Bounds on the Tail Distributions of Markov-Modulated Stochastic Max-Plus Systems

Zhen LIU[1*] Philippe NAIN[1†] and Don TOWSLEY[2‡]

[1]INRIA, B.P. 93, 06902, Sophia Antipolis Cedex, France
[2]Department of Computer Science, University of Massachusetts, Amherst, MA 01003, USA

In: Proceedings of CDC'95, New Orleans, LA, Dec. 1995, pp. 1395-1399

Abstract

In this paper we consider a particular class of linear systems under the max-plus algebra and derive exponential upper bounds for the tail distribution of each component of the state vector in the case of Markov modulated input sequences. Our results are then applied to tandem queues with infinite buffers, Markov modulated arrivals and deterministic service times.

1 Introduction

Consider a stochastic discrete-event system with max-plus dynamics:

$$X_{n+1,k} = \max_{j\in\mathcal{K}} (X_{n,j} + A_{n,jk}(Y_n))^+, \qquad \forall n \geq 0 \tag{1}$$

where $(a)^+ = \max(0,a)$ and $\mathcal{K} = \{1,2,\cdots,K\}$ $(K < \infty)$. The state variables $X_n = (X_{n,k},\ k \in \mathcal{K})$ are K-dimensional nonnegative random variables (r.v.'s), $X_n \in [0,\infty)^K$. The sequence $(Y_n)_n$ is a finite-state, irreducible, aperiodic and homogeneous Markov chain on the finite set $\mathcal{S} = \{1,2,\cdots,S\}$. For each $s \in \mathcal{S}$, $n \geq 0$, introduce the matrix $\mathbf{A}_n(s) = [A_{n,jk}(s)]_{j,k\in\mathcal{K}}$. We assume that

(A1) $(\mathbf{A}_n(1))_n, \ldots, (\mathbf{A}_n(S))_n$ form S mutually independent renewal sequences of random matrices with $-\infty \leq A_{n,jk}(s) < \infty$ almost surely (a.s.);

(A2) for each $s \in \mathcal{S}$, the matrices $(\mathbf{A}_n(s))_n$ are independent of the Markov chain $(Y_n)_n$.

Note, however, that for any given $n \geq 0$ and $s \in \mathcal{S}$, the entries of the matrix $\mathbf{A}_n(s)$ may be dependent r.v.'s. We further assume that $X_0 = (X_{0,1},\cdots,X_{0,K})$ is a nonnegative and a.s. finite random vector.

Using the max-plus algebra operators, (1) can be written in vector form as follows

$$X_{n+1} = X_n \otimes \mathbf{A}_n(Y_n) \oplus 0, \qquad \forall n \geq 0 \tag{2}$$

where $X_n = (X_{n,1},\cdots,X_{n,K})$. In (2) the "plus" operator $\otimes$ and the "max" operator $\oplus$ replace the usual matrix multiplication and matrix addition, respectively.

Our objective is to derive exponential upper bounds on the tail distribution of $X_{n,k}$, namely, to find positive constants b and θ such that $P(X_{n,k} > x) \leq b\exp(-\theta x)$ for all $x \geq 0$, $k \in \mathcal{K}$, $n \geq 0$.

There is now a substantial body of work ([1, 3, 6, 9, 10, 13, 15, 14, 16, 19, 20, 23, 24] among others) on the problem of deriving exact, approximate or asymptotic bounds for the tail distribution of the workload, queue-length, and delay in a queue in isolation. Bounds of this type have proved to be helpful for analyzing ATM multiplexers fed by different classes of traffic such as data, voice, video, etc. (see e.g. [12]).

This paper is one of the first attempts (see also [4, 5, 8, 17, 22] for related work) to derive bounds for more general structures, including acyclic queueing networks. Bounding quantities such as the probability distribution of the end-to-end delay seen by individual sessions along their route in the network will allow network designers to get a better understanding of the network behavior and to come up with more efficient admission control schemes as opposed to schemes based only on the performance at isolated nodes.

