

CONJECTURES ON DECAY RATES OF TAIL PROBABILITIES IN GENERALIZED JACKSON AND BATCH MOVEMENT NETWORKS

Masakiyo Miyazawa
Tokyo University of Science

(Received Received April 25, 2001)

Abstract Asymptotic decay rates are considered for the stationary joint distributions of customer populations in a generalized Jackson network and a batch movement network. We first define an asymptotic decay rate for a multi-dimensional distribution concerning a tail set and a direction to decrease. Then, for the stationary joint distributions of customer populations, the decay rates are conjectured to be obtained through max-min linear optimizations with convex constraints. Their validity is checked for some known results. Furthermore, the conjectured decay rates are shown to be useful to see how the decay rates are changed according to modeling parameters and the direction to decrease in a two node tandem queue.

Keywords: Queue, generalized Jackson network, joint queue length, decay rate, tail probability, stationary distribution

1. Introduction

Our primary interest is to get asymptotic decay rates of tails in the stationary joint distributions of population vectors in queueing networks under a light tail assumption on service time distributions, where the population vector is the random vector whose i -th entry is the number of customers at node i . Complete answers are obtained for the product form networks, where the product form is meant that the stationary distribution of the whole network state is of the product of its marginal distributions with respect to nodes in the network. A typical example is a Jackson network. However, even for the Jackson network, the product form solution is easily destroyed by small changes of modeling assumptions. For instance, if service times of one server are changed to have a non-exponential distribution and if customers who are served by this server join another node with a positive probability, then we can not have the product form solution any more (e.g., see Corollary 11.11 of [5]). Here, FIFO (First-In First-Out) service discipline is assumed for the Jackson network. Unfortunately, for this class of networks, it is hard to get the stationary distributions of the population vectors. So it is interesting to see their tail behaviors.

We are concerned with two network models. The first model is a generalized Jackson network of single servers and a single class of customers, and we consider the asymptotic decay rates of the stationary tail probabilities of its population vector. The generalized Jackson network is a modification of the standard Jackson network in such a way that arrival processes at nodes are replaced by renewal processes and service time distributions at nodes are replaced by general distributions, while keeping all independence assumptions. Another model of our interest is a batch movement network such that its population vector is described by a continuous-time Markov chain. This model has been recently restudied, and the product form distribution is obtained under various situations (see, e.g., [5, 28]).

However, similar to the Jackson network, those product form solutions are easy to fail by small changes.

In this paper, we are only concerned with the asymptotic decay rate, so we frequently omit "asymptotic". The decay rates have been well studied for one dimensional distributions in queueing models and their networks with acyclic routing. Specially, for single server queues, they are obtained in terms of the moment generating functions of the accumulated input processes for the queue length and waiting time distributions (see, e.g., [6, 11]). The large deviation technique has been widely used for them as well as for networks with acyclic routing (e.g., [3, 4]). However, it seems to be hard to apply the large deviation approach to joint distributions in queueing networks. To the contrary, there have been appeared some results on the decay rates for joint distributions in the recent years. Those consider product form upper bounds for the tail probabilities (e.g., see [5, 9, 12–15, 21]). Particularly, linear combination bounds in [14] is notable. Also certain necessary and sufficient conditions for the product form bounds in [15] are suggestive. Those results help us to conjecture the asymptotic decay rates.

Since there seems to be no standard definition for the decay rate of a multi-dimensional distribution, we start with its definition, introducing a tail set and a direction vector along which the tail set decreases. We then consider the decay rates in the generalized Jackson network. For the case of tandem queues, Fujimoto and Takahashi [8] conjectured similar decay rates based on numerical experience. We use some heuristics but more rely on theoretical considerations, which enable to refine their conjectures and to consider more complicated situations. A key idea is to represent the decay rate in a parametric form using the direction vector. Based on this parametric form, we conjecture product form bounds and the decay rates, which can be determined through max-min linear optimizations with certain convex constraints. The conjectures are verified for the Jackson network, and are consistent with known results for a two node tandem queue with exponential servers. We also consider this tandem queue to see how the decay rates are changed according to modeling parameters and the direction to move.

We next consider the batch movement network. This model is originally of a continuous-time, but it is also described by a discrete-time Markov chain embedded at departure-arrival instants. In this network, customers at different nodes may simultaneously depart and arrive at different nodes, where batches of departures are first requested. A basic assumption is that the requested departure sizes and the arrival sizes are independent of the network state, but are subject to a generic joint distribution. Here, all the requested departures may not be realized if there are less customers. Similar to the case of the generalized Jackson network, we conjecture the decay rates as well as their product form bounds.

Through these two network models, an interesting observation is that the decay rates belong to the boundaries of bounded convex sets under certain regularity conditions, if they exist, while those convex sets are not hard to be identified. We work on how to choose appropriate decay rates among them.

This paper is made up by seven sections. In Section 2, we give definitions of an asymptotic exponential form and its decay rates. Section 3 illustrates a derivation of the asymptotic exponential form for a queue with a single waiting line. In Section 4, the generalized Jackson network is considered, and four conjectures on the decay rates are presented. Those conjectures are examined for three examples in Section 5. Section 6 discusses the batch movement network. We finally give some concluding remarks in Section 7.

2. Decay Rates in Queueing Models

Let X be a nonnegative random variable. If there are positive constants ω, b such that

$$\lim_{x \rightarrow \infty} P(X \geq x)e^{\omega x} = b, \quad (2.1)$$

then, the distribution of X is said to have an *asymptotically exponential tail* with decay rate ω and coefficient b . We also write (2.1) as

$$P(X \geq x) \sim be^{-\omega x}, \quad x \rightarrow \infty.$$

Thus, " \sim " in this paper is slightly stronger than the same order in the conventional usage. A weaker but still useful information is the rate ω itself. This is identified as

$$\lim_{x \rightarrow \infty} \frac{1}{x} \log P(X \geq x) = -\omega. \quad (2.2)$$

We refer to this ω as an *asymptotic decay rate* or simply *decay rate*.

In many queueing applications, X is a stationary characteristic such as the waiting time or queue length in the steady state. Suppose that the distribution of the X is obtained through a function of an additive process $Y(t)$ that represents a net input process of the queue. Then, the ω may be obtained from the function ψ defined as

$$\psi(\theta) = \lim_{t \rightarrow \infty} \frac{1}{t} \log E \exp(\theta Y(t)). \quad (2.3)$$

For instance, if $Y(t)$ is the total input minus the total potential output up to time t measured from time 0 and if X is the stationary workload, then the decay rate ω is obtained as a root of the equation $\psi(\theta) = 0$ satisfying $\psi'(\theta) > 0$, provided appropriate regularity conditions (see [11]).

The function ψ is used to determine a rate function in Large deviation principle, LDP in short. In this way, the decay rate is connected to LDP, and much work has been done on this matter (see, e.g., [6, 11]). A significant feature of LDP is its applicability without independence assumptions that typically appears in the classic queueing models such as $GI/GI/1$ queue. It has been applied to one dimensional distributions in acyclic queueing networks (see, e.g., [3, 4, 23, 24, 31]). However, LDP is more difficult to apply as dynamics of a queue becomes more complicated, in particular, for queueing networks with feedback routes. In this paper, we rather stick with the classical models, assuming various independence assumptions, while allowing a general network structure.

We are concerned with a multi-dimensional distribution. So we first need to clarify what is the decay rate of such a distribution. Let \mathbf{X} be a k -dimensional nonnegative random vector, and let $\langle \mathbf{x}, \mathbf{y} \rangle = \sum_{i=1}^k x_i y_i$ for vectors $\mathbf{x}, \mathbf{y} \in \mathbb{R}^k$, where \mathbb{R}^k is the k -dimensional Euclidean space with metric $|\mathbf{x}| \equiv \langle \mathbf{x}, \mathbf{x} \rangle^{\frac{1}{2}}$. We may define a decay rate for \mathbf{X} in the following way.

Definition 2.1 A nonnegative vector $\mathbf{c} \in \mathbb{R}^k$ is said to be a direction vector if $|\mathbf{c}| = 1$. For a convex set V and direction vector \mathbf{c} , the distribution of \mathbf{X} is said to have an asymptotic exponential tail if there exist positive numbers ω and b such that

$$\lim_{\mathbf{x}/|\mathbf{x}| \rightarrow \mathbf{c}, |\mathbf{x}| \rightarrow \infty} P(\mathbf{X} \in \mathbf{x} + V)e^{\omega |\mathbf{x}|} = b, \quad (2.4)$$

where $\mathbf{x} + V = \{\mathbf{x} + \mathbf{y} \in \mathbb{R}^k; \mathbf{y} \in V\}$. In this case, the set V and the scalar ω are said to be a tail set and a decay rate with respect to \mathbf{c} .

Remark 2.1 In general, ω and b depend on direction \mathbf{c} . When we need to specify this, we write them as $\omega(\mathbf{c})$ and $b(\mathbf{c})$, respectively.

Another definition of the decay rate is

$$\lim_{x \rightarrow \infty} P(\mathbf{X} \in x\mathbf{c} + V)e^{\omega x} = b. \quad (2.5)$$

This is a slightly weaker than what we need for our discussions. For the convex set V , we mainly take $\mathbb{R}_+^k \equiv \{\mathbf{y} \in \mathbb{R}^k; \mathbf{y} > \mathbf{0}\}$ and let \mathbf{x} run over integer valued vectors. Namely,

$$\lim_{\mathbf{n}/|\mathbf{n}| \rightarrow \mathbf{c}, |\mathbf{n}| \rightarrow \infty} P(\mathbf{X} \geq \mathbf{n})e^{\omega|\mathbf{n}|} = b.$$

Here, an inequality of vectors is of component-wise, i.e., $\mathbf{x} \geq \mathbf{y}$ if and only if $x_i \geq y_i$ for all i . In a conjecture in Section 6, we also consider the decay rate defined as

$$\lim_{n \rightarrow \infty} P(\langle \mathbf{X}, \mathbf{c} \rangle \geq n)e^{\omega n} = b.$$

This is equivalent to choosing $V = \{\mathbf{y} \in \mathbb{R}^k; \langle \mathbf{y}, \mathbf{c} \rangle \geq 0\}$ in (2.5).

3. Queues with Single Waiting Lines

To illustrate our approach, we consider the $GI/G/1$ queue with FIFO service. For simplicity, we first consider the waiting time, whose stationary distribution is well known to have the asymptotic exponential tail under some regularity conditions. Let S and T be random variables subject to the service time and inter-arrival time distributions, respectively. We assume the stability condition, $E(S) < E(T)$, and there exists a $\theta_0 < 0$ such that $E(e^{-\theta S}) < \infty$ for $\theta > \theta_0$ and $\lim_{\theta \downarrow \theta_0} E(e^{-\theta S}) = \infty$. The stability condition implies the existence of the stationary waiting time distribution. Let W be a random variable subject to this distribution. As is well known, if W is independent of $S - T$, then

$$P(W > x) = P(W + S - T > x), \quad x \geq 0. \quad (3.1)$$

We tentatively assume that W has an asymptotic exponential tail with rate ω , i.e., for some $b > 0$,

$$\lim_{x \rightarrow \infty} P(W > x)e^{\omega x} = b \quad (3.2)$$

We rewrite (3.1) as

$$e^{\omega x} P(W > x) = \int_{-\infty}^{+\infty} e^{\omega(x-u)} P(W > x-u) e^{\omega u} dP(S-T \leq u), \quad x \geq 0. \quad (3.3)$$

Since $e^{\omega x} P(W > x) \leq 1$ for all $x \geq 0$ by Kingman's inequality (see [16]), the bounded convergence theorem yields

$$\lim_{x \rightarrow \infty} e^{\omega x} P(W > x) = \int_{-\infty}^{+\infty} b e^{\omega u} dP(S-T \leq u).$$

