_1})^{c_2} & for \ -c_1/c_2 \le \sigma(b_1^{h_1}, \hat{b}_2^{h_1}), \\ \exp\{v(c_1, c_2)\} & for \ -c_1/c_2 \ge \sigma(b_1^{h_1}, \hat{b}_2^{h_1}). \end{cases} \tag{6.15}$$

(e) *If* $\max\{b_2^{h_1}, b_2^{h_2}\} \le b_2^{\mathcal{K}_1}$ *and* $b_1^{h_2} > b_1^{\mathcal{K}_2}$ $\left[(b_1^{\mathcal{F}_1}, b_2^{\mathcal{F}_2}) = (b_1^{\mathcal{K}_1}, b_2^{h_2}) \in \mathcal{R}_2^2\right]$, *then*

$$\overline{\eta}(c_1, c_2) = \begin{cases} \exp\{v(c_1, c_2)\} & for \ -c_1/c_2 \le \sigma(\hat{b}_1^{h_2}, b_2^{h_2}), \\ (\hat{\eta}_1^{h_2})^{c_1}(\eta_2^{h_2})^{c_2} & for \ -c_1/c_2 \ge \sigma(\hat{b}_1^{h_2}, b_2^{h_2}). \end{cases} \tag{6.16}$$

(f) *If* $b_2^{h_1} \le b_2^{\mathcal{K}_1}$ *and* $b_1^{h_2} \le b_1^{\mathcal{K}_2}$ $\left[(b_1^{\mathcal{F}_1}, b_2^{\mathcal{F}_2}) = (b_1^{\mathcal{K}_1}, b_2^{\mathcal{K}_2}) \in \mathcal{R}_3^2\right]$, *then*

$$\overline{\eta}(c_1, c_2) = \exp\{v(c_1, c_2)\} \tag{6.17}$$

Remark 6.1 Note that the function $\overline{\eta}(c_1, c_2)$ takes of the form $\eta_1^{c_1}\eta_2^{c_2}$ in any case. Hence $\overline{\eta}(c_1 n + d_1, c_2 n + d_2) = \eta_1^{d_1}\eta_2^{d_2}\{\overline{\eta}(c_1, c_2)\}^n$, and the result of Theorem 6.2 can be understood as

$$\limsup_{n_1, n_2 \to \infty} \left\{\frac{p(n_1, n_2)}{\overline{\eta}(n_1, n_2)}\right\}^{c_1/n_1} \le 1 \tag{6.18}$$

when n_1 and n_2 get large along line $l(c_1, c_2, d_1, d_2)$.

20

Remark 6.2 We can easily check that Theorem 5.1 holds for case $c_1 = 0$ and $c_2 > 0$. However, $\overline{\eta}(0, c_2) = \overline{\eta}_2^*$ for any c_2 and thus the upper bound for $\eta^*(0, c_2, d_1, d_2)$ is trivial.

Remark 6.3 In the proof of Lemma 4.1, we use Lemma 4.2 instead of handling the rate matrix $\boldsymbol{R}$ directly. To derive a lower bound of $\eta^*(c_1, c_2, d_1, d_2)$ in the same manner, we need to find a vector $\boldsymbol{q}$ satisfying $\boldsymbol{q}\boldsymbol{R} \geq \xi\boldsymbol{q}$ for some constant ξ. However, we presently have no way to construct such a vector, and this is left as a future work.

7 Examples and discussions

Example 7.1 (Tandem queueing system 1) Fujimoto et al. [3] discussed decay rate of the joint queue-length probabilities in a two-stage tandem queueing system $PH/PH/s_1 \to /PH/s_2$. We compare our result with theirs for the single server type $s_1 = s_2 = 1$. The tandem configuration requires that $\lambda_2 = 0$, $r_{12} = r_{20} = 1$ and $r_{10} = r_{21} = 0$. Hence $\phi_2(a_2) \equiv 0$, $h_1(a_1) \equiv 0$ and $h_2(a_2) = a_2$. Then $\kappa(a_1, a_2) = \phi_1(a_1) + \psi_1(-a_1 + a_2) + \psi_2(-a_2)$, and b-values satisfy relations $b_1^{h_1} = b_1^0$, $b_2^{h_1} = 0$ and $b_1^{h_2} = b_2^{h_2}$. Notice that the exact decay rate η_1^* of the marginal queue-length distribution of Node 1 is given by $\exp\{b_1^0\}$.