As in [2], we define the *communicating graph* $G = (V,E)$ of the system as follows. The set of vertices is $V = \mathcal{K}$. For each pair of vertices $j,k \in \mathcal{K}$, there is an arc from j to k if and only if there is at least one state $s \in \mathcal{S}$ such that $A_{n,jk}(s) \neq -\infty$ with strictly positive probability. Let $\mathcal{G} = \{G_1, G_2, \cdots, G_g\}$ be the maximal decomposition of G into strongly connected subgraphs, where $G_i = (V_i, E_i)$, $1 \leq i \leq g$, are strongly connected subgraphs of G. Without loss of generality, we assume that vertices of G_i have smaller indices than those of G_{i+1}.

Under such a decomposition, the matrices $\mathbf{A}_n(s)$, $s \in \mathcal{S}$, have the form:

$$\mathbf{A}_n(s) = \begin{pmatrix} \mathbf{B}_{n,1}(s) & \mathbf{R}_{n,1,2}(s) & \cdots & \mathbf{R}_{n,1,g-1}(s) & \mathbf{R}_{n,1,g}(s) \\ -\infty & \mathbf{B}_{n,2}(s) & \cdots & \mathbf{R}_{n,2,g-1}(s) & \mathbf{R}_{n,2,g}(s) \\ -\infty & -\infty & \cdots & \mathbf{R}_{n,3,g-1}(s) & \mathbf{R}_{n,3,g}(s) \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ -\infty & -\infty & \cdots & \mathbf{B}_{n,g-1}(s) & \mathbf{R}_{n,g-1,g}(s) \\ -\infty & -\infty & \cdots & -\infty & \mathbf{B}_{n,g}(s) \end{pmatrix} \tag{3}$$

where $\mathbf{B}_{n,i}(s) = [A_{n,jk}(s)]_{j,k\in V_i}$ and $\mathbf{R}_{n,i_1,i_2}(s) = [A_{n,jk}(s)]_{j\in V_{i_1},\ k\in V_{i_2}}$, $1 \leq i \leq g$, and $1 \leq i_1 < i_2 \leq g$.

Let $\mathbf{P} = [p_{ij}]$ be the transition matrix of the Markov chain $(Y_n)_n$ and let $\pi = (\pi(1),\ldots,\pi(K))$ be its invariant measure. Define $\pi_n(i) = P_{\pi_0}(Y_n = i)$ for $i \in \mathcal{S}$, $n \geq 1$, and let $\pi_0 = (\pi_0(1),\ldots,\pi_0(K))$ be the initial probability distribution. In the following we shall drop the subscript π_0 in P_{π_0}.

*e-mail address: liu@sophia.inria.fr
†e-mail address: nain@sophia.inria.fr
‡e-mail address: towsley@gonzo.cs.umass.edu. This author was supported in part by NSF under grant NCR-9116183

Define $F_{jk,s}(x) = P(A_{n,jk}(Y_n) < x \,|\, Y_n = s)$ and let $\phi_{jk,s}(\theta) = E[\exp(\theta A_{n,jk}(Y_n)) \,|\, Y_n = s]$ where $\phi_{jk,s}(\theta) = 0$ if $A_{n,jk}(s) = -\infty$ a.s.

Throughout this paper we will assume that

(A3) $\Theta = \{\theta > 0 \,:\, \phi_{jk,s}(\theta) < \infty$ for all $(j,k,s) \in \mathcal{K} \times \mathcal{K} \times \mathcal{S}\}$

is a nonempty set. This technical assumption is satisfied in most cases of practical interest which includes r.v.'s with phase-type distributions.

Introduce $h_{jk,st}(\theta) = \phi_{jk,s}(\theta) p_{st}$. Let $\mathbf{H}_{jk}(\theta) = [h_{jk,st}(\theta)]_{s,t\in\mathcal{S}}$ be S-by-S matrices, and let $\mathbf{H}(\theta)$ be the KS-by-KS matrix defined as $\mathbf{H}(\theta) = [\mathbf{H}_{jk}(\theta)]_{j,k\in\mathcal{K}}$.

