Stochastic Concavity of Throughput in Series of Queues with Finite Buffers
Author(s): Venkat Anantharam and Pantelis Tsoucas
Source: *Advances in Applied Probability*, Vol. 22, No. 3 (Sep., 1990), pp. 761-763
Published by: Applied Probability Trust
Stable URL: http://www.jstor.org/stable/1427471
Accessed: 06/05/2010 22:03

STOCHASTIC CONCAVITY OF THROUGHPUT IN SERIES OF QUEUES WITH FINITE BUFFERS

VENKAT ANANTHARAM,* *Cornell University*
PANTELIS TSOUCAS,** *IBM Research Division*

Abstract

Stochastic concavity of the output process with respect to buffer sizes is established in a series of $\cdot/M/1/B$ queues with loss at the first node.

BLOCKING; FINITE CAPACITY QUEUES; MONOTONICITY

1. Introduction

This note is concerned with a model that has received much attention in the literature. Consider $K \cdot/M/1/B$ queues in series and denote the collection by $\mathcal{N} = \{\cdot/M/1/B_i\}_{i=1}^{K}$. 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 $(D(t))_{t \geq 0}$ the departure process from the first node.

The aim is to show that $(D(t))$ is stochastically concave with respect to the vector of buffer sizes $B = (B_1, \cdots, B_K)$. This formulated with the aid of the following definition.

Definition 1. A function $\rho : Z_+^K \to R$ is concave if for all $B \in Z_+^K$, for all $b = (b_1, \cdots, b_K) \in Z^K$ such that $|b_1|, \cdots, |b_K|$, are relatively prime, and for all $l, m \in Z_+$ such that $B + mB$ and $B - lb$ are in Z_+^K,

$$(1) \qquad (l+m)\rho(B) \geq l\rho(B+mb) + m\rho(B-lb).$$

The next result restricts the range on which (1) has to be verified.

Lemma 1. Inequalities (1) hold if for all $b \in Z^K$ as described in Definition 1, and for all $B \in Z_+^K$,

$$2\rho(B) \geq \rho(B+b) + \rho(B-b).$$

Proof. By straightforward induction on l and m.

For a vector $b \in Z^K$ as required in Definition 1, denote by $(D^+(t))$ (respectively, $(D^-(t))$) the departure process from the first node of a network identical to $\mathcal{N}$ but with buffer size vector $B + b$ (respectively, $B - b$). Denote that network by $\mathcal{N}^+$ (respectively, $\mathcal{N}^-$). We show that

$$(2) \qquad 2D(t) \geq_{st} D^+(t) + D^-(t).$$

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

The proof relies on a straightforward sample path argument which by now is routinely used in the literature in order to establish various monotonicity properties. In particular, monotonicity properties of the model examined here are studied in [4]. A version of this note

appeared as [1]. In the originally submitted version of this note the construction of Theorem 1 was used to prove concavity of the throughput only with respect to each buffer size. It has since been pointed out to us that in the same construction and with a small modification of the original proof, the throughput is seen to be concave with respect to the vector of buffer sizes in the sense of Definition 1. Meester and Shantikumar [2] have independently obtained this stronger result. They employ the same construction but they derive an expression for the number of jobs that have departed from the network as a function of time. They also provide an application of the result by proving one of the many 'folk' conjectures in optimal buffer allocation.

Second-order properties of networks have important algorithmic implications for problems of optimal allocation. See Shantikumar and Yao [3] and references therein for applications and other work on stochastic convexity. Our results stop short of a characterization of the optimal buffer allocation in a series of $\cdot/M/1/B$ queues. Further research is suggested in Section 3.

2. The main result

The proof consists in constructing the three processes corresponding to the three networks $\mathcal{N}$, $\mathcal{N}^+$ and $\mathcal{N}^-$, so that (2) 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 not empty, a customer departs at each point of the virtual service process. For network $\mathcal{N}$ and for $i = 1, \cdots, K$ the following quantities are defined:

S_n^i: the nth service time in the virtual service process at node i,

$N_i(t)$: the number of jobs in the ith queue at time t.

We consider the network at the discrete time instants

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

Also, we shall find it useful to set

$$X_0(t) = D(t),$$

$$X_k(t) = D(t) + \sum_{i=1}^{k} N_i(t), \qquad k = 1, \cdots, K.$$

Similar processes are defined for networks $\mathcal{N}^+$ and $\mathcal{N}^-$. Queueing processes in networks $\mathcal{N}^+$ and $\mathcal{N}^-$ are constructed by setting

(3) $$S_n^{+,i} = S_n^{-,i} = S_n^i, \qquad n = 1, 2, \cdots, i = 1, \cdots K,$$

(4) $$a_n^+ = a_n^- = a_n, \qquad n = 1, 2, \cdots$$

It is assumed that networks $\mathcal{N}^+$ and $\mathcal{N}^-$ also start empty at time 0.

