To appear in *Journal of Applied Probability*, **35**, *1998*

A MONOTONICITY RESULT FOR THE WORK-LOAD IN MARKOV-MODULATED QUEUES

Nicole Bäuerle
Department of Mathematics VII
University of Ulm
D-89069 Ulm, Germany

Tomasz Rolski[1]
Mathematical Institute
University of Wroclaw
50384 Wroclaw, Poland

Abstract

We consider a single server queue where the arrival process is a Markov-modulated Poisson process and service times are independent and identically distributed and independent from arrivals. The underlying intensity process is assumed ergodic with generator cQ, $c > 0$. We prove under some monotonicity assumptions on Q that the stationary work-load $W(c)$ is decreasing in c with respect to the increasing convex ordering.

Keywords: Markov-modulated queue, stationary work-load, increasing convex ordering, supermodular function.

AMS 1991 subject classification: primary 60K25

[1] Work supported in part by KBN under grant 1026/803/95/08.

1. Introduction

In this paper we consider a single server queue with FCFS queueing discipline. The arrival process is a Markov-modulated Poisson process whose intensity process $\{\lambda(t), t \geq 0\}$ is a Markov chain with a finite state space and generator cQ, $c > 0$. The service times are supposed to be independent and identically distributed, independent from arrivals. We are interested whether changing the parameter c has an effect on the stationary work-load of the queue. Since the stimulating paper of Ross [9] there have been several articles dealing with this topic, like for example Rolski [7], [8], Chang, Chao and Pinedo [5] in continuous time and Szekli [12], Bäuerle [2] in discrete time.

In the paper we consider two Markov-modulated queues with the same service process. We will prove, under the assumption of stochastic monotonicity of the generator Q and the time-reverse generator Q^* that if Q and $Q' = cQ$ with $c \in (0, 1]$ are underlying generators, then the corresponding stationary work-loads fulfil $W \leq_{\text{icx}} W'$. This theorem generalizes results obtained for Markov-modulated queues by Chang, Chao and Pinedo [5] who used a special generator to study this effect.

The paper is organized as follows. Section 2 contains the model formulation and Section 3 the main theorem. Proofs are presented in Section 4.

2. Preliminaries

The model under consideration is defined in the following way. Denote by $\boldsymbol{N}$ the point process of arrivals to the queue, i.e. $N(t) = N(0, t]$ is the number of customers who arrived before time t. We assume that $\{N(t), t \geq 0\}$ is a Markov-modulated Poisson process. That is for mutually disjoint intervals $I_1, \ldots, I_k$

$$P(N(I_1) = j_1, \ldots, N(I_k) = j_k) = E\left[\prod_{l=1}^{k} \frac{(\Lambda(I_l))^{j_l}}{j_l!} e^{-\Lambda(I_l)}\right],$$

where $\Lambda(B) = \int_B \lambda(t)\, dt$ for Borel set B denotes the random intensity measure of $\{N(t), t \geq 0\}$ defined by the random intensity function $\{\lambda(t), t \geq 0\}$, which we assume to be an ergodic Markov chain with states $\lambda_1 \leq \ldots \leq \lambda_m$ and generator Q. By π we denote the unique stationary distribution of $\{\lambda(t), t \geq 0\}$, i.e. $\pi Q = 0$ and by $Q^* = (q^*_{ij})$ we denote the generator of the time reverse process, where $q^*_{ij} = \frac{\pi_j}{\pi_i} q_{ji}$. The sequence of service times $\{S_n\}$ consists of independent and identically distributed random variables as the generic service time S, with finite second moment and independent from arrivals. For stability we assume that the traffic intensity $\rho = E(S) \sum_{i=1}^m \pi_i \lambda_i < 1$. Furthermore we denote $M(t) = \sum_{i=1}^{N(t)} S_i$, $t \geq 0$. Hence the work-load process $\{V(t), t \geq 0\}$ of the queue fulfils

$$V(t) = M(t) - t + \int_0^t 1(V(s) = 0) ds, \quad t \geq 0.$$

Therefore, we obtain $\lim_{t \to \infty} P(V(t) > x) = P(W > x)$, $x > 0$, where

$$W = \sup_{t \geq 0} (M(t) - t)$$

denotes the stationary work-load (see e.g. Borovkov [4]).