Then, in view of (3), $\mathbf{H}(\theta)$ can be rewritten as

$$\mathbf{H}(\theta) = \begin{pmatrix} \mathbf{J}_1(\theta) & \mathbf{J}_{1,2}(\theta) & \cdots & \mathbf{J}_{1,g-1}(\theta) & \mathbf{J}_{1,g}(\theta) \\ 0 & \mathbf{J}_2(\theta) & \cdots & \mathbf{J}_{2,g-1}(\theta) & \mathbf{J}_{2,g}(\theta) \\ 0 & 0 & \cdots & \mathbf{J}_{3,g-1}(\theta) & \mathbf{J}_{3,g}(\theta) \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & \mathbf{J}_{g-1}(\theta) & \mathbf{J}_{g-1,g}(\theta) \\ 0 & 0 & \cdots & 0 & \mathbf{J}_g(\theta) \end{pmatrix} \quad (4)$$

where $\mathbf{J}_i(\theta) = [h_{jk,st}(\theta)]_{j,k\in V_i,\ s,t\in\mathcal{S}}$ and $\mathbf{J}_{i_1,i_2}(\theta) = [h_{jk,st}(\theta)]_{j\in V_{i_1},k\in V_{i_2},\ s,t\in\mathcal{S}}$ for $1 \le i_1 < i_2 \le g$.

Let us introduce some additional notation. We will say that a vector v is positive (resp. nonnegative) if its components are all larger than (resp. larger than or equal to) zero. The notation $v > 0$ (resp. $v \ge 0$) will indicate that the vector v is positive (resp. nonnegative). More generally, for any pair of vectors v and w, the notation $v \le w$ will indicate that v is less than or equal to w componentwise.

Since the transition matrix $\mathbf{P}$ of the Markov chain $(Y_n)_n$ is irreducible, the matrices $\mathbf{J}_i(\theta)$ are all irreducible. Thus, according to Perron-Frobenius Theorem [11, Theorem 8.4.4], the spectral radius $\rho_i(\theta)$ of matrix $\mathbf{J}_i(\theta)$ $(1 \le i \le g)$ is a simple eigenvalue of the matrix and there is a positive vector $z_i(\theta) = (z_{k,t}(\theta))_{k\in V_i,\ t\in\mathcal{S}}$ such that

$$z_i(\theta)\,\mathbf{J}_i(\theta) = \rho_i(\theta)\,z_i(\theta). \quad (5)$$

For each $1 \le i \le g$, we normalize the left-eigenvector $z_i(\theta)$ so that

$$\sum_{k\in V_i,\ t\in\mathcal{S}} z_{k,t}(\theta) = 1. \quad (6)$$

Let $\rho(\theta)$ be the spectral radius of $\mathbf{H}(\theta)$. Since the matrix $\mathbf{H}(\theta)$ is block triangular, its spectrum is the union of the spectra of matrices $\mathbf{J}_1(\theta), \ldots, \mathbf{J}_g(\theta)$, which in turn implies that

$$\rho(\theta) = \max_{1\le i\le g} \rho_i(\theta). \quad (7)$$

It then follows [11, Theorem 8.3.1] that $\rho(\theta)$ is also an eigenvalue of $\mathbf{H}(\theta)$ and that there is a nonnegative vector $y(\theta) = (y_{k,t}(\theta))_{k\in\mathcal{K},\ t\in\mathcal{S}}$ such that

$$y(\theta)\,\mathbf{H}(\theta) = \rho(\theta)\,y(\theta). \quad (8)$$

We also normalize $y(\theta)$ so that

$$\sum_{k\in\mathcal{K},\ t\in\mathcal{S}} y_{k,t}(\theta) = 1. \quad (9)$$

We are now in position to state the main results of this work.

2 Exponential Upper Bounds

In this section we derive exponential upper bounds for the tail distribution of $X_{n,k}$ under Assumptions (A1)-(A3).

Proposition 1 (Exponential Upper Bound I) *Assume that $\theta \in \Theta$. If $\rho(\theta) \le 1$, $y(\theta) > 0$, and if for all $k \in \mathcal{K}$, $t \in \mathcal{S}$,*

$$P(X_{0,k} > x, Y_0 = t) \le b(\theta)\,y_{k,t}(\theta)\,e^{-\theta x}, \qquad \forall x \ge 0 \quad (10)$$

then, for all $n \ge 1$, $k \in \mathcal{K}$, $x \ge 0$,

$$P(X_{n,k} > x) \le b(\theta)\,e^{-\theta x} \quad (11)$$

where

$$b(\theta) = \sup_{\substack{x\ge 0\\ n\ge 0\\ k\in\mathcal{K}\\ t\in\mathcal{S}}} \frac{\displaystyle\sum_{j\in\mathcal{K},\,s\in\mathcal{S}} p_{st}\,\pi_n(s)\,(1 - F_{jk,s}(x))}{\displaystyle\sum_{j\in\mathcal{K},\,s\in\mathcal{S}} p_{st}\,y_{j,s}(\theta) \int_x^\infty e^{\theta(u-x)}\,dF_{jk,s}(u)} < \infty. \quad (12)$$

$\diamondsuit$

It is worth noting that condition (10) is automatically satisfied (in particular) if $X_{0,k} = 0$ a.s.