Their results are roughly summarized as follows[1].

(i) Under the condition $b_2^{\mathcal{K}(1)} < 0$, the decay rate $\eta^*(c_1, c_2, d_1, d_2)$ along line $l(c_1, c_2, d_1, d_2)$ is given by $(\eta_1^{h_1})^{c_1}(\hat{\eta}_2^{h_2})^{c_2}$ when $-c_1/c_2$ is sufficiently small (Theorem 3.1 of [3]).

(ii) Under the condition $\max\{b_1^{h_1}, b_1^{\mathcal{K}(2)}\} < b_1^{h_2}$, the decay rate is given by $(\hat{\eta}_1^{h_1})^{c_1}(\eta_2^{h_2})^{c_2}$ when $-c_1/c_2$ is sufficiently close to 0 (Theorem 3.2 of [3]).

Consider the case (i) and assume that $b_2^{\mathcal{K}(1)} < 0$. Then, when $-c_1/c_2$ is very small, the boundary curve $\mathcal{V}$ between regions $\mathcal{W}^+$ and $\mathcal{W}^-$ (see Figure 5.3) is near the upper left arc of $\mathcal{K}_{loop}$ from $(b_1^0, 0)$ to $(b_1^{\mathcal{K}_1}, b_2^{\mathcal{K}_1})$ as indicated in Figure 5.5. Since $\eta_1^* = \exp\{b_1^0\}$, this implies that the point $(b_1^*, b_2^*) = (\log \eta_1^*, \log \eta_2^*)$ is in $\mathcal{W}^+$ irrespective of the value of η_2^*. Hence our upper bound for $\eta^*(c_1, c_2, d_1, d_2)$ in Theorem 5.1 is given by $\exp\{c_1 b_1^* + c_2 \hat{b}_2^*\} = (\eta_1^*)^{c_1}(\hat{\eta}_2^*)^{c_2} = (\eta_1^{h_1})^{c_1}(\hat{\eta}_2^{h_2})^{c_2}$ from Lemma 5.3, and this coincides with the exact decay rate given in [3]. For the bound using $\overline{\eta}_k^*$, among the eight types of Theorem 6.2, five types, (a-0), (a-1), (a-2), (b) and (d), may occur under the condition $b_2^{\mathcal{K}(1)} < 0$. In types (a-1), (a-2), (b) and (d), we can easily check that our upper

---

[1]Theorem 3.1 of [3] is rigorously stated in the following manner. Here we use our notations. Note that i_2, the phase of the exogenous arrival process of Node 2, is always equal to 0 from the tandem assumption. If $b_2^{\mathcal{K}(1)} < 0$, for fixed n_2, i_1, j_1 and j_2, the stationary state probability decays geometrically with rate $\eta_1^{h_1}$ as $n_1 \to \infty$:

$$p(n_1, n_2)_{i_1, 0, j_1, j_2} \sim G_1(n_2; i_1; j_1, j_2)\,(\eta_1^{h_1})^{n_1}.$$

The multiplicative constant $G_1(n_2; i_1; j_1, j_2)$ decays geometrically with rate $\hat{\eta}_2^{h_2}$ as $n_2 \to \infty$:

$$G_1(n_2; i_1; j_1, j_2) \sim G_2\, C_0(i_1)\, C_1(j_1)\, C_2(j_2)\,(\hat{\eta}_2^{h_2})^{n_2},$$

where $C_0(i_1)$, $C_1(j_1)$ and $C_2(j_2)$ are constants determined from the vector $\boldsymbol{\nu}$ and G_2 is a constant independent of n_2, i_1, j_1 and j_2. The original condition of the theorem is the one similar to $\hat{\eta}_2^{h_2} < 1$, but their definition of the corresponding quantity to $\hat{\eta}_2^{h_2}$ is slightly different from ours and the condition is equivalent to $b_2^{\mathcal{K}(1)} < 0$.