Hence, using the notation $\hat{f}(\theta) = E(e^{-\theta T})$ and $\hat{g}(\theta) = E(e^{-\theta S})$, the ω must satisfy

$$\hat{f}(\omega)\hat{g}(-\omega) = 1. \quad (3.4)$$

By the convexity of the function $\psi(\theta) \equiv \hat{f}(\theta)\hat{g}(-\theta)$ and the stability condition $E(S) < E(T)$, we can indeed find a unique root $\omega > 0$ of equation (3.4). Furthermore,

$$\psi'(\omega) = E\left((S - T)e^{\omega(S-T)}\right) > 0. \quad (3.5)$$

The following arguments due to Feller [7] verify (3.2). Let $\{U_n; n \geq 1\}$ be a sequence of independent random variables that have the same distribution as $U \equiv S - T$, and let $\{Y_n; n \geq 0\}$ be the random walk generated by $\{U_n; n \geq 1\}$, i.e., $Y_n = U_1 + \dots + U_n$ with $Y_0 = 0$. We modify this random walk in the following way. Firstly change $U \equiv S - T$ to U^* subject to the distribution:

$$dP(U^* \leq x) = e^{\omega x} dP(U \leq x), \quad (3.6)$$

which is indeed a distribution because of (3.4). This distribution and the ω are said to be *twisted* and a *twisting factor*, respectively. Secondly let $\{U_n^*\}$ be a sequence of independent random variables that have the same distribution as U^* , and define $Y_n^* = U_1^* + \dots + U_n^*$ with $Y_0^* = 0$. The random walk $\{Y_n^*\}$ has a positive mean drift given by (3.5). Let H_n and H_n^* be the n -th ascending ladder height distributions of $\{Y_n; n \geq 0\}$ and $\{Y_n^*; n \geq 0\}$, respectively. Then,

$$e^{\omega(x_1 + \dots + x_n)} d_{x_1} \dots d_{x_n} P(Y_1 \leq x_1, \dots, Y_n \leq x_n) = d_{x_1} \dots d_{x_n} P(Y_1^* \leq x_1, \dots, Y_n^* \leq x_n).$$

implies

$$e^{\omega x} dH_1(x) = dH_1^*(x).$$

As is well known, W has the same distribution as $\max_{n \geq 0} Y_n$. Then, the distribution of W can be represented as

$$P(W \leq x) = \sum_{n=0}^{\infty} H_n(x)(1 - H_1(+\infty)),$$

since the random walk never goes above the present level with probability $(1 - H_1(+\infty))$. Since H_n and H_n^* are n times convolutions of H_1 and H_1^* , respectively, this yields

$$dP(W \leq y) = (1 - H_1(+\infty)) \sum_{n=0}^{\infty} e^{-\omega y} dH_n^*(y).$$

Hence, integrating over $[x, \infty)$ and multiplying $e^{\omega x}$, we have

$$e^{\omega x} P(W > x) = (1 - H_1(+\infty)) \int_0^{\infty} \sum_{n=0}^{\infty} e^{-\omega y} dH_n^*(y + x).$$

Since Y_n^* has the positive drift by (3.5), the key renewal theorem concludes

$$\lim_{x \rightarrow \infty} e^{\omega x} P(W > x) = \frac{(1 - H_1(+\infty))}{E(U^*)} \int_0^{\infty} e^{-\omega y} dy = \frac{(1 - H_1(+\infty))}{\omega E(U^*)}.$$

This proves (3.2) as claimed. Similar results are obtained for batch arrival queues in [25].

The above arguments are known to be hard to apply for other systems such as many server queues. However, using the matrix geometric approach, similar results have been

obtained for $GI/PH/s$ queues, i.e., the s server queue with renewal arrival process and phase type service time distributions. We here simply refer to the results due to Takahashi [30]. We use the same notation for random variables T and S and the Laplace-Stieltjes Transform (LST) \hat{f} and \hat{g} for the interarrival and service times. Let L , R_0 and R_1, \dots, R_s be random variables subject to the stationary joint distribution of the queue length, the remaining interarrival time and the remaining service times at servers 1 to s , respectively, where we assume the stability condition $ES < sET$. It is easy to see that there exists a unique positive root ω of the equation:

$$\hat{f}(s\omega)\hat{g}(-\omega) = 1. \quad (3.7)$$

Let $\eta = \hat{f}(s\omega)$. Then, it is obtained in [30] that, for some $b > 0$,

$$\lim_{n \rightarrow \infty} \eta^n P(L = n, R_i \leq x_i, i = 0, \dots, s) = bF_e^*(x_0)G_e^*(x_1) \dots G_e^*(x_s), \quad (3.8)$$

where F_e^* and G_e^* are nondecreasing functions such that

$$\frac{dF_e^*(x)}{dx} = \frac{\omega s}{1 - \hat{f}(s\omega)} e^{-\omega s x} P(T > x), \quad \frac{dG_e^*(x)}{dx} = \frac{\omega}{\hat{g}(-\omega) - 1} e^{\omega x} P(S > x).$$

Note that these distributions correspond with the twisted distribution (3.6) with respect to F_e and G_e , respectively, where they are defined as

$$F_e(x) = \frac{1}{ET} \int_0^x P(T > y) dy, \quad G_e(x) = \frac{1}{ES} \int_0^x P(S > y) dy.$$

Similar exponential forms are also obtained for heterogenous servers (see [2, 22, 26, 27]).

From the queues with single waiting lines, we can observe the following facts. First, the exponential decay rates is expected to exist under mild regularity conditions. Secondly, once the existence of the exponential decay rates is assured, they are identified by solving the stationary equations at off boundaries. Thirdly, the remaining arrival and service times are asymptotically independent of the queue length. The second observation may not be true for queueing networks, since the stationary distributions are multi-dimensional. Thus, we may need to choose appropriate rates among those satisfying off-boundary stationary equations. In what follows we consider this scenario.

4. Generalized Jackson Networks

In this section, we present conjectures on the asymptotic decay rates for the generalized Jackson network. We first describe this network. The network has k nodes of single servers, numbered $1, 2, \dots, k$. The outside is denoted by node 0. At each node $i \neq 0$, exogenous customers arrive according to a renewal process with interarrival time distribution F_i , which is either a proper distribution or a defective distribution such that $F_i(+\infty) = 0$. The latter represents the case that no exogenous customers arrive at node i . All nodes have infinite waiting capacity, all arriving customers are served in the FIFO manner, and their service times at node i are subject to common distribution G_i . A customer who completes service at node $i \neq 0$ immediately goes to node j , including $j = 0$, with probability r_{ij} . Hence it is assumed that

$$\sum_{j=0}^k r_{ij} = 1, \quad i = 1, \dots, k.$$

We assume that stochastic matrix $\{r_{ij}; i, j = 0, 1, \dots, k\}$ is irreducible, where r_{0i} 's are chosen in such a way that $\sum_{j=0}^k r_{0j} = 1$ with $r_{00} = 0$ and $r_{0i} > 0$ if there are exogenous arrivals at node i . For example, we can choose them such that r_{i0} is proportional to the mean arrival rate at node i . However, we only use them for the irreducible condition, and their actual values are of no importance for us. If all the interarrival time and service time distributions are exponential, then this model is the Jackson network with single servers and a single class of customers. So, we refer to it as a *generalized Jackson network*. We shall use the following notation. Let T_i and S_i denote random variables subject to F_i and G_i , respectively. Their LST's are denoted by $\hat{f}_i(\theta)$ and $\hat{g}_i(\theta)$, i.e.,

$$\hat{f}_i(\theta) = E(e^{-\theta T_i}), \quad \hat{g}_i(\theta) = E(e^{-\theta S_i}), \quad i = 1, \dots, k.$$

We assume that the network is stable, i.e., an appropriate stationary distribution exists. To give conditions for this, we need to solve the traffic equations. Let $\lambda_i = 1/E(T_i)$ if $F_i(\infty) = 1$, and $\lambda_i = 0$ otherwise. Then, the traffic equations are

$$\alpha_i = \lambda_i + \sum_{j=1}^k \alpha_j r_{ji}, \quad i = 1, \dots, k. \quad (4.1)$$

These equations have a unique nonnegative solution $(\alpha_1, \dots, \alpha_k)$ because of the irreducible assumption. It is known that the network is stable if the following conditions hold together with a mild regularity condition, called spread out, on F_i or G_i (e.g., see [18, 29]).

$$\rho_i = \alpha_i E(S_i) < 1, \quad i = 1, \dots, k. \quad (4.2)$$

We assume these stability conditions throughout the paper. Similar to the case of the $GI/G/1$ queue, we also assume the light tail assumption that, for each i and some $\theta_i^0 < 0$, which may be $-\infty$,

$$\hat{g}_i(\theta) < \infty, \text{ for } \theta > \theta_i^0 \quad \text{and} \quad \lim_{\theta \downarrow \theta_i^0} \hat{g}_i(\theta) = \infty. \quad (4.3)$$

This condition is meant that the tail probabilities of service times decay exponentially fast. More specifically, it allows the decay such as $x^\beta e^{-\alpha x}$ with $\alpha > 0$ for $\beta \geq -1$ but may not for $\beta < -1$. It includes a large class of the conventional distributions such as a phase type with finite a phase space, uniform distributions and their convolutions and mixtures.

We describe the network state at time t by the following notation. Let $X_i(t)$ be the number of customers in node i , and define the k -dimensional random vector:

$$\mathbf{X}(t) = (X_1(t), \dots, X_k(t)).$$

Let $R_i^A(t)$ and $R_i^S(t)$ be the remaining arrival and service times, respectively, and define

$$\mathbf{R}^A(t) = (R_1^A(t), \dots, R_k^A(t)), \quad \mathbf{R}^S(t) = (R_1^S(t), \dots, R_k^S(t)).$$

Then, the network state at time t is given by

$$\mathbf{Y}(t) = (\mathbf{X}(t), \mathbf{R}^A(t), \mathbf{R}^S(t)).$$

Clearly, $\{\mathbf{Y}(t)\}$ is a Markov process, and it has the stationary distribution under the stability condition (4.2). So we can assume without loss of generality that $\{\mathbf{Y}(t)\}$ is a stationary

process. In what follows we omit arguments "t" for stationary random vectors such as \mathbf{X} for $\mathbf{X}(t)$.

We next derive the balance equations in the steady state. To this end, we use the following notation.

$$\begin{aligned}\varphi(\mathbf{n}, \boldsymbol{\theta}, \boldsymbol{\eta}) &= E \left(e^{-\langle \boldsymbol{\theta}, \mathbf{R}^A \rangle - \langle \boldsymbol{\eta}, \mathbf{R}^S \rangle}; \mathbf{X} \geq \mathbf{n} \right), \\ \varphi_i^A(\mathbf{n}, \boldsymbol{\theta}, \boldsymbol{\eta}) &= E_i^A \left(e^{-\langle \boldsymbol{\theta}, \mathbf{R}^{A-} \rangle - \langle \boldsymbol{\eta}, \mathbf{R}^{S-} \rangle}; \mathbf{X}^- \geq \mathbf{n} \right), \\ \varphi_i^D(\mathbf{n}, \boldsymbol{\theta}, \boldsymbol{\eta}) &= E_i^D \left(e^{-\langle \boldsymbol{\theta}, \mathbf{R}^{A-} \rangle - \langle \boldsymbol{\eta}, \mathbf{R}^{S-} \rangle}; \mathbf{X}^- \geq \mathbf{n} \right),\end{aligned}$$

where E_i^A and E_i^D stand for the expectations concerning Palm distributions with respect to the arrival and departure point processes at node i , respectively, and \mathbf{R}^{A-} , \mathbf{R}^{S-} and \mathbf{X}^- are \mathbf{R}^A , \mathbf{R}^S and \mathbf{X} , respectively, just before the arrival or departure instants. Then, applying the rate conservation law (c.f. [19]), we have

$$\begin{aligned}\left(\sum_{i=1}^k (\theta_i 1(\lambda_i \neq 0) + \eta_i) \right) \varphi(\mathbf{n}, \boldsymbol{\theta}, \boldsymbol{\eta}) &= \sum_{i=1}^k \lambda_i \left(\varphi_i^A(\mathbf{n}, \boldsymbol{\theta}, \boldsymbol{\eta}) - \varphi_i^A(\mathbf{n} - \mathbf{e}_i, \boldsymbol{\theta}, \boldsymbol{\eta}) \hat{f}_i(\theta_i) \right) \\ &\quad + \sum_{i=1}^k \sum_{j=0}^k 1(i \neq j) \alpha_i \left(\varphi_i^D(\mathbf{n}, \boldsymbol{\theta}, \boldsymbol{\eta}) - \varphi_i^D(\mathbf{n} + \mathbf{e}_i - \mathbf{e}_j, \boldsymbol{\theta}, \boldsymbol{\eta}) \hat{g}_i(\eta_i) \right) r_{ij} \\ &\quad + \sum_{i=1}^k \alpha_i \left(\varphi_i^D(\mathbf{n}, \boldsymbol{\theta}, \boldsymbol{\eta}) - \varphi_i^D(\mathbf{n}, \boldsymbol{\theta}, \boldsymbol{\eta}) \hat{g}_i(\eta_i) \right) r_{ii}, \quad \mathbf{n} > \mathbf{0},\end{aligned}\tag{4.4}$$

where \mathbf{e}_i is the unit vector whose i -th entry is 1. This equation can be considered as the balance equation that the rate of continuous state changes are equated with those of jump transitions due to arrivals and departures.