Proof. Fix $\theta \in \Theta$ such that $\rho(\theta) \le 1$ and $y(\theta) > 0$. Let $\{\gamma_{j,s}(x), j \in \mathcal{K}, s \in \mathcal{S}\}$, $\gamma_{j,s} : [0,\infty) \to [0,\infty)$, be a set of functions satisfying

$$\sum_{\substack{j\in\mathcal{K}\\ s\in\mathcal{S}}} p_{st}\left[\int_{-\infty}^{x} \gamma_{j,s}(x-u)\,dF_{jk,s}(u) + (1 - F_{jk,s})(x)\,\pi_n(s)\right] \le \gamma_{k,t}(x). \quad (13)$$

The first step of the proof consists in proving that

$$P(X_{m,k} > x, Y_m = t) \le \gamma_{k,t}(x) \quad (14)$$

for all $m \ge 0$, $x \ge 0$, $k \in \mathcal{K}$, $t \in \mathcal{S}$. For this, we use an induction argument on m. Note that (14) holds for $m = 0$ from (10) and assume that (14) is true for $m = 0, 1, \ldots, n$. Let us show that (14) is still true for $m = n+1$.

For $x \ge 0$, we have

$$P(X_{n+1,k} > x, Y_{n+1} = t)$$
$$= P\left(\max_{j\in\mathcal{K}}(X_{n,j} + A_{n,jk}(Y_n)) > x, Y_{n+1} = t\right)$$

$$\leq \sum_{j\in\mathcal{K}} P(X_{n,j} + A_{n,jk}(Y_n) > x, Y_{n+1} = t)$$

$$= \sum_{j\in\mathcal{K}, s\in\mathcal{S}} p_{st}\pi_n(s) P(X_{n,j} + A_{n,jk}(s) > x \,|\, Y_n = s, Y_{n+1} = t)$$

$$= \sum_{j\in\mathcal{K}, s\in\mathcal{S}} p_{st}\,\pi_n(s) \left[\int_{-\infty}^{x} P(X_{n,j} > x - u \,|\, Y_n = s)\, dF_{jk,s}(u) + 1 - F_{jk,s}(x)\right]$$

$$\leq \sum_{j\in\mathcal{K}, s\in\mathcal{S}} p_{st} \left[\int_{-\infty}^{x} \gamma_{j,s}(x-u)\, dF_{jk,s}(u) + \pi_n(s)(1 - F_{jk,s}(x))\right] \quad (15)$$

$$\leq \gamma_{k,t}(x)$$

where (15) follows from the induction hypothesis and the latter inequality follows from the definition of the functions $\gamma_{j,s}$ given in (13). This concludes the proof of (14).

Before getting to the second step of the proof, let us first show that $b(\theta) < \infty$ if $y(\theta) > 0$. Indeed, we have in this case

$$b(\theta) \leq \sup_{\substack{x\geq 0\\ n\geq 0\\ k\in\mathcal{K}\\ t\in\mathcal{S}}} \frac{\sum_{s\in\mathcal{S}, j\in\mathcal{K}} p_{st}\,\pi_n(s)\,(1 - F_{jk,s}(x))}{\sum_{s\in\mathcal{S}, j\in\mathcal{K}} p_{st}\, y_{j,s}(\theta)\,(1 - F_{jk,s}(x)} \quad (16)$$

$$\leq \sup_{\substack{n\geq 0\\ j\in\mathcal{K}\\ s\in\mathcal{S}}} \frac{\pi_n(s)}{y_{j,s}(\theta)}$$

the latter quantity being finite if $y(\theta) > 0$ since both sets $\mathcal{K}$ and $\mathcal{S}$ are finite (we used the inequality $\int_x^\infty \exp(\theta(u - x))\, dF_{jk,s}(u) \geq 1 - F_{jk,s}(x)$ to derive (16)).