Table 7.1: Corresponding types of models listed in Table 1 of [2]

no.	model	ρ_1	ρ_2	type	no.	model	ρ_1	ρ_2	type
11	$M/H_2/1 \to /E_2/1$	0.60	0.40	(b)	15	$E_4/M/1 \to /H_2/1$	0.60	0.40	(b)
12	$E_2/E_2/1 \to /E_2/1$	0.60	0.40	(b)	16	$M/E_2/1 \to /E_2/1$	0.60	0.35	(b)
13	$M/H_2/1 \to /H_2/1$	0.60	0.40	(b)	17	$H_2/E_2/1 \to /E_2/1$	0.60	0.40	(b)
14	$E_2/H_2/1 \to /E_2/1$	0.60	0.40	(b)	18	$M/E_2/1 \to /H_2/1$	0.60	0.20	(b)

bound for small $-c_1/c_2$ is given by $(\eta_1^{h_1})^{c_1}(\hat{\eta}_2^{h_1})^{c_2}$. In type (a-0), we see that $\hat{b}_2^{h_1} = b_2^{h_2}$. Hence the upper bound is given by $(\eta_1^{h_1})^{c_1}(\eta_2^{h_2})^{c_2} = (\eta_1^{h_1})^{c_1}(\hat{\eta}_2^{h_1})^{c_2}$. Thus in each of the five types, our bound is equal to $(\eta_1^{h_1})^{c_1}(\hat{\eta}_2^{h_1})^{c_2}$ and coincides with the exact decay rate.

Next consider the case (ii) and assume that $\max\{b_1^{h_1}, b_1^{\mathcal{K}(2)}\} < b_1^{h_2}$. When $-c_1/c_2$ is close to 0, the boundary curve $\mathcal{V}$ is near the lower right arc of $\mathcal{K}_{loop}$ from $(0, b_2^0)$ to $(b_1^{\mathcal{K}_2}, b_2^{\mathcal{K}_2})$. However, in this type, we cannot discuss our upper bound given in Theorem 5.1 since we don't know the value of η_2^*. For the upper bound using $\overline{\eta}_k^*$, the condition $\max\{b_1^{h_1}, b_1^{\mathcal{K}(2)}\} < b_1^{h_2}$ may not be violated in five types, (a-0), (a-1), (a-2), (c) and (e), among the eight types of Theorem 6.2. In types (a-1), (a-2), (c) and (e), we can easily check that our upper bound for $-c_1/c_2$ being near to 0 is given by $(\hat{\eta}_1^{h_1})^{c_1}(\eta_2^{h_1})^{c_2}$. In type (a-0), we see that $\hat{b}_2^{h_2} = b_1^{h_1}$. Hence the upper bound is given by $(\eta_1^{h_1})^{c_1}(\eta_2^{h_2})^{c_2} = (\hat{\eta}_1^{h_1})^{c_1}(\eta_2^{h_1})^{c_2}$. Thus in each of the five types, our bound is equal to $(\hat{\eta}_1^{h_1})^{c_1}(\eta_2^{h_1})^{c_2}$ and coincides with the exact decay rate.

Example 7.2 (Tandem queueing system 2) Fujimoto et al. [2] reported results of an extensive numerical experiment on the joint queue-length probability $p(n_1, n_2)$ for tandem queueing systems PH/PH/1→/PH/1, and gave a conjecture on the decay rate. We shall see their results from our point of view of types given in Theorem 6.2 with numerical results for the function

$$\gamma(n_1, n_2) = \frac{p(n_1, n_2)}{\overline{\eta}(n_1, n_2)}. \tag{7.1}$$

Note that the function $\gamma(n_1, n_2)$ is the quantity in the braces of (6.18).

Models of type (b): In Table 1 of [2], numerical results were presented for eight models listed in Table 7.1[2] on the ratio

$$g(n_1, n_2) = \frac{p(n_1, n_2)}{(\eta_1^{h_1})^{n_1}(\hat{\eta}_2^{h_1})^{n_2}} \tag{7.2}$$

when n_1 and n_2 run along lines $l(20, 5, 15, 5)$ and $l(5, 20, 5, 15)$. The results of [2] shows that the ratio converges to a common limit in each model. By the classification in Theorem 6.2, all the models are of type (b) and $\overline{\eta}(c_1, c_2) = (\eta_1^{h_1})^{c_1}(\hat{\eta}_2^{h_1})^{c_2}$. Hence the function $g(n_1, n_2)$ coincides with