3. The result

To state the main result in Theorem 1 we need the following definition (see e.g. Shaked and Shanthikumar [10] or Stoyan [11]), where "increasing" is used in the nonstrict sense throughout this paper.

Definition 1

a) *Two real-valued random variables X and Y are ordered with respect to the* **stochastic ordering** *($X \leq_{\mathrm{st}} Y$) if*

$$E[f(X)] \leq E[f(Y)], \tag{1}$$

for all increasing functions $f : \mathbb{R} \to \mathbb{R}$ for which the expectations exist. In this case we will also write for the corresponding probability distributions that $P^X \leq_{\mathrm{st}} P^Y$.

b) *Two real-valued random variables X and Y are ordered with respect to the* **increasing convex ordering** *($X \leq_{\mathrm{icx}} Y$) if (1) holds for all increasing, convex functions $f : \mathbb{R} \to \mathbb{R}$ for which the expectations exist. In this case we will also write $P^X \leq_{\mathrm{icx}} P^Y$.*

c) A homogeneous Markov chain $\{X(t), t \geq 0\}$ with probability transition functions $(p_{ij}(t))$ is said to be **stochastically monotone** *if for all $t \geq 0$ the sequence*

$$\sum_j p_{ij}(t) f(j), \;\; i \in \mathbb{Z}$$

is increasing for all increasing functions $f : \mathbb{Z} \to \mathbb{R}_+$. In this case we will also say that the generator Q of the Markov chain is stochastically monotone.

d) A homogeneous Markov chain $\{X_n\}$ is said to be **stochastically monotone** *if for the corresponding one-step transition matrix P, the sequence*

$$\sum_j p_{ij} f(j), \;\; i \in \mathbb{Z}$$

is increasing for all increasing functions $f : \mathbb{Z} \to \mathbb{R}_+$. In this case we will also say that the transition matrix P of the Markov chain is stochastically monotone.

Remark: Note that monotonicity of the generator is equivalent to the statement that for two initial distributions fulfilling $\beta \leq_{\mathrm{st}} \gamma$ we have $P_\beta(X(t) \in .) \leq_{\mathrm{st}} P_\gamma(X(t) \in .)$ for all $t \geq 0$ and monotonicity of the Markov chain is equivalent to that $\beta \leq_{\mathrm{st}} \gamma$ implies $\beta P^n \leq_{\mathrm{st}} \gamma P^n, \; n \in \mathbb{N}$.

Suppose now that we consider two queueing systems as defined above, which differ only on their intensity process. Let the first intensity process be $\{\lambda(t), t \geq 0\}$ with generator Q and the second intensity process be $\{\lambda'(t), t \geq 0\}$ with generator Q', where $Q' = cQ$, with some fixed $c \in (0, 1]$. Denote by W and W' the stationary work-loads of the two systems respectively. Note that $\rho < 1$ implies the stability of both the systems, since the stationary distribution π is independent of c. The parameter c allows us to control the variability of the intensity process. In the model without prime the intensity process changes faster.

Theorem 1 If Q and Q^* are stochastically monotone, then for the respective stationary work-loads

$$W \leq_{\mathrm{icx}} W'.$$

Remarks:

a) Theorem 1 includes Example 4.4 of Chang, Chao and Pinedo [5], where the generator $Q = (q_{ij})$ was defined by:

$$q_{ij} = \begin{cases} \alpha_j & , i \neq j \\ -\sum_{\nu \neq j} \alpha_\nu & , i = j \end{cases} \qquad \alpha_\nu \in \mathbb{R}_+, \ \alpha := \sum_{i=1}^{m} \alpha_i.$$