We tentatively assume that there exists $\boldsymbol{\tau} > \mathbf{0}$ such that, for large \mathbf{n} ,

$$\varphi(\mathbf{n}, \boldsymbol{\theta}, \boldsymbol{\eta}) \sim \tilde{\varphi}(\boldsymbol{\theta}, \boldsymbol{\eta}) e^{-\langle \boldsymbol{\tau}, \mathbf{n} \rangle}, \quad i = 1, 2, \dots, k,\tag{4.5}$$

where $\tilde{\varphi}$ is a joint LST of a $2k$ -dimensional distribution. This is similar to (3.8). However, we do not assume the asymptotic independence of R_i^A and R_i^D , but the key issue is the asymptotic independence of the queue length and these characteristics. Since $\boldsymbol{\tau}$ may depend on the direction of increasing \mathbf{n} , we make (4.5) precise in the following way. For each direction vector \mathbf{c} , there exist positive vector $\boldsymbol{\tau}$, $b(\mathbf{c})$ and LST $\tilde{\varphi}(\boldsymbol{\theta}, \boldsymbol{\eta})$ of a $2k$ -dimensional distribution such that

$$\lim_{\mathbf{n}/|\mathbf{n}| \rightarrow \mathbf{c}, |\mathbf{n}| \rightarrow \infty} \varphi(\mathbf{n}, \boldsymbol{\theta}, \boldsymbol{\eta}) e^{\langle \boldsymbol{\tau}, \mathbf{n} \rangle} = b(\mathbf{c}) \tilde{\varphi}(\boldsymbol{\theta}, \boldsymbol{\eta}),\tag{4.6}$$

where $\boldsymbol{\tau}$ may depend on \mathbf{c} . Let $\omega = \langle \boldsymbol{\tau}, \mathbf{c} \rangle$, then $\langle \boldsymbol{\tau}, \mathbf{n} \rangle \sim \omega |\mathbf{n}|$ for large $|\mathbf{n}|$ such that $\mathbf{n} \sim \mathbf{c} |\mathbf{n}|$. Hence, (4.6) implies that \mathbf{X} has the decay rate ω for direction \mathbf{c} (see Definition 2.1). In our arguments, $\boldsymbol{\tau}$ will have an important role, and it will be clear the reason why we use vector $\boldsymbol{\tau}$ rather than scalar ω .

For our arguments, we also need similar asymptotics for φ_i^A and φ_i^D , i.e., the embedded distributions at arrivals and departures. We note the following fact, which is proved in Appendix A.

Lemma 4.1 The decay rates at departure and exogenous arrival instants at node i are identical with the decay rate at an arbitrary point in time if the distributions of T_i and S_i have failure rates that are bounded above and away from 0 in the extended sense, which is meant for S_i that, for each bounded continuous and positive function v , there exist positive constants a, b such that, for any $y \geq 0$,

$$a \int_y^\infty v(x)(1 - G_i(x))dx \leq \int_y^\infty v(x)dG_i(x) \leq b \int_y^\infty v(x)(1 - G_i(x))dx. \quad (4.7)$$

For example, phase type distributions with finite phase spaces and uniform distributions satisfy condition (4.7). This lemma suggests that the common decay rate can be expected for φ_i^A and φ_i^D . So, we conjecture

Assumption 4.1 For each direction vector $\mathbf{c} > \mathbf{0}$, there are nonnegative vector $\boldsymbol{\tau}$, constants $b_i^u(\mathbf{c})$, $u = A, D$, and a function $\tilde{\varphi}_i^u(\boldsymbol{\theta}, \boldsymbol{\eta})$ such that

$$\lim_{\mathbf{n}/|\mathbf{n}| \rightarrow \mathbf{c}, |\mathbf{n}| \rightarrow \infty} \varphi_i^u(\mathbf{n}, \boldsymbol{\theta}, \boldsymbol{\eta}) e^{\langle \boldsymbol{\tau}, \mathbf{n} \rangle} = b_i^u(\mathbf{c}) \tilde{\varphi}_i^u(\boldsymbol{\theta}, \boldsymbol{\eta}), \quad u = A, D, \quad i = 1, 2, \dots, k. \quad (4.8)$$

Remark 4.1 (i) Similar to (4.6), (4.8) implies that the decay rate of \mathbf{X} is $\langle \boldsymbol{\tau}, \mathbf{c} \rangle$ for direction \mathbf{c} . Note that we here restrict \mathbf{c} to be positive.

(ii) The $\boldsymbol{\tau}$ in (4.8) is independent of i but may depend on \mathbf{c} , so it may be better to write it as $\boldsymbol{\tau}(\mathbf{c})$. However, the latter notation is rather inconvenient for our arguments, so we keep $\boldsymbol{\tau}$ as it is.

(iii) Obviously, $\boldsymbol{\tau}$ can be represented by \mathbf{t} whose i -th entry is $t_i = e^{-\tau_i}$. Sometimes the latter is convenient for tail decay rates, but we work on $\boldsymbol{\tau}$ to prevent possible confusions.

Note that Assumption 4.1 implies (4.6). This is verified using (4.4). Multiply both sides of it with $e^{\langle \boldsymbol{\tau}, \mathbf{n} \rangle}$ and let $|\mathbf{n}|$ go to infinity in such a way that $\mathbf{n}/|\mathbf{n}|$ goes to \mathbf{c} , then the limit in (4.6) is uniquely determined by Assumption 4.1. It is not hard to see that this limit can not be identically zero, by choosing sufficiently large η_i 's. We shall use Assumption 4.1 as a starting assumption. This is the reason that we do not term it as a conjecture but as an assumption, although it is nothing but a conjecture.

Let us consider implications of Assumption 4.1. We first let

$$\sum_{j=1}^k (\theta_j 1(\lambda_j \neq 0) + \eta_j) = 0. \quad (4.9)$$

Then, multiplying both sides of (4.4) with $e^{\langle \boldsymbol{\tau}, \mathbf{n} \rangle}$ and applying (4.8), we have

$$\begin{aligned} \sum_{i=1}^k \lambda_i b_i^A(\mathbf{c}) \varphi_i^A(\boldsymbol{\theta}, \boldsymbol{\eta}) (1 - e^{\tau_i} \hat{f}_i(\theta_i)) \\ + \sum_{i=1}^k \alpha_i b_i^D(\mathbf{c}) \varphi_i^D(\boldsymbol{\theta}, \boldsymbol{\eta}) \left(1 - \hat{g}_i(\eta_i) e^{-\tau_i} \sum_{j=0}^k r_{ij} e^{\tau_j} \right) = 0, \end{aligned} \quad (4.10)$$

where $\tau_0 = 0$. We choose θ_i such that

$$1 = e^{\tau_i} \hat{f}_i(\theta_i), \quad \lambda_i > 0, \quad i = 1, \dots, k, \quad (4.11)$$

then the first summation of (4.10) is dropped. Because of the restriction (4.9), we next choose η_i except one i , say i_0 , such that

$$e^{\tau_{i_0}} = \hat{g}_{i_0}(\eta_{i_0}) \left(r_{i_0 0} + \sum_{j=1}^k r_{i_0 j} e^{\tau_j} \right), \quad i = 1, \dots, k, \quad (4.12)$$

but (4.10) implies that (4.12) has to hold also for $i = i_0$. Thus, the conditions (4.9), (4.11) and (4.12) are necessary if the asymptotic property (4.8) holds true. We define

$$\mathcal{L}_0 = \{\boldsymbol{\tau} = (\tau_1, \dots, \tau_k) \geq \mathbf{0}; \text{ (4.9), (4.11) and (4.12) hold}\}.$$

Remark 4.2 If $k = 1$, then \mathcal{L}_0 is a singleton, so τ_1 is uniquely determined. This τ_1 agrees with the well known decay rate for the $GI/G/1$ queue. \square

Lemma 4.2 Let $\bar{\mathcal{L}}_0$ be the set of all $\boldsymbol{\tau} \geq \mathbf{0}$ which satisfy, for some $\eta_1, \dots, \eta_k \in \mathbb{R}$,

$$\sum_{j=1}^k (\hat{f}_j^{-1}(e^{-\tau_j})1(\lambda_j \neq 0) + \eta_j) \leq 0, \quad (4.13)$$

$$\hat{g}_i(\eta_i)e^{-\tau_i} \left(r_{i0} + \sum_{j=1}^k r_{ij}e^{\tau_j} \right) \leq 1, \quad i \in J \equiv \{1, 2, \dots, k\}, \quad (4.14)$$

where $\hat{f}_i^{-1}(t) = \inf\{\theta; \hat{f}_i(\theta) > t\}$, i.e., the inverse function of \hat{f}_i , for i such that $\lambda_i > 0$. Then, $\bar{\mathcal{L}}_0$ is convex, and includes \mathcal{L}_0 on its boundary. Furthermore, if all $\eta_i = \inf\{\eta \in \mathbb{R}; \hat{g}_i(\eta) < \infty\}$ are finite, then $\bar{\mathcal{L}}_0$ is bounded.

PROOF. Since $\hat{f}_i(\theta)$ is a decreasing convex function on the region that $\hat{f}_i(\theta) < \infty$, $\hat{f}_i^{-1}(t)$ is also a decreasing and convex for $t > 0$, and (4.11) is equivalent to

$$\theta_i = \hat{f}_i^{-1}(e^{-\tau_i}), \quad \lambda_i > 0, \quad i = 1, \dots, k,$$

Hence, (4.13) and (4.14) with equalities are equivalent to (4.9), (4.11) and (4.12) concerning $\boldsymbol{\eta}$ and $\boldsymbol{\tau}$, where $\boldsymbol{\eta} = (\eta_1, \dots, \eta_k)$. This also implies that \mathcal{L}_0 is on the boundary of $\bar{\mathcal{L}}_0$. Since the left-hand sides of (4.13) and (4.14) are convex functions of $\boldsymbol{\eta}$ and $\boldsymbol{\tau}$, (4.13) and (4.14) constitute convex sets in \mathbb{R}^{2k} , so $\bar{\mathcal{L}}_0$ is a convex set in \mathbb{R}^k . The remaining part is proved in the following way. Since all η_i is bounded below and $\hat{f}_j^{-1}(e^{-\tau_j})$ is increasing in τ_j , (4.13) imply that τ_i 's are bounded if $\lambda_i > 0$. Furthermore, (4.13) simultaneously implies that η_i are all bounded above, so $\hat{g}_i(\eta_i)$ are all bounded below. Let $J_0 = \{i \in J; \lambda_i > 0\}$, and $J_n = \{j \in J \setminus J_{n-1}; r_{ij} > 0, \exists i \in J_{n-1}\}$ for $n = 1, 2, \dots$. If $j \in J_1$, then (4.14) with $i \in J_0$ satisfying $r_{ij} > 0$ implies that τ_j is bounded because $\hat{g}_i(\eta_i)$ is bounded below. Similarly, we can inductively show that τ_i is bounded for $i \in J_n$ with $n \geq 2$. Since $\cup_{n=0}^{k-1} J_n = J$ by the irreducibility of $\{r_{ij}\}$, we see that $\bar{\mathcal{L}}_0$ is bounded. \square

Note that \mathcal{L}_0 is the set of the $\boldsymbol{\tau}$'s which may be the decay rates for appropriate directions. However, all the $\boldsymbol{\tau}$'s may not be so. Hence, we need to exclude infeasible $\boldsymbol{\tau}$'s. To this end, we consider the decay rates for the marginal distributions.