The second step of the proof consists in checking that the functions $\{\gamma_{j,s} = b(\theta)\, y_{j,s}(\theta)\, \exp(-\theta x),\ j \in \mathcal{K}, s \in \mathcal{S}\}$ satisfy (13). This proof is analogous to the proof of Proposition 2 in [18] and is therefore omitted. Let us simply point out for later use that this proof only uses the property that $y(\theta)\,\mathbf{H}(\theta) \leq y(\theta)$ if $\rho(\theta) \leq 1$ (which follows from (8)) and does not require the identity (8).

Substituting now $\gamma_{k,t}(x)$ in (14) for $b(\theta)\, y_{k,t}(\theta)\, \exp(-\theta x)$, then summing up both sides of the inequality over all the values t in $\mathcal{S}$ and using the normalizing condition (9) yields (11). ∎

The condition that $y(\theta) > 0$ in Proposition 1 may be dropped in the case that $\rho(\theta) < 1$. More precisely, we have the following result:

Proposition 2 (Exponential Upper Bound II) *Assume that $\theta \in \Theta$ with $\rho(\theta) < 1$. Then, there exist strictly positive constants $a_1, \ldots, a_g$ such that the positive vector $v(\theta) = (v_{k,t}(\theta))_{k\in\mathcal{K}, t\in\mathcal{S}}$ defined as $v(\theta) = (a_1 z_1(\theta), \ldots, a_g z_g(\theta))$ satisfies $v(\theta)\,\mathbf{H}(\theta) \leq v(\theta)$ and $\sum_{k\in\mathcal{K}, t\in\mathcal{S}} v_{k,t}(\theta) = 1$.*

Furthermore, if for all $k \in \mathcal{K}$, $t \in \mathcal{S}$,

$$P(X_{0,k} > x, Y_0 = t) \leq c(\theta)\, v_{k,t}(\theta)\, e^{-\theta x}, \qquad \forall x \geq 0 \quad (17)$$

then, for all $n \geq 1$, $k \in \mathcal{K}$, $x \geq 0$,

$$P(X_{n,k} > x) \leq c(\theta)\, e^{-\theta x} \quad (18)$$

where

$$c(\theta) = \sup_{\substack{x\geq 0\\ n\geq 0\\ k\in\mathcal{K}\\ t\in\mathcal{S}}} \frac{\sum_{j\in\mathcal{K},\, s\in\mathcal{S}} p_{st}\,\pi_n(s)\,(1 - F_{jk,s}(x))}{\sum_{j\in\mathcal{K},\, s\in\mathcal{S}} p_{st}\, v_{j,s}(\theta) \int_x^\infty e^{\theta(u-x)}\, dF_{jk,s}(u)} < \infty. \quad (19)$$

◇

Proof. Assume first the existence of the vector $v(\theta)$. Then, the proof of (18) is identical to the proof of (11) after substituting $y(\theta)$ for $v(\theta)$. This follows from the observation (cf. step 2 in the proof of Proposition 1) that only the inequality $y(\theta)\,\mathbf{H}(\theta) \leq y(\theta)$ and the fact that $y(\theta) > 0$ are used in the proof of (11).

Let us now address the existence of constants $a_1, \ldots, a_g$ such that $v(\theta)\,\mathbf{H}(\theta) \leq v(\theta)$. By using the definition of the eigenvectors $z_i(\theta)$, $1 \leq i \leq g$, it is seen that the vector inequality $v(\theta)\,\mathbf{H}(\theta) \leq v(\theta)$ translates into the set of inequalities

$$\sum_{i=1}^{j-1} a_i\, z_i(\theta)\, J_{i,j}(\theta) + a_j\,(\rho_j(\theta) - 1)\, z_j(\theta) \leq 0, \qquad \text{for } 1 \leq j \leq g. \quad (20)$$