It is easy to check that Q is stochastically monotone and since $\pi = (\frac{\alpha_1}{\alpha}, \dots, \frac{\alpha_m}{\alpha})$ is the stationary distribution, we obtain $Q^* = Q$. Therefore Theorem 1 can be applied.

b) The assumptions on Q and Q^* are fulfilled for example by every birth-and-death process (cf. Stoyan [11]). In particular every 2 state continuous time Markov chain is monotone and reversible. The monotonicity of $EW(c)$ with respect to $c > 0$ for the governing two state intensity process was conjectured by Ross [9] and it was solved by Chang, Chao and Pinedo [5].

c) We do not know whether the assumption that the generator Q and in particular Q^* are monotone are necessary. In a recent paper of Asmussen, Frey, Rolski and Schmidt [1], the authors proved (among other results) that $W'' \leq_{\mathrm{st}} W$, where W'' is the stationary work-load in the standard $M/G/1$ queue with arrival rate $E\lambda(0)$ and service time distribution S, also under assumption that Q is stochastically monotone.

4. The proof

In order to prove the result, we need the following preparatory definition (see e.g. Meester and Shanthikumar [6]):

Definition 2

a) *A function $f : \mathbb{R}^n \to \mathbb{R}$ is called* **supermodular**, *if for all $x, y \in \mathbb{R}^n$:*

$$f(x \vee y) + f(x \wedge y) \ \geq \ f(x) + f(y)$$

where $x \vee y$ $(x \wedge y)$ denotes the coordinatewise maximum (minimum) of x, y.

b) *A function* $f : \mathbb{R}^n \to \mathbb{R}$ *is called* **directionally convex**, *if for all* $x_1, \ldots, x_4 \in \mathbb{R}^n$ *with* $x_1 \leq x_2, x_3 \leq x_4$ *and* $x_1 + x_4 = x_2 + x_3$*:*

$$f(x_1) + f(x_4) \;\geq\; f(x_2) + f(x_3)$$

c) *Two random vectors* X, Y *are ordered with respect to the* **supermodular ordering** *(*$X \leq_{\mathrm{sm}} Y$*) if for all measurable supermodular functions* $f : \mathbb{R}^n \to \mathbb{R}$

$$E[f(X)] \leq E[f(Y)], \tag{2}$$

whenever the expectations exist.

d) *Two random vectors* X, Y *are ordered with respect to the* **increasing directional convex ordering** *(*$X \leq_{\mathrm{idcx}} Y$*) if (2) holds for all increasing, directionally convex functions* $f : \mathbb{R}^n \to \mathbb{R}$ *whenever expectations exist.*

e) *For two stochastic processes* $\{X(t), t \geq 0\}$ *and* $\{Y(t), t \geq 0\}$ *we write*

$$\{X(t), t \geq 0\} \;\prec\; \{Y(t), t \geq 0\}$$

if for all $0 \leq t_1 \leq \ldots \leq t_n < \infty,\ n = 1, 2, \ldots$

$$(X(t_1), \ldots, X(t_n)) \;\prec\; (Y(t_1), \ldots Y(t_n)),$$

where $\prec$ *is one of the above defined orderings.*

Remarks:

a) Note that a function f is directionally convex if and only if f is supermodular and convex in each coordinate (see Meester and Shanthikumar [6]).

b) A useful inequality is the so so-called Lorentz inequality, which states that for identically distributed random variables $X_1, \ldots, X_n$ we obtain $(X_1, \ldots, X_n) \leq_{\mathrm{sm}} (X_1, \ldots, X_1)$.

We proceed the proof of Theorem 1 by two lemmas of independent interest. Lemma 1 provides a generalization of Lemma 3.3 in Bäuerle [2]. Let P and P' be two $m \times m$ irreducible stochastic matrices. Let $Z_0, Z_1, \ldots$ be the stationary Markov chain with one-step transition matrix P and $Z'_0, Z'_1, \ldots$ be the stationary Markov chain with one-step transition matrix P'. In the next lemma both the chains have the same stationary distribution π. In the sequel by I we denote the identity matrix.