---

[2]As for model representation, we use a similar notation to the Kendall's one. For inter-arrival and service time distributions, M stands for an exponential distribution, E_2 for an Erlang distribution with 2 phases, and H_2 for a hyper exponential distribution with 2 phases having distribution function of the form $F(x) = 0.2(1 - e^{-3.4\nu x}) + 0.8(1 - e^{-0.85\nu x})$.
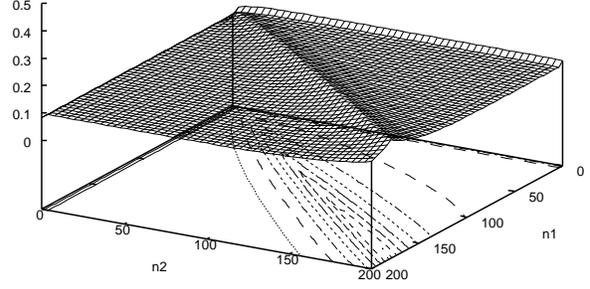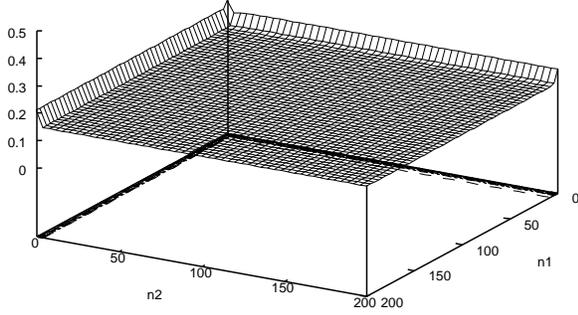
Figure 7.1: Graph of $\gamma(n_1, n_2)$ for the Model 16

Figure 7.2: Graph of $\gamma(n_1, n_2)$ for the Model 21

Table 7.2: Corresponding types of models listed in Table 2 of [2]

(group 1)

no.	model	ρ_1	ρ_2	type	$(\hat{b}_2^{h_1} - b_2^{h_2})/(b_1^{h_1} - \hat{b}_1^{h_2})$
21	$M/H_2/1 \rightarrow /E_2/1$	0.60	0.80	(a-1)	-0.901
22	$E_2/E_2/1 \rightarrow /E_2/1$	0.60	0.70	(a-0)	NAN
23	$M/H_2/1 \rightarrow /H_2/1$	0.60	0.75	(a-1)	-0.912
24	$E_2/H_2/1 \rightarrow /E_2/1$	0.60	0.85	(a-1)	-1.050
25	$E_4/M/1 \rightarrow /H_2/1$	0.60	0.77	(a-1)	-0.965

(group 2)

no.	model	ρ_1	ρ_2	type	$\sigma(b_1^{h_1}, \hat{b}_2^{h_1})$	$\sigma(\hat{b}_1^{h_2}, b_2^{h_2})$
26	$M/E_2/1 \rightarrow /E_2/1$	0.60	0.71	(a-2)	-2.506	-0.216
27	$H_2/E_2/1 \rightarrow /E_2/1$	0.60	0.70	(a-2)	-3.403	-0.050
28	$M/E_2/1 \rightarrow /H_2/1$	0.60	0.70	(a-2)	-2.786	-0.593

$\gamma(n_1, n_2)$. The authors draw graphs of $\gamma(n_1, n_2)$ for these models, and see that all the graphs are almost flat except near the axes as designated in Fig. 7.1 for Model 16. From the numerical results and the graphs, we may expect that, in a model of type (b), there exists a common limit C such that the joint state probability $p(n_1, n_2)$ decays geometrically in the sense that, for any positive integers c_1, c_2, d_1 and d_2,

$$\lim_{n \to \infty} \frac{p(c_1 n + d_1, c_2 n + d_2)}{(\eta_1^{h_1})^{c_1 n + d_1} (\hat{\eta}_2^{h_1})^{c_2 n + d_2}} = C. \tag{7.3}$$

Of course this is only a conjecture, and it is an open problem whether this property generally holds or not.