Let us consider the case that one entry of \mathbf{n} , say n_i , is increased. In this case, the decay rate would be compensated by the twisted service time distribution of node i with the twisting factor $-\eta_i > 0$, while the other nodes behave normally, i.e., $\eta_j > 0$ for $j \neq i$. Thus, we define

$$\tau_i^0 = \sup\{\tau_i \geq 0; \boldsymbol{\tau} \in \mathcal{L}_0 \text{ and } \eta_j \geq 0 \text{ for all } j \neq i\}, \quad (4.15)$$

where the equality for η_j is included so as to always attain the supremum. One may wonder whether this τ_i^0 gives the decay rate for the marginal distribution at node i . Unfortunately, this is not true in general, since other τ_j^0 may further restrict \mathcal{L}_0 as we shall see in Example 5.2 (see case (b)). Thus, our first conjecture is

Conjecture 4.1 The decay rate for the marginal queue length distribution at node i is bounded by τ_i^0 .

Remark 4.3 Consider the case that $k = 2$, $\lambda_2 = 0$ and $r_{12} = r_{20} = 1$, i.e., the two node tandem queue. Suppose that $S_1 \equiv 0$, so arriving customers at node 1 immediately go to node 2. This model reduces to the single node queue, and the decay rate is known as mentioned in Section 3. We here consider whether conditions (4.9), (4.11) and (4.12) are compatible with this known result. It is easy to see that they are equivalent to $\tau_1 = \tau_2 = \log(\hat{g}_2(\eta_2))$ and

$$\hat{f}_1(-\eta_1 - \eta_2)\hat{g}_2(\eta_2) = 1. \quad (4.16)$$

Hence, $\tau_2 = \log(\hat{g}_2(\eta_2))$ can be arbitrarily large by choosing sufficiently small η_1 , so the result is not compatible with the single node case. On the other hand, if η_1 is restricted to $\eta_1 \geq 0$, then τ_2 is maximized under (4.16) when $\eta_1 = 0$. This is indeed compatible with the single node case. Thus, the restriction by τ_i^0 in Conjecture 4.1 is really important. This example also shows why the finiteness of η_i^0 is assumed for the boundedness of $\bar{\mathcal{L}}_0$ in Lemma 4.2 since $\eta_1^0 = -\infty$ and $\bar{\mathcal{L}}_0$ is unbounded in this case.

We next consider a uniform bound for the tail probability in such a way that, for some positive C and some $\boldsymbol{\tau} \in \mathcal{L}_0$,

$$P(\mathbf{X} \geq \mathbf{n}) \leq Ce^{-\langle \boldsymbol{\tau}, \mathbf{n} \rangle}, \quad \mathbf{n} \geq \mathbf{0}. \quad (4.17)$$

In general, such a $\boldsymbol{\tau}$ may not serve a tight bound. However, if we can find the $\boldsymbol{\tau}$ satisfying (4.17) such that $\langle \boldsymbol{\tau}, \mathbf{c} \rangle$ is maximized on the convex set \mathcal{L}_0 , then the $\boldsymbol{\tau}$ must be the decay rate for the direction \mathbf{c} . As we shall see later, this is not a rare case.

Let us consider a possible class of $\boldsymbol{\tau}$ for (4.17). Since the τ_i can not be greater than the decay rate of the marginal distribution of node i , we restrict \mathcal{L}_0 to

$$\mathcal{L}_{\{i\}} = \{\boldsymbol{\tau} \in \mathcal{L}_0; \tau_i \leq \tau_i^0\},$$

and to $\cap_{i=1}^k \mathcal{L}_{\{i\}}$. We make further restrictions by higher dimensional marginals. For each subset $B \subset J$, let $\boldsymbol{\tau}(B)$ be the $|B|$ -dimensional vector whose i -th entry is τ_i for $i \in B$. Then, let

$$\mathcal{T}(B) = \{\boldsymbol{\tau}(B) \geq \mathbf{0}(B); (4.9), (4.11) \text{ and } (4.12) \text{ hold for } \eta_j \geq 0, \forall j \in J \setminus B\},$$

where $\mathbf{0}(B)$ is the $|B|$ -dimensional zero vector. Note that $\mathcal{T}(J) = \mathcal{L}_0$. We define

$$\mathcal{L}_B = \{\boldsymbol{\tau} \in \mathcal{L}_0; \boldsymbol{\tau}(B) \leq \boldsymbol{\tau}'(B), \exists \boldsymbol{\tau}'(B) \in \mathcal{T}(B)\}.$$

Clearly this definition is compatible with $\mathcal{L}_{\{i\}}$ for $B = \{i\}$. Then, we let

$$\mathcal{L}_J^* = \cap_{B \subset J} \mathcal{L}_B.$$

Intuitively, this restriction for $\boldsymbol{\tau}$ is necessary to get the product form bound (4.17), since the $\boldsymbol{\tau}$ should bound all marginal distributions. As we shall see in Section 6, this type of restrictions also appears in the batch movement network, and is known to be necessary and sufficient (see (i) of Remark 6.1). This suggests us to conjecture

Conjecture 4.2 For each $\boldsymbol{\tau} \in \mathcal{L}_J^*$, there exists positive constant C for (4.17). In particular, $\tau_i^1 \equiv \sup\{\tau_i; \boldsymbol{\tau} \in \mathcal{L}_J^*\}$ is the decay rate of the marginal distribution.

Remark 4.4 Makimoto [17] also suggested a similar one to Conjecture 4.2. For $k = 2$, Katou and Makimoto [12] (see also [13]) consider a slightly different version of Conjecture 4.2 for a phase type model, in which interarrival and service times have phase type distributions. In our notation, they introduce

$$\mathcal{F}_1 = \{\boldsymbol{\tau} \in \mathcal{L}_0; \eta_2 > 0, \tau_2 \leq \tau_2^0\}, \quad \mathcal{F}_2 = \{\boldsymbol{\tau} \in \mathcal{L}_0; \eta_1 > 0, \tau_1 \leq \tau_1^0\},$$

and show that $\boldsymbol{\tau} \in \mathcal{F}_1 \cup \mathcal{F}_2$ implies (4.17). In the recent paper [14], it is conjectured that $\bar{\tau}_i \equiv \sup\{\tau_i; \boldsymbol{\tau} \in \mathcal{F}_i\}$ is the decay rate of the marginal distribution at node i . It is not difficult to see from (4.14) that $\mathcal{F}_1 \cup \mathcal{F}_2$ is identical with \mathcal{L}_J^* for $k = 2$. Hence, $\bar{\tau}_i = \tau_i^1$, and Conjecture 4.2 is compatible with the conjecture of [14].

As we mentioned, Conjecture 4.2 only gives product form upper bounds, and it may not include the best $\boldsymbol{\tau}$ that gives exact decay rates. Recently Katou and Takahashi [14] found a very interesting bound for the tail probability for a 2-node generalized Jackson network with phase-type distributions, i.e., for $k = 2$. They show that, if $\boldsymbol{\tau}_1 \in \mathcal{L}_{\{1\}}$ and $\boldsymbol{\tau}_2 \in \mathcal{L}_{\{2\}}$, there exist positive constants C_1 and C_2 such that

$$P(\mathbf{X} \geq \mathbf{n}) \leq C_1 e^{-\langle \boldsymbol{\tau}_1, \mathbf{n} \rangle} + C_2 e^{-\langle \boldsymbol{\tau}_2, \mathbf{n} \rangle}, \quad \mathbf{n} \geq \mathbf{0}. \quad (4.18)$$

This would suggest the next conjecture.

Conjecture 4.3 Let $\mathcal{L}_B^* = \cap_{B' \subset B} \mathcal{L}_{B'}$ for $B \subset J$. Then, for each partition \mathcal{A} of J and each $\boldsymbol{\tau}_B \in \mathcal{L}_B^*$ for $B \in \mathcal{A}$, there exist positive constant C_B 's such that

$$P(\mathbf{X} \geq \mathbf{n}) \leq \sum_{B \in \mathcal{A}} C_B e^{-\langle \boldsymbol{\tau}_B, \mathbf{n} \rangle}, \quad \mathbf{n} \geq \mathbf{0}, \quad (4.19)$$

Remark 4.5 For $\mathcal{A} = \{J\}$, this conjecture reduces to Conjecture 4.2, while, for $k = 2$ and $\mathcal{A} = \{\{1\}, \{2\}\}$, it reduces to (4.18).

Suppose that we have listed up all possible bounds and the corresponding sets of $\boldsymbol{\tau}$. Since the exact decay rate should be the minimum among all upper bounds, we arrive at

Conjecture 4.4 For each direction \mathbf{c} , the decay rate, denoted by $\omega(\mathbf{c})$, is obtained as

$$\omega(\mathbf{c}) = \max_{\mathcal{A} \in \mathcal{P}(J)} \min_{B \in \mathcal{A}} \sup\{\langle \boldsymbol{\tau}, \mathbf{c} \rangle; \boldsymbol{\tau} \in \mathcal{L}_B^*\}, \quad (4.20)$$

where $\mathcal{P}(J)$ is the set of all partitions of J .

To find the decay rate $\omega(\mathbf{c})$ in Conjecture 4.4, a key step is to find $\omega_B(\mathbf{c})$ such that

$$\omega_B(\mathbf{c}) = \sup\{\langle \boldsymbol{\tau}, \mathbf{c} \rangle | \boldsymbol{\tau} \in \mathcal{L}_B^*\}. \quad (4.21)$$

Let us introduce some convenient notation to find this $\omega_B(\mathbf{c})$. A subset V of \mathbb{R}^k is said to be a positive lower set if

$$\mathbf{x} \in V, \mathbf{0} \leq \mathbf{y} \leq \mathbf{x} \quad \text{implies} \quad \mathbf{y} \in V.$$

Let $\underline{M}(\mathcal{L})$ be the minimum lower set that includes a set \mathcal{L} , which always exists since a common set of positive lower sets are also a positive lower set. Let

$$D(\mathcal{L}) = \partial(\underline{M}(\mathcal{L})) \cap \mathcal{L},$$

where $\partial(\underline{M}(\mathcal{L}))$ is the boundary of $\underline{M}(\mathcal{L})$. Then, we clearly have

$$\omega_B(\mathbf{c}) = \sup\{\langle \boldsymbol{\tau}, \mathbf{c} \rangle | \boldsymbol{\tau} \in D(\mathcal{L}_B^*)\}. \quad (4.22)$$

Since $D(\mathcal{L}_B^*)$ is included in the boundary of the convex set $\bar{\mathcal{L}}_0$ and $z = \langle \boldsymbol{\tau}, \mathbf{c} \rangle$ is a line with respect to variable $\boldsymbol{\tau}$ for each fixed z , the z can be maximized when the line, $z = \langle \boldsymbol{\tau}, \mathbf{c} \rangle$, contacts the curve $D(\mathcal{L}_B^*)$, i.e., becomes a tangent. Thus, there is a unique $\boldsymbol{\tau}$ that attains $\omega_B(\mathbf{c})$ in (4.22) (hence in (4.21)). Figure 1 illustrates this computation.

$$e^{\tau_i^{(i)}} - 1 = \sum_{j=1}^k r_{ij}(e^{\tau_j^{(i)}} - 1) + \sum_{j=1}^k \frac{\lambda_j}{\mu_i}(e^{\tau_j^{(i)}} - 1)e^{\tau_i^{(i)}}. \quad (5.5)$$

Let \mathbf{u} be the column vector whose j -th entry is $e^{\tau_j^{(i)}} - 1$, $\boldsymbol{\lambda}$ be the row vector whose j -th entry is λ_j and \mathbf{e}_j be the unit vector whose j -th entry is unit and the other entries are zero. Let R be the $k \times k$ matrix whose ij entry is r_{ij} . Then, equations (5.4) and (5.5) can be written as

$$(I - R)\mathbf{u} = \frac{e^{\tau_i^{(i)}}}{\mu_i} \langle \boldsymbol{\lambda}, \mathbf{u} \rangle \mathbf{e}_i.$$

Since R is substochastic, $I - R$ is invertible,

$$\mathbf{u} = \frac{e^{\tau_i^{(i)}}}{\mu_i} \langle \boldsymbol{\lambda}, \mathbf{u} \rangle (I - R)^{-1} \mathbf{e}_i. \quad (5.6)$$

Let $\boldsymbol{\alpha}$ be the row vector whose j -th entry is the mean arrival rate α_j for $j = 1, 2, \dots, k$. Since $\boldsymbol{\alpha} = \boldsymbol{\lambda}(I - R)^{-1}$ and $\langle \boldsymbol{\lambda}, \mathbf{u} \rangle \neq 0$, multiplying both sides of (5.6) with $\boldsymbol{\lambda}$ from the left yields

$$e^{-\tau_i^{(i)}} = \frac{\alpha_i}{\mu_i}. \quad (5.7)$$

Comparing the i -th components of (5.6), we also have

$$\langle \boldsymbol{\lambda}, \mathbf{u} \rangle = \frac{(\mu_i - \alpha_i)}{n_{ii}}, \quad (5.8)$$

where n_{ij} is the ij -entry of $(I - R)^{-1}$, which is the mean visiting number at node j when a customer departs at node i . If $i = j$, then it counts the departing node $j = i$ as well. From (5.4) and (5.7), \mathbf{u} and hence $\boldsymbol{\tau}^{(i)}$ is uniquely determined. In particular, all the components of $\boldsymbol{\tau}^{(i)}$ are positive.