Lemma 1 If P and P^* are stochastically monotone, and if $P' = (1-c)I + cP$, then

$$(Z_0, \ldots, Z_n) \ \leq_{\rm sm} \ (Z'_0, \ldots, Z'_n), \ n = 0, 1, 2, \ldots. \tag{3}$$

Proof: Let $\{X_n\}$ be a Markov chain (nonhomogeneous), with transition matrix P' for the first $i < n$ transitions and transition matrix P for the next $(n-i)$ transitions, and let $\{Y_n\}$ be a Markov chain (nonhomogeneous) with transition matrix P' for the first $(i+1)$ transitions and transition matrix P for the next $(n-i-1)$ transitions. Moreover, we suppose that $X_0 \stackrel{d}{=} Y_0 \stackrel{d}{=} \pi$. We show first that for all $n = 0, 1, 2 \ldots$

$$(X_0, \ldots, X_n) \ \leq_{\rm sm} \ (Y_0, \ldots, Y_n). \tag{4}$$

Let $f : I\!\!R^n \to I\!\!R$ be a supermodular function. We have

$$\begin{aligned} &Ef(Y_0, \ldots, Y_n) = \\ &= \ cEf(X_0, \ldots, X_n) + (1-c)Ef(X_0, \ldots, X_i, X_i, X_{i+1}, \ldots, X_{n-1}). \end{aligned}$$

Hence it suffices to prove $(X_0, \ldots, X_n) \ \leq_{\rm sm} \ (X_0, \ldots, X_i, X_i, X_{i+1}, \ldots X_{n-1})$. Let $U = (\ldots, U_{-1}, U_0, U_1, \ldots)$ be a sequence of independent and identically distributed random variables uniformly distributed over $[0,1]$ and independent from everything else. Define $g, h : \{\lambda_1, \ldots, \lambda_m\} \times [0,1] \to \{\lambda_1, \ldots, \lambda_m\}$ as follows

$$h(\lambda_i, x) = \lambda_j, \qquad \text{if} \quad \sum_{\nu=1}^{j-1} p_{i\nu} \leq x < \sum_{\nu=1}^{j} p_{i\nu}$$

$$g(\lambda_i, x) = \lambda_j, \qquad \text{if} \quad \sum_{\nu=1}^{j-1} p'^*_{i\nu} \leq x < \sum_{\nu=1}^{j} p'^*_{i\nu}.$$

For $u \in I\!\!R^{Z\!\!\!Z}$, denote $\tilde{f}_u(x, y) : \{\lambda_1, \ldots, \lambda_m\}^2 \to \{\lambda_1, \ldots, \lambda_m\}^{n+1}$ in the following way

$$\tilde{f}_u(x, y) = (\ldots, g(g(x, u_{-1}), u_{-2}), g(x, u_{-1}), x, y, h(y, u_2), h(h(y, u_2), u_3), \ldots)$$

where x stands at position $i+1$ in the vector. Now

$$(X_0, \ldots, X_n) \stackrel{d}{=} \tilde{f}_U(Z_0, Z_1)$$

where $Z_0 \stackrel{d}{=} \pi$, $P(Z_1 = j \mid Z_0 = i) = p_{ij}$ and

$$(X_0, \ldots, X_i, X_i, X_{i+1}, \ldots, X_{n-1}) \stackrel{d}{=} \tilde{f}_U(Z_0, Z_0).$$

Due to the stochastic monotonicity of P and P'^*, the functions g and h are increasing in the first coordinate. (Obviously they are also increasing in the second component). Therefore, for every supermodular function $f : I\!R^n \to I\!R$ the composition $f \circ \tilde{f}_u$ is again supermodular. Applying Lorentz inequality we have

$$\tilde{f}_u(Z_0, Z_1) \leq_{\rm sm} \tilde{f}_u(Z_0, Z_0)$$

for every fixed $u \in I\!R^{\mathbb{Z}}$. We complete the proof of (4) by taking expectation with respect to U . Now the proof of (3) follows inductively. $\square$

We fix $c \in (0, 1]$ as in Theorem 1 and define

$$P = I + \frac{1}{\alpha}Q \qquad \text{and} \qquad P' = (1-c)I + cP = I + \frac{c}{\alpha}Q, \tag{5}$$

where $\alpha \geq \sup_i q_i$.