Since $\rho_i(\theta) < 1$ for all $i = 1, 2, \ldots, g$ under the assumption that $\rho(\theta) < 1$ (cf. (7)), the following simple procedure can be derived from (20) to generate positive constants $a_1, \ldots, a_g$: pick an arbitrary $a_1 > 0$, then pick $a_2 > 0, \ldots, a_g > 0$ successively so that

$$a_j \geq \max_l \left\{ \frac{\sum_{i=1}^{j-1} a_i\,[z_i(\theta) J_{i,j}(\theta)]_l}{(1 - \rho_j(\theta))\,[z_j(\theta)]_l} \right\}, \qquad \text{for } 2 \leq j \leq g$$

where $[v]_l$ denotes the l-th component of any vector v. The proof is concluded by normalizing $a_1, \ldots, a_g$ so that the components of $v(\theta)$ sum up to 1. ∎

3 Application to Tandem Queues

Consider an open tandem queueing network consisting of K single-server queues with infinite buffers. We assume that customers may only enter the network from the outside at node 1 and that they all leave the network upon completion of their service time at node K (see Fig. 1). Let $(Z_n)_n$ be an irreducible, aperiodic and homogeneous Markov chain on a finite set $\mathcal{T} = \{1, 2, \ldots, T\}$, with transition matrix $\mathbf{Q} = [q_{ij}]$. Let $\sigma_{n,k}(Z_n) < \infty$ be the service time required by customer n at node k ($1 \leq k \leq K$) and let $\tau_n(Z_n)$ be the interarrival time in queue 1 between customers n and $n+1$. In other words, we consider a system where the interarrival and service times are modulated by the Markov chain $(Z_n)_n$. Define $X_{n,k}$ as the cumulated waiting time (excluding the service times) of customer n through its sojourn in queues $1, 2, \ldots, k$.

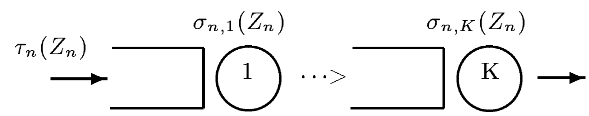

Figure 1: Tandem queues

Define

$$U_{n,j}(z_1, z_2) = \sum_{i=1}^{j} \sigma_{n,i}(z_1) - \sum_{i=1}^{j-1} \sigma_{n+1,i}(z_2) - \tau_n(z_1).$$

It is easy to check that

$$X_{n+1,k} = \max\left(X_{n,k} + U_{n,k}(Z_n, Z_{n+1}), X_{n+1,k-1}\right)^+ \quad (21)$$

$$= \max_{1 \leq j \leq k} \left(X_{n,j} + U_{n,j}(Z_n, Z_{n+1})\right)^+ \quad (22)$$

for all $n \geq 0$ and $k \in \mathcal{K}$ (by convention $X_{n+1,0} = 0$ in (21)).

The stochastic recursion (22) can be written in the form (1) by defining the Markov chain $(Y_n)_n$ and the matrix $\mathbf{A}_n(s)$ as follows:

$$Y_n = (Z_n, Z_{n+1}) \quad (23)$$

$$A_{n,jk}(s) = \begin{cases} U_{n,j}(z_1, z_2) & \text{for } 1 \leq j \leq k \\ -\infty & \text{for } k < j \leq K \end{cases} \quad (24)$$

with $s \in \mathcal{S} = \{(i,j) \,|\, q_{ij} > 0, i, j \in \mathcal{T}\}$.

In this case, $\mathbf{B}_{n,k}(s)$ and $\mathbf{R}_{n,k,j}$ in (3) are all 1-by-1 matrices given by $\mathbf{B}_{n,k}(s) = U_{n,k}(z_1, z_2)$ for $k \in \mathcal{K}$ and $\mathbf{R}_{n,k,j}(s) = \mathbf{B}_{n,k}(s)$ for $j > k$, for all $s = (z_1, z_2)$, respectively. In particular, the matrix $\mathbf{J}_i(\theta)$ ($i = 1, 2, \ldots, K$) is given by

$$\mathbf{J}_i(\theta) = \left[E[\exp(\theta\, U_{n,i}(z_1, z_2))]\, p_{st}\right]_{s=(z_1,z_2)\in\mathcal{S}, t=(z_3,z_4)\in\mathcal{S}}$$

with $p_{st} = q_{z_2 z_4}$ if $z_2 = z_3$ and $p_{st} = 0$ if $z_2 \neq z_3$, for all $s = (z_1, z_2) \in \mathcal{S}$, $t = (z_3, z_4) \in \mathcal{S}$.

