SRC TR 89-30

Concavity of Throughput in Series of Queues with Finite Buffers

By

V. Anantharam & P. Tsoucas

Concavity of Throughput in Series of Queues with Finite Buffers

by

Venkat Anantharam

School of Electrical Engineering
Cornell University
Ithaca, NY 14853

and

Pantelis Tsoucas

Systems Research Center
University of Maryland
College Park, MD 20742

ABSTRACT

Concavity of the output process with respect to buffer size is established in a series of $\cdot/M/1/B$ queues with loss at the first node. Similarly, one shows concavity of the throughput with respect to the number of servers and the buffer sizes in a node belonging to a series of $\cdot/M/s/B$ queues.

Keywords: Blocking; concavity; finite capacity queues; monotonicity; throughput.

1. Introduction

This paper is concerned with a model of a production line that has received much attention in the literature. Consider $K \cdot/M/1/N$ queues in series and denote the collection by $\{\cdot/M/1/B_i\}_{i=1}^{K}$. For convenience, the numbering of the nodes is in the reverse order of service. This notation means that queue i has one server with i.i.d. exponential service times and waiting room of size B_i. Let the service rate in node i be denoted by μ_i and consider a deterministic sequence of arrivals (a_n). An arriving job that finds buffer K full is lost. The server in node i idles whenever node $i-1$ is full. This discipline is called "communication blocking". Assume that the system is initially empty and denote by $(D^0(t))_{t\geq 0}$ the departure process from the first node.

The aim of this paper is to show that $(D^0(t))$ is stochastically concave with respect to each buffer size. Denote by $(D^1(t))$ (respectively $(D^2(t))$) the departure process from the first node when one (respectively 2) space(s) have been added to the buffer of the kth node $k=1,\ldots,K$. The three corresponding networks will be denoted by $\mathcal{N}^0$, $\mathcal{N}^1$ and $\mathcal{N}^2$ respectively. Note that the dependence on k is suppressed in this notation. We show that

$$2D^1(t) \geq_{st} D^0(t) + D^2(t). \tag{1}$$

(Recall that $X \geq_{st} Y$ if $P\{X \geq x\} \geq P\{Y \geq x\}$ for all $x \in \mathrm{R}$.)

To our knowledge, only the simple case where $K=1$, with a Poisson arrival process, and when the system is stationary, has been treated analytically. Our proofs rely on straightforward sample path arguments. Further extensions of the results are discussed in Remark 2.1.

Similar arguments have been used to establish various monotonicity results (cf. [3]). In particular, monotonicity properties for the model examined here are studied in [2]. Second order properties of networks have important algorithmic implications for problems of optimal allocation (see, e.g., [1]).

2. The main result

The proof of the result consists in constructing the three processes corresponding to the three different systems referred to above, such that (1) holds almost surely.

Recall that the virtual service process of an exponential server with rate μ is a Poisson process with rate μ. When the queue is non-empty, a customer departs at each point of the virtual service process.

For network $\mathcal{N}^0$ and for $i=1,2$ the following quantities are defined.

$S_n^{0,i}$: The nth service time in the virtual service process at node i.

$N_i^0(t)$: The number of jobs in the $i$the queue at time t.

We will consider the network at the discrete time instants

$$(T_n)_n = \{(S_n^{0,i})_n\}_{i=1,2} \cup (a_n)_n, \qquad T_n = 0.$$

Also, we will find useful to set

$$\begin{aligned} X_0^0(t) &= D^0(t), \\ X_k^0(t) &= D^0(t) + \sum_{i=1}^{k} N_i^0(t), \qquad k=1,\ldots,K. \end{aligned}$$

Similar processes are defined for networks $\mathcal{N}^1$ and $\mathcal{N}^2$.

Queueing processes in networks $\mathcal{N}^1$ and $\mathcal{N}^2$ are constructed by setting

$$S_n^{2,i} = S_n^{1,i} = S_n^{0,i} \equiv S_n^i, \qquad n=1,2,\ldots, k=1,2, \tag{2}$$

$$a_n^2 = a_n^1 = a_n^0 \equiv a_n, \qquad n = 1, 2, \ldots. \tag{3}$$

It is assumed that networks $\mathcal{N}^1$ and $\mathcal{N}^2$ start empty at time 0.