The validity of (2) is a corollary of the following.

Theorem 1. In the construction of (3) and (4) one has, for $t \geqq 0$,

(5) $$2X_i(t) \geqq X_i^+(t) + X_i^-(t), \qquad \text{a.s.}, \quad i = 0, 1, \cdots, K.$$

Proof. Inequalities (5) are trivially true for $l = 0$ and assume that they hold for $l = 1, \cdots, n$. We prove that they hold for $l = n + 1$. For some $m > 0$ we distinguish between the following cases.

(a) $T_{n+1} = S_m^1$. Inequalities (5) can be violated at T_{n+1} only for $i = 0$. There are two possibilities.

(i) $2X_0(T_n) = X_0^+(T_n) + X_0^-(T_n)$. Then (5) can be violated only if

(6) $$N_1(T_n) = 0 \quad \text{and either} \quad N_1^+(T_n) > 0 \quad \text{or} \quad N_1^-(T_n) > 0.$$

But this implies $2X_1(T_n) < X_1^+(T_n) + X_1^-(T_n)$, a contradiction.

(ii) $2X_0(T_n) = X_0^+(T_n) + X_0^-(T_n) + 1$. This possibility leads to a similar contradiction.

(b) $T_{n+1} = S_m^{k+1}$. Inequalities (5) can be violated at T_{n+1} only for $i = k$. There are two possibilities.

(i) $2X_k(T_n) = X_k^+(T_n) + X_k^-(T_n)$. Then (5) can be violated in either of two situations. First, if

(7) $\qquad N_k(T_n) = B_k$ and either $N_k^+(T_n) < B_k + b_k$ or $N_k^-(T_n) < B_k - b_k$,

then $N_k^+(T_n) + N_k^-(T_n) < 2B_k$ and hence $2X_{k-1}(T_n) < X_{k-1}^+(T_n) + X_{k-1}^-(T_n)$, a contradiction. Second, the situation where

(8) $\qquad N_{k+1}(T_n) = 0$ and either $N_{k+1}^+(T_n) > 0$ or $N_{k+1}^-(T_n) > 0$,

obtains a contradiction as in case (a).

(ii) $2X_k(T_n) = X_k^+(T_n) + X_k^-(T_n) + 1$. This possibility leads to a similar contradiction.

(c) $T_{n+1} = a_m$. This case is similar to case (b).

The only modification necessary in the proof of cases (a)(ii) and (b)(ii) is the replacement of 'either $\cdots$ or' stagements in (7), (8) and (9) by 'both $\cdots$ and' statements.

Remark 1. Other types of blocking are also possible (e.g., manufacturing blocking).

3. Conclusions

If the throughput of the network considered above is denoted by $\rho(B)$ then we have shown that $\rho(\cdot)$ is concave with respect to vector B in the sense of Definition 1. It follows in particular that $\rho(\cdot)$ is a concave function on the line $\sum_{i=1}^K B_i = M$ for any integer M. Set

$$\rho^*(M) = \max\left\{\rho(B) : \sum_{i=1}^K B_i = M\right\}.$$

By B_M^* denote the allocation that achieves the optimal throughput. One expects (Mitra, personal communication, 1987) that

$$\rho^*(M + 1) = \rho(B_M^* + e_i), \quad \text{for some} \quad i = 1, \cdots, K,$$

where e_i is the unit vector corresponding to the ith coordinate. Establishing this property should have important algorithmic consequences.

Acknowledgement

References

[1] ANANTHARAM, V. AND TSOUCAS, P. (1989) Concavity of throughput in series of queues with finite buffers. SRC Tech. Rep. # 89-30, University of Maryland, College Park.

[2] MEESTER, L. E. AND SHANTIKUMAR, J. G. (1990) Concavity of the throughput tandem queueing systems with finite buffer storage space. *Adv. Appl. Prob.* **22**, 764–767.

[3] SHANTHIKUMAR, J. G. AND YAO, D. D. (1989)) Second-order stochastic properties in queueing systems. *Proc. IEEE* **77**, 162–170

[4] TSOUCAS, P. AND WALRAND, J. (1989) Monotonicity of throughput in non-Markovian networks. *J. Appl. Prob.* **26**, 134–141.