For each $i \in J$, let

$$\mathcal{M}_i = \{\boldsymbol{\tau} \in \mathcal{L}_0; \eta_\ell \geq 0, \text{ for all } \ell \neq i\}.$$

We next show that $\tau_\ell^{(i)} = \overline{m}_\ell^{(i)} \equiv \sup\{\tau_\ell; \boldsymbol{\tau} \in \mathcal{M}_i\}$, for $\ell \neq i$, which we call that $\boldsymbol{\tau}^{(i)}$ is the right-end point of \mathcal{M}_i . On the contrary, $\boldsymbol{\tau}^{(i)}$ is the left-end point of \mathcal{M}_i if $\tau_\ell^{(i)} = \underline{m}_\ell^{(i)} \equiv \inf\{\tau_\ell; \boldsymbol{\tau} \in \mathcal{M}_i\}$, for $\ell \neq i$. Let \mathcal{T}_i be the set of all $\boldsymbol{\tau} \geq \mathbf{0}$ such that

$$e^{\tau_\ell} \leq r_{\ell 0} + \sum_{j=1}^k r_{\ell j} e^{\tau_j}, \quad \ell \neq i. \quad (5.9)$$

In the view of (5.2), the inequalities (5.9) are equivalent to $\eta_\ell \geq 0$, for all $\ell \neq i$. Let $R^{(i)}$ be the submatrix of R that the i -th row and column are deleted. Since $(I^{(i)} - R^{(i)})^{-1}$ is a nonnegative matrix, where $I^{(i)}$ is the $k - 1$ -dimensional identity matrix, (5.9) can be rewritten as

$$e^{\tau_\ell} \leq \sum_{j \neq i} n_{\ell j}^{(i)} r_{j0} + e^{\tau_i} \sum_{j \neq i} n_{\ell j}^{(i)} r_{ji}, \quad \ell \neq i,$$

where $n_{\ell j}^{(i)}$ is the ℓj entry of $(I^{(i)} - R^{(i)})^{-1}$. Hence, $\sup\{\tau_\ell; \tau \in \mathcal{M}_i\}$ is attained on the boundary of \mathcal{T}_i . Thus, $\tau^{(i)}$ is either the right-end or the left-end point. We now let $\bar{\tau}^{(i)}$ be the τ vector such that its i -th entry is $-\log(\lambda_i/\mu_i)$ and the other entries are all zero. Then, it is easy to see that $\bar{\tau}^{(i)} \in \mathcal{L}_0$. Furthermore, from (5.2) with $\tau = \bar{\tau}^{(i)}$,

$$\eta_\ell = \mu_\ell r_{\ell i} \left(\frac{\mu_i}{\lambda_i} - 1 \right) > 0, \quad \ell \neq i.$$

Hence, $\bar{\tau}^{(i)}$ is in \mathcal{M}_i . Obviously, it must be the left-end point, since all its entries except the i -th entry are zero (see Figure 2). Hence, $\tau^{(i)}$ must be the right-end point. Define τ^* as

$$\tau^* \equiv (\tau_1^{(1)}, \tau_2^{(2)}, \dots, \tau_k^{(k)}).$$

It is easy to check that this τ^* satisfies (5.3), so it is on the boundary \mathcal{L}_0 . We shall show that

$$D(\mathcal{L}_J^*) = \{\tau^*\}. \quad (5.10)$$

First consider the case that $\tau_i^{(i)} = \bar{m}_i^{(i)} \equiv \sup\{\tau_i; \tau \in \mathcal{M}_i\}$ for all $i \in J$. In this case, τ^* clearly bounds all $\tau \in \mathcal{L}_J^*$. This typically happens (see Figure 2 for the case that $k = 2$). We next consider the case that there is an $i \in J$ such that $\tau_i^{(i)} \neq \bar{m}_i^{(i)}$. In this case, we have

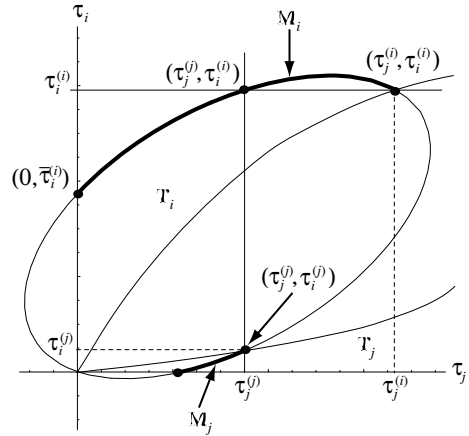
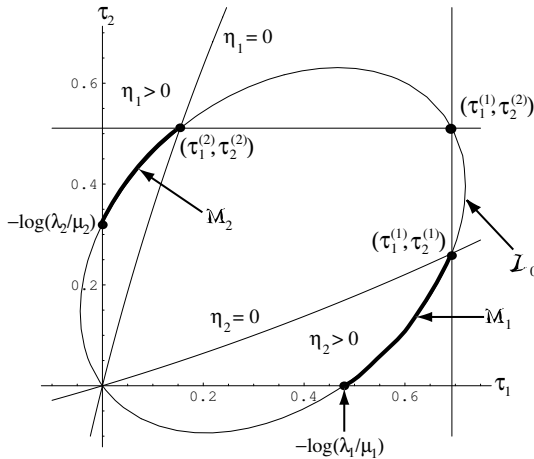


Figure 2: τ for Jackson network ($k = 2$) Figure 3: The case of $\tau_j^{(j)} < \tau_j^{(i)}$ on τ_j - τ_i plane

$\tau^* \leq \tau^{(i)}$ and $\tau^* \neq \tau^{(i)}$ since $\tau^{(i)}$ is the right-end point of \mathcal{M}_i and both of $\tau^{(i)}$ and τ^* are on the boundary \mathcal{L}_0 of the convex set $\bar{\mathcal{L}}_0$ (see Figure 3). Hence, $\tau_j^{(j)} = -\log(\alpha_j/\mu_j) < \tau_j^{(i)}$ for $j \neq i$. Then, on the τ_j - τ_i plane, we can see that three points $(0, \bar{\tau}_i^{(i)})$, $(\tau_j^{(j)}, \tau_i^{(i)})$ and $(\tau_j^{(i)}, \tau_i^{(i)})$ are in the projection of \mathcal{M}_i , and that

$$(0, \bar{\tau}_i^{(i)}) < (\tau_j^{(j)}, \tau_i^{(i)}) \leq (\tau_j^{(i)}, \tau_i^{(i)}).$$

See Figure 3 for these points. This obviously implies that

$$\tau_i^{(i)} = \sup\{\tau_i; \tau \in \mathcal{M}_i, \tau_j \leq \tau_j^{(j)}\}.$$

Combing those two cases yields that $\cap_{i \in J} \mathcal{L}_{\{i\}} = \{\tau^*\}$. Since $\tau^* \in \mathcal{L}_B$ for any $B \subset J$, we have (5.10).

From (5.10), Conjecture 4.2 implies the product form bound with τ^* , which is fully compatible with the product form solution of the Jackson network. Furthermore, for any partition \mathcal{A} of J , it can be shown that $\min_{B \in \mathcal{A}} \sup\{\langle \tau, \mathbf{c} \rangle; \tau \in \mathcal{L}_B^*\}$ is attained by τ^* . Hence, $\omega(\mathbf{c})$ of (4.20) is given by

$$\omega(\mathbf{c}) = \langle \tau^*, \mathbf{c} \rangle.$$

This is also compatible with the product form solution of the Jackson network. \square

Example 5.2 (2-node tandem queue) We next consider a two-node tandem queue. For a general two node network, which we also discuss in the next example, Katou and Takahashi [14] gave an interesting classification. They introduced

$$\hat{\tau}_1 = \sup\{\tau_1; \tau \in \mathcal{L}_0; \tau_2 = \bar{\tau}_2\}, \quad \hat{\tau}_2 = \sup\{\tau_2; \tau \in \mathcal{L}_0; \tau_1 = \bar{\tau}_1\},$$

and classified the set of τ that gives the decay rate ω of (4.20) into three types according to three possible cases: (1) $\hat{\tau}_1 = \bar{\tau}_1$ or $\hat{\tau}_2 = \bar{\tau}_2$, (2) $\hat{\tau}_1 < \bar{\tau}_1$ and (3) $\hat{\tau}_1 > \bar{\tau}_1$. We here consider another type of classifications, which are directly related to modeling parameters.

We describe a two node tandem queue in the following way. Node 1 has exogenous arrivals subject to the renewal process with a generic inter-arrival time distribution F_1 with LST \hat{f}_1 and mean $1/\lambda_1$, and the service time distributions G_1 and G_2 with LST \hat{g}_1 and \hat{g}_2 and means $1/\mu_1$ and $1/\mu_2$, respectively. Obviously, the stability condition is $\lambda_1 < \min(\mu_1, \mu_2)$, which is assumed here.

We first consider the simpler case that both servers have exponentially distributed service times. This case has been studied in [10] for the decay rates of the marginal distributions, so it is convenient to see how our conjectures, in particular, Conjecture 4.1, work. Furthermore, it turns out that this simpler model is essential to study the case of generally distributed service times. Similar to Example 5.1, condition (4.12) becomes

$$\eta_1 = \mu_1(e^{-\tau_1 + \tau_2} - 1), \quad \eta_2 = \mu_2(e^{-\tau_2} - 1),$$

while (4.9) and (4.11) yield

$$\hat{f}_1(\mu_1(1 - e^{-\tau_1 + \tau_2}) + \mu_2(1 - e^{-\tau_2})) = e^{-\tau_1}. \quad (5.11)$$

Let γ_1 (γ_2) be the positive solution τ_1 (τ_2) of these equations when $\eta_2 = 0$ ($\eta_1 = 0$, respectively). By the stability condition, these γ_1 and γ_2 are uniquely determined as the solution of the equations

$$\hat{f}_1(\mu_1(1 - e^{-\gamma_1})) = e^{-\gamma_1}, \quad \hat{f}_1(\mu_2(1 - e^{-\gamma_2})) = e^{-\gamma_2}.$$

Since \hat{f}_1 is a decreasing function, it is easy to see that $\gamma_1 \leq \gamma_2$ if and only if $\mu_1 \leq \mu_2$. We next let (u_1, u_2) be the contact point of (5.11) with the horizontal line $\tau_2 = u_2 > 0$. This point can be obtained as the solution (τ_1, τ_2) of (5.11) and the contact condition:

$$-\mu_1 \hat{f}_1'(\mu_1(1 - e^{-\tau_1 + \tau_2}) + \mu_2(1 - e^{-\tau_2})) = e^{-\tau_2}. \quad (5.12)$$

Similarly, let (v_1, v_2) be the contact point of (5.11) with the vertical line $\tau_1 = v_1 > 0$. Let σ_1 be a solution τ_1 of (5.11) with $\tau_2 = \gamma_2$ such that $\tau_1 \neq \gamma_2$, and let σ_2 be a positive solution of τ_2 of (5.11) with $\tau_1 = \gamma_1$ such that $\tau_2 \neq \gamma_1$. These solutions uniquely exist. Then, there arise five different cases, where $J = \{1, 2\}$.