Use will be made of the following lemma. It can be proved by straightforward induction.

Lemma 2.1: In the construction of (2) and (3) one has, for $t \geq 0$,

$$X_i^2(t) \geq X_i^1(t) \geq X_i^0(t), \qquad \text{a.s.}, \quad i = 0, \ldots, K.$$

The validity of (1) is a corollary of the following.

Theorem 2.1: In the construction of (2) and (3) one has, for $t \geq 0$,

$$X_i^2(t) - X_i^1(t) \leq X_i^1(t) - X_i^0(t), \qquad \text{a.s.}, \quad i = 0, 1, \ldots, K.$$

Proof: Set $\Delta_i^2(t) = X_i^2(t) - X_i^1(t)$ and $\Delta_i^1(t) = X_i^1(t) - X_i^0(t)$, and assume that the additional buffers are at the kth queue, $k = 1, \ldots, K$. It will shown that

$$\Delta_i^2(T_l) \leq \Delta_i^1(T_l), \qquad k = 1, 2, \ldots \text{a.s.} \tag{4}$$

This is trivially true for $l = 0$ and assume it holds for $l = 1, \ldots, n$. For some $m > 0$ we distinguish between the following cases.

(*i*) $T_{n+1} = S_m^1$. Inequalities (4) can be violated at T_{n+1} only for $i = 0$. One must then have $\Delta_0^2(T_n) = \Delta_0^1(T_n)$ and $N_1^2(T_n) > 0$, $N_1^1(T_n) = 0$. But this implies that $\Delta_1^2(T_n) > \Delta_1^1(T_n)$, a contradiction.

(*ii*) $T_{n+1} = S_m^l$, $l \neq k+1$. Inequalities (4) can be violated at T_{n+1} only for $i = l-1$. Then, necessarily, $\Delta_{l-1}^2(T_n) = \Delta_{l-1}^1(T_n)$. If $N_{l-1}^2(T_n) > 0$ and $N_{l-1}^1(T_n) = 0$, then one has $\Delta_l^2(T_n) > \Delta_l^1(T_n)$, a contradiction. If $N_{l-1}^2(T_n) < B_{l-1}$ and $N_{l-1}^1(T_n) = B_{l-1}$, then one concludes that $\Delta_{l-1}^2(T_n) > \Delta_{l-1}^1(T_n)$, again a contradiction.

(*iii*) $T_{n+1} = S_m^{k+1}$ ($= a_m$ if $k = K$). Inequalities (4) can be violated at T_{n+1} only for $i = k$ and if $\Delta_k^2(T_n) = \Delta_k^1(T_n)$. The possibility $N_{k+1}^2(T_n) > 0$ and $N_{k+1}^1(T_n) = 0$ leads to a contradiction as in the case above, and if $N_k^2(T_n) \leq B_k + 1$ and $N_k^1(T_n) = B_k + 1$, then one obtains the contradiction $\Delta_{k-1}^2(T_n) > \Delta_{k-1}^1(T_n)$, since $N_k^0(T_n) \leq B_k$.

(*iv*) $T_{n+1} = a_m$. Inequalities (4) can be violated at T_{n+1} only for $i = K$ and if $\Delta_K^2(T_n) = \Delta_K^1(T_n)$, $N_K^2(T_n) \leq B_K + 1$, and $N_K^1(T_n) = B_K + 1$. This implies a contradiction as in the case above.

We have proved that inequalities (4) hold for $l = n+1$.

□

Remark 2.1. Similarly, one shows concavity of the output process with respect to the number of servers and the buffer sizes in a node belonging to a series of $\cdot/M/s/B$ queues. Other types of blocking are also possible (e.g., manufacturing blocking).

Remark 2.2. Our technique fails when Bernoulli feedback is added to the series of $\{\cdot/M/1/B_i\}_{i=1}^K$ queues.

3. References

[1] Shantikumar, G. and Yao, D. (1987). Optimal server allocation in a system of multi-server stations. To appear in *Management Science.*

[2] Tsoucas, P. and Walrand, J. (1988). Monotonicity of throughput in non-Markovian networks. To appear in *Journal of Applied Probability* , March 1989.

[3] van der Wal, J. (1985). Monotonicity of the throughput of a closed exponential queueing network in the number of jobs. Eidhoven University of Technology Memo. COSOR 85-21.