In order to apply the results obtained in Section 2 to the components of the vector X_n we must ensure that Assumptions (A1)-(A3) are satisfied. These assumptions will hold, in particular, if the service requirements are deterministic (but not necessarily all equal), namely, for every fixed $k \in \mathcal{K}$, $z \in \mathcal{T}$, $\sigma_{n,k}(z)$ is constant for all $n \geq 0$, if $(\tau_n(1))_n, \ldots, (\tau_n(T))_n$ are mutually independent renewal sequences, further independent of the Markov chain $(Z_n)_n$, and if $\tau_n(z)$ has a phase-type distribution for each $z \in \mathcal{T}$ (e.g., the arrival process is a Markov modulated Poisson process). In this case, Propositions 1 and 2 give exponential upper bounds on the tail distribution of the cumulated backlogs.

In particular, the unique positive solution θ^* of the equation $\rho(\theta) = 1$ gives the best exponential decay for $P(X_{n,k} > x)$ if $y(\theta^*)$ is positive.

It is worth noting that the sequences of matrices $((\mathbf{A}_n(s))_n)_{s\in\mathcal{S}}$ are not mutually independent under only the assumption that the service and interarrival times are all mutually independent r.v.'s. This follows from the fact that matrices $\mathbf{A}_n(s)$ and $\mathbf{A}_{n+1}(t)$ both depend on the r.v.'s $\sigma_{n+1,1}(z_2), \ldots, \sigma_{n+1,K-1}(z_2)$ in the case that $s = (z_1, z_2) \in \mathcal{S}$ and $t = (z_2, z_3) \in \mathcal{S}$. Furthermore, if $q_{zz} > 0$, then the matrices $\mathbf{A}_1(s), \mathbf{A}_2(s), \ldots$ are not mutually independent for $s = (z, z)$ for the same reason. This implies, in particular, that Propositions 1 and 2 do not apply to M/M/1 queues in series.

Note that certain forms of non-deterministic service times are permitted by augmenting the state space of the original Markov chain $(Z_n)_n$.

4 Concluding Remarks

In this paper, we have extended our work for single queues [18, 19] to stochastic linear systems under the max-plus algebra, as defined by (2), and have derived exponential upper bounds on $P(X_{n,k} > x)$. The max-plus structure defined in (2) appears to be a particular case of the structure considered by Chang in [4]. We believe that our method based on an extension of Kingman's method for bounding the tail of the waiting time distribution in a GI/GI/1 queue, gives sharper upper bounds than the corresponding bounds in [4], based on Chernoff's bound. We also point out that the i.i.d. assumption placed on the input sequence in [4] is stronger than ours. Our approach will also allow us to derive exponential lower bounds for the tail distribution of $X_{n,k}$. The latter results will be reported in a forthcoming paper.

References

[1] J. Abate, G. L. Choudhury and W. Whitt, "Asymptotics for Steady-State Tail Probabilities in Structured Markov Queueing Models", *Commun. Statist. Stoch. Models*, Vol. 10, No. 1, pp. 99 – 143, 1994.

[2] F. Baccelli and Z. Liu, "On a Class of Stochastic Recursive Sequences Arising in Queueing Theory", *The Annals of Probability*, Vol. 20, No. 1, pp. 350 – 374, 1992.

[3] C.-S. Chang, "Stability, Queue Length and Delay of Deterministic and Stochastic Queueing Networks", *IEEE*

Trans. Aut. Contr., Vol. 39, No. 5, pp. 913 – 931, May 1994.

[4] C.-S. Chang, "On the Exponentiality of Stochastic Linear Systems under the Max-Plus Algebra". Preprint, Jul. 1994. Submitted to *IEEE Trans. Aut. Control*.

[5] C.-S. Chang and J. Cheng, "Computable Exponential Bounds for Intree Networks with Routing". Preprint, Jul. 1994.

[6] R. L. Cruz, "A Calculus for Network Delay, Part I: Network Elements in Isolation", *IEEE Trans. Inf. Theory*, Vol. 37, No. 1, pp. 114 – 131, Jan. 1991.