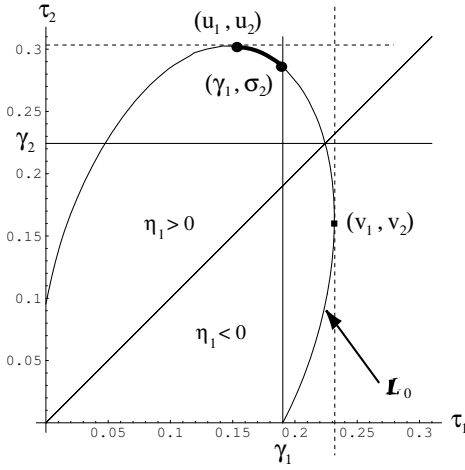


Figure 4: Tandem case (a)

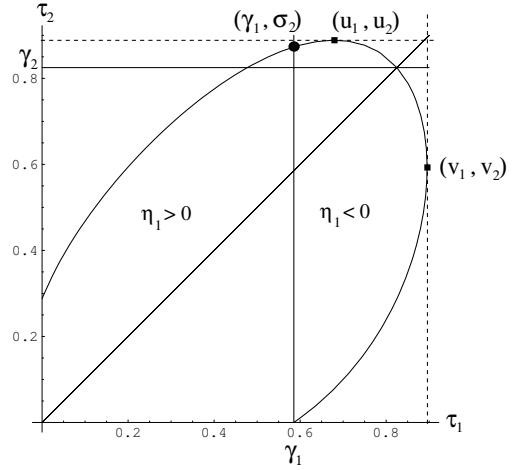


Figure 5: Tandem case (b)

- (a) $\gamma_1 \leq \gamma_2$ and $u_1 \leq \gamma_1$: $\tau_2^1 = u_2$, $D(\mathcal{L}_J^*) = \{\tau \in \mathcal{L}_0; \tau_1 \in [u_1, \gamma_1]\}$,
 $D(\mathcal{L}_{\{1\}}^*) = \{\tau \in \mathcal{L}_0; \tau_1 \in [u_1, \gamma_1]\}$ and $D(\mathcal{L}_{\{2\}}^*) = \{\tau \in \mathcal{L}_0; \tau_1 \in [u_1, v_1]\}$.
- (b) $\gamma_1 \leq \gamma_2$ and $u_1 > \gamma_1$: $\tau_2^1 = \sigma_2$, $D(\mathcal{L}_J^*) = \{(\gamma_1, \sigma_2)\}$,
 $D(\mathcal{L}_{\{1\}}^*) = \{(\gamma_1, \sigma_2)\}$ and $D(\mathcal{L}_{\{2\}}^*) = \{\tau \in \mathcal{L}_0; \tau_1 \in [u_1, v_1]\}$.
- (c) $\gamma_1 > \gamma_2$ and $u_1 \leq \gamma_2$: $\tau_2^1 = u_2$, $D(\mathcal{L}_J^*) = \{\tau \in \mathcal{L}_0; \tau_1 \in [u_1, \gamma_1]\}$,
 $D(\mathcal{L}_{\{1\}}^*) = \{\tau \in \mathcal{L}_0; \tau_1 \in [u_1, \gamma_1]\}$ and $D(\mathcal{L}_{\{2\}}^*) = \{\tau \in \mathcal{L}_0; \tau_1 \in [u_1, v_1]\}$.
- (d) $\gamma_1 > \gamma_2$, $u_1 > \gamma_2$ and $\sigma_1 \leq \gamma_1$: $\tau_2^1 = \gamma_2$, $D(\mathcal{L}_J^*) = \{\tau \in \mathcal{L}_0; \tau_1 \in [\sigma_1, \gamma_1]\}$,
 $D(\mathcal{L}_{\{1\}}^*) = \{\tau \in \mathcal{L}_0; \tau_1 \in [\sigma_1, \gamma_1]\}$ and $D(\mathcal{L}_{\{2\}}^*) = \{\tau \in \mathcal{L}_0; \tau_1 \in [\sigma_1, v_1]\}$.

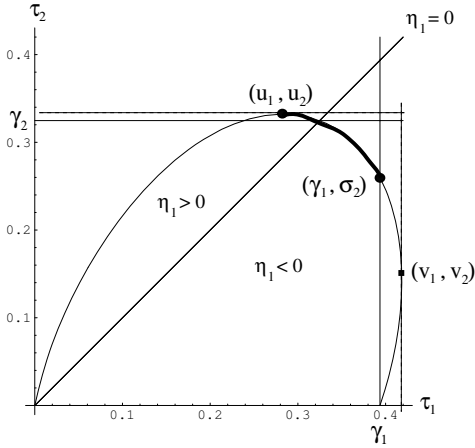


Figure 6: Tandem case (c)

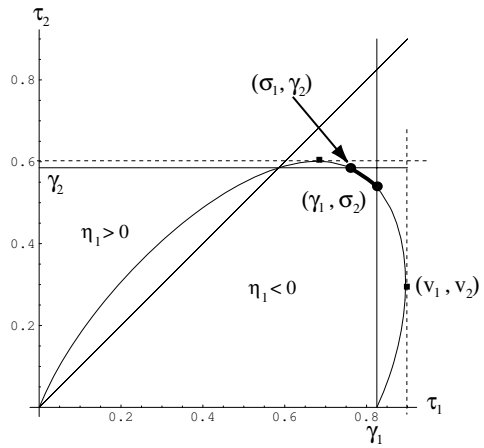


Figure 7: Tandem case (d)

- (e) $\gamma_1 > \gamma_2$, $u_1 > \gamma_2$ and $\sigma_1 > \gamma_1$: $\tau_2^1 = \gamma_2$, $D(\mathcal{L}_J^*) = \{(\gamma_2, \gamma_2), (0, \gamma_1)\}$,
 $D(\mathcal{L}_{\{1\}}^*) = \{\tau \in \mathcal{L}_0; \tau_1 \in [u_1, \gamma_1]\}$ and $D(\mathcal{L}_{\{2\}}^*) = \{\tau \in \mathcal{L}_0; \tau_1 \in [\sigma_1, v_1]\}$.

Note that the decay rate can be found in $D(\mathcal{L}_J^*)$ except case (e). Figures 4-8 illustrate these five cases, for which mixtures of Erlang distributions are used for the interarrival times.

Since $\sigma_2 = \gamma_1 + \log(\mu_2/\mu_1)$, or equivalently,

$$e^{-\sigma_2} = \frac{\mu_1}{\mu_2} e^{-\gamma_1},$$

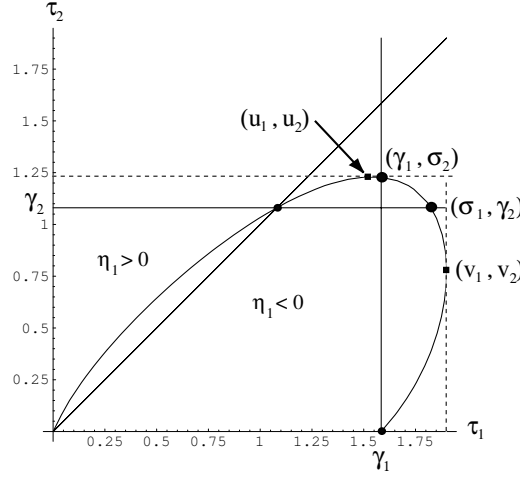


Figure 8: Tandem case (e)

which is obtained from (5.11) in the following form:

$$\hat{f}_1(\mu_1(1 - e^{-\tau_1}) + (1 - e^{-\tau_2})(\mu_2 - \mu_1 e^{-\tau_1 + \tau_2})) = e^{-\tau_1},$$

the decay rates τ_2^1 of the marginal distribution at node 2 are fully compatible with those obtained by Ganesh and Anantharam [10], in which three different cases, (a)+(c), (b) and (d)+(e), appear. For instance, from those classifications, we can observe that the decay rate is $(c_1 + c_2)\gamma_1 + c_2 \log(\mu_2/\mu_1)$ for direction \mathbf{c} in the case (b), while, in the case (e), the decay rates are classified in the following two cases. Let $\boldsymbol{\tau}(\mathbf{c})$ be the $\boldsymbol{\tau}$ that gives the exact decay rate. Then,

- (i) $c_1 < \frac{\sigma_2 - \gamma_2}{\sigma_1 - \gamma_1} c_2$: $\boldsymbol{\tau}(\mathbf{c}) = (\sigma_1, \gamma_2)$.
- (ii) $c_1 \geq \frac{\sigma_2 - \gamma_2}{\sigma_1 - \gamma_1} c_2$: $\boldsymbol{\tau}(\mathbf{c}) = (\gamma_1, \sigma_2)$.

Types 1, 2 and 3 of [14] correspond with cases (b), (e) and (a)+(c)+(d), respectively.

From these figures, we can answer to the question when the feasible set of the decay rates is a singleton, i.e., the two queues are asymptotically independent. To this end, we change μ_1 , while fixing μ_2 and F_1 . We first observe that case (b) occurs, i.e., $\gamma_1 < \min(u_1, \gamma_2)$, if $\mu_1 \leq \underline{\beta}$ for some $\underline{\beta} \leq \mu_2$, since γ_1 goes down to 0 as μ_1 goes down to λ_1 . Similarly, case (d) with $(\gamma_1, \gamma_2) \in \mathcal{L}_0$ occurs if $\mu_1 \geq \bar{\beta}$ for some $\bar{\beta} \geq \mu_2$, since γ_1 goes to infinity as μ_1 goes to infinity. These observations are compatible with the intuition that the asymptotic independence is expected if the first server is sufficiently slow or sufficiently fast. In the former case, the interdeparture times are distributed as the service times with a high probability, while, in the latter, they are distributed almost the same as the interarrival times of exogenous customers.

We now consider the generic service time case. The above arguments can be extended for this case. Since $1/\hat{g}_i(\eta)$'s are increasing in η , their inverse functions exist. Denote these functions by h_i 's. Then, conditions (4.9), (4.11) and (4.12) reduce to

$$\hat{f}_1(h_1(e^{-\tau_1 + \tau_2}) + h_2(e^{-\tau_2})) = e^{-\tau_1}, \quad (5.13)$$

which corresponds with (5.11). Furthermore, $\eta_1 \geq 0$ if and only if $\tau_2 \geq \tau_1$, and $\eta_2 \geq 0$ if and only if $\tau_2 \leq 0$. Thus, let γ_i be the solution of the equation:

$$\hat{f}_1(h_i(e^{-\tau_i})) = e^{-\tau_i},$$

and let (u_1, u_2) and (v_1, v_2) be the contact points of (5.13) with the horizontal and vertical lines, respectively, then we have the same classifications for the decay rates as in the case of Example 5.2. \square

Example 5.3 (2-node network with feedback routes) Our third example is also a two node network, but allows feedback routs, i.e., $r_{12}, r_{21} > 0$. We denote r_{12} by p_1 and r_{21} by p_2 . For simplicity, we assume that $r_{11} = r_{22} = 0$, but this is not essential. We use the same notation as in the previous example. In addition to it, we need the interarrival distribution F_2 for node 2. We denote its LST and mean by \hat{f}_2 and $1/\lambda_2$, respectively. The solution of traffic equations (4.1) is

$$\alpha_1 = \frac{\lambda_1 + \lambda_2 p_2}{1 - p_1 p_2}, \quad \alpha_2 = \frac{\lambda_2 + \lambda_1 p_1}{1 - p_1 p_2},$$

so the stability conditions of this model that we assume are

$$\frac{\lambda_1 + \lambda_2 p_2}{1 - p_1 p_2} < \mu_1, \quad \frac{\lambda_2 + \lambda_1 p_1}{1 - p_1 p_2} < \mu_2.$$

Katou and Makimoto [13] studied product form bounds for this model, assuming phase type distributions for interarrival times and service times (see Conjecture 4.2). We here consider classifications of the decay rates similar to the previous example. The exponential server case is also instructive for this model.