Models of type (a-1): In Table 2 (a) of [2], numerical results were presented for five models listed in Table 7.2 (group 1) on the ratio $g(n_1, n_2)$ in (7.2) when n_1 and n_2 run along line $l(20, 5, 15, 5)$

23
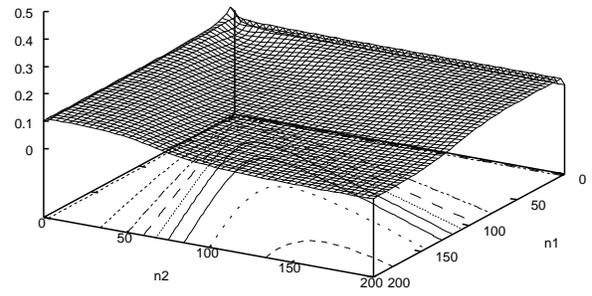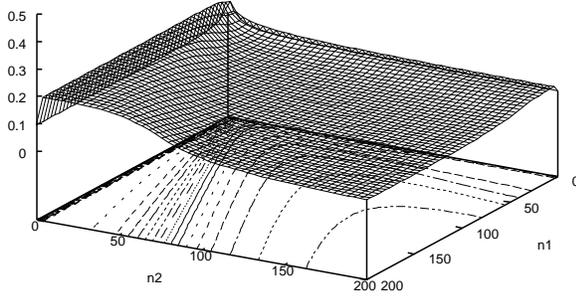
Figure 7.3: Graph of $\gamma(n_1, n_2)$ for the Model 26

Figure 7.4: Graph of $\gamma(n_1, n_2)$ for the Model 28

and on the ratio

$$\overline{g}(n_1, n_2) = \frac{p(n_1, n_2)}{(\hat{\eta}_1^{h_2})^{n_1}(\eta_2^{h_2})^{n_2}} \tag{7.4}$$

when n_1 and n_2 run along line $l(5, 20, 5, 15)$. Both ratios seem to converge. Figure 7.2 shows a graph of $\gamma(n_1, n_2)$ for Model 21. Note that this model is of type (a-1) of Theorem 6.2, and hence $\gamma(n_1, n_2) = g(n_1, n_2)$ for $-n_1/n_2 \leq u$ and $\gamma(n_1, n_2) = \overline{g}(n_1, n_2)$ for $-n_1/n_2 \geq u$, where $u = (\hat{b}_2^{h_1} - b_2^{h_2})/(b_1^{h_1} - \hat{b}_1^{h_2})$. The value of u for this particular model is -0.901 as presented in Table 7.2. The graph of $\gamma(n_1, n_2)$ in Fig. 7.2 is almost flat in the region $-n_1/n_2 < u$ and is also almost flat with another value in the region $-n_1/n_2 > u$ except for the neighborhood of the boundary. Hence we may expect that, for $-c_1/c_2 \leq u$, the geometric convergence (7.3) holds, and for $-c_1/c_2 \geq u$, the geometric convergence

$$\lim_{n \to \infty} \frac{p(c_1 n + d_1, c_2 n + d_2)}{(\hat{\eta}_1^{h_1})^{c_1 n + d_1}(\eta_2^{h_1})^{c_2 n + d_2}} = \overline{C} \tag{7.5}$$

holds with another constant $\overline{C}$. Models 23, 24 and 25 of Table 7.2 (group 1) exhibit similar behaviors in $\gamma(n_1, n_2)$. On the contrary, Model 22 is of type (a-0), and its behavior in $\gamma(n_1, n_2)$ is rather similar to those of type (b).

Models of type (a-2): In Table 2 (b) of [2], numerical results were presented for $g(n_1, n_2)$ in (7.2) when n_1 and n_2 run along line $l(20, 5, 15, 5)$ and for $\overline{g}(n_1, n_2)$ in (7.4) when n_1 and n_2 run along line $l(5, 20, 5, 15)$ for Models 26, 27 and 28 of Table 7.2 (group2). The paper [2] said that the values of these functions seemed converging but the convergence speed was very slow. In the classification of Theorem 6.2, these three models are all of type (a-2), and the values of thresholds $\sigma(b_1^{h_1}, \hat{b}_2^{h_1})$ and $\sigma(\hat{b}_1^{h_2}, b_2^{h_2})$ are as shown in Table 7.2 (group 2). For the line $l(20, 5, 15, 5)$, $-c_1/c_2$ is equal to -4.0, and for the line $l(5, 20, 5, 15)$ it is equal to -0.25. Hence in Models 26 and 27, the latter gradient $-c_1/c_2 = -0.25$ is in between $\sigma(b_1^{h_1}, \hat{b}_2^{h_1})$ and $\sigma(\hat{b}_1^{h_2}, b_2^{h_2})$. Therefore, the estimated decay rate $(\hat{\eta}_1^{h_2})^{n_1}(\eta_2^{h_2})^{n_2}$ used in (7.4) is clearly larger than the exact decay rate (if it exists), and the ratio cannot converge to a positive limit. In Figs. 7.3 and 7.4, the graph of $\gamma(n_1, n_2)$ is presented for Models 26 and 28, respectively, and they seem curving. Our numerical results also show some slow convergence or non-convergence in these models. So, to understand the decay property of $p(n_1, n_2)$ for models of type (a-2), we need further study.