[7] R. L. Cruz, "A Calculus for Network Delay, Part II: Network Analysis", *IEEE Trans. Inf. Theory*, Vol. 37, No. 1, pp. 132 – 141, Jan. 1991.

[8] R. L. Cruz and H.-L. Liu, "Non-Recursive Identities for Tandem Queueing Networks". Preprint, Mar. 1993.

[9] N. G. Duffield, "Exponential Bounds for Queues with Markovian Arrivals", Technical Report DIAS-APG-93-01, 1993.

[10] A. I. Elwalid and D. Mitra, "Effective Bandwidth of General Markovian Traffic Sources and Admission Control of High Speed Networks", *IEEE/ACM Trans. on Networking*, Vol 1, No. 3, pp. 329 – 343, Jun. 1993.

[11] R. A. Horn and C. R. Johnson, *Matrix Analysis*. Cambridge University Press, 1985.

[12] *IEEE J. in Selected Areas in Commun.*, Special Issue on "Advances in the Fundamentals of Networking - Part I: Bridging Fundamental Theory and Networking", Vol. 13, No. 6, Guest Eds. R. G. Gallager, B. Hajek, F. P. Kelly, D. Mitra and P. P. Varaiya, Aug. 1995.

[13] A. Jean-Marie, S. Lefebvre-Barbaroux and C. Chaouiya, "Real-Time Scheduling of Periodic Tasks", *Scheduling Theory and its Applications*, Ph. Chrétienne, E. G. Coffman, Jr., J. K. Lenstra and Z. Liu, Eds, J. Wiley & Sons, 1995.

[14] F. P. Kelly, "Effective Bandwidths at Multi-Class Queues" *QUESTA*, Vol. 9, pp. 5 – 15, 1991.

[15] G. Kesidis, J. Walrand and C.-S. Chang, "Effective Bandwidths for Multiclass Markov Fluids and Other ATM Sources", *IEEE/ACM Trans. Networking*, Vol. 1, No. 4, pp. 424 – 428, Aug. 1993.

[16] J. F. C. Kingman, "Inequalities in the Theory of Queues", *J. Roy. Stat. Soc.*, Series B, Vol. 32, pp. 102 – 110, 1970.

[17] J. F. Kurose, "On Computing per-Session Performance Bounds in High-Speed Multi-Hop Computer Networks", *Proc. ACM SIGMETRICS and PERFORMANCE'92*, Newport, RI, pp. 128 – 139, Jun. 1992.

[18] Z. Liu, P. Nain, and D. Towsley, "Exponential Bounds for a Class of Stochastic Processes with Application to Call Admission Control in Networks", in *Proc. of the 33rd Conf. on Decision and Control*, pp. 156 – 162, Lake Buena Vista, FL, Dec. 1994.

[19] Z. Liu, P. Nain, and D. Towsley, "Exponential Bounds with an Application to Call Admission". Technical Report 94-63, Dept. of Computer Science, University of Massachusetts, Oct. 1994.

[20] A. K. Parekh and R. G. Gallager, "A Generalized Processor Sharing Approach to Flow Control in Integrated Service Networks: The Single Node Case", *IEEE/ACM Trans. on Net.*, Vol. 1, No. 3, pp. 344 – 357, Jun. 1993.

[21] A. K. Parekh and R. G. Gallager, "A Generalized Processor Sharing Approach to Flow Control in Integrated Service Networks: The Multiple Node Case", *IEEE/ACM Trans. on Net.*, Vol. 2, No. 2, pp. 137 – 150, Apr. 1994.

[22] G. de Veciana, C. Courcoubetis and J. Walrand, "Decoupling Bandwidths for Networks: A Decomposition Approach to Resource Management". UCB/ERL Technical Report M93/50, University of California at Berkeley, Jun. 1993.

[23] O. Yaron and M. Sidi, "Performance and Stability of Communication Networks Via Robust Exponential Bounds", *IEEE/ACM Trans. Networking*, Vol. 1, No. 3, pp. 372 – 385, Jun. 1993.

[24] Z.-L. Zhang, D. Towsley and J. Kurose, "Statistical Analysis of the Generalized Processor Sharing Scheduling Discipline", *IEEE J. in Selected Areas in Commun.*, Vol. 13, No. 6, pp. 1071 – 1080, Aug. 1995.