Assume that both servers have exponentially distributed service times. Similar to the previous example, (4.11) and (4.12) become

$$\begin{aligned} \eta_1 &= \mu_1 \left((1 - p_1) e^{-\tau_1} + p_1 e^{-\tau_1 + \tau_2} - 1 \right), \\ \eta_2 &= \mu_2 \left((1 - p_2) e^{-\tau_2} + p_2 e^{-\tau_2 + \tau_1} - 1 \right), \end{aligned}$$

while (4.9) and (4.11) yield

$$\sum_{i=1}^2 \left(\hat{f}_i^{-1}(e^{-\tau_i}) + (1 - p_i) \mu_i e^{-\tau_i} \right) + (\mu_1 p_1 + \mu_2 p_2) e^{-\tau_1 + \tau_2} = \mu_1 + \mu_2. \quad (5.14)$$

This equation describes the boundary of a convex set. Let γ_1 (γ_2) be the largest positive solution τ of these equations when $\eta_2 = 0$ ($\eta_1 = 0$, respectively). By the stability condition, these γ_1 and γ_2 exist. For instance, γ_1 is a solution τ of the equations:

$$\hat{f}_1^{-1}(e^{-\tau_1}) + \hat{f}_2^{-1}(e^{-\tau_2}) + \mu_1 e^{-\tau_1} ((1 - p_1) + p_1 e^{-\tau_2}) = \mu_1, \quad (5.15)$$

$$(1 - p_2) + p_2 e^{\tau_1} = e^{\tau_2}. \quad (5.16)$$

Note that (5.15) is the boundary of a bounded and convex set, and τ_2 of (5.16) is an increasing and concave function of τ_1 . Since $\mathbf{0} = (0, 0)$ is a trivial solution, there exist at least one and at most three positive solutions (τ_1, τ_2) of (5.15) and (5.16). Let $\mathbf{u} \equiv (u_1, u_2) > \mathbf{0}$ and $\mathbf{v} \equiv (v_1, v_2) > \mathbf{0}$ be the contact points of the horizontal and vertical lines with (5.15), respectively. We also let $\sigma_1(y)$ be a solution τ_1 of (5.14) with $\tau_2 = y$ such that $\tau_1 > y$, and let $\sigma_2(x)$ be a positive solution of τ_2 of (5.14) with $\tau_1 = x$. Then, the decay rate τ is obtained according to how points $\gamma_1, \gamma_2, \mathbf{u}, \mathbf{v}, (\sigma_1(\gamma_{22}), \sigma_2(\gamma_{11}))$ are located. For instance, if $v_1 \leq \gamma_{11}$ and $u_2 \leq \gamma_{22}$, then $D(\mathcal{L}_j^*) = \mathcal{L}_0 \cap \{\tau > \mathbf{0}; \tau_1 \in [u_1, v_1]\}$.

For the case of generally distributed service times, (5.14) will be more complicated, but the decay rates τ can be classified in the same way. \square

6. Queueing Networks with Batch Movements

We now consider the batch movement network. This network is described by a discrete-time Markov chain. Similar to the standard Jackson network, the network is composed of k nodes, numbered as $1, 2, \dots, k$. So the state of the Markov chain is a k -dimensional vector of nonnegative integers. Let \mathbf{Y}_n be the state just after the n -th departure-arrival transition, let \mathbf{D}_n be the requested departures at this transition instant and let \mathbf{A}_n be the arrivals at the same instant subsequent to the departures. Then,

$$\mathbf{Y}_{n+1} = \max(\mathbf{0}, \mathbf{Y}_n - \mathbf{D}_n) + \mathbf{A}_n, \quad n = 0, 1, \dots \quad (6.1)$$

Hence, if \mathbf{Y}_n has a stationary distribution, we have

$$\mathbf{Y} \stackrel{d}{\simeq} \max(\mathbf{0}, \mathbf{Y} - \mathbf{D}) + \mathbf{A}, \quad n = 0, 1, \dots, \quad (6.2)$$

where $\stackrel{d}{\simeq}$ stands for the equality in distribution, \mathbf{Y} is a random vector subject to the stationary distribution, and \mathbf{D} and \mathbf{A} are random vectors subject to a generic joint distribution of the requested departure and arrival sizes. Here, we tacitly assume that the arrival sizes do not depend on the actual departure sizes which may be different from the requested departure sizes. In general, there are various policies to specify how the arrival sizes are changed when the requested departures are not satisfied. Those may affect stochastic bounds for the stationary distribution, but the decay rates may be independent of them. So we here adopt the simple policy.

Clearly, if $E(D_i) > E(A_i)$ for all i , then the stationary distribution for \mathbf{Y} exists. We assume this stability condition from now on. Suppose that the stationary distribution has the exponential decay with respect to direction \mathbf{c} . Namely, as $\mathbf{n} \sim |\mathbf{n}|\mathbf{c}$ ($|\mathbf{n}| \rightarrow \infty$), there exist $\boldsymbol{\tau}$ and constant $C_0 > 0$ such that

$$P(\mathbf{Y} \geq \mathbf{n}) \cong C_0 e^{-\langle \boldsymbol{\tau}, \mathbf{n} \rangle}. \quad (6.3)$$

Then, from (6.2), we have, for large \mathbf{n} ,

$$\begin{aligned} P(\mathbf{Y} \geq \mathbf{n}) &\cong E(P(\mathbf{Y} + \mathbf{A} - \mathbf{D} \geq \mathbf{n} | \mathbf{A}, \mathbf{D})) \\ &\cong C_0 E(e^{-\langle \boldsymbol{\tau}, \mathbf{n} - \mathbf{A} + \mathbf{D} \rangle}) \\ &= C_0 e^{-\langle \boldsymbol{\tau}, \mathbf{n} \rangle} E(e^{-\langle \boldsymbol{\tau}, \mathbf{D} - \mathbf{A} \rangle}). \end{aligned}$$

Hence, we get

$$E(e^{-\langle \boldsymbol{\tau}, \mathbf{D} - \mathbf{A} \rangle}) = 1. \quad (6.4)$$

Denote the joint distribution of \mathbf{D} and \mathbf{A} by H and its LST by \hat{h} . (6.4) can be written as

$$\hat{h}(\boldsymbol{\tau}, -\boldsymbol{\tau}) = 1. \quad (6.5)$$

Similar to the case of the generalized Jackson network, we let $\mathcal{K}_0 = \{\boldsymbol{\tau} \geq \mathbf{0}; \hat{h}(\boldsymbol{\tau}, -\boldsymbol{\tau}) = 1\}$ and $\bar{\mathcal{K}}_0 = \{\boldsymbol{\tau} \geq \mathbf{0}; \hat{h}(\boldsymbol{\tau}, -\boldsymbol{\tau}) \leq 1\}$. We also assume that there exist a $\theta_0^{(i)} < \infty$ and a finite vector $\mathbf{d}^{(i)} \geq \mathbf{0}$ for each i such that

$$\lim_{\theta \uparrow \theta_0^{(i)}} E(e^{\theta A_i}; \mathbf{D} \leq \mathbf{d}^{(i)}) = \infty, \quad i = 1, 2, \dots, k. \quad (6.6)$$

Lemma 6.1 $\bar{\mathcal{K}}_0$ is convex and bounded.

PROOF. Since $\hat{h}(\boldsymbol{\tau}, -\boldsymbol{\tau})$ is a convex function of $\boldsymbol{\tau}$, $\bar{\mathcal{K}}_0$ is convex. For each direction \mathbf{c} , define a function φ by

$$\varphi(t) = E\left(e^{-t\langle \mathbf{c}, \mathbf{D} - \mathbf{A} \rangle}\right), \quad t \geq 0.$$

Since $E(D_i) > E(A_i)$, $\varphi(t)$ has the right-hand derivative at $t = 0$ that is negative. On the other hand, since

$$\varphi(t) \geq \min_i e^{-t \sum_{j=1}^k c_j d_j^{(i)}} E\left(e^{tc_i A_i}; \mathbf{D} \leq \mathbf{d}^{(i)}\right), \quad t \geq 0,$$

(6.6) implies that $\varphi(t) \rightarrow \infty$ as $t \uparrow \max_i c_i \theta_0^{(i)}$. This implies that $\varphi(t) = 1$ has a unique positive solution since φ is convex. Hence, $\hat{h}(\boldsymbol{\tau}, -\boldsymbol{\tau}) = 1$ constitutes a closed surface in the positive quadrant $\boldsymbol{\tau} \geq \mathbf{0}$, so $\bar{\mathcal{K}}_0$ is bounded. \square

We next consider to choose feasible $\boldsymbol{\tau}$ from \mathcal{K}_0 for the decay rate. Although the situation is much simpler, we use the same idea as in the case of the generalized Jackson network. That is, feasible $\boldsymbol{\tau}$ should not be greater than the corresponding marginals if it gives a product form upper bound. Let $\bar{\boldsymbol{\tau}}(B)$ be $\boldsymbol{\tau} \geq \mathbf{0}$ such that $\tau_i = 0$ for $i \notin B$. Then, for each subset $B \subset \{1, 2, \dots, k\}$, we define

$$\mathcal{K}_B = \{\boldsymbol{\tau} \in \mathcal{K}_0; \hat{h}(\bar{\boldsymbol{\tau}}(B), -\bar{\boldsymbol{\tau}}(B)) \leq 1\}.$$

This set is the first step to restrict $\boldsymbol{\tau}$ so as to bound the marginal distribution with respect to B . We need more steps, and, for a full restriction, we define

$$\mathcal{K}_B^* = \cap_{B' \subset B} \mathcal{K}_{B'}.$$

Note that \mathcal{K}_B and \mathcal{K}_B^* corresponds with \mathcal{L}_B and \mathcal{L}_B^* , respectively, in Section 4. Then, we may expect the following two conjectures.

Conjecture 6.1 Let \mathcal{A} be a partition of J . Then, for each $\boldsymbol{\tau}_B \in \mathcal{K}_B^*$ for $B \in \mathcal{A}$, there exist positive constants C_B 's such that

$$P(\mathbf{Y} \geq \mathbf{n}) \leq \sum_{B \in \mathcal{A}} C_B \exp(-\langle \boldsymbol{\tau}_B, \mathbf{n} \rangle), \quad \mathbf{n} \geq \mathbf{0}. \quad (6.7)$$

Conjecture 6.2 For each direction vector \mathbf{c} and each $B \in J$, the exact rate is obtained as

$$\omega(\mathbf{c}) = \max_{\mathcal{A} \in \mathcal{P}(J)} \min_{B \in \mathcal{A}} \sup\{\langle \boldsymbol{\tau}, \mathbf{c} \rangle; \boldsymbol{\tau} \in \mathcal{K}_B^*\}, \quad (6.8)$$

where $\mathcal{P}(J)$ is the set of all partitions of J .

Remark 6.1 (i) Kella and Miyazawa [15] verifies Conjecture 6.1 for $\mathcal{A} = \{J\}$ with $C_J = 1$ for a fluid network corresponding with the batch movement network when the arrivals occur independent of the requested departures and their amounts are functions of a random variable subject to an exponential distribution (see Theorem 3.1 of [15]).

(ii) If the arrivals and departures can not simultaneously occur at more than one node in the batch movement network, then Miyazawa and Taylor [21] showed that (6.7) is obtained for $\mathcal{A} = \{J\}$ with $C_J = 1$ when all unsatisfied departure requests no arrivals (see also [5] for complex cases).

Because of the simpler structure, we can say more about the decay rate in the batch movement network. We first observe that, for each direction vector \mathbf{c} , (6.1) implies

$$\langle \mathbf{Y}_{n+1}, \mathbf{c} \rangle \geq \max(0, \langle \mathbf{Y}_n, \mathbf{c} \rangle - \langle \mathbf{D}_n, \mathbf{c} \rangle) + \langle \mathbf{A}_n, \mathbf{c} \rangle.$$

Hence, if we define \underline{Y}_n inductively by

$$\underline{Y}_{n+1} = \max(0, \underline{Y}_n, \langle \mathbf{A}_n, \mathbf{c} \rangle - \langle \mathbf{D}_n, \mathbf{c} \rangle).$$

Then, \underline{Y}_n is stochastically lower than $\langle \mathbf{Y}_n, \mathbf{c} \rangle$. Clearly, the stationary distribution of \underline{Y}_n has the same exponential decay rate as the waiting time of the $G/G/1$ queue with inter-arrival times $\{\langle \mathbf{D}_n, \mathbf{c} \rangle\}$ and service times $\{\langle \mathbf{A}_n, \mathbf{c} \rangle\}$. Here, $\langle \mathbf{D}_n, \mathbf{c} \rangle$ and $\langle \mathbf{A}_n, \mathbf{c} \rangle$ may depend on each other, but it is not essential. Hence, the decay rate, denoted by $\omega(\mathbf{c})$, is given by a unique solution $t > 0$ of the equation:

$$\hat{h}(t\mathbf{c}, -t\mathbf{c}) = 1. \quad (6.9)$$

Thus, the decay rate of $\langle \mathbf{Y}, \mathbf{c} \rangle$ is bounded below by $\omega(\mathbf{c})$. On the other hand, if $\omega(\mathbf{c})\mathbf{c} \in \mathcal{K}_J^*$, then, using the same arguments as in Theorem 4.1 of [15], it can be seen that the decay rate of $\langle \mathbf{Y}, \mathbf{c} \rangle$ is not less than $\omega(\mathbf{c})$. Hence, Conjecture 6.1 implies that

Conjecture 6.3 If $\omega(\mathbf{c})\mathbf{c} \in \mathcal{K}_J^*$, then \mathbf{Y} has the decay rate $\omega(\mathbf{c})$ for direction \mathbf{c} concerning tail set $V = \{\mathbf{y} \geq \mathbf{0}; \langle \mathbf{y}, \mathbf{c} \rangle \geq 0\}$.