Lemma 2 If $\{\lambda(t), t \geq 0\}$ and $\{\lambda'(t), t \geq 0\}$ are two intensity processes as defined in the model, both with initial distribution π and generator Q and Q' respectively and if Q and Q^* are stochastically monotone, then

$$\{\lambda(t), t \geq 0\} \ \leq_{\rm sm} \ \{\lambda'(t), t \geq 0\}.$$

Proof: According to definition 2e) we have to show

$$(\lambda(t_1), \ldots, \lambda(t_n)) \ \leq_{\rm sm} \ (\lambda'(t_1), \ldots \lambda'(t_n)),$$

for all $0 \leq t_1 \leq \ldots \leq t_n < \infty$, $n = 1, 2, \ldots$. We look at the uniformized versions of the processes. Let P and P' be defined as in (5). Furthermore let $\{\sigma_n\}$ be a sequence of independent and identically distributed random variables with $\sigma \sim exp(\alpha)$, $\{Z_n\}$, $\{Z'_n\}$ are Markov chains which are governed by transition matrix P and P' respectively, and $Z_0 \stackrel{d}{=} Z'_0 \stackrel{d}{=} \pi$. (Note that π is again a stationary distribution for P and P'.) Now the process $\{\lambda(t), t \geq 0\}$ can be constructed in the following way (cf. Stoyan [11])

$$\begin{aligned}\lambda(0) &= Z_0 \\ \lambda(t) &= Z_n, \quad \text{if } \textstyle\sum_{j=1}^n \sigma_j \leq t < \sum_{j=1}^{n+1} \sigma_j.\end{aligned}$$

The construction of $\{\lambda'(t), t \geq 0\}$ is analogous. Since Q is stochastically monotone, we have that for two initial distributions $\beta \leq_{\text{st}} \gamma$

$$\beta P^n \leq_{\text{st}} \gamma P^n, \qquad n = 1, 2, \ldots$$

Therefore P is again stochastically monotone. With the same argument we obtain the stochastic monotonicity of $P'^* = (p'^*_{ij})$, where $p'^*_{ij} = p'_{ji}\frac{\pi_j}{\pi_i}$. Now by Lemma 1

$$(Z_0, \ldots, Z_n) \ \leq_{\text{sm}} \ (Z'_0, \ldots, Z'_n). \tag{6}$$

From Bäuerle [3], Lemma 2.2 we have that also for all supermodular functions $f : I\!R^n \to I\!R$:

$$Ef(Z_{k_1}, Z_{k_1+k_2}, \ldots, Z_{k_1+\ldots+k_n}) \leq Ef(Z'_{k_1}, Z'_{k_1+k_2}, \ldots, Z'_{k_1+\ldots+k_n})$$

for all $k_1, \ldots, k_n = 0, 1, \ldots$ and $n = 1, 2, \ldots$. Hence

$$\begin{aligned}&Ef(\lambda(t_1), \ldots, \lambda(t_n)) = \\ &= \sum_{k_1=0}^{\infty} \cdots \sum_{k_n=0}^{\infty} \frac{(\alpha t_1)^{k_1}}{k_1!} \cdots \frac{(\alpha(t_n - t_{n-1}))^{k_n}}{k_n!} e^{\alpha t_n} Ef(Z_{k_1}, Z_{k_1+k_2}, \ldots, Z_{k_1+\ldots+k_n}) \\ &\leq \sum_{k_1=0}^{\infty} \cdots \sum_{k_n=0}^{\infty} \frac{(\alpha t_1)^{k_1}}{k_1!} \cdots \frac{(\alpha(t_n - t_{n-1}))^{k_n}}{k_n!} e^{\alpha t_n} Ef(Z'_{k_1}, Z'_{k_1+k_2}, \ldots, Z'_{k_1+\ldots+k_n}) \\ &= Ef(\lambda'(t_1), \ldots, \lambda'(t_n)),\end{aligned}$$

which completes the proof. □

Now we are able to prove Theorem 1.