24

Table 7.3: Two-node Markovian queueing system M/E_2–M/E_2

| no.	model	ρ_1	ρ_2	type	$\sigma(b_1^{h_1}, \hat{b}_2^{h_1})$	$\sigma(\hat{b}_1^{h_2}, b_2^{h_2})$
31	M/E_2-M/E_2	0.7	0.7	(a-2)	-2.216	-0.4500
32	M/E_2-M/E_2	0.65	0.5	(d)	-0.537	$-$
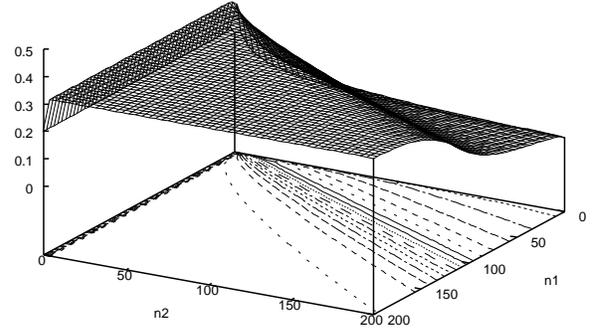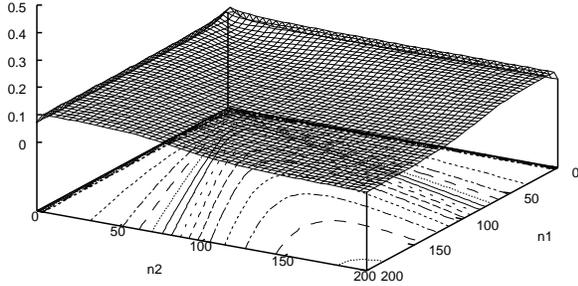
Figure 7.5: Graph of $\gamma(n_1, n_2)$ for the model 31   Figure 7.6: Graph of $\gamma(n_1, n_2)$ for the model 32

Example 7.3 (Two-node Markovian queueing system M/E_2–M/E_2) Finally, we show two examples of a two-node Markovian queueing system which is not a tandem queueing system. The models are listed in Table 7.3, and in both models $\lambda_1 = \lambda_2 = 1$ and $r_{12} = r_{21} = 0.4$. The graph of $\gamma(n_1, n_2)$ is presented in Figs. 7.5 and 7.6. Model 31 is of type (a-2), and the graph is similar to the one for Model 28 in Fig. 7.4. On the other hand, Model 32 is of type (d), and the graph is almost flat in the region $-c_1/c_2 < \sigma(b_1^{h_1}, \hat{b}_2^{h_1})$ and is curving in the region $-c_1/c_2 > \sigma(b_1^{h_1}, \hat{b}_2^{h_1})$. For details of the decaying behavior of $p(n_1, n_2)$, we need further study.

References

[1] N. G. Bean, J. M. Li and P. G. Taylor: Some asymptotic properties of two-stage tandem networks of PH/PH/1 queues, in: Advances in Matrix Analytic Methods for Stochastic Models, ed. Attahiru S. Alfa and Srinivas R. Chakravarthy (Notable Publications, New Jersey, 1998) pp.171-194.

[2] K. Fujimoto and Y. Takahashi: Tail behavior of the steady-state distribution in two-stage tandem queues: Numerical experiment and conjecture, Journal of the Operations Research Society of Japan 39 (1996) 525-540.

[3] K. Fujimoto, Y. Takahashi and N. Makimoto: Asymptotic Properties of Stationary Distributions in Two-stage Tandem Queueing Systems, Journal of the Operations Research Society of Japan 41 (1998) 118-141.

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

[5] K. Katou, N. Makimoto and Y. Takahashi: Upper bound for the decay rate of the marginal queue-length distribution in a two-node Markovian queueing system, Journal of the Operations Research Society of Japan 47 (2004) 314-338.

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

[7] 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#B-326 in Mathematical and Computing Sciences, Tokyo Institute of Technology, Japan (1997).

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

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

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