7. Concluding Remarks

The conjectures in Sections 4 and 6 may be extended for more complex models. For the generalized Jackson network, suppose that single servers at nodes are changed to many servers. Then it would be easy to extend the conjectures. For example, if node i has s_i heterogenous servers with LST $\hat{g}_{i,\ell}(\eta_{i,\ell})$, $\ell = 1, \dots, s_i$, for the service time distributions, then (4.9) and (4.12) are changed to

$$\sum_{j=1}^k \left(\theta_j 1(\lambda_j \neq 0) + \sum_{\ell=1}^{s_j} \eta_{j,\ell} \right) = 0. \quad (7.1)$$

$$e^{\tau_i} = \hat{g}_i(\eta_{i,\ell}) \left(r_{i0} + \sum_{j=1}^k r_{ij} e^{\tau_j} \right), \quad i = 1, \dots, k, \ell = 1, \dots, s_i. \quad (7.2)$$

while (4.11) is unchanged. Similar to the case of single servers, for $k = 1$, the decay rate agrees with the existence result (see, e.g., [2, 22]). Another variation is to extend the renewal arrivals to their superposition. This can be implemented by using fictitious nodes in the current model. For example, let node 1 be such a node with $S_0 = 0$. Then, node $j \neq i$ has arrivals subject to the superposition of two renewal processes if $r_{ij} > 0$.

For the batch movement network, we can consider its fluid version, and the fluid versions of Conjectures 6.1 and 6.2 may be also conjectured.

As we remarked, some of the conjectures are partially verified for some special cases. However, any major part of the conjectures has not yet been verified. To attack this problem, we are recently developing an approach using Markov additive processes with general background states. Since this is still under a way, we here roughly explain its basic idea. A key observation is that we can describe the network process in terms of a Markov additive process along the direction vector if we incorporate sufficient information into the

background states. Then, the decay rate of the tail probabilities can be obtained through the Markov renewal theorem for this additive process (see [20] for more details). In other words, the renewal arguments in Section 3 may be extended for the network models.

Acknowledgements This paper originally came out of my talk at a special session of the autumn meeting of Operations Research Society of Japan, 2000. The author is grateful to the organizer, Hirotaka Sakasegawa of Waseda University. He is also indebted to Naoki Makimoto of Tsukuba University for stimulative discussions and to Ken'ichi Katou of Yamagata University for his comments and providing his paper [14] coauthored with Yukio Takahashi of Tokyo Institute of Technology. In fact, the author originally did not have Conjecture 4.4 in the present form, which owes to their result (4.18). The author is also grateful to the anonymous referee and associate editor for their helpful comments. This research is supported in part by the Japan Society for the Promotion of Science under grant No. 13680532.

References

- [1] J. Abate, G. L. Choudhury and W. Whitt: Asymptotics for steady-state tail probabilities in structured Markov queueing models. *Stochastic Models*, **10** (1994) 99-143.
- [2] S. Asmussen and C. A. O'Cinneide: Representations for matrix-geometric and matrix-exponential steady-state distributions to many-server queues. *Stochastic Models*, **14** (1998) 369-387.
- [3] D. Bertsimas, I. C. Paschalidis and J. N. Tsitsiklis: On the large deviations behavior of acyclic networks of $G/G/1$ queues. *Annals of Applied Probability*, **8** (1998) 1027-1069.
- [4] C. S. Chang: Sample path large deviations and intree networks. *Queueing Systems*, **20** (1995) 7-36.
- [5] X. Chao, M. Miyazawa and M. Pinedo: *Queueing Networks; customers, signals and product form solutions* (Wiley, Chichester, 1999).
- [6] G. de Vaciara and J. Walrand: Effective bandwidths: Call admission, traffic policing and filtering for ATM networks. *Queueing Systems*, **20** (1995) 37-59.
- [7] W. Feller: *An Introduction to Probability Theory and Its Application*, Vol. II (Wiley, New York, 1971).
- [8] K. Fujimoto and Y. Takahashi: Tail behavior of the steady-state distribution in two-stage tandem queues: numerical experiment and conjecture. *J. Opns. Res. Soc. Japan*, **39** (1996) 525-540.
- [9] K. Fujimoto, Y. Takahashi and N. Makimoto: Asymptotic properties of stationary distributions in two stage tandem queueing systems. *J. Opns. Res. Soc. Japan*, **41** (1998) 118-141.
- [10] A. Ganesh and V. Anantharam: Stationary tail probabilities in exponential server tandems with renewal arrivals *Queueing Systems*, **22** (1996) 203-247.
- [11] P. Glynn and W. Whitt: Logarithmic asymptotics for steady state tail probabilities in a single-server queue. *J. Appl. Prob.*, **31A** (1994) 131-156.
- [12] K. Katou and N. Makimoto: On the decay rate for a two node queueing network with MAP arrivals. *The Proceedings of Performance Models for Information Communication Networks* (1999) 189-198, Hamanako, Japan (in Japanese).
- [13] K. Katou and N. Makimoto: Upper bounds for the decay rates of the joint distribution in tow-node Markovian queues. *The Proceedings of Performance Models for Information Communication Networks* (2000) 1-10, Sendai, Japan.

- [14] K. Katou and Y. Takahashi: Exponential decay of the stationary joint queue length distribution in two-node Markovian queueing system. Preprint (2001).
- [15] O. Kella and M. Miyazawa: Parallel fluid queues with constant inflows and simultaneous random reductions. *J. Appl. Prob.*, **38** (2001) 609-620.
- [16] J.F.C. Kingman: Inequalities in the theory of queues. *J. Roy. Stat. Soc. B*, **32** (1970) 883-909.
- [17] N. Makimoto: Private communications (2000).
- [18] S. P. Meyn and D. Down: Stability of generalized Jackson networks. *Annals of Applied Probability*, **4** (1994) 124-148.
- [19] M. Miyazawa: Rate conservation laws: a survey. *Queueing Systems*, **15** (1994) 1-58.
- [20] M. Miyazawa: A paradigm of Markov additive processes for queues and their networks. *The Proceedings of the Fourth International Conference on Matrix-Analytic Methods in Stochastic Models: Matrix-Analytic Methods, Theory and Applications* (2002) 265-289 (World Scientific, Singapore).
- [21] M. Miyazawa and P. G. Taylor: A geometric product-form distribution for a queueing network with nonstandard batch arrivals and batch transfers. *Adv. Appl. Prob.*, **29** (1997) 523-544.
- [22] M. F. Neuts and Y. Takahashi: Asymptotic behavior of the stationary distributions in the $GI/PH/c$ queue with heterogeneous servers. *Zeitschrift Wahrscheinlichkeitstheorie*, **57** (1997) 441-452.
- [23] N. O'Connell: Large deviations for departures from a shared buffer. *J. Appl. Prob.*, **34** (1997) 753-766.
- [24] N. O'Connell: Large deviations for queue lengths at a multi-buffered resource. *J. Appl. Prob.*, **35** (1998) 240-245.
- [25] J. C. W. Ommeren: Exponential expansion for the tail of the waiting-time probability in the single-server queue with batch arrivals. *Adv. Appl. Prob.*, **20** (1988) 880-895.
- [26] J. S. Sadowsky and W. S. Szpankowski: The probability of large queue lengths and waiting times in a heterogeneous multiserver queue I: Tight limits. *Adv. Appl. Prob.*, **27** (1995) 532-566.
- [27] J. S. Sadowsky: The probability of large queue lengths and waiting times in a heterogeneous multiserver queue I: Positive recurrence and logarithmic limits. *Adv. Appl. Prob.*, **27** (1995) 567-583.
- [28] R. F. Serfozo: *Introduction to Stochastic Networks* (Springer-Verlag, New York, 1999).
- [29] K. Sigman: The stability of open queueing networks. *Stochastic Process. Appl.*, **35** (1990) 11-25.
- [30] Y. Takahashi: Asymptotic exponentiality of the tail of the waiting-time distribution in a $PH/PH/c$ queue. *Adv. Appl. Prob.*, **13** (1981) 619-630.
- [31] D. Wischik: The output of a switch, or, effective bandwidths for networks. *Queueing Systems*, **32** (1999) 383-396.

A. Proof of Lemma 4.1

We prove Lemma 4.1 for the departure instants at node i . We use (4.4) with $\theta_j = 0$ for $j \neq i$ and $\eta_j = 0$ for all j . We first evaluate its left-hand side. For this, let A be the attained service time of a customer being served at node i at time 0 when $\mathbf{X} \geq \mathbf{n} > \mathbf{0}$ in the steady state. Then, the left-hand side with $\theta = \theta_i$ can be written as

$$\theta E(e^{-\theta R}; \mathbf{X} \geq \mathbf{n}) = E(\theta E(e^{-\theta(S_i - A)} | S_i > A) | \mathbf{X} \geq \mathbf{n}) P(\mathbf{X} \geq \mathbf{n}).$$

From (4.7), applying a partial integration, we have, for any $y > 0$,

$$\begin{aligned}
\theta E(e^{-\theta(S_i-y)} | S_i > y) &= \frac{1}{1 - G_i(y)} \int_y^\infty \theta e^{-\theta(x-y)} dG_i(x) \\
&\leq \frac{b}{1 - G_i(y)} \int_y^\infty \theta e^{-\theta(x-y)} (1 - G_i(x)) dx \\
&= \frac{b}{1 - G_i(y)} \left(\left[-e^{-\theta(x-y)} (1 - G_i(x)) \right]_y^\infty - \int_y^\infty e^{-\theta(x-y)} dG_i(x) \right) \\
&= b \left(1 - \int_y^\infty e^{-\theta(x-y)} \frac{dG_i(x)}{1 - G_i(y)} \right) < b.
\end{aligned}$$

Similarly, we have

$$\theta E(e^{-\theta(S_i-y)} | S_i > y) \geq a \left(1 - \int_y^\infty e^{-\theta(x-y)} \frac{dG_i(x)}{1 - G_i(y)} \right).$$

Hence,

$$a \leq \liminf_{\theta \rightarrow \infty} \theta E(e^{-\theta(S_i-y)} | S_i > y) \leq \limsup_{\theta \rightarrow \infty} \theta E(e^{-\theta(S_i-y)} | S_i > y) \leq b.$$

This implies, by the bounded convergence theorem,

$$a \leq \liminf_{\theta \rightarrow \infty} \theta E(e^{-\theta R} | \mathbf{X} \geq \mathbf{n}) \leq \limsup_{\theta \rightarrow \infty} \theta E(e^{-\theta R} | \mathbf{X} \geq \mathbf{n}) \leq b.$$

We next let θ_i go to infinity in the right-hand side of (4.4), then it becomes $\alpha_i \varphi_i(\mathbf{n}, \mathbf{0}, \mathbf{0})$. Hence, we get

$$aP(\mathbf{X} \geq \mathbf{n}) \leq \alpha_i P_i^D(\mathbf{X}^- \geq \mathbf{n}) \leq bP(\mathbf{X} \geq \mathbf{n}), \quad \mathbf{n} > \mathbf{0}.$$

A similar proof can be done for the arrival instants. These complete the proof.

Masakiyo Miyazawa
 Department of Information Sciences
 Tokyo University of Science
 Yamazaki 2641, Noda-shi
 Chiba 278-8510, Japan
 E-mail: miyazawa@is.noda.tus.ac.jp