Proof of Theorem 1: From Lemma 3.10 of Meester and Shanthikumar [6] and Lemma 2 we obtain for $0 \le t_1 \le t_2 \le \ldots \le t_n < \infty$, $n = 1, 2, \ldots$,

$$(N(t_1), N(t_2) - N(t_1), \ldots, N(t_n) - N(t_{n-1})) \le_{\text{idcx}}$$
$$(N'(t_1), N'(t_2) - N'(t_1), \ldots, N'(t_n) - N'(t_{n-1}))$$

and due to Lemma 3.14 in Meester and Shanthikumar [6]

$$(M(t_1), M(t_2) - M(t_1), \ldots, M(t_n) - M(t_{n-1})) \le_{\text{idcx}}$$
$$(M'(t_1), M'(t_2) - M'(t_1), \ldots, M'(t_n) - M'(t_{n-1})). \tag{7}$$

Let $f : I\!\!R \to I\!\!R$ be convex and increasing. Define $t_j(i) = \frac{j}{2^i}$ and $k(i) = i2^i$. With the same argument as in Lemma 1 of Rolski [7] , we get

$$Ef(W) = \lim_{i \to \infty} Ef(\max\{0, M(t_1(i)) - t_1(i), M(t_2(i)) - t_2(i), \ldots, M(t_{k(i)}(i)) - t_{k(i)}(i)\}).$$

Define $g_s(x) = \max\{0, x_1 + s_1, x_1 + x_2 + s_2, \ldots, x_1 + \ldots + x_k + s_k\}$. Rolski [7] showed that $g_s(.)$ is increasing, supermodular and convex in each coordinate, hence increasing and directionally convex. Therefore $f \circ g_s$ is again increasing and directionally convex. Now from (7) we obtain $W \le_{\text{icx}} W'$ and the proof of Theorem 1 is complete. □

References

[1] Asmussen, S., Frey, A., Rolski, T. and Schmidt, V. (1995) Does Markov-modulation increase the risk ? *ASTIN Bull.* **25**, 1995, 49–66.

[2] Bäuerle, N. (1997) Monotonicity results for MR/GI/1 queues. To appear in *J. Appl. Probab.* **34**.

[3] Bäuerle, N. (1997) Inequalities for stochastic models via supermodular orderings. To appear in *Commun. Statist. - Stochastic Models.*

[4] Borovkov, A. A. (1972) Random Processes in Queueing Theory. *Nauka, Moskva.*

[5] Chang, C., Chao, X. and Pinedo, M. (1991) Monotonicity results for queues with doubly stochastic Poisson arrivals: Ross's conjecture. *Adv. Appl. Prob.* **23**, 210-228.

[6] Meester, L. and Shanthikumar, J. (1993) Regularity of stochastic processes. *Prob. Eng. and Inf. Sci.* **7**, 343-360.

[7] Rolski, T. (1986) Upper bounds for single server queues with doubly stochastic Poisson arrivals. *Math. Oper. Res.* **11**, 442-450.

[8] Rolski, T. (1989) Queues with nonstationary inputs. *Queueing Systems* **5**, 113-130.

[9] Ross, S. , (1978) Average delay in queues with non-stationary Poisson arrivals. *J. Appl. Probab.* **15**, 602-609, (1978).

[10] Shaked, M. and Shanthikumar, J. (1994) *Stochastic Orders and Their Applications.* Academic Press, New York.

[11] Stoyan, D. (1983) *Comparison Methods for Queues and Other Stochastic Models.* Wiley, Chichester.

[12] Szekli, R., Disney, R. L. and Hur, S. (1994) MR/GI/1 queues with positively correlated arrival stream. *J. Appl. Probab.* **31**, 497-514.